Sensor Networks for Ambient Intelligence

Eric Pauwels, Albert Ali Salah, Romain Tavenard

Sensor Networks for Ambient Intelligence

Eric J. Pauwels
CWI
Kruislaan 413
1098 SJ Amsterdam, NL
eric.pauwels@cwi.nl

Albert A. Salah
CWI
Kruislaan 413
1098 SJ Amsterdam, NL
a.a.salah@cwi.nl

Romain Tavenard
IRISA/ENS de Cachan
Av. du Général Leclerc
35042 Rennes Cedex
romain.tavenard@irisa.fr

***Abstract*— Due to rapid advances in networking and sensing technology we are witnessing a growing interest in sensor networks, in which a variety of sensors are connected to each other and to computational devices capable of multimodal signal processing and data analysis. Such networks are seen to play an increasingly important role as key enablers in emerging pervasive computing technologies. In the first part of this paper we give an overview of recent developments in the area of multimodal sensor networks, paying special attention to ambient intelligence applications. In the second part, we discuss how the time series generated by data streams emanating from the sensors can be mined for temporal patterns, indicating cross-sensor signal correlations.**

I. Introduction: Ambient Intelligence

Ambient Intelligence (AmI) is a term coined by Philips management to conjure up a vision of an imminent future in which persons are surrounded by a multitude of fine-grained distributed networks comprising sensors, computational devices and electronics that are unobtrusively embedded in everyday objects such as furniture, clothes, and vehicles, and that together create electronic habitats that are sensitive, adaptive and responsive to the presence of people [1], [2]. It is envisaged that these environments will be able to identify the people that dwell in them, recognize their actions and even their emotions and intentions, and assist them according their individual preferences and needs.

This vision was subsequently endorsed by the European Commission and supported by a sizable chunk of the FP6 ICT program budget. The long-term ambition is to create an AmISpace in which the seamless interoperation between different environments (home, work, vehicle, public spaces, etc.) facilitates participation by the individual in a multiplicity of social and business communities [3]. Clearly, the AmI vision is part of a much wider emerging technological trend (dating back to Mark Weiser) that goes by the name of *ubiquitous* (or *pervasive*) computing and that aims to take full advantage of the relentless miniaturization in the semiconductor arena and the convergence of consumer electronics, (wireless) networking and mobile communication.

To progress towards the AmI goals, three major functionalities need to be realised:

- *Context awareness:* The data streams from a plethora of multimodal sensors embedded in the environment need to be processed and fused to derive a semantically accurate interpretation of the users identity and actions. It should be noted that sensing spectrum is rapidly expanding in step with progress in consumer electronics.
- *Ubiquitous access:* Extensive and self-configuring networks guarantee seamless and trusted communication between a host of devices and support for ubiquitous access to communication, information and services.
- *Natural interaction:* The pervasive computing becomes invisible as a new communication semantics is established between the AmI system and its human users.

Due to space limitations it is impossible to include a detailed list of all major research efforts. Suffice it to say that highly visible initiatives include the European FP6-IPs *Multi-modal Services for Meetings and Communications* (CHIL [4]), *Cognitive Robot Companion* (COGNIRON [5]), *Context Aware Vision using Image-based Active Recognition* (CAVIAR [6]), *Ambient Intelligence for the Networked Home Environment* (AMIGO [7]), *Human-Computer Interfaces Similar to Human-Human Communication* (SIMILAR [8]), *Ambient Intelligence for Mobile Communications through Wireless Sensor Networks* (e-SENSE) [9], as well as MIT's Project Oxygen [10], MERL's Ambient Intelligence for Better Buildings [11] and Georgia Tech Aware Home [12]. All these projects focus on one or more aspects of the AmI experience. They include (among others) adaptive houses that automatically adjust lighting and temperature settings to achieve optimal user satisfaction at minimal energy costs [13], [14], smart meeting rooms that provide late-comers with a summary of the arguments so far [4], and smart beds that unobtrusively monitor the sleep pattern, heart and breathing rate of seniors.

II. Sensor Networks and Perceptual Technologies

The role of sensor networks in an AmI environment is to furnish the higher levels of the system with answers to the *W5+* questions: **Who:** Tracking and identifying persons and pets, i.e. the actors of the AmI environment; **Where and When:** Providing a time frame for location and object associations to determine context; **What:** Recognizing activities, interactions, spatio-temporal relations, but also linguistic and non-linguistic messages, signals, and signs; **Why:** Association of actions with action semantics, scripts and plans, identification of tasks and behaviour patterns; **How:** Tracing the information flow through multiple modalities, recognizing expressions, movements, gestures.

The information provided by the sensors is used to drive systems that automatically analyse human behaviour [15]. As it is impossible to give a comprehensive overview of the sensors most commonly used in AmI application, we will restrict ourselves to the most important classes.

A. Audiovisual Sensors

Traditionally, research has focused mainly on audio-visual observations as these modalities provide almost all signals that humans make use of in interpersonal communication [16], [17]. Most prominent in the visual domain are face detection and identification [18], person and object tracking [19]–[21], facial expression recognition [22], body posture recognition [23], [24], attention direction sensing [25], hand tracking and hand gesture recognition [26].

Audio has been used most extensively to detect speech and identify and localize speakers (a.o. through microphone arrays [27]), speech recognition [28], and estimation of auditory features relevant to communication. Other uses of audio modality include determining properties of the source, e.g. detecting head orientation of speaking persons in the room. Recently more attention has been paid to general auditory signal analysis in order to identify broader classes of sounds. These include non-speech vocalizations, e.g. laughter or shouts for their obvious relevance for activity and emotion recognition, as well as sounds associated with various activities [29]. A striking example is the ShotSpotter system, which utilises connected roof-mounted microphones to detect gunshots and triangulate the location of origin. The system can even direct cameras to try and record the scene.

B. Passive Infrared Sensors (PIR)

Passive infrared sensors (PIR) register the infrared emissions from objects (notably humans, animals and vehicles) and since their spectral signature may be quite distinct from the visual one, they provide complementary features for motion and object classification. With a network of PIR devices deployed in the AmI environment, it becomes possible to detect fire and smoke, but also to learn movement patterns of the inhabitants, and novel correlations.

C. RFID

Radio-frequency identification (RFID) technology is making inroads in the AmI setting as it allows both sensing of object proximity and identification. Passive RFID tags are small, flexible and do not require endogenous power to operate. They can therefore be placed on every-day objects, woven into fabric, or even injected into animals or people (for a more in-depth discussion we refer to [30]). Strategically positioned RFID readers will be able to identify the objects or persons that pass in their proximity. Collecting the data from different readers allows the system to piece together the whereabouts of the main actors in an AmI scenario. With the right technology, it is even possible to infer activity patterns, as demonstrated by the so-called *iBracelet*, a device developed by Intel resembling a wrist-watch which contains a small, short-range RFID reader. This reader registers the ID-tags on all objects the user touches. By sticking a large number of RFID tags on a variety of objects and appliances (e.g. kettle, faucet, coffee mugs, etc.) in the household or office, it becomes possible to infer what activity a person is involved in (e.g. making a cup of coffee). Tagging the users of an AmI system with RFID is also very useful for collecting ground truth data for face, gesture, body posture and speech recognition applications, which usually incorporate statistical models that require large amounts of data for robust operation.

D. Multimodal Wearables

Thanks to the advances in solid state electronics it is now possible to integrate a variety of sensors into compact wearable devices. One example is the experimental SenseCam [31] developed by Microsoft Research Cambridge that combines a digital camera with a number of other sensors, including: light-intensity and light-color sensors, a PIR detector (for body heat), a temperature sensor, and a multiple-axis accelerometer. In addition there is an audio level detector, an audio recorder and a GPS. Certain changes in sensor readings can be used to automatically trigger the camera: e.g. a significant change in light level, or the detection of body heat in front of the camera can cause the camera to take a picture. The SenseCam has been used in an archival experiment to create a “lifetime store of everything” (*MyLifeBits* [32]). SUN Microsystems has launched a similar initiative with their SUNSPOT wearable sensor [33].

Researchers are also beginning to take advantage of the possibilities offered by modern cell-phones that are increasingly packed with a variety of sensors. The iPhone is a case in point: it comes equipped with an accelerometer, an ambient light sensor and an infrared sensor. Moreover, a combination of GPS, Bluetooth devices and basestation triangulation allows for an approximate location determination. For a more in-depth description of one such experiment we refer to section III-D.

III. Mining for Temporal Patterns

A. Introduction

In the preceding section we have outlined the wide and expanding range of sensors that are being deployed in AmI environments. In the wake of these technological developments we are witnessing an equally vigorous expansion in the arsenal of data analysis tools used to process the data streams generated by these networks. Most research has focused on sensor networks that have been set up with a specific scenario in mind with respect to which the incoming data streams can be checked. They are therefore relatively specific (“narrow-minded” if you like) about the data they monitor and the corresponding events they detect and report upon. However, with sensor availability and connectivity increasing by leaps and bounds, it has become viable to take a more experimental and “open-minded” stance and allow a network of sensors and computational devices to pro-actively inspect the many data streams that impinge on it in an effort to uncover unanticipated meaningful spatio-temporal patterns or associations.

Although interesting in their own right, such associations might in fact have a significant practical value as they can contribute to the robustness of the sensing process. For instance, if it is observed that large readings from sensor A are usually accompanied by a strong signal from sensor B, then the firing of sensor B might add support to a less than convincing peak in the signal from A that would otherwise have been missed. Furthermore, once reliable temporal patterns have been established, they can be used by the system to predict forthcoming events. Such predictions are extremely useful when planning future actions (e.g. an AmI house could switch on the heating 15 mins before the occupants are expected to arrive) or in data-driven attention mechanisms: attention levels are increased when the actual events turn out to differ significantly from the predicted ones (e.g. in an assisted living scenario for senior citizens). These developments have therefore spurned a renewed interest in data mining and prediction algorithms for time series [34].

A large body of work in multimodal signal processing for modeling temporal patterns deals with applications of Hidden Markov Models (HMM) which, building on their success in speech recognition (e.g. see [35] for a nice overview), have become the mainstay of spatio-temporal segmentation (e.g. see [36]). However, for the applications we have in mind (clustering of time series emanating from sensors) the classical HMM approach has two main disadvantages. First of all, the standard estimation algorithms (i.e. evaluation, decoding and learning) assume that the topology of the HMM-structure (in terms of states and transitions) is known. Clearly, when it comes to data mining, finding the structure is the crux of the problem. Rao and Cook [37] try to remedy this by defining high-level inhabitant activity states as clusters of elementary actions. A task-based Markov model represents the inhabitant activities as states in a simple hidden Markov model, which is then used to predict the next user activity. Secondly, (as pointed out in [38]) Markov models have difficulty incorporating temporal patterns across different timescales. For these reasons we will briefly highlight four alternative approaches that seem better suited to deal with the problem of detecting temporal patterns in sensor data streams (for additional references we refer to [34]).

B. T-patterns

The first approach we inspect is proposed by Magnusson [39]. He proposes an exhaustive search for recurring temporal patterns (dubbed *T-patterns*) in symbolic time series, where each symbol represents the onset of a particular event or activity. The principle here is to investigate possible relationships between pairs of symbols and then build trees of such temporal dependencies in a hierarchical fashion. To this end, the notion of *Critical Interval (CI)* is introduced: $[d_1, d_2]$ is considered as a CI for the pair of symbols (A, B) if an occurrence of A at time t implies that B is more likely to occur in the time interval $[t + d_1, t + d_2]$ than in a random interval of the same size. The standard p-value is used as a measure for how exceptional the occurrence regularity of the combination under scrutiny is. Patterns that turn out to be significant (with respect to a pre-defined p-value) are assigned a new symbol, whereupon the search resumes.

C. Clustering data streams from ultra-low resolution sensors

The second approach is a methodology for clustering the time series generated by the low-resolution sensors (e.g. binary interruption or motion sensors) [40]. In their work, the authors extol some of the virtues of these sensors. First, as they are very cheap, it is possible to install a very dense networks of sensors at minimal cost. In addition, they are seen to be far less intrusive and privacy-critical than high-resolution sensors like surveillance cameras. Finally, for simple applications (e.g. monitoring movements of people inside a building), low-resolution sensors achieve results comparable to the ones obtained with high-resolution sensors.

The proposed method starts by segmenting the training sequence into small subsequences. Similarity between these subsequences is defined in terms of their generating HMMs. More precisely, for each subsequence s_i a single HMM H_i is trained. Similarity between s_i and s_j is then defined by computing how likely s_i is with respect to HMM H_j and vice versa. The similarities thus obtained are then used to incrementally build a hierarchical cluster tree: whenever two nodes (each comprising one or more subsequences) are merged into one cluster, a new updated HMM is trained using all the subsequences in the proposed cluster.

This method is demonstrated with an experiment, in which a department comprising several offices and common rooms was wired with binary motion sensors. Peaks in the sensors' cross-correlation functions are indicative of their physical separation. Long-term observations of the sensors' activation patterns therefore generate an inter-sensor distance matrix, which can be used to reconstruct an approximate physical layout (e.g. using standard algorithms such as multi-dimensional scaling). This shows how careful analysis of multiple data streams can obviate the need for manual calibration.

D. Eigenbehaviors

Analysis of human behaviour patterns is made easy with the introduction of sensing capabilities to cell phones. Working with a large collection of data captured from Bluetooth-equipped mobile phones continuously logging location, proximate people, and communcation of 100 subjects at MIT during the course of nine months, Eagle and Pentland [38] develop methods to uncover daily human behaviour routines. Encoding the historic data for each subject in an activity matrix (where each row represents a chronology of one day's location logging) it becomes possible to uncover important temporal patterns by performing PCA and focusing on the most prominent eigenvectors (dubbed *eigenbehaviors*). These eigenbehaviors often have clear semantic interpretation (e.g. weekday versus weekend) and are able to capture long-delay correlations (e.g. sleeping late in the morning is a good predictor for being out late that night). In fact, it turns out that these low-dimensional models are sufficiently accurate to predict (with nearly 80% accuracy) the afternoon's activities based on the data collected in the morning.

E. Active LeZi: Compression-based Pattern Extraction

A final approach to detecting temporal patterns is to use the LZW compression algorithm as a pattern extractor [34]. This algorithm can easily accommodate streaming data and real-time learning as the pattern table (the so-called dictionary) is easily updated each time a symbol is added to the stream. The patterns based on the dictionary thus obtained can be post-processed by removing patterns that are either very rare, or exhibit a high variability in their duration.

IV. Conclusion

Finding temporal patterns in multiple streams of sensor data is essential for automatic analysis of human behaviours and habits in the ambient environment. In this paper we have reviewed the context of this problem as it pertains to recent sensor technology, as well as four different approaches that have different merits and weak points. As with many problems in computer science, there is no single best approach, and choice rests on the system designer.

Acknowledgment

This work was done while R.T. was at CWI. The authors gratefully acknowledge partial support by EC's NOE MUSCLE (FP6-507752), and the Dutch BSIK/BRICKS project.

References

[1] E. Aarts, H. Harwig, and M. Schuurmans, "Ambient intelligence," in *The Invisible Future*, J. Denning, Ed. New York, NY, USA: McGraw Hill, 2001.

[2] E. Aarts and J. Encarnaçao, *The Emergence of Ambient Intelligence*. Berlin, Germany: Springer, 2006.

[3] European Commission IST Advisory Group (ISTAG). Ambient Intelligence: from vision to reality. Consolidated Report 2003. [Online]. Available: ftp://ftp.cordis.lu/pub/ist/docs/istag-ist2003_draft_consolidated_report.pdf

[4] CHIL – Computers in the Human Interaction Loop . [Online]. Available: http://chil.server.de/servlet/is/101/

[5] COGNIRON: The Cognitive Robot Companion. [Online]. Available: http://www.cogniron.org/Home.php

[6] (2002-5) CAVIAR – Context Aware Vision using Image-based Active Recognition . [Online]. Available: http://homepages.inf.ed.ac.uk/rbf/CAVIAR/

[7] AMIGO – Ambient Intelligence for the Networked Home Environment. [Online]. Available: http://www.hitech-projects.com/euprojects/amigo/index.htm

[8] SIMILAR – Human-Machine Interfaces Similar to Human-Human Communication. [Online]. Available: http://www.similar.cc

[9] M. Presser, A. Gluhak, D. Babb, L. Herault, and R. Tafazolli, "e-SENSE – Capturing Ambient Intelligence for Mobile Communications through Wireless Sensor Networks," in *15th IST Summit - Greece Workshop*, 2006.

[10] MIT Project Oxygen: Pervasive Human-Centered Computing. [Online]. Available: http://oxygen.csail.mit.edu/Overview.html

[11] Mitsubishi Electric Research Laborotories: Ambient Intelligence for Better Buildings. [Online]. Available: http://www.merl.com/projects/ulrs/

[12] C. Kidd, R. Orr, G. Abowd, C. Atkeson, I. Essa, B. MacIntyre, E. Mynatt, T. Starner, and W.Newstetter, "The aware home: A living laboratory for ubiquitous computing research," *Proceedings of CoBuild 99*, 1999.

[13] M. C. Mozer, "Lessons from an adaptive home," in *Smart Environments: Technologies, Protocols, and Applications*, D. J. Cook and S. K. Das, Eds. Wiley Series on Parallel and Distributed Computing, 2005, pp. 273–294.

[14] D. Cook, M. Youngblood, E. Heierman, K. Gopalratnam, S. Rao, A. Litvin, and F. Khawaja, "Mavhom: An agent-based smart home," in *Proc. IEEE PerCom*, 2003, pp. 521–524.

[15] M. Pantic, A. Pentland, A. Nijholt, and T. Huang, "Human computing and machine understanding of human behavior: A survey," *Lecture Notes in Artificial Intelligence*, vol. 4451, pp. 47–71, 2007.

[16] J.L. Crowley, J. Coutaz, and F. Berard, "Things that see: Machine perception for human computer interaction," *Communications of the ACM*,.

[17] J.L. Crowley, "Context driven observation of human activity; keynote presentation," in *European Symposium on Ambient Intelligence*, Eindhoven, Nov 2003.

[18] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," 2001.

[19] T. B. Moeslund and E. Granum, "A survey of computer vision-based human motion capture," *Computer Vision and Image Understanding: CVIU*, vol. 81, no. 3, pp. 231–268, 2001.

[20] W. Wolf, B. Ozer, and T. Lu, "Smart cameras as embedded systems," *IEEE Computer*, vol. 35, no. 9, pp. 48–53, 2002.

[21] T. Zhao and R. Nevatia, "Tracking multiple humans in complex situations," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 9, pp. 1208–1221, 2004.

[22] Y.-I. Tian, T. Kanade, and J. Cohn, "Recognizing action units for facial expression analysis," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 2, pp. 97–115, 2001.

[23] C. Wren, A. Azarbayejani, T. Darrell, and A. Pentland, "Pfinder: real-time tracking of human body," *IEEE Transactions Pattern on Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 780–785, 1997.

[24] R. Cucchiara, C. Grana, A. Prati, and R. Mezzani, "Probabilistic posture classification for human-behavior analysis," *IEEE Trans. On Systems, Man, And Cybernetics-Part B: Cybernetics*, vol. 35, no. 1, 2005.

[25] R. Stiefelhagen, J. Yang, and A. Waibel, "Estimating focus of attention based on gaze and sound," in *Workshop on Perceptive User Interfaces (PUI '01)*, 2001.

[26] V. Pavlovic, R. Sharma, and T. S. Huang, "Visual interpretation of hand gestures for human-computer interaction: A review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, no. 7, pp. 677–695, 1997.

[27] D. Ward and M. Brandstein, *Microphone Arrays: Signal Processing Techniques and Applications*. Springer, 2001.

[28] F. Jelinek, *Statistical Methods for Speech Recognition*. MIT Press, 1998.

[29] G. Wichern, H. Thornburg, B. Mechtley, A. Fink, K. Tu, and A. Spanias, "Robust multi-feature segmentation and indexing for natural sound environments," in *Proc. CBMI'2007*, pp. 69–76.

[30] S. Shepard, *Radio Frequency Identification*. McGraw-Hill, 2005.

[31] SenseCam by Microsoft Research. [Online]. Available: http://research.microsoft.com/sendev/projects/sensecam/

[32] G. B. J. Gemmell and R. Lueder, "Mylifebits: A personal database for everything," *Communications of the ACM*,, vol. 49(1), pp. 88–95, January 2006.

[33] SUN Project Small Programmable Object Technology (SUN SPOT). [Online]. Available: http://www.sunspotworld.com/

[34] D. J. Cook, "Prediction algorithms for smart environments," in *Smart Environments: Technologies, Protocols, and Applications*, D. J. Cook and S. K. Das, Eds. Wiley Series on Parallel and Distributed Computing, 2005, pp. 175–192.

[35] L. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of IEEE*, vol. 77, no. 2, pp. 257–285, 1989.

[36] M. Bicego, M. Cristani, and V. Murino, "Unsupervised scene analysis: a hidden markov model approach," *Comput. Vis. Image Underst.*, vol. 102, no. 1, pp. 22–41, 2006.

[37] S. Rao and D. Cook, "Predicting inhabitant action using action and task models with application to smart homes," *International Journal on Artificial Intelligence Tools*, vol. 13, no. 1, pp. 81–99, 2004.

[38] N. Eagle and A. Pentland, "Eigenbehaviors: Identifying structure in routine," in *Proc. of Ubicomp'06*, 2006.

[39] M. S. Magnusson, "Discovering hidden time patterns in behavior: T-patterns and their detection." *Behav Res Methods Instrum Comput*, vol. 32, no. 1, pp. 93–110, February 2000.

[40] C. Wren, D. Minnen, and S. Rao, "Similarity-based analysis for large networks of ultra-low resolution sensors," *Pattern Recognition*, vol. 39, pp. 1918–1931, 2006.