



MAESTRO: Conducting Simulation

Team Members: Christopher Baur, Wendi Ding, Eleanor Epskamp-Hunt, Prakash Kumar, Dana Nehls, Stuti Haresh Patel, Madeleine Wilson, Linying Xie

Faculty Mentor: Dr. Gregory H. Wakefield, Electrical Engineering and Computer Science

Faculty Sponsor: Dr. Andrea Brown, Department of Conducting

CONTEXT AND BACKGROUND

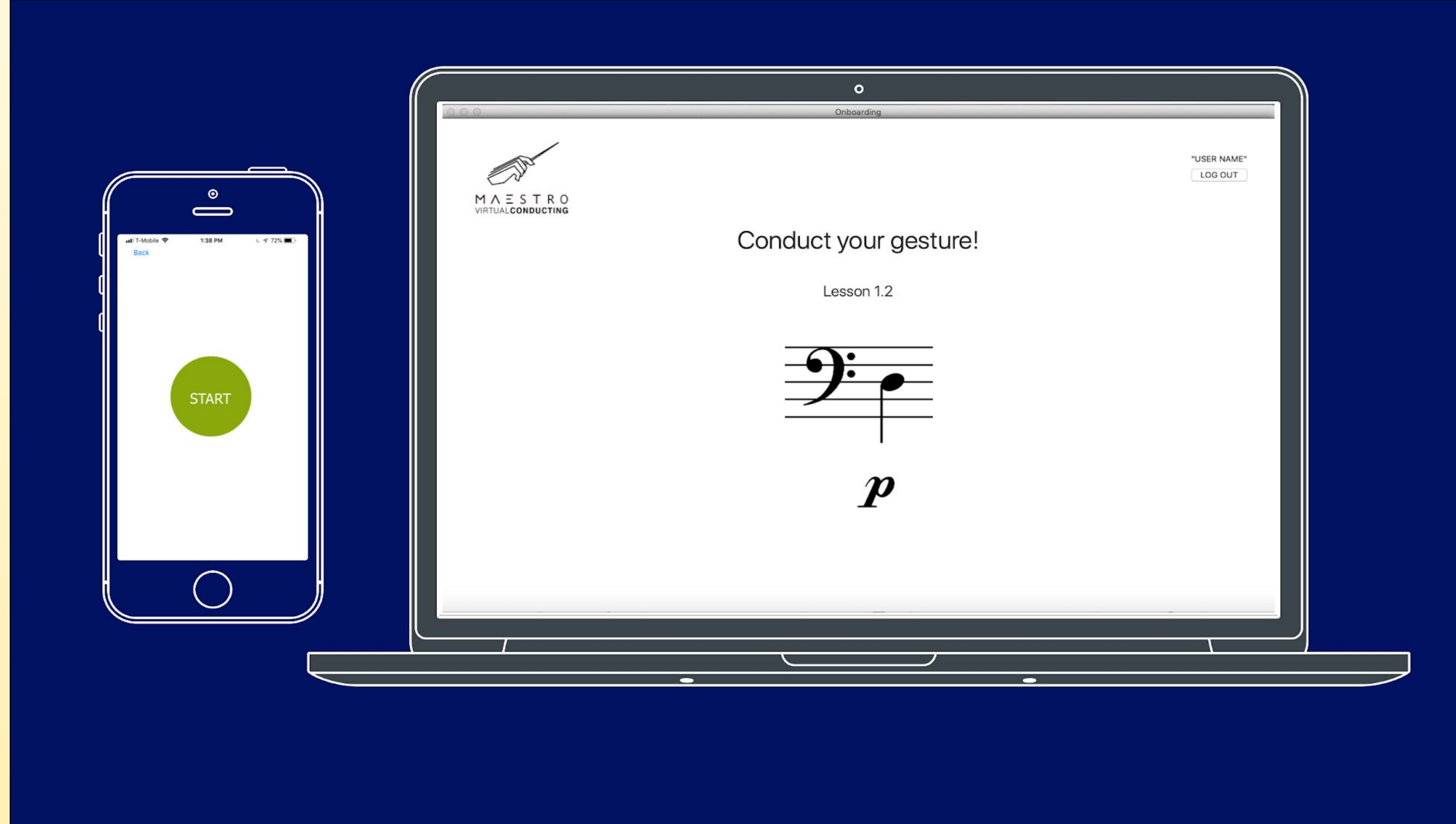


The use of technology in music conductor training is a growing area of interest for faculty at schools of music across the country. Our sponsor, Dr. Andrea Brown, also shares this interest. She was the Assistant Director of Marching and Athletic Bands at the University of Michigan from 2013 to 2018, and now holds a director position at the University of Maryland. Her research on conducting pedagogy has resulted in the virtual conducting system we call *Maestro*.

Traditionally, new conductors attend classes where an instructor watches them conduct and provides feedback. To practice outside the classroom, they either practice silently in front of a mirror, or must recruit the help of musicians. With *Maestro*, conductors practice with a computer-simulated ensemble. This allows beginning conducting students to practice their art with ease; it eliminates reliance on live musicians and any stress associated with standing in front of a crowd.

The success of this project is expected to impact conducting education in music schools throughout the country. It establishes our sponsor as a leader in the intersection of technology and conducting pedagogy. As a tool for assessing and refining student performance, the integration of *Maestro* in the classroom has the potential to revolutionize the standard conducting curriculum.

FINAL DESIGN



The *Maestro* system integrates an iPhone application and a Mac application to give users a virtual conducting experience using devices they may already own. Information about a user's gesture is sent via Bluetooth from phone to computer. Then the gesture is classified using team-developed algorithms and a corresponding auditory response is synthesized in real-time. A simple interface on the phone allows conductors to advance through a variety of lessons. They can conduct at various dynamics, articulations, and tempos and get critical feedback from the system both visually and aurally.

VALIDATION METHODOLOGIES AND RESULTS

Single Source Gesture Detection

Can the system accurately classify the dynamics and articulation of the user's gesture?

To get a realistic expectation for the performance of the system, we surveyed 20 individuals with ensemble music experience. Each user watched 48 recorded gestures and classified their perceived dynamic level and articulation. We ran similar tests with our student-developed algorithms. The accuracy from both the respondents and *Maestro* are summarized below:

Dynamic Classification

	User	System
Piano	40.48%	68.75%
Mezzo Piano	34.92%	93.75%
Mezzo Forte	57.94%	87.5%
Forte	46.34%	93.75%
Fortissimo	46.34%	100%
Average	45.78%	86.25%
Kappa*	0.327	0.959

Articulation Classification

	User	System
Staccato	80.95%	68.75%
Standard	66.07%	31.25%
Legato	61.31%	56.25%
Average	69.44%	52.08%

*Kappa is a measure of how well the system performed compared to how well it would have performed due to chance on a scale from 0 to 1. A higher value indicates that it is less likely that our results are due to chance.

The dynamic classification of the system is reliably better than individuals with ensemble experience, but the articulation classification of the system is slightly worse than these individuals.

Tempo Detection

Can the system accurately detect the tempo of the user's gesture?

To validate the tempo detection of a 4/4 pattern, we sampled 16 test files from five conducting students and one non-conducting musician. Each file contains a varying number of measures of a 4/4 pattern, for a total of 74 measures conducted. The accuracy of the algorithm is summarized below:

Total downbeats detected: 96.1%

Total beats detected: 93.2%

The beats in a user's gesture are accurately detected by the system, meaning tempo calculations based on the detected beats are reliable.

Timely Audio Feedback

Can the system promptly produce audio feedback?

By conducting our usability tests, we observed that the audio feedback is almost instantaneous as compared to the user's gesture. For this reason, we did not conduct our previously designed test methodology for this particular requirement.

Quality Audio Feedback

Can the system produce audio feedback appropriate to the user's gesture?

To validate the quality of the audio feedback, we administered a test to determine students' ability to distinguish between standard, legato, and staccato articulation types using audio clips of real instruments and clips generated by our sound synthesis engine. The accuracy of users in both tests is displayed here:

Articulation Classification

	Real Sounds	Generated Sounds
Staccato	80.5%	94.4%
Standard	29.4%	38.8%
Legato	N/A	66.6%

Students more accurately detected the articulation of generated sounds versus real sounds.

Intuitive and Comprehensible System

Are test subjects able to navigate the system with ease?

To validate the user experience, we evaluated 1) the ease of preparing the system by asking participants to follow the onboarding instructions provided by the system, 2) the ease of navigability through Lesson mode by asking participants to navigate to and describe their understanding of the Lesson screen, and 3) the discoverability of gestural interactions within Lesson mode by instructing participants to discover and express gestures of different dynamic levels, dynamic changes, and release types.

The evaluation resulted in a 100% task completion rate for all 3 tasks across all participants, meaning the system is navigable by the naive user.

STAKEHOLDER OBJECTIVES

Gesture Detection

Defines the success of the system in recognizing the type of gesture that the user conducts.

- Can the system accurately classify the dynamics (volume) and articulation (style) of the user's gesture?
- Can the system accurately detect the tempo (speed) of the user's gesture?

Sound Production

Defines the success of the system in producing an audio response to the user's gesture.

- Can the system produce audio feedback appropriate to the user's gesture?
- Can the system promptly produce audio feedback?

Usability

Describes the success of the system in being accessible to a naive user.

- Are test subjects able to navigate the system with ease?

CONCLUSIONS AND RECOMMENDATIONS

Conclusions

- 1) The new *Maestro* algorithm is able to accurately detect action point, dynamics, articulation, and tempo. It is able to detect dynamics at a far better rate than the average musician.
- 2) The system can promptly produce appropriate audio feedback when a user conducts various gestures.
- 3) The system is intuitive and easy to use, utilizing the Apple ecosystem that many music students are used to.
- 4) Lesson mode helps guide conducting students through structured learning, useful for potential curriculum-based integration.

Recommendations

- 1) Improve accuracy of articulation detection algorithm.
- 2) Add lessons for more tempo patterns, including 3/4 and 2/4.
- 3) Give audio feedback for tempo lessons through songs or MIDI files of multiple notes.
- 4) Add visual feedback using live video or a plot of the gesture performed.
- 5) Implement tracking statistics in Lesson mode to provide user with further information on how the algorithms interpreted their gestures.
- 6) Implement crescendo and decrescendo gesture detection.
- 7) Add the ability to save user data to an account for future review.

