Proceedings of the 5th International Workshop on Wearable and Implantable Body Sensor Networks, in conjunction with
The 5th International Summer School and Symposium on Medical Devices and Biosensors
The Chinese University of Hong Kong, HKSAR, China. Jun 1-3, 2008

Probabilistic Decision Level Fusion for Real-Time Correlation of Ambient and Wearable Sensors

D. G. McIlwraith, J. Pansiot, S. Thiemjarus, B. P. L. Lo and G. Z. Yang

***Abstract*— Fusing data from ambient and wearable sensors when performing in-home healthcare monitoring allows for high accuracy activity inference due to the complementary nature of sensing modalities. Where residences may house multiple occupants, we must automatically identify related data streams before fusion may occur, a process known as sensor correlation. In this paper a multi-objective variant of the Bayesian Framework for Feature Selection (BFFS) is used to construct small inter-sensor redundant feature sets which train efficient per-sensor activity classifiers. Probabilistic decision level fusion is then used to deal with noisy and erroneous sensor data and perform real-time correlation. The potential value of the proposed algorithm for pervasive sensing is demonstrated with both simulated and experimental data.**

I. Introduction

The general trend of an increasingly aged population has placed significant burdens on existing healthcare systems. This has motivated the recent development of effective home healthcare environments. By continuously monitoring patient behaviour and activity using wearable and ambient sensors, it is possible to gain an insight into the wellbeing of the elderly and identify transient abnormalities that may lead to major adverse events. For the patient, this can provide a higher level of independence and a better quality of care than has been previously possible.

Whilst the fusion of ambient and wearable sensors provides improved accuracy in activity recognition for those living alone [1], it becomes difficult to automate when dwellings may contain multiple occupants; as we must first establish a link between related sensors. This raises many technical challenges. Firstly, for issues of privacy and usability we must perform correlation without a-priori knowledge of identity. Secondly, the approach needs to be computationally efficient, allowing correlation to occur on-node and in real time. Finally, the technique must be able to operate in sub-optimal conditions when data obtained from subjects may be unreliable or contaminated by noise.

The purpose of our work is to provide a probabilistic sensor correlation framework operating at the *decision level*. By using per-sensor activity classifiers and calculating the probabilistic agreement in the underlying activity of the subject we can correlate sensors without prior knowledge of subjects and deal with potentially noisy data. Within the framework a feature selection algorithm is proposed to identify small feature sets from independent sensors. This facilitates feature generation and activity inference techniques that can be embedded into sensor hardware and execute in real time. These features sets are constructed with discriminatory overlap in mind – such that they may be utilised to detect activities independently *and simultaneously.*

Manuscript received February 25, 2008. Douglas G. McIlwraith, Julien Pansiot, Surapa Thiemjarus, Benny P. L. Lo and Guang-Zhong Yang are with the Department of Computing and Institute of Biomedical Engineering at Imperial College London, UK. Douglas G. McIlwraith may be contacted at the Department of Computing, Imperial College London, 180 Queens Gate, London. England; dm05@doc.ic.ac.uk

II. Background

To acquire ambient sensing data, it is common to use reed switches for door entry and appliance activation [2]. This can provide occupancy data at relatively low cost. Humidity, pressure, light, bearing and sound levels have also been used extensively [3]. Recently, video based sensors have been employed which can unobtrusively monitor the environment and the subjects themselves [4].

It has been shown [1] that ambient sensing can provide complementary contextual data to wearable sensors [5, 6]. There has also been significant research in the use of wearable cameras to provide information regarding the attention of the wearer [7] and how they interact with others [8].

The fusion of complementary data streams allows for high accuracy activity and behaviour profiling based on pattern recognition techniques such as Bayesian Networks [9], Hidden Markov Models (HMMs) [10] and Gaussian Mixture Models (GMMs) [1]. Due to an increase in dimensionality there is also a requirement for data reduction including feature selection [11], Principal Component Analysis (PCA) and manifold embedding [12].

III. Feature Generation & Selection

A. Ear Worn Activity Recognition (e-AR) Sensor

The wearable device used in this study is the *e-AR* ear-mounted sensor developed at Imperial College [13]. It is based on the Nordic nRF24E1 integrated chipset comprising the nRF2401 2.4GHz RF transceiver, MCU, analogue-to-digital converter and supporting hardware. A 3-axis accelerometer is also housed within the casing. During use, the *e-AR* sensor periodically samples from the accelerometer and transmits this data to a nearby *base station* where features such as head-tilt, mean, median and variance of the three dimensional signal are extracted.

B. Ambient Vision Based Blob Sensor

For environmental sensing, the *ambient blob-based vision sensor* [4] is used, consisting of video sensor, processor, wireless communication link and on-board battery. 'Blobs'

978-1-4244-2253-1/08/$25

that represent subjects under surveillance are extracted using a statistical model of the background [14] and the optical flow within is then calculated [15]. Whilst geometric features of the 'blobs' can reveal information regarding pose, optical flow can aid in the detection of local motion such as that of the limbs.

C. *Identifying Inter-Sensor Redundant Feature Sets*

Feature sets are extracted such that they exhibit classification redundancy across sensing modalities. Consequently, activities can be detected accurately and independently by per-sensor classifiers *at the same time*. Figure 1 demonstrates this concept. The area enclosed by a circle is directly proportional to the *discriminatory power* of the feature set. The cross-hatched area is proportional to *discriminatory overlap* - the proportion of data points that are correctly classified when either feature set is used independently and in the absence of the other. In the example presented here, *e-AR* feature set Ψ_A and blob set Ω_B are the most suitable pair, since they both have excellent independent discriminatory power as well as overlap.

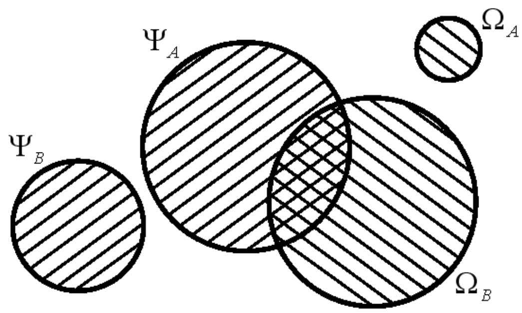

Figure 1: Demonstrating discriminatory overlap.

To reduce feature set search space we first apply BFFS [11] to rank sensor-specific features which, when used singularly and independently, yield high accuracy classification. Table I presents the top six features for each sensor. These results are based upon data from six subjects asked to perform the nine activities listed in Pansiot *et al.* [1], for duration of one minute each.

To automatically select sets with the highest redundancy across sensors we utilize the multi-objective accuracy/redundancy score $D(\Psi,\Omega)$ to rank feature sets [16];

$$D(\Psi,\Omega) = -(1-\theta)\times(E_{AUC}(\Psi\cup\Omega) - E_{AUC}(\Omega)) + \theta\times(E_{AUC}(\Psi)+E_{AUC}(\Omega))$$

Candidate *e-AR* feature sets are represented by Ψ whilst *blob* sets are denoted by Ω. Factor θ allows us to express the relative importance of per-sensor accuracy and inter-sensor redundancy for those two sets. To limit the size of feature sets constructed we use a maximum of three features - taken from the top ranked for that sensor. Table II presents the top five *e-AR/blob* feature-set pairs, ranked using value 0.6 for θ.

TABLE I
TOP PER-SENSOR FEATURE S

	e-AR Features		**Blob Features**
1	Tilt Forwards/Backwards	1	Bounding Box Aspect Ratio
2	Y Acceleration	2	Bounding Box Eigenvector (Horizontal Comp.)
3	Tilt Left/Right	3	Bounding Box Eigenvector (Vertical Comp.)
4	Mean X Acceleration	4	Average Optical Flow
5	Median X Acceleration	5	Optical Flow Eigenvector (Horizontal Comp.)

IV. PROBABILISTIC DECISION LEVEL CORRELATION

With the appropriate *e-AR* and *blob* feature sets selected, they are exposed to pre-trained per-sensor activity classifiers. We note that, whilst from an *e-AR* sensor we may generate a single feature *stream*, several may be generated from a *blob* sensor each relating to a 'blob' in its field of vision.

TABLE II
TOP INTER-SENSOR REDUNDANT FEATURE SETS

	1	**2**	**3**	**4**	**5**
Ψ	1, 2, 3	2, 3	1, 3	1, 2, 3	1, 2, 3
Ω	1, 2, 3	1, 2, 3	1, 2, 3	1, 2	1, 3
D	1.1085	1.0931	1.0880	1.0745	1.0684

For each per-sensor classifier, pre-calculated confusion matrices are used to determine $P(A\mid C)$, the probability of an activity A given classification C. By noting that

$$P(A) \propto P(A\mid C^{e-AR})\times P(A\mid C^{Blob})$$

we calculate the probability of a match in subject activity given two classifications based on data from different sensors. In order to recognise that recent activity matches are more important in correlation, and mitigate the effect of classification noise, we further calculate a correlation score, M, that linearly weights the probabilities of an activity match from the last *n* samples.

$$\mathrm{M}(i) = \sum_{j=0}^{n-1}(n-j)P_{match}(i-j)$$

For each combination of *e-AR* and *blob* classifier outputs we calculate M over time and analyze their relative values, noting that the combination of classifications based upon data from the same subject will be higher. The proportion of time that a pair of streams is matched over others is used to provide resilience to occasions where subjects are performing similar activities.

V. Experiments & Results

A. Experiment Setup

Data from 5 subjects performing 9 activities [1] is collected over a period of approximately 9 minutes and Gaussian Mixture Models are cross-trained using Expectation-Maximization (EM). Random *e-AR* and *blob* classification stream pairs are generated to add confusion to our experiment and simulate a number of different people within a household.

Table III summarises our findings and presents two accuracy metrics. *Correlation accuracy* is the percentage of time the correct sensors were matched throughout the experiment, whereas *final accuracy* is the percentage of sensors that were correctly matched at the end of the observed period. For each subject, correlation accuracy is averaged over 10 iterations and with up to 5 simulated people. This is further averaged over all 5 subjects.

All feature sets demonstrate a high accuracy of pairing. As we may expect, *correlation accuracy* is marginally lower than *final accuracy* due to periods within the experiment where subject activities are indistinguishable.

B. Real Time Operation

To demonstrate system operation we present 2 scenarios involving 2 subjects monitored over approximately 4 minutes. In both instances, per-sensor classifiers were trained using feature set 5 from Table II.

During scenario 1, subject 1 is asked to perform a circuit of activities, starting with s*itting* through *reading, eating* and *standing*. Subject 2 is asked to get up from a *lying* position, *sitting* on their bed, then *stand* and *walk*. For the second scenario, subject 1 was asked to *walk* to the bed from an initially *seated* position, and *lie* down. The other subject was instructed to *sit, eat*, and *read* at different locations within the room, *walking* between them.

For each scenario, Figures 2 and 3 *illustrate the correlation accuracy* as a function of time for both the related *e-AR/blob* pair (subject 1, in blue) and the same *e-AR* sensor paired with an erroneous *blob* sensor (subject 2, in purple). A correct match occurs when the blue trace is above the purple, i.e. a sensor couple is paired together more often than the alternatives. The activities returned from the *e-AR* and *blob* classifiers for subject 1 are also presented in each instance.

Figure 2, shows that correct *e-AR* and *blob* sensors are matched immediately and for the entire duration of scenario 1. For the first 30s, the correct sensors are always matched, since *sitting/slouching* classes returned from the second subject are never confused with *eating*. After a small reduction, an increase occurs as r*eading* is more likely paired with *eating* than s*louching,* returned from the blob sensor monitoring subject 2. At 3m28s a drop in accuracy is observed due to increased confusion with the second subject, who is now *standing* and *walking*. During the final 10s of the experiment, accuracy begins to increase as *head tilt* is observed by both sensors monitoring subject 1.

TABLE III
Accuracy Obtained With Up To 6 Subjects

	1	2	3	4	5
Correlation Accuracy (%)	95	95	94	90	95
Final Accuracy (%)	97	100	100	100	100

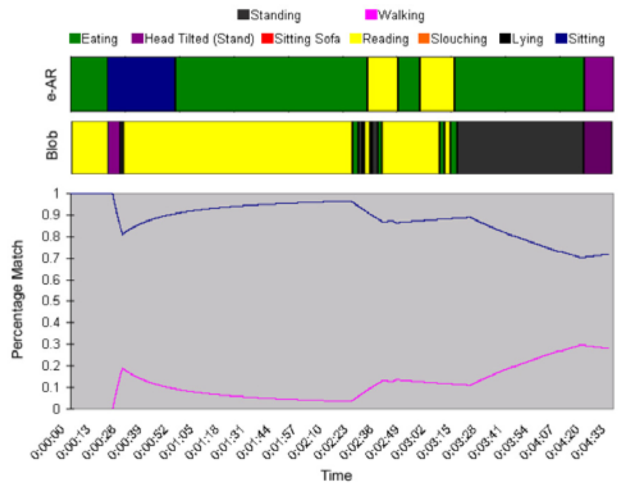

Fig. 2: Percentage match of e-AR classification against blob classification from same subject (blue) and different subject (purple) in scenario 1.

During the first 40s of the second scenario, the *e-AR* sensor worn by subject 1, who is getting up to prepare for bed, is more frequently paired with the blob sensor monitoring subject 2, since activity *head tilted* is matched similarly during a period of motion. Between 40s and 2m38s, subject 1 becomes seated on the bed and this can be readily seen by the classifications returned. During this period, the second subject has been primarily classified using data from the *blob sensor* as *walking*, with brief periods of *eating*. This

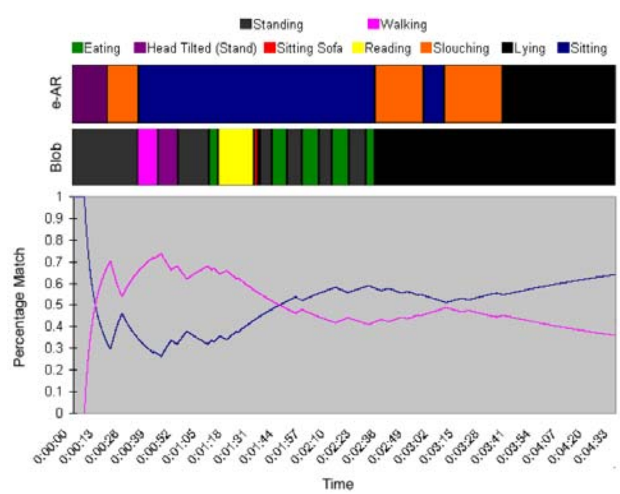

Fig. 3: Percentage match of e-AR classification against blob classification from same subject (blue) and different subject (purple) in scenario 2.

creates low confusion with the incorrect subject and leads to a steady increase in classification accuracy – with 50% first crossed at 1m57s. As subject 1 reclines, *slouching* is detected by the *e-AR* sensor, and *lying* from the associated blob sensor, distinguishing subjects well. Correlation accuracy further increases from 3m15s onwards, resulting in a match of the correct *e-AR* and *blob* sensor at the end of the experiment.

VI. CONCLUSION

In this paper we have presented a decision-level approach for real-time automatic correlation of ambient and wearable sensors that requires no knowledge of subject identity or location. Whilst small feature sets have been built with discriminatory overlap in mind we note that, for probabilistic correlation, this may not be as important as inter-sensor agreement with regard to the most probable activity performed. We plan to extend our algorithm to consider this as a parameter in feature selection. It should also be noted that the proposed system assumes the *blob* sensor may reliably extract per subject data and issues of occlusion are ignored. Whilst problematic to feature extraction, we may combine this system with temporal tracking to alleviate its effect on classification accuracy.

REFERENCES

[1] J. Pansiot, D. Stoyanov, D. McIlwraith et al. (2007) Ambient and Wearable Sensor Fusion for Activity Recognition in Healthcare Monitoring Systems, IFMBE Proceedings of the 4th Int. Workshop on Wearable and Implantable Body Sensor Networks (BSN 2007), Aachen, Germany, 2007, pp 208-212, DOI 10.1007/978-3-540-70994-7_36

[2] A. Arcelus, M. H. Jones, R. Goubran et al. (2007) Integration of Smart Home Technologies in a Health Monitoring System for the Elderly, IEEE Proceedings of the 21st Int. Conference on Advanced Information Networking and Applications Workshops (AINAW 2007).

[3] J. Lester, T. Choudhury, N. Kern et al. (2005) A Hybrid Discriminative/Generative Approach for Modeling Human Activities, Proc. Of the 19th Int. Joint Conference on Artificial Intelligence (IJCAI 2005), Edinburgh, Scotland, 2005

[4] B. P. L. Lo, J. L. Wang and G. Z. Yang (2005) From Imaging Networks to Behavior Profiling: Ubiquitous Sensing for Managed Homecare of the Elderly. Proc. International Conference on Pervasive Computing (PERVASIVE 2005), pp. 101-104

[5] K. V. Laerhoven, H. W. Gellersen, and Y. G. Malliaris (2006) Long Term Activity Monitoring with a Wearable Sensor Node, International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2006)

[6] L. Dunne, P. Walsh, B. Smyth et al. (2007) A System for Wearable Monitoring of Seated Posture in Computer Users, IFMBE Proceedings of the 4th Int. Workshop on Wearable and Implantable Body Sensor Networks (BSN 2007), Aachen, Germany, 2007, pp 203-207, DOI 10.1007/978-3-540-70994-7_35

[7] C. A. Rothkopt and J. B. Pelz (2004) Head Movement Estimation for Wearable Eye Tracker, Proc. of the 2004 Symposium on Eye Tracking Research & Applications, San Antonio, Texas, USA, 2004, pp 123-130

[8] S. Krishna, G. Little, J. Black et al. (2005) iCARE Interaction Assistant: A Wearable Face Recognition System for Individuals with Visual Impairments, Proc. of the 7th Int. ACM SIGACCESS Conference on Computers and Accessibility, Baltimore, USA, 2005

[9] D. Chen, R. Malkin and J. Yang (2004) Multimodal Detection of Human Interaction Events in a Nursing Home Environment, Proceedings of the 6th International Conference on Multimodal Interfaces (ICMI 2004), PA, USA, 2004, pp 82-89

[10] D. Sanchez, M. Tentori and J. Favela (2007) Hidden Markov Models for Activity Recognition in Ambient Intelligence Environments, Proc of the 8th Mexican Conf. on Current Trends in Computer Science (ENC 2007), pp 33-40

[11] G. Z. Yang (Ed.) (2006) Body Sensor Networks. Springer-Verlag, London, pp 262-274

[12] A. Agovic, A. Banerjee, A. Ganguly and V. Protopopescu, Anomaly Detection in Transportation Corridors using Manifold Embedding, 1st International Workshop on Knowledge Discovery from Sensor Data (Sensor-KDD) 2007

[13] O. Aziz, B. Lo, R. King, A. Darzi and G. Z. Yang (2006) Pervasive Body Sensor Network: An Approach to Monitoring the Postoperative Surgical Patient, International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2006)

[14] Dar-Shyang Lee (2005) Effective Gaussian Mixture Learning for Video Background Subtraction, IEEE Transactions on Pattern Analysis and Machine Intelligence ,vol. 27, no. 5, pp. 827-832

[15] B. K. P. Horn and B. G. Schunck (1981) Determining Optical Flow. Artificial Intelligence 17:185-203

[16] S. Thiemjarus (2007) PhD Thesis, Imperial College London, A Framework for Contextual Data Fusion in Body Sensor Networks, pp. 91-92