

A Framework For Automation Using Networked Information Appliances

NTTMIT 9904-05

Progress Report: July 1, 1999—December 31, 1999

MIT: Srinivas Devadas, Larry Rudolph NTT: Satoshi Ono

Project Overview

The proliferation of diverse and heterogeneous information appliances has produced a new set of problems, challenges, and opportunities. We are building a development infrastructure for information automation that combines the synergistic use of sensors, standard hardware interfaces, intelligent software, and adaptive algorithms in order to rapidly produce intelligent, networked information appliances. Using our infrastructure, we will automate many information-intensive tasks in office and home automation.

We are building a computing platform that consists of locally networked computational elements attached to sensors and actuators. The sensors provide image, audio, pressure or position input. The computers control actuators within appliances and also interface with other appliances. Not only is intelligent application software fully customizable and adaptable to user preferences, but it also makes use of our middleware that seamlessly integrates new custom or commodity sensors and actuators. This framework of computers, sensors and intelligent software allows for the rapid deployment of a wide variety of information automation problems. We believe that the deployment of such a solution should take no more than a week, including customizing hardware and writing software applications.

Progress Through December 1999

Software:

As a first step in building our automation platform we focused on building interface infrastructure for voice control of appliances (physical objects such as a phone, remote control) and software applications (information objects such as a web browser, email program). We are working with the SLS Group and using SLS-Lite for speech recognition and synthesis.

The speech server is first configured by automation software for the particular task it is required to perform. If the task is that of reading email, then a certain set of commands such as open, close, and read message are used. If the task is that of contacting a particular person, the database of names along with a few actions is used to configure the speech server. The speech server determines what command has been spoken from the list of commands included in the configuration, and the automation software performs the remaining tasks of understanding the command, maintaining system state, and controlling the physical or information object.

We have setup a networked Linux workstation with a speaker, microphone, infrared controller and camera, and are running our automation software on this workstation. Interface software to the speech server is almost complete, and input/output interfaces for the audio and video appliances have been written. Simple automation software that determines which number to call when calling a person from amongst his many numbers based on the person's schedule is operational. Full voice-based control of appliances that can be controlled by the infrared (IR) controller is the short-term target.

Hardware:

Eventually, the automation software will run, not on a workstation, but on a MASC computer card, which can be embedded directly into appliances. This computer card with a ARM processor, 8MB of DRAM and 2MB of Flash

has an interface FPGA on it that allows it to electrically control different appliances. This card can be embedded into a variety of different appliances and software running on the processor can control the appliance directly rather than through low bandwidth IR links. We have two prototype cards working, and are in the process of redesigning the card and fabricating 20 new cards. One of the prototype cards has been embedded into a universal remote and used to control a TV. Programming can be downloaded into the card, and the TV can be controlled completely automatically.

Research Plan for the Next Six Months

In order to control appliances such as TV, or VCR, we have to implement the IR interface drivers on Linux. The appropriate codes for different vendors have to be collected into a database that can be accessed by the automation software. Once this is done, we can control these appliances as well as receive inputs from them. The configuration of the speech server prior to interacting with these appliances will enable full voice-based control of these appliances.

We will target applications such as intelligent navigation and intelligent Web/database search and implement automation scripts for these applications. Automation scripts will sequence through combinations of automation commands while interacting with the user to perform complex tasks. The automation software will first be tested and debugged on a Linux workstation, and then on MASC cards.