Projection mapping of articulation for education in phonetic science

Takayuki Arai* and Rei Uchida
Sophia University

(Received 28 September 2015, Accepted for publication 28 December 2015)

Keywords: Projection mapping, Articulation, Ultrasonic imaging, Education

PACS number: 43.10.Sv [doi:10.1250/ast.37.175]

1. Introduction

Projection mapping (PM) is widely applied to many fields to “see through” to the inside or to the other side of objects. In medical fields, the PM technique allows us to see inside a patient’s organs from projected images on the surface of the patient’s body [e.g., 1]. This technique must be useful for the fields of speech and phonetic sciences, as well, since the articulators and vocal tract within the human head are not visible from outside. Therefore, in this study, we propose and discuss how PM can allow one to visualize articulation on the surface of the human head and neck.

2. What to project

In this section, we propose three different image objects to be projected on the surface of a speaker’s (user’s) head and neck.

2.1. Illustration of the human vocal tract

Figure 1 shows the schematic diagram of the system. An LCD projector projects a midsagittal cross-sectional illustration of the human vocal tract on to the surface of the head and neck of a speaker, as shown in Fig. 2. The illustration can be either a still image or a movie. In the latter case, the speaker has to synchronize his/her movements to the audio of the clip. The projected articulators move as if the speaker were actually articulating.

2.2. Vocal-tract model with a flexible tongue

Arai (2015) [2] showed that users of the vocal-tract model with a movable tongue can adjust the position of the tongue by listening to the sounds, receiving a tactile sensation of the model, and watching the vocal-tract configuration to produce target sounds. This study further pointed out that such models can be applied for the pronunciation training as a hands-on tool for phonetic education based on the consideration that tongue and finger movements are related in terms of motor control [2]. We, therefore, augmented the visual feedback by projecting the vocal-tract configuration on the user’s head and neck.

Figure 3 shows the schematic diagram of the system. A camera captures the vocal-tract model with a flexible tongue [3]. An LCD projector projects the images captured by the camera on to the surface of the head and neck, as shown in Fig. 4. The images can be moving pictures from a camera “on line.” In other words, the moving pictures projected on the human head and neck can be synchronized with the movement of an actual articulation by manipulating the tongue configuration of the vocal-tract model in real-time.

2.3. Ultrasonic images

The ultrasonic imaging technique is widely used in the field of articulatory phonetics [e.g., 4], and we applied it to PM. Figure 5 shows the schematic diagram of the system. An ultrasound device captures images inside the speaker’s head from a probe attached under the jaw as shown in Fig. 5. An LCD projector projects the images captured by the ultrasound device onto the surface of a human head and neck, as shown in Fig. 6. The images from the ultrasound device could be moving pictures that are already synchronized with the

*e-mail: arai@sophia.ac.jp
movement of the actual articulation, because the video signal is directly derived from the ultrasound device during the speaker’s articulation in real-time.

3. Discussion and conclusion

In this study, we proposed and tested three ways of applying the PM technique for visualizing articulation on the surface of the human head and neck: 1) illustrations of the human vocal tract, 2) images of the vocal-tract model with a flexible tongue, and 3) images from an ultrasound device. Projection images can be either still or moving pictures. With still pictures, we expect a benefit for education because speech organs and articulators are not visible from outside, and displaying them in this way is highly intuitive for users. For video images of illustrations (Sect. 2.1) and video images of the vocal-tract model (Sect. 2.2), the issue of synchronization still remains. However, to synchronize his/her movements with the audio, a speaker can practice an utterance many times to perfect synchronicity. With the vocal-tract model with a flexible-tongue, we can expect a synergy effect of visual, auditory and tactile sensations as an advantage as pointed out in [2].

With ultrasound devices, such practice of the synchronicity is not necessary because the video images from the device are already synchronized with the articulation of the utterance. Thus, if an ultrasound device is available, PM is effective for users who want improve their pronunciation. Based on our informal testing, PM with the ultrasound device is the most effective because it seems like the speaker’s articulators are visible through his/her cheek. Visualizing articulation is extremely difficult for language learners and patients with articulation disorders; therefore, our proposed method should be useful for both language learning and clinical purposes, so that the speaker could be either an instructor, a pathologist, a student or a patient. In the future, we should discuss how our proposed method might combine with biofeedback to improve pronunciation and articulation [5].
Acknowledgments
This work was partially supported by JSPS KAKENHI Grant Number 15K00930. I would also like to thank Terri Lander for her support.

References