

# The Role of Timing in Shaping Information Processing in Neural Systems

Ovidiu Dan Iancu

BS, Mathematical Sciences, Oregon State University, 1997

BS, Computer Science, Oregon State University, 1997

MS, Mathematics, Oregon State University, 1999

Presented to the Division of Biomedical Engineering within  
The Department of Science & Engineering  
and the Oregon Health & Science University  
School of Medicine  
in partial fulfillment of  
the requirements for the degree of  
Doctor of Philosophy  
in  
Biomedical Engineering

September 2008

---

Patrick Roberts, Ph.D.  
Thesis Advisor  
Adjunct Professor, Department of Physiology and Pharmacology  
Adjunct Assistant Professor, Biomedical Engineering  
Oregon Health Sciences University

---

Curtis Bell, Ph.D.

---

Tamara Hayes, Ph.D.  
Assistant Professor, Biomedical Engineering  
Oregon Health Sciences University

---

William Roberts, Ph.D.  
Assistant Professor, Biomedical Engineering  
Oregon Health Sciences University

---

Dan Hammerstrom, Ph.D.  
Professor, Electrical and Computer Engineering  
Portland State University

# Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Synaptic response to precisely timed inputs . . . . .	3
1.2	System level integration of timing information and its role on sensory processing and motor corollary signals . . . . .	4
1.3	Topographical representation of complex vocalizations in the telencephalon of zebra finch . . . . .	4
<b>2</b>	<b>Synaptic response to precisely timed inputs</b>	<b>6</b>
2.1	Importance of timing in the electrosensory system . . . . .	7
2.2	Methods . . . . .	9
2.2.1	Construction of the compartmental model . . . . .	9
2.2.2	Tuning the model using a genetic algorithm . . . . .	12
2.3	Results . . . . .	13
2.3.1	Behavior of the genetic algorithm tuning procedure . . . . .	13
2.3.2	Reproducing biophysical properties of granular cells . . . . .	13
<b>3</b>	<b>Response attenuation to paired sounds in the primary auditory cortex</b>	<b>17</b>
3.1	Relationship between interstimulus interval and response magnitude . . .	18
3.1.1	Experimental setup . . . . .	19
3.2	Methods . . . . .	20
3.2.1	Methods for the analysis of <i>LFP</i> traces . . . . .	20
3.2.2	Methods for the analysis of spike trains . . . . .	26
3.3	Results . . . . .	29
3.3.1	Different neural representations of sound stimuli and somatosensory inputs . . . . .	29
<b>4</b>	<b>Topographical representation of complex vocalizations in the telencephalon of zebra finch</b>	<b>41</b>
4.1	Temporal structure of song reflected in spatial patterns of cell activation	42

4.2	Methods . . . . .	44
4.2.1	Automatic detection of nuclear and cytoplasmic <i>IEG</i> expression .	44
4.2.2	Selective temporal or spectral degradation of songs . . . . .	68
4.3	Results . . . . .	71
4.3.1	Visualization of the density of activated cells . . . . .	72
<b>5</b>	<b>Conclusions and future directions</b>	<b>79</b>
5.1	Timing and coincidence detection at different levels of neural processing .	79
5.2	Further refinements of the methods . . . . .	81
5.2.1	Translation of timing into spike count . . . . .	81
5.2.2	Wavelet transform for time-scale analysis . . . . .	83
5.2.3	Voronoi tessellations and a measure of similarity between overlapping sets of points . . . . .	84
	<b>References</b>	<b>87</b>

## List of Tables

2.1	Biophysical parameter values of the model. . . . .	11
2.2	Compartmental dimensions of the model. . . . .	12
2.3	Final parameter values of the model. . . . .	14
4.1	Concordance coefficients for manual and automatic cell counts. . . . .	61
4.2	Concordance coefficients for ZENK labeling of cell nuclei. . . . .	65
4.3	Confusion matrix of true and assigned stimulus category of each image. .	74

## List of Figures

2.1	A. Current-voltage relationship from slice data. B. Two types of relationship between holding potential and EPSP size. . . . .	8
2.2	Model morphology. A myelinated afferent axon terminates into a terminal, which is connected to the granular cell soma by an electrical gap junction. The granular cell consists of a soma connected to the axon by a high resistance hillock. In addition, the granular soma is connected to a dendritic compartment with a large membrane surface (not shown in figure). . . . .	9
2.3	A. Fitness of best individual. B. Euclidian distance between vectors representing parameter models (individuals). C. Average population fitness. . . . .	14
2.4	A. Data and model current-voltage relationship. B. Dependence of EPSP on the postsynaptic potential. . . . .	15
2.5	Voltage traces generated by simulating an afferent spike in the afferent axon. A large afferent spike is attenuated at the terminal, presumably due to the inactivation of Na channels. Granular soma postsynaptic response has small amplitude but a long duration. This voltage deflection generates the spike in the granular axon. . . . .	16
3.1	Two alternative forced choice task training paradigm: water is associated with two tone presentation, and no water is presented following one tone. <i>Courtesy of Paulo Rodrigues</i> . . . . .	20
3.2	Performance improvement as over the course of training. Filled circles: percentage lick for two tone trials. Empty circles: percentages lick for one tone trials. <i>Courtesy of Paulo Rodrigues</i> . . . . .	21

3.3 A. Two different but overlapping distributions are labeled as N:negatives (left bell-shaped curve) and P: positives (right bell-shaped curve). The distributions are separated using a variable threshold (vertical line). TP: true positives, FP: false positive, TN: true negatives. B. Plot of true positives versus false positives, expressed in percentages, as the threshold value is varied over the whole range of data values. P(TP) is proportion of true positives, P(FP) is the proportion of false positives resulting from a particular threshold. Green arrow illustrates how the *ROC* curve is traced by varying the threshold. . . . . 22

3.4 Information content along *ROC* curves. Black curves: iso-information, marked with information in bits. Red curves: *ROC* curves. A: Iso-information curves for  $P(\text{positives}) = 0.5$ . The maximum information content point is the intersection of the *ROC* curve and the highest iso-information line. B:  $P(\text{positives}) = 0.2$ . The iso-information curves are shifted by the change in the a priori probabilities. . . . . 23

3.5 A: Illustration of the *LDA* dimensionality reduction procedure, for two dimensions. Data points are projected on the direction most discriminative. *Modified from (Park and Fry, 2004)*. B: Performance (*AUC*) in discriminating between the two categories is computed at each time point. The projection on *LDA* direction has superior performance to any individual time point. . . . . 24

3.6 Illustration of a series of steps that transforms the spike pattern on top (A) to the pattern on the bottom (B). Red spike in is deleted during the first step, green spike is added during 7th step (row). Blue arrows mark the spikes shifted for a cost of qt. . . . . 27

3.7 A. LFP responses to either one (top) or two tones (bottom) Trials where the animal licked are represented on the right. B. Spike rasters collected after either one or two tones. Blue: two tones, lick. Green: One tone, lick. Red: two tones, no lick. Cyan: one tone, no lick. . . . . 30

3.8 Firing rate (spikes/sec) versus time (sec). A: one tone, no lick. B: one tone, with lick. C: two tone, no lick. C: two tones, lick. Firing rates were computed using a Gaussian averaging window with width of 20 ms. Blue trace represents mean firing rate, green traces show the standard deviation. . . . . 31

3.9 A: *ROC* analysis of the power spectrogram contrasting one tone, no lick with the two tones, lick trials. B: *ROC* analysis of the coherograms contrasting the same data as in A. . . . . 32

3.10	A: <i>Isomap</i> projection of the tone (red) and no tone spike trains for $q = 0$ . The tone (red) and no tone (blue) are not well discriminated based on spike count alone. Axes represent two dimensional projections of the data (arbitrary units on axes not represented). B: Discrimination information versus shift cost $q$ . Larger $q$ 's attach more weight to spike timing and improve discrimination. C: <i>Isomap</i> projection for $q = 100$ . The discrimination is much improved (compare to A). . . . .	33
3.11	Boxplots quantifying differences between response to first and second tone. The boxes have lines at the lower quartile, median, and upper quartile values. The whiskers extend an addition 1.5 interquartile range to show the extent the data. Non-overlapping notches indicate that the medians of the two groups differ at the 5% significance level. A: example of a recorded <i>LFP</i> channel. B: aggregate spike counts from all recorded animals and channels. . . . .	34
3.12	Amount of response attenuation as a function of the intertone interval. Longer intertone intervals lead to a smaller attenuation of the response to second tone. . . . .	35
3.13	A: Decreasing response to the first tone over the course of the experiment, as measured by the projection on the LDA axis. B: Response to the second tone, as in A. No significant trend detected. . . . .	36
3.14	A: Correlation between total spectral power in 80-130 Hz range and the amount of response attenuation. B: Correlation coefficient between level response attenuation and power at each frequency. . . . .	36
3.15	A: Response attenuation in correct versus incorrect trials in <i>LFP</i> 's. B: Spike count differences between correct and incorrect trials. . . . .	38
3.16	Relative increase in intertone <i>LFP</i> power relative to pre-tone baseline. A: average ratio of power in correct (blue) and incorrect (red) trials. B: p-values in comparing the values in A. C: <i>AUC</i> based on the quantities in A. D: maximum information content based on <i>ROC</i> 's in C. . . . .	39
4.1	Example of a <i>ZENK</i> labeled brain slice, photographed under 40X magnification under an optical microscope. . . . .	45
4.2	A: Typical nuclear <i>ZENK</i> expressing nuclei. The red spots represent different loci of transcription. B: Typical cytoplasmic expression of <i>ZENK</i> . The fluorescent labelling consists of a concentration of color nearly surrounding the nuclei. . . . .	46

4.3	Graylevel images from the blue channel. A: the image prior to opening. The edges are blurry and irregular. B: Image after the opening operation. The shapes are more rounded, the graylevel inside the nuclei is more uniform. . . .	47
4.4	Example of the Canny edge detection algorithm applied to the edge image. Overlapping cell nuclei that partially occlude each other still retain their nearly circular shape on the non-overlapping portions of the nuclei. . . . .	49
4.5	Procedure for finding the center of circle starting from $(x, y)$ coordinates. Circles with radius $r$ , centered at each $(x, y)$ , will intersect at the desired center $(a, b)$ . . . . .	50
4.6	A: Illustration of the effect of computing the Hough transform for a radius slightly smaller than the true radius. The circles drawn using the $(x, y)$ points in the edge image will trace the small red circle, instead of intersecting on a point. B: The Hough transform is computed for a set of different radii $r$ , resulting in a conical surface spanned by the points of intersection. The apex of this cone correspond to the correct $r$ . . . . .	50
4.7	Sections through the Hough accumulator array, corresponding to increasing radii. A: $r = 7$ pixels. B: $r = 8$ pixels. C: $r = 9$ pixels. D: $r = 10$ pixels, closest value to the true radius. . . . .	51
4.8	A: raw image. B: Hough transform for a radius $r$ . Each nuclear shape in image above gives rise to a high intensity point in the Hough transform, surround by an artefactual circle of radius double the radius of the cell nucleus. . . . .	52
4.9	Local intensity information associated with candidate cell nuclei. A: Histogram of average intensity values inside detected circles. B: Histogram of ratios between the average intensity inside the candidate cell nucleus and the average intensity in the region immediately adjacent. . . . .	53
4.10	Comparison of clustering results. $x$ -axis: average intensity inside the nucleus. $y$ -axis: ratio inside versus outside the candidate nucleus. A: Information clustering. B: k-means clustering. Note the labeling of more points in red at the boundary between clusters. . . . .	54
4.11	Example of detected cell nuclei. A: original image. B: detected cell nuclei are marked with a green circle. . . . .	55
4.12	A: Example of the attraction forces generated by the edge points. Arrows indicate the direction of the force, with the length proportional with the magnitude. B: Effects of the force on the original snake (blue curve). The edge information is represented by the red curve. The deformed contour is represented by the green curve. . . . .	56

4.13	Example of fully labeled image. Top: raw image with blue and red channels. Bottom: image with nuclear labeling (red points, see also yellow arrow) and cytoplasmic labeling (red circles, see also green arrow). Red arrow: example of double-labeled cell. . . . .	59
4.14	<b>A.</b> Illustration of the counts performed by humans and the counts performed by the automatic procedures. <b>B.</b> Direct comparison between the Hough and the LoG automated procedures. Perfect agreement would place all the points on the main diagonal. . . . .	61
4.15	<b>A.</b> Image of 100 cells with variability in size and intensity. The images do not overlap, their borders touch due to the averaging filter that blurs the edges. <b>B.</b> Image of 100 cells which are allowed to overlap by 15 pixels, and with Gaussian noise added. This was the most challenging of the artificial images. . . . .	63
4.16	<b>A.</b> Comparison of automated counting methods using image sets with variable cell size and intensity. <b>B.</b> Counting results for images with same number of cells as A, but with added noise. Difference between dashed lines and solid lines represent the false positives, as identified by manual inspection. . . . .	64
4.17	Results of manual counts and our automated procedure, applied to the 14 images in the data set. First three images were generated from the silent controls, images 4-6 from birds presented with conspecific songs, 7-9 from heterospecific songs stimuli, 9-12 from white noise and 13-14 from pure tone stimulation. . . . .	65
4.18	Illustration of the Delaunay triangulation and Voronoi tessellation of a set of points. Original points are denoted by black dots. Thick lines represent the Delaunay triangulation. The definition implies that no triangle sides intersect. Dashed lines outline the Voronoi polygons. . . . .	66
4.19	<b>A:</b> Two neighboring images with overlap on the borders. Delaunay triangles are identified in each image (filled triangles). <b>B:</b> The images are stitched together using the overlap suggested by the triangles. . . . .	67
4.20	Illustration of a Voronoi tessellation applied to the whole NCM. The areas of the polygons are inversely proportional with the intensity of the color, resulting in areas of high density in warmer colors. . . . .	68
4.21	Results of image registration. Overlapping lines show the contours of same brain regions from different animals. <b>A:</b> <i>CMM</i> , prior to registration. <b>B:</b> <i>CMM</i> , post registration. <b>C:</b> <i>NCM</i> , prior to registration. <b>D:</b> <i>NCM</i> , post registration. . . . .	69

4.22	Decomposition of original sound using partially overlapping filters. Each filter covers a specified frequency band. The output of these filters can be used to reconstruct the signal. The filters used covered the range between 500 and 8000 Hz. <i>Modified from (Theunissen and Doupe, 1998).</i> . . . . .	70
4.23	Spectrograms of song stimuli. A: unmodified song. B: temporally preserved song. C: spectrally preserved song. . . . .	71
4.24	Image processing steps. Starting from the positions of activated cells (red dots in top left image), we registered the images to a common standard, we combined them by common stimulus class, and then we displayed them as heatmaps. . .	72
4.25	Density of ZENK expressing cells in NCM. A: conspecific songs; B: canary songs' C: white noise; D: pure tones; E: unstimulated controls. . . . .	73
4.26	A: Two dimensional Isomap projection of the pairwise distances between images. B: Boxplots comparing the distances between same-category stimuli images and different category images. . . . .	74
4.27	The effect of reducing the number of points on the Hausdorff distance. Each curve represent the distance between a selected image and the rest of the images. As we decrease the number of points in the selected image, the distances to the other images remain relatively constant. . . . .	75
4.28	Excitatory (top row) and inhibitory (bottom row) density of activated cells in <i>NCM</i> , in response to auditory stimuli. . . . .	76
4.29	Spatial pattern of <i>ZENK</i> expressing cells in <i>NCM</i> . A-C: spectrograms of intact or modified songs. D-F: heatmaps representing the density of activated cells. .	77
4.30	Excitatory (top) and inhibitory (bottom) spatial distribution in response to original (left) and modified (middle, right) versions of the same song. . . . .	78
5.1	A. Simulation results illustrate translation from latency to spike count. Top: time of arrival of the second input. White portions represent the stimulus on. B: Phase space illustration of the canonical model. Parabola and straight lines are nullclines, short curve is the trajectory. Color along the trajectory represents the speed. . . . .	82
5.2	A. Spectrogram of a trial where the animal was presented with two tones and licked. The spectrogram was computed using overlapping windows 100 ms in length. B. Wavelet transform using the sym2 wavelet (Daubechies, 1994). Note that small scales (low on the <i>y</i> axis) approximately correspond to high frequencies (high on <i>y</i> axis in A) . . . . .	83

## Abstract

# The Role of Timing in Shaping Information Processing in Neural Systems

Ovidiu Dan Iancu, B.S., M.S.

Doctor of Philosophy

Division of Biomedical Engineering within  
The Department of Science & Engineering  
and the Oregon Health & Science University

School of Medicine

September 2008

Thesis Advisor: Patrick Roberts

In many biological systems, information about the environment is detected by a large array of sensory receptors; it reaches more central regions of the nervous system as parallel streams of spike trains. How this flow of information is processed and which features are most salient to the organism is a central problem in neuroscience. A comprehensive and systematic approach to the analysis of the relationships between stimuli and their neural representations is illustrated. This approach is complementary to the hypothesis driven research paradigm, where the investigator states a hypothesis and then performs experiments in order to validate or invalidate the stated hypothesis. This alternative methodology enlarges the possible relationships between stimuli and their representations beyond one specific hypothesis and aims to evaluate all possible relationships between neural stimuli and their neural representations. For a specific set of data, both the set of stimuli and the set of neural response are characterized either by membership in discrete categories or as a continuous space with a similarity measure. In order to quantify these input-output relationships, existing methods of analysis are adapted or new ones developed where necessary. In the specific data sets analyzed here, timing emerges as a critical parameter for the description of neural stimuli and their representations.

# Chapter 1

## Introduction

In many neurophysiological experiments, the investigator starts with a specific hypothesis which guides the experimental design, the collection of the data, and the analysis of the results. However, the resulting data sets could contain additional information, which, even though not part of the original hypothesis, can reveal important information about the neuronal system under study. In many scientific or economic fields, rapidly accumulating data sets have motivated the development of rigorous and systematic methods of exploring the data, methods generally known as *data mining*. The overall goal of such a comprehensive review and analysis of the collected data is "*the nontrivial extraction of implicit, previously unknown, and potentially useful information from data*" (Frawley *et al.*, 1992), for a review see also (Piatetsky-Shapiro and Frawley, 1991).

In the present work, we apply data analysis techniques to experimental data resulting from three neurophysiological experiments. Our general approach is to expand the possibilities of input-output relationships beyond those stated in the original hypothesis generated by the experimentalists. The goals of this project are twofold. First, we seek to find scientifically meaningful relationships between the inputs to the neuronal system and the resulting outputs, the neuronal activity. This approach is inspired from the classic information theory paradigm, which is usually applied to a communication channel in order to quantify how much the received message tells us about what was being sent originally (Shannon *et al.*, 1998). In the case of the neuronal system, the stimulus is the original message and the neuronal activity is the received message. Here, we aim to identify which specific parameters and which temporal resolution results in the most informative or predictive relationship between the stimuli and the resulting neuronal activity.

A second and equally important goal of our work is to develop analysis methods which are appropriate for each of the data sets and at the same time have the potential

to be of wider applicability in the field of neuroscience. In implementing our approach for each of the data sets, the analysis was necessarily specific to the details of the experimental results. However, in each case we paid particular attention to the issues regarding timing in the relationship between neuronal stimuli and the resulting neuronal activity. For example, firing rates are computed using a sliding window, the temporal width of which has to be chosen appropriately. Another example of the importance of timing can be illustrated by decomposing auditory inputs in the spectrogram, which can be accomplished at different levels of time or frequency precision.

An important constraint needs to be considered in order to fully understand and contextualize the relationship between the sensory stimuli and neural activity. This parameter is the *saliency* of the stimuli for the subsequent behavior of the animal. In mammalian auditory cortex data analyzed in *Chapter 3*, we have direct access to the behavioral choices of the animal, which are based on a discrimination task, and therefore the stimuli are highly salient to the animal. When we take into account the subsequent behavior of the animal, the analysis of the neural activity is faced with additional challenges. First, the time window for analysis is shortened drastically, since the behavioral choice succeeds the sensory presentation at a very short latency. In addition, the behavior itself can result in neural activity, even in neural structures considered primarily sensory. In this case, high temporal resolution of the analysis window is necessary to separate the sensory response from behavioral correlates.

In the songbird system analyzed in *Chapter 4*, we can infer a measure of the saliency of the sensory inputs by comparing their neural representations with those of stimuli known to be of high importance, such as conspecific songs. In this data set, high temporal resolution is available only for the space of inputs. Therefore, specific manipulations of the input parameters are performed, with the goal of identifying which manipulations have the biggest effect in the resulting neural activity.

While striving for general methods with wider applicability, we evaluate the validity of our analysis methods primarily by the biological relevance of the results which they achieve. Biological data in the form of neural recordings and spatial patterns of gene expression have guided our computational and modeling investigations. *In vitro* recordings from granular cells of the electric fish *Gnathonemus petersii* (Zhang *et al.*, 2007) have guided our biophysical model of synaptic integration of precise spiking inputs. Observations about the relative importance of spike latency *versus* spike count at subsequent levels of sensory processing spurred our investigation of a possible cellular mechanism of translation from spike latency to spike count. At the system level, we used recordings from the primary auditory cortex of awake behaving rats (*Paulo Rodrigues, personal*

*communication*) to investigate the importance of spiking precision in the representation of auditory information. Fluorescent images of immediate-early gene expression in the brain of songbirds (Velho, 2008) have inspired our development of automated imaging techniques which resulted in high resolution information about the distribution of activated neuronal cells. This topographical distribution of excitatory and inhibitory cells which are activated by a stimulus was further analyzed in order to draw inferences about possible local effect of inhibitory neuronal activity.

## 1.1 Synaptic response to precisely timed inputs

Our first set of data consisted of *in vitro* recordings (Zhang *et al.*, 2007) from neuronal slices of a cerebellar structure, the *electrosensory line lobe (ELL)* of the electric fish *Gnathonemus petersii*. The *ELL* is a laminar structure, with the granular layer containing small compact cells similar in morphology to the cerebellar granule cells (Zhang *et al.*, 2007). The granular cells receive two precisely timed inputs: a centrally originating signal which can be used as a timing reference, followed by a variable latency afferent spike encoding stimulus intensity. Current injections revealed the high sensitivity of the postsynaptic response to the holding potential of the cells, which we hypothesize is in turn determined by the time elapsed from the first input.

In *Chapter 2*, we systematically explore the relationship between the specific biophysical properties of a model granular cell and its synaptic response properties, in the context of high sensitivity of the cells to the interspike interval of the input spikes. A related goal is the development of an automated search method that can explore the space of biophysically plausible parameters for values capable of reproducing the experimental results. Our space of inputs is the space of all biologically plausible parameters of the model neuron, and the dissimilarity measure on this space is simply the Euclidean distance. The space of outputs is characterized by the appropriately defined distance between the behavior of the model and the behavior of the real neuron. In order to search the space of inputs for the set of parameters with the closest behavior to the real neuron, we adapt a search procedure capable searching high-dimensional spaces, namely a *genetic algorithm (GA)*.

## 1.2 System level integration of timing information and its role on sensory processing and motor corollary signals

Our second set of data consisted of *local field potentials* (LFP's) and neuronal spikes recorded from the primary auditory cortex of rats trained in a behavioral task based on auditory clues (*Paulo Rodrigues, personal communication*). In *Chapter 3*, our goals are to explore the relationship between the neural representation of simple auditory tones and the subsequent behavior of the animal. Once again, our general approach is to comprehensively analyze all the possible connections between the set of inputs and the set of neural responses. Our space of inputs consisted of two categories of sound stimuli, which the animals have been trained to discriminate. The space of outputs consisted of the two categories of responses - correct and incorrect. An intermediate space is the space of neural responses recorded from the primary auditory cortex while the animals were stimulated with the sounds and reacted with either a correct or incorrect response. The methods developed or adapted for this project aimed to achieve the best characterization of the space of recorded neural responses, consisting of spike trains and field potentials. The biological aim consisted of determining if any characteristics of the neural response in a sensory area (primary auditory cortex) have predictive power about a subsequent behavioral choice.

## 1.3 Topographical representation of complex vocalizations in the telencephalon of zebra finch

Our final set of data was comprised of spatial patterns of activated cells in the telencephalon of the zebra finch. The cells were activated in response to auditory stimulation with complex sounds such as conspecific songs. In *Chapter 4*, we analyze the relationship between the structure of sound stimuli presented to songbirds and the spatial patterns of activated cells resulting from presenting these stimuli. The space of inputs consisted of several categories of sounds played to the zebra finches: conspecific and heterospecific songs, white noise and pure tones. The space of outputs consisted of spatial patterns of activated cells, together with a similarity measure quantifying similarity between two dimensional sets of points. One specific question addressed is whether the spatial distribution of cells, in addition to the number of activated cells, is specific to each stimuli. The importance of the fine temporal structure of the inputs was revealed by specific modifications to the spectral and temporal parameters of the songs.

In order to make the above analysis feasible, we developed a method for automatically

identifying activated cells in labeled tissue sections. We test this procedure against manual counts performed by humans and also against a different automated procedure, both on images collected from zebra finches brain slices and also using sets of artificial images. Our procedure proves accurate and robust in both cases.

## Chapter 2

### Synaptic response to precisely timed inputs

Communication in neuronal systems primarily takes the form of discrete signals or spikes. As spikes arrive at the pre-synaptic terminals of neuronal cells, their effect on their targets results in *post-synaptic potentials (PSP's)*. The cumulative effect of a collection of *PSP's* can push the target neuron over the threshold and make it produce a spike. In many systems, neurons conform to the above model and generally act as *integrators*, modulating their firing rates or spike count in proportion to the strength of the input. However, there are examples of neurons that produce just one spike in response to appropriate stimulation and the exact timing of this spike encodes information, for instance by decreasing the time to the first spike in response to stronger stimulation (Bell, 1990).

In this chapter, we investigate the postsynaptic integration of a precisely timed spike arriving after a centrally originating timing signal. Electric field intensity on the mormyrid skin is encoded in the afferents by the latency from the *electric organ discharge (EOD)* to the first afferent spike. Primary afferent axons from electroreceptors terminate in the deep layers of *electrosensory line lobe (ELL)*, where they form electrical synapses on granular cells. Granular cells are the first stage of the sensory pathway to also receive descending input which marks the exact time or the *EOD*. The granular cells are a clear example of neurons for which the relevant inputs are the precise relative timing of *PSP's*. They are acting as coincidence detectors since their post-synaptic activity is dependent on these timing signals (Zhang *et al.*, 2007).

We applied our general approach of input-output analysis to the problem of finding parameters for a biophysically detailed model for the granular cells. Formally, our inputs consisted of specific electrophysiological properties of the cells, such as the current-voltage relationship and the size of the synaptic response. The outputs consisted of sets of parameters consistent with these measured properties of the granular cells. Once a set of parameters that conforms to experimental observations has been identified, the

resulting model is used in simulations intended to test and refine specific hypotheses resulting from the electrophysiological experiments.

## 2.1 Importance of timing in the electrosensory system

The mormyrid electric fish senses its environment by emitting an *EOD* and detecting the perturbations that nearby objects cause in its self-generated electric field. Specialized mormyromast electroreceptors sense the self-generated field and its distortions. Electric field intensity on the mormyrid skin is encoded in the afferents by the latency from the *EOD* to the first afferent spike. These responses are conveyed to the cortex of the *ELL* where the afferent fibers terminate in the mormyromast zone, affecting the granular cells through electrical synapses (Bell, 1990).

With each *EOD*, the *ELL* cortex is affected not only by input from the periphery but also by the *electric organ corollary discharge (EOCD)* signals that originate centrally. These *EOCD* signals are time-locked with the *EOD* motor command, which elicits the *EOD* (Zipser and Bennet, 1976). Therefore, the synaptic response of the granular cells needs to be responsive to the precise relative timing of the two *PSP*'s.

The mormyrid *ELL* has a highly regular structure similar to the organization of the cerebellum (Bell *et al.*, 1997). Work on the mormyrid *ELL* is, thus, relevant to other structures, such as the gymnotid *ELL* (Bastian, 1995), the octavolateral nucleus of sharks and rays (Montgomery and Bodznick, 1994), and the dorsal cochlear nucleus of mammals (Oertel and Young, 2004). An advantage of studying the mormyrid *ELL* is that the information carried by the centrally generated signals is well understood: these signals inform the *ELL* when an *EOD* has taken place. The *EOCD* signals originate in central structures such as the *juxtalobar nucleus (JLN)* and provide a fixed latency signal to the deeper layers of the *ELL*. Interaction between the variable afferent input and the fixed latency corollary discharge signal is thought to be the means by which afferent latency is decoded as a measure of stimulus intensity (Bell *et al.*, 1992).

*In vitro* patch clamp recordings from the granular cell soma have revealed large (5-25 mV), all-or-none *electrical excitatory postsynaptic potentials (eEPSPs)* in response to afferent fiber stimulation (Zhang *et al.*, 2007). These *eEPSPs* are consistent with the gap junctions observed morphologically (Bell, 1989). Additionally, in previous experiments, antidromic spikes were recorded in the afferent axon, suggesting that the conductance of the gap junction is significant in both directions (Bell, 1990). Experimental data shows that the amplitude, and even the occurrence of the afferent evoked *eEPSP*'s depends on the granular cell's own membrane potential. The *eEPSP* is absent at depolarizations

above  $-40$  mV. As the membrane is hyperpolarized below this level, the *eEPSP* grows in a roughly linear manner, but then disappears abruptly with hyperpolarization beyond  $-90$  mV. The *eEPSP* dependence on the holding membrane potential appears to be stronger than the expected, ohmic effect of the reversal potential. In a subset of granular cells, the disappearance of the *eEPSP* proceeds in a more gradual manner (See Figure 2.1).

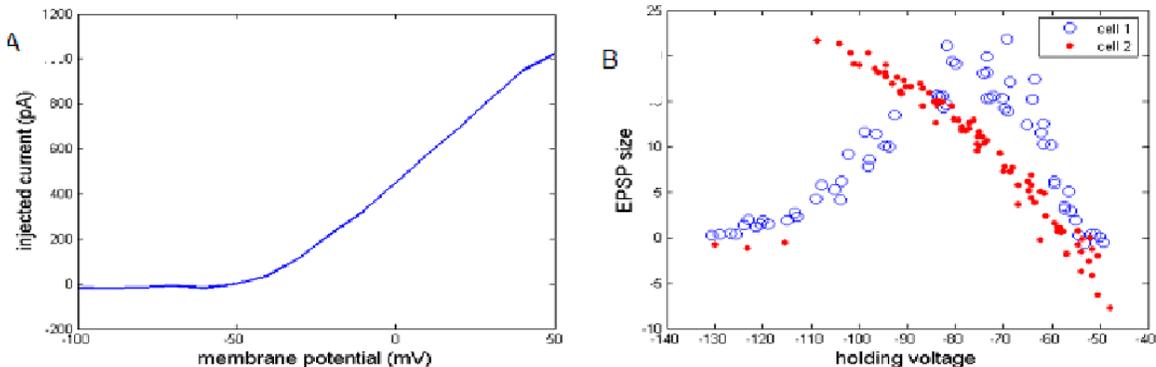


Figure 2.1: A. Current-voltage relationship from slice data. B. Two types of relationship between holding potential and EPSP size.

One hypothesis generated by these experimental results was that the activation and inactivation levels of the *Na* channels in the afferent terminals are sensitive to the membrane potential of the granular cells. Such a mechanism of bidirectional interaction would provide an explanation for the sensitivity of the postsynaptic response to the holding voltage, which we hypothesize to depend on the time elapsed since the *EOD*. The sequence of events at the terminal-granular cell gap junction would proceed as follows: *EOD* input would arrive at the granular cell in the form of release of chemical neurotransmitter, affecting the voltage of the granular cell. This voltage then propagates through the gap junction to the afferent terminal, where it affects the inactivation levels of *Na* channels. Therefore, the elapsed time from the *EOD* is the determinant factor for the amount of current injected into the granular cells when the afferent spikes arrives at the terminal. At that time, the available *Na* channels in the terminal are activated and current is injected into the terminal and, through the gap junction, into the granular cell. An alternative hypothesis would be that the granular cell itself has *Na* channels. This alternative is not consistent with the experimental data, since the size of the spikes recorded at the granular soma is relatively small (Zhang *et al.*, 2007), which precludes the presence of *Na* channels in the granular soma itself.

A biophysically realistic model of the granular cell and afferent terminal complex

would allow for simulations intended to test and potentially corroborate the presumptive mechanisms of synaptic interactions outlined above. Therefore, our next step was to integrate available information about these cells into a mathematical model. This model was based on the observed morphology and the pharmacological evidence about the presence of specific ionic channels in these cells (Zhang *et al.*, 2007).

## 2.2 Methods

The morphology and connectivity between the afferent neurons and the granular cells are relatively well-described (Zhang *et al.*, 2007). Based on this data, we created a detailed biophysical model of the afferent cell terminal and granular cell circuit, using the *NEURON* simulation environment (Hines and Carnevale, 1997). We tuned this compartmental model using a *genetic algorithm (GA)* approach, using custom software developed in *Matlab (MathWorks, Inc.)*. This model, in turn, was used to test our hypotheses outlined above, and could also be used to inform further electrophysiological experiments.

### 2.2.1 Construction of the compartmental model

Our model included the afferent axon with two myelinated sections separated by *nodes of Ranvier*, an afferent terminal and a gap junction. The gap junction was connected to the granular cell, which was comprised of a small compact soma, a dendrite and an axon (see Figure 2.2). We tuned the model parameters, within a biologically plausible range encountered in cells with similar morphology (Hille, 2001), (D'Angelo *et al.*, 2001) to reflect the measured electrical properties of the neuronal circuit.



Figure 2.2: Model morphology. A myelinated afferent axon terminates into a terminal, which is connected to the granular cell soma by an electrical gap junction. The granular cell consists of a soma connected to the axon by a high resistance hillock. In addition, the granular soma is connected to a dendritic compartment with a large membrane surface (not shown in figure).

Three significant features of the physiological recordings guided our construction

of the model, features which we aimed to replicate in simulations. First, the spikes recorded at the soma were small (10 – 20 mV) and the membrane potential was not reset to the resting potential following a spike. This implied to us that the spike was generated at a significant electrotonic distance from the soma thus reducing its recorded amplitude and preventing the reset of the soma membrane potential. In the model, a high-resistance hillock was inserted in between the granular soma and axon. Second, the cells show outward rectification due presumably to non-inactivating, depolarization activated outward current. Third, the cells have a long time constant (greater than 40 ms) and high input resistance ( $\sim 2 \text{ G}\Omega$ ). The long time constant implies a relatively large membrane surface which was reflected in the model by the inclusion of a dendritic compartment, consistent with the observed morphology of the granular cells (Zhang *et al.*, 2007).

The design of the currents in the soma relied on the second and third features of the physiological data described above. We inserted a non-inactivating potassium current,  $I_{KV}$ , into the soma and we allowed the *GA* tuning procedure to modify the maximum conductance, since the reversal potential of this type of channel is likely responsible for the deflection of the *IV* curve, illustrated in Figure 2.1. The input resistance was measured experimentally (Zhang *et al.*, 2007) and the capacitance was determined from the resistance and the time constant. The size of the spikes recorded at the soma was therefore dependent on the hillock resistance. We simulated the voltage time course at the granular cell soma as the *JLN* chemical *EPSP* is succeeded, at variable latencies, by the afferent spike. The construction of our model assumed a single compartment for the granular soma. For this type of model, the time course of the voltage can be approximated by the equation

$$c_m \frac{dV}{dt} = -\bar{g}_L(V - E_L) + \frac{I_E}{A} + \frac{I_{KV}}{A}$$

where  $c_m$  is the specific membrane capacitance,  $\bar{g}_L$  is the maximum leak conductance per unit area,  $E_L$  is the leak current reversal potential,  $A$  is the total area of the cell and  $I_E$  is the sum of currents entering the granular cell soma from the granular hillock and the gap junction.

Following the general framework of the Hodgkin-Huxley model (Hodgkin and Huxley, 1952), each ionic conductance is approximated using the formula:

$$I = \bar{g}_L n^k h (V - V_{rev})$$

where  $\bar{g}_{KV}$  is the maximum conductance,  $n$  is the activation variable,  $k$  is the number of gating events necessary to open and therefore activate a channel,  $h$  is the inactivation

Table 2.1: Biophysical parameter values of the model.

	$\bar{g}(\frac{\text{mS}}{\text{mm}^2})$	$k$	$V_{rev}(\text{mV})$	$\alpha(\text{sec}^{-1})$	$\beta(\text{sec}^{-1})$
$I_{Na,terminal}$	0.12 to 2.4	3	50	$\frac{0.1(V+40)}{1-e^{-0.1(V+40)}}$	$4e^{-0.056(V+65)}$
inactivation		1		$0.07e^{-0.05(V+65)}$	$\frac{1}{1+e^{-0.1(V+35)}}$
$I_K,terminal$	0.036 to 0.72	1	-77	$\frac{0.01(V+55)}{1-e^{-0.1(V+55)}}$	$.125e^{-0.125(V+65)}$
$I_{leak,soma}$	0.001 to 0.1		-54		
$I_{KV,soma}$	0.001 to 0.5	4	-70	$\frac{0.13(V+25)}{1-e^{-0.1(V+25)}}$	$1.69e^{-0.0125(V+35)}$
$I_{Na,axon}$	0.12	3	50	$\frac{0.1(V+40)}{1-e^{-0.1(V+40)}}$	$4e^{-0.056(V+65)}$
inactivation	0.036	1		$0.07e^{-0.05(V+65)}$	$\frac{1}{1+e^{-0.1(V+35)}}$

variable,  $V$  is the instantaneous voltage and  $V_{rev}$  is the reversal voltage of the channel.

The probability of each of the  $k$  gating units being open is dependent on the instantaneous voltage  $V$  and is governed by the equation

$$\frac{dn}{dt} = \alpha(V)(1 - n) + \beta(V)n$$

with  $\alpha(V)$  and  $\beta(V)$  being specific for each ionic channel. The  $h$  variable dependence on voltage is similar.

Based on measured values in cerebellar granule cells (D'Angelo *et al.*, 2001) and their morphological similarity with the granular cells of the *ELL* (Zhang *et al.*, 2007), the parameters of the granular cells and afferent terminal were allowed to vary within physiological ranges which have been measured in channels of these types (Hille, 2001). We present the range of parameters used in our model in Table 2.1.

The specific dimensions of each compartment, together with the axial resistances, are also necessary for a full description of the model. The axial resistance is inversely proportional to the area of the segment and grows linearly with the length. The constant of proportionality is the axial resistivity, which is also a parameter in the model. In addition, the long time constant measured of the granular cells required a high total capacitance. Since the capacitance of neuronal membranes has a fairly constant value per unit area, (Hille, 2001), the large capacitance is likely a result of a large surface area. Experimental observations (Zhang *et al.*, 2007) revealed that the granular soma itself has a small size. Therefore we hypothesize that the large dendritic processes connected to the soma are likely the source of the additional membrane surface. We summarize below the size and the axial resistivity of each compartment. Some of the parameters were allowed to vary during the tuning procedure We we present the admissible ranges for their values in Table 2.2.

Table 2.2: Compartmental dimensions of the model.

	length ( $\mu\text{m}$ )	diameter ( $\mu\text{m}$ )	axial resistivity ( $k\Omega \cdot \text{mm}$ )
afferent axon	500	2	100
afferent terminal	1 to 3	1	100
granular soma	8	8	100
granular dendrite	600 to 1000	1	100
hillock	10	0.1	100 to 300
axon	20	0.5	100

## 2.2.2 Tuning the model using a genetic algorithm

Finding a suitable set of parameters which reproduces basic properties of the cells required searching a large dimensional parameter space. A flexible method to search the range of parameters is provided by a *genetic algorithm* approach (Keren *et al.*, 2005). The procedure involves constructing a set of models, each with different parameters, and evaluating their fitness using a *fitness value*. In our case, the fitness value  $F$  reflects the difference of desired behavior (experimental data) and the behavior of the model. Minimizing this difference results in fitting the voltage-current relationship (*IV curve*), the *eEPSP* dependence of the holding voltage, and also the relatively long time constant of the granular cell:

$$F(\theta) = F_{IV}(\theta) + F_{EPSP}(\theta) + w \times (\tau_{cell} - \tau_m(\theta))^2$$

where  $F_{IV}$  is the *mean squared error* contribution from the *IV curve*,  $F_{EPSP}$  the contribution from the *eEPSP* and  $w \times (\tau_{cell} - \tau_m(\theta))^2$  reflects the error contribution from the difference between the time constant of the cell  $\tau_{cell}$  and the time constant of the model cell  $\tau_m(\theta)$ . The set of parameters, collectively represented by  $\theta$ , consisted of the dimensions of the cellular compartments in the model (afferent axon and terminal, gap junction, granular soma, dendrite, hillock and axon) together with the electrophysiological properties of the membranes, such as capacitance and specific conductance.

In the context of *GA* operation, an *individual* consists of a model constructed using a specific set of parameter values. A *population* consists of a collection of individuals, with each individual characterized by a unique set of biophysical parameters. After a random initial population of individuals is created, each individual is evaluated using its *fitness value*, which measures how well the model conforms to the experimental data. In our implementation, the initial population was generated by independently choosing a random value for each parameter, drawn from a uniform distribution bounded by the range derived from experimental data (Zhang *et al.*, 2007). The next step consists of computing the fitness value of each individual and assigning to each a *rank*, reflecting

their position in the list of ordered fitness values. The rank is used to assign to each individual a probability of being selected for reproduction.

There were three choices for reproduction or creation of children, or members of the next generation. *Elite* children were the individuals kept unchanged from the current generation. They were the individuals with the best fitness values, and their preservation guaranteed a monotonic decrease in the error from one generation to the next. *Crossover* children were created by combining the parameter values of two parents. For the first generations, our implementation of the genetic algorithm used random selection, where each parameter value will be chosen at random from one of the two parents. For the later generations, we used an intermediate value for the value of the parameter. In this case the parameter value of the child was chosen from a uniform distribution bounded by the values of the parameter values of the two parents. *Mutation* children were created by introducing random changes, or mutations, to some of the parameters of a single parent. In our implementation, the random mutation consisted of choosing a parameter value by drawing from the same uniform distribution used to construct the first generation.

Criteria for termination of the algorithm was based on examining the diversity of the population. This was accomplished by measuring the distance between the individuals in each generation and measuring the average fitness value of all individuals. When these quantities have decreased significantly the algorithm was terminated and the best parameter values were reported.

## 2.3 Results

### 2.3.1 Behavior of the genetic algorithm tuning procedure

The behavior of this tuning procedure is displayed in Figure 2.3. We see a relatively rapid convergence in the best individual fitness value, which reaches a stable value in about 10 generations. In generations 10 – 20, the improvement in best individual fitness (Figure 2.3A) is much smaller, but the population diversity and average generation fitness decrease further (Figure 2.3 B and C). After further 10 generations, our criteria for stopping the search procedure is satisfied and we select the best individual model for further exploration.

### 2.3.2 Reproducing biophysical properties of granular cells

The best individual resulting from our *GA* tuning procedure are listed in Table 2.3.

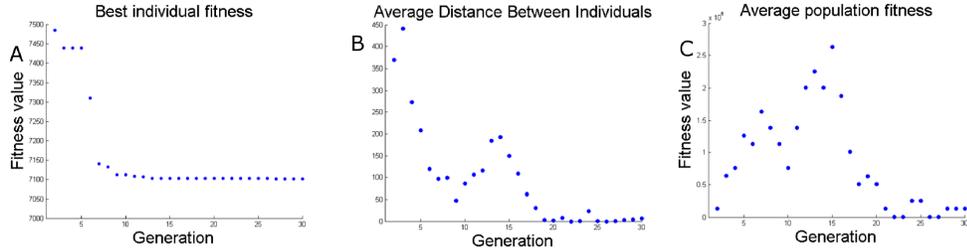


Figure 2.3: A. Fitness of best individual. B. Euclidian distance between vectors representing parameter models (individuals). C. Average population fitness.

Table 2.3: Final parameter values of the model.

$I_{Na}$ , terminal	$2.15 \frac{\text{mS}}{\text{mm}^2}$
$I_K$ , terminal	$0.41 \frac{\text{mS}}{\text{mm}^2}$
$I_{leak}$ , soma	$0.06 \frac{\text{mS}}{\text{mm}^2}$
$I_{KV}$ , soma	$0.42 \frac{\text{mS}}{\text{mm}^2}$
afferent terminal	$2, 6 \mu\text{m}$
granular dendrite	$874 \mu\text{m}$
hillock axial resistivity	$212k\Omega \cdot \text{mm}$

We note the large value of the  $Na$  conductance, which is probably required in order for a sufficient amount of charge to be injected into the granular soma. The large  $K$  conductance in the terminal is required for the rectifying phase of the afferent terminal action potential. The  $I_{KV}$  conductance is responsible for the knee in the  $IV$  curve illustrated in Figure 2.1 A. The afferent terminal size, in conjunction with the density of active  $Na$  conductance, generates the electrical charge entering the granular soma through the gap junction. The size of the granular dendrite is large, reflecting the large capacitance which underlies the long time constant of the granular cells. The relatively high hillock resistivity ensures that spikes generated in the granular axon do not backpropagate into the soma.

The parameter values outlined above are consistent with biophysical mechanisms likely to underlie the behavior of the granular-afferent cell complex and also with the intuitive outline of the sequence of events generated by an afferent spike. However, finding specific values for the 7 parameters which fine tune the biophysical responses has proven very difficult to achieve manually. In contrast, our search procedure using the  $GA$  has been proven to be able to find parameter values which are balanced and satisfy all the competing requirements.

Specifically, our model cell reproduced the current-voltage relationship measured in granular cells, as illustrated in Figure 2.4 A. In addition, the model replicated many

of the properties of the electrical *EPSP*, in particular its modulation by the membrane potential of the granular cell. Our simulation results were therefore consistent with the hypothesized mechanism of pre-synaptic control of the *EPSP* size, namely the electric coupling through the gap junction (Zhang *et al.*, 2007). In Figure 2.4 B we illustrate the relative contribution of the *Na* channel inactivation vs. simple ohmic dependence on the holding potential. For different holding potentials, we injected a constant amount of current, approximately equivalent with the *Na* current due to an afferent spike at a holding potential of  $-90$  mV, into the terminal. The resulting *EPSP* relation to holding potential is due solely to ohmic factors. This procedure revealed that the electric coupling between the axon terminal and the granular cell has a significant effect on the post-synaptic *EPSP*.

The model also reproduced the abrupt disappearance of the *EPSP* at about  $-90$  mV. This effect was presumed to be due to inhibition of the presynaptic spike by the hyperpolarizing current from the gap junction. (Zhang *et al.*, 2007), hypothesis consistent with our modelling results.

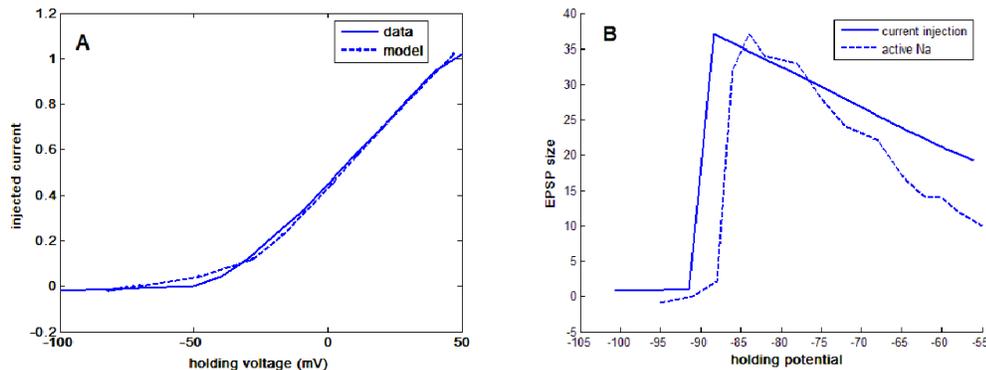


Figure 2.4: A. Data and model current-voltage relationship. B. Dependence of EPSP on the postsynaptic potential.

Physiological recordings under voltage clamp (Zhang *et al.*, 2007) showed that currents of a magnitude of about 1 nA underlie the *eEPSP*. To simulate the *eEPSP*, we injected a brief current into the afferent, such that a spike was initiated. The afferent spike propagates to the afferent terminal and, through the gap junction, propagates further into the granular cell soma, as illustrated in Figure 2.5.

The results of the simulation illustrated in Figure 2.5 revealed that an afferent spike generates a sharp deflection in the voltage recorded at the granular soma. This deflection is smaller than a deflection generated by a typical spike, as expected since the granular soma does not have active *Na* channels and it is connected to the terminal by a gap

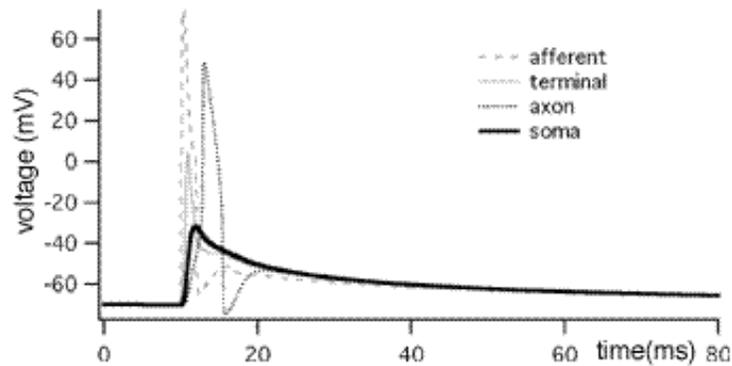


Figure 2.5: Voltage traces generated by simulating an afferent spike in the afferent axon. A large afferent spike is attenuated at the terminal, presumably due to the inactivation of Na channels. Granular soma postsynaptic response has small amplitude but a long duration. This voltage deflection generates the spike in the granular axon.

junction with a relatively high axial resistance. In addition, the soma voltage relaxed toward its original resting voltage in a slow manner, consistent with the electrophysiological recordings (Zhang *et al.*, 2007). Nevertheless, a spike was generated in the granular cell axon, which is connected to the soma by the high-resistance hillock.

The importance of this model lies the expansion the possibilities of investigation. The availability of the model allows us access to parameters such as the level of *Na* activation in the pre-synaptic compartment, a parameter which is inaccessible during *in vitro* experiments. We note that, due to the high dimensionality of the search space, we have no guarantee about the uniqueness of the parameter set found. In the strict sense, we have shown the existence of a solution, but we have no guarantees about its uniqueness. However, the fact that our mathematical model reproduces in a biologically realistic manner so many of the characteristics of the granular cells and its interaction with the afferent terminal is notable. Such a level of concordance between the experimental and modeling results makes it unlikely that the model operates in a drastically different manner from the granular cells of the *ELL*.

## Chapter 3

# Response attenuation to paired sounds in the primary auditory cortex

In complex neuronal systems such as the mammalian primary auditory cortex, large numbers of neurons contribute to the representation of sensory information, decision making and motor action. In this chapter we analyze cortical responses to pairs of identical tones presented as part of a behavioral task. Neural responses consist of recordings of simultaneous field potentials and spikes from multiple sites. We show that awake behaving rats display differences between the responses to the first and second tone, a phenomenon known as *response attenuation*, consistent with previous findings in several species (Näätänen and Picton, 1987), (Mäntysalo and Näätänen, 1987), (Gottlieb *et al.*, 1989), (Javitt *et al.*, 1992), (Brosch *et al.*, 1999). We investigate how the response attenuation changes during the course of the experiments, which consist of several hundred trials. We also inquire whether the amount of response attenuation is predictive of the performance in a behavioral task.

As mentioned in *Introduction*, our aim is to explore the data in a comprehensive manner in order to find significant relationships between the stimuli, neural representations and subsequent behavioral choices. In this set of data, the neural stimuli consisted of two simple categories: either a single tone or a pair of tones. The behavioral choices consisted of either a lick or the absence of a lick in response to stimuli. In contrast, the space of neural responses recorded consisted of traces *LFP's* and spike trains recorded simultaneously from multiple electrodes. Consequently, we identify and evaluate a set of methods which can be used for describing these high dimensional responses and compare them across different stimulus classes or behavioral responses.

In order to evaluate the statistical significance of our findings, we adapted established pattern recognition procedures such as *linear discriminant analysis (LDA)* (Duda *et al.*, 2000) to the specific characteristics of our *LFP* data. Subsequently, we employed a

non-parametric bootstrapping procedure (Chernick, 1999), which allowed an evaluation of the statistical significance of the findings with only minimal assumptions about the specific probability distribution. For analysis of the spike trains resulting from auditory stimulation, we used a *spike metric* approach, which allows quantitative evaluation of the amount of information that spike patterns convey about a given set of stimuli, taking into consideration both the spike counts and spike timings.

In this series of experiments, we had direct access to the behavioral saliency of different features of neural representation, since the animals were awake and performing a task related to the stimuli presented. We observed that the behavioral action of the animal succeeded the stimulus presentation at a very short latency, presenting challenges for the use of analysis techniques such as spectral decomposition of the recorded signals. This observation motivates the choice of analysis techniques used.

### 3.1 Relationship between interstimulus interval and response magnitude

In many species, the brain responds to closely spaced identical sounds with diminishing neural activity (Näätänen and Picton, 1987), (Mäntysalo and Näätänen, 1987), (Gottlieb *et al.*, 1989), (Javitt *et al.*, 1992), (Brosch *et al.*, 1999). While this effect is consistently found across species and recording paradigms, its role in sensory processing and in subsequent behavioral choices remains unclear (Werner-Reiss *et al.*, 2006). This dependence on sound history is thought to be involved in auditory perception phenomena at various timescales, including the precedence effect and forward masking at short timescales, and speech comprehension and loudness adaptation at longer timescales (Werner-Reiss *et al.*, 2006). Consequently, a number of studies have investigated and quantified the diminishing neural response to the second of two closely spaced but separate auditory stimuli: anesthetized cat (Hoehnerman and Gilat, 1981), (Calford and Semple, 1995), (Brosch and Schreiner, 1997), (Lu *et al.*, 2000), (Reale and Brugge, 2000), awake baboon (Gottlieb *et al.*, 1989), anesthetized monkey (Brosch *et al.*, 1999), and awake rabbit: (Fitzpatrick *et al.*, 1999). A more recent study has detailed the long persistence (5 s or even longer) and inter-stimulus interval dependence of the diminishing effect of successive stimuli in awake behaving monkeys (Werner-Reiss *et al.*, 2006).

High frequency oscillations (gamma range: 30 to 100 Hz) in the neocortex have been correlated with different aspects of sensory processing, and have been proposed as a possible sensory binding mechanism or as correlates of behavioral states such as attention and arousal (Lakatos *et al.*, 2004), for a review see also (Lee *et al.*, 2003).

We analyzed the responses of awake behaving rats to sequences of two closely spaced, identical sounds. Our data consisted of recordings from several channels of a multiple electrode array placed over the primary auditory cortex. Our stimuli consisted of either one or two 10 kHz tones, each 20 ms long, with the intertone interval varied between 100 and 400 ms. Some of the animals were trained to recognize and respond to two tones by licking a water spout, which dispensed water only in the two tones trials. The training regimen allowed us to evaluate the behavioral effects of the neural responses to sound, by quantifying whether we can predict the future behavior of the animal based on characteristics of the neural response to these behaviorally relevant auditory stimuli.

Our experimental paradigm allowed the observation of the response attenuation in the auditory cortex of awake behaving rats, the quantification of this effect and comparison with previous studies, specifically in relation to the magnitude of the attenuation as a function of inter-stimulus interval duration (Werner-Reiss *et al.*, 2006). We have also investigated the evolution of the relationship between the neural response to the first and second tone over the course of our experiment, which consisted typically of several hundred trials. In addition, the inter-stimulus interval of 200 ms was long enough to permit the estimation of the power spectrum with an acceptable level of accuracy, given our sampling rate  $f_s = 1\text{KHz}$ . We investigated the relationship between the spectral power of gamma-frequency oscillations in the *LFP*, measured in the interval between the tones, and the amount of response attenuation to the second tone. Finally, we investigated whether the response attenuation and gamma-frequency power have predictive power in relation to the future behavior of the animal.

### 3.1.1 Experimental setup

For several weeks preceding the recording experiments, rats were trained to distinguish between one-tone *versus* a two-tone sound presentation. The animals were presented with 2 kHz tones, 20 ms long. During this training or acquisition phase, animals learned to associate two tones with a reward, which was lacking when just one tone was presented. The interval between the tones was varied randomly between 25 ms and 400 ms, with the intention of conditioning the animal to the presence of two tones, as opposed to learning a specific intertone interval. Water was released through the spout at a latency of 800 ms following the second tone. Neural recordings in the form of *LFP*'s and spikes were extracted using standard neurophysiologic laboratory procedures (*Paulo Rodrigues, personal communication*).

During the course of several weeks of training, the animals gradually learned to

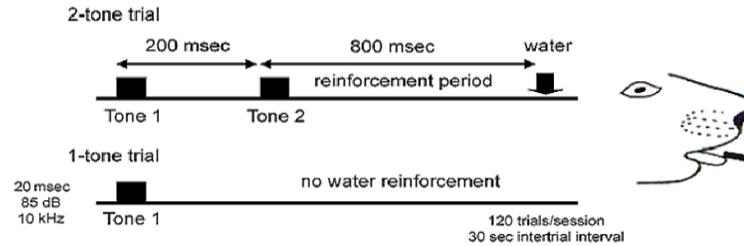


Figure 3.1: Two alternative forced choice task training paradigm: water is associated with two tone presentation, and no water is presented following one tone. *Courtesy of Paulo Rodrigues.*

associate two tones with the expected reward and started to lick the water spout in anticipation of the reward (see Figure 3.1). The animals were considered trained when the frequency of licking in response to two tones was significantly higher than the frequency of licking in response to one tone. The animal had the opportunity to lick and was judged to have a correct or incorrect trial during a window of time of 800 ms following the second tone. A trial is judged correct if the animal licks in response to two tones, and if it does not lick following a one tone presentation. The evolution of the performance during several weeks of training is illustrated in Figure 3.2.

## 3.2 Methods

### 3.2.1 Methods for the analysis of *LFP* traces

#### Dimensionality reduction and statistical analysis

We have collected *LFP* waveforms in a time window starting 10 ms before the onset of each tone and ending 50 ms after the onset of the tone, resulting in a collection of 60 dimensional vectors. In order to quantify differences between these vectors, one approach employed was *receiver operator characteristic (ROC)* analysis (Green and Swets, 1966). Each of the 60 time points (dimensions) was initially compared independently, in order to find significant differences between the responses to the first and second tone. The process is illustrated in Figure 3.3. A threshold is chosen to separate the values in two categories. The amount of overlap between the assigned categories and the true categories of the data is computed. The process is repeated for all possible values of the

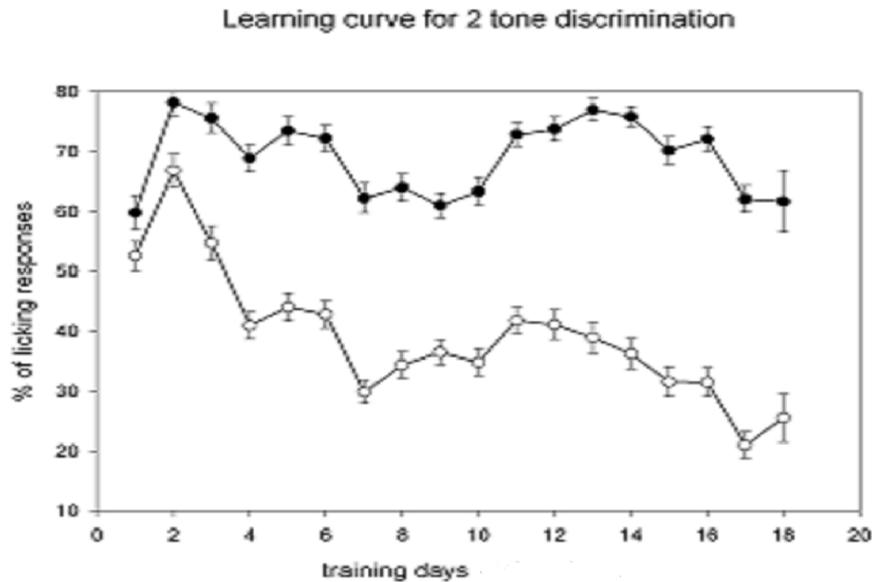


Figure 3.2: Performance improvement as over the course of training. Filled circles: percentage lick for two tone trials. Empty circles: percentages lick for one tone trials. *Courtesy of Paulo Rodrigues.*

separating threshold. Each such value of the threshold results in proportions of correctly and incorrectly labeled data points. Plotting of true positives *versus* the false positives traces a curve between the points  $(0, 0)$  and  $(1, 1)$ . For a perfect classifier, the curve will trace the left and top segments of the unit square, resulting in an area under the curve (*AUC*) of one. We also note that the diagram in Figure 3.3 is symmetric with respect to labeling in A and with respect with the main diagonal in B, respectively: if the labeling of the  $T$  and  $N$  data points is switched, the *ROC* curve will be traced in the other side of the main diagonal, resulting into an *AUC* between 0 and 0.5.

The reasons for employing the *ROC* analysis are twofold. First, this method is very sensitive to small differences between two distributions. A deviation from the 0.5 value signifies a potential difference between the two sets of values, which can subsequently be evaluated using more rigorous statistical tests. Second, computing the *AUC* is relatively fast, which is convenient when a very large number of comparison need to be computed. Examples of such large data sets include collections of spectrograms where statistical comparisons are performed on each coefficient in the spectrogram separately.

The *AUC* has a clear interpretation when used in the context of two-alternative forced choice task: it is the probability of making a correct choice (Green and Swets, 1966).

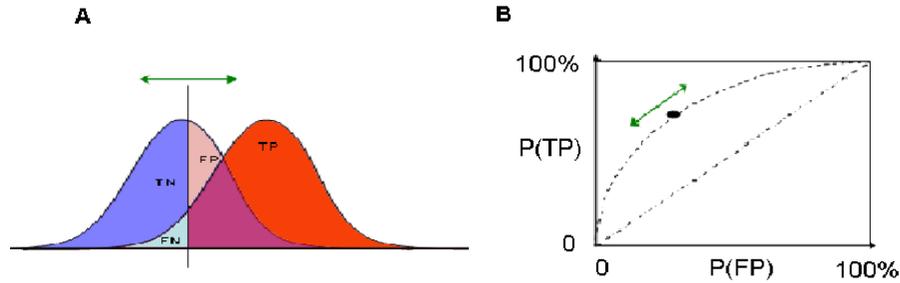


Figure 3.3: A. Two different but overlapping distributions are labeled as N:negatives (left bell-shaped curve) and P: positives (right bell-shaped curve). The distributions are separated using a variable threshold (vertical line). TP: true positives, FP: false positive, TN: true negatives. B. Plot of true positives versus false positives, expressed in percentages, as the threshold value is varied over the whole range of data values.  $P(TP)$  is proportion of true positives,  $P(FP)$  is the proportion of false positives resulting from a particular threshold. Green arrow illustrates how the *ROC* curve is traced by varying the threshold.

However, this measure can be ambiguous when comparing two curves from different experiments, where the probability of a positive signal is different (Metz *et al.*, 1973). In this case, a more appropriate measure of the performance of a classifier is provided by the *maximum information content* of a particular *ROC* curve. A specific point on the *ROC* curve, called an *operating point*, has associated with it the amount of *transmitted information* or *Shannon information* (Shannon, 1946).

We note that the *transmitted information* measure, in contrast with the *AUC* measure, makes use of the *a priori* probabilities  $P(T)$  and  $P(F)$ . The transmitted information defines a set of curves in the unit square, called *iso-information curves* (Metz *et al.*, 1973). Along each of these curves, the amount of transmitted information is constant (See Figure 3.4). We evaluate each *ROC* curve by the operating point on the *ROC* curve which belongs with the maximal iso-information curve. This procedure allows meaningful comparisons among *ROC* curves generated by different *a priori* probabilities  $P(T)$  and also among *ROC* curves which have same *AUC* but intersect in just one point.

In a subsequent step in our analysis, we used *LDA* in order to reduce the dimensionality of the recorded neural responses. This procedure is similar to *principal component analysis*, but where the latter attempts to find an ordered basis based on the values of the eigenvalues, the former attempts to project the data onto an ordered orthogonal basis with the first vector corresponding to the dimension which is the most discriminative between a set of categories (Duda *et al.*, 2000).

In order to illustrate the procedure, we consider the problem of projecting an  $n$ -

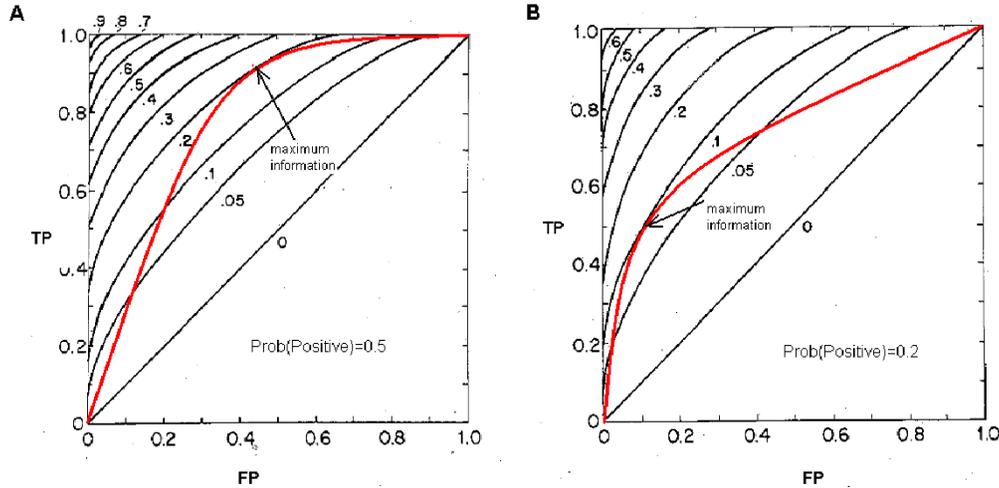


Figure 3.4: Information content along *ROC* curves. Black curves: iso-information, marked with information in bits. Red curves: *ROC* curves. A: Iso-information curves for  $P(\text{positives}) = 0.5$ . The maximum information content point is the intersection of the *ROC* curve and the highest iso-information line. B:  $P(\text{positives}) = 0.2$ . The iso-information curves are shifted by the change in the a priori probabilities.

dimensional set of points  $\mathbf{x}$ , divided into two categories  $D_1, D_2$ , onto a single direction. We can accomplish this projection by choosing a set of weights  $w_i, i = 1..n$  and computing the inner product  $\mathbf{w} \cdot \mathbf{x}$ :

$$y = \mathbf{w} \cdot \mathbf{x} = w_1x_1 + \dots + w_nx_n$$

Choosing the set of weights  $w$  is a subject of a set of constraints. The samples  $y_1, y_2$  correspond to the two categories in our data, and let  $m_i = \frac{1}{n_i} \sum_{y \in D_i} y$  be the means of the projected two categories of data points. The *LDA* procedure consists of making the difference in the projected means  $|m_1 - m_2|$  small relative to the variance of the data. This requirement can be formally represented as a generalized eigenvalue problem (Duda *et al.*, 2000). The solution can also be interpreted as the set of weights  $\mathbf{w}$  which minimizes the *within group variance* relative to the *between group variance* of the two sets of projected data points  $y_1, y_2$ . A simple illustration of this process is displayed in Figure 3.5 A.

We applied this dimensionality reduction procedure to our data. Restricting the representation of the data to the most discriminative component, we obtained a value  $y$  for each *LFP* trace. For each trial, the amount of response attenuation was quantified as the difference of the scalar value corresponding to the first tone and second tone.

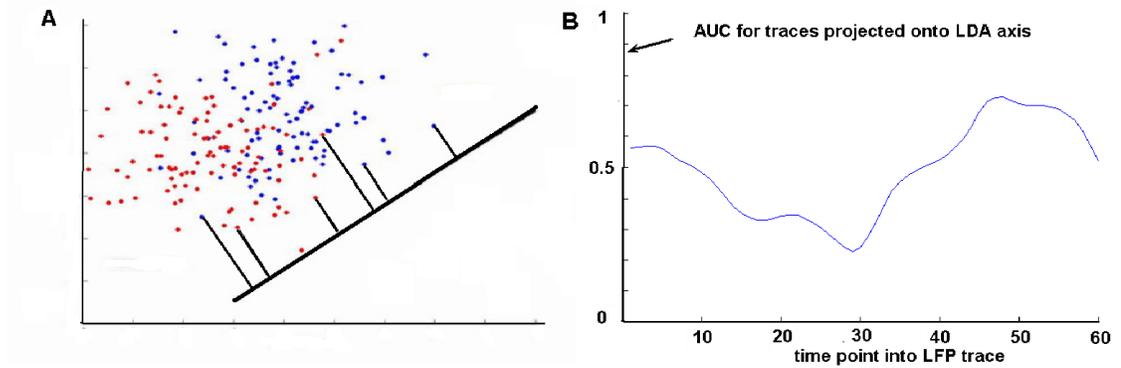


Figure 3.5: A: Illustration of the *LDA* dimensionality reduction procedure, for two dimensions. Data points are projected on the direction most discriminative. *Modified from (Park and Fry, 2004)*. B: Performance (*AUC*) in discriminating between the two categories is computed at each time point. The projection on *LDA* direction has superior performance to any individual time point.

In choosing the appropriate length of our *LFP* trace, we first evaluated the individual performance of each time point, as measured by *AUC*, in discriminating between first and second tone (see Figure 3.5 B). As differences between traces become informative, the plot departs from the  $AUC = 0.5$  value, which represents chance performance. The discrimination information becomes nearly zero ( $AUC = 0.5$ ) around 60 ms after the tone, providing us with a criterion for choosing the appropriate length of *LFP* trace to be selected for *LDA* analysis. When the scalar value which is the output of *LDA* projection is used for discrimination, its performance is superior to that of any individual time point (Figure 3.5 B).

When investigating the statistical significance of our results, we chose statistical tests that are not dependent on normality assumptions for the data. For two sample comparison test, we used the *Wilcoxon rank sum test for equal medians* (Gibbons and Chakraborti, 2003). When computing the strength of correlation between two random variables, we used the non-parametric *Spearman  $\rho$  coefficient* (Spearman, 1987). Since our  $\rho$  coefficient values were often relatively small, we evaluated their statistical significance using a non-parametric bootstrap procedure (Press *et al.*, 2007). Unless otherwise noted, all our statistical comparisons and correlation inferences were significant at  $p < 0.05$ , as evaluated using these non-parametric statistical techniques.

### Multitaper methods for spectral estimates

*LFP* data consisted of electrode recordings that were low-pass filtered at 500 Hz in order to prevent contamination from spiking activity (*Paulo Rodrigues, personal com-*

communication). Spectral analysis of the continuous data in *LFP*'s was performed using multitaper methods, which provide an optimal balance between the bias and variance of estimated spectral quantities (Mitra and Pesaran, 1999), and consequently have gained acceptance on the neuroscience community. In our particular case, the relatively short length of the traces used to compute spectral quantities (150 ms for a sampling frequency of 1000 Hz) made it imperative to use low-bias, low variance methods of spectral estimation. In the multitaper approach to spectral estimation, a segment of data is multiplied by a taper as a preprocessing step in advance of Fourier transformation. Slepian functions (Slepian and Pollak, 1961) are used as data tapers, due to their property of concentrating the frequency power in the interval  $[-W..W]$ , for each choice of time length  $T$ . The improvement in the estimation of spectral quantities is mainly due to the use of multiple Slepian functions as data tapers (Slepian and Pollak, 1961).

Formally, our multitaper spectral estimate will be expressed as an average over several tapered estimates (Mitra and Pesaran, 1999) <sup>1</sup>:

$$S_{MT}(f) = \frac{1}{K} \sum_{k=1}^K |\tilde{x}_k(f)|^2$$

where

$$\tilde{x}_k(f) = \sum_1^N w_t(k) x_t \exp(-2\pi i f t)$$

with  $N$  being the data points in the time domain  $x_1..x_N$  and  $w_t(k)$ , ( $k = 1, 2, \dots, K$ ) being  $K$  orthogonal taper functions with characteristics determined by the bandwidth parameter  $W$ . We note that the bandwidth parameter  $W$  is still subject to the fundamental time-frequency estimation trade-off which limits the frequency resolution available for a taper of length  $N$  (Jarvis and Mitra, 2001). Following considerations outlined in (Mitra and Pesaran, 1999), we choose the bandwidth  $W$  equal to a small multiple of the best theoretically attainable frequency resolution, the Raleigh frequency  $\frac{f_s}{N}$ . where  $f_s = 1KHz$  is the sampling frequency of the data.

For simultaneously recorded *LFP* traces, we can compute the average *coherency* between two simultaneous signals. To compute the coherency, one starts from the expression for the cross-spectrum:

---

<sup>1</sup>For the implementation of the spectral estimation methods, we used *Matlab* (*MathWorks, Inc.*) libraries freely available at [www.chronux.org](http://www.chronux.org), as well as custom code specifically tailored to our data sets

$$S_{yx}(f) = \frac{1}{K} \sum_{k=1}^K \tilde{y}_k(f) \tilde{x}_k^*(f)$$

and then normalizes by the respective spectra:

$$C_{yx}(f) = \frac{S_{yx}(f)}{\sqrt{S_x(f)S_y(f)}}$$

The absolute value of the coherency, the *coherence*, is a quantity between 0 and 1 that quantifies the amount of variation in one signal that can be explained as a linear filter of the other signal.

### 3.2.2 Methods for the analysis of spike trains

A method of incorporating the burst count as well as spike timing into a decoding scheme that aims to find the amount of information present in a set of neural responses is given by the spike metric approach (Victor, 2005)<sup>2</sup>.

A metric is an abstract notion of distance between points belonging to a set. Our points consist of spike trains  $S_a$ ,  $S_b$  and their collection is the set of all recorded spike trains. A spike metric must satisfy the following properties:

*Symmetry:*  $D(S_a, S_b) = D(S_b, S_a)$

*Triangle inequality:*  $D(S_a, S_c) \leq D(S_a, S_b) + D(S_b, S_c)$

*Non-negativity:*  $D(S_a, S_b) > 0$  unless  $S_a = S_b$ .

We note that the non-negativity condition is relaxed in reference to spike trains, in the sense that distinct spike trains can still have zero distance. Formally, our notion of distance is then transferred to distance between equivalence classes of spike trains.

The spike distance is the sum of a series of costs corresponding to steps by which the first spike train can be transformed into the second spike train - see Figure 3.6. Insertion or deletion of a spike into the spike train always has a cost of one. Changing the timing of a spike by an amount of time  $t$  has an associated cost of  $qt$ , where the parameter  $q$  results in different metrics and is useful in providing a balance between the spike count and the spike timing in a burst: if  $q$  is very small, then the metric is equivalent to a spike count. If  $q = 0$ , spike trains with the same number of spikes belong to the same equivalence class. If  $q$  is large, temporal differences in spikes become dominant.

---

<sup>2</sup>Information theoretic analyses in this study were conducted with the *Spike Train Analysis Toolkit*—a neuroinformatics resource funded by the NIH’s *Human Brain Project*, which is freely available at <http://neuroanalysis.org>.

An efficient dynamic programming algorithm for computing pairwise distances between spike trains was provided by (Victor *et al.*, 2007).

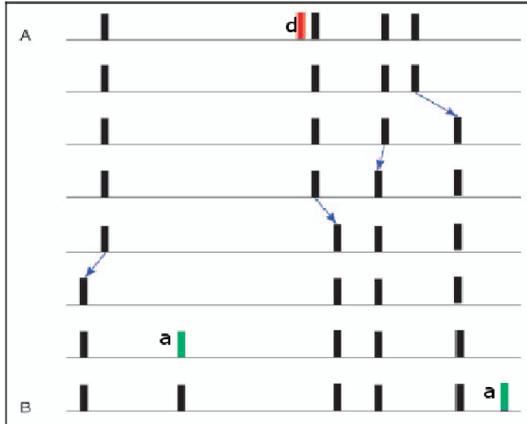


Figure 3.6: Illustration of a series of steps that transforms the spike pattern on top (A) to the pattern on the bottom (B). Red spike in is deleted during the first step, green spike is added during 7th step (row). Blue arrows mark the spikes shifted for a cost of  $qt$ .

Based on these pairwise distances between spike patterns, the neural responses can be categorized in clusters. The level of overlap between these clusters and stimulus categories in a particular experimental paradigm quantifies the information that spike patterns convey about the stimulus.

Our procedure followed the general method outlined in (Victor, 2005). We start with  $N$  spike trains, in response to a set of stimuli. For each spike response, we calculate average distance to stimuli of each category:

$$d(S, s_\gamma) = [\langle (D_q(S, S'))^z \rangle_{s'}]^{1/z}$$

where  $s_\gamma$  is the set of responses to a particular stimulus  $\gamma$  and  $S$  is a particular spike train. Next, each spike train is assigned to the stimulus category corresponding to the smallest  $d$ . The above procedure results in a *confusion matrix*, with the off-diagonal elements corresponding to mislabeled spike trains. This confusion matrix is transformed into a two dimensional frequency histogram, which allows evaluation of the amount of information. Here, by *information* we mean the *decrease in uncertainty* about the stimulus that is achieved by observing the neural response. Intuitively, we can consider the case when there are two stimulus categories, strong and weak, which typically result in two different neural responses, for instance in a high or a low firing rate. However, occasionally the strong stimulus will result in a low rate and the weak stimulus will

generate a high rate. If these events are rare, then we can say that the response conveys a high amount of information about the stimulus. As such, the relationship between the stimulus and neural response can be analyzed similarly with any communication channel (Shannon *et al.*, 1998), (Rieke *et al.*, 1997), (Dayan and Abbott, 2005).

The concept of entropy has been associated with how "interesting" a set of responses is (Dayan and Abbott, 2005). For the random variable  $r$  associated with the response, we define the entropy of its probability distribution  $P(r)$ :

$$H = -\sum_r P[r] \log P[r]$$

The logarithm function is used in the above sum due to several desirable properties. First, it is a decreasing function, which satisfies the intuitive concept that events with low probability have higher measure of surprise. Second, products get transformed into sums by the log function, satisfying the condition that witnessing a response consisting of two independent events should sum their respective entropies (Dayan and Abbott, 2005). The logarithm function satisfies these conditions and therefore it is used in computing the entropy. Low-probability events have a high amount of surprise, but their weight in the sum is low. In contrast, high probability events have high weight but their logarithm is nearly zero.

The entropy of the responses can have two separate sources. The neural responses could have an inherent source of entropy, unrelated to the stimulus, which we call *noise entropy*, which is present even when the stimulus is held constant. For a specific stimulus  $s$ , held constant, the response entropy can be calculated:

$$H_s = -\sum_r P[r|s] \log P[r|s]$$

The total noise entropy is calculated by a weighted averaging the above entropy  $H_s$  over all stimulus conditions:

$$H_{noise} = -\sum_s P[s] H_s = -\sum_{s,r} P[s] P[r|s] \log P[r|s]$$

The *mutual information* between stimuli and responses is the difference between the entropy due to variations in the stimulus and the entropy that is present in the responses regardless of the entropy of the stimulus:

$$I_m = H - H_{noise} = -\sum_r P[r] \log P[r] + \sum_{s,r} P[s]P[r|s] \log P[r|s] = \sum_{s,r} P[r,s] \log \frac{P[r,s]}{P[r]P[s]}$$

where the last equality makes use of the fact that  $P[r,s] = P[s]P[r|s] = P[r]P[s|r]$  by Bayes rule. The last expression reveals that the mutual information is symmetric, in the sense that the information gained about the stimulus by observing the neural responses is equal with the information known about the neural responses if one has access to the stimuli presented to the system. Our discussion above closely follows the outline provided in (Dayan and Abbott, 2005), *chapter 4*.

Often times it is desirable to have a visually intuitive way of displaying the data. A set of pairwise distances can be inspected visually by using a multidimensional scaling procedure such as *Isomap* (Tenenbaum *et al.*, 2000). This procedure finds the best two dimensional representation of the data that preserves as much as possible of the pairwise relationships between the data points. Therefore, even though the data is formally residing in a high dimensional space, in some cases a two dimensional representation is still intuitively helpful, especially in cases in which a small number of factors explain much of the structure of the data. In the case of pairwise distances between spike trains, *Isomap* is especially convenient since it accepts as input pairwise distances.

### 3.3 Results

#### 3.3.1 Different neural representations of sound stimuli and somatosensory inputs

The *LFP* traces recorded from the primary auditory cortex reveal a sharp, precisely timed deflection immediately following the presentation of a sound stimulus (see Figure 3.7 A). This response is characteristic of recordings from the auditory cortex. We also observed a precisely timed series of spikes immediately following each tone presentation (see Figure 3.7 B).

Visual inspection of the spike rasters in Figure 3.7 B suggested that there was an increase in spiking during the time of the lick (compare blue and green spike rasters *versus* red and cyan). In order to quantify this observation, we computed the average firing rate for each animal, pooling all the spikes recorded across different electrodes. The number of well-isolated electrodes from each animal varied between 8 and 14 across the 4 animals used in the experiments. All animals/experiments revealed similar results;

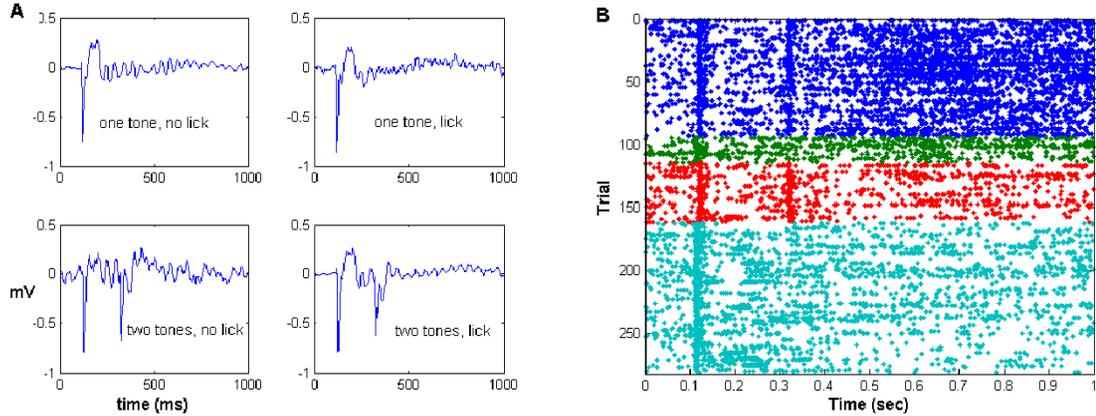


Figure 3.7: A. LFP responses to either one (top) or two tones (bottom) Trials where the animal licked are represented on the right. B. Spike rasters collected after either one or two tones. Blue: two tones, lick. Green: One tone, lick. Red: two tones, no lick. Cyan: one tone, no lick.

we present a typical example in Figure 3.8.

The increase in spiking corresponding to the licking action is somewhat surprising considering that our recordings come from the primary auditory cortex. While our experimental paradigm does not permit us to assess with absolute certainty whether the licking itself generated some low level sound, audibly undetectable to the experimenter (*Paulo Rodrigues, personal communication*), we believe that a more likely explanation is a direct modulation of the firing rate by the motor action. This is consistent with recent work in a variety of mammalian systems which describe modulation of the firing in the auditory cortex by a variety of non-auditory stimuli, including motor action (Kayser *et al.*, 2008), (Bizley *et al.*, 2007), (Noesselt *et al.*, 2007), (Martuzzi *et al.*, 2007).

Spectral analysis of the oscillations in the *LFP* has revealed an increase in the power in a broad range of frequencies (*Paulo Rodrigues, personal communication*), in response to both auditory input and motor action. However, while there was a similar increase in power for these qualitatively different inputs (auditory tone and motor action), there was a notable difference in the coherence computed between pairs of recording electrodes.

For each trial, we computed the spectrogram of the recorded *LFP*. Each spectrogram resulted in a time-frequency plot represented as a matrix. For each element in these matrices, we performed an *ROC* analysis contrasting the trials where the animal was presented with two tones and licked, *versus* the trials where the animal was presented with one tone and did not lick. In Figure 3.9 A we illustrate the results of this comparison. In order to have sufficient time resolution to distinguish between the response to

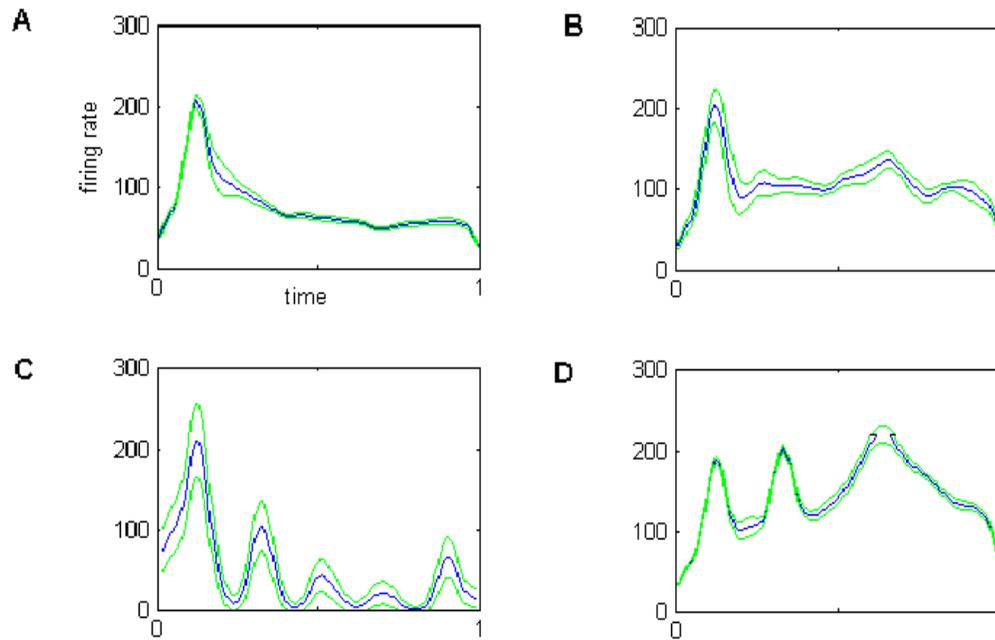


Figure 3.8: Firing rate (spikes/sec) versus time (sec). A: one tone, no lick. B: one tone, with lick. C: two tone, no lick. C: two tones, lick. Firing rates were computed using a Gaussian averaging window with width of 20 ms. Blue trace represents mean firing rate, green traces show the standard deviation.

the second tone (0.3 to 0.4 s into the trial) from the response to the lick (approx. 0.5 s into the trial), we had to use a short time window in computing the spectrogram (100 ms), which resulted in the loss of frequency resolution. Therefore, we emphasize that the plots in Figure 3.9 are not intended to identify specific frequencies and do not allow the interpretation of the apparent break in the frequency plot at around 40 Hz. We will return to the issues related to time-frequency resolution in *Chapter 5*.

The colors in the plots represent  $AUC$  values, with dark red ( $AUC = 1$ ) illustrating the fact that the power in trials with two tones and lick was greater than the power in trials with one tone and no lick. Cyan color corresponds to  $AUC = 0.5$ , which means that there is no significant difference in power between the two conditions. For the time between 0.1 and 0.2 s, corresponding to the first tone, there is no difference between the two conditions, as expected since the first tone was present in both conditions. For the time between 0.3 and 0.4 s into the trial, the presence of the second tone is apparent

by the increase in power. After a brief pause, the response to the lick is signified by the increase in power corresponding to the time when the animal licks the water spout.

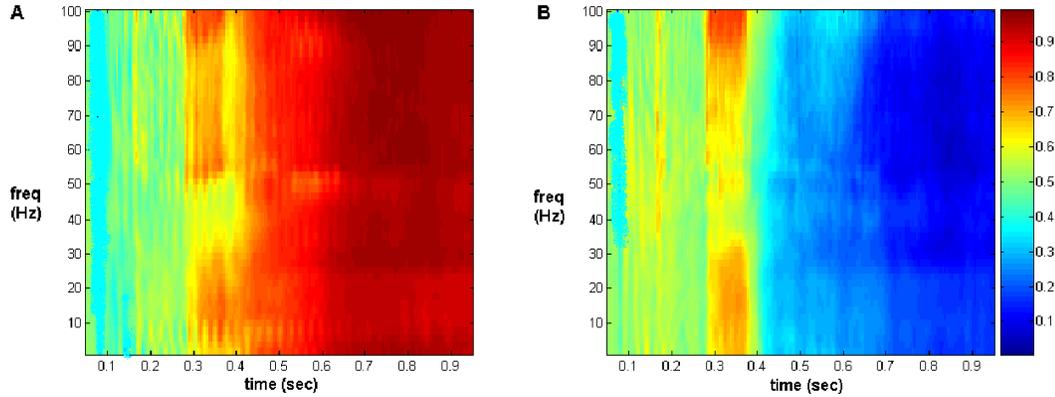


Figure 3.9: A: *ROC* analysis of the power spectrogram contrasting one tone, no lick with the two tones, lick trials. B: *ROC* analysis of the coherograms contrasting the same data as in A.

We performed the same *ROC* analysis on the coherograms - see Figure 3.9 B. For the response to the tone, we see a similar increase in the coherence, corresponding to the presence of the second tone between 0.3 and 0.4 s into the trial. However, during the lick, we see a marked decrease in the coherence between the *LFP* signals, represented by the blue color ( $AUC < 0.5$ ), between 0.5 and 1 s into the trial.

The underlying causes and the functional roles of this marked difference between power and coherence are unclear at this point. While a number of recent studies have detected neural correlates of motor activity in the auditory cortex, (Kayser *et al.*, 2008), (Bizley *et al.*, 2007), (Noesselt *et al.*, 2007), (Martuzzi *et al.*, 2007), a precise quantification of the neural encoding of non-auditory signals by the auditory cortex is not available as of yet. One immediate question raised by our observation concerns a contrast between representation of motor signals in the auditory cortex, compared with the somatosensory cortex. As noted above, the primary auditory cortex responds with increased activity to both auditory stimulation and also motor corollary signals. However, our results suggest that coherent neural activity is specific to auditory inputs and it actually decreases during motor activity.

Spiking activity was also correlated with both auditory stimulation and motor action (See figure 3.7 B). The tones are represented by a well-timed increase in spiking, immediately after the tone. In contrast, representation of motor action resulted in a more persistent but less precise increase in the firing rate. In order to quantify the relative importance of the spike count *versus* spike timing in the neural representation

of these different kind of stimuli, we applied a spike metric analysis to spike trains (see Figure 3.10). The spikes were collected from a time window covering  $50$  ms, beginning immediately before the time of the second tone, if second tone was present.

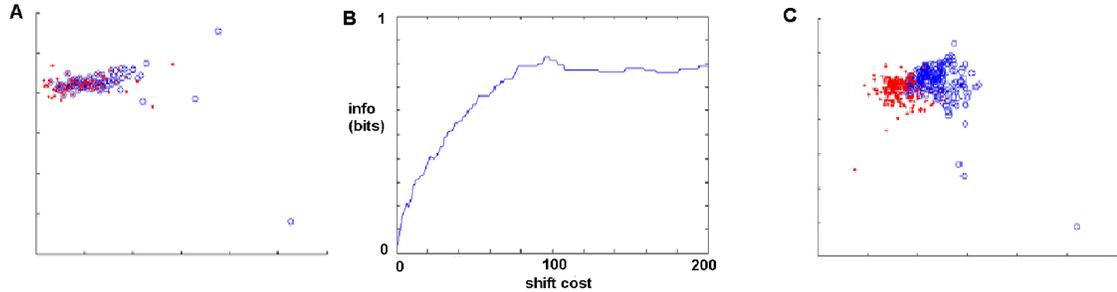


Figure 3.10: A: *Isomap* projection of the tone (red) and no tone spike trains for  $q = 0$ . The tone (red) and no tone (blue) are not well discriminated based on spike count alone. Axes represent two dimensional projections of the data (arbitrary units on axes not represented). B: Discrimination information versus shift cost  $q$ . Larger  $q$ 's attach more weight to spike timing and improve discrimination. C: *Isomap* projection for  $q = 100$ . The discrimination is much improved (compare to A).

The spike metric analysis produces a set of pairwise distances between spike trains. In order to visualize the data, we used a multidimensional scaling procedure (*Isomap*), which can operate on pairwise distances (Tenenbaum *et al.*, 2000). We collected spike trains from a window which spans 300 to 350 ms into the trial. In the trials where a second tone was presented at  $t = 300$  ms, a well timed series of spikes were generated in response. In Figure 3.10 A, the relatively low amount of information contained in the spike counts is illustrated by the overlapping clusters of spike patterns. As the temporal sensitivity of the spike metric is increased by larger values for the  $q$  parameter, the amount of information increases accordingly (Figure 3.10 B), resulting in clusters of spike patterns which are well separated (Figure 3.10 C). In computing the information illustrated in Figure 3.10 B, we followed the outline detailed in the *Methods* section of this chapter: we computed the pairwise distances between spike and then we computed the confusion matrix. The confusion matrix was used to compute a two dimensional histogram, followed by the computation of the mutual information.

The benefits of using the spike metric approach become now apparent. In Figure 3.7 B, the spike rasters suggest that a well-timed increase in the firing rate represents the neural response to the auditory tone. When a firing rate is computed in Figure 3.8, one has to choose a window for averaging over the spikes represented as  $\delta$  functions. The choice of the window width is often arbitrary. By inspecting Figure 3.10 B, a systematic way of choosing the best window width becomes apparent. For  $q = 100$ , the

information between stimuli and responses is maximal. This value of the parameter  $q$  is equivalent to considering two spikes separated by a distance of less than  $\frac{2}{q} = 20$  ms as being equivalent. In conclusion, a window width of 20 ms should be chosen to provide a smooth approximation to the instantaneous firing rate. For shorter window widths, one runs the risk of not including all the spikes produced in response to the tone. For larger window widths, there is increased risk of including spikes removed temporally from the brief tone, in this case possibly spikes generated by the motor action. For instance, a large window would result in a less distinct peak in the firing rate corresponding to the second tone (see Figure 3.8), which would merge with the peak in firing rate caused later by the motor action.

The average *LFP* traces plotted in Figure 3.7 A suggested a smaller response to the second of the tones. Initially, this was quantified simply as the difference between the lowest point and the next highest point in a window of 20 ms immediately following each tone. In all the channels recorded the difference was consistent; we present an example of one channel in Figure 3.11 A. An equivalent result was quantified in the spike counts collected in a window of 50 ms immediately following each tone. The number of spikes in response to the second tone was smaller than the number of spikes in response to the first tone. Individual recorded channels had an insufficient spike count to quantify this result, but pooled spikes from all the channels recorded in an animal/experiment produced a statistically significant result. Figure 3.11 B presents the pooled data from all channels/animals/experiments.

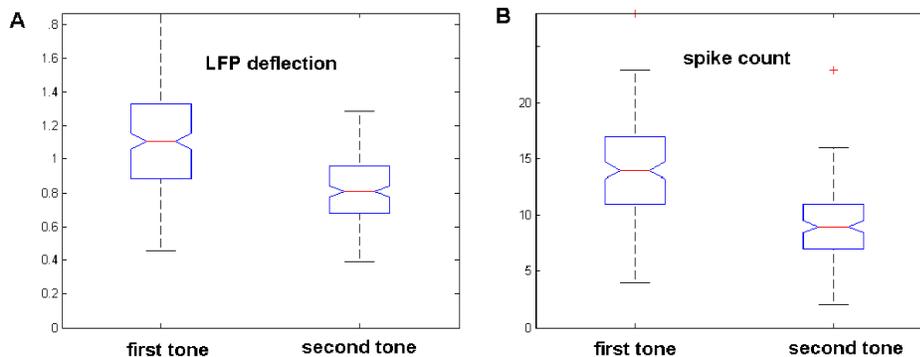


Figure 3.11: Boxplots quantifying differences between response to first and second tone. The boxes have lines at the lower quartile, median, and upper quartile values. The whiskers extend an additional 1.5 interquartile range to show the extent of the data. Non-overlapping notches indicate that the medians of the two groups differ at the 5% significance level. A: example of a recorded *LFP* channel. B: aggregate spike counts from all recorded animals and channels.

We also asked whether the attenuation in the response to the second tone is depen-

dent on the time passed from the first tone. We grouped together trials with 100, 200 and 400 ms intertone intervals, and we used *LDA* analysis to identify the direction most discriminative between the *LFP* traces recorded after the first and second tone. Consistent with previous studies in several mammalian preparations (Hoehnerman and Gilat, 1981), (Calford and Semple, 1995), (Brosch and Schreiner, 1997), (Reale and Brugge, 2000), (Fitzpatrick *et al.*, 1999), (Werner-Reiss *et al.*, 2006) we found that the amount of the attenuation decreased for longer intertone intervals, as shown in Figure 3.12.

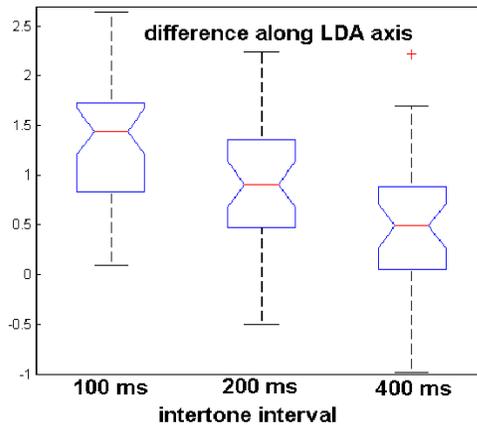


Figure 3.12: Amount of response attenuation as a function of the intertone interval. Longer intertone intervals lead to a smaller attenuation of the response to second tone.

During the recording session, we also observed a systematic decrease in the response to the first tone (Figure 3.13). Since this change was quantified as the projection on the *LDA* axis that discriminates between the response to the first and second tones, and the second tone response was generally smaller, we consider the change in response to be a decrease.

A similar trend was not detected in the responses to the second tone, which generally did not display a decreasing trend when the first tone did (14 out of 40 analyzed channels). This phenomenon has similarities to the general adaptation of neural responses to repeated stimuli. However, in our preparation this adaptation was selectively expressed for just the first tone in a succession of two tones.

The amount of response attenuation appears to be correlated with the amount of high-gamma power (80 – 130 Hz) measured in the interval between the two tones (see Figure 3.14 A).

The relationship between the lower frequencies and the amount of response attenuation was opposite that of the high gamma (see Figure 3.14 B). However, the interpreta-

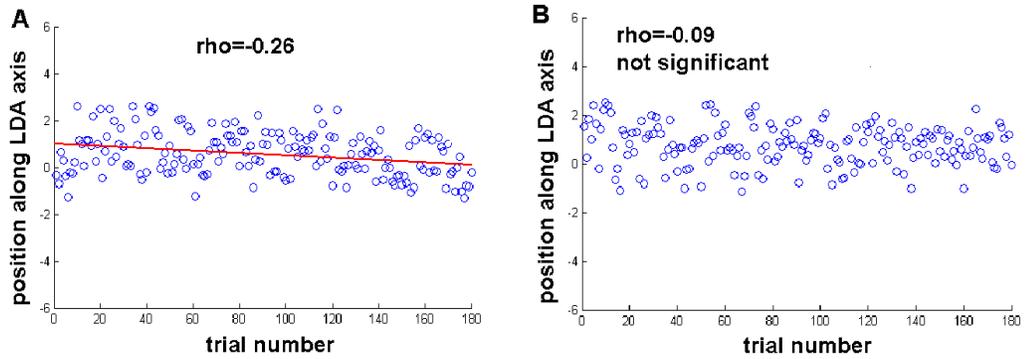


Figure 3.13: A: Decreasing response to the first tone over the course of the experiment, as measured by the projection on the LDA axis. B: Response to the second tone, as in A. No significant trend detected.

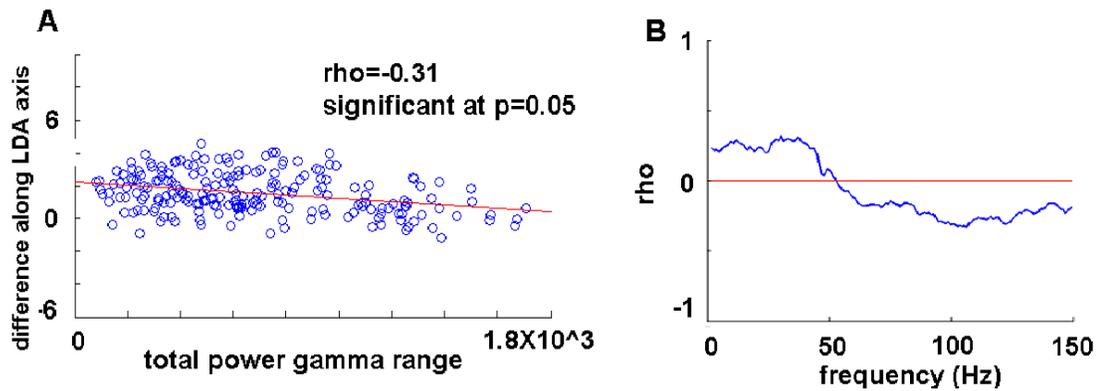


Figure 3.14: A: Correlation between total spectral power in 80-130 Hz range and the amount of response attenuation. B: Correlation coefficient between level response attenuation and power at each frequency.

tion of the lower frequencies power and the response attenuation is complicated by the autocorrelation present in the *LFP* traces. As mentioned above, the *LFP* traces used in the computation of the *LDA* direction end about 50 ms after the tone and do not overlap with the traces used for computation of the spectral power. However, the increase in lower frequencies which is directly attributable to the tone will last longer than the 50 ms. A large response to the first tone will be quantified as a large response attenuation (assuming the second tone response does not increase as well). The same large response to the first tone will also lead to a larger low-frequency spectral power, which we believe is leading to the positive correlation measured between the low frequency power and the amount of response attenuation. Under this assumption, the negative correlation

quantified in Figure 3.14 is likely underestimated due to frequency spillover effects from lower frequencies, and the correlation between frequencies immediately under 80 Hz are skewed by the strong positive correlation between low frequencies and the amount of response attenuation. While theoretically the Fourier decomposition uses a set of linearly independent basis vectors, corresponding to the different frequencies, spectral estimation methods, including the multitaper method, are affected by the time-frequency trade-off which limits the resolution of the spectral parameters and cannot eliminate frequency spillover effects (Mitra and Pesaran, 1999).

Regardless of the issues regarding the lower frequencies outlined above, the relationship between the high frequencies and the difference between the first and second tone, computed along the *LDA* axis, is clear: more high-frequency spectral power in the interval between the tones leads to a lower level of response attenuation.

We have also investigated the possible involvement of response attenuation to the subsequent behavioral choice of the animal. In a subset of our recordings, we observed a significant difference in the amount of response attenuation between the correct and incorrect trials.

In the *LFP*'s, more response attenuation was observed in the incorrect trials (Figure 3.15 A). In the spiking activity, the spike count in response to the second tone, but not that in response to the first tone, was significantly different in the two conditions (Figure 3.15 B). For the *LFP* data, the projection on the *LDA* axis of the first tone or second tone traces, individually, did not reveal a significant effect, which was only visible in analyzing the difference, for each trial, between the response to the first and second tone. The reasons for this discrepancy between the *LFP* and spike data remain unclear; we hypothesize that trial variability affects the overall amplitude of the *LFP* traces but affects less their difference.

The tone presented to the animal resulted in an increase in the *LFP* oscillatory amplitude across a broad range of frequencies. We compared the ratio between the spectral power of the *LFP* traces in a range of 150 ms immediately before the first tone, and traces starting 50 ms after the tone. The amount of the increase varied as a function of the frequency (see Figure 3.16). We quantified this difference in several different ways. In Figure 3.16 B, we plot the *p-values* of a *ranksum* test of equality of the medians. In 3.16 C, we plot the *AUC* determined by discriminating between the correct and incorrect trials. Finally, in D we plot the maximum information content along the *ROC* curve. These measures illustrate the fact that, for frequencies in the range of approximately 80 – 130 Hz, the *LFP* oscillations in the interval between the tones have more power in the correct trials. While the power ratio was highest for frequencies below 30 Hz,

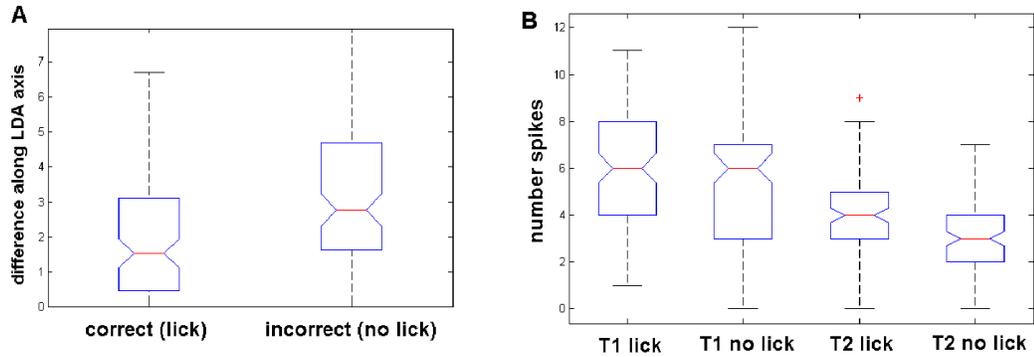


Figure 3.15: A: Response attenuation in correct versus incorrect trials in LFP's. B: Spike count differences between correct and incorrect trials.

a significant difference between the correct and incorrect trials was observed only for 80 – 130 Hz.

To our knowledge, this is the first example of a change in behavioral response being correlated to attributes of the sensory response such as the level of response attenuation. Previous work in awake behaving monkeys did not detect a change in the response attenuation based on the attentional state of the animal (Werner-Reiss *et al.*, 2006). While this result is certainly intriguing, we remain cautious since a significant difference was observed consistently in just one experiment (6 out of 8 recorded *LFP* channels), and in the spiking from the pooled channels in the same experiment. An additional *LFP* channel in a second experiment showed this effect. Additional experiments, preferably using carefully placed electrodes over a specific region/layer of the auditory cortex, would be required. In particular, the frequency tuning of the individual neurons would be a possible determining factor.

While the amount of response attenuation has not been previously correlated with behavioral performance, gamma-band oscillations have been proposed as mechanisms for a varied collection of sensory or cognitive functions (Steriade *et al.*, 1996), (Llinas and Ribary, 1993), (Tiitinen *et al.*, 1993), (Pantev, 1995). Several studies have associated increased gamma band oscillatory power with the formation of percepts and memory (Singer, 1990), (Engel *et al.*, 1991), (Pantev *et al.*, 1995). Linguistic processing has also been associated with gamma band oscillations in recorded *EEG* potentials (Pulvermüller *et al.*, 1995).

Our analysis has shown a correlation between the gamma-band power in *LFP* recordings between the *two tones* and the amount of response attenuation. This effect was detected in nearly all the channels/experiments analyzed. In a subset of our recordings,

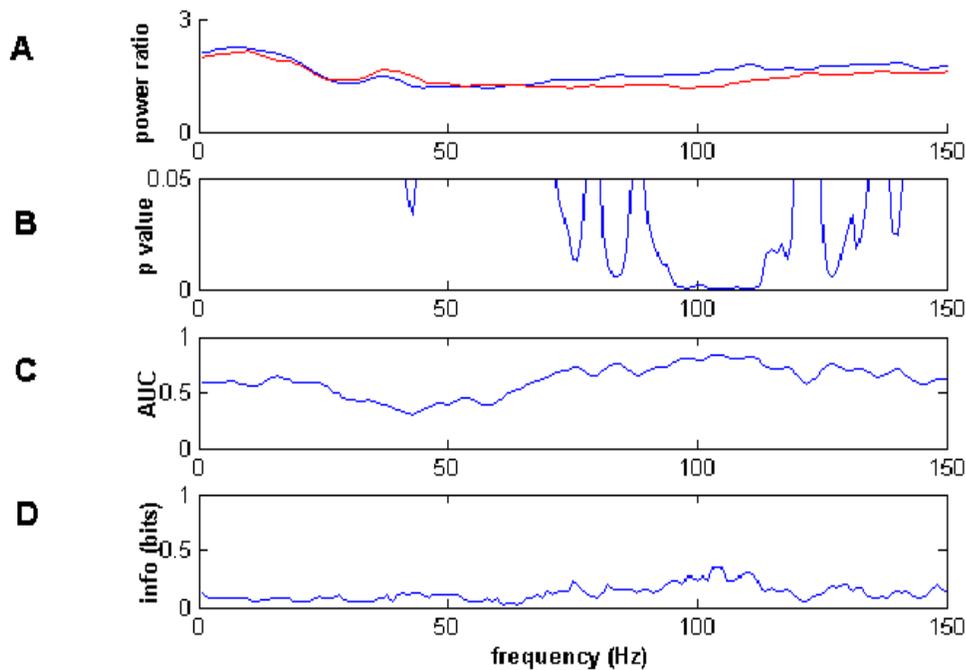


Figure 3.16: Relative increase in intertone *LFP* power relative to pre-tone baseline. A: average ratio of power in correct (blue) and incorrect (red) trials. B: p-values in comparing the values in A. C: *AUC* based on the quantities in A. D: maximum information content based on *ROC*'s in C.

we have also been able to predict the behavioral performance of the animal based on the amount of gamma band power in anticipation of the second tone. A possible interpretation for this effect is an increase in attention that is associated with an increase in gamma band activity. Attention has been related to modulations in the gamma band activity in a number of different studies: (Vidal *et al.*, 2006), (Bauer *et al.*, 2006), (Gruber *et al.*, 1999), (Tallon-Baudry *et al.*, 2005), (Doesburg *et al.*, 2008). Our experimental results are consistent with an attentional effect that is detected as an increase in the gamma band power. Attention could increase the saliency of the second tone, therefore decreasing the amount of response attenuation.

The analysis performed on the recordings from the primary auditory cortex illustrate the benefits of our comprehensive approach of data analysis. We started by computing spectrograms and coherograms for all the *LFP* recordings and comparing the average spectrograms across conditions: correct *versus* incorrect, one tone *versus* two tones, lick *versus* no lick. The comparisons were accomplished by a *ROC* analysis performed individually on each spectrogram coefficient. Significant differences between these coeffi-

cients revealed novel and biologically relevant aspects of neural activity, such as response attenuation and its relationship with future behavior, and also the decrease in coherence in the auditory cortex during motor action. These effects were not stated as hypothesis prior to the data analysis, either by the experimental investigator, or by the author. In addition to uncovering these novel effects, our data analysis procedure also confirmed a hypothesis generated by the experimentalist (*Paulo Rodrigues, personal communication*), namely that increased gamma range oscillatory activity between the tones leads to an increase in behavioral performance.

Our second goal for this project was the identification and evaluation of appropriate methods for analysis of the data set. Examining the regions in the spectrograms with significant differences between the experimental conditions led to the refinement of the analysis by the use of more sensitive methods. The differences between the response to the first and second tones were quantified using the *LDA* approach, which is more appropriate due to the fact that the short duration of the tone response (at most 60 ms) precludes the use of Fourier analysis, inappropriate for such short duration signals. In contrast, the estimation of the gamma frequency power can use the whole interval between the tones (200 ms), which is sufficiently long for an estimate of the frequencies in the gamma range.

## Chapter 4

# Topographical representation of complex vocalizations in the telencephalon of zebra finch

Songbirds are vocal learners that generate songs, patterns of sound with a highly complex spectral-temporal pattern. Conspecific songs are very important to the birds, and are used to attract mates or mark territory. Several distinct areas in the zebra finch brain respond to song exposure by an elevated level of expression of activity-dependent genes (Mello *et al.*, 1992), (Velho *et al.*, 2005). The overall level of response, as measured by the total number of cells expressing activity dependent genes, is dependent on the stimulus type. For example, conspecific song activates a larger number of cells than white noise. An immediate question regards the kind of information contained in the exact topographical distribution of activated cells in a particular brain region. This kind of analysis requires an automated method of detecting activated cells, a way to describe their topographical distribution, and also quantitative ways of distinguishing between different patterns of spatial distribution. In addition, the stimulus space of conspecific songs can be manipulated in ways designed to reveal the most salient features in the acoustical components of song. As in our analysis of synaptic integration and the system level processing, we observe a marked sensitivity to the fine temporal features of the input, in this case the temporal aspects of the song syllables.

In a similar manner to the previous chapters, we explore, in a comprehensive way, the space of inputs and its relationship with the neural responses recorded during the experiment. In this series of experiments, the behavioral saliency of different sensory stimuli is inferred indirectly, by comparing with the neural response of conspecific songs, a highly salient stimulus class.

## 4.1 Temporal structure of song reflected in spatial patterns of cell activation

Neural recordings from single neurons or small groups of neurons clearly show the temporal sensitivity of neural sensory systems, as we detailed in the above chapters. Current neural recording techniques suffer from a lack of spatial resolution, since even multi-electrode recordings can access at most hundreds of units at a time (Kralik *et al.*, 2001), (Nicolelis and Ribeiro, 2002). A different method of accessing the spatial distribution of neural representations is offered by the mapping of *immediate-early-genes (IEGs)* in response to appropriate stimulation (Pinaud and Tremere, 2006), (Terleph *et al.*, 2006), (Mello, 2002). Using exclusively mapping, information about the specific temporal patterns of the neural response (firing patterns of the neurons involved) is not available. However, we can still examine the saliency of the fine temporal structure of the songs. We accomplish this by modifying the songs in specific ways that preserve different amounts of the fine temporal and spectral structure, and then we quantify the changes in the spatial distribution of activated cells.

Complex acoustical signals constitute an important communication mechanism for a variety of species. Mapping the neural representation of complex sounds is essential for the understanding the mechanisms of sensory/perceptual processing. While significant progress has been made in understanding the neural basis of sound localization in owls and echolocation in bats, the study of complex sounds in the context of vocal communication is in a more incipient phase (Ribeiro *et al.*, 1998), (Tchernichovski *et al.*, 2001).

Songbirds represent one of the most accessible animal models for the study of neural representation of auditory communication signals. The brain regions involved in song processing and discrimination have been identified and their electrophysiological responses are relatively well characterized. Exposure to conspecific songs, a naturally learned behavior commonly used for communication, induces the expression of a number of activity dependent genes in auditory processing brain areas (Mello *et al.*, 1992), (Velho *et al.*, 2005). The identification of the brain regions involved in song can be helped by the mapping of the neuronal cells expressing genes such as *ZENK*, which encodes a transcription factor associated with neuronal depolarization and therefore has been previously used as a marker for neuronal activation (Mello, 2002). While in some areas of the songbird brain there is evidence for electrophysiological activity which does not induce expression of activity-dependent genes (Velho *et al.*, 2007), as of yet we are unaware of upregulation of these genes in the absence of neuronal activation.

A combination of electrophysiological, pharmacological, lesions and gene expression evidence suggests that the caudomedial nidopallium (*NCM*) is essential for the processing of song. Among the functions associated with this structure are perceptual processing, discrimination and the formation and storage of song-related auditory memories (Mello *et al.*, 1992), (Gobes and Bolhuis, 2007). In canaries, expression of the *ZENK* gene has been shown to have a spatial structure dependent on the acoustic features of the songs used as stimuli, with different song syllables generating expression in different *NCM* spatial subdomains (Ribeiro *et al.*, 1998).

The spatial distribution of *IEG* expressing cells has been leveraged to study a variety of systems (Pinaud and Tremere, 2006). The study of the auditory pathways of songbirds has made extensive use of these mapping techniques (Ribeiro *et al.*, 1998), (Terleph *et al.*, 2006). The high level of interest in the avian song system is due to several factors. First, complex vocalizations learned from a tutor is a trait found in just a few species besides songbirds: humans, cetaceans and bats (Brenowitz *et al.*, 1997), (Nottebohm, 1991). Therefore the study of the neural representation and memorization of complex auditory inputs can give us insight into the perceptual aspects of vocal communication, perhaps including in humans (Doupe and Kuhl, 1999). Second, the songbird neural circuits, in contrast with the other vocal learners, are relatively well characterized (Brenowitz *et al.*, 1997).

The spatial pattern of activated cells is stimulus specific, in the sense that, after a certain threshold, the amount of stimulation (number of repetitions) has no effect on the number of activated cells, while the type of stimulation (for instance white noise versus conspecific songs) is a determining factor for the spatial distribution of activated cells (Mello *et al.*, 1992). Additionally, the induction of *IEGs* is rapid and transient, being detected first in the transcription sites within the nucleus, and then becoming detectable in to the cytoplasm in the form of *mRNA* products. This rapid transient expression is essential for the purposes of comparing responses to different stimuli in the same animal. Differentiating between nuclear and cytoplasmic expression is equivalent to differentiating between stimuli presented at different times. Therefore, an automated procedure capable of detecting and differentiating nuclear and cytoplasmic cell labeling is of high interest, and creating such an algorithm constitutes an important component of the present work.

Once the exact positions of the activated cells are determined, either manually or using our automated procedure, the next step is the analysis of the resulting spatial patterns. Principal component analysis (*PCA*) has been previously used to quantify the similarity between the spatial representations of songs in the avian brain (Ribeiro *et al.*,

1998). Due to the small number of images available, the use of *PCA* is not appropriate for our data set. Instead, we compute pairwise distances between images consisting of two dimensional patterns of activated cells. Based on these pairwise distances, we generate a confusion matrix, in a manner similar to the classification of the spike trains generated in *Chapter 3*.

Electrophysiological experiments have revealed that the brain areas involved in song processing contain neurons responding to very complex auditory features. These “song-selective” neurons respond more strongly to conspecific songs or to the bird’s own song than to other sounds (Margoliash, 1983), (Margoliash and Fortune, 1992). In the context of electrophysiological recordings from the song areas in the bird brain, specific methods of song manipulation have been developed (Theunissen and Doupe, 1998), These methods allow for selective degradation of specific spectral and temporal parameters of the songs and they also have been used in functional *MRI* studies of the bird brain (Boumans *et al.*, 2007). We employ these techniques of song manipulation and selective degradation of song features in the context of mapping activated cells. Next, we evaluate the resulting changes in the spatial patterns of *IEG* expression compared with the patterns of intact songs. In turn, the results of these studies could be used to guide electrophysiological experiments, for instance the optimal placing of recording electrodes within a brain structure.

## 4.2 Methods

### 4.2.1 Automatic detection of nuclear and cytoplasmic *IEG* expression

The neuronal cells displaying *IEG* expression can be mapped using *fluorescent in-situ hybridization (FISH)*. This procedure yields a series of pictures consisting of all the nuclei in a brain slice (blue channel in Figure 4.1 bellow) and also the fluorescent markers for *IEG* expression (red channel).

There are two distinct patterns of *ZENK* expression. Nuclear expression consists typically of small, intense fluorescent markers contained inside the nucleus (Figure 4.2 A). Cytoplasmic expression consists of a "halo" nearly surrounding the cell nucleus (Figure 4.2 A).

The goal of an automated procedure for quantifying the gene expression consist of detecting each cell nuclei and also detecting the cells expressing nuclear or cytoplasmic *ZENK*. The challenges facing this task are immediately apparent. Figure 4.1 reveals that

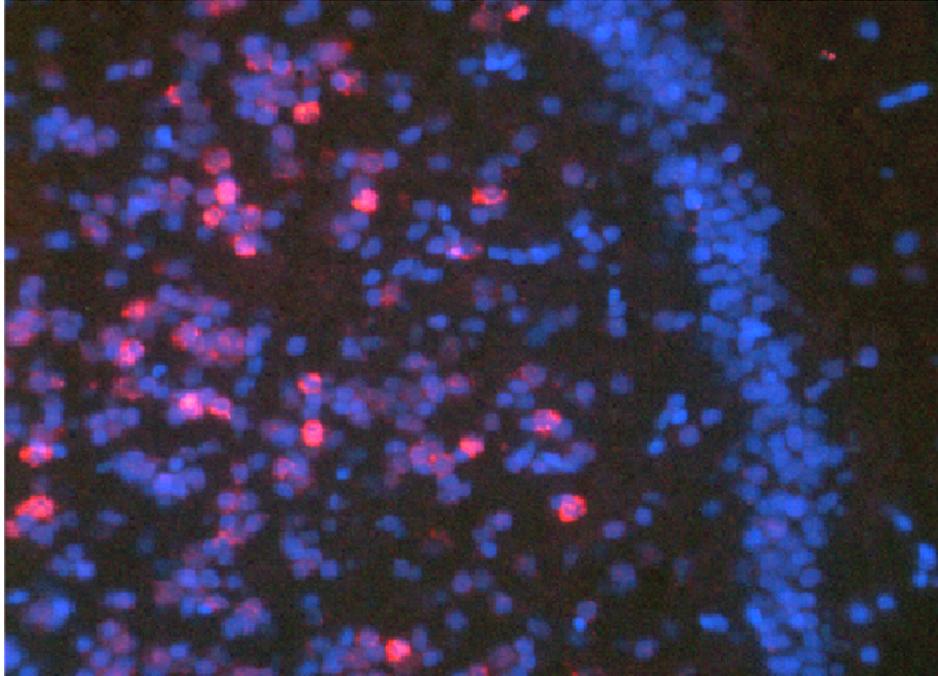


Figure 4.1: Example of a *ZENK* labeled brain slice, photographed under 40X magnification under an optical microscope.

the nuclear shapes are often diffuse and occlude each other. The close proximity of most of the cells implies that the cytoplasmic expression of one cell can be easily mislabeled as nuclear expression for a nearby cell.

Previous attempts to automatically detect and classify labeled cells have been based on confocal image stacks (Lin *et al.*, 2003), (Lin *et al.*, 2007). However, the images used for this work consist of two dimensional images as collected from an optical microscope. To our knowledge, an automated procedure for segmenting and labeling two dimensional images, capable of overcoming the challenges listed above, is not available.

### **Preprocessing steps**

Accurate and robust detection of the cell nuclei is the essential first step toward a quantification of the *ZENK* patterns. Our approach was to use the edge information to detect circular or nearly circular shapes in the blue channel. The initial step was a series of mathematical morphological operations applied to the graylevel images in the blue channel. The goal of these operations is to remove small, low intensity features of the image, which are not cell nuclei, and also enhance the sharpness of the boundaries of the remaining shapes.

In digital image processing, morphological operations consist of functions applied

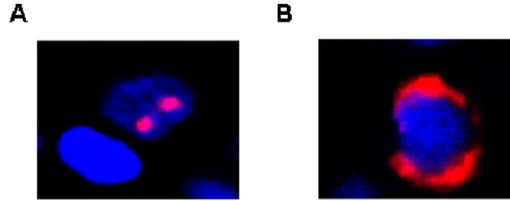


Figure 4.2: A: Typical nuclear *ZENK* expressing nuclei. The red spots represent different loci of transcription. B: Typical cytoplasmic expression of *ZENK*. The fluorescent labelling consists of a concentration of color nearly surrounding the nuclei.

to a pair of inputs. The first input is the original image, while the second input is a *structuring element*. For the purposes of the present work, the structuring element used is a disk with the radius of 3 pixels. A typical nucleus in the analyzed pictures has a radius of 10 pixels. The structuring element is translated over the whole range of the image, and our morphological operations are applied at each point.

Morphological operations are defined on sets of pixels. They are based on the fundamental concepts of Minkowski addition and subtraction (Gonzalez and Woods, 2007). For two sets  $A$  and  $B$ , we define Minkowski addition and Minkowski subtraction, respectively:

$$A \oplus B = \bigcup_{\beta \in B} (A + \beta)$$

$$A \ominus B = \bigcap_{\beta \in B} (A - \beta)$$

In the context of digital image processing, we use these operations to define operations of *dilation* and *erosion*. We start by considering the case of binary images. Image dilation is equivalent with Minkowski addition, with the structuring element being translated over the whole range of the image and the addition being performed at each pixel. If the structuring element intersects the original image even at just one pixel, then the whole surface of the structuring element is added to the original image.

Image erosion is defined formally as

$$E(I, SE) = I \ominus (-SE) = \bigcap_{\beta \in SE} (I - \beta)$$

where  $I$  is our original image,  $SE$  is the structuring element and  $-SE = \{-\beta | \beta \in SE\}$ . Intuitively, as we translate the structuring element over the whole image, any portion of the image that does not contain the whole of the structuring element is set

to zero.

The above definitions apply to binary images. The concepts can be easily generalized to grayscale images. The essential difference consists of the fact that the input image is almost never zero, but instead takes a whole range of (grayscale) values. The structuring element, for the purposes of our work, has a flat value over its extent. As the structuring element is moved over the input image, the intersection between the image and the currently positioned structuring element is set to the smallest value of the intersection in the case of erosion, respectively the highest value in the case of dilation. The net effect is the fading of small specs of noise, up to the size of the structuring element, the uniformization of the intensity values inside high or low intensity regions, and the smoothing of jagged borders separating regions of overall different intensity.

We applied to our images the operation of *opening*, which is simply an erosion followed by a dilation using the same structuring element. Typical results of opening are the removal of small specks and small invaginations or extensions on the border of the image. This results can be described as a general rounding of the borders, an important step towards our goal of finding and labeling the cell nuclei. The results of this operation are visible in Figure 4.3 A, B.

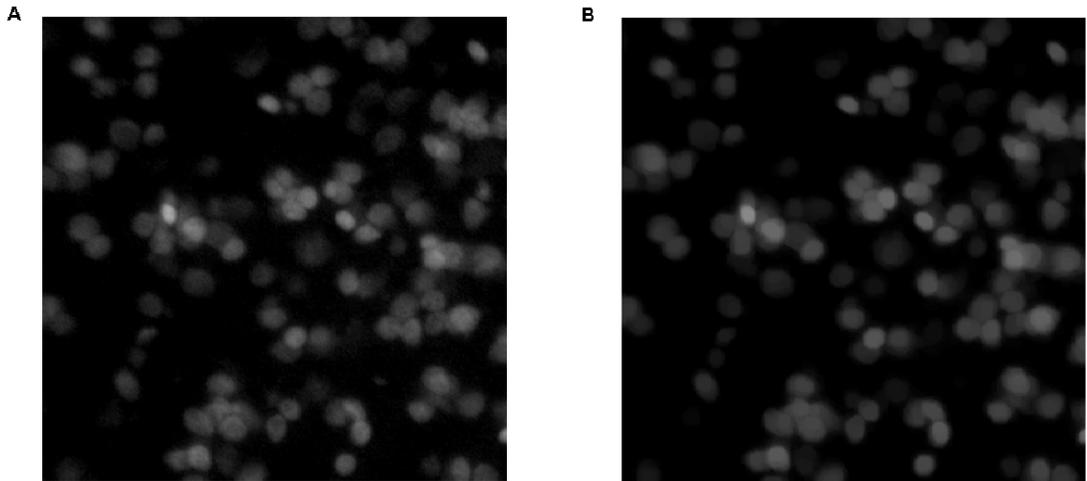


Figure 4.3: Graylevel images from the blue channel. A: the image prior to opening. The edges are blurry and irregular. B: Image after the opening operation. The shapes are more rounded, the graylevel inside the nuclei is more uniform.

## Edge detection

Next, an detection routine was applied to the image with the goal of finding the nuclei borders. We have selected the Canny edge detection procedure (Canny, 1986), due to several characteristics. First, the Canny edge detector makes explicit use of the orientation of the detected edges. The algorithm starts by computing the gradient of the input image, which in the discrete case of image pixels is more conveniently expressed using the central difference operators in the  $x$  and  $y$  directions :

$$\begin{aligned} L_x(x, y) &= -\frac{1}{2}L(x-1, y) + \frac{1}{2}L(x+1, y) \\ L_y(x, y) &= -\frac{1}{2}L(x, y-1) + \frac{1}{2}L(x, y+1) \end{aligned}$$

where  $L(x, y)$  represents the pixel intensity at position  $(x, y)$ . The magnitude and orientation of the presumptive edge are estimated at each pixel:

$$\begin{aligned} |\nabla L| &= \sqrt{L_x^2 + L_y^2} \\ \theta &= \arctan(L_x, L_y) \end{aligned}$$

Using the above information, edge strength values points are evaluated by estimating if they are local maximum in the gradient direction. Based on these criteria, each point in the image gets assigned an edge strength value.

The second important characteristic of the Canny edge detection is the use of *threshold hysteresis*. A stringent threshold is used to detect initial edge points. Starting from these initial points, neighboring points are evaluated using a less stringent threshold. This strategy implements the intuitive concept that edges should be continuous curves through the image. An example of this procedure being applied to our image is presented in Figure 4.4. We note several characteristics of this edge image. First, the edges are incomplete and occasionally small isolated edges appear superfluously. Second, if cell nuclei are overlapping, the edge information is lost over their intersection, but is still present on the non-overlapping contours. The most important characteristic of the edge image, which is pivotal for our subsequent analysis, is the fact that even incomplete borders retain their nearly circular shape.

## Hough transform

In the next step of the analysis we will use the curvature information for the detection of the centers of the cell nuclei in the original image, which coincide to the centers of

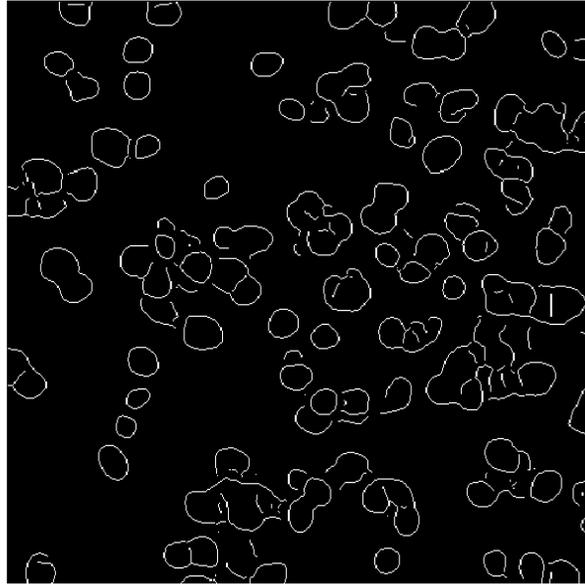


Figure 4.4: Example of the Canny edge detection algorithm applied to the edge image. Overlapping cell nuclei that partially occlude each other still retain their nearly circular shape on the non-overlapping portions of the nuclei.

the fragmented but nearly circular shapes in the edge image. The method was initially used for finding lines in an image (Hough, 1959). It has since been extended for curves and adapted for modern image processing applications (Duda and Hart, 1972) and later generalized to more general shapes (Ballard, 1981).

In practical terms, for a fixed radius  $r$ , the procedure of finding circles in the edge images is straightforward. Each edge point  $(x, y)$  can be parameterized by the position of its center  $(a, b)$  and the radius  $r$ :

$$\begin{aligned} x &= a + r \cdot \cos(\theta) \\ y &= b + r \cdot \sin(\theta), \quad \theta \in [0, 2\pi] \end{aligned}$$

The position of the edge points  $(x, y)$  is known, and the goal of the algorithm is to find the  $(a, b)$  points corresponding to the center of the circle. We observe that if a circle is drawn, centered at each  $(x, y)$  and with radius  $r$ , the intersection of all these circles would intersect at the point  $(a, b)$ , as illustrated in Figure 4.5.

In the *Hough* transform image, the high intensity points represent likely circle centers. The large circles surrounding the centers represent artifacts from computing the *Hough*

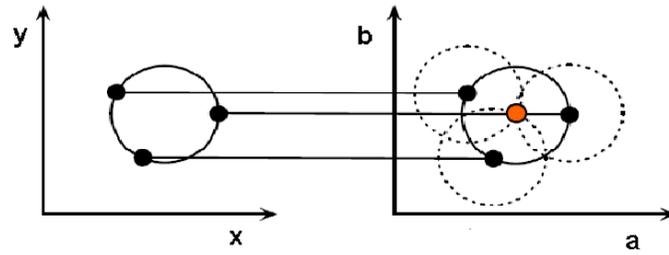


Figure 4.5: Procedure for finding the center of circle starting from  $(x, y)$  coordinates. Circles with radius  $r$ , centered at each  $(x, y)$ , will intersect at the desired center  $(a, b)$ .

transform, and they have twice the radius of the real circle, as illustrated in Figure 4.5. We can also observe some small circles of high intensity. These small circles result from the fact that the *Hough* transform is computed for a set of candidate radii, resulting in a three dimensional matrix known as a *Hough accumulator array* (Gonzalez and Woods, 2007). The candidate radius closest to the real radii of the cell nucleus results in a high intensity point surrounded by a large circle, while the other radii considered will result in a small circle surrounded by a larger circle. The true radius can then be selected based on the fact that it corresponds to relatively high intensity points in *Hough* transform image. The phenomenon is detailed in Figure 4.6.

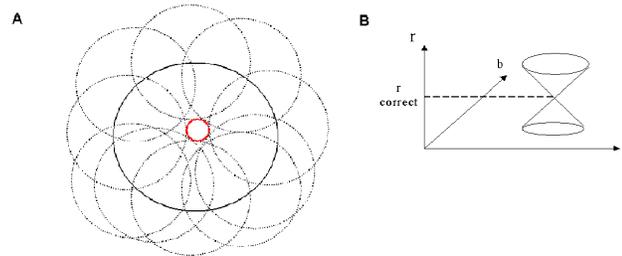


Figure 4.6: A: Illustration of the effect of computing the Hough transform for a radius slightly smaller than the true radius. The circles drawn using the  $(x, y)$  points in the edge image will trace the small red circle, instead of intersecting on a point. B: The Hough transform is computed for a set of different radii  $r$ , resulting in a conical surface spanned by the points of intersection. The apex of this cone correspond to the correct  $r$ .

The arguments outlined above can be illustrated by presenting several sections in the *Hough* accumulator array computed using an example image in our data set. In Figure 4.7, we present sections in the accumulator array corresponding to increasing radii. Values of the radius that are close but not equal to the true radius will result in

small circles, with the intensity of these small circles (arrows in Figure 4.7, A, B, C) being smaller than the intensity of the point corresponding to the true radius (arrow in Figure 4.7 D).

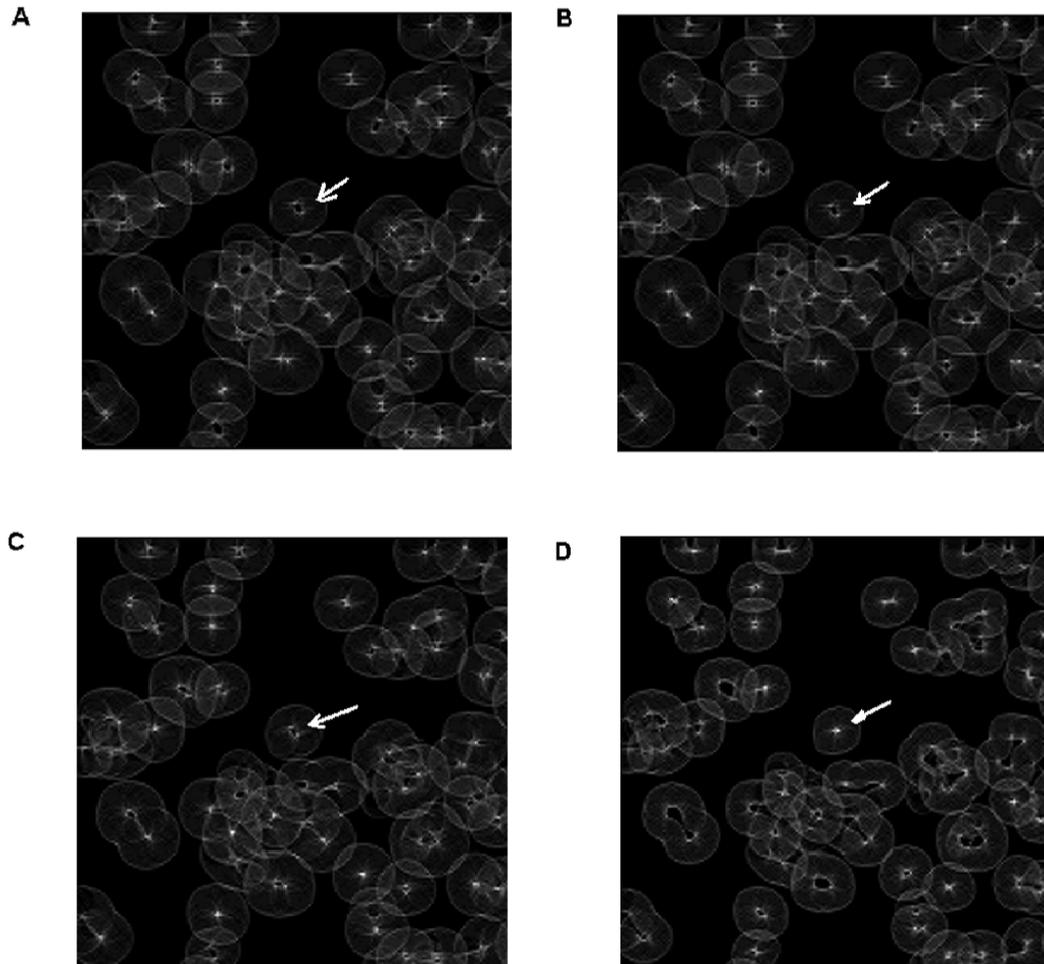


Figure 4.7: Sections through the Hough accumulator array, corresponding to increasing radii. A:  $r = 7$  pixels. B:  $r = 8$  pixels. C:  $r = 9$  pixels. D:  $r = 10$  pixels, closest value to the true radius.

Using the *Hough* transform, partial occlusion of cell nuclei is overcome. The procedure is robust to occlusion since the curvature of the edges is not affected by their intersection. A representative example of the results of the transform applied to our images is presented in Figure 4.8.

At this point, our procedure resulted in a three dimensional *Hough* accumulator array, where local maxima corresponding to the centers of the nuclei. In order to label

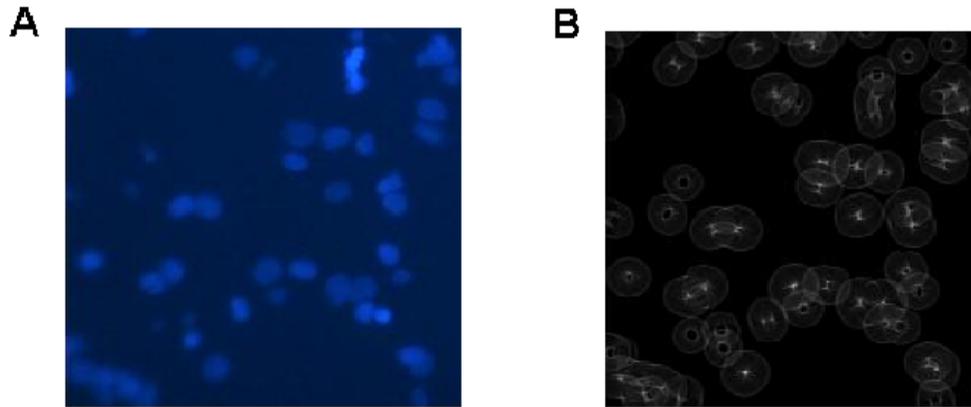


Figure 4.8: A: raw image. B: Hough transform for a radius  $r$ . Each nuclear shape in image above gives rise to a high intensity point in the Hough transform, surround by an artefactual circle of radius double the radius of the cell nucleus.

the cell nuclei, we needed to choose among these peaks the ones that truly correspond to cell nuclei, *versus* spurious peaks. We started with the highest point in the accumulator array and we retained it as a candidate cell nucleus. Next, for a small radius around this point, we set the accumulator array to zero, which implements the requirement that no cell nuclei can be overlapping by more than half their diameter. The accumulator array was set to zero for all the radii considered, since we cannot have overlapping cell nuclei, with different radii, centered on the same point.

### Unsupervised clustering of empty circles

A criterion for terminating the previous procedure, effectively deciding how many nuclei we have in our image, cannot be automatically generated. To overcome this difficulty, we decided to use an unsupervised clustering algorithm to separate true cell nuclei from false ones. The information used for clustering consisted of the local intensity values in the neighborhood of each candidate cell nucleus. The *Hough* accumulator array provide us with the centers and radii of the candidate cell nuclei. We collected the average intensity values inside these circles, since false nuclei are expected to have, on average, smaller intensity values. However, this information is not sufficient to separate the cell nuclei, because the nuclei intensity can vary - see Figure 4.3 A, B. To add additional information, we collected more information from the immediate neighborhood of each candidate cell nucleus. We expanded the cell nucleus and we collected ratio of the intensity values inside the candidate nuclei and the intensity value in the neighborhood

immediately outside the candidate nucleus. Histograms of these quantities are presented in Figure 4.9.

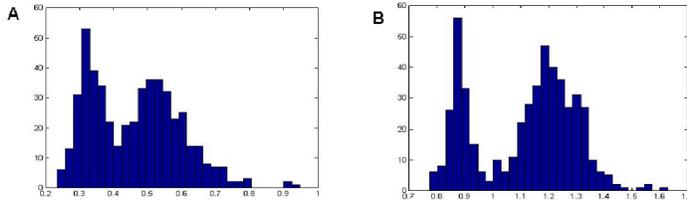


Figure 4.9: Local intensity information associated with candidate cell nuclei. A: Histogram of average intensity values inside detected circles. B: Histogram of ratios between the average intensity inside the candidate cell nucleus and the average intensity in the region immediately adjacent.

The bimodal distributions can be used for an unsupervised clustering algorithm that separates the true nuclei from the peaks in the *Hough* accumulator array not corresponding to true nuclei. Since we have no a priori knowledge about the true number of nuclei in a particular image, we collected twice the number of nuclei in the most dense image in our data set, which results in 500 candidate cell nuclei for each analyzed image.

We evaluated two unsupervised clustering algorithms. The first one is the well-known *k-means* algorithm, which uses an iterative procedure in order to minimize the variance between the points assigned to the same category (Duda *et al.*, 2000). The error function to be minimized can be expressed as:

$$E = \sum_{i=1}^k \sum_{x_j \in S_i} (x_j - \mu_i)^2$$

where there are  $k$  clusters  $S_i$ ,  $i = 1, 2, \dots, k$ , and  $\mu_i$  is the centroid or mean point of all the points  $x_j \in S_i$ .

A second clustering algorithm evaluated was *information clustering algorithm* (Slonim *et al.*, 2005), based on classic rate-distortion theory (Shannon *et al.*, 1998). In this formulation, a limited bandwidth is available for the representation of the data. This bandwidth (the rate), in our case is represented by two clusters and is equivalent to one bit. A distortion function is also defined, measuring how much points assigned to the same cluster differ from each other. Formally, following (Slonim *et al.*, 2005), we define the following quantities:

Average similarity between elements in the same cluster:

$$\langle s \rangle = \sum_{C=1}^{N_C} P(C) s(C)$$

where  $P(C)$  is the probability of a cluster (proportion of total elements), and  $s(C)$  is the average similarity between the elements in the cluster. The input to the algorithm consist of all the pairwise similarity measures between the data points.

The mutual information between the cluster identities and the data points is defined as:

$$I(C; i) = \frac{1}{N} \sum_{i=1}^N \sum_{C=1}^{N_C} P(C|i) \ln \left[ \frac{P(C|i)}{P(C)} \right]$$

which is subject to the constraint that we have a limited number of clusters to represent the data. The clustering problem can then be formulated as a constrained maximization problem. Finding the clusters is equivalent with maximizing the functional:

$$F = \langle s \rangle - T \cdot I(C; i)$$

where  $T$  is a Lagrangian multiplier (Arfken and Weber, 2005). Results from applying these procedures are displayed in Figure 4.10.

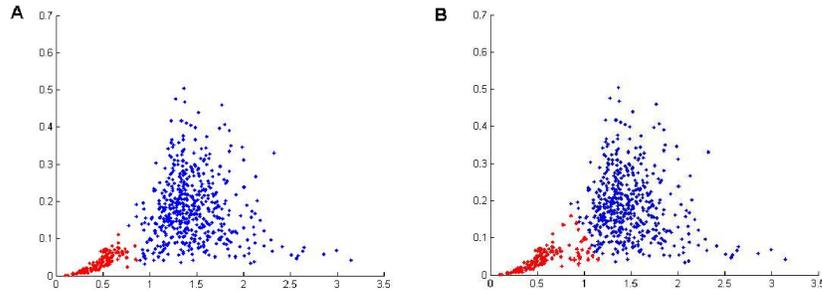


Figure 4.10: Comparison of clustering results.  $x$  – axis: average intensity inside the nucleus.  $y$  – axis: ratio inside versus outside the candidate nucleus. A: Information clustering. B: k-means clustering. Note the labeling of more points in red at the boundary between clusters.

Generally, the information clustering procedure produced better results, especially in the reduction of false negatives. The false nuclei (red cluster) is including more points (false negatives) for the k-means clustering (Figure 4.10 B). The advantage of the information clustering procedure can be attributed to the fact that it implements the intuitive notion that "tight groups are hard to join" (Slonim *et al.*, 2005). One example

of the results of this clustering procedure is presented in Figure 4.11. The robustness of our procedure in the face of partial occlusion of cell nuclei can be evaluated by examining the bottom left portion of the images in Figure 4.11. In spite of the fact that the nuclei are closely packed together, they can be reliably detected - see also Figure 4.8 for a section through the *Hough* accumulation array. The *Hough* transform provides the high sensitivity of our procedure, while the local intensity information provides the information necessary for discriminating the true cell nuclei.

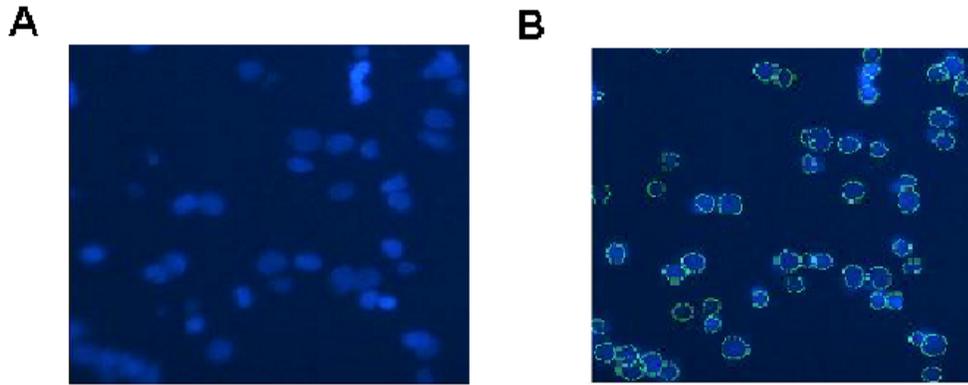


Figure 4.11: Example of detected cell nuclei. A: original image. B: detected cell nuclei are marked with a green circle.

### Deformation of circles

As a further refinement, we used the edge information to refine the borders of the detected nuclei. Where the edge information is available, we used a deformable model (Xu and Prince, 1997) to bend the marker circles, parameterized as a planar curve  $x(s)$ , toward the detected edge. In essence, each detected edge point generates an attraction force  $\gamma$ , with magnitude that decays with the distance, similar to a gravitational force. At the same time, the circles used to mark the cell nuclei are endowed with rigidity, adjustable through the parameter  $\alpha$ , and elasticity through the parameter  $\beta$ . The net result of these forces acting on the marker circles is a molding of the circles toward the true edges of the cell nuclei, which, as a result of our preceding work, is in the immediate vicinity of the true edges. We can summarize the resulting forces in the expression:

$$x(s, t) = \alpha \cdot \frac{\partial x}{\partial s} + \beta \cdot \frac{\partial^2 x}{\partial s^2} + \gamma$$

Iterative application of the above expression results in small movements of each point comprising the original circle. Using this approach, a more accurate representation of the cell nuclei becomes available - see Figure 4.12.

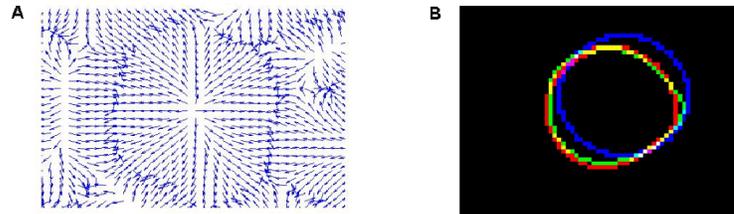


Figure 4.12: A: Example of the attraction forces generated by the edge points. Arrows indicate the direction of the force, with the length proportional with the magnitude. B: Effects of the force on the original snake (blue curve). The edge information is represented by the red curve. The deformed contour is represented by the green curve.

### Labeling of nuclear and cytoplasmic *ZENK*

Once the cell nuclei were detected, we implemented a series of steps for detecting nuclear and cytoplasmic *ZENK*, which was present in the red spectral channel (see Figure 4.1). The detection of *ZENK* markers followed a different approach compared to the detection of cell nuclei. The detection of cell nuclei used a probabilistic approach, collecting a large number of potential nuclei and subsequently discriminating the correct ones. For nuclear or cytoplasmic *ZENK*, a probabilistic approach is not feasible. Many of the images do not contain any nuclear or cytoplasmic markers and, in the images that contain *ZENK* markers, their number is too small for probabilistic approaches.

The detection of *ZENK* markers was accomplished using a rule-based approach. Initially, high intensity regions in the red channel were identified and then a series of tests were applied to these candidate *ZENK* labels. A cell was labeled as expressing nuclear or cytoplasmic *ZENK* only if all the criteria were satisfied. We start with the rules for detecting nuclear *ZENK*.

#### Select high intensity, compact and relatively small regions in the red channel

The nuclear *ZENK* marker will consist of compact regions of high intensity (see Figure 4.2 A). These regions of compact, high intensity regions are detected using the *top-hat transform* (Meyer, 1979), (Gonzalez and Woods, 2007). The top-hat transform consists of an image opening, followed by subtraction from the original

image. The image opening flattens the intensity of the original image, which in the case of nuclear *ZENK* consists of peaks of high intensity, generally convex (highest value is in the middle and they gradually taper toward the edges). The subtraction has the role of adjusting for different levels of overall red (*ZENK*) intensity, for instance when the nuclear *ZENK* is located on top of more diffuse levels of cytoplasmic *ZENK* (Figure 4.1, also Figure 4.2 B). The top-hat transform returns a set of locations where the intensity image has a local maximum on the scale of the structuring element used. If these locations fall inside a detected cell nuclei, we collect a list of these cell nuclei, which are candidates for nuclear *ZENK*, and we denote the set as  $\Omega_{\text{bright spots}}$ .

**Select cell nuclei with high red range** The above list of candidate cell nuclei contains typically contains a number of false positives, generally because cytoplasmic *ZENK* expression from neighboring nuclei. To eliminate these errors, we implemented the intuitive concept that nuclear expression consists of small, high intensity regions. The *range operator* (Gonzalez and Woods, 2007) measures large changes in intensity within a small neighborhood. For all cell nuclei detected, we compute the red channel range values inside the cell nuclei. Cells marked with nuclear *ZENK* are expected to have high range values, compared with cells overlapping with cytoplasmic *ZENK* from neighboring cells. We required that each cell nuclei have a range value of at least one standard deviation above the mean red range inside the set of all cell nuclei. The set of cell nuclei satisfying this condition is denoted as  $\Omega_{\text{high range nuclei}}$ .

Formally, our final selection of nuclear *ZENK* marked cells is the intersection of the above sets:

$$\Omega_{\text{nuclear}} = \Omega_{\text{bright spots}} \cap \Omega_{\text{high range nuclei}}$$

In conjunction, the above rules return a more accurate list of nuclear *ZENK* markers than either rule by itself. We note that both the top-hat and the range operator are robust in relation to the overall intensity from image to image. In addition, the fact that we require coincident satisfaction of both rules serves as an effective measure for the case where no nuclear *ZENK* is present, in which case very few locations will be returned from the top-hat transform, and application of the second rule eliminated most of them, resulting in very few or no false positives.

A different set of rules guided the detection of cytoplasmic *ZENK*.

**Select circular shapes in the red channel** Cytoplasmic *ZENK* consist of relatively large regions of red of moderate intensity, surrounding the cell nuclei (see Figure

4.2 B). We used the *Hough* transform to identify circular shapes in the edge image derived from the raw red channel. For each circular shape, we collected the local information: the ratio of average red intensity just outside the circular shape and the red intensity inside the circular shape. For true cytoplasmic *ZENK* expression, we expect a concavity corresponding to the position of the blue cell nuclei, and therefore the intensity inside the circle should be less than the intensity outside. Out of all the circular shapes collected using the *Hough* transform, we select the ones that have a ratio higher than the average ratio and we denote this set as  $\Psi_{red\ circles}$ .

**Identify circular shapes close to detected nuclei** The cytoplasmic *ZENK* must be in the immediate vicinity of a cell nuclei. For each candidate circular shape, we find the position of the closest detected nuclei. If this position is within one nuclei radius of the center of the circular shape, we select the circular shape and we denote the set of these circular shapes as  $\Psi_{nucleus\ attached}$ .

**Evaluate the cell nuclei border** For each cell nuclei close to a circular shape in the red channels, as identified in the above rule, we collect the average red intensity on the border. We then compare this intensity with the red intensity on the border of all nuclei detected. We select the circular shapes for which the attached nucleus has an intensity greater than the mean intensity. We denote this set as  $\Psi_{red\ nucleus\ border}$ .

Formally, our final selection of cytoplasmic *ZENK* cells is the intersection of the above sets:

$$\Psi_{cytoplasmic} = \Psi_{red\ circles} \cap \Psi_{nucleus\ attached} \cap \Psi_{red\ nucleus\ border}$$

We illustrate an example of a fully analyzed picture in Figure 4.13.

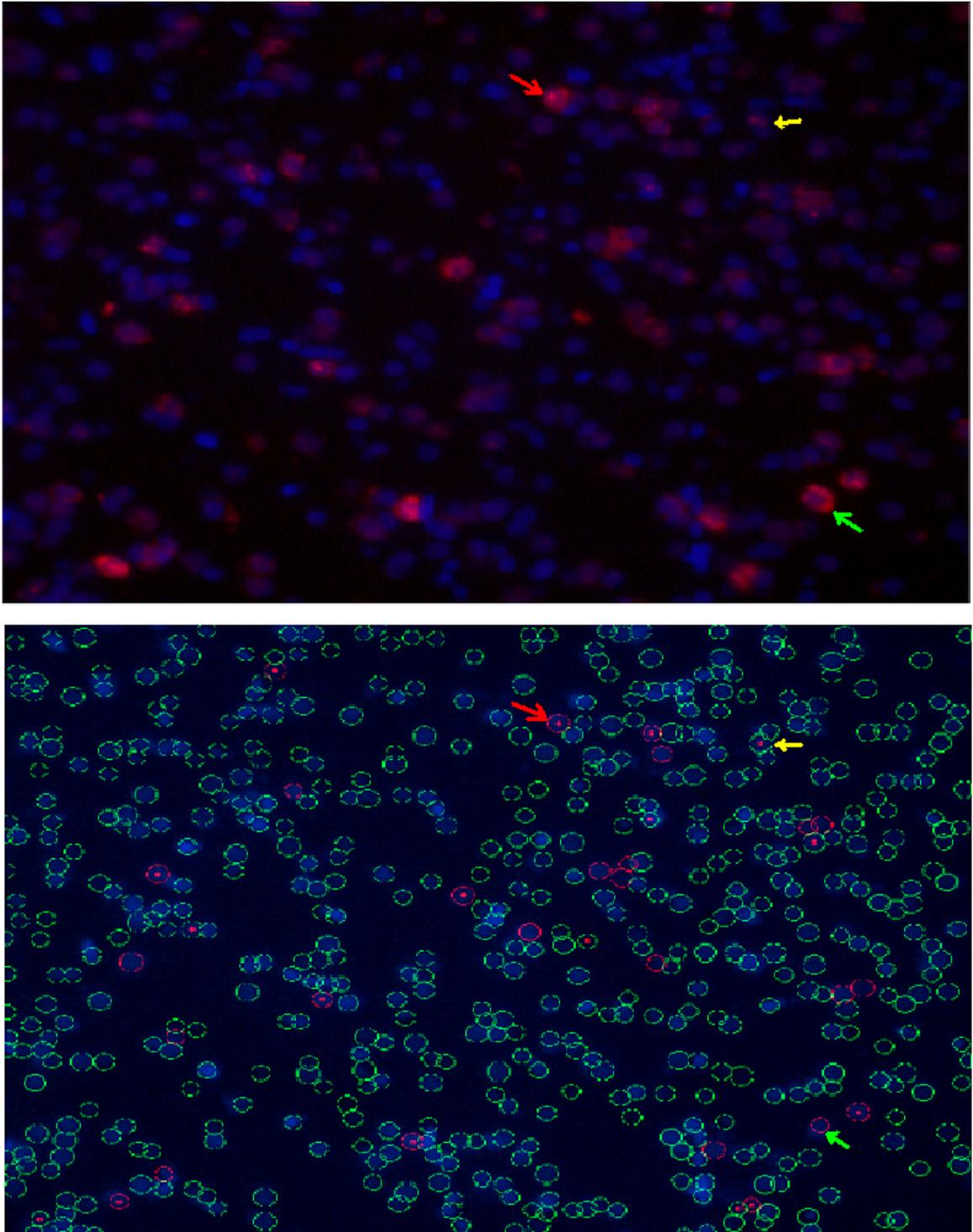


Figure 4.13: Example of fully labeled image. Top: raw image with blue and red channels. Bottom: image with nuclear labeling (red points, see also yellow arrow) and cytoplasmic labeling (red circles, see also green arrow). Red arrow: example of double-labeled cell.

## Evaluating and validating the automated cell counting

In order to quantify the results of our automated cell detection procedure, we compared its performance with that of several human observers and to the performance of a different cell counting procedure (Byun *et al.*, 2006). This cell counting procedure is based on convolving the intensity image with a Laplacian of *Gaussian* (*LoG*) filter (Gonzalez and Woods, 2007), and then detecting local maximums in this filtered image. We note two important differences between our approach, based on the *Hough* transform, and the *LoG* filter approach. First, the *Hough* transform places a heavy emphasis on the *qualitative* information present in the image, namely the roundness of the cell borders. We note that it is not essential for the cells to be perfectly round; a partial border that is round is sufficient to create a peak in the *Hough* transform. In addition, our image opening procedure increases the roundness of the borders (Gonzalez and Woods, 2007). The *LoG* filter is also somewhat sensitive to the roundness of the cells, since the filter shape is circular in the  $(x, y)$  plane, but our comparisons outlined below suggest that occlusion and local noise in the images affects its performance. Second, while both *Hough* transform approach make explicit use of the local intensity of the image, the *Hough* transform approach is more specific, since it compares the intensity inside a candidate cell with the intensity immediately outside its border. We will relate these observations with the performance of these automated procedures in the following paragraphs.

We used a set of 32 images to compare the performance of human observers and the performance of the two automated procedures. The results of all the counters are shown in Figure 4.14 A. We observe the large variability in the results of different observers and the results of the automatic procedures. In addition, the *LoG* procedure appears to systematically undercount the images, as illustrated in Figure 4.14 B, where the counts from the *LoG* procedure are significantly lower than the counts from the *Hough* transform procedure.

In order to compare the performances of all the different counters, we computed Lin's *concordance correlation coefficient* (Lin, 1989) between repeated counts of the same image. This measure is specifically designed to compare the performance of different observers quantifying the same data (Lin, 1989) and is commonly used to quantify inter-observer variability in cell counts (Warke *et al.*, 2001), (vanSchaik G. *et al.*, 2002). As detailed in (Lin, 1989), if two observers are in perfect agreement, then all the pair observations would fall on the 45° line. The concordance correlation coefficient, as defined in (Lin, 1989), implements the following measure:

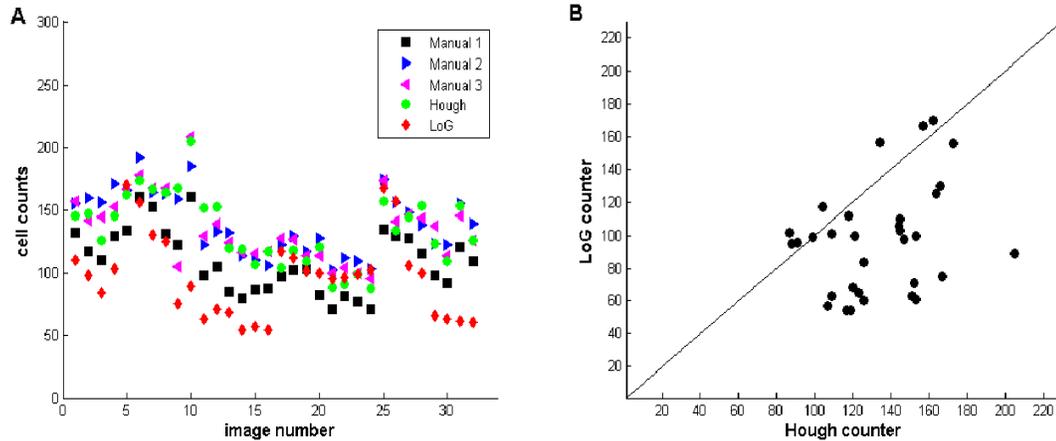


Figure 4.14: **A.** Illustration of the counts performed by humans and the counts performed by the automatic procedures. **B.** Direct comparison between the Hough and the LoG automated procedures. Perfect agreement would place all the points on the main diagonal.

Table 4.1: Concordance coefficients for manual and automatic cell counts.

	<i>Manual 1</i>	<i>Manual 2</i>	<i>Manual 3</i>	<i>Hough</i>	<i>LoG</i>
<i>Manual 1</i>	1	.55	.60	.62	.48
<i>Manual 2</i>	.55	1	.86	.84	.25
<i>Manual 3</i>	.60	.86	1	.85	.26
<i>Hough</i>	.62	.84	.85	1	.17
<i>LoG</i>	.48	.25	.26	.17	1

$$\rho_c = 1 - \frac{\text{Expected squared deviation from } 45^\circ}{\text{Expected squared deviation from } 45^\circ \text{ for uncorrelated data}}$$

As applied to our measurements, the correlation coefficient returned the results listed in Table 4.1.

Based on these concordance values, we can draw several conclusions. First, as illustrated in Figure 4.14 A, there is large variability between different counters, including the manual ones (concordance as low as 0.55). Second, the *LoG* counter has lower concordance with the humans than *Hough* counter (.48, .25 and .25 *versus* .62, .84 and .85). Third, the lowest concordance is between the two automated procedures, which we believe is due in large part to undercounting by *LoG* (see also Figure 4.14 B).

In order to identify the exact causes of undercounting by the *LoG versus Hough* counters, we generated two sets of artificial images. These images were constructed based on the several issues which we believe affect the performance of the counters.

These issues include partial occlusion (overlapping cells), variability in the intensity of the fluorescent labeling for each cell, variability in the size of the cells, and the presence of noise. Consequently, we generated 400x400 size images, and we generated 100 randomly positioned artificial cells in each of these images. In order to replicate in a controlled fashion the issues affecting the detectability of the cell nuclei (see Figure 4.11), we allowed variability in both the size and the intensity of the cells. The cell diameter was generated from a uniform distribution over the set  $\{13, 14, 15, 16, 17\}$ . The intensity of the cells was also drawn from a uniform distribution of ten levels, such that the highest intensity was tenfold the intensity of the lowest. The borders of the cells were blurred by convolving the images with an averaging filter in the form of a disk with diameter of 3 pixels. Next, the images were convolved with a two dimensional *Gaussian* filter with standard deviation of 6 pixels, which was intended to reproduce the intensity spillover effects between neighboring cells observed in real images. Finally, low-passed *Gaussian* noise was added to the images. The noise image added to each cell image was generated using the following procedure: for each pixel, we generated a random value from a *Gaussian* distribution with mean 5 (the average intensity of a cell) and standard deviation 5. This noise image was then filtered using a two-dimensional *Gaussian* with diameter 10 pixels and standard deviation 6 pixels. These values were arrived at by visually inspecting the real images such as Figure 4.11.

The first set of artificial images consisted of images with cells of varying sizes and intensities. The second set of images also included the noise. For each of these sets of images, we allowed progressively more occlusion, starting from 0 pixels (no overlap) to 14 pixels of overlap. Examples of the resulting images are presented in Figure 4.15.

Using these two sets of artificial images, we performed counts using the *Hough* counter and also the *LoG* counter. We present the results of this procedure in Figure 4.16. In the absence of noise, manual inspection revealed no falsely identified cells. When low-pass *Gaussian* noise was added, manual inspection revealed the apparition of falsely identified cells for both the *Hough* and the *LoG* counting procedures (see Figure 4.16 B).

The results illustrated in Figure 4.16 suggests to us the following conclusions. First, there is inherent variability in image-to-image counting performance for both methods. Performance generally decreases with the amount of overlap, but due to variability the trend is not strictly monotonous. Second, the decrease in performance is due to an accumulation of factors. Overlap or heterogeneity in intensity or size, by themselves, were not sufficient to degrade the performance. We note that when the images contained just one of these confounding factors in isolation, both methods returned a perfect count. The undercounting of cells is due to a combination of all the above factors as illustrated

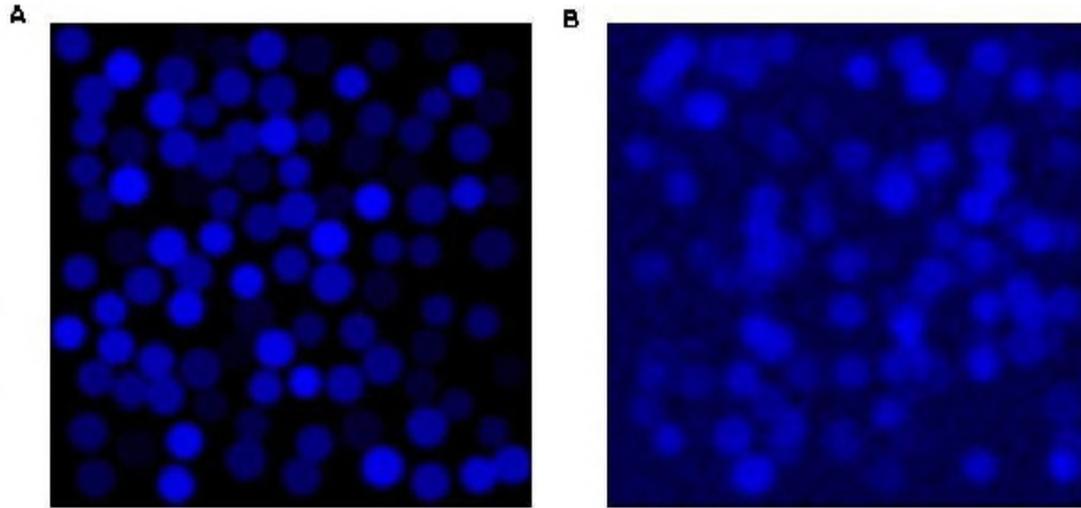


Figure 4.15: **A.** Image of 100 cells with variability in size and intensity. The images do not overlap, their borders touch due to the averaging filter that blurs the edges. **B.** Image of 100 cells which are allowed to overlap by 15 pixels, and with Gaussian noise added. This was the most challenging of the artificial images.

in Figure 4.16 A.

The introduction of noise had the most dramatic effect in the performance of both automated procedures (see Figure 4.16 B). In addition to loss of sensitivity, we also detected by manual inspection the apparition of large numbers of falsely identified cells (difference between the solid and dashed lines in Figure 4.16 B). The *Hough* counter was especially vulnerable to the introduction of noise, and this effect allows us to identify the limits of our detection procedure. We note that the amount of *Gaussian* noise introduced in our artificial images (Figure 4.15 B) is much bigger than the amount of noise that is typically present in the images collected from zebra finches (see Figure 4.8, top). We introduced such a large amount of noise in order to identify the point where our procedure exhibits a drastic loss in performance, and we conclude that the images collected from the birdsong brains are well within the operating range where we have confidence in the automatic procedure.

The low-pass *Gaussian* noise introduced into the artificial images is roughly equivalent to small spots of fluorescence, which nevertheless are not cells and should not be counted. The size of these spots of noise is dependent on the spatial extent of the *Gaussian* filter used to attenuate the noise added to the images, in this case a standard deviation of 6 pixels. Based on the above considerations, in practical terms, the

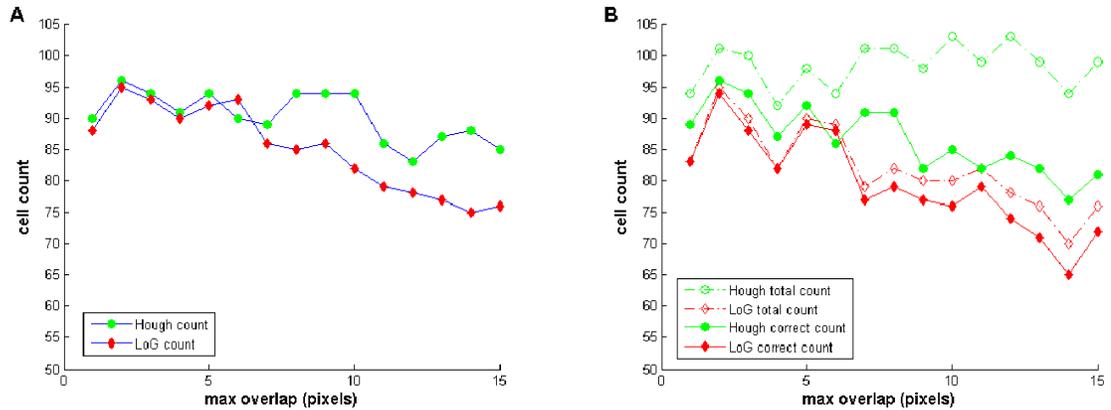


Figure 4.16: **A.** Comparison of automated counting methods using image sets with variable cell size and intensity. **B.** Counting results for images with same number of cells as A, but with added noise. Difference between dashed lines and solid lines represent the false positives, as identified by manual inspection.

*Hough* cell detector introduced here should be used with caution when the images to be analyzed contain spots of noise with a diameter close to a third of the diameter of the real cells to be detected. If the spots of noise are smaller, we expect that our image opening procedure (erosion followed by dilation) is likely to eliminate them (Gonzalez and Woods, 2007).

### Evaluating and validating the automated detection of nuclear labeling

We evaluated the accuracy of our detection of nuclear labeling in a manner similar to the evaluation of the cell counting procedure in the preceding section. We compared the results of four manual counts with the results of our automatic procedure and we display the results in Figure 4.17.

We note that in the case of nuclear labeling, we can consider the first manual count (*Manual 1*) to be the "gold standard", since it was performed by the experimentalist with years of experience analyzing fluorescent images (*Tarciso Velho, personal communication*). Figure 4.17, suggests that our automated procedure (*Hough*) is relatively close to the best performance available. In order to quantify this observation, we computed the concordance coefficient for the counts and we present the results in Table 4.2:

The concordance between the *Hough* counter and the most accurate count, *Manual 1*, is .95. This is the highest concordance in Table 4.2, and therefore we consider our automated procedure to have high reliability on this particular set of images<sup>1</sup>.

<sup>1</sup>The concordance coefficient was computed using only the counts from images 4-14. Images 1-3 came

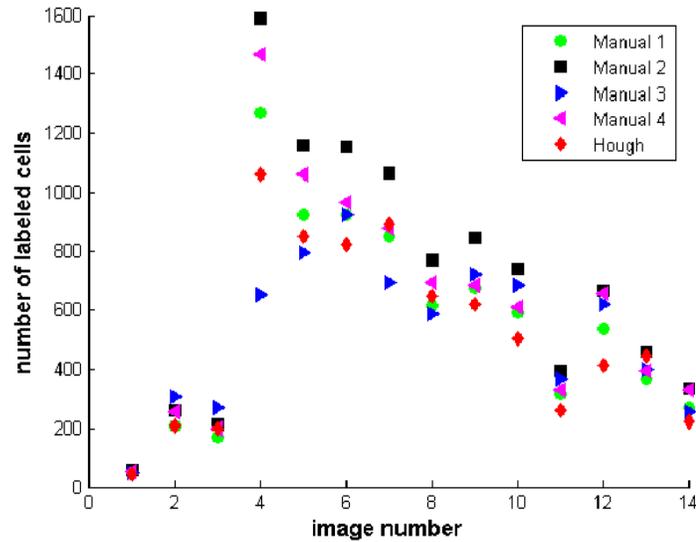


Figure 4.17: Results of manual counts and our automated procedure, applied to the 14 images in the data set. First three images were generated from the silent controls, images 4-6 from birds presented with conspecific songs, 7-9 from heterospecific songs stimuli, 9-12 from white noise and 13-14 from pure tone stimulation.

Table 4.2: Concordance coefficients for ZENK labeling of cell nuclei.

	<i>Manual 1</i>	<i>Manual 2</i>	<i>Manual 3</i>	<i>Manual 4</i>	<i>Hough</i>
<i>Manual 1</i>	1	.85	.93	.96	.95
<i>Manual 2</i>	.85	1	.75	.92	.75
<i>Manual 3</i>	.93	.75	1	.87	.83
<i>Manual 4</i>	.96	.92	.87	1	.87
<i>Hough</i>	.95	.75	.83	.87	1

## Quantifying the spatial distribution of cells

Once we have found the exact positions of the activated cells, their topographical distribution was quantified using the computational geometry techniques of *Delaunay triangulation* (Delaunay, 1934), (Okabe *et al.*, 2000) and *Voronoi tessellation* (Voronoi, 1907), (Okabe *et al.*, 2000).

**Definition 1** *Let  $P$  be a set of points in the plane. The Voronoi polygon generated by a point  $p \in P$  is the set of all points  $x$  in the plane for which the distance from  $x$  to  $p$  is*

---

from silent controls and had very few numbers of activated cells (100-300) compared to 900-1200 for conspecific stimuli. Computing the concordance coefficient assumes that the true counts are drawn from the same probability distribution (Lin, 1989). That assumption is violated if we use counts that vary widely in the number of cells, as they do between silent controls and conspecific songs.

smaller than or equal to the distance from  $x$  to  $q$ , for all other points  $q$  in  $P$ . The Voronoi tessellation of  $P$  is the collection of Voronoi polygons of all  $p \in P$ .

**Definition 2** The Delaunay triangulation of a set  $P$  of points in the plane consist of the set of all triangles spanned by points  $p \in P$ , such that no triangle contains a point  $p \in P$ .

We illustrate these concepts in Figure 4.18.

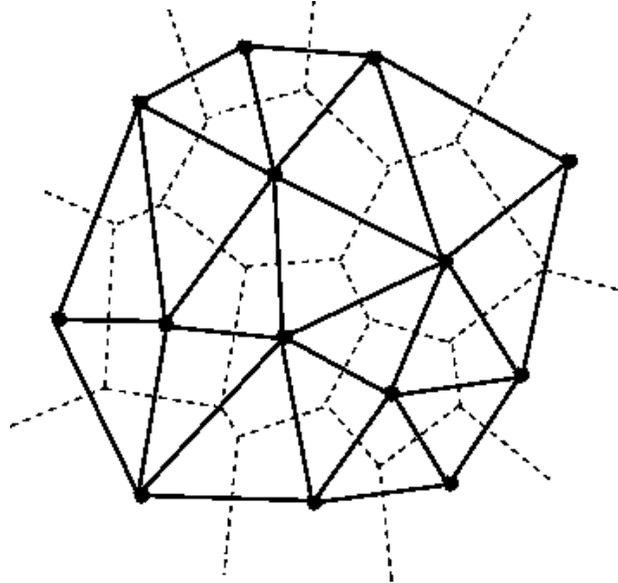


Figure 4.18: Illustration of the Delaunay triangulation and Voronoi tessellation of a set of points. Original points are denoted by black dots. Thick lines represent the Delaunay triangulation. The definition implies that no triangle sides intersect. Dashed lines outline the Voronoi polygons.

An application of the Delaunay triangulation was the stitching of the images. A neuronal slice was photographed under the microscope, resulting in about 100 partially overlapping images. Each image was analyzed individually, resulting in the coordinates of all identified cell nuclei, as outlined in the previous section. In order to properly combine the images without counting the overlapping border areas twice, we used the side length and area of the Delaunay triangles detected in two neighboring images. In conjunction, these measures provide unique identifiers for overlapping image borders. This procedure is robust to the possible errors in identifying all the cells in each image, since for practical purposes just three Delaunay triangles identified in both images will suffice to provide a good stitching. The process is illustrated in Figure 4.19.

Once the images were stitched and aligned, a *Gaussian* spatial filter was applied each aligned image, resulting in an average density map for each stimulus type. This density

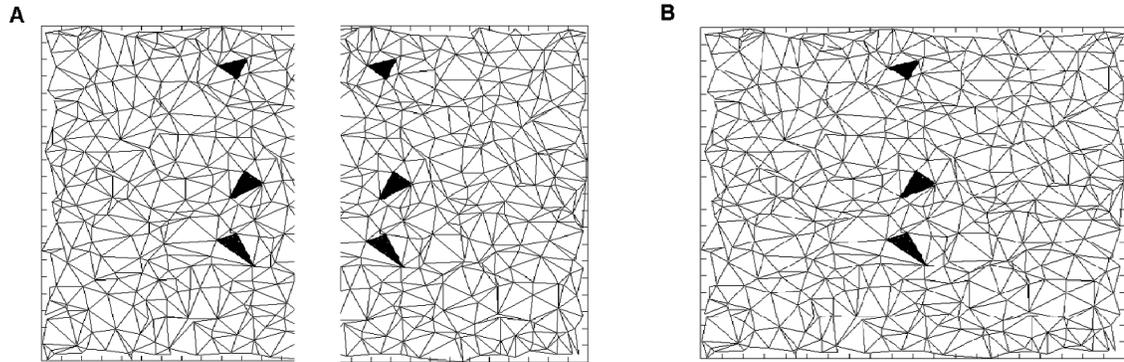


Figure 4.19: A: Two neighboring images with overlap on the borders. Delaunay triangles are identified in each image (filled triangles). B: The images are stitched together using the overlap suggested by the triangles.

information was displayed in the form of a heatmap, with the intensity of the color proportional with the local density of activated cells. An alternative visual representation can be achieved by plotting the Voronoi polygons centered at each cell, with the display color inversely proportional to the area of each polygon, as illustrated in Figure 4.20.

Our image segmentation procedure results in a set of coordinates for each labeled cell in a particular brain slice. In order to compare this topographical distribution across birds, we performed an image registration procedure aligning different brain sections to a common set of anatomical landmarks. This procedure is necessary in order to correct for histological distortions and individual variability. As landmarks, we manually chose a set of points common to the anatomy of the zebra finch brain, such as the ventricle wall (Velho, 2008). Next, we applied an *image registration* procedure (Gonzalez and Woods, 2007), which finds a spatial transformation which moves the individual landmark points, manually selected in each image, to a set of global landmark points. Once this spatial transformation is found, it is applied to the whole image, in effect aligning each image to the global standard. We used a common registration procedure, known as *projective image registration* (Gonzalez and Woods, 2007). The results of this procedure are displayed in Figure 4.21.

After image registration, the remaining variability in the distribution of cells could be attributed largely to stimulus variability. Therefore, a measure of similarity between two sets of points is required. In image processing applications, a frequently used measure of distance between two sets of points is the Hausdorff distance metric (Rote, 1991).

**Definition 3** For two sets of points  $A = \{a_1, a_2, \dots, a_n\}$  and  $B = \{b_1, b_2, \dots, b_n\}$  is defined

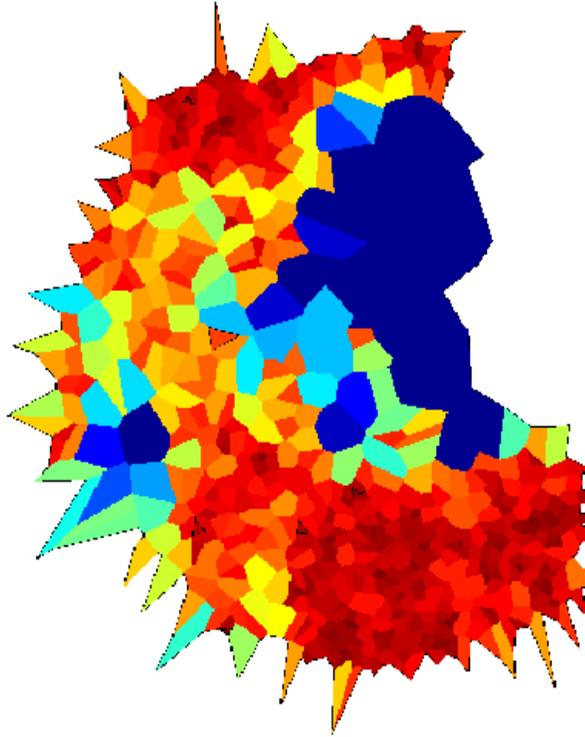


Figure 4.20: Illustration of a Voronoi tessellation applied to the whole NCM. The areas of the polygons are inversely proportional with the intensity of the color, resulting in areas of high density in warmer colors.

as

$$H(A, B) = \max(h(A, B), h(B, A))$$

where  $h(A, B) = \max_{a \in A} \min_{b \in B} \|a - b\|$  is the directed distance from  $A$  to  $B$ .

The directed distance is not symmetric and therefore not a metric. The Hausdorff distance identifies the point  $a \in A$  that is farthest from any point of  $B$ , and retains the distance from this element  $a$  to the closest  $b \in B$ .

## 4.2.2 Selective temporal or spectral degradation of songs

In this section we present the methodology used for systematically altering the songs presented as stimuli (Theunissen and Doupe, 1998). This method is based on a time–frequency decomposition of sound commonly used in speech analysis (Flanagan, 1980), and can describe any song completely using a relatively small set of parameters. This parametrization can be explicitly related to the spectral and temporal structure of the

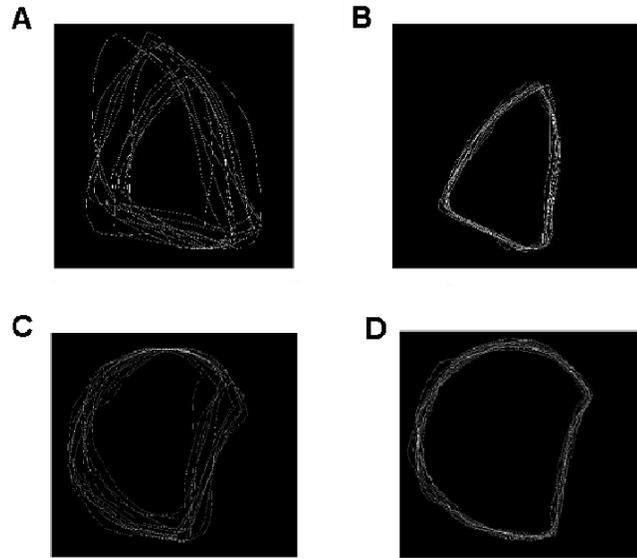


Figure 4.21: Results of image registration. Overlapping lines show the contours of same brain regions from different animals. A: *CMM*, prior to registration. B: *CMM*, post registration. C: *NCM*, prior to registration. D: *NCM*, post registration.

songs. Access to these parameters also allows a systematic degradation of selective features of the song, in particular the temporal or spectra resolution.

The original song signal  $s(t)$  is first divided into  $n$  bandpassed component signals  $s_n(t)$ , illustrated in Figure 4.22. To completely describe the signal with these, the filters in the filter banks need to collectively be equivalent to a filter transform that is flat over the spectral extent of  $s(t)$ . Additionally, the phase distortion of each filter needs to be minimized. Under these conditions, the original sound can be reconstructed as the sum of the bandpassed signals:

$$s(t) = \sum s_n(t)$$

The output of each filter is next used to computed the *analytical signal* in each frequency band (Cohen, 1994), (Flanagan, 1980):

$$s_n(t) = A_n(t) \cdot \cos[\theta(t)]$$

The instantaneous amplitude of the signal  $s_n(t)$  is represented by the factor  $A_n(t)$ , while the instantaneous phase is represented by  $\theta(t)$ . The instantaneous phase is further decomposed into its derivative, the instantaneous frequency  $w(\tau)$  and an absolute phase

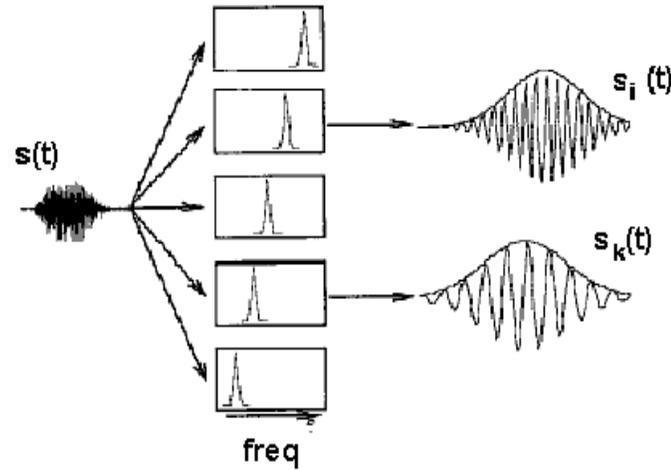


Figure 4.22: Decomposition of original sound using partially overlapping filters. Each filter covers a specified frequency band. The output of these filters can be used to reconstruct the signal. The filters used covered the range between 500 and 8000 Hz. *Modified from (Theunissen and Doupe, 1998).*

$\theta_n$

$$s_n(t) = A_n(t) \cdot \cos \left[ \int_0^t w(\tau) d\tau + \theta_n \right]$$

Finally, the instantaneous frequency  $w(\tau)$  is expressed as a modulation around the center frequency for the spectral band,  $w_n$  :

$$s_n(t) = A_n(t) \cdot \cos \left[ w_n \cdot t + \int_0^t w_{FM}(\tau) d\tau + \theta_n \right]$$

The information about the original signal  $s(t)$  is partially contained in the individual amplitude envelopes  $A_n(t)$ . If a large number of filters are used, the frequency band covered by each one is small. With this increased frequency resolution comes the loss of temporal resolution. Choosing the number of filters, or equivalently the bandwidth of each filter, is the first and most critical step in the process of degrading a song. The amplitude envelopes  $A_n(t)$  still contain spectral information, but limited to frequencies below the bandwidth of the filter.

The second source of information about the signal  $s(t)$  is the collection of instantaneous phases  $\theta_n(t)$ . These instantaneous phases are related across different frequency bands, and the information carried by their relative magnitude is known as the fine

temporal structure of the signal. A more subtle way of degrading the song is to diminish the information carried by the fine temporal structure of the song. This was accomplished by low-pass filtering the modulations  $w_{FM}$  around the center frequencies in each frequency band,  $w_n$ .

Following suggestions from the authors of the package implementing these methods (Frederick Theunissen, personal communication), we created two synthetic songs, starting from the same original. The first song, which we called *temporally preserved*, was obtained by choosing the upper limit of the temporal frequency (modulations in the amplitude envelope) of up to 500 Hz and setting the spectral frequency cutoff to be 0.0005 (or 0.5 cycles/kHz). The second song, which we called *spectrally preserved*, was obtained by choosing the upper limit of the temporal frequency (modulations in the amplitude envelope) of up to 3 Hz and setting the spectral frequency cutoff to be 0.016 (or 16 cycles/kHz). The spectrograms of the original song and its modified versions are presented in Figure 4.23.

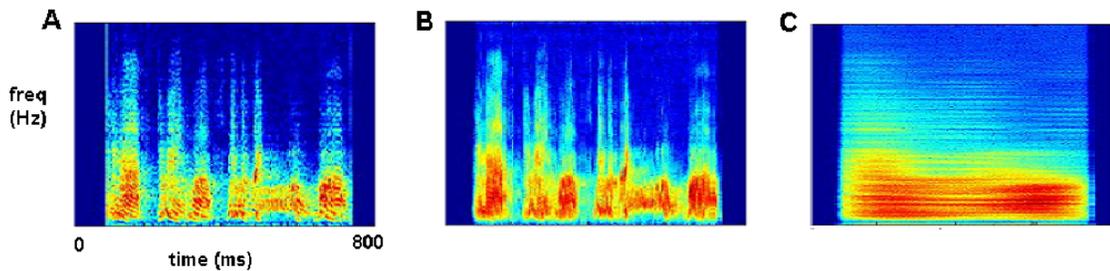


Figure 4.23: Spectrograms of song stimuli. A: unmodified song. B: temporally preserved song. C: spectrally preserved song.

We note that the temporally preserved song (Figure 4.23 B) has kept the pauses between the syllables, but fine frequency resolution is lost. In contrast, the spectrally preserved song (Figure 4.23 C) has good spectral resolution, but the syllable partition is lost.

### 4.3 Results

The images analyzed in this section were generated from brain slices from a total of 14 different birds (Velho, 2008). Prior to the *FISH* experiment, the birds were exposed for about 30 min to repeated presentation of auditory stimuli. These stimuli consisted of conspecific song (*CON*,  $n = 3$  adult females) and other sounds that have very

different spectro-temporal organizations, including heterospecific (canary) song (*HET*,  $n = 3$ ), white noise (*WN*,  $n = 3$ ), and tonal sequences (*TONE*,  $n = 2$ ), in addition to unstimulated controls (*UNSTIM*;  $n = 3$ ).

### 4.3.1 Visualization of the density of activated cells

Using the mapping tools developed in our *Methods* section, we investigated the spatial distribution of activated cells following stimulus presentation. Starting from the unprocessed images containing the cell nuclei and *ZENK* fluorescent markers (for an example see Figure 4.1), we mapped the position of each activated cell, either manually or using our automated process. Next, we used the common anatomical markers to register the images to a common standard. For each stimulus type, all images were combined resulting in an image containing all activated cells from several animals. This image was smoothed using a *Gaussian* filter and visualized as a heatmap. The processing steps are summarized in Figure 4.24.

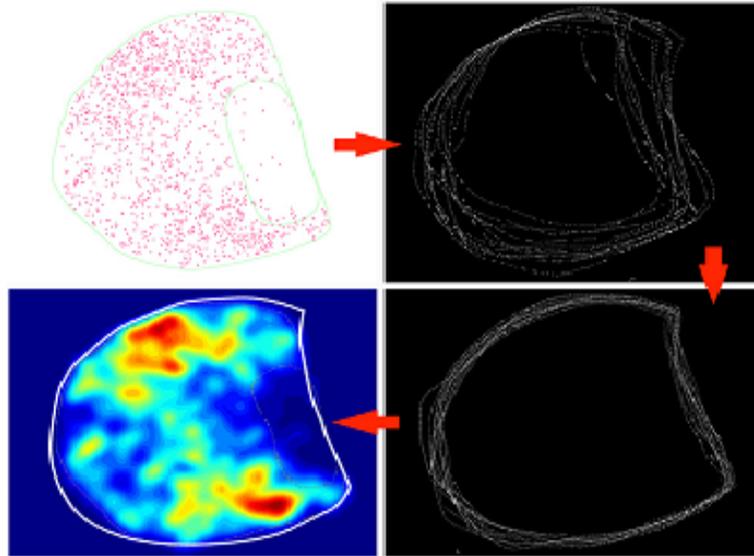


Figure 4.24: Image processing steps. Starting from the positions of activated cells (red dots in top left image), we registered the images to a common standard, we combined them by common stimulus class, and then we displayed them as heatmaps.

Individual density maps of *ZENK* expression were registered and averaged. The resulting average density maps revealed overall increases in *ZENK* labeled cells in the *NCM* of birds hearing *CON*, *HET* and *WN* compared to unstimulated controls. Additionally, the spatial pattern was not uniform and varied significantly across stimulus

type. In *CON* stimulated birds, *ZENK* expressing cells were concentrated mostly in dorsal and ventral *NCM*, with a lower density in the more central region. A moderate but significant activation was observed also in the caudal *NCM*. In contrast, *ZENK* expressing cells in *HET*- and *WN* stimulated birds were mostly located in central *NCM* with some labeled cells detected also in dorsal *NCM*. *ZENK* expressing cells in caudal *NCM* were very nearly absent for all stimuli except *CON*, suggesting that this area may be the most selective *NCM* subregion.

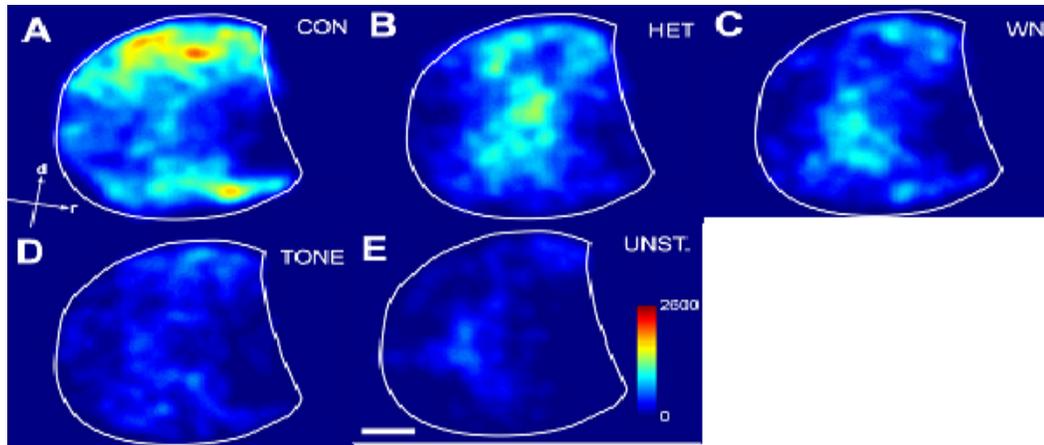


Figure 4.25: Density of *ZENK* expressing cells in *NCM*. A: conspecific songs; B: canary songs; C: white noise; D: pure tones; E: unstimulated controls.

In order to quantify the discriminative power in regards to stimulus type that is present in the topographical distribution of activated cells in *NCM*, we computed the pairwise distances between all images. Our hypothesis is that images generated by the same stimulus type will be more similar than images generated using different stimulus types. We have found a significant difference (*ranksum test*,  $p < 0.05$ , see Figure 4.26 B), showing that same category stimuli produced more similar spatial distributions of activated cells than different category stimuli. Our similarity measure was the Hausdorff metric introduced in the *Methods* section. A multidimensional scaling of this set of pairwise distances is shown in Figure 4.26 A.

Using the distances between images, we performed a clustering procedure that assigned each image to a stimulus category, based on the lowest average distance of each image to images resulting from similar stimuli. This procedure is identical with the clustering method presented in *Chapter 3*, where clustering was performed using spike metric distances. Regardless of the neural responses available in a particular experiment (spike trains or spatial patterns of activated cells), the approach we use conforms to the guiding principles and goals outlined in *Introduction*: we aim to quantify how much the

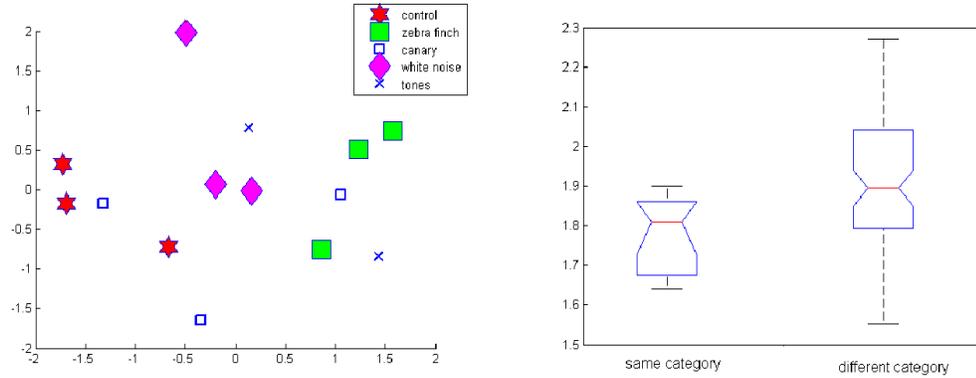


Figure 4.26: A: Two dimensional Isomap projection of the pairwise distances between images. B: Boxplots comparing the distances between same-category stimuli images and different category images.

Table 4.3: Confusion matrix of true and assigned stimulus category of each image.

	<i>UNSTIM</i>	<i>ZF</i>	<i>HET</i>	<i>WN</i>	<i>TONE</i>
<i>UNSTIM</i>	3	0	0	0	0
<i>ZF</i>	0	3	0	0	0
<i>HET</i>	0	1	2	0	0
<i>WN</i>	0	0	1	2	0
<i>TONE</i>	0	0	0	1	1

neural responses tell us about the stimuli presented. The neural responses are clustered and then each one is assigned to a stimulus category. If these assigned categories are close to the true stimulus categories, we can conclude that having access to the responses is informative in regard to the stimuli. We present the results of this procedure in the confusion matrix outlined in Table 4.3.

In Table 4.3, rows represent the true stimulus categories and columns represent the category assigned to each image by our clustering procedure. In a perfect assignment, all the entries not on the main diagonal would be zero. In our case, 11 out of 14 images are assigned to the correct stimulus category, based solely on the spatial patterns of activated cells. We note that the Hausdorff distance measure used to compute distances between images is not taking into consideration the numbers of cells in each image. We also illustrate this observation empirically. We selected one of our images, consisting of 1200 activated cells. We computed the Hausdorff distance between this image and the other 13 images in the data set. Next, we undersampled the image, by selecting, at random, progressively smaller numbers of points from the original points in the image. For each of these undersampled images, we computed the distance to the other 13 images in the

data set (which remained intact), in order to detect what effect the number of points has in computing the Hausdorff distance. The results of this procedure are presented in Figure 4.27.

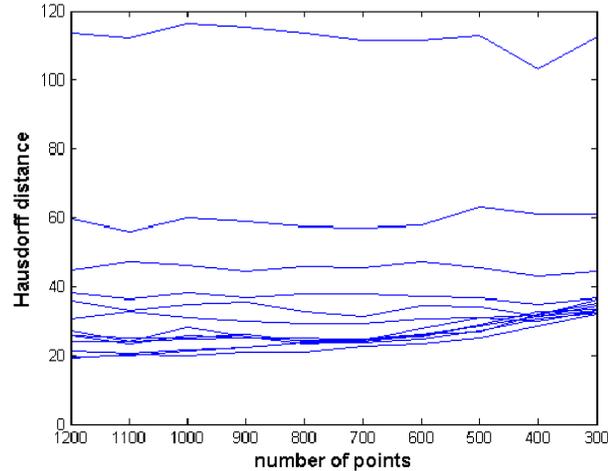


Figure 4.27: The effect of reducing the number of points on the Hausdorff distance. Each curve represent the distance between a selected image and the rest of the images. As we decrease the number of points in the selected image, the distances to the other images remain relatively constant.

By examining the relatively flat curves in Figure 4.27, we conclude that changing the number of points in the image, but leaving the overall spatial distribution the same does not significantly affect the Hausdorff measure. Based on this observation, we conclude that the specific position of the activated cells in the images contains additional information related to the specific stimulus presented to the birds, in addition to the strong effect in the number of cells activated by each stimulus class.

The number of cells activated is also dependent on the stimulus class used to stimulate the birds (Velho, 2008). A distance measure to combine the information present in the number of cells with the information about their relative distribution will be introduced in *Chapter 5*.

Auditory stimulation elicits activation of both gabaergic (inhibitory) and non-gabaergic (probably excitatory) cells (Pinaud *et al.*, 2004). To investigate the separate contributions of these cell types to the ensemble of song-responsive neurons in *NCM*, we have identified and mapped cells expressing *ZENK* and *gad65*, a gabaergic marker, using double-labeling in situ hybridization (Velho, 2008).

We can observe an effect of stimulus type, as the number of *ZENK*+/*gad65*+ cells was significantly higher for *HET* than for the other stimuli. Indeed, the central region

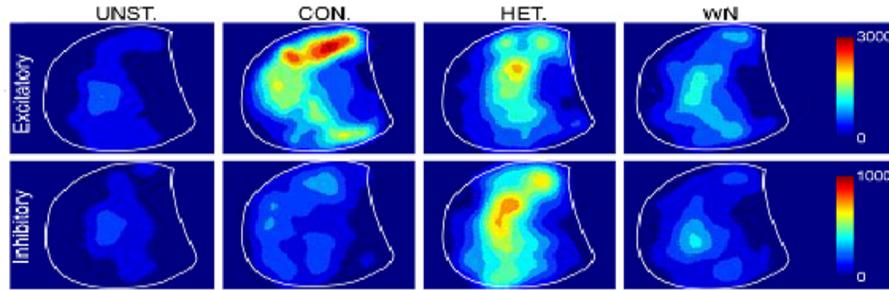


Figure 4.28: Excitatory (top row) and inhibitory (bottom row) density of activated cells in *NCM*, in response to auditory stimuli.

of *NCM* in the *HET* group had a high density of *ZENK*+/*gad65*+ cells compared to the other stimuli, which had rather low levels of this cell type throughout *NCM* (Figure 4.28, lower panels). This contrasted with *ZENK*+/*gad65*- cells, which were more numerous for *CON* and *HET* than for *WN* and suggested higher values for *CON* than *HET*.

The classes of stimuli presented so far (conspecific songs, heterospecific, white noise, tones) differ *qualitatively* in their origin and presumed saliency for the bird. In order to gain an understanding about the relative importance of the frequency and time compositions of conspecific song, we compared a normal song (Song A) to the same song after a decomposition that degraded its frequency and time components. We generated time-preserved (Song Atp) and frequency-preserved (Song Afp) songs respectively, as described in the *Methods* section of this chapter.

The expression pattern evoked by song A (Figure 4.29 D) seemed to differ markedly from those elicited by the modified songs (Figure 4.29 E-F): the modified songs appeared to have much higher densities of labeled cells in central *NCM* and lower ones in dorsal, caudal and ventral regions compared to intact song A.

We also displayed the separate contributions of gabaergic and non-gabaergic cells to these patterns using dFISH for *ZENK* and *gad65*. The density maps indicated that the spectrally and temporally modified versions of song A resulted in higher densities of *ZENK*+/*gad65*- (non-gabaergic) cells in central *NCM* as compared to the intact song. A similar effect was observed for *ZENK*+/*gad65*- (gabaergic) cells (Figure 4.30 top, compare middle and right to left panel).

The techniques introduced in this study demonstrate the capacity of activity-dependent gene expression to be used as a mapping tool for neural activation. This methodology allows for brain activation mapping at cellular resolution in awake behaving animals. This offers the possibility of automated mapping neural activity at a spatiotemporal

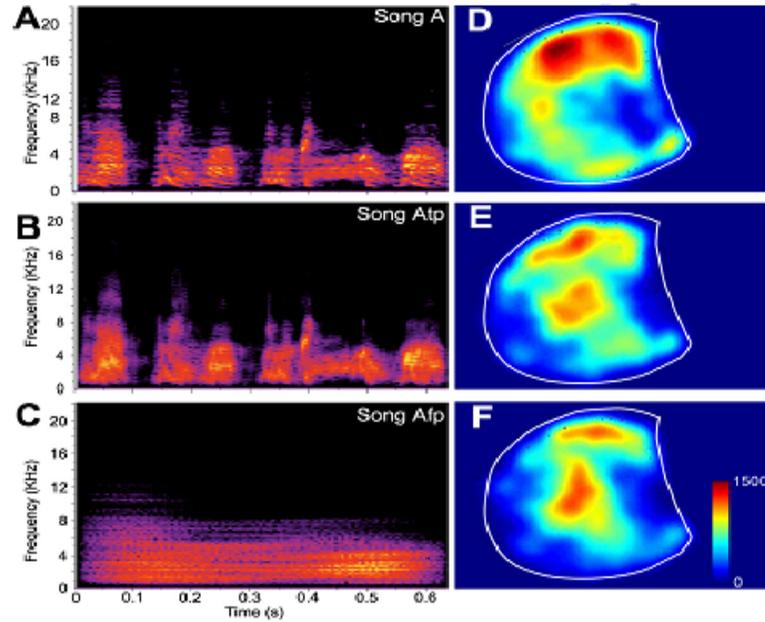


Figure 4.29: Spatial pattern of *ZENK* expressing cells in *NCM*. A-C: spectrograms of intact or modified songs. D-F: heatmaps representing the density of activated cells.

level that is inaccessible to other neurophysiological techniques. Electrode recordings, even using multielectrode arrays, do not offer the possibility of accessing the activity of large, spatially distributed collections of neurons. Other imaging techniques such as fMRI lack the spatial resolution required to isolate specific brain regions such as the *NCM* (Boumans *et al.*, 2007). Our automated method of cell nuclei segmentation and labeling is extensible to three-dimensional stacks of images, creating the possibility of visualizing and quantifying the full spatial extent of neuronal activation in response to various stimuli. Such a method could potentially fill a gap in the set of methods available to neurophysiologists.

New methods such as calcium imaging techniques do increase the spatial resolution of the data collected during an experiment, but they are limited in other ways such as the requirement that the structure under study be accessible to a microscope. In contrast, *IEG* fluorescent labeling accesses neuronal activity in a live, behaving and intact animal.

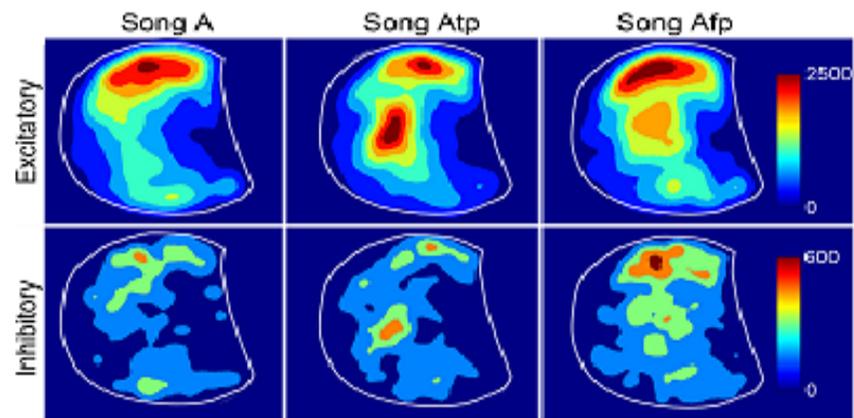


Figure 4.30: Excitatory (top) and inhibitory (bottom) spatial distribution in response to original (left) and modified (middle, right) versions of the same song.

## Chapter 5

### Conclusions and future directions

#### 5.1 Timing and coincidence detection at different levels of neural processing

In the previous chapters, we have analyzed neuronal systems where the role of precise timing in the transmission and processing of information was important. Our results suggest that coincidence detection of converging signals in relatively short time windows plays a significant role in the functioning of neural assemblies. An incoming sensory signal is evaluated at subsequent levels of neural processing, and its saliency or relevance to the organism is determined in large part by its arrival time. We can exemplify this concept in each of our model systems.

At the level of synaptic integration detailed in *Chapter 2*, we observed how a centrally originating timing signal, the corollary discharge, has a profound effect at the very early stages of sensory perception. The granular cells of *ELL* are just a synapse away from the first neurons responding to sensory stimulation, the afferents. At the level of the granular cells, the integration of the centrally originating corollary discharge signal and the afferent spikes is largely determined by their relative timing (Zhang *et al.*, 2007). Our simulations reproduced the cellular properties of the granular cells and provided a candidate biophysical mechanism for implementing interactions at such short time scales. Overall, these results illustrate how a short window of time is created by the arrival of the corollary discharge signal at the granular cell. Afferent spikes that arrive at lower holding potentials have a considerably larger effect than afferent spikes that arrive at higher holding potentials (see Figure 2.1 B). We hypothesize that the holding potential is a result of multiple inputs, including the corollary discharge and possible inhibitory inputs from large gabaergic cells observed in the immediate vicinity of the granular cells (Han *et al.*, 2000). These inhibitory cells are activated very rapidly, possibly playing a role in the value of the membrane potential of the granular cell at the time of afferent

spike arrival.

At the system level, in *Chapter 3* we analyzed the neural activity recorded in the primary auditory cortex while the animal was presented with auditory stimuli relevant for performing a task. The importance of signal timing was revealed in several ways. First, the level of response attenuation was dependent on time interval between stimuli, such that identical inputs resulted in different outputs, depending on their timing (see Figure 3.12). Second, the level of fine temporal coordination of *LFP* signals, as measured by the coherence, was consistently increased on the presence of the stimulus (see Figure 3.9). The spikes emitted in response to the tone were extremely well timed (see spike rasters in Figure 3.7), and taking into account their timing dramatically increased the capacity of an observer to detect the presence of a tone (see Figure 3.10). Finally, the relatively brief interval between the two tones revealed a transitory increase in gamma frequency power related to the performance of the animal (see Figure 3.16). Taken together, these results illustrate that at the system level, millisecond scale coordination between multiple streams of information (coherence) and a well-timed succession of events (increase in gamma frequency preceding the second tone) are important parameters relating sensory performance and subsequent behavior.

The neural recording available from the two model systems described above had extremely fine temporal resolution, but their spatial resolution was very low. In contrast, in the songbird system we analyzed in *Chapter 4*, we had access to extremely high spatial resolution neural data, but the temporal resolution was on the order of minutes which is the time necessary for gene expression to become visible in the nucleus of the activated neurons. However, in this case the importance of timing was revealed by manipulating the inputs presented to the animals. Since conspecific songs are the stimuli with the highest relevance for the organism, manipulating the fine temporal structure of the songs resulted in markedly different spatial patterns of activated cells over the extent of the same neural structure (see Figure 4.29). This illustrates the high importance of subtle cues in the relative timing of different frequencies composing the zebra finch song.

In conclusion, our results have illustrated how the analysis of complex spatiotemporal patterns of neural responses can be guided by two important constraints. The first such constraint is the temporal extent of the stimuli and/or neural responses. An example is the extreme sensitivity of some neurons in the bird brain to extremely brief features in the conspecific songs. Often, determining the best temporal scale of analysis is an important step in the selection of appropriate methods.

An additional, related constraint is the saliency of the stimuli or neural responses for the subsequent stages of processing. In many cases, the temporal window available for

processing the incoming signals can be determined by considering the latency of response at subsequent stages. For instance, in the primary auditory cortex, the motor activity follows the presentation of the second tone by less than 250 ms, implying that within that time frame the animal has made a decision.

## 5.2 Further refinements of the methods

The development of appropriate methods for the analysis and modeling of the sets of data available for our study has been a central component of the present work. We outline below future steps meant to improve or complement these analysis methods. Providing to the neuroscience community intuitive, user-friendly versions of the software developed during our work is also an important goal.

### 5.2.1 Translation of timing into spike count

In *Chapter 2*, we developed a biophysically detailed model of the granular cell. The model was sensitive to the relative timing of two inputs, responding with a large *EPSP* (and presumably more spikes) to a shorter inter-stimulus interval. We have shown the critical importance of timing in a highly specific, biophysically detailed model.

This concept can be generalized using a canonical model of a neuron, a two dimensional system capable of reproducing the capabilities of real neurons, such as spiking, bursting and rebound inhibition. In several sensory systems, the first level of sensory encoding makes use of the latency of the first spike for encoding the strength of the sensory stimulation. At subsequent stages of sensory processing, spike counts appear to be significantly modulated by the same sensory stimulation. Our aim is to determine whether the dynamics of two-dimensional system resembling a single neuron are capable of translating the latency information into a spike count, or a more complex network based interaction is necessary for such a translation.

A two dimensional dynamical system consists of two equations describing the rate of change of each variable:

$$\begin{aligned}\frac{dV}{dt} &= f(U, V) \\ \frac{dU}{dt} &= g(U, V)\end{aligned}$$

The *nullclines* of the system are the loci of the points defined by  $\frac{dV}{dt} = 0$  or  $\frac{dU}{dt} = 0$ , respectively. The intersections of the nullclines are called *critical points*, and their num-

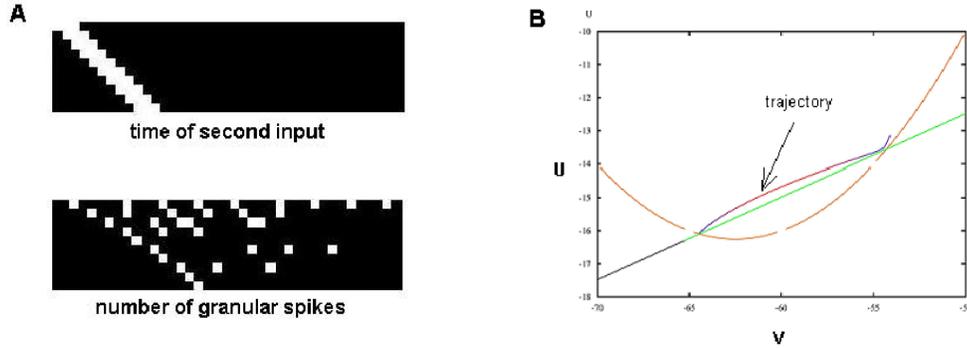


Figure 5.1: A. Simulation results illustrate translation from latency to spike count. Top: time of arrival of the second input. White portions represent the stimulus on. B: Phase space illustration of the canonical model. Parabola and straight lines are nullclines, short curve is the trajectory. Color along the trajectory represents the speed.

ber, nature and relative position can characterize the dynamical system in a qualitative way (Izhikevich, 2000). This type of mathematical analysis of neuronal excitability has a rich tradition in neuroscience, beginning with the abstract models of neuronal spike generation (Fitzhugh, 1961), constructed based on detailed measurements of channel conductances (Hodgkin and Huxley, 1952). In more recent years, many neuronal properties such as different models of spike generation (integration *versus* resonance) have been described using this approach (Izhikevich, 2000). We believe that this general method could be useful in describing the biophysical underpinnings of coincidence detection. A preliminary version of a model neuron, behaviorally similar to the granular cells but without biophysical details, has shown promising results (see Figure 5.1).

Two identical pulses, separated by variable a short interval of time, represent the inputs to our model. The output of the system is a spike burst of variable count, with shorter interstimulus intervals resulting in more spikes (see Figure 5.1 A). Our proposed mechanism for this translation can be intuitively visualized in Figure 5.1 B. The first input takes the system from rest (the critical point to the left) to a point close to the second critical point (to the right and higher). As time progresses, the system returns toward the rest point along the trajectory depicted in 5.1 B. The speed along this trajectory is non-uniform, and this is pivotal for our proposed mechanism. The second input will catch the system somewhere along the trajectory and take it to a new point in phase space (not shown in figure). The return from this new point entails a spiraling trajectory towards the rest point, with each spiral corresponding to a spike. In summary, the non-uniform speed on the return trajectory ensures that the second input takes the system to points well separated in the  $(U, V)$  plane. In turn, these new starting points

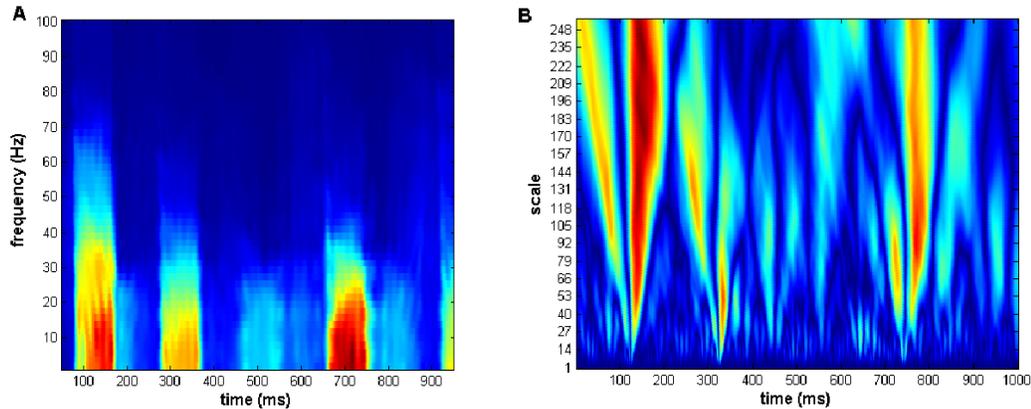


Figure 5.2: A. Spectrogram of a trial where the animal was presented with two tones and licked. The spectrogram was computed using overlapping windows 100 ms in length. B. Wavelet transform using the sym2 wavelet (Daubechies, 1994). Note that small scales (low on the  $y$  axis) approximately correspond to high frequencies (high on  $y$  axis in A)

will generate different number of spikes by starting the spiraling return trajectory in different positions.<sup>1</sup>

### 5.2.2 Wavelet transform for time-scale analysis

Our analysis of *LFP* recordings from the primary auditory cortex had two basic components. First, we isolated the signal immediately following a tone, a high amplitude but brief deflection, and we applied *LDA* analysis. Second, we isolated the signal between the tones, longer in extent, and we performed a Fourier spectral decomposition of this signal. We believe that a comprehensive approach which has the potential to operate at different scales is possible using the wavelet transform (Daubechies, 1994). A comparison of the results of a short time Fourier transform and the wavelet transform is presented in Figure 5.2)

The advantages of using the wavelet transform are immediately apparent. At small scales, we have very good temporal resolution, such that we can identify visually the times of tone presentation (100 ms and 300 ms into the trial, see Figure 5.2 B). Even though more signal power is present at higher scales, the wavelet transform isolates the time of onset for the neural response to tone presentation. In addition, we can also observe neural activity probably related to the licking action, between 700-800 ms

<sup>1</sup>As depicted above in Figure 5.1 B, the system is not capable of a spiraling trajectories, since this would intersect the trajectory depicted. The mechanism described above is truly realizable in three dimensions (Izhikevich, 2000), or alternatively in a two dimensional model with hard reset (Izhikevich, 2003). Our model uses a hard reset.

in Figure 5.2 B. In contrast, the spectrogram in Figure 5.2 A is subject to the time-frequency trade-offs inherent in using the short time Fourier transform. The temporal resolution does not allow the clear identification of the time of tone onset. The high power present at low frequencies leaks into the higher frequencies in a more dramatic fashion than in the wavelet transform.

A more precise identification of the timing of neural events related to tone presentation or licking action opens new avenues of investigation, extending the analysis presented in *Chapter 3*. It is possible to evaluate whether the latency of the tone responses is related to the behavioral performance. We could also quantify the response attenuation at different scales of the tone responses. The precise identification of the licking time would allow for the analysis of an additional piece of information present in the data set: the exact time the animal licks is identified by the breaking of a light switch placed in front of water spout.

### 5.2.3 Voronoi tessellations and a measure of similarity between overlapping sets of points

Previous studies of the auditory response in the birdsong brain (Velho, 2008) have shown that the total number of neurons activated in structures such as the *NCM* is strongly modulated by the stimulus type, with conspecific songs activating the largest number of cells. In *Chapter 4*, we have shown that the relative distribution of the activated neurons is also informative about the stimulus type. In order to quantify differences between patterns of activated cells, we used the Hausdorff distance, a choice consistent with our goal of determining if the spatial distribution of the activated cells is informative about the stimulus, above and beyond the information contained in the number of cells.

The relative insensitivity of the Hausdorff distance to the numbers of points (activated cells) in the images, combined to the fact that the number of activated cells is actually informative about the stimulus type, (Velho, 2008), reveals the need for a distance measure between point sets which is sensitive to both. The reasons for which such a distance measure is desirable can be illustrated with the case of the spike metric, which is sensitive to both spike count and to spike timing (see Figure 3.10). As more information about the spike timing is taken into account, the neural response becomes more and more informative about the stimulus type, as illustrated by the dramatically increased clustering in Figure 3.10 C *versus* Figure 3.10 C. In a similar manner, combining the information in the number of cells and the information about their relative

distribution would reveal the true strength of the relationship between the stimulus type and the neural response. We also believe that a method for quantifying distances between point sets would have immediate applicability to images generated by calcium imaging, which also consist of patterns of cells activated by different stimuli. Based on these considerations, we have constructed a preliminary version of a distance measure that takes into account both the number of cells and their distribution.

The image in Figure 4.20 suggests that the global distribution of cells can be described using the area associated with each Voronoi polygon. Based on this insight, we defined a measure of similarity between the spatial distribution of two different but globally overlapping sets of points. Each point  $p \in A$  has associated with it the area of its own Voronoi polygon,  $Area_p^A$ . The same point in image  $A$  is also located in one and only one polygon in the Voronoi tessellation generated by image  $B$ ,  $Area_p^B$ . Without loss of generality, assume that the number of points in  $A$  is greater than the number of points in  $B$ . Thus, the  $n$  points of image  $A$  generate two vectors, each of length  $n$ :

$$\begin{aligned} Area_{p \in A}^A &= [Area_1^A, Area_2^A \dots Area_{n_1}^A] \\ Area_{p \in A}^B &= [Area_1^B, Area_2^B \dots Area_{n_1}^B] \end{aligned}$$

If two images have a distribution of points that is similar, then they will have Voronoi polygons with small areas in the same regions. The strength of association can be measured by the *Spearman footrule distance* between the vectors representing the polygon areas (Diaconis and Graham, 1977).

$$v(Area_{p \in A}^A, Area_{p \in A}^B) = \sum_{i=1}^n |a_i - b_i|$$

where  $a_i$ ,  $b_i$  are the rankings of  $Area_{p \in A}^A$  and  $Area_{p \in A}^B$  respectively. Transforming the raw scores, in our case the polygon areas, into rankings, guards against too much weight being given to isolated points that can have Voronoi polygons with large areas.

**Definition 4** *The Voronoi tessellation based dissimilarity  $\Delta$  between two overlapping collections of points in  $R^2$  is defined as*

$$\Delta(A, B) = v(Area_{p \in A}^A, Area_{p \in A}^B) + v(Area_{p \in B}^A, Area_{p \in B}^B)$$

where  $\mathbf{a}$ ,  $\mathbf{b}$  are the ranking vectors of Voronoi polygon areas and  $A, B \subset R^2$ .

If  $A = B$ , both  $v$  distances will equal zero and therefore  $\Delta = 0$ . We note that the above measure of dissimilarity takes into account both the number of points in the data

sets and also their relative densities. If the number of points is vastly different, then one of the  $Area_{p \in A}^A, Area_{p \in A}^B$  will have a large number of duplicates. This will be the dominant part of the sum. On the other hand, if the number of points are equal, then the distance between the sets will be completely determined by the relative densities.

**Proposition 5** *Let  $\Phi$  be a set whose elements are sets of points  $A \subset R^2$ . Then  $(\Phi, \Delta)$  is a semi-metric space, satisfying the properties of non-negativity and symmetry. For the appended set  $\widehat{\Phi}$  defined bellow, the triangle inequality is satisfied and  $(\widehat{\Phi}, \Delta)$  is a full metric space.*

**Non-negativity:** Immediate since definition of  $\Delta$  involves sums of squares and  $n \geq 2$ .

**Symmetry:** Interchange of  $A, B$  in definition yields equal expression, since we choose the image with the larger number of points.

**Triangle inequality:** We need to show that

$$\Delta(A, B) + \Delta(A, C) \geq \Delta(B, C)$$

In the case where the sets  $A, B, C$  have the same number of points, the triangle inequality is an established result. (Diaconis and Graham, 1977). At this point, we introduce some restrictions on the set  $\Phi$  over which our measure  $\Delta$  can operate as a metric. We require that there exist some set of points  $M \in \Phi$  which has a finite but maximum number of points compared with any other set of points  $X \in \Phi$ . For these other sets of points  $X \in \Phi$ , we append the Voronoi tessellation of  $X$  by duplicating the Voronoi polygons in the tessellation of  $X$  which contain multiple points from  $M$ , and we call this set  $\widehat{X}$ , resulting in a collection  $\widehat{\Phi}$ . In the strict mathematical sense, it is  $(\widehat{\Phi}, \Delta)$  which satisfies the triangle inequality  $\Delta(\widehat{A}, \widehat{B}) + \Delta(\widehat{A}, \widehat{C}) \geq \Delta(\widehat{B}, \widehat{C})$ . To see this, note that the appended sets  $\widehat{A}, \widehat{B}, \widehat{C}$  all have the same number of points equal to  $|M|$ .

In order to use the above dissimilarity measure, we will search for a version that satisfies the triangle inequality. Failing to find such a metric, we would clearly identify the issues arising from using a non-metric dissimilarity measure. Non-metric similarity measures are useful in evaluating pairwise relationships between images (Tan *et al.*, 2006). A review of the statistical issues related to non-metric similarity measures can be found in (Fasham, 1977).

## References

- George B. Arfken and Hans J. Weber. *Mathematical Methods For Physicists*. Academic Press, 6th edition, June 2005.
- D. H. Ballard. Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognition*, 13:111–122, 1981.
- J. Bastian. Pyramidal-cell plasticity in weakly electric fish: A mechanism for attenuating responses to reafferent electrosensory inputs. *J. Comp. Physiol. A*, 176:63–73, 1995.
- Markus Bauer, Robert Oostenveld, Maarten Peeters, and Pascal Fries. Tactile spatial attention enhances gamma-band activity in somatosensory cortex and reduces low-frequency activity in parieto-occipital areas. *J. Neurosci.*, 26:490–501, 2006.
- Curtis C. Bell, Kirsty Grant, and Jacques Serrier. Sensory processing and corollary discharge effects in the mormyromast regions of the mormyrid electrosensory lobe: I. Field potentials and cellular activity in associated structures. *J. Neurophysiol.*, 68:843–858, 1992.
- Curtis C. Bell, D. Bodznick, J. Montgomery, and J. Bastian. The generation and subtraction of sensory expectations within cerebellum-like structures. *Brain. Beh. Evol.*, 50:17–31, 1997. (suppl.1).
- Curtis C. Bell. Sensory coding and corollary discharge effects in mormyrid electric fish. *J. Exp. Biol.*, 146:229–253, 1989.
- Curtis C. Bell. Mormyromast electroreceptor organs and their afferent fibers in mormyrid fish. II. Intra-axonal recordings show initial stages of central processing. *J. Neurophysiol.*, 63:303–318, 1990.
- Jennifer K. Bizley, Fernando R. Nodal, Victoria M. Bajo, Israel Nelken, and Andrew J. King. Physiological and anatomical evidence for multisensory interactions in auditory cortex. *Cereb. Cortex*, 17:2172–2189, September 2007.

- Tiny Boumans, Frederick E Theunissen, Colline Poirier, and Annemie Van Der Linden. Neural representation of spectral and temporal features of song in the auditory forebrain of zebra finches as revealed by functional MRI. *The European Journal of Neuroscience*, 26, November 2007.
- E. A. Brenowitz, D. Margoliash, and K. W. Nordeen. An introduction to birdsong and the avian song system. *Journal of Neurobiology*, 33:495–500, November 1997.
- Michael Brosch and Christoph E. Schreiner. Time course of forward masking tuning curves in cat primary auditory cortex. *J Neurophysiol*, 77(2):923–943, February 1997.
- M. Brosch, A. Schulz, and H. Scheich. Processing of sound sequences in macaque auditory cortex: response enhancement. *Journal of neurophysiology*, 82(3):1542–1559, September 1999.
- J. S. Byun, Verardo M R, Sumengen B, Lewis G P, Manunath B S, and Fisher S. Automated tool for detection of cell nuclei in digital microscopic images: Application to retinal images. *Molecular Vision Res.*, 12:949–960, 2006.
- M. B. Calford and M. N. Semple. Monaural inhibition in cat auditory cortex. *J Neurophysiol*, 73(5):1876–1891, May 1995.
- J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Mach. Intell.*, 8:679–698, 1986.
- Michael R. Chernick. *Bootstrap Methods: A Practitioner's Guide (Wiley Series in Probability and Statistics)*. Wiley-Interscience, September 1999.
- Leon Cohen. *Time Frequency Analysis: Theory and Applications*. Prentice Hall PTR, 1994.
- E. D'Angelo, T. Nieuw, A. Maffei, S. Armano, P. Rossi, V. Taglietti, A. Fontana, and G. Naldi. Theta-frequency bursting and resonance in cerebellar granule cells: Experimental evidence and modeling of a slow K<sup>+</sup>-dependent mechanism. *J. Neurosci.*, 21:759–770, 2001.
- Ingrid Daubechies. Ten lectures on wavelets. *SIAM*, 61:198–202 and 254–256, 1994.
- Peter Dayan and L. F. Abbott. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems*. The MIT Press, September 2005.

- B. Delaunay. Sur la sphere vide. *Bull. Acad. Sci. USSR*, 7:793–800, 1934.
- Persi Diaconis and R. L. Graham. Spearman’s footrule as a measure of disarray. *Journal of the Royal Statistical Society. Series B (Methodological)*, 39:262–268, 1977.
- Sam M. Doesburg, Alexa B. Roggeveen, Keiichi Kitajo, and Lawrence M. Ward. Large-scale gamma-band phase synchronization and selective attention. *Cereb. Cortex*, 18:386–396, February 2008.
- A. J. Doupe and P. K. Kuhl. Birdsong and human speech: common themes and mechanisms. *Annual Review of Neuroscience*, 22:567–631, 1999.
- Richard O. Duda and Peter E. Hart. Use of the Hough transformation to detect lines and curves in pictures. *Commun. ACM*, 15:11–15, 1972.
- Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification*. Wiley-Interscience, 2nd edition, November 2000.
- A. K. Engel, P. Konig, A. K. Kreiter, and W. Singer. Interhemispheric synchronization of oscillatory neuronal responses in cat visual cortex. *Science (New York, N.Y.)*, 252:1177–9, May 1991.
- M. J. R. Fasham. A comparison of nonmetric multidimensional scaling, principal components and reciprocal averaging for the ordination of simulated coenoclines, and coenoplanes. *Ecology*, 58:551–561, 1977.
- Richard Fitzhugh. Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.*, 1(6):445–466, July 1961.
- D. C. Fitzpatrick, S. Kuwada, D. O. Kim, K. Parham, and R. Batra. Responses of neurons to click-pairs as simulated echoes: auditory nerve to auditory cortex. *The Journal of the Acoustical Society of America*, 106(6):3460–3472, December 1999.
- J. L. Flanagan. Parametric coding of speech spectra. *Acoustical Society of America Journal*, 68:412–419, August 1980.
- W. Frawley, G. Piatetsky-Shapiro, and C. Matheus. Knowledge discovery in databases: An overview. *AI Magazine*, 13:57–70, 1992.
- Jean D. Gibbons and Subhabrata Chakraborti. *Nonparametric Statistical Inference*. CRC, 4th edition, May 2003.

- Sharon M.H. Gobes and Johan J. Bolhuis. Birdsong memory: A neural dissociation between song recognition and production. *Current Biology*, Vol 17:789–793, May 2007.
- Rafael C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Prentice Hall, 3rd edition, August 2007.
- Y. Gottlieb, E. Vaadia, and M. Abeles. Single unit activity in the auditory cortex of a monkey performing a short term memory task. *Exp Brain Res*, 74(1):139–148, 1989.
- David M. Green and John A. Swets. *Signal Detection Theory and Psychophysics*. Wiley, 1966.
- T Gruber, Muller M.M., Keil A., and Elbert T. Selective visual-spatial attention alters induced gamma band responses in the human EEG. *Clinical Neurophysiology*, 110:2074–2085, December 1999.
- Victor Han, Kirsty Grant, and Curtis C. Bell. Rapid activation of GABAergic interneurons and possible calcium independent GABA release in the mormyrid electrosensory lobe. *J. Neurophysiol.*, 83:1592–1604, 2000.
- Bertil Hille. *Ion Channels of Excitable Membranes*. Sinauer Associates, 3rd casebound edition, July 2001.
- M. Hines and T. Carnevale. The neuron simulation environment. *Neural Comput.*, 9:1179–1209, 1997.
- S. Hocherman and E. Gilat. Dependence of auditory cortex evoked unit activity on interstimulus interval in the cat. *Journal of neurophysiology*, 45(6):987–997, June 1981.
- A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol. (London)*, 108:37–77, 1952.
- P. V. C. Hough. Machine analysis of bubble chamber pictures. *International Conference on High Energy Accelerators and Instrumentation*, pages 554–556, 1959.
- Eugene M. Izhikevich. Neural excitability, spiking and bursting. *International Journal of Bifurcation & Chaos*, 10(6):1171–1266, June 2000.

- E. M. Izhikevich. Simple model of spiking neurons. *IEEE Transactions on Neural Networks*, 14:1569–1572, 2003.
- M. R. Jarvis and P. P. Mitra. Sampling properties of the spectrum and coherency of sequences of action potentials. *Neural Comput.*, 13(4):717–749, April 2001.
- D. C. Javitt, C. E. Schroeder, M. Steinschneider, J. C. Arezzo, and H. G. Vaughan. Demonstration of mismatch negativity in the monkey. *Electroencephalography and clinical neurophysiology*, 83(1):87–90, July 1992.
- Christoph Kayser, Christopher I. Petkov, and Nikos K. Logothetis. Visual modulation of neurons in auditory cortex. *Cereb. Cortex*, 18:1560–1574, July 2008.
- N. Keren, N. Peled, and A. Korngreen. Constraining compartmental models using multiple voltage recordings and genetic algorithms. *J Neurophysiol*, 94(6):3730–3742, December 2005.
- J. D. Kralik, D F Dimitrov, D J Krupa, D B Katz, D Cohen, and M A Nicolelis. Techniques for long-term multisite neuronal ensemble recordings in behaving animals. *Methods (San Diego, Calif.)*, 25:121–50, October 2001.
- Peter Lakatos, Nóra Szilágyi, Zsuzsanna Pincze, Csaba Rajkai, István Ulbert, and György Karmos. Attention and arousal related modulation of spontaneous gamma-activity in the auditory cortex of the cat. *Cognitive Brain Research*, 19(1):1–9, March 2004.
- Kwang-Hyuk Lee, Leanne M. Williams, Michael Breakspear, and Evian Gordon. Synchronous gamma activity: a review and contribution to an integrative neuroscience model of schizophrenia. *Brain Research Reviews*, 41(1):57–78, January 2003.
- Gang Lin, Umesh Adiga, Kathy Olson, John F Guzowski, Carol A Barnes, and Badrinath Roysam. A hybrid 3d watershed algorithm incorporating gradient cues and object models for automatic segmentation of nuclei in confocal image stacks. *Cytometry. Part A: The Journal of the International Society for Analytical Cytology*, 56:23–36, November 2003.
- Gang Lin, Monica K Chawla, Kathy Olson, Carol A Barnes, John F Guzowski, Christopher Bjornsson, William Shain, and Badrinath Roysam. A multi-model approach to simultaneous segmentation and classification of heterogeneous populations of cell nuclei in 3d confocal microscope images. *Cytometry. Part A: The Journal of the International Society for Analytical Cytology*, 71:724–36, September 2007.

- L. I. Lin. A concordance coefficient to evaluate reproducibility. *Biometrics*, 45(1):255–268, 1989.
- R. Llinas and U. Ribary. Coherent 40-hz oscillation characterizes dream state in humans. *Proceedings of the National Academy of Sciences of the United States of America*, 90:2078–81, March 1993.
- You Ming Lu, Isabelle M. Mansuy, Eric R. Kandel, and John Roder. Calcineurin-mediated LTD of GABAergic inhibition underlies the increased excitability of CA1 neurons associated with LTP. *Neuron*, 26:197–205, 2000.
- S. Mäntysalo and R. Näätänen. The duration of a neuronal trace of an auditory stimulus as indicated by event-related potentials. *Biological psychology*, 24(3):183–195, June 1987.
- D. Margoliash and E. S. Fortune. Temporal and harmonic combination-sensitive neurons in the zebra finch’s hvc. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 12:4309–26, November 1992.
- D. Margoliash. Acoustic parameters underlying the responses of song-specific neurons in the white-crowned sparrow. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 3:1039–57, May 1983.
- Roberto Martuzzi, Micah M. Murray, Christoph M. Michel, Jean-Philippe Thiran, Philippe P. Maeder, Stephanie Clarke, and Reto A. Meuli. Multisensory interactions within human primary cortices revealed by bold dynamics. *Cereb. Cortex*, 17:1672–1679, July 2007.
- C. V. Mello, D S Vicario, and D F Clayton. Song presentation induces gene expression in the songbird forebrain. *Proceedings of the National Academy of Sciences of the United States of America*, 89:6818–22, August 1992.
- C V. Mello. Mapping vocal communication pathways in birds with inducible gene expression. *Journal of Comparative Physiology. A, Neuroethology, Sensory, Neural, and Behavioral Physiology*, 188:943–59, December 2002.
- C. E. Metz, D J Goodenough, and K Rossmann. Evaluation of receiver operating characteristic curve data in terms of information theory, with applications in radiography. *Radiology*, 109:297–303, November 1973.

- F. Meyer. Iterative image transformations for an automatic screening of cervical smears. *J. Histochem. Cytochem.*, 27:128–135, 1979.
- P. P. Mitra and B. Pesaran. Analysis of dynamic brain imaging data. *Biophys J*, 76(2):691–708, February 1999.
- John C. Montgomery and David Bodznick. An adaptive filter that cancels self-induced noise in the electrosensory and lateral line mechanosensory systems of fish. *Neurosci. Letters*, 174:145–148, 1994.
- R. Näätänen and T. Picton. The n1 wave of the human electric and magnetic response to sound: a review and an analysis of the component structure. *Psychophysiology*, 24(4):375–425, July 1987.
- Miguel A. L. Nicolelis and Sidarta Ribeiro. Multielectrode recordings: the next steps. *Current Opinion in Neurobiology*, 12:602–6, October 2002.
- Toemme Noesselt, Jochem W. Rieger, Mircea Ariel Schoenfeld, Martin Kanowski, Hermann Hinrichs, Hans-Jochen Heinze, and Jon Driver. Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *J. Neurosci.*, 27:11431–11441, October 2007.
- F. Nottebohm. Reassessing the mechanisms and origins of vocal learning in birds. *Trends in Neurosciences*, 14:206–11, May 1991.
- D. Oertel and E. D. Young. What’s a cerebellar circuit doing in the auditory system? *Trends. Neurosci.*, 27:104–10, 2004.
- Atsuyuki Okabe, Barry Boots, Kokichi Sugihara, and Sung Nok Chiu. *Spatial Tessellations: Concepts and Applications of Voronoi Diagrams*. Wiley, 2 edition, August 2000.
- Christo Pantev, Thomas Elbert, and Bernd Lütkenhöner. *Oscillatory Event-Related Brain Dynamics*. Springer, 1 edition, 1995.
- C. Pantev. Evoked and induced gamma-band activity of the human cortex. *Brain Topography*, 7:321–30, 1995.
- A. Park and C. Fry. Statistical modeling of user switching behavior based on reward histories. [http://web.mit.edu/9.29/www/brett/ca\\_model.html](http://web.mit.edu/9.29/www/brett/ca_model.html), 2004.

- Gregory Piatetsky-Shapiro and William J. Frawley. *Knowledge Discovery in Databases*. AAAI Press, December 1991.
- Raphael Pinaud and Liisa Tremere. *Immediate Early Genes in Sensory Processing, Cognitive Performance and Neurological Disorders*. Springer-Verlag, New York, 2006.
- Raphael Pinaud, Tarciso A. F. Velho, Jin K Jeong, Liisa A .Tremere, Ricardo M. Leao, Henrique von Gersdorff, and Claudio V. Mello. Gabaergic neurons participate in the brain's response to birdsong auditory stimulation. *The European Journal of Neuroscience*, 20:1318–30, September 2004.
- William H. Press, Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes: The Art of Scientific Computing*. Cambridge University Press, August 2007.
- F. Pulvermüller, W. Lutzenberger, H. Preissl, and N. Birbaumer. Spectral responses in the gamma-band: physiological signs of higher cognitive processes? *Neuroreport*, 6:2059–64, October 1995.
- R. A. Reale and J. F. Brugge. Directional sensitivity of neurons in the primary auditory (ai) cortex of the cat to successive sounds ordered in time and space. *Journal of neurophysiology*, 84(1):435–450, July 2000.
- S. Ribeiro, G. A. Cecchi, M. O. Magnasco, and C. V. Mello. Toward a song code: evidence for a syllabic representation in the canary brain. *Neuron*, 21:359–71, August 1998.
- F. Rieke, D. Warland, R. R. d. R. v. Stevenick, and W. Bialek. *Spikes: Exploring the Nerual Code*. MIT Press, Cambridge, MA, 1997.
- Gunter Rote. Computing the minimum hausdorff distance between two point sets on a line under translation. *Inf. Process. Lett.*, 38:123–127, 1991.
- Claude E. Shannon, Warren Weaver, and Shannon. *The Mathematical Theory of Communication*. University of Illinois Press, September 1998.
- W. Singer. The formation of cooperative cell assemblies in the visual cortex. *The Journal of Experimental Biology*, 153:177–97, October 1990.
- D. Slepian and H. O. Pollak. Prolate spheroidal wavefunctions fourier analysis and uncertainty. i. *Bell Sys. Tech. J.*, 40(1):43–63, 1961.

- Noam Slonim, Gurinder Singh Atwal, Gasper Tkacik, and William Bialek. Information-based clustering. *Proceedings of the National Academy of Sciences of the United States of America*, 102:18297–302, December 2005.
- C. Spearman. The proof and measurement of association between two things. *The American journal of psychology*, 100(3-4):441–471, 1987.
- M. Steriade, F. Amzica, and D. Contreras. Synchronization of fast (30-40 hz) spontaneous cortical rhythms during brain activation. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 16:392–417, 1996.
- Catherine Tallon-Baudry, Olivier Bertrand, Marie-Anne Henaff, Jean Isnard, and Catherine Fischer. Attention modulates gamma-band oscillations differently in the human lateral occipital cortex and fusiform gyrus. *Cereb. Cortex*, 15:654–662, May 2005.
- Xiaoyang Tan, Songcan Chen, Jun Li, and Zhi-Hua Zhou. Learning non-metric partial similarity based on maximal margin criterion. In *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1*, pages 168–145. IEEE Computer Society, 2006.
- Ofer Tchernichovski, Partha P. Mitra, Thierry Lints, and Fernando Nottebohm. Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science*, pages 2564–2569, March 2001.
- Joshua B. Tenenbaum, Vin de Silva, and John C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319–2323, December 2000.
- Thomas A. Terleph, Claudio V. Mello, and David S. Vicario. Auditory topography and temporal response dynamics of canary caudal telencephalon. *Journal of Neurobiology*, 66:281–92, February 2006.
- F. E. Theunissen and A. J. Doupe. Temporal and spectral sensitivity of complex auditory neurons in the nucleus hvc of male zebra finches. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 18:3786–802, May 1998.
- H. Tiitinen, J. Sinkkonen, K. Reinikainen, K. Alho, J. Lavikainen, and R. Näätänen. Selective attention enhances the auditory 40-hz transient response in humans. *Nature*, 364:59–60, July 1993.

- vanSchaik G., Lotem M., and Schukken Y. H. Trends in somatic cell counts, bacterial counts, and antibiotic residue violations in new york state during 1999-2000. *Journal of Dairy Science*, 85(4):782–789, 2002.
- Tarciso A. F. Velho, Raphael Pinaud, Paulo V. Rodrigues, and Claudio V. Mello. Co-induction of activity-dependent genes in songbirds. *The European Journal of Neuroscience*, 22:1667–78, October 2005.
- Tarciso A. F. Velho, Peter Lovell, and Claudio V. Mello. Enriched expression and developmental regulation of the middle-weight neurofilament (nf-m) gene in song control nuclei of the zebra finch. *The Journal of Comparative Neurology*, 500:477–97, 2007.
- Tarciso Velho. Auditory representation of complex sounds and hierarchical selectivity in the zebra finch brain. *Doctoral Thesis, Oregon Health & Science University*, 2008.
- Jonathan D. Victor, David H. Goldberg, and Daniel Gardner. Dynamic programming algorithms for comparing multineuronal spike trains via cost-based metrics and alignments. *Journal of Neuroscience Methods*, 161(2):351–360, April 2007.
- J. D. Victor. Spike train metrics. *Curr Opin Neurobiol*, 15(5):585–592, October 2005.
- Juan R. Vidal, Maximilien Chaumon, J. Kevin O’Regan, and Catherine Tallon-Baudry. Visual grouping and the focusing of attention induce gamma-band oscillations at different frequencies in human magnetoencephalogram signals. *J. Cogn. Neurosci.*, 18:1850–1862, November 2006.
- Georgy Voronoi. Nouvelles applications des paramètres continus à la théorie des formes quadratiques. *Journal für die Reine und Angewandte Mathematik*, 133:97–178, 1907.
- T. J. Warke, Kamath S., Fitch P. S., Brown V., Shields M. D., and Ennis M. The repeatability of nonbronchoscopic bronchoalveolar lavage differential cell counts. *European Respiratory Journal*, 18:1009–1012, 2001.
- Uri Werner-Reiss, Kristin Porter, Abigail Underhill, and Jennifer Groh. Long lasting attenuation by prior sounds in auditory cortex of awake primates. *Experimental Brain Research*, 168(1):272–276, January 2006.
- Chenyang Xu and Jerry L. Prince. Gradient vector flow: A new external force for snakes. In *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR ’97)*, page 66. IEEE Computer Society, 1997.

- Jianmei Zhang, Victor Z. Han, Johannes Meek, and Curtis C. Bell. Granular cells of the mormyrid electrosensory lobe and postsynaptic control over presynaptic spike occurrence and amplitude through an electrical synapse. *J Neurophysiol*, Jan 2007.
- B. Zipser and M. V. L. Bennet. Interaction of electrosensory and electromotor signals in lateral line lobe of a mormyrid fish. *J. Neurophysiol.*, 39:713–721, 1976.