JOMO KENYATTA UNIVERSITY OF AGRICULTURE AND TECHNOLOGY (JKUAT)

DEPARTMENT OF COMPUTING

ICS 2406: COMPUTER SYSTEMS PROJECT

PROJECT PROPOSAL REPORT

TITLE: AN IMPLEMENTATION OF SPEECH RECOGNITION IN AN EMERGENCY APPLICATION

Author:

Name:	NGIGI SHALON NJERI	Reg. No.:	CS281-1566/2014
Submission Date:	_____	Signature:	_____
Course:	B. SC. COMPUTER SCIENCE		

Supervisor 1:

Name:	PROF. WAWERU MWANGI	Signature:	_____
Date:	_____		

Supervisor 2:

Name:	_____	Signature:	_____
Date:	_____		

PERIOD: OCTOBER 2017

ABSTRACT:

Speech recognition is one from the fastest growing technologies. It has a wide range of application and has been effectively deployed in mobile and embedded devices, providing functionalities including personal assistants such as Siri, Ok! Google and Cortana.

Speech recognition applications include voice user interfaces such as voice dialing e.g. (call home), call routing, search, simple data entry, speech-to-text processing.

The aim of this project is to design and implement a speech recognition based android application, purposed for use during emergency situations. It will allow a user to interact with it through voice commands, using preset keywords to activate it, and to navigate and seek assistance, for example, voice commands to enable calling an operator or send an automatic text to a listed emergency contact.

PERIOD: OCTOBER 2017

CONTENTS:

PERIOD: OCTOBER 2017

INTRODUCTION

Emergency response is a sector with a number of controls set up such emergency call operators, stand-by firefighting crew and equipment, ambulances and emergency rooms. However technological advancements in this area are not fully utilized. Emergency situations such as road accidents, fires, crime scenes and medical conditions including heart attacks, burns, chocking, fainting spells and seizures happen frequently around the world, and can be extremely taxing and often confusing for an individual, sometimes resulting in incapacitation and it therefore cannot always be guaranteed that a user in red-alert situation would be in a condition of handling physical input to a device to seek assistance.

Speech recognition is the inter-disciplinary sub-field of computational linguistics, that develops methodologies and technologies which enable the recognition and translation of spoken language into text by computers; Also known as "automatic speech recognition" (ASR), "computer speech recognition", or just "speech to text" (STT). It incorporates knowledge and research in the linguistics, computer science, and electrical engineering fields (Wikipedia).

It is one of the proposed methods of computer interactions, involving software and hardware, that act together to audibly detect speech and convert it into a string of words. Speech recognition is achieved by breaking down sounds the hardware detects into smaller non-divisible sounds called phonemes. “Those” forming “th”, “o” and “s”. Speech recognition attempts to match the detected phonemes with known words from a stored dictionary.

An enhancement to speech recognition systems, is to process detected words by analyzing their grammatical relationship and relative context. Natural language processing used concert with speech recognition provides a powerful tool for operating a device.

However a drawback of a conventional natural language processing, is that it may fail to determine the correct meaning of the words detected by the speech recognition system.

There is a need to create an application that integrates voice recognition into an application to provide a quick and effective way to call or ask for help during emergencies; and in cases of disablement or unconsciousness, paramedics and helpful strangers can obtain relevant medical information and emergency contacts to call from the application.

Drawbacks of speech recognition and natural language processing

One of the downsides of speech recognition includes its inability to capture words due variations in pronunciation, its lack of support for most languages outside English and its inability to sort through background noise.

Natural Language Processing may fail to determine the correct meaning of the words detected by the speech recognition system.

Performance of speech recognition is measured by accuracy and speed. Accuracy is measured by degree of error in transcription, whereas speech is measured with the real-time factor. A variety of factors can affect computer speech recognition performance, including pronunciation, accent and pitch**.**

PROBLEM STATEMENT

The main objective of this project is designing an android application that uses speech recognition to activate and navigate through, using certain voice commands. One speaks into a phone and it uses what was said as input to trigger some sort of reaction such as activating the application, calling an operator or automatically sending a pre-recorded text to contacts listed by a user.

How the application works:

The speech recognition based application will be running in the user's background, using minimal resources possible. A user can easily call for help during an emergency or when they feel like they are in danger. The application can be opened, or activated during an emergency by uttering keywords which are preset during the first configuration of the application on their phones. (The user gets to set their keywords of choice, so to keep the application from activating unnecessarily when certain words are used)

The app will recognize the keywords as either: just open the application with no further action, call the emergency call operator (such as 999 for Kenya) or send a distress message with appended details such as: GPS coordinates, a link to Google Maps, and a timestamp by SMS or email, to a previously specified contact list.

Research Questions to be addressed

1. What are the fundamentals of Speech recognition?
2. How can speech recognition be implemented?

Objectives

The main objectives addressed within this project are:

1. To understand the process of speech recognition and its fundamentals
2. To study various ways of implementing speech recognition
3. To compare various tools for implementing speech recognition

JUSTIFICATION

The main aim of this project is to design an emergency response application that can be operated using speech; to account for cases when a user is not physically able to handle their phones.

There exist applications for the purpose of alerting people of a user's call for help. Current emergency applications within the market include: Emergency+ (a national app developed by Australia's emergency services and government, which uses a phone's GPS functionality to provide emergency call-takers), ICE (In Case of Emergency – provides a way to store information needed by rescuers and doctors), Medical ID, bugle (allows user to send free notification services to family if they don't check in when they're expected to).

However, none of these applications incorporate speech recognition the module. They are usually cumbersome to use and not as user-friendly; beating the entire logic of using them. They are dependent on a user being in a capable state of handling phones and providing physical input, for the application to work, which is not always the case during an emergency.

A speech recognition based application would be a great milestone for users who may be incapacitated, and therefore not have the ability to provide physical input during an emergency situation, but would require to make a call or send a distress message requesting for assistance.

A venture into this topic is aimed at understanding how speech recognition works and tools necessary to build an automatic speech recognition system.

LITERATURE RELEVANT TO THE PROPOSAL

1. X.D. Huang, Y. Ariki, and M. A. Jack, *Hidden Markov Models for speech recognition* (University of Eidenburg)
2. Jaime Diaz and Raiza Muniz, *Voice Recognition System,* a journal article, [May 2007]
3. Bhiksha Raj, Rita Singh, *Design and Implementation of Speech Recognition Systems*, a journal article, [2011]
4. Kumbharana, Chanderesh k, 2007, *"Speech Pattern Recognition for Speech to Text Conversion"*, thesis PhD, Saurashtra University
5. Dr. Joseph Picone, *Fundamentals of Speech Recognition: A Short Course,* Institute for Signal and Information Processing Department of Electrical and Computer Engineering Mississippi State University.

RESEARCH METHODS AND DESIGN

Research methods

- **Experimental Method and Thesis Proposal.** Development of applications embedding speech recognition is a known problem with existing solutions. However, the plan is to study these existing solutions and incorporate speech recognition into the application to allow a user to use the application hands free in the case of emergencies.

  These methods will also prove useful in identifying models, methods and algorithms (Including: Hidden Markov models, Dynamic time warping (DTW)-based speech recognition, Neural networks, End-to-end automatic speech recognition), studying, implementing and testing them against the actual environment of an emergency to check for efficiency and how factors such as pitch, accents and noise affect the applications performance.

- **Build Method.** This method will consist of building an artefact, a software, to show that it is possible to design a speech recognition based emergency android application

Design

The basic design is an application implementing speech recognition to activate and navigate through it.

SCHEDULE

ACTIVITY	DATE	DURATION
Project Identification and Proposal Writing	28th September – 6th October, 2017	~ 2 weeks
Literature Review	10th October – 20th October, 2017	~ 2 weeks
Requirements Analysis	November – December, 2017	2 months
Design	December – January, 2018	2 months
System Implementation	February – April, 2018	3 months
Project Submission	April 2018	1 day

BUDGET

RESOURCE	COST
PRINTING AND BINDING	KSHS 1000
AIRTIME AND BUNDLE PURCHASE	KSHS 4000
FLASH DISK FOR STORAGE	KSHS 2000
TOTAL	**KSHS 7000**

CONCLUSION

Understanding the limitations as well as the strengths of speech recognition, is important for effective use of speech input and output in design of a user interface. An application with the capability of using speech for control, might be the definition of life and death; a critical moment for a person without the physical capability to use their phones effectively during an emergency or a situation of danger.

REFERENCES

[1] *Introduction to speech recognition* – http://www.sestek.com/2014/11/introduction-to-speech-recognition/

[2] Nikam Tanmay, Gangurde VIdya, Patil Chaitalee, Vidya Kawtikwar, *Personal Safety App using Speech Recognition API,* International Journal of Application or Innovation in Engineering and Management

[3] Pratibha Channamsetty, Shruthi Samba, *Speech recognition for mobile systems*

[4] Jose Nelson Amaral, *About Computing Science Research Methodologies* https://webdocs.cs.ualberta.ca>readings

[5] Speech Recognition, From Wikipedia https://en.wikipedia.org/wiki/Speech_recognition