What Was That Bird? Birdsong Query-by-Humming Using Asymmetric Set Inclusion of Pitch-Curve Segments

Kendra Oudyk[*1], and Michael Cochez[2,3,4]

[*]Corresponding Author: kendra.oudyk@gmail.com
[1]Department of Music, Art, and Culture Studies, University of Jyväskylä, Jyväskylä, Finland
[2]Fraunhofer Institute for Applied Information Technology FIT, Augustin, Germany
[3]Information Systems and Database Technology, RWTH Aachen University, Aachen, Germany
[4]Faculty of Information Technology, University of Jyväskylä, Jyväskylä, Finland

Birdsong query-by-humming presents a challenging classification problem. Although advances have been made in classifying birdsong queries (e.g., Rassak & Nachamai, 2016), little work has been done in classifying human imitations of birdsong. The physiological differences between birds and humans give rise to different vocal abilities (Marler & Slabbekoorn, 2004), and humans are incapable of imitating some features of bird vocalizations. Additionally, bird vocalizations may vary both between and within individual birds of a species (Sakata, Hampton, & Brainard, 2008). Therefore, birdsong query-by-humming must consider variable queries and variable birdsong examples; thus it may not be appropriate to adapt methods used for birdsong classification and for music query-by-humming. The purpose of this methodological project is to propose a novel algorithm for birdsong query-by-humming. In this algorithm, similarity between a human imitation (i.e., the query) and example birdsong from different species is measured with asymmetric set inclusion. The sets are the segmented pitch curves of the query (Q) and of the birdsong examples for a given species (Es). The similarity between Q and Es is taken as the number of shared segments over the total number of query segments. A segment is considered 'shared' if the Q-segment is correlated above a threshold to an Es-segment, after dynamic time warping to account for small temporal variations. Thus, the exact version of this method involves pairwise comparisons of every segment in the Q and E sets. This exact method is currently being tested, and a time-saving approximation of the exact method may be implemented using a locality-sensitive hashing forest to search for candidate Q-E segment pairs to be compared. This project has potential methodological implications for classifying time-series data, as well as practical implications for birdsong query-by-humming technology. Additionally, it may further our understanding of how humans imitate the sounds of other species.

Marler, P. R., & Slabbekoorn, H. (2004). *Nature's music: The science of birdsong*. San Diego, CA: Elsevier Academic Press.

Rassak, S., & Nachamai, M. (2016). Survey study on the methods of bird vocalization classification. In *Current Trends in Advanced Computing (ICCTAC), IEEE International Conference on* (pp. 1-8). IEEE. doi:10.1109/ICCTAC.2016.7567337

Sakata, J. T., Hampton, C. M., & Brainard, M. S. (2008). Social modulation of sequence and syllable variability in adult birdsong. *Journal of neurophysiology*, *99*(4), 1700-1711. doi:10.1152/jn.01296.2007