

# Maestro Conducting Project

Charu Dwivedi, Da Fang, William Kacsur, Emily Kirven, Fidelia Lam, Nilay Muchhala, Daphna Raz, J. Nick Smith  
Faculty Mentor: Gregory H. Wakefield, Electrical Engineering and Computer Science  
Faculty Sponsor: Andrea Brown, Department of Conducting

## CONTEXT AND BACKGROUND



Among the variety of studies that are ongoing around the country in various School's of Music, research in conducting gesture and pedagogy remains one of the most popular. In 2011, Dr. Andrea Brown, currently Assistant Director of Marching and Athletic Bands at the University of Michigan, began a study into a technology-based elementary conducting application designed to aid students in practicing various conducting lessons outside of the classroom laboratory and without the need for live musicians. Born at Georgia Institute of Technology, this technology (*Maestro 1.0*) provided the first of its kind in gesture recognition, but proved somewhat unreliable and one-dimensional. Dr. Brown, sponsor of this Multidisciplinary Design Project, secured funds from the TLTC Quick Wins Discovery Project, which allowed for the continued development of what has become *Maestro 2.0*, a system that provides a more accurate and user-friendly experience to conducting students.

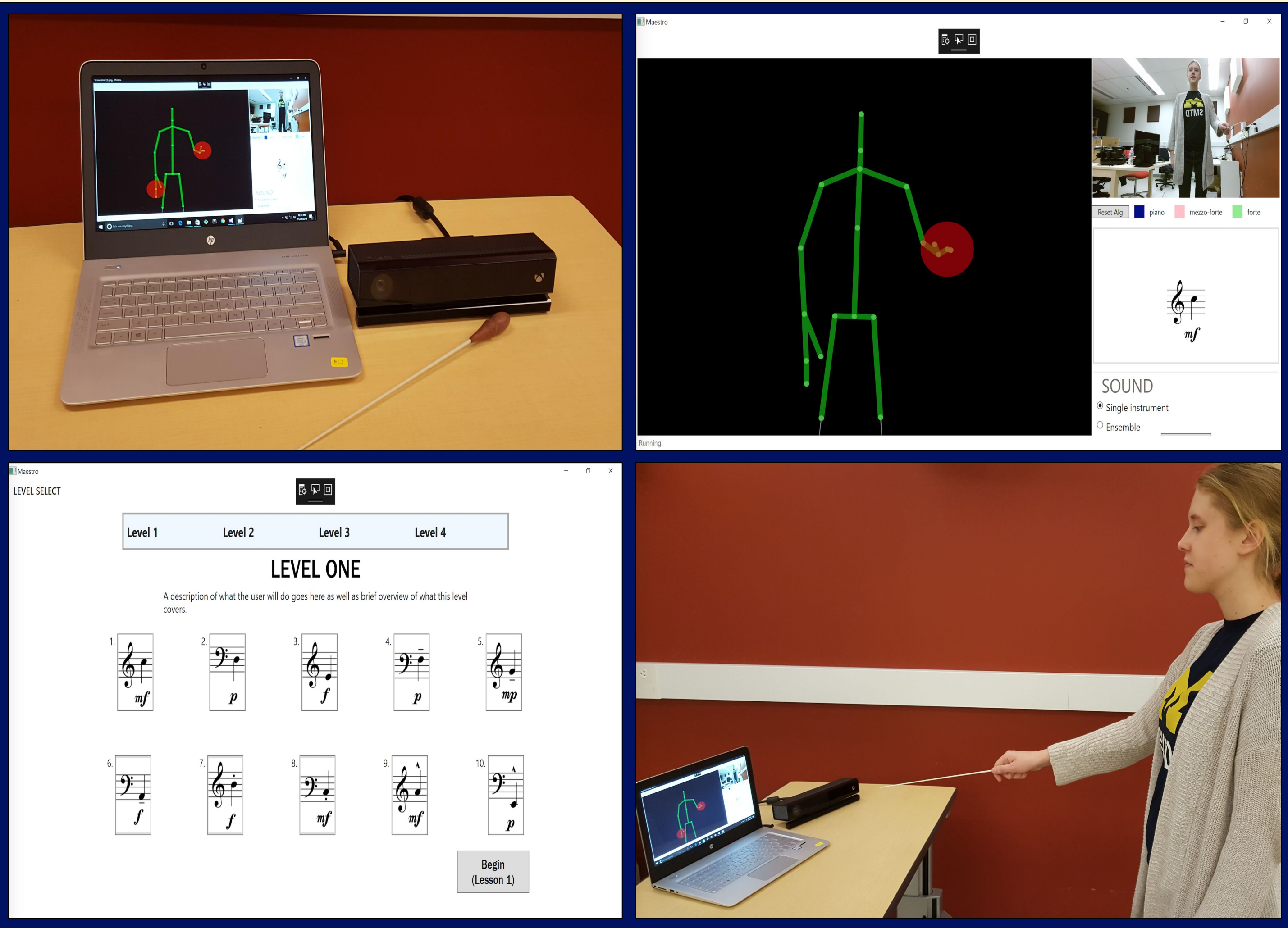
Work on Maestro 2.0 began in January 2016 with an investigation into the various sensors that could be used to acquire conducting gesture data. IMU sensors, computer vision using OpenCV, and the Microsoft Kinect were all contenders in the search for the best fit to this project. Ultimately, the Microsoft Kinect was chosen because of its ease of use, both for the developer and the user, especially considering its intent to be used at the collegiate educational level. Because of its potentially much better bandwidth, a sub-team was also tasked with pursuing IMU sensing as a future technology for integration into Maestro.

## STAKEHOLDER OBJECTIVES AND ENGINEERING REQUIREMENTS

Stakeholder Requirement	Relative Priority	Specification	Measurement Methodology
Accurately detect the action point of a gesture	1	<b>Success rate of 80% or higher</b>	Calculate success rate of detection based on a sample consisting of Dr. Brown's COND 315 students.
Accurately detect gestures across multiple subjects	1	<i>Success rate of 80% or higher</i>	Calculate success rate of detection based on a sample consisting of Dr. Brown's COND 315 students.
Responsive audio feedback	2	<b>System response time of 30ms or less on average</b>	Run multiple tests of our device using Emily and Nick and have them evaluate the responsiveness
Attractive audio feedback	2	<u>At least 75% respond with "attractive"</u>	Survey Dr. Brown's COND 315 class: attractive / not attractive
Intuitive and simple-to-use UI	3	<b>Average of 3.5 on Likert scale</b>	Survey Dr. Brown's COND 315 class using a Likert scale of 1-5 on intuitiveness and ease of use of UI

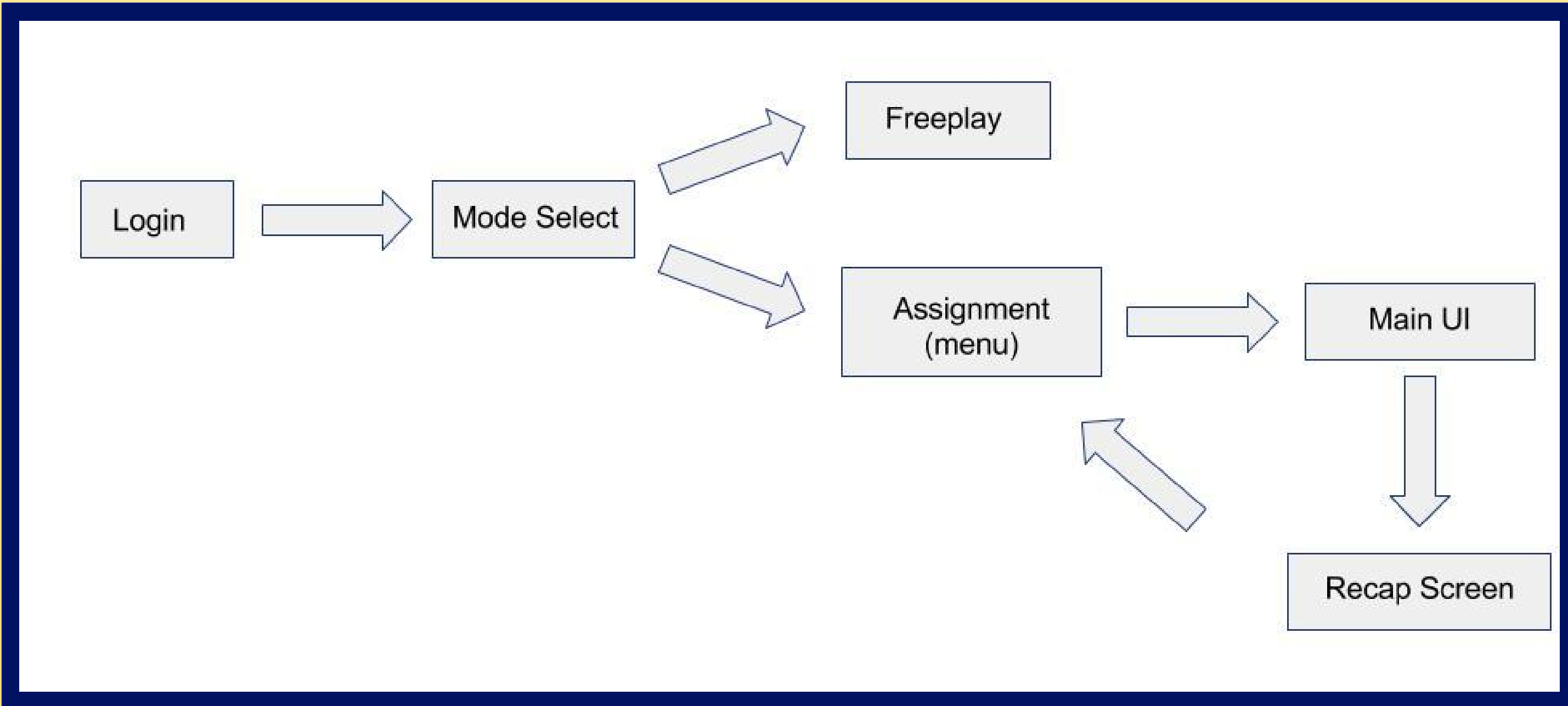
**Bold:** Goal met  
*Italicized:* Goal Not met  
Underlined: Goal very close to being met

## FINAL DESIGN



The Maestro 2.0 system integrates the Microsoft Kinect with an intelligently designed user interface to offer any user a one-of-a-kind gestural analysis experience. A conductor may choose to follow a series of prescribed lessons or experiment with the "free play mode." In both cases, the audio and visual feedback provides necessary information and suggested next steps for balanced and structured improvement.

## TYPICAL USAGE CASE/ PROCESS FLOW



The Maestro 2.0 system is a user based system--users with verified accounts can login and select either a freeplay mode or assignment mode. In either mode, the main UI is prompted (seen in Final Design). In assignment mode, when users are finished, they are redirected to a recap screen that helps them reflect on the last lesson. They can then continue with Maestro, or logout.

## VALIDATION METHODOLOGIES AND RESULTS

Our Validation Experiments comprised three different aspects:

- 1) **Validation of Action Point Accuracy**
- 2) **Validation of Dynamic Classification**
- 3) **Validation of User Interface**

To validate these three desired attributes of our product, we conducted a test with seven conducting students from the School of Music, Theatre, and Dance. Two tests were conducted. The first test evaluated the "Freeplay" mode by instructing participants to express different dynamic level (piano, mezzo-piano, mezzo-forte, forte, and fortissimo). Each gesture was judged independently by an expert (a graduate student in conducting). Each dynamic level was produced five times and Maestro/expert scores were recorded.

Results for Maestro are reported below as average classification across trials and participants. In all cases, the conducting students were judged by the expert to have properly expressed the appropriate dynamic level.

The results are summarized below:

- 90% of Action Points were correctly detected
- 52% of Piano gestures were correctly identified
- 68% of Mezzo-Piano gestures were correctly identified
- 60% of Mezzo-Forte gestures were correctly identified
- 36% of Forte gestures were correctly identified
- 32% of Fortissimo gestures were correctly identified

To validate the User Interface, participants were asked to complete Lesson 4 by completing the following instructions:

- 1) **Log in**
- 2) **Complete Lesson 4**
- 3) **Log out**

After each participant completed the lesson, he or she was asked to complete a survey that measured their reactions to the UI, using a 5-pt. Likert scale (1 - strongly disagree, 3 - neutral, and 5 - strong agree). The average scores over the seven participants are summarized below:

- It was easy to log in: 4.167**
- The system was easy to navigate: 3.333**
- I understand how to complete the assignment: 2.667**
- The main UI (with the skeleton) displayed relevant information: 4.167**
- Calibration was clear: 2.167**
- It was easy to logout: 5**
- Average Score: 4.3002**

## CONCLUSIONS AND RECOMMENDATIONS

### Conclusion

- 1) The new Maestro **algorithm** is able to:
  - Detect the action point (exact point where the beat occurs) of each gesture 90% of the time
  - Can partially classify the expression of dynamics - piano, mezzo piano, mezzo forte, forte, or fortissimo
- 2) Our **user interface** is:
  - Working towards being intuitive and easy to use for users, and has been thoroughly tested by our sponsor, conducting specialists, and outside focus groups.
- 3) A separate **IMU** based system has been investigated as a potentially more responsive and accurate detection source
  - This system could improve upon the limited frame-rate of the Kinect
- 4) A basic **sound synthesis** library has been developed to provide aural feedback

### Recommendation

The current system is ready for further user testing and evaluation. Meanwhile, refinements to the system should address classification accuracy and an expansion of types of gestures.

