Voice Recognition System for Massey University Smarthouse

A thesis presented in partial fulfilment of the requirements for the degree of

Master of Engineering
in
Information Engineering

at

Massey University, Auckland, New Zealand

Rafik Gadalla

2006

Acknowledgment

I would like to express my sincere gratitude to my project supervisor, Dr Tom Moir, for his guidance and support throughout the duration of this project.

I would also like to thank my colleagues Leon Kok, Grettle Lomiwes and Vaitheki Yoganathan and other members of the smarthouse projects for their effort and dedication towards the formation of the smarthouse.

I'm grateful to my family for all their support and encouragement. In particular I would like to thank my parents for their love, patience and encouragement, my fiancé Manal, my brothers and friends whom without their assistance this project would not have been made possible.

And above all, I would like to thank God for making it all happen.

Abstract

The concept of a smarthouse aims to integrate technology into houses to a level where most daily tasks are automated and to provide comfort, safety and entertainment to the house residents. The concept is mainly aimed at the elderly population to improve their quality of life.

In order to maintain a natural medium of communication, the house employs a speech recognition system capable of analysing spoken language, and extracting commands from it. This project focuses on the development and evaluation of a windows application developed with a high level programming language which incorporates speech recognition technology by utilising a commercial speech recognition engine. The speech recognition system acts as a hub within the Smarthouse to receive and delegate user commands to different switching and control systems.

Initial trails were built using Dragon Naturally Speaking as the recognition engine. However that proved inappropriate for use in the Smarthouse project as it is speaker dependent and requires each user to train it with his/her own voice.

The application now utilizes the Microsoft Speech Application Programming Interface (SAPI), a software layer which sits between applications and speech engines and the Microsoft Speech Recognition Engine, which is freely distributed with some Microsoft products. Although Dragon Naturally Speaking offers better recognition for dictation, MS engine can be optimized using Context Free Grammar (CFG) to give enhanced recognition in the intended application. The application is designed to be speaker independent and can handle continuous speech. It connects to a database

oriented expert system to carry out full conversations with the users. Audible prompts and confirmations are achieved through speech synthesis using any SAPI compliant text to speech engine.

Other developments focused on designing a telephony system using Microsoft Telephony Application Programming Interface (TAPI). This allows the house to be remotely controlled from anywhere in the world. House residents will be able to call their house from any part of the world and regardless of their location, the house will be able to respond to and fulfil their commands.

Table of Contents

List of Figures

List of Abbreviations

SAPI	Speech Application Programming Interface
CFG	Context Free Grammar
TAPI	Telephony Application Programming Interface
TCP	Transfer Control Protocol
IP	Internet Protocol
PCM	Pulse Code Modulation
LPC	Linear Predictive Coding
FFT	Fast Fourier Transform
MFCC	Mel Frequency Cepstral Coefficient
HMM	Hidden Markov Model
ASR	Automatic Speech Recognition
TTS	Text To Speech
API	Application Programming Interface
DDI	Device Driver Interface
XML	eXtensible Mark-up Language
COM	Component Object Model
SPI	Service Provider Interface
DTMF	Dual Tone Multiple Frequency
SALT	Speech Application Language Tags
IVR	Integrated Voice Response
RMS	Root Mean Square
SNR	Signal to Noise Ratio

Chapter 1: Introduction

The purpose of this project was to develop a voice recognition system, that can be used in Massey University Smarthouse to respond to the occupants needs and desires simply by taking their voice requests and transforming them into actions. The system acts as a hub that services and delegates all voice requests to other control systems within the house.

1.1 Massey Smart house

Massey University Smarthouse is a collaborative research and development project among the Institute of Information and Mathematical Sciences and the Institute of Technology and Engineering and other industry partners. The goal of the project is to create a house where technology and appliances in the house help make life easier, safer and more enjoyable for its occupants. It responds to the needs and desires of occupants by, for example, monitoring their health, adjusting lighting, temperature, or even ambient music to their personal preferences, and wherever possible assists, them in all their daily tasks. The Smarthouse main aims are:

- Monitor the health and safety of its occupants, by using the latest in information systems and biotechnology.
- Automate common house management tasks, thus allowing inhabitants to have a more enjoyable and comfortable life.
- Provide information and entertainment to the occupants upon their demand.

It should hide the technique and details of how it works and be completely intuitive to use (Human Centred Design).

The main beneficiaries of this project will be the elderly population who want to retain their independence, and their families and friends who can be secure in the knowledge that they are safe, well and comfortable. The health sector will benefit by being able to more effectively help and monitor people in their care. There will also be a number of other benefits for the construction industry, appliance industry, and for other people who wish to improve their quality of life.

The Massey University Smarthouse [1] will be a world-class showcase for the integration of house automation, health care and smart appliance technology. Figure 1.1 provides an overview of the different components of the Smarthouse that are discussed below

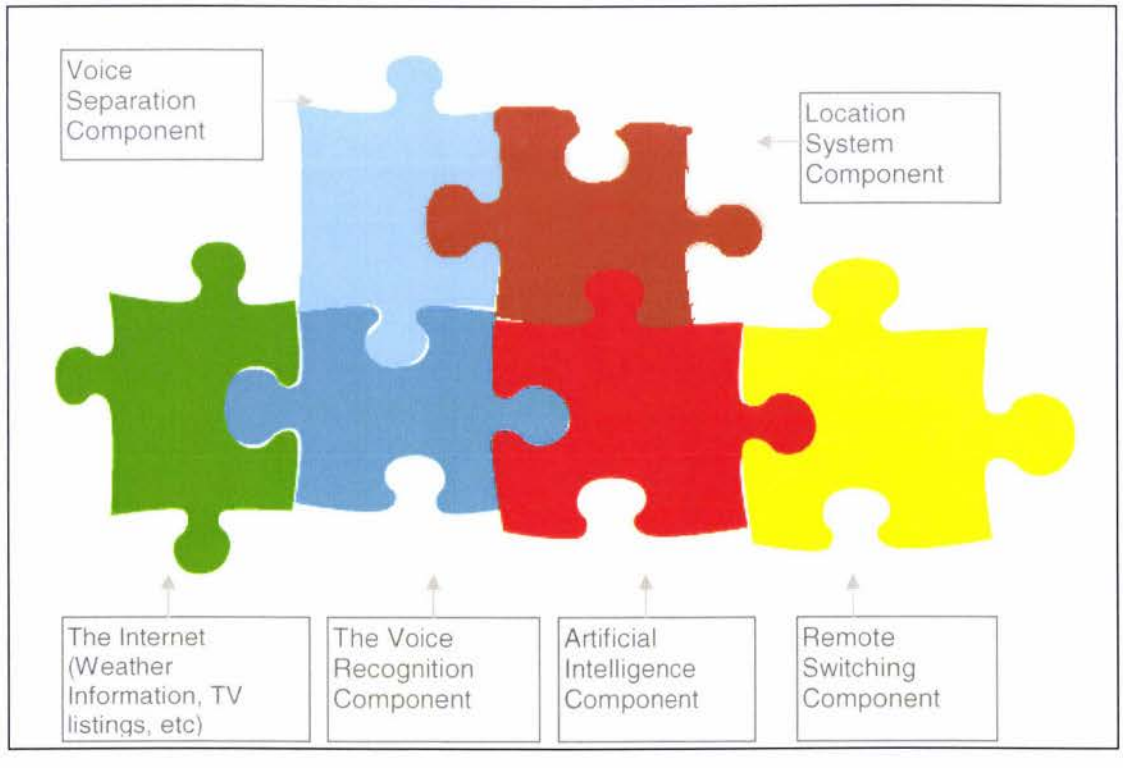

Figure 1.1 The different components of the Massey University Smarthouse and how they integrate together

1.1.1 Location and Positioning System

Tracking the position of occupants and devices within the house is essential to attain smart control and monitoring. The smarthouse therefore will be equipped with a Bluetooth ubiquitous network that consists of transceiver nodes that span across the roof of the entire house. The occupants of the house will be wearing a Bluetooth transmitting watch that contains their uniquely identifiable code that lets the house know who they are, and exactly where they are within the house.

Figure 1.2: Bluetooth watch worn by Massy Smarthouse occupants

1.1.2 Voice separation system

To enable the smarthouse to be controlled by voice, two approaches can be taken. The first is for all occupants to wear a voice capturing device, in the form of headset or watch or other. The second is to use wall or roof mounted microphones to allow for distant speech recognition. Because the first approach is restrictive to the occupants, Massey University's Smarthouse will be using beamformer microphone arrays. The development of the beamformer arrays utilise some well known beamforming algorithms to minimize noise, and provide clean, high quality speech for the speech

recognition system. The main algorithm used will be a modified version of the Griffiths-Jim beamformer.

1.1.3 House Management System

The house management system is a PC based software that contains all the rules that govern the operation of the house. It will act as the central control unit that will be communicating all the necessary information to and from other components within the house. The application will be equipped with an expert system implemented in the form of a database. The system will collect information from the location system, the speech recognition system and the different sensors within the house to manage the daily operations of the house in an intelligent manner.

1.1.4 Remote Switching System

Switching and control of appliances is made possible by a TCP/IP switching system built using an embedded system that, although capable of being used as a single and stand-alone device to aid in home-automation, also integrates into the smarthouse environment, allowing a number of smart appliances to be networked and controlled by the house management system. The device offers a simple web browser interface to show the status of connected devices at any given moment.