

Performing Computer Operations with Voice Controlled Mouse

Sarita

Abstract— The goal of Human Computer interaction (HCI) is to design a system that improves the interaction between computers and users, mainly for people with disabilities. For this reason, this research describes how to make computers more usable by using speech based interaction. Proposed system will use speech and non speech characteristics of human voice for hands free computing. System will use human vocal characteristics to command mouse pointer. Words and sounds will be used to perform mouse 2-D movement and performing operating system's operations. By doing this people with disabilities can use the computer systems normally as the normal users use it. Current speech recognition systems provide discrete interaction because they process user's vocal utterances at word level, proposed system will be able to perform mouse-like operations fluidly.

Index Terms— Human computer interaction, Speech recognition, Linear Predictive Coding, Artificial Neural Network, Hidden Markov Model

1 INTRODUCTION

Human computer interaction (HCI) field has much recent advancements. HCI field has been evolved to make easier and enjoyable computer access. In the past several decades great amount of research has been done in speech recognition technology. Advances in speech recognition have enriched HCI area but human speech characteristics still remains largely unexploited. In modern computer interactions, mouse control has become a crucial aspect.

Recently, Artificial neural networks have been considered for speech recognition, there performance is better than the traditional approaches. Mouse is one of the best and pervasive input devices since the emergence of GUI. There are many people in the world who are suffering from physical disabilities. Only in the United States there are 700,000 people with disabilities of spinal cord [1]. These people can't use the computers the same as the normal users do. They can't use the standard input devices. So voice based system will make their life easier.

Current systems work well when user needs to execute a command or enter textual information but when user wants to perform operations like double clicking, picking file and dropping to another folder, selection of unnamed objects or arbitrary points on a screen, continuous path following(e.g., in games and drawing), dragging, scrolling and so forth, they do not work efficiently.

Proposed system will be different from the existing systems because this system will provide immediate processing of vocal input. The key benefit is that it will instantaneously and continuously process the vocal inputs. System will use vowel sounds which will make it possible to move mouse cursor instantaneously as vocal input is provided to the system. It will process continuous vocal characteristics and transform them into mouse cursor movement. It will provide highly responsive interaction, when user's vocal characteristics are changed then corresponding results will be reflected immediately upon the interface. The system will be conceptually different than the other systems in several respects. It will allow the user to make instantaneous directional changes using voice. In the existing systems when entire word is recognized then mouse cursor start movement but in our system mouse will start moving as soon as some vocal input arrives.

2 LITERATURE REVIEW

Jeff A. Bilme, Jonathan Malkin, Xiao Li, Susumu Harada [1] in 2006 proposed a system called "vocal joystick" using non-speech voice features such as pitch, energy and vowel quality to control mouse movements. Their system was beyond the capabilities of discrete speech sounds, it exploits the human vocal system characteristics such as loudness, vowel quality etc, which are then mapped to the continuous mouse control parameters. This system was developed in C++ and works on windows and Linux. No hardware cost involves, only a good quality microphone was required. In this system pattern recognition was done by dynamic Bayesian network and vowel classification using Multilayer Perceptron (MLP). For discrete sound recognition they used Hidden Markov Model which help

in reducing the false positives. It was better from existing systems because it doesn't wait for entire word to be

• Sarita is currently pursuing masters degree program in computer science engineering in Lovely Professional University, India, E-mail: sarita90sharma@gmail.com

spoken instead as soon as system receives the vocal sound mouse movement start.

Frank Loewenich and Frederic Maire [III] in 2008 combine two techniques called as face recognition and speech recognition. In the previous paper all non-vocal characteristics needed to be exploited but here mouse movement taken the benefit of face recognition also. Additional hardware was required in this system, a web-camera. System was based on the image processing algorithms and speech recognition technologies. The algorithm used for the face detection was based on AdaBoost Classification. Head tracking was done with the implementation of Lucas-Kanade optical flow algorithm. For speech recognition they used the Microsoft speech SDK 5.1 which was freely available. This system provided the users to perform click-drag and drag-drop operations which were not available in the existing systems.

R.L.K.Venkateswarlu, R. Vasantha Kumari, G.Vani JayaSri [IV] in 2012 used the Artificial Neural Network (ANN) techniques for speech recognition. They designed the voice controlled computer interface that enables the individuals with motor impairments to use the system efficiently. This system was divided into six main blocks: keyboard and mouse layouts, speech recognizer, keyboard and mouse circuit, microphone, microcontroller and mouse control structure. To overcome the limitations of the present systems they considered the hybrid approach here, in this user control the mouse using voice and using voice characteristics. Various NN techniques were used in the proposed system. Radial Basis Function (RBF) used for the pattern classification and decoder used the Hidden Markov

Model graphs to recognize the text. This system does the same work as explained in the previous two papers but the difference was that it used the neural networks techniques which include the advanced technology. Due to their massively parallel structure they provide robust systems and better and fast classification of the patterns using various techniques for pattern classification for example, Multilayer perceptron and Radial Basis Function.

3 PROBLEM FORMULATION

Our idea is to construct a system which will take human voice as input, using this voice system will move mouse pointer. Further this feature can be used to perform computer operations such as double clicking, picking file and dropping into another folder, selection of unnamed objects or arbitrary points on a screen, continuous path following(e.g., in games and drawing), dragging, scrolling and so forth. Our primary focus will be on non speech characteristics of human voice such as pitch, loudness and vowel quality to make mouse movements more effectively and rapidly. Mouse pointer movement will be occurring instantaneously and continuously as soon as voice input received by the system. This system will help physically disabled people as well as general users to use the enhanced system functionality. Various operating system operations will be controlled and done by human voice. Non speech characteristics will be used because they result in rapid movement of the mouse. They have various benefits over spoken words because using spoken words speech recognition will be done after the entire word has been arrived but using vowel sounds and non speech characteristics of the human voice mouse movement begins instantaneously.

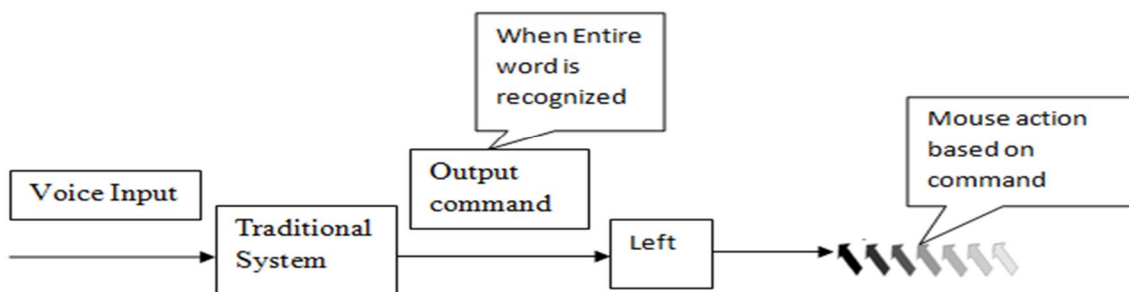


Fig1: Traditional System Working

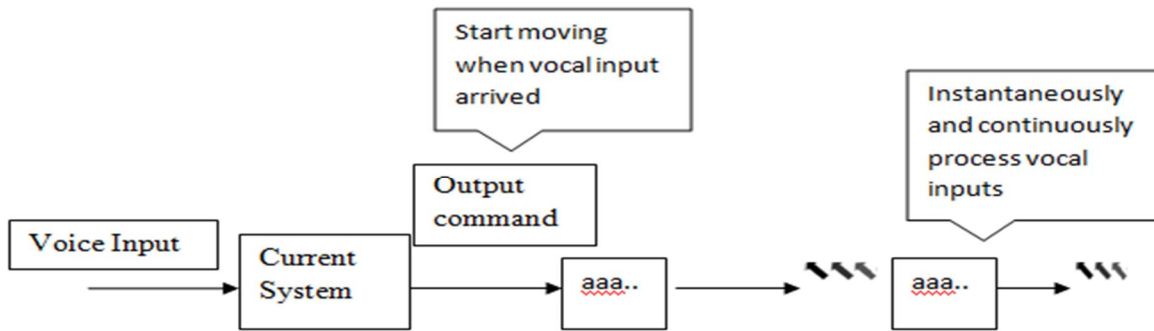


Fig2: Proposed System Working

4 OBJECTIVES

Main objectives of our research are given below:

4.1 Voice Controlled Mouse pointer

Controlling mouse with human voice is main motive which will provide the benefit to the people with disabilities to use computer in more effective and efficient way.

4.2 To use speech and non speech characteristics of vocal system

Existing systems are developed using speech or on the word recognition, proposed system will provide extra functionality to the system by controlling mouse pointer using vowel sound and cursor's direction and velocity will be decided by the vocal characteristics such as pitch, energy, loudness etc.

Using both speech and vowel sound commands to the system, needed action should be performed. Our main concern will be with non speech characteristics of human vocal system to perform operating system operations. Because limited research has been done in the area of non speech characteristics of human voice so we will use non speech voice characteristics such as pitch, vowel sound and energy to efficiently commanding computer to perform operations

For example:

- Double click
- Dragging
- Picking file from one folder and drop to another
- Continuous path following(e.g., in games and drawing)
- Scrolling

5 RESEARCH METODOLOGY

The proposed system will consists of various modules. Each individual module consists of different techniques and algorithm. Previous module output becomes input for the next module.

Following modules will be in our system:

Module1:- Signal Processing

Module2:-Feature Extraction

Module3:- Pattern Recognition

Module 4:- Motion Control

Module 5:- Mouse Movement

Module6:-Performing Operations

These are the general modules in our system. First three modules are core module of our system. Output of these modules will result in the mouse pointer movement. Desired computer operations will be performed after motion control module is ready.

Each module will use different algorithms and techniques to perform their operations. Signal processing module's output will become the input to the feature extraction module; output of feature extraction module will be used as input to the feature recognition module; feature recognition module will be used as input to the motion module which will perform the mouse motion. Its output will become input to the module 6 which is performing computer operations. User will give voice input which will be processed by the system and then extract the speech characteristics of the system.

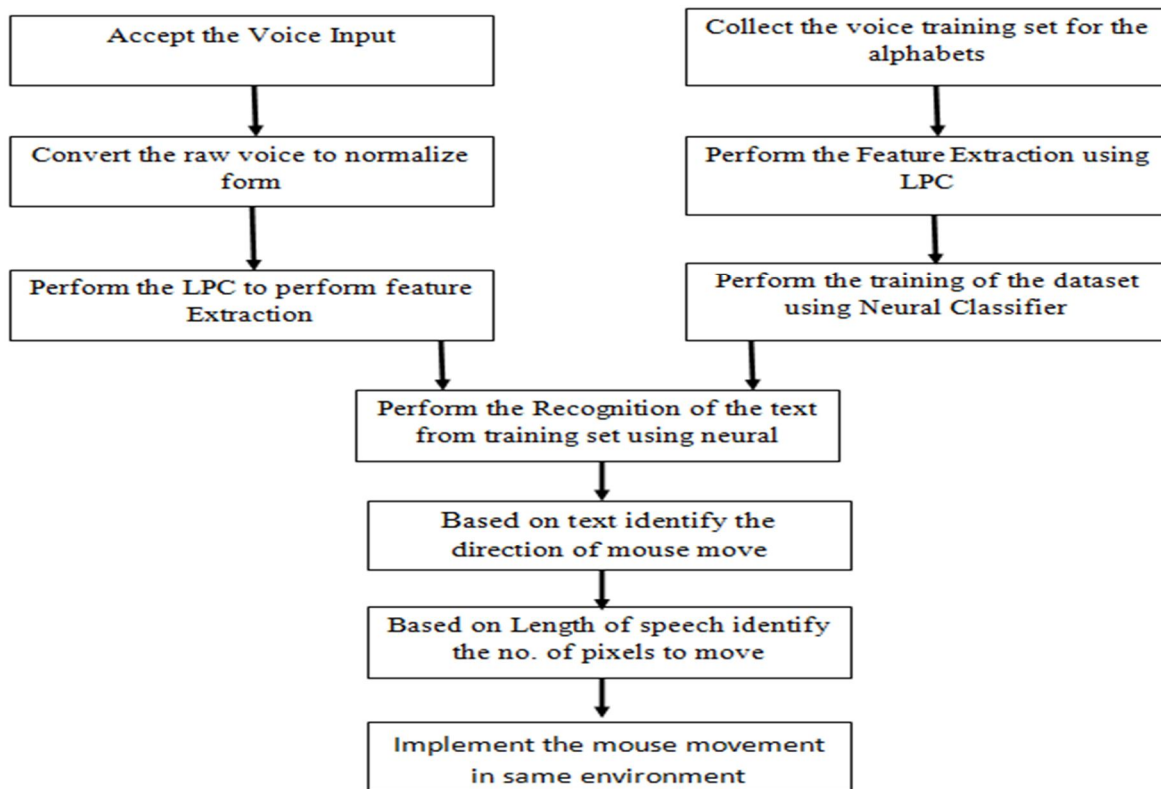


Fig3: Work Flow

6 CONCLUSION & FUTURE SCOPE

Proposed system will provide instantaneous and continuous mouse movement. This will minimize the time to perform various operations which need continuous input.

Implementation of this system can be continued further by using combination of discrete and continuous commands, which will help to minimize more time to perform operations.

REFERENCES

- [I] Jeff A. Bilmes, Jonathan Malkin, Xiao Li, Susumu Harada, Kelley Kilanski, Katrin Kirchhoff, Richard Wright, Amarnag Subramanya, James A. Landay, Patricia Dowden and Howard Chizeck (2006) " *The vocal Joystick*", in Proc. IEEE Intl. Conf. on Acoustics, Speech, and Signal Processing.
- [II] Sandeep Kaur (2012) " *Mouse Movement using Speech and Non Speech Characteristics of Human Voice*", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249 – 8958, Volume-1, Issue-5.
- [III] Frank Loewenich and Frederic Maire (2008) " *Motion-Tracking and Speech Recognition for Hands-Free Mouse-Pointer Manipulation*", Queensland University of Technology Australia.
- [IV] R.L.K.Venkateswarlu, R. Vasantha Kumari, G.Vani JayaSri (2012) " *Speaker Independent Recognition System with Mouse Movements*", International Journal of P2P Network Trends and Technology- Volume-2, Issue-1.
- [V] Tyler Simpson, Colin Broughton, Michel J. A. Gauthier, and Arthur Prochazka (2008) " *Tooth-Click Control of a Hands-Free Computer Interface*", IEEE, Transactions on biomedical engineering Volume-55, NO. 8.
- [VI] Piotr Dalka and Andrzej Czyzewski (2010) " *Human-computer interface based on visual lip movement and gesture recognition*", International Journal of Computer Science and Applications, Techno mathematics Research Foun dation Volume- 7 No. 3, pp. 124 – 139.