

SE 705 Human Computer Interaction

Project Phase One Report

Speech Controlled PACS

703272 - Hüseyin Candan

1579044 – Atakan Küçükatalay

Instructor: Dr. Tolga CAN

Motivation

Using speech for Human-Computer Interaction is a research area since the inventions of the computers. There are many applications to use speech for supplying information exchange between computer and human. This communication has two directions: computer to human and human to computer. Speaking computer applications seems to be easier to develop than speech recognition which makes it more common in the industry. As an example, today it is possible to listen the news in Turkish from the web site of a newspaper (e.g. Sabah). On the other hand, there are also many studies and also successful applications to convert human voice to text. "Dikte" and "GVZ" are most popular speech recognition applications in Turkish. Especially, "Dikte" has a version called as "Tibbi Dikte" which is widely known and used by radiologists to prepare radiology reports by just talking to a microphone.

Our Application

In our project, we want to use Dikte's tool to provide a speech controlled user interface for a web-based Picture Archiving and Communication System (PACS) application. PACS's are applications to retrieve medical images from various medical imaging instruments (e.g. ultrasound, magnetic resonance, mammograms, etc.), store and present these images in a required format. We have already developed a PACS which has a web-based user interface and in this project we will make it available to control some functions of this application by speech commands.

Potential Users

Potential Users of our tool is same as the users of the PACS which includes radiologists and technicians. Although the potential users usually have very little experience on computer usage, many of them are using PACS's for their work. Especially the radiologists either using "Tibbi Dikte" or have knowledge about it, that makes our job easier to inform them about the conceptual model of our implementation.

Hardware and Software

There is a demo application available called "duybeni", which makes it possible to control some functions of Windows XP by speech commands. In our case, platform will be web-based and we will embed Dikte API into the web browser using ActiveX technology. We will also use Javascript to provide functional support for the commands retrieved by the ActiveX and to provide communication between client side requests and application server.

Development environment includes MS Visual Studio for building web application, Borland Delphi for ActiveX and MS SQL Server for RDBMS.

Actually, there is no need for a special hardware to use the application. But, "Dikte" built a special microphone set to increase the success of speech recognition. This device includes two microphones, first one is naturally for the user and the second one is to get environmental noise and subtract it from the voice recorded from the first microphone. This subtraction mechanism which is used to eliminate the environmental noise from human voice provides input of high quality to the speech recognition engine which increases the success rate of speech recognition.

Objectives

In fact, the goal of our project is not to make the radiologists and technicians use PACS systems via speech commands, rather we want to show that it is possible to use speech recognition for Human-Computer Interaction even for web applications, and it is easy to make the users to get used to control the software via speech. Today, the usage of the speech recognition applications are so limited especially in Turkey, but it is clear that, as the success of the speech recognition tools and user-friendliness of the applications using such tools increase, more and more users will control their applications including Operating System by sound commands.