Area: APPLICATIONS Start Date: June 1, 1999 Status: Renewal Date: Mar 1, 1999

Title: Ubiquitous Video (and Audio) in the Home.

Investigators: Irfan Essa, G. Abowd, and C. Atkeson
College of Computing,
Georgia Institute of Technology

Summary:

The ubiquitous video/audio project explores potential implementations of video and audio technology that seeks to become transparent to everyday activities. As a part of this project, we are developing video and audio processing techniques and exploring their applications within the context of an "Aware Home." During the first year of the project, we developed prototypical systems for high and low bandwidth video and audio processing. In the second year we propose to explore and prototype a spectrum of applications in the home that uses these systems.

Impact on Technologies and Applications:

We expect the use of video in the home to grow from isolated applications such as videophones, smart front doors, and surveillance devices to become pervasive. Such pervasive use of video and audio technology will make the entire volume of the house viewable, and permit synthesis of desired views from existing views. Once we have established an infrastructure to support such a pervasive use of video and audio technology, a homeowner will be able to install video hardware for one purpose and use it for many others. For this reason, we expect an explosion of demand for video applications. This proposal attempts to anticipate that demand and explore the technologies to provide ubiquitous sensing of video and audio and explore candidate video applications now. Making video and audio sensing transparent to everyday activities and distributing it everywhere in an environment is essential in the development of an "aware home."

There are three different classes of video and audio processing technologies that we are exploring:

1. Low cost, low bandwidth technologies that emphasize the transmission of single images or low frame rates. These applications could use power lines or wireless links to transmit images. Camera does little processing besides reformatting data for transmission.

   In the first year of the project, we have worked with the Rhino Board (http://www.cc.gatech.edu/fce/uvid/rhino.html) for this purpose. We have used the CEBus architecture and powerline modem technology for communication and a V25 microprocessor - Superset of Intel 8086 instruction set. We are also considering using other kinds of microprocessors to support this effort. These devices can be used for applications like security systems and automatic door entry.

2. High frame rate applications that require the processing and transmission of images at 30 Hz or better. This is specifically important in applications where high fidelity information is required to track activities.

   In the first year of the project, we have relied on Pentium II PCs with analog color video cameras and frame-grabbers to develop these applications. We have developed various computer vision algorithms that run on these desktop workstations to track people and recognize their activities. Due to high bandwidth requirements, these applications will probably require next generation home cabling. Some options of this include IEEE 1394 (firewire), gigabit ethernet, high bandwidth wireless, etc.

3. Camera and sensors with local onboard processing. These applications use task specific image processing right at the camera. This increases the cost of each installed camera, but decreases the requirement for high bandwidth information transmission.

   So far, we have kept away from this technology with the hopes that good commercial solutions will be available soon. The methods used in high bandwidth applications will directly apply to these specialized cameras.

Example applications we will explore within the house this year are:

Video intercom: Allow users to communicate within a house using a videophone like device. We would also like to develop a version of this device that did not require the user to orient himself with respect to a camera and pick up a handset. Appropriate views of a mobile user can be synthesized from multiple cameras, and sound pickup and generation can be localized using phased array techniques.

Voice or gesture command: Allow users to control devices in the home using voice or gestures without requiring a hand held device or orientation to a particular camera.

Surveillance: Monitoring children, elderly, and pets anywhere in the house. Simple techniques can be used to identify people (out of a very limited set of approx. 5 people making up a household). Views can be displayed on correspondingly ubiquitous display or projection technology for the observer.

Intelligent surveillance: To the extent we can make the system intelligent, we can (1) detect when an individual has entered a particular area (keep pets off couch, keep children out of garage or shop or away from stove). (2) detect potential problems (loud noises, silence/non-movement for too long, body in unnatural position on floor, variations in normal behavior patterns, pot boiling over on stove).

Entertainment: There are many interesting possibilities of making the house an extension to a children's toy or video-game.

The work proposed here has direct impact to telecommunication technologies and for applications in home automation and surveillance. The recent growth in computational power, plus an added infrastructure that exists on regular computers to deal with audio, video and other multi-modal signals has brought the applications of automatic analysis of these signals to the forefront in the recent times. The potential impact of this research has led several major industrial research labs to open dialogue with us on these projects.

We have already discussed this project with researchers at Microsoft (Steve Schafer, Turner Whithead), Intel (Bob Liang, Mark Holler), IBM (Mark Lucente), Bell South (Lee Friedmann, Scott Stillman), MERL (Bill Freeman), DEC-CRL (Jim Rehg), TI, British Telecom, several other leading places.

Progress from last year:

We undertook a two tiered approach towards building ubiquitous video devices. One was aimed at low end, low bandwidth devices. A report for this work is available from http://www.cc.gatech.edu/fce/uvid/. The other was aimed at higher-bandwidth, faster processing systems based on desktop PCs. The progress on this end is available from http://www.gvu.gatech.edu/perception/projects/UVA/. The publications describing the techniques developed to aid in achieving ubiquitous video and audio are available from the above WWW pages.

Staffing:

- Irfan Essa (Assistant Professor, College of Computing)
- Drew Steedly (Ph.D. Student, College of Computing)
- Darnell Moore (Ph.D. Student, ECE Dept)
- Scott Stillman (Ph.D. Student, ECE Dept and Staff, Bell South)

Relation to Other Areas and Centers:

Large part of this effort is of interest to the GVU Center and the Center for Signal and Image Processing.

Budget (Tentative):

Personnel [approximate]

Irfan Essa	1 month salary

Drew Steedly	1 year GRA @ 50% plus tuition
Darnell Moore	1 year GRA @ 50% plus tuition

Equipment:	
4 PCs for Signal Input and Analysis	$16,000
2 Sony EVI-G20, pan-tilt-zoom cameras	$ 2,000
2 Omni-directional Cameras	$ 8,000
1 Microphone Arrays with Mixers	$12,000
Contact Sensors from Infusion Systems	$ 2,000

Total (Equipment)	$40,000