Journal of Universal Computer Science, vol. 16, no. 12 (2010), 1650-1665
submitted: 29/12/09, accepted: 5/5/10, appeared: 28/6/10 © J.UCS

Mobile Intelligence

Yang Cai
(Carnegie Mellon University, Pittsburgh, USA
ycai@cmu.edu)

Abstract: Analyzing human motion in a building has been an active subject in Ambient Intelligence and Universal Design for many years. In this study, we present a rapid prototype of a mobile and interactive sensing platform for smart buildings. The biologically inspired robot can follow the moving person around, memorize the motion patterns in form of sequences of symbols, and detect surprising events, based on similarity between the priori and posteriori probability distributions. The key modules in this study have been prototyped and tested with real-world data, such as the two–month sensory data in a building. Furthermore, the author believes that simple and recursive algorithms would enable mobile robots to simulate natural ethological intelligence.

Keywords: sensor network, motion, smart environment, instinctive computing, anomaly detection
Categories: L.7, J.0, H.1

1 Introduction

For over a decade, Ambient Intelligence (AmI) developers have aimed to build smart environments with sensor networks to pervasively monitor human activities in a building for elderly and disabled people. Recent studies include designing a home for elderly people or people with disabilities. Healthcare systems are also looking for an easy and cost effective way to collect and transmit data from a patient's home. For example, a study shows that the wireless network used by most major cell phone companies was the best for sending data to hospitals from a patient's home.

Emerging active sensing technologies bring broad opportunities to AmI for healthcare. For example, 1) Implantable Sensors, including insulin pumps for monitoring and injections, implantable RFID, implantable neural chips, and so on. 2) Pill Sensors, including body imaging and body temperature measurement, 3) Wearable Sensors, including Bodymedia™armband sensor and transmitter [Farringdon, et al, 2005] that tracks body temperature, galvanic skin response, heat flux, and other data, and 4) Mobile Sensors, including mobile robotic development platforms such as the Roomba and SONY's robotic dog. The implantable and pill sensors are normally developed within medical communities. Those devices require special approvals by authorities such as FDA (Federal Drug Administration) of USA. Wearable and mobile sensors, on the other hand, are less regulated, enabling innovations from broader communities, from sports to domestic robotics. Figure 1 illustrates a spectrum of active sensors.

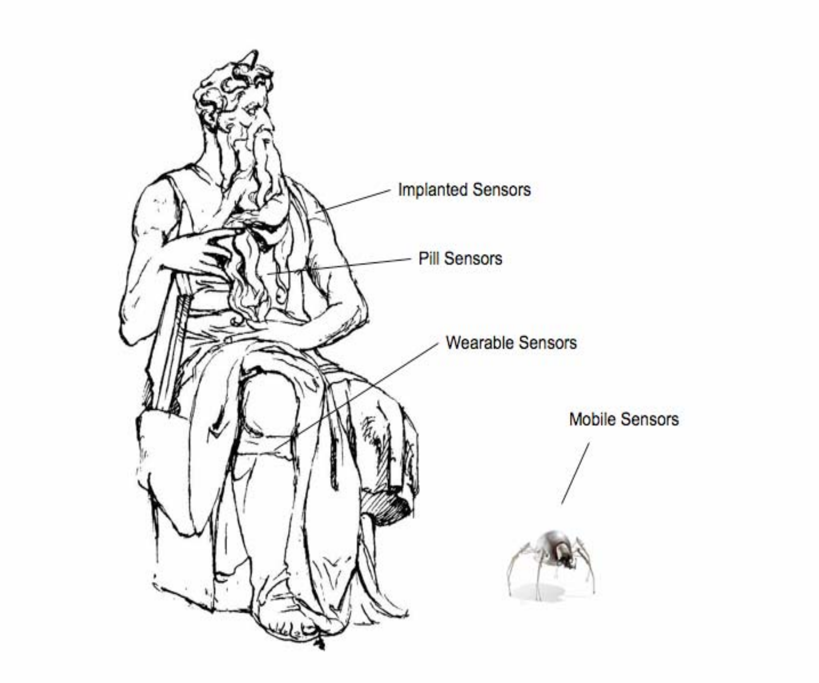

Figure 1: Active-sensing devices for healthcare

1.1 Active-Sensing vs. Passive-Sensing

In the past ten years, there has been a wave of building 'smart environments' across universities and corporations that launched labs to explore the healthy living environment, such as LiveNet, HomeNet, and Philips' HomeLab. Those experimental systems contain networked passive or stationary sensors from a docent to several hundreds. Unfortunately, the passive-sensing networks have several technical challenges:

- *Scalability*. Although an individual sensor is inexpensive, as the number of sensors increase, the costs and data traffic grow exponentially [Cai, 2006].
- *Fusion.* How to abstract human motions and recognize them in different environments is still a hard problem. As the number of passive sensors increases, data fusion becomes a bottleneck to many passive-sensing systems.
- *Robustness*. Conventional sensor networks often have dead zones, which are hard to calibration, and are vulnerable to environmental changes.

On the other hand, active-sensing systems provide complimentary solutions to the existing problems. For example, for scalability, active-sensing systems are independent from the size and number of rooms. Because the active–sensing systems integrate sensors into one platform, the sensory fusion and dada mining and transferring are relatively easier than passive and stationary sensing. Besides, as active-sensing systems are integrated into one piece, they make calibration and maintenance easier, excepting the implanted ones. Furthermore, mobile sensing platforms are able to eliminate blind spots simply by moving around the subject.

However, active-sensing systems have their own weaknesses: energy, transmission and speed. Most active-sensing systems are powered by battery. The duration of the sensing is limited by the life of the battery. For example, a pill camera's battery life is only 8 hours. It takes 24 hours for a pill sensor to pass through a human body. Therefore, the sensor is in sleeping state until it moves to the target area. To save the battery life, data transmission is also a challenge. Short range and low powered wireless transmission is desirable, such as Bluetooth. To further cut down the power consumption, a lower data bit rate transmission is needed, such as two frames per second for pill cameras. In addition, a mobile active-sensing platform sometimes has perceptual delay caused by motion. It is a trade-off between the spatial resolution and temporal resolution. A mobile sensor may improve spatial resolution by moving around but sacrifice the observational speed due to its sequential process.

1.2 Ethologically Inspired Mobile Sensing

Ethology is the study of animal behavior. It tries to explain the causation, development, survival, and evolution of behavior patterns within animals. Perhaps the most famous ethologists are Niko Tinbergen, Konrad Lorenz and Karl von Frisch. Their hierarchical behavior models have often been quoted by Artificial Intelligence communities [Tinbergen, 1951] and [McFarland, 1985]. Their theories have inspired an early movement of behavioral mobile robotics [Brooks, 1999] that connects perception directly to actuation.

Delphine de Gerardin said, "Instinct is the nose of the mind." Dogs are perhaps the most instinctual helper to humans. Dogs are vigilant, empathic and royal. For thousands of years, dogs have been domestic assistants for the elderly and disabled, especially visually impaired people. In addition to dogs' outstanding capacity for olfaction, their motion intelligence is also remarkable. Dogs can sense complex human activities, ranging from gestures, facial expressions, down to micro movements such as shivering.

Dogs use mobile and interactive sensing in contrast to the passive sensing in Ambient Intelligence (AmI) systems. For example, machine learning based methods are not connected with the rule-based systems. A rule-based system is similar to the Internal Release Mechanism (IRM) in animals and insects that an action is devoted to a particular stimulus. Those rules can be encoded into the DNA of a dog, so-called instincts.

A robotic dog model appears to be desirable here because of its canine senses, mobility, scalability and multimodal interactions with humans. Instead of building a large sensor network for a smart environment, we intend to explore the potential of a mobile and interactive sensing platform. In this paper, we aim to use Dog 1.0 as a metaphor for interactive and mobile sensing for smart buildings. Figure 2 shows a diagram of the ethological perception of motion.

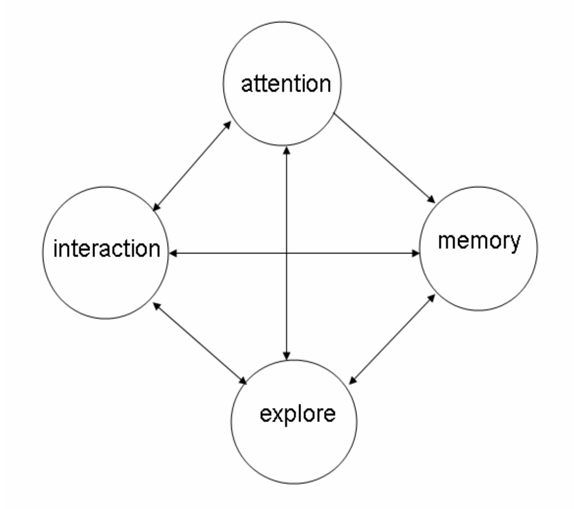

Figure 2: A framework for ethological perception of motion

2 Attention

A dog often gives undivided attention to its owner. A dog has the instincts to recognize and follow its host. Ethologists called this phenomenon imprint [Tinbergen, 1951]. Dog 1.0 starts with a simplified version of human following algorithm. Based on just one infrared sensor and two servo motors, the robot can search for the target and follow the human within a predefined distance. Figure 3 shows the robot in action. The pseudo code of the control algorithm is following.

Figure 3: Human following with infrared sensing and servo control

This crude prototype is just a working model of ethological imprint. It provides a mobile sensing platform for smart environments. In addition to heat sensing, the robot can mount color and camera sensors for more reliable tracking of targets. By recording the trajectory of the mobile robot, we can analyze human motion patterns in spatial and temporal dimensions. For the indoor environments, encoded infrared beacons can be used as positioning and homing landmarks, similar to iRobot's homing system [iRobot, 2010]. In addition, the visual features on the floor and ceiling may also be used for navigation. However, visually guided navigation is more expensive compared to infrared.

Some people won't like a dog following them around the house. In this case, a cockroach behavior can be simulated to hide the robot in shadows so it stays away from humans.

3 Memory

Dogs have a good memory for navigation, identification and general pattern cognition. How a dog represents motion patterns is still unknown. But we are certain that it must involve neurons, space and time. Cellular Automata (CA) has been a computational model for studying the spatiotemporal dynamics of biological and social systems, ranging from cells to a subway system. In a CA model, each cell is coded with identical rules that are triggered by stimuli from neighborhood. It is similar to the sensory cells on animal skins. However, a CA approach appears to be less efficient in representing a large area with a fine resolution of human activities. In this study, we hypothesize that human motion in a building can be represented as a sequence of actions in form of a state machine.

Figure 4 shows a simplified home model, where the space is divided into four areas: kitchen, bedroom, bathroom and living room. The correspondent activities are Eat (E), Sleep (S), Toilet (T) and Play (P). A sequence daily activities can be represented with a string sequence, similar to the DNA strand: SETPETPS…

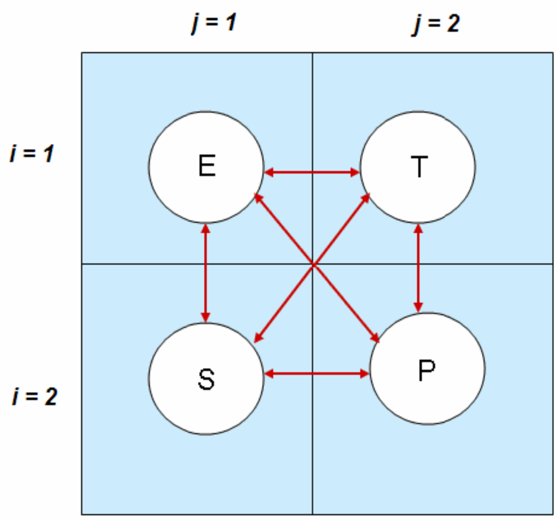

Figure 4: Home model of four areas with four activities

Linear encoding: divide the space into m by n cells with rows i = 1, … , m and columns j = 1, … , n, can linearly map the two-dimensional motion into a one-dimensional string L(k):

$$L(k) = (i - 1) * n + j \qquad (1)$$

Assume the robot can store all the position sequences. Then how does a computer retrieve the sequential patterns from a database? This problem is similar to a DNA sequence search. But DNA only contains four letters. Here we want to generalize the number of unique positions or activities to an arbitrary amount. To realize such a pattern matching algorithm, we apply the fast text search with errors. Given a pattern string ***P***, find an approximate match from a string ***T*** within a number of errors ***k***, measured by Levenstein Distance, including insertion, deletion and replacement of letters [Mander and Wu, 1992]. For examples, the following cases are matching cases:

```
dog → dag (replacement)
dog → dg (deleetion)
dog → dogs (insertion)
```

The algorithm in this study extends the original text length to infinite length and number of symbols. To enable a fast string match, the algorithm builds a binary map for the matching operation. For example, an approximate match with an insertion error:

	a	**a**	**b**	**a**	**a**	**c**	**a**	
a	1	1	1	1	1	1	1	
a	0	1	1	1	1	1	1	
b	0	0	1	1	0	0	0	
a	0	0	0	1	1	0	0	
c	0	0	0	0	0	1	0	R^1 (with one insertion)

Given this string matching algorithm, we can find unique strings from the database and give them an importance score. While we calculate the importance score for each unique string, we can, at the same time, find its *K-Nearest Neighbors*. We can then test for anomalies using the distance based KNN model.

4 Explore

Mobile sensors can be used to explore the human activities in a building, including anomalous or surprising event detection. In this study, we look into a case of a large building with over 10 stops for the mobile sensor Dog 1.0. At each stop, the user identification and time is recorded. Given a real-world database of movement sequences over a two-month period, we want to gather information about the data and make inferences from it. Figure 5 shows a layout of the access points.

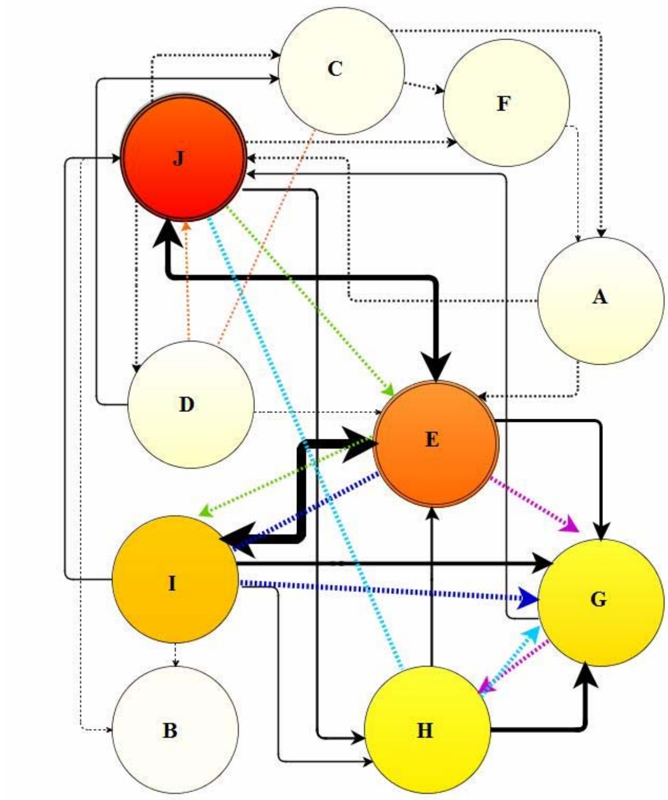

Figure 5: Floor plan layout of 10 mobile sensing stop locations

4.1 Sequence Anomaly Detection

First we want to look at which sequences of movements are popular. Afterwards, we want to make a model to find anomalies in the data as well as to test whether a given input sequence would be considered normal or abnormal based on the history of movements. In this case, we use the approximation string-searching algorithm to sort the motion sequences at different pattern length, such as 1, 2, 3, 4 and 5.

In Figure 6, the network models the activities of the rooms. The more frequently visited rooms are labeled with darker colors. The size of the black arrow paths represent how often an employee goes between the two rooms. The colored dashed paths represent commonly traveled paths of length > 2.

The data matches what was said in Figure 6; sensor F is the most commonly triggered sensor. It is also triggered repeatedly often (J = 7227, JJ = 2884, JJJ = 782). Strings of length 2 containing E are also really frequent (EI = 973, JE = 304, EG = 232, EJ = 199, HE = 118), suggesting that E branches off to a lot of other locations. Also, most occurrences of J also involve E, suggesting that employees at position J usually transition to E, suggesting that J is a major chokehold in the building, which might be problematic if employees could not access J, since it's likely they are cut off from the rest of the building, since E is the center of the building. Certain anomalies were also detected. Some sequences described routes that were roundabout and ~~was~~ not the closest way to get from point A to B. For instance, DCI was detected as an anomaly. Looking at the floor plan, we found that the path from D directly to I is much shorter than D, C and I. Furthermore, sometimes errors might happen and a scanner fails to record an employee walking back. For example, EB was detected as an anomaly and looking at the floor plan, it seems unlikely to get from E to B without triggering any other sensors. Thus, it's possible that this sequence was produced because the sensor didn't work that time (events like this would occur rarely). Other suspicious sequences including walking in a circle (IHEI, GHJEI, CFAC) and walking back and forth for long periods of time (HGHGGHGGHGHGH).

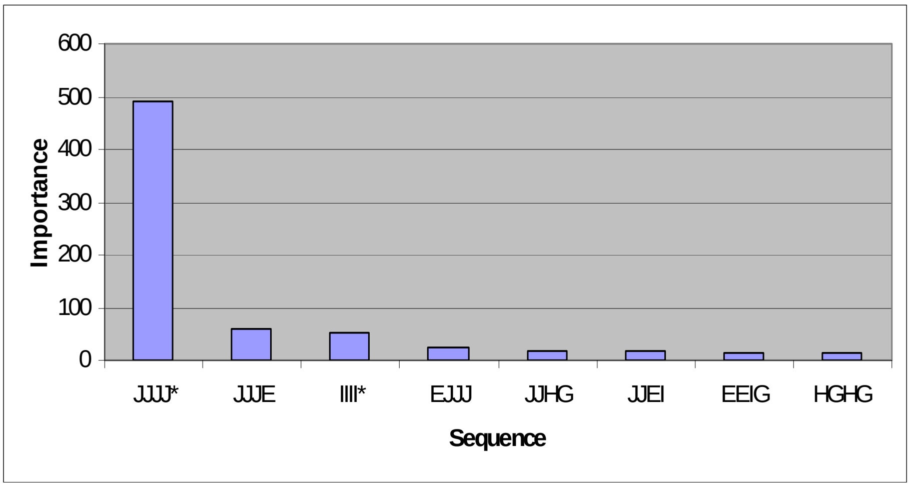

Figure 6: Frequency ranks of sequences of length = 4 or more. Note: JJJJ includes all sequence of length 4 or greater consisting only J's.*

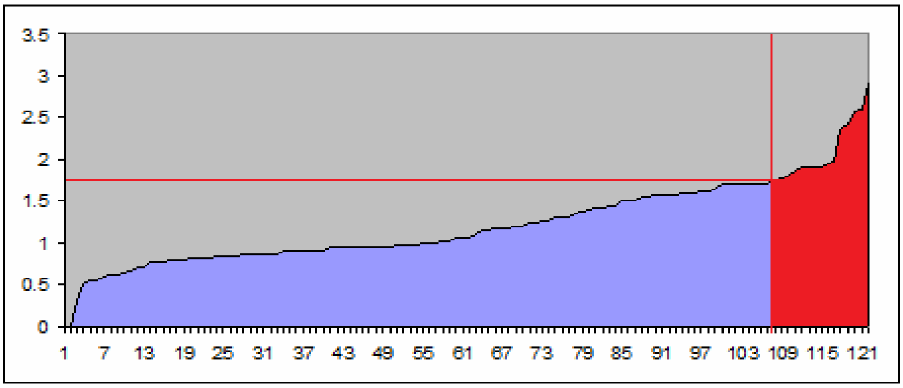

Figure 7: Outlier score distribution for sequences of length 4
(error = 1.75, anomaly rate = 11.6%)

We found that many anomalous events are conditional or local. For example, our algorithm picked up a unique anomalous event sequence: JJJJJJJ. However, when we consider the time of day when the events happened, the sequences become JJ, JJJ, JJ,.., which are normal. On the other hand, a few cases seem normal. But in fact they are anomalous events, if we consider the 'time of day' variable. For example, frequent accessing living after middle night is often viewed as an anomalous event.

In fact, all anomalies are determined by conditional probability distribution functions. Song et al proposed a Gaussian Mixture Model (GMM) to represent these conditional anomalous events [Song et al 2007]. Unfortunately, learning the conditional probability distribution functions appears to be non-trivial. So far, only small sizes of the data sets (up to 205 x 100) have been experimented. Can a system inherit knowledge from other systems? The answer is yes! For example, the famous OpenCV's face detection uses the trained feature database from previous systems that has been inherited for many generations. However, the data structures of such a knowledge have been deeply hidden inside the length code layers, which is hard to understand, even though they are open source.

To classify and standardize those conditional probability distribution functions is a way toward shared and reusable common sense across borders of intelligent systems. It is certainly a bridge between a mainly hand coded rule-based system and a purely machine learning system. In our Dog 1.0 system, we start to use well-known common sense to guide the data mining and represent those rules in forms of conditional probability functions, for example, the usage of kitchen and living room at a 24-hour cycle.

4.2 Similarity versus Surprise

The concepts of similarity and surprise are central to sensory processing, adaptation, learning, and attention. There are many mathematical theories to quantitatively characterize similarity, but very few about surprise elicited by a stimulus or event.

What is a similarity between two sets of data in terms of an information model? Given a message at one end, can we reproduce approximately the same message from another end? Formally, the similarity or mutual information of two discrete random variables *X* and *Y* can be defined as:

$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log \left(\frac{p(x,y)}{p_1(x)\, p_2(y)} \right) \quad (2)$$

where $p(x,y)$ is the joint probability distribution function of *X* and *Y*, and $p_1(x)$ and $p_2(y)$ are the marginal probability distribution functions of *X* and *Y* respectively. As we can see, the more similarity between two sets of data, the less surprising we are. Unsurprising data yields little difference between posterior and prior distributions of beliefs over models, while surprising data yields a large shift: in mathematical terms, an event is surprising when the distance between posterior and prior distributions of beliefs over all models is large.

Itti and Baldi provide a solution of the surprise under minimal axiomatic assumptions, by measuring the difference between posterior and prior beliefs of the observer [Itti and Baldi, 2005]. Instead of measuring information carried by data, it measures the difference between prior and posterior probability distribution functions. Using the relative entropy or Kullback-Liebler (KL) divergence [Kullback, 1959], we are able to measure the level of surprise in the historical database in a building.

Figures 8 to 12 show results of our experiments of the surprise values for single space access and cross space motion events over 24 hours, based on the collected data of two months in a sensory building. Figure 8 shows a strong surprise happens when a person shows up at space J between 0 am and 5 am because normally there is no body there. Figure 9 shows less surprise value at space I because there are some people in that area, more or less. Figure 10 shows a huge spike of surprise because the historical data was zero or missing at 9 am to 10 am window. Similar trends also apply to Figures 11 and 12.

Why does the model respond to missing data so violently? From the KL model, we are certain that missing data also contains information! Ethologists discovered that even animals can perceive the information behind the missing data. For example, pigeons can predict the missing patterns in a sequence of imagery stimuli [Hearst, 1984]. However, missing data may not be reliable negative evidence unless the quality of the sensory channel is checked. Therefore, we need to add a conditional probability to the missing data case. In reality, land mine experts frequently tune the sensitivity of the metal detector to ensure the silence of the sensor means no metal will be detected. We will discuss how to implement this function in Dog 1.0 in the following section.

We also need to add a 'forgetting' function to the surprise detection model. Through Bayesian updates, the model would yield a large shift between the prior and posterior distributions of beliefs, with the posterior 'adaptively' favoring a 'surprising' event model, as it happens frequently in a recent short period.

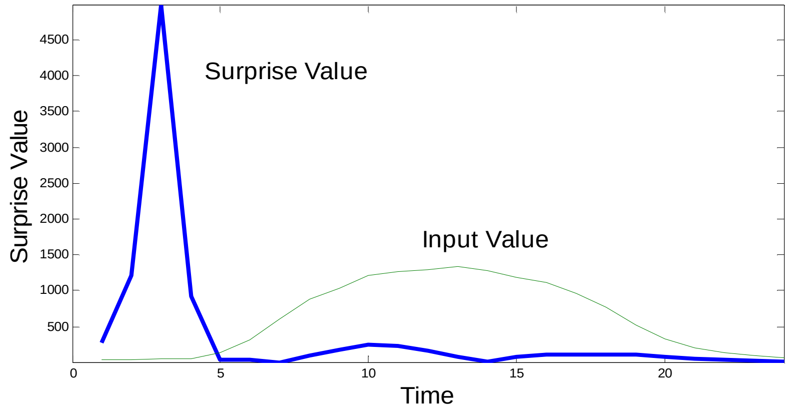

Figure 8: Surprise value at space J in a building with 2 months historical data

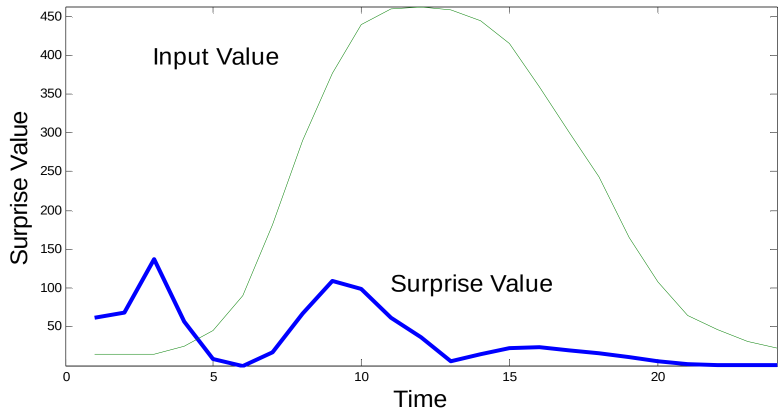

Figure 9: Surprise value at space I in a building with 2 months historical data

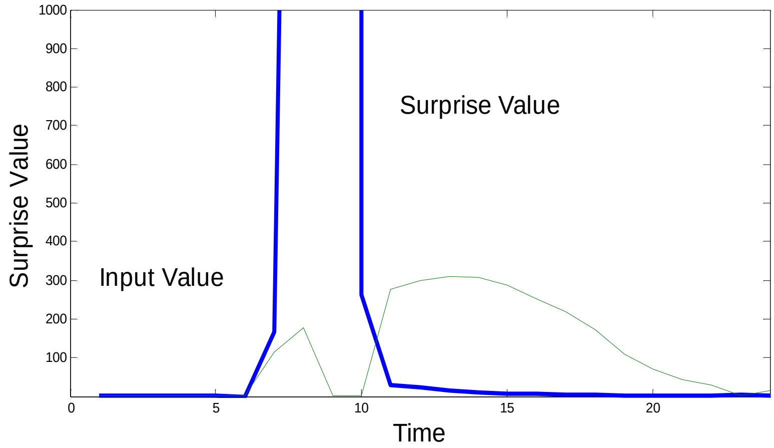

Figure 10: Surprise value at space E in a building with 2 months historical data

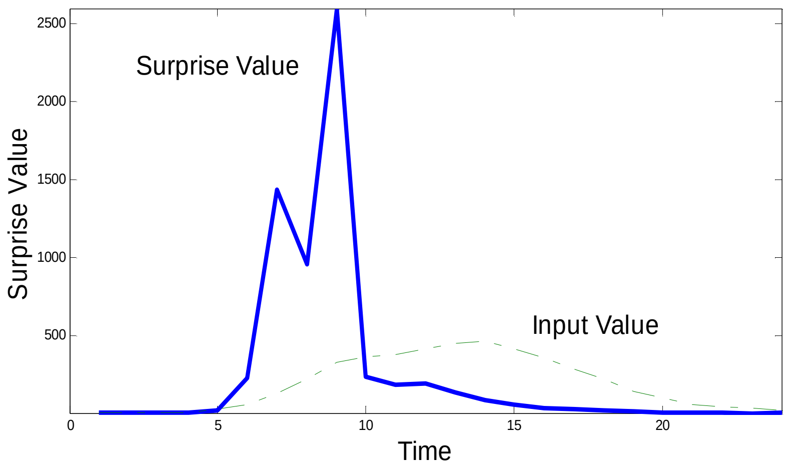

Figure 11: Surprise value for the motions from J to E in a building

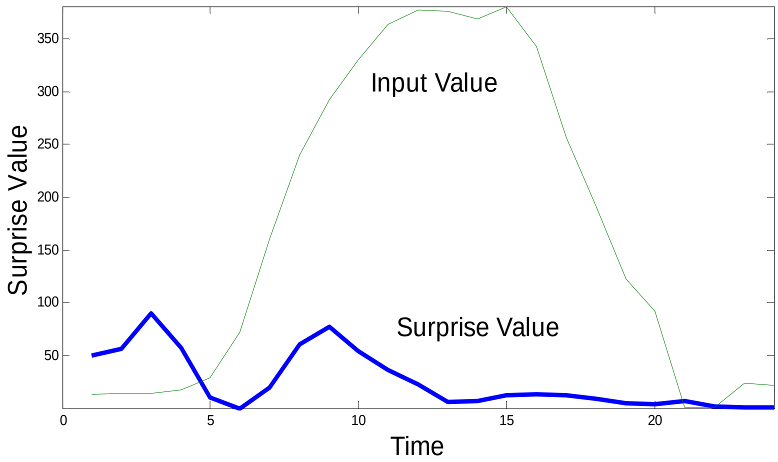

Figure 12: Surprise value for the motions from E to J in a building

5 Interaction

One of great advantages of a dog is interaction with humans. From a sensing point of view, it creates an active and interactive sensing system. In many smart environments, there are always blind spots and ambiguous judgments. For example, detecting when people fall down on the floor. To many patients who suffer a stroke or a heart attack, the 'golden minutes' for the detection and rescue means live or death. Wearable sensors can pick up the sudden falling signals. However, it is easy to be confused with other activities such as rising from a chair [Cai, 2006][Rowe, 2006]. Infrared imaging can clearly reveal the warm body on the floor. However, the camera is not always pointed to the target, and it's expensive as well. In this study, we design interaction functions to sense human activities and diagnose problems. Instead of stationary sensing, the robot moves close to the person and examines the situation.

Active fall detection is the first conceptual model for Dog 1.0. With three infrared sensors mounted vertically, the robot is able to measure the perimeter of a human body on the floor. The movements of the robot form a set of vectors that can be represented as a Chain Code [Fu, 1978]. From the Chain Code, we can filter out the noises and recognize the fallen body shape. This is a trade-off between sensor resolution and processing speed. Experiments show that the currently version takes 20 seconds to trace a human body on the floor. With an improved pattern recognition algorithm, we are able to recognize partial contours of a human body and cut down the time from 1/3 to ½.

Another function we are experimenting with is multimodal interactions, such as visual, auditory, and tactile stimuli. For many elderly people, visual is perhaps the preferred signal. A flash of light may attract the person to turn around and gaze back. However, a visual signal only works when the person pays attention to the robot. When the visual or auditory stimuli fail to contact the person, the robot would drive close to the person and physically contact to see the response. This action happens a lot in many real world cases when a dog wants to know the state of a motionless body.

6 Conclusions

In this study, we present a plausible model of a mobile and interactive sensing platform for smart buildings. The biologically inspired robot can follow the moving person around, memorize the motion patterns and detect similar or surprising events. Finally the interactive sensing algorithm is discussed to detect a fallen person on the floor.

A dog cannot be built overnight. Although modules in this study have been prototyped and tested with real-world data, such as the 2–month sensor network data from a building, the overall integrated systems performance has yet to be evaluated at a grant scale. Furthermore, most of the computational models are still at the PC level, which could be overkill to a mobile robot. By making the models recursive and minimalist, we believe that we are making our robot closer to natural ethological intelligence. *A dog is instinctual because it inherits and learns conditional probabilities about the stimuli.* The advantage of the ethological perception lies far beyond a dog. For example, we can integrate infrared imaging sensors and olfaction sensors to be a symbiotic system that can send text messages back to people.

Acknowledgement

The author would like to thank Bosch Research Center in Pittsburgh, Army Research Office (ARO) and Cylab for their sponsorship. The author is also in debt to Cheong Kin Ng, Peter Liang and Emily Durbin for their assistance.

References

[Bayes, 1763] Bayes, Thomas (1763). "An Essay towards solving a Problem in the Doctrine of Chances.". Philosophical Transactions of the Royal Society of London **53**: 370–418. doi:10.1098/rstl.1763.0053

[Berg, 2009] Berg, D.J., S. E. Boehnke, R. A. Marino, D. P. Munoz, L. Itti, Free viewing of dynamic stimuli by humans and monkeys, *Journal of Vision*, Vol. **9**, No. 5:19, pp. 1-15, May 2009.

[Berg, 2006] Berg, D.J., S. E. Boehnke, R. A. Marino, P. F. Baldi, D. P. Munoz, L. Itti, Characterizing Surprise in Humans and Monkeys, **In:** *Proc. Vision Science Society Annual Meeting (VSS06),* May 2006.

[Brooks, 1999] Brooks, R. Cambrian Intelligence: The Early History of the New AI, The MIT Press, 1999

[Cai, 2006] Cai, Y. Empathic Computing, in Cai, Y. and Abascal, J. (eds.), Ambient Intelligence for Everyday Life, Lecture Notes in Artificial Intelligence (LNAI 3864), Springer, 2006

[Chandola, 2009] Chandola, V., Banerjee, A., and Kumar, V. Anomaly detection: a survey, ACM Computing Surveys, September 2009

[Dewsbury, 2001] Dewsbury, Guy, Taylor, Bruce, Edge, Martin: Designing Safe Smart Home Systems for Vulnerable People, In Proceedings of the Dependability and Healthcare Informatics Workshop, Edinburgh, March 2001, 2001.

[Delin, 2005] Delin, K. A., Shannon P. Jackson, David W. Johnson, Scott C. Burleigh, Richard R. Woodrow, J. Michael McAuley, James M. Dohm, Felipe Ip, Ty P.A. Ferré, Dale F. Rucker, and Victor R. Baker, Environmental Studies with the Sensor Web: Principles and Practice, Sensors, Vol. 5, 2005, pp. 103-117

[Farringdon, 2005] Farringdon, Jonathan, and Sarah Nashold. "Continuous Body Monitoring." Ambient Intelligence for Scientific Discovery (2005): 202-223.

[Figueiredo, 1999] Figueiredo, M, et al, Expanding NASA's Data Processing to Spacecraft, Computer, June, 1999

[Galstyn, 2004] Galstyan, A. Distributed online Localization in sensor networks using a moving target. University of Southern California. 2004

[Geddes, 2009] Geddes, L. Model of surprise has 'wow' factor built in, *New Scientist*, Jan 2009

[Gigerenzer, 2007] Gigerenzer, G. Gut Feelings, Penguin Books, 2007

[Gist, 2004] Gist, Y. J., and Lisa I. Hetzel. Department of Commerce. We are the People: Aging in the United States, 2004.

[Gunarathne, 2002] Gunarathne, G.P. Presmasiri, Gunarathne, Tharaka R.: "Arterial Blood-Volume Pulse Analyser." IEEE, Instrumentation and Measurement Technology Conference, AK, USA, May. (2002) 1249-1254

[Hearst, 1984] Hearst, E. Absence as information: some implications for learning, performance, and representational process, in Animal Cognition, edited by H.L. Roitblat, T.G. Bever, and H.S. Terrace, LEA, 1984

[iRobot, 2010] iRobot: http://www.ibobot.com

[Itti, 2005] Itti, L., P. F. Baldi, A Principled Approach to Detecting Surprising Events in Video, **In:** *Proc. IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 631-637, Jun 2005.

[Itti, 2009] Itti, L., P. F. Baldi, Bayesian Surprise Attracts Human Attention, *Vision Research*, Vol. **49**, No. 10, pp. 1295-1306, May 2009.

[Kullback, 1959] Kullback, S. Information theory and statistics, Wiley, New York, 1959

[McFarland, 1985] McFarland, D. Animal behavior, Benjamin/Cummings, Menlo Park, California, 1985

[Manber, 1992] Manber, Udi and S. Wu. Fast text search allowing errors. Communications of the ACM, 35(10): pp. 83–91, October 1992.

[Mundhenk, 2009] Mundhenk, T. N., W. Einhaeuser, L. Itti, Automatic Computation of an Image's Statistical Surprise Predicts Performance of Human Observers on a Natural Image Detection Task, *Vision Research*, Vol. **49**, No. 13, pp. 1620-1637, Jun 2009.

[Navarro, 2001] Navarro, G. A guided tour to approximate string match, ACM Computing Surveys, Vol. 33, Issue 1, March 2001

[Petrovskiy, 2003] Petrovskiy, M.I. Outlier Detection Algorithms in Data Mining Systems. Programmirovanie, Vol. 29, No. 4, 2003

[Song, et al, 2007] Song, X., etc. Conditional Anomaly Detection, IEEE Trans. on Knowledge and Data Engineering, Vol. 19, No.5, May 2007

[Tinbergen, 1951] Tinbergen, N. The study of instinct, Oxford University Press, Oxford, UK, 1951