A GESTURAL INTERFACE FOR ORCHESTRAL CONDUCTING EDUCATION

Lijuan Peng, Dr. David Gerhard
Computer Science, University of Regina, Regina, Saskatchewan, Canada
peng200l@uregina.ca, david.gerhard@uregina.ca

Keywords: computer-based conducting system, drill and practice, gestural interface, pedagogy, visual representation

Abstract: Over the past few years, a number of computer-based orchestral conducting systems have been designed and implemented. However, only a few of them have been developed to help a user learn and practice musical conducting gestures. This paper is intended to address research related to this area. It utilizes an infrared baton and an acceleration sensor to track the standard conducting gestures. The infrared baton is similar to a conducting baton and has little influence on the conducting. A drill and practice instructional strategy has been applied in this gestural interface. Five options are implemented. Once an option is chosen, users must conduct according to the supported conducting gestures. While a student is conducting, his/her gestures are identified and followed by the system using an accurate and relatively simple process. The conducting is interpreted using a few visual items that clearly show a conducting gesture and reveal its quality. In addition, aural representation informs students of beats or errors when eyes are busy.

1 INTRODUCTION

Since the 1980s, various computer-based conducting systems have been developed (Nintendo, 2006) (Satoshi, 1998) to allow a student to conduct a piece of music using a digital system. Most of these systems focus on the act of conducting instead of gestures or education. Visual representation as a straightforward interpretation for a gesture has only been implemented in one system (Guy, 1999). The research described here is intended to present a gestural interface that is designed and implemented for pedagogy. It presents both visual and aural representations for students.

2 RELATED RESEARCH

2.1 Instructional strategies

As computer and electronic instruments spread, computer-based musical systems have been a supplement to traditional teaching approaches (e.g. printed music notation). Several instructional strategies, such as programmed learning and drill and practice, have been used. A system supporting programmed learning presents some questions and gives feedback to students' answer according to expected one. Drill and practice let students do some pre-designed activities repeatedly. (Brandao, 1999)

2.2 Visual representation of musical parameters

Although it is natural for music education systems to provide aural responses, since music is based on hearing, aural responses can interfere with music being used as an exercise or target. Therefore, visual representation is also supported in many music education systems. It is important to note, however, that visual feedback can interfere with the learning of visual aspects of music in the same way.

In an example of visual feedback, *pianoFORTE* (Stephen, 1995), a system for piano education, utilizes different colors and shapes on the original score to show the difference between the performance of teachers and students.

2.3 Computer-based conducting systems

A few current conducting systems have a pedagogical purpose. For example, *Wireless sensor interface and gesture follower* (Frederic, 2007) was designed to find problems in a student's gesture compared to a teacher's gesture.

Various sensors have been used in computer-based conducting systems. Acceleration sensors (Satoshi, 1998) can be equipped on baton-like devices, which may change the weight and balance of the controller. Cameras (Paul, 2004) capturing the front view of a conductor can show a 2-dimensional trajectory of a conductor's motions. Infrared sensors (Guy, 1999) only track the movement of infrared light sources thus avoiding the influence of background or other confounding visual objects. In addition, other sensors, such as the Wii Remote (Nintendo, 2006), have been used for conducting.

3 INSTRUCTIONAL STRATEGIES

The gestural interface presented in this paper is designed for learning and practicing conducting gestures. Currently, drill and practice has been used. Once an option (there are five options in all.) is chosen, students can repeat a certain conducting gesture to practice it. The feedback from the system lets students know the accuracy of the gestures.

4 DESIGN AND IMPLEMENTATION

This gestural interface has five aspects: tracking, analysis, recognition, following, and response.

The implementation is on an iMac personal computer using Max/MSP, Jitter, and Java. Once the system is run, a main window (Figure 1) is shown on the screen. It displays the menu, the conducting window, and the information related to tempo and dynamics. The visual feedback is also displayed on this main window.

For the right hand, there are two modes: the option selection mode and the conducting mode. The "Menu" bar is used to go back to the menu area (option selection mode). In the menu, if students stay on an option, for example 2-Beat, for a period of time, the focus will be moved back to the conducting window (the conducting mode). An infrared baton is used as both a mouse and an infrared light source. Thus, a lot time is saved on the switch between system manipulation and conducting.

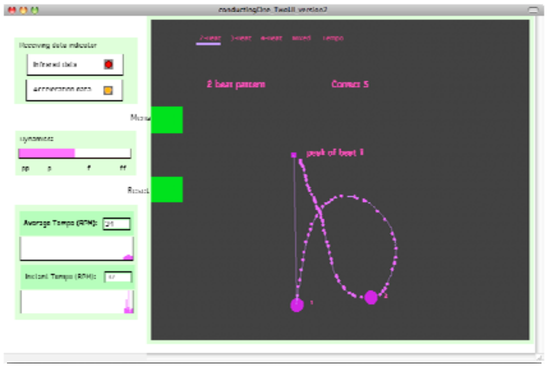

Figure 1: A snapshot of the main window.

4.1 Gestures

When conducting, the movements of both the right hand and the left hand are located in a chest-high virtual rectangle named the conducting window.

For the right hand, this gestural interface focuses on expressive legato gestures (continuous, curved) as shown in Figure 2 (Joseph, 2000) (Brock, 1989). It can be extended to support other beat patterns.

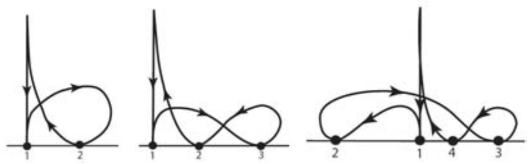

Figure 2: 2-beat pattern, 3-beat pattern, and 4-beat pattern.

For the left hand, the gestures to show dynamics are supported. When the left hand is held with the palm facing up, it means louder. The palm facing out means softer playing.

4.2 Gesture tracking

The system described in this paper uses an infrared sensor because it has higher sensitivity and less computation time compared to systems using a video camera, and data captured by an infrared sensor is easier for visualization compared to those collected by an acceleration sensor.

The infrared baton (Figure 3) used in this gestural interface consists of a conducting baton, an infrared LED (110 degrees viewing angle), a button, and a battery. During conducting, a student holds the infrared baton in the right hand like holding a real conducting baton and presses the baton using his/her thumb. Thus, it is not difficult for a student to learn to use. A Wii Remote (Figure 4) is employed as an infrared camera in front of a student. An acceleration sensor named WiTilt v2.5 (Figure 4) is applied to capture the movement of the left hand.

Figure 3: The infrared baton.

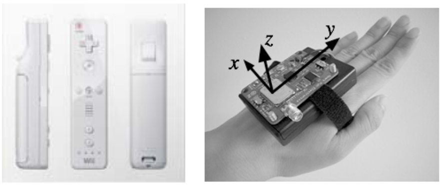

Figure 4: Wii Remote (from (Wii, 2008)) and WiTilt v2.5.

4.3 Gesture analysis

The features for the right hand consist of coordinates and beats. Coordinates of the tip of a baton at each time are used to generate beats, fundamental components of a beat pattern. Figure 2 shows that a beat always occurs at the vertical minimum of a movement. A beat detection algorithm is developed to detect the peaks and troughs of a trajectory.

The results of gesture analysis are displayed as shown in Figure 1. All coordinates are small dots. A trough (beat) is represented with a larger dot. Thus, it is easy for a student to visually differentiate one feature from another. A curve connects these dots and represents the trajectory of the movement. The result is a quantitative interpretation which shows students exactly what their gesture looked like, and can be compared to a reference gesture from a teacher or textbook.

For the left hand, 3-dimensional acceleration data are features and will be used directly in gesture recognition/following.

4.4 Gesture recognition

4.4.1 Task-target gestures

The first gesture recognition is able to tell whether a beat pattern is conducted correctly or not on the basis of an assumption that the beat pattern is known beforehand. It is reasonable because the conductor and performers in an orchestra have the score that indicates the time signature. Currently, it only supports 2-beat, 3-beat, and 4-beat per measure. It is not difficult to extend to other patterns.

Initially, the downbeat is detected as a downward motion based on horizontal coordinate comparison and represented visually by a vertical line. Subsequent beats are then detected according to their coordinate position relative to the downbeat. The recognized result is displayed on the upper left side of the conducting window as shown in Figure 1.

The accuracy of this recognition depends on the downbeat identification and the quality of other beats. If the student performs the gesture correctly, the system recognizes it and rewards the student.

4.4.2 Free-form conducting gestures

The second recognition is a more general one, allowing the student to conduct any beat pattern. The downbeat is detected first, as for the task-target gestures, and then the number of beats is accumulated until the next downbeat is found. The amount of beats reveals the beat pattern performed. Figure 5 shows an example of 12-beat patterns.

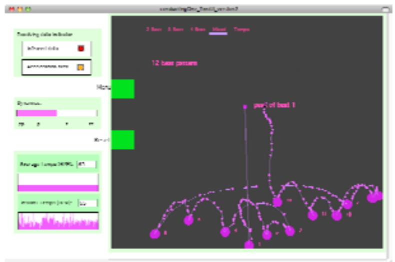

Figure 5: An example of 12-beat pattern.

The downbeat of a subsequent gesture affects the recognition accuracy because an incorrect downbeat does not stop the beat counter. As a result, the number of beats increases until a beat is identified as a downbeat.

4.5 Gesture following

4.5.1 Tempo tracking

Tempo tracking reveals the speed of conducting, enabling a student to practice consistent timing. Each time a beat occurs, the value of the tempo will be calculated using beats per minute (BPM). Both average and instant values are supported. The average value is a moving average and estimated based on the past 10 beats. Therefore, it can follow the changes, especially the significant changes, in tempo quickly and also show the speed trend of conducting clearly. The instant value is intended to show current speed of conducting.

There are two representations for the value of the tempo (Figure 5). The numerical tempo value is a value at a certain time. The diagram clearly shows the changes of the tempo in terms of remaining steady, the increase, and the decrease.

An experiment was performed to demonstrate the accuracy of the tempo tracking system. Individuals were asked to conduct a piece at a specific tempo measured in real time. Results are shown in Table 1.

Table 1: Comparison between the calculated average tempo and the real average tempo.

Amount (2-beat)	Calculated tempo by the system (BPM)	Real tempo by a stopwatch (BPM)
15	115.38	115.68
30	99.45	100.19
45	128.27	128.85

4.5.2 Dynamics tracking

Dynamics tracking is for the left hand and done on the basis of 3-dimensional accelerometer values measuring tilt. As the orientation of the left hand changes, the tilt values change corresponding to the intended change in dynamics. A particular hand position equates to a specific dynamic, and the recognized result is shown using a slider, which is a conventional visual representation for the volume and easy to understand for students.

4.6 Response

In this gestural interface, visual and aural feedback are presented once a gesture is recognized.

Visual representation is intended to present a more direct interpretation to gestures and can be compared to that of a teacher's or a diagram on a textbook. It may be easier for students to adjust and improve their gestures.

Aural representation consists of playing a certain tone corresponding to the recognition of a certain beat. For example, C4 will be played when the downbeat is found. This kind of aural representation gives students the feedback they need while conducting, but does not require them to keep their eyes on the screen. Correct gestures and errors are identified this way. This aural representation also allows professionals and instructors to use the system as they conduct a real orchestra.

5 CONCLUSIONS

This gestural interface aims to help conducting students learn and practice conducting gestures. The Wii Remote is not expensive and easy to acquire. Students do not need to spend time learning how to manipulate the whole system because real conducting gestures are employed with an infrared baton, which is similar to a real baton. Both visual and aural representations are presented to students. The process of gesture recognition and following is simple, fast, and accurate.

REFERENCES

Brandao M., Wiggins G., Pain H., 1999. Computers in music education. *In Proceedings of the AISB'99 Symposium on Musical Creativity*.

Brock McElheran, 1989. *Conducting technique for beginners and professionals revised edition*, Oxford University Press.

Frederic Bevilacqua, Fabrice Guedy, Norbert Schnell, Emmanuel Flety, Nicolas Leroy, 2007. Wireless sensor interface and gesture-follower for music pedagogy. *In Proceedings of the 7th international conference on New interfaces for musical expression*. Pages 124-129.

Guy E. Garnett, Fernando Malvar-Ruiz, Fred Stoltzfus, 1999. Virtual conducting practice environment. *In Proceedings of the International Computer Music Conference*. ICMA. Pages 371-374.

Joseph A. Labuta, 2000. *Basic conducting techniques, fourth edition*, Prentice Hall.

Nintendo, 2006. *Wii Music Orchestra*. http://www.gamespot.com/wii/puzzle/wiimusicorchestra/index.html. Retrieved May 18, 2008.

Paul Kolesnik, 2004. *Conducting gesture recognition, analysis and performance system*. Master's thesis, McGill University.

Satoshi Usa, Yasunori Mochida, 1998. A multi-modal conducting simulator. *In Proceedings of the International Computer Music Conference*. ICMA. Pages 25-32.

Stephen W. Smoliar, John A. Waterworth, Peter R. Kellock, 1995. pianoFORTE: a system for piano education beyond notation literacy. *In MULTIMEDIA'95: Proceedings of the 3rd ACM International Conference on Multimedia*. ACM Press. Pages 457-465.

The Wii Remote, http://www.nintendo.com/wii/what/controllers/remote, Retrieved May 18, 2008.