INTERSPEECH 2017: Show & Tell Contribution
August 20–24, 2017, Stockholm, Sweden

MoPAReST - Mobile Phone Assisted Remote Speech Therapy Platform

Chitralekha Bhat[1], *Anjali Kant*[2], *Bhavik Vachhani*[1], *Sarita Rautara*[2]
Ashok Kumar Sinha[2], *Sunil Kopparapu*[1]

[1]TCS Innovation Labs, Mumbai, India
[2]Ali Yavar Jung National Institute for Speech and Hearing Disabilities, Mumbai, India
bhat.chitralekha@tcs.com

Abstract

Through this paper, we present the Mobile Phone Assisted Remote Speech Therapy Platform for individuals with speech disabilities to avail the benefits of therapy remotely with minimal *face-to-face* sessions with the Speech Language Pathologist (SLP). The objetive is to address the skewed ratio of SLP to patients as well increase the efficacy of the therapy by keeping the patient engaged more frequently albeit asynchronously and remotely. The platform comprises (1) A web-interface to be used by the SLP to monitor the progress of their patients at a time convenient to them and (2) A mobile application along with speech processing algorithms to provide instant feedback to the patient. We envision this platform to cut down the therapy time, especially for rural Indian patients. Evaluation of this platform is being done for five patients with mis-articulation in Marathi language.

Index Terms: Remote speech therapy, mis-articulation, automatic assessment, speech disabilities

1. Introduction

Internet of things (IoT) and pervasive frameworks involving smart devices have revolutionized the health care domain [1, 2]. Rehabilitation is a typical application of IoT-based smart systems that targets problems related to shortage of health care professionals and ambient assisted living for elderly [3]. Research in the speech signal processing domain is gearing up towards building IoT-based systems for rehabilitation of patients with speech and hearing disorders. Serious attempts are being made towards understanding the application of speech signal processing and machine learning techniques to pathological speech [4]. Automatic assessment of pathological speech has been studied extensively by speech researchers [5, 6]. In [7], automatic speech recognition (ASR) technology in conjunction with multimedia gaming is used to provide computer-based therapy to patients with dysarthric speech.

In this work, MoPAReST (Mobile Phone Assisted Remote Speech Therapy) Platform we use IoT-based framework to rehabilitate patients with speech disorders. Automatic assessment of pathological speech using speech signal processing is an integral part of this smart system. Other major components of the framework are (1) A web interface to be used by the speech language pathologist (SLP) which enables the SLP to digitize their interactions with their patients and monitor patient progress remotely at a convenient time and (2) A mobile application for an Internet enabled smart phone to be used by the patients to continue with therapy drills or practice exercises.

2. MoPAReST platform design

Figure 1 shows the system architecture for the MoPAReST platform. Each component of the platform is described in the subsequent sections. This platform is envisioned to cut down the therapy time and cost as well as increase the reach of the SLP. In the current scenario, especially rural Indian patients who are in need of speech therapy cannot avail the benefits of an SLP through *face-to-face* sessions as frequently as desired by the SLP, due to the travel and opportunity costs involved. Also, the patients are expected to practice exercises or drills prescribed by the SLP either on their own or under care-giver supervision. The infrequent visits and lack of supervised practice sessions impacts the therapy adversely, often leading to regression of patients in subsequent sessions. MoPAReST platform attempts to bridge the gap of continuous and supervised training for the patients. A typical life-cycle of therapy using MoPAReST plaform is described below:

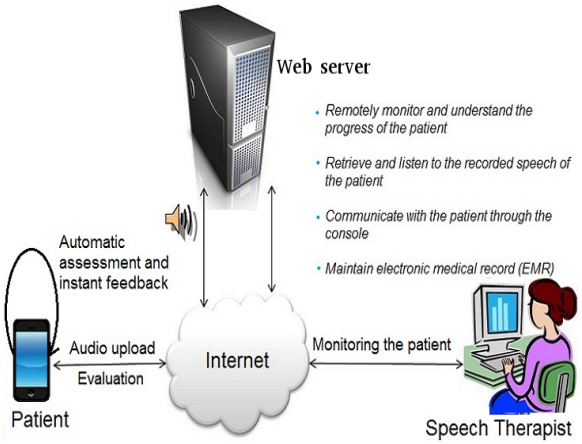

Figure 1: *System architecture of MoPAReST platform*

- Initial *face-to-face* sessions with an SLP.
- Based on the SLP's recommendation, a patient is enrolled on to the platform, wherein the patient uses the mobile application on a smart phone to practice drills.
- Patient speech is assessed using speech processing algorithms that run natively (no Internet needed) on the mobile device and provide instant feedback to the patient.
- Patient speech and the automatic assessment provided by the speech algorithm are transmitted to the web server as and when the Internet is available.
- SLP is intimated regarding new patient data for monitoring wherein the SLP can listen to the patient speech, view the automatic assessment and correct the assessment if required through the web console.

The novelty of this platform lies in the fact that although therapy process is fine tuned using speech algorithms and instant

feedback to the patient, it retains the SLP involvement and intervention at regular intervals which is crucial to a therapy process. In its current form the MoPAReST platform caters to the speech disorder - misarticulation, wherein a patient is unable to pronounce correctly a particular sound. The SLP first establishes the speech sound that is in error and then transfers the patient onto the platform for further practice via drills. For each speech sound we define a specific repertoire of drills to be followed. Drills for each speech sound are built into the mobile application, which is used by the patient.

2.1. SLP Interface

The web interface is the SLP touch point to the MoPAReST platform. SLP enrolls a patient into the MoPAReST platform through the web interface, which then enables the SLP to monitor the patient progress remotely. During the enrollment, the SLP selects the phoneme or speech sound that needs to be practiced as well as the order in which the practice needs to be done in case of multiple phonemes are in error. The personal and therapy information for a patient is saved into the MoPAReST database for further use. The patient monitor page is as shown in Figure 2

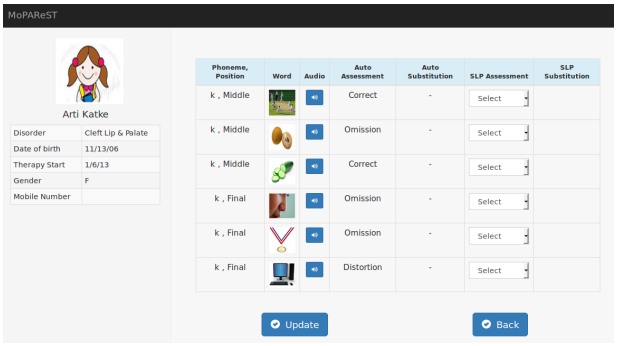

Figure 2: *MoPAReST platform - web interface*

2.2. Mobile Application

Mobile application is designed to enable a patient to practice therapy drills as determined by the SLP. Post enrollment of the patient into the MoPAReST platform by the SLP, the patient is provided login credentials for the mobile application. Patient's information on the MoPAReST database is synchronized with the mobile application using the patient login credentials. In order for the patient to practice drill lessons for a particular phoneme, visual stimuli are provided for words with the error phoneme in the initial or middle or final position of the word. Patient needs to identify the word from the visual stimulus and record the utterance and then submit the utterance for evaluation for correctness. On submission, the speech algorithm similar to the one described in [8] runs natively on the mobile device. An instant visual feedback is provided to the patient regarding the utterance. Based on the feedback and patient fatigue, multiple attempts for the same word is allowed. An audio cue for every word as well as an instructional video for every phoneme are provided to help the patient with the therapy drills. The drill itself is divided into multiple levels and the patient needs to complete the drill exercises level-wise. The multiple levels and instant feedback simulate a game-like environment that appeals to the young patients. Figure 3 shows the screen for the Marathi word */kulup/* corresponding to phoneme /k/ in the initial position.

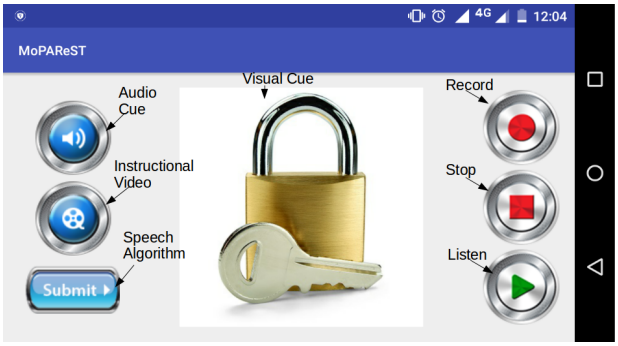

Figure 3: *MoPAReST platform - Andoid application screen shot*

3. Conclusions

IoT-based systems are taking the forefront in the health care domain. We describe one such framework for providing speech therapy remotely. Evaluation of this platform is being done for five patients with mis-articulations in Marathi language. MoPAReST framework can be extended to provide speech therapy to multiple speech disorders in many different languages. Further the system also serves as electronic medical record for patient progress.

4. Acknowledgements

This research was supported in part by Science for Equity Empowerment and Development (SEED), Department of Science and Technology, Government of India.

5. References

[1] S. M. R. Islam, D. Kwak, M. H. Kabir, M. Hossain, and K. S. Kwak, "The internet of things for health care: A comprehensive survey," *IEEE Access*, vol. 3, pp. 678–708, 2015.

[2] Z. Pang, "Technologies and architectures of the internet-of-things (iot) for health and well-being," Ph.D. dissertation, KTH Royal Institute of Technology, 2013.

[3] Y. J. Fan, Y. H. Yin, L. D. Xu, Y. Zeng, and F. Wu, "Iot-based smart rehabilitation system," *IEEE Transactions on Industrial Informatics*, vol. 10, no. 2, pp. 1568–1577, May 2014.

[4] R. Gupta, T. Chaspari, J. Kim, N. Kumar, D. Bone, and S. S. Narayanan, "Pathological speech processing: State-of-the-art, current challenges, and future directions," in *In Proc (ICASSP)*, Mar. 2016.

[5] C. Middag, J.-P. Martens, G. Van Nuffelen, and M. De Bodt, "Automated intelligibility assessment of pathological speech using phonological features," *EURASIP J. Adv. Signal Process*, pp. 3:1–3:9, Jan. 2009.

[6] A. Maier, T. Haderlein, U. Eysholdt, F. Rosanowski, A. Batliner, M. Schuster, and E. Nöth, "PEAKS A system for the automatic evaluation of voice and speech disorders," *Speech Communication*, vol. 51, no. 5, pp. 425 – 437, 2009.

[7] M. Ganzeboom, E. Yilmaz, C. Cucchiarini, and H. Strik, "On the development of an asr-based multimedia game for speech therapy: Preliminary results," in *In Proc ACM Workshop*, ser. MMHealth '16, 2016, pp. 3–8.

[8] C. Bhat, B. Vachhani, and S. Kopparapu, "Automatic assessment of articulation errors in hindi speech at phone level," in *TENCON 2015 - 2015 IEEE Region 10 Conference*, Nov 2015, pp. 1–4.