

Design of a Sound Sensitive Motorola IPTV Set-Top Box

By: Meng-Chieh Ho, David Lewis, Devon Yang

Advisor: Priya Narasimhan

Motivation

This system is designed to provide the ultimate experience for a typical "couch potato." It eliminates the need for a remote and performs dynamic actions on the television based on sounds occurring near the set-top box. These features provide a comfortable viewing experience for the viewer.

The system is able to simulate the actions from a typical remote control. It can also detect sounds and volume changes in the environment to dynamically change the television settings based on what is going on in the room. For example, if the room becomes noisy, the volume will be adjusted accordingly.

Example Scenario:

1. A television is on and a conversation is taking place.
2. The noise level in the room starts to grow.
3. After some time, the volume of the television goes down because the system realizes people are talking.
4. The conversation ends.
5. The volume returns to normal since the conversation is over.

Development Environment

Hardware:

Motorola VIP 1970 Set-Top Box (STB)
High Quality Microphone
Linux-Based Computer acting as an Infocast Server



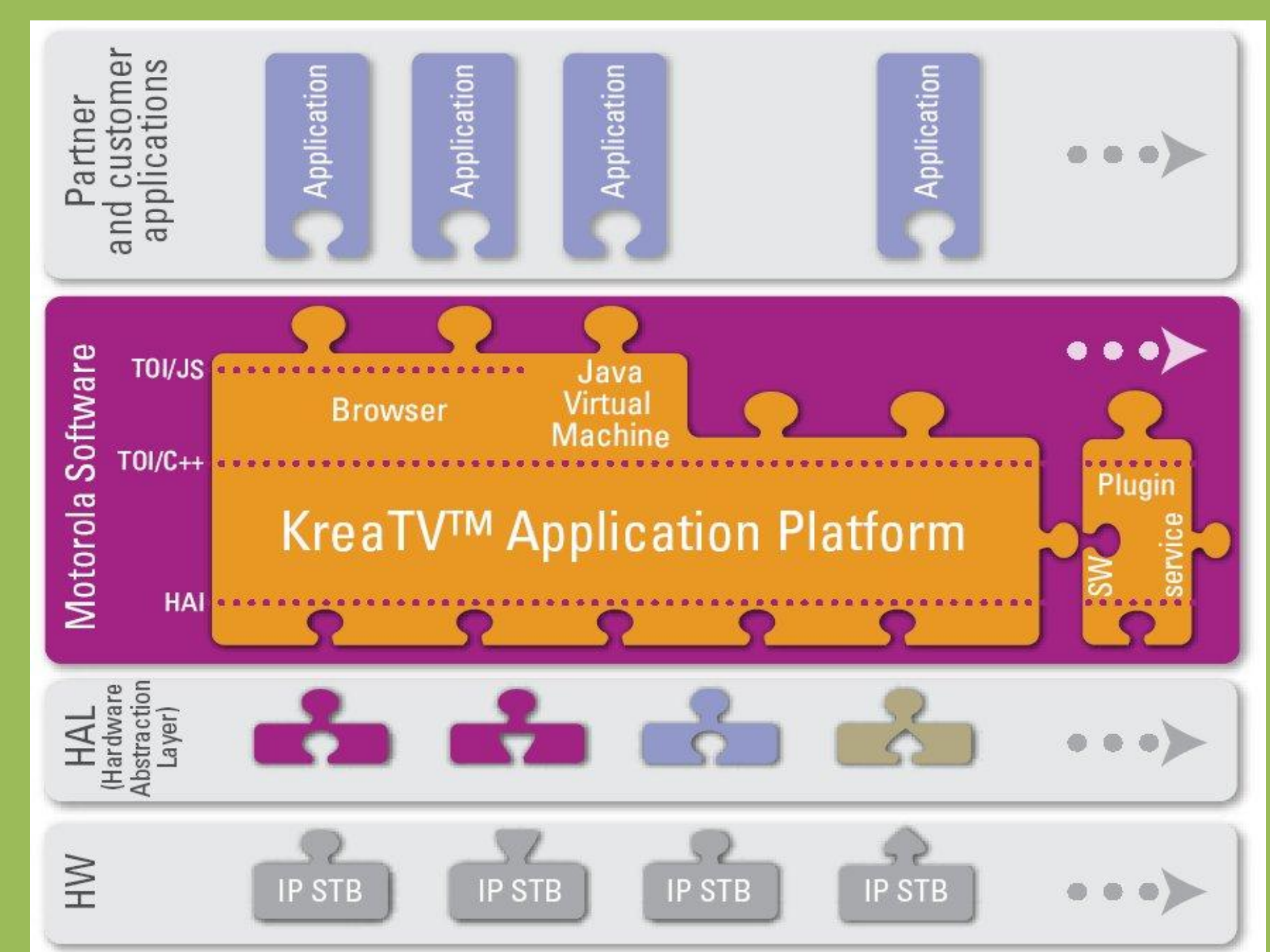
Software:

Java Development Environment
CMU Sphinx-4 Speech Recognition Framework
JavaScript Language with AJAX
KreaTV Application Platform

Existing

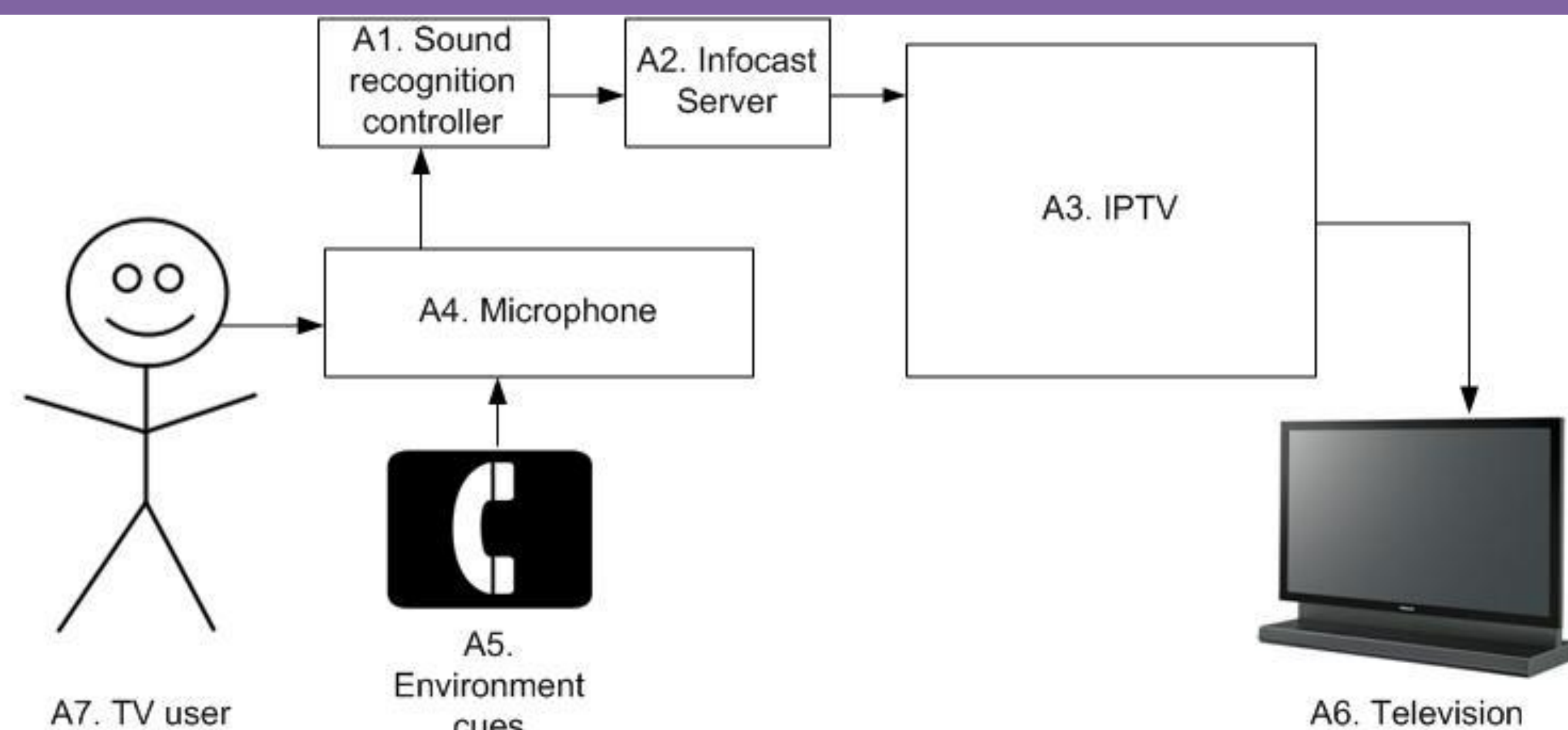
Infrastructure:

Network with DHCP server



[http://www.linuxdevices.com/files/misc/motorola_kreatv_arch.jpg]

Architecture

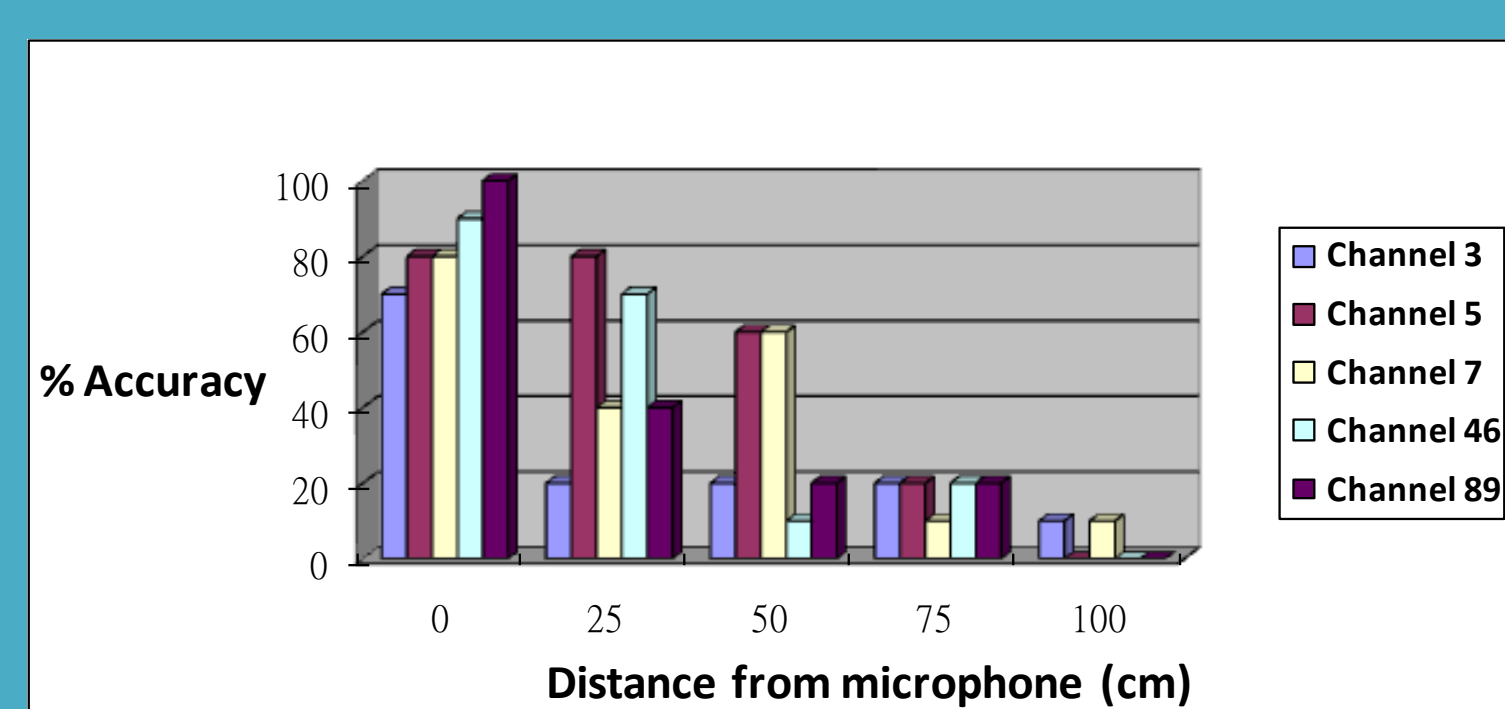
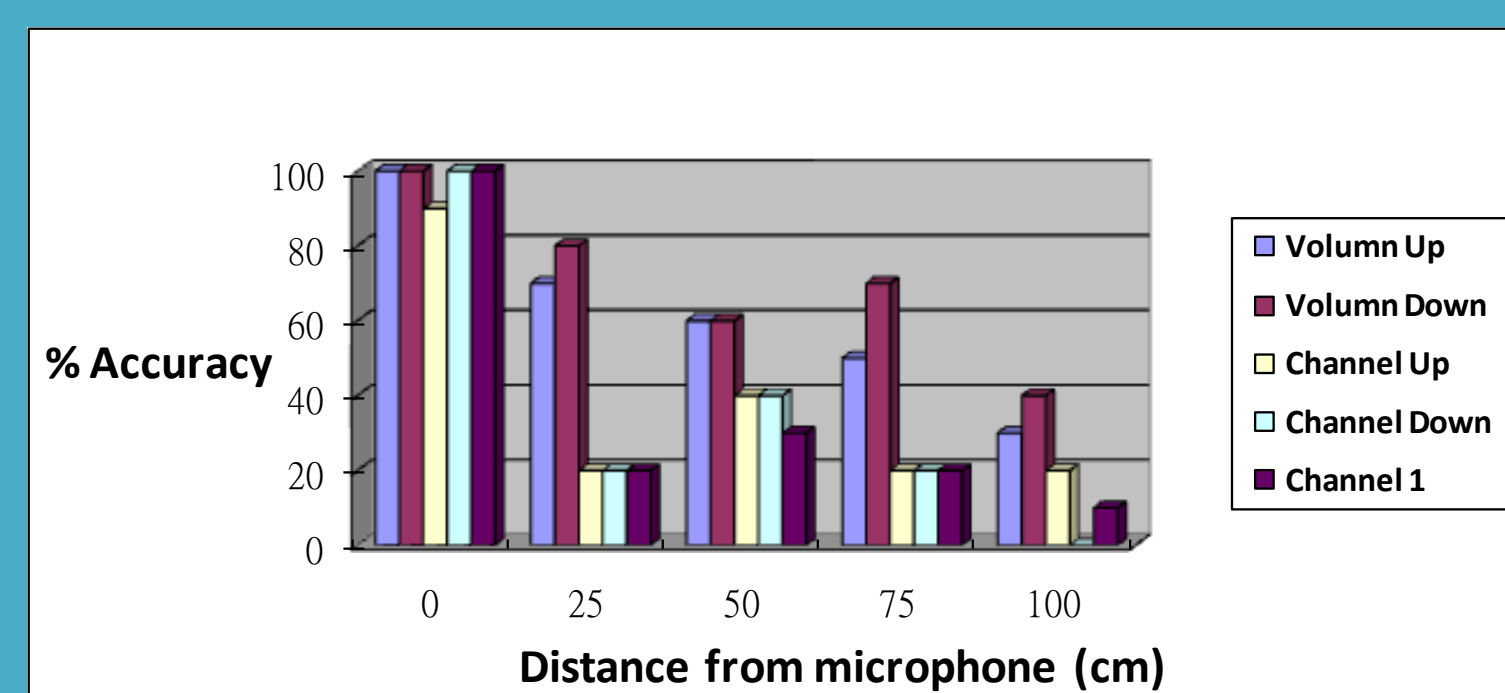
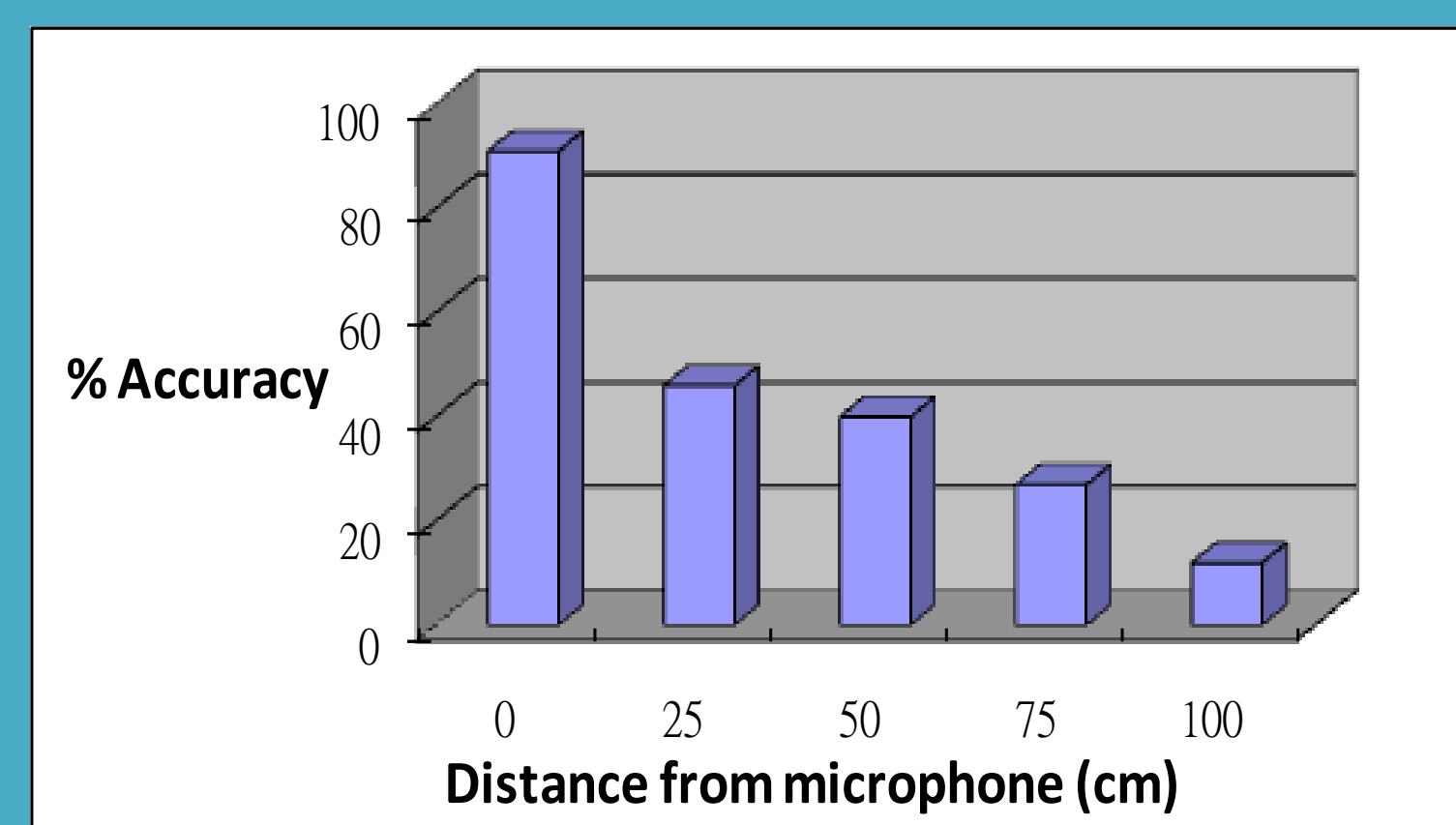


- A1. Sound recognition controller to dynamically control the TV settings
- A2. Infocast server to broadcast information to STB
- A3. Set top box used to drive television controls
- A4. User microphone to recognize human voice commands and ambient noise
- A5. Environmental cues to drive the ambient noise queues
- A6. Television to deliver output commands to
- A7. Television watcher to drive the voice commands

User and environmental cues are sent through a microphone to a sound recognition controller. This controller uses the Sphinx-4 Speech Recognition Framework to communicate desired commands to the Infocast Server. The Infocast Server communicates with the STB through the network to provide video stream locations and system information on boot up. Once the STB is running, commands are streamed to the IPTV via AJAX through a Portal Page that has been created to support this system.

Results

The following are graphs generated from test data to gauge the accuracy of the system.



The graphs on the left show that our system yields very high accuracy when commands are given within a short distance from where the microphone is located. Testing was done with one male and one female voice sampled ten times per command, each.

Response time from the time a command is issued to the time that the command is acted upon by the STB is comparable to those issued by a standard IR remote.



Carnegie Mellon

