An Optimal Control Model of Zebra Finch Vocalization

Mike Schachter *
Helen Wills Neuroscience Institute
University of California, Berkeley
Berkeley, CA 94720
mike.schachter@gmail.com

Abstract

In this work, a nonlinear oscillator modeling the syringeal folds of the Zebra Finch is controlled by the state of higher level linear dynamical system. We formulate an optimal control cost function and solution for the learning of bird song.

1 Introduction

First we'll discuss the Zebra Finch vocalization system and it's mathematical formulation, and then the control of that system. All code used to generate figures and results in this paper can be found at:

http://github.com/mschachter/birdy

1.1 The Zebra Finch Vocalization System

Like human speech, Zebra Finch song is generated at it's source by oscillations in airflow generated by vibrating vocal cords. A model for these oscillations is given as:

$$\begin{aligned} \dot{x} &= v \\ \dot{v} &= \gamma^2\alpha + \gamma^2\beta x - \gamma^2 x^3 - \gamma x^2 v + \gamma^2 x^2 - \gamma x v \end{aligned}$$

The control parameters of the model are α and β. γ = 23500 is a constant.

1.2 Control of the Syrinx Model

The goal is to control the parameters α and β in order to produce an observed vocalization. Let $\boldsymbol{\phi}(t) = [\alpha(t)\ \beta(t)]^T$ be the state vector. Let $F_f(t)$ be the observed fundamental frequency at time t.

Assume that the temporal evolution of $\boldsymbol{\phi}(t)$ is defined by a controlled linear dynamical system:

$$\dot{\boldsymbol{\phi}} = A\boldsymbol{\phi} + B\boldsymbol{u}$$

where A is a 2x2 matrix, chosen to make the passive dynamics of the control system decay to some physiologically relevant rest state. B is a 2x2 control matrix, chosen to be the identity matrix for simplicity.

*With much help from Hédi Soula: hsoula@gmail.com

In optimal control theory, a cost function is specified and is to be minized over time to produce an optimal control law. The time-scale of the simulation will be much finer than that of the control. Let $\Delta\tau$ be the time step for simulation of the control system, and define $t_k = t - k\Delta\tau$. We then define the instantaneous cost function as:

$$C\left(\boldsymbol{\phi}(t_{k-1}), F_f(t_k), \boldsymbol{u}\right) = \boldsymbol{\phi}(t_k)^T Q \boldsymbol{\phi}(t_k) + \boldsymbol{u}^T R \boldsymbol{u} + C_f(F_f(t_k), \boldsymbol{\phi}(t_k))$$

where Q and R are 2x2 penalty matrices, $\boldsymbol{\phi}(t_k)$ is given as a forward Euler step:

$$\boldsymbol{\phi}(t_k) = \left(A\boldsymbol{x} + B\boldsymbol{u}\right)\Delta\tau + \boldsymbol{\phi}(t_{k-1})$$

and the cost of getting the right controls for the desired fundamental frequency is defined as:

$$C_f(F_f, \boldsymbol{\phi}) \propto \frac{1}{p\left(F_f|\boldsymbol{\phi}\right)}$$

where $p\left(F_f|\boldsymbol{\phi}\right)$ is the probability of observing a fundamental frequency F_f given a control $\boldsymbol{\phi}$.

1.3 Temporal Hierarchy of Representation

Generically, we want to determine a joint probability between $\boldsymbol{\phi}$ and a set of variables that are related to the acoustic representation. These features may span several time scales. For example, instead of just relating $F_f(t)$ and $\boldsymbol{\phi}$, i.e. looking at the joint distribution $p\left(F_f(t), \boldsymbol{\phi}\right)$, we might want to look at the running variance of $F_f(t)$ for a specified time window.

Let $\{\sigma_1, ..., \sigma_m\}$ be a set of statistics for some acoustic variables. Construct this set so that they are ordered by time scale. By this, we mean that computing σ_i requires a larger window of time than computing σ_j if $i < j$.

Finding the optimal control requires maximizing a conditional probability:

$$\operatorname*{argmin}_{\boldsymbol{u}} C_f = \operatorname*{argmax}_{\boldsymbol{\phi}(\boldsymbol{u})} p\left(\boldsymbol{\phi}(\boldsymbol{u})|\sigma_1, ..., \sigma_m\right)$$

There may be a "telescoping" algorithm to maximizing $p\left(\boldsymbol{\phi}(\boldsymbol{u})|\sigma_1, ..., \sigma_m\right)$ quickly. Start with σ_1, which is the statistic with the longest time scale. There should be many instantaneous values of $\boldsymbol{u}(t)$ that give a nonzero probability of occurance with σ_1. But not all the values will, so restrict the search for all $\boldsymbol{u}$ to that space. Do the same for σ_2, which will reduce the size of the space even further. Continue this process until the space of actual $\boldsymbol{u}$ is small enough to do a more efficient optimization, and then peform that optimization to find the optimal control.

Acknowledgments

Thank you to Hédi Soula, who created a C++/Python implementation of this model in our lab and has provided ideas, expertise, and feedback on this project.

References