A Novel Self-adjusting Ultrasound Transducer Stabilized Elastic System Realized by 3D Printing

Dan Wu¹, Wenhuan Lu¹*, Yuxuan Li², Haiqiang Shi³ and Jianguo Wei⁴

¹ College of Intelligence and Computing, Tianjin University, Tianjin, Tianjin, 300350, China
¹* College of Intelligence and Computing, Tianjin University, Tianjin, Tianjin, 300350, China
² College of Intelligence and Computing, Tianjin University, Tianjin, Tianjin, 300350, China
³ Information Center, The Party School of Qinghai, Xining, Qinghai, 810007, China
⁴ College of Intelligence and Computing and School of Computer Science, Tianjin University and Qinghai Nationalities University, Tianjin, Tianjin, 300350, China
*E-mail: wenhuan@tju.edu.cn

Abstract. The human tongue is a significant organ for swallowing and speaking. The measurement of tongue function is challenging because it is hidden inside the mouth. Ultrasonic imaging is often used to study and record tongue movement. In order to ensure the validity and authenticity of the collected data, numerous methods are proposed to stabilize the transducer. However, in the long-term data acquisition, few devices consider issues such as whether the jaw is in compression during speech and whether the participant can autonomously complete the adjustment, which leads to data frame discarding and experimenter discomfort. This paper describes a novel self-adjustable ultrasound transducer stabilization elastic system using 3d printing technology that allows free head movement and flexible jaw movement. The system performance is evaluated quantitatively. It shows that in the 95% confidence interval, the average translational offset of the transducer relative to the head is 2.44mm and the rotation angle is 3.21º. Furthermore, the small head constraint makes the voice close to its natural state and more suitable for scientific studies of voice and swallowing.

1. Introduction
The tongue is an essential part of all oropharyngeal behaviors and aiding pronunciation. Due to the characteristics of the tongue, studying its movement is an extremely challenging task. Therefore, a non-invasive ultrasound tongue imaging technology (UTI is ensuring that the probe’s 2D plane of analysis, once obtained, is consistent.[1]) is introduced, which can record and observe the position, shape, and movement of the tongue in real-time. Ultrasound tongue imaging consists of an ultrasound transducer emitting high-frequency ultrasound waves through the mandible through the tongue, and the imaging high-frequency ultrasound waves returned to the ultrasound transducer through refraction, reflection, and attenuation of the ultrasound waves, which finally form an image in the body of the ultrasound instrument (shown in Figure 1). In recent decades, despite the rapid development of UIT, it has been widely used in speech teaching [2,3], speech therapy [4,5], etc. However, the effects produced by
transducer rotation or translation and tongue base tissue compression will lead to the problem that the acquired data are not informative or increase the workload of correcting the data later. The usability of the UIT is extremely reduced [6]. Hence, the stability of the transducer with respect to the head and uncompressed state of the jaw are prerequisites to ensure that the UIT accurately reflects the tongue movement.

Over the years, various methods have been proposed to stabilize the ultrasound probe to observe tongue movement, but each has its advantages and disadvantages. Researchers initially handheld ultrasound transducer stabilizers [7] (see Figure 2a), to using a dental chair to fix the head and transducer in combination [8] (see Figure 2b). Despite successfully resolved the stability of the transducer to collect data, but it also brings long wear so that the subject’s head unconsciously slip occurs on the probe, which causes the loss of the ultrasound image problem. Later, the application of the Head and Transducer Support System (HATS) [9] (see Figure 2c) and some of its evolved structures [10,11] solved the problem of unintentional head sliding on the transducer probe, but it also caused painful to keep the head stationary for a long time. A promising technique to overcome the above challenges is to fix the transducer within the structure of the helmet. The helmet modification is being developed at Queen Margaret University College in Edinburgh, Scotland. It is more feasible. However, it is less universally applicable because it needs to be designed for different people.

The Haskins Optically Corrected Ultrasound System (HOCUS) [12] (see Figure 2d) incorporated an optical tracking system, solving the problem of not being able to correct the ultrasound image position as a way to filter out frames that deviate from the standard. HOCUS remains relatively well used, and it also determines the extent to which the ultrasound probe deviates from the sagittal midplane in displacement and rotation. But the inclusion of metal will collect data with other signal devices, which can interfere with data acquisition. As a secondary concern, prolonged wear (approximately 30-45 minutes at a time) can lead to discomfort, as tested by Scobbie et al. [1] for a 20-minute run time.

With the rapid development of 3D technology, in recent studies, Donald Derrick’s (2018) design of a Three-dimensional printable ultrasound transducer stabilization system [15] (see Figure 3a), Lorenzo Spreafico et al. (2018) proposed UltraFit [16] (see Figure 3b), with a mass of 200g-350g, which dramatically alleviates the pain given to the head by the helmet, the wearing process is more comfortable, and the production is easy and cheap. The disadvantages are its rigid structure and non-elastic bandage structure at the fixation from the head to the jaw, resulting in restricted jaw movement and compression of the tissue at the base of the tongue during speech. M. Hamed Mozaffari and Won-Sook Lee (2019) developed a hand-held 3D printed instrument UltraChin [17] (see Figure 3c), and to verify stability, a magnetic tracking sensor was attached to the UltraChin. Although the stability was not as good as the
previous helmet, the error was still within the acceptable deviation. The weakness is that unknowingly holding the bracket for a long time, there will be a similar problem of the head can not be secured, resulting in the loss of data frames.

Figure 3. 3D printing series: a) Three-dimensional printable ultrasound transducer stabilization system [15]. b) UltraFit[16]. c) UltraChin[17]

This article describes a system in which participants can adjust themselves according to their comfort level. Also, an adjustable elastic structure is used on the side of the helmet instead of the previous rigid structural connectors, weakening the upward force received by the jaw. The system combines OpenCV and Dilb based head posture estimation techniques [18] and HSV color model to track video transducer marker positions. Thus it can be used to evaluate and correct the probe and head out of acceptable offset range during use, as a way to accomplish probe repositioning relative to the head. It was demonstrated that the relative stability of the probe and head remained excellent under prolonged wear, with a translation of 2.44 mm and a rotation of 3.21° at the 95% confidence interval. It also allowed free movement of the head and jaw without adverse effects in the experimenter.

2. Development of the system

Our goal was to combine the advantages of earlier designs and the stability of traditional helmets, and at the same time, we conducted a targeted study of head geometry anatomy [19] to provide logical modelling data for this system and complete the design. Modelling uses Solidwork to complete the design and joins the prototype iteration mode. Modifications were made through the wearing feedback of the experimenters.

The system consists of upper roof-mouth slot with an elastic pad (shown in Figure 4A), a bracket around the forehead of the head (shown in Figure 4B), a holder with a bi-directional rotating structure at the back of the head (shown in Figure 4C), and some fixation bolts and nuts.

Part B (Figure 4 B) and part C (Figure 4C) are connected with a non-elastic bandage to complete the head fixation. Part B-a (Figure 4 B) can be retrofitted with a camera or microphone in the future through the designed multi-functional slot. Part A-a (Figure 4A) places a memory foam pad to increase the comfort level. The two-way frame C-a (Figure 4C) gear part C-b (Figure 4C) fits together for easy adjustment by the experimenter and allows different people to wear it. The elastic band used for the connection between part A (Figure 4A) and part C (Figure. 4C) makes the movement of the jaw during speech unrestricted by the rigid structure, and the adjustable bandage allows people with different head shapes to wear it. Adjustability and good flexibility are designed to bring the speaker closer to a comfortable state, reflecting more authentically on the tongue and producing better image quality.

3. Experiments

To assess the accuracy of the system, we recorded two non-linguist participants, one male, and one female, in two sets of experiments:
a) Repeating the mouth as wide as possible 5 times
b) Reading screen data (collecting ultrasound images): vowels /a/, /i/, /u/, vowel-consonant combinations /pa/, /ka/, /ta/, Chinese sentences, English phrases, swallowing status, drinking water, each action repeated three times

The participants were free of speech, hearing, swallowing disorders, and comprehension, and all signed a voluntary letter.

Figure 4. System development components: A) upper roof-mouth slot. B) Bracket around the forehead of the head. C) Double rotary actuator.

Figure 5. Experimental equipment and wearing conditions: a) Front view. b) Articulate Instruments. c) Side view and the protruding red dot of the probe holder bracket.

Figure 5 shows the combination of equipment used to acquire the above data, the main equipment used: Articulate Instruments (Figure 5b), the research system comes with software for experimental research, including precise time alignment of audio and ultrasound streams (frame rate for the ultrasound stream is 60-120 Hz). [20]. The USB port of Machenike’s late 2015 model, which has a 2.5 GHz, quad-core i5, 8 GB RAM, and camera ultrasound probe with an external ECM8000 microphone.

Combined with recent research [21, 22], it automatically tracks the 3D movement of the head and completes a large number of facial feature training in the Dilb library, so it can be applied to different facial contours. We used the left eye corner, right eye corner, nose, left mouth corner, right mouth corner, and chin, a total of 6 points to mark the head motion.

The marker points are positioned at the transducer’s probe location. The two red marker points with the fixation clips are parallel to the bottom of the probe. The Euclidean distance between the probe and each marker point gives reasonable three-dimensional data (as shown in Figure 5c). A Python script was written using the HSV color model to track the coordinates of the video sensor marker locations, combined with OpenCV and Dilb-based head pose estimation techniques. Head motion was estimated from the global coordinate system of the camera or room, and then the probe motion was transformed from a global coordinate system to a local coordinate system relative to the head (as shown in Figure 6).

Before each start of the experiment, a small amount of ultrasound gel is applied to the probe pads to improve contact with the skin and the probe, to provide captured image sharpness. Whereby the center of the probe recorded six degrees of freedom X (transverse), Y (coronal), Z (sagittal) displacement, and roll, pitch, yaw.
4. Result & Discussion
The maximum errors for the two participants in the group experiment are shown in Table 1.

Table 1. Maximum offset and offset angle for 2 speakers.

| Displacement (mm) (95%) | Angle (degrees) (95%) |
|-------------------------|-----------------------|
| x                       | y                     | z                     | roll | pitch | yaw |
| Male                    | a) 1.68               | 1.53                  | 1.21 | 2.89  | 1.13 | 1.31 |
|                         | b) 3.22               | 3.17                  | 1.95 | 3.55  | 2.55 | 2.56 |
| Female                  | a) 1.24               | 0.74                  | 1.68 | 1.51  | 1.50 | 2.03 |
|                         | b) 1.22               | 2.42                  | 2.27 | 2.93  | 3.37 | 3.76 |
| Mean                    |                       |                       |      |       | 2.44 | 3.21 |

In this study, we present a 3D printed self-adjusting ultrasound transducer stabilization elastic system to ensure that the head and transducer remain relatively stable. It can adjust the system according to its comfort level to complete data acquisition for ultrasound imaging of the tongue. We use an elastic structure instead of a rigid structure on the side so that the ultrasound probe moves with the jaw. It reduces the degree of pressure on the tongue tissue and allows for more natural pronunciation. It allows UIT to adapt to more areas of research, such as second language acquisition in children or the treatment of speech disorders. The system was evaluated quantitatively in terms of accuracy. By tracking the position of the video tag, measuring transducer head movement relative to six degrees of freedom. The results show that our device can remain relatively stable for a long time with an average deviation of 2.44 mm and a rotation angle of 3.21° within the 95% confidence interval in all three directions, well within the acceptable range of 2 to 4 mm offset and 5° to 7° rotation angle [1].

5. Conclusion
The effectiveness and convenience of the system, as well as the printability of the material, make it a feasible solution. No tedious instruction steps are required to assemble the device. It is necessary to complete the above experiments on other adapted fixation clips at a later stage. So that in the future, it can be more widely combined with various new technologies, so that it can record the articulatory organs more comprehensively and study speech and swallowing more rigorously and scientifically.
Acknowledgment
This work was supported by National Natural Science Foundation of China (No.61876131, U1936102) and the National Key R&D Program of China (No.2020YFC2004100).

References
[1] James M Scobbie, Alan A Wrench, and Marietta van der Linden. (2008) Head-probe stabilisation in ultrasound tongue imaging using a headset to permit natural head movement”. In Proceedings of the 8th International seminar on speech production.
[2] Abel, J., Allen, B., Burton, S., Kazama, M., Kim, B., & Noguchi, M., et al. (2015). Ultrasound-enhanced multimodal approaches to pronunciation teaching and learning. Jennifer
[3] Sonya Bird and Bryan Gick. (2018) Ultrasound biofeedback in pro-nunciation teaching and learning. In Proc. International Symposium on Applied Phonetics, pages 5–11.
[4] Bacsfalvi, P., & Bernhardt, B. M. (2011). Long-term outcomes of speech therapy for seven adolescents with visual feedback technologies: ultrasound and electropalatography. Clin Linguist Phon, 25(11-12), 1034-1043.
[5] Bernhardt, B., Gick, B., Bacsfalvi, P., & M Adler-Bock. (2005). Ultrasound in speech therapy with adolescents and adults. Clinical Linguistics & Phonetics, 19(6-7), 605.
[6] Gick, Bryan & Bird, Sonya & Wilson, Ian. (2005) Techniques for field application of linguistic ultrasound imaging. Clinical linguistics & phonetics. 19. 503-14.
[7] Stone, and Maureen. (1990) A three-dimensional model of tongue movement based on ultrasound and x-ray microbeam data. Journal of the Acoustical Society of America 87.5:2207.
[8] Stone, M., Shawker, T., Talbot, T., & Rich, A. (1988) Cross-sectional tongue shape during the production of vowels “. Journal of the Acoustical Society America, 83 (4), 1586-1596.
[9] Stone, M. & Davis, E. (1995) A head and transducer support (HATS) system for use in ultrasound imaging of the tongue during speech”. Journal of the Acoustical Society of America, 98, 3107-3112.
[10] C L Peng P G Jost-Brinkmann R R Miethke C T Lin. (2000) Ultrasonographic measurement of tongue movement during swallowing. Journal of Ultrasound in Medicine, 19:1, 15-20.
[11] Davidson, Lisa & Decker, Paul. (2005) Stabilization techniques for ultrasound imaging of speech articulations. The Journal of the Acoustical Society of America. 117. 2544.
[12] Stone, M., Sutton, M. Parthasarathy, V., Prince, J., Li, M., & Kambhamettu, C. (2002) Effects of upright and supine orientation on tongue position during silence. Joint meeting of the Acoustical Society of America. Iberoamerican Congress of Acoustics and Mexican Congress on Acoustics. Cancun, Mexico. December 2-6.
[13] Whalen, D. H., Iskarous, K., Tiede, M. K. (2005) The Haskins Optically Corrected Ultrasound System (HOCUS). Journal of Speech, Language, and Hearing Research 48, 543–553.
[14] Stone, and Maureen. “A guide to analysing tongue motion from ultrasound images. “ Clinical Linguistics & Phonetics 19.6-7,2005.
[15] Derrick, Donald, et al. (2018) Three-dimensional printable ultrasound transducer stabilization system. The Journal of the Acoustical Society of America 144.5:EL392-EL398.
[16] Spreficco, Lorenzo, M. Pucher, and A. Matosova. (2018) UltraFit: A Speaker-friendly Headset for Ultrasound Recordings in Speech Science. Interspeech 2018. 1517-1520.
[17] Moazzafi, M. Hamed, and W. S. Lee. (2019) Real-time Ultrasound-enhanced Multimodal Imaging of Tongue using 3D Printable Stabilizer System: A Deep Learning Approach.
[18] Dr. David Kriegman, Kevin Barnes and Avatar. (Sept. 26 2016) [Online] Available:https://learnopencv.com/head-pose-estimation-using-opencv-and-dlib/.
[19] Dean Lung Professor, (2020) Systematic Anatomy, the 9th edition of the textbook for human health [Online] Available: https://www.med2.cn/view/6084.html.

[20] Bird, Sonya & Gick, Bryan. Ultrasound biofeedback in pronunciation teaching and learning. ISAPh.2018-2,5-11. 10.21437/.

[21] Mozaffari, M., Wen, S., Wang, N., & Lee, W. S. (2019) Real-time Automatic Tongue Contour Tracking in Ultrasound Video for Guided Pronunciation Training. 14th International Conference on Computer Graphics Theory and Applications. Volume 1:302–309.

[22] M Hamed Mozaffari, Shenyong Guan, Shuangyue Wen, Nan Wang, and Won-Sook Lee. (2018) Guided learning of pronunciation by visualizing tongue articulation in ultrasound image sequences. IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications, pages 1–5.