

Ubiquitous Sensing for Smart and Aware Environments: Technologies towards the building of an Aware Home.

Irfan A. Essa

Future Computing Environments | GVU Center
College of Computing, Georgia Institute of Technology
Atlanta, GA 30332-0280
irfan@cc.gatech.edu | www.cc.gatech.edu/~irfan

Introduction:

As computing technology increasingly becomes a part of our daily activities, we are required to consider what is the future of computing and how it will change our lives. To address this question, we are interested in developing technologies that would allow for ubiquitous sensing and recognition of daily activities in an environment. Such environments will be aware of the activities performed within it and will be capable of supporting these activities without increasing the cognitive load on the users in the space. Towards this end, we are prototyping different types of smart and aware spaces, each supporting a different part of our daily life and each varying in function and detail. Our most significant effort in this direction is the building of the “Aware Home” at Georgia Tech.

In this position paper, we outline the research issues that we are pursuing towards the building of such smart and aware environments, and especially the Aware Home. We are interested in developing an infrastructure for ubiquitous sensing and recognition of activities in environments. Such sensing will be transparent to the everyday activities, while providing the embedded computing infrastructure an awareness of what is happening in a space. We expect such a ubiquitous sensing infrastructure to support different environments, with varying needs and complexities. These sensors can be mobile or static, configuring their sensing to suit the task at hand while sharing relevant information with other available sensors. This configurable sensor-net will provide high-end sensory data for sensor processing and fusion. We are also building signal understanding methods to process the sensory data captured from these sensors to recognize the activities in the space.

Research Goals:

A significant aspect of building an aware environment is to explore easily accessible and more pervasive computing services than available via the traditional desktop computing. Computing and sensing in such environments has to be reliable, persistent (always remains on), easy to interact with and transparent (user does not know it is there and does not need to search for it). The environment has to be aware of the users it is interacting with and be capable of unencumbered and intelligent interaction. Building of a computing infrastructure that supports all these needs is therefore one of our primary goals. Additionally, we will also develop techniques for processing and analysis of the captured sensory streams that will provide the interpretation of the data and make the environment aware. We expect to rely on rich multi-modal sensory data to provide high-end awareness of what is happening in the environment.

Following are the significant aspects of our research effort in the areas of ubiquitous sensing and recognition.

Ubiquitous Sensing: The concept of ubiquitous computing seeks to develop a distributed and networked computing infrastructure to support user activities, while remaining transparent to the users. Ubiquitous sensing, supports this concept by exploring potential implementations of sensing technology that provides sensing abilities to the ubiquitous computing infrastructure without increasing the burden on the users. A network of sensors that is configured with a network of processing devices can yield a rich multi-modal stream of sensory data. Sensory data, in turn is analyzed to determine the specifics of the environment and provide context for interaction between co-located and distributed users and environments. Such analysis is also useful to determine what is happening in an environment, so that it is supported effectively.

We are specifically interested in high-end multi-modal sensors that provide rich spatio-temporal information about the environment. We are interested in instrumenting the environment and the user with cameras and microphones to extract such level of sensory data. In the next section, we discuss how content is extracted from the data streams from these high-end sensors. Here, we discuss how these types of sensors can be used to instrument a space. In addition to the video and audio sensors, we are also working on augmenting the environment and the user with other forms of sensors that will share information with each other.

Following are some of the features that are important in developing ubiquitous sensing for an aware environment.

Self-calibration: The sensors in an aware environment need to be able to calibrate themselves and adapt to the environment as needed. All the sensors in an environment need to communicate their state and their coverage area to each other and develop a model of the environment. Once the static sensors are calibrated to a given environment, they can then communicate with the mobile sensors and provide them with information that they perceive and allow them to be calibrated too.

Networking: Combinations of processors and sensors needed to build aware environments require an elaborate networking infrastructure. This infrastructure needs to support both high-bandwidth and low-bandwidth data transmission as determined by context and sensor/processor abilities – sometimes video and audio needs to be transmitted, while sometimes only extracted label need to be transmitted. It would be ideal to have these sensors in the space to set up a network dynamically as they are installed. Once a dynamic network is established then the sensors can transparently capture relevant streams and share with other sensors or processing engines. We are working with researchers in networking to develop an infrastructure to support such computing and networking needs.

Distributed Computing: In order to install all the above-mentioned ubiquitous computing services in our aware spaces, we need to study and develop a computing infrastructure to support these services. This infrastructure will serve as the brain for the environment where all the information regarding the space is processed. We will develop an abstraction of a virtual processing center for the space. The virtual center will connect various processors distributed throughout the space and allow for transparency in terms of processing and responsiveness.

Optical and audio sensors: We are interested in using video and audio sensors as high-content sensors. Traditionally, these sensors are considered as recording devices, however, these sensors carry a lot of content that is essential for interpretation of the activities in an environment. If context and the task permit, these sensors can also serve the purpose of recording interactions and allow for face-to-face interactions with spatially separated users.

Mobile sensors: In addition to the static cameras and microphones, we also envision mobile sensors. These could be cameras and microphones either worn by the users or distributed on mobile platforms allowing the cameras to move, pan, tilt and zoom. These sensors will provide a mobile viewpoint in the environment and allow for focusing on the important events and activities as needed.

Embedded sensors: We are working in computer engineering researchers to develop small embeddable cameras that will be mounted in the ceiling. A large number of such cameras will be distributed in a scene allowing for an elaborate coverage of the space. We are also working in instrumenting the spaces with phased-array microphones that will be embedded into the walls and ceilings. Such microphones will allow for accurate location of the speaker and will provide a higher quality audio stream for speech recognition.

Other sensors: In addition to the video and audio sensors, we are also studying others types of sensors to augment the user and the environment. These include simple contact sensors to detect which furniture is in use to a touch sensitive carpet to track walking people. Recently, some new biosensors that have been developed to measure biomedical data. The users could wear these and the data transmitted to the environment's sensor-net for higher-level content interpretation.

Computational Perception for Recognition of Activities: Using the ubiquitous sensors, the space will be able to sense the activities of its occupants and "learn" their routines. Using computer vision and audio processing techniques, the environment will locate and identify its occupants, and recognize their activities. These audio and video sensors will also help with establishing a natural interface with the space as the space will be able to interpret speech and gestures. These same sensors will allow for face-to-face interactions with other users at other locations.

Following are a few aspects of perceptual processing of video and audio streams that can lead to awareness

Identity: Using speaker identification and face recognition techniques, the users can be identified. User identity is a strong cue for context awareness. Both speaker ID and face recognition technologies are commercially available at present.

Location: Using visual and tactile tracking, users can be located. Where the users are will serve as a strong cue of what they are doing. Phased-array microphones can also be used to locate a speaking person.

Activity: By processing a video stream, we can track the users as they move. These movements could be over short terms (in minutes and seconds) or over long terms (days). Each type of movement provides important cues about what is happening in a scene. Recognition of short-term activities could be used for determining what a user wants at a specific moment, while long-term activities could be learned as routines and then deviations could be measured.

Pose tracking: Recognizing where a user is looking is a significant cue as to what a user is doing. This can also be achieved in an aware space by tracking the body and face as it moves.

Expression/Gesture Recognition: Visual tracking and recognition of facial expression and hand gesticulation is essential in determining the state of the users. Facial expressions help with interpreting the users' affective state, while gesture recognition will aid in analyzing

the communicative intent of the users. We are working on various methods for this form of spatio-temporal analysis of the video streams.

Audio processing: Speech recognition is at present a commercially available technology. Though not perfect, these commercially available systems can be used for recognition of spoken words. Using the phased array microphones, we can extract reliable speech signal for such speech recognition systems. In addition to speech recognition, we are also studying non-verbal analysis of the speech signal to extract stylistic information that is encoded in speech.

All of the above technologies are based on our research in the area of signal processing and interpretation. We are prototyping these technologies in our aware and smart spaces.

Applications:

We are interested in developing several smart and aware environments. These environments include a classroom, a living room, a meeting room and an office. These spaces will serve as “living laboratories” for development and experimentation. We will incorporate into these environments various applications. These include anytime, anywhere videophones (so that the occupant does not have to search for a communications device), intelligent surveillance and monitoring (to monitor children, elderly, and those in need of health care, and prevent unauthorized access), speech and gestural interfaces (to support natural interfaces), education (connection to school, distance learning), entertainment (support collaborative game-playing), *etc.*

Our most significant step in this direction is the building of an “Aware Home” on the Georgia Tech campus. This Aware Home is our attempt at building a large scale intelligent environment that is aware of the activities of its inhabitants. Initially, this home will serve as a living laboratory while inhabited by our students, however, on the long run we expect this space to become an environment to support the elderly and the sick. An aware home, with embedded abilities to perceive its inhabitant’s activities can provide a richer quality of life to the elderly (while being cheaper than an assisted living facility). Awareness of daily routines and deviations from some typical behaviors can trigger warnings. Crisis intervention would also be possible as the aware home could detect dangerous circumstances. The aware home can also serve to provide a familial connection between distant family members and provide a ubiquitous support structure.

Summary:

We are building various types of intelligent environments that are aware of their inhabitants. Each aware environment has varying needs and complexity and therefore requires a unified sensing and recognition infrastructure. We are building a large-scale “Aware Home” as a prototypical intelligent space. This space will provide us with a test-bed for implementing our technologies of ubiquitous sensing to make the environment aware of different types of activities. We are pursuing various applications in this aware home. Our most significant effort in this direction is to make this Aware Home an assisted living facility for the elderly.