Chapter 2
Development of Human Speech Signal–Based Intelligent Human–Computer Interface for Driving a Wheelchair in Enhancing the Quality–of–Life of the Persons

Uvanesh Kasiviswanathan
Indian Institute of Technology (BHU), India

Abhishek Kushwaha
Indian Institute of Technology (BHU), India

Shiru Sharma
Indian Institute of Technology (BHU), India

ABSTRACT

For the past few decades, an increase in experimental research has been carried out in enhancing the quality-of-life of the persons with different levels of disabilities. To enhance the lifestyle of differently disabled in terms of their mobility or movement or transportation, a proper aid with appropriate human-computer interface system is needed. So, in this chapter, a hybrid classification model is proposed, which combines and uses hM-GM and ANN models, for classifying human speech signal, especially the word for driving a wheelchair for helping the people, who seek transportation. For classifying the correct word from the phase of sentence (i.e., the human speech signal) to corresponding trigger command for an electrically powered wheelchair prototype, under the certain experimental condition, the hM-GM model yields good recognition of words, but they suffer major limitations as it relies on strong statistical properties and probability. Hence, by combining hM-GM and ANN model-based classifier for enhancing the accuracy of classifying the word to corresponding trigger command.

1. INTRODUCTION

The real-time processing of speech signal generally deals with analysing the nature and characteristics of the signal(s) produced to the corresponding observable outputs (Schultz et al., 2017). The signals used can either be discrete or analog/continuous in nature and either it can be stationary or non-stationary signals (Gannot & Burshtein, 2001). If the signals are stationary then, the signal's statistical properties won't change with respect to time, but if it is non-stationary signal then, the signal's statistical properties will get change depending upon the time. By analyzing this, we are able to develop a "signal model" (Mermelstein, 1973; Huang et al., 2001). The signal model developed is nothing but a theoretical description of the system. Speech signals are basically one-dimension signals which vary over the time period and can be related to a source-filter system (Bell-Berti & Harris, 1981).

The source filter system (Galas & Rodet, 1991) consists of respiratory mechanism as a source and filtering are caused due to change in air pressure when the air starts to flow from lungs with pulmonary pressure. The air pressure gets changed as it reaches the larynx in the throat and produces vibration, and the hissingness. Vibration and hissingness is the primary cause of speech production (Mermelstein, 1973; Deller et al., 1993).

1.1. The Anatomical Structures Responsible for Speech Production

- Alveolar Ridge: A bony ridge located on the roof of the mouth (a half-inch behind upper teeth).
- Glottis: A hole/space present in between the vocal folds of the throat.
- Larynx (or Adam's Apple): Made of cartilage that holds the vocal folds of the throat.
- Lips: Are used to form consonants.
- Palate: The roof, divided into hard (front region) and soft (back region) palate present in the mouth.
- Pharynx: A tube that connects the larynx to the mouth region (oral cavity), located at backside in the throat.
- Tongue: Where the speech is originated. It is a large muscle, which can change its shape to generate different sounds.
- Uvula: Present at soft palate (the velum) that is responsible for the articulation of consonants.
- Vocal Folds/Cords: A small flap of muscle in the larynx, vibrates to generate different sounds (Isshiki 1989).

1.2. The Flow of Speech Production

A steady flow of air is supplied from the lungs with some pulmonary air pressure through the trachea, larynx, and pharynx. Then, the vocal folds/cord present in the larynx gets vibrated and generates a fluctuations/variations in air pressure, due to which sound waves generate (Mermelstein, 1973). The significate changes in the vocal tract modify the wave according to the position and shape of the speech organs to produce voiced/tone sound waves. Thus, through the mouth and nose openings, the sound waves radiate out into the environment (Mermelstein 1973; Rabiner & Juang 1993; Bogdashevsky et al., 1999).

Related Content

An Hebraic Alternative to Mind-Body Dualism
Sylvia Herold Olney (2024). *Perspectives on Social and Material Fractures in Care (pp. 184-205).*
www.irma-international.org/chapter/an-hebraic-alternative-to-mind-body-dualism/339017

Global Telemedicine and eHealth: Advances for Future Healthcare – Using a Systems Approach to Integrate Healthcare Functions
S. A. Davis (2015). *Healthcare Administration: Concepts, Methodologies, Tools, and Applications (pp. 1570-1586).*
www.irma-international.org/chapter/global-telemedicine-and-ehealth/116293

The Buying Decision-Making Process in the Selection of Device Suppliers in Reproductive Medicine: A Preliminary Model
Paulo Botelho Pires, António Correia Barrosand Filipe Taveira Santos (2021). *Management and Marketing for Improved Competitiveness and Performance in the Healthcare Sector (pp. 1-23).*
www.irma-international.org/chapter/the-buying-decision-making-process-in-the-selection-of-device-suppliers-in-reproductive-medicine/285799

The Need for a Socio-Technical Analysis in E-Health: The Case of the PCEHR
Imran Muhammad, Say Yen Teohand Nilmini Wickramasinghe (2015). *Healthcare Administration: Concepts, Methodologies, Tools, and Applications (pp. 1123-1138).*
www.irma-international.org/chapter/the-need-for-a-socio-technical-analysis-in-e-health/116268

A Survey on Health Care Services Using Wireless Sensor Networks
Sunilkumar S. Manviand Manjula R. B. (2015). *Healthcare Administration: Concepts, Methodologies, Tools, and Applications (pp. 132-151).*
www.irma-international.org/chapter/a-survey-on-health-care-services-using-wireless-sensor-networks/116212