Available online at www.sciencedirect.com
ScienceDirect
ELSEVIER
Pervasive and Mobile Computing 3 (2007) 74–94
pervasive and mobile computing
www.elsevier.com/locate/pmc

Multi-modal emotive computing in a smart house environment

Simon Moncrieff*, Svetha Venkatesh, Geoff West, Stewart Greenhill

Department of Computing, Curtin University of Technology, Kent Street, Bentley 6102, Western Australia, Australia

Received 30 September 2005; received in revised form 29 June 2006; accepted 16 July 2006
Available online 11 September 2006

Abstract

We determine hazards within a smart house environment using an emotive computing framework. Representing a hazardous situation as an abnormal activity, we model normality using the concept of anxiety, using an agent based probabilistic approach. Interactions between a user and the environment are determined using multi-modal sensor data. The anxiety framework is a scalable, real-time approach that is able to incorporate data from a number of sources, or agents, and able to accommodate interleaving event sequences. In addition to using simple sensors, we introduce a method for using audio as a pervasive sensor indicating the presence of an activity. The audio data enabled the detection of activity when interactions between a user and a monitored device didn't occur, successfully preventing false hazardous situations from being detected. We present results for a number of activity sequences, both normal and abnormal.

Keywords: Assisted living; Audio; Surveillance and monitoring; Sensor network; Hazard detection

1. Introduction

This paper is concerned with the development of smart environments for the assisted living of elderly people. A particular aspect of smart environments relevant to the care of

* Corresponding author. Tel.: +61 8 9266 7680; fax: +61 8 9266 2819.
E-mail addresses: simonm@cs.curtin.edu.au (S. Moncrieff), svetha@cs.curtin.edu.au (S. Venkatesh), geoff@cs.curtin.edu.au (G. West), stewartg@cs.curtin.edu.au (S. Greenhill).

1574-1192/$ - see front matter
doi:10.1016/j.pmcj.2006.07.003

the elderly is the determination of hazards within the environment in real time. Typical hazards are appliances being left unattended, such as unattended electrical devices, or abnormal environmental states such as a front door being left open and unattended. It is important to note that estimation of a hazard must be made in the context of the user's normal behaviour, and it is this requirement that makes the problem difficult.

Most systems for activity monitoring are interested in event detection after the fact, for example every couple of hours [1]. Although this is valid in many cases, real time response in needed for many hazards at home e.g. leaving the bath to run over. Other approaches to such problems is to use temporal models of activity recognition with hidden Markov models [2,3] and build representations for normality from which abnormality can be inferred. However, the variability in behaviour patterns both within and across individuals makes it difficult to detect consistent patterns. Additionally the computational complexity makes it difficult to use in real time. Furthermore, when multiple as well as multi-modal sensors are used, a typical probabilistic based method of activity recognition starts to fail because the state space grows exponentially with the number of devices and states.

Potential hazards can be determined in many ways across a broad spectrum of modalities. At one end of the spectrum, we can have a fully monitored home with multiple cameras augmented by other sensors in the environment. At the other end of the spectrum, the home can be augmented with simple sensors such as pressure mats, simple microphones (for sound level sensing) and reed switches. All these sensors have modalities that have advantages and disadvantages. For example pressure pads give simple discrete on/off events usually with 100% reliability. Video processing is highly flexible and consequently can be used to detect many different aspects of activity. However the current state of the art in video processing is not reliable enough and the processing required is complex. Although microphones can be regarded as simple sensors when used as a sound level device e.g. for detecting loud noises, it can also be used to detect more complex audio events in various ways. They can also be used to localise sound events [4] but this requires complex processing. Ultimately, the challenge is the integration of most if not all of these different devices seamlessly into a framework for activity and hazard monitoring. When using audio and video sensors, one must be aware of the privacy issues raised and hence it is important to use these devices to *only* detect events and not record the video and sound for later playback.

In this paper we explore the integration of sound, and simple sensors such as pressure pads, reed switches and X10 devices into a multi-modal framework for hazard detection in assisted living. This is based on a recent model proposed in [5] where each device, with the potential to be hazardous, is represented as an agent. A measure of *anxiety* is associated with each agent, representing the potential hazard represented by the device. The proposed agent based approach is device centric and the state space is automatically factored, thus making the approach scalable and applicable in small as well as large pervasive environments. Whilst the previous work concentrated on using simple sensors such as pressure pads and reed switches, we explore a more comprehensive integration of the multi-modal sensor data, in the form of simple sensor and audio data, into the anxiety framework. We introduce a method for using the audio as a pervasive sensor using foreground sound events to determine audio activity. We argue that the use of sound