The Demonstration of MVP Support System as an AR Realtime Pitch Feedback System

Yasumasa Yamaguchi[1], Taku Kawada[2], Toru Nagahama[3] and Tatsuya Horita[3]

[1] Sendai University
[2] Sendai Shirayuri Gakuen Elementary School
[3]Graduate School of Information Sciences, Tohoku University
ys-yamaguchi@sendai-u.ac.jp

Abstract. This demo paper presents and explains the system of MVP (Musical pitch Visualization Perception) support system as an AR (Augmented Reality) real-time pitch feedback system for instrumentalists who must play with correct intonation. The pitch feedback system itself uses a machine learning system called "ml5.js" and utilizes Google Glass as a feedback indicator. The system will assist not only in musical performance as a support system but also in investigating the cognitive process of intonation and musical performance as an experimental application.

Keywords: Realtime Pitch Feedback, ICT, Performance Support, System Development, Augmented Reality

1 Introduction

The proliferation of ICT (Information and Communication Technology) and the advancement of information sciences have brought about innovation in musical performance. Playing a musical instrument can be considered a perceptual-motor skill. Several studies have explored the process of perceptual-motor skill acquisition and learning. Perceptual-motor learning can be divided into three stages: the cognitive stage, associative stage, and autonomous stage [1]. On the other hand, the Acquisition of Cognitive Skill (ACT) theory consists of two stages with one section: the declarative stage, knowledge compilation, and the procedural stage [2]. In these contexts, feedback is one of the most crucial concepts in the realm of perceptual-motor learning [1].
Within the context of musical performance, there has been extensive discussion about intonation. In the fields of cognitive science and musical education, Kreitman defines a listening loop for the intonation of musical performance by instrumentalists [3]. According to Kreitman, the cognitive process of instrumentalists can be classified into four sections. First, the student begins with a concept of the music in their inner ear. Second, their brain sends messages to the body to create actions. Third, these actions produce sound from the instrument. Fourth, the sound enters the ear and is sent to the

brain for analysis [3]. In this process, pitch feedback from a tuner greatly assists in performing with correct pitch and intonation.
Due to the benefits of using a tuner during performances, some instrumentalists tend to use a tuner placed on their music stand to receive real-time pitch feedback. While this action may contribute to good intonation, it can limit their ability to freely gaze at their music sheet, the conductor, and other musicians. Furthermore, this situation can cause a downward head posture, which is considered detrimental for instrumentalists. To address these issues, we have been developing the MVP (Musical pitch Visualization Perception) support system for smart glasses as an AR (Augmented Reality) pitch feedback system [4-5].

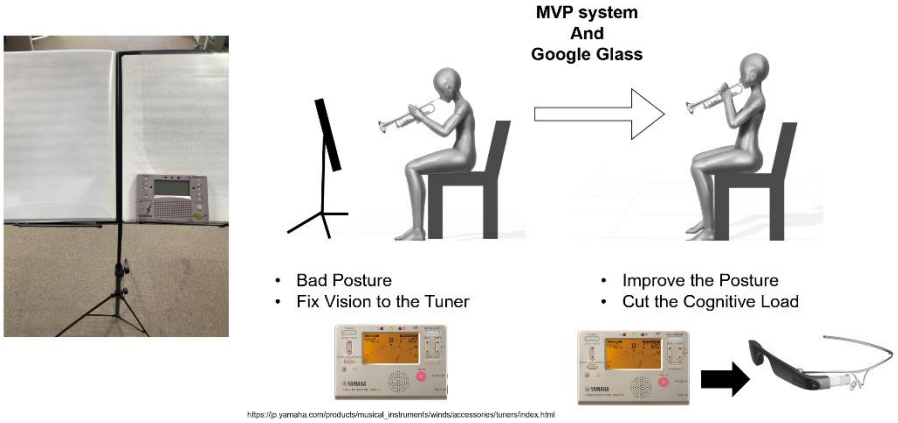

Fig. 1. The picture of the conventional tuner on the music stand and the aim of the MVP support system with a Google Glass

2 Development Background and System

To develop our pitch feedback system, we focused on four key aspects. Firstly, we prioritized timeliness, ensuring that the system can provide real-time pitch feedback. Secondly, we aimed for the feedback indications to be easily understandable and clear. Traditional tuners often use complex scale and needle displays, which can be overwhelming for performers during their play. Therefore, we designed our system to utilize color indicators instead. Thirdly, we incorporated recordability, a feature that sets our system apart from conventional tuners. This aspect enables the system to offer new methods and teaching materials to enhance musical education and improve performance skills. Lastly, we aimed for stability that is independent of the type of ICT equipment and environment used. Our goal is to develop a system that can work on various ICT equipment by leveraging web browsers.
The fundamental structure of our pitch feedback system is depicted in Figure 2. The system captures the sound from the participants' instrument using a condenser microphone connected to a computer. The information is then analyzed, and the pitch is

estimated using the pitch detection system. The estimated value is colorized on a web page based on the browser's capabilities, and the Google Glass displays the colorized webpage for the participant to view.

The computation for determining the tonal pitch is performed using the "ml5.js" library, which runs in TensorFlow [6]. This library incorporates the "Pitch Detection" package, which employs the deep-learning-based "CREPE" algorithm [7-8]. We have implemented this user-friendly system using the TensorFlow backend. In this system, the participant receives a rating based on a three-tier scale: "correct," "higher," or "lower" in comparison to the correct pitch. The feedback is sent to the participant's Smart Glass device. The display shown on the participant's Smart Glass changes interactively based on the received rating. A green display indicates a correct pitch, purple indicates a higher pitch, and blue indicates a lower pitch. This visual feedback helps the participant adjust their pitch accordingly. For our study, we defined the correct pitch range as the target, expressed in Hertz, with a tolerance of ±1%. This range determines the threshold for determining whether the participant's pitch is correct, higher, or lower. We specifically utilized the "Glass Enterprise edition 2" device by Google for this study. It was selected as the platform for delivering the tonal pitch feedback. Notably, we ensured the reliability of the Google Glass system as a musical tuner by verifying its performance with a professional musician.

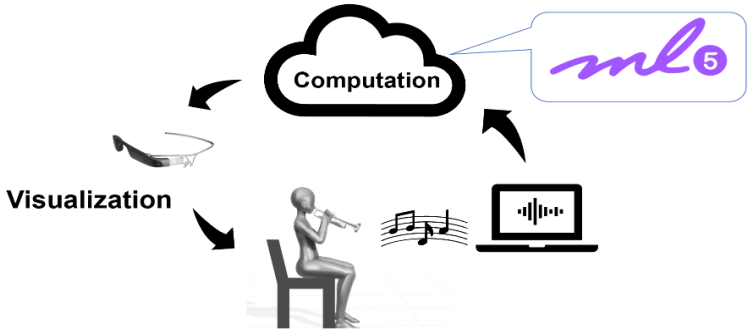

Fig. 2. The Schematic view of the MVP support system

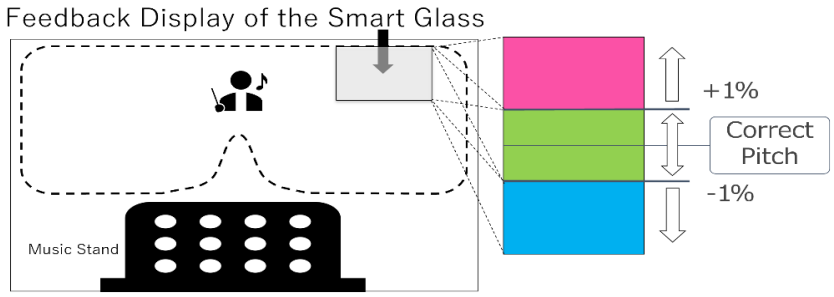

Fig. 3. The sight of user and function of MVP support system

3 Conclusion

3.1 Future Development

In the early stages of development, we initially adopted FFT algorithms; however, the resolution rate and computation time proved insufficient for real-time pitch feedback in musical performance situations. Consequently, we decided to utilize the machine learning package provided by ml5.js, as it proves capable enough to achieve our goals. On the other hand, we encountered an issue with the smart glass. The microphone attached to the glass was not robust enough to accurately capture the tones of musical instruments. It occasionally recognized overtones of the instrument. To address this, we need to incorporate a low-pass or high-pass filter into the system.

3.2 Information

The system's effectiveness was successfully presented as a pilot study [5] at the EdMedia + Innovate Learning 2022 conference in New York, held in June 2022 and further research [4] continued to display this effectiveness. In these experimental evaluations of the system, participants provided positive comments, and the system demonstrated superior performance compared to conventional tuners. These experiments highlight the potential of the browser-based system as an experimental tool for studying musical performance and pitch intonation. This demo paper is a description of the MVP support system utilizing Google Glass in [4-5] for demonstration for CMMR 2023.

References

1. Fitts, P.M.: Perceptual-motor skill learning. In A. W. Melton (Ed.), Categories of human learning, 243–292. New York Academic Press (1964).
2. Anderson, J.R: Acquisition of cognitive skill. Psychological Review, 89,369-406(1982).
3. Kreitman, E., Teaching with an Open Heart. Western Springs School of Talent Education (2010).
4. Yamaguchi, Y., Kawada, T., Nagahama, T., & Horita, T. Development and Evaluation of a Musical Instrument Performance Support System Using Smart Glasses: from the Subjective Evaluation of Form, Gaze, and Performance. *Japan journal of educational technology*, 46(Suppl.). 185-188. (2023).
5. Yamaguchi, Y., Kawada, T., Nagahama, T., & Horita, T. A Pilot Study of the MVP Support System using Google Glass. In EdMedia+ Innovate Learning.17-28. Association for the Advancement of Computing in Education (AACE). (2022, June).
6. NYU. ITP *"ml5js· Friendly Machine Learning for the Web.*" ml5js website. Accessed March 16, 2022. https://ml5js.org/, last accessed 2023/5/4.
7. Mathieu, B., Essid, S., Fillon, T., Prado, J., & Richard, G. *Yaafe, an easy to use and efficient audio feature extraction softwarem* In Ismir. (2010).
8. Kim, J. W., Salamon, J., Li, P., & Bello, J. P. *Crepe: A convolutional representation for pitch estimation.* In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 161-165). IEEE. (2018).