

## Efficiency Enhancing Parameters in Speaker Recognition System

Sriyeda Tanu\*

Department of Computer Science, RAK  
Medical & Health Sciences University,  
UAE

### Abstract

In this paper, I present high-level speaker specific feature extraction from speech signals that takes into account intonation, linguistics rhythm, linguistics stress, and prosodic features. I believe the rhythm is linked to linguistic units like syllables and manifests itself as changes in measurable parameters like fundamental frequency, duration, and energy. The syllable type features are used as the basic unit for expressing prosodic features in this research. Automatically locating the vowel starting point approximates the segmentation of continuous speech into syllable units. The knowledge of high-level speaker's specific speakers is used as a reference for extracting the prosodic features of the speech signal. High-level speaker-specific features extracted using this method may be useful in applications such as speaker recognition where explicit phoneme boundaries are not readily available. On TIMIT and HTIMIT corpora that were initially sampled in the TIMIT, the efficiency of the specific characteristics of the specific features used for automatic speaker recognition was evaluated. In summary, the experiment, the basic discriminating system, and the HMM system are formed on TIMIT corpus with a set of phonemes. For TIMIT utterances, the proposed ASR system shows efficiency gains over the traditional ASR system.

**Keywords:** Automatic speaker; Account intonation; ASR system

\*Corresponding author:

Sriyeda Tanu

✉ tanu131279@gmail.com

Department of Computer Science, RAK  
Medical & Health Sciences University, UAE

**Citation:** Tanu S (2021) Efficiency Enhancing Parameters in Speaker Recognition System. Int J Inn Res Compu Commun Eng. Vol.6 No.3:09

**Received:** September 07, 2021, **Accepted:** September 21, 2021, **Published:** September 28, 2021

### Introduction

In everyday life, the language is mostly used to transfer information from one person to another. It is broadcast using a network of licensed sound units. The limits set by the language must be respected in this sequence. As a result, speech and language are complementary to one another and cannot be separated. Speaker-specific variables are also included into the speech signal because each speaker has distinct physiological aspects of speech and speech generating style. As a result, the voice signal includes not only the desired message but also language and speaker characteristics [1].

Additionally, the speaker's emotional condition is conveyed through words. The speech message is primarily expressed as a collection of lawful sound units, each of which corresponds to the mode and place of speech output by a specific sound unit. Several levels of functionality are used to derive the language, emotions, and speaker elements of the information included in the speech signal. Existing speaker, language, emotion, and speech recognition systems rely on short-term spectral analysis-derived features. The spectral features, on the other hand, are influenced by channel and noise characteristics [2].

As a result, researchers are looking into the usage of additional traits that could provide more proof of a spectrum-based system. Voice processing research strives to develop machines that can perform automatic speech recognition, speech synthesis, speaker recognition, and a variety of other speech processing tasks, including human-like speech recognition. The researchers were successful in designing speech systems that could operate in a limited area [3]. Many of these systems rely only on acoustic models built from spectral data. These audio models are missing a lot of the higher-level data that humans employ for the same task [4].

Prosody, context, and lexical understanding are among the greatest layers of information. It is thought that incorporating knowledge of prosody into the automated speaker recognition (ASR) system of the vocal systems will make them more intelligent and human-like. The importance of prosodic qualities for speech processing applications has been proven by a number of studies in the past. Unfortunately, including prosody into speech systems necessitates addressing a number of difficulties.

### Conclusion

The automatic extraction and representation of prosody, as well

as its application in speaker recognition to improve ASR efficiency, is an important topic. Our basic understanding of the processes in most of the speech perception modules is rudimentary at best, but it is widely accepted that the human brain contains some physical correlates of each of the steps in the speech perception model, making the entire model useful for thinking about the processes that occur.

## References

1. S Singh (2018) Forensic and Automatic Speaker Recognition System. Int J Applied Eng Res 5: 2804-2811.
2. Singh S (2018) Evaluation of Scarification algorithm and Its Application in Speaker Recognition System. J Applied Eng Res 13: 13015-13021.
3. Shriberg E, Stolcke A, Hakkani-Tur D, Tur G (2000) Prosody-based automatic segmentation of speech into sentences and topics. Speech Communication 32: 127-154.
4. Iliou T, Anagnostopoulos CN (2009) Statistical evaluation of speech features for emotion recognition. In Proceedings of Fourth International Conference on Digital Telecommunications 1: 121-126.