

Automated Extraction of Mouse Vocalizations from Noisy Recordings

Matthew Moreno*, Adam A. Smith*

*University of Puget Sound



Abstract

Quantification of the type and frequency of mouse ultrasonic vocalizations (USVs) can serve as an effective assay of mouse social and affective state. Identifying and characterizing USVs from raw recordings by hand is a slow and resource-intensive task. However, software packages developed for automated extraction and characterization of USVs that are currently available are difficult to use and are confounded by the presence of non-vocalization noise. To address these issues, an algorithm employing hidden Markov models (HMMs) has been developed by Dr. Smith of the University of Puget Sound in conjunction with collaborators from OHSU and UC Santa Cruz. The algorithm provides enhanced call-identification accuracy by recognizing and excluding time intervals with significant non-USV noise. Unfortunately, mouse vocalizations that occur in these time intervals, which could be readily identified by manual processing, are excluded. Thus, an algorithm to identify USV in the presence of non-USV noise might complement the HMM-based approach developed by Dr. Smith. To this end, several filtering algorithms inspired by the Sobel Edge detection method were developed and tested. Convolution of spectrographic data with a best fit Sobel matrix generated from manual annotations achieved some discrimination between true-call signal and noise signal. However, a flexible generic double-edge-detecting matrix proved to be the most successful filter. It is hoped that the Sobel-inspired methods might be employed in conjunction with other USV-detection schemes to provide biological researchers a more complete and reliable insight into the vocalizations of their experimental subjects.

Background

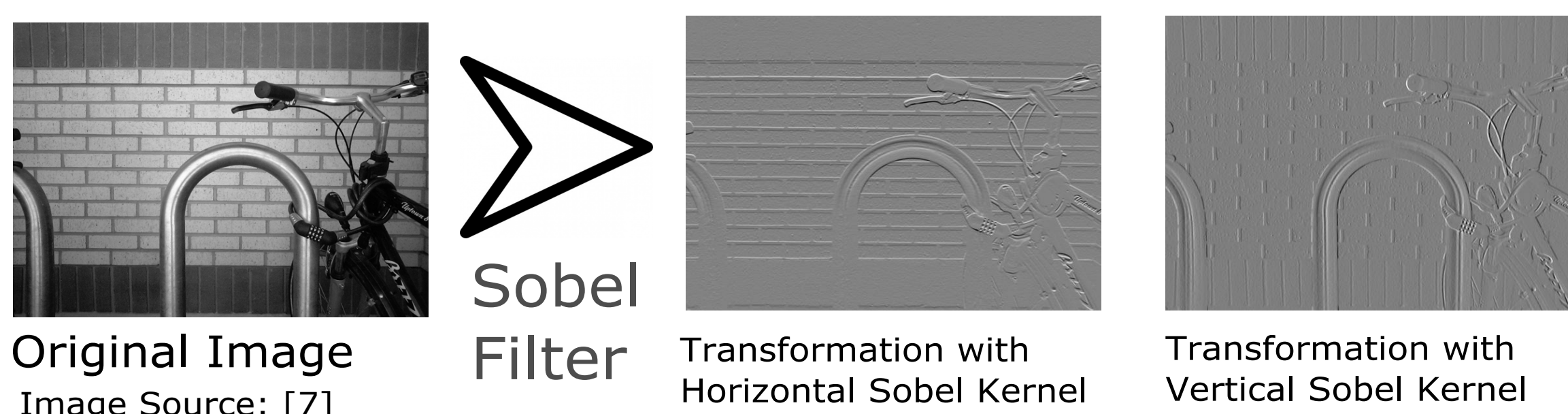
What is the Sobel filter?

+ The Sobel operation is a common tool in computer vision that was developed by Irwin Sobel and Gary Feldman in 1968 ([5], [6]). The filter is designed to highlight edges within images and shows sensitivity to edge orientation (i.e. horizontal versus vertical).

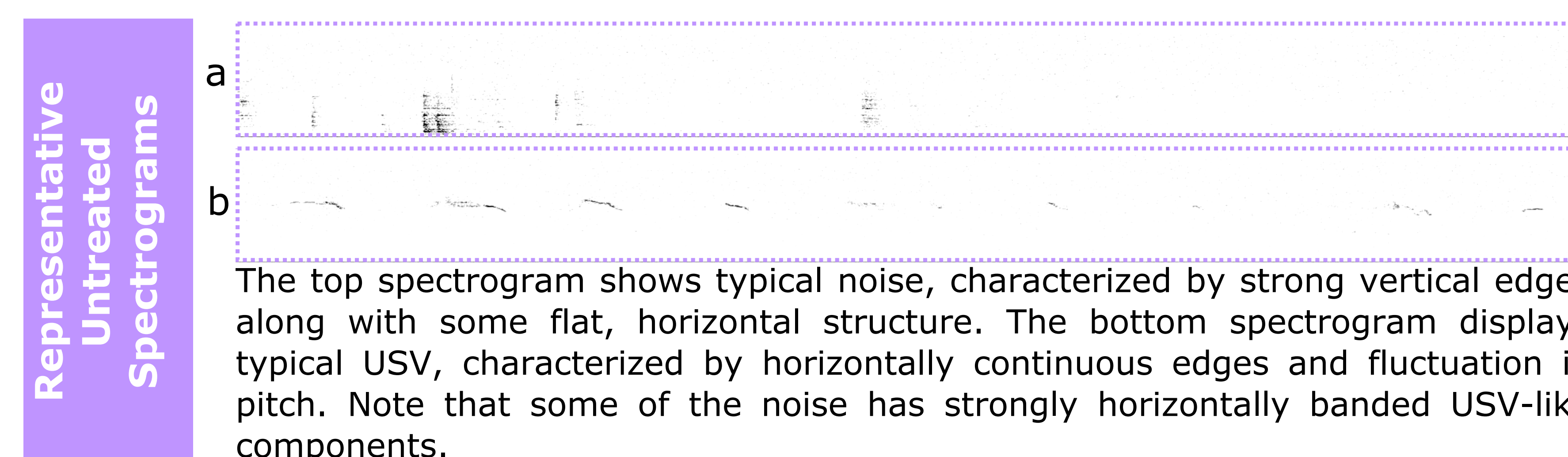
+ Traditionally, the Sobel edge detection method processes an image by applying a specific 3x3 kernel matrix to a two-dimensional matrix representation of an image by a convolution operation.

+ In a convolution operation, a comparatively small kernel matrix is swept along a stepwise path over the entire image; discounting the image edges, each position of the kernel matrix corresponds to a single pixel in the original image.

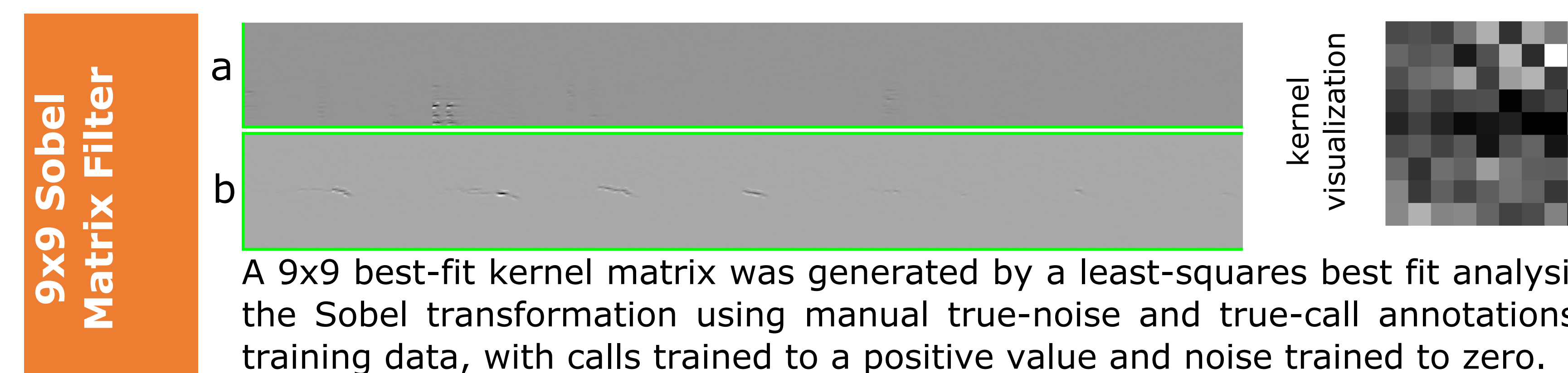
+ At each pixel position, a simple sum of products of the superimposed image and kernel matrix values is calculated. This scalar value, which essentially represents the level of geometric similarity between the compared matrix regions, is represented as intensity in the transformed image.



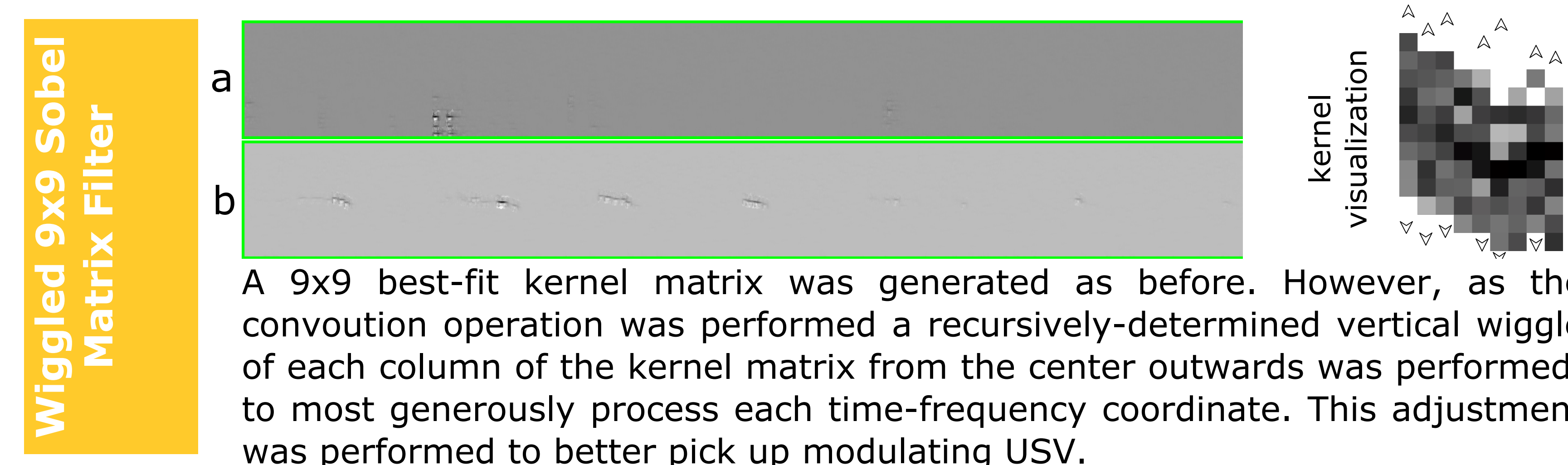
Results



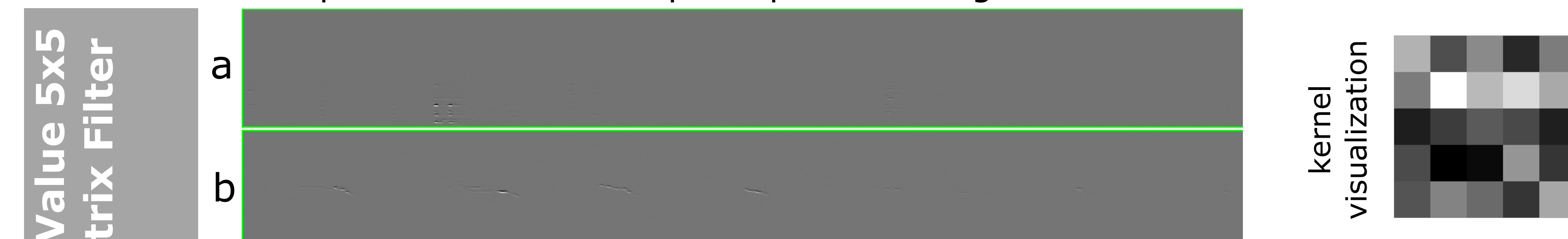
The top spectrogram shows typical noise, characterized by strong vertical edges along with some flat, horizontal structure. The bottom spectrogram displays typical USV, characterized by horizontally continuous edges and fluctuation in pitch. Note that some of the noise has strongly horizontally banded USV-like components.



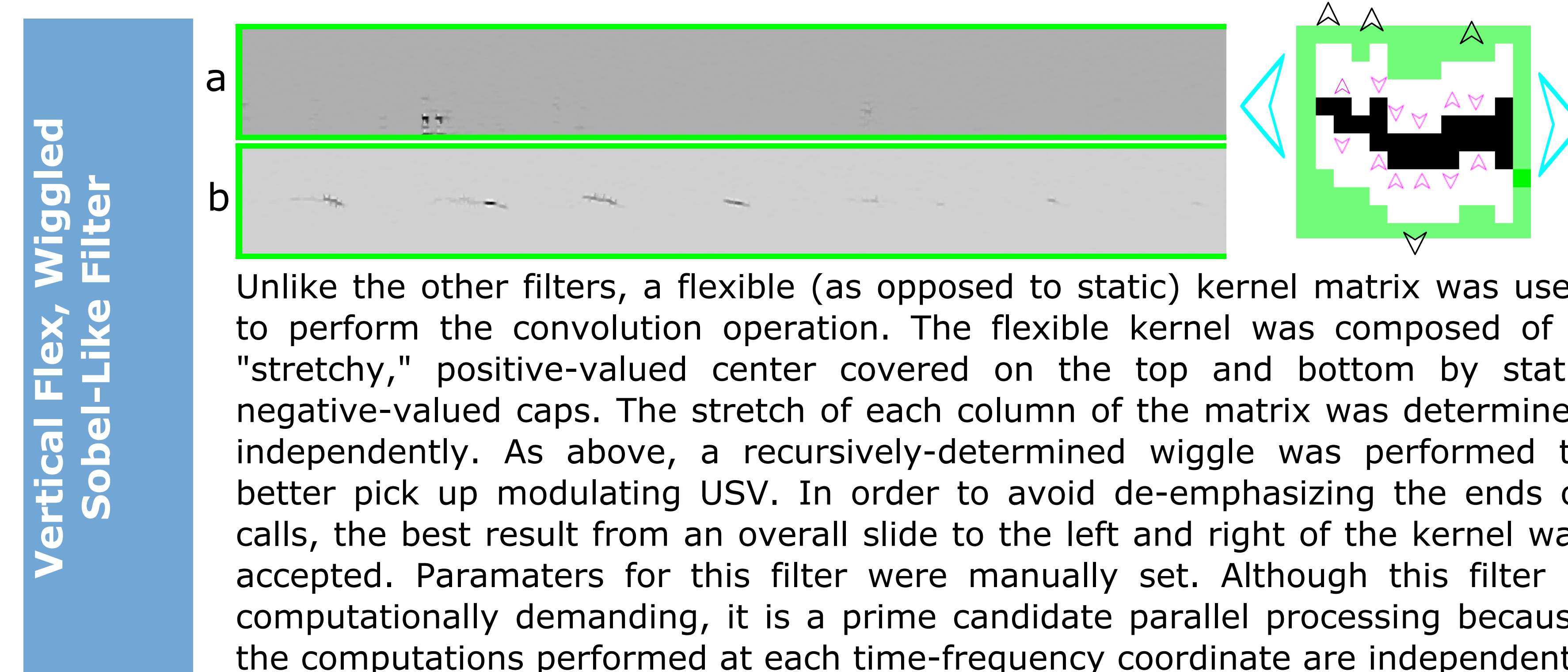
A 9x9 best-fit kernel matrix was generated by a least-squares best fit analysis of the Sobel transformation using manual true-noise and true-call annotations as training data, with calls trained to a positive value and noise trained to zero.



A 9x9 best-fit kernel matrix was generated as before. However, as the convolution operation was performed a recursively-determined vertical wiggle of each column of the kernel matrix from the center outwards was performed to most generously process each time-frequency coordinate. This adjustment was performed to better pick up modulating USV.



A 5x5 best-fit kernel matrix was generated by a least-squares best fit analysis of the sobel transformation using manual annotations as training data. Time-frequency coordinates on the "bottom" of calls were trained to a negative value, those on the "top" of calls to a positive value, and noise time-frequency coordinates were trained to zero.



Unlike the other filters, a flexible (as opposed to static) kernel matrix was used to perform the convolution operation. The flexible kernel was composed of a "stretchy," positive-valued center covered on the top and bottom by static negative-valued caps. The stretch of each column of the matrix was determined independently. As above, a recursively-determined wiggle was performed to better pick up modulating USV. In order to avoid de-emphasizing the ends of calls, the best result from an overall slide to the left and right of the kernel was accepted. Parameters for this filter were manually set. Although this filter is computationally demanding, it is a prime candidate for parallel processing because the computations performed at each time-frequency coordinate are independent.

Background (cont.)

Why Mouse Ultrasonic Vocalizations (USVs)?

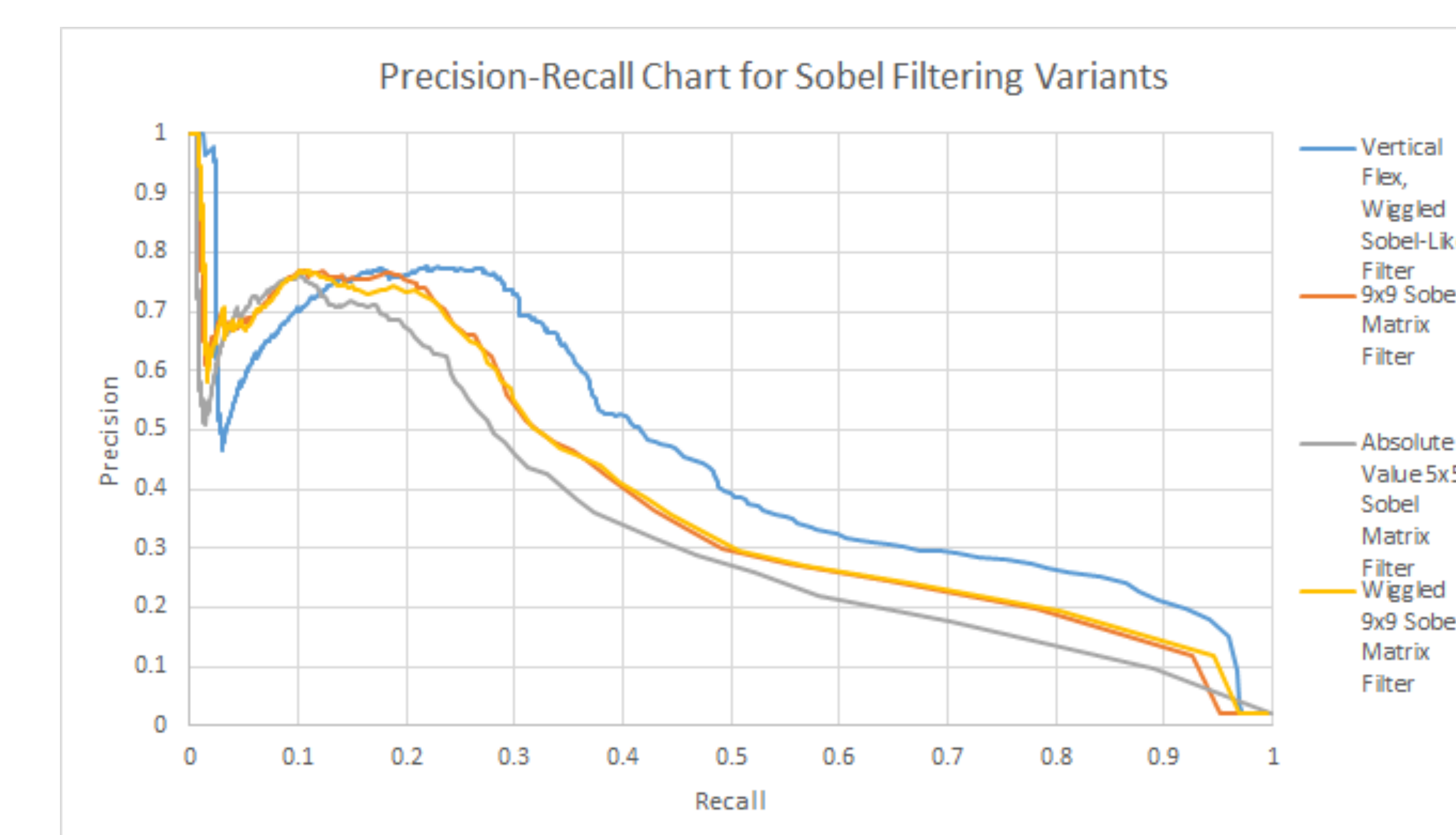
+ Mice serve as a model organism in a wide variety of biological research projects because they are economical, easy to work with, well-characterized, and share many biological similarities with humans.

+ Mouse models are used to study autism spectrum disorders, the phenomenal regulation of sexual courtship and the characterization of the cholinergic system require a metric to characterize the affective and social behaviors of mice ([1], [2], [3], [4]).

+ Mice are known to emit USVs in social, sexual, and infantile contexts ([1]).

+ USVs provide a quantifiable assessment of the the social and affective characteristics of mice. In particular, USV frequency, length, or modulation pattern can be of experimental interest.

Results (cont.)



The Sobel-like vertical flex, wiggled filter displays overall superior precision compared to other variants over nearly the entire range of relevant recalls.

Acknowledgement

I am grateful the University of Puget Sound Summer Research grant panel for funding the project in conjunction with the NASA Research Scholar program. Thanks also to Garet Lahvis of OHSU for providing mouse USV data to test and evaluate filtering algorithms. Most of all, I would like to thank my mentor Adam Smith. His patience and pragmatic advice is much appreciated.

Works Cited

- [1] M. Wöhr and R. K. W. Schwarting, "Affective communication in rodents: ultrasonic vocalizations as a tool for research on emotion and motivation," *Cell and Tissue Research*, vol. 354, no. 1, pp. 81–97, Oct. 2013.
- [2] J. Nyby, C. J. Wysocki, G. Whitney, and G. Dizinno, "Pheromonal regulation of male mouse ultrasonic courtship (*Mus musculus*)," *Anim Behav*, vol. 25, no. 2, pp. 333–341, May 1977.
- [3] I. Branchi, P. Campolongo, and E. Alleva, "Scopolamine effects on ultrasonic vocalization emission and behavior in the neonatal mouse," *Behav. Brain Res.*, vol. 151, no. 1–2, pp. 9–16, May 2004.
- [4] M. L. Scattoni, J. Crawley, and L. Ricceri, "Ultrasonic vocalizations: a tool for behavioural phenotyping of mouse models of neurodevelopmental disorders," *Neurosci Biobehav Rev*, vol. 33, no. 4, pp. 508–515, Apr. 2009.
- [5] R. Fischer, S. Perkins, A. Walker, and A. Wolfart, "Feature Detectors - Sobel Edge Detector." [Online]. Available: <http://homepages.inf.ed.ac.uk/rbf/HIPR2/sobel.htm>. [Accessed: 27-Feb-2015].
- [6] S. Gong and S. Gong, "Pixel Classification Algorithms for Noise Removal and Signal Preservation in Low Pass Filtering," *masters*, Concordia University, 2013.
- [7] Wikipedia contributors, "Sobel operator," *Wikipedia, The Free Encyclopedia*, 16 August 2015, 04:31 UTC, <https://en.wikipedia.org/w/index.php?title=Sobel_operator&oldid=676310685> [accessed 31 August 2015].