

5. Analysis and Design

5.1. Introduction

This chapter will discuss the analysis and design aspects of the implementation part of this research. Design is explained using use cases and their decomposition. System architecture section discuss the overall design aspects of the research.

5.2. Abstract of the System Design

This research is basically focused for recognize and interpret the human voice in to text. As the first part system is designed to identify the digits speak in Sinhala language. Microphone is use to capture the human voice

System also has set of recognized words which are used as command such as start, exit, stop etc...

Basically system contains two major components, one use to capture the user data and process the acoustic signal. Other component is meant to interpret the processed signal into matching digits

5.3. Functional overview of the system

As mentioned in the abstract system contains two major components at first. When we are going to detail design system has three major functional behaviors under those two major components

First component takes the audio input and process it, extract the features from input signal to help the recognition. This is basically the front end of the system. Most of the user interactions are happening with this component

Second component is the most important part of this system design. It is the decoder, decoder use the output of the first component the match the output with the knowledge base and performs search to match the correct digit according to the vocal signal

Next is the knowledge base, it basically contains three parts

1. Dictionary - Matching the Digit with Pronunciation
2. Language Model – This contains the model vocals, this will help to refine the search

3. Acoustic Model – This is the acoustic database of statistical models

5.4. System Architecture and structure

This phase describes the system architecture and the interconnection between the modules within and outside. This recognition system comprises of several modules, so it is very important function all those interconnection properly in order to function the recognition as whole.

5.4.1. Architecture

System contains three major components as describes above

Front end captures the user inputs thru a microphone and analyze and break the vocal in to frames and then to features to pass into the decoder

Row data captured using a microphone sends to a data framer to frame the row data, afterwards frames will be extracts in to signal features. Then processes the extracted features and pass them to decoder

Following image describes front end behavior

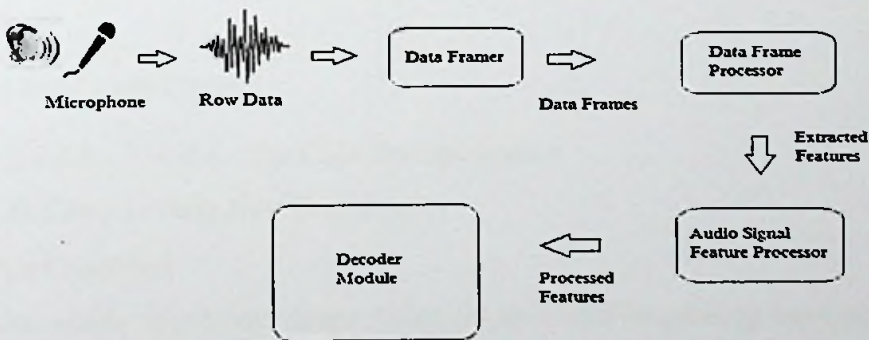


Figure 5-1- Frontend Behavior of the System

After the signal features send to the decoder, decoder uses the knowledge base to interpret the received features. Once decoder finishes the decoding output or the identified digit sends to the frontend for display.

5.5. Detail System Software Design

5.5.1. Use Cases

There are two major use cases to this recognition system. Since this is developing to a call center, always only authorized personnel can access to the system to retrieve customer data by giving their account number, so first major use case is to create authorized profile to call center agents. Next use case is the speech recognition use case

5.5.1.1.1. Use Case - Level 0

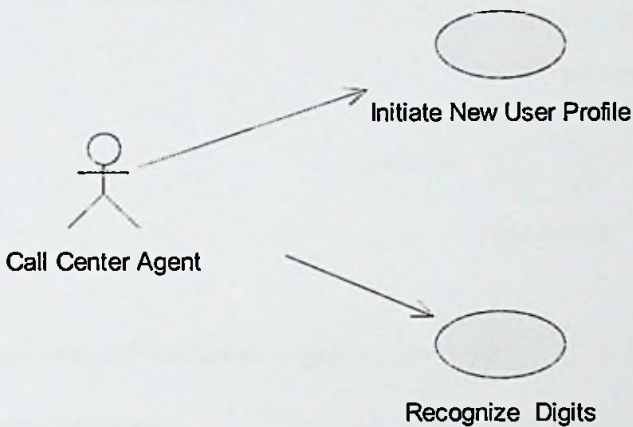


Figure 5-2- Level 0 Use case

5.5.1.1.2. Level 0 - Use Case Decomposition

Use Case: Initiate New User Profile

Pre-Condition

User needs to provide sample voice cut to a well-functioning microphone with low and consistent noise level

Description

This is the use case performs the user profile creation, once a person is allocated as a call center agent he/she needs to create their own profile in order to use the system. In order to create a user profile user needs to enter his/her sample voice cut to the system.

Then the profile handler analyses the input voice and obtain specific features for the given user and create new user profile by saving those data to a storage (to Database or to a Flat file).

Level 1: Use Case for “Initiate New User Profile”

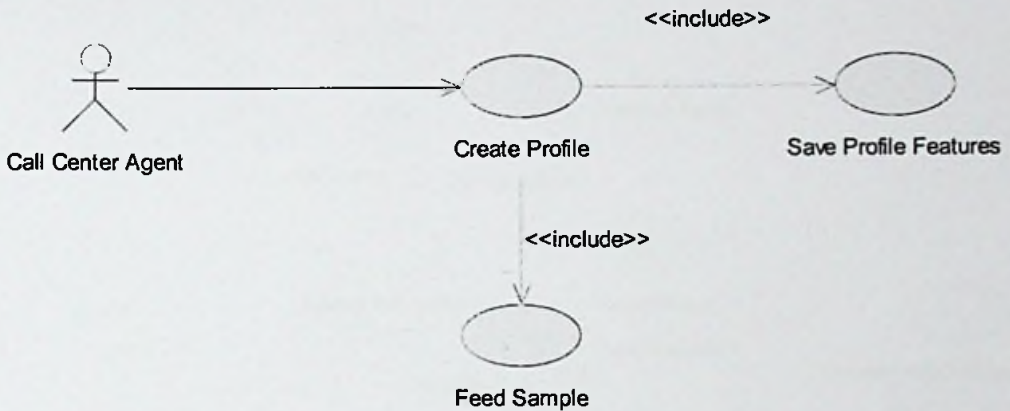


Figure 5-3- Level 1: Use Case for “Initiate New User Profile”

Post Conditions

New user profile needs to create against the user

Use Case: Recognize Digits

Pre-Conditions

User needs to have proper user profile and successfully spoke to the microphone

Description

This is the most important use case of the system, this will get the vocal from user and get the user profile from first use case and do the analytical and recognizing functions using system resources

This use case is responsible of following areas

1. Processing the input signal
2. Match the outcome of the signal processing with the knowledgebase
3. Retrieve the correct digit from knowledge base and return it to front end

4. Display the digit

Level 1: Use Case for "Recognize Digits"

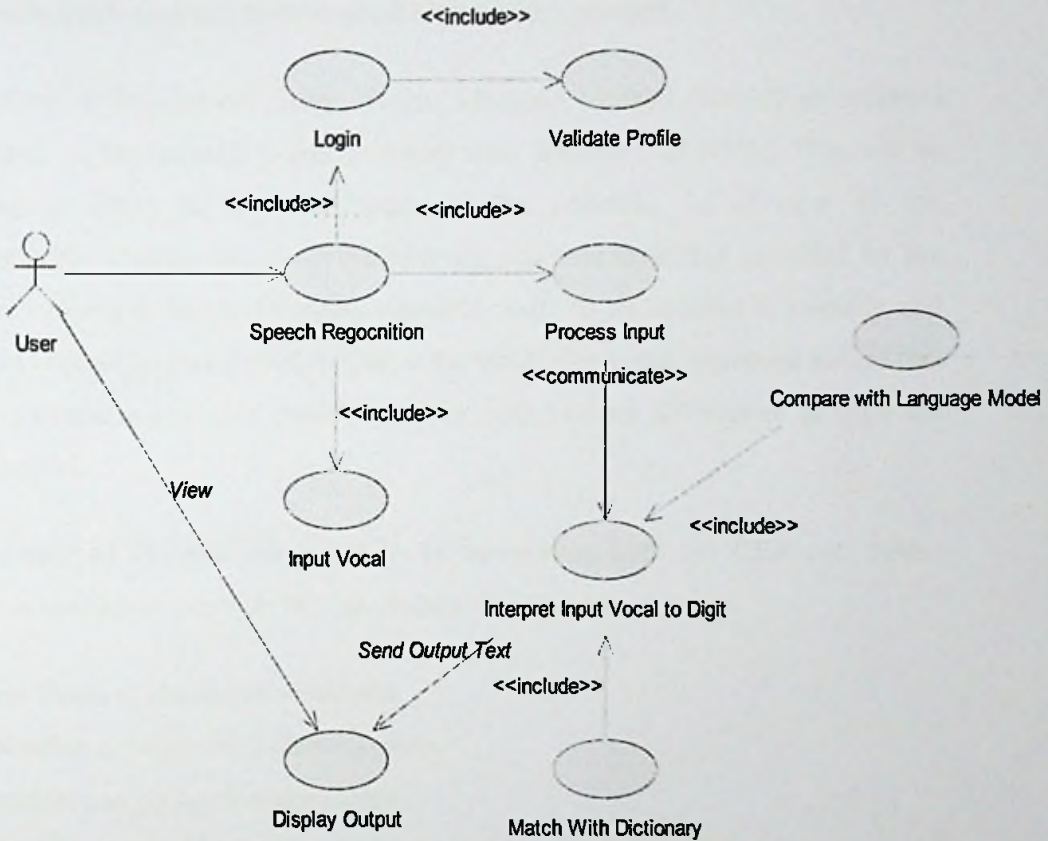


Figure 5-4 -Level 1: Use Case for "Recognize Digits"

Post Conditions

System needs to display correct digit according to the users' vocal input

Exception: If user input is invalid or login failed, system needs to provide appropriate error messages