HMM-based phoneme speech recognition system for control and command of industrial robots

Adwait Naik[1,2]

[1]EPFL
[2]KJ Somaiya College of Engineering

May 27, 2020

Abstract

In recent years integration of Human-Robot interaction with speech recognition has gained a lot of pace in the manufacturing industries. Undoubtedly, this bridges the large gap created between the operator and robot by communication's point of view. Although there are numerous ways in which communication can be established between a human operator and the robot-like, controlling with a teaching pendant, or a joystick. Currently, the robots are controlled semi-autonomously by means of a computer. However, speech and touch [16] are natural ways of communication in humans, where speech recognition, being the best, is heavily researched technology. In this study, we aim at developing a stable and robust speech recognition system to allow humans to communicate with machines (Robotic-arm) in a seamless manner. This paper intends to investigate the potential of the Linear Predictive coding technique to develop a stable and robust HMM-based phoneme speech recognition system for robotics applications. Our system is divided into three segments: a microphone array, a voice module, and a 3-DOF robotic arm (Figure 1). To validate our approach, we have performed tests with simple and complex sentences for various robotics activities like manipulating a cube and pick and place tasks. Moreover, we also analyzed the test results to rectify problems like accuracy, recognition score, etc. Also the paper briefly enumerates the future prospects and applications of our approach.

Hosted file

Improved_paper.pdf available at https://authorea.com/users/313322/articles/454543-hmm-based-phoneme-speech-recognition-system-for-control-and-command-of-industrial-robots

Recognition score

Recognition score

Word Error Rate(%)

Speech
Signal

Frame
Blocking

Windowing

Autocorrelation
Analysis

LPC

LPC
Analysis

Training

Speech

Feature
Extraction

Feature

Vectors

Hidden Markov
Model Trainer

Grammar
Estimator

Phonetic
Models

Lexicon

Recognition

Training

Grammar

Word Models

Speech

Input

Feature
Extraction

Feature

Vectors

Recognition
Search

Most
Likely
Sentences

4

Word Accuracy Rate(%)