Constructing a personalizable gesture-recognizer infrastructure for the K-Bow

R. Fiebrink¹, M. Schedel², B. Threw³
¹ Princeton University
² Stony Brook University
³ Keith McMillen Instruments

Keith McMillen Instruments sells a series of properly weighted and balanced Bluetooth sensor bows for violin, viola, cello and bass—the K-Bow. The K-Bow has eight sensors, the values of which are communicated via OSC: x-, y-, and z-axis accelerometers, hair tension, grip pressure, bow location along frog to tip and bridge to fingerboard, and tilt. We have been constructing a software system to perform recognition of traditional and non-traditional bowing gestures using the K-Bow. In constructing a supervised-learning-based gesture recognizer, we have employed both a traditional supervised learning workflow (i.e., constructing a gesture database and optimizing model performance on that database) and a less traditional, interactive workflow in which we iteratively add training examples, re-train, subjectively evaluate, and modify the gesture recognizer in real-time. Notably, while a K-Bow performer (our end user) may run our own trained gesture recognizer “out of the box,” he or she may also interact explicitly with the learning system in order to easily create, personalize and augment a library of known motions.

We will describe the trajectory of our experience working together as a software engineer, a composer, and a music technology researcher to implement a robust system for interactively re-trainable gesture recognition, using the standardized K-Bow hardware and building on the Wekinator system for real-time, interactive learning. In our experiences, both off-line and real-time machine learning are integral to constructing a robust and flexible gesture recognition system, and appropriate interfaces for these tasks can aid designers and users alike. We will situate this work with respect to prior research on bowing gesture recognition. We will describe the infrastructure built for our own work, including a method for systematically evaluating features computed from the bow sensor outputs, as well as a real-time gesture recognizer prototyping interface optimized for a string player who uses both hands to play. We will discuss our goal of providing the end user (composer or performer) with an interface for modifying the learned gesture recognition model according to her musical expertise and intuition, while not requiring prior machine learning expertise; we will offer an evaluation of how our current user interface supports that goal. Informed by our experiences with the K-Bow, we will then discuss how applying machine learning for gesture recognition differs from more traditional machine learning applications in interesting and non-trivial ways. For example, we may be interested in multiple overlapping labels (e.g., up bow or down bow, fast or slow bow, marcato or détaché); we are also ultimately concerned with not only acquiring labels for gestures, but understanding the degree of confidence the model has that a given label is correct. Importantly, the ability of the performer to add more examples (rather than change the learning algorithm) to improve labeling accuracy distinguishes this task from more traditional machine learning applications.
Our presentation will conclude with a live demonstration of a performer using the K-Bow with a pre-trained gesture recognizer, modifying the gesture recognizer to accommodate her personal style, and adding a new gesture to the recognizer’s vocabulary.