Multimodal Sensing for Pediatric Obesity Applications

M. Annavaram[†],N. Medvidovic[†], U. Mitra[†], S. Narayanan[†] G. Sukhatme[†],
Z. Meng[‡], S. Qiu[‡], R. Kumar[†], G. Thatte[†], D. Spruijt-Metz[§]
[†] Viterbi School of Engineering, University of Southern California, Los Angeles, CA 90089, USA
Email: {annavara,ubli,thatte}@usc.edu, {neno,gaurav}@cs.usc.edu, shri@sipi.usc.edu
[‡] Tsinghua University, Beijing China
Email: {third,forth}@institution.edu
[§] Keck School of Medicine, University of Southern California, Los Angeles, CA 90089, USA
Email: metz@usc.edu

Abstract

In this paper, a wireless body area network comprised of heterogeneous sensors is developed for wearable health monitoring applications. The ultimate application space is in the context of pediatric obesity. The specific task examined herein is activity detection based on heart rate monitor and accelerometer data. Based on statistical analysis of experimental data for different key states (lying down, sitting, standing, walking and running), a multi-modal detection strategy is proposed. The resulting detector can achieve 85-95% accuracy in state detection. It is observed that the accelerometer is more informative for the active states, while the heart rate monitor is more informative for the passive states.

I. Introduction

Wearable health monitoring systems coupled with wireless communications are the bedrock of an emerging class of sensor networks: wireless body area networks (WBAN). The objectives of such WBANs are manifold from diet monitoring [14], activity detection [3], [4], and health crisis support[6]. These new networks demand significant technological advances from sensor development to novel software engineering, signal processing, wireless communications and networking. Importantly, WBANs must be designed with application-specific design and end-use requirements in mind. These advancements are necessary to cope with the unique challenges introduced by deployment **on** people, such as: unpredictable mobility, heterogeneous sensor nodes, new wireless channels, very low power requirements, non-invasive sensing and the need for sensors with small footprints. Furthermore, drawing robust inference from sensor streams requires information from multiple, often disparate, sources. In the current work, we provide preliminary results from the construction of a WBAN which we will use to drive the development of assessments and interventions for pediatric obesity applications.

Pediatric obesity has emerged as a major national and international health crisis. National collected data from 2003-2006 show 11.3% of adolescents aged 12 - 19 years by some measures could be designated as obese; a further 16% would be classified as overweight and 32% considered at risk for being overweight [13]. While physical activity (PA) is tightly related to lower obesity rates in children [11], [7], there are additional factors leading to obesity. The increasing environmental stress may promote both general obesity (through lifestyle behaviors such as decreased physical activity) and visceral obesity (through hypothalamic-pituitary-adrenal axis activation and increased cortisol secretion)[5]. Current monitoring systems validated for research in children typically monitor physical activity only (such as the much-used Actigraph accelerometer). However, in order to truly understand and reverse childhood obesity, we need a multimodal system that will track stress levels, PA levels, blood glucose levels and other vital signs simultaneously, as well as anchor these levels to context such as time of day and geographical location. Our preliminary KNOWME network is a first step towards such a system.

A key aspect of our work is the unified design and evaluation of multimodal sensing and interpretation, for automatically recognizing, predicting and reasoning about human physical activity and socio-cognitive behavior states. On the one hand, this meets the needs of traditional observational research practices in the obesity and metabolic health domain (based on, and validated through, careful expert human coding of data) while on the other, this enables new analysis capabilities that have not been possible before such as providing information on user emotional state in conjunction with physical activity and energy expenditure.

Many aspects of human behavior are inherently multimodal or require multimodal processing. For example, measuring and understanding energy expenditure and its etiology requires processing not only activity from accelerometers but other data such as pulse rate, ECG, oxygen intake, as well as contextual information such as emotions that are marked by humans through their voice, body posture and through physiological signals skin conductance measures (electro dermal response). Hence, to model human behavior and task-specific activity, both in terms of what people do, how they do it, and *why they do it*, it is critical to understand and capture the interplay between such multimodal streams. Multi-modal coverage of our approach enables cross-channel comparison and verification (allowing us, for example, to capture relationships between increased heart rate, increased emotional activity, and changes in physical activity). Our approach to this problem is grounded in statistical signal processing.

In the current work, we summarize preliminary results on activity assessment. We consider a mix of low mobility (lying down, sitting, standing) and higher mobility (walking, running) states. Features of our problem and approach do appear in the prior literature. Much work on activity detection appears to center on accelerometer data alone (*e.g.*[8], [3], [10]) with some systems employing many accelerometer packages. On the other hand, multi-sensor WBANs have been implemented and deployed (see *e.g.* [12], [9], [6]); however in those works, the emphasis was on the higher layer communication network processing and hardware design – signals from each sensor were transmitted directly to a central decision making unit. Our focus is on a modest number of heterogeneous sensors and the utilization of multi-modal signal processing methods; we wish to design decision making and data interpretation methods that will reside within the WBAN and allow for interaction with the WBAN wearer. For our pediatric obesity application, activity detection is an indirect measure of energy expenditure quantification as discussed above. In [4], multi-modal classification **is** considered. There are some key differences to the approach taken herein. First, while different sensors are employed, they are similar in the types of measurements taken (*e.g.* accelerometers, gyroscopes and tilt measurements), herein we use sensors which measure fundamentally different quantities that are correlated, but the statistical relationships are unclear *a priori*. The goal of [4] is to determine a sampling scheme (with respect to frequency of sampling and sleeping/waking cycles) for multiple sensors to minimize power consumption. The authors show that their new methods achieve reduced power relative to classical joint schemes. Our goal is on classifier performance with heterogeneous sensors – future versions of our methods could incorporate power minimization strategies of [4]. An important question to address is how the correlation between measurements affects power minimization. We conjecture that the sensors employed in [4] have more highly correlated observations with regards to the states of interest than our sensors and thus greater power minimization is possible through the use of their methods.

As our WBAN must be used for a diverse set of decision making processes, all sensors may not be uniformly useful for each task. We, in fact, see this with the activity detection problem considered herein.

II. KNOWME Network Architecture

The basic foundation of the KNOWME network is our three tier network architecture as depicted in Figure 1. The first tier's goal is data collection based on the heterogeneous sensors that are coupled to a mobile phone which acts as a "base station," equipped with data transmission and processing capabilities. The second tier is a web server that receives data and can perform additional processing; the web server transmits the data to the final tier: a back-end database server that stores the information. In the sequel, we shall discuss the specific sensors employed.

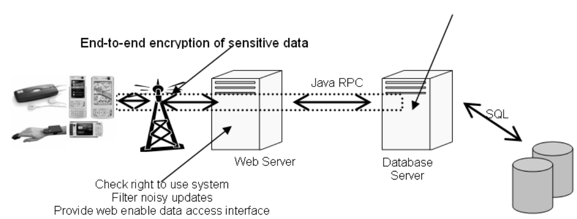

Fig. 1. Three-tier architecture overview of wireless body area network sensor system.

Currently, the primary focus of this research is to perform multi-modal sensing and interpretation of data to serve some of the end-user needs. As such, significant effort has been spent in integrating heterogeneous sensors to a mobile phone. One challenge in integrating heterogeneous sensors is that these sensors have different APIs, packaging, and data collection methods. In addition to integrating multiple sensors, synchronization of the data received from multiple sensors in the phone is critical for statistical correlation of sensor data and to perform the multi-modal data processing. Sensor information is continuously recorded on the local storage on the mobile phone. Our mobile device platform has a 8GB in-built flash memory that can be used for storing sensor information. Sensor data rates vary from 300bps for the accelerometers to 100 bps for the heart rate monitor. Using these data rates, we estimate that our 8GB local storage can store 1000 days worth of data. As the Bluetooth wireless link is a bottleneck for our current data collection, we use time-division multiple-access to schedule the data from different sensors (equal time share).

The software development phase uses well-known unit testing to extensively test the mobile software suite. In order to minimize errors in configuring the software, our software has several built-in checks to advise the user if any of the sensor readings do not match expected sensor behavior. Since the mobile device has to transmit the data to the backend servers, we are currently developing an opportunistic data transfer mechanism that uses an open WiFi network where available to transfer data both efficiently and cheaply. In the absence of WiFi networks, the mobile software is configured to automatically use the cellular data network to transmit the data. Our initial deployment is mostly with graduate and undergraduate student test subjects with limited (on-going) pilot experiments with children in the Exercise Physiology Lab at the USC Keck School of Medicine.

A. Sensor Systems

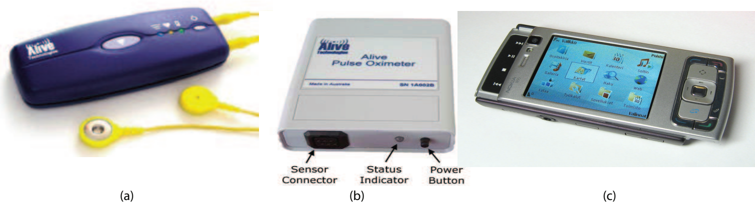

Fig. 2. (a) ECG monitor, (b) pulse oximeter, (c) Nokia Smartphone (GPS and accelerometer).

The sensor layer is a collection of off-the-shelf devices that measure features which can provide insight about metabolic activity; most (with the exception of galvanic skin response) are also capable of wirelessly transmitting this data over a Bluetooth interface. The current study employs an Alive Technologies[1] electrocardiograph (ECG). The ECG is a single channel device with 8 bit resolution and a peak sampling rate of 300 samples/second. The pulse-oximeter, also from Alive, provides non-invasive monitoring of oxygen saturation (SpO2) and pulse rate. The oximeter is a Bluetooth slave device that supports the Bluetooth Serial Port Profile (SPP). We also have BodyMedia WMS sensors [2] to measure Galvanic Skin Response (GSR) [1] and motion estimation using accelerometers. We use feature rich Nokia N95 as the mobile phone platform. N95 supports Bluetooth 2.0 + EDR for quick pairing with external Bluetooth sensors, and has 3G and WiFi radios for high bandwidth data transfer. In addition to the high bandwidth radio capabilities, the N95 mobile phone platform has a highly accurate built-in assisted GPS unit that uses a combination of GPS satellites, cellular tower and WiFi scanning to obtain a GPS position lock in less than 10 seconds. The stated location accuracy of GPS unit is 30 meters. We have observed accuracy at less than 3 meters in practice. The data collected from multiple sensors is geo-tagged using the location data collected from the in-built GPS. Furthermore, our system is also capable of audio and video tagging to assist users to supplement the automatically collected sensor data (as in [14]). Some WBAN components are depicted in Figure 2.

III. Activity Modeling

Data collected from our experimental system setup can be used in multiple contexts, for instance by the users to regularly monitor their physical well being as well as by medical practitioners in assessing the physical health of their patients. Here, we describe one such application of using the data to automatically derive the activity of a person with data collected from multiple sensors. Statistical modeling of various test subject states was undertaken based on the data collected from the WBAN. We examined five different states: lying down, sitting, standing, walking and running. Again, to reiterate, activity detection has been previously considered with an emphasis on the use of many accelerometers, yielding a cumbersome network to wear. We conjecture that multimodal data analysis will enable the achievement equal or even better accuracy and robustness in activity detection with fewer sensors.

[1] The data of the WMS GSRs are not currently included due to issues with time synchronization.

In this research, multiple distributions were considered to fit the data which for each sensor was predominantly unimodal in nature. After extensive experimentation, the use of the pulse oximeter sensor was abandoned due to limited change in readings for any of the states of interest for our activity detection problem. Thus, we focused on ECG and accelerometer data. The distributions under consideration were: T location-scale, Gaussian, log-normal, logistic, log-logistic, one-side Gaussian and Laplacian. Where possible, Gaussian distributions were selected to facilitate the determination of joint densities. The ECG data were preprocessed as follows: peak detection was performed and the inter-peak time collected. The inter-peak time was modeled as a Gaussian random variable. An average of the empirical variance for each of the axes over a prespecified window of time for the accelerometer data was employed. The walking and running state data were modeled as Gaussian; however, the lower-activity level data (lying down, sitting and standing) was modeled as a Laplacian to achieve a better fit. Figure 3 (L) and (R) shows the ECG and accelerometer data for the running and sitting modes, respectively. We see that both states are relatively well distinguished from each other with significant differences in the accelerometer data.

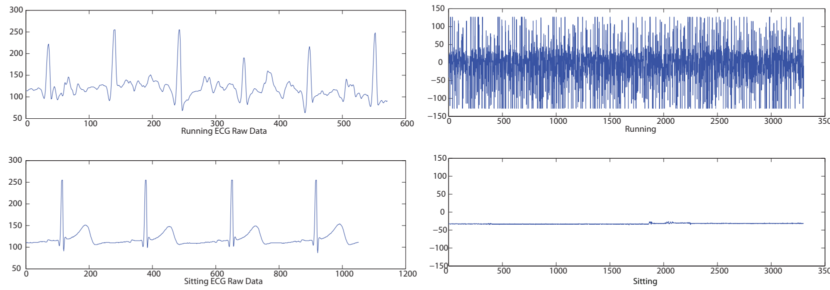

Fig. 3. (L) ECG and (R) accelerometer data from the heart-rate monitor for sitting and running.

Not surprisingly, ECG and accelerometer data had different discriminatory properties for the various states, underscoring the benefits of multi-modal sensing and signal processing. In Figure 4, we see the statistical fits for the accelerometer data for high activity states and the statistical fits for the ECG data for low activity fits. To develop bivariate models (joint densities) for the ECG and accelerometer data, additional processing (resampling) was required to determine the correlation between the ECG statistic and the accelerometer statistic in the high-activity levels.

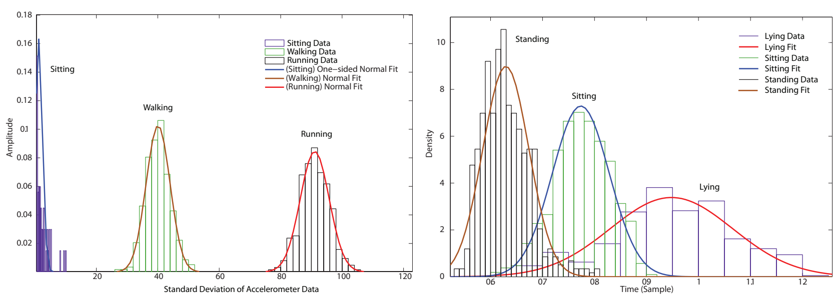

Fig. 4. (L) Statistical fitting for higher activity states (accelerometer data): sitting, walking, and running. (R) Statistical fitting for lower activity states (ECG data): lying down, sitting, and standing.

In the low-activity level cases, the ECG and accelerometer statistics were assumed to be independent. The resulting bivariate densities for each of the five hypotheses are shown in Figure 5(L) and (R). For clarity, the low activity states are shown separate from the higher activity states. Bivariate testing yielded state detection rates on the order of 85% to 95% – achieving detection rates with two heterogeneous sensors comparable to the rates found in [3], where nine single mode (accelerometer) sensors were employed.

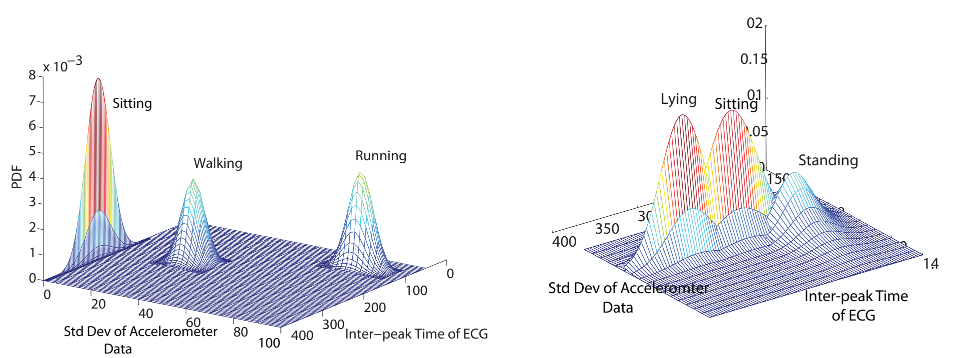

Fig. 5. Bivariate distributions for (L) running, walking and sitting and for (R) lying down, sitting and standing.

IV. Observations and Ongoing Work

Our preliminary system successfully collects data and transmits it to the cellular phone. We conjecture from our experiments that a few heterogeneous sensors may offer better discrimination and robustness than many homogeneous sensors. Our preliminary data for activity detection in comparison to [3] appears to bear this out this conjecture. There are however important engineering challenges associated with WBANs, especially for activity detection. For our particular set up, we are limited by the mobile phone platform which can only accommodate a maximum of eight different sensors. If all sensors sample at their maximum sampling rate, the expected throughput would exceed the capabilities of the Bluetooth link leading to dropped packets. The battery power of the cellular phone is another bottleneck for the system. Finally, for activity detection, high activity/mobility can impair a sensor's ability to sense. This fact can be viewed two ways: it is detrimental in that we lose sensor accuracy, on the other hand, new features are introduced into the signal which are still indicative of high activity. Our preliminary results suggest that sensor selection and prioritization will be important to ensure that packets are not lost; furthermore energy aware sensor management will be critical.

We have recently conducted a pilot study with two pre-adolescent girls following an observation protocol typical for pediatric obesity studies. We are currently analyzing this data, including designing multi-modal detection algorithms for deciding between the various states. We hope to share those findings at the workshop. Finally, introducing contextual cues for use of the WBAN in everyday life will be extremely important; to this end, the image processing and analysis methods of DietSense [14] will prove very useful. Finally, as noted earlier, power minimization is of high importance for WBANs and their attendant applications; we expect the methods of [4] will have promise when properly adapted to our context.

Acknoweldgements

The research is funded in part by Qualcomm. Additional support is from Nokia for the Nokia Smartphones used in our studies.

References

[1] Wireless Health Monitors from Alive Technology, www.alivetec.com, Retrieved on July 24, 2008.
[2] Sensewear BMS available at http://www.sensewear.com/solutions_bms.php
[3] S. Biswas and M. Quwaider. Body Posutre Identification using Hidden Markov Model with Wearable Sensor Networks. *Proc. of BodyNets Workshop 2008*, Tempe, AZ, March 2008.
[Sensys 2007:]
[4] A. Benbasat and J. Paradiso , A framework for the automated generation of power-efficient classifiers for embedded sensor nodes. *Proceedings of Sensys*, pp. 219-232, Novemberm 2007, Sydney, Australia.
[5] P. Bjorntorp, R. Rosmond. Neuroendocrine abnormalities in visceral obesity. *International Journal of Obesity & Related Metabolic Disorders: Journal of the International Association for the Study of Obesity* 2000;24 Suppl 2:S80-5.
[6] T. Gao, C. Pesto, L. Selavo, *et. al.*. Wireless Medical Sensor Networks in Emergency Response: Implementation and Pilot Results, *Proc. 2008 IEEE International Conference on Technologies for Homeland Security*, May, 2008.
[7] M. Goran, K. Reynolds, C. Lindquist. Role of physical activity in the prevention of obesity in children. *International Journal of Obesity & Related Metabolic Disorders: Journal of the International Association for the Study of Obesity*. 1999;23:S18-33.
[8] S. Jiang, Y. Cao, S. Iyengar, *et. al.*. CareNet: An Integrated Wireless Sensor Networking Environment for Remote Healthcare. *Proc. of BodyNets Workshop 2008*, Tempe, AZ, March 2008 (work in progress paper).
[9] E. Jovanov, A. Milenkovic, C. Otto and P. C. de Groen. A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation. *Journal of NeuroEngineering and Rehabilitation*, 2:6, March 2005.
[10] A. Kalpaxis. Wireless Temporal-Spatial Human Mobility Analysis Using Real-Time Three Dimensional Acceleration Data. *Proc. of International Multi-Conference on Computing in the Global Information Technology* , March 2007.
[11] S. Kimm, N. Glynn, E. Obarzanek, *et. al.* Relation between the changes in physical activity and body-mass index during adolescence: a multicentre longitudinal study. *The Lancet* 2005;366:301.
[12] D. Konstantas, A. Van Halteren, R. Bults, *et. al.* MobiHealth: Ambulant Patient Monitoring over Public Wireless Networks. *Proc. Mediterranean Conf. on Medical and Biological Engineering (MEDICON)*, August 2004, Ischia, Italy.
[13] C. Ogden, M. Carroll, and K. Flegal. High Body Mass Index for Age Among US Children and Adolescents, 2003-2006. JAMA 2008:299:2401.
[14] S. Reddy, A. Parker, J. Hyman, *et. al.* Image Browsing, Processing, and Clustering for Participatory Sensing: Lessons From a DietSense Prototype. *Proc. EmNets'07*, June 2007, Cork, Ireland.