

Chapter 42

Speaker Recognition With Normal and Telephonic Assamese Speech Using I-Vector and Learning-Based Classifier


Mridusmita Sharma

Gauhati University, India

Rituraj Kaushik

Tezpur University, India

Kandarpa Kumar Sarma

 <https://orcid.org/0000-0002-6236-0461>

Gauhati University, India

ABSTRACT

Speaker recognition is the task of identifying a person by his/her unique identification features or behavioural characteristics that are included in the speech uttered by the person. Speaker recognition deals with the identity of the speaker. It is a biometric modality which uses the features of the speaker that is influenced by one's individual behaviour as well as the characteristics of the vocal cord. The issue becomes more complex when regional languages are considered. Here, the authors report the design of a speaker recognition system using normal and telephonic Assamese speech for their case study. In their work, the authors have implemented i-vectors as features to generate an optimal feature set and have used the Feed Forward Neural Network for the recognition purpose which gives a fairly high recognition rate.

DOI: 10.4018/978-1-7998-2460-2.ch042

INTRODUCTION

Speaker or voice recognition is the task of automatically recognizing the identity of the person speaking. The task of speaker recognition considers the individual behavior as well as the characteristics of the vocal cord as features to identify a speaker. Speaker recognition is however very much different from speech recognition. In speech recognition, the speech signals which convey much information to the listeners are being detected with the help of the features that has the ability to represent the speech content. The primary level of speech recognition is to recognize the speech content spoken by the speaker. But at a higher level, the spoken utterances or speech samples contains the information regarding the gender, emotions and also the individuality of the speaker speaking (Reynolds, 2002).

With the rapid growth of techniques in the signal processing and model building and empowerment of computing devices, significant progress has been made in the speech recognition area. Research in the field of speaker recognition has also evolved at par with speech recognition and speech synthesis because of the similar characteristics and challenges associated with it. Research and development in the field of speaker recognition dates back to over four decades and this area is still an active topic for research. The application of speaker recognition technology has been continually growing in various fields of application such as forensic applications, dynamic signatures, gait, keystroke recognition, data encryption purpose, user validation in contact centers, etc (Beigi, 2011).

From the literature, it has been found that recent speaker recognition tasks have implemented i-vector based features for their purpose. The characteristics of a voice sample or about the speaker can be obtained by the features extracted from the speech sample. Recently the i-vector paradigm is widely used for the speaker recognition systems. The i-vector based feature extraction is based on Joint Factor Analysis (JFA) approach which provides an elegant way to convert the Mel Frequency Cepstral Coefficient (MFCC) feature vector of a variable length utterance into a low dimensional vector representation (Dehak et. al., 2011). I-vectors are one example of subspace modeling approaches that can be used to reduce the dimensionality of data before training and applying to classifiers for recognition purpose. The dimensionality reduction should make training of the classifier less computationally expensive which could enable us to train the system with more data. In this work, we have considered Assamese speech for our experimental work as the number of work done in this language is relatively less. Also, Assamese language has a very rich and diverse phonological content which shows notable variations from region to region as well as speaker to speaker. The diversity and variations present in this language increases the need to develop a language specific tool for speech and speaker recognition system (Sharma & Sarma, 2015).

The organization of the chapter is as follows. The second section provides an overview of the basic considerations and theoretical background that are associated with various speaker recognition tasks. A detailed literature survey of the previous work done related to speaker recognition problems and the features used for the purpose is presented in the third section. The fourth section gives the proposed system model, the speech database and the basic algorithm used for our proposed work. Experimental details and results derived from our work is shown in Section five. Section six concludes the chapter.

THEORETICAL CONSIDERATIONS

The study of the basic theoretical considerations of speaker recognition and other such related topics gives a proper understanding of the speaker recognition problems. In this section a brief overview of

23 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:
www.igi-global.com/chapter/speaker-recognition-with-normal-and-telephonic-assamese-speech-using-i-vector-and-learning-based-classifier/252058

Related Content

A Novel Bat Algorithm for Line-of-Sight Localization in Internet of Things and Wireless Sensor Network

Mihoubi Miloud, Rahmoun Abdellatif and Pascal Lorenz (2019). *Machine Learning and Cognitive Science Applications in Cyber Security* (pp. 213-239).

www.irma-international.org/chapter/a-novel-bat-algorithm-for-line-of-sight-localization-in-internet-of-things-and-wireless-sensor-network/227583

ISOPM: Framework for IT/IS Outsourcing Project Management

Maria da Glória Fraga, João Varajão and Paula Cristina Oliveira (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 1859-1881).

www.irma-international.org/chapter/isopm/252115

Social Media in Accelerating Mobile Apps

Asta Bäck and Päivi Jaring (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 1513-1526).

www.irma-international.org/chapter/social-media-in-accelerating-mobile-apps/252095

Speaker Recognition With Normal and Telephonic Assamese Speech Using I-Vector and Learning-Based Classifier

Mridusmita Sharma, Rituraj Kaushik and Kandarpa Kumar Sarma (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 805-829).

www.irma-international.org/chapter/speaker-recognition-with-normal-and-telephonic-assamese-speech-using-i-vector-and-learning-based-classifier/252058

A Cyber-Physical Photovoltaic Array Monitoring and Control System

Gowtham Muniraju, Sunil Rao, Sameeksha Katoch, Andreas Spanias, Cihan Tepedelenlioglu, Pavan Turaga, Mahesh K. Banavar and Devarajan Srinivasan (2020). *Cognitive Analytics: Concepts, Methodologies, Tools, and Applications* (pp. 978-1000).

www.irma-international.org/chapter/a-cyber-physical-photovoltaic-array-monitoring-and-control-system/252066