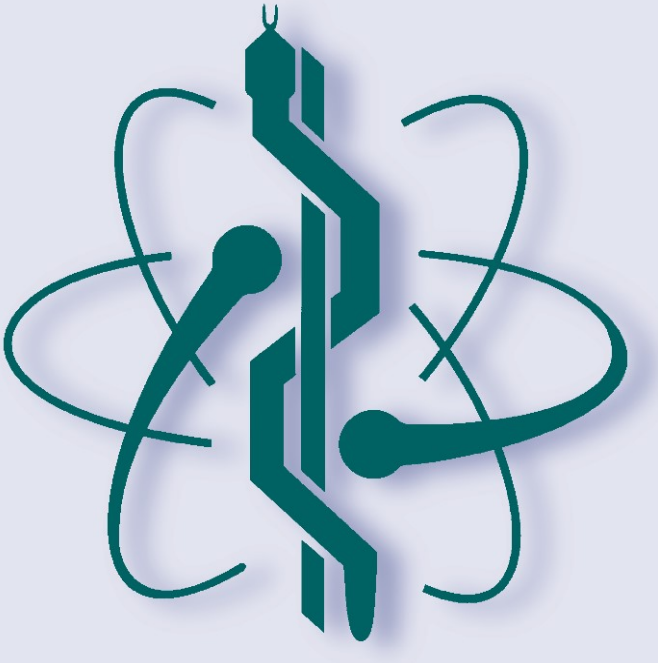
IFMBE Proceedings

S. Leonhardt · T. Falck · P. Mähönen (Eds.)

Volume 13

4th International Workshop
on Wearable and Implantable
Body Sensor Networks (BSN 2007)
March 26 – 28, 2007
RWTH Aachen University,
Germany



Springer

IFMBE Proceedings Volume 13

Editor: R. Magjarevic and J. H. Nagel

The International Federation for Medical and Biological Engineering, IFMBE, is a federation of national and transnational organizations representing internationally the interests of medical and biological engineering and sciences. The IFMBE is a non-profit organization fostering the creation, dissemination and application of medical and biological engineering knowledge and the management of technology for improved health and quality of life. Its activities include participation in the formulation of public policy and the dissemination of information through publications and forums. Within the field of medical, clinical, and biological engineering, IFMBE's aims are to encourage research and the application of knowledge, and to disseminate information and promote collaboration. The objectives of the IFMBE are scientific, technological, literary, and educational.

The IFMBE is a WHO accredited NGO covering the full range of biomedical and clinical engineering, healthcare, healthcare technology and management. It is representing through its 58 member societies some 120.000 professionals involved in the various issues of improved health and health care delivery.

IFMBE Officers
President: Makoto Kikuchi, Vice-President: Herbert Voigt, Past-President: Joachim H. Nagel,
Treasurer: Shankar M. Krishnan, Secretary-General: Ratko Magjarevic
http://www.ifmbe.org

Previous Editions:

IFMBE Proceedings ICBMEC 2005 "The 12th International Conference on Biomedical Engineering", Vol. 12, 2005, Singapore, CD

IFMBE Proceedings EMBEC'05 "3rd European Medical & Biological Engineering Conference, IFMBE European Conference on Biomedical Engineering", Vol. 11, 2005, Prague, Czech Republic, CD

IFMBE Proceedings ICCE 2005 "The 7th International Conference on Cellular Engineering", Vol. 10, 2005, Seoul, Korea, CD

IFMBE Proceedings NBC 2005 "13th Nordic Baltic Conference on Biomedical Engineering and Medical Physics", Vol. 9, 2005, Umeå, Sweden

IFMBE Proceedings APCMBE 2005 "6th Asian-Pacific Conference on Medical and Biological Engineering", Vol. 8, 2005, Tsukuba, Japan, CD

IFMBE Proceedings BIOMED 2004 "Kuala Lumpur International Conference on Biomedical Engineering", Vol. 7, 2004, Kuala Lumpur, Malaysia

IFMBE Proceedings MEDICON and HEALTH TELEMATICS 2004 "X Mediterranean Conference on Medical and Biological Engineering", Vol. 6, 2004, Ischia, Italy, CD

IFMBE Proceedings 3rd Latin – American Congress on Biomedical Engineering "III CLAEB 2004", Vol. 5, 2004, Joao Pessoa, Brazil, CD

IFMBE Proceedings WC2003 "World Congress on Medical Physics and Biomedical Engineering", Vol. 4, 2003, Sydney, Australia, CD

IFMBE Proceedings EMBEC'02 "2nd European Medical and Biological Engineering Conference", Vol. 3, Parts 1 & 2, 2002, H. Hutten and P. Kroesl (Eds.), Vienna, Austria

IFMBE Proceedings 12NBC "12th Nordic Baltic Conference on Biomedical Engineering and Medical Physics", Vol. 2, 2002, Stefan Sigurdsson (Ed.) Reykjavik, Iceland

IFMBE Proceedings MEDICON 2001 – "IX Mediterranean Conference on Medical Engineering and Computing", Vol. 1, Parts 1 & 2, 2001, R. Magjarevic, S. Tonkovic, V. Bilas, I. Lackovic (Eds.), Pula, Croatia

For ordering Proceedings from the IFMBE Proceedings Series, please contact the IFMBE Secretariat at: office@ifmbe.org.

IFMBE Proceedings Vol. 13
Steffen Leonhardt, Thomas Falck,
Petri Mähönen (Eds.)

4th International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2007)

March 26 – 28, 2007
RWTH Aachen University, Germany

Springer

Editors

Prof. Dr.-Ing. Dr. med. Steffen Leonhardt
Chair of Medical Information Technology
Helmholtz Institute of Biomedical Engineering
RWTH Aachen University
Pauwelstr. 20
52074 Aachen, Germany
E-Mail: leonhardt@hia.rwth-aachen.de

Prof. Dr. Petri Mähönen
Chair of Wireless Networks
RWTH Aachen University
Kackertstr. 9
52072 Aachen, Germany
E-Mail: pma@mobnets.rwth-aachen.de

Dipl.-Inform. Thomas Falck
Connectivity Systems Group
Philips Research Europe
Weisshausstr. 2
52066 Aachen, Germany
E-Mail: thomas.falck@philips.com

Library of Congress Control Number: 2007922364

ISSN: 1680-0737
ISBN-13: 978-3-540-70993-0 Springer Berlin Heidelberg New York

The IFMBE Proceedings is an Offical Publication of the International Federation for Medical and Biological Engineering (IFMBE)

Springer is a part of Springer Science+Business Media
springer.com

Typesetting: Data supplied by the authors
Production: Le-Tex Jelonek, Schmidt & Vöckler GbR
Cover design: deblik, Berlin

Printed on acid-free paper SPIN 11980759 7/3100/YL – 5 4 3 2 1 0

Introduction

About IFMBE

The International Federation for Medical and Biological Engineering (IFMBE) was established in 1959 to provide medical and biological engineering with a vehicle for international collaboration in research and practice of the profession. The Federation has a long history of encouraging and promoting international cooperation and collaboration in the use of science and engineering for improving health and quality of life.

The IFMBE is an organization with membership of national and transnational societies and an International Academy. At present there are 52 national members and 5 transnational members representing a total membership in excess of 120 000 worldwide. An observer category is provided to groups or organizations considering formal affiliation. Personal membership is possible for individuals living in countries without a member society The International Academy includes individuals who have been recognized by the IFMBE for their outstanding contributions to biomedical engineering.

Objectives

The objectives of the International Federation for Medical and Biological Engineering are scientific, technological, literary, and educational. Within the field of medical, clinical and biological engineering it's aims are to encourage research and the application of knowledge, and to disseminate information and promote collaboration.

In pursuit of these aims the Federation engages in the following activities: sponsorship of national and international meetings, publication of official journals, cooperation with other societies and organizations, appointment of commissions on special problems, awarding of prizes and distinctions, establishment of professional standards and ethics within the field, as well as other activities which in the opinion of the General Assembly or the Administrative Council would further the cause of medical, clinical or biological engineering. It promotes the formation of regional, national, international or specialized societies, groups or boards, the coordination of bibliographic or informational services and the improvement of standards in terminology, equipment, methods and safety practices, and the delivery of health care.

The Federation works to promote improved communication and understanding in the world community of engineering, medicine and biology.

Activities

Publications of IFMBE include: the journal *Medical and Biological Engineering and Computing*, the electronic magazine *IFMBE News*, and the Book Series on Biomedical Engineering. In cooperation with its international and regional conferences, IFMBE also publishes the IFMBE Proceedings Series. All publications of the IFMBE are published by Springer Verlag. The Federation has two divisions: Clinical Engineering and Health Care Technology Assessment.

Every three years the IFMBE holds a World Congress on Medical Physics and Biomedical Engineering, organized in cooperation with the IOMP and the IUPESM. In addition, annual, milestone and regional conferences are organized in different regions of the world, such as Asia Pacific, Europe, the Nordic-Baltic and Mediterranean regions, Africa and Latin America.

The administrative council of the IFMBE meets once a year and is the steering body for the IFMBE: The council is subject to the rulings of the General Assembly, which meets every three years.

Information on the activities of the IFMBE are found on its web site at: http://www.ifmbe.org.

BSN 2007 Conference Committees

Conference Chairs

Thomas Falck, Philips Research Europe, Germany
Steffen Leonhardt, RWTH Aachen University, Germany
Petri Mähönen, RWTH Aachen University, Germany

Program Committee

Paolo Bonato, Harvard Medical School, USA
Terry Dishongh, Intel, USA
Leonard Fass, GE Healthcare, UK
Thomas Gries, RWTH Aachen University, Germany
Bert Gyselinckx, IMEC, Belgium
Holger Karl, University of Paderborn, Germany
Malte Kelm, RWTH Aachen University Clinic, Germany
David Krebs, MIT, USA
Torsten Linz, Fraunhofer IZM, Germany
Nigel Lovell, University of New South Wales, Australia
Paul Lukowicz, University of Passau, Germany
Andreas Lymberis, European Union, Belgium
Winfried Mokwa, RWTH Aachen University, Germany
Wolfgang Niederlag, Krankenhaus Dresden Friedrichstadt, Germany
Joe Paradiso, MIT, USA
Adrian Perrig, Carnegie Mellon University, USA
Rolf Rossaint, RWTH Aachen University Clinic, Gemany
Ivo Salmärä, Microsoft, USA
Thomas Schmitz-Rode, RWTH Aachen University Clinic, Germany
Frank Siegemund, European Microsoft Innovation Center, Germany
Morris Sloman, Imperial College, UK
Wilhelm Stork, Karlsruhe University, Germany
Chris Toumazou, Imperial College, UK
Gerhard Tröster, ETH Zurich, Switzerland
Martin Westhofen, RWTH Aachen University Clinic, Germany
Hans-Jürgen Wildau, Biotronik, Germany
Paul Wright, University of California at Berkeley, USA
Guang-Zhong Yang, Imperial College, UK
Yuan-Ting Zhang, The Chinese University of Hong Kong, China

Additional Reviewers

Akram Alomainy, Oliver Amft, Louis Atallah, Heribert Baldus, Stacy Morris Bamberg, Fernando Brunetti, Javier Espina, Oscar Garcia Morchon, Daniel Garrison, Sye Loong Keoh, Jens Mühlsteff, Matthias Steffen, Olaf Such, Stephan Tobies, Lei Wang

Local Organizers

Lisa Beckmann, RWTH Aachen University, Germany
Margret Frey, Philips Research Europe, Germany
Evgeny Osipov, RWTH Aachen University, Germany
Janne Riihijärrvi, RWTH Aachen University, Germany
Matthias Steffen, RWTH Aachen University, Germany

Contents

Communication I

Smart Textiles

Posters

On-Body Sensors

Applications

Wearable Computing

Communication II

In-Body Sensors

1st Session
Communication I

Modelling and Characterisation of a Compact Sensor Antenna for Healthcare Applications

Akram Alomainy[1], Yang Hao[1] and Frank Pasveer[2]

[1]Department of Electronic Engineering, Queen Mary, University of London, London, United Kingdom
[2]Healthcare Devices and Instrumentation, Philips Research Europe, Eindhoven, The Netherlands

***Abstract*— The paper presents a planar compact antenna structure used in sensors aimed at healthcare applications. Antenna performance is numerically investigated with regards to impedance matching, radiation patterns, gain and efficiency. The compact size of the sensor causes the antenna to be susceptible to variable changes caused by the presence of lumped components. The study illustrated the importance of including full sensor details in determining and analysing the antenna performance. The body-worn sensor performance is also demonstrated and effects on antenna parameters are analysed, specifically radiated power, efficiency and front to back ratio of radiated energy. Radio propagation characterisation of the sensor operation in stand alone and on-body scenarios are introduced. Improvements are necessary in antenna design, matching circuitry and also sensor layout for better coverage and for obtaining maximum achievable communication range to produce efficient and reliable medical telemetry and monitoring systems.**

***Index Terms*— Healthcare sensor, compact antenna, wearable device, numerical modelling, efficiency.**

I. Introduction

Wireless body area networks (WBAN) provide promising applications in medical sensor systems and personal entertainment technologies [1]-[2]. They present the apparent option for efficient, flexible systems with constant availability, reconfigurability, unobtrusiveness and true extension of a human's mind. The idea of a body area network was initiated for medical purposes in order to keep continuous record of patient's health at all times. Sensors are placed on the human body to measure specified parameters and signals in the body, e.g. blood pressure, heart signals, sugar level, temperature, etc.

Antennas are essential part of WBAN and their complexity depends on the radio transceiver requirements and also on the propagation characteristics of the surrounding environments. For the conventional long to short wave radio communication, conventional antennas have proved to be more than sufficient to provide desired performance minimising the restrains on cost and production time. On the other hand, for todays and tomorrows communication devices, the antenna is required to perform more than one task or in other words the antenna needs to operate at different frequencies to account for the increasing new technologies and services available to the user, e.g. systems on pills [5] and miniaturised sensors [6]. For wireless sensor applications, antennas need to be efficient and immuned from frequency and polarisation detuning. Understanding the antenna radiation pattern for wireless sensors, specifically when applied for body-worn applications, is vital in determining the sensor performance. It is also important to specify how coupling into the propagation mode which may be a surface wave or free space wave or a combination of both occurs.

This paper presents sensor designs and modules developed by the group Healthcare Devices and Instrumentation, Philips Research, for operation in the unlicensed ISM band (2.4 GHz). Maximum achievable coverage range is to be delivered by the sensor with respect to the transceiver sensitivity levels. Modelling and characterisation of the antenna deployed in the full sensor are discussed and investigated with regards to surrounding components and data connectors (full sensor details are included). Antenna enhancement techniques are demonstrated to improve impedance matching and hence antenna efficiency. Radio channel characterisation of propagation from a wearable sensor is also presented and cross-referenced with numerical simulation results for validation.

II. Antenna Requirements

The sensor antenna design is restricted by many factors including the sensor size, chips placement, lumped element locations and flexibility of the sensor structure to be shuffled with minimum cost and changes to antenna performance. Figure 1 shows photographs of the sensor transceiver layer and the prototype module fabricated. A schematic design of the modelled sensor antenna and also an exploded diagram of the proposed sensor is illustrated in Fig. 2.

The current antenna deployed in the sensor design is a printed quarter wavelength monopole, etched on the edge of the circular PCB board. Hence the antenna is designed with the printed wire wrapping the transceiver chip and other components. The antenna is derived from the circumference monopole, which is derived from the bent and inverted L antenna [7]. The sensor antenna performance is sensitive to lumped components, pins and copper routings presence. The surrounding and adjacent components are modelled as a perfect conductor block around which the antenna is printed. The PCB board includes ground plane and supply voltage copper sheets.

When a monopole is placed or printed on a dielectric material with permittivity other than 1, the antenna dimensions have to be modified in order to achieve performance at the frequency of interest. This leads to redefinition of antenna impedance by the approximation,

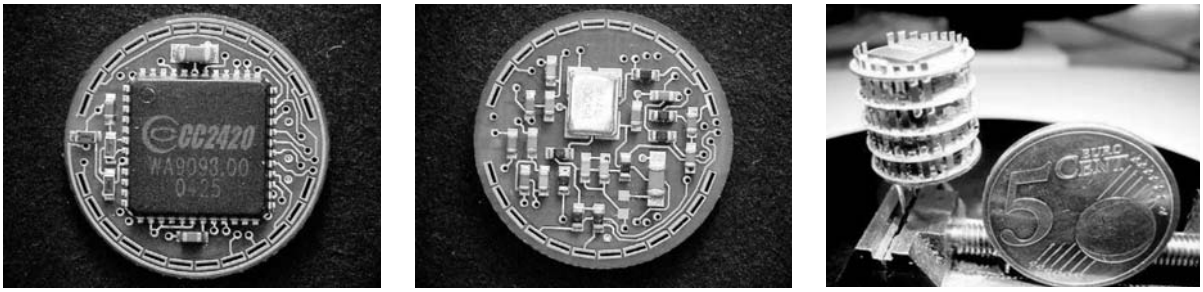


Fig. 1. Photographs of (a) top view of the transceiver layer, (b) bottom view of the transceiver layer and (c) the manufactured prototype sensor

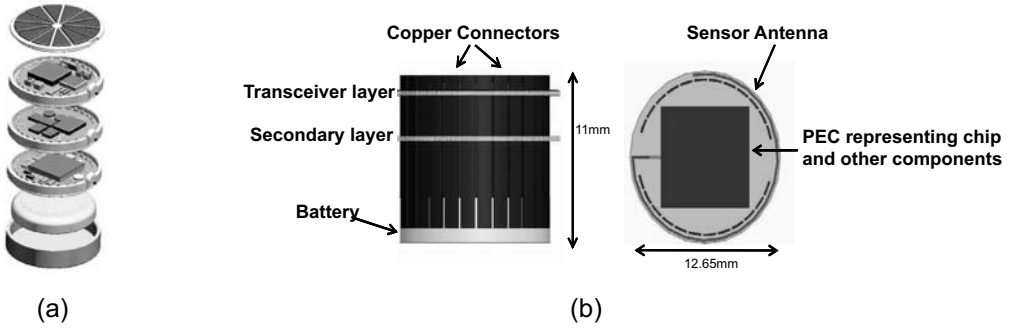


Fig. 2. (a) Exploded diagram of the proposed healthcare sensor and (b) dimensions and detailed structure of the modelled sensor including the antenna.

$$Z\left(\omega, \epsilon_{eff}\right) = \frac{1}{\sqrt{\epsilon_{eff}}} Z\left(\sqrt{\epsilon_{eff}}\omega, \epsilon_0\right) \tag{1}$$

where ϵ_{eff} is the effective permittivity and ϵ_0 is the vacuum permittivity. This leads to the antenna length being shortened by a factor of $1/\sqrt{\epsilon_r}$ to ensure the resonance is at the desired frequency. The metal pins and copper lines provide an extension to the antenna which increases the impedance value expected, however, due to the capacitive coupling between the antenna and surrounding components the antenna impedance tends to decrease comparing to the general monopole impedance of 36 Ω. Therefore, the radiation characteristics of the antenna are directly affected; hence changes in antenna gain and efficiency are expected.

III. Numerical Analysis of Antenna Performance

A. Initial Design

The antenna design deployed in the proposed sensor is numerically analysed using the Finite Element Method (FEM) utilised in the High Frequency Structure Simulator (HFSS), Ansoft™. The printed circular monopole antenna is modelled on FR4 substrate (ϵ_r=4.6 and thickness of 0.3mm). The printed antenna thickness is $35\mu m$ and the width of the line is $150\mu m$. The ground and supply voltage layers added have a diameter of 5.5mm, thickness of $17.5\mu m$ each and separation between the layers of $80\mu m$. The actual antenna length is 31.5mm (approximately quarter wavelength of the required frequency, 2.4GHz). The complex impedance at the RF transceiver differential output is 115+j180 , therefore a matching circuit is applied in order to match the output to the single-ended monopole (matching to 50Ω) [10].

Figure 3 presents the return loss of the sensor antenna when only one layer is modelled in comparison to full sensor modelling. The one layer model includes the printed antenna, the transceiver chip and PCB board. The figure illustrates the significance of considering full structure modelling in characterising small antenna designs integrated with radio systems. The return loss of the antenna demonstrates the effect of the connectors on reducing the resonance frequency due to increase in electric length of the antenna caused by connectors current distribution. The calculated antenna gain is -1.2 dB with radiation efficiency of 48%.

B. Performance Enhancement

Investigating the impedance of the current sensor antenna, the real part of the antenna impedance is calculated as 11Ω. In general monopole antennas have impedance of 36Ω; however, this impedance is dependent on shape and size of the antenna. The bend shown in current antenna design causes the inductive reacting part to increase in comparison to capacitive reactance and reduces the real impedance part. Therefore, one solution is to modify the current matching circuit to match complex transceiver impedance $115 + j180\Omega$ to 11Ω. Figure 3 shows the return loss of one layer and full sensor antenna models when a modified matching circuit is applied and the improved impedance matching is apparent for both cases at

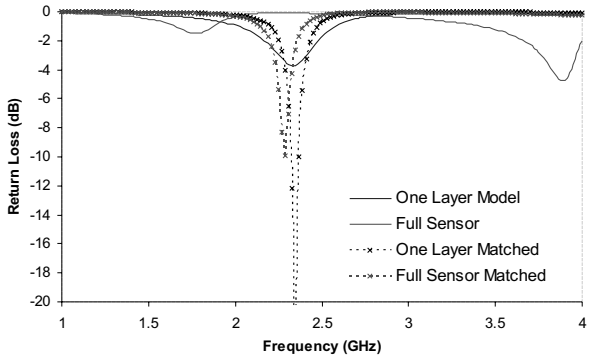


Fig. 3. Comparison of sensor antenna return loss for one layer and full sensor models. Comparison between original antenna matching and modified matching circuits.

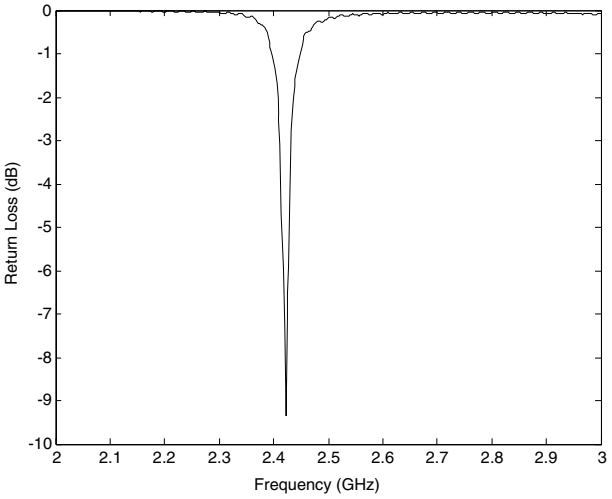


Fig. 4. Simulated return loss of sensor antenna in free space with resonance around the desired frequency.

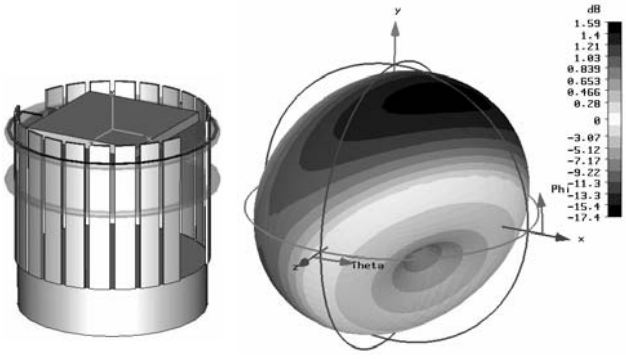


Fig. 5. Sensor antenna radiation pattern numerically calculated at 2.4GHz.

the frequency of interest.

Initial analysis of the monopole impedance at 2.4GHz demonstrated the effect of adjacent components on antenna performance and hence antenna impedance. The antenna has a calculated complex impedance of $35 + j320\Omega$ at 2.4GHz. Another solution of the impedance matching problem is to introduce A 0.2pF capacitor at the antenna input to match the antenna to the 50Ω impedance seen at the current matching circuit output (recommended by transceiver data sheet [10]). Figure 4 presents the sensor antenna return loss and the resulting narrowband response caused by capacitor addition. The 3D antenna radiation pattern at 2.4GHz is shown in Fig. 5. The pattern is different from that of a conventional vertical monopole due to introduced bend and also the effect of surrounding elements. The antenna gain calculated numerically is around 1.6dB with efficiency of 77% which demonstrates improvements in both impedance matching and antenna radiation and hence total antenna efficiency in comparison to the unmatched antenna design discussed above. This illustartes the potential extended coverage area served by the sensor with simple and reliable performance enhancement techniques. The modelled sensor structure includes two layers of modules in comparison to four illustrated in the prototype photograph in Fig. 1. Two layers proved to be sufficient enough for obtaining reliable data.

IV. Body-Worn Sensor Modelling

Body-worn sensor antenna performance is numerically investigated using the finite integral technique (FIT) utilised in CST Microwave Studio™. The sensor is placed on the human chest, as shown in Fig. 6, with the antenna radiating element normal to the body. The used model is the commonly available detailed multi-layer human model, namely the visible male model developed by the US air force (http://www.brooks.af.mil/AFRL/HED/hedr/). The simulated return loss of the sensor shows the slight detuning due to presence of lossy tissues with sensor placed 2mm away from the body, Fig. 7. Although the detuning effect is minimum, the narrowband nature of the antenna (with capacitor used for matching) signifies such changes and its direct influence on radiated energy.

The antenna gain when placed on the body is increased to 2.4dB caused by reflections from the human body, which at this high frequency is considered as a large reflector due to high losses and also very small penetration depth. The azimuth plane radiation patterns shown in Fig. 8 illustrate the effect of human body on the antenna performance and the reduced radiated power in the backward direction with front to back ration of around 25-30 dB. This is clearly demonstrated in Fig. 9, which presents the electric field distribution around the human body at 2.4 GHz in both vertical and horizontal planes. The field distribution defines clearly the creeping waves created by diffraction of propagating waves around the body curvature and also the power absorption with regards to front to back power ratio.

V. Radio Propagation Characteristics of the Sensor Antenna

The spatial performance of the sensor antenna is anaylsed experimentally in free space (anechoic chamber measurement).

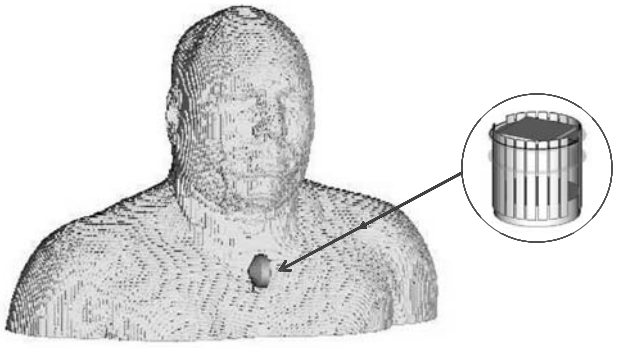


Fig. 6. Sensor placed on the male model provided by the visible US human project.

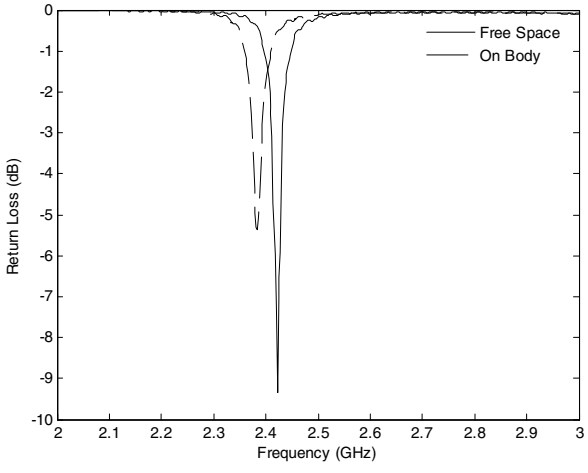


Fig. 7. Return loss for capacitor matched sensor antenna with resonance at 2.4 GHz when place in free space and on the body.

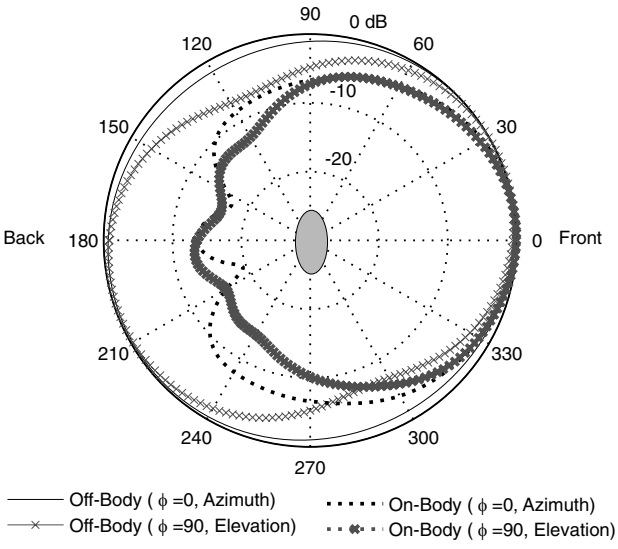


Fig. 8. Radiation pattern of sensor antenna when placed in free space and on the body in both aziuth and elevation planes (Fig. 5).

The sensor is placed on a turn-table with angular variations of 0^0 to 360^0. The sensor is used as transmitting unit and a microstrip patch antenna connected to a spectrum analyser acts as a receiving node. The patch antenna is placed 88 cm away from the transmitting sensor. The sensor is placed on the turn-table in horizontal and vertical orientations to investigate polarisation effect on propagation channel. Similar setting is repeated when the sensor is worn by the user (similar position presented in Fig. 6). The spectrum analyser received spectrum demonstrated a picked signal at 2.405GHz, which is covered by both sensor and patch antennas.

Figure 10 shows measured co-polar and cross-polar patterns of the sensor antenna radiation performance in free space and also when placed on the body with the antenna normal to

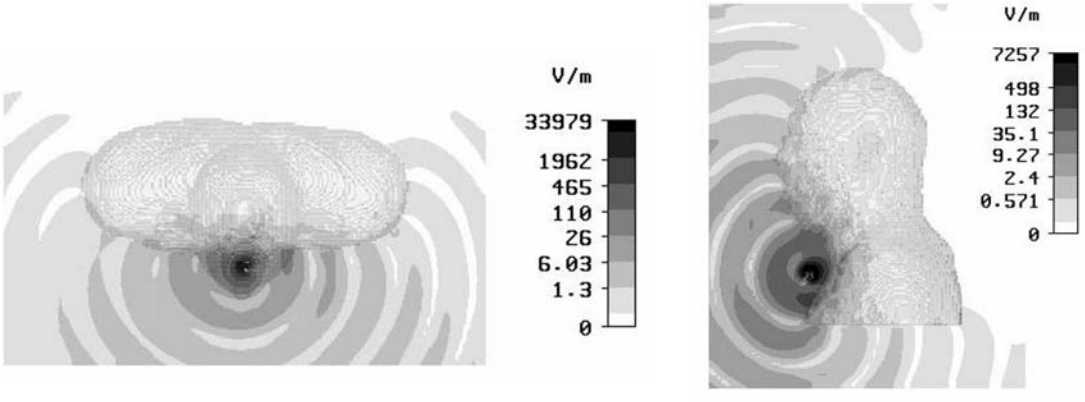


Fig. 9. Electric field distribution around the human body induced by the sensor antenna in (a) Horizontal (top view) and (b) Vertical (side view).

the body. When the body shadows the communication link between Tx-Rx nodes, at around 180^0 the loss due to the body shadowing is around 30dB, which agrees with results aforementioned from numerical analysis. The angular patterns in Fig. 9 present reasonable omni-directional behaviour of the sensor antenna with maximum variation of 8-10dB for free space cases (off-body).

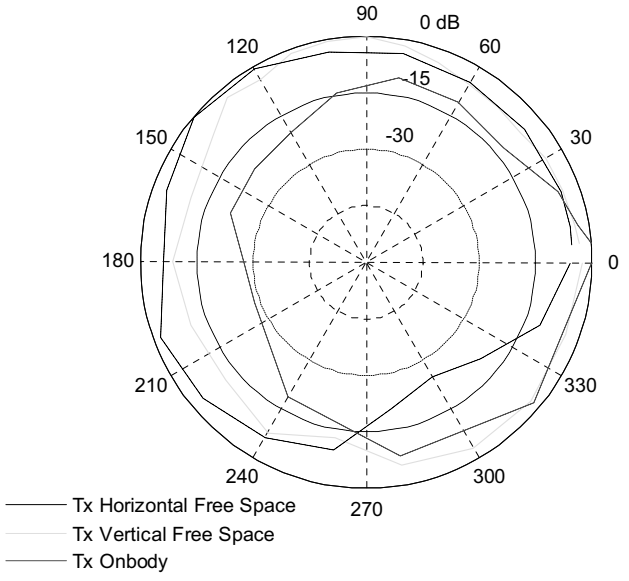


Fig. 10. Received power pattern when Tx (Sensor) is placed 88cm from a receiving patch antenna for horizontal and vertical sensor placements.

Following the set-up described above, path loss analysis of the radio channel between the transmitting sensor and the patch for cases where the sensor is placed in free space and on the body in the anechoic chamber and in indoor environment is performed. Figure 11 demonstrates the path loss measured in indoor environment. As predicted from Frii's path loss formula, the exponent is lower than that of free space with a value of 1.3 when the sensor is placed on the body due to multipath components from the different scatterers. For similar distances the loss is higher for NLOS cases. The directivity of the antenna increases when the antenna is placed on the body due to high losses at 2.4GHz of the human tissues which leads to an increase in the received power for the same distances applied in the stand alone sensor scenario.

VI. Conclusion

A study of a compact antenna used in sensors aimed at healthcare applications in the ISM band (2.4 GHz) was presented. The antenna performance was investigated numerically considering the effect of full sensor structure in comparison to antenna printed on PCB board model. The investigation demonstrated the significance of considering the detailed sensor design in analysing the antenna and radio propagation performance due to the compact size of the proposed sensor. The compact size of the sensor and the careful placement of components surrounding the antenna

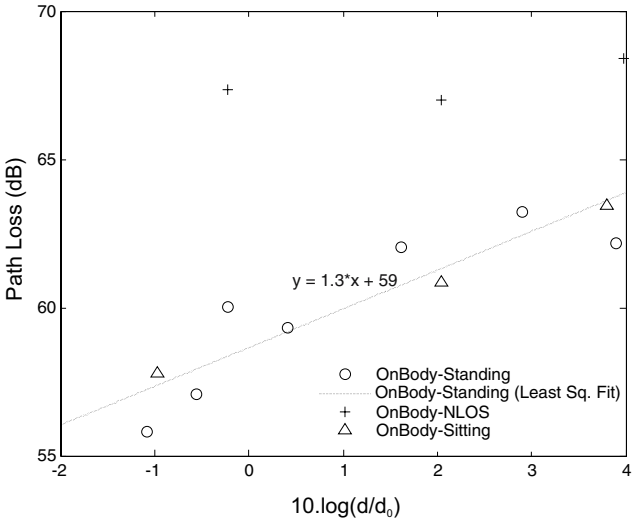


Fig. 11. Indoor measured path loss when sensor placed off and on body with modelled path loss using the least fit square technique (line related body standing skip off body).

introduced many challenges in determining and improving the antenna performance with regards to impedance matching, gain, efficiency and front to back ratio of radiated energy. The antenna performance evaluation and radio propagation characterisation provided promising indications of potential optimum performance sensor designs.

References

[1] J. Bernardhard, P. Nagel, J. Hupp, W. Strauss, and T. von der grun, "BAN-Body area network for wearable computing", 9th Wireless World Research Forum Meeting, Zurich, Jul, 2003.

[2] E. Jovanov, A. ODonnell-Lords, D. Raskovic, P. Cox, R. Adhami, and F. Andrasik, "Stress monitoring using a distributed wireless intelligent sensor system", IEEE Engineering in Medicine and Biology Magazine, vol. 22, issue 3, pp. 49-55, May/Jun, 2003.

[3] C. Kunze, U. Grossmann, W. Stork, and K. Muller-Glaser, "Application of ubiquitous computing in personal health monitoring systems", in Biomedizinische Technik: 36th Annual meeting of the German Society for Biomedical Engineering, 2002, pp. 360-362.

[4] N.F. Timmons, W.G. Scanlon, "Analysis of the performance of IEEE 802.15.4 for medical sensor body area networking", 2004 First Annual IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON), 4-7 Oct, 2004, pp: 16-24.

[5] Erik A. Johannessen, Lei Wang, Cathy Wyse, David R. S. Cumming,and Jon M. Cooper, "Biocompatibility of a Lab-on-a-Pill Sensor in Artificial Gastrointestinal Environments", IEEE Transactions on Biomedical Engineering, volume 53, issue 11, Nov. 2006, pp:2333-2340.

[6] Jani Miettinen, Matti Mantysalo, Kimmo Kaija and Eero 0. Ristolainen, "System design issues for 3D system-in-package (SiP)", 54th Electronic Components and Technology Conference 2004 Proceedings, Volume 1, 1-4 June 2004, pp:610-615.

[7] Jen-Yea Jan, Liang-Chih Tseng, Wen-Shyang Chen, Yuan-Tung Cheng, "Printed monopole antennas stacked with a shorted parasitic wire for Bluetooth and WLAN applications", IEEE 2004 Antennas and Propagation Society International Symposium, Vol. 3, June 2004 pp: 2607 - 2610.

[8] Akram Alomainy, Yang Hao, Abdus Owadally, Clive G. Parini, Yuri Nechayev, Peter S. Hall and Costas C. Constantinou, "Statistical Analysis and Performance Evaluation for On-Body Radio Propagation with Microstrip Patch Antennas", IEEE Transactions on Antennas and Propagation, January 2007 (in press).
[9] P. S. Hall, Y. Hao, "Antennas and Propagation for Body-Centric Wireless Networks", Artech House Publisher, September 2006.
[10] Chipcon CC2420 transceiver chip, 2.4 GHz IEEE 802.15.4 / ZigBee-ready RF Transceiver, URL: http://www.chipcon.com

Address of the corresponding author:

Dr. Yang Hao
Electronic Engineering
Queen Mary, University of London
Mile End Road
London E1 4NS
United Kingdom
Tel: +44 (0) 2078825341
E-mail: y.hao@elec.qmul.ac.uk

Performance Analysis of 802.15.4 and 802.11e for Body Sensor Network Applications

Dave Cavalcanti, Ruediger Schmitt and Amjad Soomro

Philips Research North America, Wireless Communications and Networking Department, Briarcliff Manor, NY, USA

Abstract— **This paper studies the energy efficiency and QoS performance of 802.15.4 and 802.11e MAC protocols for body sensor network applications. We simulated a stand-alone body sensor network, as well as co-existence scenarios where the body sensors operate in the presence of voice, video and IT traffic. Our results indicate that although 802.15.4 and 802.11e can provide an acceptable compromise between power consumption and QoS in some scenarios, there are situations (e.g. co-existence with video and heavy data traffic) in which both performance criteria can not be met simultaneously. This highlights the need for improving existing MAC protocols or designing new solutions that can provide both extremely low power and QoS for body sensor networks (BSNs).**

Keywords— **BSN, Energy efficiency, QoS, 802.15.4, 802.11**

I. Introduction

Recently, there has been increasing interest in Body Sensor Networks (BSN) formed by low-power wireless devices placed in, on or around the body. BSNs enable new healthcare and wellness applications, such as long-term monitoring of physiological signs and control of medical implants. The two fundamental design challenges in BSNs are energy efficiency and Quality of Service (QoS). Energy efficiency is key to extend the life-time of battery-powered sensors, reduce maintenance costs and avoid invasive procedures to replace battery in the case of implantable devices. In particular, implantable sensors will require power consumption as low as tens of μW. On the other hand, QoS provisioning is fundamental to achieve reliability and timely data delivery, which are strict requirements for healthcare applications, such as patient monitoring.

Co-existence with other wireless networks is another important requirement for BSNs. A typical usage scenario could be a user wearing sensors to monitor physiological signs (e.g. ECG, blood pressure, glucose level, etc…) while talking on a Voice over IP (VoIP) WiFi phone or downloading a video over the home Wireless LAN (WLAN). In order to meet the requirements of BSNs, a wireless technology should be able to achieve extremely low power consumption with guaranteed QoS not only in controlled or well planed environments, but also in co-existence scenarios with other applications and wireless networks.

Among the several wireless standards available today, Zigbee [1] with the IEEE 802.15.4 MAC/PHY [2] has been considered as the technology of choice for several BSN applications. Although 802.15.4 consumes very low power, the figures may not reach the levels required by implantable sensors. Even for less critical applications, such as physiological monitoring with external sensors, 802.15.4 has several QoS issues [3]-[6]. For instance, it is well known that 802.15.4 devices operating in 2.4 GHz band suffer serious performance degradation due to interference from IEEE 802.11 WLANs [7] operating on the same band. Although the UWB-based 802.15.4a mode is expected to provide higher rates and additional robustness, it is still under standardization and it is not expected to replace the wide range of Zigbee/802.15.4 products operating on the 2.4 GHz band. Therefore, co-existence between 802.15.4 and other technologies is still an issue for BSN healthcare applications.

On the other hand, 802.11 has been considered to consume too much power to be useful for BSNs. However, advanced IC designs have lead to several 802.11 low power solutions [8]. In addition, the power-save and QoS enhancements introduced by 802.11e [9] can also be exploited to improve energy efficiency. These enhancements to the 802.11 protocol may bring about new opportunities. However, it is unclear whether low-power 802.11e solutions can meet low power and QoS required by BSN applications.

Clearly, there are issues associated with 802.15.4 as well as with 802.11 that lead to the question of whether these technologies are appropriate for BSNs. In this paper, we address this question through a simulation-based analysis of the energy efficiency and QoS of both 802.15.4 and 802.11e for BSNs. Although other technologies are being considered for BSN applications, such as Bluetooth [10], in this paper, we focus on 802.15.4 and 802.11e only. In particular, we consider the 802.11e Enhanced Distributed Channel Access (EDCA) with the 11b PHY mode operating at 11 Mbps, and the non-slotted version of the 802.15.4 MAC operating with the 250 Kbps PHY mode in the 2.4 GHz band. Moreover, we define two generic BSN applications involving a body sensor and a gateway node, namely: *i*) a wave-form real-time stream; and *ii*) a real-time parameter measurement stream. Then, we consider four networking scenarios where a BSN is expected to operate in real-life, and analyze,

through simulations, whether 802.15.4 and 802.11e would meet the energy efficiency and QoS requirements of the BSN in each scenario. Firstly, we consider a stand-alone BSN, which provides insight on the best performance that could be achieved with both 802.15.4 and 802.11e. Then, we consider scenarios in which the BSN operates in the presence of WLAN interference that includes IT (FTP, e-mail, http and telnet), VoIP and video. Our results indicate that although 802.15.4 is more energy efficient under light interference, it has high packet drop rate under video and heavy IT interference. In interference scenarios, 802.11e with AC_VO category is an alternative to achieve better QoS, but at the expense of higher power consumption. Therefore, neither solution could provide extremely low power consumption while ensuring QoS under all co-existence scenarios. This highlights the need for improving existing MAC protocols or designing new energy efficient solutions that support QoS.

II. Background

A. BSN Applications

Healthcare and wellness monitoring through sensors attached to the body are driver applications for BSNs. One possible application shown in Fig. 1 includes two sensors wirelessly connected to a bedside hub to enable continuous monitoring of a patient. The other application in Fig. 1 is a wellness/fitness application including two sensors and a portable gateway that could provide connectivity to a back-end monitoring service or display the ECG and other fitness related data (e.g. speed, distance, calories, etc.) to the user.

The sensors in Fig. 1 are few examples of a large number of bio-sensors that could be used in healthcare and wellness applications. In fact, the development of bio-sensors for different body signs and conditions constitutes another important research area. Also, future applications would include implanted sensors to monitor physiological parameters and/or control implanted medical devices. Since it is not feasible to study a system with every possible type of bio-sensor available, without lack of generality, we define two generic models that we believe to accuratly represent a wide range of healthcare and wellness applications:

- *Wave-form Sensor*: generates a data stream of one 185 bytes packet at every 0.25 seconds to model a real-time wave-form between the sensor and a gateway device.
- *Parameter Sensor*: generates one packet (37 bytes) per second that carries real-time information about one or more physiological parameters.

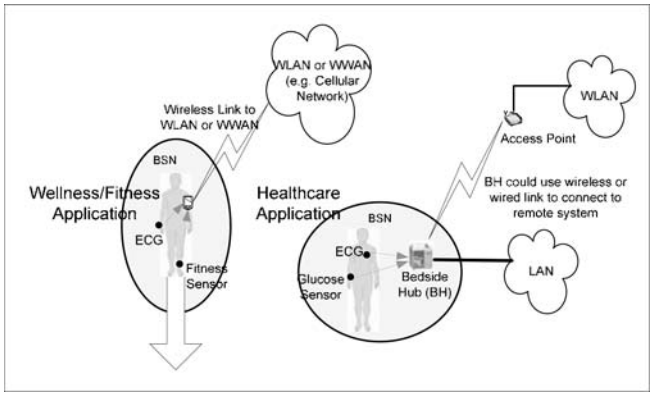


Fig. 1 Healthcare and Wellness applications

We consider a BSN formed by a wave-form sensor, a parameter sensor and a gateway node, and we study its performance in different scenarios as described next.

B. Networking Scenarios and Requirements

Firstly, we consider a stand-alone BSN as depicted in Fig. 2(a), which is a best-case situation for the BSN, since there is no other traffic contending for the channel. However, in real-life, BSNs will most likely operate in overlapping areas with other wireless networks. For example, 802.11 WLANs are being widely deployed in hospitals and homes, where BSNs may also be used as well. Therefore, it is fundamental to analyze how BSNs technologies operate in co-existence scenarios with WLANs. For that, we have selected the three scenarios depicted in Fig. 2, in which the BSN co-exists with VoIP, IT and Video applications. The motivation and details about each scenario are as follows:

- *Stand-alone BSN* (Fig. 2a): The BSN operates without any interference. The dashed arrows in Fig. 2 indicate the direction of the data flow, which is from the sensors to the gateway. In our simulations, the sensors start generating data between 1 and 2 seconds and remain active until the end of the simulation. This scenario provides valuable information on the best performance that could be achieved with a given technology for the BSN.
- *BSN + VoIP* (Fig. 2b): this is a typical example of a user wearing monitoring sensors while talking on a VoIP/WiFi phone. The VoIP phone is placed at 1 m from both sensors, and we use the GSM-quality traffic model available in OPNET for the voice call, which randomly starts between [10, 15] secs and runs until the end of the simulation.
- *BSN + VoIP + IT* (Fig. 2c): in this scenario we add an IT interferer device to the BSN and VoIP/WiFi phone. This situation could happen in a hospital, where the BSN could

be used to monitor a patient while a clinician uses a VoIP/WiFi phone near the bed and downloads patient's information (e.g. MRI, Xray, Ultrasound images, etc.). This could also be a typical home scenario, where the BSN's user could be talking on the VoIP/WiFi phone, while other nearby person uses a laptop to browse the Web or to read e-mail. Therefore, to cover these realistic situations, we model the IT traffic as a set of applications including FTP, http, e-mail and telnet. Most of the load is due to FTP, which includes uploads and downloads of 5 MB files, with inter-request time exponentially distributed with 10 sec average. The IT traffic starts at 1 second and runs until the end of the simulation.

- *BSN + Video* (Fig. 2d): in this case, we consider a video device placed at 1 m from the sensors. This could be a typical scenario at home or in the hospital, where the BSN's user could also be downloading a video through the home/hospital WLAN infrastructure. We use a standard OPNET VCR quality video model, which generates a downlink traffic load of ≈ 6 Mbps starting between [10, 15] sec until the end of the simulation.

In particular, for the three co-existence scenarios, we consider the worst-case situations in which the BSN and the interfering traffic operate on the same channel (when both use 802.11) or in overlapping channels (when the BSN uses 802.15.4). Obviously, if the BSN and the WLAN are configured to operate in non-interfering channels, the impact of the WLAN traffic on the BSN's performance would be minimal or even non-existent depending on the channels selected. However, a careful frequency planning may not be always possible in practice. Therefore, the performance in a worst-case situation gives an insight on the feasibility of a given wireless technology for BSNs.

Given the envisioned applications and usage scenarios, the main BSNs' requirements can be summarized as follows:

1. *Extremely low power consumption*: the power consumption could vary from tens of μW (for implantable sensors) to few mWs (for external body sensors).
2. *QoS support*: extremely low or possibly no packet loss is desirable. Typical delay bounds for physiological vital sign monitoring can be 250 msec or higher, which makes packet loss ratio the main QoS parameter.
3. *Co-existence*: the above power consumption and QoS requirements should be satisfied in co-existence scenarios with other applications and wireless networks as shown in Fig. 2.

Our objective is to determine whether 802.15.4 and 802.11e satisfy the BSN's requirements.

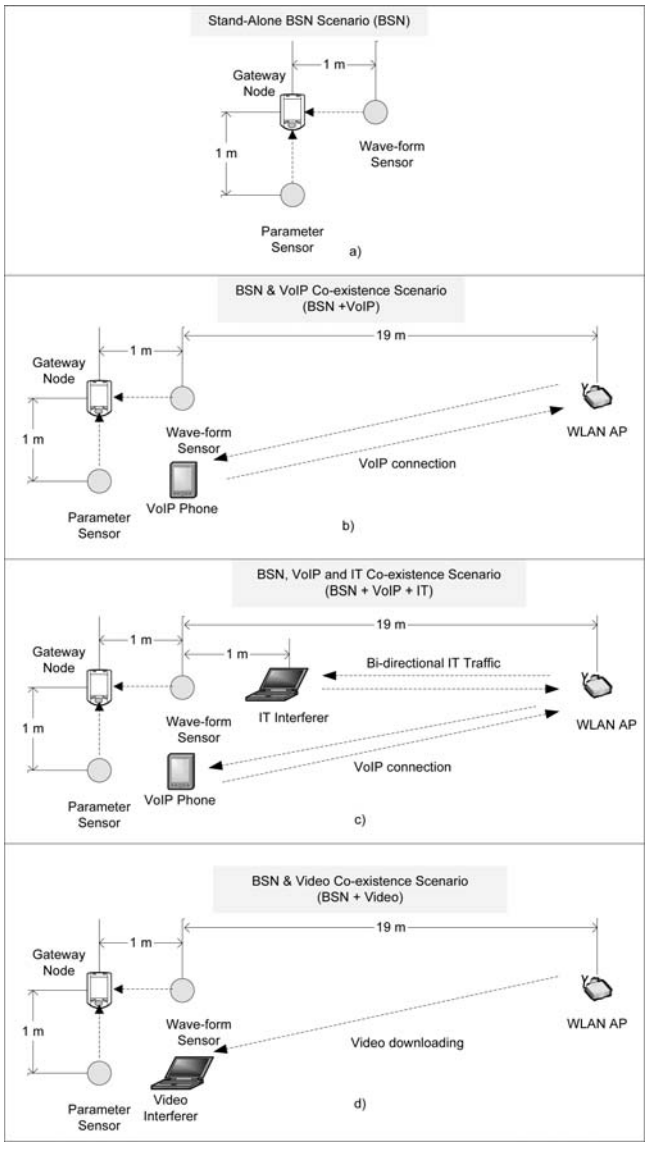


Fig. 2 Networking Scenarios

C. Overview of 802.15.4 MAC

The IEEE 802.15.4 standard [2] defines two MAC layer alternatives, a slotted and un-slotted version both based on CSMA/CA. We consider only the unslotted CSMA/CA MAC due to its simplicity, which allows the devices to save power, which is important in BSNs. In this case, a device that has data to transmit, initially defers for a backoff period randomly selected in the range 0 to 2^{BE}-1, where *BE* is the backoff exponent. After a bakcoff period, the device performs a Clear Channel Assessment (CCA) procedure that indicates whether the channel is busy or idle. We assume the CCA mode in which the medium is considered busy

upon detection of any energy above the energy detection threshold [2]. Then, if CCA indicates medium is idle, the device transmits the data, otherwise, it increments the number of backoffs performed (*NB*) as well as the *BE* (up to a maximum value *aMaxBE*). If *NB* is smaller than the maximum number of backoffs allowed (*MaxCSMABackoff*), the device performs another backoff before trying to access the channel again; otherwise, it indicates a failure in the transmission, which causes the frame to be dropped. We also enable acknowledgements, since reliability is an important requirement for BSNs.

D. Overview of 802.11e EDCA

We consider the EDCA mechanism specified in the 802.11e amendment [9], which defines up to four backfoff entities (queues) that can operate independently in a single station. Each backoff entity is associated with an access category and contends for the channel with different priorities. The four distinct access categories are: voice (AC_VO), video (AC_VI), best-effort (AC_BE) and background (AC_BK) traffic. The MAC layer parameters that are adjusted to define the access categories are: *i*) arbitration interframe space (*AIFS*), which is the interval of time the medium has to be idle before a backoff entity initiates a frame transmission or starts down counting its backoff-counter; *ii*) minimal and maximal contention window (CW_{min} and CW_{max}), which determines the initial and maximum range of random backoff counter, respectively; and *iii*) transmission opportunity limit (TXOP limit), which is the duration during which the TXOP holder maintains uninterrupted control of the medium [9].

E. Related Work

Co-existence issues between 802.15.4 and 802.11 have been studied in [3]-[6]. Results in [3], [4] and [5] suggest that the packet loss in an 802.15.4 link could be as high as 90% when a WLAN interferer is operating in close proximity and using an overlapping channel. But these works focus on QoS aspects, without analyzing the power consumption, which is key factor for BSNs.

In [11], the authors study the power consumption of a medical 802.15.4 BSN consisting of an external gateway and up to 10 implatable sensors. According to results in [11] 802.15.4 could be used for the implantable devices in its unslotted mode with a power save mechanism. Beacon frames in the slotted CSMA/CA protocol imposes an extra-burden in terms of power consumption, therefore reducing the sensors' lifetimes. However, no interference traffic or co-existence scenarios were considered in [11].

III. Performance in a stand-alone BSN

First, we simulate the stand-alone BSN depicted in Fig. 2(a) with three different choices of medium access mechanisms: *i*) 802.15.4 (un-slotted CSMA/CA), *ii*) 802.11e AC_BE with default parameters [9]; and *iii*) 802.11e AC_VO with default parameters [9]. In addition, we consider a simple power-save operation in which the sensors switch to power-save mode when there is no data to transmit and they stay in this mode until a new data packet is generated. We apply this scheme to 802.15.4 and 802.11e, since both standards support power-save mode. In this way, the sensors will maximize their energy savings, and the energy efficiency will only depend on the MAC layer protocol and characteristics of the chipset used.

We developed a customized 802.11e OPNET simulation model and integrated it with the NIST ZigBee/802.15.4 OPNET model [12], which includes detailed MAC and PHY characteristics. We considered the 802.11b mode operating at 11 Mbps on the channel 1 (centered at 2412 MHz) with the power consumption characteristics given in Table 1. The transmit power is 15 dBm. For 802.15.4, we assume the non-slotted CSMA/CA MAC operating at 250 Kbps on a channel centered at 2412 MHz, which overlaps with the 802.11 channel 1, but in this scenario the two technologies are not used simultaneously, and hence there is no interference. The power consumption characteristics for 802.15.4 are the same as in [11] and are given in Table 1. The physical layer is modeled as in [4], but the details are omitted here due to space constraints. Moreover, we do not consider the energy consumption in the bio-sensing process, since the main objective here is to study only the communication aspects. All simulations run for 300 secs and all results are averaged over 10 runs with different seeds.

Table 2 shows the average packet delivery ratio (PDR) for each application and the average power consumption for each sensor node. The PDR is computed at the application layer as the fraction of packets successfully delivered at the gateway. As can be seen in Table 2, all three alternatives achieved 100% PDRs for the wave-form stream. For the parameter stream, only a few packets were dropped due to retry limit with 802.15.4, but the PDR was close to unity in all three cases.

With respect to power consumption, both 802.15.4 and 802.11e used few mWs in this scenario, with 802.15.4 having the lowest consumption. For the wave-form sensor, 802.15.4 consumed approximately half of the power consumed by 802.11e with AC_VO category. The difference in power consumption was bigger for the parameter sensor, which consumed approximately 10 times less power with 802.15.4 than with 802.11e AC_VO.

Table 1 Simulation parameters.

802.11	P_{rx}, P_{tx}, P_{PS} (mW)[1]		330, 550, 2
	T_{PS-A} (μsec)[2]		10
	PHY rate (Mbits/sec)		11
	AC_BE	AIFS (μsec)	170
		CW_{min}	31
		TXOP limit[3]	0
	AC_VO	AIFS (μsec)	70
		CW_{min}	7
		TXOP limit	3.264 msec
802.15.4	P_{rx}, P_{tx}, P_{ps} (mW)		36, 31.3, 0.0018
	aMaxBE		5
	MaxCSMABackoff		4
	Retry limit		3

Table 2 Average PDR and power consumption in a stand-alone BSN.

Sensor		**802.15.4**	**802.11e AC_BE**	**802.11e AC_VO**
Wave-form	PDR	100%	100%	100%
	Power (mW)	1.82	4.01	3.57
Parameter	PDR	99.77%	100%	100%
	Power (mW)	0.26	2.88	2.77

In order to compare the energy efficiency of the protocols, we also analyzed the average energy consumed per bit delivered at the gateway (Fig. 3). As shown in Fig. 3, 802.15.4 is the most efficient in this scenario. Furthermore, the less energy consumption per bit for the wave-form sensor compared to parametric sensor is due to larger packet sizes when compared to the parameter application. The advantages of 802.15.4 compared to 802.11 are clear in this scenario, mainly because there is no other contending traffic and due to the low traffic load generated by the sensors. In this case, the power consumption characteristics of the chipset play a major role in determining the overall power consumption.

In summary, the results show that 802.15.4 can provide QoS in a stand-alone BSN, while consuming a few mWs. On the other hand, 802.11e can also provide QoS, but at the cost of higher power consumption, although the power consumption is on the order of few mWs as well. However, these power figures are still high for BSN applications that use implantable sensors. Neither 802.15.4 nor 802.11 are scalable in terms of power consumption, and they could not be used as a single solution for all BSN applications.

1 P_{tx}, P_{rx} and P_{ps} are the values for power consumed in transmit, receive and power save mode, respectively, for the BWG 211 low power chipset [8].

2 T_{PS-A} is the turnaround transition time between PS and Active modes for the 802.11 device.

3 TXOP Limit 0 indicates that a single data frame may be transmitted in each TXOP.

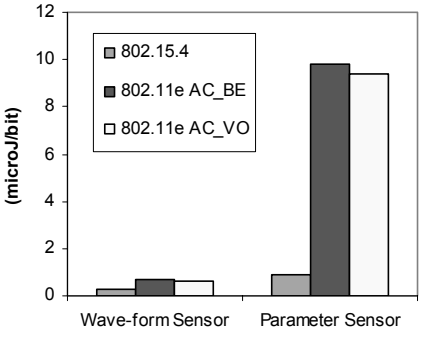


Fig. 3 Energy consumed by sensor nodes per bit delivered at the gateway in the Stand-alone BSN scenario

IV. Performance in Co-existence Scenarios

In this section we study the three co-existence scenarios where the BSN operates in presence of WLAN interference (see Fig. 2). The simulation parameters for the BSN are identical to the stand-alone BSN (see Table 1). The other WLAN devices operate at 11 Mbps with 15 dBm transmit power. The VoIP phone uses the AC_VO; the IT interferer uses the AC_BE, and the Video interferer uses AC_VI access category, all with default parameters [9].

The average energy consumption per bit delivered for the wave-form and parameter sensors are shown in Fig 4 and Fig 5, respectively. For the sake of comparison, we included stand-alone BSN (BSN) results in the figures as well as the corresponding average PDR (%) on top of each column.

It can be seen in Fig 4 that the WLAN voice call (scenario BSN+VoIP) did not significantly impact the performance of the wave-form sensor with respect to the stand-alone BSN, regardless of the protocol used. However, the performance changed drastically, with IT (BSN+VoIP+IT), where 802.11e was more energy efficient, as well as provided better QoS, since 100% of the packets were successfully delivered, against only 36.58% with 802.15.4. The poor performance of 802.15.4 was mainly due the fact that the sensors had to backoff more frequently to the contending WLAN traffic, which resulted in packets being dropped due to the *MaxCSMABackoff* limit. Also, more packets were lost at the gateway due to collisions and channel errors caused by WLAN transmissions. Similarly, with video traffic (BSN+Video), the 802.15.4's PDR was drastically impacted (48.32%). The lowest energy consumption per bit delivered was achieved with 802.11e AC_VO. Moreover, the performance with 802.11e AC_VO did not change much as the interferers were introduced, but this was achieved at the cost of higher power consumption (see Table 3).

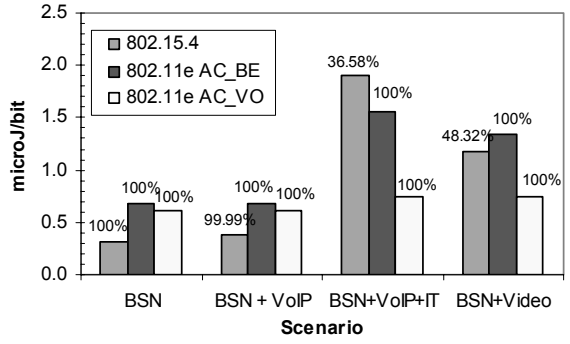


Fig 4 Energy consumed by wave-form sensor per bit delivered at the gateway and PDR in co-existence scenarios

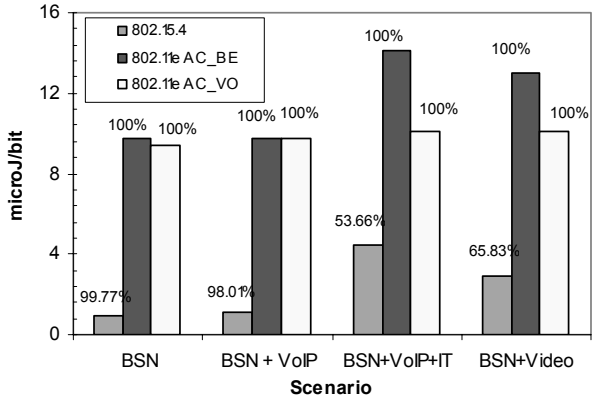


Fig 5: Energy consumed by parameter sensor per bit delivered at the gateway and PDR in co-existence scenarios

Fig 5 shows that 802.15.4 provided the lowest energy consumption per bit delivered for the parameter sensor in all scenarios. However, it could not provide QoS in co-existence with IT and video, where the average 802.15.4 PDRs were only 53.66% and 65.83%, respectively. Again, 802.11e provided better QoS in this case, but it consumes more power (see Table 3).

Table 3 Average power consumption (mW) for wave-form and parameter sensors

Sensor/Coexistence Scenario		**802.15.4**	**802.11e AC_BE**	**802.11e AC_VO**
Wave-form Sensor	BSN+VoIP	2.27	4.05	3.59
	BSN+VoIP+IT	3.77	9.22	4.35
	BSN+Video	3.38	7.91	4.43
Parame-ter Sensor	BSN+VoIP	0.30	2.89	2.89
	BSN+VoIP+IT	0.67	4.18	2.97
	BSN+Video	0.55	3.84	2.99

V. Conclusions

This paper provides a comprehensive simulation analysis of 802.15.4 and 802.11e as connectivity solutions for BSNs. Our results confirm that 802.15.4 is more energy efficient in controlled environments, i.e., without interference, but it is fails in supporting QoS in some co-existence scenarios, which is a serious issue for healthcare applications. In co-existence scenarios, 802.11e with AC_VO category could provide a compromise between power consumption and QoS. However, the power consumption characteristics of 802.11 is the main bottleneck to achieve extremely low power, which is desirable to improve the lifetime of body sensors, and avoid frequent battery replacements. Neither option fully satisfy all the requirements for BSN, therefore there is a need for new PHY and MAC protocols that provide scalable power consumption for different types of sensors, support QoS and co-exist with other networks.

References

1. ZigBee Alliance, Zigbee Specification, Rev. r13 December 2006.
2. IEEE 802.15.4 Std, “Wireless medium access control (MAC) and physical layer (PHY) specifications for low-rate Wireless Personal Area Networks (LR-WPANs),” 2003 Edition.
3. A. Sikora, and V. Groza, “Coexistence of IEEE 802.15.4 with other Systems in the 2.4 GHz ISM-Band,” IEEE IMTC Proc., May 2005.
4. N. Golmie, D. Cypher, and O. Rebala, "Performance Analysis of Low Rate Wireless Technologies for Medical Applications," Computer Communications, Vol. 28, NO. 10, June 2005, pp 1255-1275.
5. N. Chevrollier, N. Montavont and N. Golmie, "Handovers and Interference Mitigation in Healthcare Environments," IEEE MILCOM Proc., Oct 2005.
6. I. Howitt, and J. Gutierrez, “IEEE 802.15.4 Low Rate–Wireless Personal Area Network Coexistence Issues,” IEEE WNCN Proc. 2003.
7. IEEE 802.11 Std, “Wireless LAN medium access control (MAC) and physical layer (PHY) specifications,” 1999 (Reaff 2003).
8. BGW 211 Low Power WLAN SiP, http://www.nxp.com
9. IEEE 802.11e Std “Amendment to Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Medium Access Control Quality of Services Enhancements,” November 2005.
10. Bluetooth SIG. www.bluetooth.com
11. N. Timmos and W. Scanlon, “Analysis of the Performance of IEEE 802.15.4 for Medical Sensor Body Area Networking,” IEEE SECON Proc., 2004.
12. NIST Zigbee Model for OPNET, available at: www.opnet.com.

Address of the corresponding author:

Author: Dave Cavalcanti
Institute: Philips Research North America
Street: 345 Scarborough Road
City: Briarcliff Manor, NY
Country: USA
Email: dave.cavalcanti@philips.com

Low Energy On-Body Communication for BSN

Hoi-Jun Yoo, Seong-Jun Song, Namjun Cho and Hye-Jeong Kim

Semiconductor System Lab, Dept. of EE, KAIST, Daejeon, Korea
hjyoo@ee.kaist.ac.kr

Abstract—**Low power on-body communication is introduced. The human body surface is examined for the communication channel and 10KHz-100MHz frequency band, 'bodywire', is found to be effective in the wireless on-body communication. DCI is proposed to avoid any intentional ground electrode for the capacitive coupling. A CMOS transceiver chip for the on-body communication is fabricated and can achieve 2Mbps with 0.2mW power consumption. The architecture of the BSN controller is proposed and fabricated with CMOS. It has a 16b RISC and a schedule director with TCAM. It can separately control 254 sensor nodes and consumes 14uW in normal mode and 160uW in alert mode including leakage current. The fabricated chip is used to transmit MP3 data from the finger tip to the earphone to enjoy the music. In addition, the BSN controller can detect the emotion of the user by using the data from the sensor nodes transmitted through on-body communication channel.**

Keywords— **On-Body Communication, BSN transceiver, BSN controller, Low Power, High Speed.**

I. INTRODUCTION

Body sensor network opens up many new potential applications in the medical, healthcare and entertainment areas[1]. Although many researches have been directed to the in/through-body wireless communication or off-body wireless communication of Fig 1 very seldom research on the wireless on-body communication has been reported for BSN application[2]. The off-body, BSN has used the frequency band of 2.4GHz for its compatibility with the existing wireless communication standards such as Bluetooth or UWB[3]. It can transmit the signal to the relatively long distance, ~10m. But, its energy consumption for the communication is relatively high, the order of 10nJ/bit. The <50Cm distance in-body communication system supports low speed data rate, ~100bps, with 430MHz carrier frequency. The short distance, <5Cm, through-body communication between implanted device and external read-out terminal can support Mbps data rate using inductive coupling of 415MHz carrier. But its available power is less than milliwatts due to biological safety levels of 10mW/Cm^2 and again it gives a limitation in the channel capacity.

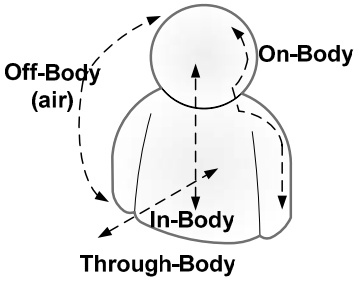


Fig. 1 Types of Body Channel Communication

This work is sponsored by Korean Ministry of Information and Communication and Institute of Information Technology Assessment.

In this paper, the on-body communication with less than 200MHz frequency band is introduced for low energy and high speed body sensor network. The proposed on-body communication covers < 2m distance on the skin of the human body. Its usage is relatively safe and convenient compared to other communication methods. For the frequency range ~ 2.4GHz, most of the signal energy is outside of the human body due to the radiation and skin effect[3]. Therefore, the signal is not transmitted on the body or in the body. The proposed on-body communication uses 10KHz – 200MHz band. Early researches of the on-body communication studied <1MHz frequency band but its signal loss is too large to get the practical application and its signal transfer rate is relatively low[4,5,6]. Compared with other body communication band, the frequency band of our on-body communication has been relatively less studied.

We examine the channel characteristics of the human body. Based on the results of the study, the high speed low power transceiver and intelligent controller ICs are designed and fabricated. Their details will be explained in the following sections. In addition, the applications of the on-body communication to the BSN will be introduced.

II. MODEL OF THE BODY CHANNEL

The blood has the lowest resistivity, 1.5 Ωm, in the human body and transmits most of the signal. The average conductivity of the human body is in the range of 0.1 – 0.5 S/m for the frequency band of 100KHz – 300MHz.

The Fig. 2 shows the schematic diagram of the body channel communication. The external signal is capacitively coupled to the conducting material below the skin. Then, the receiver recovers the signal by capacitive coupling at a distant location.

The coupling capacitor can be quasi statically calculated and the human body is assumed as an ideal conductor, a point node[4]. However, as the frequency increases, the effect of the coupling capacitor decreases and the impedance of the human body cannot be ignored. A distributed RC model of Fig. 3 is proposed to describe the signal transmission through the human body for the 1MHz-300MHz band.

The unit RC circuit is obtained when the body is segmented into 10Cm length as an unit block, and the circuit model of the human body can be constructed by the string of the unit blocks. The resistance of the unit block in this study is 15Ω and its coupling capacitance to the earth ground is 8pF[7]. The air coupling capacitances between the transceiver ground and the body is highly affected by the body configurations and has the values of 10fF – 100fF. By cascading multiple RC blocks, a complete circuit model of the human body can be constructed as shown in Fig. 3. The frequency characteristics of the human body are simulated by using the distributed RC model. Its results are compared with the measured values as shown in Fig. 4.

From the Fig 4, at the low frequency region, the body channel behaves like a high pass filter because the coupling capacitor in the return path is dominant under 8MHz. However, above 20MHz, the resistance of the human body plays more important role than the coupling capacitor. Although the return path through the capacitive coupling between the body and the external ground becomes stronger and induces larger signal with frequency, more signal power is radiated out and wasted. As a result, its S_{21} curve is bent gradually downward like the convex curve of the band pass filter. For the longer distance of the channel, the convex point is placed at the lower frequency. This is because the area exposed to the air increases resulting in more radiation with the channel length. For the step input, the output is a pulse signal with the width of about 8ns, corresponding to the bandwidth of 125MHz.

From the frequency characteristics, the human body is modeled as a band-pass filter with the bandwidth of about 100MHz and shows about 5dB attenuation. According to this investigation, the suitable frequency for the body channel communication in the human body channel is found to exist in the range of 10kHz to 120MHz. It is named as "*the Bodywire channel*" which can give the vast bandwidth for data communications in the body area environment.

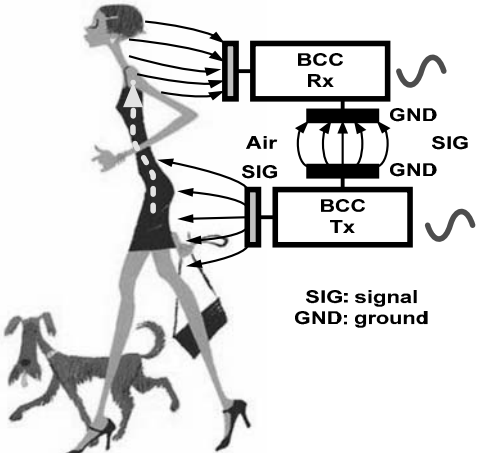


Fig. 2 Schematic Diagram of On-Body Communication System

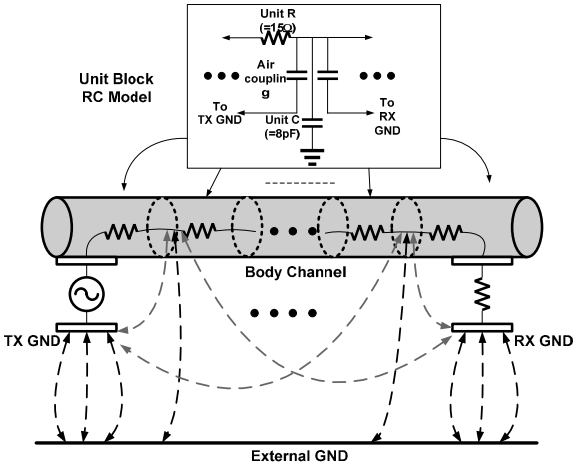


Fig. 3 Electrical Model of the Human Body

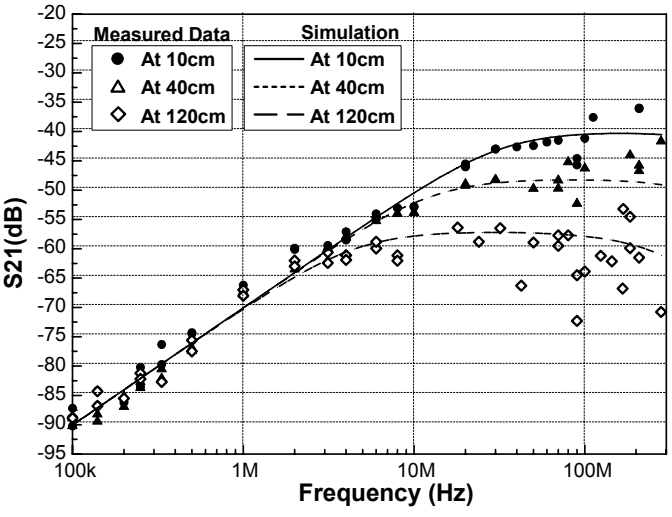


Fig. 4 Frequency Characteristics of the Human Body

III. LOW POWER TRANSCEIVERS

Based on the study of the channel characteristics of the human body, we propose a new method, the wideband pulse signaling (WBS) transceiver with a direct-coupled interface (DCI) to achieve lower power consumption and higher data rate. DCI can remove the purposeful ground electrode from the transmitter and receiver leading to simple implementation. For example, an extra ground electrode contacting the skin or a special electro-optic sensor to detect the feeble electric-field is unnecessary.

A test IC is designed, fabricated and tested in the real environment to demonstrate its feasibility and find the performance requirements for the transceiver, such as the signal detection sensitivity and power consumption[2]. The fabricated transceiver adopts the WBS technique through the *Bodywire* communication channel with DCI interface scheme.

Figure 5 shows the block diagram of the transceiver that comprises a NRZ data transmitter and a CDR-based WBS receiver. The AFE accomplishes three tasks[8]: amplifying, triggering, and level shifting. A wide bandwidth preamplifier provides sufficient amplification to a pulse signal although it is corrupted by the channel. This enables the stable triggering to positive and negative states in the following schmitt trigger circuit. Subsequently, the signal is shifted up to ground level.

A test DCI is implemented with a PCB including a single Ag/AgCl electrode powered by an alkaline battery. Figure 6 shows the measured eye diagrams of input data, recovered clock and data of the CDR circuit for 2Mb/s 2^7-1 PRBS, exhibiting the recovered clock jitter to be 1.4ns rms. The WBS transceiver chip shown in Figure 7 is fabricated using 0.25-μm standard CMOS technology and its core area is 0.85 mm^2. Its power consumption is less than 0.2mW at a single 1V supply. The AFE consumes most of the total power[8]. The measured bit error rate is less than 10^{-7} for 2^7-1 PRBS. Figure 8 shows the power loss along the distance on the body and the signal degradation at 1m distance is -20dB which is within the receiver detection level. The transmission of the MP3 file from the finger tip to the ear,headphone, is possible and will be explained in Section V.

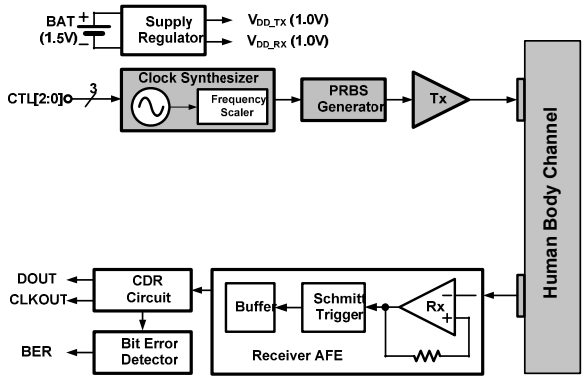


Fig. 5 Schematic Diagram of WBS Transceiver

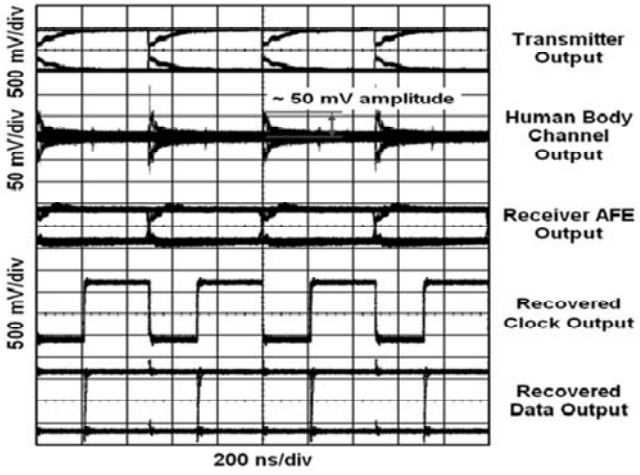


Fig. 6 Eye Diagram of WBS Transceiver Chip

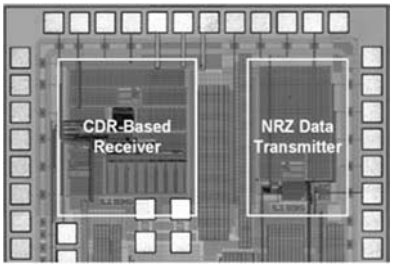


Fig. 7 Microphotograph of the WBS Transceiver Chip

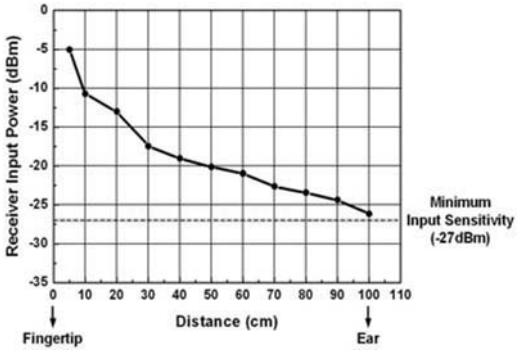


Fig. 8 Power Loss along the Distance

IV. INTELLIGENT BSN CONTROLLER

Collecting and analyzing the vital signals from the various places spread on the human body are the key issues of BSN. An intelligent network controller is required to control data traffic from the multiple simple sensor nodes. Fig. 9 shows a BSN system with one intelligent base-station for the control of all the sensor nodes for the sensing of the vital signals. Data are transmitted as a packet format which consists of 16bit header and 16×n bit payload.

In this paper, we propose an intelligent BSN controller and fabricate it into ultra low power CMOS IC. Figure 10 shows the structure of the proposed bio processor[9]. It has two modes, normal and alert. The normal mode gathers bio-signals from the sensor nodes with programmed periods and checks whether the incoming signal is in alert situation or not. The alert is issued when the signal values exceed the pre-set range. If the value is over the range, the mode is changed to alert mode and branch to the corresponding algorithm to resolve the problem.

In the normal mode, the Schedule Director (SD) requests and analyzes incoming packets, and stores bio-signal data to the Sensed Data Memory (SDM). In the mode, system clock is deactivated, and only SD and SDM work with 16.384kHz CLKSD.

If the incoming packet has the alert information, SD enables the system clock and wakes the RISC up by providing appropriate program counter value. And the RISC controls the whole system by executing the program starting from the new PC address. After the problem is resolved, the RISC powers off itself, deactivates the system clock and the system returns to the normal mode.

The SD checks the schedule for data request in every second and generates the request packet for maximum 254 sensor nodes which have independently programmed request-period. 256x16bit TCAM block is used to store and issue the schedule information instead of using 256 independent timer circuits[10].

To manage 1-day scheduling by second-to-second for 254 independent devices, using the traditional timer needs 21,945,600 clock transitions and it consumes high power. The proposed CAM scheduler, however, requires 86,400 transitions and corresponding search operations, only 0.4% of transitions. It provides 15 kinds of schedule periods, 2/5/10/20/30-seconds, 1/2/5/10/20/30-minutes and 1/2/4/8-hours, which cover most of the measuring period for bio-signal monitoring.

Each bit of the TCAM words indicates validation bit, 2s, 5s ~ 2h, 4h, 8h, respectively. The timer block generates 15bit search data according to the real-time clock. ID generator checks Match Line(ML)s starting from the CAM Word Line. If the match line is activated, it generates corresponding 8bit ID number of the sensor node and SD generates the request packet for the sensor. It checks the next ML after SD gets the requested data successfully. If the result is mismatch, it skips generating the request packet and checks the next ML[10]. Packet Analyzer gets incoming packets, detects alert information and stores the payloads to SDM.

The RISC analyzes the sensor data, executes algorithms and compresses data[11]. It has 16bit data width to handle most of bio-signal data, 3-stage pipelined Harvard architecture and the lossless data compression engine. The data compression engine reduces the amount of the sensing data such as vital signal waveforms without losing its information and saves memory space and power

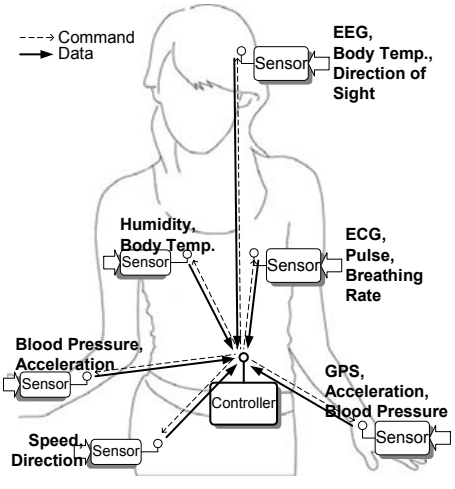


Fig. 9 Concept of BSN using On-Body Communication

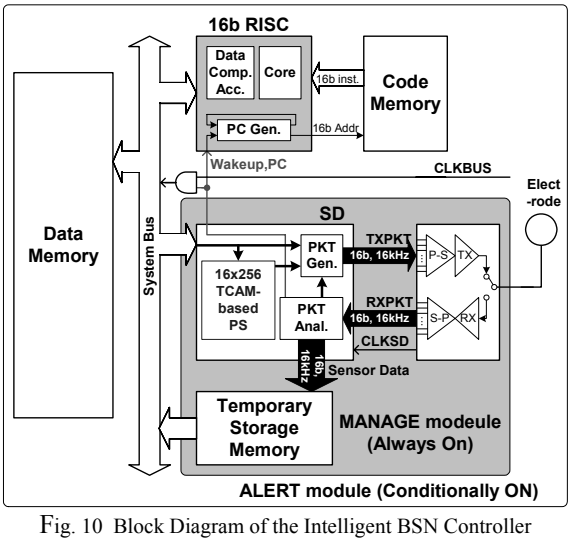


Fig. 10 Block Diagram of the Intelligent BSN Controller

consumption for data transmission. It consists of 16 x 16bit memory block which can be accessed both in vertical and horizontal ways. With the lossless data compression engine, the RISC spends 5% of clock cycle for the data compression compared with general RISCs. It consumes 2.0nJ energy to compress 16x16bit data while the state of the art low power processor requires 32.4nJ.

Figure 11 is the chip microphotograph of the proposed Bio Processor. It consists of 16bit RISC, SD, Human body communication transceiver and 3 memories, 128kb CM, 512kb CDM, 128kb SDM. The 16bit system bus with 4.194304MHz clock connects RISC, CDM, SDM and SD. RISC works as a master, accesses memories and sets SD registers. SD, HBT and write path of SDM works with CLKSD clock and the frequency can be set to 8.192/16.384/32.768kHz according to the required bandwidth mode, slow/normal/fast mode, respectively. With the normal bandwidth mode, the processor consumes 14uW in normal mode and 160uW in alert mode including leakage current.

V. BSN APPLICATIONS

There are many potential applications of the proposed on-body communication to BSN. One is the entertainment application, replacing the wire of the MP3 headphone. Fig. 12 shows a MP3 player with on-body communication transmitter and an earphone with on-body communication receiver. The user can enjoy the high quality MP3 music without any discomfort. The only short coming of this communication method is that the user should contact the electrodes of the transmitter and receiver.

Another interesting application area is the affective computing, especially detection of the emotion[12]. The physiological information such as body temperature, blood pressure and heart rate can be gathered and analyzed to get not only the medical and health information but also the emotional state and behavioural state of the user for the machine intelligences.

In this study, the temperature, blood pressure and ECG are monitored at 10 different spots on the body to get the bio information using the on-body communication, and the information is processed by the intelligent network controller to detect the emotional state of the user.

Fig. 13 shows the photographs of the intelligent controller and many sensor nodes on a human body. The BSN is composed of the 8 temperature sensor nodes(T), 1 ECG(C) or 1 blood pressure sensor node(B) and the measured data are transmitted to the controller through on-body communication link. The sensors are activated and its data are captured at the predetermined time by the controller.

The controller pre-processes the collected data $\{T_i, C, B\}$ to get the 5 statistical data such as range(min/max), means, standard deviations, slopes and comparison with predetermined reference for each sensor data. Then, the obtained 50 physiology dependent features are transferred to outside computer and reduced to 5 principal features to represent most of the features of the data sets. The training data have been gathered from multiple subjects and grouped into their emotion types with afterward user interviews. The testing boundaries are obtained from the training sets and applied to the classification of the emotion. We tentatively categorize the emotion into 4 typical states, Joy, Sadness, Fear and Anger. The bottom of Fig. 13 is an example of the emotion detection predicted with the data collected through on-body communication. Currently, more detail studies incorporating the classification algorithm on the intelligent controller and increasing the accuracy of the emotion detection are under going[13] and will be reported.

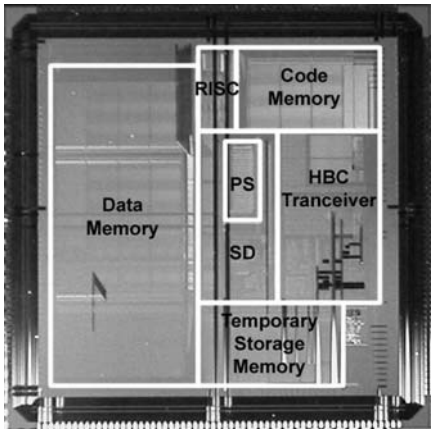


Fig. 11 Photomicrograph of the BSN Controller Chip

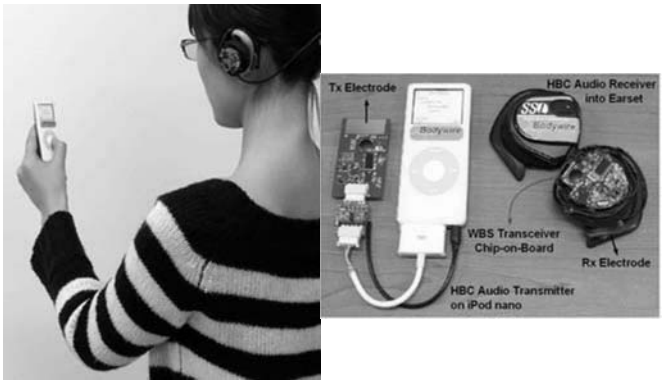


Fig. 12 MP3 File Transfer using On-Body Communication

VI. CONCLUSIONS

We introduced and explained a BSN architecture, a communication chip, a BSN controller chip and its applications. The electrical characteristics of the human body surface are investigated for the use of the communication channel. 10KHz-100MHz frequency band, named as 'bodywire', is found to be most effective in the wireless on-body communication. We did not use any intentional ground electrode for the capacitive coupling and this simple structure is called DCI. A CMOS transceiver chip with the on-body communication is designed and fabricated. It can achieve 2Mbps with 0.2mW power consumption. The architecture of the intelligent BSN controller is proposed and fabricated with CMOS. It has a 16b RISC and a schedule director with TCAM. It can separately control 254 sensor nodes and consumes 14uW in normal mode and 160uW in alert mode including leakage current. The fabricated chip is used to transmit MP3 data from the finger tip to the earphone to enjoy high fidelity music. In addition, the proposed BSN system is used to detect the emotion of the user by classifying the data from the sensor nodes transmitted through on-body communication channel.

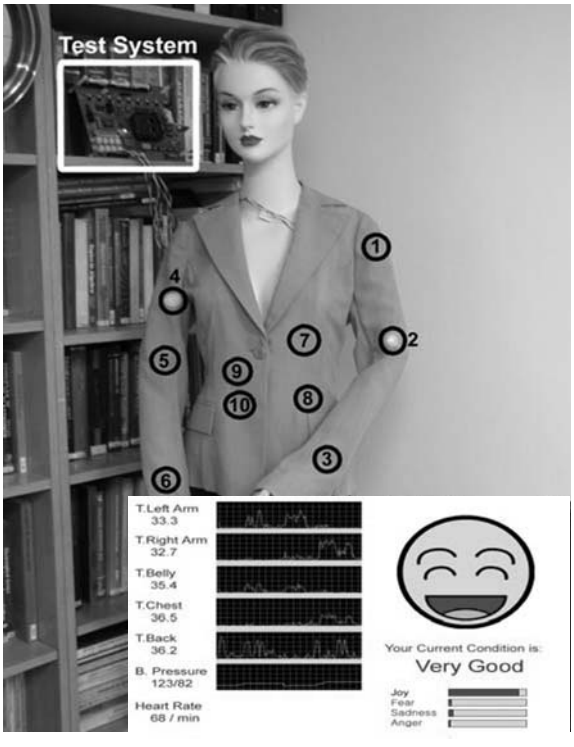


Fig. 13 BSN of On-Body Communication and an Example of Emotion Detection

The on-body wireless communication is convenient and effective for the BSN, and has many potential application areas such as entertainment, healthcare and even affective computing.

ACKNOWLEDGMENT

The authors thank to Dr. Sungdae Choi at Tokyo University and Mr. Jerald Yoo at KAIST for supporting useful data.

REFERENCES

[1] R. Ashok and D. Agrawal, "Next-Generation Wearable Netwoks", IEEE Computer, pp. 31-39, Nov. 2003.
[2] S.-J. Song, et al., "A 2Mb/s Wideband Pulse Transceiver with Direct-Coupled Interface for Human Body Communications," ISSCC Dig. Tech. Papers, pp. 558-559, Feb., 2006.
[3] J. Ryckaert, et al., "A 16mA UWB 3-to-5GHz 20Mpulses/s Quadrature Analog Correlation Receiver in 0.18um CMOS," *ISSCC Dig. Tech. Papers*, pp. 114-115, Feb. 2006
[4] T. G. Zimmerman, "Personal area networks: Near-field intrabody communication," *IBM Syst. J.*, vol. 35, no. 3-4, pp. 609-617, 1996..
[5] K. Hachisuka, et al., "Development and Performance Analysis of an Intra-Body Communication Device," *Transducers '03*, pp. 1722-1725, June 2003
[6] M. Shinagawa, et al., "A Near-Field-Sensing Transceiver for Intrabody Communication Based on the Electrooptic Effect," *IEEE Trans. Instrum. Meas.*, vol. 53, pp. 1533-1538, Dec. 2004.
[7] N. Cho, et al, "The Human Body Chracteristics as a Signal Transmission Medium for Intra Body Communication," Trans. Microwave Teory and Techniques, Submitted.
[8] S.-J. Song, et al., "A 4.8mW 10Mb/s Wideband Signaling Receiver Analog Front-End for Human Body Communications," Proc. ESSCIRC, pp. 488-491, Sep, 2006.
[9] Sungdae Choi et al, "A 24.2-μW Dual-Mode Human Body Communication Controller for Body Sensor Network" ESSCIRC Dig. Tech. Papers, pp.227-230, Sep. 2006.
[10] Sungdae Choi et al, "A TCAM based Periodic Event Generator for Multi-node Management in Body Sensor Network," Proc. ASSCC 2006, pp.307-310, Nov. 2006.
[11] Heyjeong Kim, et al, "A Low Power 16bit RISC with Lossless Compression Accelerator for Body Sensor Network System," Proc. ASSCC 2006, pp.207-210, Nov. 2006.
[12] R. W. Picard, et al, "Toward Machine Emotional Intelligence: Analysis of Affective Physiological State," IEEE Trans. Pattern Analysis and Machine Intelligence, pp. 1175-1190, Oct. 2001.
[13] Heyjeong Kim, et al, "A Low Power Sensor Node Controller for Human Body Monitoring," Journal of Korean Institute of Next Generation Computing, pp.14-20, Vol.2, No. 3, Sep. 2006.

Address of the corresponding author:

Prof. Hoi-Jun Yoo
Department of EE & CS, KAIST,
373-1, Guseong-dong, Yuseong-gu, Daejeon, 305-701, Korea.
Tel: +82-(42)869-3468
Fax: +82-(42)869-3410
Email: hjyoo@ee.kaist.ac.k

2nd Session
Smart Textiles

Bioimpedance Spectroscopy with textile Electrodes for a continuous Monitoring Application

G. Medrano[1], L. Beckmann[1], N. Zimmermann[2], T. Grundmann[2], T. Gries[2] and S. Leonhardt[1]

[1] Philips Chair for Medical Information Technology (MedIT), Aachen, Germany
[2] Institute für Textiltechnik der RWTH Aachen (ITA), Aachen, Germany

Abstract— **Bioimpedance spectroscopy (BIS) enables the determination of the human body composition (e.g. fat content, water content). From this data, it is possible to draw conclusions about the person's health condition. The measurement is carried out with at least four electrodes placed on the body. Nowadays, positioning and wiring of the electrodes can only been conducted by qualified personnel. Even the latest systems on the market are uncomfortable to wear and their use for mobile purposes is highly limited. The commercial BIS electrodes, not suitable for a long term use, may cause allergic reactions. Textile integration plays a role not just concerning the manufacturing of long term electrodes, but also concerning the integration of cables and other electrical components into a wearable and comfortable application. In this article, a portable BIS system combined with textile electrodes is presented as a possible future application. The first validation results of the portable BIS-device and textile electrodes are analysed and the suitability of application is discussed in order to develop a wearable bioimpedance spectroscopy system.**

Keywords— **bioimpedance spectroscopy, textile electrodes, long-term-monitoring, wearable**

I. Introduction

The textile integration of sensors and electronic components will play an important role in the future within the medical-technical area [1]. As a consequence of the natural ageing process, elderly people gradually lose the physiological attraction of eating and drinking, which can lead to age anorexia. Tumor patients also frequently suffer from severe body weight loss (Kachexia), in particular after chemotherapy sessions. These examples point to the necessity of an exact monitoring of nutritional condition and water balance. The continuous monitoring of elderly humans living on their own could be facilitated through the integration of monitoring systems into everyday textiles, e.g. clothing. This monitoring takes place with the help of the bioimpedance spectroscopy (BIS), which determines the composition (e.g. fat content and water content) of the human body. This data enables conclusions about the person's health condition.

So far, dilution methods have been the gold standard for body composition analysis [2]. Nevertheless, they are time consuming and not suitable for a continuous mobile application. In contrast, BIS allows a fast measurement and its implementation in a portable application [2] but has suffered from non-credibility because of its reported limited accuracy when compared with the gold standard. Recent research, however, has shown that BIS could be as precise as the traditional dilution methods or even better [3].

Although BIS is a suitable technology in terms of simplicity and portability for a continuous monitoring application, it still lacks a wearable character. The actual BIS procedure demands the measurements to be undertaken under controlled conditions. The exact placement of the electrodes, the necessary wiring and control of other factors (e.g. room temperature, body position) cause the measurement to be only accomplished by technical personnel. Nowadays electrodes consist of aluminium and are coated with a hydrogel, which both serve as the adhesive of the electrodes on the skin and as an electrolytic medium. The use of gel coating may produce allergic reactions, which limits its long-term use. A wearable system should allow an easy use, so that the user can wear the system without supervision, but at the same time it should have enough precision to detect small physiological changes (i.e. a mild dehydration means only a loss of 5 % of body water [4]). The system also has to take the influence of external parameters into account. From our experience and from that of other authors, the influence of the body position on the measured impedance alone can be around ± 3-6 % [5], [6]. In the last years, several researchers have reported on the development of portable BIS systems [7], [8], which do not satisfy the mentioned requirements. Vuorela [7] uses a textile system, which seems to offer good wearing comfort. However, it does not take into account the influence of external parameters (e.g. temperature or body position) and uses just one frequency. It has already been demonstrated that the use of a single frequency is not sensitive enough to detect small physiological changes [9]. Yang et al. [8] neither takes the comfort factor into account, nor the influence of external factors.

Beckmann et al. [10] have recently reported on the development of a portable BIS device which uses 50 frequencies ranging from 5 KHz to 1 MHz. The results presented

by Medrano et al. [11] using textile electrodes in BIS measurements show the suitability and the possible implementation of a truly wearable system using a modified version of this portable device.

II. Theoretical Background

A. Basics of Bioimpedance Spectroscopy

The determination of the body composition through BIS is based on the fact that the electrical characteristics of the human body change according to the relative amount of body fluid and tissues. Materials such as blood or muscle have a higher conductivity in comparison to bones or fat [12].

The content of water in a human tissue can be divided into intracellular (ICW) and extracellular water (ECW), which are separated by the cellular membrane. The ECW and ICW are predominantly electrical resistive entities, whereas the cellular membrane, due to its lipid layer, has an isolating (capacitive) behaviour. According to that the behaviour of an injected current will be different for low and high frequencies: low frequency current only flows around the cells trough the ECW, whereas a high frequency current will also pass through the cell membrane and the ICW (see figure 1 left).

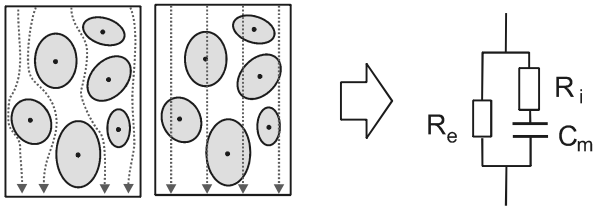


Fig. 1 Low and high frequency current flow through body tissue

This phenomenon can be represented by the electrical model given in figure 1 (right), known as the Cole-Cole model [13]. The values of the electrical model R_e, R_i and C_m can be determined by measuring the body impedance at 0 (R_0) and ∞ (R_∞) frequencies and solving the equation for the parallel circuit (see equations 1 and 2). In fact, frequencies between 5 KHz and 1 MHz are used instead and curve fitting methods are used to calculate the parameters of the Cole-Cole model (R_e, R_i and C_m).

$$R_0 = R_e \tag{1}$$

$$R_\infty = \frac{R_e R_i}{R_e + R_i} \tag{2}$$

Using the Cole-Cole parameters, the basics of the Hanai Theory [14] and devising the human body as a cylindrical volume, we get the following equations for the calculation of the extracellular water volume (V_{ECW}), the intracellular water volume (V_{ICW}), and the total body Volume (V_{TBW}) [15]:

$$V_{ECW} = (K_b \rho_{ECW} \frac{H^2}{R_e})^{2/3} V_b^{1/3}\ , \tag{3}$$

$$(1+\frac{V_{ICW}}{V_{ECW}})^{5/2} = \frac{R_e + R_i}{R_i}(1+K_b \frac{V_{ICW}}{V_{ECW}})\,, \tag{4}$$

$$V_{TBW} = V_{ICW} + V_{ECW}\,. \tag{5}$$

Where K_b is a dimensionless shape factor, calculated from statistical anatomical measurements, V_b and H are the body volume (m^3) and body height (m) respectively.

B. Electrode Theory

Electrodes constitute an interface between the electronic current in the measurement circuit and the ionic current in the tissue.

In a routine BIS procedure, a four point measurement will be used (see fig. 2). Two electrodes are used for the current injection I(t), and the other two for the voltage measurement V(t). This method assumes that the impedance skin-electrode ($Z_{Skin\text{-}Electrode}$) is the same for the four electrodes, so that there is no influence on the measurement.

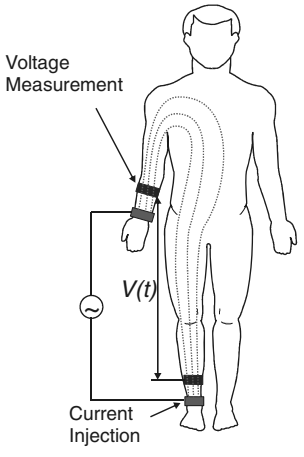


Fig. 2 Position of the electrodes in a Bioimpedance Spectroscopy measurement

A simplified electrical equivalent circuit of the skin and the electrode can be seen in figure 3 (note that the DC voltage source is omitted because of the frequency range). In this model, $Z_{Skin\text{-}Electrode}$ is composed of three impedances in a serial configuration ($Z_{Skin\text{-}Electrode} = Z_{Skin}+Z_{Contact}+Z_{Electrode}$).

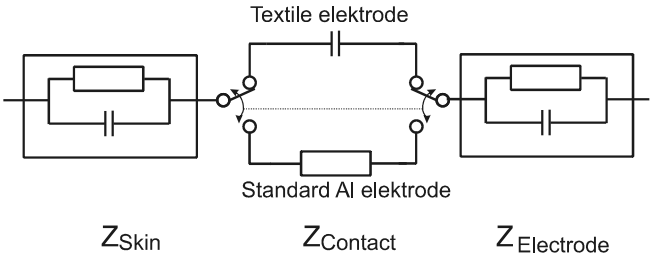


Fig. 3 Simplified electrical equivalent circuit for the skin-electrode impedance

The standard BIS electrodes are composed of aluminium (Al) and a hydrogel (with a resistive behaviour), which acts as adhesive and electrolytic medium ($Z_{Contact} = R_{Gel}$). Due to the lack of electrolytic medium, textile electrodes present a strong capacitive impedance $Z_{Contact}$, which is dependent on the contact between electrode and skin. Small contact changes can then have a big influence in the measurements. High skin impedance affects the common mode voltage in the measuring system, which leads to additional errors in the measurement.

III. Hardware Realization and Results

A. Hardware Design

The portable BIS system developed during this work has an approximated volume of 20 x 16 x 6 cm and a weight of 700 g. The system can be worn on a belt, is battery-powered and measures the body impedance at 100 frequencies between 5 kHz and 1 MHz.

During a measurement, the alternating current I(t) is produced by a signal generator followed by a current source and is injected into the body. The current and voltage measurement are processed by an analogue quadratic demodulator to extract the real and imaginary part of the impedance. These impedance values are sent to the microcontroller via an analogue to digital converter. The microcontroller itself controls the current generation and stores the measured data in a SD card. The hardware also contains a safety switch to guarantee adherence to the IEC 60601-1 electrical safety standard.

After a measurement the data can be read by a PC. The PC uses the values to calculate the parameters of the Cole-Cole model (R_e, R_i, C_m) with an algorithm implemented in Matlab®. After that the amount of ECF, ICF and TBW can be derived from the calculated Cole-Cole parameters using eq. 3, 4 and 5. A block diagram of our whole system as well as a picture of the developed system is shown in figure 4 and 5.

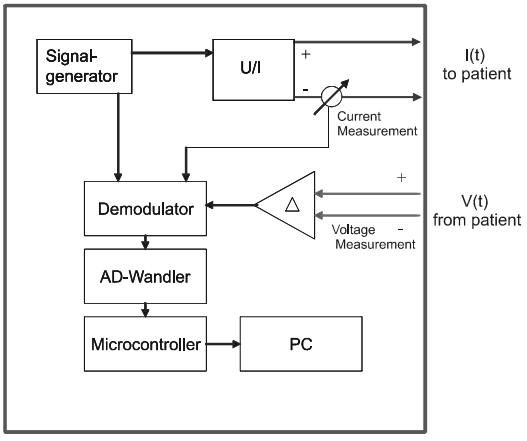


Fig. 4 Block diagram of the portable BIS System

Fig. 5 Portable BIS System

B. Results

The BIS system was tested with different test circuits as well as with several experiments on human volunteers. The commercially available BIS System Xitron Hydra 4200 served as a gold standard.

As one test circuit, the official Xitron test dummy was used, which is used to periodically check the accuracy of the medical certified BIS system (Xitron Hydra 4200). Its circuit diagram represents the resistive and capacitive component of the skin in the points where the electrodes are placed as well as the Cole-Cole Model for a 70 Kg 175cm person. Figure 6 shows the impedance of the test circuit

measured by the new developed BIS System versus the Xitron Hydra 4200.

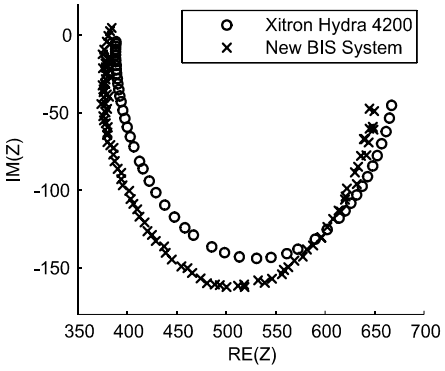


Fig. 6 Impedance Measurement of the new BIS System versus the Xitron Hydra 4200

Using the measured data from figure 6 the Cole-Cole Parameter, the Volume of ECW, ICW und TBW and the deviation of the measured data from the original values are calculated and shown in table 1.

Table 1 Measured Cole Cole Parameter and Water Volumes by comparison

Parameter	Original value	Measured (Xitron)	Measured (New BIS)	Difference [%]
Re (Ω)	681 ± 1%	675.7	686	1.5
Ri (Ω)	909 ± 1%	904.5	880	2.7
Cm (nF)	3.3 ± 5%	3.26	3.39	3.9
ECW (l)	15.81– 16.18	16.07	15.87	1.2
ICW (l)	29.39– 30.74	30.13	31.17	3.4
TBW (l)	45.2– 46,92	46.20	47.0	1.7

The difference in the measurement results between our BIS system and the certified medical BIS device with the dummy model was less than 5%. The results were especially good for Re and ECF (≤ 1.5%) and are within the limits provided by the manufacturer for this model.

First measurement experiences on human volunteers were conducted using aluminium electrodes attached to wrist and ankle. The calculated parameters were all in the correct measurement range, but up to now they did not have the desired accuracy as compared to the Xitron 4200. To solve these problems, a detailed analysis of the system is being undertaken at the moment. Two possible explanations of the deviation of the measurement data could be the influence of the electrode skin intersection as well as the cable connection of the electrodes to the system.

IV. Fabrication of the Textile Electrodes and Results

A. Material and Method used

The yarn for the production of the electrodes must possess a very high conductivity, whereby an exact evaluation of the results of measurement is made possible with bio impedance spectroscopy. Furthermore, it should be biocompatible due to the constant skin contact. That means, it must be adapted in its chemical and physical surface properties to the skin, in order to ensure a good compatibility [16]. Due to the antibacterial effect and the fulfilment of all criteria specified above, silvered (Ag) yarn is well suitable. The production of the textile electrodes takes place on a circular knitting machine. Some example of manufactured textiles electrodes are shown together with the standard BIS electrodes in figure 7.

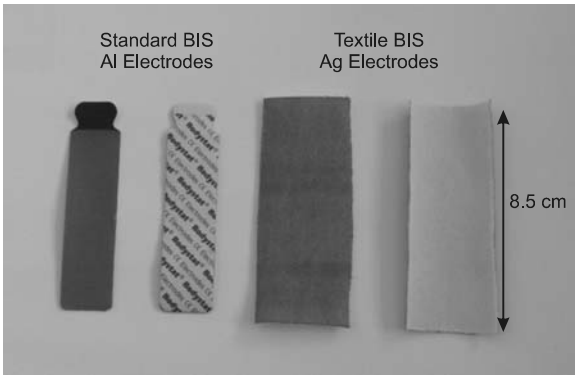


Fig. 7 Standard aluminium (left) and textile silver (right) BIS electrodes

The BIS measurements were performed using a commercial BIS device (Hydra 4200 from Xitron Technologies Inc.), in a room with controlled temperature (24°C).

All experiments were done in a lying position (4 minutes after lying) between wrist and ankle, positioning the standard BIS-electrodes on the left side and the textile electrodes on the right side (see figure 8). The textile BIS electrodes and the BIS device where connected using the standard device connectors (“crocodile clips”). The first attempt was done with textile electrodes without gel. The contact pressure of the electrodes on the skin was controlled through elastic bands, which allowed a loose (low contact pressure) and a fixed, but not uncomfortable, position (higher contact pressure). In the second trial, the textile electrodes were wetted with gel normally used for electrocardiographic measurements.

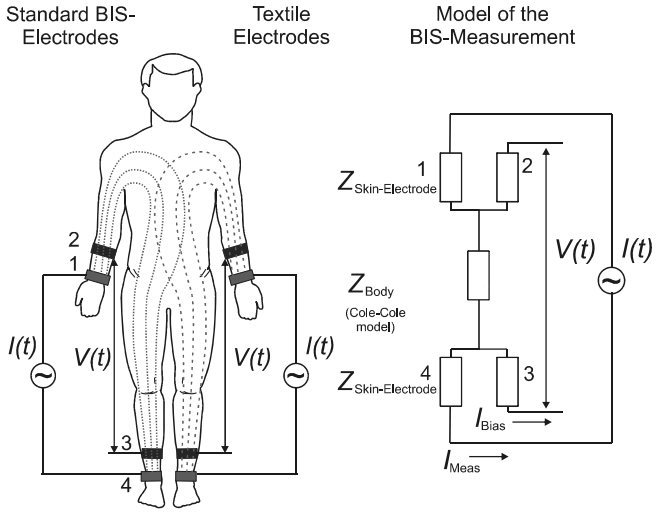


Fig. 8 BIS measurement with standard (left side of body) and textile electrodes (right side of body) and model of the measurement (right)

B. Results

The measured BIS curves are shown in figure 9. The best results, in comparison with the standard BIS electrodes, were achieved with use of textile electrodes and gel. In this case, Z_{Gel} constitutes a resistive component and presents, as a consequence, a similar result with the Standard BIS electrodes ($Z = Z_{Body}$).

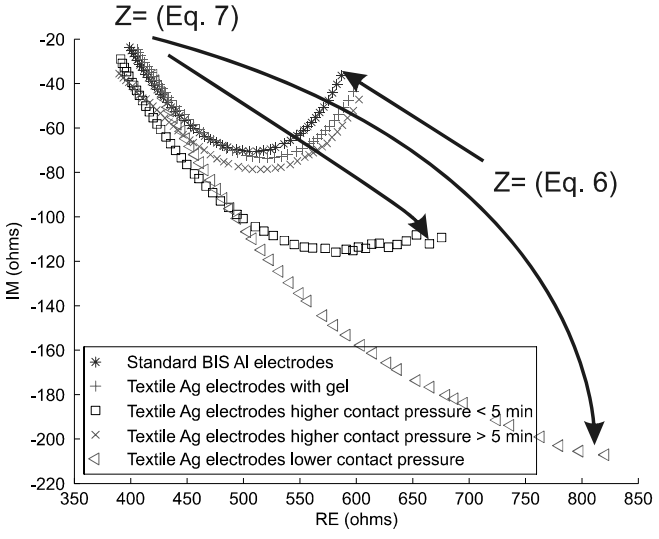


Fig. 9 BIS measurement with standard and textile electrodes (measuring over a period of two minutes)

The measurements with textile silver-electrodes without gel show a higher impedance at lower frequencies ($Z \neq Z_{Body}$), which could be attributed to the capacitive and different electrode coupling (influence of $Z_{Skin\text{-}Electrode}$). A higher contact pressure (better electrode-skin contact) reduces Z_{Gel} and the same impedance values are achieved on every single electrode which actually improves the result.

The results can be explained using an equivalent circuit for the four point measurement (see figure 8 right). An equal $Z_{Skin\text{-}Electrode}$ in the four points (e.g. case of standard BIS electrodes, or textile electrodes with gel) will eliminate the influence of I_{Bias} in the measurement and will allow the following result:

$$Z = \frac{I_{meas} \cdot Z_{Body}}{I_{meas}} \tag{6}$$

In the case of using textile electrodes without gel, the contact between electrode and skin is not the same in all the points. In this case, I_{Bias} will have a big influence in the measurement and the resultant impedance will be:

$$Z = \frac{\left(I_{meas} - I_{Bias}\right) \cdot Z_{Body} + I_{meas} \cdot 2 \cdot Z_{Skin-Electrode}}{I_{meas}} \tag{7}$$

This effect can be better observed with the use of textile electrodes without gel, with lower contact pressure and with higher contact pressure < 5 min. Even the time spent wearing the electrodes improved the result (what we call coupling time). After about five minutes the behaviour of the electrodes improved and the observed measured impedance reproduces equation 6 better than 7 without the necessity of using gel. This effect was most probably due to sweating underneath.

The summary of the calculated parameters R_e, R_i, C_m, ECW, ICW and TBW for the measurements with standard BIS-electrodes, textile electrodes with gel and textile electrodes after 5 minutes use (high contact pressure) is shown in table 2.

Table 2
Calculated Cole-Cole parameters and Water Volumes in measurements using standard and textile BIS electrodes

Parameter	Standard BIS-Electrodes	Textile Electrodes + Gel	Difference (%)	Textile Electrodes	Difference (%)
Re (Ω)	617.1	636.0	3.06	638.1	3.40
Ri (Ω)	1067.5	1056.8	1.00	947.1	11.28
Cm (nF)	2.64	2.73	3.41	2.93	10.98
ECF (l)	18.53	18.16	1.99	18.12	2.21
ICF (l)	27.88	28.29	1.47	31.11	11.59
TBW (l)	46.41	46.45	0.08	49.23	6.08

The difference between the values obtained using the standard electrodes and the textile with gel is lower than 2 % for the physiological parameters (ECW, ICW, TBW). The measurements using the textile electrodes without gel show good results for the calculation of ECW, but a big difference in comparison with standard electrodes by the calculation of ICW and TBW. The observed behavior indicates that a longer coupling time could improve the results.

V. Discussion

The textile integration of the developed portable BIS system can offer the accuracy and simplicity required for a wearable application. The number and range of used frequencies will allow enough sensitivity to detect small changes in physiological measurements [8]. A textile integration of the system allows the development of a comfortable system, where the position of the electrodes will be a part of daily worn clothing.

Nevertheless, further work in the hardware and textile area will be necessary. First, the electronic measurement accuracy of the portable system must be improved. The addition of sensors to monitor external factors like temperature, body position or movement, electrodes position should also be included in the future work. The system should be able to perform measurements only when the optimal conditions are present (e.g. no movement, lying body position, required temperature, required electrode position), which is not always the case in daily life, so that the modelling of the influence of some of this factors may be required. The suitability of the models for each individual and the combination of the new hardware with the textile electrode will be the next major challenges. The problematic of the interconnection of electronic and textile conductors for a wearable application is still unsolved. The different phenomena implicated in a possible textile integration of cables, the interface between circuits and textiles and its influence on the BIS measurements will be further investigated. Also the influence of sweating, time and contact pressure on the textile electrodes should be tested. The use of other structures and other materials in order to diminish the impedance between electrodes and skin is also a necessary future activity.

Acknowledgment

This work was partially sponsored by the Ministry of Science and Research of the state of North Rhine-Westphalia and conducted within the project "Technik-Kultur-Alter. Zukunftstechnologien für Mobilität und Gesundheit: TEKLA" (http://www.tl.rwth-aachen.de/forschung/tekla/index.html)

References

1. Lymberis A, Olsson S (2003) Intelligent Biomedical Clothing for Personal Health and Disease Management: State of the Art. Telemedicine Journal and e-Health. Vol. 9, Nr. 4.
2. Jossinet J (2005) Bioimpedance and p-Health. Personalised Health Management Systems (2005) IOS Press. Nugent C. 35-41.
3. Moissl U, Wabel P, Chamney P, et al. (2006) Body fluid volume determination via body composition spectroscopy in health and disease. Physiol. Meas 27: 921-933.
4. Winters, R. The body fluids in pediatrics (1977) Little, Brown and company. Boston. USA.
5. Medrano G, Beckmann L, Leonhardt S (2006) Einfluss der Körperlage auf Bioimpedanz-Spektroskopie Messungen. Dreiländertagung BMT 2006, ETH Zürich, 6 – 9 September 2006.
6. Fenech M, Jaffrin MY (2004) Extracellular and intracellular volume variations during postural change measured by segmental and wrist-ankle bioimpedance spectroscopy. IEEE Transactions on Biomedical Engineering. 51(1): 166-175.
7. Vuorela T (2005) Textile electrode usage in a bioimpedance measurement, Intelligent Ambience and Well-Being, Finland, September 2005.
8. Yang Y, Wang J, Yu G, et al. (2006) Design and preliminary evaluation of a portable device for the measurement of bioimpedance spectroscopy. Phyiol. Meas. 27: 1293-1310.
9. Gudivaka R, Schoeller R, Kushner F, et al. (1999) Single and multifrequency models for bioelectrical impedance analysis of body water compartments. J Appl. Physiol. 87(3):1087-196.
10. Beckmann L, Medrano G, Leonhardt S (2006) Tragbares Bioimpedanz-Spektroskopie Gerät. Dreiländertagung BMT 2006, ETH Zürich, 6 – 9 September 2006.
11. Medrano G, Zimmermann N, Gries T, Leonhardt S (2006) Bioimpedanz-Spektroskopie (BIS) mit textilen Elektroden. Dreiländertagung BMT 2006, ETH Zürich, 6 – 9 September 2006.
12. Gabriel C, Gabriel S, Lau R W (1996) The dielectrical properties of biological tissues: II. Measurements in the frequency range 10 Hz to 20 GHz. In: Physics in Medicine and Biology 41.
13. Grimnes S, Martinsen O (2000) Bioimpedance and bioelectricity basics.1st ed. Academic Press, London.
14. Hanai, T (1968) Electrical properties of emulsions. In: Sherman DH, ed. Emulsions Science. London: Academic, 354-477.
15. Fenech M, Maasrani M, Jaffrin M (2001) Fluid volumes determination by impedance spectroscopy and hematocrit monitoring: application to pediatric hemodialysis. artificial organs. 25(2): 89-98, Blackwell Science Inc.
16. Wintermantel, E.; Ha, S.W (1998) Biokompatible Werkstoffe und Bauweisen, Implantate für Medizin und Umwelt, 2. Auflage, Springer-Verlag, Heidelberg.

Address of the corresponding author:

Author: Guillermo Medrano, M. Sc.
Institute: Philips Lehrstuhl für Medizinische Informationstechnik
Street: Pauwelsstrasse 20
City: Aachen
Country: Germany
Email: medrano@hia.rwth-aachen.

Contactless EMG sensors embroidered onto textile

Torsten Linz[1], Lena Gourmelon[2] and Geert Langereis[2]

[1] Fraunhofer IZM, Berlin, Germany
[2] Philips Research, Eindhoven, The Netherlands

Abstract— **To obtain maximum unobtrusiveness with sensors for monitoring health parameters on the human body, two technical solutions are combined. First we propose contactless sensors for capacitive electromyography measurements. Secondly, the sensors are integrated into textile, so complete fusion with a wearable garment is enabled. We are presenting the first successful measurements with such sensors.**

Keywords— **surface electromyography, capacitive transducer, embroidery, textile electronics, interconnect**

I. Introduction

An excellent embodiment of a health monitoring device is a wearable system. Clothes are natural possessions and are part of the processes and routines in our daily life. The technological drive is to integrate sensors and electronics into textiles in such a way that the usage and advantages of cloths are maintained. Therefore a high level of textile integration has to be combined with aspects of reliability, comfort and washing resistance.

Sensors suitable for integration in clothes should be non-invasive and must be capable of monitoring health and wellness parameters. The common approach is to take relatively simple measuring techniques and to use signal processing and multi-parameter analysis to derive the physiological parameters of interest. In the ConText project [1-3], the universal measurement method of surface electromyography (sEMG) is used from which information about fatigue is derived by signal processing. Interpretation of the sEMG signal may be assisted by electrocardiography (ECG) signals and the output of movement/position sensors.

In ConText we develop a vest that contactlessly measures an sEMG signal with textile integrated sensors. So the vest can be worn over other clothes and will still extract electrophysiological signals like fatigue. The vest can be used by untrained individuals and therefore enables 24 hour monitoring at home and at work.

In ConText a number of different integration technologies are investigated: weaving of conductive yarn, printing of conductive material onto fabric, lamination of conductive and non-conductive fabric layers and embroidery of conductive yarn. This paper presents the efforts to make textile circuits and interconnections using embroidery.

II. Contactless surface EMG

All living cells are surrounded by membranes. These membranes are selectively permeable for various ions and may actively transport them through the membrane resulting into a membrane potential. Nerve cells and muscle fibres are depolarized when activated by a certain threshold voltage. The result is the propagation of a depolarization wave along the nerve and muscle fibre [4]. Such an electrical wave over the muscle fibre is the direct cause of muscular contraction and is subsequently followed by relaxation. The quick combination of contraction and relaxation of a muscle fibre is referred to as “twitch”. Since all muscle fibres in a muscle do not twitch simultaneously, the overall observed potential over a muscle is the random summation of multiple single fibre action potentials. This random signal is conducted to the surface of the skin by means of volume conduction.

Surface electromyography (sEMG) electrodes are placed on the skin in order to record the muscle potentials. A common configuration is a set-up of two electrodes along a muscle contacting the skin using conductive gel. A problem with this set-up is that the interface potential between the skin and the solid electrode is undefined. In addition, it is uncomfortable to tape such electrodes onto the skin. Therefore, contactless electrodes are proposed in literature to monitor signals of the heart [5-7], which is also a muscle. These electrodes detect an electric displacement current by coupling capacitively to the body instead of detecting a Nernstian current; therefore, they require no electrical contact with the skin. The possibility to avoid direct skin contact reduces the skin irritations problems.

III. Contactless EMG measurements using PCB electrodes

Fig. 1 shows the bipolar set-up with two contactless electrodes with which the first EMG experiments are successfully performed [8]. The first measurements are done using electrodes which are not yet integrated into textile, but consist of 12 mm circular shapes on a printed circuit board.

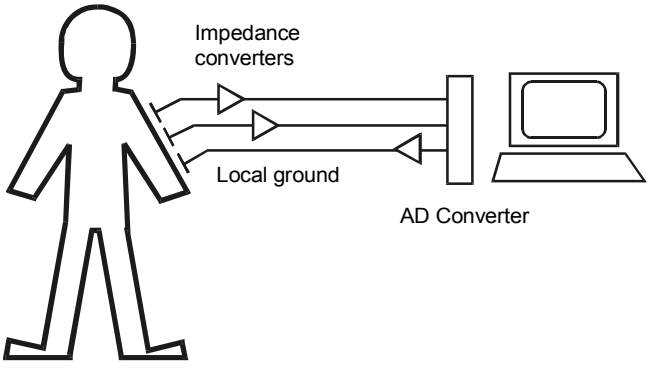


Fig. 1 Set-up for contactlessly sensing EMG signals

Because of the capacitive coupling, the impedance of the sensor is extremely high. The result is that environmental noise is easily picked up. This problem is solved by placing an impedance converter directly on top of the electrode. The electrode is actively shielded by feeding back the output signal of the local amplifier to a metal cap over the electrode. The individual sensor output signals are fed into an analogue to digital converter after anti-aliasing filtering.

Fig. 2 shows the measured EMG signal on the biceps while lifting a weight of 2.5 kg with a 90 degrees bended arm. The contactless sensors have an electrode spacing of 37 mm and a gain of 11. It is compared to a commercial active sEMG electrode (B&L Engineering type BL-AE-N) having a spacing of 20.6 mm and a gain of 346. In Fig. 2, the signals are normalized by the gains to give the skin surface voltage. Note that it is not possible to perform the two recordings simultaneously.

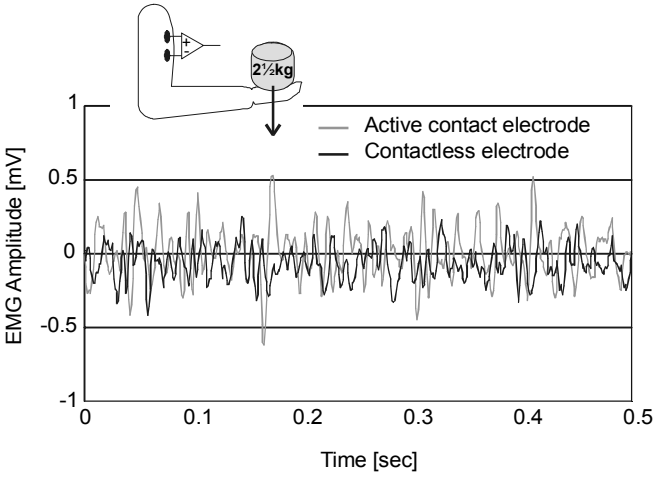


Fig. 2 Recorded EMG signals with both an active- and a contactless electrode

In fig. 3, the spectra are shown for the two measurement methods using the same data set as in fig. 2. As a reference, the spectra during rest are plotted as well. We can see that the contactless electrodes and the commercial active sEMG electrodes provide similar signal levels and shapes. Only the bottom noise during rest is a little bit higher. The bandwidth of the contactless electrode set-up is adequate and comparable to the reference measurement.

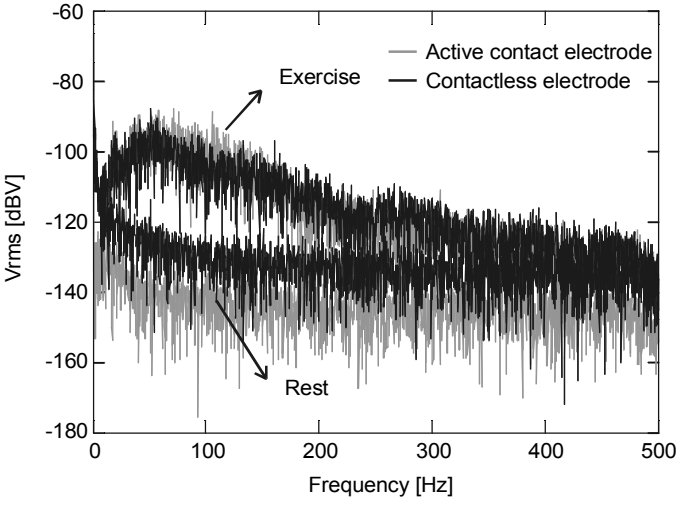


Fig. 3 Spectra of the recorded EMG signals with both an active- and contactless electrode

IV. The electronic module

Textile structures are orders of magnitude larger than electronics structures. Therefore a dimension adapter is required. As proposed in [9] we used a so called interposer for that purpose. This is a flexible polymer substrate which carries the electronics and has conductive pads to connect the conductive textile structures.

Although textile structures are big it is advisable to miniaturize the electronics to archive an overall textile character and to improve the reliability. E.g. thinned silicon chips are flexible and therefore in a mechanically stressed environment more reliable. However miniaturization is very expensive and it is often difficult to find bare dies. Therefore we have only produced a limitedly miniaturized sensor electronics module for the first tests with an amplifier in an SO8 package as shown on the picture.

A 25µm polyimide flex foil was used as substrate and was structured with a 25µm thin layer of copper-nickel-gold metallisation and coated with a 15µm layer of solder-resist mask. Components were only placed on the top side. The substrate was folded to form a ground shielding around the

electronics to protect them form RF noise. All the areas that were not used have been covered with the ground plane to make this shielding most effective.

Unfortunately the whole package is rather thick as some of the components were not available in smaller size. The largest part is the amplifier with a thickness of 1.7mm.

To make an embroidered interconnection to external textile structures as described in [10] the pads on the substrate must be conductive on the top side. As the module has to be folded before embroidery, the folded part has openings above the metallised pads so that the embroidery needle punches through the pads only.

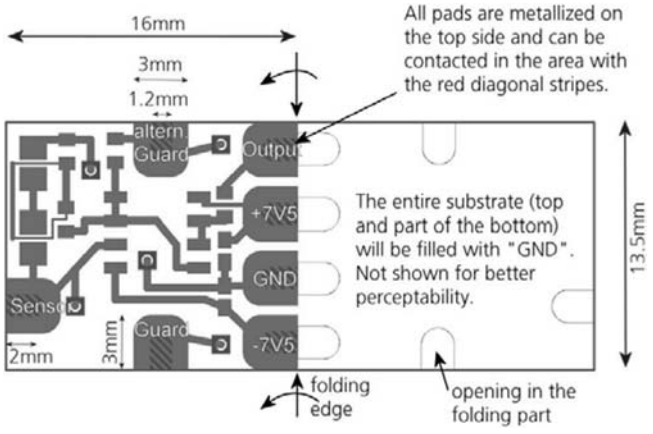


Fig. 4 Design of the flexible substrate

This module is still about as big as the whole sensor when it is manufactured from one piece of flexible substrate (incl. electronics and the sensor disk) and rather thick, which is certainly not desired for the final product. At this first stage of the project an expensive flip chips miniaturization was not worth the risk as the design still needed to prove that it works in textile generally. Generally a bipolar EMG-lead consists always of two disks. It is thinkable that one module of approximately the same size as this one serves as amplifier for both. This will already make the package appear smaller.

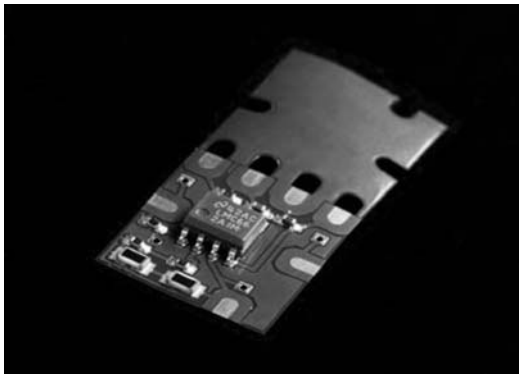


Fig. 5 Flexible substrate

V. Embroidering the senor electrode and the interconnection

In [11] embroidery with conductive yarn is described as a means to produce conductive textile structures on fabric. Additionally in [12] embroidery was found to be an excellent technology to interconnect such embroidered circuits with electronic modules.

The fabrication of embroidered contactless EMG sensors additionally requires a multilayer design. In this project it was found that isolating layers of embroidery can be constructed. As shown in Figure 7 three layers of non-conductive embroidery are required to make sure that the layer below (the disk) and layer above (the guard cap) are isolated from another. To achieve a successful isolation it is essential that each new layer is embroidered perpendicular to the layer below (be it an isolating one or a non-isolating one). Furthermore for the construction of this special sensor it is fundamental that the top layer – the guard cap – is embroidered with a non-conduction bobbin-thread and with a conducting top-thread. Otherwise the guard cap would entirely shield the sensor and it wouldn't sense anything.

Note that the layer one (the disk) is embroidered with a step stitch that means that the needle connects top and bottom thread every 2mm. This is not possible in the layers above because these threads must not go through the conductive areas of layer one. Therefore the other layers are embroidered with a satin stitch which means that the needle goes through the fabric only at the endpoints of a desired structure (e.g. on the circumference of the guard cap in fig. 8).

The first conductive layer shown in figure 6 connects the embroidered circuit with the flexible substrate. This requires that at least the top thread is conductive. However, better results have been achieved with conductive top and bobbin thread.

The sensor connects to the computer via snap fasteners which are crimped through the embroidery as shown in figure 9.

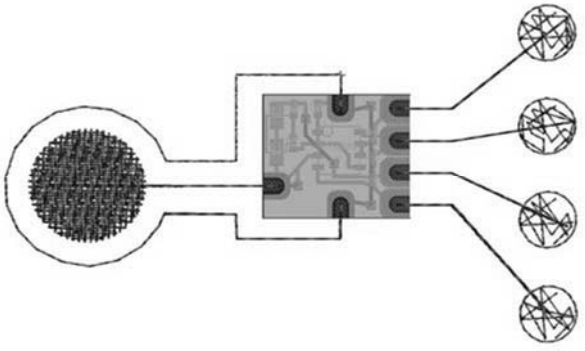


Fig. 6 Folded substrate with first embroidery layer; conductive yarn for the guard connection and the capacitive disk

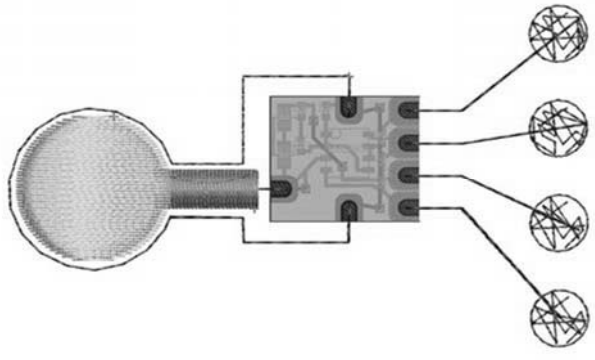


Fig. 7 The second, third and forth layer are non-conducting to isolate the sensor disk from the guard cap

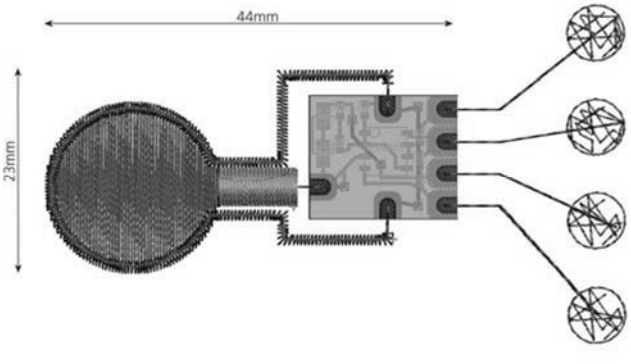


Fig. 8 Fifth layer with conductive yarn for the guard cap

Embroidery is a very demanding process for the yarn. The needle thread is bent by 180° around the needle. The thread is pulled through the eye of the needle and other small parts at a high speed. This can result in fussiness of the yarn or even a breaking. Yarn must be designed for embroidery.

Currently only a very limited number of conductive yarns are available for embroidery. One product series comes from Statex. Good results have been achieved with Shieldex 117/f17 2-ply. However the conductivity of around 500Ohm/m is rather low. A better conductivity can be expected from the ELITEX series by TITV which is based on Shieldex and further galvanized. Good results have been achieved with ELITEX PA/Ag 110/f34 2-ply.

For the interconnection as described in [11] it is necessary that the top-thread is conductive on the surface (everywhere). This means the technology cannot make use of spun yarn that consists of non-conductive fibres spun together with a copper wire as the position of the wire is arbitrary. These yarns cannot be embroidered anyway.

Another pre-condition for this technology is the use of flexible electronic substrates because the needle punches a hole through the substrate. This is not possible with thick FR4 substrates.

An encapsulation of these interconnects is required and will be investigated in the coming months.

Currently we believe that the actual contact mechanism is purely a mechanical one. The thread lies on the metallisation of the pad and is trapped in the gap of the wrenched substrate. Furthermore the contact could be improved with encapsulation, pressing the thread down on the substrate.

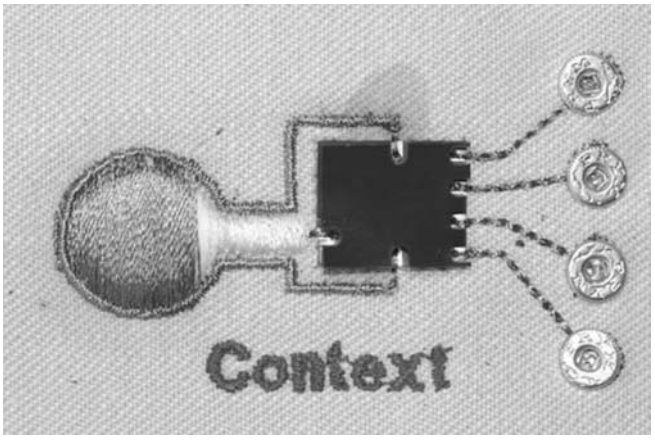


Fig. 9 Embroidered Sensor with interconnections to the electronic module and snap fasteners as interface to the computer

VI. Mesurements with the embroidered sensors

To evaluate the embroidered capacitive transducers, without having the problems of motion artefacts and noise of the human body, an artificial muscle model was used. Fig. 10 shows the hardware model. The muscle itself is emulated by a strip of moderately resistive paper. By two aluminium beams, an electrical current can be forced through this muscle. On top of the muscle, a leather chamois is used which mimics the human skin. On the chamois we can put several types of textile on which in its turn the sensor is placed. So, the model does imitate the contactless behaviour and the distributed shape of a buried muscle, but does not include the human tissue volume conductor properties.

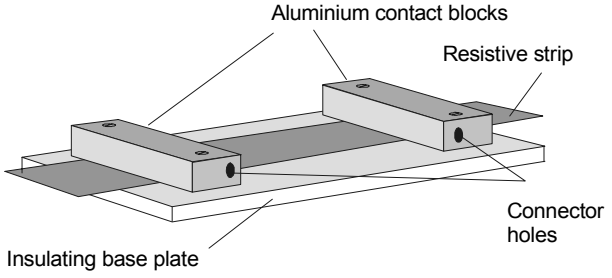


Fig. 10 Artificial muscle model

A waveform generator was connected to the artificial muscle. A square wave of 1 V_{pp} with a frequency of 20 Hz was generated. Fig. 11 shows the recorded signal by using a single embroidered sensor on the artificial muscle. Note that the envelope of the recorded square wave shows a 50 Hz noise signal. This is the result of the single-sensor approach. By using a single sensor with respect to a grounded reference, we will see 50 Hz noise as picked up capacitively from the environment. This will be cancelled when using the set-up of Fig 1.

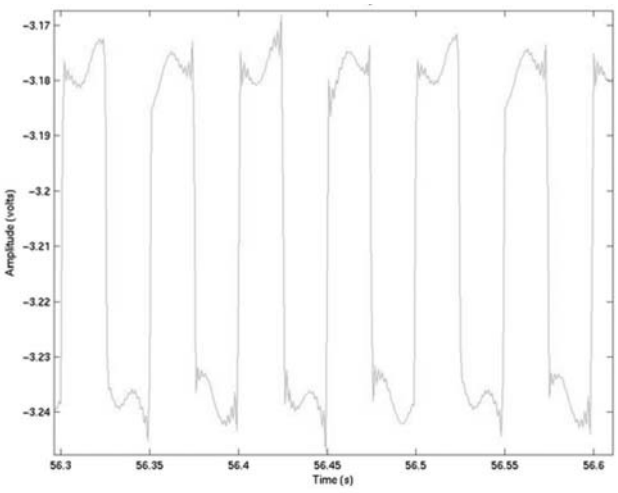


Fig. 11 Capacitively recorded square wave

In Fig. 12, the human EMG is measured on the biceps using two electrodes of the type of Fig. 9. The set-up is similar as used in section III. At t = 0 sec and t = 100 sec, a contraction of the biceps was applied. We can see that muscular activity is clearly detected by the textile embroidered sensors. Some motion artefacts are visible as spikes on the signal.

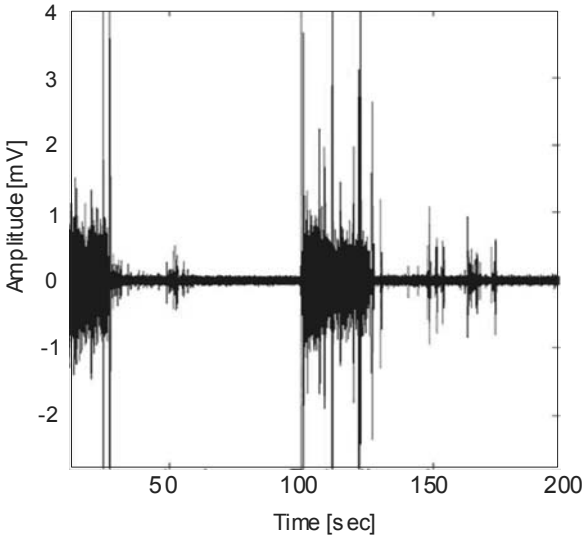


Fig. 12 Capacitively recorded bipolar EMG on a human biceps

VII. Conclusions

It could be demonstrated that contactless EMG electrodes deliver similar signals as contact electrodes and allow at least the differentiation of the states "rest" and "exercise". Furthermore a textile integration based on embroidery was presented and approved by electrical measurements in a test environment.

In the next steps we will try to reduce motion artefacts and improve the robustness of the EMG signal. Once this works the encapsulation and the reliability of the sensor will be tested.

Acknowledgment

We thank the European Commission for funding part of this work under contract IST-027291 (ConText).

References

1. ConText project webpage: http://www.context-project.org/
2. Langereis GR, de Voogd-Claessen L, Sipliä A, Illing Günther H, Spaepen A, Linz T (2006), ConText Contactless sensors for body monitoring incorporated in textiles, FiberMed06; Conf. on Fibrous Products in Medical & Health Care, Tampere, Finland, June 7-9, 2006
3. de Voogd-Claessen L, Sipilä A, Illing-Günther H, Langereis G, Linz T, Spaepen A (2006), Contactless Sensors Integrated into Textiles, Nanotechnologies and Smart Textiles for Industry and Fashion, 11-12 October 2006, The Royal Society, London, UK
4. Merletti R, Parker PA (2004), Electromyography, Physiology, Engineering and Noninvasive Applications, Wiley Interscience, ISBN 0-471-67580-6
5. Richardson PC (1968), The insulated electrode: A pasteless electrocardiographic technique, Proc. Annu. Conf. Eng. Med. Biol., 9:15.7.
6. Smith WJ, LaCourse JR (2004), Non-contact biopotential measurement from the human body using a low-impedance charge amplifier, Proceedings of the IEEE 30th Annual Northeast Bioengineering Conference, 17-18 April 2004, pp 31 - 32
7. Ko Keun Kim, Yong Kiu Lim, Kwang Suk Park (2005), Common Mode Noise Cancellation for Electrically Non-Contact ECG Measurement System on a Chair, Proceeding of the IEEE 27th Annual Conference in Medicine and Biology, Shanghai, China, September 1-4, 2005
8. Gourmelon L, Langereis G (2006), Contactless sensors for surface electromyography, IEEE-EMBC '06, 28th Annual International Conference IEEE Engineering in Medicine and Biology Society (EMBS), New York, August 30 – September 3, 2006
9. Kallmayer C et al. (2003), "New Assembly Technologies for Textile Transponder Systems", Electronic Components and Technology Conference, New Orleans, May 2003.
10. Linz T et al. (2005), "Embroidering Electrical Interconnects with Conductive Yarn for the Integration of Flexible Electronic Modules into Fabric", *IEEE ISWC*, 2005
11. Post ER, Orth M et al. (2000), "E-broidery: Design and fabrication of textile-based computing", *IBM Systems Journal*, IBM, 2000, Vol. 39, pp. 840-860.
12. Linz T et al. (2005), "New Interconnection Technologies for the Integration of Electronics on Textile Substrates", *Ambience 2005*, Tampere, September 2005.
13. T.Linz et al. (2006) "Fully Integrated EKG Shirt based on Embroidered Electrical Interconnections with Conductive Yarn and Miniaturized Flexible Electronics", *IEEE BSN*, Boston, 2006

Address of the corresponding author:

Author: Torsten Linz
Institute: Fraunhofer IZM
Street: Gustav-Meyer-Allee 25
City: 13355 Berlin
Country: Germany
Email: Torsten.Linz@izm.fhg.de

Bio-sensing textiles - Wearable Chemical Biosensors for Health Monitoring

Shirley Coyle[1], Yanzhe Wu[1,2], King-Tong Lau[1], Sarah Brady[1], Gordon Wallace[1,2], Dermot Diamond[1]

[1]Adaptive Sensors Group, School of Chemical Sciences, National Centre for Sensor Research,
Dublin City University, Dublin 9, Ireland

[2]Intelligent Polymer Research Institute (IPRI) and
ARC Centre of Excellence for Electromaterials Science (ACES), University of Wollongong, Wollongong, Australia)

ABSTRACT

In recent years much progress has been made in the integration of physical transducers into clothing e.g. breathing rate, heart rate and temperature [1]. The integration of chemical sensing into textiles adds a new dimension to the field of smart clothing. Wearable chemical sensors may be used to provide valuable information about the wearer's health, monitoring the wearer during their daily routine within their natural environment. In addition to physiological measurements chemical sensors may also be used to monitor the wearer's surrounding environment, identifying safety concerns and detecting threats. Whether the clothes are looking into the wearer's personal health status or looking out into the surroundings, chemical sensing calls for a novel approach to sensor and textile integration. In contrast to physical sensors, chemical sensors and biosensors depend on selective reactions happening at an active surface which must be directly exposed to a sample. Therefore issues of fluid handling, calibration and safety must be considered. This paper discusses the constraints in integrating chemical sensors into a textile substrate. Methods of fluid control using inherently conducting polymers (ICPs) are discussed and a pH textile sensor is presented. This sensor uses colorimetric techniques using LEDs controlled by a wireless platform. Some of the potential applications of wearable chemical sensors are discussed.

KEY WORDS

Wearable sensors, patient monitoring, biosensors, pervasive healthcare

1. Introduction

Wearable sensors provide personalised healthcare through monitoring the wearer in their natural environment, providing a far more realistic outlook than in a clinical setting. An important benefit is that the wearer becomes more aware of their personal health status which has a huge impact for preventative healthcare. Wearable sensors for healthcare have been implemented by many research groups for various applications including cardiovascular disease [2, 3] and ambulatory monitoring of the elderly [4]. Such applications typically monitor physiological signals such as breathing rate, heart rate, ECG and temperature. These are all physical sensors, i.e. they convert physical properties into electrical signals. Chemical sensing has not yet been implemented in these applications although it has the potential of offering much information about the wearer's health. Chemical biosensors have numerous applications in clinical analysis and may offer complimentary information to the physical sensors.

A major issue in monitoring biological samples in vivo is sensor placement and sample delivery. As a wearable device a non-invasive sensing device is essential. Urine, saliva, sweat, tears and breath are possible samples that may be acquired non-invasively [5]. Sweat is the most accessible specimen within a garment, and there are many developments within the textile industry to accommodate the movement of sweat through fabrics for sports performance clothing. This work is part of the EU-funded BIOTEX project which aims to develop real textile sensors embedded in a garment allowing direct collection and analysis of sweat. The paper discusses the development of a textile pH fabric sensor for measuring sweat pH, and discusses how conducting polymers may be used for control fluid handling for sample delivery and calibration.

2. Textiles and Chemical Sensors

Integration of chemical sensors into textiles is not a straightforward task. Consider the known problems of wearable physical sensors which are subject to physiological, environmental noise and motion artefacts. Chemical sensors are subject to similar problems but there are additional constraints due to the nature of the sensing mechanisms. Chemical sensors respond to a particular analyte in a selective way through a chemical reaction. In order for this to happen, changes in the sensor surface or bulk characteristics must occur to give rise to the signal. Therefore the device surface must change and the analyte

must be in contact with the sensor. The following issues must therefore be considered for integration:

Fluid movement/control - The garment must collect samples and deliver samples to sensor. If the sensor is detecting the external environment, e.g. acidic gas plumes, the sample is likely to be volatile, whereas if the sensor is monitoring the body's physiology the sample e.g. sweat must be delivered to the sensor. In either case there may be a need for reagent handling, fluid control and waste storage, for which, the dynamic control of fluid is preferred. The development of fluid handling systems requires efficient microfluidic pumps. Some micropumps may require extremely high voltage to operate [6-8], some are slow to complete an actuation cycle [9-11], some are complicated in configuration and only suitable for particular applications [12-15]. One class of emerging actuation materials for micro-pumps are the inherently conducting polymers (ICPs). Their novel actuation mechanism is based on the reversible ion doping / dedoping process to electrical stimu lation at low voltage ~ 1 V [16] (Figure 1). It is capable of producing at least 10 times more force than skeletal muscle, and potentially 1000 times more [17] and a comparable strain ~ 26% [18] are also practically achievable.

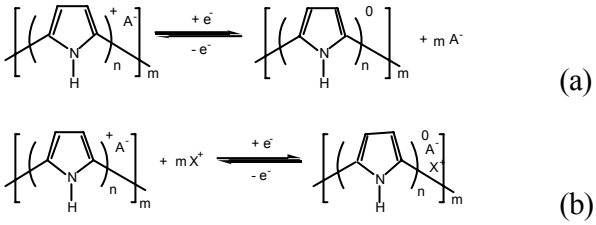


Figure 1. Schematic diagram showing the ion doping /dedoping process in response to the electrical reduction/oxidation of polypyrrole, where A- is the dopant anion incorporated into the PPy during synthesis, X- represents a cation from the electrolyte, n is the number of pyrrole units for each A- incorporated, and m is the numbers of PPy repeat units that determine the molecular weight of polymer, (a) A is small and mobile anionic dopant, (b) A is bulky and immobile anionic dopant.

A micropump called TITAN (means the 'tube in tube aligned node') has been constructed and reported in an earlier paper [19]. The intrinsic resistance of ICP can be utilised to create a peristaltic pumping regime with great potential for integration with small devices such as wearable chemical sensors.

Calibration - Regeneration of original surface characteristics is extremely difficult and devices have to be constantly calibrated. This links in with fluid handling and control, the overall concept is illustrated in figure 2.

Wearability – The sensor itself (including the fluid handling components) must be robust, miniature, flexible, washable in order to be integrated into a textile. Materials that have similar mechanical properties to the textile are required.

Safety – The overall assembly of sensor must be safe for the wearer's health. Non-toxic or hazardous chemicals must be avoided or well isolated from the wearer. Practically, encapsulation is required to separate its potential harm from body. Components that pose a risk of electric should be avoided, e.g. ICP based devices use low operation voltage ~ 1 V. whereas some piezoelectric micropumps use high AC voltage

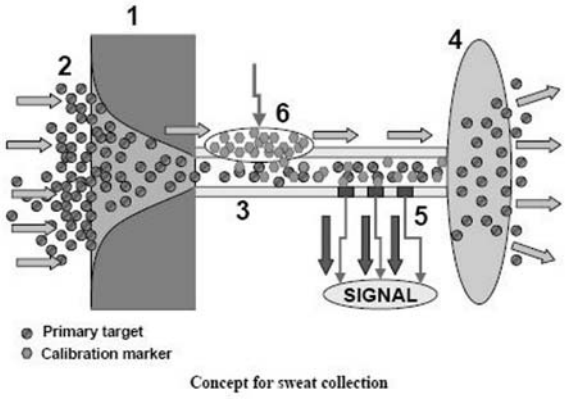


Figure 2. Concept of wearable chemical biosensors for health monitoring using electroactive functional materials, Functionalised material (1) in contact with the skin draws sweat with primary target species, (2) into the fibre tubules or bundles such as the TITAN micropump, (3) channels for the fluid handling to desired locations, (4) waste collection region, (5) surface with sensing capabilities to specifically interact with the target species and generate a signal through LEDs.

3. Colorimetric pH Fabric Sensor

A colorimetric approach is taken to develop a textile-based pH sensor. This involves using pH sensitive dyes that have different absorption properties depending on their pH. The change in pH can therefore be detected using optical components. The pH indicator may be immobilised within the sensor either onto the surface of the components or onto the textile substrate itself and the colour may be monitored using either a transmission or reflectance mode configuration, as shown in figure 2. LEDs have been chosen to perform this optical sensing as they are versatile components that have been demonstrated to operate as detectors as well as light sources. Operating LEDs as the light source and detector lends to a low-cost and low-power solution which is desirable for any wearable application.

Immobilisation of the dye onto the textile is the chosen approach, as this allows sweat samples to be delivered more easily to the textile sensor when the textile

is in direct contact with the skin. Moisture wicking textiles are used to aid this process. Bromothymol blue was chosen as a suitable indicator as this dye exhibits a colour change from yellow to blue between 6.0 and 7.6 pH units. The fabric was coated in solution using ethyl cellulose to bind the dye to the surface and tetraoctyl ammonium bromide was used as a fixative agent.

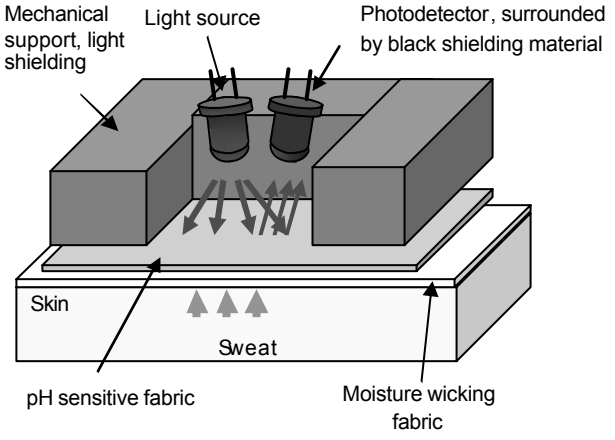


Figure 3 Textile pH colorimetric sensor with optical sensing configuration

To detect the colour change of the sensing textile a paired emitter-detector LED configuration was used. A red LED (660nm) was used as the light source (Kingbright, L934SRCG). This absorbance at this wavelength varies as the sensor changes from yellow to blue. An LED is also used as a light sensor. A red LED similar to the emitter was used for the detector, as it detects light at all wavelengths below its own spectral peak [21]. To function as a detector the LED is reverse biased at a specific voltage to generate photocurrent upon incident light. This photocurrent then discharges the LED at a rate that is proportional to the intensity of light reaching the detector. A simple threshold detection/timer circuit is used to indirectly measure the photocurrent at the detector LED to give digital output. This configuration is described in detail elsewhere [20]. For this application the driving electronics and data acquisition is carried out using Crossbow Mote platform. A Mica2dot mote is used to control the paired LED detectors and send the detected signal intensity to the basestation (Mica2) connected to a laptop.

The results are shown for the calibration of a fabric coated with bromothymol blue. The fabric was calibrated by being submerged in a range of buffer solutions. The dye exhibits a colour change (yellow to blue) in the region of pH 6.0 to pH 7.6. As the pH increases, more light from the emitting red LED is absorbed by the dyed fabric, therefore less light falls on the detector LED which takes a longer time to discharge. The detected light intensity is plotted against the pH value, figure 4a. A best-fit sigmoid curve was fitted to the data. The model for the sigmoid function used was:

$$I = \frac{a}{(1+e^{b(pH\ z)}} + c$$

Where I is the detected light intensity, a is the peak height, b is the slope coefficient, z is the point of inflection and c accounts for a baseline offset [22]. The point of inflection, the maximum of the first derivative of the sigmoid, is illustrated in figure 4b. The pKa for this sensor is estimated to be pH 6.6, slightly lower than the reported value of 7.1 [23]. The pKa shift is due to the dye being in solid form once immobilized onto the textile.

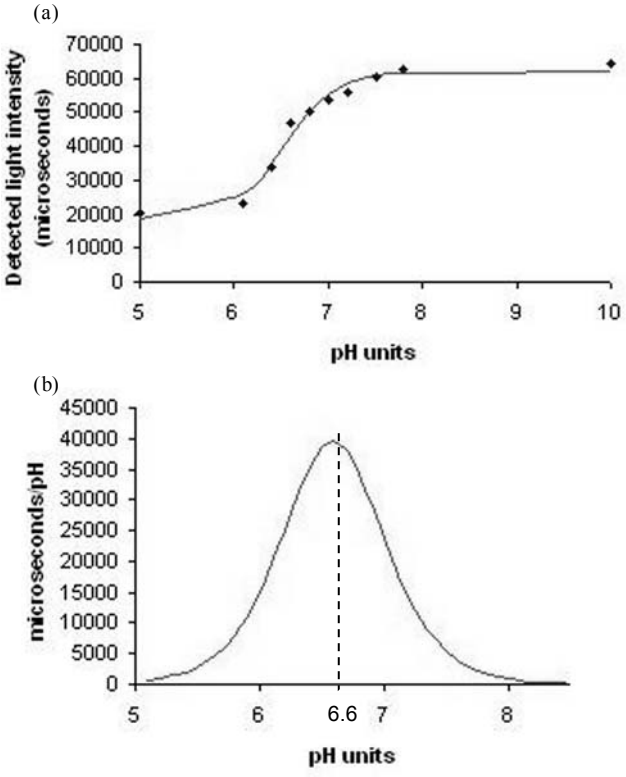


Figure 4 Characteristics of fabric pH sensor a) Detected light(μs) vs pH b) First derivative of detected light signal, to determine pKa

4. Fluid handling/calibration

Chemical sensors need to be frequently calibrated using known markers, and a possible approach to perform this within a textile is to use a conducting polymer, polypyrrole (PPy). This has been demonstrated to perform controlled release of chemical markers.. Sulforhodamine B (SB) was utilised as a model chemical marker and was incorporated either into a PPy coated hollow porous fibre [24] or doped into PPy film during synthesis. In both cases, small anions were used as dopant and the PPy backbone provides a positively charged structure to allow the exchange of negatively charged

anions. The rate of ion exchange can be substantially promoted by electrical stimulation by means of doping and de-doping processes. When the dye was stored in the hollow porous fibre, the release rate can be electrically enhanced by more than 7 times from 0.018 g / min to 0.15 g / min. The hollow fibre structure was able to hold larger amount of SB and the enhanced release can be electrically switched on and off. Continuously enhanced release at an average rate of 0.12 g / min has been demonstrated over 3 hours (Figure 5).

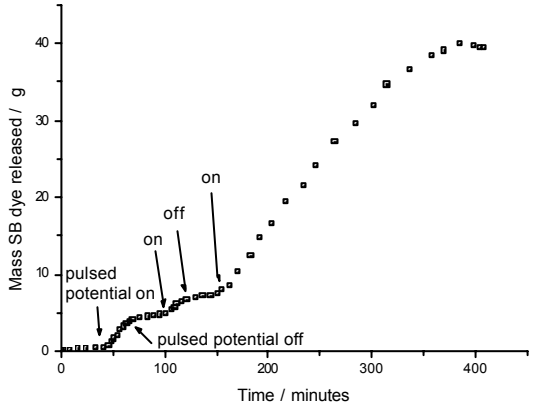


Figure 5. Controlled release pattern of anionic SB from a PPy coated hollow porous fibre in response to a pulsed electrical stimulation, the applied potential was switched between – 500 mV and + 600 mV (vs. Ag/AgCl) at 30 s intervals.

5. Discussion

It is widely accepted that sweat can serve as a useful diagnostic tool. The sweat test is used as the gold standard for CF diagnosis [25]. A wearable sensor to perform the sweat test by measuring sodium and chloride concentrations in situ has recently been demonstrated[26]. While this is a one-off diagnostic test, there are a number of potential applications where continuous sweat monitoring may aid in clinical assessment including remote patient monitoring and sports applications. In sports applications there is a great demand for wearable sensors for in-field testing which is used to check the athlete's health and also assess the effect of training programmes. Also, the amount of sweat lost in a training session gives an indication of fluid and electrolyte losses which is an important parameter to consider in fluid re-hydration strategies. The great advantage of analysing sweat is its non-invasive nature, and a wearable sensing platform would also provide a non-obtrusive means of performing measurements. However, sweat composition has not been widely studied given the impracticalities of collecting sweat. Current collection and measurement techniques include Minor's method, humidity sensors, direct observation of sweat drops using microscopes or CCD, "wash-down" techniques, parafilm patches for analysis at the end of trials, and more recently non-occlusive patches for drug testing [19, 27-29]. An important part of sensor integration into textiles is the development of methods to control fluid movement though the textile for sample delivery and calibration. This will allow real-time measurements in-vivo which is currently not possible with the methods mentioned above.

6. Conclusions

This manuscript aimed to raise the issues intrinsically associated with chemical sensors and exploit the possibility and potential benefits using functional materials particularly for wearable chemical sensors. This includes issues for measurement, fluid handling and calibration. The components considered includes the use of LEDs in colorimetric measurement, PPy micropumps and controlled release of chemical marker for calibration process. LEDs are low power components which is an important factor for wearable sensing. They are also very flexible in terms of their deployment within an optical sensing configuration which allowing them to be integrated in the most suitable fashion within a textile. Various arrangements may be investigated using transmission mode or reflectance mode operation. The use of surface mount devices and in the future organic LEDs may allow seamless integration within the textile. PPy based micro-pump and controlled release devices have advantages in mechanical flexibility, valve-less fluid handling and safe operation. This is a recently emerging, but promising area of research, and there is much work still in progress.

The concepts described in this paper follow the true impression of "smart fabrics" where the textile itself becomes the sensor. This extends the capabilities of the garment with wireless networks allowing discreet interaction with the wearer and environment.

Acknowledgements

We gratefully acknowledge the financial support of the European Union (Biotex FP6-2004-IST-NMP-2), Science Foundation Ireland (SFI 03/IN.3/I361) and the Australian Research Council.

References

[1] P. Bonato, "Advances in wearable technology and applications in physical medicine and rehabilitation," *Journal of NeuroEngineering and Rehabilitation*, vol. 2, pp. 2-5, 2005.

[2] R. Paradiso, G. Loriga, and N. Taccini, "Wearable System for Vital Signs Monitoring," *Studies in Health Technology and Informatics*, vol. 108, pp. 253-259, 2004.

[3] R. Paradiso, G. Loriga, and N. Taccini, "A Wearable Health Care System Based on Knitted Integrated Sensors," *IEEE Transactions on*

Information Technology in Biomedicine, vol. 9, pp. 337-344, 2005.

[4] P. Lukowicz, U. Anliker, J. Ward, G. Tröster, E. Hirt, and C. Neufelt, "AMON: A Wearable Medical Computer for High Risk Patients," presented at Proceedings of the 6th International Symposium on Wearable Computers, 7-10 October, 2002.

[5] S. Brady, L. Dunne, A. Lynch, B. Smyth, and D. Diamond, "Wearable Sensors? What is there to sense?," *Stud Health Technol Inform*, vol. 117, pp. 80-88, 2006.

[6] M. C. Carrozza, N. Croce, B. Magnani, and P. Dario, "Piezoelectric-driven stereolithography-fabricated micropump," *Journal of Micromechanics and Microengineering*, vol. 5, 1995.

[7] S. Bohm, W. Olthuis, and P. Bergveld, "A plastic micropump constructed with conventional techniques and materials," *Sensors and Actuators A: Physical*, vol. 77, 1999.

[8] A. Manz, E. Verpoorte, D. E. Raymond, C. S. Effenhauser, N. Burggraf, and H. M. Widmer, "Micro Total Analysis Systems," P. Bergveld, Ed.: Dordrecht: Kluwer Academic Publishers, 1995, pp. 5-27.

[9] F. C. M. Van de Pol, F. M. G. Van Lintel, M. Elwenspoek, and J. H. J. Fluitman, "A thermopneumatic micropump based on micro-engineering techniques," *Sensors and Actuators A: Physical*, vol. 21, pp. 198-202, 1990.

[10] M. C. Acero, J. A. Plaza, J. Esteve, M. Carmona, S. Marco, and J. Samitier, "Design of a modular micropump based on anodic bonding,"*Journal of Micromechanics and Microengineering*, vol. 7, pp. 179-182, 1997.

[11] A. Wego and L. Pagel, "A self-filling micropump based on PCB technology," *Sensors and Actuators A: Physical*, vol. 88, pp. 220-226, 2001.

[12] S. Bohm, B. Timmer, W. Olthuis, and P. Bergveld, "Micro Total Analysis Systems," A. van den Berg, Ed.: Dordrecht: Kluwer Academic Publishers, 2000, pp. 347-350.

[13] S. Bart, M. Mehregany, L. S. Tavrow, and J. H. Lang, "Microfabricated electrohydrodynamic pumps," presented at Transducers, Montreux, Switzerland, 1989.

[14] A. Richter and H. Sandmaier, "An electrohydrodynamic micropump," presented at Proceedings of the MEMS '90, 11-14 February 1990, Napa Valley, 1990.

[15] A. Richter, A. Plettner, K. A. Hoffmann, and H. Sandmaier, "Electrohydrodynamic pumping and flow measurement," presented at Proceedings of the MEMS '91, Nara, Japan, `.

[16] R. H. Baughman, I. W. Shacklette, and R. L. Elsenbaumer, "Microelectromechanical actuators based on conducting polymers," in *Topics in Molecular Organization and Engineering, Molecular Electronics*, vol. 7, P. I. Lazarev, Ed.: Kluwer, Dordrecht, 1991, pp. 267.

[17] J. D. Madden, R. A. Cush, T. S. Kanigan, and I. W. Hunter, "Fast contracting polypyrrole actuators," *Synthetic Metals*, vol. 113, pp. 185-192, 2000.

[18] S. Hara, T. Zama, W. Takashima, and K. Kaneto, "TFSI-doped polypyrrole actuator with 26% strain," *J. Mater. Chem*, vol. 14, pp. 1516-1517, 2004.

[19] Y. Wu, D. Zhou, G. M. Spinks, P. C. Innis, and G. G. Wallace, "TITAN: A Conducting Polymer Based Microfluidic Pump," *Accepted and to be published by Smart Marterials and Structures*, 2005.

[20] K. T. Lau, S. Baldwin, R. L. Shepherd, P. H. Dietz, W. S. Yerzunis, and D. Diamond, "Novel fused-LEDs devices as optical sensors for colorimetric analysis," *Talanta*, vol. 63, pp. 167-173, 2004.

[21] K. T. Lau, W. S. Yerzunis, R. L. Shepherd, and D. Diamond, "Quantitative colorimetric analysis of dye mixtures using an optical photometer based on LED array," *Sensors and Actuators B*, vol. 114, pp. 819-825, 2006.

[22] J. Frisby, "Monitoring food quality using sensor technology from harvest to home.," Dublin City University, 2005.

[23] F. J. Green, *The Sigma Aldrich Handbook of Stains, Dyes and Indicators*.

[24] Y. Wu, "Chapter 6: Controlled release using polypyrrole coated hollow fibre," in *PhD thesis, Department of Chemistry*. Wollongong: University of Wollongong, 2006.

[25] J. Massie, K. Gaskin, P. Van Asperen, and B. Wilcken, "Sweat testing following newborn screening for cystic fibrosis.," *Pediatric Pulmonology*, vol. 29, pp. 452-456, 2000.

[26] A. Lynch, D. Diamond, and M. Leader, "CF Diagnosis using a Potentiometric Ion-Selective Electrode Array," *Analyst*, vol. 125, pp. 2264-2267, 2000.

[27] S. M. Shirreffs and R. J. Maughan, "Whole body sweat collection in humans: an improved method with preliminary data on electrolyte content.," *Journal of Applied Physiology*, vol. 82, pp. 336-341, 1997.

[28] T. Ohhashi, M. Sakaguchi, and T. Tsuda, "Human Perspiration measurement," *Physiological Measurement*, vol. 19, pp. 449-461, 1998.

[29] M. Patterson, S. Galloway, and M. A. Nimmo, "Variations in regional sweat composition in normal human males," *Experimental Physiology*, vol. 85, pp. 869-876, 2000.

3rd Session
Posters

Wearable Joint-Angle Measurement with Modulated Magnetic Field from L/C Oscilators

Michael Barry[1] Agnes Grünerbl[1] and Paul Lukowicz [1,2]
[1]UMIT, Hall in Tirol, Austria
[2]University of Passau, Germany
Email: michael.barry@umit.at

January 10, 2007

Abstract

We demonstrate how modulated magnetic field generated by an simple LC oscilator can be used to measure joint angles, which are a key element in posture recognition and many motion analysis applications. Our method uses the same physical principle as large stationary motion trackers, however it applies the principle in a way suitable for a small, low power wearable system. It has the potential to be more accurate while being smaller and cheaper then inertial tracking (MARG) approaches that today are state of the art in wearable motion tracking. The paper describes the principle behind our method, discusses the advantages and problems and presents our prototype implementation. On a large data set with an hour of recording, hundreds of motions and tens of thousands of measurement-points we demonstrate, that even with an initial crude system implementation reasonable accuracy (between 6 and 9 percent average error) can be achieved.

1 Introduction

Posture and gesture tracking has a range of important applications such as activity recognition [5, 3], sports, wellness, arts (e.g.[7]), user interfaces (e.g. [4]) and rehabilitation. A good overview of tracking technologies can be found in [8]. This overview is titled "No silver bullet but an respectable arsenal". Our work proposes one more addition to this arsenal, that, while not a silver bullet itself, has a combination of properties so far not found in other systems.

In many areas tracking is needed in a mobile environment where no infrastructure other than small, energy efficient on-body sensors can be assumed. Thus conventional motion tracking systems such as for example the ascention "flock of birds" (http://www.ascension-tech.com/) system are not usable (although a wearable use of the ascention system placed in a large backpack has been proposed [1]). Instead the most common approach today is based on inertial tracking, often combined with magnetic field sensors to acquire absolute orientation (MARG for Magnetic Rotation Gravity). In addition to a large body of research work ranging from experimental systems (e.g. [2]) to complex so called inverse kinematic models (e.g. [7]) first commercial systems have also been introduced (http://www.xsense.com). An alternative method is the use of elongation or bend sensors integrated in the user's garment (e.g. [6]). We propose a novel method based on magnetic coupling between two oscillating circuits. The coupling allows short-range distance-measurement between body parts from which the corresponding joint-angles can be inferred. The technique uses the same physical principle that is found in stationary motion capture systems such as the ascencion 'flock of birds' device and emerging PAN (personal area network) technology such as the auracom (www.auracomm.com) system.

Contributions with Respect to Related Work The main contribution of our work is to demonstrate how this well-known physical principle can be applied in a novel way. Instead of exact 6 DOF (degrees of freedom) absolute tracking of body parts over a distance of a few meters provided by conventional stationary magnetic motion trackers we restrict ourselves to relative distance measurement between body parts. This allows us to use small, low power sensor nodes that can be easily attached to to user's outfit. At the same time the joint angle information that our system

provides is sufficient for a wide range of applications and comparable with the type of information provided by other wearable systems.

Compared with such systems the proposed method has the following advantages:

1. MARG systems rely on earth magnetic field for absolute orientation. Thus they are easily disturbed by ambient magnetic fields (e.g. from AC currents). In addition they use double integration over acceleration and rotation to arrive at position. Thus errors accumulate rapidly leading to considerable accuracy problems. Although statistical filtering can help reduce such errors, the reliability of MARG systems is at best mixed. In contrast, as will be explained later, the proposed system is much less sensitive to disturbances and does not have the error accumulation problem. Finally MARG systems require three types of sensors that are more expensive, larger and consume more power then a simple L/C oscilator.

2. Elongation sensors in general need to cover a large area of the body and really only make sense if they can be fully integrated into the garment. This may or may not be possible. In addition such sensors often have problems with accuracy, hysteresis and sampling rate.

Note that our method can not only replace but also supplement the above approaches leading to increase accuracy and reliability.

In the rest of the paper we first describe the physical principle behind our system and the design considerations that follow from the principle. We then present our implementation and describe results of an evaluation of this implementation on a large data set. The data set contains an hour of recording, hundreds of motions and tens of thousands of measurement-points. We demonstrate, that even with an initial crude system implementation reasonable accuracy (between 6 and 9 percent average error) can be achieved.

2 Measurement Principle

Physical Principle A circuit consisting of a capacitor and an inductor (coil) constitutes an electrical oscilator. When energy is injected into such a circuit (e.g. by a short electrical pulse) it periodically flows between the elctric field of the capacitor and the magnetic field of the coil. The period is given by the resonant frequency of the circuit which in turn is determined by the choice of the capacitor and the inductor In an ideal lossless circuit the oscilation would continue forever. In reality the oscilation is damped and can only be kept going on by re-injecting energy into the circuit with a periodic pulse tuned to the resonant frequency.

Whilst the electric field is constrained to the inside of the capacitor, the oscilation of the magnetic field of the coil is measurable outside the circuit. If another circuit with an identical resonant frequency (receiver circuit) is placed in this field, the magnetic field injects energy into the circuit because of Farradays Law and an oscillating voltage is induced. This voltage is determined by the strength of the magnetic field, which in turn depends on the distance between the two circuits. Thus the induced voltage can be used for a distance measurement, which is how the proposed system works.

An important thing to note above the above principle is the difference to a radio transmission based system (RF System). If an oscilator circuit is connected to an antenna then it transmits energy in form of electromagnetic waves, which is the principle behind all radio communication system. In our system there is no antenna and thus no RF transmission [1]. The magnetic field build up by the coils is fully 'retracted' as the energy flows back into the electric field of the capacitor during the next oscillation cycle. Apart from the damping losses, the only energy that is extracted from the system is the energy needed to induce the voltage in the receiver's coil.

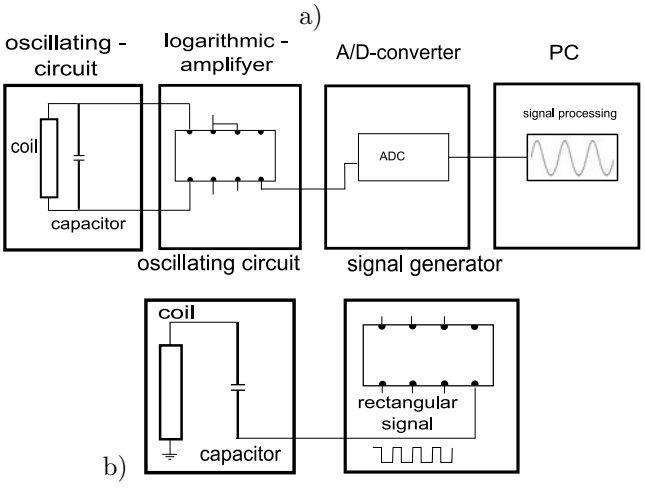


Figure 1: Receiver (a) and Transmitter (b) circuits.

[1] In reality cables and data lines act as parasitic antennas and some energy is indeed radiated.

Design Considerations In terms of suitability for a wearable posture tracking system the above principle has a number of advantages

1. Both the transmitter and the receiver circuits are very simple: essentially a coil, a capacitor, a pulse generator for the transmitter and an amplifier for the receiver. Thus they can be build to be small and inexpensive.

2. As described above, our system does not radiate energy and (except for in-circuit losses) only the energy actually used to generate the signal in the receiver is 'lost' to the transmitter. This means that even if fairly strong fields are used to achieve good signal quality or larger range, the system is energy efficient.

3. The average power densitiy of the magnetic field P_{mag} decreases with the sixth power of the distance d. This may seam to be a disadvantage as the range of the system is limited (around 1m the type of small coils used in our experiments). However for on-body distance measurement this is not a problem. At the same time this means that reflections and multi-path propagation that causes problems in RF and ultrasonic distance measurements is not an issue.

4. Only a magnetic field oscilating with the corresponding resonant frequency will have any measurable effect on the receiver circuit. Thus the system is largely immune to interference from environmental magnetic fields (e.g. from cables in the wall). This is in sharp contrast with systems that use magnetic sensors to determine the absolute orientation of the limbs using a combination of earth magnetic field and inertial tracking.

5. Short of large, massive metallic objects magnetic field modulated with the frequencies used in our experiment is fairly immune to obstructions. In particular the human body has little influence on the signal. Note that since we look at short on body distances large metallic objects are much less of an issue then for stationary magnetic tracking systems such as the flock of birds.

The main difficulties that a practical implementation of the system needs to address are the following:

1. The form of the magnetic field depends on the shape of the coil, but is in general spatially highly inhomogeneous. This means that the intensity of the signal depends not only on the distance but also on the relative orientation of the receiver and transmitter coils. However as long as we are looking at joints with only one degree of freedom and devices firmly fixed to the adjoining body parts this only means that the field intensity relationship needs to be trained rather then derived from the simple $1/d^6$ relationship (see below).

2. The $1/d^6$ decrease of the power density means that there is a large absolute difference between the signal from the minimum and maximum angle. This means that the Amplifier and the A/D converter need to have a large dynamic range. Since sensor nodes typically use on-chip A/D converters integrated on microcontrollers such a dynamic range is not available. To solve this problem our system uses a logarithmic amplifier.

3. The $1/d^6$ decrease of the power density also means that we need to use a very sensitive amplifier. This in turn means that it will pick up noise from the environment unless the circuit is very well designed and shielded. The lack of such shielding (which is a costly and time consuming issue that makes little sense for an initial investigation) is one of the major sources of error in our prototype and will be addressed in the next generation system.

Prototype Implementation The transmitter of our prototype is an oscillating circuit consisting of a capacitor of 300 pF and a coil of 68 mH. The driving pulse for the oscilator of the transmitter is generated by a commercial wide-spread 555-timer-chip. The receiver consists of a capacitor of 330 nF and a coil of 25 μH (figure 2). The resonant frequency of both circuits is 34.35 khz [2]. The voltage induced in this oscillating circuit by the transmitter is measured using a commercially available logarithmic Amplifier [**?**]. The Output of this AD8307-chip is then sampled by a 12-bit-ADC attached to a Personal Computer.

Using the System for Posture Tracking The basic idea behind the system is to infer joint-angles from distance-measurements. In our experiments we use the distance between the upper and lower arm (leg) to infer elbow (knee) angles. In principle if the

[2]The reason for using different setups for the receiver and the transmitter is technical and has to do with 1) the fact that a sequential hookup of the coil and the capacitor is better for the transmitter while a parallel one is better for the receiver and 2) the quality of a sequential oscilator depends on the capacitance and the inductance in a different way (the inverse of the L/C constant) than that of a parallel one.

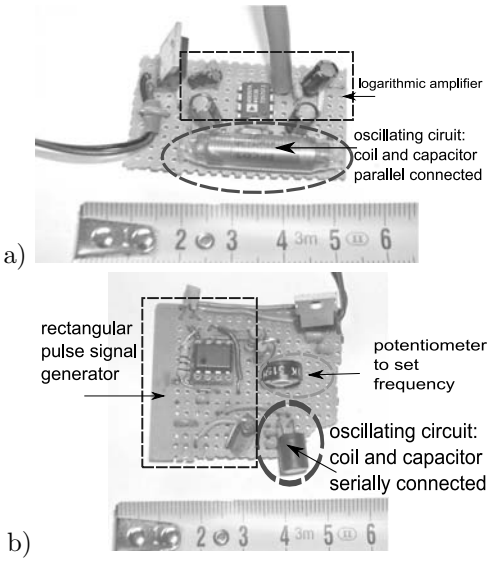


Figure 2: Receiver (a) and Transmitter (b).

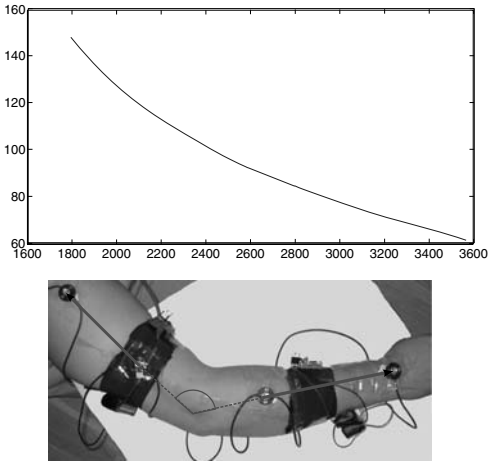


Figure 3: Dependence of joint-angle on ADC-value (top) and the measurement setup for the elbow angle (bottom).

distance of the transmitter and the receiver from the joint is known then the angle can be computed from the measured distance using the cosine law (figure 3 bottom). However due to the inhomogeneity of the magnetic field and the fact that the orientation of the coils with respect to each other changes as the joint rotates, such a straight forward calculation is not practicable. Instead to be able to convert the ADC-readings to the corresponding angle, piecwise linear approximation was used on a set of training data (see figure (figure 3 top). To improve the measurement's quality the raw data is smoothened using a Moving-Average-Filter.

3 System Evaluation

To evaluate our concepts the elbow and the knee joint-angles were measured for two subjects (one male, one female). For each subject and limb about 16 min of data from bending-motions performed at different speeds were collected. The speeds were fast (about 1 sec per full bend), medium (few sec per full bend), slow (more then ten sec per full bend) and paused when just few bends with longer (several seconds) breaks in between were performed (see figure 4). The speed of bending was not used in any quantitative analysis, we just wanted to ensure a certain variation, so that the qualitative instructions given to the subject on speed were sufficient. Altogether over one hour of data with several hundreds individual bend motions were recorded. Depending on the size of the smoothing window this resulted in some 30'000 to 300'000 individual angle measurements.

To get a reliable ground truth, the position of 4 markers, 2 on the lower and 2 on the upper arm was determined using a commercial, stationary optical motion tracking system system (www.lucotronic.at).

3.1 Results

The results of the above experiment are summarized in table 3. It shows the average error, the error standard deviation the error 90 and 95 percentile (values below which lie 90 or 95 percent of the errors) for different window-sizes for just the arm, the leg and both together. . Depending on the window size the average error for the arm is between 6% and 7%, for the leg between 6% and 9%. This in itself is reasonable for a number of applications such as activity recognition or non professional sports trainers. This is even more so as, unlike in systems that use inertial tracking, the error does not accumulate over time.

At this stage it is important to note, that the above experiments were undertaken to establish the general viability of the proposed method. As a consequence we have been working with a fairly crude initial ("breadboard" see figure 2) prototype of the hardware and a simple velcro band put over the user's trousers for sensor attachment. In addition no other signal-processing than the sliding window smoothing was undertaken. Thus the results presented above must be seen as an indication of the potential of the proposed technique rather then a performance limit.

Windowsize:	5	11	21	25	35	51	75
leg							
average error	9,80	9,61	9,06	8,79	8,02	6,96	6,22
error std deviation	7,81	7,53	6,97	6,72	6,06	5,00	4,23
95 percentile	22,97	22,53	21,14	20,46	18,48	15,59	13,44
90 percentile	19,34	19,01	17,93	17,40	15,83	13,30	11,72
arm							
average error	7,26	7,15	6,94	6,86	6,66	6,42	6,57
error std deviation	6,57	6,25	5,70	5,48	5,01	4,64	4,16
95 percentile	18,96	18,47	17,24	16,67	15,36	14,62	13,97
90 percentile	14,04	13,68	12,86	12,49	11,89	11,46	11,50
overall							
average error	8,53	8,38	8,01	7,83	7,34	6,69	6,4
error std deviation	7,19	6,89	6,34	6,10	5,54	4,87	4,20
95 percentile	20,97	20,51	19,20	18,57	16,92	15,11	13,71
90 percentile	16,69	16,34	15,40	14,95	13,87	12,38	11,61

Table 1: Evaluation results for different smoothing window sizes for the arm, leg and both combined

Error Discussion The above becomes more evident after a closer look at the errors. In figure 4 an example of an excellent, good, bad and a very bad measurement is shown. Each is a single 2 min sequence where a specific user was performing arm(leg) bends with one of the 4 possible speeds. The measurements have an average error of 1.43%, 2.83%, 4.6% and 15,7% and a 95 percentile of 3.55%, 5.41%, 11.3% and 35.4%. It can be seen that even in the bad measurements the shape of the measured signals follows the ground truth pretty well. The problem is, that the measured signal is scaled and shifted by a reasonably constant factor. This can be traced to the sensors being mechanically shifted, due to the lack of robust attachment (which has been observed during the experiment). Obvious geometric considerations show that if the distance between the transmitter/receiver and the joint changes, then the same angle will produce a different distance. In addition if the coils are rotated with respect to each other the inhomogeneity of the magnetic field will lead to deviations from the trained values. If we remove the data sets in which the transmitter and/or reiever was shifted during the experiments (6 out of 32 data sets) then the overall results (using a smoothing window size of 10 samples which is appropriate for all speeds) improves to an average error of 4.9 for the arm, 8.4 for the leg and 6.7 overall with a 95 percentile of 15.5, 16.4 and 16.5 respectively.

4 Conclusion and Future Work

The main conclusion from this work is that joint angle measurements with L/C oscilator generated magnetic field have considerable potential as wearable posture, gesture, and motion capture sensors. We believe that in particularly applications such as activity recognition, sports training and rehabilitation could greatly benefit from this technology.

While we believe that in many applications our method has clear advantages over inertial tracking or bend sensors (see section 1) it should be noted that it can also easily be combined with those modalities for an even higher accuracy.

Clearly this paper constitutes only an initial investigation. Based on the experience from this investigation we are currently working on the following issues:

1. A more advanced implementation of the hardware on a carefully designed PCB with a better driving circuit and robust shielding against external interference.

2. Robust attachment methods as well as methods for dynamic re-calibration of the system in case of shifting sensors.

3. Using more complex signal processing methods (e.g Kalman filters) to further improve accuracy.

4. Enhancing the functionality by using two or three perpendicular, coils transmitting in a time multiplexed mode to get not just scalar distance but also orientation information. Such a system will

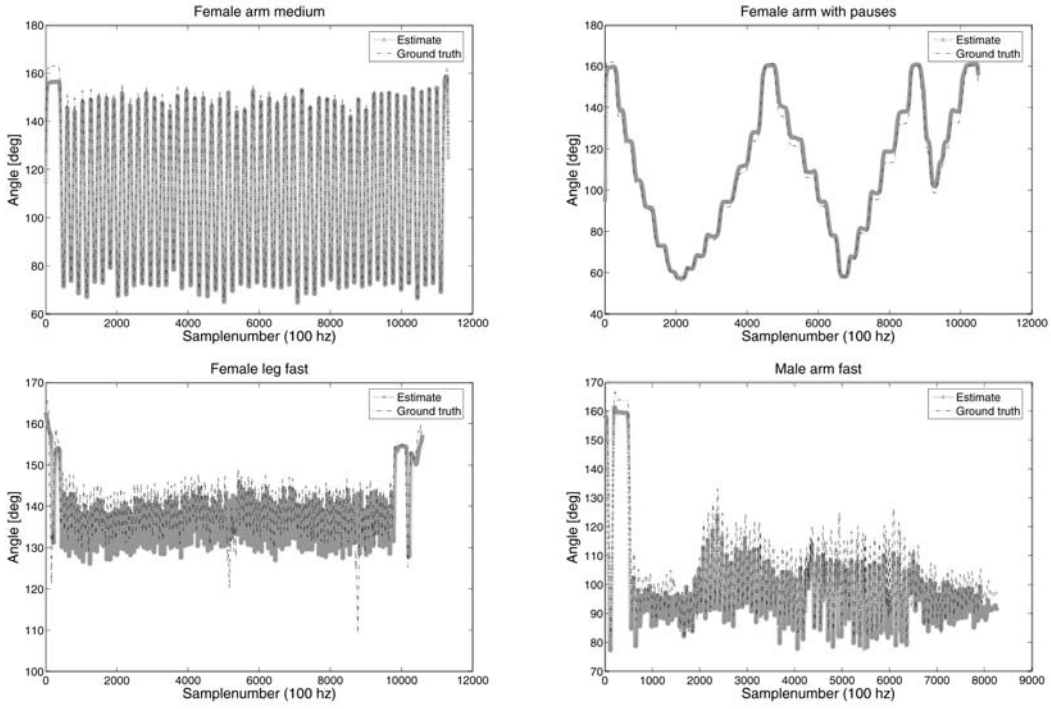


Figure 4: Examples of measurements. Top left: excellent from medium motion of female arm; Top right: good from paused motion of female arm; Bottom left: bad from fast motion of female arm; Bottom right: very bad from fast motion of male arm.

enable us to work with joints that have more than 1 degree of freedom.

References

[1] Masayuki Kanbara Akihiro Hamaguchi and Naokazu Yokoya. User localization using wearable electromagnetic tracker and orientation sensor. *Proc. Tenth Internartional Symposium on Wearable Computers ISWC 2006, Montreux, Switzerland*, 00, 2006.

[2] Eric R. Bachmann, Robert B. McGhee, Xiaoping Yun, and Michael J. Zyda. Inertial and magnetic posture tracking for inserting humans into networked virtual environments. In *VRST '01: Proceedings of the ACM symposium on Virtual reality software and technology*, pages 9–16, New York, NY, USA, 2001. ACM Press.

[3] Jonny Farringdon, Andrew J. Moore, Nancy Tilbury, James Church, and Pieter D. Biemond. Wearable sensor badge and sensor jacket for context awareness. *Proc. Internartional Symposium on Wearable Computers ISWC 99*, 00:107, 1999.

[4] Juha Kela, Panu Korpip, Jani Mantyjarvi, Sanna Kallio, Giuseppe Savino, Luca Jozzo, and Di Marca. Accelerometer-based gesture control for a design environment. *Personal Ubiquitous Comput.*, 10(5):285–299, 2006.

[5] Paul Lukowicz, Jamie A. Ward, Holger Junker, Mathias Stäger, Gerhard Tröster, Amin Atrash, and Thad Starner. Recognizing workshop activity using body worn microphones and accelerometers. In *Pervasive Computing: Proceedings of the 2nd International Conference*, pages 18–22. Springer-Verlag Heidelberg: Lecture Notes in Computer Science, April 2004.

[6] Corinne Mattmann, Tünde Kirstein, and Gerhard Tröster. A method to measure elongations of clothing. In *Proc. 1st International Scientific Conference Ambience 05*, September 2005.

[7] Matja; Mihelj. Inverse kinematics of human arm based on multisensor data integration. *J. Intell. Robotics Syst.*, 47(2):139–153, 2006.

[8] Greg Welch and Eric Foxlin. Motion tracking: No silver bullet, but a respectable arsenal. *IEEE Comput. Graph. Appl.*, 22(6):24–38, 2002.

In-body Wireless Communication Made Real.

H Higgins
Zarlink Semiconductor, Caldicot, NP26 5YW, UK.

Abstract—***This paper take the principle of wireless communication with body implants to the implementation and testing stage. This paper describes a module that can fit into an implant case and results of testing a wireless link with a body model.***

Keywords—***Clinical applications, low-power communication, miniaturisation, wireless.***

I. Introduction

Communication with implanted medical devices is key to effective diagnosis and therapy. There are several critical requirements for an implanted communication system, including size, power consumption and data rate. At BSN 2005 a paper was presented [1] that described the essential features of a communication system. More details are contained in Body Sensor Networks edited by Guang-Zhong Yang [2]. This paper shows the size of real hardware and presents the results of tests using a human body substitute.

The system uses an integrated circuit (IC) that incorporates a media access controller (MAC), a 2.45GHz wake-up function and 403MHz two-way radio frequency (RF) data communication. The 403MHz frequency used for in-body communication is known as the medical implant communication system (MICS) [3].

II. Hardware

The implant comprises the therapy or diagnostic function and the communication module. Small size is essential, as there is very little room to fit implants into the human body and some locations pose severe challenges. The hardware needs to be robust to withstand the day-to-day knocks and bumps of human body movement and must also survive transport by land, sea and air.

A wireless communication module can be built using a printed circuit board (PCB), or a ceramic thick film hybrid, as the substrate with small active and passive components attached with solder or conductive epoxy. Large active components, with many connections, are typically fixed to the PCB with epoxy and the connections made with bond wire to the substrate. See Figure 1. To protect the wire bonds a non-conductive encapsulant ("glob") can be added.

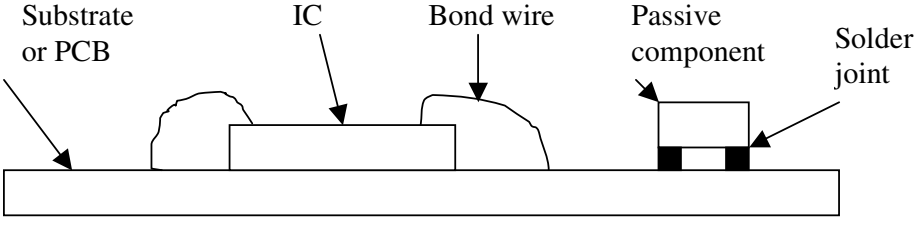


Fig. 1 Module Assembly

The wireless communication module is electrically connected to the other functions of the implant and is part of the same enclosure. Connection to the other parts of the implant can be made with a flexible PCB that is electrically connected to the parent board. This will allow some relative movement between the two PCBs. An alternative approach is to fit solder bumps to the communication PCB such that it can be soldered down to the parent board. The antenna connection must be very short to minimise losses and to keep the impedance at the implant constant.

Figure 2 shows a typical implant board; the communication IC covered in glob, the matching network is to the right and the antenna connections are the two pads in from the edge of the board. Also on the module is the crystal, on the left, that provides the reference for the communication and data handling. In this example the IC is the Zarlink ZL70101. The communication IC has all the functions necessary to establish a wireless link. It is typically controlled by a microprocessor on the parent PCB.

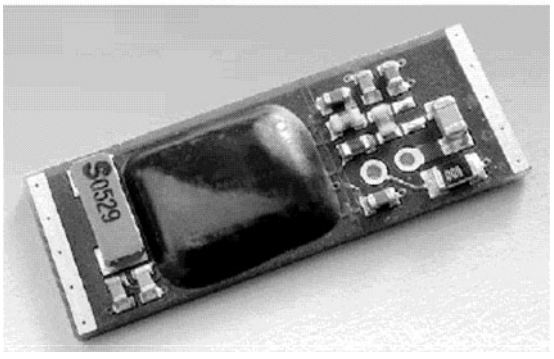


Fig. 2 Implantable RF Communication Module

III. Testing

To test the performance of the communication link a reference module was powered by a battery. The communication IC (ZL70101) was controlled by a dedicated microprocessor to enable the RF functions and data transfer. The ZL70101 and microprocessor were mounted in a metal enclosure with a 19.5 x 32 mm patch antenna attached; see Figure 3. This was either sealed to make it water tight or enclosed in a thin nonconductive latex bag.

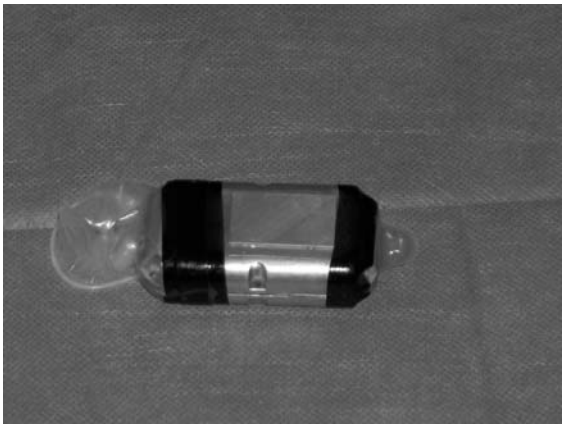


Fig. 3 Implant Case With Patch Antenna Enclosed in Latex Bag

The human body model is defined by ETSI [4] and is made of 10mm thick Perspex. The Perspex model is filled with a liquid that mimics the electrical properties of the human body. For these tests the liquid is valid up to 1GHz, and the recipe is defined by Wojclk [5]. For different frequencies other recipes are available. The implant is suspended in the liquid attached to a PTFE holder. PTFE is chosen, as it will cause the least interference to the radiated signal.

The base-station is controlled by a laptop PC and will generate the signals needed to communicate with the implant. The base-station generates a “wake-up” signal centred at 2.45GHz to get the implant powered up and ready to communicate. The base-station also transmits and receives signals within the MICS band. The wake-up control and communication uses the same IC as the implant.

The power limit for the wake-up transmitter is typically 100mW (country dependant). The power limit for the MICS transmission is 25µW (-16dBm) effective radiated power – this takes account of antenna gain. This 25µW applies to the implant but only at the skin surface. Even if the raw power produced by the communication IC is in the order of 1mW (0dBm), losses through the body will typically reduce the power level to well below the 25µW limit.

The test environment was the anechoic chamber at the University of Bristol. This comprised a screened room that has absorbent cones on the inside to minimise any reflections from walls or the floor that could distort the results. In a real life environment there will be reflections from walls, desks and other equipment and hardware. The body model was mounted on a wood stand (non-conductive). The MICS base-station dipole antenna was mounted on a PTFE stand. The wake-up antenna was on a length of cable to provide freedom of movement. See Figure 4.

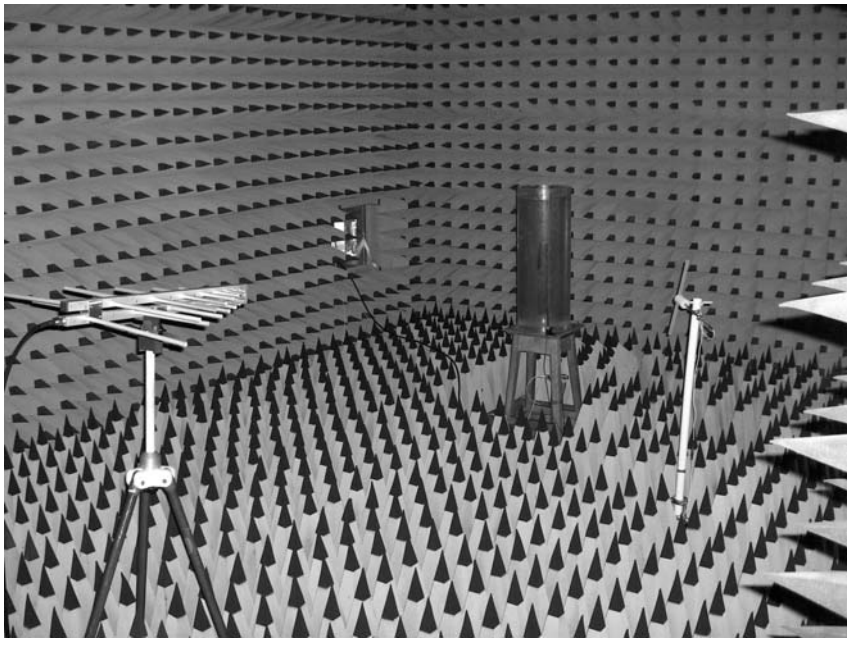


Fig. 4 Anechoic Chamber Showing Body Model (on Wood Stand), Log Periodic Test Antenna (Foreground) and Base-station Dipole (Right)

IV. Results

Once a communication session is established it is possible to perform measurements on the link. These include measurement of effective radiated power from the implant, a measure of the received signal at the implant from the base-station and a measure of the link quality. It is possible to set the implant to transmit a continuous wave (CW) signal so the effective radiated power can be measured. Also the base-station can be set to transmit a CW signal and the implant can measure the relative signal level. These measurements were performed at various depths of the implant in the liquid.

The distance between the implant and either the test antenna or base-station was 3m. The test antenna was a broadband directional log periodic with known characteristics that are necessary to calculate the power radiated from the body model. Depth is the distance from the front of the tank and the implant antenna. Vertical polarisation of the implant is when the long edge of the patch antenna is vertical.

IV. A. Implant Transmitted Power vs. Depth.

Figure 5 shows the effective radiated power (ERP), at the tank surface, versus depth. The ERP is calculated from the measured signal and the test antenna characteristics. As seen, the signal level is dependent on polarisation, and apart from when both implant and test antenna were vertical the power declines with depth.

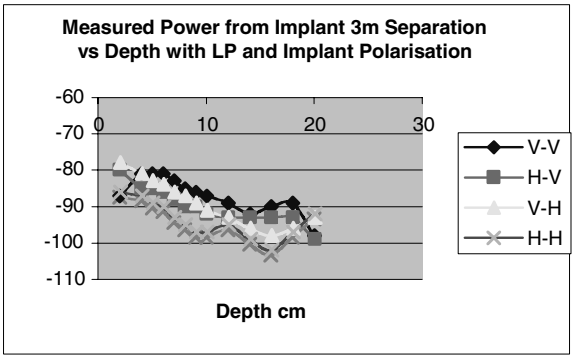


Fig. 5 Effective Radiated Power from the Implant vs. Depth. The Key on the Right Refers to the Test Antenna and Implant Polarisation Respectively

IV. B. Power Received by the Implant vs. Depth.

For this test only the base-station was use, not the antenna. The base-station was set to transmit CW for approximately 30 seconds during which time the implant performed a signal level measurement. This measurement was then transmitted to the base-station. The power at the implant is measured with an internal received signal strength indication (RSSI) function that will give a relative measure of the signal at antenna. Figure 6 shows the RSSI vs. depth for the implant. It can be seen that there is a peak in signal level at about 3-4cm. This is close to the peak in signal level seen in Figure 5.

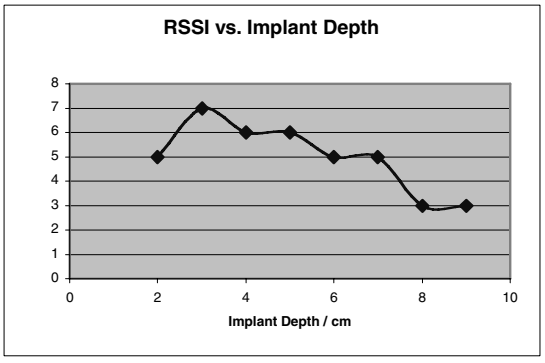


Fig. 6 Implant RSSI vs. Implant Depth

IV. C. Data Transmission.

A link had to be maintained for the above tests or the tests would not have been possible. Signal level is not meaningful unless it can be related to the transfer of data. One way to measure the link is to plot the number of times the error correction code (ECC) or cyclic redundancy code (CRC) needs to be invoked to produce 100 good blocks of data. This test used the base-station and the implant at various depths. The lower the count the better the link is. Figure 7 shows a typical plot.

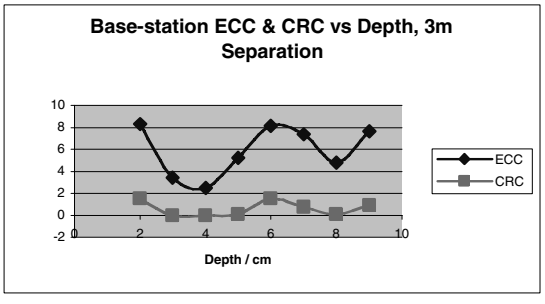


Fig. 7 ECC and CRC Count vs. Implant Depth for 100 Message Transactions

Figure 8 shows that the best performance is obtained when the implant is at 3-4 cm depth that corresponds with the results of the implant power and RSSI results.

V. Conclusions

The results here are measured and not simulated. The body model provides a reasonable representation of the human body. A data link can be maintained at an implant depth of over 15cm. The link is maintained with a base-station transmit power within the MICS limit. At an optimum depth the body can enhance the signal and thereby the data link, meaning communication with an implant is practical over a distance of 3m with a high data rate.

References

1. Higgins H, Radio Frequency Technology and In-Body Communications Systems, BSN 2005 Conference Proceedings.
2. Yang G-Z, Body Sensor Networks, Springer 2006, pp 117-143.
3. MICS Band, Australian Communications Agency, Paper SP 6/03, 2003.
4. ETSI EN 301 839-1 v1.1.1, 2002-4. “Electromagnetic Compatibility and Radio Spectrum Matters (ERM); Radio Equipment In the Frequency Range 402 to 405MHz for Ultra Low Power Medical Implants and Accessories; Part 1; Technical Characteristics, Including Electromagnetic Compatibility Requirements and Test Methods”. Page 42.
5. Wojclk J et.al.; “Tissue Recipe Calibration Requirements, SSI/DRB-TP-D01-003”. Spectrum Sciences Institute RF Dosemetry Research Board, 51 Spectrum Way, Nepean, Ontario, K2R 1E6, Canada.

Author Contact Details.

Henry Higgins
Zarlink Semiconductor
Castlegate Business Park
Portskewett
Caldicot
Monmouthshire
NP26 5YW
UK

henry.higgins@zarlink.com

The PSI Board: Realizing a Phone-Centric Body Sensor Network

Trevor Pering[1], Pei Zhang[2], Rohit Chaudhri[3], Yaw Anokwa[4], and Roy Want[1]

[1] Ubiquity Group, Intel Research, Santa Clara, CA, USA
[2] Princeton University, Princeton, NJ, USA
[3] Motorola Labs, Schaumburg, IL, USA
[4] University of Washington, Seattle, WA, USA

Abstract—**When designing a body sensor network, the mobile phone is a natural design point for data aggregation services. However, there is a missing bridge that allows sensors to communicate with existing commercial phones, even if they already possess sufficient storage and processing capabilities. The PSI board is a small expansion module that interfaces body sensor networks with commercially available cell-phones through a standard MMC/SD slot. Adding this expansion capability allows researchers to extend phone devices with additional capabilities, allowing the devices to easily serve as the hub of a wearable body sensor network. The PSI board has an integrated switching mechanism allowing transparent access to an MMC/SD card, an embedded microcontroller, accelerometer, expansion connector, and an IEEE 802.15.4 radio that can connect to a variety of commonly available wireless sensors. The capabilities afforded by the PSI board enable several new applications that would be difficult to implement given the sensors and capabilities currently found in existing cell-phone platforms.**

Keywords— **Body sensor networks, mobile phones, MMD/SD cards, IEEE 802.15.4, low-power systems.**

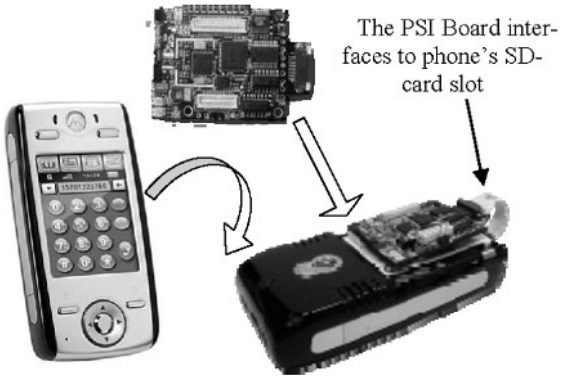


Fig 1: The PSI board attaches to the back of commercially available cell-phones, creating an integrated sensing platform. For deployment the battery panel can be replaced by a module that integrates with the PSI board. .

I. Introduction

Mobile phones provide researchers with a new opportunity to provide computation and storage capabilities to wearable body sensor networks without adding significant amounts of specialized hardware. In the context of ambulatory medical monitoring, reducing the number of devices being carried and reusing devices that a patient is already motivated to carry for personal communication is a considerable advantage. However, phones are typically closed systems from a hardware perspective and do not have much in the way of general purpose expansion capabilities, which might impact the streamlined design necessary for commercial success. Because of this, it is often difficult to experiment with phones when new peripherals/sensors are needed to be integrated in the system.

The Phone System Interface (PSI) module is a bridge between wireless sensor networks and commercially available mobile phones. The PSI module allows a phone platform to be extended to perform integrated sensing (sensors combined with the phone), or to wirelessly communicate with a body area (sensor) network.

The module is a small board designed to interface to commercially-available cell-phone platforms through the phone's MMC/SD-card socket, as shown in Figure 1. An MMC/SD-interface is supported for the phone connection, along with an MSP430 embedded microcontroller, a 3-axis accelerometer, and an integrated CC2420 802.15.4 radio. Additionally, there are expansion connectors that can be used for a variety of purposes, such as adding Near Field Communication (NFC) capability. The prototype system combines the PSI board with Motorola's ROKR E2 and E680i/g mobile phone, both running the Linux operating system. Programmatic access to the PSI board is provided through custom kernel drivers.

One important consideration used when designing the PSI module was its impact the phone experience. Since the phone is a highly personal device, users have high expectations and will not be willing to adopt the technology if the module interferes with their use of the device as a phone. The three main considerations here are battery life, phone operation, and physical size. By relegating many of the core sensing functions to a separate low-power embedded processor sub-system, the PSI board is able to manage sensing with minimal impact on the device's power consumption,

while only adding a small amount of physical size/weight to the phone body, and not directly impacting the phone's functionality.

This paper describes the architecture of PSI board, details several of the subsystems critical to system operation, and provides an overview of three applications models supported by PSI board capabilities. The design and implementation of the PSI board shows that it is possible to extend some commercial cell-phone platforms to create wearable sensing systems, and highlights several design challenges for such systems.

II. RELATED WORK

The PSI board touches upon two major bodies of related work: the phone platform and wearable sensor networks.

The Green Phone [4] is an open-source phone platform that has been recently released and allows developers to update and download any software they choose to the platform. This is a new capability because typical phone platforms do not allow changes to the core software. The Green phone, however, still does not enable easy hardware modifications that would be necessary to enable such technologies as IEEE 802.15.4 or adding accelerometers or other sensors. So while the open *software* capability of the device enables new kinds of experimentation for mobile platforms, it still is closed to hardware expansion.

iStuff Mobile [2] is a prototyping environment that allows designers to easily add sensors and other capabilities to mobile phones. Unlike the PSI board, this system communicates all of the data captured from sensors to the nearby infrastructure where it is processed. So, although it works very well for prototyping applications in controlled environments and exploring new areas of interaction, it is not intended as a viable long-term solution for deployments of mobile phone applications.

The CodeBlue project [7] is just one example of a wearable sensor system built around several different mote platforms, several of which support the 802.15.4 standard. A dedicated bridge [6] between a custom wireless sensor network and GSM system demonstrates the feasibility and attractiveness of the basic concept, but does not demonstrate integration with a phone, instead requiring the user to carry and manage a separate unit. Integrating sensors using Bluetooth [1] would allow using commercially available phone devices, but this architecture does not integrate very well with common sensor platforms and would place an increased power burden on the mobile device.

IEEE 802.15.4, and the closely related Zigbee standard, are specifically designed to provide lower-power operation for sensor networks and are the radio of choice for the majority of such projects. Using Bluetooth in a sensor network capacity is likely to have an adverse power impact on the sensing nodes. Furthermore, tightly integrating the sensor communication with the existing phone Bluetooth subsystem requires that the phone be active during all communication, which will significantly drain the phone's battery. Efforts are being made in the industry to address Bluetooth's power impact on mobile devices with a new standard called Wibree; however, since Bluetooth is inherently a point-to-point or Master-Slave protocol, it complicates some sensor network architectures where the sensors nodes need to communicate to peers in addition to communicating with the Master node. 802.15.4 does not impose this limitation and so is better suited for such networks.

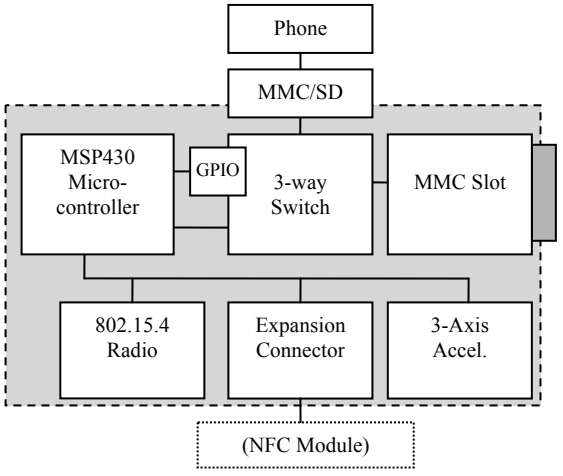


Fig 2: PSI Block Diagram. The basic module connects to a phone through its MMC/SD-card slot.

III. ARCHITECTURE

The basic PSI-board architecture connects together three major system components: host phone, MMC/SD storage card, and sensor bridge. One key contribution of the PSI system is an understanding of the signaling and switching necessary to connect the PSI board to the host phone platform, without effecting basic phone operation.

A. Host Platform

The PSI system is designed to work with a series of commercially available Linux-based cell-phone platforms, primarily the Motorola E680i/g and ROKR E2 phones. These phones are based on the Intel/Marvel PXA-27x family of ARM-architecture processors, running at speeds of up to 400 MHz. Internal to the phones there are 32 MB of DRAM

and 32 MB of FLASH. MMC/SD expansion cards are also available in sizes of at least 2 GB. The phones run an embedded version of the Linux 2.4 operating system.

The PSI board uses the same hardware interface as a MMC/SD card. The Linux PSI interface is implemented as a dynamically loadable kernel module that replaces the default MMC/SD kernel module. Custom modifications were required to the phone startup sequence to prevent the standard drivers from loading.

The kernel PSI driver appears to user space applications as a special interface that allows basic communication between the user programs and the PSI firmware. The basic abstraction is a channel metaphor between virtual software channels and the individual hardware resources (e.g., radio, accelerometer, and expansion connector). A dedicated software daemon, PSID, runs on the phone and provides a socket interface for other programs. Many of the primary prototype applications are written as Java Midlets, using a standard socket connection to the PSID.

B. Mechanical Attachment

The basic dimensions of the PSI board are 47x35x8 mm, while the battery for both the E680i and ROKR E2 measure around 55x35x5 mm (with a battery capacity of approximately 780 mAh). This size allows the PSI board to be attached to the back of the phone, on top of the battery. Conceptually, the PSI board could be built into a combined battery & expansion pack – a solution that would require more extensive industrial engineering, but permit the electronics to be supported within an integrated plastic housing.

To interface into the MMC/SD card slot on the phone, a small ribbon cable is used to feed from the slot to the card as shown in Figure 1 (e.g., mounted on the back of the device where the battery is). This cable can be easily constructed out of a standard MMC/SD-card to trans-flash adaptor card, which provides a compatible mechanical design that fits into the MMC/SD card slot.

C. Switch detector, GPIO module, Signaling

The core of the PSI board is a switching mechanism that arbitrates connections between the host phone platform and either the embedded microcontroller or MMC/SD card. There are three main operating modes provided by the switch: MMC, MSP, and CROSS. Each mode is entered through a separate signaling mechanism, described below, and is always initiated by the phone device. The switching mechanism itself uses a set of analog switches between the various components, allowing for transparent bi-directional communication with minimal signal impact.

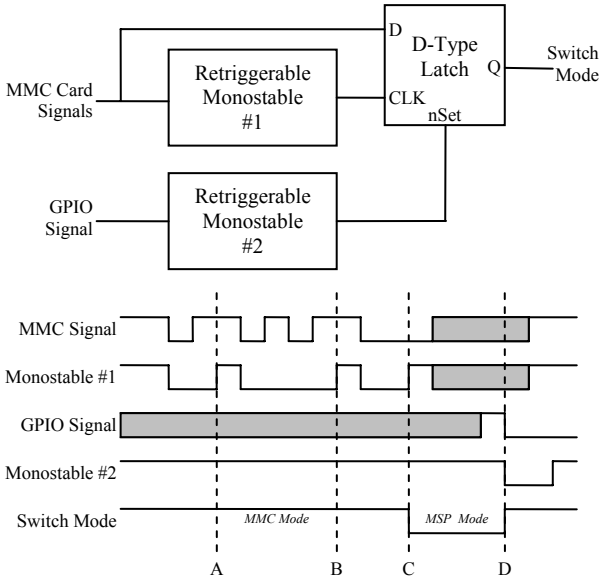


Fig 3: Simplified circuit diagram of switching mechanism, with timing diagrams for set & reset pulses. A) Normal MMC short pulse is ignored because signal is high when monostable fires. B) Multiple pulses are ignored because monostable is retriggerable. C) Single long pulse clocks low signal into latch, flipping switch. D) Falling edge on GPIO line sends limited duration pulse, switching back to MMC mode.

When the system is in MMC mode (Switch Mode HIGH), all signals from the phone are routed to the MMC card, allowing transparent access to the card's contents. To trigger a switch to MSP mode, a retriggerable monostable device, which turns an edge transition into a pulse of fixed length, is used to detect an out-of-spec signal condition on the phone's MMC/SD control lines, highlighted in Figure 3. This signaling is similar to the standard BREAK signaling on a serial UART line, wherein a data signal is held in a constant state for an extended length of time. Since the switching mechanism is out-of-spec for normal MMC/SD operation, this condition is only triggered by a specific intent to switch modes and will not adversely affect the card's normal operation.

While in MSP mode (Switch Mode LOW), the phone can use the MMC/SD interface to communicate with both the MSP and a dedicated GPIO module. Normal operation consists of communication with the MSP module, while the GPIO module is used for debugging (control over LEDs), in-circuit programming (described below), and switching out of MSP mode back to MMC mode. This is shown in Figure 3. While in MSP mode, the MMC/SD card is disconnected from the phone's data-bus, allowing it to retain

any transaction state and quickly resume operation. In addition to the basic MSP mode, there is a special case used to program the MSP processor's program FLASH, detailed below. While in MSP mode, any spurious switch detect signals (used to switch *into* MSP mode) are implicitly ignored. Furthermore, a second monostable device is used to turn the GPIO switch-to-MMC-mode signal into a finite duration reset pulse, since as soon as the device switches back to MMC mode the phone can no longer communicate with the GPIO module.

In the current PSI implementation, the basic switch from MMC to MSP mode takes approximately 1 ms while switching MSP to MMC takes approximately 32 us. The speed of the switch into MSP mode is governed by the time of a single data-bit while running at the slowest allowable clock speed (100 kHz), which could theoretically be as low as 100 us. The speed of the switch into MMC mode is limited by the time necessary to issue a 2-byte SPI write to the GPIO module to change the state of the switching GPIO line. These switching times affect the efficiency with which the system can switch between modes, along with any software driver overhead on the phone side required to multiplex access to the shared resource.

D. Microprocessor Subsystem

The microprocessor subsystem is based around the MSP430 microcontroller, and is very similar in design to the many standard wireless sensors, such as the Telos-B [5]. Similar to many of these platforms, the PSI board also provides an expansion connector useful for attaching expansion boards, which can be used to experiment with new sensors or other capabilities.

The main advantage of the MSP processor in this context is its relative high-capability with low-power consumption. This is an important requirement so the PSI can collect and process data while keeping the phone in sleep-mode and thus avoiding draining the phone's battery. Although similar in hardware to a wireless sensor node, the phone/PSI system is different in that the PSI board shares the battery with the phone host, which means the user only has to remember to charge a single device.

The MSP runs a custom embedded program that provides basic communication between the host processor and the underlying data sources/sinks (such as the accelerometer and radio). Since the PSI board is designed to work in conjunction with the phone's host processor, most computation tasks can be relegated to the more capable host, removing the need for task-specific programming of the embedded platform.

IV. PHONE-PSI INTERFACE MODES

This section describes the operating conditions of the major modes of the PSI system, while the previous section described how the system was able to switch between the various modes. A Linux device driver handles arbitration between the various modes by using software mutexes to control access to the single physical resource and presents multiple interfaces to higher-level software layers.

A. MMC Mode

The MMC mode places the system in a default access mode where the phone can access the contents of the MMC/SD card just as if the card were inserted directly into the phone. Data is accessed through the standard device drivers for the phone and is completely transparent to higher-level software devices. Furthermore, this mode provides a mechanism for the phone to efficiently access data stored directly on the card by the CROSS mode, described below (for example, while the phone was in a sleep state).

There are two main limitations for the communication channel between the MSP processor and the host. First, MSP cannot signal the host when in MMC mode because all the available signals on the MMC/SD connection are used for active communication; therefore, in order for the phone to detect/receive any signal from the MSP subsystem, it must either remain in MMC mode whenever idle or at least occasionally switch to MMC mode to detect if any data is available. Second, the MSP cannot wake the host processor while it is asleep. This is a limitation of the specific processor hardware pins used to communicate between the host processor and MMC/SD subsystem.

B. MSP Mode

In MSP mode, the phone can talk directly to the MSP processor or the dedicated GPIO module. In this mode, it can participate in bi-directional data-transfer with the MSP device using the SPI serial protocol. Additionally, a signaling IRQ line can be used to indicate an interrupt condition back to the host. The MSP mode could be the default idle mode for the system, allowing the phone to quickly switch to MMC mode if it needs to access card data. This would allow the MSP to signal an interrupt condition, such as incoming data ready, to the phone if necessary.

One limitation presented by the current implementation is a restriction on the data-transfer bandwidth between the phone and MSP devices. The MSP processor only operates at 6MHz, which, although allows it to have a relatively low power consumption, does not allow it to process the data from the PXA quickly. Additionally, the minimum data-rate

supported by the PXA hardware is 300 kHz, which is still fast enough that it would overwhelm the MSP FIFO hardware. Therefore, the data-transfer mechanism for the MSP uses a slower bit-bang interface on the PXA that allows the data-rate to be appropriately throttled. There are several possible technical solutions to this problem, ranging from a faster embedded processor (which might increase its power consumption), or a dedicated hardware FIFO module (increasing system complexity and cost).

C. CROSS Mode

In this mode, the PXA effectively disconnects from the MMC bus and allows the MSP to talk directly to the MMC/SD card. This mode is useful because it allows the PSI board to continue operation even though the host phone system may be shutdown or asleep, and it also provides a more efficient data-path for storing data. Without it, the system would have to transfer data from the MSP to the PXA, and then from the PXA to the MSP, creating a communication bottleneck between the Phone and PSI board. Additionally, this capability can be used to store data directly to the storage card without having to activate the higher-power phone platform. In order to manage the data storage on the MMC card, standard filesystem partitioning can be used to create separate storage areas for MSP data and regular phone data.

In CROSS mode, the MSP talks directly to the MMC/SD card, and the phone is disconnects from the channel. Two signals remain active between the MSP and phone to facilitate signaling for returning to MSP mode (and then eventually to MMC mode, if necessary). Returning from this mode to MMC mode requires that the phone reinitialize its MMC card interface since the previous state of the phone's transactions will have been lost. CROSS mode is entered by the phone first switching to MSP mode, if necessary, and then issuing a command to the MSP to enter CROSS mode, at which point the phone disconnects from the MMC/SD electrical interface.

D. Inline programming

A sub-mode of the MSP mode is used to directly program the MSP from the phone device. This is useful because it allows the phone to update the MSP device without requiring it to be removed and plugged into a dedicated programming device. This could even be done in the field by delivering the code update through a GPRS connection. To implement this mode, the phone communicates with the embedded GPIO module to place the MSP in a reset state and applies the necessary control sequence to place the device in the programmable state. The phone can then program the device, and reset it to execute the new firmware.

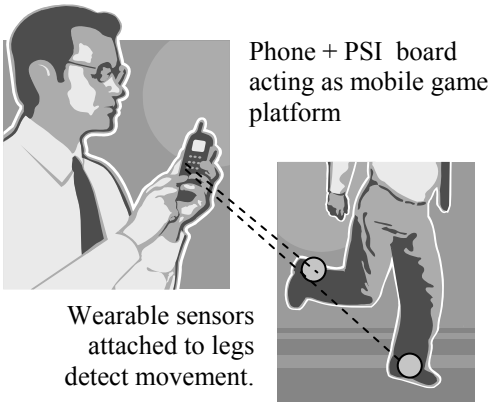


Fig 4: The PSI/Phone device and wearable sensors could provide the platform for a new class of mobile games that enable physical movement, such as a walking motion, to control actions in the virtual game world.

Once placed in the programmable state, the MSP needs to receive programming data using a standard serial protocol at 9600 BAUD, which is not supported by the phone hardware on the relevant pins. Therefore, similar to the technique used for implementing a slow SPI connection, the phone uses a bit-bang serial driver to send data at the necessary 9600 BAUD to the MSP for programming., (9600 BAUD is slow enough that the phone can manually generate the necessary signals within the required tolerance.)

V. APPLICATIONS

The PSI board is a flexible platform that can support a wide variety of mobile and wearable sensing applications. Several of these target applications are described below.

A. Wearable Activity Monitor

One primary use of the PSI board is as a wearable activity monitor: a device that could monitor and report on the user's daily physical activity patterns. Some current work on activity monitoring [3] uses a separate sensing component, which, although it has some advantages, ultimately requires users to manage a separate mobile device. Since the PSI is tightly integrated with a user's cell-phone, it would be very easy for them to keep it with them most of the day. By making the wearable sensor ecosystem more accessible to end users, this design should increase the availability of lightweight sensing applications to the general population.

Technical considerations of note are the ability to use the CROSS mode to store sensor data directly to the MMC card. In this case, off-line processing would be used (e.g., a desktop inference engine triggered when the phone is docked with a power source) that would be too processing intensive and power hungry to run on the mobile device itself. For some applications, the phone's processor could be used to perform the activity inferencing on the device, allowing for real-time interaction based on activity – possibly reminding users when they should, or should not, be doing some specific activity.

B. Physical Gaming

A network of wearable sensors attached, for example, to a user's arms and/or legs enables a new class of physical game that would allow people to interact with the phone without being restricted to the phone's limited input capabilities. A racing game, for example, could be controlled by how fast somebody can shuffle their feet up and down (Figure 4), or arm and leg movements could control a fighting game. This capability would be similar to systems that use a wireless joystick to control a PC game – except the PSI system would enable a more mobile sensor mobile experience outside of the home.

The technical considerations of these kinds of systems requires management of the wireless sensor network, since the wireless nodes need to be powered and wouldn't continually connected to the cell-phone's battery (unlike the PSI board itself). Furthermore, the system would need to be configured to associate the sensors with the mobile device. These basic power and association problems are similar to those found for wearable healthcare systems, but fortunately a wearable game would likely be less battery-critical because it would not necessarily be used all the time.

C. Gesture Recognizer

Another application of the Phone + PSI platform is to use an attached Near Field Communication (NFC) reader combined with accelerometer readings to facilitate connections between devices. The system would use NFC, which enables close proximity based communication between devices, to allow a user to specify *which* device to connect to with their phone, and then a gesture, detected by the accelerometers on the PSI board, indicating *what* should happen when the devices are connected. For example, the user could scan their home stereo system with the mobile phone, and then do a "play jazz" gesture with their device to initiate the music playing.

Currently, neither NFC nor accelerometers are commonly available in mobile phone platforms. Some platforms are starting to have NFC capability, but they still generally do not have accelerometers. The PSI board enables researchers and developers to experiment with capabilities and interaction models that otherwise would be very difficult to explore. In time, research will (ideally) demonstrate the usefulness and viability of these technologies and they can then be incorporated into the standard phone platforms. Until that time, the PSI board is necessary to enable the underlying experimentation.

VI. CONCLUSION

The PSI board provides an integrated mechanism for extending the sensing and communication capabilities of commercially available cell-phone platforms. The device integrates seamlessly with existing phone applications and usage models enabled by hardware and kernel switching abstractions. This switching mechanism requires a custom detection mechanism that automatically detects a special out-of-band switching signal that does not interfere with standard MMC/SD card operation.

By adding embedded processing, an accelerometer, and wireless-network communication capability without interfering with the normal use of the MMC card slot, the PSI board enables a variety of new applications that cannot be easily prototyped with existing phone platforms. The PSI board, designed for extensibility, thereby enables new classes of applications that would otherwise require users to carry a separate device with them.

REFERENCES

1. Baker, J. P., Bones, P. J., Lim, M. A.. *Wireless Health Monitor*. In Proceedings of the Electronics New Zealand Conference (ENZ Con 2006), November 13-14, 2006, Christchurch, New Zealand..
2. Ballagas R., Memon F., Reiners R., and Borchers J., iStuff Mobile: Prototyping Interactions For Mobile Phones In Interactive Spaces
3. Consolvo S., Everitt K, Smith I., Landay J.: Design requirements for technologies that encourage physical activity. CHI 2006
4. http://www.trolltech.com/products/qtopia/greenphone
5. http://www.xbow.com/Products/productsdetails.aspx?sid=126
6. Kumar A., Fazlur R. System for Wireless Health Monitoring. Sensors for Industry Conference, New Orleans, Louisiana, 27 January 2004.
7. Shnayder V., Chen B., Lorincz K., Fulford-Jones T, Welsh M.,, Sensor Networks for Medical Care, Harvard University Technical Report TR-08-05, April 2005.

Address of the corresponding author:

Author: Trevor Pering
Institute: Intel Research
Street: 2200 Mission College Blvd. MS RNB-6-61
City: Santa Clara, CA
Country: USA
Email: trevor.pering@intel.com

Groggy Wakeup - Automated Generation of Power-Efficient Detection Hierarchies for Wearable Sensors

Ari Y. Benbasat[1] and Joseph A. Paradiso[1]

[1] The Media Laboratory, MIT, Cambridge, MA USA

Abstract—**We present a framework for the automated generation of power-efficient state detection in wearable sensor nodes. The core of the framework is a decision tree classifier, which dynamically adjusts the activation and sampling rate of the sensors (termed groggy wakeup), such that only the data necessary to determine the system state is collected at any given time. This classifier can be tuned to trade-off accuracy and power in a structured fashion. Use of a sensor set which measures the phenomena of interest in multiple fashions and with various accuracies further improves the savings by increasing the possible choices for the above decision process.**

An application based on a wearable gait monitor provides quantitative results. Comparing the decision tree classifier to a Support Vector Machine, it is shown that groggy wakeup allows the system to achieve the same detection accuracy for less average power. A simulation of real-time operation demonstrates that our multi-tiered system detects states as accurately as a single-trigger (binary) wakeup system, drawing substantially less power with only a negligible increase in latency.

Keywords— **power-efficient sensing, tiered wakeup, cost-based classification, wearable gait monitoring.**

I. Introduction

Wearable sensor nodes are currently being used in a wide variety of applications. These include, but are certainly not limited to, implantable cardiac defibrillators[1], everyday activity loggers[2] and gait analysis systems[3]. Such systems are part of a new class of sensor-driven applications, leveraging the decrease in both price and size of the components to allow rich, multimodal data streams to be captured by very compact systems. However, wearable sensing applications are constrained to short lifespans by high power usage and limited battery size. For example, long-term medical monitoring is currently hindered by its power consumption. Both fixed environmental sensors – which cannot provide a full picture of an active patient's movements – and body-worn sensors – which require large battery packs and/or frequent replacement – are inadequate.

By concentrating our design efforts on the sensors themselves, rather than on the networking or processing, it is possible to construct sensor systems that achieve their goal(s) while drawing significantly less power. By reducing the power consumption, it is possible to reduce the size of the device and/or increase its lifespan. Both of these parameters directly improve marketability and user acceptance, allowing many more applications to make the transition from laboratory to marketplace and thereby benefit a wider population.

II. Framework overview

The main goal of this work is the reduction of energy usage in wearable sensor systems through the creation and demonstration of new tools and algorithms for the design and construction of power-efficient sensor systems. We start from a fundamental: the *raison d'etre* of these devices is to collect and process data and therefore the design of the sensors should be central. Therefore, we concentrated on reducing the energy usage of the sensors within the nodes. This metric was chosen since both general and tractable, though it is important to note that any power savings in the form of reduced sensing also correspond to further power savings through a reduction in:

- data to process.
- data to transmit or store.
- data to analyze (particularly for a human expert).

Any gains through this work can be considered independently from the large body of work exploring power savings through improvements to *ad-hoc* networking protocols and processor efficiency.

The framework presented is centered on the concept of "groggy", or tiered, wake up. This is in contrast with the more common binary wake up systems, which have only two modes: fully active, collecting all possible data and drawing maximal power, or fully asleep, collecting no data and drawing virtually no power. Instead we envision a system with a number of different levels of activity and associated power usages. Each of these comprises the currently active sensors for state determination and their sampling rate, and algorithms to describe the levels transitions. Our goal is to determine the system state at any given point in time for the smallest outlay of energy. The response thereto is not constrained

The form of the solution is such that the sensor sampling rates, as well as the transitions between them, are generated in a semi-autonomous fashion and can easily be embedded in

hardware. Therefore, the work should be applicable to a wide variety of applications.

Given a desired application, the design process will proceed as follows. Hardware for the individual application will be configured. In the initial deployment for testing and sample data collection, it is assumed that the system will include any sensors that could possibly be of value for state determination. A training data stream is collected and is annotated by the application designer. It is then used to determine the sensor data necessary to differentiate between the states. This information allows the final form of the hardware to be built (possibly with a pared down sensor set) and the state determination algorithm to be implemented on it. Each of these tasks is discussed in the proceeding sections.

III. Related works

There are a number of tiered wakeup systems, mostly built in an *ad-hoc* fashion, which have recently appeared in the literature. A group at UC Berkeley[4] examined the problem of detecting and identifying civilians, soldiers and vehicles traveling through a dense grid of independent sensor nodes. The nodes were awakened using a passive infrared detector, which then activated a microphone and magnetic sensor to identify the source of the trigger. This project did not meet its lifespan goals for a number of reasons. Key among them was the much higher incidence of false alarms than predicted, mostly caused by moving flora. This illustrates the danger of designing systems without the use of real-world sample data streams, instead relying on hand-scripted thresholds[5]. Also, the system draws more power than necessary since it turns on all sensors after a trigger, while the acoustic sensor alone can be used to distinguish between humans and vehicles and draws far less power than the magnetic sensor. A more analytic approach would likely have both identified these flaws sooner and made them easier to correct.

Van Laerhoven[6] designed an activity monitor based on a cluster of tilt switches and a single two-axis accelerometer. The tilt switches are used both for pose detection and to determine when the activity level is high energy to merit turning on the accelerometers. While this design is quite clever, the techniques presented have all been determined and hard-coded for a single application and sensor set. There is no apparent way to extend or generalize these techniques.

Non-state-based (unsupervised) systems have also been reported. Jain and Chang's[7] work on adaptive sampling for sensor networks is one example. They use the innovation of a Kalman filter[8] as a measure of the entropy rate of the data stream and adjust the sampling rate accordingly. Note that this is a purely entropic approach - more data is collected because the phenomena are varying at a faster rate. While this technique generated good results in a sample application, it assumes that data should be collected in all states and cannot differentiate between them. Further, the Kalman filter itself is a fairly structured (and computationally expensive) model that would not be appropriate for all systems.

IV. Hardware

The prototype hardware for this framework is implemented using a modular sensor platform we have designed. This platform is based around a series of circuit boards (or panes), each of which instantiates a specific sensing modality – *e.g.* inertial sensing, tactile sensing or ambient sensing – with the goal of reducing needless reimplementation of common components. Each board encapsulates the best practices in a given field, thereby saving substantial design time. Further, these boards can be arbitrarily combined and recombined, allowing for rapid prototyping and testing of proposed sensor combinations. For a given application, the designer can use this platform to quickly put together a sensor node with which to collect training data. We discuss two important design characteristics below. A more detailed discussion of this platform can be found in [9].

Since much of the power savings of the framework is predicated on power-cycling the various components, reducing the wake up time is key to minimizing the power wasted during that interval. For sensors, this parameter can vary widely both between different sensing mechanisms for a given phenomenon (*e.g.* effectively nil for a phototransistor to 40ms for a IR rangefinder) and individual parts (*e.g.* 8ms for the ADXL202 MEMS accelerometer to 100ms for the pin compatible MXR2312 thermal accelerometer). The wake up time sets the upper limit of how quickly a sensor can be cycled while still offering power savings over continuous activation. It should be noted that the availability of this information is spotty at best – given on some datasheets while completely ignored on others. Further, no information is given about the power draw during wakeup. In many cases it is likely the same as normal, though for some sensors (*e.g.* that need to charge internal capacitor or equilibrate filters) it may well be noticeably more. Further, since most microcontroller-based analog to digital converters can sample far faster than sensors can be activated, the energy used to wake up the sensor can be considered to be the energy used per wakeup cycle.

A second key design technique will be the use of multiple sensors to measure a single parameter of interest. For example, the inertial board uses both passive tilt switches and accelerometers to measure motion. The vast majority of sensor systems limit themselves (usually in the interests of simplicity or compactness) to a single sensor for each modality of interest. No matter how efficient such an implementation is for extracting information, it is guaranteed to be power inefficient

in states were less (or more) data is necessary to determine the transitions. A system which can tailor its sensing in real-time to the current state of the device can draw far less power on average. While it seems counterintuitive that we can make a system more power-efficient by adding complexity (and/or redundancy), the key is that the system has been given a new, lower energy source of information.

V. Pattern recognition

A. Data Collection

We have chosen a supervised training approach based on the assumption of fairly constrained applications and clarity of designer intent. The main benefit of this approach is the ability to ascribe specific meaning to the chosen states (*e.g.* walking), to combine states that might otherwise be separated (*e.g.* fast and slow gaits) and to ignore altogether portions of the data stream that could potentially be considered interesting (*e.g.* skipping).

Training data is collected using the following two-part procedure to acquire the most relevant data. The first stream will be reasonably short and contains the active (high energy/variance) states – both those considered interesting (*i.e.* to be detected) and not – that are known to the designer. The selection of states is left to the discretion of the application designer. The second contains a long-term background recording, to provide a baseline for the uninteresting cases and to catch states that were not considered by the designer above. Both streams are captured at the maximum useful data rate. For wearable/human applications, we use 200Hz.

While a long-term data stream could be used as the sole source of the training data, this tends to be inefficient for a number of reasons. Firstly, the length of the recording necessary to acquire good examples of all complex states can be quite long and their duration may be quite short. Secondly, the vast majority of the uninteresting data collected will be of little to no value in the classifier training.

B. Feature Extraction

A set of simple first order functions is used to calculate the features – the windowed mean, variance, minimum and maximum. These features have been used successfully on time series of human motion, both with inertial[10] and video[11] data. These are good general statistics for two reasons. First, they are algorithmically efficient to implement in a point-wise fashion such that they can be calculated quickly and for little power in an embedded microcontroller. Second, since the data stream will be wide sense stationary over a window size equal to the period of the data (within any given state), the mean and variance will be constant (with the exception of additive noise). This converts a sequence of time varying values to one that varies with state alone.

There are two free parameters in the calculation of these features – window size and sampling frequency. Window size is fixed as the period of the motion of interest, for reasons given above. On the other hand, multiple sampling frequencies are used to allow the classifier to choose the lowest rate (and power usage) that achieves its accuracy goals. However, to avoid having to power cycle a microprocessor on a complex and irregular schedule (which would vary with state), we limit the frequencies to the Nyquist frequency and power of two fractions thereof.

C. Classifier Design

The specific goal of this work is to create a hierarchy of activation states to allow the system to make a state determination using as little information (*i.e.* energy) as possible. Therefore, the classifier used should be able to make decisions in the same fashion - using more or less data as needed. Most classifiers do not have this ability, and instead require that all data be present to make any state determination.

In contrast, decision trees structure classification in the form of a series of successive queries, with each response leading to a following query until a state is determined[12]. In this way, the tree uses different sets of features to classify different states (or subsets thereof). In the case of an unbalanced tree, some classifications are made with fewer decisions (and therefore less energy) than others. Overall, the desire for hierarchical activation requires a hierarchical classifier. Further, the query-based structure allows for very fast implementation in an embedded microcontroller, since only simple comparisons are required to evaluate the classifier (beyond the calculation of the features themselves). Therefore, decision trees with are used in this framework.

In general, our system follows the CART algorithms for constructing decision trees as detailed in [13]. The key difference is that we take the cost of the features – which we define as the energy necessary to collect and calculate them – into account. In the case of sensor data, this almost always reduces to the cost to power up the sensor itself. The hierarchical structure of the decision – where certain sensors may already have been used to answer a query – adds some complexity. Any sensor used at the same (or higher) sampling rate higher in the tree has its cost reduced to zero, since the data has already been collected. If the sensor was used at a lower sampling rate, the cost of use is discount by that already paid, since some of the necessary data is already being collected.

Energy usage is taken into account by reducing the splitting criterion used to determine the query to use at any given node i by a factor of $(1+ TC_i)^W$, where TC_i is the test cost for the sensor at that node and W is a parameter used to adjust the trade-off between power and accuracy. This factor does not

distinguish between active sensors and requires unused sensor to achieve a certain quality of split (depending on W) before they will be activated. To avoid non-linearities, test costs should be scaled such that the minimum value is greater than ten. The ratio of the value of the splitting criterion for a perfect split and a meaningless one sets the maximum useful value for W – in the case of the Gini index used in CART, this value is 0.29.

VI. Embedded software

Having trained an appropriate classifier with the desired power and accuracy characteristics, it is fairly straightforward to implement it on a wearable sensor node. The decision tree is encoded in array format, with each entry containing a sensor feature to test on, the threshold for that test, and the index of the next node given a positive or negative result. Entries corresponding to leaf nodes will simply contain the state determination. In either case, a listing of the sensors necessary to get to that point is also given. The code itself keeps track of the currently active sensors as well as an accounting of the recent requests for inactive sensors.

The code execution cycles through a fairly simple loop. The processor awakens at a predetermined time and awakens and collects the data from the active sensors and updates their features. Constant time algorithms for the features are well known and will not be presented here. The classifier itself is then run. If a state is determined, the processor executes the desired response (set by the application designer). If the information necessary to determine a state is not available, the system is considered to be in an indeterminate state. Next, the sensor needs of the tree are examined, with an update of the accounting of both the sensors that are needed but are inactive, and the sensors that are active but are not needed. The system then goes to sleep until the next cycle. Cycle length is determined by the highest update rate among the currently active sensors.

Effectively dealing with sensor activation is key to the efficient operation of this system. Instantly turning on sensors when requested or turning them off when not used to make a decision will leave the system highly susceptible to noise. Waiting too long to activate the sensors will increase latency, while waiting too long to turn off a sensor will waste power. In the current simulations, a sensor must be requested or unused for 5 cycles before its activation state will be altered. This parameter has not yet been optimized. It may be that a higher value is necessary, since whenever a sensor is reactivated there is a delay equal to the window length (or one cycle of the state of interest) before enough data will be available to calculate the features and run the classifier again. Therefore, it may be necessary to sacrifice some power to keep the latency to a reasonable level.

VII. Limitations

We note two important limitations imposed on this framework. First, we have limited the systems examined to the use of passive sensors – those which measure the environment without affecting it. Active sensors (*e.g.* sonar, radar) - those which control the transmission as well as the reception of the measured signal - are excluded both because of their power usage (one to two orders of magnitude greater than passive sensors) and their complex power management (both the output power and the sampling rate can be adjusted individually).

Structurally, any wake-up based system has the potential to miss anomalous events (*i.e.* those with no precursor) - either entirely or during the state determination procedure. While this is a problem in general, it should be minimized given our concentration on wearable sensing. Since most activities in this domain take place on the order of seconds (at minimum) and state determination requires at most a second, the chance of missing an event is minimal.

VIII. Analysis

A. Data Set and Collection

For testing, a data stream containing a wide variety of different ambulatory activities was collected, using a shoe mounted node containing three axes of gyroscopes, three axes of accelerometers and a four-way passive tilt switch. The set of activities chosen was: normal gait, walking uphill and downhill, ascending and descending stairs, and shuffling gait. This data set allows us to create classifiers that attempt to separate a single (complex) ambulatory activity from the rest. Such a system would be valuable for Pakinson’s Disease patients, where a doctor would be most interested in collecting information about the frequency and parameters of the patient's shuffling episodes[14]. For patients with total knee replacement, activities such as ascending stairs (where the knee flexion is >90°) are the most important to measure[15]. Such cases allow for far more complex and richer classifiers than those used to simply separate ambulatory from non-ambulatory (roughly: still) states.

Data was collected using the process described above. The active data stream contained two separate segments of approximately 30 seconds in length of each of the non-walking motions. These motions were each bookended by a segment of normal gait. This data was collected in a single session lasting approximately 30 minutes and was recorded with a video camera. Annotations were later added to the data stream based on camera's time stamp.

The long-term data stream was designed to collect data representing the everyday activities of an office-bound worker. Two hours of data were collected with the subject

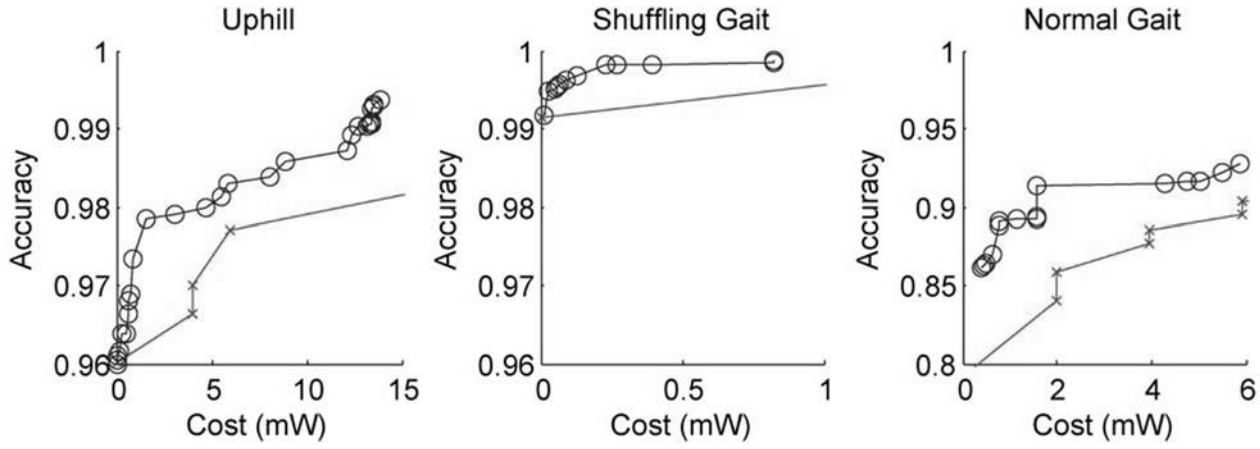


Fig. 1: Power/Accuracy trade-off curves for decision tree and SVM classifiers

sitting at his desk performing a number of basic activities -- typing, reading, searching for papers, *etc.* Non-ambulatory motions such as adjusting the position of the feet, moving from one desk to another and waving the feet under the desk were collected. The subject recorded his activities in a simple diary, with annotations later added to the data stream using the diary as a rough guide and visual inspection of the data stream to more precisely mark the times.

B. Classifier Performance

The above data stream was used to test the ability of our framework to create classifiers that trade off power and accuracy. Classifiers were trained to detect each of the six walking motions amongst the others. Only the active data set was used to avoid placing a large positive skew on the accuracy (since non-walking motions are trivially excluded using the tilt sensors). The sensors and their power usage at the various frequencies used in the classifiers are shown in Table 1. Note that the gyroscopes cannot be power cycled above 33Hz and the accelerometers above 125Hz. It is assumed that ambulatory motion is 80% normal gait and 4% of each of the other gaits. While these are estimates, we note that decision trees are fairly robust to errors in prior probabilities.

Table 1: Power usage of sensors for various frequencies

	25 Hz	50 Hz	100 Hz	200 Hz
Gyroscope	22.5mW	30mW	30mW	30mW
Accelerometer	0.396mW	0.792mW	1.58mW	1.98mW
Tilt Switch	0.41μW	0.83μW	1.6μW	3.3μW

Figure 1 shows the power-accuracy trade-off curves for detecting uphill, shuffling and normal gait (other classifiers omitted due to space constraints). The final point in each graph is from the decision tree grown without power constraint. We note that while the accuracy always increases with power, there is a knee point at which the marginal gain becomes quite low. This data is given in table 2 below.

Table 2: Asymptotic power behavior of trained tree classifiers

	Inflection Point		Maximum	
	Power	Accuracy	Power	Accuracy
Uphill	5.66mW	0.9824	33.82mW	0.9905
Shuffling	125uW	0.9976	1.585mW	0.9985
Normal	5.89mW	0.9272	10.02mW	0.9325

Figure 1 also shows the results of a Gaussian kernel Support Vector Machine (SVM)[16] trained with various subsets of the sensors (always activated). Only the monotonically increasing hull of the points is shown (hence the two point graph for shuffling). While the eventual maximum accuracy (beyond the right-hand edge of the graph) is greater than that of the decision trees, we note that the trees outperform in the low power domain. This confirms the benefits of the tiered structure of the decision tree, even when compared with a (nominally) more powerful technique.

C. Embedded Simulation

Since the tree classifiers were trained without regard to time information, it is important to confirm that they perform well in real-time conditions. The embedded software described above was simulated in MATLAB. Figure 2 shows the results for uphill gait classifier at the inflection point. Both the ground truth (user annotation) and smoothed detection of the classifier (state changes are only acknowledged if their duration is greater than 0.25 seconds) are shown. Table 3 gives the duration of false positives (FP) and false negatives (FN) compared to the true positives (TP) and true negatives (TN) for the tasks above. Total length is 323 seconds. For

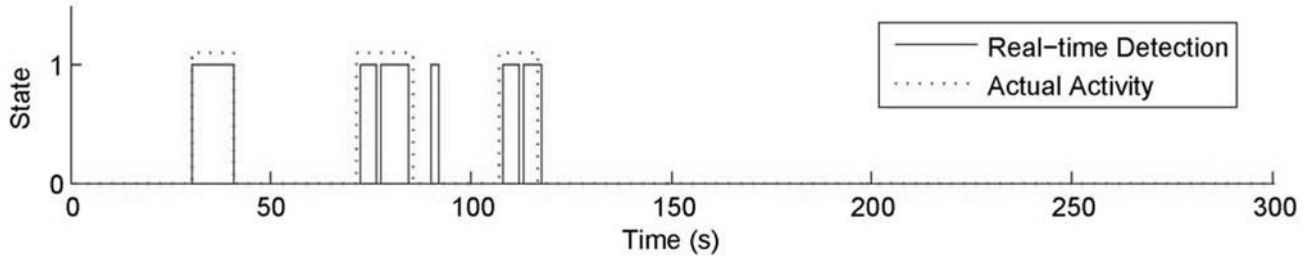


Fig. 2: Model of Real-time Operation for Detection of Uphill Gait

shuffling gait, most of the FPs come for small triggers likely caused by readjusting the feet. For normal gait, most come from overly conservative activity labeling (which is beneficial when training the classifier).

Table 3: Incidence of false positives and negatives on trained trees

	FN	TP	FP	TN
Uphill	5.81s	28.23s	3.11s	285.57s
Shuffling	0.77s	34.93s	15.82s	271.20s
Normal	7.66s	99.62s	50.97s	164.48s

As a final test, we constructed binary classifiers using the root node of the chosen decision trees (*i.e.* if the first query is met, the system is fully activated). This allowed us to estimate the power advantage of a multistage wakeup. For uphill and normal gait, the savings were 56% and 46% (respectively). The best solution for shuffling was found to be a binary classifier.

IX. Conclusions and future work

We have described a three-component framework for power-efficient detection in wearable sensors. The first is modular hardware platform for ease of application prototype. The second and key component is a semi-autonomous classifier construction algorithm. Given an annotated data stream, the system uses a version of the CART algorithms, modified to take sensor power into account, to produce a family of decision trees with various power/accuracy parameters. The final component is an embedded implementation of this classifier for use with wearable sensors nodes. Numerical analysis was presented which supports our contention that a tiered wake up approach can reduce power usage with a minimal reduction in accuracy.

Avenues for future work on this framework center around extending from sensor nodes to networks. For the case of a body-worn sensor network, there are two approaches assuming the goal of detecting the current state while minimizing overall power usage. If in any given state one node is the network will be superior to the others, those nodes can share their current state with the other nodes, allowing them to reduce their data collection or turn off entirely. If data from multiple physical locations make some state determinations easier than if done at a single node, the data in question can be shared. In either case, this information can easily be added to the decision tree construction algorithm with a cost based on the energy expended to wirelessly transmit and receive the data.

Acknowledgments

The authors would like to thanks Paolo Bonato for fruitful discussions. The authors also acknowledge the support of the Things that Think Consortium, as well as all the sponsors of the MIT Media Lab.

References

1. J. P. DiMarco. Inplantable cardioverter-defibrillators. New England Journal of Medicine, 349(19): 1836–1847, 2003.
2. Brian Clarkson. Life Patterns: structure from wearable sensors. PhD thesis, MIT, 2002.
3. Ari Y. Benbasat, Stacy J. Morris, and Joseph A. Paradiso. A wireless modular sensor architecture and its application in on-shoe gait analysis. IEEE Sensors 2003.
4. P. Dutta et al. Design of a wireless sensor network platform for detecting rare, random and ephemeral events. IPSN 2005.
5. P. Dutta, University of Calfornia Berkeley. IPSN ’05 talk.
6. K. Van Laerhoven, H.W. Gellersen, and Y.G. Malliaris. Long-term activity monitoring with a wearable sensor node. BSN 2006.
7. A. Jain and E. Chang. Adaptive sampling for sensor networks. Data Management for Sensor Networks 2004.
8. A. Gelb, editor. Applied Optimal Estimation. MIT Press, 1974.
9. Ari Y. Benbasat and Joseph A. Paradiso. A compact modular wireless sensor platform. IPSN 2005.
10. Stacy J Morris. A Shoe-Integrated Sensor System for Wireless Gait Analysis and Real-TimeTherapeutic Feedback. ScD thesis, MIT, 2004.
11. Lily Lee. Gait Analysis for Classification. PhD thesis, MIT, 2002.
12. A. Webb. Statistical Pattern Recognition. Wiley, 2002.
13. L. Breiman et al. Classification and Regression Trees. Wadsworth,, 1984.
14. H. Zhu, J. Wertsch, et al. Foot pressure distribution during walking and shuffling. Arch Phys Med Rehabil, 72:390–397, May 1991.
15. D.E. Krebs et al. Biomotion community-wearable human activity monitor: Total knee replacement and healthy control subjects. BSN 2006.
16. R. Duda, P. Hart, and D. Stork. Pattern Classification. Wiley, 2001.

A Low Power Compression Processor for Body Sensor Network System

Hyejung Kim, Sungdae Choi, and Hoi-Jun Yoo

Semiconductor System Laboratory
Department of Electrical Engineering and Computer Science, KAIST
Daejeon, Korea

Abstract— **A low power 16-bit RISC is proposed for body sensor network system. The RISC is designed of basic 3 stage pipeline architecture which has 28 instruction sets. Some special instructions are proposed for efficient applications. The lossless compression accelerator is embedded in the RISC to support the low energy data compression. The accelerator consists of 16x16-bit storage array which has vertical and horizontal access path. By using the accelerator the energy consumption of the lossless compression operation is reduced by 95%. The RISC is implemented by 1-poly 6-metal 0.18um CMOS technology with 16k gates. It operates at 4MHz and consumes 24.2uW at 0.6V supply voltage.**

Keywords— **Body Sensor Network, RISC, Lossless Compression**

I. INTRODUCTION

Recently, in according with the interests about the healthcare increase, people desire to check their vital signal or health condition at anytime and anywhere. To solve this request, the wireless sensor network (WSN) system has been studied to apply for the body, which called body sensor network (BSN) system [1]. The BSN system offers health condition monitoring, vital signal collecting, collected data analysis and diagnostics. To provide these various operations, BSN system should consists of lots of sensors, the data processing unit and the efficient network system. Various BSN systems have been proposed to provide the services [1]-[6].

BSN system requires the ultra low energy operation for stable long time operation and the small footprint for wearability. To achieve these requirements, the size of hardware components has to be small. Moreover, they also have limited power supply, bandwidth for communication, processing speed, and memory space. Various researches have been conducted so far focuses on how to achieve the maximum utilization of limited source. The data compression is one of the most effective methods [7]. Since the transmitting power consumption is much more than the data processing power, minimizing data size before transmitting can reduce total system power consumption. Moreover, the low power hardware design and the efficient algorithms are also important for limited resource.

This paper presents a low power RISC for low energy BSN system. The efficient compression algorithms optimized to bio signals is also proposed. We verify low energy consumption of the proposed RISC by implementation of real silicon and test board.

II. 16-BIT RISC ARCHITECTURE

A. Top Architecture of 16-bit RISC

The 16-bit RISC is designed based on a basic 3-stage pipeline which is optimal selection for low power operation [3]. The figure 1 shows the pipeline flow diagram of the proposed 16-bit RISC. The first stage fetches the instruction from code memory, and the second stage decodes the fetched instruction. The last stage executes ALU operations, memory access, and write-back to the register file. Since the both operations of read and write the register file occurs in the same stage, the data hazard is eliminated. The branch is performed with a 2-cycle penalty.

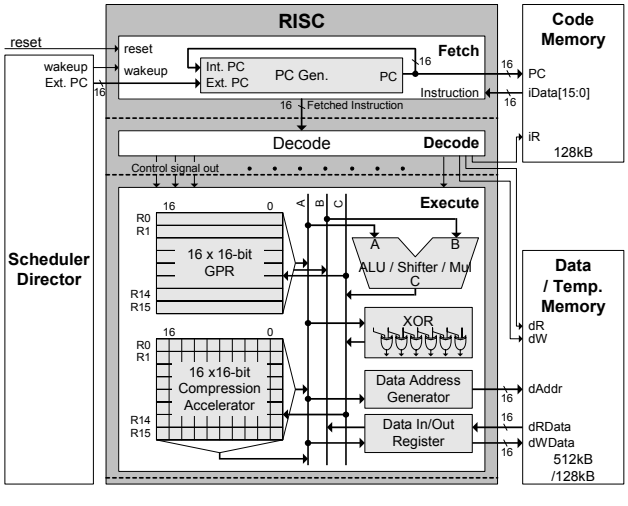


Figure 1. Top Architecture of the Processor

The RISC has 2 kinds of register files which are the 16 general register files and the compression accelerator for the proposed lossless compression algorithm described in next section.

TABLE I. RISC INSTRUCTION SET ARCHITECTURE

ALU Register	MOV, CMP, ADD, SUB, AND, ORR, XOR, BIC, MUL, MVN, NEG
	MGX, MXG, REV, XRC, XRD
ALU Immediate	MOV, ADD, SUB, CMP
Shift	LSR, LSL, ASR, ROR
Memory	LDR, STR, STRX
Conditional Branch	BEQ, BNE, BGT, etc… (16 Conditions)

B. Instruction Set Architecture

The processor implements 16-bit Instruction Set Architecture (ISA). The RISC has 28 instructions, the table I shows the implemented instruction sets. Some special instructions are proposed for the compression algorithm.

XRC, XRD : XRC instruction operates bitwise XOR operation with right next bit. XRD instruction performs reverse operation of XRD. Figure 2 shows the block diagram of the XRC operation, where *Rs* represents the source register and *Rd* represents the destination register.

$$Rd[i] = Rs[i] \oplus Rs[i+1]$$

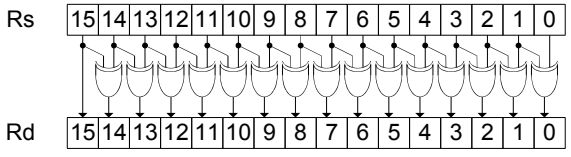


Figure 2. Block Diagram of XRC operation

MGX, MXG : MGX moves the data from general purpose register to special purpose register. MXG moves the data from special purpose register to general purpose register.

REV : REV moves the MSB to LSB, and LSB to MSB. Figure 3 shows the block diagram of the REV operation.

$$Rd[i] = Rs[15-i]$$

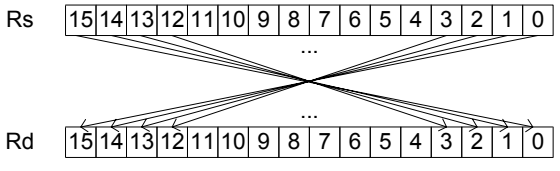


Figure 3. Block Diagram of REV operation

STRX : STRX loads the data from special purpose register, and store them to the data memory if the header bit is one. After store the data, the memory address is increased automatically. It helps to reduce the repeated store operation, when the 256 bit data are stored. *Rm* represents the memory address for data memory.

$$if\ (!head)\ Mem[Rm] = \{Rd\},\ Rm = Rm + 0x01$$

III. PROPOSED LOSSLESS COMPRESSION ALGORITHM

A. Compression for BSN

The sensor nodes gather the bio signal data, process the collected data, and transfer the processed data to base station. If a single sensor node sends n-bit data, the total power consumption (P_{Total}) is derived as following:

$$P_{Total} = n \times \{P_{Comp} + (1-\alpha)(P_{TX} + P_{RX} + P_{DeComp})\}$$

where,
n = *Data bit width*
a = *Compression rate*
P_{Comp} = *power consumption for Data compression*
P_{DeComp} = *power consumption for Data de-compression*
P_{TX} = *power consumption for Data transmission*
P_{RX} = *power consumption for Data receive*

We have quantitative value of P_{TX}=2.5nJ/b, P_{RX}=2.5nJ/b, P_{Comp}=0.003nJ/b and P_{DeComp}=0.003nJ/b from previous work [3-5]. Assuming that the sensor node gathers 1Mb data, the total power consumption would be changed under variable compression rates, the figure 4 shows them. The high compression rate helps greatly to reduce the total power consumption.

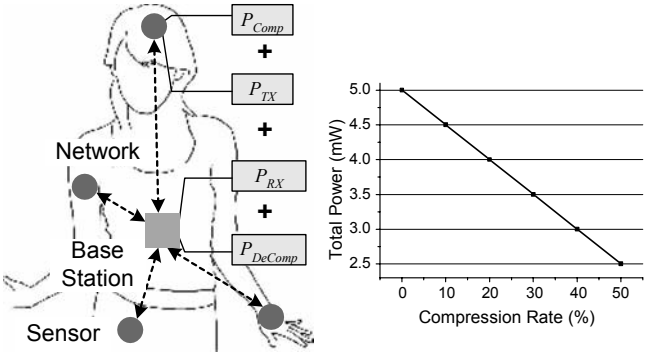
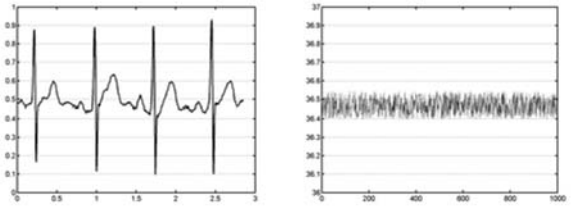


Figure 4. Comparison of Power Consumption

Therefore, it is necessary to employ a data compression algorithm for body sensor network system. There are some limitations to apply the compression algorithm to BSN system. First, since the sensor node currently has limited

resources such as battery energy, CPU performance and the memory capacity, the algorithm size must be as small as possible. The second one is the operating frequency. Usually the processor of sensor node operates only less than 4MHz, therefore it is necessary to design a low complexity algorithm which is enough to operate at low frequency. Because the hardwired accelerator and the special instructions help to reduce the operating cycles and the code size, the proposed compression algorithm is satisfied these requirements.



(a) ECG : 16bit (b) Temperature : 10bit

Figure 5. The Bio Signal with Variable Resolution

The proposed compression algorithm, which name is BiPAC, has three features. First, BiPAC provides the lossless compression. The lossly compression algorithm is efficient to minimize data size, but all loss algorithms have some degree of quantization error, resulting in a possible loss of diagnostic information. Thus, the lossless compression algorithm is studied to preserve all the information of original data. Second, BiPAC is optimized to continuous signal data. Most of bio signals consist of continuous and periodic waveform differ to other randomly media data. The last one is various precision coding. Each kind of sensor data has variable precision. For example, the ECG signal consists of 16 bit resolution, otherwise, the temperature signal consists of only 10 bit resolution. The figure 5 shows that 2 kinds of bio signals. Since the width of the system bus is 16bit, a lot of leading zeros can be generated when the resolution is less than 16 bit. BiPAC eliminates these redundant leading zeros simply so that it is suitable for the all kinds of various precision data compression.

B. BiPAC - Lossless Compression Algorithm

The algorithm1 and figure 6 describes BiPAC which is the proposed lossless compression algorithm. BiPAC represents the Bi-PAth compression.

Algorithm 1 : BiPAC

Input : 16x16b TM data (TM[i], i=0..15)

Output : nx16b compressed DM data (DM[i] , i=0..n, n≤15)

```
1. for i=0 to 15 loop
     d[i] = TM[addr1+i];
   endloop
2. h0=d0; spr[0][*]=0; n=0;
3. for i=1 to 15 loop
     spr[i][*] = d[i] - d[i-1];
4. for i=0 to 15 loop
     for j=0 to 14 loop
          spr[i][j] = spr[i][j] xor spr[i][j+1];
     endloop
   endloop
5. for i=0 to 15 loop
     for j=0 to 15 loop
          h1[i] = h1[i] or spr[j][i];
     endloop
   endloop
6. for i=0 to 15 loop
     if h1[i]!=0
          DM[addr2+j] = spr[i][*];
          n = n +1;
     endif
   endloop
7. return DM[0]..DM[n]
```

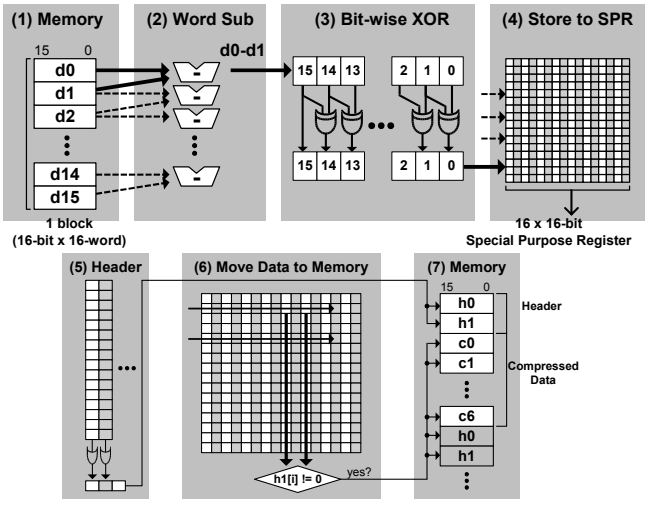


Figure 6. The Lossless Compression Algorithm

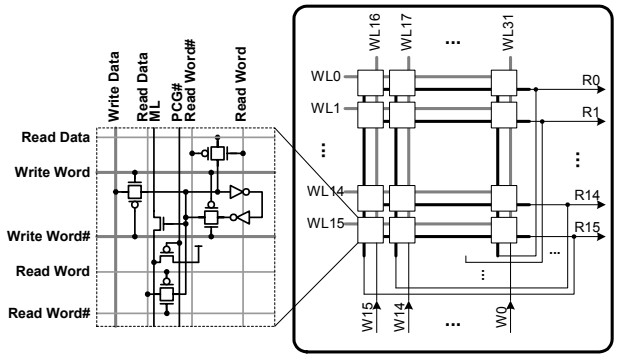


Figure 7. Lossless Compression Accelerator

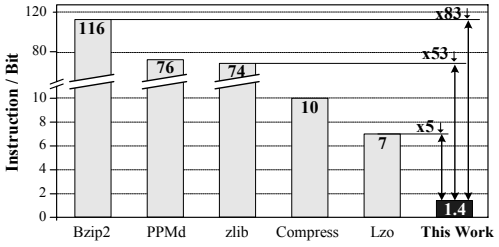


Figure 8. Required Instruction per Removed A Bit

C. Lossless Data Compression Accelerator

The data compression accelerator is embedded in the RISC to support the low energy data compression / decompression operation and it consists of 16x16 storage array shown in figure 7. Its vertical and horizontal accessibility reduces the required execution cycle to 95% compared with conventional RISC operation. The interface of accelerator is same as the register file, so all data transition takes 1 cycle by special instructions. In addition, it can be used as a general purpose registers when the compression program is not executed.

The proposed algorithm compresses the 1 block data (16x16-bit) into 2 headers and the compressed data and it takes up to 114-cycle. It is recalculated to 1.425-instructions/bit. Figure 8 shows comparison result of the number of instruction required by removing a single bit. By using the proposed algorithm with accelerator, the performance is improved by maximum 83 times compare to other conventional algorithms [7].

IV. Results

A. Simulation Results

The ECG records data from the MIT/BIH [9] are used to verify BiPAC. The sampling rate and the resolution are 360 samples/s and 12 bits, respectively. The figure 9(a) shows the ECG record waveform and the figure 9(b) shows the simulation result of compression operation. The data memory write enable signal goes high if the column data are not zero. The compression rate is good if the number of enable signal is small. The compression rate is 25% and 56.3% at steep and slow slop, respectively. The better compression rate is obtained with stable signals. The average compression rate is 38.7% for 10 sec amount data. It consumes only 641us and 0.69nJ to compress 1 sec amount ECG data with 4MHz operation frequency. Table-II shows that the energy consumption of this work is much smaller than that of the conventional low-power processors [4], [5], [6] when 16x16bit data compression is executed. The proposed compression algorithm can operate sufficiently real time lossless compression/decompression with ultra low energy consumption.

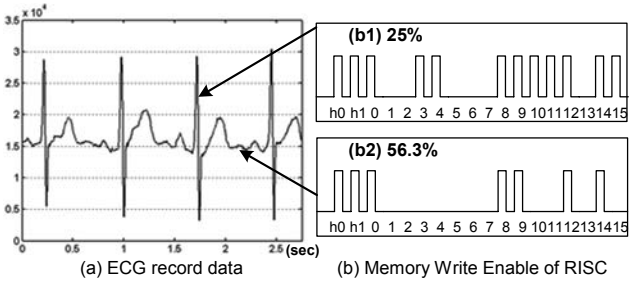


Figure 9. Simulation Result of Proposed Compression Algorithm

TABLE II. Energy Comsumption for 16x16b Data Compression

	[4]	[5]	[6]	Proposed
VDD	1.0V	0.66V	0.23V	0.6V
Clock	500kHz	4MHz	833kHz	4MHz
Energy	32.4nJ	64.8nJ	11.2nJ	0.69nJ

B. Implementation Results

With the proposed architecture, the RISC is implemented into a chip by using 1-poly 6-metal 0.18um CMOS technology. A demonstration board system is also implemented. They are shown in figure 10. The RISC size is 400umx1000um with the compression accelerator. The memory size of code memory, data memory and temporary memory is 128kb, 512kb and 128kb, respectively. The RISC operates maximum frequency of 200MHz with 1.8V and 22MHz with 0.6V. However, the RISC generally operates at 4MHz with 0.6V supply voltage for low energy consumption and reliable operation. The power consumption is 24.2uW with 4MHz at 0.6V supply voltage.

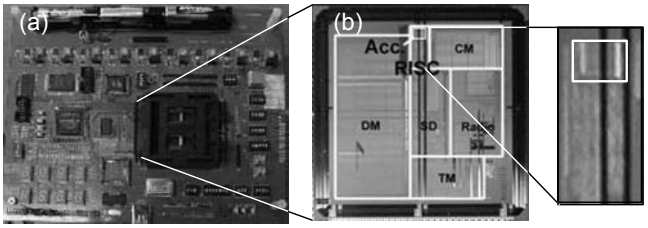


Figure 10. Implementation Result (a) Demonstraion Board (b) Chip Microphotograph

V. Conclusion

A 16-bit RISC of base station is proposed for body sensor network system. The RISC is designed based on basic 3-stage pipeline architecture. The lossless compression accelerator is embedded for low energy data compression. By using accelerator which consists of 16x16-bit storage array with vertical and horizontal access path, the energy consumption of the lossless compression operation is reduced by 83 times. The RISC is implemented by 1-poly 6-metal 0.18um CMOS technology with 16k gates. It operates at 4MHz and consumes 24.2uW at 0.6V supply voltage. The evaluation results clearly indicate the proposed RISC and compression algorithm are suitable for the body sensor network system.

References

1. Sungdae Choi, Seong-Jun Song, Kyomin Sohn, Hyejung Kim, Jooyoung Kim, Namjun Cho, Jeong-Ho Woo, Jerald Yoo and Hoi-Jun Yoo, “A Multi-Nodes Human Body Communication Sensor Network Control Processor,” Proc. CICC, Sep.2006.
2. Benton H.Calhoun, Denis C.Daly, Naveen Verma, Daniel F.Finchelstein, David D.Wentzloff, Alice Wang, Seong-Hwan Cho, and Anantha P.Chandrakasan, “Design Considerations for Ultra-Low Energy Wireless Microsensor Nodes.” IEE Tran.Computer, vol.54, no.6, pp.727-740, Jun.2005.
3. Mark Hempstead, Nikhil Tripathi, Patrick Mauro, Gu-Yeon Wei, David Brooks, “An Ultra Low Power System Architecture for Sensor Network Applications,” Proc. ISCA 2005.
4. Brett A.Warneke, Kristofer S.J.Pister, “An Ultra-Low Energy Microcontroller for SmartDust Wireless Sensor Networks,” Proc.ISSCC, Feb.2004.
5. Virantha N.Ekanayake, Clinton Kelly, IV, and Rajit Manohar, “BitSNAP: Dynamic Significance Compression For a Low-Energy Sensor Network Asynchronous Processor,” Proc. ASYNC, pp.144-154, Mar.2005.
6. Leyla Nazhandali, et al, “A Second-Generation Sensor Network Processor with Application-Driven Memory Optimizations and Out-of-Order Execution,” Proc. ACM CASES, pp.249-256, Sep.2005.
7. Naoto Kimura and Shaharm Latifi, “A Survey on Data Compression in Wireless Sensor Networks,” Proc.ITCC 2005.
8. S.Song, et al., “A 2Mb/s Widband Pulse Transceiver with Direct-Coupled Interface for Human Body Communication,” Proc. IEEE ISSCC 2006.
9. http://www.physionet.org/physiobank/database/mitdb/

Address of the corresponding author:

Author: Hyejung Kim
Institute: KAIST
Street: Guseondong 373-1, Yuseonggu
City: Daejeon
Country: KOREA
Email: seeseah@eeinfo.kaist.ac.kr

Path Loss For Short Range Telemetry

Yong Liu[1], Kairaz Contractor[1] and Yunjuan Kang[2]

[1] Wireless Communication Group, Philips Research East Asia, Shanghai, China
[2] Dept. of Electronic Engineering, East China Normal University, Shanghai, China

Abstract— **The popular Friis transmission formula is used to evaluate the amount of path loss in free space between the transmit and receive antennas for the design of wireless transceivers. It is could be also used to estimate the path loss for the link of the on-body network when there is no barrier and no body surface in between. Paying special attention to short-range communication, this paper develops composite expressions for path loss between two dipoles. Theoretical analysis is used to prepare an impedance-based model and an S-parameter-based model to extract the path loss. These expressions support the validity of Friis formula up to a certain separation distance, below which an error is induced, possibly resulting in miscalculation of required transmission power, sensitivity and margins, during the design of transceivers. Simulations and experiments were carried out to verify the proposed model, and their results match up well. The difference in values of path loss obtained by the proposed analytical model and Friis path loss formula in free space is highlighted. This is important to be aware of while choosing design parameters for transceivers that will be used in applications operating at short ranges as compared to wavelength at the frequency of operation.**

Keywords— **Attenuation, Scattering parameters measurement, Impedance, Dipole Antennas, Mutual Coupling**

I. Introduction

Radio communication enables on-body biomedical devices, cable-less medical devices, etc. to provide diagnosis, monitoring and therapy in a flexible and convenient way. For optimal design of the transceivers in such devices, the knowledge of precise path loss between on-body devices or between an on-body device and an external device is necessary.

Hall et.al. investigated different potential propagation modes of radio wave in on-body communications extensively and provided the measurement results at selected microwave frequencies [1][2][3]. For surface wave propagation and shadowed free space paths, statistics and measurement are the most efficient method to evaluate path loss. In the case of free space mode, e.g. from the belt to wrist path when the hand is in front of the body, the analytical method based on Friis formula might be adopted.

An on-body device and an external device linked by a Line-of-Sight(LOS) path or two on-body transceivers linked by a free space path without barriers often operate in the near-field zone of free space due to the choice of frequency and communication distance. The ISM bands that are present in the lower UHF range may be used as the communication bands for biomedical devices. Due to the longer wavelengths at those frequencies than at higher frequencies in the GHz range, the communication range is often within the near field zone of the transmit and receive antennas.

The well-known Friis transmission formula is widely used in radio design for long distance or mobile communication in free space. It is not intended for use in the near field. Using the Friis formula to estimate the path loss for the above mentioned biomedical applications could result in errors below certain distances of separation between the transmit and the receive antennas. This makes the investigation, to generate a uniform expression for path loss in both the near field as well as the far field, an essential study.

David et.al. investigated the near field channel between two coils [4]. Their paper showed that the near field path loss is inversely proportion to the sixth order of distance and proportional to the square of frequency of operation. Schantz studied near field propagation in [5]. In his propagation equation, the path loss decreases more rapidly in the near field than it does in the far field region. The inverse proportion to the sixth order of separation distance is also found in his paper. FitzGerrell investigated the insertion loss between pairs of thin, linear antennas, while considering the effect of proximity to the ground [6]. His work focused on the effect on insertion loss caused due to variation of operation frequency, and only three distinct separations between two dipoles had been investigated.

When there is a high probability of operation in the near field region of the transmit and receive antennas, mutual coupling between them should be considered. Derneryd and Kristensson [7] introduced an impedance matrix to characterize the mutual coupling between two antennas, but their focus was on the complex correlation coefficient between two transmit antennas of a Multi-Input Multi-Output (MIMO) system, and not on the field attenuation loss between them.

The next section of this paper develops the impedance matrix of arbitrary antennas into an expression for path loss between two dipoles. Another path loss expression based

on S-parameter principles is also presented in Section II. Section III compares the path loss models developed and described in Section II, and verifies them with simulations and experiments. Additional recommendations are provided for evaluation of path loss for arbitrary antennas. The conclusion is stated in section IV.

II. Models to express path loss

A. Friis transmission formula

Path loss is usually expressed as the ratio of received power to the transmitted power. The well-known Friis transmission formula defines the free space path loss along a Line-of-Sight(LOS) channel as:

$$\frac{P_{received}}{P_{transmitted}} = \frac{G_T G_R c^2}{(4\pi fd)^2} \tag{1}$$

In equation(1), G_T and G_R are the gains of the transmit and receive antennas respectively, along the specific direction that links them; c is the velocity of light in free space in meters per second; f is the frequency in hertz and d is the straight-line distance between the two antennas, in meters.

The Friis formula assumes that the receive antenna is perfectly conjugate matched. If otherwise, the mismatch factors should be considered and added to (1) [8].

However, Friis formula is constrained to the far field. If the separation d, in (1) is continually decreased, eventually the received power appears greater than the transmitted power. Obviously impossible in reality, since free space is not an amplifier.

B. Path loss expressed via the impedance matrix

Assume that two arbitrary antennas are separated by distance d in free space, as depicted in Fig. 1, the transmit antenna is fed by a source with an internal impedance $Z_S=R_S+jX_S$ and that the receive antenna is terminated by a load impedance $Z_L=R_L+jX_L$.

This dual antenna system can be equated to a two-port network, also depicted in Fig. 1, provided that port 1 of the network has the same port voltage V_1 and current I_1 as the input terminals of the dual antenna system and V_2 and current I_2 on port2 of the network are same as those at the output terminal of the dual antenna system.

The port parameters of the network are related by an impedance matrix [Z] [7]

$$\begin{pmatrix} V_1 \\ -Z_L I_2 \end{pmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix} \tag{2}$$

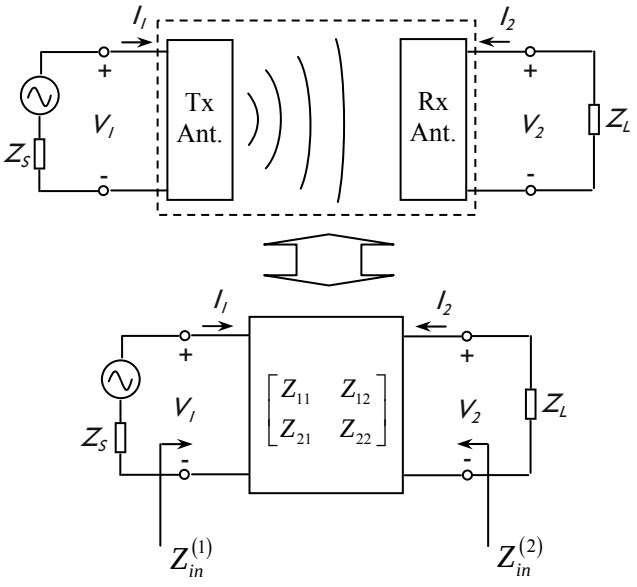


Fig. 1 Equivalent two-port network of a dual antenna system

$Z_{11}=R_{11}+jX_{11}$ and $Z_{22}=R_{22}+jX_{22}$ are the self-impedances of the network at ports 1 & 2 respectively. Z_{12} and Z_{21} are the mutual impedances of the two antennas as well as of the network between ports 1 & 2.

Because of the equivalent relation between the two port network and the dual antenna system as shown in Fig. 1, the power that flows into port 1 of the 2-port network, $\mathrm{Re}(V_1 I_1^*)$, is equal to the power absorbed at the input terminal of the dual antenna system. Similarly, the power that is delivered to the load through port 2 of the 2-port network, $|I_2|^2\mathrm{Re}(Z_L)$, is equal to the power received and delivered to the load of the dual antenna system by the receive antenna. In addition, assuming that there is no ohmic loss on either antenna of the dual antenna system, all the power absorbed by the transmit antenna is radiated. Hence,

$$P_{transmitted} = \mathrm{Re}\left(V_1 I_1^*\right) = \mathrm{Re}\left(Z_{11} - \frac{Z_{12}^2}{Z_L + Z_{22}}\right)|I_1|^2 \tag{3}$$

$$P_{received} = |I_2|^2 \,\mathrm{Re}(Z_L) = \mathrm{Re}(Z_L)\frac{|Z_{12}|^2}{|Z_L + Z_{22}|^2}|I_1|^2 \tag{4}$$

Thus the ratio of received power to transmitted power (in other words, the path loss if considered in decibels) is

$$\frac{P_{received}}{P_{transmitted}} = \frac{R_L\left(R_{12}^2 + X_{12}^2\right)}{A - B - C} \tag{5}$$

where,

$$A = R_{11}\left[\left(R_L + R_{22}\right)^2 + \left(X_L + X_{22}\right)^2\right] \tag{6}$$

$$B = \left(R_{12}^2 - X_{12}^2\right)\left(R_L + R_{22}\right) \tag{7}$$

$$C = 2R_{12}X_{12}\left(X_L + X_{22}\right) \tag{8}$$

Amongst other factors, the amount of power received will vary with different values of load impedances Z_L (R_L and X_L) in equation(5). In accordance with the assumption in Friis formula, Z_L is chosen to be the conjugated value of input impedance at port 2, $Z_{in}^{(2)}$, to ensure maximum power reception.

The input impedance at port 2, $Z_{in}^{(2)}$, can be derived from the impedance matrix (2) when port 1 is terminated by source impedance Z_S.

$$Z_{in}^{(2)} = \frac{V_2}{I_2} = Z_{22} - \frac{Z_{12}^2}{Z_S + Z_{11}} \tag{9}$$

Thus, load impedance Z_L is the conjugated value of (9)

$$Z_L = \left[Z_{in}^{(2)}\right]^* = \left(Z_{22} - \frac{Z_{12}^2}{Z_S + Z_{11}}\right)^* \tag{10}$$

According to equation (5), even though Z_L is known, we still need to evaluate the self and mutual impedances of the transmit and receive antennas. However, it is not easy to have analytical expressions of self and mutual impedance of an arbitrary antenna in practice, except for a few simple antennas.

To illustrate the theory of expression (5), this paper will work with two identical dipoles, whose parameters are obtainable and well known. We need to obtain expressions for Z_{21}, Z_{12}, Z_{11} and Z_{22}.

The mutual impedance Z_{21} of two dipoles has been modeled in literature for three different placement positions, viz. side-by-side, collinear and staggered [9]. When two dipoles of length L are placed side-by-side in parallel and separated by distance d, the mutual impedance Z_{21} is given by

$$R_{21} = 30\left\{2\mathrm{Ci}(\beta d) - \mathrm{Ci}\left[\beta\left(\sqrt{d^2 + L^2} + L\right)\right] - \mathrm{Ci}\left[\beta\left(\sqrt{d^2 + L^2} - L\right)\right]\right\} \tag{11}$$

$$X_{21} = 30\left\{2\mathrm{Si}(\beta d) - \mathrm{Si}\left[\beta\left(\sqrt{d^2 + L^2} + L\right)\right] - \mathrm{Si}\left[\beta\left(\sqrt{d^2 + L^2} - L\right)\right]\right\} \tag{12}$$

where C_i and S_i represent cosine and sine integration respectively; $\beta = 2\pi/\lambda$ is the wave number in free space. The expression for Z_{12}, i.e. for R_{12} and X_{12} would be the same as (11) and (12) respectively, since the dual antenna system is reciprocal.

Next, we evaluate Z_{11} and Z_{22}. The lengths of our two dipoles are identical and are set to half of the working frequency wavelength (i.e. They are resonant at the working frequency). In this case, the typical value of the input impedance is *73+j42.5* ohms [9]; hence both R_{11} and R_{22} were chosen to be *73ohms*, and X_{11} and X_{22} are set to *-42.5 ohms*.

C. Path loss expressed by S-parameters

The S parameter results of a two-port network obtained from simulators or a Network Analyzer (NA) are valid under the condition that the requisite ports are matched by the characteristic impedance of the transmission line. However, this assumption is not equivalent to the conjugate-match assumption of Friis transmission formula. In order to align the two obtained power delivery ratios, the S parameters obtained from a simulator or a NA have to be processed as follows.

The S_{11} of a two-port reciprocal network is

$$\left|S_{11}\right|^2 = \left.\frac{P_{reflected}^{(1)}}{P_{incident}^{(1)}}\right|_{Z_{Load}^{(2)} = Z_C} \tag{13}$$

where $P_{reflected}^{(1)}$ is the reflected power at port 1, $P_{incident}^{(1)}$ is the incident power on port 1, Z_C is the characteristic impedance of the transmission line and $Z_{Load}^{(2)}$ is the impedance that terminates port 2. Similarly,

$$\left|S_{22}\right|^2 = \left.\frac{P_{reflected}^{(2)}}{P_{incident}^{(2)}}\right|_{Z_{Load}^{(1)} = Z_C} \tag{14}$$

and

$$\left|S_{21}\right|^2 = \left|S_{12}\right|^2 = \left.\frac{P_{received}^{(2)}}{P_{incident}^{(1)or(2)}}\right|_{Z_{Load}^{(2)or(1)} = Z_C} \tag{15}$$

The transmitted power from port 1 expressed in term of S-parameters is

$$P_{transmit}^{(1)} = P_{incident}^{(1)} - P_{reflected}^{(1)} = P_{incident}^{(1)}\left(1 - \left|S_{11}\right|^2\right) \tag{16}$$

The received power at port 2 is the value that is delivered to the load by the receive antenna after considering the loss due to reflection caused by the mismatch between the receive antenna and the transmission line at the receive end. The received power at port 2 is then given by

$$P_{received}^{(2)} = P_{conjugate\ matched}^{(2)}\left(1 - \left|S_{22}\right|^2\right) \tag{17}$$

Thus, the ratio of the maximum received power to the transmitted power is

$$\frac{P_{conjugate\ matched}^{(2)}}{P_{transmitted}^{(1)}} = \frac{\left|S_{21}\right|^2}{\left(1 - \left|S_{11}\right|^2\right)\left(1 - \left|S_{22}\right|^2\right)} \tag{18}$$

Readings obtained from an NA or a simulator must be processed using the above relation to obtain a value of path loss that has assumptions coherent with Friis formula.

III. Comparison of Described Path Loss Models

The path loss between a receive and transmit antenna has been described using three different methods in the previous section. In order to verify the consistency between them, they were applied and compared via simulations and experiment-based verifications.

Two parallel half-wavelength dipoles in free space were modeled in XFDTD™ with suitably configured absorbing boundary conditions. Simulations were carried out at different separations between the dipoles, ranging from 1mm to around 85cm, at frequencies between 868MHz and 3GHz. The S-parameter results of the simulations were processed using (18).

Experiments were carried out to verify the accuracy of the analytical and simulated results. Two coaxial probes were mounted on a slider apparatus and connected to a NA for measurement of self and mutual coupling in free space. Measurements were conducted from 300 MHz up to 3 GHz, at different spacing, varying from 35mm to around 300mm. The resulting S-parameter readouts from the NA were processed using the expression in (18) to obtain the theoretical path loss. In addition, a uniform offset was included in the experiment results across the distance scale to account for the difference in gain between the measurement probes and the radiating dipole structures used in the analytical models and simulations.

The processed path loss obtained from the simulator was plotted on a graph using asterisks, as seen in Fig. 2. The path loss obtained from the measurement exercise was plotted as a trace formed by crosses. The hyphenated trace represents plots of unmodified Friis formula at the same frequencies, using the typical value of half-wavelength-dipole gain, $G_T=G_R=1.64$. The solid trace is the plot of the analytical path loss model as developed in (5).

The simulation and experiment results matched very closely to the analytical model at both frequencies, and at both the far field as well as the near field.

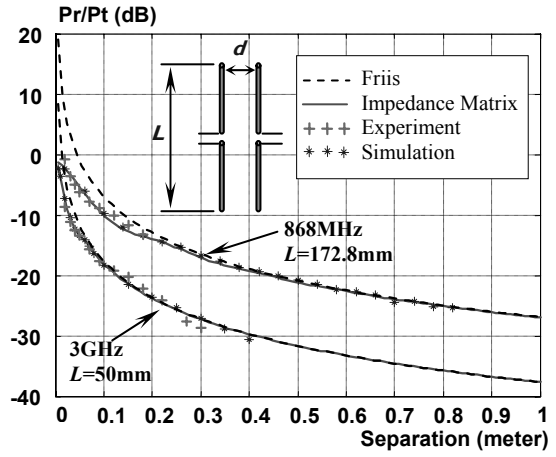


Fig. 2 Path loss between two dipoles at two frequencies

It was noted as expected, that while the path loss described by Friis formula converged with all the other curves at higher frequencies and in the far field, it had a difference in the near field when compared to all the other verification methods. At a fixed frequency, the error in Friis formula increased as the separation distance decreased. At a fixed separation distance, the error in Friis formula increased as the frequency decreased. Notably, all the methods converged in the far field at any frequency.

The analytical impedance method as well as the simulation results as modified by the S-parameter method converge in both the near field and the far field.

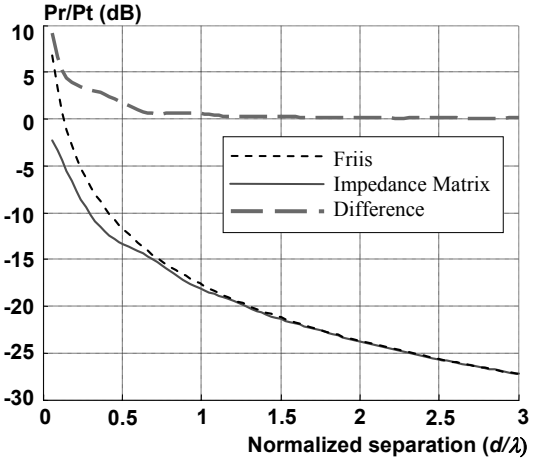


Fig. 3. Difference between Friis and other path loss methods

So the question arises - until what distance could the Friis formula be safely used? The Friis formula was plotted in the form of a hyphenated trace, along with the proven analytical expression plotted as a solid trace, on a scale of d/λ, as seen in Fig. 3; d being the distance of separation

between the transmit and receive antennas and λ the wavelength of the frequency of operation, in free space. The difference in value between the two traces is depicted by a thicker hyphenated trace. It was noticed that for the dipole structures we used for simulation and analysis, the difference between Friis formula and the actual path loss was negligible beyond a d/λ value of around 1.5.

As mentioned earlier, using Friis formula even to estimate the path loss in the near field could be a fatal design flaw, during selection of specification for design of the transceiver for communication in free space, since below a certain separation distance in wavelength terms, a significant difference is noted in the path loss as expressed by Friis and as proven by experiments and theory.

For arbitrary antennas being used in a communication system that will predominantly operate in its near field zone in free space (such as in many medical sensor systems), it is recommended that the transmit and receive antennas be designed first. An accurate simulation with correctly configured boundary conditions and parameters should then follow, with its results being processed using equation (18). This will allow for design using the optimum transmission power requirements, and facilitate correct selection of other parameters such as receive sensitivity and related margins.

IV. Conclusion

An analytical method to express path loss in terms of the impedance parameters has been presented. This method is unique in that it is a single expression that is valid in the near field as well as the far field. It was shown that at a fixed frequency, the error in Friis formula increases as the separation distance between the transmit antenna and receive antenna decreases. At a fixed separation distance, the error in Friis formula increases as the frequency decreases. This is important to be aware of while choosing design parameters for transceivers that will be used in applications operating at short ranges as compared to wavelength at the frequency of operation.

The analytical impedance method as well as the S-parameter method is valid in both near and far field.

If arbitrary antennas are used, an accurate simulation with correctly configured boundary conditions, and further processing of its S-parameters using the provided model would be the best way to find out the path loss.

Acknowledgment

The authors would like to offer sincere thanks to Xuecheng Qian for his many helpful suggestions and discussions. Thanks are also due to Yang Liu, Jianliang Wu and Jiwen Chen for their support during the measurement campaign.

References

1. Hall, P.S.; Ricci, M.; Hee, T.W., Characterization of on-body communication channels, Int. Conf. Microwave Millimeter Wave Technol., 2002, pp.770 - 772
2. Kamarudin, M.R.; Nechayev, Y.I.; Hall, P.S., Performance of antennas in the on-body environment, APS Int Symp Dig, Vol. 3A, July 2005, pp. 475 - 478, DOI 10.1109/APS.2005.1552290
3. Kamarudin, M.R.; Nechayev, Y.I.; Hall, P.S., Antennas for on-body communication systems, Proc. IEEE Int. Workshop Antenna Technol. Small Antennas Novel Metamater, 2005, pp:17 – 20, DOI 10.1109/IWAT.2005.1460985
4. Yates, D.C., Holmes, A.S., Burdett, A.J., Optimal transmission frequency for ultralow-power short-range radio links, IEEE Trans. Circuits Syst. Regul. Pap., Vol. 51, No. 7, 2004, pp. 1405 – 1413, DOI 10.1109/TCSI.2004.830696
5. Schantz, H.G., Near field propagation law & a novel fundamental limit to antenna gain versus size; APS Int Symp Dig, Vol. 3A, July 2005, pp.237 – 240, DOI 10.1109 /APS.2005.1552223
6. FitzGerrell, R., Standard linear antennas, 30 to 1000 MHz, IEEE Trans Antennas Propag, Vol. 34, No. 12, 1986, pp.1425 - 1429
7. Derneryd, A.; Kristensson, G., Signal correlation including antenna coupling, Electron. Lett., Vol. 40, No. 3, Feb. 2004, pp.157 – 159, DOI 10.1049/el:20040140
8. Warren L. Stutzman, Antenna Theory and Design, John Wiley & Sons
9. John D. Kraus and Ronald J. Marhefka, Antenna for All Applications, Third Edition, McGraw Hill

Address of the corresponding author:

Author: Yong Liu
Institute: Philips Research East Asia
Street: 218 Tian Mu Xi Road
City: Shanghai
Country: P.R.China
Email: wallace.liu@philips.com

ACOUSTIC MONITORING OF LUNG SOUNDS FOR THE DETECTION OF ONE LUNG INTUBATION

[1]Tejman-Yarden S, [2]Zlotnik A, [3]Weizman L, [3]Tabrikian J, [3]Cohen A, [2]Weksler N and [2]Gurman GM

[1]Division of Pediatrics, Soroka Medical Center, Beer-Sheva, Israel
[2]Division of Anesthesiology and Critical Care, Soroka Medical Center, Beer-Sheva, Israel
[3]Electrical and Computer Engineering Department, Ben Gurion University of the Negev, Beer-Sheva, Israel

Abstract— **Patient monitoring methods for the early diagnosis of one lung intubation (OLI) are non-specific and controversial. The aim of this study is to evaluate a new acoustic monitoring system for the detection of OLI. Lung sounds were collected from 24 adult surgical patients scheduled for routine surgical procedures. Four piezoelectric microphones attached to the patients' back were used to sample lung sounds during induction to anesthesia and tube positioning. To achieve OLI, the endotracheal tube was inserted and advanced down the airway so that diminished or no breath sounds were heard on the left side of the chest. The tube was then withdrawn stepwise until equal breath sounds were heard. Fiberoptic bronchoscopy confirmed the tube's final position. Acoustic analyses were preformed by a new algorithm which assumes a Multiple Input Multiple Output (MIMO) system, in which a multi-dimensional Auto-Regressive (AR) model relates the input (lungs) and the output (recorded sounds) and a classifier, based on a Generalized Likelihood Ratio Test (GLRT), indicates the number of ventilated lungs without retrieving the original lung sounds from the recorded samples. This algorithm achieved an OLI detection probability of 95.2% with a false alarm probability of 4.8%. Higher detection values can be achieved at the price of a higher incidence of false alarms.**

I. Introduction

Today, inadvertent endobronchial intubation or one lung intubation (OLI) is the most common complication of endotracheal intubation (1). OLI is particularly common during pediatric anesthesia or laparoscopic surgery (2,3) and it can lead to complications such as atelectasis, pneumothorax, hypoxemia, cardiac arrhythmia, and hypotension (4).

Lung auscultation, which is the most common method for the assessment of the position of the ET tube and capnography were found unreliable for the detection of OLI (4-7). Pulse oxymetry, by which most cases of OLI are detected, sets off its alarm only after the patient has become symptomatic and does not specifically indicate the actual cause of hypoxemia.

Analysis of lung sound signals to diagnose and monitor the respiratory system is well recognized. (12-15). Based on these works we conducted an initial study in which acoustic lung sounds were sampled from human patients undergoing lung procedures that required the insertion of a double lumen tube (16). In that study we sampled each patient for only right, only left and bilateral lung ventilation, using two piezoelectric microphones, one on each side of the chest. The sounds samples were band-pass filtered and for every respiration, the left and the right energy envelopes were calculated from each microphone separately, assuming that each microphone samples its ipsilateral lung. The diagnosis of the ventilation status was based on the energy ratios between the left and the right. This analysis algorithm yielded 92% recognition of selective right one lung ventilation with a sensitivity of only 90%.

In that study we found detection of lung sound bilaterally, during one lung ventilation. The use of a double lumen tube with inflated balloons demonstrated that significant air leak is not the cause of the phenomenon and that each microphone attached to the chest actually samples both lungs but with different ipsilateral and contralateral contributions. We concluded in that study that each lung does not produce a uniform sound; it produces various different sounds at different times and simultaneously receives transmitted sounds from the contralateral lung (16).

Hence we proceeded to develop a new algorithm which assumes a MIMO (Multi Input Multi Output) system, in which a multi-dimensional Auto-Regressive (AR) model relates the input (lungs) and the output (recorded sounds) and a classifier, based on a Generalized Likelihood Ratio Test (GLRT), indicates the number of ventilated lungs without retrieving the original lung sounds from the recorded samples (18,19).

II. Experiment structure

In order to examine the suggested method and to check the possibility of developing a monitoring tool in the future, a database of recorded breathings was established. The database was composed of 24 patients who were recorded in a surgery room in both situations: during correct ventilation,

when the tip of the tube is placed above the carina and during a situation of OLI.

During this study the microphones were attached to the patients' back, two on each side in parallel position, recorded the breathing sounds of the patients in both situations. The ventilations were performed manually and not mechanically, in order to achieve higher signal-to-noise ratio in the recorded sounds. The trials were performed in the main surgery room of medical center Soroka - Israel, during the anesthesia part in the beginning of the surgery.

III. Problem formulation

In order to determine the number of lungs that are ventilated from analysis of the recorded sounds, an AR (Auto-Regressive) model that relates the lungs and the microphones was assumed. In the first stage, we assume that each lung is a point source, and develop an OLI detector for this case. In the second stage the point source assumption is removed and the algorithm is modified accordingly, as will be discussed in section V.

Our goal is to estimate the number of sources from the received signals by the sensors. Let K and L denote the number of sources (lungs) and sensors (microphones), respectively ($K<L$). Therefore, the vector of source signals, $\mathbf{x}[n]$, is defined as a $K\times 1$ vector as follows:

$$\mathbf{x}[n]=\begin{bmatrix}x_1[n] & x_2[n] & \cdot & \cdot & \cdot & x_K[n]\end{bmatrix}^T. \quad (1)$$

And $\mathbf{y}[n]$, the $L\times 1$ vector of measurements is defined by:

$$\mathbf{y}[n]=\begin{bmatrix}y_1[n] & y_2[n] & \cdot & \cdot & \cdot & y_L[n]\end{bmatrix}^T \quad (2)$$

The relation between the source signals and the measurements is given by a MIMO (Multiple-Input Multiple-Output) AR model:

$$\mathbf{y}[n]=\mathbf{A}\mathbf{u}[n]+\mathbf{C}\mathbf{x}[n]+\mathbf{e}[n], \quad (3)$$

where $\mathbf{u}[n]$ is a $ML\times 1$ vector defined as follows:

$$\mathbf{u}[n]=\begin{bmatrix}\mathbf{u}_1^T[n] & \mathbf{u}_2^T[n] & \cdot & \cdot & \cdot & \mathbf{u}_L^T[n]\end{bmatrix}^T \quad (4)$$

and $\mathbf{u}_i[n]$ is a $M\times 1$ vector which contains the past values of the i-th sensor, $y_i[n]$, up to sample M:

$$\mathbf{u}_i[n]=\begin{bmatrix}y_i[n-1] & y_i[n-2] & \cdot & \cdot & \cdot & y_i[n-M]\end{bmatrix}^T. \quad (5)$$

$\mathbf{A}$ is an $L\times ML$ matrix defined as:

$$\mathbf{A}=\begin{bmatrix}\mathbf{a}_{11}^T & \cdot & \cdot & \cdot & \mathbf{a}_{1L}^T\\ \cdot & & & & \cdot\\ \cdot & & & & \cdot\\ \cdot & & & & \cdot\\ \mathbf{a}_{L1}^T & \cdot & \cdot & \cdot & \mathbf{a}_{LL}^T\end{bmatrix}, \quad (6)$$

where $\mathbf{a}_{ij}$ is a $M\times 1$ vector, which relates the n-th sample of the i-th sensor, $y_i[n]$, with the past values of the j-th sensor, $y_j[n\text{-}1]$, …, $y_j[n\text{-}M]$.

$\mathbf{C}$ is a $L\times K$ matrix whose elements, c_{ij}, relate the sensor i with the source j.

Finally, $\mathbf{e}[n]$ is a $L\times 1$ vector that representing the additive white Gaussian noise. We assume that the noise and source signals are independent, zero mean, Gaussian with covariance matrices $\sigma^2\mathbf{I}$ and $\mathbf{I}$ respectively. The last assumption is done without loosing of generality, because the covariance of the sources is controlled by the matrix $\mathbf{C}$, as it can clearly be seen from (3). As a result, we obtain that the conditional distribution of $\mathbf{y}[n]/\mathbf{u}[n]$ is Gaussian: $\mathbf{y}[n]/\mathbf{u}[n]\sim N(\mathbf{A}\mathbf{u}[n],\mathbf{R})$, where $\mathbf{R}=\mathbf{C}\mathbf{C}^T+\sigma^2\mathbf{I}$. We can note that the unknown parameters: $\mathbf{A},\mathbf{R},\sigma^2,M$ and K are need to be estimated from a set of N measurements, $\mathbf{y}[1],\ldots,\mathbf{y}[N]$, and its past values, $\mathbf{u}[1],\ldots,\mathbf{u}[N]$.

IV. The ML estimator

In order to determine the number of sources, K, we need first to estimate the unknown matrices, $\mathbf{A}$ and $\mathbf{R}$, from the N samples of the data: $\mathbf{y}[1],\ldots,\mathbf{y}[N]$. For this purpose, the Maximum–Likelihood (ML) estimator is used. The ML estimator of the matrices $\mathbf{A}$ and $\mathbf{R}$, is obtained by maximizing the conditional probability density function (pdf) of the output samples given its past values, which is:

$$f(\mathbf{y}[1],\ldots,\mathbf{y}[N]/\mathbf{u}[1],\ldots,\mathbf{u}[N];\mathbf{R},\mathbf{A})= \quad (7)$$
$$\frac{1}{(2\pi)^{LN/2}|\mathbf{R}|^{N/2}}\prod_{n=1}^{N}\exp\{-\tfrac{1}{2}(\mathbf{y}[n]-\mathbf{A}\mathbf{u}[n])^T\mathbf{R}^{-1}(\mathbf{y}[n]-\mathbf{A}\mathbf{u}[n])\}$$

The log-likelihood function can be maximized by equaling its derivations with respect to $\mathbf{R}$ and $\mathbf{A}$ and solving two matrix equations. This process yields:

$$\hat{\mathbf{A}}_{ML}=\left(\sum_{n=1}^{N}\mathbf{y}[n]\mathbf{u}^T[n]\right)\left(\sum_{n=1}^{N}\mathbf{u}[n]\mathbf{u}^T[n]\right)^{-1} \quad (8a)$$

and

$$\hat{\mathbf{R}}_{ML}=\tfrac{1}{N}\sum_{n=1}^{N}\mathbf{y}[n]\mathbf{y}^T[n]-\hat{\mathbf{A}}_{ML}\tfrac{1}{N}\sum_{n=1}^{N}\mathbf{u}[n]\mathbf{y}^T[n] \quad (8b)$$

Substituting the maximum likelihood estimates (8) in the log-likelihood, with some straightforward manipulations, we obtain:

$$\log\left[f\left(\mathbf{y}[1],..., \mathbf{y}[N]/\mathbf{u}[1],..., \mathbf{u}[N]\right)\right] = \quad (9)$$
$$-\tfrac{LN}{2}\log(2\pi) - \tfrac{N}{2}\log\left|\hat{\mathbf{R}}\right| - \tfrac{1}{2}NL$$

In order to be avoided from over estimating of *K* and *M,* a penalty function that depends on the free parameters of the model should be added to (9).

V. Modification of the Solution for Realistic Model

It can be noted that the only expression that depends on *K* and *M* in (9) is $|\hat{\mathbf{R}}|$, which is actually the product of the eigenvalues of $\hat{\mathbf{R}}$. According to (22), some of these eigenvalues represent the energy of the source signals (marked as l_i) and some of them represent the noise level (marked as σ^2). According to the selected model and [9], under each hypothesis (OLI or TRI) the eigenvalues of $\mathbf{R}$ are expected to be:

$$\text{TRI}: l_1, l_2, \underbrace{\sigma^2, \sigma^2, ..., \sigma^2}_{L-2}, \quad (10a)$$

$$\text{OLI}: l_1, \underbrace{\sigma^2, \sigma^2, \sigma^2, ..., \sigma^2}_{L-1}. \quad (10b)$$

As it can be seen from (10), the second highest eigenvalue of $\mathbf{R}$ can be used as an indicator for OLI situation.

Because of the expected mismatches between the selected model and the real one, even during cases of OLI the second highest eigenvalue of $\mathbf{R}$ is higher than the noise level, σ^2. The reason for this, is that the lungs are not point sources as they were treated in the selected model. In fact, each lung is composed of several distributed and independent point sources whose energies are distributed over all of the eigenvalues of $\mathbf{R}$. Therefore, even during OLI situation the second highest eigenvalue is above the noise level.

During OLI situation only one lung is ventilated and therefore, the point sources are less scattered than in case of TRI where both lungs are ventilated. This fact causes the energy of these point sources to be less scattered over the eigenvalues of $\mathbf{R}$. Therefore, in OLI case the second highest eigenvalue of $\mathbf{R}$ value is smaller than in TRI case. This is the reason why choosing the second highest eigenvalue as a detector provides a reliable detection of OLI situations, as will be shown in the next section.

VI. Results

Fig. 1 shows a few breathing cycles, of both OLI and TRI situations, recorded by the four microphones. As it can be seen from this figure, determination between OLI and TRI cases by only the amplitude of the recorded sounds is not a simple task, as explained in section II.

The breathing signals were limited to cut-off frequency of 4kHz and the data were divided into windows of 2000 samples each, with 80% of overlapping. An arbitrary AR order of 15 was set, and the unknown matrices **A** and **R** were estimated for each window. Fig. 2 shows the second highest eigenvalues of $\hat{\mathbf{R}}$ as a function of time. As it can clearly be seen from Fig. 2, we can easily decide between OLI and TRI cases, by the value of the second highest eigenvalue in every breathing. The results of the proposed algorithm were consistent over 24 experiments made on patients as explained in section II.

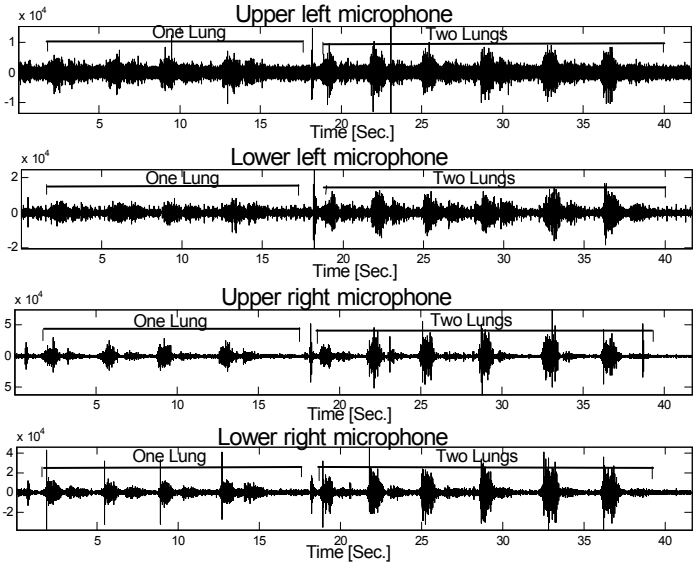


Fig. 1 Some recorded breathing cycles by the four microphones, OLI and TRI cases are included.

The curves that represent the distribution of the second highest eigenvalue under each hypothesis are shown in Fig. 3. From these curves we can draw the Detection Error Tradeoff (DET) curve, which is the probability of miss, P_{miss}, as a function of probability of false-alarm, P_{FA}. The DET curve provides information about the receiver performance, where each point on the curve shows the P_{FA} and P_{miss} for a given threshold. In Fig. 4 we can see the DET curve of the proposed decision system. The Equal Error Rate (EER) point which is defined as the point on the DET curve where $P_{miss} = P_{FA}$ is 4.8%.

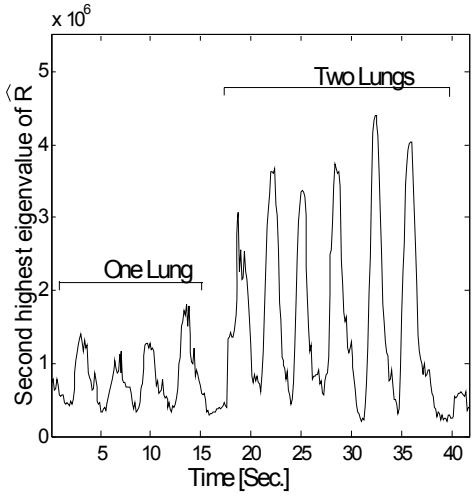


Fig. 2 Second eigenvalue of $\hat{\mathbf{R}}$ as a function of time. OLI and TRI cases can clearly be determined.

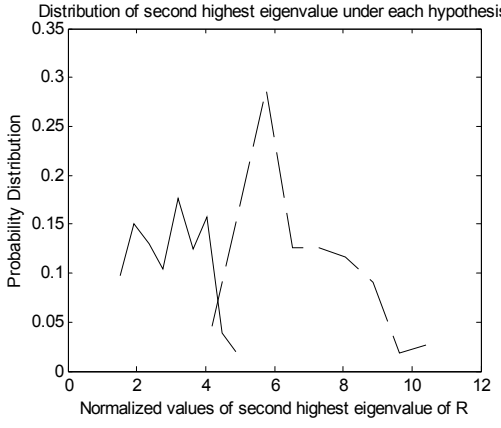
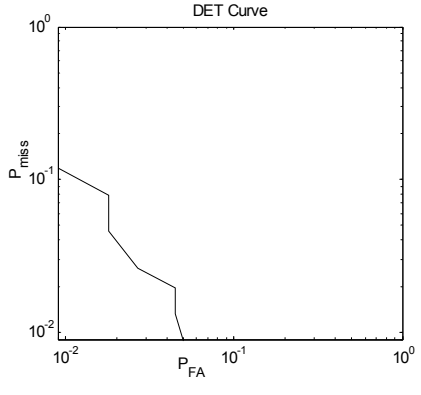


Fig. 3 The distribution of the second eigenvalue. The left curve represents the distribution of the normalized values when OLI situation occurs, and the right curve represents the values on TRI situation.

Fig. 4 The DET of a classifier based on the second highest eigenvalue of estimated **R**:



VII. Discussion and conclusions

This paper endeavored to take the OLI detection research one step forward by showing the concept of a new automatic system for detection of OLI. An algorithm for detection of OLI by monitoring lungs sounds was developed and tested. It has been shown that assuming a MIMO AR model and selecting the second highest eigenvalue of the residual covariance matrix as a feature proves itself as a reliable method for detection of OLI.

References

1. Szekely SM, Webb RK, Williamson JA, Russell WJ. The Australian Incident Monitoring Study. Problems related to the endotracheal tube: an analysis of 2000 incident reports. Anesth Intensive Care 1993;21:611-6.
2. Rivera R, Tibballs J. Complications of endotracheal intubation and mechanical ventilation in infants and children. Crit Care Med 1992;20:193-9.
3. Joshi GP. Complications of laparoscopy. Anesthesiol Clin North America 2001;19:89-105.
4. Brunel W, Coleman DL, Schwartz DE et al. Assessment of routine chest roentgenograms and the physical examination to confirm endotracheal tube position. Chest 1989;96:1043-5.
5. Klepper ID, Webb RK, Van der Walt JH et al. The Australian Incident Monitoring Study. The stethoscope: applications and limitations--an analysis of 2000 incident reports. Anesth Intensive Care 1993;21:575-8.
6. Schwartz DE, Lieberman JA, Cohen NH. Women are at greater risk than men for malpositioning of the endotracheal tube after emergent intubation. Crit Care Med 1994;22:1127-31.
7. Sugiyama K, Yokoyama K. Reliability of auscultation of bilateral breath sounds in confirming endotracheal tube position. Anesthesiology 1995;83:1373.
8. Heneghan CP, Scallan MJ, Branthwaite MA. End-tidal carbon dioxide during thoracotomy. Its relation to blood level in adults and children. Anaesthesia 1981;36:1017-21.
9. Johnson DH, Chang PC, Hurst TS et al. Changes in PETCO2 and pulmonary blood flow after bronchial occlusion in dogs. Can J Anaesth 1992;39:184-91.
10. Webb RK, van der Walt JH, Runciman WB et al. The Australian Incident Monitoring Study. Which monitor? An analysis of 2000 incident reports. Anesth Intensive Care 1993;21:529-42.
11. Webster TA. Now that we have pulse oximeters and capnographs, we don't need precordial and esophageal stethoscopes. J Clin Monit 1987;3:191-2.
12. Cohen A, Berstein AD. Acoustic transmission of the respiratory system using speech stimulation. IEEE Trans Biomed Eng 1991;38:126-32.
13. Cohen A, Landsberg D. Analysis and automatic classification of breath sounds. IEEE Trans Biomed Eng 1984;31:585-90.
14. Sod-Moriah G. Ventilation monitoring during anesthesia and respiratory intensive care Department of Electrical and Computer Engineering. Beer-Sheva, Israel: Ben-Gurion University of the Negev, 1995.
15. O'Connor CJ, Mansy H, Balk RA, Tuman KJ, Sandler RH. Identification of endotracheal tube malpositions using computerized analysis of breath sounds via electronic stethoscopes. Anesth Analg. 2005 Sep;101(3):735-9.
16. Tejman-Yarden S, Lederman D, Eilig I, Zlotnik A, Weksler N, Cohen A Gurman GM. Acoustic monitoring of double lumen ventilated lungs for the detection of selective unilateral lung ventilation. Anesth Analg. (in press)
17. Heinonen J, Takki S, Tammisto T. Effect of the Trendelenburg tilt and other procedures on the position of endotracheal tubes. Lancet 1969;1:850-3.

18. Weizman L. ,*Detection on one-lung intubation incidents* ,M.Sc. Thesis Ben-Gurion University of the Negev, Department of Electrical and Computer Engineering , Israel, August 2004.
19. Weizman L, Tabrikian J, Cohen A. Detection of one lung intubation by monitoring lung sounds IEEE In Proc. 2004 Intl. Conference of the Engineering in Medicine and Biology Society (EMBC 2004), San Francisco, CA, pp. 917-920
20. Porat B. Digital processing of random signals, theory and methods Englewood Cliffs, N.J.: Prentice-Hall, 1994.
21. Proakis JG, Manolakis DG. Digital signal processing: principles, algorithms, and applications: MacmIllan, 1996.
22. Wax M, Kailath T. Detection of signals by information theoretic criteria. IEEE Trans Accoust Speech, Signal Processing 1994;ASSP-33:387-92.

Automatic Step Detection in the Accelerometer Signal

H. Ying[1], C. Silex [1], A. Schnitzer[1], S. Leonhardt[2], and M. Schiek[1]

[1] Central Institute for Electronics, Research Centre Jülich, Germany
[2] Philips Chair for Medical Information Technology, RWTH Aachen University, Germany

Abstract— **The automatic step detection is a crucial component for the analysis of vegetative locomotor coordination during monitoring the patients with Parkinson's disease. It is aimed to develop the algorithms for automatic step detection in the accelerometer signal, which will be integrated in sensor networks for neurological rehabilitation research. In this paper, three algorithms (Pan-Tompkins method, template matching method and peak detection based on combined dual-axial signals) are detailed described. Finally, these methods will be discussed by means of dis- and advantages.**

Keywords— **Step detection, accelerometer, Pan-Tompkins method, template-matching method, neurological rehabilitation research**

I. Introduction

Parkinson's disease (PD) is associated with reduced coordination between respiration and locomotion [1]. For the neurological rehabilitation research, it requires a long-time monitoring system, which enables the online analysis of the patients' vegetative locomotor coordination. The online analysis allows the later integration of bio-feedback protocol for the rehabilitation purpose. This system will facilitate the identification of the therapeutic effect and measuring the patients' health status. To monitor the phase synchronization between respiration and locomotion, the accurate time of step event has to be determined. The commercial pedometers, which count each step a person makes, can not fulfill this requirement. For this purpose, the online algorithm is developed for the automatic step detection. Since the characteristics of the PD patients' gait are pronounced different from the normal gait [2], the database for algorithm development in our work is acquired only from the PD patients.

The first experiment on the locomotion measuring is carried out using accelerometer. The accelerometer signals show considerable difference in morphology and amplitude among the individuals. We attempt to design a fast and robust algorithm, which should be suitable for individual patients, without any user-specified parameters. So far, three methods have been investigated for this project:

- Pan-Tompkins Method
- Template-Matching Method
- Peak-detection method based on combined dual-axial signals.

These three algorithms will be described in detail below, as well as their preliminary results on the patients' data. Finally, the dis- and advantages of these methods will be discussed.

II. Methodology

A. Subjects and Equipment

Subjects: In total, eight patients suffering from PD were measured in the Clinic Ambrock for Neurology, Centre for Sleep- and Rehabilitation Research, Germany. The patients are at different stage from 1 through 4 according to unified rating scale (UPDRS). To evaluate the algorithms, data in total duration of 115 minutes were collected from eight recordings. They are acquired at a sample rate of 200 Hz. The patients were walking on a treadmill with accelerometers attaching on the lateral side of the left and right feet during the measurement. Meanwhile, the abdominal and thoracic respiration signals were measured and acquired with a sample rate of 20 Hz.

Equipment: The horizontal and vertical acceleration in x- and z-axis generated from the locomotion during steps is measured by ADXL322, a low power dual-axis accelerometer, produced by Analog Devices, Inc. The schematic circuit diagram of the accelerometer for activity measuring is provided in Fig.1 below. The accelerometer is supplied by battery voltage through a low-dropout voltage regulator (TPS77027) to maintain constant 2.7 V level. To improve transient response and noise rejection, a 1 μF ceramic capacitor (C_{DC}) is connected between V_{RIN} and GND. The voltage regulator requires a 4.7 μF capacitor (C_{ROUT}) connected between V_{ROUT} and GND. The output of regulator is brought into ADXL 322 through pin Vs. The ADXL322 contains two 32 $k\Omega$ resistors (R_X and R_Y) in each output inside IC, respectively. The 0.1 μF capacitors (C_X and C_Y) are added at the pins X_{OUT} and Y_{OUT} to implement passive low-pass filtering of 50Hz, first order, to eliminate high frequent noise. The output signals: V_{AOUTX} and V_{AOUTY} are analog voltages that are proportional to acceleration.

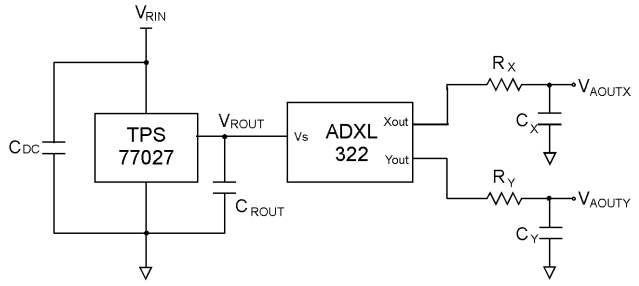


Fig. 1 Schematic diagram of the accelerometer application circuit

Fig. 2 depicts the acceleration of three patients on the left feet in the horizontal (x-axial) and vertical (z-axial) direction during walking. Each column of figures shows the data of one patient. The morphology of the acceleration, which depends upon the posture and gait of the individuals, shows great difference among the patients. The accelerometer signal on the left foot is similar with the signal on the right foot. Observe from Fig. 2 that normally, one step consists of two positive peaks, which occur when the foot lifts off the ground and heel strikes the ground, respectively. Either of them is the significant feature of steps in the time domain. The peaks of varied amplitude and morphology are attempted to be detected using the following three algorithms.

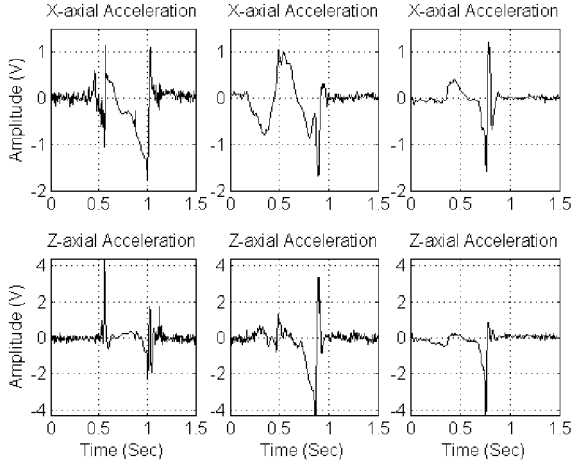


Fig. 2 The acceleration signal measured on the left feet of three patients in the horizontal and vertical directions

B. Algorithm Overview

I Pan-Tompkins method: Pan and Tompkins proposed a real-time algorithm for detection of R peak in ECG signal [3],[4]. An experiment is made to investigate, whether this algorithm can also be used for step detection in the accelerometer signal. This algorithm includes a series of filters and methods that perform low-pass, derivative, squaring, integration for preprocess and adaptive thresholds for peak-searching. Fig. 3 illustrates the steps of the algorithm in schematic form. And the results of each step processed on a signal segment of 6 s as shown in Fig. (a) are plotted in Fig. 4 (b) through (g) in sequence.

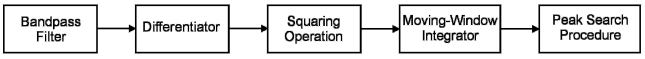


Fig. 3 Block diagram of the Pan-Tompkins algorithm

Bandpass-filter: The bandpass filter reduces the influence of artifacts in the signal. In this work, the highpass filter is not applied. The digital lowpass filters with small integer coefficients are designed for fast execution. The following lowpass filter with a cutoff frequency of 20 Hz was applied:

$$H(z) = \frac{1}{16}\frac{(1-z^{-4})^2}{(1-z^{-1})^2} \qquad (1)$$

Derivative operator: The derivative operator is specified as:

$$y(n) = \frac{1}{8}[2x(n) + x(n-1) - x(n-3) - 2x(n-4)] \qquad (2)$$

Note that eqn. (2) approximates the ideal derivative operator up to 30 Hz. Fig. 4 (b) illustrates the effect of the derivative, i.e., to suppress the low-frequency components and enlarge the high frequency components from the high slopes.

Squaring: The squaring operation leads to positive result and enhances large values more than small values. As shown in Fig. 4 (c), the squaring operator increases the high-frequency components further.

Integration: The output resulted from the preceding operation in Fig 4 (c) exhibits multiple peaks and hence needs to be smoothed. It is smoothed through a moving-window integration filtered:

$$y(n) = \frac{1}{N}[x(n-(N-1)) + \dots \\ x(n-(N-2)) + \cdots + x(n)] \qquad (3)$$

where N is chosen to be 20 empirically. The Fig. 4 (e) depicts the preprocessed signal beneath the original one in Fig 4 (a) for comparison. Each step cycle yields double peaks with monotonic ascend and descend.

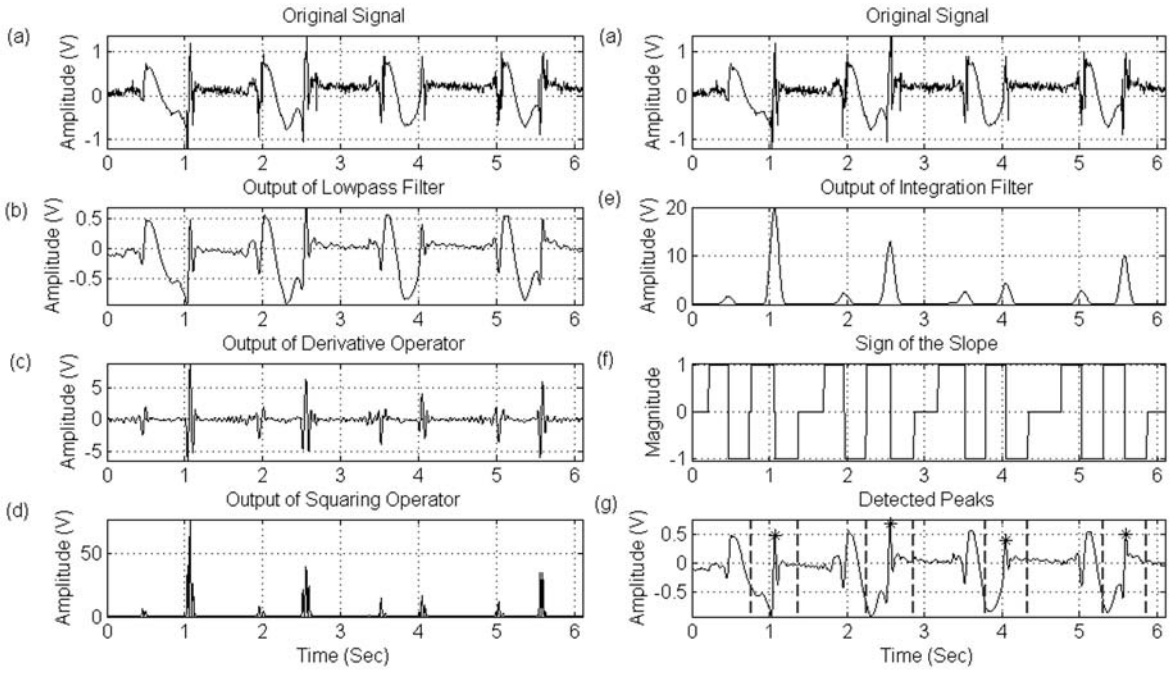


Fig. 4 Results of the Pan-Tompkins method on the x-axial acceleration of 6 s measured on the left foot in the step procedure

Fiducial mark: The so-called fiducial point, which is defined as location of the peak, is detected using an adaptive threshold within the Pan-Tompkins method. Unfortunately, the magnitude of the peaks in the accelerometer signal varies with time, stance of walking and the struck intensity of the foot on the ground, etc. Furthermore, the mathematical transform amplifies the difference in magnitude remarkably. Therefore, the peak-searching algorithm using threshold is not adequate for this solution. In this work, the property of the successive ascends and descends in the transformed signal is utilized for peak detection. The ascend and descend result in positive and negative slope, respectively. Fig. 4 (f) demonstrates sign of the slope arising from the preprocessed signal, in which the positive slope is equal to 1, whereas the negative one is transformed to -1. In this way, one step cycle signal is converted into a pair of consecutive 1 and -1 value, i.e., [1 -1 1 -1] in Fig. 4 (f). The second [1 -1] interval, corresponding to the onset and offset of the second peak in Fig 4 (e) is defined as peak-searching interval, which is plotted with dashed line in Fig. 4 (g). The local maximum within the peak-searching interval on the filtered signal is detected as step, which indicates the struck of feet on the ground, and marked with asterisk in Fig. 4 (g).

II Template-matching method: The main concept of the template-matching method is to generate a template, which represents a typical step cycle. In the unknown signal, an event is declared to be detected when there is a match between the signal and the template to certain degree. Fig. 5 illustrates a flow chart in schematic form explaining the template matching method.

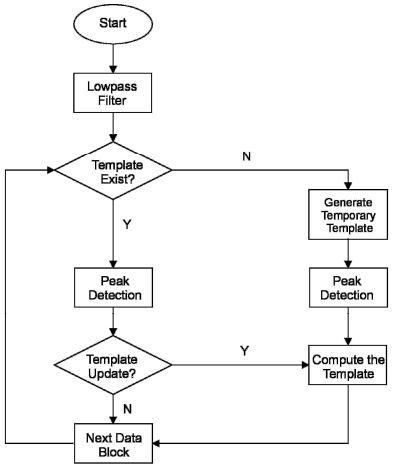


Fig. 5 Flowchart of the template matching method

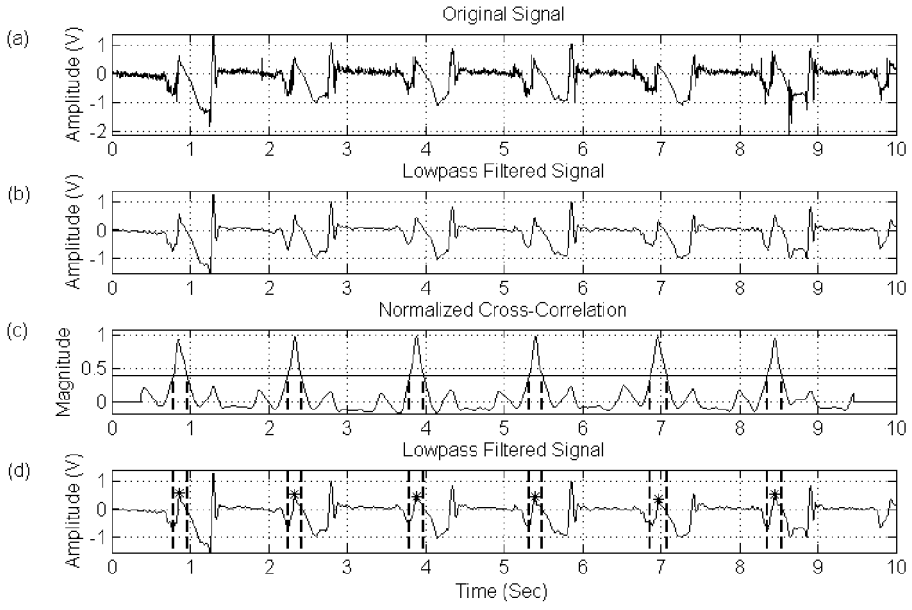


Fig. 6 Results of template matching on a data block of x-axial acceleration on the left foot in step procedure

Initially, the whole recording is broken into several non-overlapping data blocks of 10 seconds each. Fig 6 (a) depicts a data block measured on the left foot in the x-axis of 10 seconds, containing six step cycles. Then the signal is filtered by a lowpass filter with cutoff frequency of 20 Hz as shown in Fig. 6 (b).

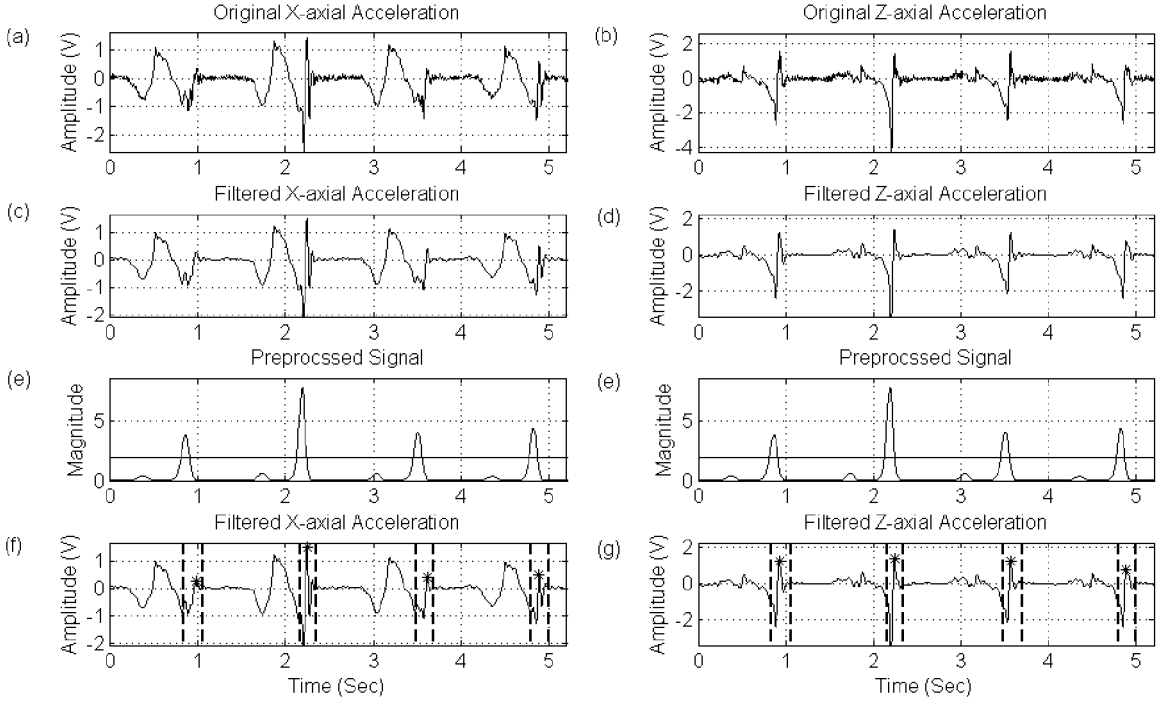
Next, the algorithm will check whether there is any template present in memory. If not present, the first step cycle is extracted as a temporary template. All the parameters are approximated, since the characteristic of the step is unknown. This template is slid across the whole data block and the normalized cross-correlation, is calculated between the template and signal, shown as in Fig 6 (c). The normalized cross-correlation indicates the similarity between two vectors X and Y, which is given in the eqn. (4) [5]:

$$R_N[k] = \frac{< X, Y >}{\sqrt{\|X\| \cdot \|Y\|}} = \frac{R_{XY}(k)}{\sqrt{R_{XX}(0) \cdot R_{YY}(0)}} \qquad (4)$$

where $< X \cdot Y >$ is the inner product of X and Y, $\|X\|$ is the norm of the Vector X, $R_{XY}(k)$ is cross-correlation of X and Y for arbitrary k, and $R_{XX}(0)$ is autocorrelation of X[n] at point of zero.

The bound value for the maximum is 1 for absolute identity, which allows setting an uniform threshold for all the data despite different amplitudes. The peaks in the normalized cross-correlation in Fig. 6 (c) indicate great similarity between the step template and signal segment, thus the occurrence of the event. The interval, in which the cross-correlation exceeds the threshold 0.4, is defined as peak-searching interval. It is marked with dashed line in Fig 6 (c) and (d). The threshold is plotted as solid straight line in Fig 6. (c). The local maxima falling within the peak-searching interval in the filtered signal are marked as fiducial points of the steps, and symbolized with asterisk in Fig 6 (d). Using the template matching method, the first positive peaks of steps, which occur when the feet lift off the ground, are detected. A more representative template should be generated instead of the temporary template, when the peaks in the first data block are detected. These six step cycles are aligned with the peaks and averaged together to generate a new template, which will be applied for the further data processing.

If it is decided that a template has already been present, the peaks are detected using the same solution procedures stated above. Before the next data block is processed using the same template, a determination need to be made as to whether the template will be updated. The step signal may change dynamically with time; accordingly the template may not represent the current step signal. If the major peaks in the normalized cross-correlation are lower than 0.55, the new template is generated using the steps cycles in the current data block. Otherwise the peak detection is done in the next data block. The algorithm stops when the size of the recording is reached.



F*ig. 7* Results of the method based on the combined signals processed on the x- and z-axial acceleration.

III Peak-detection method based on combined dual-axial signals: Observe from the acceleration signal of each step that the negative wave in the x-axial signal occurs coincident with the negative wave in the z-axial signal. Sometimes, the negative peak due to the deceleration before the foot hitting the ground within one step is more apparent than the positive peak. The third method is based on the coincident negative wave in the x- and z-axial acceleration signal. Fig. 8. depicts the entire procedure of the algorithm. And this block diagram will be explained accompanied with an example of the acceleration signal segments of 5 s measured on the left foot in the x- and z-axial direction, which are shown in Fig 7 (a) and (b), respectively.

F*ig. 8* Block diagram of the peak detection algorithm on the combined dual-axial signal

The preprocess procedure consisted of bandpass filtering, summing up negative elements, a moving-window integrator and a squaring operation. At first, both signals are passed through a lowpass filter with cutoff frequency 20 Hz. The out-coming results are depicts in Fig 7 (c) and (d), respectively. According to the results of spectral analysis, highpass filter is not applied. Then the positive elements in both arrays are set to zero, whereas the negative elements remain. Both arrays are summed up entry-by-entry. Next, the intermediate results are smoothed using the moving-window integration filter, specified in eqn. 3. Then the squaring operator boosts the large value, which is associated with the deceleration before the feet striking on the ground, and suppresses the other waves of small amplitudes. The preprocessed signal is shown in Fig. 7 (e). For comparison purpose, the preprocessed signal is plotted at the same location in the left and right columns.

The preprocess procedures transform the complex signal of each step to a sloping wave of great amplitude. The location of the peak within this wave is very close to the second positive peak in the filtered signal for the short temporal distance between the both events. Shortly after the maximal deceleration, the impact force, exerting on the feet by the ground during the struck, results in the second positive peak in the step signal. The peaks in the preprocessed signal are detected using a threshold, one fourth of the maximum in array, which is marked with solid straight lines in Fig 7 (e).

The location of these peaks is defined as the onset of the peak-searching interval in the filtered signal and twice half-width of the sloping waves is defined as the width of the interval. The onset and offset of the interval is marked with dashed line in Fig 7 (f) and (g). The maxima within the peak-searching interval in the filtered signal are marked with asterisk in Fig 7 (f) and (g). This method allows detection of the second positive peak of each step signal.

III. Conclusions

Accelerometers are commonly integrated into body sensor networks. And the automatic step detection in the accelerometer signal provides useful information for the patients monitoring. Three algorithms have been developed for this purpose. Pan-Tompkins method is easy to implement, but the fluctuation in the signal, yielding the positive and negative slopes as the useful feature, can result in false peak-searching interval. The great advantage of the template-matching method is that the algorithm is capable of detecting the steps self-adaptively and generating the representative template according to the current step signal, under the condition that the first template should be correct. But the first template is estimated since the parameters are unknown; hence it may not approximate the real step signal. As a sequence, all the processing on the further data is false. Peak detection based on combined dual-axial signals is the fastest and easiest among the three algorithms. So far, it provides good performance in the first experiment. Moreover, it can be easily written in fixed-point algorithm, which is suitable for integration into microprocessor of limited computing power. This method still needs to be verified on more patients' data. In the future, the three algorithms will be compared quantitatively by means of complexity and precision, namely how many peaks are detected correctly in contrast to annotation on the database. According to the result of the comparison, one algorithm will be chosen, optimized and implemented in embedded system. Additionally, pressure sensors will be tested for step detection in the nearly future, which is commonly used for gait analysis [1],[6].

Acknowledgment

We wish to acknowledge the contribution of Prof. M. E. Schläfke at Clinic Ambrock for Neurology, Centre for Sleep- and Rehabilitation Research, Germany, and PD. Dr. D. Schäfer from Medical Department, Ruhr-University Bochum. Without their efforts of collecting the original data, the data processing would not have been possible.

References

1. Schiermeier S and Schäfer. D and Schläfke. ME, Breathing and locomotion in patients with Parkinson's disease (2001), Eur. J. Physiol., 443:67–71
2. Melnick ME and Radtka S, Gait analysis and Parkinson's disease (2002), Rehab. Manag., 58: 46-48
3. Pan. J and Tompkins. W, A real-time QRS detection algorithm (1985), IEEE Transactions on Biomedical Engineering, 32: 230-236
4. Rangayyan R. M, Biomedical Signal Analysis (2002), IEEE Press Series on Biomedical Engineering
5. Kincke. U and Jäkel. H, Signale und Systeme (2002), Oldenbourg Verlag, München
6. Hegewald, G, Ganganalytische Bestimmung und Bewertung der Druckverteilung unterm Fuß und von Gelenkwinkelverläufen (2000), Dissertation, Humboldt Universität zu Berlin

Address of the corresponding author:

Author: Hong Ying
Institute: Central Institute for Electronic (ZEL)
Street: Research Centre Jülich
City: Jülich
Country: Germany
Email: h.ying@fz-juelich.de

EMFi in wearable audio applications

S. Kärki[1], T. Salpavaara[1] and J. Lekkala[1]

[1] Institute of Measurement and Information Technology, Tampere University of Technology, Tampere, Finland

Abstract— **Flexible microphone and earphone prototypes for wearable applications were developed by using ElectroMechanical Film (EMFi™)*. A suitable application for the developed headset can be, for instance, as accessory of rescue services or sport enthusiasts. Due its versatile properties, EMFi can be used both as microphone and earphone material. The sensor operation is based on thickness changes caused by an external force or pressure, generating charge and thus voltage on the electrodes. EMFi also works conversely, converting electrical energy to vibration and hence functioning as an actuator. In addition, measurement electronics for the microphone and earphone were implemented. Preliminary test measurements were carried out: the frequency response of the EMFi microphone was compared with the one of a reference B&K microphone. The EMFi microphone provides rather good response. Also subjective listening tests were done. For these measurements, the EMFi headset was integrated inside a neoprene hood used by the surface rescuers. Both with the microphone and earphone the quality of voice was sufficient. Based on the results, EMFi seems to be a promising material for some wearable audio applications.**

Keywords— **EMFi, film, microphone, earphone.**

I. Introduction

The Electromechanical Film (EMFi) is a thin and cellular polypropylene film that can be used as an electret [1]. Generally speaking, electret is a material that retains its electric polarization after being subjected to a strong electric field during manufacturing [2]. Many of these materials are based on piezoelectricity. According to the piezoelectric theory, a material changes its dimensions when subjected to an electrical field, and conversely, generates electrical charge when mechanically deformed [3]. The advantage of EMFi over solid polymer electrets is in its flexibility due to the open internal structure combined with strong permanent charge [4].

The cellular structure of EMFi is made by stretching the polypropylene film preform during manufacturing both in longitudinal and transversal direction [5]. The film is charged in high electric field and finally metallized on both sides to provide electrodes. More detailed information about the manufacturing process can be found in [5].

The EMFi material consists of three layers: smooth and homogenous surface layers and a dominant, thicker midsection (see Fig. 1a). The midsection is full of flat voids separated by leaf-like polypropylene layers [4]. The total thickness of the EMFi material is only few dozens of micrometers. The EMFi is commercially available through a Finnish company Emfit Ltd. Emfit film specifications sheet can be found in the company homepage [3].

The EMFi material is sensitive to dynamic forces exerted normal to its surface. The sensor operation has a capacitive nature and it is based on thickness variations in the midsection of the film caused by an external force. The operation of the actuator is reciprocal [1]. The film can be used both as a transmitter and a receiver at audio frequencies [6]. The EMFi material has a very wide frequency range [3] and it is thus suitable for audio applications.

Lot of research has been done with electrets [1] by using different polymer materials. The wide spectrum of applications and low base material cost are the main advantages of the EMFi material. For example, the EMFi material has been used in force transducers [7], panel loudspeakers [8], active noise control [9], keypads [10], acoustic guitar pickups [11] and ballistocardiographic measurements [12]. It is also lightweight and easy to cut almost in any shape and size [1]. The material is useful especially in physiological applications, where the material is in contact with skin or clothing [1]. Thus it was chosen to be a proper material for this study.

The purpose of this study was to implement a thin and flexible headset for wearable applications, involving a separate microphone and earphone. The headset is particularly meant to be integrated inside clothing. For instance, the suitable application areas for the developed headset can be in rescue services and sport. As a preliminary test application, the headset was integrated inside a neoprene hood used by the surface rescuers.

The environment of the surface rescue operation is demanding. During the surface rescue operation, the headset will be subjected to mechanical stress and moisture. The neoprene hood is very tight-fitting and thus mechanical stress is directed at the headset especially when putting on the hood. However, the EMFi material is bendable and it stands some bending, twisting and stretching. In addition, the EMFi sensors and actuators can be sealed hermetically.

* EMFi is a registered trademark of Emfit Ltd.

The content of this paper is following. In Sections 2 and 3, the structures and operation principles of EMFi microphone and earphone are presented. The measurement electronics are described in Section 4. In Section 5, the preliminary test measurements are reported and discussed. Section 6 represents the conclusions of this study.

II. EMFI MICROPHONE

Sound is a longitudinal wave in a medium [13]. It can be treated as a dynamic force interacting with a film-like EMFi sensor. The operation of the microphone is based on the changes in thickness of the film caused by the external force. The change in thickness generates a corresponding charge, and hence, a voltage to appear on the electrodes [3]. Thus the force applied to the film can be detected by measuring the voltage over the sensor.

The structure and operation principle of the EMFi sensor is shown in Fig. 1.

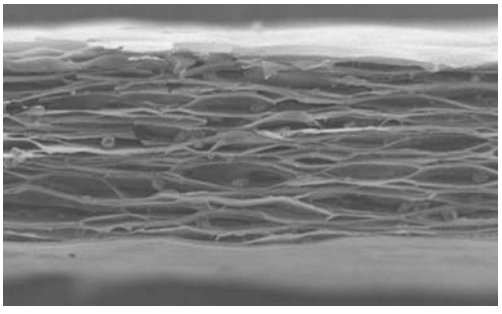


Fig. 1a Scanning electron microscope picture of the structure of the EMFi material. The EMFi material consists of three layers: smooth and homogenous surface layers and a dominant, thicker midsection. Photo: the courtesy of VTT.

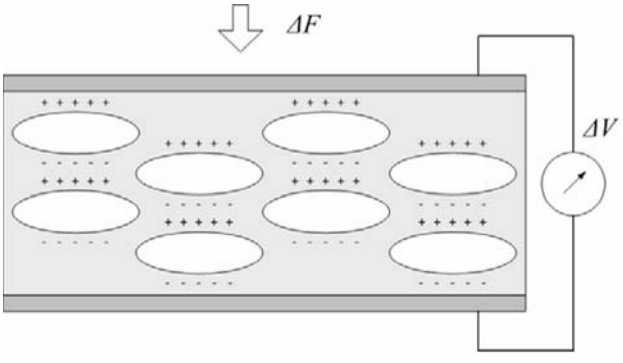


Fig. 1b Simplified structure of EMFi material. The operation of EMFi sensor is based on the changes in thickness of the charged micro bubbles inside material caused by an external force. The force ΔF generates a charge and hence, a corresponding voltage ΔV to the electrodes. The structure and operation principle is described more precisely in [1] and [14].

In this study, an EMFi microphone prototype was designed for wearable applications. Consequently, the microphone is integrated inside clothing. Since the microphone will be in contact with skin, the external force will be the vibration of tissue caused by speaking. Therefore the microphone can be regarded as a bone or tissue conduction microphone measuring mechanical vibrations rather than a conventional microphone that is based on detecting pressure waves in the air.

The microphone was constructed by using commercial EMFi material that was metallized on both sides by sputtering to provide electrodes [1]. The thickness of the electrode layers is only few dozens of nanometers. Conductivity of the electrodes can by ensured by gluing a strip of aluminum foil in the material. This also provides a better mechanical protection.

The EMFi sensors can be made almost in any shape and size – in this case, the microphone was chosen to be rectangular, size 6 x 9 cm. A 6 x 18 cm piece of material was cut from the EMFi sheet. This piece was folded twice over to implement the 6 x 9 cm two-layer EMFi sensor. The metal coating was removed from the inner surface edges with cotton stick dipped in alcohol to prevent short circuits. In Fig. 2 these steps are illustrated.

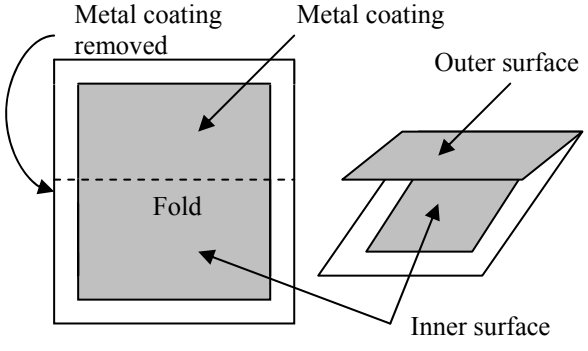


Fig. 2a Making of the EMFi sensor. The outer metal surface of the sensor acts as a grounded electric shield and the inner surface as a signal electrode.

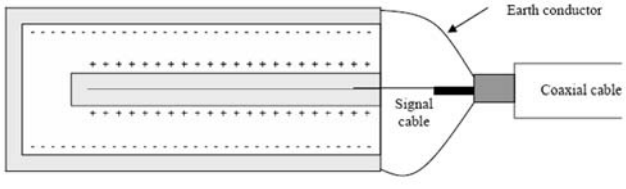


Fig. 2b Simplified cross-section of the EMFi microphone. The white area describes the EMFi material and grey areas the aluminum foil used to ensure the conductivity.

Connections were made by coaxial cables. Too high temperatures will harm the structure of the EMFi material permanently and therefore the wire has to be connected to the EMFi sensor by using tin-plated copper tape: the wire is soldered to a tin-plated copper tape and it is then taped on the EMFi sensor. The signal wire of the coaxial cable was connected to inner surface of the sensor and the earth conductor to the outer surface. The wire connections were ensured with conductive carbon paint. In the folded sensor constructed, the outer metal surface acts as a grounded electric shield and the inner surface as a signal electrode.

The microphone implemented here has a rather large size, because the larger microphone is easier to construct manually. However, it is also possible to make the sensor smaller.

The microphone was finished by gluing it on a thin piece of rubber to make the structure stronger. The coaxial cable connecting the microphone to the electronics was clamped to the rubber base by stapling. Fig. 3 introduces the constructed microphone. Finally, the EMFi microphone was sewed inside a textile pocket which can be attached inside clothing. To protect the microphone against moisture, it can be covered with plastic sheeting or with a thin layer of silicon rubber. The capacitance of the microphone, 1.21 nF, was measured with HP 4262A LCR meter.

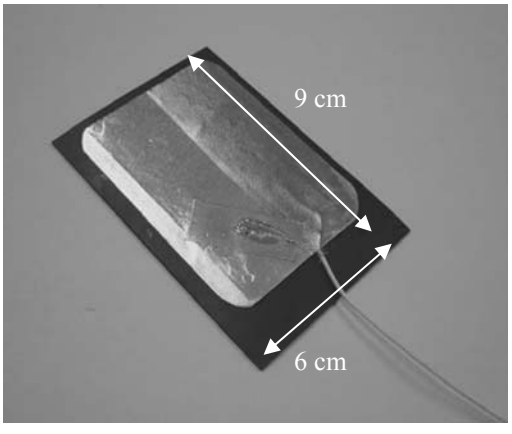


Fig. 3 EMFi microphone.

III. EMFI EARPHONE

The operation principle of the EMFi actuator is opposite to the EMFi sensor. The operation is based on the changes in the film thickness when external voltage is applied to the electrodes. The external sinusoidal voltage generates corresponding thickness variations in the film, producing a sound wave. The operation principle of the EMFi actuator is shown in Fig. 4.

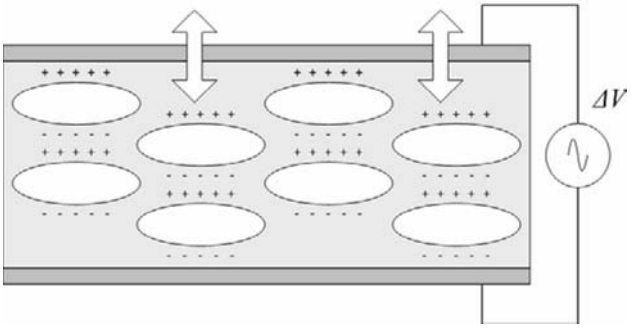


Fig. 4 Operation principle of the EMFi actuator. An external voltage ΔV is applied to the electrodes, causing a change in the film thickness.

The earphone has a shape of a rectangle with size 10 x 6.5 cm. The earphone is large enough to cover the ear when integrated inside clothing and thus it provides higher sound pressure. Instead of coaxial cable, normal wires were used in the earphone. Fig. 5 represents the earphone constructed.

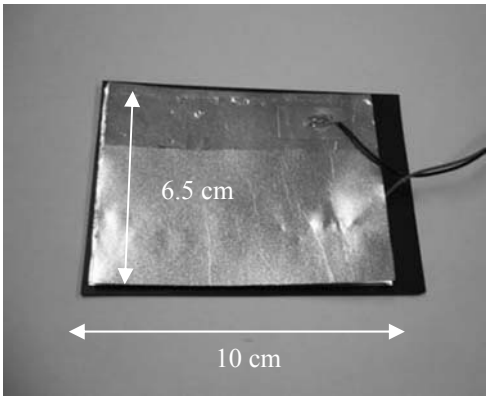


Fig. 5 EMFi earphone.

The sound pressure level of a single layer of EMFi material is low; by stacking several layers of the film the sound pressure level can be increased [15]. The increase in amplitude has been found to be nearly linear [15]. The prototype of EMFi earphone was constructed by mainly the same way as the microphone. However, the microphone has only two layers of EMFi material whereas the earphone needs to have more layers to improve its efficiency. In this case, six layers of EMFi material were used.

Also the earphone was attached to a piece of rubber and sewed inside a textile pocket. The capacitance of six-layer earphone was measured to be 5.40 nF.

IV. Measurement electronics

The headset prototype has a separate 5.5 cm x 7.5 cm x 2.2 cm plastic housing for the amplifier electronics and power source, shown in Fig. 6. The headset is powered with a 9 V alkaline battery, enclosed inside the plastic covering.

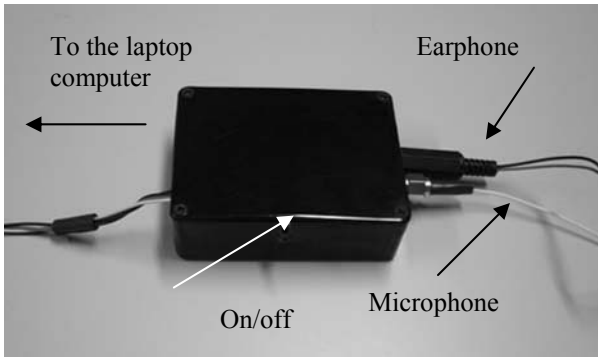


Fig. 6 Plastic casing for the amplifier electronics and power source. The microphone and earphone are connected to the electronics. The electronics can be connected to a laptop computer. Also the on/off switch is marked in the figure.

The microphone preamplifier is realized with a transimpedance amplifier. It is followed by an AC-coupled voltage amplifier able to produce line level signal. The frequency response of the amplifier is represented in the Fig. 7.

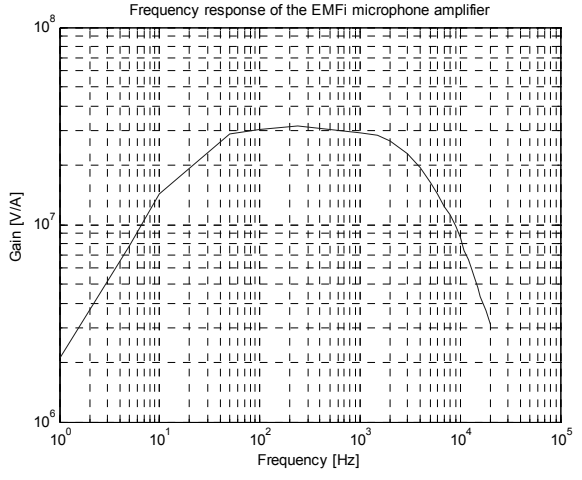


Fig. 7 Frequency response of the EMFi microphone amplifier.

The frequency response was measured by feeding a 0.1 V_{pp} sinusoidal voltage to the amplifier input through a 10 MΩ resistor. The output voltage was determined with Hameg HM507 oscilloscope. The lower cut-off frequency (-3 dB) was found to be 20 Hz and the upper one 3.2 kHz.

The emitted sound pressure of the EMFi earphone depends on its driving voltage. The voltage of the line level input voltage signal is amplified with a rail-to-rail operational amplifier LM7301 in order to achieve maximum output with the 9 V supply voltage. The capacitive load of the EMFi earphone is compensated with a 470 Ω resistor in series with operational amplifier output.

The frequency response of the EMFi earphone amplifier was determined by feeding a 0.1 V_{pp} sinusoidal signal to the earphone amplifier input. During the output voltage measurements, the EMFi earphone was connected to the measurement electronics. Based on the results, the lower cut-off frequency (-3 dB) was found to be 20 Hz and the upper one 32.8 kHz.

The electronics was connected to a laptop computer. With the laptop computer it is possible to evaluate the headset properties, that is, to feed a line level signal to the earphone and to store the recorded microphone signal. The test measurements are described more precisely in Section 5.

V. Test measurements

Preliminary test measurements have been carried out. The subsection A represents the test measurement done with EMFi and reference microphones in anechoic chamber. The subsection B introduces subjective listening tests carried out. In the first test measurement section, the measurements were done with EMFi microphone sewed inside a textile pocket as it would be real applications. In the second section, the microphone and earphone were integrated inside surface rescuer's neoprene hood.

A. Measurements in anechoic chamber with the EMFi microphone

The first part of the test measurements was carried out in the anechoic chamber of our Institute. In order to find out the frequency response of the EMFi microphone, the following test measurement was carried out even though this measurement setup is not optimal for the constructed EMFi microphone. The frequency response of the EMFi microphone was compared with a B&K Prepolarized Free-field ½" Microphone Type 4189 reference microphone. The EMFi microphone was attached to a vertical wooden plate by taping. Both microphones were positioned side by side opposite to the Genelec 1029 loudspeaker which was used as a sound source. The Genelec loudspeaker has lower cut-off frequency (-3 dB) 68 Hz and upper cut-off frequency 20 kHz [16]. The distance between the microphones and the loudspeaker was 1.65 m. The temperature of the anechoic room was 12°C. Both the microphones with amplifier elec-

tronics and the loudspeaker were connected to a laptop computer.

Pseudorandom noise (MSL, maximum-length sequence) was replayed with the Genelec loudspeaker connected to the laptop computer. The microphones have equal distance to the loudspeaker and hence the both microphones were exposed to a same relative sound pressure level. The outputs from the microphones were stored and analyzed with Matlab®. In Fig. 8 the responses of the reference B&K and EMFi microphones are shown. The vertical axis represents the response in dB and the horizontal axis the frequency in Hz.

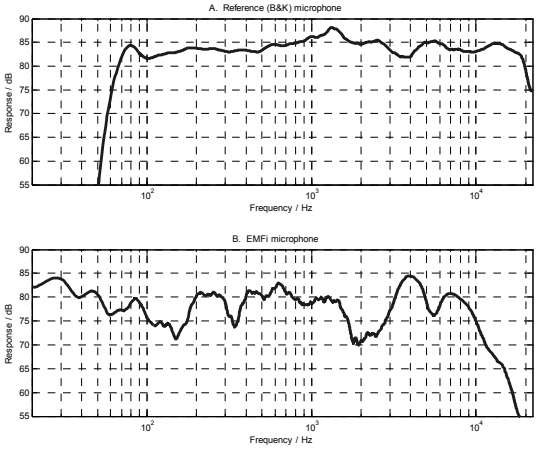


Fig. 8 Frequency responses of a) reference B&K and b) EMFi microphones.

Fig. 8 shows that the EMFi microphone provides rather good response compared with the reference microphone. The frequency response is relatively flat to 10 kHz (± 7 dB). However, the origin of the ripples seen in the Fig. 8b did not become evident during this study. Similar ripples were found also with different fixings and back materials. In addition, the EMFi microphone recognizes signal on frequencies below the lower cut-off frequency of the Genelec loudspeaker. The reason for this may be found from the operation principle of the EMFi material. The EMFi measures acoustic pressure waves as well as the mechanical vibrations. Some mechanical vibrations may have been appeared due to the attaching technique, and this vibration can be seen in Fig. 8b in lower frequencies. It is also possible that part of the signal shown in Fig. 8b has a mechanical vibration origin at higher frequencies.

With the earphone the corresponding test measurements were not done. The earphone is especially designed for the wearable applications and therefore the sound level is not very high. The sound level is adequate when the earphone is held against ear and thus it is not reasonable to measure the frequency response in air. In the future, the frequency response of the earphone will be measured by using an artificial head.

B. Subjective listening test

The second part of the test measurements was carried out for the microphone and the earphone by subjective listening tests. The aim of the test measurement was to evaluate the performance of the headset. For these measurements, the headset was integrated inside the surface rescuer's neoprene hood. The earphone was located against ear to provide adequate sound level and the microphone against cheek to recognize the vibrations caused by the speaking. In Fig 9, the headset layout is shown.

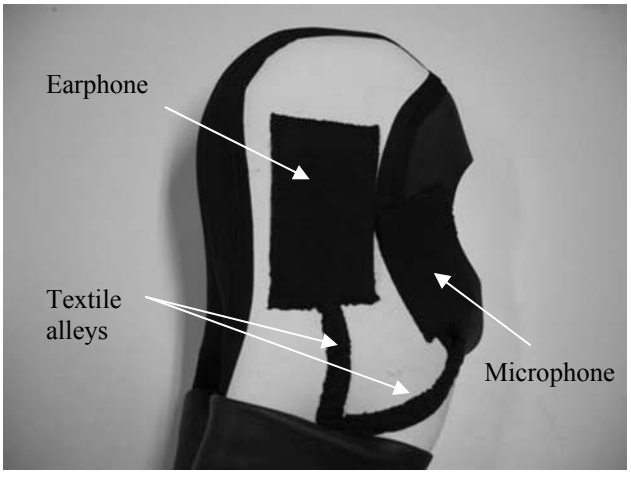


Fig. 9 Layout of the EMFi headset. The positions of the microphone and earphone are marked in the figure. Also the textile alleys protecting the cables are shown.

A functional testing was performed by listening and recording spoken sentences and vocals with the EMFi headset. A laptop computer was used to feed a line level signal to the earphone and to store the recorded microphone signal.

In microphone testing a test person read out a short piece of text and the microphone signal was stored. The spoken text was replayed with a laptop computer and the quality of sound was evaluated. As a result, the voice recorded with the microphone was understandable and the tone of voice didn't change, even though some minor crunch and rustle occurred.

In the earphone testing the test person listened to music from the earphone and evaluated the loudness of the music. The music was played with a MP3 player. The earphone is able to produce clearly audible and understandably voice in laboratory environment. However, the sound level may be

inadequate for noisy environment. On the other hand, the thick neoprene hood attenuates effectively noise outside the hood. The sound level can be increased by adding more layers of EMFi material to the earphone.

VI. Conclusions

In this study, a thin and flexible headset was implemented for wearable applications. In wearable applications, the headset is integrated inside clothing that presses the headset tightly against head. For instance, the clothing can be a fire fighter's helmet or rescue diver's neoprene hood. The headset has separate microphone and earphone, both realized with EMFi material. In addition, the measurement electronics for the headset was implemented.

In test measurements, the frequency response of the EMFi microphone was compared with the one of the reference microphone (B&K). The EMFi microphone provided rather good results. However, the frequency range of the EMFi microphone is not as wide as the reference microphone has but adequate for measuring frequencies appearing in a normal speaking.

Also some subjective listening tests were carried out to evaluate the performance of the headset. To conduct these measurements, the headset was integrated inside the neoprene hood used by the surface rescuers. Both with the microphone and the earphone the quality of voice was sufficient. Nevertheless, the sound level of the earphone was rather low.

Since the EMFi microphone is sensitive to both acoustic and mechanical vibrations, it also detects the head movements. This causes some artifacts to the signal, even though the influence is not very remarkable.

The headset implemented in this study is only an introductory prototype. The purpose was to realize a simple and practical headset with potential for further development. The headset integrated inside the hood is comfortable and easy to wear and use. All the problems generated by the challenging operation environment were not solved during this study, for example, the headset is not yet totally waterproof. More test measurements are still required to evaluate the headset properties. However, the results achieved in the study were encouraging. To conclude, the EMFi is a promising material for wearable audio applications.

References

1. M. Paajanen, J. Lekkala and K. Kirjavainen, ElectroMechanical Film (EMFi) – a New Multipurpose Electret Material, Sensors and Actuators A: Physical, 84 (2000), pp. 95-102, 2000.
2. Encyclopedia Britannica, http://www.britannica.com.
3. Homepage of Emfit Ltd, http://www.emfit.com.
4. M. Paajanen, J. Lekkala and H. Välimäki, Electromechanical modeling and properties of the electret film EMFi, IEEE Transactions on Dielectrics and Electrical Insulation, Vol. 8, No. 4, pp. 629-636, 2001.
5. K. Kirjavainen, Electromechanical Film and procedure for manufacturing same, U.S. Patent No. 4654546, 1987.
6. J. Backman and M. Karjalainen, Audio and ultrasonic transducers based on Electrothermomechanical Film (ETMF), Acoustics, Speech, and Signal Processing International Conference on 3-6 Apr. 1990, 1173-1776.
7. G. Evreinov and R. Raisamo, One-directional position-sensitive force transducer based on EMFi, Sensors and Actuators A: Physical, 123-124 (2005), pp. 204-209.
8. M. Antila, H. Nykänen and K. Saarinen, Multichannel Electromechanical Film panel loudspeakers, AES 16th International Conference on Spatial Sound Reproduction, 1999.
9. H. Nykanen, M. Antila, J. Kataja, J. Lekkala, and S. Uosukainen, Active control of sound based on utilizing EMFi technology, ACTIVE 99, pp. 1159–1170.
10. Homepage of Screentec Ltd, http://www.screentec.com.
11. Homepage of B-Band Ltd, http://www.b-band.com.
12. T. Koivistoinen, S. Junnila, A. Varri and T. Koobi, A new method for measuring the ballistocardiogram using EMFi sensors in a normal chair, Engineering in Medicine and Biology Society, 2004. pp. 2026-2029.
13. H. Young and R. Freedman, University Physics, 9th edition, Addison-Wesley Publishing Company, Ch. 21, pp. 646-668.
14. M. Paajanen, H. Välimäki and J. Lekkala, Modelling the Electromechanical Film (EMFi), Journal of Electrostatics, 48 (2000), 193-204.
15. J. Lekkala, M. Paajanen and K. Kirjavainen, A New dimension in loudspeakers, Materials World, 2001, 22-24.
16. Genelec 1029A Active monitoring system datasheet, www.genelec.com.

Address of the corresponding author:

Author: Satu Kärki
Institute: Measurement and Information Technology
Street: P.O. Box 692
City: 33101 Tampere
Country: Finland
Email: satu.karki@tut.fi

Policy-based Management for Body-Sensor Networks

Sye Loong Keoh[1], Kevin Twidle[1], Nathaniel Pryce[1], Alberto E. Schaeffer-Filho[1], Emil Lupu[1], Naranker Dulay[1], Morris Sloman[1], Steven Heeps[2], Stephen Strowes[2], Joe Sventek[2] and Eleftheria Katsiri[1]

[1]Deparment of Computing, Imperial College London, UK
[2]Department of Computing Science, University of Glasgow, UK

Abstract—**Body sensor networks e.g., for health monitoring, consist of several low-power on-body wireless sensors, higher-level devices such as PDAs and possibly actuators such as drug delivery pumps. It is important that such networks can adapt autonomously to changing conditions such as failures, changes in context e.g., user activity, or changes in the clinical condition of patients. Potential reconfiguration actions include changing the monitoring thresholds on sensors, the analysis algorithms or the configuration of the network itself. This paper presents a policy-based approach for autonomous management of body-sensor networks using the concept of a Self-Managed Cell (SMC). *Ponder2* is an implementation of this approach that permits the specification and enforcement of policies that facilitate management and adaptation of the response to changing conditions. A Tiny Policy Interpreter has also been developed in order to provide programmable decision-making capability for BSN nodes.**

Keywords—**Autonomic management, adaptive sensing, policy-based adaptation, reconfigurable networks.**

I. Introduction

On-body and implantable sensors have the potential to be used in hospitals or homes for monitoring physiological parameters of post-operative and chronically ill patients (e.g., those suffering from diabetes mellitus). These sensors use wireless communications to form body sensor networks (BSNs) [1] and can interact with wearable processing units such as PDAs, mobile phones and the fixed network infrastructure. Health monitoring using BSNs enables early release of patients from hospital and facilitates continuous monitoring of clinical condition in the home or at work. Additionally, healthcare personnel can be automatically alerted to obtain assistance if the patient's condition deteriorates.

Whilst much of the current work focuses on the development of new sensors and processing the data acquired from them [1], we focus on providing adaptation and self-management at both sensor level and for body sensor networks. For example, there is a need to adapt the frequency of measurements on a sensor depending on the activity and clinical condition of the patient. This enables optimising power consumption whilst ensuring that important episodes are not missed. Similarly, the use of variable thresholds for transmitting sensor readings reduces the need for communication and thus power consumption. Typically, sensor configuration may also change depending on the user's context, e.g., location, current activity and medical history. Physiological parameters such as heart rate thresholds then need to be configured and customised accordingly. Policy-based techniques have been used for over a decade in network and systems management in order to define how the system should adapt in response to events such as failures, changes of context or changes in requirements. By specifying the policies (i.e., what actions should be performed in response to an event) declaratively and separately from the implementation of the actions, it is possible to dynamically change the adaptation directives without changing the implementation or interrupting the functioning of the device. Thus, policy-based mechanisms provide feedback control over the system and a constrained form of programmability.

In this paper, we present a policy-based architecture that supports autonomic management for body sensor networks, based on the concept of a Self-Managed Cell (SMC) [2]. A SMC consists of an autonomous set of hardware and software components that represent an administrative domain such as a body area network of physiological sensors and controllers. We introduce *Ponder2* [3], a toolkit that supports the specification and enforcement of policies in the form of event-condition-action rules, the grouping of a SMC's components in domains for management purposes and the dynamic loading of new functionality and new communication protocols. Policies can be defined with respect to the SMC's components, for interactions with the other SMCs and for the management (i.e., loading, removal, activation) of the policies themselves. We also describe the implementation of a policy interpreter for BSN nodes [4].

The paper is organised as follows: Section II describes the SMC architectural pattern. Section III presents the Ponder2 policy interpreter as the core component for adaptation and feedback control. Section IV describes the implementation of a policy interpreter for the BSN nodes. Section V discusses our prototype implementation of Ponder2 and Tiny Policy Interpreter. Sections VI and VII present the related work and conclude the paper with directions for future work.

II. Self-Managed Cells (SMC)

As an example, an SMC represents a body-sensor network consisting of several sensors, i.e., *glucose, blood pressure, heart-rate, ECG monitor* and actuators such as an insulin pump to administer appropriate dosage of insulin in a diabetes Type II patient monitoring system. Context sensors such as accelerometers may be used to sense the patient's activities in order to facilitate the adjustment of drug delivery and monitoring threshold (e.g., for heart-rate).

Figure 1 illustrates the architectural pattern of a SMC that manages a set of heterogeneous components (i.e., managed resources) such as those in a body-sensor network, a room or even a large-scale distributed application. Resource adapters are instantiated to provide a unified view for interaction with the resources as they may use different interfaces or communication protocols. For example, interactions with BSN nodes occur via IEEE 802.15.4 wireless links, while interactions with PDAs, mobile phones or Gumstix[1] typically occur over Wifi or Bluetooth.

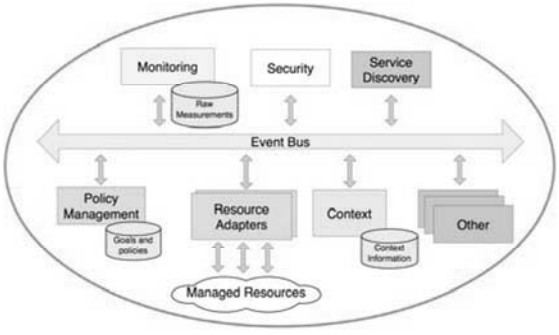


Fig. 1 The SMC architecture pattern

A SMC can load other components and services for detecting context changes, monitoring component behaviours or for security (authentication and access control). However, the *event bus*, the *policy service,* and the *discovery service* work in conjunction with each other and form the core functionality of a SMC that must always be present.

As most pervasive systems are event-driven, the services of a SMC interact using a common publish/subscribe event bus, although we do not constrain all communication to be event-based. The event bus [5] forwards event notifications from services onto any interested parties within the SMC who have subscribed to receive the event. This has the advantage of decoupling the services since an event publisher does not need to know about the recipients, thus permitting the addition of new services to the SMC without disrupting the behaviour of existing services. Secondly, multiple services in the SMC can respond concurrently and independently to the same notification with different actions. Finally, the event bus can be used for both management and application data such as alarms indicating that threshold have been exceeded. In order to lower communication overhead, sensors typically only transmit events when an unusual situation arises rather than transmitting all sensor readings.

The policy service implements a local *feedback control loop* to achieve adaptation and self-management. It caters for two types of policies: *obligation policies* (event-condition-action rules), which define the actions that must be performed in response to events, and *authorisation policies* which specify what actions are permitted on which resources and services. The discovery service is used to discover new components which are capable of becoming members of the SMC, e.g., sensors and other SMCs in the vicinity. It establishes a profile describing the services they offer and generates an event describing the addition of the new device for other SMC components to use as appropriate. The discovery service also manages the SMC's membership as it is necessary to cater for transient failures, which are common in wireless communications, and to detect permanent departure (e.g., device out of range, switched off, or failure).

Complex environments can be realised through the federation and composition of several SMCs. This permits exchanging policies between SMCs and thus programming in a restricted way the behaviour of collaborating or composed SMCs [6].

III. Ponder2 Policy Service

Ponder2 [3] is the policy service for the SMC and has been inspired by the lessons learnt in the development of Ponder [7], a policy language and toolkit developed at Imperial College over a number of years. In contrast to Ponder, which was designed for general network and systems management, Ponder2 has been designed as an entirely extensible framework that can be used at different levels of scale from small embedded devices to complex environments.

Ponder2 combines a general-purpose object management system with a *domain service*, *obligation policy interpreter* and a *command interpreter*. The Domain Service provides a hierarchical structure for managing objects. The Obligation Policy Interpreter handles Event, Condition, Action rules (ECA). The Command Interpreter accepts a set of commands in XML form via several communication interfaces. These commands can be used to interrogate the Domain Service and perform invocations on the managed objects.

Managed objects (also called adapter objects) represent sensors and other SMC devices, services within those devices and remote SMCs. Domains and policies are managed

[1] www.gumstix.com

objects in their own right on which actions can be performed, e.g., adding/removing an object from a domain, enabling or disabling a policy. Obligation policies are also used to decide which adapters should be created when new components appear and in which domain they should be placed. Other policies specified for that domain will then automatically apply to the new component.

As shown in Figure 2, the overall architecture of the policy service comprises the domain structure, the triggering mechanism matching events to obligation policies and the execution invocation engine which is used to make the calls to the objects inside the domain structure. Conceptually the policy service has an event interface through which event notifications are received from the external event bus, an invocation interface through which external invocations are received and an action interface for invocations on external objects.

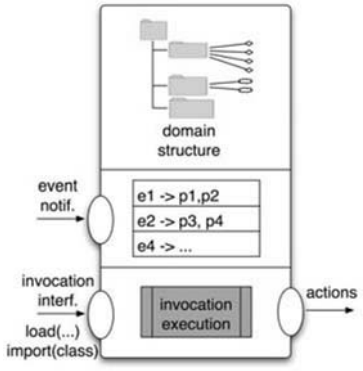


Fig. 2 The policy service architecture

Ponder2 uses XML to encode policies, event specifications and invocation of actions. It interprets XML as a sequence of statements that identify the managed object to be used and parameters or sub-elements within the XML that are to be sent to the managed object. For example, the following snippet identifies the root domain ("/") and sends it an *add* command. The *add* command has its own structure and information saying what is to be added to the domain structure. In this example, the managed object to be added to the root domain will be called *newobject*.

```
<use name="/">
  <add name="newobject">
    ...
  </add>
</use>
```

XML is verbose and not easy to use, so we are developing a higher-level declarative language from which the XML can be generated.

A. Instantiation of Managed Objects

Ponder2 has the ability to load all the code needed on-demand. This enables us to use it across a wide variety of applications and devices with different capabilities by only loading those components that are necessary in each case. By default, a domain hierarchy rooted at / will be created. A built-in *domain factory* is used to create new domain objects within the domain hierarchy. All other managed objects are instantiated through factories that can be loaded on demand. Ponder2 has defined several other *factory objects* such as *policy (authorisation and obligation), basic managed object,* and *event.* This provides the flexibility to dynamically create policies and managed objects for communicating with various sensors and devices with their respective communication protocol such as UDP datagram, Bluetooth [8], IEEE 802.15.4, or Zigbee [9]. For example, the following XML snippet uses a BSN *glucose factory* object to create a managed object for a BSN based glucose sensor and places it into the `/bsn` domain under the name *GlucoseSensor*. This managed object acts as an adapter to the actual sensor and implements the high-level interface in terms of interactions with the sensor via IEEE 802.15.4 radio.

```
<use name="/bsn">
   <add name="GlucoseSensor">
      <use name="/factory/BSNglucose">
         <create addr="0:0:0:3" />
      </use>
   </add>
</use>
```

Once instantiated, managed objects receive commands in the form of XML structures. Typically, the main XML element is the `<use name="x">` construct to identify a managed object, with one or more operations or sub-commands represented by child elements. A command to a managed object takes the form:

```
<use name="managedobject" arg1="foo" arg2="bar">
  <op1 op1arg1="doh"/>
  <op2 op2arg1="argh"/>
  ...
</use>
```

where the operations/commands `op1` and `op2` could have child elements of their own.

B. Integration with Event Buses

An *event factory* implements the interface with an external event bus and encapsulates the protocols necessary to communicate with it. Multiple event factories can be used to integrate with different external event buses if required. The *event factory* can be used to create new Event Types that issue a subscription to the external content-based event bus and define the event type name and its arguments. When an

event corresponding to the subscription expression occurs on the event bus, the event factory is notified of its occurrence and raises the corresponding event type in the policy interpreter in order to trigger the policies.

```
<use name="/Event">
  <add name="highGlucoseEvent">
    <use name="/factory/GlasgowEvent">
      <create>
        <and>
         <sub name="level" op="GEQ" value="90" />
         <sub name="type" op="EQ" value="glucose"/>
        </and>
        <arg name="type"/>
        <arg name="level"/>
        <arg name="daytime"/>
      </create>
    </use>
  </add>
</use>
```

In the above example, an Event Type called */event/highGlucoseEvent* that has three arguments, sensor *type,* glucose *level* and *context* is created. The subscription expression to which this event type corresponds is matched by the external event bus against published events on the bus. In this example all events with type *glucose* and level ≥ 90 are matched.

This example assumes that the sensor transmits glucose readings periodically as events to the event bus. The frequency with which this is done is determined by the sensor and may be configurable. However, in order to minimise power consumption linked to the communication of readings, it is desirable that the sensor itself be programmable in terms of basic policies that can determine when actions such as raising external notifications should be done. This has led us to develop a basic policy interpreter for BSN nodes which is detailed in Section IV.

C. Policy Specification

We are primarily concerned with two types of policies: *authorisation policies* define what actions are permitted under given circumstances and *obligation policies* define what actions to carry out when specific events occur if the specified condition is true. An obligation policy specifies the Event Type that will trigger the policy together with arguments expected of that event, optional conditions that must be satisfied and a set of actions to be performed.

The example below shows the XML encoding of an obligation policy that will respond to events of type */event/highGlucoseEvent*. The policy becomes active as soon as it is created. It makes use of three named arguments provided by the event and their value is substituted in the actions and constraint by enclosing the name inside the two characters "!" and ";", e.g., *!level;* and *!daytime;*. The condition can contain simple Boolean statements comparing string and integer values. In the example above, Ponder2 checks whether the glucose level is greater than *125* and the current *time* is later than 20:00. Conditions can contain any combination of *and*, *or*, *not*, *eq*, *ne*, *gt*, *ge*, *lt* or *le*. *And* and *or* take any number of XML sub-elements, *not* takes one, while the others all take two. Note that the arguments for the comparisons have been separated by XML comments to ensure that they are separate XML elements. The action part consists of Ponder2 XML commands, which are only executed after the event has occurred and if the conditions are true. In this example, if the conditions are satisfied, the policy invokes an action on the insulin pump to increase the night dosage by 10%. Other actions defined in a policy could be to send an alarm SMS message to a medical service or tell the patient to perform some actions.

```
<use name="/policy">
  <add name="AdjustDosagePolicy">
    <use name="/factory/policy">
      <create type="obligation"
       event= "/event/highGlucoseEvent"
       active="true">
        <arg name="type"/>
        <arg name="level"/>
        <arg name="daytime"/>
        <condition>
          <and>
            <gt>!level;<!-- -->125</gt>
            <gt>!daytime;<!-- -->20:00</gt>
          </and>
        </condition>
        <action>
          <use name="/actuators/ip">
            <modify ctx="night" value="+10"/>
          </use>
        </action>
      </create>
    </use>
  </add>
</use>
```

IV. Tiny Policy Interpreter

Although Ponder2 confers a level of programmability and adaptation to the SMC, it is equally desirable to be able to introduce similar abstractions at the sensor level in order to endow sensors with programmable local decision capability. Consequently, we have implemented a simplified policy interpreter for TinyOS [10] that can be deployed to BSN nodes [4]. It is implemented as a NesC [11] component library that can dynamically add and remove policies specified either as textual scripts or as data structures. Figure 3 shows the component architecture of the Tiny Policy Interpreter. The Policy Script Controller is responsible for loading, adding and removing policies, while the policy interpreter is invoked to execute the actions according to the deployed obligation policies when events occur.

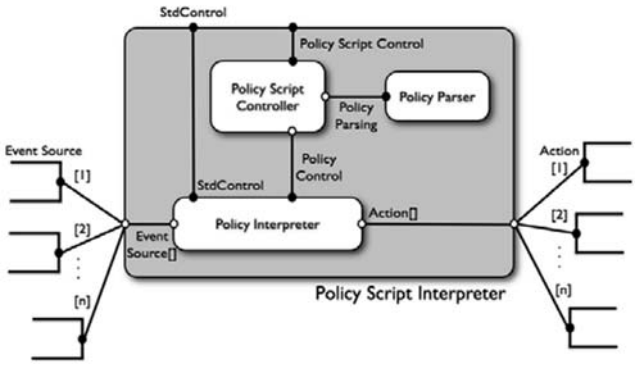


Fig. 3 Configuration of the Tiny Policy Interpreter

Tiny Policies are also specified as event-condition-action rules. Events represent samples from sensors and carry a 32-bit parameter that is the sampled value. Conditions are represented as inclusive ranges: if the value lies within the range, the condition is met. Event sources and actions are represented by NesC interfaces and they are pre-defined as arrays bound to the Tiny Policy Interpreter.

```
policy = event condition "->" action

Event = uint8 "?"

condition = equals_condition
          | in_range_condition
          | less_than_condition
          | greater_than_condition
          | always_condition

equals_condition = uint32

in_range_condition = "["min:uint32 ".." max:uint32"]"

less_than_condition = "<=" max:uint32

greater_than_condition = ">=" min:uint32

always_condition = "always"

action = do_action | set_property_action

do action = uint8 "!" action_arg?

action_arg = "(" uint32 ")"

uint8 = <unsigned 8-bit integer>

uint32 = <unsigned 32-bit integer>
```

Fig. 4 Policy syntax of Tiny Policy Interpreter

The Tiny Policy Interpreter implements a simple script syntax shown in Figure 4. This syntax is much simpler and compact than the Ponder2 XML policy syntax, which can be translated into the Tiny Policy syntax before deployment to sensors. For example, the policy `1? <= 4 -> 2!` specifies that if event 1 fires with a parameter less than or equal to 4, then perform action 2, while the policy `2? [5..9] -> 1!` means that if event 2 fires with a parameter between 5 and 9 inclusive, then perform action 1. The `always` clause is used to execute an action without needing to evaluate any policy condition, e.g., the policy `1? always -> 3!` means that when event 1 occurs, perform action 3.

A. Actions Library

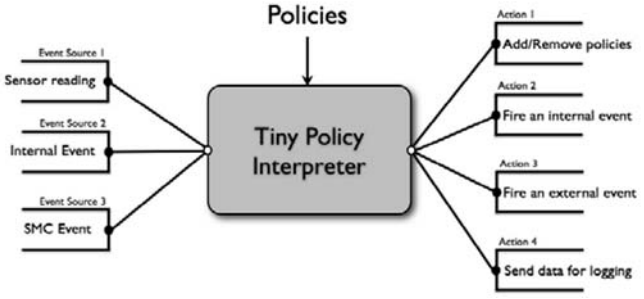


Fig. 5 Event source and action libraries

A set of general-purpose actions for the Tiny Policy Interpreter has also been defined as shown in Figure 5. This includes actions to add and remove policies from the Tiny Policy Interpreter itself. This facilitates adaptation to context changes as policies can be used to decide which set of policies applies according to the user's activity. For example, we can load a policy with a higher heart-rate threshold when the accelerometer has detected higher levels of physical activity. Additional actions include the ability to fire an internal event which triggers another policy within the interpreter in order to allow for the execution of sequential actions, or an external event to the SMC's event bus in order to trigger management actions at the SMC level. Finally, an additional pre-defined action can be used to log measurement values, e.g., to a medical database in the SMC. The event sources for triggering policies within the Tiny Policy Interpreter are either sensor readings, pre-defined internal events or events received from the SMC's event bus, e.g., events generated by other sensors or managed objects in the SMC.

Ponder2 is typically running on a PDA or Gumstix device can load and remove policies from the Tiny Policy Interpreter, thus enabling it to dynamically configure the sensing parameters and behaviour. This provides the flexibility to reconfigure the sensor behaviour without needing to re-program them.

V. Implementation And Evaluation

Ponder2 has been implemented in Java 1.4 and a version has been ported to J2ME. It is deployable to any computing platforms that have a Java Virtual Machine and we have used it in PDAs, Gumstix and mobile phones. We have also implemented a content-based publish/subscribe event notification system [5] for the SMC. Other event systems such

as XMLBlaster [12] and JMS [13] can also be used in conjunction with Ponder2.

In terms of evaluation, we have deployed Ponder2 onto a Gumstix which has a 400 MHz Intel XScale PXA255 processor with 16 MB flash memory and 64 MB SDRAM, running Linux. We observed that the evaluation of policy constraints incurs the most overheads as this involves parsing of the constraints, string comparisons and arithmetic operations. Constraints are stored in the policies in XML format and substantial gains in performance would be possible by pre-compiling them. The time taken to execute a policy without a condition and with an empty action is only 13.57 ms, while it takes 30.05 ms to execute a policy with a simple condition and an action to publish a new event. We also observed that it takes 23.88 ms to execute a policy (with no condition) to invoke an action to issue a command to BSN node via IEEE 802.15.4. In a medical scenario such as diabetes monitoring, many operations take place over time frames of minutes or even hours, so these performance figure are more than adequate in many cases. Some applications such as heart ECG monitoring and analysis, require much faster processing. In these cases, processing can be done directly in the managed objects and sometimes on the BSN sensors themselves.

Each policy is instantiated as a Java object which consumes 3.214 kB. This however includes the policy type (obligation or authorisation), the list of events which may trigger the policy, the actions to be performed and the constraints that need to be evaluated. At the moment we are using XML for internal policy representation, which carries a significant memory overhead. More compact representations could be used for devices with limited memory.

As for the Tiny Policy Interpreter, its current implementation can be installed on BSN nodes [4] and the size of its codebase is 11.61 kB.

VI. Related Work

Work on policy-driven systems has been on-going for over a decade in various application areas. Traditional approaches rooted in network and systems management include PCIM [14], PDL [15], NGOSS Policy [16], Ponder [7] and PMAC [17]. They all make use of event-condition-action rules for adaptation but are aimed at the management of distributed systems and network elements and do not scale down to small devices and sensors.

There are a number of pervasive systems that define frameworks for realising pervasive spaces. Gaia [18] and Aura [19] introduce the notion of *active space* and *smart space* respectively in order to provide a "meta-operating system" to build pervasive applications. These projects focus on spaces of relatively fixed size, e.g., a room or a house and on specific concerns such as context-related applications, user presence and intent and foraging for computational resources. We consider a SMC as an architectural pattern that applies at different levels of scale and we focus on generic adaptation mechanisms through policies.

The Pervasive Information Community Organisation (PICO) [20] is a middleware platform to enable effective communication and collaboration among heterogeneous hardware and software entities in pervasive computing. A community is a grouping of hardware entities and software agents that work together to achieve goals. The notion of community is similar to our SMC, but our focus is to facilitate self-configuration and self-management using policies.

CodeBlue [21, 22] is an ad-hoc sensor network infrastructure for emergency medical care. It integrates low-power, wireless vital sign sensors, PDAs and PC-class systems to provide a combined hardware and software platform for medical sensor networks. CodeBlue also provides protocols for device discovery, publish/subscribe, multi-hop routing and a simple data query interface for medical monitoring. CodeBlue investigates the data rates, node mobility, patterns of packet loss and route maintenance of the wireless sensor network, while the SMC framework focuses on the management of body-sensor networks using policies.

The co-operative artefacts concept [23] is based on embedded domain knowledge, perceptual intelligence and rule-based inference in movable artefacts. Measurements from the sensors are translated into observational knowledge, which is then being evaluated against pre-defined rules that are defined using Horn logic and domain knowledge. This allows the inference engine to derive the artefact's behaviour in response to the changes in the environment.

Smart-Its Context Language (SICL) [24] is a high-level description language for developing context-aware applications on embedded systems. It is integrated with a tuple-space-based communication abstraction that enables inter-object collaboration. The language provides a means of specifying sensor interfaces, inference rules for fusing sensors readings, adaptation rules and basic application behaviour. However, this approach does not support dynamic re-configuration of sensors as re-configuration implies reprogramming of sensors.

VII. Conclusions And Future Work

We have proposed the SMC abstraction as a basic architectural pattern that provides local feedback control and autonomy. Policies in the form of event-condition-action rules, provide a simple and effective encoding of the adaptation strategy required in response to changes of context or in

application requirements. The ability to dynamically load, enable and disable the policies together with the ability to use policies in order to manage other policies caters for a wide variety of application needs. Ponder2 has been designed as an extensible framework where all required components can be loaded on demand. This enables us to scale the pattern down to relatively small devices and customise the interpreter for specific application requirements and tasks. The ability to provide a constrained form of programming such as policies is equally important at the individual sensor level. It enables adaptive behaviour of the sensor according to context and thus to also adapt computational requirements and communication and hence power consumption. However, it also provides flexibility to reprogram the sensor with new adaptation strategies without requiring installation of new code.

The requirements we have used have been derived from the need for self-configuration and adaptation in e-Health applications. However, the resulting principles and framework developed are equally applicable to other application areas such as unmanned vehicles, ad-hoc networks, virtual collaborations as well as network and systems management.

The implementation of authorisation policies in both Ponder2 and Tiny Policy Interpreter is currently under way, as we are concerned with trust, security and privacy issues particularly for health-based applications. We are currently investigating security issues for body sensor networks in the CareGrid project [25].

Acknowledgment

We gratefully acknowledge financial support from the UK Engineering and Physical Sciences Research Council (EPSRC) through grants GR/S68040/01, GR/S68033/01 (Amuse Project) and EP/C547586/1 (Biosensornet Project). We also thank Ralf Damaschke for enhancing the Tiny Policy Interpreter during his UROP internship.

References

1. G.-Z. Yang (Ed.) *Body Sensor Networks,* Springer-Verlag, 2006.
2. N. Dulay, E. Lupu, M. Sloman, J. Sventek, N. Badr and S. Heeps. Self-Managed Cells for Ubiquitous Systems. In the *Proceedings of the 3rd International Conference on Mathematical Methods, Models and Architecture for Computer Network Security*, 2005.
3. Ponder2 Documentation Available: www.ponder2.net
4. The BSN Node Specification developed as part of UbiMon project. Available: http://vip.doc.ic.ac.uk/bsn/index.php?m=206
5. S. Strowes, N. Badr, N. Dulay, S. Heeps, E. Lupu, M. Sloman, and J. Sventek. An Event Service Supporting Autonomic Management of Ubiquitous Systems for e-Health, In *Proc. of the 5th Int. Workshop on Distributed Event-based Systems*, Lisbon, Portugal, July 2006.
6. E. Lupu, N. Dulay, M. Sloman, J. Sventek, S. Heeps, S. Strowes, K. Twidle, S.L. Keoh, and A. Schaeffer-Filho. AMUSE: Autonomic Management of Ubiquitous e-Health Systems. *Special Issues of the Journal of Concurrency and Computation: Practice and Experience*, Wiley (In Press).
7. N. Damianou, N. Dulay, E. Lupu, and M. Sloman, In *Proceedings of the International Worksop on Policies for Distributed Systems and Networks (POLICY 2001),* Bristol, UK, Jan 2001.
8. Bluetooth SIG, Inc., Available: http://www.bluetooth.org/
9. Zigbee Alliance, Available: http://www.zigbee.org/
10. TinyOS, Available: http://www.tinyos.net/
11. D. Gay, P. Levis, R. Behren, M. Welsh, E. Brewer, and D. Culler. The nesC Language: A Holistic Approach to Networked Embedded Systems, In *Proceedings of Programming Language Design and Implementation (PLDI)*, San Diego, June 2003.
12. XMLBlaster, Available: http://www.xmlblaster.org/
13. JMS, Available: http://java.sun.com/products/jms/.
14. B. Moore, E. Ellesson, J. Strassner, and A. Westerinen, Policy Core Information Model Version 1 Specification, Network Working Group, RFC2060, http://www.ietf.org/rfc/rfc3060.txt, 2001.
15. J. Lobo, R. Bhatia, and S. Naqvi, A Policy Description Language. In *Proceedings of the 16th National Conference on Artificial Intelligence,* Orlando, Florida, July 1999.
16. J. Strassner, *Policy-based Network Management*, Morgan Kaufmann, 2004.
17. D. Agrawal, S. Calo, J. Giles, K.W. Lee, and D. Verma, Policy Management for Networked Systems and Applications, In *Proceedings of the 9th IFIP/IEEE International Symposium on Integrated Network Management,* Nice, France, May 2005.
18. M. Roman, C. Hess, R. Cerqueira, A. Ranganathan, R. Campbell and K. Nahrstedt, A Middleware Infrastructure for Active Spaces, *IEEE Pervasive Computing,* 1(4):74-83, 2002.
19. D. Garlan, D. P. Siewiorek, A. Smailagic and P. Steenkiste, Aura: Toward Distraction-Free Pervasive Computing, *IEEE Pervasive Computing,* 1(2), 2002, pp. 22 - 31.
20. M. Kumar, B.A. Shirazi, S.K. Das, B.Y. Sung, D. Levine, and M. Singhal, PICO: A Middleware Framework for Pervasive Computing, *IEEE Pervasive Computing,* 2(3):72-79, 2003.
21. D. Malan, T. Fulford-Jones, M. Welsh and S. Moulton. CodeBlue: An Ad Hoc Sensor Network Infrastructure for Emergency Medical Care. In *Proc. of the International Workshop on Wearable and Implantable Body Sensor Networks*, April 2004.
22. V. Shnayder, B.-R. Chen, K. Lorincz, T.R.F. Fulford and M. Welsh. Sensor Networks for Medical Care. *Havard University Technical Report TR-08-05,* April 2005.
23. M. Strohbach, H-W. Gellersen, G. Kortuem and C. Kray. Cooperative Artifacts; Assessing Real World Situations with Embedded Technology. In *Proc. of the 6th Intl. Conference on Ubiquitous Computing.* Nottingham, UK, September 7-10, 2004
24. F. Siegemund. A Context-Aware Communication Platform for Smart Objects. In *Proc. of the 2nd International Conference on Pervasive Computing (Pervasive 2004), Vienna, Austria, April 2004.*
25. EPSRC CareGrid Project. Available: http://www.caregrid.org/

Address of the corresponding author:

Author:	Sye Loong Keoh
Institute:	Department of Computing, Imperial College London
Street:	South Kensington Campus,
City:	London, SW7 2AZ
Country:	United Kingdom
Email:	sye.keoh@imperial.ac.uk

Design Issues and Implementation of Query-Driven Healthcare System Using Wireless Sensor Ad-hoc Network

Wan-Young Chung[1], Gaurav Walia[1], Young-Dong Lee[1] and Risto Myllyla[2]

[1] Graduate School of Design and IT, Dongseo University, Busan 617-716, South Korea
[2] Dept. of Electrical and Information Engineering, 90014 University of Oulu, Finland

Abstract— **Application specific design requirements for wireless sensor network have posed new challenges and need for revision of existing designs for its implementation in healthcare systems. These requirements are different from that needed for environmental, agricultural and industrial purposes. The proposed paper discusses the design issues in implementation of Query driven Healthcare monitoring system using wireless sensor network. Various MAC layers designs have been studied for their usefulness and compatibility with the requirements in the healthcare system. Topology and network layer design have also been discussed. Further the low and computational capabilities of wireless sensor nodes itself adds to the design complexity. Therefore some compromises have to be made in both the domains. We aimed to implement a healthcare system using wireless sensor node based on the above study. A hardware platform was designed for the use as wireless sensor node. In this system, a wireless sensor node attached on the human body provides ECG (Electrocardiogram) and body temperature from multiple patients in an ad-hoc network. A mote-based 3-lead ECG and body temperature monitoring system operates in wireless sensor network. Health data from multiple patients can be relayed wirelessly using multi-hop routing scheme to a base-station, following IEEE 802.15.4 standard for wireless communication. Unique id assigned to each mote is used to identify each patient in the network.**

Keywords—**Wireless sensor network, Media Access Protocols, latency, Vital signal, healthcare.**

I. Introduction

With the rapid day by day advancement in miniaturization and low-power design techniques there have been fast active research in large-scale, highly distributed small-size, wireless sensors applications. In past few years various small sized computation and communication capable applications have emerged in almost every field including civil, military, environment monitoring, industries, security and disaster management. Healthcare and medical field also cannot remain unaffected by this new technology. The need for introducing wireless technology in healthcare is further supported by the fact that future healthcare management will need ubiquitous wireless monitoring of health with least actual interaction of doctor and patients. This would be highly useful for providing health monitoring in out-of-hospital conditions for older people and patients who requires regular supervision. Medical sensor networks must have support for ad-hoc routing topologies, mobility, wide ranges of data rates and high degree of reliability. Some of the healthcare system today uses Wi-Fi and other wireless LAN technologies. The European community's MobiHealth System (2002-2004) demonstrates the Body Area Network (BAN) [1]. Code blue [2] is a wireless infrastructure for deployment in emergency medical care. Another health monitoring system is Coach's Companion [3], which allows the monitoring of physical activity. CardioNet employs PDA to collect data from ECG monitor and send it over a cellular network to a service center [4]. Medtronic uses a dedicated monitor connected to the Internet to send pacemaker information to a medical professional.

However we think that the present systems have treated different healthcare data (i.e. from different sensors) with same view ignoring their reliability requirements, with little or no discussion regarding the issues needed to address in wireless sensor network for healthcare scenario. Waveform-independent healthcare data (example body temperature, blood pressure and oxygen content) which need low sampling rate and less monitoring should be differentiated from waveform-dependent (examples ECG, EKG) those that require high sampling rate and more reliability. Unlike typical wireless sensor environment where many sensors can sense same event simultaneously, in healthcare domain each sensor individually senses different events therefore the query model for healthcare system should be less complex in design.

This paper discusses the design issues for query driven healthcare system and its implementation using wireless sensor node in an Ad-hoc network in Section II. Our system design and approaches are discussed in Section III covering architecture, hardware, software and query model. Sections IV demonstrate some of our experimental results. Finally,

Section V concludes the paper and provides the future directions in this field.

II. Desgin issues

The primary objective of wireless sensor network is to maximize the node/network lifetime while performance metrics are secondary objectives. On the other hand the primary aim of wireless healthcare system should be the data transfer with minimum delay as patients health status is the utmost important. Therefore careful selection among available network routing and MAC (Media Access Control) protocols is needed for the combination that performs closer to our requirements and achieves both of the above mention objectives.

One of the main reasons of energy wastages is retransmission for packet lost during collusion, overhearing, idle listening and over-emitting [5]. In healthcare scenario recovery of lost packet is not needed because here data have to be real time, any lost packet will be replaced by most recent update. Also it will require more storage at the motes, which is not generally available. Since each node transfers data independently of other nodes in the network, protocols like SMAC [6], TMAC [7] and DSMAC [8] which needs RTS/CTS packet exchanges for lost packet recovery and have synchronized sleep and wake periods are ignored. Although SIFT [9] achieves low latency on cost of some increase in energy consumption it needs system-wide synchronization for slotted contention window and will be complex to use with protocols not using synchronization.

Since the choice of MAC layer protocol also depends on the application requirements, topology used and network layer routing protocol it worth to discuss this aspect along with the MAC layer issues. Healthcare system generally requires a single monitoring base station which collects data from all the sensors in the network and directs them for monitoring and further analysis. Therefore convergent type communication patterns can be wise choice. The paths from sources to sink can be represented as data gathering trees. For convergent type communication DMAC [10] protocol is the best choice as it achieves very good latency compared to other sleep/listen periods. Although DMAC does not avoid collusion detection but the possibilities of collusion is reduced because in healthcare system, events will not trigger all nodes simultaneously. BMAC [11] can also be possible candidate as it avoids packet collusion and have power management scheme via low power listening, but unlike SMAC does not performs link level retransmission or hidden terminal avoidance using RTS/CTS schemes[12]. Instead BMAC uses snooping on traffic over the broadcast medium for extracting information about the surroundings topology. Best thing about BMAC is that it offers control to the protocols above it, allowing the routing and application layer to change control parameters. This cross layer communication is highly demanded in healthcare and other applications.

As stated earlier the type of data also plays important role in the design of Data delivery model [13]. Waveform independent data do not need continuous data transfer therefore data transfer should be initiated only when desired by the monitoring side or in periods of longer durations. On the other hand waveform dependent data requires long time data transfer, in critical and emergency situations. Therefore data-centric approaches would be better for waveform dependent data, where a query or command from the base station initiates the data transfer. The waveform independent data can use either the data-centric approach or the event-driven approach where data is transferred when an event is sensed. It is generally observed that waveform independent data changes gradually without any periodic sequence. So only the change in amplitude is sometime required to detect an abnormal event. Therefore to limit the amount of data transferred a threshold limit can be set and if the value of sampled data goes beyond the threshold then only the data transfer is initiated. Also packet loss is not a major issue for these kinds of data because even if one packet is missed it can be replaced by the most current update after some time. For waveform dependent data any packet lost can be a big problem as it will result is loss of useful information or can also give false impression of abnormality.

Routing layer protocol is also very crucial in the overall performance of the complete system. Routing of message packets from or to moving nodes is more challenging since stability becomes an important optimization factor, in addition to energy and bandwidth etc. Since fast data delivery is prime requirement in healthcare proactive routing protocols will be efficient where path to be chosen is known in advance, before data transmission. Also any change in topology is updated by timely updates. Among all proactive protocols DSDV [14] seems less complex in implementation and supports flat network. The key advantage of DSDV is that it guarantees loop freedom.

Traditional methods of data aggregation cannot be applied in healthcare network as reading generated from different sensors will be at different rates, subject to different quality of service constraints and following multiple data delivery model. Also energy saving techniques that use data compression cannot be applied. Therefore such a heterogeneous environment makes data routing more challenging. Scalability is also major issue because single gateway architecture is not scaleable for large set of sensors.

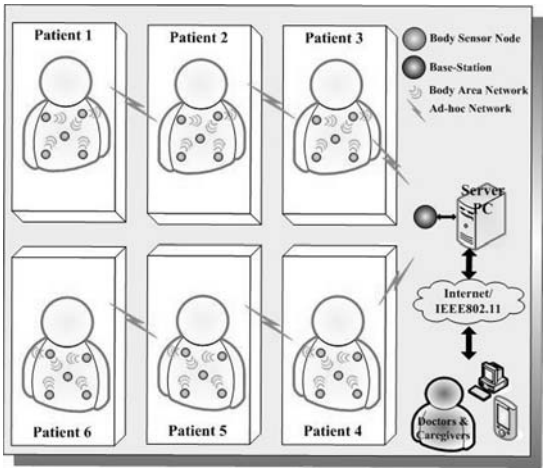


Fig.1 Architecture for Healthcare system.

III. SYSTEM DESIGN

Keeping the above mentioned issues in mind the ubiquitous healthcare prototype system for hospitals and home environment is designed. The concept of Ubiquitous healthcare system is to place unobtrusive wireless sensors on a person's body to form a wireless network which can communicate the patient's health status with base station connected to the monitoring PC. Among two types of health-care data mentioned above (i.e. waveform dependent and waveform independent) we have chosen one each from both types. Among wave dependent data ECG was chosen as it still remains one of the most commonly monitored vital signs in clinical and trauma care. Body temperature was selected as candidate among waveform independent data as it is easily sensed and needs low sampling rate and less frequent monitoring. Any other healthcare parameter can also be used depending on its category and design of interface.

A. Architecture

Fig. 1 shows the architecture of our system. In our prototype design we have used a wireless network of motes that provides sensor data from biomedical sensors placed on a patient's body. The wireless data communication follows bidirectional radio frequency communication with ad-hoc routing thus enabling every mote attached to patients, to send data to base-station even if they are not in direct radio range of base-station. The base-station comprises of developed USN node (excluding the sensors), for receiving and broadcast packets via node's USB. The USB port was set to 57600 bits per second.

Considering the issues discussed in previous section the desired energy saving is done by implementing on-command 'Sleep' and 'Wake_up' feature from PC side, which enables the motes to transfer data only when desired and sleep for the rest of the time [15], thus saving an useful amount of energy. During 'Sleep' mode only the sensing and associated data transfer is stopped but the radio of node is still active to receive, listen and route the data or query from the network. Here the cost of keeping the radio active have to be afforded because in healthcare system the priority of reliable data transfer is more than the lifetime of batteries. The query system provides the ability to access specific health parameter of any patient through a simple query feature. The signal from the electrodes of 3-leads system will first be amplified with low noise instrument amplifier, and then passed through 0.1Hz high pass filter (HPF) and 35 Hz low pass filter (LPF).The ECG signal will be sampled and filled in form of packets for transmission. The patient will be identified by the unique id assigned to each mote. Among various existing routing protocols, presently simple variant of Destination Sequence Distance Vector algorithm with a single destination node (the base-station) and active two-way link estimation has being used for routing and transmission. Base-station periodically broadcast its identity, which is used by receiving nodes for updating its routing information table. They then rebroadcast the new routing update to any nodes in their range. Thus a hierarchy of nodes is formed with base-station at top level and all nodes target their data to the nodes which are just above them. Any change in topology is conveyed to the parent nodes, which can update their table.

The MAC layer is designed on the using the cross layer communication features similar to BMAC. The application layer interacts with BMAC to listen weather the packet transfer was successful or not. Each mote in network simultaneously acts as transmitter, receiver and a router.

The base-station receives data packets from all the motes in the network and directs them to the attached PC acting as server. At server side user can view a GUI based window showing waveform for ECG and other parameters. The same GUI can also used to send the commands and query to the network. Thus medical professional can monitor the patients remotely, in a mobile and real-time environment. The health parameters can be stored for future reference, making a health report as desired by doctor. In case of emergency doctors or care givers can provide medical assistance well before time. This provides a better medical environment for near future with least inconvenience to patients.

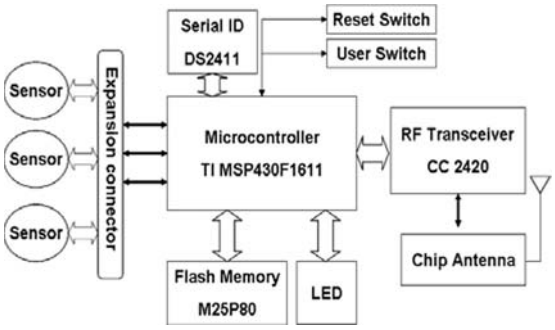


Fig.2 Block Diagram for USN platform

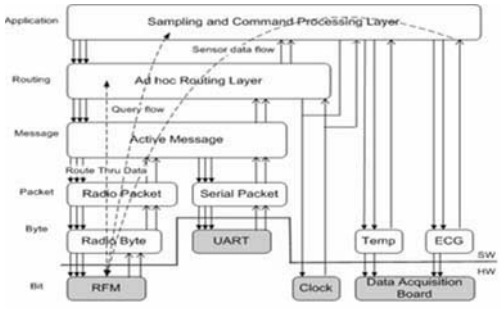


Fig.3 Application' component graph.

B. Hardware design

The designed architecture of USN (Ubiquitous Sensor Network) node platform is shown in Fig. 2. The essential part of USN node is the ultra low power Texas Instruments MSP430F1611 microcontroller, main features of which includes 10kB RAM, 48kB Flash, 128B of information storage and 8 channels of 12-bit A/D converter. Low current consumption (less than 1mA in active mode and ~1μA in standby mode) feature enables the node to run for longer duration. ChipCon CC2420 is radio chip for wireless communication and is IEEE 803.15.4 compliant. It provides PHY and some MAC layer functions. The radio chip is controlled by the microcontroller through SPI port and series of digital I/O lines. The M25P80 is a 4Mb (512 x 8) Serial Flash memory with write protection mechanism, accessible from SPI bus. To minimize the size of node we have made the programming interface as a separate module which is needed only when nodes is connected to the PC either for application download or when node act as base station.

IV. SOFTWARE ARCHITECTURE

The software architecture for the designed system can be broadly classified into two categories. ***Network programming model*** that deals with the software architecture for motes used in the network and ***query model*** that is associated software architecture for issuing command and queries from server to the motes.

A. Networking Programming model

Wireless sensor nodes run application software for sampling, query processing and routing, that was developed using ‘nesC language' [16] which runs on TinyOS [17]. The component-based architecture and event-driven execution model of TinyOS enables fine- grained power management while minimizing code size keeping in view the memory constraints in sensor network. Our architecture for software is based on Active Message communication model [18]. Fig. 3 shows the application component graph. There is component that provide asynchronous interface to each sensor and other component implements networking on the radio. The Lower Layer transmits or receives bytes bit by bit over the radio, providing phase and rate control. The packet level component spools the incoming bytes and delivers the packet receive event. Sampling component is periodically interrupted by clock, following which it acquires the sensor data, fill the packets and transmit them toward the base station in a multihop network. Query/command processing at the mote side is also implemented at the application level. Application level components have handlers connected directly to hardware interrupts, which can be external interrupts, timer events, or counter events. Transmission rate control is implemented within application component. If packet send is requested when radio is busy (i.e. either transmitting or receiving), the request is not granted. Once packet component has accepted a packet for transmission it will work until it acquires the channel and transmits it. During the busy state of radio the data capture is not affected and the incoming packets are stored in buffer queue. To ensure that waiting queue does not grow larger the older waiting packets are dropped to accommodate new packets.

B. Query Model

The Query sending system was implemented in Java and runs on the PC connected to the base-station. Queries send from monitoring system interacts with the sensor network. To keep the design simple we have used only one-time snap short data acquisition queries [19] instead of long-running (or continuous) queries. Also aggregation queries [20] are avoided because, as stated earlier in healthcare scenario all data channel and nodes work independently. Presently the

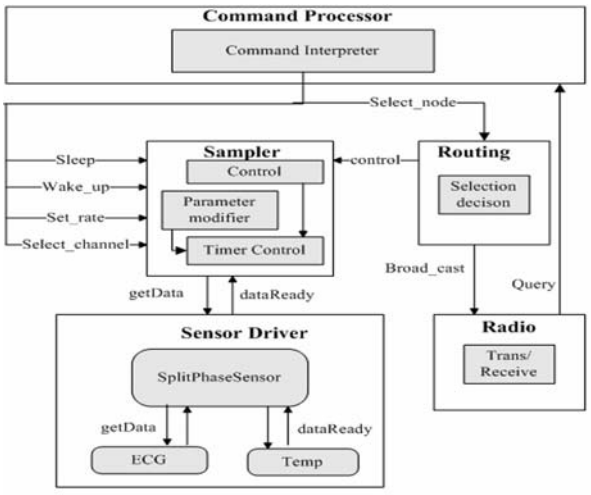


Fig.4 Query processing model on the node.

query system has four associated action or commands. The data of interest specified in query are collected from the nodes and send to the base station. User need to issue the 'select_channel' query, before which he must specify the target node. This will insure the data will obtained from the specific sensor of specified target node. Also the sampling rate can be altered by 'set_rate' command. For conserving the energy the motes can be send to sleep state individually or all in one time by issuing 'sleep' command. Any particular mote could be wakened by 'wake_up' command.

Fig. 4 shows our query model on the node side. Command Interpreter process the received command/query. If command is for the processing node then control activates 'Sampler' and parameters are modified accordingly. The modified parameters control the access of data from Sensor Driver. However if the command is for some other node then it is broadcasted via radio.

V. Experimental results

A. Hardware Platform

The designed USN platform is circular shape of 40mm diameter as shown in Fig. 5. It is in close approximation to serve as a platform for developing wearable biosensor. Also the stackable design of USN node and the available interface channel make it easy to integrate with different sensors. To add a new sensor, only interfacing of that sensor circuit with designed platform and a wireless command to set sampling rate is required. In our tests, ECG data was obtained from sensors attached to real human body, via ECG interface circuit.

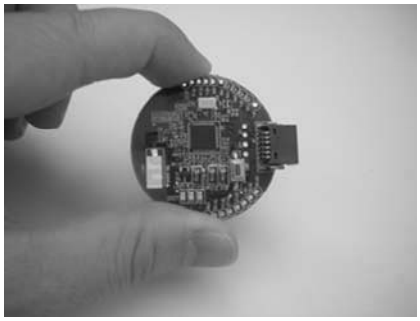


Fig.5 Developed USN platform measuring 40mm in diameter.

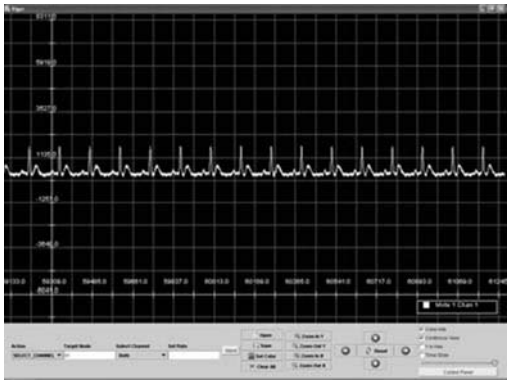


Fig.6 ECG data obtained at Base Station.

B. Obtained Data

Fig. 6 shows the ECG signal obtained at the remote base-station wirelessly on 'Vigor' tool. This tool was developed using 'java' language. The GUI of this tool enables to view the graphically, the obtained data, and query/commands can be send by selecting combination of buttons and checkboxes. This obtained signal is nearly identical to a theoretically calculated ECG signal.

Since packet loss information will help in evaluation of the performance of the network our software on the PC side kept vigilance on packet loss in the network. Fig.7 shows the snap shot view of the packet lost monitoring software. Each single field shows the source mote id, channel, time of arrival of packet and the total packet lost observed for that node. Here we observed that there is some packet loss during the starting phase of network discovery. Remedy for this packet loss is to broadcast initially the network discovery packets from parent nodes at faster rate to reduce the time that node takes to identify its parent. However since this may increase network traffic this rate should be brought back to normal after some time.

Fig.7 Packet Loss monitoring program.

VI. Conclusions

It is very difficult to address all the needs and constrains of healthcare systems with single MAC protocol, network protocol or topology, few compromises have to be accepted. Keeping discussed issues in mind we attempted to develop a healthcare system for hospital and home environment. The software for the query, sampling and data routing was developed using nesC language in TinyOS. The designed USN platform is ultra low powered and small enough to be used as wearable sensor node. External commands by the medical doctor or care givers for activating and deactivating the nodes can help in reducing the power consumption and traffic congestion during long time continuous monitoring. Simple query based access of data can be convenient for doctors and caregivers. By changing the interfacing circuits and sensors other healthcare parameters can also be measured. The data obtained can be analyzed by doctors and care providers to monitor a health status of patient in real time environment.

Acknowledgment

This research was partially supported by the program for the Training of Graduate Students in Regional Innovation which was conducted by the Ministry of Commerce Industry and Energy of Korean Government.

References

1. D. Konstantas, (2004) The Mobihealth Project. IST Project. IST-2001-36006, European Commission: Deliverable 2.6, http://www.mobihealth.org
2. Victor Shnayder, Bor-rong Chen, Konrad Lorincz, Thaddeus R. F. Fulford-Jones, and Matt Welsh (2005). Sensor Networks for Medical Care. Technical Report TR-08-05, Division of Engineering and Applied Science, Harvard University.
3. Lim, L. and B. Yee, Coach's Companion. Athlete's Health Monitoring System.University of California, Berkeley. http://www.limlloyd.com/coach/
4. P.E. Ross (December 2004). gManaging Care Through the Air. IEEE Spectrum, pp 14-19.
5. Ilker Demirkol, Cem Ersoy, and Faith Alagoz (2005). MAC Protocol for Wireless Sensor Network: a Survey. IEEE communication magazine.
6. W. Ye, J. Heidemann and D. Estrin (June2004). Medium Access Control With Coordinated Adaptive Sleeping for Wireless Sensor Networks. IEEE/ACM Transaction on Networking, Volume 12, Issue: 3, pp 493-506.
7. T.V. Dam and K. Langendoen (2003). An Adaptive Energy-Efficient MAC Protocol for Wireless Sensor Networks. The First ACM Conference on Embedded Networked Sensors Systems, Los Angeles, CA, USA.
8. P. Lin, C. Qiao and X. Wang (2004). Medium access control for dynamic duty cycle for sensor networks. IEEE Wireless Communication and Networking Conference, Volume: 3, pp 1534-1539.
9. K. Jamieson, H. Balakrishnan and Y.C. Tay(2003). Sift: A MAC Protocol for Event-Driven Wireless Sensor Networks. MIT laboratory for Computer Science , Tech. Rep. 894
10. G. Lu, B. Krishnamachari, C.S. Raghavendra (2004). An adaptive energy-efficient and low-latency MAC for data gathering in Wireless Sensor Networks. Proceedings of 18th International Parallel and Distributed Processing Symposium, pp 224
11. J. Polastre, J.Hill and D. Culler (2004). Versatile Low Power Media Access for Wireless Sensor Networks. ACM SenSys 2004
12. Umberto Malesci (2005). A Measurement-based Analysis of interaction among MAC, Network and Application Layers in Wireless Sensor Networks. Master Thesis, Department of Electrical and Computer Science, MIT, May 6 ,2005
13. S. Tilak et al. (2002). A Taxonomy of Wireless Microsensor Network Models. In ACM Mobile Communication and Computation Review (MC2R).
14. Charles E. Perkins and Parvin Bhagwat (1994). Highly Dynamic Destination Sequenced Distance-Vector Routing (DSDV) for mobile computers. In proceeding of SIGCOMM '94 Conference on Communication Architecture, Protocols and Applications, pp 234-244
15. Curt Schurgers, Mani B Shrivastava. (2001). Energy Efficient Routing in Wireless Sensor Network. MILCOM
16. David Gay, Philip Levis, David Culler and Eric Brewer (2003). nesC 1.1 Language Reference Manual. May 2003.
17. TinyOS at http://www.tinyos.net
18. Philip Buonadonna, Jason Hill, David Culler (2001). Active Message Communication for Tiny Networked Sensor. In Proceedings of the 20th Annual Joint Conference of the IEEE, 2001.
19. Philippe Bonnet, Jahannes Gaehrke and Praveen Seshadri (2000). Querying the Physical World. IEEE Personal Communications, Vol. 7, No. 5
20. Jeffery Considine, Feifei Li, George Kollios and John Byers (2004). Approximate Aggregation Techniques for Sensor Databases. ICDE

Address of the corresponding author:

Author: Wan-Young Chung
Institute: Dongseo University
Street: Jurye 2-dong, Sasang-gu
City: Busan
Country: South Korea
Email: wychung@dongseo.ac.kr

The Development of an In-Vivo Active Pressure Monitoring System

C.K. Lin[1], David Jea[2], Foad Dabiri[3], Tammara Massey[3], Robert Tan[1] Majid Sarrafzadeh[3], Mani Srivastava[2], Peter Schulam[4], Jacob Schmidt[1] and CD Montemagno[5]

[1] University of California/Bioengineering Department, Los Angeles, CA, 90032 USA
[2] University of California/Electrical Engineering Department, Los Angeles, CA, 90032 USA
[3] University of California/Computer Science Department, Los Angeles, CA, 90032 USA
[4] University of California/Urology Department, Los Angeles, CA, 90032 USA
[5] University of Cincinnati/College of Engineering, Cincinnati, OH 45221 USA

***Abstract*—Medical examinations often extract localized symptoms rather than systemic observations and snap shots rather than continuous monitoring. Using these methodologies, one cannot discretely analyze how a patient's lifestyle affects his/her physiological conditions and if additional symptoms occur under various stimuli. We present a minimally invasive implantable pressure sensing system that actively monitors long-term physiological changes in real-time. Specifically, we investigate pressure changes in the upper urinary tract per degree of obstruction. Our system integrates three components: a miniaturized sensor module, a lightweight embedded central processing unit with battery, and a PDA. Our tether-free system measures pressure continuously for forty-eight hours and actively transmits an outgoing signal from an implanted sensor node to a remote PDA twenty feet away. The software in this in-vivo system is remotely reconfigurable and can be updated when needed. Preliminary experimental results of the in-vivo pressure system demonstrate how it can wirelessly transmit pressure readings measuring 0 to 1 PSI with an accuracy of 0.02 PSI. The challenges in biocompatible packaging, transducer drift, power management, and in-vivo signal transmission are discussed. This research brings researchers a step closer to continuous, real-time systemic monitoring that will allow one to analyze the dynamic human physiology.**

Keywords— **Biomedical monitoring, implantable sensor, body sensor network.**

I. Introduction

Localized obstructions between the kidney and the urethra cause elevated pressures in both the upper urinary track (kidney, renal pelvis, and ureter) and lower urinary track (bladder, prostate, and urethra). Overlooked acute and chronical elevated pressure in the urinary track often leads to increased risk of infection, the formation of kidney stones, and irreversible damage to the bladder and the kidneys (if untreated for an extensive period of time) [8].

Kidney stones, ureter strictures, tumors, and uretopelvic junction obstruction block the upper urinary track. A symptom of urinary blockage is elevated pressure. In the lower tract, elevated pressures often results in involuntary bladder spasms and a lack of bladder control. Lack of bladder control and involuntary spasms are often sources of embarrassment and lifestyle impediment for the millions of men and women with urinary incontinence [8].

An urodynamics exam of the lower urinary tract is the customary diagnostic test for urinary track blockage. This exam is performed by placing a transurethral catheter into the bladder and the rectum. While supine, the patient's bladder is filled with saline over the course of two hours. This artificial stimulus causes considerable discomfort to the patient and often results in an inaccurate snapshot of pressure in the urinary track. If one could accurately determine the level of pressure within the urinary track, surgical and pharmaceutical measures can reduce the pressure and avoid harmful complications that result from prolonged elevated pressure [6][9].

To address the problem of localized pressure monitoring, an implantable *active* pressure sensor has been developed for the continuous measurement of elevated pelvic and ureteral renal pressures. This ambulatory sensor evaluates patients with ureteral obstruction and has applications in additional pressure applications, such as intracranial. The main benefits of the in-vivo pressure monitoring system include:

1. The continuous, *active* monitoring of pressure within the upper urinary track to determine how one's lifestyle affects physical symptoms;
2. The real-time automated distribution of data to a patient's PDA;
3. The aggregation of data to a online database that stores information for later analysis of the symptoms, and
4. The remote in-vivo reconfiguration of the software.

The overall goal of the in-vivo pressure monitoring system is to actively gather and distribute information on the pressure within the upper urinary track in a manner that is extremely fault tolerant. The technical challenges in devel-

oping the system are the design and fabrication of the transducer, the design and construction of the catheter, and the development of a reconfigurable program that can remotely update the implanted sensor node with future code modifications.

II. Background

Extensive work has been done in passive pressure monitoring where an incoming radio signal induces the pressure measurement. The earliest pressure capsules were developed by [7] and [2] in 1957 using a passive telemetering capsule that used the motion of the iron near the coil to determine internal pressure and temperature. Later, an implantable passive eye transistor measured pressure [1]. [3] investigated passive pressure sensors that did not contain a power supply or an active circuitry for high temperature environments. Recently, [4] developed a passive pressure sensor for acute uses with liquid crystal polymers and chronic uses using ceramic fabrication.

Passive measurements enable the construction of extremely miniature devices, but have a low range and must be close to an emitting radio signal in order for measurements to be collected. Compared with the above mentioned research, our in-vivo active pressure monitoring system has several unique features.

1. *Active Pressure Monitoring*. Unlike passive sensors, our active sensor has its own internal power supply that generates an outgoing signal. Active pressure monitoring has the benefits of increased transmission range and ubiquity in deployment. Our system can more easily communicate with current devices already being carried, such as PDAs. The drawbacks of battery lifetime and size are justifiable in applications where active communication of medical conditions are necessary.
2. *Collection Mechanism.* The data is transmitted to a PDA that uploads the data to an online database. This allows for the later analysis of the data to determine patterns in how one's lifestyle affects the internal urinary track pressure and the long-term progression of urinary track blockage.

III. An Active Pressure Monitoring System

The system design for continuous in-vivo monitoring of intraluminal pressures in the urinary tract includes integrated novel sensors, wireless miniaturized sensor nodes, and proven implant packaging technologies pioneered by Minimed Medtronic (Figure 1). The research focus is to safely implant these devices for forty eight hours of usage in animals and subsequently humans.

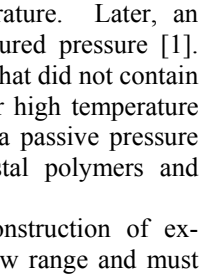


Fig. 1 Pigtail in-vivo catheter

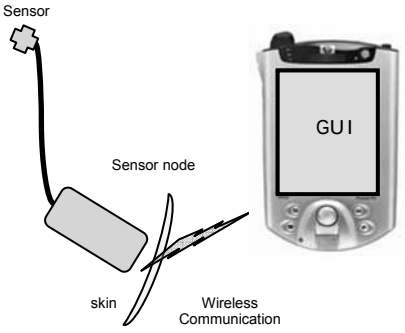


Fig. 2 System architecture of the active pressure monitoring system.

A. System Architecture

The system architecture is composed of three components:

1. *Sensor Node*: a computational unit (Mica2Dot) with a signal conditioning circuit and battery.
2. *Pressure Sensor*: ultra-small, low-cost OEM pressure die.
3. *Base Station*: a pocket PC that gathers and processes data from the sensor node.

The Mica2Dot measures 25mm and consists of a Atmega 128L microcontroller and wireless transmitter. The microcontroller has a 10-bit analog to digital converter, 128K of flash memory, and uses 8 mA of power. The multi-channel transmitter is has a center frequency of 868/916 Mhz with a data rate of 38.5 kbps. The radio uses 27mA of power while transmitting and 10mA of power while receiving. The Mica2Dot sends pressure readings to a PDA/Pocket PC that is the base station terminal of the system. The PDA visualizes the data and can perform more complex data processing, such as decision support. A general diagram of the system architecture is shown in Figure 2.

B. Pressure Transducers

A pressure transducer is a transducer that converts pressure into an analog electrical signal. Although there are various types of pressure transducers, one of the most common is the strain-gauge based transducer (Figure 3a). The conversion of pressure into an electrical signal is achieved by the physical deformation of strain gages that are bounded

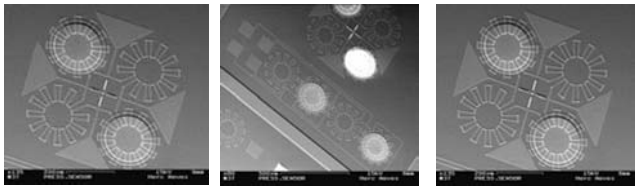


Fig. 3a First generation pressure transducers.

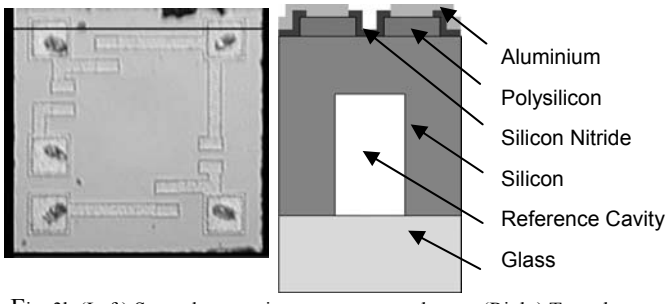


Fig. 3b (Left) Second generation pressure transducers. (Right) Transducer cross section

to the diaphragm of the pressure transducer and wired into a Wheatstone bridge configuration. Pressure applied to the pressure transducer produces a deflection of the diaphragm which introduces strain to the gages. The strain will produce an electrical resistance change proportional to the pressure.

A miniature OEM pressure die measuring 0.65 mm x 0.65 mm senses the pressure (Figure 3b). The die is an extremely small silicon micro-machined piezoresistive pressure sensing chip optimized to provide accurate readings for its minute size. The accurate precision is achieved through careful resistor placement and mechanical configuration. The sensor is configured as a resistive Wheatstone bridge. Bridges offer an attractive alternative to measuring small resistance changes accurately. The bridge consists of four connected resistors that form a quadrilateral, a source of current or voltage excitation across one of the diagonals. The voltage detector is connected across the other diagonal and measures the difference between the outputs of two voltage dividers across the excitation.

The desired range of pressure sensing in our application was 0 to 1 PSI with accuracy of 0.02 PSI. Resistance change corresponds to this range and results in a very small peak to peak voltage difference, 6mV. Since the pressure range in our application is 0 to 1 PSI, the voltage span is approximately 6mv. The small degree of voltage can be considerably affected by environmental noise. Even the fluctuation of the power supply voltage can cause considerable error in the circuit if it is not carefully designed. The precise signal conditioning circuit is responsible for:

1. *Stabilizing excitation voltage.* The sensor excitation voltage needs to be steadied at 3V.

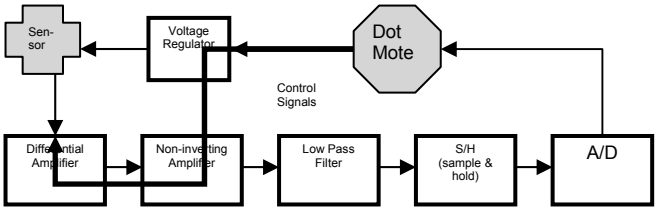


Fig. 4. Signal conditioning circuit in conjunction with the sensor and Mica2Dot.

2. *Electrical noise reduction.* Considerable noise is generated within the circuit and must be accurately filtered out.
3. *Voltage amplification.* The voltage range must be amplified to fully utilize the 0 to 3 voltage range.
4. *Offset voltage removal*: The sensors bridge is not symmetric when no pressure is applied. This lack of symmetry results in an offset voltage that reduces the voltage swing between 0 to 1 PSI.

An overall diagram of the sensor node, including the signal conditioning sensor and pressure sensor, is shown in Figure 4. The differential amplifier buffers the sensor from the rest of the circuit and partially removes the offset voltage. The second amplifying stage magnifies the voltage by a factor of approximately fifty-three. The high frequency noise is significantly reduced with the low-pass filter. As seen in the diagram, the voltage regulator is responsible to drive both the sensor and signal conditioning circuit with a stable voltage. The circuit detects a 1/50 PSI pressure change.

C. Catheter Design

The diameter, materials, and stiffness of the catheter has been carefully designed for usability, biocompatibility, and maneuverability for in-vivo operation (Figure 5). This is a 7.5 French catheter constructed of four platinum-iridium conductors wound in a helix around a high-tensile polyester core. The conductors are individually PTFE insulated and the catheter body is sheathed in silicone rubber. An inner

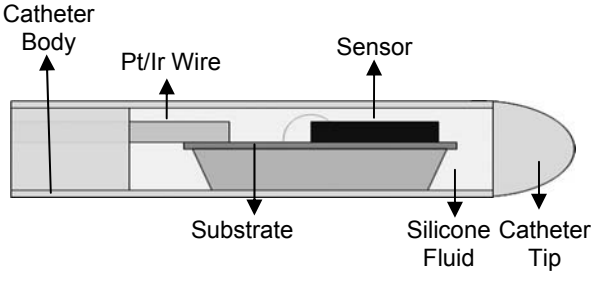


Fig. 5 Catheter Design Cross Section

Fig. 6 System components of the active pressure system: a PDA, a catheter, a wireless transmitter, and a lightweight microprocessor. A detailed picture of the catheter is shown in the upper right hand corner.

rubber tube of barium filled material improves radiopacity. A molded and strain-relieved coaxial connector (tip and ring) is affixed at one end of the catheter. The conductors are soldered to bonding pads on a PCB substrate and encapsulated with biocompatible epoxy. In addition, to relieve the anticipated forces exerted on the system during packaging and implantation, the monitoring system includes stress relief measures to ensure additional robustness (Figure 6).

D. Power and communication

The two main consumers of power in this system are the

1. Sensor bridge and signal conditioning circuit and the
2. Wireless communication.

The power source in the implantable platform is a li-polymer battery that cannot be recharged or changed after it has been implanted. Therefore, a simple power management strategy was adopted for the voltage regulator. The voltage regulator drives the sensor and the signal conditioning circuit through control signals from the central processing unit. The system excites the sensor for 20 ms during each sampling period (1 Hz frequency). After the sampling period, the device goes in inactive mode and power dissipation is greatly reduced when the circuit is completely shut off. This simple "pulse excitation" method reduces the power consumption significantly.

The multi-channel radio has a data rate of 38.5kbps and consumes 89 mW when transmitting data and 33 mW when receiving data. Sampled data is stored locally on the Mica2Dot and is transmitted to the PDA every thirty seconds to reduce power consumption and header overhead. The pulse excitation method and accumulated data transmissions increase the lifetime of the system from six hours to forty-eight hours.

E. Remote In-vivo Software Reconfiguration

After the sensor is implanted, the software may need to be modified. Therefore, an operating system designed specifically for embedded systems called Sensor Operating System (SOS) was used [5]. This message passing system severs the ties between the core operating system and individual applications or modules. A module measuring the pressure was created that could be loaded or removed at run time without interrupting the core operating system. A lightweight medical processor can transmit a program or module to be executed on the in-vivo processor. In-vivo code updates are essential in creating a system that can be reconfigured after deployment.

F. Biocompatible Packaging

Unsatisfactory packaging degrades the performance of the sensor and can result in device failure or a severe immunogenic response from the subject. Most researchers package their implantable sensors with only silicone dipping or parylene. However, most silicones are not designed to be used internally and parylene has been shown to attract immunogenic cells. With the knowledge of the pitfalls of conventional packaging schemes, we specifically designed the package to be a modular platform for catheter-based implantable sensors.

We packaged our system with two dual passivation layers, parylene and medical grade encapsulant (Figure 7). To combat the volatile environments in the body and the harsh chemicals in post processing, we evaporated a thin layer of parylene to protect the sensors. With its low surface energy, parylene deposition is not limited by line-of-sight and therefore conforms to any geometry on the micro-scale. Parylene forms a pin-hole free conformal coating of a thickness

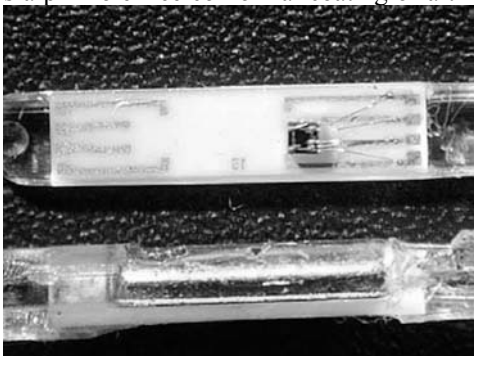


Fig. 7 Top and bottom view of the packaged sensor.

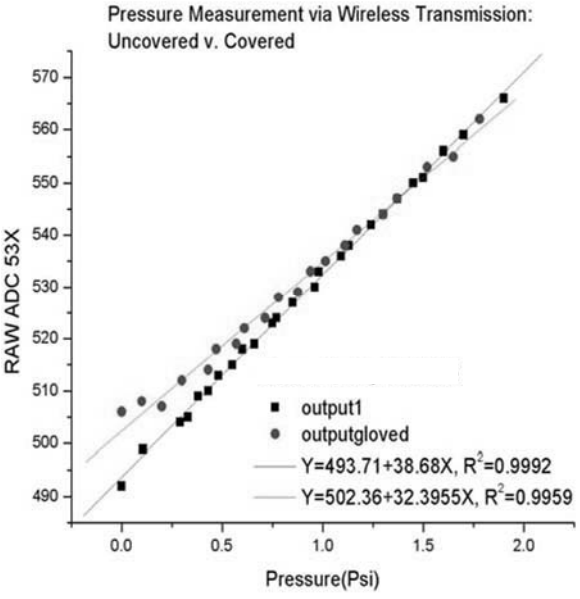


Fig. 8 Pressure measurements collected with wireless transmission in a lab setting.

as small as 0.03 um. Parylene also has high resistance to permeation and solvent absorption and strengthens wire-bonds at interfaces. Medical grade silicone, on the other hand, has excellent bulk property and malleability to surface properties. Silicone also has a long history of biological and biomedical applications. The pliability, surface property flexibility, and biocompatibility make silicone particularly attractive for long term medical device packaging material.

IV. Experimental Results

We tested the system in dry-lab setting and implanted it inside a pig. We first tested the pressure sensor to verify its accuracy and to determine the repeatability of the system. Figure 8 shows the first set of experimental results. We put the pressure sensor in a pressure chamber and modified the pressure between 0 to 2 PSI. The black line represents a linear fit to the data when the sensor node had no cover. We also covered the sensor node and the data also shown a linear fit with the red line. As seen in the figure, voltage-pressure dependency can be modeled as a linear function with less than 0.004 in standard deviation.

To combat drift, the signal was changed to a square wave form from a DC voltage. Then the duty cycle was minimized to a short time period that allowed the pressure to be accurately measured. The drift in each sensor is characterized by a trend line and the absolute values from various parameters are compared. As shown in Figure 9, our system addresses drift by lowering the duty cycle of the excitation voltage input. There is a 10% drift within ~800+ hours using the above methods (Table 1).

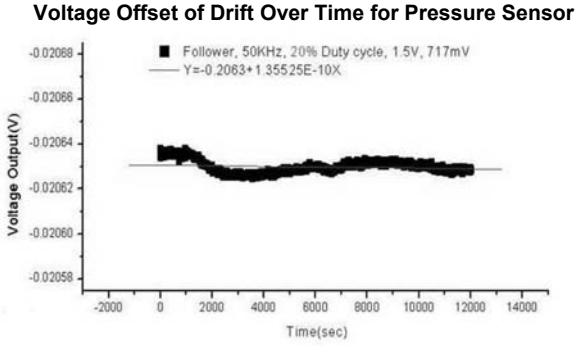


Fig 9. Analysis of voltage offset of drift over time using 20% duty cycle and 50KHz.

Table 1: Analysis of Drift Tolerance with Various Frequencies and Duty Cycles

Freq	Duty Cycle	ABS(Drift Rate)	Time to 10%(hours)
50K	50	7.086E-09	15.68
50K	20	1.355E-10	820.01
100K	50	2.000E-09	55.56
200K	50	1.660E-09	66.93
200K	20	1.760E-09	63.13

V. Discussion

We demonstrated the development of a minimally invasive, implantable system capable of continuous in vitro and in vivo data collection. This implantable telemetry system can collect comprehensive lifestyle data from a patient for an extended period of time, and therefore is a useful research and clinical management tool.

Several obstacles presented themselves during our investigation. During our characterization process, we discovered that upon voltage excitation our raw sensor output has a slight drift. Though the drift within forty-eight hours was minimal, during a longer implanted period the drift may reduce the usefulness of the system in the clinic.

To meet the forty-eight hour specification, we used a high capacity, miniature profile lithium polymer cell and used power conservation techniques. To further extend the lifetime of the system, we are investigating mechanisms to recharge the cell, to further reduce the duty cycle, and to optimize software excitation or circuit components.

Furthermore, we will also expand the range of sensor function to measure pulse, saturated oxygen, and temperature. This will be done through the in-house fabrication of these sensors on the same substrate as the pressure sensor.

This sensor platform will eventually enable the collection of massive physiological data in multiple parts of the body.

III. Conclusions

Over 200 million people worldwide have severe urinary incontinence or bladder complications. Although incontinence is not life-threatening, it can lead to various complications, such as kidney stones and infections. Since elevated pressures have strong correlations to the physical health of the urinary system, physiological pressure monitoring is critical. Conventional urodynamics examinations impose artificial filling and other risk factors influence how the urinary tract behaves. Therefore, we developed a miniaturized implantable system to accurately assess physiological pressures. These direct, continuous, and minimally invasive pressure measurements shed light on the diagnosis of conditions resulting from elevated pressure in the urinary track.

Pinpointing when irreversible changes in renal blood flow and function occur will help doctors to develop new strategies for treating these symptoms. Our in-vivo pressure monitoring system enables the creation of therapeutic guidelines for the lower urinary track. Moreover, the future coupling with an actuator system would enable timely delivery of local therapy and revolutionize the treatment of conditions, such as bladder instability.

Acknowledgments

This project is funded by USAMRMC TATRC. The authors would like to thank Ed Chernoff, Robert Shaeffer, and Andy Chen for their generous support and valuable input.

References

1. Collins CC, "Miniature Passive Pressure Transensor for Implanting in the eye." *IEEE Trans on Biomedical Engineering*, vol. BME-14, no. 2, 1967.
2. Farrar J, Zworykin V, and Baum J. "Pressure-sensitive telemetering capsule for study of gastrointestinal motility". *Science.* 1957, 126: 975-976.
3. Fonseca M, Kroh J, White J, and Allen M, "Flexible Wireless Passive Pressure Sensors for Biomedical Applications," *Tech. Dig. Solid-State Sensor, Actuator, and Microsystems Workshop (Hilton Head 2006)*, June 2006.
4. Fonseca M, English J, von Arx M, and Allen M. "High Temperature Characterization of Ceramic Pressure Sensors." *Proc Transducers 2001*. Vol. 1, p 486-489, 2001.
5. Han C., Rengaswamy R., Shea R., Kohler E., and Srivastava M. "SOS: A Dynamic Operating System for Sensor Networks". *Third International Conference on Mobile Systems, Applications, And Services (Mobisys)*, Seattle, Washington, 2005.
6. Haga Y. Biomedical micro-systems for minimally invasive diagnosis and treatment. *Proceedings of the IEEE*, 92(1), 98-114.
7. Mackay R, Jacobson B. "Endoradiosonde". *Nature*. 1957; 179: 1239-1240.
8. Piper I. "Urethral pressure measurement". Scand *J Urol Nephrol Suppl*, 207, 61-66.
9. Rose B. "Diagnosis of Urinary Tract Obstruction and Hydronephrosis" at www.uptodate.com. Accessed December 4, 2004.

Address of the corresponding authors:

Author: Chihkang (C.K) Lin
Institute: University of California, Los Angeles
Street: 7523 Boelter Hall
City: Los Angeles
Country: USA
Email: cklin@ucla.edu

Author: David Jea
Institute: University of California, Los Angeles
Street: 1762 Boelter Hall
City: Los Angeles
Country: USA
Email: dcjea@ee.ucla.edu

Author: Foad Dabiri
Institute: University of California, Los Angeles
Street: 3514 Boelter Hall
City: Los Angeles
Country: USA
Email: dabiri@cs.ucla.edu

Author: Tammara Massey
Institute: University of California, Los Angeles
Street: 3256N Boelter Hall
City: Los Angeles
Country: USA
Email: tmassey@cs.ucla.edu

Author: Robert Tan
Institute: University of California, Los Angeles
Street: 7523 Boelter Hall
City: Los Angeles
Country: USA

Author: Majid Sarrafzadeh
Institute: University of California, Los Angeles
Street: 3532C Boelter Hall
City: Los Angeles
Country: USA

Author: Mani Srivastava
Institute: University of California, Los Angeles
Street: 6731H Boelter Hall
City: Los Angeles
Country: USA

Author: Peter Schulam
Institute: University of California, Los Angeles
Street: UCLA Medical Center Box 951738
City: Los Angeles
Country: USA

Author: Jacob Schmidt
Institute: University of California, Los Angeles
Street: 7523 Boelter Hall
City: Los Angeles
Country: USA
Email: schmidt@seas.ucla.edu

Author: Carlos Montemagno
Institute: University of Cincinnati
Street: 801 Engineering Research Center
City: Cincinnati
Country: USA

Optimizing On-Chip Piezoelectric Energy Scavenging for Integration of Medical Sensors with Low-Power Wireless Networks

E. K. Reilly[1], L. M. Miller[1], P. K. Wright[1]

[1]University of California Berkeley, Department of Mechanical Engineering, Berkeley, CA, USA

Abstract— **Vibrational energy scavenging using piezoelectric material is a viable method to provide sufficient energy for low-power wireless sensor networks. The applications for such devices in hospital settings as well as *in vivo* are abundant. Current devices are limited by both their design and material selection. This paper will address optimizing the design of microscale devices by showing how the device strains under input vibrations are directly proportional to its power output, and by proposing alternate designs which increase the strain distribution over more of the device volume. Finite element modeling (ANSYS®) was used to determine the strain distribution in a cantilever, modified cantilever, trapezoid, and spiral shaped piezoelectric microscale energy scavenging system. The increase in strain under uniform acceleration was determined to be 0, 29.2, 37.8, and 87.0%, respectively, over that of a simple cantilever.**

Keywords— **Energy scavenging, MEMS, piezoelectric, sensor network**

I. Introduction

Wireless sensor networks have extraordinary potential for use in medical applications ranging from information management in hospitals and remote-monitoring of patient vital signs to wearable and embedded sensors [1]. However, many of these applications are currently limited by the inability to reduce the size of the sensor nodes or make them maintenance-free without eliminating the need to use a finite power source [2]. Batteries currently represent approximately 90% of the sensor node's volume. Our research focuses on developing microscale energy scavengers that use environmental vibrations to generate power, i.e. eliminating the need for external power sources. Integration of these scavengers with wireless sensor nodes will allow miniaturization of the nodes, as well as making them self-sufficient.

A variety of research and product development has been done on energy scavenging at the mesoscale (centimeter scale) to eliminate batteries as a power source, but microscale scavenging research is still very new [2-3]. Although miniaturization is not crucial for every sensor use, some especially promising applications, such as embedded sensors, will only be feasible if the sensors are as small as square micrometers in area [4]. Also, the sensors must not solely use batteries as a power source since the maintenance required for battery replacement is very undesirable in the case of embedded sensors. Energy can be scavenged by taking advantage of a number of physical phenomena, including the photovoltaic, electromagnetic, and piezoelectric effects. Many medical applications involve motion and vibration, making the piezoelectric effect the most useful for our research [5].

Our previous research has resulted in successful prototypes of MEMS energy scavenging devices, but the power output from these devices is not yet satisfactorily high [6]. Before the microscale energy scavengers can be integrated with the wireless sensor nodes, they need to be able to produce enough energy to power the node. The power output of the energy scavengers can be increased by optimizing the geometry of the vibrating structures to increase usage of the piezoelectric. The first piezoelectric beam shape used in this research was a cantilever beam. Finite element analysis (FEA) models of this geometry show only induced strain in the very uppermost part of the beam. This situation leaves a large amount of the piezoelectric unutilized for power generation. Other geometric beam shapes, such as modified cantilevers, trapezoids, and spirals, allow more of the piezoelectric area to be strained, according to finite element analysis. This paper will examine the corresponding change in power output from the modified energy scavenging structures.

II. Background

The constitutive equations for thin film piezoelectric cantilever beams are expressed in reduced tensor form below:

$$S_1^P = s_{11}^P T_1 - d_{31} E_3 \quad (1)$$

$$-D_3^P = d_{31} T_1^P - \varepsilon_{33}^T E_3 \quad (2)$$

Where S^x_i = strain, s_{ij} = mechanical compliance, T^x_i= stress, d_{ij}= piezoelectric coupling coefficient, E_i= electric field, D_i= dielectric displacement, and ε_{ij}= electrical permittivity. The superscript indicates the effected element, with E= elastic

layer, P= piezoelectric layer, and T= layer under constant stress, and the subscript indicates the direction.

When the beam is strained anti-parallel to the polarization the piezoelectric element creates an electric field across the thickness of the beam (Fig. 1).

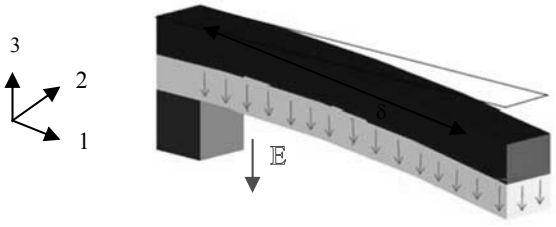


Fig.1 Schematic of piezoelectric cantilever under input acceleration

Ideally, the upper non-piezoelectric element will not interact with the induced field or resist deformation of the active layer under input mechanical deformation, but will serve as an electrode to harness the surface charge created. At equilibrium it may be assumed the there is no movement of the individual layers at the interface, i.e., the layers are tightly joined. Under the aforementioned assumptions the total electric field created under input vibration is:

$$E_3 = \frac{1}{d_{31}}\left[\frac{F^P s_{11}^P}{w h_p} + \frac{F^E s_{11}^E}{w h_e} + \frac{h_p M^P}{2E^P I^P} + \frac{h_e M^E}{2E^E I^E}\right] \tag{3}$$

Where F^x = applied forced due to vibration, M^x = applied moment, I^x= moment of inertia, h_x=thickness of the element, and w = width of the beam. The superscript or subscript "x" will represent the elastic layer, *e*, or the piezoelectric layer, *p*.

Several simplifications can be made due to the geometry of the system. First the film layers are sufficiently thin so that they have a common radius of curvature.

$$R = \frac{M^P}{E^P I^P} = \frac{M^E}{E^E I^E} \tag{4}$$

For a cantilever beam:

$$M^E = M^P\left(\frac{h_E}{h_P}\right)^3\left(\frac{s_{11}^P}{s_{11}^E}\right) \tag{5}$$

Secondly at equilibrium there is no net displacement in the 1-direction and the moments about the free end are zero.

$$F^E = F^P = F \tag{6}$$

$$\sum M_x = 0 \tag{7}$$

$$\therefore M^P + M^E = F\left(\frac{h_p}{2} + \frac{h_e}{2}\right) \tag{8}$$

Plugging Eq. 3, 5, and 8 into Eq. 1 gives the expression for electric field (E_3) as a result of input strain (S_1^P).

$$E_3 = \frac{S_1^p B}{(1 - d_{31}B)} \tag{9}$$

$$B = \frac{4s^p s^e h_e h_p^3 + 4s_{11}^p s_{11}^e h_e^3 h_p + (s_{11}^p)^2 h_p^4 + 6s_{11}^p s_{11}^e h_e^2 h_p^2}{4s_{11}^e h_p^3 + s_{11}^p h_e^3 - 6s_e^{11} z_p (h_p h_e + h_p^2) + 3s_{11}^3 h_p^2} \tag{10}$$

$$where\ (0 \le z_p \le h_p)$$

The electric field generated by the piezoelectric effect produces a charge, Q, over the length of the strained material. The voltage, V, due to the charge obeys the following equation:

$$Q = E_3 C h_p \tag{11}$$

$$V = \frac{Q}{C} \tag{12}$$

where C is the capacitance of the piezoelectric layer.

Assuming the capacitance is constant over the geometry of the design, the output power can than be estimated by

$$P = \frac{V_{\max}^{\ 2}}{R} \tag{13}$$

$$R_{opt} = \frac{1}{\omega C_o} \tag{14}$$

$$P = \left(\frac{S_1^p B}{(1 - d_{31}B)} h_p\right)^2 C_o \omega \tag{15}$$

Typical capacitance values for a MEMS piezoelectric cantilever range from 0.1- 1 nF [6]. The power equation demonstrates the dependence of the output power on the strain induced in the geometry. Increasing the amount of the structure that experiences appreciable strain will increase the power density of the device.

III. Modeling

ANSYS® finite element software was used to model the strain concentrations in various shapes in order to determine alternative configurations that result in higher strain under input vibrations. Designs explored include a modified cantilever (Fig. 2a, b), trapezoid (Fig. 3), and spiral (Fig. 4). These geometries were chosen to keep the resonant frequencies below 10,000 Hz, which is higher than what is available from most ambient vibrations but lower than the resonant frequency of many MEMS structures. Although these devices resonate higher than the first nodal resonant frequency of most environmental vibration sources, they have the potential to harness the subsequent nodal frequencies [7]. The first nodal resonant frequencies of the device geometries are located in Table 1.

Table 1 Resonant Frequency

Geometry	Resonant Frequency (Hz)
Cantilever	1000
Modified Cantilever	1113
Trapezoid	5583
Spiral	6247

The frequencies tend to increase as the geometry becomes more exotic. The resonant frequencies may be reduced through the application of a proof mass and changing the cross section (length, width) of the shapes. For modeling purposes the thicknesses of the various film layers were kept constant (1 μm PZT, 400 nm elastic layer) and the density of PZT used was 7800 g/cm^3. The strain was determined by applying an acceleration of 2.25 m/s^2, which is typical of those found in the residential environments, [7].

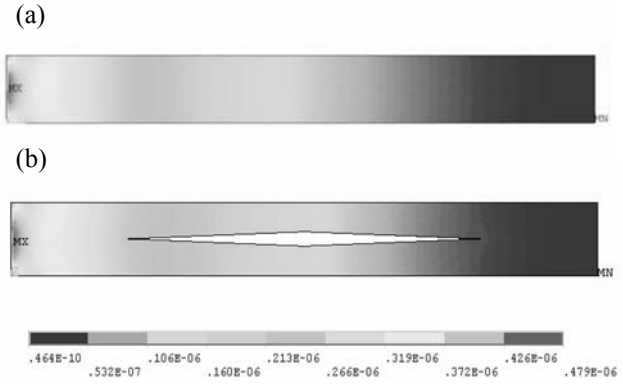


Fig. 2a Finite element analysis of a) cantilever, b) modified cantilever design

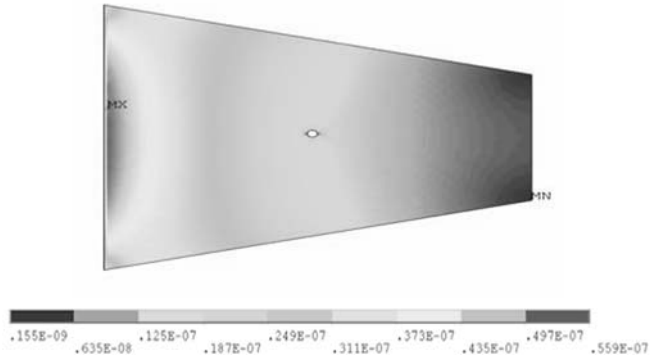


Fig. 3 Finite element analysis of trapezoid design

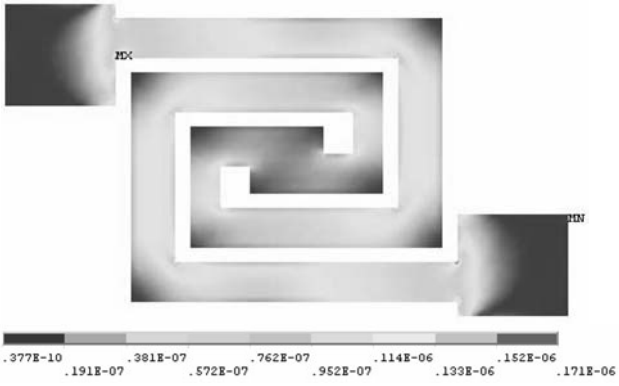


Fig. 4 Finite element analysis of spiral design

The strain in the structures was compared by calculating the percentage of the geometry that underwent at least 20% of the maximum strain present in each of the structures under input acceleration. Table 2 compares the percent increase in strain for different geometries over that of a simple cantilever.

Table 2 Geometry Strain Results and Percent Increase

Geometry	Percent of Area Strained (%)	Percent Increase (%)
Cantilever	40.4	0
Modified Cantilever	50.2	29.2
Trapezoid	55.6	37.7
Spiral	74.5	87.0

IV. Experimental results

The piezoelectric film, $PbZi_{0.47}Ti_{0.53}O_3$ (PZT), was grown via pulsed laser deposition on a Si wafer coated with 20 nm $SrTiO_3$ (STO) (provided by Motorola). The films were grown epitaxially using a bottom oxide electrode of $SrRuO_3$ (SRO) with a configuration as shown in Figure 5 [8].

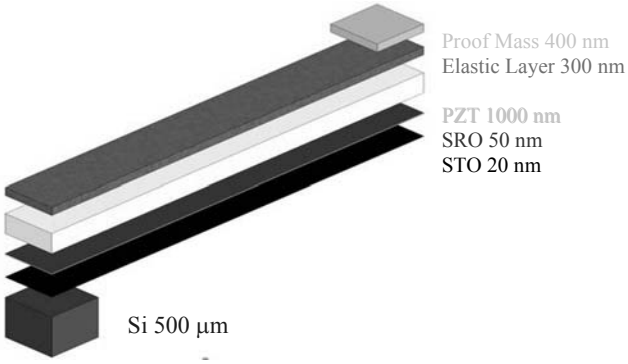


Fig. 5 Configuration of thin film layers

The metallic layer is deposited using electron beam and thermal evaporation, and the device is then released from the substrate using xenon difluoride (XeF_2) gaseous etch. Because significant residual stresses were present in the film, neutral ion bombardment was used to balance the residual stresses of the film with application of an induced compressive stress layer [9]. Fig. 6, 7, and 8 show scanning electron microscope photographs of the modified cantilever, trapezoid, and spiral films, respectively, fabricated for this study.

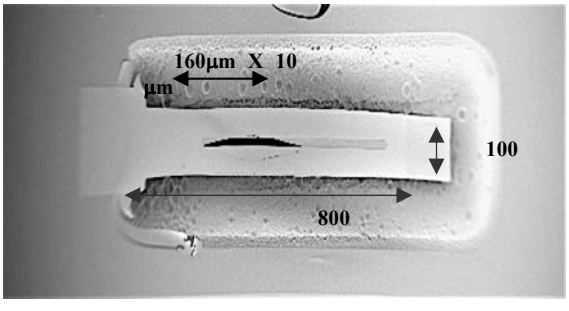


Fig. 6 SEM photograph of modified cantilever

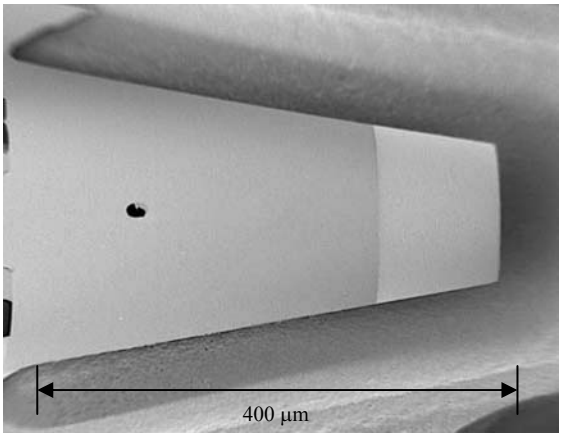


Fig. 7 SEM photograph of trapezoid

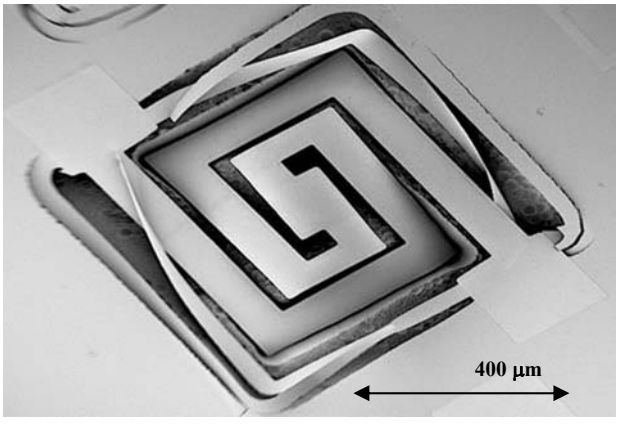


Fig. 8 SEM photograph of spiral

In order to help drive the deflection of the devices, proof masses were added to the free end, and, in the case of the spiral, the center. Testing is currently underway to determine the electrical characteristics of the devices. The mechanical quality factor of the standard piezoelectric cantilever is about 200 [6], and these new designs are expected to approach or exceed that result.

V. Conclusions

Microscale vibrational energy scavenging has gained popularity in recent years due the advancement in microfabrication techniques and ability to grow high quality piezoelectric films on silicon. Design of these devices becomes critical due to the small power density inherent in these devices. Maximizing the functionality of these devices

using the restricted dimensions of a MEMS system creates a challenging problem. Geometries which demonstrate distributed strain concentrations over the entire device will markedly improve the power output. The modified cantilever, trapezoid, and spiral configuration all represent improvements in that they make use of between 50 – 75 % of the available piezoelectric versus the standard cantilever which only uses about 40%. Devices of these designs have been successfully modeled and initial fabrication attempts have been promising. Testing is underway to determine the functionality of such devices in wireless sensor networks.

Acknowledgment

The authors would like to recognize the California Energy Commission, the Luce Foundation, and Berkeley-ITRI Research Center funding this work.

References

1. Aziz, O., Lo, Benny P. L., King, R., Yang, Darzi, A., Guang-Zhong, (2006) "Pervasive Body Sensor Network: An Approach to Monitoring the Post-operative Surgical Patient," Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks, Cambridge, MA, USA, April 3-5 2006, 13-16.
2. Glenn-Jones, P., Beeby, S., and White, N., (2001) "Towards a piezoelectric vibration-powered microgenerator, *IEE Proc. Sci., Meas. Technol.,* vol. 148, 2001, 69-72.
3. Kymissis, J. et. al.,(1998) "Parasitic power harvesting is in shoes," Proceedings of the Second IEEE International Conference on Wearable Computing ISWC, Pittsburg, PA, USA 1998.
4. Roundy, S., Otis, B., Chee, Y-H., Rabaey, J., and Wright, P.K., (2003) "A 1.9 GHz Transmit Beacon using Environmentally Scavenged Energy." ISPLED, Seoul, Korea, August 25-27, 2003.
5. Mitcheson, P., Reilly, E., Wright, P., and Yeatman, E., (2006) "Transduction Mechanisms and Power Density for MEMS Inertial Energy Scavengers," Proceeding of the Fourth International Conference on Power MEMS, Berkeley, CA, USA, 2006, *in press.*
6. Reilly, E., and Wright, P., (2006) "Thin Film Piezoelectric Energy Scavenging Systems for an On-Chip Power Supply" Proceeding of the Fourth International Conference on Power MEMS, Berkeley, CA, USA, 2006, *in press.*
7. Roundy, S., Wright, P. K., and Rabaey, J., (2003) "A Study of Low Level Vibrations as a Power Source for Wireless Sensor Nodes", *Computer Communications*, vol. 26, no. 11, 1131 - 1144.
8. Reilly, E. (2004) "Design and Fabrication of a Thin Film Piezoelectric Vibrational Energy Scavenging System," Masters Thesis, University of California Berkeley, Berkeley CA, 2004.
9. T. Bifano, H. Johnson, P. Bierden, and R. Mali, (2002) "Elimination of Stress-Induced Curvature in Thin-Film Structures," *J. Microelectromech. Sys*., 11, 5, (2002) 592-597.

Address of the corresponding author:

Elizabeth K. Reilly
University of California Berkeley
2111 Etcheverry Hall
Berkeley, CA 94703
USA

Email: beth@kingkong.me.berkeley.edu

A Smart Phone-based Personal Area Network for Remote Monitoring of Biosignals

M. J. Morón, J. R. Luque, A. A. Botella, E. J. Cuberos, E. Casilari and A. Diaz-Estrella

Dpto. Tecnología Electrónica, University of Málaga, Spain

Abstract— **This paper presents a system whose purpose is to monitor a patient continuously from indoor or outdoor environments. The system is based on a Bluetooth PAN, carried by the patient, whose central node, a smart phone, compiles information about patient's location and health status. These data are encrypted to be sent to a server through Wifi or GPRS/UMTS. The system provides facilities to access to patient's data, even from a smart phone by a J2ME application. It also allows to configure remotely the threshold values used to detect emergency situations.**

Keywords— **Bluetooth, Smart Phone, pulsioximeter, PAN, J2ME.**

I. Introduction

The increase in the processing and integration capacity of electronic devices, as well as the advances of low power wireless communications and, in general, wireless networking has enabled the development of unwired intelligent sensors for a wide set of applications. One of the most promising application fields is the medical telemonitoring of patients. Even a new concept: M-Health (or Mobile Health) has been proposed for the integration in a sanitary system of both mobile technologies and wireless sensors [1]. The generic goal of an M-health system is to increase the patient mobility and allow the enable the sanitary agents to access the medical information seamlessly and with independence of the physical location of both (patient and sanitary staff).

The incorporation of mobile technologies to the medical care has diverse benefits: mobile solutions have been shown to help improve patient safety, decrease the risk of medical errors and increase physician productivity and efficiency. Similarly, wireless sensors enable the patients' freedom of movements and promote new ways of patient monitoring such as home monitoring [2]. This clearly improves patients' quality of life. Missed days of school or work are reduced and health related restrictions on normal daily activities are minimized. Moreover, it allows the "service continuity" of the healthcare attention, as it defines a constant link to medical professionals who are able to assist in the disease management process. Patient telemonitoring facilitates the extension of healthcare services to remote and sparse-populated areas avoiding the need of expensive medical premises. On the other hand, providing physicians and clinicians wireless access to patient information and medical references largely eliminates the need to locate and read through patient charts or search for lab results from other departments. Telemonitoring also increments the medical presence in emergency scenarios and makes possible remote diagnosis.

As a consequence of the advances in wireless technologies, the architecture of a classical telemedicine service has strongly evolved [3]. Initial telemetry systems just contemplated the simple retransmission of the biosignals captured by wired sensors via a POTS modem or ISDN. Today new networking technologies as well as new communication paradigms (such as 'Context Aware' or 'Always Best Connected') enable new possibilities in telemonitoring services, ranging from limited indoor ('short distance') scenarios (e.g.: a care units) to outdoor applications without any mobility restrictions.

The paradigm of present (and future) medical monitoring systems is the definition of a W-PAN (*Wireless Personal Area Networks*) or W-BAN (*Wireless Body Area Networks*). The PAN (or BAN) must integrate a set of wearable wireless devices capable of sensing and transmitting biosignals. In most cases, the PAN/BAN is coordinated by a node which in turn may retransmit the signals to a remote central monitoring unit [4]. Thus, the general architecture of a PAN/BAN considers three components [2]:

- A medical sensor network (which can be completed with movement and positioning sensors), equipped with a low power, short range wireless interface, mainly based on Bluetooth or ZigBee technologies.
- A coordinator central node that communicates with sensors and, simultaneously, acts as an Internet gateway to other networks (GSM/GPRS/UMTS or WLAN) in order to transmit sensors biosignals (or medical alarms) and to receive control information.
- A central node (or a distributed central system) in charge of storing the sensors signals, detecting possible alarms and distributing the patients' information (e.g: via Internet, SMS, e-mail, etc) to the medical staff.

Optionally, the medical staff may receive the patients´ information in a handheld (PDA, phone) specifically designed or programmed for this purpose.

This article presents an architecture that defines a monitoring network of Bluetooth biosensors connected to a commercial 3G cell phone with a WLAN interface. The main goal of this architecture is to evaluate in the future the performance of smart phones when used as the gateway/master in a PAN of Bluetooth sensors. The evaluation especially will take into consideration the limitations of the extended J2ME (Java Mobile Edition) as a designing tool for this type of monitoring applications.

The rest of the work is structured as follows: Section 2 describes related works. Section 3 summarizes the specifications of the architecture and briefly describes the general structure of the prototype. Section 4 comments in more detail the implementation. Finally section 5 presents some conclusions and project's current status.

II. Related works on wireless PAN

The first projects that introduced the concept of Personal Area networks (PAN) [6][7] employed proprietary systems for wireless transmissions. The work in [6], for example, described a PAN that integrated the information generated by different intelligent sensors. The wireless interface of the PAN utilised a 916 MHZ RF transceptor from Link Technologies which provided a bi-directional bit-rate of 33.6 Kbps in a range of 50 m. Similarly, in [7] authors develop a PAN system which interconnects low-power sensors to PDAs and PCs by means a radio transceptor that operates at 916 or 433 MHz with a bit rate of 76.8 kbps and a coverage radio of 20-30 m.

In spite of these initial works, the present tendency in the field of medical PANs (and telemedicine in general) is the utilisation of standards for the different wireless communications that the PAN requires. The use of standards notably reduces the development cost while easing the deployment of telemedicine systems and product interoperability [8].

In case of requiring to reduce as much as possible the sensor consumption (which is basic aspect when dealing with implanted sensors whose batteries cannot be easily replaced), Zigbee/802.15.4 based transmission may be the best choice (see, for example, the works in [2] or [9]). However, in applications where battery restrictions are not so exigent, Bluetooth (BT) is by far the most utilised technology by designers to interconnect PAN sensors with the monitoring system unit (and or signal gateway). Among the main reasons that make this popular standard an attractive candidate to dynamically monitor physiological parameters the literature usually mentions the small size, reduced cost and low power consumption of the BT radio modules, the BT technique of frequency hopping (which increases security and privacy in radio transmissions) [4] [10], the capability of BT to conform scatternets and ad hoc networks, the potentiality of BT for wearable systems [11][8][12], and specially the present penetration of BT in the market (and its related commercial support) when designing solutions for short range transmission. Additionally the bandwidth supported by BT (up to 1 Mb/s) is enough to convey (and still multiplex) any real time biosignal.

Diverse prototypes of BT sensors have been developed for different biosignals, including ECG, glucometers, tensiometers, pulse-oximeters and even stethoscopes [10]. In fact, during last three years several vendors (Corscience, Nonin Medical, A&D Medical, etc.) have launched numerous homologated wireless biosensors with Bluetooth interfaces which can be easily integrated in a Body Area Network. The apparition of these commercial Bluetooth terminals permits a straightforward integration of general purpose devices (PDAs, embedded PCs, Mobile Phones,…) in the design of the PAN/BAN networks (which is the main goal of this work). So, the system can benefit from the computing power of these devices just by simply programming Bluetooth communications through conventional programming libraries (e.g. BlueZ). This increases the versatility and capability of reconfiguration of the network, reducing its deployment time and its final cost.

Different examples of networking architectures have been proposed in the literature to solve the problem of medical telemonitoring.

The AMON Project [13] has implemented a portable equipment in the wrist of the patient. This equipment is capable of measuring several biosignals simultaneously (SPO2, ECG, blood pressure, the temperature and the patient movements). The developed equipment processes the signals and, in case of medical alert, communicates with the medical center by means of the cellular network (sending a SMS or creating a connection of switched circuit through GSM). However, as it is difficult to integrate multiple biosensors in a single device, today most medical telemetry systems include some kind of 'wearable piconet' of independent wireless sensors. In most cases, the architecture consists of a (wearable or not) 'multiplexing' PAN node which collects the biosignals from one or more BT (or wired) biosensors. Once the signals are received, the node directly shows them on an embedded display or (more commonly) it retransmits them to a central node (normally located on medical premises) by means of a medium (Wifi) or long range (GPRS, UMTS) transmission technology.

For example, in the works [14] and [15] an ECG signal is transmitted via BT to a PALM device and a smart phone, respectively. These devices in turn forward the signal to a

server through GPRS. A similar system for ECG monitoring, presented in [16], adds to the GPRS transmitted flow the compressed information from a GPS. Analogously, the architecture in [4] allows a wearable unit (PDA or cell phone) to receive via BT the signals from a ECG but also from a pulse monitor and a tensiometer. In [17] a pulsioximeter and a tensiometer are integrated in the same device which transmits the signals to mobile phone via BT. A J2ME programmed application in the phone allows to process the signals and detect medical alarms. In that case, a SMS is sent to the medical staff. If the mobile service is not operative, the sensor (which can remotely configured) is able to store the signals temporally until GPRS service is recovered. The architecture in [18] also includes a mobile wearable equipment which act as gateway between the BT sensors and a GPRS/UMTS connected central unit. The main particularity of the system is the utilisation of XML for the configuration of the sensors.

Authors in [5] develop a piconet of BT smart sensors. The sensors can be programmed by a control node which is remotely accessed by GPRS/UMTS. The same authors present in [19] another piconet of wearable sensors. The control node is situated in a cell phone or PDA with Wifi, GPRS and UMTS interfaces.

A similar system is described in [3]. In this case, the piconet is substituted by a set of wired sensors connected to a microprocessor board specifically designed for this application. The board also includes a BT module to intercommunicate with the hospital through a GPRS mobile phone. In the reception part medical staff can receive the information of the patients in portable PDAs.

In contrast with previous works (in which BT communications are programmed at HCI or L2CAP layers), the main novelty of the BT piconet described in [20] is that, in order to ease the interoperability among different vendors, BT sensors are accessed by using Bluetooth Serial Port Profile.

The work in [21] introduces a PAN architecture comprising a set of plug and play BT sensors (including ECG, pulsioximeter, humidity and temperature sensors) controlled by a wearable data logger. By using BT the logger connects to an Internet-attached base station which stores the signals in a distributed data base. The interoperability of the sensors is achieved by means of the ISO/IEEE 11073 (Medical Information Bus) standard.

Authors in [22] remark that the main goal of a telemonitoring system is to transmit eventual medical alarms. As these systems are aimed at patients and elder people who spend most time at their homes, the work proposes the use of not-wearable (fixed) BT access points that will be responsible for collecting the signals and transmitting them to a central unit (a PDA) in a hospital.

From these experiences, we can remark the growing use of 'general purpose' mobile devices (mainly PDAs or Cell Phones) to build the PAN node which acts as the gateway between the wireless sensors and the final remote control node. This can be justified by the increasing computing power, storing and communicating capabilities (BT is included in an important percentage of models) as well as by the decreasing costs of these devices. Moreover, due to the universalised use of cell phones in developed countries, the implementation of the medical BAN central node can be performed without introducing any new wearable hardware and just with a re-configuration of a familiar and quotidian element in the everyday life of the patients (the mobile phone).

As it refers to the programming language that is normally chosen to build the software in the phones (or PDAs), initial architectures utilised C or C++, which permitted an optimised design for real time processing and a better interaction with underlying operative system (e.g: Symbian in the case of most phones). However, since the apparition of J2ME (Java version for mobile devices) and in spite of its limitations (for example, with the APIs to incorporate BT), the portability of Java is notably stimulating its adoption in the field of medical BANs.

III. System description

The profile of the target user is a patient who needs continuous monitoring. The patient usually resides in his own home or in a shared environment with other patients or users, i.e. a residence for elder people. This scenario will be referred as a "controlled environment". Besides, health status is not an obstacle to leave the *controlled environment* to go for a walk, for example. In this case the patient will be in "outdoor scenario". The purpose is to carry out a smart tracking of patients to achieve the following advantages:

- To avoid patients to have to stay on bed and connected by cable to a medical monitor
- To increase the patient's mobility and comfort level.
- To allow remote monitoring in outdoor spaces.
- To reduce number of periodical check-up.

A. Functions

It is intended to provide the next functions:

1. Monitoring parameters related to health status in order to detect an emergency.

2. Recording of the outdoor location of patient.
3. To provide medical assistance in case of emergency.

The wireless transmission of patient's information will be protected by means of authentication and encryption.

B. Architecture

Regardless the patient's environment, there are three basic parts or components in the system:

1. An iBAN worn by the patient to be remotely monitored.
2. A Central Control Server (CCS), located in the controlled environment.
3. A Mobile Control and Monitoring Unit (MCMU) carried by the physicians.

The iBAN prototype is based on Bluetooth technology. Currently the iBAN comprises a commercial medical sensor, a pulsioximeter, a GPS device and an *Intelligent Node*, referred as IN. The IN integrates two wireless communications interfaces, 802.11 and Bluetooth, as well as a connection to cellular network. Bluetooth is the technology used for communications between the devices and the IN, which plays the master role in the piconet. Wifi and/or GPRS/UMTS are employed to send information from the IN to the CCS and to the MCMU.

In the following subsections, the interaction between the different components is described.

C. Patient iBAN – CCS Interaction

An application, which resides in the IN, manages the Bluetooth piconet and the CCS connections. When it is automatically started, i.e., without user request, the next sequence of steps is carried out: (1) inquiry procedure, (2) services search, (3) connection establishment with pulsioximeter and GPS devices, (4) pulsioximeter configuration and (5) server connection and user registration.

As IN receives data from iBAN nodes, it stores them during a period which can be remotely configured. When the time is out, the IN sends these data to the server by the 802.11 interface in case that the user is in a controlled environment and there is an available Wifi access point. In other case, the data will be transmitted by GPRS/UMTS.

As in [5], the Bluetooth connections are kept active as long as the devices are in the coverage area of the IN. Just as the authors of [5] point out, the advantage is that, as soon as an event occurs, the IN can notify it to the server. However, the main drawback of this policy is the increase in power consumption.

The data rate depends on the mode operation of the iBAN:

- *Default Mode (mode 1)*: The pulsioximeter sends 3 bytes/s. The information from GPS is also received and stored in the IN, but it is not retransmited to the CCS.
- *Verbose Mode (mode 2)*: This operation mode has to be remotely enabled from the server or the MCMU. In this case the pulsioximeter sends a rate of 375 bytes/s. SPO2 data and GPS parameters are delivered to the CCS.

Furthermore, if the verbose mode is activated and the GPS device does not provide valid data, the latest stored parameters are recovered.

Additionally, in both modes, the information received from the pulsioximeter is processed in order to verify that the oxygen saturation and the heart rate are not out of the security range established for the patient. If the IN detects a value lower or greater than a predefined threshold, a SMS will be sent to the physician mobile, and, if desired, to some patient's relatives. The thresholds can be set up by a physician from the server or his own MCMU.

Furthermore, periodically the CCS requests IN to switch to verbose mode in order to check battery status. This is necessary because the chosen commercial pulsioximeter only includes this indicator under this mode.

The sequence diagram displayed in figure 1 shows the initialization procedure, the default mode operation and the enabling of the verbose mode.

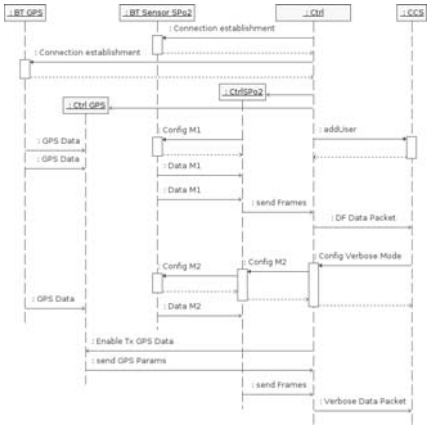


Fig. 1 Diagram of the interaction between the iBAN sensors and the CCS

D. MCMU-CCS Interaction

When a MCMU receives an alert SMS from a patient, the physicians starts a J2ME control and monitoring application. This one allows to select a patient and the sensor whose data the physician wants to display. In order to get requested data, a *midlet* establishes a TCP connection with the CCS. Through this connection, the CCS retransmits the

frames received from iBAN in quasi-real time. The transmission is not stopped until the physician selects the corresponding option.

Additionally, the application provides the possibility of configuring the operation mode, the thresholds values used to detect an emergency or even the transmission period. In order to guarantee the reliability these configuration commands are sent through a TCP connection.

IV. Implementation

The next devices has been incorporated in the iBAN prototype: (1) a smart phone, initially a Nokia 9500 model and later, a Nokia N93, as hardware platforms for the intelligent node (IN), (2) a Bluetooth pulsioximeter, from Nonin Medical [23]. This sensor implements SPP profile, and a Bluetooth GPS receiver with SirfStarIII chipset and SPP profile too.

For the Mobile Control and Monitoring Unit (MCMU) a smart phone has been also employed.

A. IN and MCMU

A J2ME application has been implemented for the IN. The first version, developed for the Nokia 9500 (Series 80 2nd.), uses the next APIs: *Bluetooth API* (JSR-82) for managing BT connections and *Wireless Messaging API* (JSR-120) for sending SMSs.

The *midlet*, in a second version implemented for the Nokia N93 (*Series 60 3rd. Edition*), extends the functionality by using: (1) *Security and Trust Services API* (JSR-177) to include data encryption; (2) the *Location API* (JSR-179) to get data from GPS with more reliability. The last version of *Wireless Messaging API* (JSR-205) has been used instead of JSR-120. Additionally, the *File Connection* API (JSR-75) has been employed to record patient's information in files that can be created into the file system residing on the mobile device or on an external memory card.

The need of privacy when transmitting and processing medical data is the main reason to select the model Nokia N93. To encrypt data in the transmission to CCS, a symmetric algorithm DES with a prefixed key, previously arranged between both communication end points, is utilised.

The control and monitoring application for the MCMU also has been developed as a *midlet*. The program also employs *Security and Trust Services API* (JSR-177).

The screenshot displayed in fig. 2 shows pulsioximeter's parameters during an execution in Nokia N93.

Fig.2. Screenshot of monitoring *midlet*

B. CCS

The CCS is an Apache web server implemented as two *servlets* under Tomcat. The first one carries out two main functions: (1) it manages the connection requests from patients to be monitored, compiling and processing medical data (2) it provides access to the IN capabilities by a web interface using applets. The second servlet serves access requests from the MCMUs carried by physicians and retransmit data from the IN to MCMU through a TCP connection.

The communication between CSS and patient's iBAN can be established using two protocols: (a) *HTTP*: data are sent in the HTTP request itself. If the server has to send a command, it will use a HTTP message. (b) *TCP-UDP/IP*: commands requested to the IN as well as the responses to commands and indications are transmitted through TCP, whereas the data are sent over UDP.

In relation to provide an easy access to physicians, two mechanisms are included:

1. *Web interface*: The physician can access the system through a PC, from the hospital or in any location with Internet access, using a navigator with RMI support, such as Firefox, by means of a *JApplet*.

   This *JApplet* is structured in two main components: (a) *Configuration Interface*, to manage the iBAN configuration and operation mode, and (b) *Display Data*, to show health parameters and location information. The *Configuration Intereface* component communicates with CSS by The RMI. *Display Data* subsystem also uses RMI to get, in real-time, the parameters extracted from decoded frames in the server.

2. *Mobile Network*: The control and monitoring *midlet* installed on the MCMU. The physician only has to start the *midlet*, which will establish a TCP connection with the server to receive the encrypted frames from the IN which were redirected by the CSS.

VI. Conclusions

The use of smart phones as central node in a PAN to compile information about patient's location and health status, is favourable because it takes advantage of the user's acquaintance with mobile device. However, the main restriction from usability perspective is that the monitoring application carries out tasks considered by operating system as risky for security. Therefore acknowledgement is requested to user before continuing certain actions. Nevertheless, this drawback could be avoided by means of a validated certificate, instead of the self certificate used.

On the other hand, the J2ME implementation for monitoring applications is advantageous since its portability is greater than Symbian. Nonetheless, it has been detected that the problems occurred in the *midlet* implementation depend on the smart phone. The phone Nokia 9500 generates Symbian errors during the SPP connection establishment. It has been checked that these errors also happen with other applications that use the SPP profile. However, the SPP connection establishment works successfully in Nokia N93 and Nokia N70 (Series 60 2nd. Edition FP3). However, only in Nokia N93, sometimes, the *midlet* freezes during Wifi access point searching. Symbian errors also take place in Nokia 9500 when a Bluetooth connection with a device is dropped and the application tries to recover it.

Finally, it has been verified with Nokia N93 that the application continues running when the cover is closed, on the contrary to other experiences [18].

Acknowledgments

This work has been supported by Projects No. TIC2003-07953-C02-01 and TEC2006-12211-C02-01/TCM.

References

1. Istepanian R S, Jovanov E, Zhang Y T (2004) Introduction to the Special Section on M-Health: Beyond Seamless Mobility and Global Wireless Health-Care Connectivity, *IEEE Trans. on Inf. Tech. in Biomedicine*, vol. 8, pp 405–414, Dec. 2004.
2. Jovanov E, Milenkovic A, Otto C et al. (2005) A wireless body area network of intelligent motion sensors for computer assisted physical rehabilitation, *Jour. of Neuroengineering Rehabilitation*, vol. 2, no. 6.
3. Rasid A, Woodward B (2005) Bluetooth Telemedicine Processor for Multichannel Biomedical Signal Transmission via Mobile Cellular Networks, *IEEE Trans.on Inf. Tech. in Biomedicine,* vol. 9, pp 35–43.
4. Hung K, Zhang, Y, Tai B (2004) Wearable medical devices for tele-home healthcare, in *Proc. of Conf. EMBC 2004*, vol. 7, pp 5384–5387.
5. Krco S (2003) Implementation solutions and issues in building a personal sensor network for health care monitoring, in *Inf. Tech. Applications in Biomedicine, 2003. 4th International IEEE EMBS Special Topic Conference on*, pp 350–353.
6. Jovanov E, Price J, Raskovic D et al. (2000) Wireless Personal Area Networks in Telemedical Environment, in *Proc. of the IEEE EMBS Intern. Conf. on Inf. Tech. Applications in Biomedicine*, pp 22–27.
7. Malan A D, Fulford-Jones T, Welsh M (2004) CodeBlue: An Ad Hoc Sensor Network Infraestructure for Emergency Medical Care, in *Proc. of MobiSys WAMES 2004, Boston, MA*, pp 12–14.
8. Warren S, Yao J, Schmitz R et al. (2003) Wearable telemonitoring systems designed with interoperability in mind, in *Proc. of IEEE Eng. in Medicine and Biology Society, 2003*, vol. 4, pp 3736– 3739.
9. Timmons N, Scanlon W (2004) Analysis of the performance of IEEE 802.15.4 for medical sensor body area networking, in *Proc. of IEEE SECON 2004*, pp 16–24.
10. Hung K, Zhang Y (2002) Usage of Bluetooth in wireless sensors for tele-healthcare, in *[Engineering in Medicine and Biology, 2002. 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society] EMBS/BMES Conference, 2002. Proc. of the Second Joint*, vol. 3, pp 1881–1882.
11. Warren S, Yao J, Schmitz R et al. (2004) Reconfigurable Point-of-Care Systems Designed with Interoperability Standards, in *Proc. of EMBC 2004*, vol. 2, pp 3270–3273.
12. Yao J, Schmitz R, Warren S (2005) A wearable point-of-care system for home use that incorporates plug-and-play and wireless standards, *Information Technology in Biomedicine, IEEE Trans. on*, vol. 9, pp 363–371.
13. Anliker U, Ward J, Lukowicz P et al. (2004) AMON: a wearable multiparameter medical monitoring and alert system, *IEEE Trans.on Inf. Tech. in Biomedicine*, vol. 8, pp 415 – 427.
14. Dong J, Zhu H (2004) Mobile ECG detector through GPRS/Internet, in *Computer-Based Medical Systems, 2004. CBMS 2004. Proc.. 17th IEEE Symposium on*, pp 485– 489.
15. Khoor S, Nieberl K, Fugedi K et al. (2001) Telemedicine ECG-telemetry with Bluetooth technology, *Computers in Cardiology 2001*, pp 585–588.
16. Liszka K, Mackin M, Lichter (2004) Keeping a beat on the heart, in *Pervasive Computing, IEEE*, vol. 3, pp 42–49.
17. Lee R G, Hsiao C C, Chen C C et al. (2006) A mobile-care system integrated with bluetooth blood pressure and pulse monitor, and cellular phone, *IEICE Trans. on Information and Systems*, vol. E89-D, pp 1702–1711.
18. Krco S, Kostic S, Sakac D et al. (2005) mSens mobile health monitoring system, vol. 1, pp 80–83, The International Conference on Computer as a Tool. EUROCON 2005.
19. Krco S, Delic V (2003) Personal wireless sensor network for mobile health care monitoring, in *Proc. of TELSIKS 2003*, vol. 2, pp 471–474.
20. Wang D, Lu Y, Zhang H et al. (2005) A wireless sensor network based on Bluetooth for telemedicine monitoring system, in *Microwave, Antenna, Propagation and EMC Technologies for Wireless Communications*, vol. 2, pp 1361– 1364.
21. Warren S, Lebak J, Yao J et al. (2005) Interoperability and Security in Wireless Body Area Network Infrastructures, in *Engineering in Medicine and Biology Society, 2005. IEEE-EMBS 2005. 27th Annual International Conference of the*, pp 3837–3840.
22. Park D, Kang S (2004) Development of reusable and expandable communication for wearable medical sensor network, in *Proc. of. EMBC 2004*, vol. 7, pp 5380– 5383.
23. Nonin Medical Inc. at http://www.nonin.com

Real-Time Pervasive Monitoring for Postoperative Care

Benny Lo[1], Louis Atallah[1], Omer Aziz[2], Mohammed El ElHew[1], Ara Darzi[2] and Guang-Zhong Yang[1]

[1] Institute of Biomedical Engineering
[2] Department of Biosurgery and Surgical Technology
Imperial College London

Abstract—**Post surgical care is an important part of the surgical recovery process. With the introduction of minimally invasive surgery (MIS), the recovery time of patients has been shortened significantly. This has led to a shift of postoperative care from hospital to home environment. To prevent the occurrence of adverse events, the care of these patients is mainly relied on routine visits by home-care nurses. This type of episodic examination can only capture a snapshot of the overall recovery process, and many early signs of potential complication can go undetected. The development of Body Sensor Networks (BSNs) has enabled the use of miniaturised wireless sensors for continuous monitoring of postoperative patients. This paper examines the potential of processing-on-node algorithms for further reducing the wireless bandwidth, and therefore the overall power consumption of the sensors. The accuracy and robustness of the technique are demonstrated with lab experiments and a preliminary clinical case study.**

Keywords— **postoperative care, real-time analysis, Bayesian classifier, multivariate Gaussian model**

I. Introduction

Recent advances in Minimally Invasive Surgery (MIS) have greatly reduced patient trauma and improved post-surgical recovery. Unlike traditional open surgical approaches, MIS procedures are conducted through small body incisions with laparoscopic or robotic assisted instruments [1]. The significant patient benefits have propelled the general acceptance of MIS for a range of surgical procedures [2]. Currently, the rapid recovery process of MIS has led to a shift of postoperative care from hospital to home environment. For instance, the traditional postoperative stay in hospital after major gastrointestinal surgery has been between 5 to 10 days [3, 4], whereas for patients undergo laparoscopic cholecystectomy they may be discharged as early as 48 hours after the operation [5-8]. Although the MIS approach can significantly reduce the chances of post-operative reflux, life-threatening complications may still occur days after the operation. As an example, an early discharged patients was diagnosed with ureterovaginal fistula several weeks after the laparoscopic hysterectomy [9]. To prevent these adverse events, the care of early discharged patients has been relied mainly on homecare services with regular pre-scheduled visits [10]. These are usually conducted by specialists or nurses, and patients are monitored with further telephone contacts and postoperative examinations [11]. The drawback of this approach is that it is difficult to accurately quantify the recovery process which is a crucial element of the postoperative care.

Thus far, a range of techniques have been proposed to quantify the recovery process. They range from paper/electronic questionnaires [12] based methods to the use of activity sensors worn by the patient [13, 14]. The acquired data is then analysed retrospectively. Current research has also identified the need for continuous, real-time monitoring for MIS post-operative care and a pilot study has been conducted by Aziz *et al* by the use of Body Sensor Networks (BSNs) [15]. Through the use of an e-AR (Ear-worn Activity Recognition) sensor, it has been demonstrated that the mobility of the patient can be used as an indicator of the general well-being and recovery rate of the patient post MIS operation.

However, to monitor patients continuously, the sensor data has to be transmitted to a local processing unit for data processing and trend analysis. The high power demand makes the miniaturisation of the device difficult, which can affect the general acceptance of the device by patients. The provision of processing on-node is therefore desirable both from the data reduction and user acceptance perspectives. In this way, the sensor data can be reduced at the node level to cut down radio transmission, and therefore the overall power consumption. This can also avoid the classification errors due to packet losses. The purpose of this paper is to provide a real-time BSN implementation framework that is suitable for post-surgical care. To enable real-time processing, relevant features are first identified using feature selection techniques, which is then followed by an efficient multivariate Bayes classifier for activity classification. Validation results based on lab-experiment are provided.

II. Real-time postoperative care

A. e-AR sensor

The e-AR sensor used for this study is based on the BSN platform that consists of a Texas Instrument MSP430 processor, a Chipcon CC2420 radio transceiver and an Atmel 512KB EEPROM [16, 17]. For capturing physiological parameters and general activities of the subject, a MCC ChipOX SpO_2 module and a 3-axis accelerometer (built by using two Analog Device ADXL202JE 2-axis accelerometers) are integrated into the device. Fig. 1 illustrates the overall design of the devices and how it is worn by the patient. Under this design, an ear-clip SpO_2 sensor is used to measure the oxygen saturation level and the heart rate of the patient. The accelerometer is embedded into the sensor for capturing the mobility and activity information of the user. Compared to other pervasive health monitoring devices, such as the ring sensor [18] and the EKG shirt [19], the e-AR sensor can capture a range of activities in addition to physiological indices. By positioning the sensor on the ear, the motion artefacts of the SpO_2 sensor can be reduced. This also avoids sensor placement and calibration problems encountered by other wearable systems [20, 21]. Although the e-AR sensor can provide both physiological parameters and activity information, this study will mainly focus on the use of motion information to infer the recovery progress of the patients.

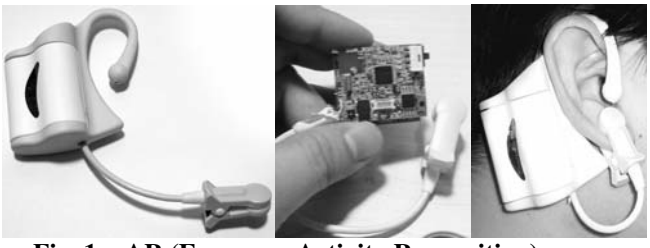


Fig. 1 e-AR (Ear-worn Activity Recognition) sensor

B. Feature selection

To quantify the mobility information and classify the activity of the subject, relevant features have to be identified. In general, selecting only relevant features can not only cut down the sensor channels but also improve the classification accuracy [22]. Most early studies for activity recognition are based on empirical feature selection techniques[23-26]. Recent studies have adopted more systematic feature selection techniques for enhancing the classification of activities [27-29]. Given the real-time constraint for feature extraction in this work, a number of features are identified:

a) Signal energy
b) Signal RMS (root mean square) value
c) Average distance between consecutive peaks
d) Troughs
e) Zero-crossing
f) Number of zero-crossing
g) Variance across the 3 axes

Other information that can be incorporated includes the orientation of the subject under static postures as captured by the inertial sensor.

The Bayesian Framework for Feature Selection (BFFS) was used to rank the relevance of features to different activity classes [30]. The result of the BFFS as applied to the e-AR sensor is shown on Fig. 2, where the expected AUC (Areas Under the ROC curve) is plotted against different features. It is apparent that the variance is the feature that provides the best discrimination among all activity classes. In addition, the variance provided a good measure on the level of activity ("activity index") of the subject.

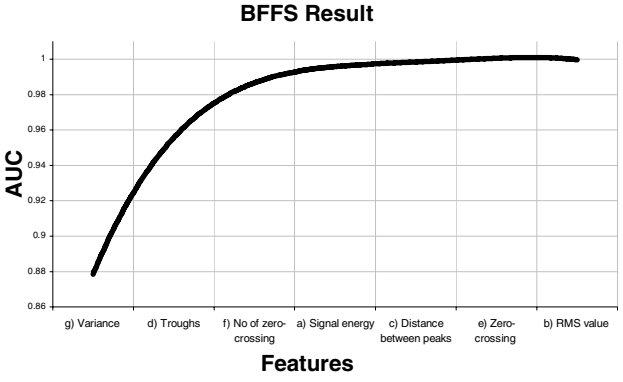


Fig. 2 BFFS results

C. Real-time classifier

Currently, most activity classification techniques are designed for retrospective analysis. For example, the hierarchical binary tree framework proposed by Mathie *et al.* for classifying daily activities requires extensive computation resources [31]. The HMM (Hidden Markov Model) approach for physical activity classification proposed by Lester *et al.* was designed to provide off-line classification of activities [32]. Although there is provision of an embedded version of the algorithm for porting to a sensor node, the feasibility of porting the computation intensive HMM onto a sensor node is difficult.

Several methods have been proposed for real-time classification on the sensor node. For instance, DeVaul and Dunn proposed a two layered model, which combines Gaussian mixture model with Markov models, for real-time motion

classification [33]. However, only the feature extraction and Gaussian class-conditional posterior code was implemented on the 200MIPS StrongArm processor of the MIThril system. Carlos *et al.* proposed a very low complexity algorithm for ambulatory activity classification [26]. A decision-tree based technique was proposed to classify different activities, such as resting, lying, walking, running and going up/downstairs. Although it has been shown that the proposed technique can be implemented on a fixed-point Philips LPC2106 30MHz microcontroller, such 32-bit ARM processor is a relatively high power processor compare to microcontrollers in typical wireless sensors. Recently, Karantonis *et al.* proposed an implementation of a scaled down version of the binary decision tree framework by Mathie *et al.* onto a TI MSP430 based sensor node [34]. However, to minimise the computation required, decisions are mainly based on thresholding the sensor readings, which can be sensitive to noise.

To enable real-time classification on a sensor node, a Multivariate Gaussian Bayes classifier is used for classifying different activities. The classifier is light-weighted in terms of computation but has an intrinsically efficient inferencing capabilities. As extensive memory is required for training the Bayes classifier, the training is conducted off-line where multivariate Gaussian density is used to model the different activity classes. Under this framework, the likelihood of multivariate feature vector x (variance across 3 axes) belong to class C_j is defined as:

$$P(x \mid C_j) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\Sigma_j|}} \exp\left[-\frac{1}{2}(x-\mu_j)^T \Sigma_j^{-1}(x-\mu_j)\right] \quad (1)$$

where n is the dimension of the feature vector, and μ_j and $\sum_j$ denote the mean and covariance of class C_j respectively. The mean μ_j and covariance $\sum_j$ are derived from the training data.

$$\mu_j = \frac{1}{N_j}\sum_{x \in C_j} x$$
$$\sum_j = \frac{1}{N_j}\sum_{x \in C_j} xx^T - \mu_j \mu_j^T \quad (2)$$

where N_j represents the number of features belong to class C_j.

According to the Bayes theorem, the posterior probability of activity class j given feature vector x is:

$$P(C_j \mid x) = \alpha P(x \mid C_j) P(C_j) \quad (3)$$

where α denotes the normalising constant and $P(C_j)$ represents the prior probably for class C_j. To minimise the bias towards more frequent activities, activity classes are assumed to have equal prior probabilities. The classification can be simplified as follows:

$$d_j(x) = -\frac{1}{2}\ln|\Sigma_j| - \frac{1}{2}\left[(x-\mu_j)^T \Sigma_j^{-1}(x-\mu_j)\right] \quad (4)$$

where x is belong to class C_j if $d_j(x) > d_i(x)$ for all $j \neq i$.

To enhance the accuracy of the classification, the signal amplitude is used to identify static activities (such as lying down or sleeping) and falls, and the classifier is used to determine other dynamic activities (such as running and walking).

III. Experiments

A. Lab-based experiment

A lab-based experiment was conducted to evaluate the proposed technique. The experiment involved 8 subjects who wore the e-AR sensor as they performed 5 different activities (reading, walking slowly, lying down, walking and running). To examine the performance of the classifier on the lab based data from the experiments, a moving window of size 4 seconds was used to extract the variance features over the three accelerometer signals. Fig. 3 shows the variances of the 3D accelerometer from the experiment. As an example, only three activities are shown on Fig. 3, which are reading, walking and running (physical activity). As shown in this graph, the variance data of different activities are scattered apart from each other which depicts the relevance of the variance feature.

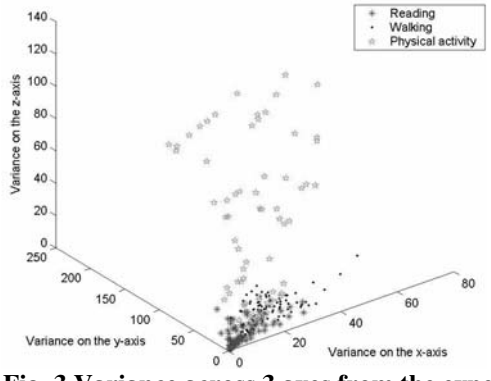


Fig. 3 Variance across 3 axes from the experiment

For activity classification, the activities are first grouped into four levels of activities, and the classifier is then applied to discriminate different activity levels. Table 1 shows

the accuracy of the classifier in labelling each activity against the corresponding activity level from the experiment.

Table 1 Activity classification results from the lab-based experiment

Subjects \ Activities	1	2	3	4	5	6	7	8
Reading	87.5	93.8	80.0	84.6	100.0	91.7	80.0	100.0
Walking slowly	100.0	91.7	100.0	80.0	93.3	100.0	100.0	90.9
Lying down	80.0	91.7	62.5	100.0	0	100.0	33.3	87.5
Walking	100.0	95.0	77.8	100.0	80.0	72.7	100.0	66.7
Running	100.0	93.8	100.0	100.0	83.3	100.0	84.6	100.0

It can be seen from Table 1 that the accuracy of the classifier is high in general, with an average accuracy of 87%. Most of the activities can be classified with high accuracy except for lying down. It is evident that by positioning the sensor on the ear, the sensor calibration and subject dependency problems are naturally eliminated, and the classifier can identify activities of different subjects accurately.

B. Patient trial

To validate the feasibility of using the e-AR sensor for postoperative care, a pilot study was conducted to evaluate the sensor and the proposed analysis technique. Five patients who were due to undergo various elective abdominal surgical procedures were approached and consented to participate in the trial. Before their operations, patient were shown how to wear the earpiece and a laptop computer with the wireless receiver was set up in the patient's home to collect and store the information captured.

A week before the operation, the patients were asked to wear the sensor for a day, and which will be used as a 'baseline' of their normal activity, where we could gauge their normal mobility level, daily activities, and body physiology. The patients were subsequently admitted for MIS operation and then discharged as per the hospital protocol. From this point onwards, patients were asked to wear the sensor for five days post-operatively. The only time they did not wear the sensor was when they had a shower or when they were sleeping at night. To minimise the disruption to the recovery of the patients, only routine telephone contacts were made by a clinician to check the status of the patients.

As an example, data for an 86-year old male who took part in the pilot study is presented. He had a large right inguinal hernia which was repaired with a mesh using an open approach. He required an overnight stay but was sent home promptly the next day. Fig. 4 shows the averaged activity index (every 4 minutes) for this subject for the pre-operative day and five post-operative days between the hours of 3pm and 7pm each day. This activity index was calculated using the techniques developed in the lab-based study as previously described (the variance across the 3 axes). The figure shows some interesting trends in this patient's activity over the 6 days he wore the sensor. Preoperatively, there are clear high activity periods which do not occur during the first few days following surgery. This is most likely due to impaired mobility because of a healing groin wound and some scrotal oedema. What is also interesting is the trend showing recovery of the activity index towards pre-operative levels during the five days after surgery. These results suggest that recovery in terms of mobility can potentially be quantified and objectively measured, perhaps into a 'recovery curve' for each individual patient. The technology would therefore be useful not only in comparing different surgical interventions (laparoscopic versus open surgery) and their effects on the recovery curve, but also to monitor a patient's recovery in real-time and pick up poor recovery early. The latter would be a particularly useful adjunct to goal-directed recovery and home district nursing.

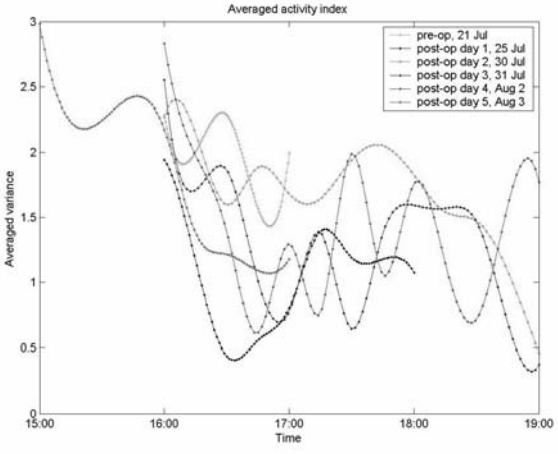


Fig. 4 The averaged activity index (every 12 minutes) of the patient for the pre-operative day and five post-operative days between the hours of 3pm and 7pm each day

IV. Conclusions

Advances in MIS have significantly shortened the hospital stay that permits early discharge of patients for home based postoperative care. The routine examinations used in the current homecare services involve extensive human interaction and cannot capture reliable transient events that may lead to critical events. A real-time analysis technique based on the e-AR sensor is proposed to objectively quan-

tify the recovery curve of the patient in a non-intrusive manner. By relying on processing on-node, the amount of radio transmission, and therefore the overall power consumption can be reduced significantly. This facilitates further minimisation of the sensors to permit pervasive monitoring. To enable real-time processing on a miniaturised sensor node, the proposed technique processes only the relevant features, and it identifies different activities by using an efficient multivariate Bayes classifier. Although the proposed technique is relatively light-weighted, its accuracy has been demonstrated from our lab-experiment. In addition, the feasibility of using the activity index for post-operative care has demonstrated the potential clinical value of the technique.

Acknowledgment

This work is carried out under the project SAPHE (Smart and Aware Pervasive Healthcare Environment) which is supported by the Department of Trade and Industry (DTI) UK.

References

1. Patti, M.G., et al. (1999) Minimally Invasive Surgery for Achalasia: An 8-Year Experience With 168 Patients. Annals of Surgery. 230(4): p. 587.
2. Cohn, L.H., et al. (1997) Minimally Invasive Cardiac Valve Surgery Improves Patient Satisfaction While Reducing Costs of Cardiac Valve Replacement and Repair. Annals of Surgery. 226(4): p. 421-428.
3. Delaney, C.P., et al. (2001) `Fast track' postoperative management protocol for patients with high co-morbidity undergoing complex abdominal and pelvic colorectal surgery. British Journal of Surgery. 88: p. 1533-1538.
4. Bokey, E.L., et al. (1995) Postoperative morbidity and mortality following resection of the colon and rectum for cancer. Diseases of the Colon & Rectum. 38(5): p. 480-487.
5. McClane, S., A. Senagore, and P. Marcello (2000) Experience-based postoperative care in laparoscopic-assisted colectomy reduces length of stay. Diseases of the Colon & Rectum. 43: p. A54.
6. Bardram, L., et al. (1995) Recovery after laparoscopic colonic surgery with epidural analgesia, and early oral nutrition and mobilisation. Lancet. 345(8952): p. 763-764.
7. Binderow, S.R., et al. (1994) Must early postoperative oral intake be limited to laparoscopy? Diseases of the Colon & Rectum. 37(6): p. 584-589.
8. Reissman, P., et al. (1995) Is early oral feeding safe after elective colorectal surgery? A prospective randomized trial. Annals of Surgery. 222(1): p. 73-77.
9. Boschert, S. (2004) Early discharge OK in laparoscopic hysterectomy: avoiding an overnight stay - gynecology. Ob. Gyn News: p. 8.
10. Booth, J.E., et al. (2004) A trial of early discharge with homecare compared to conventional hospital care for patients undergoing coronary artery bypass grafting. Heart. 90: p. 1344-1345.
11. Engelman, R., et al. (1994) Fast-track recovery of the coronary bypass patient. The Annals of Thoracic Surgery. 58(6): p. 1742-1746.
12. Myles, P., et al. (2000) Validity and reliability of a postoperative quality of recovery score: the QoR-40. British Journal of Anaesthesia. 84(1): p. 11-15.
13. Inoue, Y., et al. (2003) Is laparoscopic colorectal surgery less invasive than classical open surgery? Surgical Endoscopy. 17(8): p. 1269-1273.
14. Inoue, Y., et al. (2003) A New Parameter for Assessing Postoperative Recovery of Physical Activity Using an Accelerometer. Surgery Today 2003. 33: p. 645-650.
15. Aziz, O., et al. (2006). Pervasive Body Sensor Network: An Approach to Monitoring the Post-operative Surgical Patient In The International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2006). MIT, Boston, USA.
16. Yang, G.-Z., et al. (2004). From Sensor Networks to Behaviour Profiling: A Homecare Perspective of Intelligent Building. In The IEE Seminar for Intelligent Buildings.
17. Lo, B. and G.-Z. Yang. (2005). Architecture for Body Sensor Networks. In The Perspective in Pervasive Computing. IEE Savoy Place: pp. 23-28.
18. Rhee, S., et al. (1998). The Ring Sensor : a New Ambulatory Wearable Sensor for Twenty-Four Hour Patient Monitoring. In The 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Hong Kong.
19. Linz, T., et al. (2005). Fully untegrated EKG shirt based on embroidered electrical interconnections with conductive yarn and miniaturized flexible electronics. In International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2006). MIT Boston USA.
20. Venkatasubramanian, K., et al. (2005). Ayushman: A Wireless Sensor Network Based Health Monitoring Infrastructure and Testbed. In IEEE International Con-

ference on Distributed Computing in Sensor Systems (DCOSS). Marina del Rey, CA, USA.
21. Amft, O., H. Junker, and G. Troster. (2005). Detection of eating and drinking arm gestures using inertial body-worn sensors. In Ninth IEEE International Symposium on Wearable Computers: pp. 160-163.
22. S.Thiemjarus, et al. (2004). Feature Selection for Wireless Sensor Networks. In the 1st International Workshop on Wearable and Implantable Body Sensor Networks. Imperial, London, UK.
23. Minnen, D., et al. (2005). Recognizing and Discovering Human Actions from On-Body Sensor Data. In IEEE International Conference on Multimedia and Expo, 2005 (ICME 2005): pp. 1545-1548.
24. Bao, L. and S.S. Intille. (2004). Activity Recognition from User-Annotated Acceleration Data. In The 2nd International Conference on Pervasive Computing PERVASIVE 2004: pp. 1-17.
25. Ravi, N., et al. (2005). Activity Recognition from Accelerometer Data. In The Twentieth National Conference on Artificial Intelligence and the Seventeenth Innovative Applications of Artificial Intelligence Conference. Pittsburgh, Pennsylvania, USA: pp. 1541-1546.
26. Caros, J.S.i., et al. (2005). Very Low Complexity Algorithm for Ambulatory Activity Classification. In 3rd European Medical and Biological Conference EMBEC 2005.
27. Parkka, J., et al. (2006) Activity Classification Using Realistic Data From Wearable Sensors. IEEE Transactions on Information Technology in Biomedicine. 10(1): p. 119-128.
28. Lester, J., et al. (2005). A Hybrid Discriminative/Generative Approach for Modeling Human Activities. In The Nineteenth International Joint Conference on Artificial Intelligence. Edinburgh, Scotland.
29. Nurmi, P. and P. Floreen. (2005). Online feature selection for contextual time series data. In Subspace, Latent Structure and Feature Selection techniques: Statistical and Optimisation perspectives Workshop (SLSFS05).
30. Thiemjarus, S., et al. (2004). Feature Selection for Wireless Sensor Networks. In the 1st International Workshop on Wearable and Implantable Body Sensor Networks. Imperial, London, UK.
31. Mathie, M.J., et al. (2004) Classification of basic daily movements using a triaxial accelerometer. Biomedical and Life Sciences, Engineering and Medicine. 42(5): p. 679-687.
32. Lester, J., T. Choudhury, and G. Borriello. (2006). A Practical Approach to Recognizing Physical Activities. In PERVASIVE 2006. Dublin Ireland: pp. 1-16.
33. DeVaul, R.W. and S. Dunn (2001) Real-Time Motion Classification for Wearable Computing Applications.MIT Media Laboratory
34. Karantonis, D.M., et al. (2006) Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring. IEEE Trans on Information Technology in Biomedicine. 10(1): p. 156-167.

Address of the corresponding author:

Author: Benny Lo
Institute: Dept of Computing, Imperial College London
Street: Rm 307 Huxley Building, South Kensington campus
City: London
Country: United Kingdom
Email: benny.lo@imperial.ac.uk

Embedded Real-Time Heart Variability Analysis

Eleftheria Katsiri*, Mel Ho†,Lei Wang, Benny Lo* and Chris Toumazou†

*Department of Computing
Imperial College London
180, Queens Gate
SW7 2AZ
London UK
Email: {ek,leiwang,benlo}@doc.ic.ac.uk

†Institute of Biomedical Engineering
Imperial College London
180, Queens Gate
SW7 2AZ
London UK
Email: m.ho@imperial.ac.uk

***Abstract*— Heart Variability Analysis (HRV) is not suitable for real-time processing on a resource-limited, single sensor network node, such as a Body Sensor Network (BSN) node, due to the high sampling rate ($> 200Hz$) required to digitise ECG signals and the non-preemtable nature of operating systems such as tinyOS. Both reasons combined dictate that the processing of each sample needs to be completed withing the inter-sample period, typically 5 msec for ECG signals. This paper discusses a dual-layer real-time heart variability analysis algorithm. The top layer is invoked every time a sample arrives. This layer includes a real-time algorithm that delineates the significant part of the ECG signal, the QRS complex. The second layer, is near real-time and is invoked only when a potential QRS is detected, at a significantly lower rate that corresponds to the person heart rate. This layer is responsible for detecting R peaks, estimating the interval between two successive peaks and performs heart rate variability analysis in the frequency domain.**

Our system outperforms traditional ECG processing algorithms because the top layer completes well within the 5 msec sample inter-arrival period, ensuring that no samples are lost. The bottom layer can be delegated either to an underlying background task or a second processor. Because it is invoked less frequently than the top layer, it results in a lower interrupt rate, allowing for more flexible processing.

I. Motivation

The *BiosensorNET* [5] project assembles a multi-disciplinary team with the overall objective to lay the foundations for a new generation of intelligent, self-managing, context-aware biosensor networks for critical control of human health. In this scope, we are a team of researchers looking more specifically at how Body Sensor Networks [15] could be used as a programmable platform for modelling biomedical applications. Body Sensor Networks can be defined as networks of intelligent nodes. Each node is effectively a low-power assembly of h/w components comprising of a sensor array, a micro controller unit (MCU) and a radio interface, connected through a common bus. The current BSN nodes were designated to be low-power, miniature and cheap, but with limited computation and storage abilities. However, each node can be programmed to behave in an autonomic manner [6] and is not dedicated to any application.

The current BSN node technology is based on TinyOS. TinyOS is a single threaded operating system where all tasks run to completion without being pre-emptable. This means that the processing of a sample needs to be completed before the next sample arrives, else the next sample is lost. Buffers are not sufficient as the traffic is constantly very high-rate, not bursty.

On the other hand, ECG signals are usually digitised at rather high sampling rates, typically 200Hz and higher. This means that each sample needs to be processed, locally, on the node, within a 5 msec window or less. Else, input samples are lost. The alternative solution of transmitting every sample to a host PC for processing, is far from energy-efficient and results in signal loss, as the rate of packet assembly and transmission over the wireless interface cannot keep up with the incoming sample rate.

Real-time ECG processing is very important as several health-care applications rely on it. These include, real-time HRV, ventricular de-fibrillation and arrhythmia detection. In the wearable case, a robust ECG algorithm is particularly important as the ECG signals rarely resemble the perfect ECG waveforms that are the output from patient monitoring in a hospital. This is partly due to the fact that in the latter case, the patient is lying completely still and a 12-electrode ECG monitoring device is used instead of the 3-electrode one that is available in wearable computing. Figure 1 shows the most important functional processing entities that are involved in these applications. Although a limited number of solutions are mentioned in the literature, these are specialised to specific applications and their performance relies solely on using specialised hardware. This approach is out of the scope of our research interest. Instead we are interested in investigating

programmable, software solutions that run on inexpensive, off-the-self systems such as BSN nodes. Thus motivated, this paper discusses a dual-layer system for programmable, software real-time ECG processing and HRV analysis.

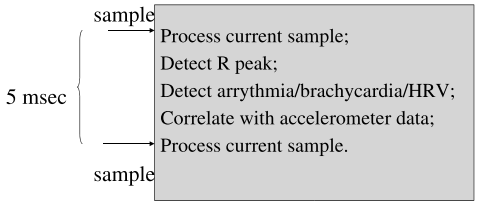


Fig. 1. Single layer architecture.

The performance of the system is based on the fact that the top layer is invoked every time a sample arrives, whether this is part of the QRS or the baseline. This layer detects and delineates the *important* part of the ECG signal, the QRS. The second layer is invoked only when a QRS has been detected, therefore at a significantly lower rate, for example, approximately twice per second for a heart rate of 120BPM. The QRS is used by the second layer for performing three additional tasks:

- detect R peaks
- estimate inter-peak intervals and
- perform HRV analysis

This separation of the computational processing into two layers has the following advantages. As the top layer is invoked with every sample, it is implemented using a well-known, popular peak detection algorithm that completes well within the 5 msec interval. There are two options for the second layer. It is delegated either to an underlying background process or a second BSN node. As it is invoked at a significantly lower rate, in the first case, the background task can utilise the idle periods of the processor for processing the QRS complex. In the second case, the second processor is interrupted much less frequently allowing for a longer processing time between successive QRSs. Last, it must be noted that the efficiency of the R peak detection algorithm that is implemented in the bottom layer depends on the amount of information that is communicated to the bottom layer by the top one, i.e. the size of the QRS section. However, we believe that our solution implements a good compromise.

II. Prior Art

Software QRS detections has been the subject of [9], [7], [2], [8]. A comparison of the most important algorithms is done in [9]. Note that our solution is very similar to the real-time algorithms discussed there (group a); Group a type of algorithms rated very high in terms of its real-time properties. One of the reasons why software QRS detection has been studied so widely is because it is integral in almost all applications in the area of cardiology and in particular monitoring and diagnostics. The works of [12], [13], [3], [2] deal with arrhythmia detection, fibrillation and real-time HRV analysis. They all depend on correct QRS detection. The work of [14] discusses a special purpose, wearable ECG processing device that is available commercially. This work does not deal with HRV analysis and the QRS detection is based on matched filters implemented in specialised hardware. This solution is application-specific, is almost certainly not inexpensive, and it does not constitute a flexible programmable platform such as the BSN node. Wearable ECG processing has been the focus of [1], [4] but those solutions typically delegate the processing to a host PC.

III. Architecture

The architecture of the system is shown in Figure 2.

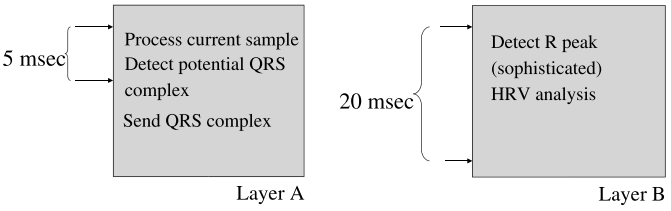


Fig. 2. Dual layer architecture.

A. Top Layer

As mentioned before, the top layer is responsible for signal conditioning and detecting potential QRS segments. It is invoked every time a sample is received. The conditioning of the signal is undertaken by a band-pass filter. Next, potential peaks are detected by means of calculating the *slope* of the ECG signal. The current sample and the previous instance of the sample are needed. The slope is equivalent to calculating the derivative dy/dt. A potential peak is characterised by a point where the value of the slope is zero. The algorithm, derived from the works in [11], [10], is illustrated in the top two graphs in Figure 3.

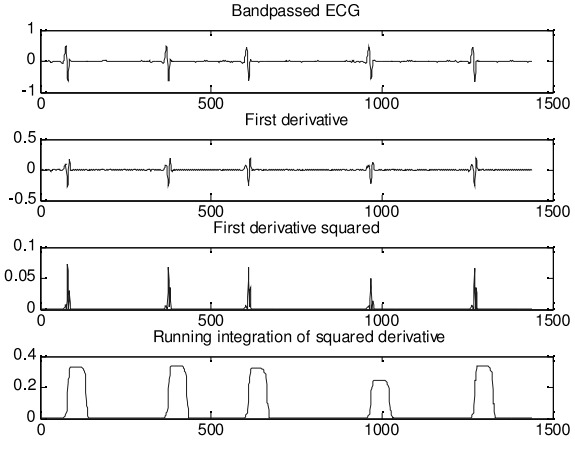


Fig. 3. Real-time ECG processing.

The value of the derivative is then squared in order to eliminate the negative parts of the signal and to produce a clearly detectable peak. Both the current R peak value and the previous R peak value are stored. This is shown in the third graph from the top, in Figure 3. Next, the algorithm performs a running integration of the squared derivative using the previous value of the R peak to calculate a threshold. This is shown in the bottom graph of Figure 3. By storing a fixed-length history of previous signal samples, the QRS complex can be identified as the segment of the original signal that falls between two successive intersections of the integrated waveform with the adaptive threshold; the first intersection indicates a threshold that has been exceeded and the second, the level where the integration waveform falls below the adaptive threshold. The delineated QRS is then saved in a buffer ready to be transmitted to the bottom layer. This is shown in Figure 4.

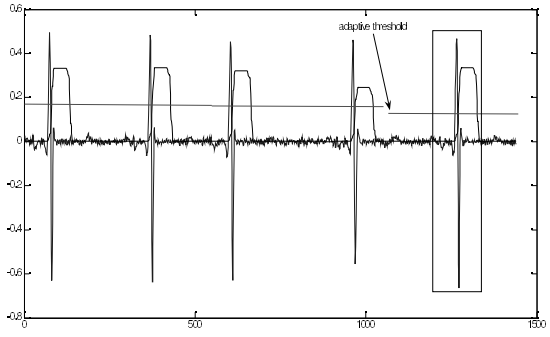


Fig. 4. Real-time QRS delineation.

B. Bottom Layer

Given a potential QRS complex , the bottom layer performs a more *sophisticated R-peak detection*, it *estimates the inter-peak intervals* and performs *spectrum-based HRV analysis*. A block diagram illustrates these tasks in Figure 5.

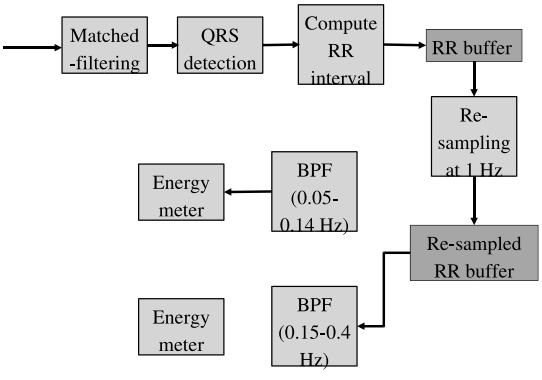


Fig. 5. HRV analysis.

IV. Implementation

The top layer of the system is implemented in a routine called *bsnqrs1* which is embedded in the interrupt handler for incoming samples. The bottom layer is implemented in the routines *bsnrr1* and *analyze_rr* that are allocated to a background task. We decided not to implement the bandpass filtering as the input that we receive from the patient simulator (see Section IV-A is already conditioned. An outline of the program illustrating the control and data flow aspects of the program is shown below in *nesC* notation.

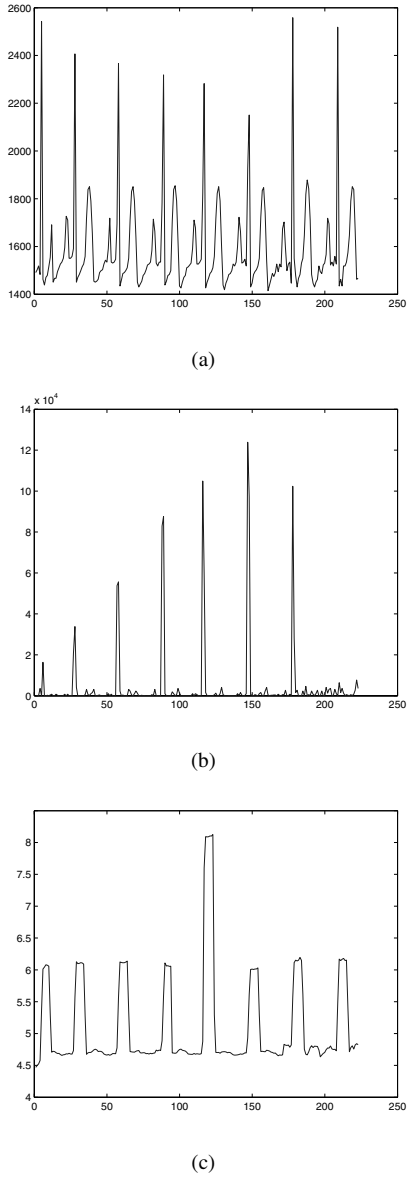
```
async event result_t ECG.dataReady
                        (uint16_t data){
  dbg(DBG_USR1,"dataReady!\n");
  atomic{
         sendToAnalyser(data1);
        }
  return SUCCESS;
 }

void  sendToAnalyser(uint16_t data1){
bsnqrs1(data1);
post estimateRRInts();
}

void bsnqrs1(uint16_t  ecg1)
{   //top layer implementation
    //return  QRS complex
}

task estimateRRInts(){

// bottom layer implementation

uint16_t rr;
uint16_t rr_accum=0;

dbg(DBG_USR1, "estimateRRints!\n");

if (ready_to_send==0){
dbg(DBG_USR1, "ready_to_send\n");
rr=bsnrr1( time_last_qrs, qrs_to_send,
                BLKLEN_QRS_TO_SEND);
buf_rr[ptr_rr++]=rr;
rr_accum+=rr;

if ( (rr_accum>(RR_WIN_SECONDS*FS))
              || (ptr_rr==BLKLEN_RR) )
   {
      analyze_rr(ptr_rr);
      ptr_rr=0;
      rr_accum=0;
   }
  time_last_qrs=time_last_qrs+rr;
  ready_to_send=0;
 }
}
```

A. Description of equipment

We implemented our system on a BSN node [15] to which we attached an ecg-sensor. We used a patient simulator to generate a test ECG signal. The sampling rate was set to 200Hz.

V. RESULTS

We tested our system with the patient simulator, set at 80BPM and we plotted the output of the ecg buffer (raw signal), the buffer that contains the first derivative squared and the buffer that contains the integrated values. Although all the processing takes place localaly on the node, strictly for the purpose of visualising the output and checking the correct behaviour of our algorithm, we transmitted the computed results to a host PC. For this purpose we used the TOS_Msg packet structure, modified to host the data format that was appropriate for each buffer type (uint32_t and uint16_t). The results from two indicative test runs are shown in Figures 6 and 7. We also tested the performance of the top-layer which falls within the 5 msec window.

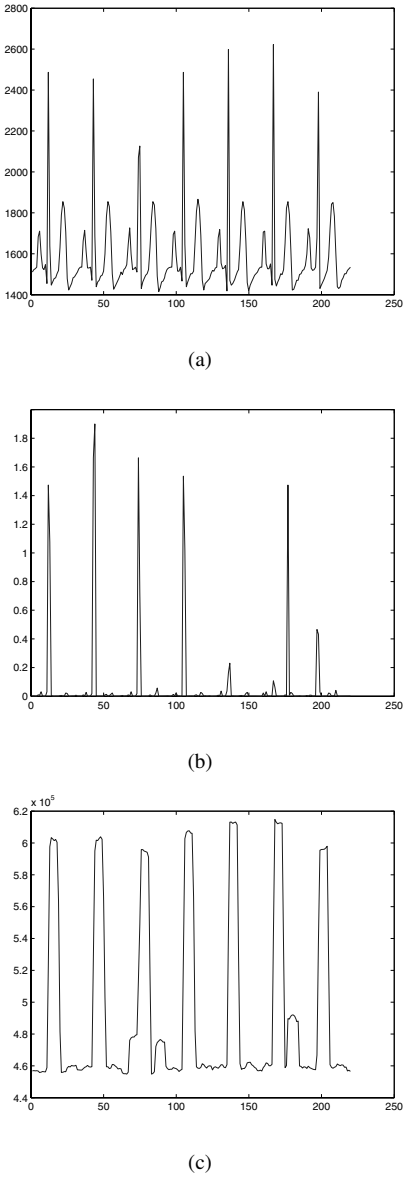


(a)

(b)

(c)

Fig. 6. ECG sample (80BPM-200Hz) (a), first derivative squared (b), integrated signal (c).



(a)

(b)

(c)

Fig. 7. ECG sample (80BPM-200Hz) (a), first derivative squared (b), integrated signal (c).

VI. A PROPOSED MULTI-LAYER ARCHITECTURE

In the case of the HRV analysis, in the near-real-time version of this algorithm, a collection of 2 minutes of R-R intervals is assembled before the analysis can take place. This means that following the same approach as before, the HRV analysis can be seen as a separate layer, which will be interrupted only when the historical information is ready, i.e. every 2 minutes. This is shown in Figure 8. It can be easily seen that this architecture can be extended to any number of layers according to the real-time properties of each specific application.

We are currently profiling the various functional entities in our system, in order to determine their individual performance.

This requires developping our own libraries in order to use efficiently the available timing infrastructure of the msp430 processor. We are also looking at hardware solutions for more accurate profiling.

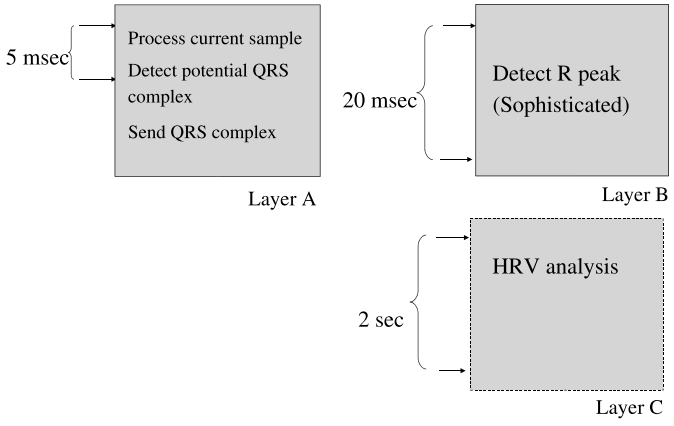


Fig. 8. A multi-layer architecture.

VII. Conclusions and future work

Our motivation has been to use BSN nodes as a flexible programming infrastructure for next-generation healthcare. In this view, we investigated the feasibility of embedding a real-time critical application such as HRV analysis into a BSN network.

In this scope, we designed implemented a dual-layered system for real-time or near real-time HRV analysis. The top layer consists of a fast, real-time software QRS detection algorithm that completes well within 5 msec. The bottom layer is invoked only when a QRS complex has been detected, at a significantly lower rate than the top layer, which is invoked each time a sample is received. The bottom layer is responsible for sophisticated QRS detection and HRV analysis.

In our current implementation, the bottom layer is implemented as a background task that can be pre-empted by the incoming samples if necessary. However, we envision that the bottom layer can be implemented by a second processor, that due to the lower interruption rate will allow for more flexible processing. Section VI discusses an extension to our architecture in multiple layers, separated by their real-time properties, i.e. their interruption rate. We envision that each layer could be implemented by a separate processor resulting into a multi-core implementation.

VIII. Acknowledgements

The authors would like to thank Morris Sloman, Guang Zhong Young and Chris Toumazou for discussion and guidance.

References

[1] J. Dong, M. Xu, H. Zhu, and W. Lu. Wearable ECG Recognition and Monitor. In *the 18th IEEE symposium on Computer-Based Medical Systems (CBMS'05)*, 2005.

[2] L. Faes, G. Nollo, R. Antolini, F. Gaita, and F. Ravelli. A Method for Quantifying Atrial Fibrillation Organization Based on Wave-Morphology Similarity. *IEEE Transactions on Biomedical Engineering*, 49(12):1504–1513, 2002.

[3] J. Felblinger and E. Yvroud. Real Time Detection of Ventricular Fibrillation and its Use in a Semi-Automatic Defibrillator. *Engineering in Medicine and Biology Society Vol.14. Proceedings of the Annual International Conference of the IEEE*, 6, Nov. 1992.

[4] R. Fensli, E. Gunnarson, and T. Gundersen. A wearable ecg-recording system for continuous arrythmia monitoring in a wireless tele-home-care situation. In *the 18th IEEE symposium on Computer-Based Medical Systems (CBMS'05)*, Dublin, 2005.

[5] EPSRC BiosensorNet: Autonomic Biosensor Networks for Pervasive Healthcare. http://ubimon.doc.ic.ac.uk/biosensornet/index.php?m=432.

[6] S.L. Keoh, K. Twidle, N. Pryce, A.E. Shaeffer-Filho, E. Lupu, N. Dulay, M.Sloman, S. Heeps, S. Strowes, J. Sventek, and E. Katsiri. Policy-based management for body sensor networks. In *submitted to BSN 2007*, 2006.

[7] B-U Kohler, C. Hennig, and R. Orglmeister. The principles of software qrs detection. *IEEE Engineering in Medicine and Biology*, pages 42–57, January/February 2002.

[8] J. P. Martinez and S. Olmos R. Almeida. A Wavelet-Based ECG Delineator: Evaluation on Standard Databases. *IEEE Transactions on Biomedical Engineering*, 51(4):570–581, March 2004.

[9] V.A Nagin and S.V. Selishchev. Implementation of Algorithms for Identification of QRS-Complexes in Real-Time ecg Systems. *Biomedical Engineering*, 2001.

[10] J. Pan and W. Tompkins. A real-time detection algorithm. *IEEE Transactions of Biomedical Engineering*, 1993.

[11] A. Ruha, S. Sallinen, and S. Nissila. A real-time microprocessor qrs detection system with 1-ms timing accuracy for the measurement of ambulatory hrv. *IEEE Transactions of Biomedical Engineering*, 44, 1997.

[12] S. R. Seydnejad and R. Kitney. Real-Time Heart Rate Variability Extraction Using the Kaiser Window. *IEEE Transactions on Biomedical Engineering*, 44(10):990–1005, 1997.

[13] M. P. Tulippo, T. H. Makikallio, T. E. S. Takala, T. Sepannen, and H. V. Huikuri. Quantitative beat-to-beat analysis of heart rate dynamics during exercise. In *Computers in Cardiology*. The American Physiological Society, Sep. 1997.

[14] J. Welch, F. Guilak, and S. D. Baker. A wireless ECG Smart Sensor for Broad Aapplication in Life Threatening Event Detection. In *the 26th Annual International Conference of the IEEE EMBS*, 2001.

[15] G.Z. Yang, editor. *Body Sensor Networks*. Springer, 2005.

Behaviour Profiling with Ambient and Wearable Sensing

L. Atallah, M. ElHelw, J. Pansiot, D. Stoyanov, L. Wang, B. Lo, and G.Z. Yang

Department of Computing, Imperial College, London, United Kingdom

Abstract—**This paper investigates the combined use of ambient and wearable sensing for inferring changes in patient behaviour patterns. It has been demonstrated that with the use of wearable and blob based ambient sensors, it is possible to develop an effective visualization framework allowing the observation of daily activities in a homecare environment. An effective behaviour modelling method based on Hidden Markov Models (HMMs) has been proposed for highlighting changes in activity patterns. This allows for the representation of sequences in a similarity space that can be used for clustering or data-exploration.**

Keywords—**body sensor networks, similarity based clustering, blob sensors, behaviour profiling.**

I. Introduction

Over the next fifty years, the proportion of people aged beyond 60 is expected to more than double. With demographic changes associated with the aging population and the increasing number of people living alone, the social and economic structure of our society is changing rapidly. In almost all countries, longevity has given rise to expensive age-related disabilities and diseases. With the steady decline of the ratio of workers to retirees, a fundamental change of the way that we care for the aging population is required.

Recent advances in the semiconductor industry have led to miniaturisation and cost reduction of both sensor and computing technologies, thus making truly pervasive monitoring of patients with chronic disease and those living alone a reality. For the elderly, home-based healthcare encourages the maintenance of physical fitness, social activity and cognitive engagement to function independently in their own homes. For care professionals, it also provides a more accurate measure of how well the elderly person is managing with his/her daily activities, thus better targeting limited human carer resources to those who need the most.

Research in pervasive healthcare has recently transcended from the traditional telecare systems and several groups have investigated 'smart homes' that can provide intelligent pervasive home monitoring. Examples include MIT's PlaceLab [1], BT's work on smart homes [2] and the Welfare Techno house in Japan [3]. Most of these projects employ a large number of sensors that are ubiquitously placed around the house. These sensors include temperature sensors, water flow and utility usage sensors, pressure sensors on furniture, as well as vital signs monitoring devices. These sensors provide data that can be analysed to observe patient behaviour or detect the occurrence of critical events such as falls.

One of the limitations of ambient sensing based on simple sensors is that it is difficult to infer detailed changes in activity and physiological changes related to the progression of disease. In fact, even for the detection of simple activities such as leaving and returning home, the analysis steps involved can be complex even by the explicit use of certain constraints. It is well known that subtle changes in behaviour of the elderly or patients with chronic disorders can provide telltale signs of the onset or progression of the disease. For example, research has shown that changes in gait can be associated with early signs of neurologic abnormalities linked to several types of non-Alzheimer's dementias. Subjects with neurologic gait abnormalities had a greater risk of developing dementia.

Another challenge associated with the current ambient sensing framework is the availability of the training data. In many scenarios, the requirement for the patient to perform specific activities in order to obtain a 'labelled training' set is not realistic. The large number of sensors involved also makes a wide-spread practical deployment difficult. The purpose of this work is twofold. First, we aim to provide a simple hardware architecture with an integrated use of e-AR (ear-worn activity recognition) sensor and blob based ambient sensors. We demonstrate that by the use of these two types of sensors, it is possible to provide rich information that can be used for analysing most types of daily activities. Second, we propose an effective visualisation framework and behaviour modelling method based on Hidden Markov Models (HMMs) for highlighting changes in activity patterns without the use of explicit labelling of data into activity categories. These two unique features of the system make it easily deployable to a range of homecare settings.

II. Combining ambient and wearable sensing

Ambient sensing refers to the use of environment sensors for the monitoring of daily activities. In this work, blob sensors based on the concept of using abstracted image blobs to derive personal metrics and perform behaviour

profiling are used [4]. With these sensors, the captured image is immediately turned into blobs that encapsulate shape outline and motion vectors of the body at the device level. No appearance data is stored or transmitted at any stage of the processing and it is not possible to reconstruct this abstracted information back to images. This ensures the privacy of the patients and also makes the device usable to all areas of a home environment. The shape of a blob (or outline) detected by the sensor depends on the relative position of the subject and the sensor. A view-independent model can be generated by fusing a set of blobs captured by respective sensors at different known positions, which can be used to generate a more detailed activity signature. Figure 1 shows several example outputs of the blob sensor. The main information derived from the blob sensors in this work is room occupancy and the number of people in the room. More elaborated analysis of blobs for posture and gait recognition is performed in a separate study [4]. The focus of this paper is directed towards overall activity patterns within the house. For N rooms, at time t, the blob sensors provide a vector of room occupancy $L_t(N)$ of size N.

Fig. 1 Outputs of the blob sensor, showing 3 different activities.

In this work, we use a combined wearable-ambient sensor framework to determine the location of the person being monitored. To allow the identification of the patient(s) being monitored in a multi-dwelling environment, the patient wears an e-AR sensor (ear worn activity recognition sensor). The e-AR is presented in more details in [5], and is based on the BSN platform that consists of a Texas Instrument MSP430 processor, a Chipcon CC2420 radio transceiver and an Atmel 512KB EEPROM [6, 7]. The e-AR sensor contains a 3-axis accelerometer and a MCC ChipOX SpO_2 that can be used to monitor the change of the physiological parameters of the patient (as investigated in [8]). However, its primary function in this work is to locate the patient being observed. Thus, at each time stamp, the signal strength transmitted from the wearable sensor to the receivers (typically co-located with the blob sensor) is recorded. For M receivers, the wearable sensors provide a vector $X_t(M)$ of signal strength of size M.

III. A Bayesian Classifier to Identify Patient Location

A Bayesian classifier is used to model the relationships between the signal strengths and the room occupancies. For each vector of signal strengths X_t at a time t, the classifier provides a likelihood of room occupancy. Training data is obtained by recording both occupancies and signal strengths of a person while being alone in the house. In this case, a simple Naïve Bayes classifier was used due to its simplicity, but Gaussian Mixture Models could also be used to model these relationships. Thus at each time stamp, the classifier provides a probability of being in a certain room (an occupancy vector), given a vector of signal strengths:

$$P\big(L(j)/X_t\big) = \frac{\prod_{i=1}^{N} P\big(X_t(i)/L(j)\big)\,P\big(L(j)\big)}{L(j)} \qquad (1)$$

The classifier is used whenever there is an ambiguity over the number of people in a certain room. However, it is not used if the ambient sensors determine that there is only one person in the house and there is no uncertainty in the number of occupants. The training data-set (of signal strengths) was obtained over 10 minutes from a single person moving in a lab-based environment. The training labels for the classifier are the values of the occupancy vector $L_t(N)$. Table 1 shows the 5-fold cross validation averaged results of this classifier on the training dataset.

Table 1 Results of correct classification using the Bayesian classifier

Classifier results\Room number	**1**	**2**	**3**	**4**	**5**
Success rates (percent)	70	94	100	66	79

IV. A Realtime Activity Grid

Given the data from both the blob and e-AR sensors, the Bayesian classifier can be used at each time step to enhance the certainty of the location of the patient. This can be used by the observers as a real-time activity grid. Figure 2 illustrates an example of two patients (colour-coded green and red) who are being observed moving around a house that has five rooms. A combined use of wearable and ambient sensors is employed to determine the position of the patient being observed. This framework is scalable for monitoring multiple patients cohabiting with normal individuals. In this case, the identities of the normal individuals are recognised

and they are simply shown as grey bullets on the activity grid. This activity grid provides a novel framework that is easy to use by healthcare professionals without intruding on the privacy of the patients because of the relative abstractness of the display used. Several behaviour patterns can be observed, including the patient's interaction with other people, and the general habits of the patient. Thus, the system can also be used for examining social interaction of the patient in terms of the frequency and pattern of receiving visitors.

It is worth noting that comparing activity grids over different periods can become a cumbersome task for an observer. In this work, a method for representing the behaviour pattern of the patient and its temporal variation is developed based on an HMM framework.

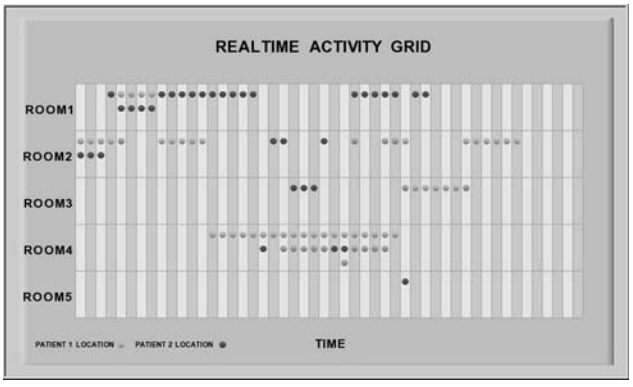


Fig. 2 A real time activity grid showing 2 people being observed as they move between 5 rooms. They are shown as the green and red dots. Other people are represented as grey dots.

V. Similarity-based behaviour clustering using HMM

Some of the relevant research to behaviour and activity monitoring includes work by Oliver *et al.* [9] who look at identifying human activities from multimodal sensor information. In [9], they use Hidden Markov Models (HMMs) and Dynamic Bayesian Networks for recognising office activities. Whereas in [10], they investigate a Bayesian framework using HMMs and Coupled Hidden Markov Models (CHMM) for recognising different human behaviours and interactions. However, their work is focused on recognition of activity rather than clustering or developing a similarity measure. In addition to that, the method is not applied to home environments. Relevant work which examines traces of activities includes the work on location based activity recognition by Liao *et al.* [11]. Although they provide methods of recognising activities from location sequences, the method requires labelling of activities, which could prove to be difficult in the case of elderly patients.

In this work, we will investigate a method of clustering behaviour sequences based on behavioural modelling using HMMs based on both wearable and ambient sensors but without explicitly defining activities, thus respecting the patients' privacy. The HMM analysis framework is based on that of Bicego *et al.* [12, 13], who introduce a similarity based clustering of sequences using HMMs.

A. Hidden Markov Models

HMMs are finite state stochastic machines that allow dynamic time warping for modelling sequential data. An HMM can be defined by the following [14]:

- $S = \{S_1, S_2, S_3, ..., S_N\}$, a finite set of hidden states.
- The transition matrix A, where each element a_{ij} represents the probability of moving from one hidden state S_j to another hidden state a_i.
- The emission matrix B where each element indicates the probability of emission of an observable symbol o.
- $\pi = \{\pi_i\}$, the initial state probability distribution.

Therefore, an HMM can be represented by the triplet $\lambda = (A, B, \pi)$. The Baum-Welsh algorithm can be used to learn the parameters of an HMM, where the likelihood of a sequence of observable states given the model is maximised.

B. Similarity Based Clustering of Sequences

Standard approaches to clustering with HMM include the training of an HMM with a sequence, then using pair-wise distance based methods to perform the clustering. However, Bicego *et al.* [12, 13] introduce a clustering approach for sequences in a feature space that is generated using HMMs. The features describing a sequence O are calculated as similarity measures $D\ O\ O_i$ between a sequence O and other reference sequences O_i. The reference sequences can be selected sequences from the whole set as chosen by experts, or the whole dataset can be used. The algorithm to represent the sequences in a new feature space is as follows:

- Select a set of reference sequences $\{P_1, ..., P_R\}$.
- Train an HMM λ_r for each sequence P_r.

- Represent each sequence in the data by a vector of similarities to the elements of the reference set. Each element of the similarities vector $D_R(O_i)$ is the log likelihood of the HMM λ_r predicting sequence O_i normalised by the length T_i of the sequence O_i :

$$D(O_i, P_r) = \frac{1}{T_i} \log P(O_i / \lambda_r) \quad (2)$$

For patient behaviour profiling, one is interested in observing the clustering of sequences in the new feature space, as well as observing the Similarity matrix D in a lower dimensional space. This allows the observation of similarities and differences between labelled behavioural sequences.

Table 2 Sequences of different lengths showing a person with different behavioural patterns between 5 rooms.

Sequence Number	Description	Group
1	Walking in all 5 rooms	1
2	Walking in the corridor, spending almost equal times in rooms 3 and 5	2
3	Similar to 2 but doing more activities in room 3	2
4	Spending time in room 3 with a short period in room 4	2
5	Doing activities between rooms 2, 3 and 5	2
6	Doing activities in all five rooms, mostly walking in room 2.	3
7	Activities in all rooms, although spending more time in room2.	3
8	Activities in all five rooms.	3
9	Moving between rooms 2 and 4 and interacting with another person in room 4.	4
10	Moving between rooms 1, 2, 3 and 4 and interacting with another person in room 4.	4
11	Interacting with another person in room 1 for a long time, but also going to rooms 2, 3 and 5 briefly.	5
12	Interacting with someone in room 1 for a very short time, and then walking in rooms 2 and 5.	5
13	Walking in all 5 rooms	1

VI. Experimental setup

A simulation study was performed in a lab based environment representing a house with five different rooms. Blob sensors were placed in the ceilings of each room as shown in Figure 3, and a BSN receiver was placed in each room. One person was asked to wear an e-AR node while performing a list of 13 activity sequences in these rooms. The person had no limitation on time or type of activity but had to follow the guidelines given in Table 2. During the experiment, there were also several other people moving around the rooms to simulate a cohabiting environment.

The behaviour patterns in Table 2 were categorised into 5 groups according to the type of behaviour done. For example, group 2 contains sequences that include activities done in room 3 (such as sequences 3 and 4) shown in Figure 4. Although other people are shown moving around in figure 4, the Bayesian Classifier explained in section 2 allows the exact location of the person being tracked (shown as a red dot).

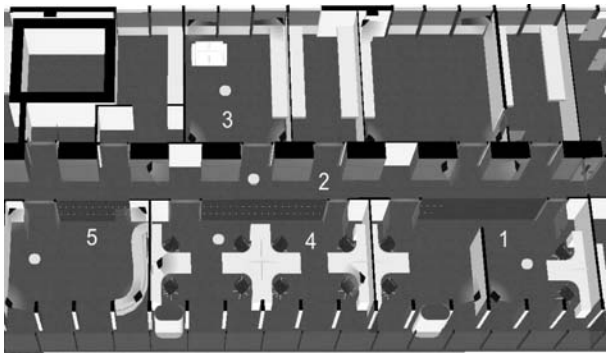


Fig. 3 A simulation environment showing the 5 rooms used (labeled 1-5). The blob sensors are shown in green and the BSN receivers are shown as yellow ellipses. Each room had 1 BSN receiver, but several blob sensors.

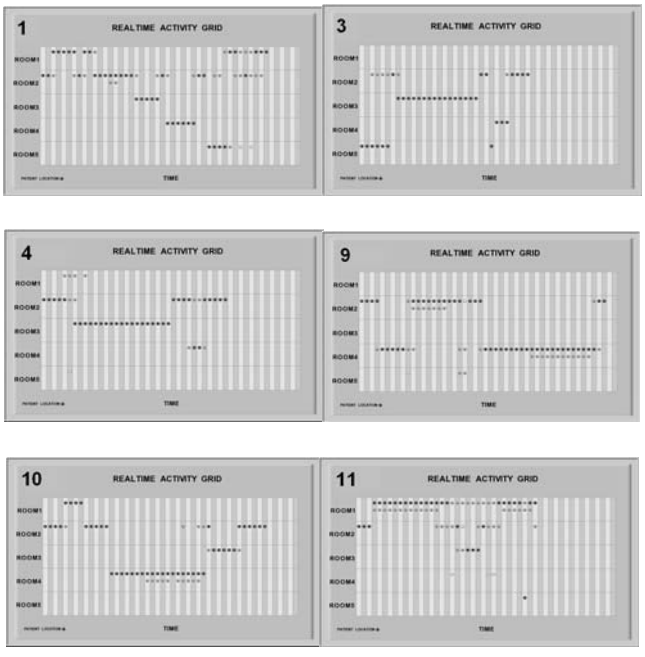


Fig. 4 Behaviour sequences from Table 2 showing different behaviour patterns. The numbers refer to the sequence number in Table 2. Activity sequences are normalized over time (although they are originally of different lengths) for display purposes.

VII. Results

A. Observing behaviour by using Transition matrices

Several observations can be made by examining the Markov state transition matrices (given in Figure 5) from the sequences in Table 2. Although sequences 3 and 4 show some similarity, it is difficult by using these traditional transition matrices alone to judge their similarity to other sequences in Table 2. Sequence 11, on the other hand, is quite different from the others as it involves the observed subject spending a long time in Room 1 interacting with another person (as can be seen in the last activity grid in Figure 4). However, it is difficult for a user to deduce that from the transition matrices in Figure 5. The HMM similarity framework proposed, however, provides easy abstraction of this information for comparing behaviour sequences of different lengths.

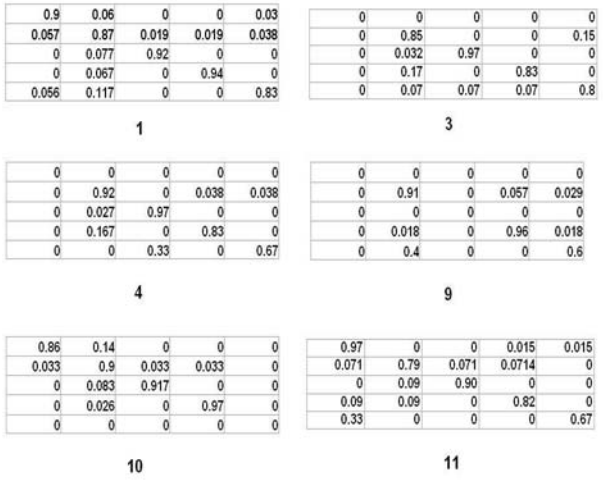


Fig. 5 Transition matrices of the activities in Table 2 of dimension 5x5 showing the probability of transition of a person from a room (vertical row) to another (horizontal row).

B. Results of the HMM Similarity Framework

The HMM framework described above was used to obtain a feature matrix to describe each sequence. The dimensionality of each feature is 13 since we have 13 different behaviour sequences. PCA was used to reduce the dimensionality of the matrix and observe similarities between different behavioural sequences. The results are shown in Figure 5 where each group is colour-coded. It is evident that the following observations can be deduced from Figure 5:

- Sequences 2, 3, 4 and 5 cluster together. These sequences mainly involve activities in Room 3. Although the movement between rooms and the types of activities done are quite varied (as shown for sequences 3 and 4 in Figure 4), they are grouped together since they describe a certain common behavioural pattern that is not present in other sequences.
- Sequences 13 and 1 are similar and describe a general motion between all 5 rooms. Sequence 8 is also quite similar.
- Sequence 10 is closer to sequence 1 than sequence 9. Figure 5 shows that in general, sequence 10 resembles 1 more, as it involves more rooms, whereas sequence 9 involves a long time in room 4 interacting with another person.
- Sequence 11 is significantly different from the others, as it involves the person spending a long time interacting with someone in room 1. This behaviour is shown to be quite different from all other sequences.

By comparing the results with the initial labelling of behaviour sequences given in Table 2, intrinsic similarities of the patterns can be observed. By assuming that certain clusters of behaviour sequences represent a normal activity of a patient over a period of time, the graph provides an effective means of observing deviations from normal behaviour patterns. If this deviation is large, such as that shown by sequence 11, further data analysis can be performed to establish why the class deviates from the 'normal behaviour' cluster. This can then trigger a scheduled home visit by the carer.

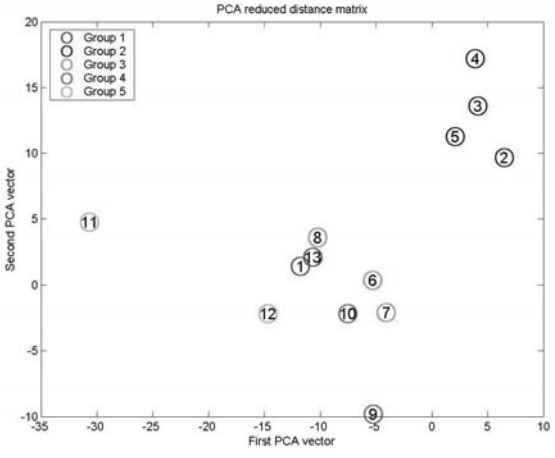


Fig.6 Each sequence described in Table 2 is represented a circle. The graph can be used to judge similarity between sequences.

VIII. Conclusions

This work presents a novel method of behaviour profiling based on ambient and wearable sensors. A behaviour modelling approach is considered since it is less intrusive to patient's privacy than identifying activities directly. A Bayesian classifier for extracting the intrinsic relationship between signal strengths from wearable sensors and room occupancy is used. The proposed method can be expanded to a more general home environment with minimal restrictions. Although a large amount of sensor data is obtained, most of the information can be abstracted immediately into high level information as the activity grid shown in this paper. The model can also deal with errors in the data resulting from data acquisition. Healthcare workers can use a reduced space grid to observe behaviour changes and social interaction. The HMM framework for behaviour clustering, could indicate an irregular event, as it would deviate from a cluster of normal behaviour. This can be used for more detailed analysis of the activity pattern or scheduling a home visit by the carers.

References

1. Intille, S.S., et al. *Using a live-in laboratory for ubiquitous computing research*. in *PERVASIVE 2006*. Berlin Heidelberg: Springer-Verlag.
2. Edwards, N., et al., *Life-Style Monitoring for Supported Independence.* BT Technology Journal, 2000. **18**(1): p. 64-65.
3. Tamura, T. *A Smart Home for Emergencies in the Elderly*. in *ICOST 2006*. 2006. Belfast: IOS Press.
4. Pansiot, J., et al. *Towards image-based modeling for ambient sensing*. in *BSN*. 2006.
5. Lo, B., et al., *Real-Time Pervasive Monitoring for Postoperative Care*, in *BSN 2007*. 2007: Aachen.
6. Yang, G.-Z., et al. *From Sensor Networks to Behaviour Profiling: A Homecare Perspective of Intelligent Building*. in *The IEE Seminar for Intelligent Buildings*. 2004: IEE.
7. Lo, B. and G.-Z. Yang. *Architecture for Body Sensor Networks*. in *The Perspective in Pervasive Computing*. 2005. IEE Savoy Place: IEE.
8. Aziz, O., B. Lo, and G.Z. Yang. *Pervasive Body Sensor Network: An Approach to Monitoring the Post-operative Surgical Patient*. in *BSN 2006*.
9. Oliver, N.M., et al. *A comparison of HMMs and dynamic bayesian networks for recognizing office activities*. in *UM 2005* 2005. Edinburgh: Springer, Berlin.
10. Oliver, N.M., B. Rosario, and A.P. Pentland, *A Bayesian computer vision system for modeling human interactions.* Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2000. **22**(8): p. 831-843.
11. Liao, L., D. Fox, and H. Kautz, *Location-based activity recognition using relational Markov networks*, in *Proc. of the International Joint Conference on Artificial Intelligence*. 2005.
12. Bicego, M., V. Murino, and M.A.T. Figueiredo, *Similarity-based classification of sequences using hidden Markov models.* Pattern Recognition, 2004. **37**(12): p. 2281-2291.
13. Bicego, M., V. Murino, and M. Figuerido. *Similarity based clustering of sequences using Hidden Markov Models*. in *MLDM*. 2003: Springer-Verlag.
14. Rabiner, L. and B. Juang, *An introduction to hidden Markov models.* ASSP Magazine, IEEE [see also IEEE Signal Processing Magazine], 1986. **3**(1): p. 4-16.

Address of the corresponding author:
Author: Dr Louis Atallah
Institute: Dept of Computing, Imperial College London
Street: Huxley Building, 180 Queen's gate
City: London
Country: UK
Email: latallah@doc.ic.ac.uk

Role of signal processing in wearable devices: application to sleep evaluation

A. M. Bianchi, O. Villantieri, M. Mendez and S. Cerutti

Politecnico di Milano, Department of Biomedical Engineering, Milano, Italy

Abstract—**In this paper we discuss the possibility of performing a sleep evaluation from the heart rate variability (HRV) and respiratory signals. This is particularly useful for non standard sleep measurements based on wearable devices or special sensors inserted in the bed. The HRV and the respiration signals were analysed in the frequency domain and the statistics on the spectral and cross-spectral parameters put into evidence the possibility of a sleep evaluation on their basis instead of the classical polysomnography in a sleep laboratory. Additional information can be achieved from the number of microarousals recognized in correspondence of typical modifications in the HRV signal.**

Keywords—**Sleep, HRV, spectral analysis, AR model**

I. Introduction

The great advances in nanotechnology, internet, wireless technologies, digital signal processing, textile sensors and sophisticated implantable devices gave rise to wide application of wearable devices (WD) and home monitoring systems for recording vital signs. The related advantages range from better possible diagnosis and monitoring for the patient during his/her daily life without being perturbed by the hospital environment, to the reduction of hospitalization costs. In addition also other non medical applications can be extended to healthy subjects, for example during sport activity. However, the increasing amount of signals and information produced by these devices needs to be properly interpreted and managed [1].

At this purpose, the role of signal processing in WD is to provide an intelligence to the system for the following functions:

- extract relevant features from the signal;
- select from the signal only informative time windows and eventually transmit only those to a remote station for the interpretation;
- calculate parameters for the synthesis of the data;
- provide a feedback to the user.

Monitoring sleep through WD or through sensors inserted in the bed, requires to face different problems, including the proper choice of the signals. In fact, the traditional parameters used in sleep evaluation (EEG, EOG, EMG, etc.) require an accurate preparation of the subject, made by clinical personnel and, more in general, a "clinical environment".

In the present paper we will discuss the possibility of performing sleep evaluation from signals non traditionally included in the polysomnography, but from signals that can be easily recorded through home devices.

II. Methods

A. Signal selection

Sleep is a physiological condition mainly involving the Central Nervous System (CNS), and is usually studied through the EEG signal associated with EMG and EOG.

In a sleep laboratory the polysomnography (i.e. the continuous recording of many signals, including EEG, EOG, EMG, Respiration, ECG, during sleep), is generally intended to the characterization of three main issues:

1. Sleep staging that requires EEG, EOG and EMG;
2. Insomnia and sleep fragmentation that require EEG, EOG, EMG and sometimes activity;
3. Sleep Apnea, that requires ECG, respiratory airflow, respiratory movements at chest and abdomen level, oxygen saturation.

However it is well known that also the Autonomic Nervous System (ANS) is strongly affected by the status of the Central Nervous System (CNS) and many researches testify the presence of autonomic changes related to different sleep stages [3]. The action of the ANS produces spontaneous fluctuations in the blood pressure and in the heart rate, that can be highlighted and quantified through the frequency domain analysis of the heart rate variability (HRV) signal. It was stated by numerous researches that three main components contribute in the HRV:

- the very low frequency (VLF) component is assumed to be related to long-term regulatory mechanismis such as humoral factors, temperature, and other slow components;
- the rhythm corresponding to vasomotor waves and present in heart period and arterial pressure variabilities is defined as the low frequency (LF) component. Since it increases during sympathetic stimulation, it can be considered a marker of sympathetic activation;

- the respiratory rhythm, synchronous with the respiration rate, defined as high frequency, (HF) is generally considered a marker of vagal modulation.

A reciprocal relation do exists between the power of the LF and HF rhythms. This balance can be quantified by the LF/HF ratio [4, 5].

B. Sleep staging

24 healthy subjects were recorded in the sleep lab through the clinical polysomnography. The frequency analysis of the HRV signal was applied during sleep after the sleep staging performed through the classical procedure based on [2]. During sleep HRV is characterized by high non stationarity in correspondence of stage transitions, microarousals, movements, ect., thus, for a reliable frequency analysis we used AR models implemented in the time-variant form [7]. The spectral parameters calculated from 24 normal subjects were fed in a classifier based on Hidden Markov models and an automatic sleep staging on REM and non REM sleep was obtained.

C. Arousals detection

10 subject with sleep fragmentation were recorded in the sleep lab. An arousal on HRV is recognizable as a series of HR falls and consequent rises with respect to a previous baseline within a time window of about 10 sec. Fig.2 shows the HRV in respect to the mean value and expressed in sec, in correspondence of an arousal. The algorithm evaluates when the signal falls more than 15% respect the previous baseline. If, 10 sec later, the signal rises more than 10% respect the initial baseline than it is recognized as an arousal and the beat of the recognized falls is considered its beginning (see Figure 2).

D. Apnea detection

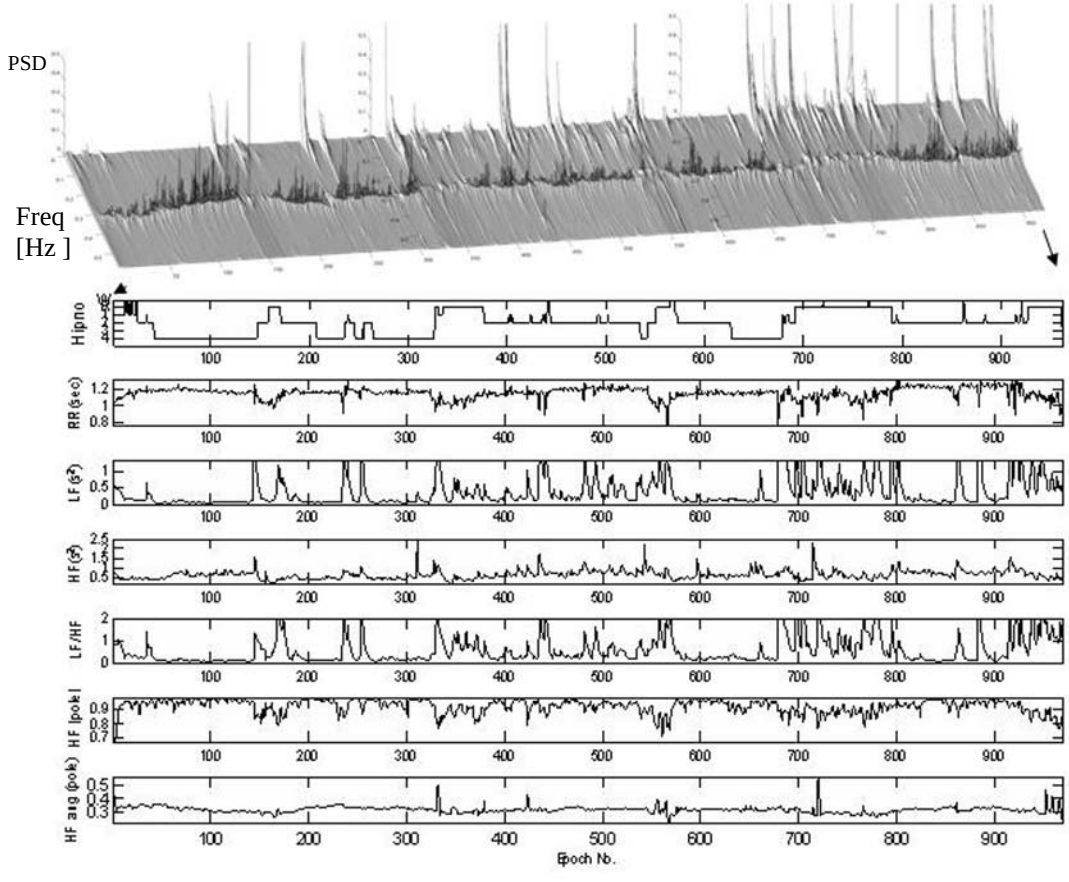


Fig.1 a) sequence of the beat-to-beat spectra obtained during a whole night recording; b) Hypnogram evaluated through the clinical polysomnography; c) RR interval sequence; d) beat-to-beat LF power; e) beat-to-beat HF power; f) beat-tobeat LF/HF

Overnight recordings coming from five healthy and five pathologic subjects with severe obstructive sleep apnoea (OSA) were analysed. Table 1 presents the mean and Standard Error of the HRV indexes used in the study. The statistics was carried out between REM and the others sleep stages for each group.

III. Results

Fig. 1 shows an example of the time evolution of PSDs and the related spectral indexes of the HRV during all night of one subject. From the PSD epoch evolution it can be noticed a clear amplitude decrement hence as a displacement toward a lower frequency of HF in correspondence to the REM stages (visible on the hypnogram plotted below), together with a high LF increment. Spectral indexes plot shows RR intervals, LF and HF power, LF/HF, module and frequency of HF pole during each epoch. Values of the indexes at each epoch were obtained by computing the mean each 30 seconds which corresponds to the classical time scaling used in the evaluation of the clinical hypnogram. From the RR plot one can appreciate a decrement.when a REM stage occur or during a change from a deeper to a lighter sleep stage, for instance, the change of stage from 4 to 2. HF showed similar variations, but not so clear, and LF behavior presented the opposite variations to the RR intervals, however with strong changes. Observing the LF/HF ratio, the combination of powers in both the relevant frequency ranges, it is remarkable a clear difference between the REM and NREM sleep stages. Furthermore, the changes in sleep stage and awakenings are very clear in this index. A classifier based on Hidden Merkov Models, with the above indices as input, classified the REM/nonREM sleep with the following scores:

- Sensitivity = 76.4 %
- Specificity = 69.2 %
- Accuracy = 89.9 %.

The arousals detected on the HRV signal were compared with the arousals detected on the EEG by an expert neurologist (gold standard) and the following results were obtained:

- Sensitivity = 81 %
- Specificity = 99 %
- Accuracy = 98.5 %.

Table 1 presents the mean and Standard Error of the HRV indexes used in the study. The statistics was carried out between REM and the others sleep stages for each group. In normal subjects, the RR intervals presented high values in light and deep than REM and wake, being significant only in light respect to REM. VLFn was lower during deep, light and wake than REM (P>0.5). LFn and LF/HFn presented a smaller value during deep and light than wake and REM. The significant difference was found respect to REM in deep sleep stage. HFn presented a increase level (P>0.5) during deep sleep. Interesting results are the statistic difference, in VLFn, between light-deep and REM in OSA patients. On the other hand most of the spectral indexes presented statistic difference in the sleep stages between both groups. VLFn, LFn and LF/HFn were higher in OSA subjects than normal ones while HFn presented lower values.

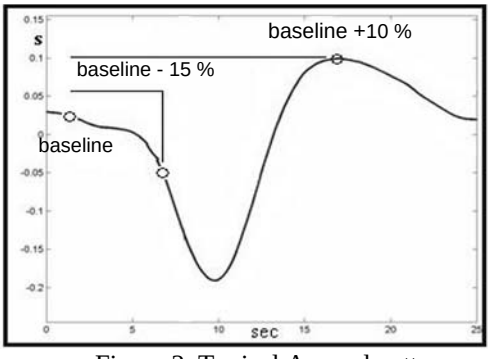


Figure 2. Typical Arousal pattern

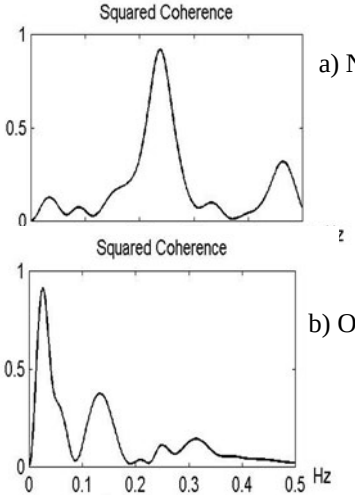


Fig.3 squared coherence function between HRV and respiration in a) normal subject and b) OSA patient

In addition bivariate analysis between HRV and respiration signals put into evidence a different mechanism of coupling. Fig.3 shows the squared coherence funtion (k^2) for RR sequence and respiration in a normal subject (a) and in a OSA patient (b). The k^2 function is calculated according the following equation:

$$k^2(f) = \frac{S_{xy}^2(f)}{S_x(f) \cdot S_y(f)}$$

TABLE 1. Mean and Standard Error of the spectral indexes of Heart Rate Variability during the Sleep Stages in both normal and OSA subjects.

Index	Normal				Obstructive Sleep Apnoea			
	Wake	Light	Deep	Rem	Wake	Light	Deep	Rem
RR (s)	0.976±0.05	1.008±0.06	0.986±0.06	0.942±0.05	0.875±0.02	0.949±0.03	0.899±0.04	0.975±0.03
VLFn (%)	0.367±0.03	0.169±0.02	0.111±0.02	0.259±0.04	0.382±0.01	0.267±0.01*	0.230±0.01*	0.428±0.04*
LFn (%)	0.428±0.01	0.503±0.45	0.393±0.02	0.499±0.02	0.515±0.01*	0.637±0.01*	0.583±0.02*	0.521±0.41
HFn (%)	0.203±0.03	0.327±0.03	0.494±0.03	0.241±0.04	0.102±0.01*	0.095±0.03*	0.186±0.03*	0.515±0.00*
LF/HFn (%)	5.020±1.05	2.802±0.66	1.399±0.36	3.347±0.54	20.33±7.81*	10.39±1.43*	6.097±2.45*	14.06±1.17*

RR = time interval between consecutive R peaks of the Electrocardiogram, LFn = low frequency component, HFn = high frequency component, LF/HFn low to high frequency ratio. * represents significant difference between corresponding sleep stages of the groups. The gray color denotes the statistic difference between REM and the other sleep stages for each group. $P < 0.05$.

Where $S_{xy}(f)$ is the cross-spectrum of the two signals, and $S_x(f)$ and $S_y(f)$ are the respective auto-spectra.

In normal subjects there is a peak in the coherence at the respiratory frequency around 0.23 Hz, while in OSA patient the peak is around 0.02 Hz, that represent the frequency of repetition of the apnea episodes.

IV. Conclusions

Dealing with wearable devices, it is important to keep in mind that the recordings must present good signal to noise ratio, must be robust to movements, and their information content should not strictly depend upon the correctness of the position of the sensors (the user should be able to wear the sensors by himself, without the help of an expert). From this point of view the HRV signal is much better then the EEG and thus we decided to base the sleep evaluation on this analysis. In this paper we demonstrated that some spectral and cross-spectral indexes from the cardio-respiratory system are significantly related to sleep stages and some sleep disturbances, thus they could be the basis of a sleep evaluation instead of the more classical EEG. That is motivated by the need of using signals that can be obtained from wearable systems, which the user can wear by himself at home, without the help of anybody. Further, HRV is less sensitive to noise then EEG. The statistical analysis put into evidence that many of them seem suitable for segregating between REM and non REM sleep, and can characterize some episodes occurring during sleep (arousals, apneas). The use of a robust classifier allows the automatic evaluation of the sleep on the basis of the selected parameters. Aim of the study is to provide a sleep evaluation, that could be of help in defining the quality of the sleep during the night and to the early identification of sleep disorders, without the need of a complete polysomnography, but at home, using simple devices.

Acknowledgment

This work was supported by the European project MY HEART and CONACYT.

References

1. R. Paradiso, G. Loriga, N. Taccini., A wearable health care system based on knitted integrated sensors. IEEE Trans Inf Technol Biomed, vol .9, pp. 337-44, 2005
2. A. Rechtschaffen , A. Kales, Manual of standardized terminology,techniques and scoring systems for sleep stages of human subjects. Los Angeles, CA: Brain Inf. Service/Brain Research Institute, 1968
3. U. Sholz, A.M. Bianchi, S. Cerutti, S.Kubicki, Vegetative backgruound of Sleep: Spectral Analysis of the Heart Rate Variability, Physiol& Behavi, vol 62, pp. 1037-1043, 1997.
4. Task Force of the European Society of Cardiology and the North American Society of Pacing and Electrophysiology Heart rate variability. Standards of measurements, physiological interpretation, and clinical use. Circulation 93, 1043–1065, 1996
5. A. Malliani, The pattern of Sympathovagal Balance Explored in the Frequency domain, News Physiol. Sci. ,vol 14, pp. 111-117, June 1999
6. A.M. Bianchi, OP Villantieri, MO Mendez, V Patruno,S Cerutti, N Montano, Different Effects of CPAP and APAP Therapies on the Autonomic Nervous System in OSA Patients, Computers in Cardiology Conference, Lyon, 2005
7. A.M. Bianchi, L. Mainardi, E. Petrucci, et al.. Time-variant power spectrum analysis for the detection of transient episodes in HRV signal," *IEEE trans. Biomed. Eng.*, vol . 40, no. 2, pp. 136144, 1993.

Address of the corresponding author:

Author: Anna M. Bianchi
Institute: Politecnico di Milano, Dep. Of Biomedical Engineering
Street: piazza Leonardo da Vinci 32
City: 20 133 Milano
Country: Italy
Email: annamaria.bianchi@polimi.it

Evaluation of a new, wireless pulse oximetry monitoring system in infants: the BBA bootee

Jean-Luc Weber, Yves Rimet, Eric Mallet, Dominique Ronayette, Caroline Rambaud, Christian Terlaud, Yves Brusquet, Christian Dageville, Marc Lubrano, David Blanc, Jérôme Silve, Olivier Lerda, Francis Klefstad-Sillonville, Larissa I. Netchiporouk

***Abstract*—A new system for wireless monitoring of pulse oximetry (SpO2) was developed on the basis of a Nonin OEM III oximetry module and a radio-frequency transceiver. The electronic unit and the power supply was integrated into a specially designed infant shoe named BBA bootee. Clinical evaluation of the system revealed a good agreement between the pulse oximetry data transmitted by the BBA bootee and those recorded simultaneously by a reference monitor. The comparative data collected in 39 babies yielded a mean (bias SD) value of (–1.1 1.9)% for SpO2 and (-2 8) beats per minute for heart rate. Use of an integrated accelerometer/actimeter to reduce the motion artifact is addressed as well as ergonomics of the sensor-supporting garment.**

***Index Terms*— apparent life threatening events, smart textiles, telemedicine, SIDS, surveillance**

I. Introduction

Each year, numerous newborns and infants present serious malaises due to major apnea and/or bradycardia, and they may even die suddenly. The causes of such accidents may vary, ranging from the consequences of prematurity to pathological malformations, metabolic disturbances or serious even fatal (and sometimes unexplained) acute syndromes [1].

When the possible onset of such events is foreseeable, the monitoring of major functions (heart rate, respiratory rate, pulse oximetry) is generally proposed. Such surveillance, which is often ensured in a hospital environment, may also be possible in the home when the causal mechanism underlying such incidents is insufficiently controlled [2].

Monitored surveillance involves the use of sensors which are fixed to the infant and linked by wires to a monitor for analysis or to trigger an alarm [3]. This technical interface hampers the comfort of the infant and the family, and may even affect the quality of parent-child relationships. It can also be a factor for poor compliance with use of the monitor [4]. Efforts have been made to reduce this discomfort, particularly by integrating sensors in clothing, such as pajamas or underwear [5]-[7]. In order to minimize this problem, we propose a novel system to monitor oximetry and heart rate ("pulse oximetry") called the BBA bootee, which is characterized by the textile base for the sensors – a bootee – and the wireless link to the monitor. This bootee also integrates an accelerometer to provide information on foot movements by the infant. We report herewith on a multicenter evaluation of the reliability and reproducibility of the data collected by the pulse oximetry sensor (SpO2) and the accelerometer integrated in the bootee, performed in both newborns and infants.

The principal end-point of this evaluation was to compare the SpO2 and heart rate (HR) values generated by the BBA bootee with those provided by a reference monitor, in a hospital setting.

The secondary end-point was to evaluate the ergonomic characteristics of the BBA bootee.

The research which enabled development of the BBA bootee and its evaluation in a hospital setting was funded by the Fondation ANAÏS DUMANOIS, working under the aegis of the Fondation de France.

Y. Rimet is and Y. Brusquet was with Hospital Center of Aix-en-Provence, 13100 France (e-mail: yrimet@ch-aix.fr)

D. Ronayette is with University Hospital Center of Limoges, 87000 France (e-mail: dominique.ronayette@wanadoo.fr)

C. Dageville is with University Hospital Center of Nice, 06200 France (e-mail: dageville.c@chu-nice.fr)

M. Lubrano and E. Mallet are with University Hospital Center of Rouen, 76000 France (e-mail: marc.lubrano@chu-rouen.fr)

C. Rambaud is with Raymond Poincaré Hospital, Garches, 92380 France (e-mail: caroline.rambaud@rpc.aphp.fr)

C. Terlaud was with Hospital Center of Saint-Junien, 87200 France (e-mail: christian.terlaud@wanadoo.fr)

L.I. Netchiporouk, D. Blanc, J. Silve, O. Lerda, F. Klefstad-Sillonville, and J.-L. Weber (corresponding author) are with TAM-Télésanté, Aix-en-Provence, 13545 France (phone:+33(0)4 42 97 11 14; fax:+33(0)4 42 97 11 15; e-mail: tam@tamtelesante.com)

II. MATERIALS AND METHODS

A. Description of the monitoring system

1) Pulse oximetry

This takes the form of a reusable sensor (Nonin, model 8008J) with shortened cable and a Nonin pulse oximetry module (ref. OEM III). Every second, this module outputs mean SpO2 and HR values concerning the last four beats. An adjustable fabric strap on the bootee keeps the sensor in contact with the skin and protects it from the influence of ambient light.

2) Accelerometer/actimeter

The system is equipped with a complete, dual access accelerometer (Analog Devices, ref. ADXL 311.)

3) *Transceiver and antenna*

The Nordic microcontroller transceiver (ref. nRF9E5) transmits oxymetric and actimetric data via a short wave radiofrequency link every second to a monitoring and recording unit situated a few meters away, via a ceramic antenna. The actual power emitted by the antenna is less than 0.5 mW.

4) Management board and battery

This controls the measurement process performed by the bootee and its energy supply. It includes the Nonin module, the Nordic transceiver, the accelerometer/actimeter, the on-off switch, the battery and its recharging circuitry. Depending on the type of battery used the bootee has an autonomy of 10 to 45 hours.

5) Electronic containing box

Made of synthetic material and measuring 0.7 cm 0.4 cm 16 cm, it is integrated into the sole of the bootee. It is a sealed, splash-proof unit (IP 64) which can be removed so that the textile parts of the bootee can be washed (Fig. 1).

Fig. 1. A sealed, splash-proof box contains all electronic components for measurements and radio transmissions, integrated in the sole of the bootee.

6) Textile used for the BBA bootee

The BBA bootee is made like ordinary baby clothes, and is available in several sizes (European sizes 13 to 19). The straps ensure contact between the SpO2 sensor and the skin and permit to adjust the bootee to the foot as the infant grows from one size to another. Openings and adjustable fixings ensure reliable pre-positioning of the SpO2 monitoring sensor (Fig. 2). The fabric is resistant to natural projections from the infant (urine, liquid stools, etc.); it can be washed by hand or in a machine at 40°C (no spinning) and dries rapidly. A single bootee in the pair is sufficient to ensure monitoring.

The weight of the equipped bootee, including its battery, is 46 g.

Fig. 2. The BBA bootee with the pulse oximetry system fixed within it.

7) The monitoring and recording unit (MRU)

A PC computer with a Celeron 2.8 Mhz processor under Windows XP and software in C#, is located close to the infant's bed. It is equipped with a Nordic radio transceiver identical to that integrated in the bootee; it receives and records all data sent by the transceiver from the bootee. During the study, it also received by a RS 232 link all data transmitted from the reference monitor.

8) Webcam

During the study, evaluation of the pertinence of the actimeter integrated in the bootee sole was ensured by a camera (e.g. Webcam) placed on the cot and focused on the infant's feet. An orthogonal reference point was marked on the sole of the bootee (Fig. 3).

9) Reference monitors

The hospital departments which participated in the evaluation program used one or other of the following monitors:

- Hewlett Packard Merlin (HP), equipped with a Nellcor SpO2 measurement module and Oxysensor single patient sensors, providing averaged measurements over 12 seconds, updated every second.
- DATASCOPE Passport II (DS), equipped with a Masimo SpO2 measurement module and LNOP-Neo single patient sensors, providing averaged measurements over 8 seconds, updated every second.

10) Signal processing and comparison of data

Initially, it was intended to apply the standard method for correlation between the values supplied by the bootee and those supplied by the reference monitor: this comparison produced satisfactory results for HR, particularly because the range of HR measurements was sufficient, but it was not satisfactory for SpO2, which had a very limited range of values.

Indeed, this comparison was one of the points which had been criticized by Bland and Altman [8] concerning the statistical methods used to evaluate the concordance between two methods for clinical measurement, when it was demonstrated that the method of correlation could indicate poor concordance, even though the methods did in fact agree.

Consequently, we applied the method recommended by Bland and Altman, which consisted in analyzing differences in the measurements obtained using two sensor systems: the BBA bootee and the reference monitor.

For each recording, the algorithm was initiated by extracting pairs of values recorded at the same time, but separately, by the two measurement systems. It was thus possible to obtain approximately one pair of values per measurement second. For each pair, the algebraic difference between the BBA bootee value and the reference monitor value was calculated. It was thus possible to calculate the mean of algebraic differences and its standard deviation for all the exploitable data recorded.

B. Monitoring protocol

The bootee was placed on the left foot of the infant, connected by radio to the MRU. The right foot was equipped with an SpO2 sensor linked by wire to the reference monitor, which in turn was connected to the MRU, which received and recorded the data from the two measurement systems simultaneously (Fig. 3). The camera was positioned opposite the infant's feet and the recording of a video sequence was triggered each time the feet moved. An observation sheet was associated with each recording.

The monitors were placed on a trolley near the infant's bed.

A control screen made it possible to view the data transmitted by the BBA bootee to the monitoring and recording unit.

C. Population of the study

To be included, the infants had to be hospitalized for a condition requiring surveillance with a monitor (pulse oximetry on the foot), and hospital treatment

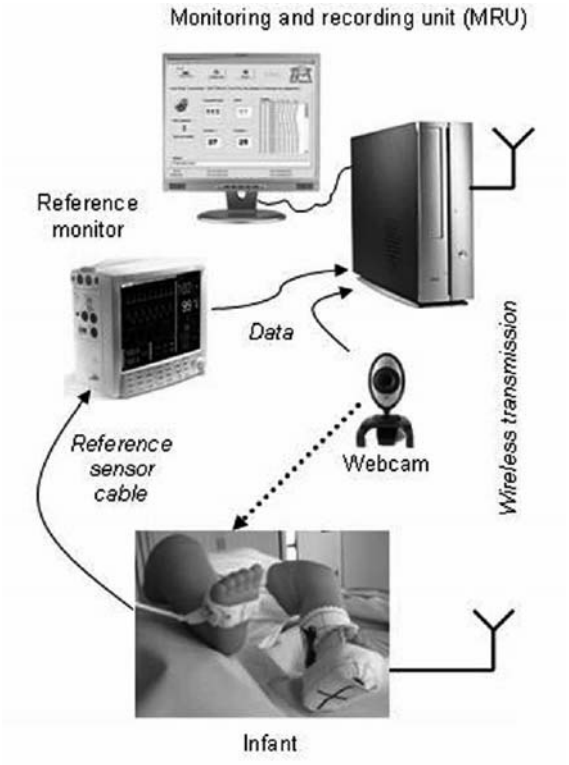


Fig. 3. Diagram showing comparative data collection setting.

would not be hampered by the conduct of recordings. A major instability of vital functions constituted a contraindication to inclusion.

Comparative recordings were made over a period ranging from 2 to 18 hours. After each recording, the BBA bootee was cleaned as follows: machine washing at 40°C using a disinfectant detergent (Anios, ref. Lessive Atomisée Désinfectante.) The SpO2 sensor was cleaned using contact disinfectant.

Approval was sought from French Ethics Committee (CCPPRB), and parents were asked to sign an informed consent form prior to inclusion of their infant.

III. Results and Discussion

19 full-term newborns, 9 premature infants and 11 infants aged from 1 to 6 months (total: 39 infants) were included between March 2005 and April 2006, at four clinical sites (*1*).

A total of 167 hours of recordings were made during 54 sessions. A summary of these recordings is shown in Table I.

TABLE I.
SUMMARY OF SPO2 AND HEART RATE (HR) MEASUREMENTS MADE WITH THE BBA BOOTEE COMPARED WITH THOSE USING THE REFERENCE MONITORS (HP: HEWLETT PACKARD MERLIN, DS: DATASCOPE PASSPORT II)

Clinical site (reference monitor)	No. of recor-dings	Mean body weight and age	Total duration of recordings (mean value)	Mean reference values of **SpO2** and Heart Rate	Mean of differences $=V_{BBA}$-Vref **SpO2** HR	Standard deviation of differences SD($=V_{BBA}$-Vref) **SD(SpO2)** SD(HR)
I (HP)	22	4 weeks 3.0 kg	90h 51 min (3h 52 min)	**97.0%** 147bpm	**-2.1%** -4 bpm	**2.2%** 10 bpm
II (DS)	17	3 weeks 2.1 Kg	41h 34 min (2h 26 min)	**93.2%** 149 bpm	**1.6%** -3 bpm	**1.6%** 7 bpm
III (HP)	8	3 months 4.5 kg	23h 40 min (2h 57 min)	**98.1%** 119 bpm	**-3.2%** 2 bpm	**2%** 9 bpm
IV (DS)	7	1 week 2.7 Kg	11h 01 min (1h 34 min)	**94.6%** 143 bpm	**-0.5%** -1 bpm	**1.9%** 4 bpm
Total	**54**	5 weeks 3.1 kg	167h 06 min (2h 42 min)	**95.7%** 140 bpm	**-1.1%** -2 bpm	**1.9%** 8 bpm

A. Metrological comparison

The measurements performed at the four clinical sites demonstrated good concordance between the values provided in the same patient at the same time by the BBA bootee and the reference monitor.

The mean of the difference between the SpO2 values measured by the BBA bootee and those of the reference was -1.1% (standard deviation 1.9 %.)

The mean of the difference between the heart rate values measured by the bootee and those of the reference was -2 bpm (standard deviation 8 bpm.)

For SpO2, calculation of the confidence intervals at 2 standard deviations for each recording showed that for all recordings, the variability in the confidence interval ranged from 87% with an SD 1.3%, to 96% with an SD 3.9%.

Taking into account that the manufacturers of the Nonin system and the reference monitors specified a standard deviation of ±3% within the range of SpO2 measurements from 70% to 95%, and that the 96% confidence interval was attained with 2 standard deviations, i.e. 6%, it was possible to conclude as to a good concordance of measurements within the limits stipulated by the manufacturers with respect to the accuracy of the systems under comparison.

The differences observed from one clinical site to another between the mean values supplied by the reference monitors could be explained by disparities in patient recruitment (depending on the type of patients treated at each site) and by the dispersion of measurements specific to each monitor. In particular, in Center III, the infants included had a higher mean age than in the other centers, which explained a significant drop in the mean HR values in this center, as measured by the reference monitor (119 versus 145, which was the mean in the three other centers).

Within the ranges of low perfusion (SpO2 at around 86%), bootee measurements were well correlated with reference measurements. For example, in two cases in Center II, the mean for the reference monitor was 86%, and the mean of differences with the bootee was (0.3 1.8) %.

SpO2 measurements appeared to be more sensitive to differences in sensor fabrication than HR measurements. For example, the mean SpO2 values recorded by reference monitors (depending on the sites) were 97% and 98.1% with the HP Merlin and 93.2% and 94.6% with the DATASCOPE, while the mean HR values were almost identical (from 143 to 149 bpm), whichever reference monitor was used, and after correction for the age of infants (cf. above). This could explain the dispersion in mean of differences for SpO2 measurements between sites, as a function of the reference monitors used. In contrast, HR measurements were very similar to the reference measurements (Fig. 4).

These studies showed that modification of the sensor application system did not affect the quality of measurements, as the errors observed were usually attributable to movements of the infant by himself or by a caregiver.

SpO2 monitors such as Nellcor and Masimo currently used in hospitals are reputed to reduce these motion artifacts by means of a sophisticated signal processing. As counterpart, their dimensions, weight and power consumption exclude their use in an

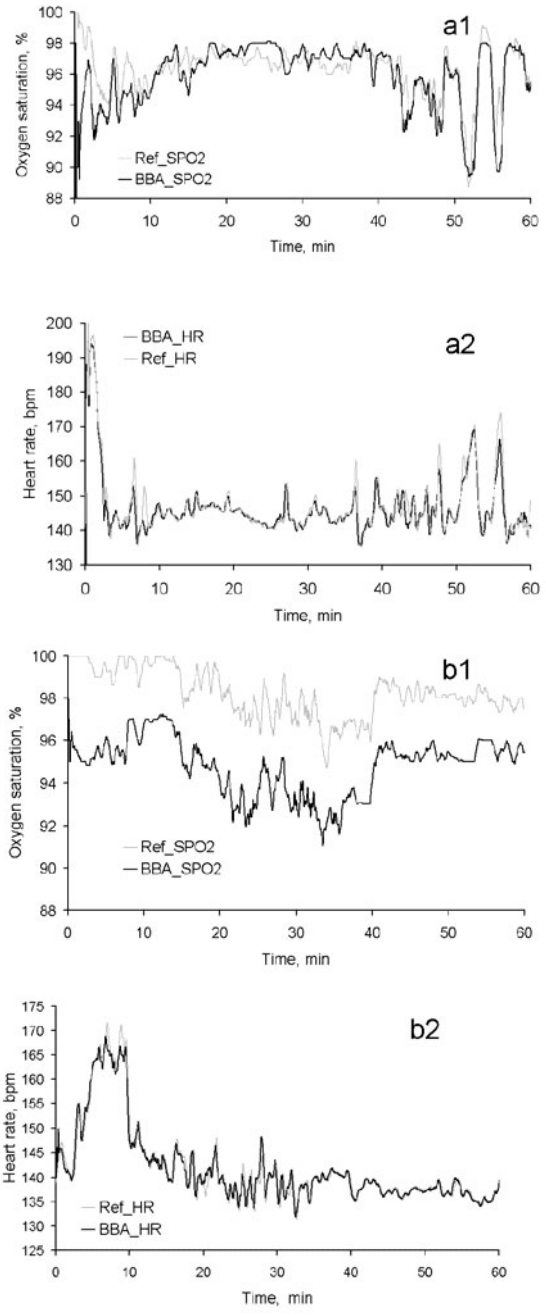


Fig. 4. Recording graphs of SpO2 values and heart rate (HR) using the BBA bootee (dark line) by comparison with the reference monitor (clear line).

Graphs a1 and a2 refer to a case where measurements made by the bootee did not show any significant difference from the reference values (mean of difference SD for SpO2 and HR is (0.5 1.2)% and (2 4)bpm, respectively).

Graphs b1 and b2 represent a case where SpO2 values follow those of the reference with a constant shift (mean of difference SD for SpO2 is (3.5 0.9)%), while the HR values are practically identical (mean of difference SD for HR is (0 2) bpm). The difference in behavior could be explained by a difference in the sensors calibration; it had no consequences on clinical interpretation, as the values fell within the acceptable range stipulated by the sensor manufacturers.

autonomous system like described here. For the moment, Nonin Inc. is the only recognized manufacturer offering a quality pulse oximeter module suitable to be integrated in a bootee for newborn infants.

We have tried another way to address the motion artifact problem by integrating accelerometers in the bootee. The idea was to match the oximeter errors against the infant's movements confirmed by the video images. There was a good correlation between rapid changes of the accelerometer signal level (that could be interpreted as baby's movements) and error messages sent by the oximeter. However, this part of analysis has not been achieved with the present evaluation protocol since few recordings were accompanied by exploitable Webcam images. Complementary data are expected from the clinical trial scheduled from February, 2007.

B.Ergonomic evaluation

It was simple to fit the bootee on the foot of an infant, as this only required the adjustment of two straps to the size and shape of the foot. Particular attention was paid to the shape of the bootee, so as to reduce the possibility of movement of the foot within the bootee, and prevent the infant from removing the bootee with the other foot (Fig. 5).

Wearing a bootee for several hours during the study sometimes induced a slight mark on the foot at the site of the sensor, but this was comparable to the effects of the reference sensors, and rapidly regressed. Without the constraint of any wires linked to the monitor, the feet could be moved freely. Baby's ability (or motivation) to move seemed not to be compromised by the weight of the bootee.

Fig. 5. Bootee in place on the foot

Preliminary studies performed using a wired system had shown that changes made to the Nonin SpO2

sensor did not affect its measurement and analytical properties. Compromises were made in order to reconcile reliability and ergonomics (simplicity and comfort of use, weight, autonomy of the battery, etc.). Thus only the results of calculations performed by microchips were transmitted once a second from the bootee to the MRU, and it was not possible to record plethysmographic curves.

If use of the BBA bootee is to be prescribed as an extension to the monitoring of infants in the home or in a hospital environment, certain points still need to be refined, such as the relevant detection of values which exceed the thresholds predetermined by clinicians.

The range of the transceiver which was validated during studies at clinical sites was approximately 2 meters. It would be useful to test an outreach of 5 meters: this range has been shown to function in a laboratory and at home, but data collection was hampered when the infant was in a cot with metal bars (Faraday cage effect).

IV. Conclusion

The BBA bootee for wireless pulse oximeter monitoring considerably improves the ergonomics of monitoring systems.

The multicenter studies performed in a wireless situation have demonstrated the quality and reliability of SpO2 and heart rate values, which are comparable to those obtained by the reference monitors widely used at present in hospitals.

The lightness and ergonomic design of the system open the way to a broad range of applications:

-monitoring of high-risk infants in the home (apparent life threatening events, malformation syndromes, chronic lung diseases, etc.) [9], [10],

-monitoring of newborns in the delivery suite during initial skin-to-skin contact with the mother, which is tending to become widespread, in the knowledge that several serious, sometimes fatal, accidents have been reported [11], [12],

-monitoring of newborn infants in the special care unit in the event of apnea and/or bradycardia and/or persistent desaturation events or other risk factors, such as major prematurity, which justify close monitoring.

The ability of the BBA bootee system to memorize alarm-triggering events so that they can be analyzed later, complies with current recommendations concerning monitored surveillance in the home [2], [10].

(1) Neonatal Care units involved in the evaluation:

Hospital Centre of Aix en Provence: Dr. Y. Rimet, Dr. C. Laisné (Site I);

University Hospital Centre of Limoges: Dr. D. Ronayette (Site II);

University Hospital Centre of Rouen: Dr. M. Lubrano, Prof. E. Mallet (Site III);

University Hospital Centre of Nice: Dr. C. Dageville, Prof. P. Boutet (Site IV).

Acknowledgement

The authors gratefully acknowledge the contribution of the late Professor André Kahn to the work which enabled this publication. Critical review of the manuscript by Dr. André Dittmar is greatly appreciated. Finally, we thank all the infants and their parents who participated in the trial.

References

[1] National Institutes of Health Consensus Development Conference on Infantile Apnea and Home Monitoring, Sept 29 to Oct 1, 1986, *Pediatrics*, 1987, vol.79, pp. 292-299.

[2] "Committee on Fetus and Newborn. American Academy of Pediatrics. Apnea, sudden infant death syndrome, and home monitoring", *Pediatrics,* 2003, vol.111, pp. 914-917.

[3] M.R. Neuman, H. Watson, R.S. Mendenhall, J.T. Zoldak, J.M. Di Fiore, M. Peucker, T.M. Baird, D.H. Crowell, T.T. Hoppenbrouwers, D. Hufford, C.E. Hunt, M.J. Corwin, L.R. Tinsley, D.E. Weese-Mayer, M.A. Sackner, Collaborative Home Infant Monitor Evaluation (CHIME) Study Group, "Cardiopulmonary monitoring at home: the CHIME monitor," *Physiol Meas*., 2001, vol. 22, no. 2, pp. 267-286.

[4] "The Collaborative Home Infant Monitoring Evaluation," *Arch. Pediatr. Adolesc. Med*., 2005, vol.159, pp.18-24.

[5] A.F. Aly. D. Afchine, P. Esser, M. Joos, H.J. Niewerth, A. Wiater, M. Meier, D. Padeken, A. Pericas, D. Schwartmann, T. Weber, V. Wendrix, M. Wirtz, "Telemetry as a new concept in long term monitoring of SIDS-risk infant", *Eur. J. Med. Res*., 2000, vol. 5, no. 1, pp.19-22.

[6] V. Gramse, A. De Groote, M. Paiva, "Novel concept for a Noninvasive cardiopulmonary monitor for infants: a pair of pajamas with an integrated sensor module," *Ann. Biomed. Eng*., 2003, vol. 31, no. 2, pp.152-158.

[7] M. Catrysse, R. Puers, C. Hertleer, L. Van Langenhove, H. van Egmond, D. Matthys, "Towards the integration of textile sensors in a wireless monitoring suit," *Sensors and Actuators A*, 2004, vol. 114, no. 2-3, pp. 302-311.

[8] J.M. Bland, D.G. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," *Lancet*, 1986, vol 1, pp.307-310.

[9] P. Leonard, T.F. Beattie, P.S. Addison, J.N. Watson, "Wavelet analysis of pulse oximeter waveform permits identification of unwell children," Emerg. Med. J., 2004, vol. 21, no. 1, pp. 59-60.

[10] A. Kahn, "European Society for the Study and Prevention of Infant Death. Recommended clinical evaluation of infants with an apparent life-threatening event. Consensus document of the European Society for the Study and Prevention of Infant Death 2003," Eur. J. Pediatr., 2004, vol. 163, no. 2, pp. 108-115.

[11] S.Espagne, I. Hamon, O. Thiébaugeorges, J.-M. Hascoet, "Sudden death of neonates in the delivery room," Arch. De pédiatrie, 2004, vol. 11, pp. 436-439.

[12] H. Gatti, C. Castel, P. Andrini, P. Durand, C. Carlus, J.L. Chabernaud, M. Vial, M. Dehan, C. Boithias, "Cardiorespiratory arrest in full term newborn infants: six case reports," Arch. De pédiatrie, 2004, vol. 11, pp. 432-435.

Electrochemical noise properties of different electrode materials in different electrolytes

J. Riistama and J. Lekkala

Institute of Measurement and Information Technology, Tampere University of Technology, Tampere, Finland

Abstract— **Signal-to-noise ratio of a biopotential measurement system is determined not only by the electronics of the measurement device but also by the electrochemical noise of the electrode-electrolyte interface. The intrinsic electrochemical noise depends on the electrode/electrolyte interface used. Noise of three metals suitable for implantation (Au, Pt and Stainless Steel) are being examined and referenced to Ag/AgCl commonly used as surface biopotential electrode material. Measurements of the electrochemical noise have been conducted in saline, PBS and SBF. Two different sizes of electrodes of every material have been used. Frequency bandwidth of 0.5 – 500 Hz was used in the measurements.**

Keywords— **Electrode, electrolyte, interface, noise**

I. Introduction

The electrode material plays an important role in the performance of a biopotential measurement system. It determines, together with the amplifier electronics, the detection limit and signal-to-noise ratio. Therefore, much attention has to be paid on the choice of material to be used as implanted electrode. There are several different materials that are suitable for measurement purposes but some of these are not biocompatible. An ideal choice for an implantable electrode is that it is highly biocompatible and shows only little intrinsic electrochemical noise arising from the unstable electrode-electrolyte interface.

Experiments on the electrochemical noise of the electrode-electrolyte interface has been made [1,2]. These experiments, however, do not take into account the electrolyte used. It would be interesting to see if the concentration of the electrolyte has any contribution to the noise voltage the electrode shows. Three different electrolytes have therefore been included in this research all of different conductivities.

Research reported here is a part of a research project which aims at an implantable ECG-monitoring device. Hence, much attention has been paid on the biocompatibility of the electrode materials.

II. Theoretical background and methods

A. Noise of the electrode-electrolyte interface

A conversion from ions to electrons occurs at the interface between electrolyte and the electrode. What is the actual charge distribution at the interface depends on the material of the electrode and the type of the electrolyte. There are at least four different models proposed on the charge distribution at the interface [3]. Similar to all of these models is that they all include a zone of charges that is constantly occupied by them. This layer is called the electric double layer. The potential over the double layer is called half cell potential and it is characteristic for every electrode/electrolyte couple [3].

When two electrodes made of the same material with equal area are placed into an electrolyte, theoretically they should not show any difference in half cell potentials. This is not, however, the case in real measurements since the electrode-electrolyte interface fluctuates, surface properties change, and some contaminant may be attached on one or both of the electrodes [3]. The interface between electrode and the electrolyte can be modeled with a resistor, R_s, that describes the conductive properties of the electrolyte. The resistor is connected in series with a constant phase element (CPE), a capacitor C_p, which is connected in parallel with another resistance, R_{ct}, refer to Fig. 1. An electrochemical noise source v_n and half cell potential source v_p have also been included into the picture. According to [1], calculating the thermal noise of the electrode-electrolyte interface using prescribed model is not enough to describe the noise behavior of the interface. This also indicates that a simple circuit model is not sufficient to model complex interface reactions and interface properties in details.

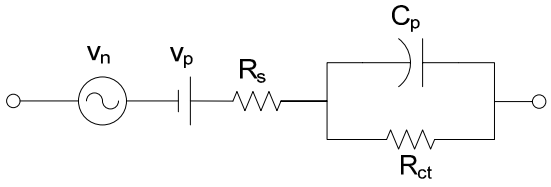


Fig. 1 A circuit equivalent for electrode-electrolyte interface together with voltage noise source and half cell potential source.

In order to obtain a high signal-to-noise ratio (S/N-ratio) in the measurements, the intrinsic noise of the measurement system including the interface ought to be minimized. Measurements of the interface noise are made to obtain more knowledge about the electrode materials and the electrode-electrolyte interface behavior.

According to [2], doubling of the electrode area will lower the voltage noise of the electrode by a factor of $\sqrt{2} \approx 1.414$.

B. Materials

Electrode materials studied are gold (Au), platinum (Pt), stainless steel (SS) and silver-silver chloride (Ag/AgCl). Only the first three of these are biocompatible. It is reported that an AgCl-coating will dissolve into the electrolyte in less than a week [4]. AgCl-coating is also toxic to tissue, as well as pure Ag. Therefore it is not suitable for implantation.

Ag/AgCl is, however, the most used material in surface biopotential electrodes. Together with electrode paste it is very stabile material which shows only a little electrochemical noise compared to many other materials [5]. It also has a remarkably rapid stabilization time, i.e. time it takes for the electrochemical noise of the electrode to stabilize to long term level, when immersed into an electrolyte. This property among others makes it usable in sensitive and rapid surface biopotential measurements [2]. The Ag/AgCl-electrode was chosen to be the reference for the other electrode materials in terms of electrochemical noise. The AgCl-coating was self applied on silver electrodes with a potential difference of 3 V held on for 2 minutes in 0.9 m-% NaCl solution.

Gold is found to be biocompatible material but it is not very commonly used as an electrode material in biopotential measurements.. This is mainly because it is expensive and suffers from corrosion much more than other noble metals when it is used for stimulating purposes. As a recording electrode, it is reported to undergo little or no corrosion at all. [6]

Platinum is widely used in electrodes, both in stimulating and recording ones. Platinum electrodes are found e.g. in cochlear implants and in pace makers as stimulating electrodes. They are found to be very biocompatible and suffer from corrosion only a little in both recording and stimulating use. [6]

Stainless steel (SS) is a common surgical material which is used e.g. in wires that close the sternum after a heart operation [7]. SS is chosen as one of the materials to be tested due to its generality and relatively low price.

The materials differ in their electric behaviour so that Ag/AgCl is a nearly ideally non-polarisable material while Pt being nearly ideally polarisable material. In a completely polarisable material there is no net transfer of charge over the electrode-electrolyte interface. Such a material can be regarded as a capacitor where there are two conducting surfaces, electrode and electrolyte, separated by an insulator, a double layer. Non-polarisable materials act contrary to polarisable materials: there is a net transfer of charge over the interface. Au and SS are located somewhere in between the two extremes. When electrodes are used for stimulation purposes, the polarisable ones are better than non-polarisables since the stimulation current is transferred capacitively over the interface and the electrode material does not release ions.

Several different electrolytes are being used in our in vitro test measurements. Saline, 0.9 m-% NaCl-solution, is commonly used although it is a quite simplified model fluid for the electrochemical environment inside the human body. More complicated solutions have been developed which simulate more accurately the various properties of the body fluids. Such solutions are among others, phosphate buffered saline (PBS, pH 7.4) and simulated body fluid (SBF) whose pH is temperature dependent being 7.4 at 37°C. All of the previously mentioned solutions are included in this research to investigate the contribution of the solution to the measurement result. Electrical conductivities of the electrolytes at DC were measured with a four wire measurement and results are reported in Table 1.

Table 1. Electrical conductivities of the electrolytes at DC.

Electrolyte	**Conductivity σ (S/m)**
PBS	1.411
Saline	1.599
SBF	1.834

C. Electrode preparation and setup

Electrode metals that were measured were ordered in two sizes, the ones having double the area of the others. The two different electrode sizes are used to be able to verify the area dependency of the electrochemical noise proposed in

section II.A. The electrodes with greater area have dimensions of 7 x 7 x 0.5 mm and the electrodes with smaller area have dimensions 7 x 3.5 x 0.5 mm.

The mid conductor of a coaxial cable of type RG174 was soldered on the electrodes where after the interface between cable and electrolyte was encapsulated with a layer of epoxy resin by Casco Inc. This encapsulation electrically insulates the centre wire from the electrolyte. The glue used to insulate the cable metal from the electrolyte covered partially also the electrode surface and not only the connection point of cable and electrode. Area caught by the glue was estimated to be approximately 6 mm^2. This electrically inactive area has to be taken into account in the analysis of voltage noise per unit area of the electrode.

The electrodes are attached on plastic plates with cable ties so that both their orientation and separation remains constant. The electrodes were placed on the plates in pairs with similar material and size together. The separation between the midpoints of the electrodes was set to 55 mm.

D. Electrochemical noise measurements

The electrode arrays were immersed into a container of dimensions 24 x 22 x 15 cm^3 (w x d x h) which was filled up to 5 cm height with the electrolyte to be studied. The container was placed inside a larger thermostated water container. Water was heated up to 42 ^oC which by thermal conduction heated the electrolyte up to 37^oC. Prior to each measurement, the electrodes were allowed to stabilize in the electrolyte for one minute. The electrolyte was not stirred under the measurement to secure stabile measurement circumstances.

The electrochemical noise measurements were performed by using a self-made amplifier and HP3561A Spectrum Analyzer. The measurement signals from the electrodes were led with coaxial cables to a battery operated measurement amplifier and further to the spectrum analyzer.

The signal was preconditioned in the measurement amplifier by high pass filtering it with a first order RC-filter with cut-off frequency at 520 mHz. The purpose is to filter off the DC-level of the measurement signal. After the filtering, the signal was fed into the instrumentation amplifier. The signal can be then amplified with one high gain amplifier without fear of clipping the signal. The gain in the amplifier is set to approximately 34 dB. The measurement is then done with a spectrum analyzer. The noise measurements were performed on bandwidth 0.5 – 500 Hz, which is often used in ECG-electrode measurements, [1,2,5]. On this frequency range the input impedance of the instrumentation amplifier is 100 MΩ at minimum which is sufficiently to secure a non loading measurement configuration.

The spectrum analyzer has a built-in low pass filter, or anti-aliasing filter, whose cut-off frequency is set automatically according to the frequency band to be measured. Therefore no separate low pass filter was needed. The measured signals were mean valued over 32 separate measurements to obtain a more stabile result.

The spectrum analyzer gives the spectral power density of the measured signal over the predetermined frequency band with units V^2/Hz. The power density spectrum is a variance spectrum where point-by-point subtraction and summation operations are possible.

From the measured power density spectrum, one can calculate the root-mean-square (rms) noise voltage over the bandwidth.

III. Results

A. Electrolyte: Phosphate Buffered Saline (PBS)

The noise measurements were performed in PBS as described in section II.D. Some problems arouse when Ag/AgCl-electrodes were being measured. Due to lack of instrumentation, e.g. scanning electron microscope (SEM), it could not be verified, what chemical processes took place on the surface of the electrode. It is suspected that some extra passivation layer was formed on the surface of the smaller Ag/AgCl-electrode during its exposure to air. The effect of the passivation layer was seen as lower voltage noise density compared to Ag/AgCl-electrodes with greater area. This is against the theory presented in section II.A.

The problem was solved by first removing the AgCl-layer from the both electrode pairs by rubbing the surface with fine sandpaper, grade 1000, avoiding to scratch the surface unnecessarily much. By scratching the surface, one would increase the effective area of the electrode which was not intended. The chloridation process was then applied again and the chlorided surfaces were further wiped with acetone before immersing into PBS.

Electrodes in the PBS-electrolyte were measured straight after measurements done in saline and it was noticed that some saline residues were left at the surfaces. Residues were deduced from the measurement results which were against the theory as explained and the noise level was extremely low. Au for example showed similar kind of behaviour as Ag/AgCl mentioned previously. The acetone cleaning was enough to clean up the surfaces and more realistic results were obtained. Results of the measurements are shown in Fig. 2.

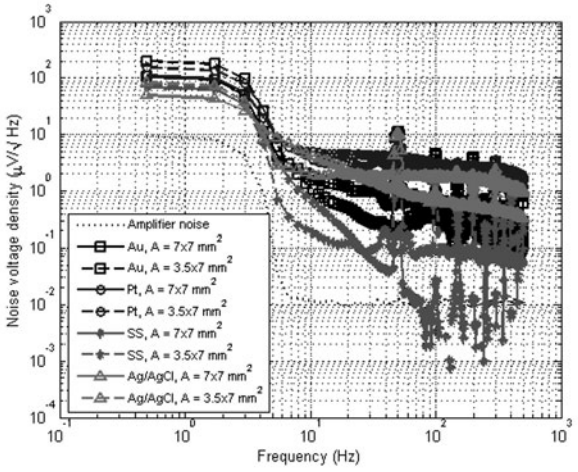


Fig. 2. Noise voltage density of the electrochemical noise at the electrode-electrolyte interface in PBS on frequency range 0.5-500 Hz. The notable spike at 50 Hz is the power line interference and its multiples at 50 Hz intervals. The voltage noise of the measurement amplifier has been subtracted from the electrode noise measurements.

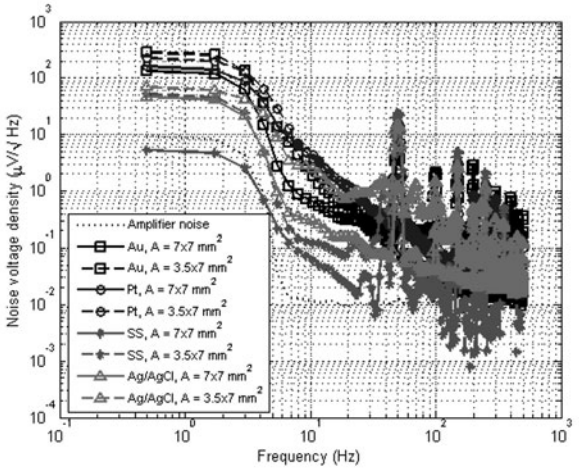


Fig. 3. Electrochemical noise voltage density in saline on frequency band 0.5-500 Hz. 50 Hz power line interference is noticeable in this measurement as well as its multiples. The voltage noise of the measurement amplifier has been subtracted from the electrode noise measurements.

The results indicate that Platinum is surprisingly noisy and SS the least noisy of all materials. Ratios of electrochemical rms-noises show good correspondence with the theory apart from the measurement results with Ag/AgCl-electrodes, see Table 2.

The root-mean-square (rms) noise voltage should be related to the area of the electrode as explained in section II.A, i.e. inversely proportional to the square root of the effective area. The rms-noise voltages, $v_{n,g}$ and $v_{n,s}$, were calculated from the spectrum together with the ratios of the noise voltages between different electrode areas. Results are reported in Table 2.

B. Electrolyte: Saline (0,9 m-% NaCl)

Similar problems were confronted with the Ag/AgCl-electrodes in saline as in PBS. The same method as in saline was used successfully and the measurements could be run through. Results of the measurements are shown in Fig. 3.

Table 2. shows the calculated rms-voltages for different electrode materials and different areas for saline. The ratio between the rms-noise voltage of the electrode pair of smaller area and greater area is shown.

The results with saline are very similar to the results of measurements made in PBS. Platinum tends to be the noisiest material except with smaller area where Gold exhibits the highest rms-noise at low frequencies. Discrepancy between the theory and measurements regarding the area dependency can be seen again in Ag/AgCl.

C. Electrolyte: Simulated Body Fluid (SBF)

SBF is a very interesting electrolyte since its pH varies as a function of temperature. A pH 7.4 is obtained at 37°C where the temperature of the electrolyte is kept during the measurements. Temperature is kept constant with the thermostatically controlled system explained in more detail in Section II.D.

Ag/AgCl was problematic also in SBF and precisely the same behaviour was observed as in PBS and NaCl. The method to clear the electrodes that was used in the two other electrolytes was successful again in SBF. Electrochemical noise spectrum is shown in Fig. 4 and corresponding rms-noise voltages are tabulated in Table 2.

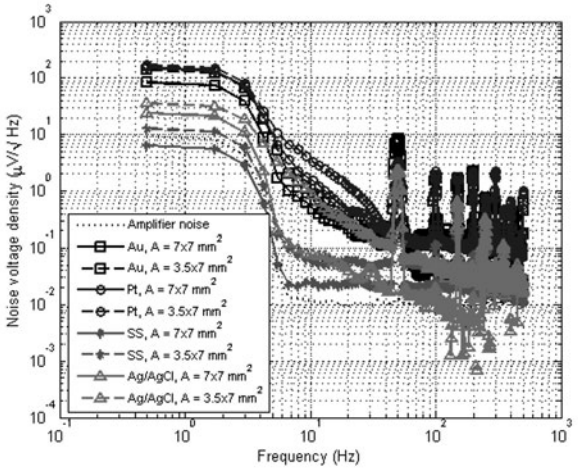


Fig. 4. Noise voltage density of the electrochemical noise at the electrode-electrolyte interface in SBF on frequency interval 0.5-500 Hz. 50 Hz power line interference can be clearly seen with its multiples. The voltage noise of the measurement amplifier has been subtracted from the electrode noise measurements.

Behaviour of the electrode materials is the same as in other electrolytes, i.e. Platinum shows the greatest noise and Stainless Steel the smallest.

D. Comparison between electrolytes

A comparison between the electrochemical noises of electrodes in different electrolytes was made by collecting the rms-noise voltages into one picture and connecting them with lines for clarity. Fig. 5 shows the result of the comparison.

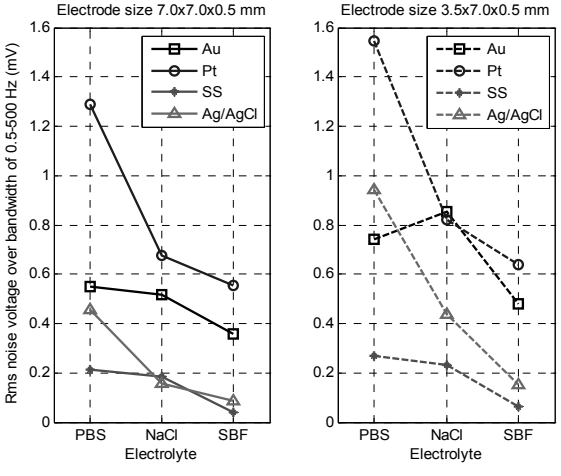


Fig. 5. Rms-noise comparison plot between the different electrode materials in different electrolytes.

Table 2. Rms-noises of different metals in different electrolytes.

Electrolyte	Material	$v_{n,g}$ (μV) (dim 7 x 7 x 0.5 mm)	$v_{n,s}$ (μV) (dim 3.5 x 7 x 0.5 mm)	Ratio $v_{n,s}/v_{n,g}$
PBS	Au	540	725	1.34
	Pt	1260	1510	1.20
	SS	212	275	1.30
	Ag/AgCl	449	922	2.06
Saline	Au	507	834	1.64
	Pt	661	804	1.22
	SS	182	227	1.25
	Ag/AgCl	152	430	2.83
SBF	Au	352	469	1.33
	Pt	542	623	1.14
	SS	39	62	1.58
	Ag/AgCl	84	148	1.76

IV. Discussion

The form of the measurement results of the different electrode materials in different electrolytes is the same in all cases: On low frequencies the noise is higher than on higher frequencies, which indicates the existence of $1/f$ –noise. It is not clear why the values of the noise spectrum on higher frequencies differ clearly from each other. According to [3], the noise voltage densities should be approximately the same on high frequencies and the noticeable difference can be seen at frequencies below 10 Hz. 50 Hz powerline interference and its multipliers can be clearly seen in the noise spectrum through Fig. 2 to Fig. 4.

Measurements made on the electrochemical noise of the electrode in various electrolytes indicate that the materials exhibit very different amount of electrochemical noise which also depends on the used electrolyte, as seen in Fig. 5. There is some discrepancy between the measurement results which depend partially on the preparation of the electrodes and partially of the environment where the measurements were conducted. The electrodes were allowed to stabilize in the electrolyte for one minute prior to the measurements but the electrolyte itself may have changed its homogeneity during the various measurements that were done. Stirring of the electrolyte between the measurements of different electrode materials could have been successful method to retain the homogeneity of the electrolyte. These measurements contain the electrochemical noise measurements done on four different materials of two different areas accompanied by the interface impedance measurements. All measurements show a common trend that suggests that the electrochemical noise diminishes with increasing DC conductivity of the electrolyte, refer to Table 1 and Fig. 5.

Electrochemical noise measurements, Fig. 5, indicate that at this selected frequency range of 0.5 – 500 Hz platinum shows the highest electrochemical noise in almost every case, gold the second greatest, then Ag/AgCl and finally stainless steel the least noise. The electrochemical noise of the electrodes of different area should, according to theory presented in [2] behave as inverse of the square root of ratio of the areas. Square root of ratios between the areas of the electrodes was 1.41 calculated with 6 mm^2 coverage of the insulating epoxy over the wires. The measurement results are in good agreement with the theory except for Ag/AgCl where the rms-noise ratios are higher than their area ratio would suggest.

The difference between the theoretical and experimental values can be caused by the imperfect AgCl-coating on the electrodes. The surface of the electrode may have been contaminated thus leading to imperfect AgCl-coating. The pure silver electrodes show higher electrochemical noise than coated ones, [5], which explains the results against theory in these measurements. The coating could be made more stabile if the coating time would be extended to several hours.

V. Conclusions

From the measurement results it can be seen that stainless steel electrodes exhibit the least electrochemical noise hence the most rapid stabilization time. This encourages use of stainless steel electrodes as biopotential measuring electrodes in implantable conditions. It would, actually, be even more useful to use stainless steel electrodes in surface biopotential measurements since according to our measurements, they seem to have even faster stabilization time than the Ag/AgCl electrodes have.

Stainless steel is not commonly used as electrode material but according to results gained in this research, its use as electrode material in implantable conditions is strongly encouraged.

The research also clearly shows that the effect the electrolyte on the electrochemical noise measurement is essential. Therefore, the noise measurements should be conducted in an electrolyte that is as close to the actual measurement environment as possible if absolute noise values are desired.

Acknowledgements

The research was supported by the Academy of Finland as a part of Future Electronics research program, project number 202758.

References

1. M. Fernández and R. Pallás-Areny, “Ag-AgCl electrode noise in high-resolution ECG-measurements”, Biomedical instrumentation & Technology, pp.125-130, March/April, 2000.
2. E. Huigen, A. Peper and C. A. Grimbergen, “Investigation into the origin of the noise of surface electrodes”, Medical & Biological Engineering & Computing, vol. 40, pp. 332-8, 2002.
3. L. A. Geddes and L. E. Baker, Principles of applied biomedical instrumentation, Wiley, 1975.
4. F. Moussy and D. J. Harrison, “Prevention of the rapid degradation of subcutaneously implanted Ag/AgCl reference electrodes using polymer coatings”, Anal Chem., Mar 1, 66(5), pp. 674-679, 1994.
5. J.Riistama and J. Lekkala, “Characteristic Properties of Implantable Ag/AgCl- and Pt-electrodes”, in Proc. 26th Annual IEEE Engineering in Medicine and Biology Society Conference, San Francisco, CA, USA, Sep. 1-5, 2004, p. 2360-2363.
6. R.W. Cahn (ed.), P. Hansen (ed.), E.J. Kramer (ed.), Materials Science and Technology – A Comprehensive Treatment – Vol. 14: Medical and Dental Materials, Vol. 14, Weinheim: VCH, 1992.
7. R. S. Jutley, D. E. Shepherd, D. W. Hukins and R. R. Jeffrey, “Preliminary evaluation of the Sternum Screw: a novel method for improved sternal closure to prevent dehiscence”, Cardiovascular Surgery, Vol. 11, No. 1, pp. 85–89, 2003.

Address of the corresponding author:

Author: Jarno Riistama
Institute: Institute of Measurement and Information Technology, Tampere University of Technology
Street: Korkeakoulunkatu 3
City: Tampere
Country: Finland
Email: jarno.riistama@tut.fi

Smart textiles for automotive: application to airbag development

Emilie Drean[1], Laurence Schacher[1], Dominique Adolphe[1], François Bauer[2]

[1] Laboratoire de Physique et Mécanique Textile FRE CNRS 2636, ENSISA

[2] Piezotech S.A.

Abstract— **The study presented here aims at developing a PVDF sensor that can be inserted within a seat-cover fabric. The final objective is to measure accurately the mass of a passenger sitting in a car. The measurement method consists in analyzing the response under compressive stress of a PVDF disc using the resonant frequency of the material. The maximal phase of the sensor at resonant frequency is taken up for each constraint applied. The influence of external parameters (vibrations, temperature variations) has been studied. The position of the sensor in the seat-cover was given starting from an ergonomic study. Tests in real case were carried out. A linear variation of the phase has been shown only taking into account the mass of the subjects' sitting position.**

Keywords— **seat-cover fabric; PVDF; resonant frequency; weight; automotive; airbag**

I. Problem statements

Smart textiles, ie textiles that are able to sense stimuli from the environment, to react to them and adapt them by integration of functionalities in the textile structure, came into force a few year ago. Such textiles are basically used for five functions, sensor, data processing, actuators, storage and/or communication [1].

Many instrumentation systems were developed by the automotive industry. But although the vehicles themselves become smarter and smarter, the level of integration of smart textiles is very low.

Early studies in automotive are oriented on security, and particularly on passive security. In fact, frontal accidents represent 42 % of the cases, 33 % of these are lethal. That's why first protection systems took care of frontal protection with driver side and passenger airbags. Real accidents analyses show that airbag alone increases survival of 32 % and airbag/seat-belt combination reduces injury hazard of 75 %.

Actually the airbag must inflate not only according to the shock violence, but also taking in account the passenger morphology and positioning. "Smart" airbags development is therefore of great interest, to avoid any risk of injury not due to the accident itself. One of the parameters most often considered, and for the measurement of which the developments are most advanced, is the mass of the passengers. If such devices already equip some vehicles models, they are not able to achieve a precise mass measurement.

The study presented here aims at developing a polyvinylidene fluoride (PVDF) sensor able to weigh car passengers and that can be inserted within a seat-cover fabric.

PVDF and its copolymers exhibit tailorable ferroelectric, piezoelectric and structural properties. PVDF polymer sensors are films able to provide nanosecond, time-resolved stress measurements, thanks to a patented polarization process, which confers reproducible properties to PVDF [2]. These film-sensors could directly be integrated in fabrics such as seat-cover, which are made of three layers.

II. Measurement method using a PVDF sensor

A. Car-seat description

A car seat is composed of different independent parts (Fig. 1).

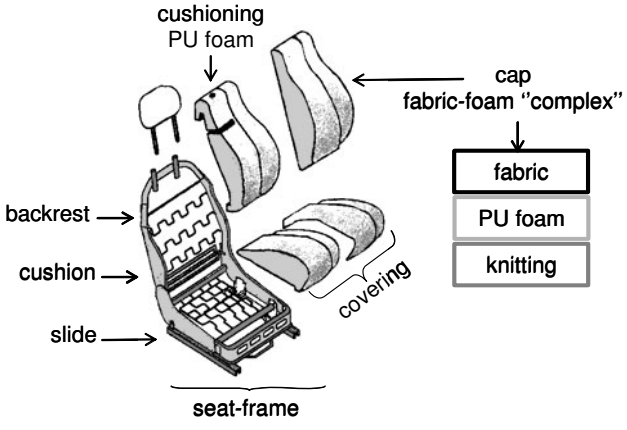


Fig. 1: Car-seat structure [3]

The first element of the seat is the frame, composed of a backrest and a cushion, slides for depth adjustment, and posture and adjustment mechanisms. On this frame, security devices, springs and wires are also fixed. The covering of the frame is divided in two parts, a polyurethane foam cushioning and a cap. The cap is a specific fabric called "complex" made of three layers: surface polyester fabric (weav-

ing, knitting, Alcantara®, leather, etc), polyurethane foam, and polyamide knitting which improves cap slip on the cushioning. This structure shows the interest of inserting sensors in the car seat; this part is directly in contact with car occupant.

B. The sensor: description and integration

PVDF is a semicrystalline polymer in which each monomer unit –(CH_2-CF_2)– has two dipole moments, one due to CF_2 and the other to CH_2. PVDF exhibits a variety of molecular conformations and crystal structures depending on the method of preparation. The most common and most studied of these are the and the phases. The phase is called the ferroelectric phase and is the most useful phase. Among ferroelectric polymers, -PVDF has the unusual property that the forces responsible for the ordering of dipoles in the ferroelectric phase are sufficiently strong that the polymer melts before it undergoes a ferroelectric to paraelectric phase transition. Copolymers of VDF and TrFE exhibit ferroelectric transitions below the melting point. They are useful in a variety of transducer and sensor applications [4].

The material used to design the sensor is a stacking of copolymer sheets of 25-µm thicknesses. Between copolymer layers, polymer sheets are inserted. Their melting point (60°C) is lower than the depolarization temperature of the copolymer (110°C). The copolymer layers are stuck by polymer using a heating press, without being depolarized. Final material is of 500-µm thicknesses.

The shape of the sensor is circular (2.5 cm²), 500 µm thick and gold sputtered on both sides (electrodes). Two thin wires soldered on each side of the sensor assume connection. The sensor has been inserted in the "complex" fabric between the external layer and the foam. The fabric and the foam are separated; the sensor is inserted and the two layers are linked again using a flexible adhesive.

C. Use of the sensor as a resonator

Measurement environment is really difficult. Considering sensor sensitivity, measurement will be disturbed by car internal electrical noises (embedded electronics, static electricity, wires…). For these reasons, the best-adapted measurement method was to use the sensor as a resonator. The method consists in analyzing the response under compressive stress using the resonant frequency of the material [5].

The sensor is connected with an impedance gain/phase analyzer, which measures frequency, phase and impedance in a frequency range of 100 Hz – 40 MHz. Raw measurements are adjusted thanks to a short circuit and an open circuit measurement, done in the same frequency range. The analyzer piezoelectrically excites the resonator, and the electrical impedance of the resonator is recorded. At the resonant frequency, the phase is maximal and the module presents an inflection point (Fig. 2).

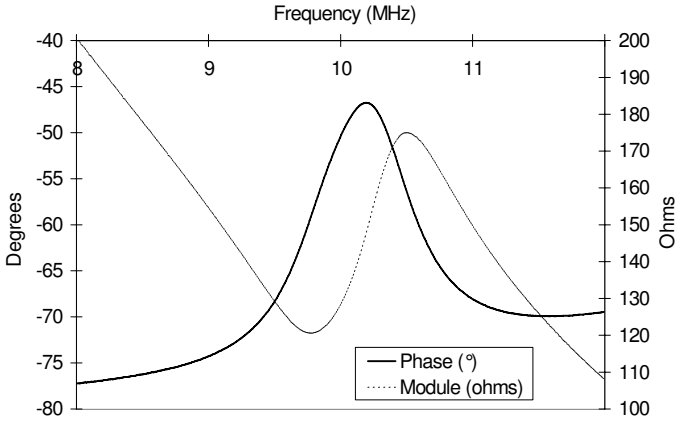


Fig. 2: Resonator electrical impedance versus frequency

The integrated resonator characterization has been achieved by using a counterweight parallelogram. This device allows applying a perfectly vertical force (Fig. 3).

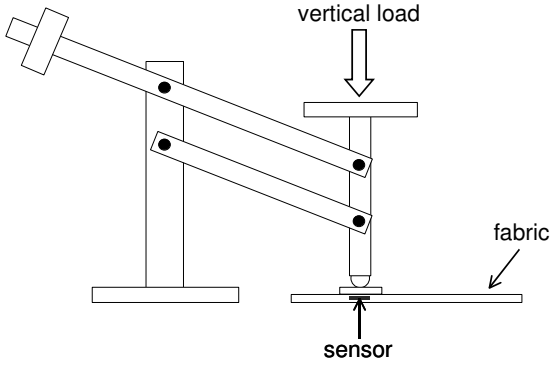


Fig. 3: Measurement device

A constraint series is applied to the fabric in the sensor area; the maximal phase at the sensor's resonant frequency is taken up for each one. An increase of the constraint induced a linear decrease of the phase (Fig. 4) [6].

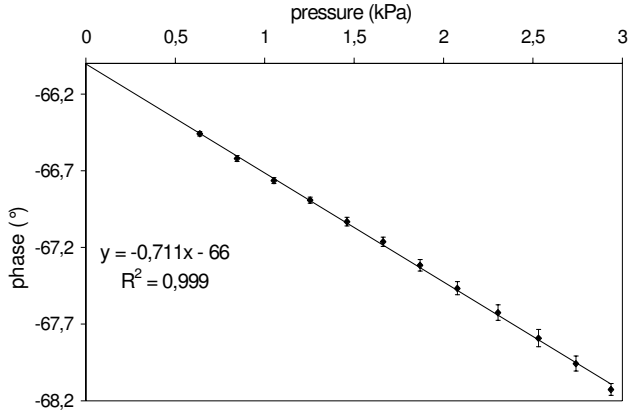


Fig. 4: Phase versus constraint for the textile integrated PVDF resonator

III. Influence of external parameters

A. Vibrations

A moving vehicle is subjected to vibrations coming from various sources, such as the engine, the tires on the road irregularities, the shock absorbers, or the electrical sources. The passenger must be weighed before the vehicle starts, when the passengers sit down in the vehicle. The sensor will only be subjected to vibrations caused by the engine. The vibration frequency is around 100 Hz, i.e. a frequency clearly lower than the resonant frequency of the resonator, which is about 2 MHz.

We have therefore considered that the vibrations related to the engine would have no influence on measurement.

B. Temperature variations

The temperature variation for automotive textiles is included between -30°C and +80°C. Material behavior in temperature is important to know as we use its mechanical properties (vibration) for measurement. In fact, a temperature decrease should induce a lower phase variation as material rigidity increases.

The whole assembly, i.e. the instrumented fabric and the counterweight device, have been placed inside a climatic chamber, which can create precise thermal conditions. The tests were carried out as previously, by application of a constraint series on the fabric between -20°C and +40°C. The range of temperature is less wide than the standard one used in the automobile industry. Indeed, the thermal loss at-30°C was too important; it was impossible to maintain a constant temperature. It is not possible to reach a higher temperature as the polymer binding the copolymer layers in the resonator melts at 60 C.

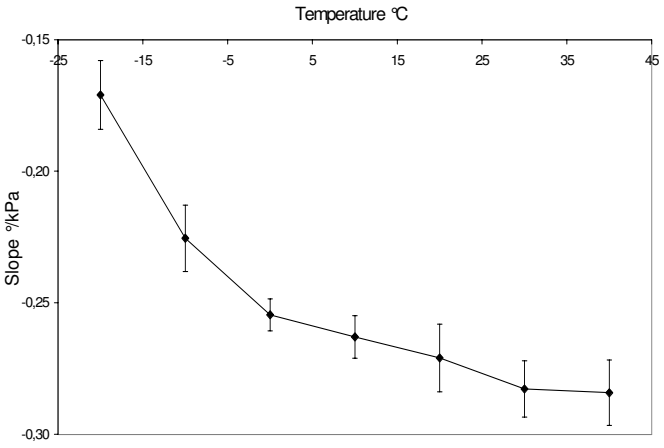


Fig. 5: Phase variation slopes versus temperature

Figure 5 presents the slopes values raised at the various considered temperatures. Each test was repeated ten times for each temperature.

The results are not constant with the temperature. The variation is much more important when the temperature decreases than when the temperature increases. The variation beyond 20°C seems not to be really important.

IV. Real case tests

A. Resonator positioning

The position of the sensor in the seat-cover was given starting from an ergonomic study.

We have first considered that the driving position did not have importance in our case, since measurement will not be taken while driving but when the passenger sits down in the vehicle. We decided not to place the resonator in a pressure peak area, i.e. in the pelvis area, because these peaks depend on the passenger's sitting position, and also on his morphology [7]. We have thus preferred to insert the resonator in the cushion area located under a thigh of the passenger.

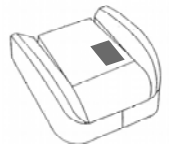


Fig. 6: Resonator position

B. Subjects data

The tests were carried out on fourteen subjects, eleven men and three women. The data are presented in Table 1, and are noted in italic for the female subjects.

Table 1: Subjects data

	standing mass (kg)	sitting mass (kg)	% sitting / standing	height (cm)	BMI
subject 1	*65*	*45*	*69*	*166*	*24*
subject 2	70	50	71	174	23
subject 3	69	49	71	183	21
subject 4	81	46	57	171	28
subject 5	75	54	72	170	26
subject 6	*64*	*41*	*64*	*160*	*25*
subject 7	*60*	*42*	*70*	*169*	*21*
subject 8	90	65	72	183	27
subject 9	70	50	71	184	21
subject 10	103	61	59	180	32
subject 11	70	49	70	182	21
subject 12	59	43	73	171	20
subject 13	55	45	82	170	19
subject 14	93	51	55	174	31
mean	**73**	**49**	**68**	**174**	**24**
standard deviation	**14**	**7**	**7**	**7**	**4**

The panel was selected to be as large as possible, from 55 to 103 kg. The mass and height were measured for each panelist. The body mass index (BMI) was calculated by the formula:

$$BMI = \frac{\text{mass (kg)}}{\text{height}^2\text{ (m)}} \quad (1)$$

To measure the mass in sitting position, we used a bathroom scale placed on a chair, and we asked the fourteen subjects to sit down above, the back leaning on the backrest and the feet leaning on the ground.

The mass in sitting position measurement does not give an exact value. We notice that in our case, the mass in sitting position represents around 70% of the total mass of the various subjects. However, according to the ergonomic studies, the load transmitted on the cushion of an automobile seat represents 40% of the total. This could be explained by the lack of support elements in comparison with a real car-seat placed in a vehicle, particularly on the backrest (in a car-seat, the backrest supports 20% of the total load transmitted to the seat) [8]. Although, the masses in sitting position we measured strongly depend on the subjects positioning on the scale, they give an order of magnitude. Moreover, it was requested from the subjects to try to preserve a positioning identical to this one for following measurements.

C. Measurement

Measurements were taken in the following way.

The car-seat cushion has been placed on a chair. The resonator, integrated in the cap, has been connected to the gain/phase analyzer (Fig. 7).

The phase value at rest is recorded; then, each subject sits down on the cushion, and the phase value under constraint is taken up. Measurement is carried out ten times for each subject.

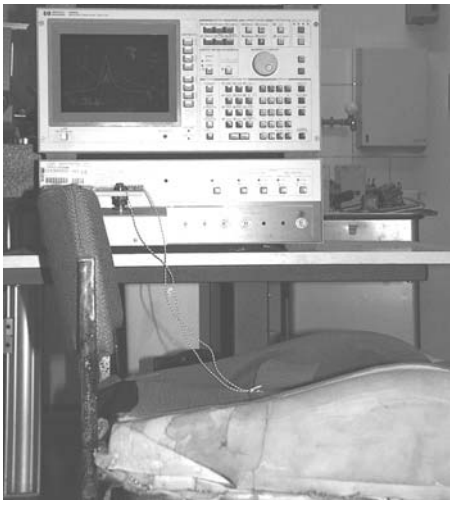


Fig. 7: Real case measurement method

D. Results and analysis

Initially, we thought that the phase variation could directly be linked with a pressure value. Calibration curves have been measured, and a slope in °/kPa is available for

each type of seat-cover fabric. The problem is that the body pressure distribution on the seat is not constant, and we cannot know which proportion of the total pressure is applied to the resonator. Thus, we cannot directly use the curves slopes to determine the mass of the subject.

We then research a correlation between the subjects' morphological characteristics and the phase variations. There is no correlation with the mass in standing position; the ratio standing mass/sitting mass is not constant, and there is thus no linearity between these two parameters. As the BMI also depends on the standing position mass of the subjects, this parameter is neither usable (Fig. 8).

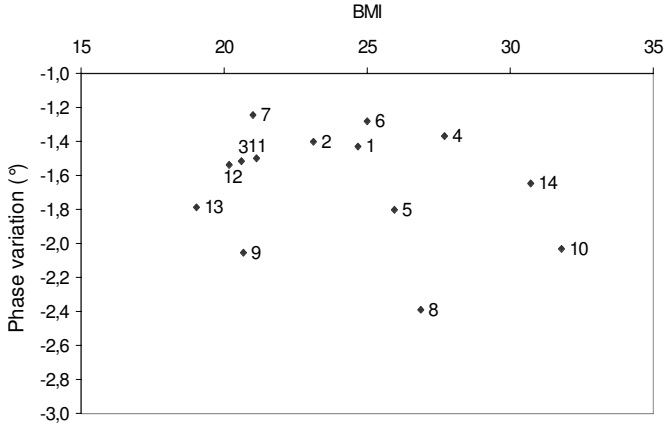


Fig. 8: Phase variation versus BMI of the 14 subjects

We focused the study on the mass in sitting position. Figure 9 shows the obtained results.

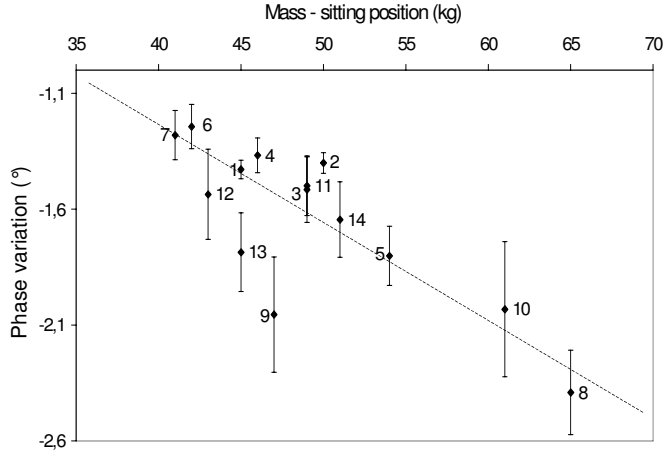


Fig. 9: Phase variation versus mass in sitting position

A linear trend can clearly be observed. This result is very interesting since a direct relation between the resonator phase variation and the passenger mass in sitting position is demonstrated. The interpretation of the information given by the resonator is thus more easily usable for the automobile application. Moreover, the mass in sitting position is more interesting than total weight since the bust movement has to be supported by the airbags.

Two subjects are particularly far away from the tendency slope. These subjects are the largest male subject, and the lightest male subject, which are both out of the average morphological criteria. In these precise cases an acurate mass would be difficult to measure.

However the automobile request is not to produce a bathroom scale, but to be able to classify the occupants of a vehicle according to ranges' of mass of about 20 kg. Thus, in spite of the system inaccuracies this objective is reached.

V. Conclusion and perspectives

The study was focused on the integration of a PVDF resonator in a seat-cover fabric, to obtain information about car occupant.

The influence of external parameters, such as vibrations or temperature variations, was first studied. We have estimated that the vibrations would not have an influence. On the other hand, we have shown that low temperatures (lower than 10°C) induce a material decrease of the resonator phase variation under constraint.

The position of the sensor in the cap of the seat was given starting from an ergonomic study. Tests in “real” case were then carried out. Taking only the mass of the subjects' sitting position into account, a linear variation of the phase has been shown.

First results are very interesting, but many improvements will be necessary to continue the study. The work presented here only proves feasibility.

It will initially be necessary to undertake a new “real” case study, by using a complete automobile seat placed in a vehicle, in order to reproduce the occupants' real sitting positions. Measurement will thus be more reproducible because the subjects will be “constrained” because of the door and the wheel (or the dashboard).

Then, the use of a proper full car seat instead of a simple cushion would more constraint the passenger position as the backrest has side supports. Moreover, its height (higher than the backrest of a chair for example) ensures an important support, which reduces the total load applied to the cushion.

We can then imagine extending the use of the “complex” instrumented to other automotive applications, such as for example the detection of driver vigilance failure. Significant factors of hypovigilance are a decrease in the respiratory rate and an increase of its regularity. The sensor should

therefore be able to measure, record and analyze the driver breathing to prevent drowsiness while driving.

REFERENCES

[1] VAN LANGENHOVE L., HERTLEER C., Smart textiles in vehicles: a foresight, Journal of Textiles and Apparel. Technology and Management., Volume 3, Issue 4, Winter 2004, 6 pages (2004)

[2] BAUER F., Method and device for polarizing ferroelectric materials, US Patent 4, 611, 260 (1986)

[3] PRUDENT O., THIONNET A., GOURGUECHON O., PAJON M., BAKACHA M., RENARD J., Simulation numérique du confort de sièges d'automobiles: comportement mécanique 3D de mousses de polyuréthane; optimisation 2D d'un profil simplifié de sièges, Mec. Ind., Vol. 1, pp 511-520 (2000)

[4] BAUER F., GRAHAM R.A., Very high pressure behavior of precisely-poled PVDF, Ferroelectrics, Vol. 171, pp 497-518 (1995)

[5] NGAMALOU, L., BENECH, P., CHAMBEROD, E., Pressure measurement with resonant PVDF or copolymer, Ferroelectrics, Vol. 171, N°1-4, pp 217-224 (1995)

[6] DREAN E., SCHACHER L., BAUER F., ADOLPHE D., "A smart sensor for induced stresses measurement in automotive textiles", The Fall 2005 Meeting of the Fiber Society, 17-19 October 2005, Newark, NJ, USA, pp 87-89

[7] YANG B.J., CHEN C.F., LIN Y.H., LIEN I.N., Pressure measurement on the ischial tuberosity of the human body in sitting position and evaluation of the pressure relieving effect of various cushions, Journal of Formosan Medical Association, N°83, pp 682-698 (1984)

[8] ANDREONI G., SANTAMBROGIO G.C., RABUFFETTI M., PEDOTTI A., Method for the analysis of posture and interface pressure of car drivers, Applied Ergonomics, Vol. 33, pp 511-522 (2002)

Available: http://www.afpyro.org (reference_available)
http://www.Crash-test.org (reference_available)

Address of the corresponding author:

Laboratoire de Physique et Mécanique Textile FRE CNRS 2636
ENSISA
11 rue Alfred Werner
F-68093 Mulhouse cedex
Tel: +33-(0)389336320
Fax: +33-(0)389336339
Email: emilie.drean@uha.fr

Celeritas – A Wearable Sensor System for Interactive Digital Dance Theatre

B. O'Flynn[1], G Torre[2], M. Fernstrom[2], T. Winkler[3], A. Lynch[1], J. Barton[1], P.Angove[1], S. C O'Mathuna[1]

[1] Ambient Electronic Systems, Tyndall National Institute, Cork, Ireland
[2] Interaction Design Centre, University of Limerick, Limerick, Ireland
[3] Brown University, Providence, RI, USA

Abstract— Celeritas is an artistic/scientific collaboration between the Tyndall National Institute (Cork), the Interaction Design Centre in Limerick, Cindy Cummings (Dance Artist, Cork) and Todd Winkler (Composer and Digital Artist, Brown University, USA). Research Teams at the Tyndall Institute are developing wireless sensor network nodes, also known as motes, and associated miniaturized sensors. Motes can be applied in many different domains, ranging from medical and environmental monitoring to everyday applications in ubiquitous computing. This project aims to apply Tyndall's sensor system to create a wireless dance costume for audio/visual performance using inertial sensor monitoring technology.

Dancers could be regarded as experts on human movement, producing accurate and expressive actions that provide a rich testing ground for human-computer interaction. The collaboration will push the boundaries of both artistic practice and wearable mote technology, as we will adapt and apply the Tyndall mote platform in a prototype body suit embedded with sensors. Software developed by the Interaction Design Centre and Todd Winkler will then convert the movement information detected by the sensors into computer generated sounds and processed video images. This mapping allows the dancer (Cummings) to fuse aspects of the physical body with the extended possibilities of the electronic body.

This paper presents the hardware platform that has been developed for the Celeritas project. The system is based around the Tyndall 25mm Wireless Inertial Measurement Unit (WIMU) node. The WIMU system is designed for integration into a body suit, which is to be worn by the dancer, whose movements are extracted from the wearable network of sensors and processed by a high-level software system that connects to the dancer wirelessly.

Keywords— Wireless Sensor System, Interactive Dance, Inertial Measurement Unit, IMU, Wearable Computing

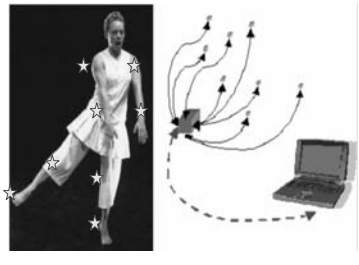


Figure 1. CELERITAS Concept

I. Introduction

Performances involving real-time signal processing are becoming increasingly important in contemporary art [1]. This approach has mainly been successfully used for multidisciplinary works, where a stream of sensor data is generated by an interactive dance performer, who has the possibility to simultaneously take the role of composer and choreographer of his/her own work. Research is being carried out world wide in the area of interactive electronics development, for instance the responsive environment group in MIT Media Lab [2]. This work reported on systems using magnetic sensors for Tangible Music Interfaces [3] and processing algorithms for large groups of dancers using wireless networks [4] as well as inertial sensing systems for dance applications [5], [6].

CELERITAS is a collaborative project combining the wireless sensor technology of Tyndall with the software processing capabilities developed by Todd Winkler at Brown University, using the Max/MSP/Jitter platform [7] to create original music and video processing in response to dance as well as the Interaction Design Centre in University of Limerick. Together they collaborated in the development of an interactive dance exhibition incorporating state of the art hardware and software developments from the research institutes.

The work is focussed on the development of an inertial sensing mechanism to be worn by a contemporary dancer whose main area of interest is the merging of contemporary dance with modern day technology. After eight years of work with a vision-based system [8], Winkler and Cummings felt that experimenting with a wearable system could open up new ideas for digital dance theatre. The proposed system is designed around the Tyndall 25mm WIMU node, which is a wirelessly enabled 6 Degree of Freedom (DOF) Inertial Measurement Unit (IMU) [9]. The system consists of six WIMU modules distributed down the trunk and at the extremities of the dancer. Each of the sensors transmits its data to a base station, using a customised address driven Master/Slave request/reply protocol, for data harvesting and processing. The concept of the CELERITAS system can be seen in Figure 1, where one can see how IMUs could be distributed around the dancer's body.

II. HARDWARE DEVELOPMENT

The 25mm WIMU has been designed based upon Tyndall's 25mm modular wireless sensor node [10] technology. The 25mm wireless node has been used to develop a platform for low volume prototyping and research in the wireless sensor network domain. It has been developed for use as a platform for sensing and actuating, for use in scalable, reconfigurable distributed autonomous sensing networks in a number of research projects currently underway at the institute [11].

The modular nature of the Tyndall hardware lends itself to the development of numerous layers for use in various application scenarios. Layers can be combined in an innovative plug and play fashion and include communication, processing, sensing and power supply layers.

The 25mm Wireless node has an integrated ATMEL ATMega128 [12] microcontroller for the networking of the modules. This feature coupled with the 2.4GHz transceiver, RF Nordic nRF2401 [13], produces a very powerful customisable wireless node. Alternative implementations enable Zigbee (IEEE 802.15.4) communications in the 25mm form factor if required. An FPGA layer can be integrated into the system where high-speed DSP processing is required such as in the implementation of Kalman filtering for conditioning of inertial measurement data.

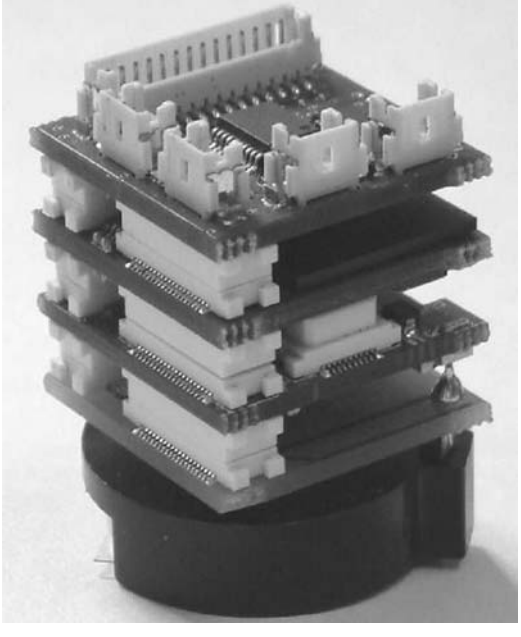


Figure 2. 25mm Stackable System Layers

The stackable configuration (Figure 2) enables ease of connectivity between layers depending on the system level requirements and deployment scenarios.

A. 25MM WIMU

The 25mm WIMU is a 6 DOF IMU module, which is made up of an array of inertial sensors coupled with a high resolution Analog to Digital converter (ADC). The 25mm WIMU module utilises the wireless communication capabilities of the 25mm Wireless Node to realise a fully autonomous WIMU module. The 25mm WIMU can be seen in Figure 3.

The inertial sensor array that makes up the 25mm IMU consists of three single axis gyroscopes, ADXRS150 [14] from Analog Devices, two dual axis accelerometers, ADXL202 [15] from Analog Devices, and two dual axis magnetometers, HMC1052L [16] from Honeywell. The sensor array was designed with a novel 3D structure, which produces the 6 DOF functionality.

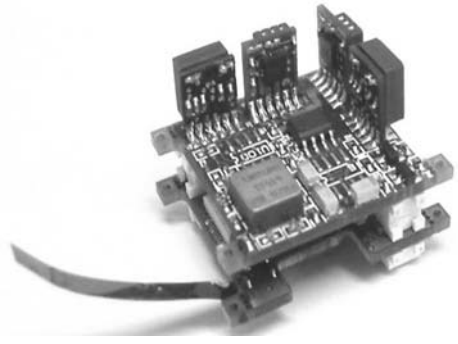


Figure 3. 25mm WIMU module

The module also has a 12-bit ADC chip, AD7490 [17] from Analog Devices, handling the data conversion. This chip has a Serial Peripheral Interface (SPI), which allows easy interfacing to the ATMEL microcontroller. The 3D structure was implemented using a motherboard/daughterboard configuration. Miniature slots were drilled in the motherboard to accept the daughterboard at 90°. The motherboard and daughterboard were designed so that their respective pads would match accurately and allow for a 90° solder connection. The motherboard was configured so the sensors would make up the relevant array to generate the 6 degrees of freedom required for the module. The interconnection method can be seen in Figure 4.

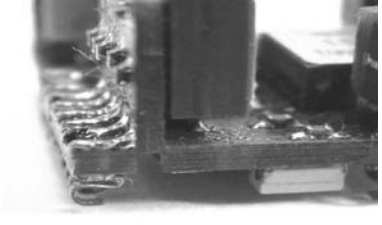


Figure 4. 90° Integration Technique

III. INTERACTIVE SYSTEM DEVELOPMENT

Based on the "Vitruvian Man", the proposed system being developed divides the space around the human figure, according to the ratio of his/her body, an approach similar to the one used by Leonardo Da Vinci in the sketch of the same name.

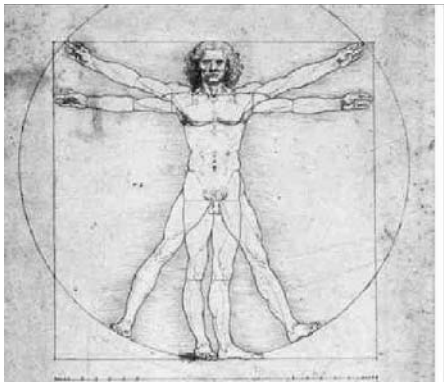


Figure 5. Leonardo da Vincis "Vitruvian Man"

The main goal of the system under development is to create a virtual 3-D musical instrument around the body of the performer. The platform is used to create an improvisatory dance performance in which the dancer can create in real-time the music to which he or she is dancing through the mapping of the performers' movement to the music being generated. One of the main advantages of using a wireless data transceiver system in such a manner is that it can enable an interactive performance between performers. The final result can then be a true polyphony created by as many independent instruments/performer as are participating in the performance.

Six motes are used in the system and they are distributed around the body of the performer as follows, Figure 6:

Mote_left_leg attached to ankle of the left leg
Mote_right_leg attached to ankle of the right leg
Mote_baricenter attached to middle of the chest
Mote_left_arm attached to wrist of the left arm
Mote_right_arm attached to wrist of the right arm
Mote_head attached to on the top of the head

A preliminary step in mapping the movement of the performer is the creation of a virtual spherical space around the body of the performer. To set initial system attributes, some of the principles described in Leonardo da Vinci notes to design his Vitruvian Man are taken into account. In his notes, Leonardo uses the height to deduce all the other proportions of the human body, setting the genital area at the middle of a square circumscribing it, thereby also representing the middle point of the human body. The circle is gained using the distance between the navel and the feet as radius. The navel location is calculated through several geometrical steps, but essentially the navel (middle point) is calculated by dividing the distance between feet and straight arms above the head by two.

In software terms, the mote_baricenter is considered as the centre of the sphere. The ray is empirically measured by asking the performer to assume a "calibration" position similar to the one shown in Figure 6. The inertial measurements from Tyndall's 25mm WIMU system can then develop the performance according to the dancer's movements.

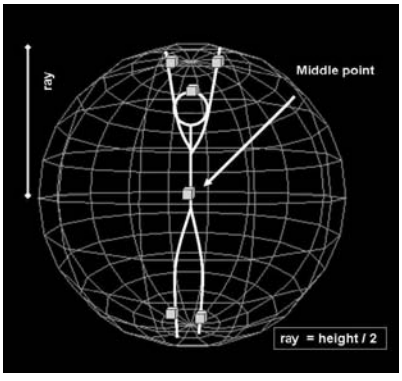


Figure 6. "Celeritas Man"

The six WIMU modules are to be deployed on the dancer's body in the final wearable implementation, placed at the ankles, chest, and wrists as described. The suit will have integrated 25mm connectors placed, on flexible substrates, at the specified locations. To enable reconfiguration and if necessary replacement of the nodes, each WIMU system has its own power supply, a low profile rechargeable lithium ion battery pack weighing approximately 20 grams and capable of providing 740mA/hr of energy.

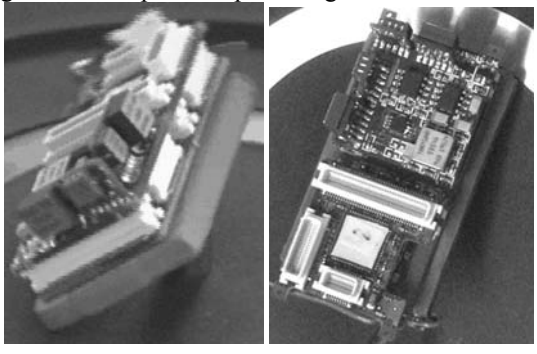


Figure 7. Tyndall 25mm wearable WIMU system

This configuration allows for easy replacement of nodes in the event of a node failure. A customised networking protocol connects the WIMU network with a base station.

IV. RF PROTOCOL DEVELOPMENT

A. Radio Transmission Mechanism

The platform uses the nRF2401 single chip 2.4 GHz transceiver from Nordic [13]. The transceiver has two active modes: direct and shockburst. In direct mode, the radio is essentially fully controlled from external hardware, while in shockburst mode several tasks are pushed into the radio hardware. This enables power saving within the system to reduce the frequency of replacing/recharging the batteries.

The shockburst mode uses an on-chip FIFO to clock in data at possibly low data rates and transmit at a very high rate. By transmitting faster, the radio can quickly return to a power saving mode and therefore energy consumption is reduced. The radio packet format of the transceiver, when in shockburst mode, is shown in Figure 8. The packet has a fixed size and the transceiver automatically adds the preamble and CRC fields releasing the microcontroller of these tasks, enabling the system to be run by a low power consumption, low cost microcontroller, and enabling the microcontroller to run at slow speed (again for power consumption considerations).

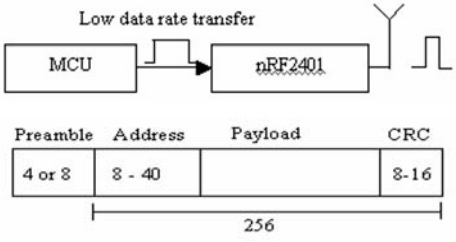


Figure 8. Shockburst mode operation of the nRF2401

When a packet is received, the radio checks the address and CRC fields. If the packet is addressed to a different radio unit or is corrupted then it is dropped. Only undamaged packets addressed to the unit generate interrupts to the microcontroller.

The transceiver is capable of transmitting the data at 1Mega Bit per Second using Gaussian Minimum Shift Keying (GMSK) Modulation and Demodulation, in the 2.4 GHz ISM radio band which is unlicensed worldwide.

B. Body Area Network Protocol

The data is harvested from the dancer using a master / slave style address driven protocol. Each of the (slave) nodes distributed on the dancer's body is prescribed with a node address. The base station unit consists of a transceiver node with a UART output connected to a serial port on the processing PC. The processing unit (PC) is connected to the base station 25mm receiver node via a serial cable.

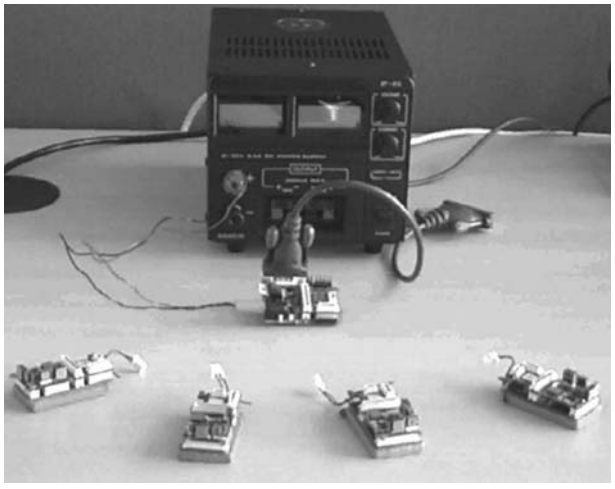


Figure 9. Master/Slave Sensor Network

On startup, a broadcast is transmitted to all slaves in the system, figure 9, to check for faulty nodes (due to low batteries for instance), so that they can be removed from the system. The master transmitter unit queries the slaves sequentially. This unit has a dedicated sequence and can be placed anywhere in the dance environment. Each of the slaves processes this request and only the queried slave replies. On receipt of a valid address, the sensor node replies with its most up to date valid data. Each of the other slaves listens to the data that is being replied and will switch back to their listening state when they receive the end of data sequence embedded in the replying slave's data.

The protocol is under characterisation and evaluation prior to deployment in the envisaged interactive dance environment. Once the protocol is verified, the time required to harvest the information, from a single node, will be optimised as to maximise the data throughput that can be achieved when requesting data from the nodes. This is a very important consideration for a real-time performance of this nature. The effect the inherent latency has on the system is currently being analysed in preliminary field tests. The system latency is governed by the time taken for the master to retrieve the information from the body nodes. With four nodes connected the smallest interval between two data packages recorded is 15 ms which we currently take as our sampling rate. As typical body movements operate in the 10's of Hz range this meets the preliminary requirements of the dance scenario.

In order to minimise this latency in the system we used a high speed (20MHz) crystal to operate the microcontroller system clock. This had the effect of noticeably speeding up the operation of the system despite the fact that the microcontroller is not specified to operate with such a high frequency clock

V. INTERACTIVE DANCE DEVELOPMENT

Celeritas is work for digital dance/theatre, with the body suit under development enabling a solo dancer to effect surround-sound, multiple video projections, and 3-D environments. In turn, the sonic results and processed video streams of the dancer influence choreographic decisions and kinaesthetic response, creating a dynamic three-way interaction that opens up new possibilities to explore the body as an agent for technological transformation. The sensing system is, in effect, an extension of the physical body into virtual (computer-generated) worlds.

Thematically, various sections of Celeritas explore concepts of speed and time, artistic interpretations of ideas suggested by experimental physics. (The speed of light in a vacuum is represented by the letter c from the Latin celeritas – swiftness). As the Tyndall sensors (literally) report speed from accelerometers and gyroscopes, that data is mapped in software to develop temporal transformations in both digital audio and in the processed video stream of the dancer using techniques such as time delays, granular synthesis, fragmentation, filters, and modulation.

VI. DISCUSSIONS AND FUTURE WORK

This paper presented the hardware platform developed for the realisation of a Body Sensor Network for digital dance theatre. The system being developed is based upon Tyndall's 25mm Wireless Inertial Measurement Unit node. The characterisation tests indicate that the unit is operating within acceptable limits of the required specification. The inertial measurement system is currently being integrated to a network scenario using a customised network protocol at the University of Limerick's Interaction Design Centre and will ultimately be integrated into a custom-designed wearable suit. The full hardware system is to be merged with Winkler's software processing environment for live performance, developed at Brown University.

ACKNOWLEDGEMENTS

The authors would like to thank all the members of the Ambient Electronic Systems team who have contributed to the project to date. We would like to extend our thanks to Kieran Harney, Analog devices Micromachining Division and Michael Coln, SAR ADC Group Analog Devices, for the provision of devices for this work. We would like to thank the European Commission for funding the MULTIPLEYE project in the IST programme (IST-2001 35108) and the rest of the MULTIPLEYE consortium, the Irish National Arts council for their support, and we would like to acknowledge the support of Enterprise Ireland in funding the D-SYSTEMS project (ATRP-02/227A). Thanks also to Science Foundation Ireland for funding the Tyndall National Access Program (NAP) under which some of this work was carried out.

REFERENCES

1. DIEM Digital Dance System: http://hjem.get2net.dk/diem/products.h
2. MIT: http://www.media.mit.edu/resenv/
3. Paradiso J. A., Hsiao K., and Benbasat A., Tangible Music Interfaces Using Passive Magnetic Tags, 2001 , ACM CHI Conference - Special Workshop on New Interfaces for Musical Expression.
4. Feldmeier M., and Paradiso J. A., Ultra Low Cost Wireless Sensors for Musical Interaction for very large groups, 2001, UbicompWorkshop on Designing Ubiquitous Computing Games
5. Sensemble: A Wireless, Compact, Multi-User Sensor System for Interactive Dance, Ryan Aylward, and Joseph A. Paradiso, in Proc. of NIME 2006, The 2006 International Conference on New Interfaces for Musical Expression, Paris, France, June 4-8, 2006, pp. 134-139.
6. Design and Implementation of Expressive Footwear." J. Paradiso, K. Hsiao, A. Benbasat, Z. Teegarden, IBM Systems Journal, Volume 39, Nos. 3 & 4, October 2000, pp. 511-529.
7. http://www.cycling74.com/products/maxmsp
8. Winkler T., Movement-Activated Sound and Video Processing for Multimedia Dance/Theatre. Proc' 2003 International Computer Music Conference.(San Francisco,: Computer Music Association, '03).
9. Barton J., Lynch A., Bellis S., O'Flynn B., Murphy F., Delaney K., O'Mathuna S. C., Repetto P., R Finizio, Carvignese C.; andLiotti L., "Miniaturised Inertial Measurement Units (IMU) for Wireless Sensor Networks and Novel Display Interfaces", In Proceedings ECTC 2005, May 31- June 3, 2005, 1402-1406
10. O'Flynn B, Bellis S, Mahmood K, Morris M, Duffy G, Delaney K, and O'Mathuna C, A 3-D Miniaturised Programmable Transceiver, Microelectronics International, Volume 22, No 2, (Feb.2005), 8-12,
11. Ambient Systems Group, Tyndall National Institute, University College Cork, Ireland. www.tyndall.ie/research/mai-group/index.html
12. Atmel Website ATMEGA Product Information, www.atmel.com/dyn/products/product_card.asp?part_id=2018
13. NordicVLSI Website nRF2401 Product Information, www.nvlsi.no/index.cfm?obj=product&act=display&pro=64
14. Analog Devices Website ADXRS150 Product Information, www.analog.com/en/prod/0%2C2877%2CADXRS150%2C00.html
15. Analog Devices Website ADXL202 Product Information, www.analog.com/en/prod/0%2C2877%2CADXl202%2C00.html
16. Solid State Electronics Centre, Honeywell HMC1052L, http://www.ssec.honeywell.com/magnetic/products.html
17. Analog Devices Website AD7490 Product Information, www.analog.com/en/prod/0%2C2877%2CAD7490%2C00.html

Address of the corresponding author:

Author:	Brendan O'Flynn
Institute:	Tyndall National Institute
Street:	Lee Maltings
City:	Cork
Country:	Ireland
Email:	brendan.oflynn@tyndall.ie

The Use of Telemetry-Evoked Compound Action Potentials (TECAP) in Cochlear Implantation

J. Ilgner, W.H. Döring, M. Westhofen

Department of Otorhinolaryngology, Head and Neck Surgery, University of Aachen RWTH, Aachen, Germany

Abstract— **For about 20 years, Cochlear Implants have successfully restored hearing in postlingual deaf or helped to acquire auditory communication in prelingual deaf patients. However, only in recent years Cochlear Implant manufacturers have implemented stimulating and recording protocols for evoked compound action potentials in order to assess auditory nerve function in relation to the implanted device. Along with the principles of stimulation and recording of TECAPs, two cases are presented which highlight their diagnostic value as well as their limitations. While the pooled data obtained from patients are yet too variable to serve as a predictor for individual stimulation strategies, TECAPs are highly useful in intra-individual follow-up of patients in whom Cochlear Implant stimulation currents may vary even after years and who require re-adjustment according to objectively registered neural responses.**

Keywords— **Cochlear Implant, Compound action potential, Inner Ear, Hearing, Rehabilitation**

I. Introduction

About 30 years ago, Cochlear Implants (CIs) were developed to electrically stimulate the auditory nerve from an intracochlear site by means of an external speech processor and a subcutaneously implanted receiver unit in adult patients with profound sensory deafness. Since the first devices implanted by House and co-workers were of single-electrode design and predominantly experimental, resulting in mere sound perception [1], the technology had entered clinical routine in the mid-1980's with further advances in speech coding and neural stimulation. From then on, open speech recognition from cochlear implants with and without lip reading became a more and more common outcome. Given these positive results in adult patients with postlingual deafness, it did not take until the beginning of the 1990's before children who were deaf at birth or before language acquisition were implanted routinely [2,3]. Although it is naturally difficult in children before pre-school age to assess hearing and speech perception, the results were equally encouraging, while it is now generally accepted that the earlier deafness is treated with a cochlear implant, the better the chances for hearing and speech acquisition are. Maturing of central auditory pathways allows successful implantation approximately to the age of 7 to 8 years. Nowadays, more than 60,000 cochlear implant wearers live worldwide [4,5], making this modality of treating profoundly deaf patients of all ages a major economic factor in health care.

As success rates in terms of speech acquisition and recognition rise continuously, the indication for implantation is expanding: Papsin [4] could show from a series of 103 children of inner ear anomalies that only those with a cochlear deformity ("common cavity") and those with a narrowing of the inner ear canal showed poorer speech acquisition results. However, these groups constituted only 19 of 103 patients, therefore anatomic malformation as such cannot be judged as contraindication against surgery [5,6]. Also former chronic inflammatory middle ear disease, which can be a cause for sensory hearing loss especially in adult patients, is still compatible with consecutive implantation, as long as the former inflammatory disease is eradicated [6,7,8].

On the other end of the patient spectrum, cochlear implantation is not exclusively reserved for patients with entire sensory deafness. In some cases cochlear implants can successfully be combined with conventional hearing aids, as long as there is residual hearing for lower frequencies. In these patients, incomplete insertion of the electrode stimulates basal sections of the cochlea, i.e. the upper and middle frequency range, while sensory hearing for lower frequencies near the cochlear apex are preserved. Furthermore it could be shown that due to the peripheral characteristics of auditory pathway stimulation, directional hearing can be restituted by bilateral cochlear implantation [9]. However, given the economic impact of bilateral implantation, audiologists and surgeons are required to assess the benefit of bilateral in contrast to unilateral implantation preoperatively as best as possible.

Currently, apart from refining speech encoding strategies and neural stimulation patterns, efforts of manufacturers and researchers head towards further improvement of hearing quality which is not only directed to verbal communication but also towards recognition and even enjoyment of music [10]. On the other hand, surgeons address minimization of insertion trauma, thereby improving implantation quality under the aspect of maintaining vestibular function in bilaterally implanted patients or patients with residual hearing [11,12]. As the risk of bacterial meningitis is increased in implanted patients in relation to the healthy population for

reasons which are yet not fully known, minimization of intraoperative trauma may play a key role to keep the rate of unwanted sequelae as low as possible [13].

It may be surprising, however, that visual confirmation under the operation microscope and the measurement of electrode impedances plus postoperative x-ray studies have been the only means of assessing cochlear implant contact in relation to host structures, i.e. the cochlear modiolus for most of the recent years. Only about 4 years ago, telemetry-evoked compound action potentials (TECAP) were incorporated in three major manufacturers' (Cochlear, Advanced Bionics and MED-EL) CI programming and testing software [14,15].

This article focuses on the potential of TECAP registration intraoperatively and postoperatively and their relevance for intraoperative CI function assessment and postoperative follow-up, especially in children.

II. Patients, material and methods

TECAP measurements are performed using individual hardware and software material provided by three Cochlear Implant manufacturers. These are

Cochlear Inc., Lane Cove, Australia (product name "Nucleus", TECAPs referred to as "Neural Response Telemetry" (NRT))

Advanced Bionics Inc., Sylmar CA, USA, (product name "Clarion", TECAPs referred to as "Neural Response Imaging" (NRI))

MED-EL GmbH, Innsbruck, Austria (product name "Pulsar", TECAPs referred to as "Auditory Response Telemetry" (ART)).

Given the individual hardware layout and stimulation / measurement strategies, stimulation values are expressed in individual "current units". As these units are chosen arbitrarily by each manufacturer, the amplitudes of the responding signal [μV] cannot be compared between implant products of different manufacturers.

The outline of a principal circuit diagram for measuring TECAPs is given in Fig. 1. Although the number of electrode contacts placed along the silicone sheath varies between 16 and 27, depending on the manufacturer, the stimulating and the recording electrode are generally chosen over a distance of two, thereby achieving a compromise between localization specificity and the ability to elicit a suitable signal from the stimulus (Fig. 2).

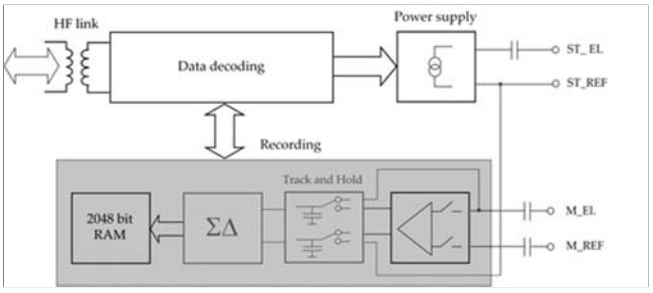


Fig. 1: Principal circuit diagram for measuring TECAPs (MED-EL Pulsar CI 100, MED-EL, Innsbruck, Austria)

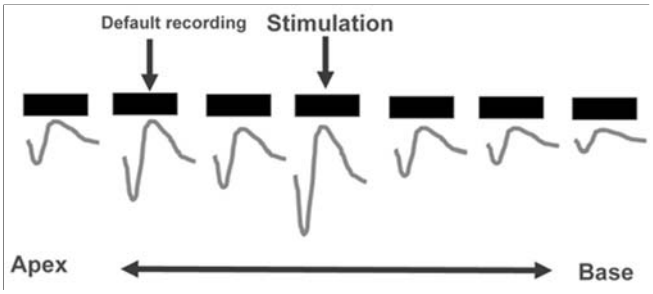


Fig. 2: Spread between stimulating and recording signal over a distance of two electrode contacts (default setting) (Advanced Bionics Inc., Sylmar, CA, USA). "Apex" and "Base" refer to the position of electrode contacts in relation to the widest part of the human cochlea (Base), which represents high frequency perception, whereas the narrowest part (Apex) represents low frequencies

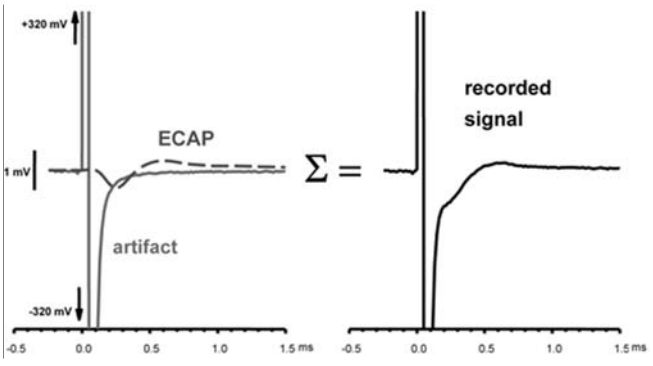


Fig. 3: Summation of stimulus potential and evoked compound action potential (ECAP) of the auditory nerve

Due to the magnitude of the stimulus (Fig.3), various filtering strategies (e.g. forward masking) are applied to elicit the evoked compound action potential (ECAP) of the auditory nerve. In Fig. 4a and 4b, two typical response curves are shown which are based on two different stimulation strategies provided by one manufacturer (Cochlear Inc.).

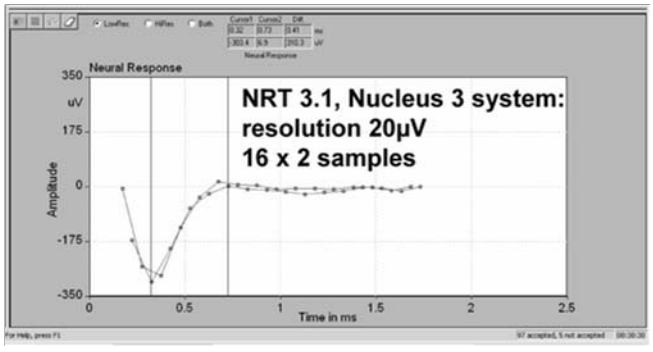


Fig. 4a: TECAP response curve resulting from 2 passes of 16 samples each; amplitude resolution 20µV (Nucleus "3" System, Cochlear Inc.)

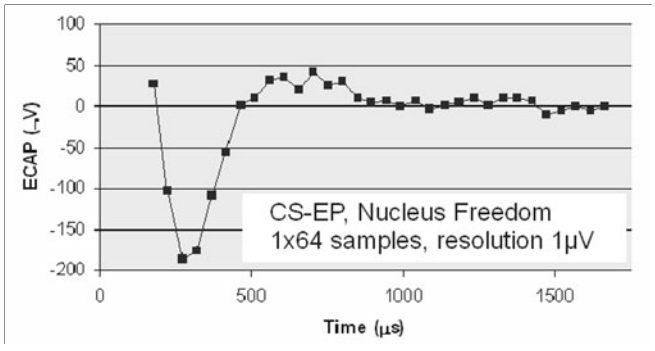


Fig. 4b: TECAP response curve, 1 pass with 64 samples; amplitude resolution 1µV (Nucleus "Freedom" System, Cochlear Inc.)

With the TECAP hardware and software setup provided by above mentioned manufacturers, we were able to examine 37 patients over 24 months intraoperatively and postoperatively. 20/37 patients received a Cochlear Implant unilaterally, while 14/37 patients received a second Cochlear Implant contralaterally after having been implanted monaurally in an earlier operation. 3/37 patients were implanted bilaterally in one operation, due to acute deafness following bacterial meningitis (2/37) or Usher's syndrome (hereditary degeneration of the inner ear hair cells and the retina). 14/37 patients were children who suffered from prelingual deafness.

III. Results

Following successful Cochlear Implantation, measuring electric impedance values plus registration of TECAPs for all electrodes requires 20 to 25 minutes. Regular TECAP responses could be elicited intraoperatively in 35/37 patients (94.6%). Fig. 5 shows a typical diagram representing growth function of TECAP responses with increasing stimulation currents.

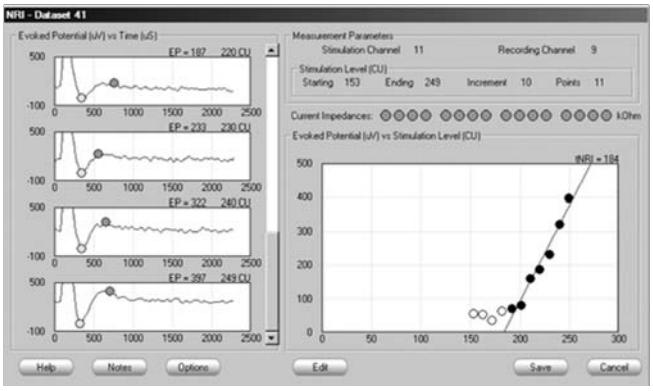


Fig. 5: Normal TECAP response curves (left) and growth function diagram (right). (Stimulating electrode #11, recording electrode #9) (Clarion HiRes 90K, Advanced Bionics Inc.)

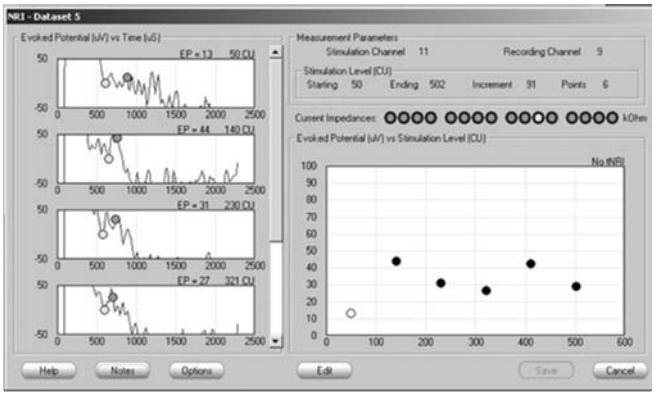


Fig. 6: No TECAP response from electrode #11 (recording electrode #9). Cochlear Implant model and manufacturer as above (Case 1)

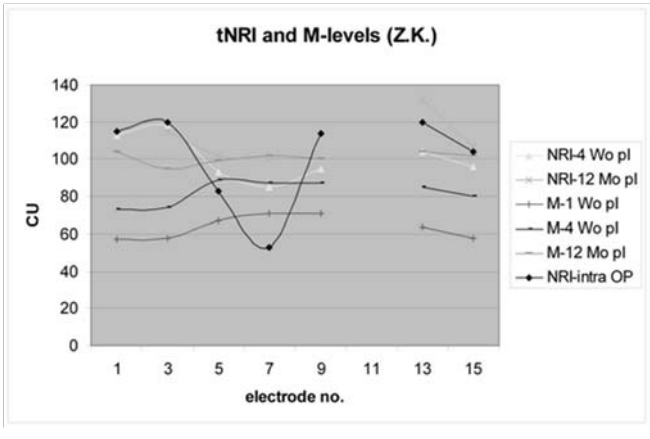


Fig. 7: Change of NRI-levels (intraoperatively, 4 and 12 weeks postoperatively) and M-levels (1, 4 weeks and 12 months postoperatively) with electrode #11 switched off (Case 1)

A. Case 1:

A 12-year-old boy (Z.K.) with gradually progressive sensory hearing loss and subsequent perilingual deafness had been implanted unilaterally at 6 years of age. Due to Cochlear Implant failure following trauma to the temporal and parietal bone, the Implant was exchanged at 12 years of age. Intraoperative TECAP recording revealed no response from electrode #11 (recording electrode #9) (Fig.6). All other electrodes, however, showed regular responses. Although the electrode was moved slightly following first TECAP registration during the operation, the result remained the same in intraoperative re-testing. When his Cochlear Implant was adjusted, M- and T-levels were markedly reduced in comparison to the other electrodes, which resulted in an unfavorable stimulation result by the Implant as a whole. Consequently, this individual electrode was switched off, which resulted in regular hearing rehabilitation with no further unwanted effects (Fig. 7).

B. Case 2:

A 7-year-old girl with prelingual deafness was implanted with a "Clarion Hi-Focus CII" Cochlear Implant at 4 years of age. 3 years past implantation, she sowed an uneventful course, with pure tone hearing thresholds (CI-aided) of 30-40dB and a speech audiogram (Goettinger I) with a word recognition score of 75% at 70dB. M-levels had been stable over 3 years. No individual adjustments had to be made to stimulation mapping and she was comfortable wearing the speech processor all day. About 3 months after her last follow-up, the child complained during an emergency consultation that even with minimum volume setting (50%), sound perception was far too loud and speech recognition was severely reduced. Consequently, she refused to wear the device any longer. Individual stimulation mapping revealed that her M-levels were reduced to 25-35% of the original values. Impedance values showed no change. However, TECAP values confirmed increased response levels for basal and medial electrodes (Fig. 8a and b). With new stimulation mapping, volume was adjusted to comfortable levels and speech perception could be restored to former values.

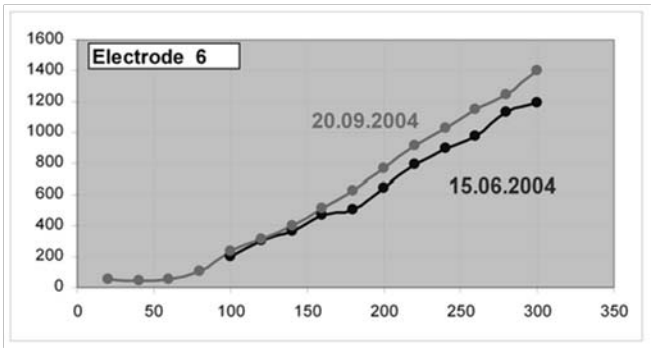


Fig. 8a: Minor change of TECAP values for (apical) electrode #6 over 3 months (Case 2)

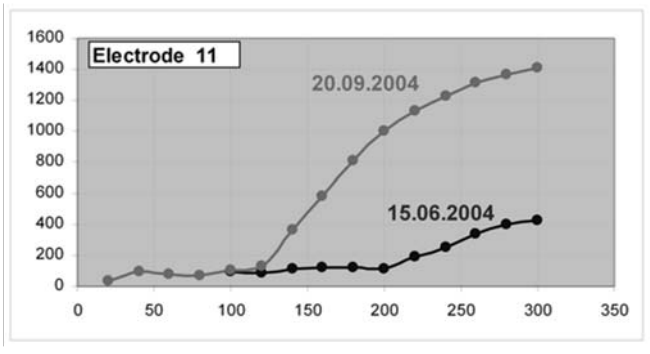


Fig. 8b: Marked increase of TECAP responses for (basal) electrode #11 over 3 months (Case 2)

IV. Discussion

From the individual setup concerning implant hardware and stimulation strategies provided by each manufacturer, stimulation currents and TECAP response amplitudes can only be compared intraindividually and, within limits, between patients fitted with the same type and series of Cochlear Implant. However, bearing these restraints in mind, it can be shown that TECAPs are a valuable tool in assessing not only Cochlear Implant function intraoperatively, but also the auditory nerve's ability to respond to external electric stimulation. Furthermore, the current threshold levels for TECAP responses can be quantified. The key question, however, is whether these current levels correspond to the current levels of hearing threshold (T-level) and of maximum volume (M-level) which are expressed by the patient when the Cochlear Implant is fine tuned individually about four weeks past implantation. From earlier reports given by a multicenter study [14] and by Dillier and co-workers [15] it is known that the current levels that are used to elicit TECAPs intraoperatively correspond to current levels that lie within the range between later T- and M-levels to about

75 to 90%.

On the other hand, missing TECAP potentials during implantation do not indicate implant failure or cochlear nerve dysfunction. Although in one of two cases, the electrode position was checked additionally by an intraoperative CT scan, both implant electrodes showed no abnormalities during insertion and normal electric impedance values. In both cases, individual hearing rehabilitation showed a normal and uneventful course.

As we could show, TECAP levels themselves do show a certain shift especially during the first four weeks following implantation, which levels off during the following months. Thus, especially in children who are implanted due to prelingual deafness and are yet too young to express comfort or discomfort in hearing adequately, TECAP registration is extremely useful to assess changes in individual stimulation levels in an objective and reproducible manner. As it could be shown, not only TECAP levels but also M- and T- values can change without apparent reason even years after CI usage. Therefore, TECAPs can help to assess neural stimulation levels objectively. The ability to quantify threshold levels will also serve further branches of research, e.g. drug delivery devices incorporated in Cochlear Implant electrodes, by which means neurotrophic factors are applied next to the modiolus [16]. By branching of the cochlear nerve towards the electrode, a gradual reduction in current levels for hearing is postulated which could be assessed by TECAP follow-up.

V. Conclusion

TECAP registration is already a reliable tool which objectively assesses auditory nerve response to electric stimulation by Cochlear Implants during follow-up, especially in those patients who are unable to communicate verbally on their individual sound and speech perception.

With further advances in stimulation and registration strategies plus the acquisition of more data, it is likely that TECAP patterns are identified which may serve as a predictive tool for individual hearing levels in Cochlear Implant patients. As yet, however, the data available up to present are too variable to allow any individual prognosis.

Acknowledgment

The authors would like to acknowledge the help of MED-EL GmbH, Innsbruck, Austria; Cochlear Inc., Lane Cove, Australia and Advanced Bionics, Sylmar CA, USA for their kind permission to reproduce part of their data and screenshots appearing in this article.

References

1. Moller AR (2006) History of cochlear implants and auditory brainstem implants. Adv Otorhinolaryngol 64:1-10
2. Geers AE (2006) Factors influencing spoken language outcomes in children following early cochlear implantation. Adv Otorhinolaryngol 64:50-65
3. Harrison RV, Gordon KA, Mount RJ (2005) Is there a critical period for cochlear implantation in congenitally deaf children? Analyses of hearing and speech perception performance after implantation. Dev Psychobiol 46(3):252-61
4. Papsin BC (2005) Cochlear implantation in children with anomalous cochleovestibular anatomy. Laryngoscope 115(1 Pt 2 Suppl 106):1-26
5. Zeng FG (2004) Trends in cochlear implants. Trends Amplif 8(1):1-34
6. Mylanus EA, Rotteveel LJ, Leeuw RL (2004) Congenital malformation of the inner ear and pediatric cochlear implantation. Otol Neurotol 25(3):308-17
7. Incesulu A, Kocaturk S, Vural M (2004) Cochlear implantation in chronic otitis media. J Laryngol Otol 118(1):3-7
8. El-Kashlan HK, Telian SA (2004) Cochlear implantation in the chronically diseased ear. Curr Opin Otolaryngol Head Neck Surg 12(5):384-6
9. Das S, Buchman CA (2005) Bilateral cochlear implantation: current concepts. Curr Opin Otolaryngol Head Neck Surg 13(5):290-3
10. McDermott HJ (2004) Music perception with cochlear implants: a review. Trends Amplif 8(2):49-82
11. Roland PS, Wright CG (2006) Surgical aspects of cochlear implantation: mechanisms of insertional trauma. Adv Otorhinolaryngol 64:11-30
12. Briggs RJ, Tykocinski M, Stidham K, Roberson JB (2005) Cochleostomy site: implications for electrode placement and hearing preservation. Acta Otolaryngol 125(8):870-6
13. Callanan V, Poje C (2004) Cochlear implantation and meningitis. Int J Pediatr Otorhinolaryngol 68(5):545-50
14. Cafarelli Dees D, Dillier N, Lai WK, von Wallenberg E, van Dijk B, Akdas F, Aksit M, Batman C, Beynon A, Burdo S, Chanal JM, Collet L, Conway M, Coudert C, Craddock L, Cullington H, Deggouj N, Fraysse B, Grabel S, Kiefer J, Kiss JG, Lenarz T, Mair A, Maune S, Muller-Deile J, Piron JP, Razza S, Tasche C, Thai-Van H, Toth F, Truy E, Uziel A, Smoorenburg GF (2005) Normative findings of electrically evoked compound action potential measurements using the neural response telemetry of the Nucleus CI24M cochlear implant system. Audiol Neurootol 10(2):105-16
15. Dillier N, Lai WK, Almqvist B, Frohne C, Muller-Deile J, Stecker M, von Wallenberg E (2002) Measurement of the electrically evoked compound action potential via a neural response telemetry system. Ann Otol Rhinol Laryngol 111(5 Pt 1):407-14
16. Gillespie LN, Shepherd RK (2005) Clinical application of neurotrophic factors: the potential for primary auditory neuron protection. Eur J Neurosci 22(9):2123-33

Address of the corresponding author:

Author: Justus Ilgner MD
Institute: Department of Otorhinolaryngology, Head and Neck Surgery, University of Aachen RWTH
Street: Pauwelsstrasse 30
City: 52057 Aachen
Country: Germany
Email: jilgner@ukaachen.de

Medical Sensors of the BASUMA Body Sensor Network

N. Loew[1], K.-J. Winzer[2], G. Becher[3], D. Schönfuß[4], Th. Falck[5], G. Uhlrich[6], M. Katterle[1] and F.W. Scheller[1]

[1] University of Potsdam, Institute of Biology and Biochemistry, Analytical Biochemistry, Golm, Germany
[2] Charité, Brustzentrum, Berlin, Germany
[3] FILT Lungen- und Thoraxdiagnostik GmbH, Berlin, Germany
[4] BST Bio Sensor Technologie GmbH, Berlin, Germany
[5] Philips Research Europe, Aachen, Germany
[6] ABS Ges. für Automatisierung Bildverarbeitung und Software mbH, Jena, Germany

Abstract— **The BASUMA (Body Area System for Ubiquitous Multimedia Applications) body sensor network will consist of several wearable or handheld wireless medical sensors and a PDA like base station. The sensors, designed in a way to minimize disturbance of the patient's everyday life, will record important diagnostic parameters which are analysed on the base station. This will allow a continuous monitoring of the patient's state of health outside the hospital. Sensors were developed for the measurement of electrocardiograms (ECGs), air and blood content of the thorax (thoracic impedance), body temperature, breathing rate and cough control, blood pressure, pulse rate, oxygen saturation, lung functions, reactive oxygen species (ROS) (exhaled H_2O_2), and lactate in breath condensate.**

Keywords— **point of care diagnostics, biosensor**

I. Introduction

BASUMA (Body Area System for Ubiquitous Multimedia Applications) [1] shall offer patients a continuous monitoring of their state of health outside the hospital. For this purpose, wearable or handheld wireless medical sensors record important diagnostic parameters which are analysed on a PDA like base station. All sensors will be designed in a way to minimize disturbance of the patient's everyday life.

In contrast to traditional telemedical systems where the raw data are sent to a medical centre for processing, this approach skips time-consuming data transfers in order to give real-time recommendations to the patient. In the BASUMA concept, a medical centre is alarmed via the mobile network in case of a critical condition.

The BASUMA platform is designed in such a way that different non-invasive sensors can be combined to meet the individual patient's needs without changing the basic set-up.

In addition to established devices, sensors were developed to measure the following parameters:

- Electrocardiogram (ECG)
- Air and blood content of the thorax (thoracic impedance)
- Body temperature
- Breathing rate and cough control
- Blood pressure
- Pulse rate, oxygen saturation
- Lung functions
- Reactive oxygen species (ROS) (exhaled H_2O_2)
- Lactate in breath condensate

II. The BASUMA System

The BASUMA approach requires low-power systems on the body, new radio standards, easy-to-use sensors and intelligent software for pre-analyses. All known technical standards for point-of-care sensors, data collection, radio transfer and analysis were verified for BASUMA usability.

Patient-device interactions should be enabled by touch screen and/or phonetic input.

We developed and tested demonstration units for phonetic input and output for the patient-machine interaction. A simple diary and a schedule for medicine intake and control of diagnostic parameters were designed.

To enable wireless communication we used the IEEE 802.15.4/ZigBee compliant Philips AquisGrain platform shown in Fig. 1 [2].

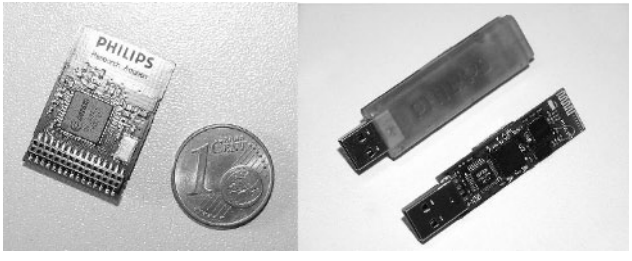


Fig. 1 IEEE802.15.4/ZigBee-based AquisGrain platform used to realize the wireless sensors and USB gateway

III. Continuous Cuff-Less Blood Pressure

Traditional blood pressure devices are mainly based on a sphygmo-manometric occlusive arm-cuff, which is clumsy, uncomfortable and allows only for intermittent measurements at intervals of several minutes.

An alternative way is to infer the arterial blood pressure using the pulse wave velocity methodology. This can be done by using a wireless sensor configuration with an electrocardiogram (ECG) derived with a textile belt from the waist as the proximal signal and a photoplethysmogram (PPG) measured at the ear [3, 4]. With this set-up we can measure the pulse arrival time (PAT) which is defined as the time delay between the R peak of the QRS wave from the ECG and the arrival of the arterial pulse wave at the periphery, i.e. the PPG. We demonstrated that with the wireless body sensors depicted in Fig. 2 the systolic blood pressure can then be estimated from the PAT with an accuracy of 6.9 mmHg (root mean square error) at beat level.

Both the ECG and the PPG signal have been sampled at 200 Hz with a 10-bit resolution. The algorithms perform the extraction of signal features with sub sample time precision thanks to an interpolation scheme to ensure a relative precision of about 1% of the PAT values. To realize the wireless sensors we integrated the Philips AquisGrain sensor platform into the ECG and PPG front ends.

Fig. 2 Wireless ECG belt (top) and PPG ear sensor (bottom)

IV. Thoracic Impedance

The thoracic impedance is one possibility to measure a signal proportional to the respiratory movement. Traditional tracings with a thoracic belt need a couple of electrodes and have a high energy consumption.

The new BASUMA solution is a 4-electrode system with a recently developed amplifier. Placed on the left thorax in the front and at the back the system enables a continuous tracing of the breath movements. The goal parameters are breathing rate and trend of tidal volume.

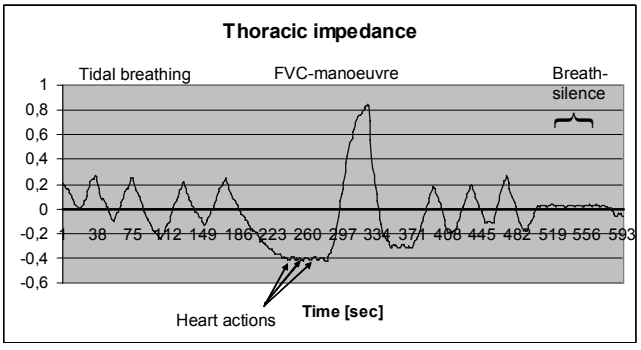


Fig. 3 Tracing of thoracic impedance

The impedance gives a robust signal not susceptible to disturbances (Fig. 3). The energy consumption is very low. The electrodes may also be used for ECG in a time splitting manner.

V. Cough Sensor

Fig. 4 Microphone capsule on male test person

The continuous monitoring of cough attacks and irritative coughs is a common problem in clinical medicine. The subjective counting of coughs is not valid. Cough attacks may be monitored by using a specially designed microphone (Fig. 4, 5) for the detection of lung sounds and thoracic noises (Fig. 6).

Fig. 5 Development steps of microphones

During a continuous monitoring of breath functions under everyday life conditions the problem remains to identify and exclude interferences with speaking and extraneous

noises. Nevertheless a robust counting of coughs in the signal obtained is possible.

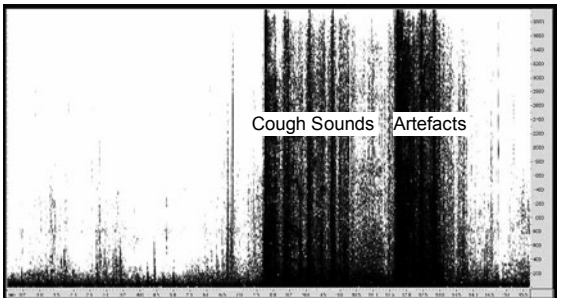


Fig. 6 Cough Sound, y-scale 200, 400, …, 3800 Hz

VI. Lactate Biosensor

BST Bio Sensor Technologie GmbH and Philips Research have developed a prototype of a hand-held device that allows the determination of the lactate concentration by only one exhalation [5].

The device is based on an electrochemical biosensor and works according to the following functional principle: A ceramic-based lactate biosensor with a condensation area is cooled by a peltier element. Exhalation on the biosensor results in the collection of one drop of breath condensate. Catalyzed by lactate oxidase lactate is metabolised and hydrogen peroxide is produced, which is detected by the electrode system. Fig. 7 shows the experimental set-up and Fig. 8 the resulting curve of the measured current which is used to determine the lactate concentration.

Fig. 7 Experimental set-up for lactate measurement

By means of AquisGrain, the handheld device wirelessly connects to the BASUMA body sensor network for further processing of the measured lactate concentration.

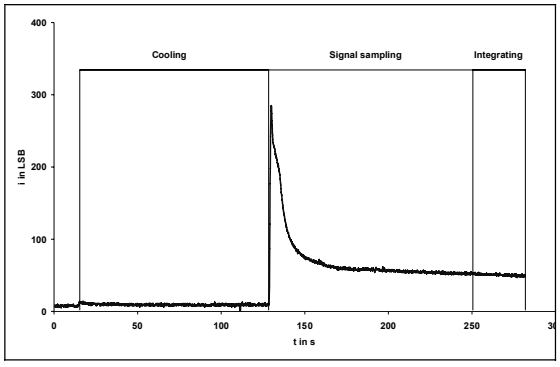


Fig. 8 Example for the lactate measurement: a current – time curve

VII. Hydrogen Peroxide Sensor

Hydrogen Peroxide (H_2O_2) can be measured electrochemically either by oxidation to oxygen (O_2) or by reduction to water (H_2O).

An exhaled H_2O_2 sensor suitable for the BASUMA system has to be very sensitive ($0 – 1000$ nmol l^{-1} H_2O_2 for breast cancer patients, $500 – 1000$ nmol l^{-1} H_2O_2 for COPD patients) and require very small sample volumes.

A. Principle 1: H_2O_2 Oxidation

H_2O_2 can easily be oxidized at a platinum electrode. An electrode similar to that for lactate detection was used for this purpose. This electrode consists of a special condensation surface to collect the sample, a platinum working electrode and a suitable reference electrode (Fig. 9).

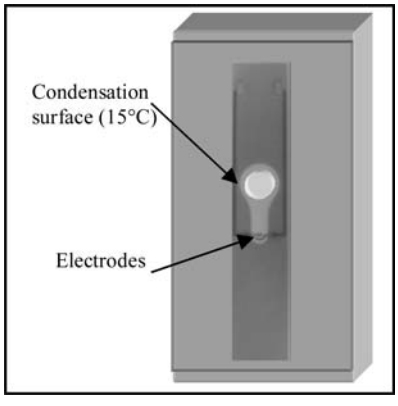


Fig. 9 Scheme for the H_2O_2/Lactate sensor

To measure the exhaled H_2O_2, the patient breathes onto the condensation surface, where the breath is cooled down to 15°C. A drop of about 2 µl breath condensate is formed

and drops down to the electrodes where the H_2O_2 is measured.

To minimize interfering capacitive currents, the H_2O_2 oxidation is measured via differential pulse amperometry (DPA). With this method, the capacitive current can be minimized in a first pulse and the oxidation current measured in a second pulse. The pulse parameters for the DPA measurement were optimized so that the oxidation current depends linearly on H_2O_2 concentration.

Fig. 10 shows the dependency of the oxidation current on the H_2O_2 concentration. The linear range extends from 8 to 250 nmol l^{-1} (Fig. 10, insert) with saturation above 250 nmol l^{-1}.

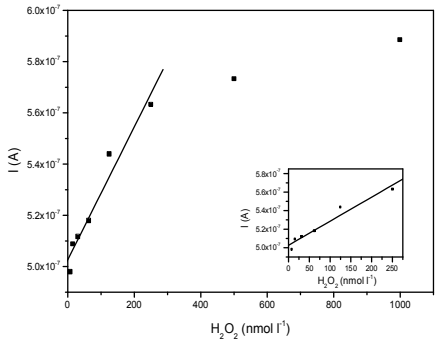


Fig. 10 Concentration dependency for an H_2O_2 oxidizing sensor

While this method is very sensitive and works with small sample volumes, the measuring range is yet too small.

B. Principle 2: H_2O_2 Reduction

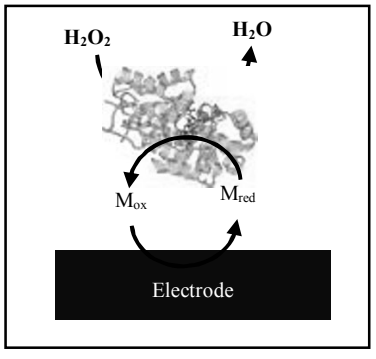


Fig. 11 Principle of a redox-mediated peroxidase biosensor

A wider measuring range can be achieved with an H_2O_2 reducing sensor. A peroxidase biosensor is such an H_2O_2 reducing sensor. An established approach to build a peroxidase biosensor is to immobilize horseradish peroxidase (HRP) together with a redox mediator on an electrode surface. H_2O_2 is then reduced by the enzyme to H_2O. Subsequently the enzyme is reduced by the mediator which then is recycled at the electrode (Fig. 11).This principle was also applied for the H_2O_2 sensor used for the clinical trial.

By changing the measurement set-up a 40 fold decrease of the required sample volume to 12.5 µl is achieved. The linear range (50 to 2000 nmol l^{-1}) is much wider than for the oxidizing sensor (Fig. 12) but with a limit of detection of 50 nmol l^{-1} the HRP sensor is less sensitive.

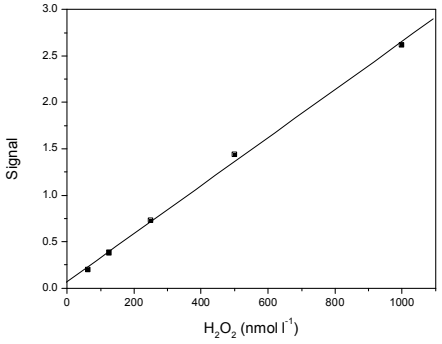


Fig. 12 Concentration dependency for a H_2O_2 reducing sensor

VIII. Clinical Trial

In this project, H_2O_2 and lactate sensors are tested for their suitability for patients suffering from

- chronic obstructive pulmonary disease (COPD)
- breast cancer, undergoing chemotherapy

In order to validate the selected diagnostic parameters for a certain group of patients, they have been compared to traditional methods in a clinical trial. As all sensors in the BASUMA system are non-invasive, the available parameters might be very different from the traditional ones.

The progress of chemotherapy is monitored by the course of the number of neutrophilic granulocytes. Being phagocytes, neutrophiles have the enzyme NADPH oxidase, which is producing large amounts of superoxide radicals by transferring one electron from cytosolic NADPH to an extra-cytosolic oxygen molecule. Superoxide radicals have a very short life-time and are converted quickly into other reactive oxygen species (ROS), e.g. hydrogen peroxide (H_2O_2). Thus, H_2O_2 serves as an important indicator mole-

cule for ROS, since it is a common intermediate in the decay of the oxygen radicals.

Therefore the H_2O_2 concentration in the exhaled breath can be a parameter for the activity of neutrophilic granulocytes in the proximity of the lungs, as was shown for lung-cancer patients [6].

Lactate is often used as a marker of anaerobic metabolism [7]. A high lactate concentration thus indicates disturbances in the oxygen metabolism.

In a first clinical trial we studied exhaled H_2O_2 and/or lactate as non-invasive parameters to monitor the chemotherapy of breast cancer patients.

This trial was carried out at the interdisciplinary breast center (IBZ) at the Charité clinic in Berlin-Mitte. As the BASUMA system and its sensors are still under development, commercially available sensors were adapted for this study and used in the clinic.

Breath condensate and blood samples were taken weekly from 18 voluntary patients. The H_2O_2 and lactate concentrations of the breath condensate and the lactate concentration of the blood were determined on site. A standard differential blood count was done of blood taken on the same day.

A. Breath Condensate

To collect the breath condensate, an EcoScreen®(FILT GmbH, Germany) device was used. The patients breathed into a mouth piece for two to fourteen minutes until a total volume of 100 l was reached. The breath was cooled down in a tube to give a total volume of 0.5 to 2 ml of condensate.

B. H_2O_2 Measurement

An EcoCheck® (FILT GmbH, Germany) measuring system was used for H_2O_2 determination. An amperometric peroxidase biosensor was inserted into the system. 400 to 500 µl of breath condensate were mixed thoroughly with an equal volume of dilution buffer within minutes after collection and injected into the sample channel. The measurement was started immediately.

C. Lactate Measurement

Two LactatProfi® (ABT GmbH, Germany) devices were used. One device was used with the standard buffer, biosensor and calibration solution for the measurement in blood. For the measurement in breath condensate buffer, biosensor and calibration solution were adapted. 50 to 100 µl of breath condensate were mixed thoroughly with an equal volume of dilution buffer prior to measuring. The actual lactate concentration was calculated considering the dilution and difference in calibration solution.

D. Results

First evaluations show that the hydrogen peroxide concentration in breath condensate before chemotherapy is 200 ± 130 nmol l^{-1}. The high scatter suggests that this value is patient specific. The different concentrations given in literature for healthy people, e.g. 569 ± 30 nmol l^{-1} [8] or 310 nmol l^{-1} [9], support this assumption.

Though the starting point for each patient differs, the H_2O_2 concentration clearly varies over the course of the chemotherapy.

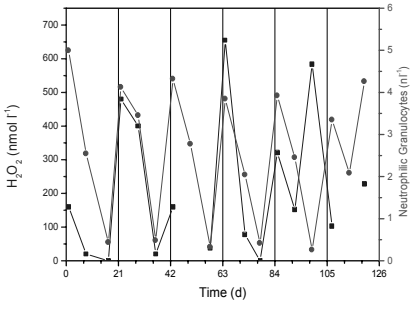


Fig. 13 Comparison of H_2O_2 in breath condensate and neutrophiles in blood for patient A, vertical lines indicate beginning of new cycle

Fig. 13 shows an example for a comparison of the H_2O_2 concentration in breath condensate and the amount of neutrophilic granulocytes in blood for one patient. Over the first four cycles both H_2O_2 concentration and amount of neutrophiles go down during the cycle and up before the following drug infusion. A cycle is characterized by a three-week period and one dose of chemotherapy on the first day.

According to first evaluations the progress of the chemotherapy is not reflected as exactly by the exhaled H_2O_2 as by the amount of neutrophiles. But a much smaller time span between measurements, which will be possible with the BASUMA system, should compensate for this lack of accuracy.

The lactate concentration in the breath condensate does not seem to correlate with the lactate concentration in blood (data not shown). A comparison of the lactate concentration in breath condensate with the amount of neutrophilic granulocytes showed a parallel curve over a maximum of three cycles (Fig. 14).

Whether lactate is an indicator for other symptoms that are not related to the change of the amount of neutrophiles, has not yet been elucidated.

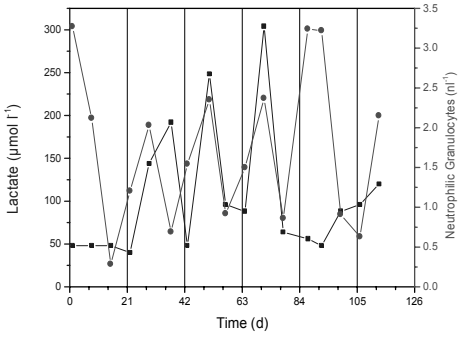


Fig. 14 Comparison of lactate in breath condensate and neutrophiles in blood for patient B, vertical lines indicate beginning of new cycle

IX. Conclusions

The components of the BASUMA system are ready to use in single-sensor systems. The usefulness of a cough sensor with a microphone, the thoracic impedance, ECG, blood-pressure, oxygen saturation as on-line measurements is proven. Off-line measurements of H_2O_2 and lactate in exhaled air, e.g. breath condensate, are possible and give additional information about the state of health of the patient.

It seems possible to give valid information by long-term measurements of simple parameters.

In the case of chemotherapy for breast cancer patients it has not yet been elucidated if the BASUMA system can replace weekly blood counts. A continuous monitoring of their state of health will help improve the patients' comfort and adaptation of medication.

The daily and/or weekly trend of the number of cough-attacks, the blood pressure or thoracic impedance does not need an absolute calibration of the sensors. For a direct intervention or a call to a medical center a deviation of a pre-set value is sufficient.

Therefore, handling of the sensors is very simple both for patients and doctors. Calibrations and standardized measurements (e.g. standard ECG tracings) are for the most part not necessary.

The cough sensor using special isolated microphones is robust and simple compared to the so-called "lung-sound analyses" in use at present.

Acknowledgment

We thank the German Federal Ministry of Economics and Technology (BMWi) for financial support (01MT308, BASUMA).

Furthermore we thank S. Terjung and A. Lehman for helpful experiments.

References

1. BASUMA at www.basuma.de
2. Espina J, Falck T, Mülhens O (2006) Network Topologies, Communication Protocols, and Standards, In: G.Z. Yang (Ed.) Body Sensor Networks, Springer, pp 145-182
3. Muehlsteff J, Aubert X, Schuett M (2006) Cuffless Estimation of Systolic Blood Pressure for Short Effort Bicycle Tests: The Prominent Role of the Pre-Ejection Period, EMBC Proc., New York, USA, 2006
4. Espina J, Falck T, Muehlsteff M, Aubert X, Wireless Body Sensor Network for Continuous Cuff-less Blood Pressure Monitoring, ISSS-MDBS Proc., 3rd International Summer School and Symposium on Medical Devices and Biosensors, Cambridge (MA), USA, 2006, pp 11-15
5. Schönfuß D, Szeponik J, Gall I; Verfahren und Vorrichtung zur Gewinnung und Analyse von Atemkondensaten; Pat. Nr. DE 102005020102
6. Wewel A R, Crusius J A M, Gatzemeier U et al. (2005) Time course of exhaled hydrogen peroxide (H_2O_2) and nitric oxide (eNO) during chemotherapy in patients with lung cancer. Eur Respir J 26: Suppl. 49:92s
7. Kellum J A, Kramer D J, Lee K et al. (2006) Release of Lactate by the Lung in Acute Lung Injury. Chest 111:1301-1305
8. Becher G, Rothe M, Decker M (2005) Breath Analysis For Clinical Diagnosis And Therapeutic Monitoring, p. 421-428, World Scientific Publishing
9. Kostikas K, Papatheodorou G, Psathakis K et al. (2003) Oxidative Stress in Expired Breath Condensate of Patients with COPD. Chest 124:1372-1380

Address of the corresponding author:

Author:	Prof. Dr. Frieder W. Scheller
Institute:	University of Potsdam
	Institute of Biology and Biochemistry
	Analytical Biochemistry
Street:	Karl-Liebknecht-Str. 24-25
City:	14476 Golm
Country:	Germany
Email:	fschell@rz.uni-potsdam.de

4th Session
On-Body Sensors

Reflective Photoplethysmograph Earpiece Sensor for Ubiquitous Heart Rate Monitoring

L. Wang, B. Lo and G. Z. Yang

Department of Computing, Imperial College London, London, United Kingdom SW7 2RH

Abstract— **This paper addresses the design considerations of a novel earpiece photoplethymograph (PPG) sensor and its *in-situ* evaluation results. The device is encapsulated with multiple LEDs and photodiodes based on a reflective PPG design. A compact and low power circuitry was developed for signal control and conditioning. PPG signals with an averaged a.c./d.c. ratio of 0.001-0.01 and 10% relative strength (compared to finger-based approach) were recorded from the superior and posterior auricular skins. PPG signal integrity and heart rate detection accuracy were evaluated and the results showed that with adequate optical shielding and motion cancellation, the device could reliably detect heart rate both during rest and moderate exercise. The proposed sensor design is low power, easy to wear compared to conventional earlobe PPG devices.**

Keywords— **Wearable sensors, photoplethysmograph (PPG), heart-rate monitoring, earpiece, body sensor networks**

I. Introduction

Continuous and non-intrusive monitoring of cardiovascular function is essential for the future development of pervasive healthcare [1]. Although extensive measurement of biomechanical and biochemical information is available in almost all clinical settings, the diagnostic and monitoring utility is generally limited to the brief time points and perhaps unrepresentative physiological states such as supine and sedated, or artificially introduced exercise tests. Transient abnormalities, in this case, cannot always be captured. Many cardiac diseases are associated with episodic rather than continuous abnormalities. These abnormalities are important but their timing cannot be predicted and much time and effort is wasted in trying to capture an "episode" with controlled monitoring. Important and even-life threatening disorders can go undetected because they occur only infrequently and may never be recorded objectively.

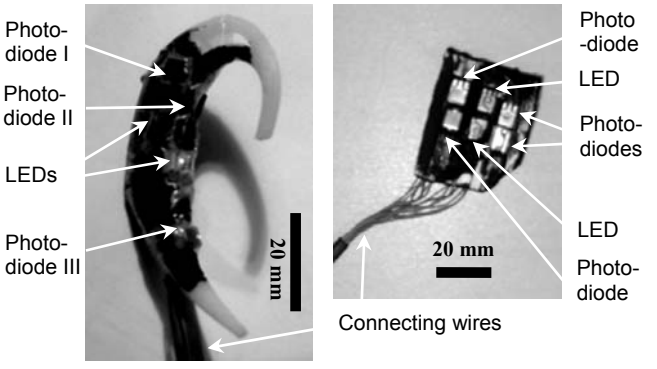
Thus far, a range of ECG monitoring devices that permits the continuous recording of heart-rate variability has been proposed. These include digital Holter devices for capturing arrhythmogenic events and chest-strip type devices for professional sports and exercise [2]. Photoplethysmograph (PPG) devices have received significant attention in recent years due to their ease of being integrated with wearable, pervasive sensing devices. PPG is based on the detection of subcutaneous blood perfusion by shining lights through a capillary bed. As arterial pulsations fill the capillary bed, the volumetric changes of the blood vessels modify the absorption, reflection or scattering of the incident lights, so the resultant reflective/transmittal lights could indicate the timing of the cardiovascular events, such as heart rate. The PPG sensor requires at least one light source (usually infrared) and one photo detector in its close proximity [3]. PPG sensors are commonly worn on fingers because of the highest signal strength that can be achieved [4]. This configuration, however, is not suitable for pervasive sensing as most daily activities involve the use of fingers.

Different positioning of the PPG sensors has been explored extensively in recent years. This includes body locations such as ring finger [5], wrist [6], brachia [7], belly [8] and esophageal [9]. For commercial clinical PPG sensors, it is also common to use earlobe and forehead [10] as the anatomical regions of interest. An ear-clip can cause pain if it is used over a long period of time, and neither approach is suitable for pervasive sensing applications. The purpose of this paper is to explore a reflective PPG sensor design that can be integrated with the ear-worn activity recognition (e-AR) platform. The device is small, discreet to wear, and thus is suitable for long-term pervasive monitoring.

II. Material and Method

The basic structure of the PPG sensor is illustrated in Fig. 1. The optical components used include LEDs DLED-660/905, DLED 660/940 from UDT® and PDI-E835 from API®, and photodiodes PIN-4.0, PIN-8.0 from UDT® and BPW34F (with daylight filter) from Siemens®. The active areas of these photodiodes were 4, 8 and 7 mm^2, respectively. For the dual-light LED components, only the infrared light channel was used in our experiments. There are two sensing planes for the e-AR PPG sensor, which are perpendicular to each other as shown in Fig. 1(a). To assess the relative signal strength from the proposed sensor, a standard patch sensor was also constructed by using the same optical components. The distances between the LEDs and the correlated photodiodes were in a range of 8-12 mm [11]. Both sensors were equipped with multiple LEDs and

photodiodes, operating by reflective lights [12, 13]. For optical shielding, all components were recessed into the base at 1 mm depth, and the adjacent components were bridged with opaque medium. The bases of both sensors were painted black to prevent multiple scatterings.



(a) (b)

Fig. 1 LED and photodiode components were encapsulated into (a) an e-AR sensor and (b) a patch, forming the reflectance earpiece PPG and patch PPG sensors. Optical shunt was avoided by a recessed design and careful component layout. For the e-AR sensor, the photodiodes I, II and III formed the reception channels 1, 2 and 3, correspondently. Reception channel 1 uses the daylight filter photodiode. When worn, the e-AR sensor has two perpendicular optical sensing planes, one being parallel to the temporal scalp (reception channel 1), another being vertical to the temporal scalp (reception channels 2 and 3).

The control circuit for the PPG sensors was based on the design of Zhang et al. [14]. In this study, one current regulator diode (SST50x from Vishay®) was in series with one LED and there were 7 current regulator diodes in total for different current requirements. The photodiode output currents were fed into the differential trans-impedance amplifiers (OP297s from Analog®). A rail-to-rail amplifier (LT1491 from Linear®) was used for different gain levels. A LM2664 from National Semiconductor® was used to provide +/-3 V power supplies for OP297s. The circuit had three amplification channels for simultaneous receptions, and the power consumption was approximately 6 mW per channel. For *in-situ* experiments, a DAQ device (USB-6009 from National Instruments®) was used to acquire data into a PC, at a sample rate of 1KSps per channel. Data processing and visualization were completed off-line using the same PC. For performance comparison of the proposed e-AR PPG sensor, a bed-side pulse oximeter (OxiMax N-560 from Nellcor®) with finger/earlobe/forehead sensor probes was used.

The earpiece and the patch PPG sensors were tested *in-situ* with 10 healthy subjects (age between 23-35, 6 males, 4 females). Fig. 2 indicates the locations of the e-AR PPG sensor and the patch sensor used for the experiment. The chosen locations have rich vascularity (i.e. superficial temporal and posterior auricular arteries/veins and adjunct capillaries) and thin epidermal layer with less skin pigmentations – this is advantageous because the total optical absorption of the epidermis depends primarily on the melanin absorptions [15]. The infrared lights therefore could reach the subcutaneous blood vessels with less attenuation. For all experiments, one of the Nellcor PPG sensors was attached to the subject and the analogue output from the Nellcor device was used as the reference signal throughout the experiments. The experiments conducted include the following two protocols:

1) *PPG sensing when the subjects are still*. Different LED driving currents were used in this protocol and the raw PPG signals were down-sampled (decimation) to be 50 Sps per channel, prior to baseline (d.c.) detection and band-pass filtering with a pass-band being 0.5Hz-4Hz. The processed signal was visualized and the averaged pulsatile portion of the processed signal (a.c.) was estimated. This is to evaluate the accuracy of the PPG signal at different sensor locations. The corresponding results are summarised in Fig. 3 and Table 1.

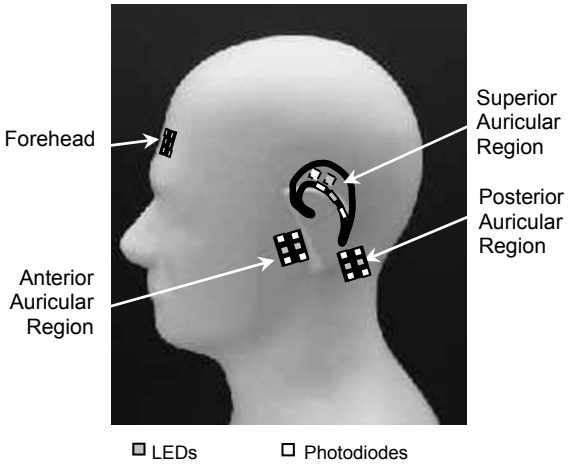


Fig. 2 Earpiece and patch PPG sensors were placed at different auricular regions and the forehead. For ambient light shielding, an opaque medium was used to cover the earpiece sensor / the auricular region. All locations were indicated with sensors here, but for every experiment, there was only one sensor to be worn.

2) *PPG sensing during a sequence of activities*. This is to assess the accuracy of the PPG sensor during general daily activities (standing, walking, and sitting) and then walking on a treadmill at a moderate speed (from 4 km/h to 7 km/h back to 4 km/h). A complete circuit lasted approximate 20 minutes and each activity was time-stamped. A moving FFT (Hanning-windowed, window length 20 seconds) was used to calculate the PPG signal spectrum density (amplitude) and the heart rate [3]. For assessing the accu-

racy of the heart rate detection algorithm, a heart rate spectrum fidelity index f_{HRS} was introduced. This is defined as the ratio of the interested frequency spectrum density over the full-spectrum strength. This index has a positive value between 0 and 1, *i.e.* for single-frequency sine wave f_{HRS} = 1; and for white noise f_{HRS} equals to the ratio of the interested frequency span over half sampling rate. With this experiment protocol, it is possible to evaluate the effects of motion artifact on heart rate detection [16]. The effect of multiple LEDs/photodiodes arrangements was also assessed and the results are summarised in Figs. 4-6 and Table 1.

III. Result

The e-AR PPG sensor was found to be comfortable to wear by all subjects. Fig. 3 demonstrates two typical PPG signals detected from the mastoid process and the super temporal region, respectively. LED driving currents between 4-8 mA were applied and no clear PPG signal-to-noise ratio variations were observed within this current range. For the e-AR sensor the experiment confirmed that the reception channel 1 (with the daylight filter photodiode) did not show noticeable difference comparing to the other two reception channels.

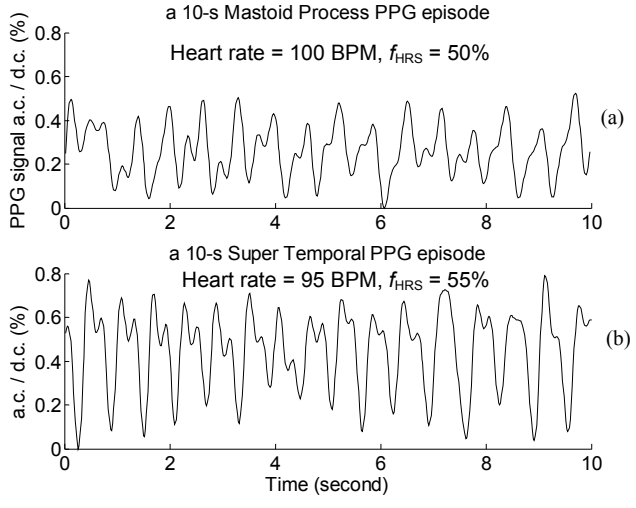


Fig. 3 Two PPG signal episodes detected from a subject in still position, using patch sensor (a) and earpiece sensor (b). PPG signals were clean and the dicrotic wave was visible. LED driving currents were 4 mA. For f_{HRS} calculations, the interested frequency span was set to be 0.2 Hz. The f_{HRS} values implied that half of the frequency spectrum energy was within 94-106 BPM frequency band for signal episode (a), and within 89-101 BPM frequency band for signal episode (b).

It was found that the PPG signal deteriorated from all locations during subject movement. This is mainly due to the relative displacements between the capillaries and the e-AR sensor. To demonstrate this effect, a frequency domain result using windowed-FFTs was displayed in Fig. 4. This subject was walking at a treadmill speed of 5.5 km/h with a corresponding pace of 115 steps per minute. From Fig. 4, it is evident that there are clearly two spectrum peaks from e-AR channel 2 (when LEDs were switched on) and the reference channel. In this case, the first peak corresponds to step frequency whereas the second peak corresponds to the real heart rate at 150 BPM (left-column plots), 155 BPM (middle-column plots) and 160 BPM (right-column plots), respectively. It can be seen that e-AR channel 1, when LEDs were switched on, is dominated by motion spectrum, suggesting its susceptibility to motion artifacts.

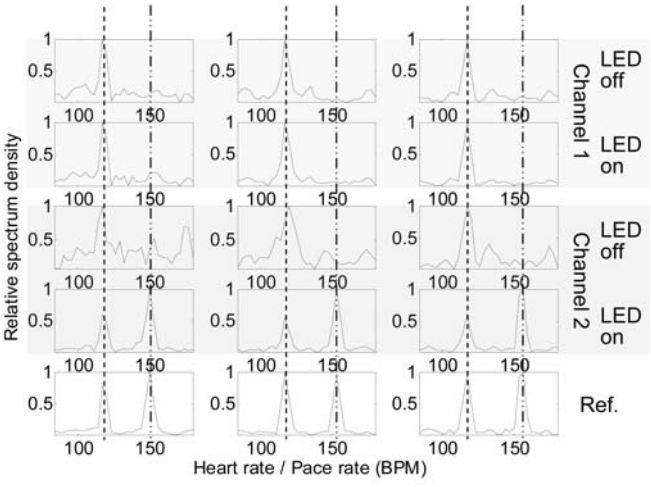


Fig. 4 Relative spectrum density plots from the e-AR channel 1 (the first two rows); the e-AR channel 2 (the third and fourth rows); with LEDs on and off, respectively; and the reference channel (the bottom-raw plots). The LEDs were steered to be on (8 mA driving current) and off with a 25% duty circle. Each plot was derived from a 20-second PPG signal episode.

When the LEDs were switched off, the outputs from the photodiodes were dominated by the motion artifact. This is because, there are induced currents to the photodiodes due to residual ambient lights and the practical difficulties of achieving 'perfect' optical shielding. The induced currents vary according to the relative displacement/angle changes of the photodiodes, which are caused by the subject's movements. This was demonstrated in the first and the third rows of the Fig. 4, where a single step frequency spectrum peak from both reception channels was recorded in every frequency domain plot.

Experiments with Protocol (2) showed the aforementioned frequency-domain patterns existed during most of the activities. An automatic motion cancellation algorithm was designated to remove the motion artifacts. This algorithm first determines the step frequency using the induced current readings from the reception channels, and then removes the step frequency band from the reception channel outputs when LEDs were switched on – this was achieved by using a notch filtering within the step frequency span (0.2 Hz, or ±6 BPM). The remaining spectrum peak therefore represents the heart rate frequency peak. Heart rate curves calculated from 3 subjects using this motion cancellation method were illustrated in Fig. 5.

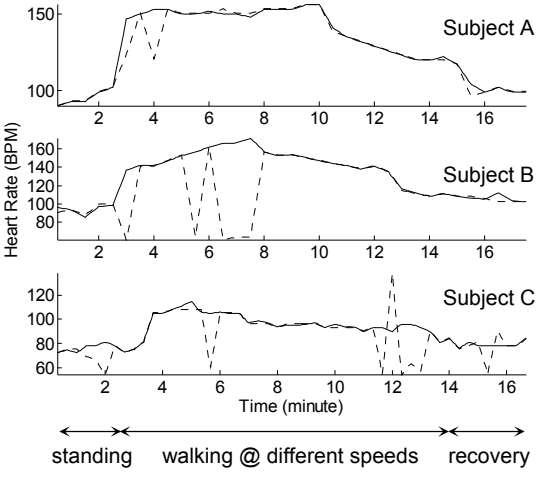


Fig. 5 Heart rate curves from 3 subjects after the motion cancellation method applied to the motion dataset. For standing and recovery stages there is no motion cancellation. Solid lines represented the Nellcor reference results. Dash lines represented results from the earpiece sensor. Different activities were time-stamped.

The heart rate curves shown in Fig. 5 matched well with the reference source for more than 80% of the overall monitoring duration. However, there were several detection errors. Possible reasons were, 1) subject's heart rate was close to his / her step frequency so the heart rate spectrum peak was removed by the motion cancellation algorithm. 2) temporary optical shielding problems.

Fig. 6 shows the dynamic f_{HRS} values from the e-AR PPG sensor when all 3 simultaneous reception channels were activated. For this experiment, the e-AR channel 1 achieved better heart rate signal fidelity than other two channels for most of the monitoring time. However, repetitive experiments from different subjects showed that the best-performed channel was not always Channel 1, *i.e.* for some experimental runs the Channel 2 or Channel 3 had a higher overall f_{HRS} value. This variation was because of the different shapes of the subjects' ears and the non-identical sensor treatments each time. In practice, the channel selection was completed during the calibration stage, *i.e.* all reception channels should be initially enabled, and then the less effective channel(s) were disabled after fidelity comparisons. In this way the power consumptions of the overall system were reduced as well.

Table 1 summarises all the experiment results. Results from anterior and posterior auricular regions were similar so only mastoid process result was presented. The first row represented the PPG signal ratio (a.c./d.c.) due to absorption in pulsatile blood (a.c.) vs. absorption in total tissue (d.c). The second row represented the relative strength of the pulsatile PPG. The finger and forehead results matched in range with literatures. Compared to the finger the PPG signals detected from the auricular skins were expected to be 1/10th in strength. Rows 3-5 compared the f_{HRS} values during different experiment stages. For earpiece sensor, during physical exercises the f_{HRS} values were approximately 10-15% lower than during still positions, because part of the PPG signal energy was 'leaked' from the heart rate frequency band into the pace rate frequency band.

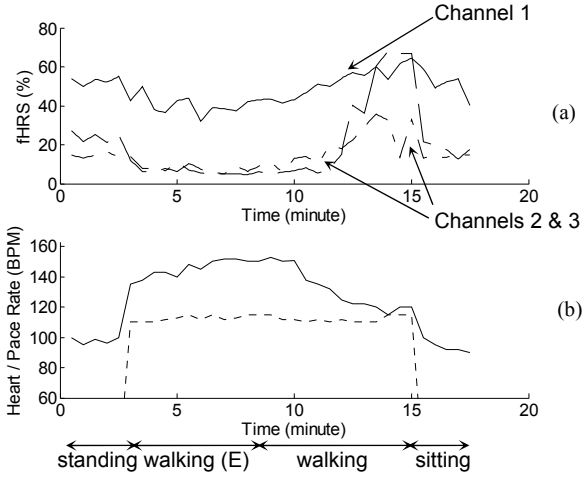


Fig. 6 f_{HRS} values from 3 simultaneous reception channels (a), and the correspondent heart rate (solid line in (b)). Pace rate was recorded around 110-115 BPM during walking periods (dashed line in (b)). Different activities were time-stamped. E denotes the treadmill was elevated to be at 10^0 in order to increase the physical exercise level (hence the heart rate) of the subject while the subject maintained a relative constant pace rate.

Table 1 Summary of experimental results

	Finger	Forehead	Mastoid Process	Superior Auricular region
PPG signal a.c./d.c. ratio at still positions	0.01-0.05	0.005-0.02	0.001-0.01	0.001-0.01
Relative PPG signal strength at still positions	1	0.2-0.3	0.05-0.2	0.05-0.1
Averaged f_{HRS} before walking*	n/a	65±5 %	40±10 %	35±15 %
Averaged f_{HRS} during walking*	n/a	50±10 %	15±10 %	25±10 %
Average f_{HRS} after walking*	n/a	65±10 %	40±10 %	40±10 %
Sensor modality	Patch	Patch	Patch	e-AR

* Averaged from the best-performed channel, not from all simultaneous reception channels.

IV. Conclusion

In this paper, a feasibility study is conducted for examining a reflectance based PPG e-AR sensor design for well-being monitoring. Experimental results have shown that the sensor is able to record PPG signals reliably from superior and posterior auricular skins with a much reduced LED driving current. With adequate optical shielding and motion cancellation, the earpiece sensor was capable of detecting dynamic heart rate changes both during rest and exercise.

Wearable sensors usually have stringent power constraints. Our design achieved reliable heart rate detections by using only 4-8 mA LED peak driving currents. By assuming a 25% lightning duty circle the equivalent continuous current was 1-2 mA per LED, this is low compared to a standard low-power commercial earlobe module, which specifies an LED peak current of 80 mA for shining through the earlobe (equivalent continuous current is 12 mA, ChipOX from MCC®, which is optimised for both SpO2 and heart rate measurements).

Preliminarily user trials have shown that the device is easy to wear and is ideally suited for general well-being monitoring. The heuristic motion cancellation scheme proposed has eliminated the effect of body motion, hence leading to a robust system design suitable for a range of pervasive sensing applications.

Acknowledgment

The authors would like to thank the Pervasive Computing research group members for helpful discussions and the voluntary experimental works. L. Wang would like to acknowledge financial supports from the UK EPSRC Biosensornet project and the UK DTI SAPHE project. The authors would also like to thank the anonymous reviewers for helping to improve the quality of this paper.

References

1. Yang GZ (2006) Body sensor networks. Springer-Verlag London Ltd
2. Thompson PD (2000) Exercise and sports cardiology. McGraw-Hill Publishing
3. Webster JG (1997) Design of pulse oximeters. Institute of Physics Publishing
4. Hertzman AB (1938) The blood supply of various skin areas as estimated by the photoelectric plethysmograph. Am J. Physiol. 124(2): 328-340
5. Rhee S, Yang BH, Asada HH (2001) Artifact-resistant power-efficient design of finger-ring plethysmographic sensors. IEEE Trans. on Biomed. Eng. 48(7): 795-805
6. Grajales L, Nicolaescu IV (2006) Wearable multisensor heart rate monitor, Proc. 3rd international. workshop on wearable and implantable body sensor networks, Cambridge, Massachusetts, USA, 2006, pp. 154-157
7. Maguire M, Ward T (2002) The design and clinical use of a reflective brachial photoplethysmograph, Technical Report NUIM/SS/--/2002/04, Nation University of Ireland, Maynooth, Ireland
8. Spigulis J (2005) Optical noninvasive monitoring of skin blood pulsations. Applied Optics 44(10): 1850-1857
9. Kyriacou PA (2006) Pulse oximetry in the oesophagus. Physiological Measurements. 27(6): R1-R35
10. Mendelson Y, Pujary C (2003) Measurement site and photodetector size considerations in optimizing power consumption of a wearable reflectance pulse oximeter, Proc. 25th IEEE EMBS, Cancun, Mexico, 2003, pp. 3016-3019
11. Reuss JL (2005) Multilayer modeling of reflectance pulse oximetry. IEEE Trans. on Biomed. Eng. 52(2): 153-159
12. Branche P, Mendelson Y (2005) Signal quality and power consumption of a new prototype reflectance pulse oximeter sensor, Proc. 31st IEEE annual NE bioengineering conference, 2005, pp. 42-43
13. Nogawa M, Tanaka S, Shibata M et al (2004) Development of a tissue oxygen consumption measurement method based on near-infrared photoplethysmography, Proc. 26th IEEE EMBS, San Francisco, USA, 2004, pp. 2227-2230
14. Wong A, Pun KP, Zhang YZ et al (2005) A near-Infrared heart rate measurement IC with very low cutoff frequency using current steering technique. IEEE Trans. on Circuits and Systems-I Regular Papers 52(12): 2642-2647
15. Drake RL, Vogl W, Mitchell AWM (2005) Gray's anatomy for students. Elsevier
16. Celka P, Verjus C, Vetter R et al(2004) Motion resistant earphone located infrared based heart rate measurement device, Proc. 2nd International Conference Biomedical Engineering, Innsbruck, Austria, 2004, pp 582-585

Address of the corresponding author:

Author: Dr. Lei Wang
Institute: Imperial College London
Street: 180 Queens Gate
City: London
Country: United Kingdom
Email: leiwang@doc.ic.ac.uk

Multi-Axis Inertial Measurement Units measuring human Posture and Motion

M. Trächtler[1], D. Hodgins[2], L. Kenney[3], M. Dienger[1], T. Link[1] and Y. Manoli[1,4]

[1] HSG-IMIT, Wilhelm-Schickard-Str.10, 78052 Villingen-Schwenningen, Germany
[2] ETB, Codicote Innovation Centre, Codicote, Herts, SG4 8WH, UK
[3] CRHPR, University of Salford, Salford M6 6PU, UK
[4] IMTEK, University of Freiburg, Georges-Köhler-Allee 102, 79110 Freiburg, Germany

Abstract— **This paper presents the development of systems monitoring human body motions and postures for clinical purposes. The hardware for these applications exploits a newly released commercial Micro-Electro-Mechanical (MEM) 3-axis accelerometer and a MEM 3-axis rate gyroscope being developed by HSG-IMIT. First the paper gives an overview of wearable 3-axis accelerometer systems and the corresponding data storage and transmission developed by ETB. The main part of the paper describes the design of the 3-axis rate gyroscope to be implemented in an Inertial Measurement Unit (IMU) for the body motion monitoring. Such health applications require IMUs of very low size to be able to fix the sensor cluster to the human being. This prevents the use of state-of-the-art IMUs implemented by three perpendicular orientated rate gyroscopes for this purpose. Thus a novel 3-axis rate gyroscope realised in one plane, on a single die, is being developed. With this device a reduction of the package size of multi-axial MEM sensors is achievable at least by a factor of ten. A further advantage of this approach is the reduced cost due to the omission of a spatial configuration and due to the small dimensions. Finally the synthesis of the 3-axis accelerometer and the 3-axis rate gyroscope to an IMU for monitoring 3D body segment orientation is addressed.**

Keywords— **human body monitoring, IMU, 3-axis rate gyroscope, inertial sensors, MEMS**

I. Introduction

The market growth of multi-axis IMUs using MEM rate gyroscopes and accelerometers is estimated to be about 15 % per year [1]. In this context the increasing number of health applications requiring inertial sensor-based systems plays a major role. The growth in these application areas including control of assistive devices [2-4], activity monitoring [5] and measurement of posture and motion [6, 7] comes along with the development of low power, low cost MEM inertial sensors, mobile computing and low power on-the-body communication systems.

This paper reports on the development of a novel inertial sensor and a range of inertial sensor based products for health applications that are emerging from the Healthy Aims project [8]. Healthy Aims is three years in to a four year FP6 project and is funded under IST. The project goal is to develop a number of intelligent medical implants and diagnostic systems integrating a range of underpinning Micro and Nano technologies. The 23 M€ project has 26 partners from ten EU countries, including seven SMEs, six clinical partners and a range of academics, LEs and research groups.

II. Accelerometer-based systems

ETB is developing wearable 3 axis accelerometer-based systems for human motion measurement and activity monitoring. In each of these applications the necessary information is inferred from accelerometer data alone. The first lab-based system, commercialised as the Libra unit (http://www.etb.co.uk/features.html), is a USB connected unit in which the accelerometer data is stored and processed on a computer.

The second system (storage activity monitor) logs the accelerometer data to a storage card and download periodically to a computer for processing.

Fig. 1 Storage activity monitor

Activity will be automatically classified as: sitting; lying; standing; walking; running; and for those that involve lo-

comotion, stride characteristics will be provided. Preliminary trials have started and results will be published in 2007.

For those applications where real time data is required, for example continuous monitoring of elderly patients who are likely to fall, a wireless activity monitor is being developed by ETB. The storage function will be replaced with a wireless chip set which will transmit the data, analysis will be done at the mobile phone and then summary data transmitted into the mobile network. This data can then be provided to remote clinics for monitoring purposes. This system is being developed under a nationally funded programme, in collaboration with Vodafone and Zarlink.

Another Healthy Aims application that makes use of only accelerometer data is a wireless sensor module for integration with an implantable Functional Electrical Stimulation system [3]. This system is being developed to assist patients with stroke in grasping objects. The arm-located sensor module incorporates a multi-axis accelerometer that provides an output dependent on the orientation and motion of the affected arm. Algorithms interpret these signals and identify appropriate points in the data at which to trigger the stimulation that opens the user's hand. As the output from the accelerometer sensor module is dependent both on location and orientation (i.e. pose) of the module on the arm, it is important to identify where to locate the sensors in order to maximise the chances of correct interpretation of the accelerometer signals. In order to support the designer in choosing the optimal pose for the sensor module, while keeping the burden on the patient to a minimum, software has been developed that allows the designer to simulate the output at different poses on the arm from a minimum number of patient trials [9].

This part of the paper has provided an overview of the work in Healthy Aims to develop applications based on the use of multi-axis accelerometers. The next part of the paper describes in detail the development of a new single die 3 axis rate gyro that opens up opportunities for the development of small, and hence clinically acceptable, devices incorporating multi-axis rate gyros.

III. 3 AXIS RATE GYROSCOPE

In body monitoring applications, in which body segment orientation (posture) as well as linear acceleration and angular rate are required an accelerometer gets combined with a gyroscope to an inertial measurement unit to gather the required information. This unit has to acquire angular rates of more than 1000 °/s and linear accelerations of more than 50 g to measure fast motions of the body and limbs accurately. The bandwidth will be up to 100 Hz depending on the body motion analysis.

Body monitoring applications require small sensor units of low cost with a minimum of performance loss due to minimisation. State-of-the-art IMU apply single-axis micromachined gyroscopes to prevent strong cross coupling and sensitivity to fabrication imperfections. However, the realisation of such an IMU requires a rectangular arrangement of single-axis gyroscopes. Problems concerning the assembly of such sensor clusters are the complexity of the packaging with its mismatch in sensor alignment, the immense costs based on the packaging, the high weight and the big dimensions of such a cluster. Regarding these aspects the future goal is the realisation of a multi-axis sensor cluster in one plane, ideally on the same silicon substrate, to eliminate the restrictions of spatial configuration of the sensors, as well as to reduce the package costs, weight and dimensions (Figure 2) [10].

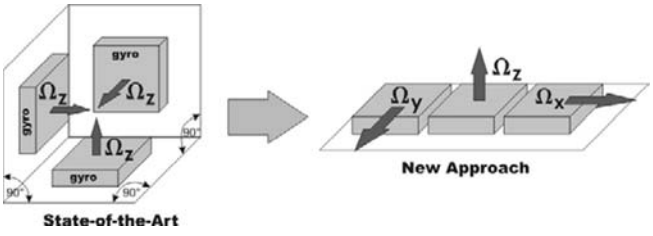


Fig. 2 Schematic view of the new approach for multi-axis Inertial Sensor Units on a single silicon die

Whereas tri-axial micromachined accelerometers are already on the market tri-axial gyroscopes are still in the development phase. This fact might mainly be caused by the higher complexity of gyroscopes compared to accelerometers. MEMS Coriolis Vibrating Gyroscopes (CVG) need two oscillations rectangular to each other to detect angular rates about their sensitive axis perpendicular to both oscillation axes: a driven oscillation (excitation, primary mode) and a detection oscillation (secondary mode). The last-mentioned oscillation is induced by Coriolis forces due to the application of an external angular rate (Ω) in presence of the primary oscillation (Figure 3). Hence, the realisation of tri-axial gyroscopes requires in-plane as well as out-of-plane oscillations to implement in-plane and out-of-plane sensitive axes. State-of-the-art CVG realise the oscillations in general 'in-plane' while the sensitive axis is 'out-of-plane'. Therefore, the main task achieving tri-axial gyroscopes was the development of a sensor concept and technology enabling not only in-plane motion but out-of-plane motion. Details of this multi-axial gyroscope, developed at HSG-IMIT are given below.

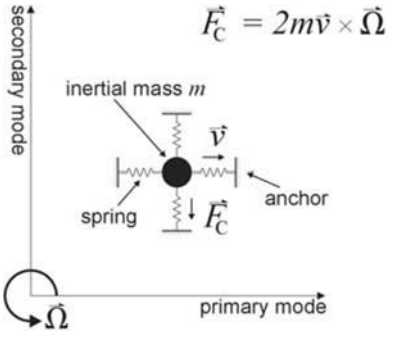


Fig. 3 General spring-mass model of a CVG

A standard Silicon-On-Insulator-(SOI)-technology prepares special SOI substrates by bonding a handle wafer with pre-etched cavities to a top wafer. The latter will be thinned to an active layer thickness of 50 µm (device wafer). The concerning wafer is normally structured with only one photo resist mask and defines later the sensor area. Wafer-scale packaging with Silicon fusion bonding or glass frit bonding encapsulates low pressure and protects the sensor against contamination and humidity. Based on this standard SOI process for realising in-plane oscillations, a double mask layer realises a second level of the movable comb structures with time-controlled Deep Reactive Ion Etching (DRIE) process. The inhomogeneous electrostatic field of these comb structures induces a force and capacitive change and will lead the movable structure to out-of-plane oscillations (Figure 4).

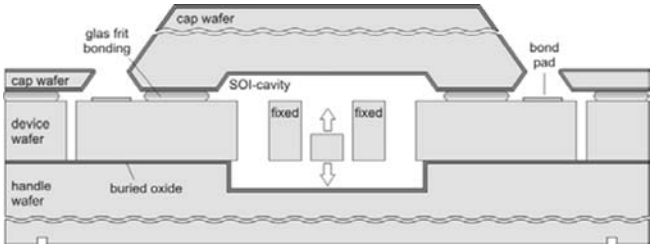


Fig. 4 Schematic cross sectional view of the modified SOI-technology

A critical technological step of this sensor principle might be the realisation of the second level of the comb structures with time-controlled DRIE. The height of the reduced finger structures depends on the homogeneity of the etching process and the etch rate variation, but also on the Total Thickness Variation (TTV) of the SOI wafer. However, simulations show that within the operating range of adequate primary oscillation amplitudes of ± 4 µm, the influence of technology induced different heights on the electrostatic force is negligible [11].

Besides the realisation of the out-of-plane drive the suspension of the drive mass plays an important role for the sensor performance especially with respect to the achievable linearity. Realising the flexure beams for the vertical motion using a thinning process to reduce the height of the beams, the time controlled etching process together with the TTV of the SOI wafer lead to high frequency shifts and therefore a quite sophisticated process flow has to be used to achieve the resonance frequencies. Using the torsion mode of the structures [12], no thinning process is needed and thus no additional shift in resonance frequency is given. Unfortunately, a hardening spring phenomenon due to the torsion mode, results in a high non-linearity of the out-of-plane oscillation. The stiffness increases in relation to the motion amplitude and in comparison to in-plane oscillations using bending effect, the non-linearity of torsion beams for vertical motion is up to two orders of magnitude higher. The realisation of non-thinned torsion beams therefore needs a beam design concept with reduced non-linearity to reach the order of magnitude of the in-plane oscillating bending beams.

By using the IMIT-patented double decoupled gyroscope concept quadrature effects, arising due to the out-of-plane motion, can be neglected [13]. Due to this arrangement of the springs only the electrostatic driven element and the seismic mass are excited to a linear oscillation along the z-axis (primary mode) and the seismic mass as well as the detection element (detection unit) execute the secondary oscillation. Ideally, in this way the secondary oscillation does not influence the driving mechanism and the primary oscillation, as well as parasitic effects of the comb drives, are suppressed.

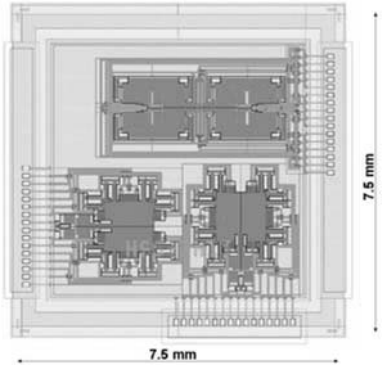


Fig. 5 Layout of a multi-axial gyroscope

With the realisation of a 3 axis gyroscope in one plane on the same silicon substrate, performance loss due to the assembly mismatch of spatial configuration of single axis sensors can be neglected. Fig. 5 illustrates the small size of

this 3-axis gyroscope sensor chip. A reduction of the package size of multi-axial MEMS sensors is achieved at least by factor 10 (2 cm³ to 0.2 cm³). In combination with high-density integrated readout electronics (ASIC) ultra-small, low weight IMUs can be build up which fulfil the high requirements of measuring human posture and motion.

IV. Inertial measurement unit (IMU)

The single die rate gyroscope developed by IMIT and discussed above, is being exploited in an inertial measurement unit, developed within Healthy Aims. The IMU, that will be made by ETB, will log sensor data to a memory storage card which will be downloaded periodically for data processing on a computer.

In contrast to some other IMUs that use a magnetometer to provide a reference aligned with the earth's magnetic field, the Healthy Aims IMU will only use data from the IMIT rate gyro and a commercially available (Kionix) 3 axis accelerometer to estimate body segment 3D orientation over time. The decision to not rely on magnetometer inputs was taken early on in the project to avoid the challenging problems introduced by the heterogenous magnetic field present in most built environments. However, such an approach cannot completely eliminate drift of the output about the gravity vector.

The work began with the premise that the unit would be designed to be used either as a single segment IMU or as part of a series of IMUs on interconnected limb segments in which software would take advantage of the joint constraints to improve accuracy. The single unit estimator is based on an extended Kalman filtering approach and simulations using realistic inputs derived from measured human motion data together with some early assumptions on sensor characteristics demonstrated good performance in terms of orientation estimation about the axes orthogonal to the gravity vector and a low drift rate about the gravity vector. Simulations suggested that in most representative situations the orientation accuracy about axes orthogonal to the gravity vector was within 3°. Drift rate error about the vertical axis in most cases was within 0.15 °/s.

The work is now being extended. In particular, a so-called "high level estimator" is being developed, which uses information from IMUs on multiple, interconnected limb segments together with knowledge of joint constraints to improve estimator accuracy. This work is focusing on the lower limb at present, but will be sufficiently generic to be extended to other body segments in the future.

V. Conclusion

The Healthy Aims project has produced a number of multi-axis accelerometer based devices for a range of health related applications, Functional Electrical Stimulation Systems and an activity monitor. In recognition of the size issues with current 6 Degrees Of Freedom (DOF) IMUs, IMIT are developing a single die rate gyro that opens up the opportunity to considerably reduce the size and hence increase the user-acceptability of future IMUs. The first of the 3 axis rate gyros will be available in Spring 2007 and tested at the University of Salford human motion labs. The IMIT gyro will be integrated with a commercial 3 axis accelerometer in a new IMU being developed by ETB.

Acknowledgment

This research project is funded by the EC project Healthy Aims of the European Information Societies Technology Programme (IST-2002-1-001837).

References

1. Yole Development, WISM'05: World Inertial Sensors Market, 2005
2. I. P. Pappas, M. R. Popovic, T. Keller, V. Dietz, and M. Morari, "A reliable gait phase detection system," IEEE Trans Neural Syst Rehabil Eng, vol. 9, pp. 113-125, 2001.
3. A. T. Willemsen, F. Bloemhof, and H. B. Boom, "Automatic stance-swing phase detection from accelerometer data for peroneal nerve stimulation," IEEE Trans Biomed Eng, vol. 37, pp. 1201-8, 1990.
4. P. Tresadern, S. Thies, L. Kenney, D. Howard, J. Goulermas, "Artificial neural network prediction using accelerometers to control upper limb FES during reaching and grasping following stroke," presented at 28th IEEE Int. Conf. Engineering in Medicine and Biology Society (EMBS), New York, USA, 2006.
5. D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," IEEE Trans Inf Technol Biomed, vol. 10, pp. 156-67, 2006.
6. D. Giansanti, V. Macellari, G. Maccioni, and A. Cappozzo, "Is it feasible to reconstruct body segment 3-D position and orientation using accelerometric data?," IEEE Trans Biomed Eng, vol. 50, pp. 476-83, 2003.
7. A. M. Sabatini, "Quaternion-based extended Kalman filter for determining orientation by inertial and magnetic sensing," IEEE Trans Biomed Eng, vol. 53, pp. 1346-56, 2006.
8. Healthy-Aims, Implantable microsensors and microsystems for ambulatory measurement and control in medical products, IP IST-2002-001837
9. Thies SB, Tresadern PA, Kenney L, Howard D, Goulermas JY et al A "virtual sensor" tool to simulate accelerometers for upper limb FES triggering J Biomech, vol 39 (S1), pp. S80, 2006
10. M. Trächtler, et al., Adapted SOI Technology for Multi-Axis Inertial Sensors, Micro System Technologies, 5.-6. Oktober 2005, München, Proceedings, pp.141-147

11. M. Trächtler, et al., A new Approach for Multi-Axis Inertial Sensor Units on a Single Silicon Die based on SOI-Technology, Proc. Symposium Gyro Technology 2006, Stuttgart, Germany, Sep. 2006
12. J. Kim, et al., A Planar, X-Axis, Single-Crystalline Silicon Gyroscope Fabricated Using the Extended SBM Process, Proc. IEEE MEMS, pp 556-9, Maastricht, The Netherlands, Jan. 25-29, 2004
13. Hahn-Schickard-Gesellschaft, Drehratensensor und Drehratensensorsystem, Offenlegungsschrift DE 100 40 418 A1

Address of the corresponding author:

Author: Martin Trächtler
Institute: Institut für Mikro- und Informationstechnik der Hahn Schickard-Gesellschaft e.V. (HSG-IMIT)
Street: Wilhelm-Schickard-Str.10
City: 78052 Villingen-Schwenningen
Country: Germany
Email: martin.traechtler@hsg-imit.de

In-Ear Acquisition of Vital Signs Discloses New Chances for Preventive Continuous Cardiovascular Monitoring

O. Brodersen[1], D. Römhild[1], D. Starke[1], A. Steinke[1] and S. Vogel[2]

[1]CiS Institut für Mikrosensorik GmbH, Erfurt, Germany
[2]Lehrstuhl für Medizinische Informationstechnik, RWTH Aachen, Germany

Abstract— **For preventive continuous cardiovascular monitoring an In-Ear-implemented system will be designed. Using the photoplethysmographic curve a personal In-Ear-sensor measures the physiological parameters. The data are transmitted to a portable data analyser via wireless network, automatically reviewed and in the case of passing a critical value, an alert will be sent out. This In-Ear acquisition is facilitated by a new remission sensor designed by CiS Institut für Mikrosensorik GmbH. This In-Ear sensor is the core component ensuring light emission and reception by planar set-up and therefore, it is capable to use the inherent advantages of this measuring site for vital signs acquisition. Initial results have supported this approach.**

Keywords— **vital signs monitoring, In-Ear sensor, remission principle**

I. Introduction

Cardiovascular diseases are one of the major contributors to the health statistics in the industrial countries and therefore an early evaluation in the general context of cardiovascular risk factors is required [1]. Currently available state-of-the-art methods of vital signs estimation are complex or need sophisticated technical equipment which limits their application to the clinical and ambulatory use.

Recent improvements in signal acquisition and processing techniques have motivated the apparition of an emerging trend in continuous vital signs measurements [2].

Our novel "In-Ear sensor" approach attempts to measure vital signs extracted from optical signals which are acquired in the ear channel and their wireless transmission with a motion-artifact cancellation mechanism.

Further oximetry has become an essential tool in the modern practice of emergency medicine, used to determine trends in patients' blood oxygen saturation, the boost of electrolytes imbalances, and thus the warning of dangerous cardiopulmonary saturation levels (i.e. hypoxia, chest pain, therapies effects, etc.). However, conventional pulse oximetry has some inherent limitations related to motion artifacts, perfusion, ambient light, vasoconstriction and others [3]. Provided that a sensor could be made available which would be able to send artifact-resistant (especially motion artifact-resistant) measuring data by implementing in the ear channel, such solution would lead to an essential improvement in clinical and ambulatory patient care.

II. In-Ear Acquisition by Means of Remission Sensors

A. Arguments for In-Ear acquisition

Applications of micro-optical sensors in the ear channel operating on the principle of remission state establish an innovative and highly interesting alternative compared to the well-known measuring methods involving fingers, toes or ear lobes [4]. The main reason for using the ear channel as a measuring site is given by providing a series of conditions which ensure reproducible and artifact-resistant measurements of vital signs in particular in motion too.

The following essential arguments can be highlighted:

- High temperature stability in the ear channel,
- no significant muscles causing motion artifacts,
- good micro-capillary tissue perfusion in state of shock, too
- no disturbing pigments in the skin and no hypodermic fat tissue,
- good fixing into the ear channel and,
- no relevant hydrostatical pressure modulation in case of measurements during motion.

The following graph (Figure 1) shows in terms of finger application how variation of distances between heart and measuring site at the finger affects the results of vital signs.

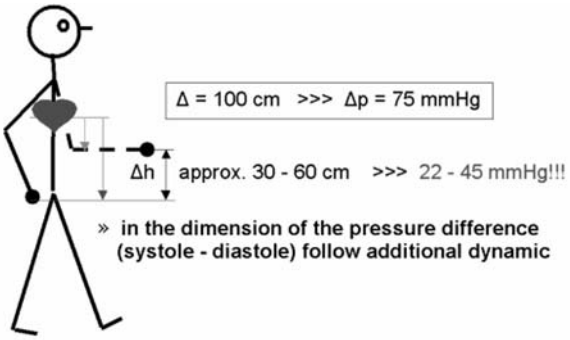


Fig. 1 Hydrostatical pressure modulation related to finger application

Using open otoplastics the In-Ear sensors leads only to marginal reduction of the physiological ear function.

Given that it is possible to manufacture adequate small remission sensors, the new In-Ear approach discloses an excellent chance for applications to make a non-invasive optical diagnosis.

B. MORES® basis modules

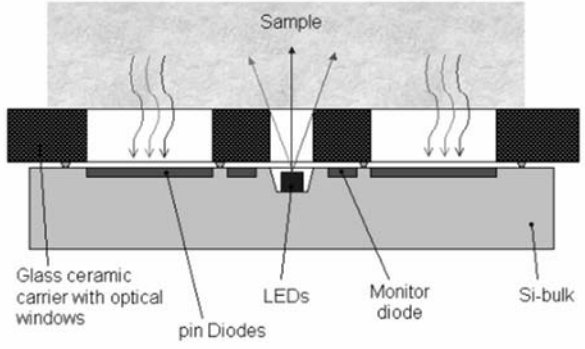


Figure 2 represents the ***M**icro-**O**ptical Remissions-/ **Re**flexions-**S**ensor system* MORES®[i] schematically.

Fig. 2 MORES® cross section implanted into silicon (Si) substrate

The silicon sensor chip contains two pin receiver diodes which are arranged symmetrically around the centre, a monitor-diode as well as an optically shielded diode, which provides chip temperature measurements. Additionally this chip has a 200 µm deep etched cavity in the centre, into which several LEDs can be installed as bare chips. Between the LEDs and the pin-diodes there is an optical barrier in the glass to minimize the optical cross-talk. Up to 3 LEDs can be integrated in such a sensor-system. The respective emitter-wavelengths are defined by LED selection. The integrated monitor-diode permits simultaneous, independent remissions/ reflections measurements of the radiation-performance of the LEDs and thereby, for example, the compensation of temperature-dependence and drift of the emitted radiation. The surface of the sensor-system consists of flat and chemically resistant glass. This generic module is suitable for using optical principles such as reflection, absorption, or fluorescence.

The entire transceiver-module can be soldered with Flip-Chip-Technology onto the glass-ceramics carrier as a quasi flat planar chip. Additionally, the reflective walls of the cavity can increase the useful light input into the object to be measured.

The glass-ceramics body has 3 functions:

- as assembly plate with optically smooth touching surface
- as connection and wiring carrier for the sensor
- for suppression the optical cross-talk from the transmitter to the receiver by optical shields.

In Figure 3 an emitter module (LED-bare chip) implanted into the silicon substrate is shown. The emitter modules are surrounded by a photo-diode-array with the aim of receiving the reflected light. This technology guarantees a high reproducibility and reliability.

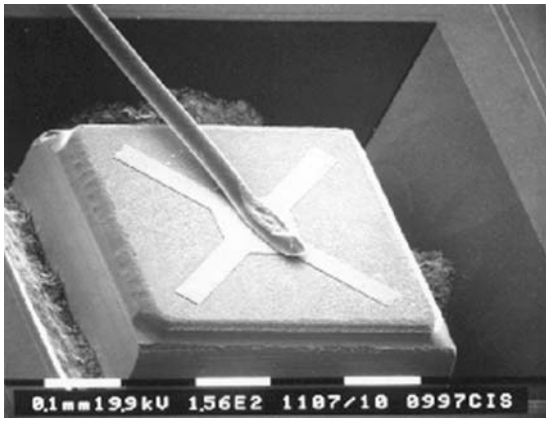


Fig. 3 A LED assembled in silicon

A basic problem of all integrated reflection / remission sensors is the coupling of scattered light in the protection and covering layers, which causes optical cross-talk and resulting negative effects in the dynamics range and proof-sensitivity. A new technology concept was developed and successfully applied with the MORES®-sensors using photo-sensitive glass. The standard Borofloat-glass medium was replaced by a specially processed, photo-sensitive glass, for example FOTURAN®[ii]. By precisely directed ultraviolet (UV) exposure and subsequent annealing ceramized volume areas were generated. The ceramised volumes

are highly absorbing and can be used as optical aperture shield between light sources and the pin-diodes (compare Figure 2).

The attainable optical suppression is both dependent on the wavelength of the radiation, as well as on the effective aperture (Figure 4).

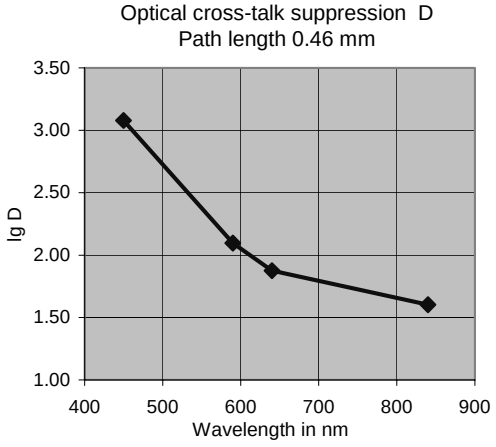


Fig. 4 Wavelength dependence of the suppression

A progressive suppression is observed at shorter wavelengths, which compensates at least partially for the simultaneous drop of the quantum yield of the Si detectors. Suppression factors $>10^3$ can be achieved for high sensitivity applications in the near UV-area (for example, fluorescence).

C. Different optical sensor modes

In accordance to Figure 5, the information gained by the emitter-receiver is based on:

- Change of the reflection at an external object by displacement, distance- or angle-change (1)
- Change of the spectral absorption in an indicator-layer caused by analyte effects (2). The beam is reflected at the upper surface of the layer (preferably by total reflection) and must pass the indicator-layer twice.
- Change of the scattering or remission in a target placed on the module (3). Advantageously, strongly absorbing areas can be positioned in the glass-medium in order to suppress large unwanted bias caused by radiation bypassing the object and reaching the receivers via total reflection (5).
- Change of the surface reflection interference in a thin layer, caused by change of its optical thickness (4). Analyte effects lead to swelling or change of the refractive index of the sensitive layer.
- Change of the limiting angle of total-reflection (5) on the upper glass surface caused by changes in the surrounding medium.

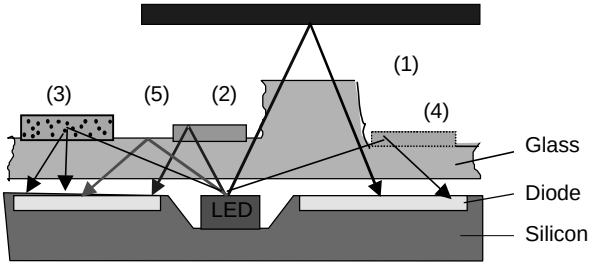


Fig. 5 Light path for various recognition systems

D. Remission principle for medical application

Beer's Law describes how the "blood oxygen saturation ratio measured by pulse oximeter" (SpO_2) is derived from extinction coefficients of fully saturated hemoglobin (oxyhemoglobin, HbO_2) and of reduced hemoglobin (deoxygenated hemoglobin, Hb) [5]. Reflectance sensors measure these extinction coefficients of tissue covered by the sensor. Therefore, light of two frequencies is directed into the tissue and the amount of reflected (scattered) light is acquired using a photo diode. The signal to noise ratio (SNR) and thus, the precision, is maximized when wavelengths are selected, for which the extinction coefficients of HbO_2 and Hb differ significantly (e.g. 660 nm and 940 nm, resp., see figure 6) [5].

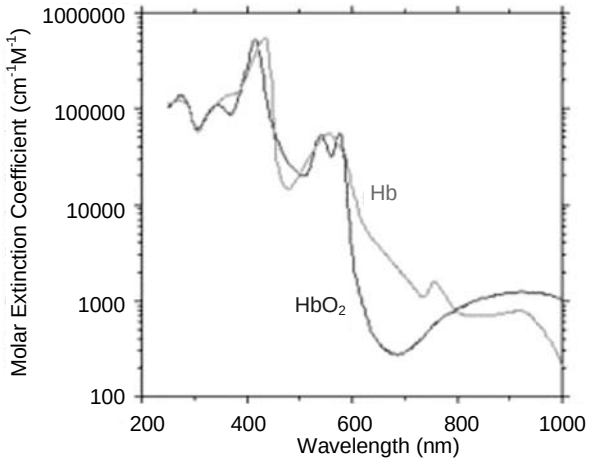


Fig. 6 Absorption spectrum of oxygenic and anoxemic haemoglobin [6]

An essential advantage of this measuring method is the fact that the application is independent of the measuring site.

A basic issue affecting the performance of remission sensors is the parasitic cross-talk between light source and receiver.

The MORES® technology developed by CiS overcomes this problem by use of light absorbing glass-ceramics.

E. Sensors for medical applications

CiS has designed and prototyped first remission sensors specified to medical applications, e.g. as Two-Beam-Photometer, sensors with different optical pathway of tissues and further, for vital signs measurements at fingers but also as In-Ear sensor [7].

The following figures show

- The vital signs signal acquisition and processing for a remission sensor based on MORES® technology for a finger application (Figure 7) and
- An In-Ear sensor prototype (otoplastic based) with miniaturised optical remission sensor (Figure 8).

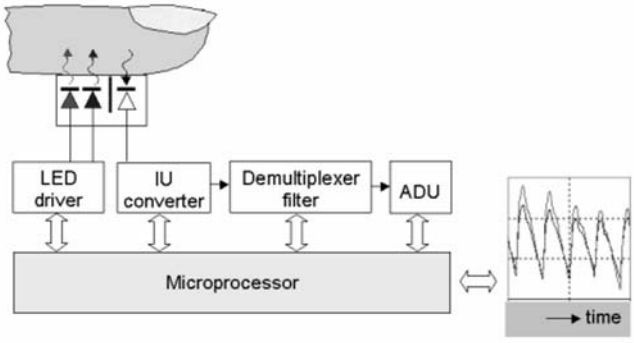


Fig. 7 Signal processing

Fig. 8 In-Ear sensor prototype

New remission sensors for further medical applications are in preparation.

F. Vital signs acquisition by the In-Ear sensor approved

Time-resolved measurements of the red and near-infrared remission acquired from the ear channel of probands have basically proved the power and capability of the In-Ear sensors.

These first In-Ear measurements were carried out by means of the MORES® sensor type POX.

The following Figures 9 and 10 show the measured photoplethysmographic curves to estimate vital signs on two levels:

1. Resting position of the proband (Figure 9) and
2. Proband under exercise (Figure 10)

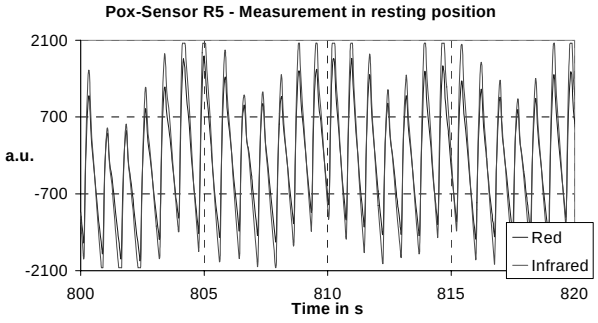


Fig. 9 Time-resolved remission measurement in resting position

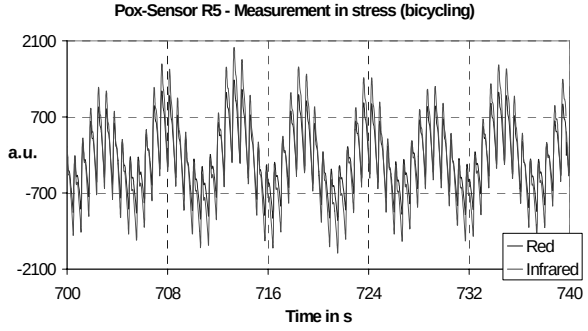


Fig. 10 Time-resolved remission measurement under exercise

G. Validation of the In-Ear sensor

In order to show that the data acquired by the In-Ear sensor are comparable to data provided by clinically used equipment, parallel measurements were accomplished with two healthy subjects.

As reference setup, an EnviteC SpO_2 finger clip sensor was attached to the left middle finger of the subject and connected to a measurement PC via an MCC ChipOx pulse oximeter (all measurement values of this reference setup are indexed by "REF"). In order to achieve maximum correlation, the In-Ear sensor was attached to the index finger as shown in figure 8 and the data were sent to the measurement PC by our measurement hardware (indexed by "device under test" = "DUT"). The data of both sensors were recorded simultaneously and stored in one file for each single measurement.

Spread over several days, 22 single measurements were performed on two healthy subjects (Subject A: female, age 51 years; Subject B: male, age 29 years). Each measurement lasted 1:30 minutes, during which the test person sat relaxed on a chair, breathed normal, did not talk and tried to stop moving.

Each single measurement was parted into three epochs of 30 seconds. For each epoch, the power spectrum density (PSD) was estimated for the ChipOx plethysmogram and for the red and infrared absorbance measured by the In-Ear sensor. Using the PSD values within the range of 0.2 – 20 Hz the following quantities were calculated:

- "Pulse": the frequency of the peak value in the spectrum (corresponds to the mean heart rate per epoch, see figure 11),
- Pearson's Correlation Coefficient r and the p-value for the PSD functions (gives a measure for similarity between acquired signal shape of the ChipOx plethysmogram compared to the MORES® red/infrared channel).

The following values were determined using the data of all epochs:

- Pearson's Correlation coefficient and p-value for comparing the mean heart rate per epoch, determined by the reference device, to the "Pulse" value described above (results: table 1),
- the mean and standard deviation of the correlation coefficients r and p-values for the PSD functions (results: table 2)

Figure 11 shows the heart rate ("Pulse") calculated from the PSD spectra (symbols +, x, o) and the average heart rate per epoch determined by the reference device (symbol *).

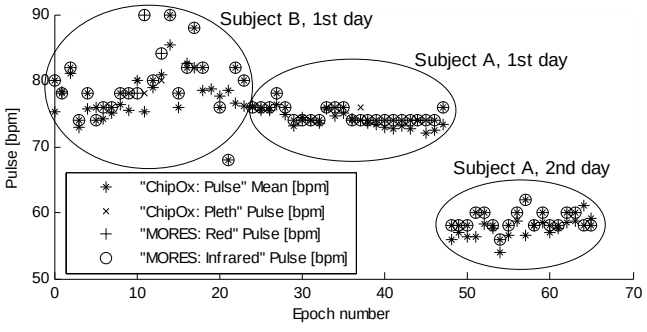


Fig. 11: Measured heart rate per epoch: * pulse signal calculated by reference device, x heart rate calculated from PSD of reference device, + heart rate from PSD of red channel of the In-Ear sensor, o heart rate from PSD of infrared channel of In-Ear sensor. Epochs 0-23: Subject A, epochs 24-

Table 1 shows the correlation coefficient r and the p-Values for comparing the mean heart rate per epoch determined by the reference device to the heart rate per epoch derived from the PSD of the red and infrared absorbance measured by the In-Ear sensor:

Table 1 Correlation of mean heart rates calculated by the ChipOx compared to the heart rate derived from red channel (";R,DUT") and infrared channel (";IR,DUT") of the device under test (In-Ear sensor).

Value	r	p
$Pulse_{REF;R,DUT}$	93.7 %	< 0.001
$Pulse_{REF;IR,DUT}$	92.2 %	< 0.001

Statistical moments of the correlation coefficient r for the PSD calculated for each epoch are shown in table 2.

Table 2 Correlation of PSD of the reference ChipOx plethysmogram compared to the PSD of the red channel (index ";R,DUT") and infrared channel (index ";IR,DUT") of the device under test (In-Ear sensor). The statistical moments are calculated over all epochs.

Value	**Mean(r)**	**σ(r)**	**p**
$PSD_{REF;R,DUT}$	98.34 %	2.04 %	< 0.001
$PSD_{REF;IR,DUT}$	97.92 %	2.50 %	< 0.001

The results presented show that the signals acquired by the In-Ear sensor correlate well to the plethysmogram measured by the reference platform. Furthermore, the heart rate determined from the PSD functions of the In-Ear sensor measurement values are similar to the heart rate determined by the reference sensor.

III. Conclusions

Based on the CiS remission sensor know-how for In-Ear measurements the partners EnviteC-Wismar GmbH, CiS Institut für Mikrosensorik GmbH Erfurt, Audia Akustik GmbH Sömmerda and RWTH Aachen will focus their efforts on the joint project IN-MONIT (www.in-monit.de) to design and manufacture the first universal and flexible In-Ear sensor adaptable for different individual ears.

Using the photoplethysmographic curve the new In-Ear sensor will be capable - in combination with other body sensors - to provide motion-resistant initial data for the vital signs, in particular *oxygen saturation, pulse frequency, perfusion, heart rhythm, heart rate variability, breathing rate and blood pressure variability* to monitor non-invasively and continuously the cardiovascular functions under on-line conditions.

Essential challenges of the new monitoring system IN-MONIT will be:

- The multiparameter data analysis,
- Evaluation of the vital signs derived from the In-Ear data,
- Artefact compensation,
- Hard- and software solutions for analysing the In-Ear data,
- An advanced energy management and
- An universal and flexible otoplastic combined with an advanced sensor integration and
- Implementation of a patient-tailored alert solution.

As a result of the development of the IN-MONIT system a new In-Ear sensor will be designed specified to the measuring conditions in the ear channel.

Acknowledgment

The IN-MONIT project is financially supported by the Federal Ministry of Education and Research in Germany under the grant number 16SV2259.

The consortium thanks the project executing organisation VDI/VDE Innovation + Technik GmbH, Mrs. Christine Weiss, for her technical assistance.

References

1. Cornélissen G, Delmore P, Halberg F (2004) Why 7-Day Blood Pressure Monitoring? Healthwatch 3, Halberg Chronobiology Center, University of Minnesota, at http://www.phoenix.tc-ieee.org
2. Meigas K, Kattai R, Lass J (2001) Continuous blood pressure monitoring using pulse wave delay. Proc. 23rd Annual International Conference of the IEEE-EMBS 4: 3171–3174.
3. Moyle J T (2002) Pulse Oximetry. BMJ, London
4. Barker S J (2002) "Motion-Resistant" pulse oximetry: A comparison of new and old models. Anesth. Analg. 95 (4): 967-972
5. Webster J G (Ed.) (1997) Design of Pulse Oxymeters. IOP Publishing, Bristol, UK
6. Ernst, N (2002) Entwicklung und Evaluierung eines mikrooptischen Sensorsystems für die nichtinvasive Erfassung von Herz-Kreislauf-Parametern im Gehörgang. Diploma thesis, University of Applied Sciences Jena
7. Römhild D, Steinke A, Waldmann J (2002) Mikrooptischer Sensor für schnelle Herz-Kreislauf-Diagnostik 11. ITG/GMA-Fachtagung Sensoren und Mess-Systeme, Ludwigsburg, Germany, 2002, pp 411-414

Address of the corresponding author:

Author: Olaf Brodersen
Institute: CiS Institut für Mikrosensorik GmbH
Street: Konrad-Zuse-Strasse 14
City: Erfurt
Country: Germany
Email: obrodersen@cismst.de

[i] MORES® is a registered trademark of the CiS Institut für Mikrosensorik GmbH

[ii] FOTURAN® is a registered trademark of the Schott AG.

Micro Capacitive Tilt Sensor for Human Body Movement Detection

L. Zhao, E. M. Yeatman

Optical and Semiconductor Devices Group, Department of Electrical & Electronic Engineering
Imperial College London, SW7 2AZ, UK
li.zhao@imperial.ac.uk

Abstract— **Tilt sensing is important for human body motion detection and measurement. Two tilt sensors are introduced in this paper, based on MEMS (Micro-electro-mechanical Systems) variable capacitors, and utilizing the gravitational effect on a suspended proof mass to detect inclinations. A symmetric comb structure with high aspect ratio is adopted to obtain high capacitance. The first device can achieve a full range (–90° to +90°) tilt angle detection and relax the high-resolution requirement of the readout system by its linear output characteristics. Based on the same concept, a novel inherently digital sensor is proposed. The digital signal can be read out without complex processing, and so low power consumption can be achieved. A fabrication process, and simulation and processing results, are presented.**

Keywords— **tilt sensor, comb drive, suspension beam, differential capacitance, MEMS.**

I. Introduction

Tilt is an important parameter in many motion detection applications, including the study of human body motion, currently a topic of wide interest within biosensor design. Tracking the movement of different parts of the body can help to provide important information such as the recovery status of joint injuries, movement patterns of athletes (Fig. 1), and the sleeping patterns of insomniacs. Because tilt sensors can be highly compact, they can be combined with other motion detectors and chemical sensors, e.g. for glucose or pH, in one package, to measure physical and biochemical changes simultaneously.

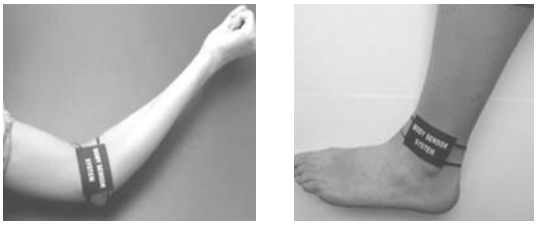


Fig 1 Examples of body movement detection; wireless sensor packs are provided at limb joints to monitor several motion parameters continuously.

Most tilt sensors reported in the literature are of capacitive type, for which we can obtain linear and analog outputs with respect to tilt angle [1]. An absolute angular encoder based on the capacitive coupling between a stator and rotor electrode has been presented in 1991 [2]. Bantien [3] invented a micromechanical tilt sensor with a moveable silicon mass mounted in a cavity using conventional micromechanical fabrication methods. This provides an alternative to conventional mercury-based sensors, avoiding the cost and toxicity inherent in mercury use. Another flat plate capacitive method for accurate angle measurement is reported in [4]. In this angle sensor, the displacement of the moveable electrode is driven by Lorentz force using a magnetic field. Changing the permittivity of the dielectric material is also a possible method for tilt angle detection [5].

While many such devices have been presented and various methods have been reported to improve their performance, the detecting range and resolution are still limited. MEMS is a good solution to extend the limitations and realize miniaturization and low cost. Although wireless body-mounted devices are the initial target, implantable variants are also a possibility, in which case minimizing size and power consumption are naturally critical requirements.

Some accelerometers can also be used as tilt sensors [6, 7]; however, the resolution is still not high enough, and the range is limited in tilt detection, due to the nonlinearity. Many accelerometers have comparatively high power consumption, which is not ideal in low power budget applications such as wireless sensor nodes. The tilt sensors introduce here are designed with low power consumption in mind.

The device fabrication is based on bonded silicon on insulator substrates, with front and back etching to release the moving parts. An inherently digital design is introduced which can simplify the read-out circuitry, or as a combined analog-digital device can maximize the range-to-resolution ratio.

II. Device Concept

The gravity driven tilt sensor consists of a central proof mass, comb drive capacitors, and suspension beams. A 3D illustration of the device concept, as modeled in CoventorWare, is shown in Fig 2.

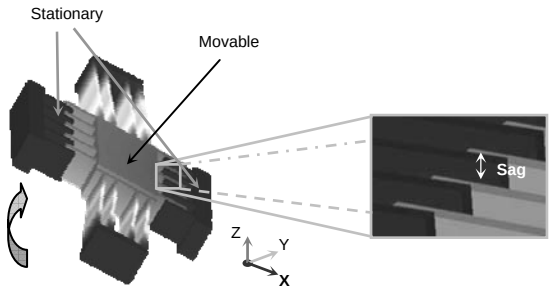


Fig 2. Operation principle of the tilt sensor.

The suspension is designed to have maximum flexibility in one in-plane direction (shown here as *x*), while being stiff in the other in-plane dimension, as well as for out-of plane motion and rotations. When tilt occurs, gravity forces the central mass to move toward one or other of the adjacent stationary comb electrodes. Hence the capacitance increases at this side, while decreasing at the other. The differential capacitance ($|C_{right}-C_{left}|$) accordingly provides a measurable parameter relating to tilt.

The comb capacitor is an efficient solution for obtaining both high capacitance and miniaturization. Operating as a displacement sensor, the key parameters of a lateral comb drive are the capacitance and lateral displacement, which are given by the following equations:

$$C = \frac{(l_o \pm \Delta x)h\varepsilon}{G} N = \frac{A_o \varepsilon}{G} N \qquad (1)$$

$$\Delta x = mg \sin(\theta) / k_x \qquad (2)$$

where C is the capacitance; l_o is the initial overlap length of the finger pairs; h is the thickness of the devices; G is the width of the gap between fingers; N is the number of finger pairs; ε is the permittivity of the medium between fingers; $A_o=(l_o \pm \Delta x)h$ is the overlap area; mg is the gravitational force on the mass; k_x is the spring constant in the x direction and Δx is the lateral displacement.

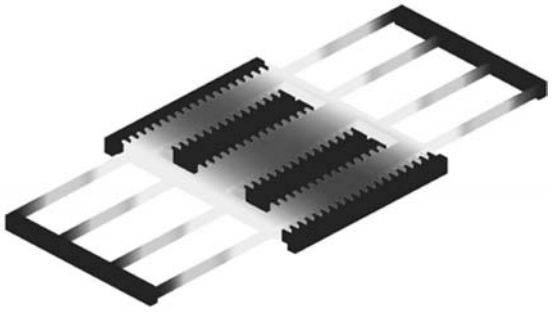
As can be seen, besides the characteristic suspension (k_x), the critical issues in achieving large capacitance variation are the aspect ratio (AR=h/G) and N. A novel etching technique described in [8] can achieve very high AR by a two step etching process. To increase the number of finger pairs, a few more sets of combs can be added without enlarging the size of the sensor.

The suspension beam design is a key issue in this sensor. Therefore, a large number of beam structures have been analysed [9, 10]. An important trade-off is that as the rigidity decreases, greater displacement (and thus greater capacitance variation) can be achieved, but the suspension will be larger, and will also allow more motion in unwanted directions, particularly sag, as shown in Fig 2.

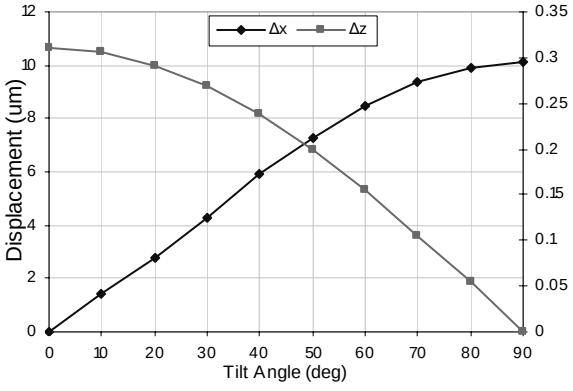
III. Device Structures

A. Initial Design

By applying the double hammock suspension to restrict the out of plane twist and adding extra comb fingers within the body of the proof mass to increase the capacitance, we arrive at the structure shown in Fig 3(a).



(a)



(b)

Fig 3 (a) Device design; (b) displacement in both x and z direction vs. tilt angle (the secondary axis is for Δz).

With the dimensions in Table 1, simulation results give the lateral and vertical displacements shown in Fig 3(b).

Table 1
Dimensions of the tilt sensor

G	2μm	w*	2μm
h	50μm	N	250 / 208
l_0	15μm	Beam	$5000\times3\times50\mu m^3$
L*	30μm	Size	$0.25\ cm^2\times50\mu m$

*w denotes finger width and L is the length of the fingers.

Fig 4 shows the variation of differential capacitance ($|C_{right}-C_{left}|$) as the sensor tilt from 0° to 90°.

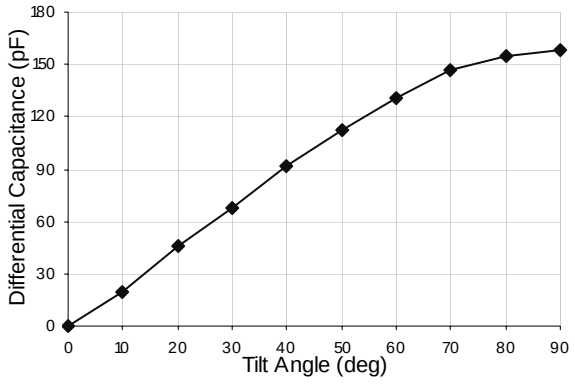


Fig 4 Relation between differential capacitance ($|C_{right}-C_{left}|$) and tilt angle.

As can be seen, the capacitance is not linearly changing with the tilt angle. This will bring complexity to the readout circuit design and cause lower resolution at large tilt angle. There are some linearization techniques which can be applied. Finger shaping is efficient to linearize the curve in Fig 4, but at the cost of lower capacitance readout. Another method is to enlarge the total sag; since the decreasing sag with tilt acts to increase capacitance as does the lateral displacement, the combined effect increases sensitivity. The summed effect gives a much more linear variation of capacitance with angle. The sag compensation has two benefits: maintenance of full resolution over the whole measurement range, and increased simplicity (and consequently reduced power consumption) of the readout circuit.

The lowest order mechanical resonant frequencies in x and z directions are 156.6 Hz and 146.9 Hz respectively. This is acceptable for body motion measurement, where a measurement bandwidth well below 100 Hz is suitable.

B. Digital Sensor

MEMS variable capacitors are used in a variety of devices, such as accelerometers, to detect displacement, and in general they achieve this using a monotonic variation of capacitance across the measured displacement range. The total capacitance variation will be limited by device size and other constraints, and so the range/resolution ratio for the detector will be determined by this total variation divided by the resolution of the capacitance measurement circuit (and so ultimately by noise). However, it is possible to alter this simple range/resolution relationship by implementing a capacitance that is periodic with displacement, a technique commonly used in optical position encoders [11]. The advantage is that the rate of change of the measured quantity with displacement can be increased; however, since the measurement repeats, the output is ambiguous. This ambiguity can be resolved by having an additional reader, for example in a "coarse" and "fine" arrangement.

Using this approach we have designed a novel comb drive structure with groups of varying width finger electrodes, each group having a different width variation pattern. One possibility this allows is to directly implement a binary reader, with different periodicity of capacitance for each bit. The value of the bits can then be obtained by a simple binary comparator, either between differential capacitors or against a threshold value.

Fig 5 illustrates a 3 bit digital comb structure of this form (with only one finger per type shown for clarity). Both the moving and fixed comb fingers have periodic protrusions, and when these come into alignment the capacitance is maximised. By setting appropriate thresholds, the three bit position can be read without additional processing.

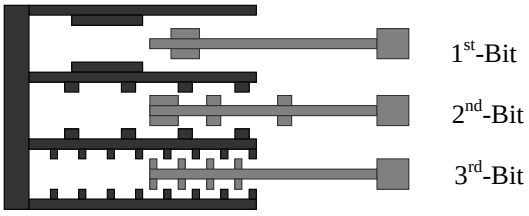


Fig 5 Schematic of a 3-bit digital comb drive.

As in the conventional design, it is important to maximise the capacitance variation. Fringing field effects will tend to "blur" the variation, and thus reduce the amplitude of the periodic capacitance change, and this effect increases as the periodicity increases (i.e. for lower order bits). However, the capacitance variation can be increased by varying the tooth-to-gap width ratio. This effect is illustrated in Fig. 6.

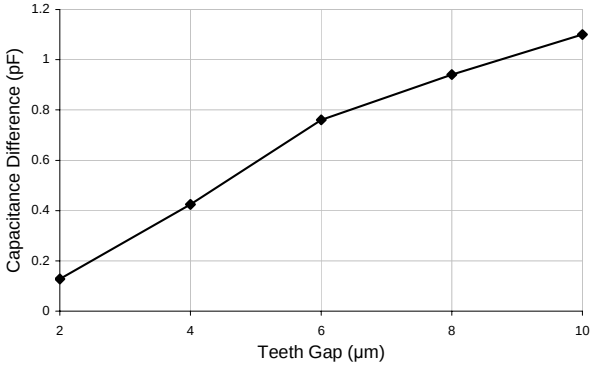


Fig 6 Capacitance variation vs. distance between teeth, for fixed tooth width of 2 µm.

The capacitance difference increases monotonically as the teeth gap width increases. However, as the gap increases, the comb fingers also have to be lengthened, and the total displacement range also increased. A large displacement range requires a large suspension, and reduces the stiffness in unwanted motion directions. Thus, a 6 μm gap width is chosen as a compromise between device size and sensitivity.

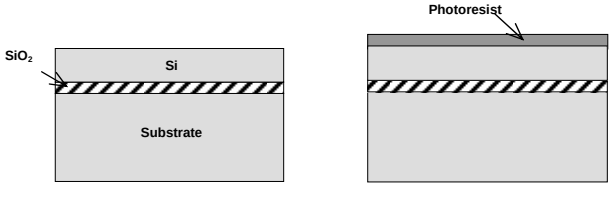
For a fully digital design, a high range/resolution ratio requires a large number of bits, and this will ultimately be limited by the minimum practical tooth width for the lowest order bit. This tooth width minimum is in turn limited by the minimum gap between fixed and moving teeth, and this depends on device stiffness as well as on fabrication precision. Meanwhile, we can still gain a resolution advantage with a combined structure as in Fig 5, by reading an analogue position from each finger type, and using the high periodicity for resolution and the low periodicity to resolve ambiguity.

IV. Fabrication

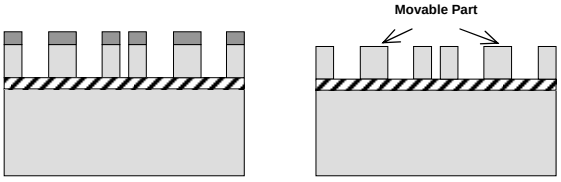
High thickness is desirable for the moving parts, both to maximize the proof mass and to increase out-of-plane stiffness. Electrical isolation between various elements is also needed, with low parasitic capacitance. SOI (silicon-on insulator) wafers provide a good platform for a device of this kind. Back etching is required in order to fully release the device. Therefore the sensor mechanical parts are fabricated using a two sided process, for the device layer etching and handle layer etching respectively.

A. Device Layer Processing

Bulk machining is applied for the device layer etching. The processing flow chart is shown in Fig 7. The top silicon layer of SOI is 50 μm thick with a 1 μm thick buried oxide. Then the photoresist S1813 is spin coated and patterned by contact photolithography, and the pattern is transferred to the wafer by high aspect ratio deep reactive ion etching (DRIE).



(a) Neat SOI wafer (b) Top photoresist coating

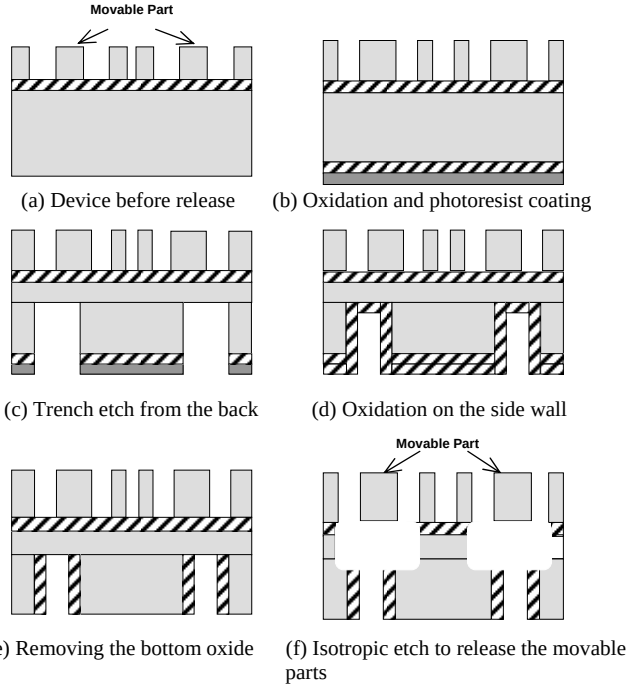


(c) High aspect ratio etch (d) Photoresist stripping

Fig 7 Device layer processing flow chart.

B. Handle Layer Processing

The processing for device release is demanding, because there are two stationary parts in the middle of the central mass, as shown in Fig 3(a). These two parts must be electrically isolated from the movable mass, while still fixed to the substrate mechanically. Therefore, normal back etching is not acceptable here. A good solution for this problem is SCREAM (Single Crystal Reactive Etch and Metallization) processing, which is efficient to produce high aspect ratio isolated island microstructures [12, 13]. The flow chart is shown in Fig 8.



(a) Device before release (b) Oxidation and photoresist coating

(c) Trench etch from the back (d) Oxidation on the side wall

(e) Removing the bottom oxide (f) Isotropic etch to release the movable parts

Fig 8 Handle layer processing flow chart.

This part of the processing starts with anisotropic trench etching. A thick masking oxide layer must be grown at the

back of the wafer to stand the long time high aspect ratio DRIE (Fig 8(b)). After the pattern transfer to the silicon (Fig 8(c)), a new layer of oxide is grown by thermal or chemical vapor deposition to protect the side wall of the trench during the isotropic release etching (Fig 8(d)). A second mask is required to remove the silicon dioxide from the bottom of the trench (Fig 8(e)). Finally, an isotropic SF_6 is applied to release the movable central mass (Fig 8(f)). The released mass is suspended by the beam, which is anchored to the silicon dioxide layer of the SOI wafer, and the isolated island in the middle is supported by the substrate.

The performance of this sensor depends highly on the flatness of the sidewall etching of the comb fingers. A dry etch process has been developed for comb drive definition, giving high verticality and smooth sidewall surfaces. A fine detail of the resulting comb structure fabricated on a test wafer is shown in Fig 9. DRIE is applied for 20 minutes, and the etching depth is around 45 μm. Better features can be achieved by applying the two-step etching method described in [8].

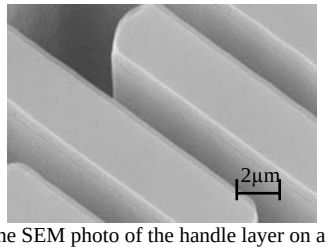


Fig 9 The SEM photo of the handle layer on a test wafer.

V. Experiments And Results

For the periodic width finger structure, the change of the capacitance is a combination of both the periodic change in tooth alignment and the total inserted length of the moving finger. In order to separate these effects, a reference comb finger without teeth is modelled. Taking the 3rd bit as an example, the capacitance variation of 400 finger pairs is shown in Fig. 11. According to Fig 6, a 6 μm tooth gap is applied. The teeth width of the 3rd bit is 2 μm, while the widths of other bits are multiples of 2 μm.

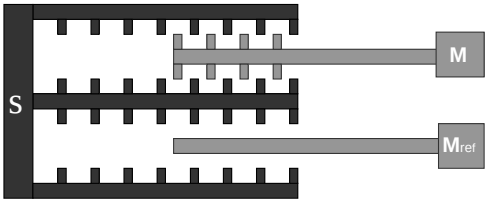


Fig 10 Digitized structure with reference finger.

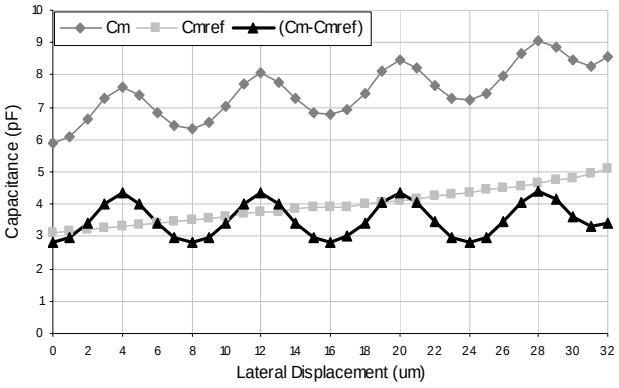


Fig 11 Capacitance variation (of 400 finger pairs) with position, for toothed and reference fingers, and the difference.

The actual capacitance read-out is (C_M-C_{Mref}). Since the lateral displacement of this sensor is large, a more flexible suspension has to be applied. For such a tilt sensor with 50 μm thickness and 400 finger pairs of each bit, the modelled relation between displacement and capacitance (simulated in CoventorWare) is shown in Fig 12, for each bit. These results show the form expected. It can be seen that for the lower order (high periodicity) bits, the capacitance contrast decreases. This is a result of the increasing effect of fringing fields

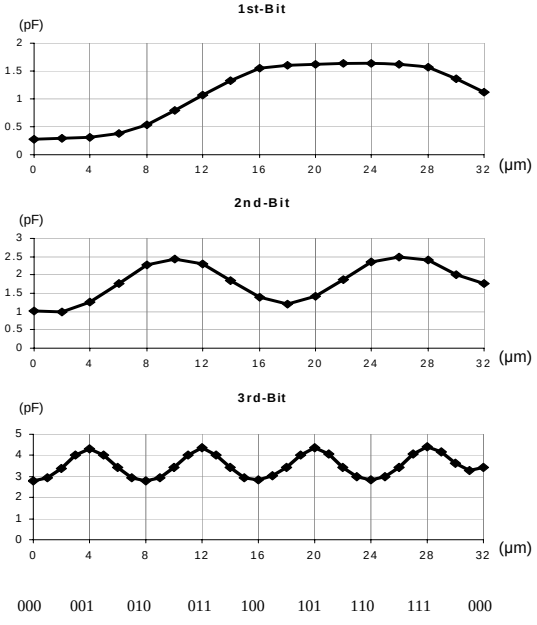


Fig 12 Output characteristic of the digital tilt sensor.

VI. Conclusions And Discussion

A tilt sensor based on inter-digitated comb capacitors has been presented. By applying proper suspension designs, the full range of tilt angle can be accurately measured. A novel design for an inherently digital sensor is presented and analysed, which is shown to be promising in tilt angle detection. A fabrication process is described for a silicon micromachined device, and high quality finger profiles, especially the sidewalls of the fingers, are demonstrated.

Although only 3 bits are implemented in this initial device, there are still some interesting applications. One example is in patient monitoring systems. When other sensors, like heart pace detectors, are working, the information about the patient's movement, such as sleeping, walking or suddenly falling over, are also important. The resolution is less importance than low power consumption, since the whole package of sensor nodes are driven by a limited power supply.

Apart from the general applications mentioned, the tilt sensor can be used to monitor motion, and thus help to correct image distortion, in PET (Positron Emission Tomography) or computed (X-ray) tomographic scanning. Besides the clinical applications, it can also be used in bionics studies, such as robot design.

References

[1] H.Kawamoto, "A Study of Capacitance Type Incline Sensor," *Institute of Electric and Electronics Engineering of Japan*, vol. R83-56, pp. 13-18, 1984.

[2] R. F. Wolffenbuttel and R. P. Van Kampen, "An Integrable Capacitive Angular Displacement Sensor with Improved Linearity," *Sensors and actuators. A, Physical*, vol. 27, pp. 835-843, 1991.

[3] F. Bantien, "Micromechanical Tilt Sensor." United States Patent, 1992.

[4] J. R. Kaienburg, M. Huonker, and R. Schellin, "Surface Micromachined Bridge Configurations for Accurate Angle Measurements," Micro electro mechanical systems, Miyazaki, Japan, 2000.

[5] H. Ueda, H. Ueno, K. Itoigawa, and T. Hattori, "Micro Capacitive Inclination Sensor Utilizing Dielectric Nano-Particles," IEEE, MEMS 2006, Istanbul, Turkey, 2006.

[6] A. Selvakumar and K. Najafi, "A high-sensitivity z-axis capacitive silicon microaccelerometer with a torsional suspension," *Journal of Microelectromechanical Systems*, vol. 7, pp. 192-200, 1998.

[7] J. W. Weigold, K. Najafi, and S. W. Pang, "Design and Fabrication of Submicrometer, Single Crystal Si Accelerometer," *Journal of Microelectromechanical Systems*, vol. 10, pp. 518-524, 2001.

[8] A. Lipson and E. M. Yeatman, "Free-Space MEMS tunable optical filter on (110) silicon," IEEE/LEOS International Conference on Optical MEMS and Their Applications Conference, 2005.

[9] V. P. Jaecklin, C. Linder, N. F. De Rooij, and J. M. Moret, "Micromechanical comb actuators with low driving voltage," *Journal of Micromechanics and Microengineering*, vol. 2, pp. 250, 1992.

[10] R. Legtenberg, A. W. Groeneveld, and M. Elwenspoek, "Comb-drive actuators for large displacements," *Journal of Micromechanics and Microengineering*, vol. 6, pp. 320-329, 1996.

[11] E. M. Yeatman, P. J. Kushner, and D. A. Roberts, "Use of scanned detection in optical position encoders," *IEEE Transactions on Instrumentation and Measurement*, vol. 53, pp. 37-44, 2004.

[12] G. K. Fedder, "MEMS fabrication," International Test Conference Proceedings, 2003.

[13] K. A. Shaw and N. C. MacDonald, "Integrating SCREAM micromachined devices with integrated circuits," IEEE, The Ninth Annual International Workshop on MicroElectroMechanicalSystems Proceedings, 1996.

5th Session
Applications

A System for Wearable Monitoring of Seated Posture in Computer Users

L.E. Dunne[1], P. Walsh[2], B. Smyth[1], and B. Caulfield[2]

[1] School of Computer Science and Informatics, University College Dublin, Dublin, Ireland
[2] School of Physiotherapy and Performance Science, University College Dublin, Dublin, Ireland

Abstract—**This work describes the evaluation of a wearable plastic optical fiber (POF) sensor for monitoring seated spinal posture, and the development of a workstation interface for the posture monitoring system. A garment-integrated POF sensor was developed and tested on nine healthy subjects. Data from the wearable sensor were compared to data taken simultaneously from a marker-based motion capture system. Peak analysis of the resulting data show a mean value error of 0.64 degrees and a mean time error of 0.53 seconds. These results show that the wearable sensor approximates the accuracy of expert visual analysis, and provides enough accuracy of measurement to reliably monitor seated spinal posture. The initial development of the system hardware and software interface are also described.**

Keywords— **Wearable technology, smart clothing, seated posture, wearable sensors.**

I. Introduction

Poor seated posture is an increasingly significant source of back problems. Although good seated posture is rare in regular computer users, no means currently exists to monitor posture and provide the user biofeedback in real time. A major reason for this is the unavailability of a simple, wearable means of long-term monitoring spinal posture in the working environment. The research presented in this paper is part of a larger project which seeks to develop a garment-integrated wearable posture monitor for use in the workplace. Specifically, this paper is focused on the evaluation of the garment-integrated sensor, and initial explorations into a minimally disruptive workstation-based user interface.

II. Background

A. Seated posture and musculoskeletal disorder

Computer work has long been associated with musculoskeletal disorders of the upper extremity [1]-[3]. Workplace design, working postures, long hours of computer work and prolonged periods of holding a static posture are some of the ergonomic factors shown to be related to an increased risk of developing work related upper limb and neck disorders [2]-[7]. The prevalence of work related upper extremity musculoskeletal disorders reported in the United States has increased dramatically during the past two decades. In 1982, they accounted for 18% of all reported occupational illnesses in the USA; in 2002, they accounted for two thirds of all reported occupational illnesses in the USA [8]. Posture-related musculoskeletal disorders are believed to cost between $15 and $20 billion each year in lost work time and medical claims.

Many ergonomic studies have focused on the postural effects of changing parts of the computer workstation such as the display height, and/or keyboard height as well as other interventions [9]. Despite widespread acceptance that work related upper limb musculoskeletal disorders among computer users can be prevented by posture modification [10] or ergonomic interventions like specific hardware (eg. adjustable chairs, forearm supports, alternative input devices), there exist very few reliable, objective and accurate methods of continually monitoring posture in the work environment to ascertain whether or not these modifications or interventions are successful.

B. Measuring seated posture

An extensive literature review by Li and Buckle [11] examines current techniques for assessing physical exposure to work-related musculoskeletal risks, with emphasis on posture-based methods. This review shows an extensive range of data collection methods, including self administered questionnaires, electromyography, inclino-meters, goniometry, electro-goniometry, professional observations, physical examinations and three-dimensional kinematics. Of the available assessment techniques, 3D kinematic systems offer the most direct and detailed motion capture data, as they alone are capable of quickly recording with a great deal of precision the simultaneous movement of a large customizable set of body landmarks, without significant discomfort for the subject.

Prior study was conducted by the authors [12] to explore the variables involved in monitoring seated posture and to inform the design of a wearable system. In this study, ten healthy subjects were monitored using an optical marker-based motion capture system, while they completed a calibration exercise and a 6-minute typing task. The data

collected were then analysed and used to develop a single-variable threshold model of seated posture. This model uses a single vector (overall spinal sagittal flexion) to evaluate seated posture. It was concluded from that study that the curvature of the spine from the C7 vertebra to the L4 vertebra (approximately base of the neck to waist) was a sufficiently accurate indicator of seated posture. That experiment informed the design of the wearable sensor evaluated here.

C. Wearable body monitoring

Three-dimensional kinematic motion capture offers a very detailed record of body movement. It is also among the least invasive and most comfortable body monitoring techniques, requiring only that markers be adhered lightly to the skin. Compared with techniques such as intra-muscular electromyography (which requires the insertion of needles directly into muscle tissue), it is a fast and easy method of detailed body motion capture. However, none of the existing clinical body monitoring techniques allows the long-term monitoring of a user in their actual work environment, nor do they provide the user with real-time biofeedback.

A long-term field analysis of seated posture thus requires that the user be fitted with a simple, easy-to-use, wearable posture monitor that requires neither alteration of the work environment (to include cameras, etc.) nor the use of a computationally complex data processor.

Bend sensing is performed in a variety of ways. The most common wearable bend sensors are usually variable resistors [13], piezo-electric materials [14], electro-active polymers [15], or optical fibers [16]. All of these can potentially be integrated into garments, although some are markedly better suited than others, and some are markedly more accurate than others.

Fiber-optic bend sensing is easily accomplished and lends itself very well to custom construction. A fiber-optic bend sensor consists of three parts: a light source, a light sensor, and a length of plastic optical fiber (POF). The POF is abraded along one side to etch away the surface of the fiber, which allows light to escape from the abraded area. The amount of light escaping depends on the bend of the fiber. The light source is placed at one end of the POF and the light sensor at the other end, and the amount of light sensed is determined by the amount of bend in the fiber. The POF can be cut to any length, and abraded only in the areas where bend is to be sensed. The sensor response is reliable, accurate, repeatable, and free of drift.

III. GARMENT-INTEGRATED WEARABLE SENSOR

In a prior study [17], the design and initial evaluation of a garment-integrated wearable posture sensor was described (see that publication for detail on the data collection methodology and analysis). In this work, we elaborate on the sensor evaluation by using more complex metrics, and describe the development of a workstation interface for computer users. The sensing garment is depicted in Figure 1, for reference.

Fig. 1 Optical sensor shirt

The initial validation of the sensor's performance used three major metrics: the correlation coefficients between the sensor data and the gold-standard (CODA) motion-capture data, the variation in CODA bend angles for a given sensor measurement, and an expert visual analysis of video stills extracted for each of the aforementioned sensor measurements. The bend angles and video stills were used to gauge the sensor's performance in the hypothetical monitoring device, which was expected to use a threshold value to identify the critical point at which a subject's posture became unhealthy. The still images and CODA bend angles were evaluated for similarity at each crossing of this threshold point. The visual analysis was conducted to lend context to the evaluation of bend angle data, which has been shown to be far more precise than expert visual analysis, which is the traditional means of evaluating posture. The precision of bend angle measured by CODA is greater than that measured by the human eye. [18]

The results of that study showed a very strong correlation between the two datasets, with a mean r^2 value of 0.913. Subjects showed an average variation in bend angle

of 6.02 degrees, or 25% of the overall range of motion. 25% is a relatively large percentage, but in visual analysis this was shown to be negligible in the evaluation of actual seated posture and was well within the level of variability in expert visual evaluation of seated posture [18].

Although correlation is important in establishing some degree of strength in the relationship between the gold-standard data and the wearable sensor data, Bland and Altman [19] have established that it is not a definitive measure when evaluating agreement between different methods of clinical measurement. Their criticisms of the use of correlation are valid, and necessitate the evaluation of sensor data by alternate means, but the method proposed by Bland and Altman is not useful for this type of data, taken in its entirety. Bland-Altman analysis is designed for measurements desired to be interchangeable, and thus does not allow for changes in scale of measurement. Additionally, it presumes some degree of regular error. As seen in Figure 2, the sensor response data is visually a reasonable approximation of the motion-capture standard. However, during the course of the test period, the sensor response displays both value (amplitude) error and time (phase-shift) error, and neither in a consistent manner. Thus, the Bland-Altman evaluation of measurement difference will be significantly affected by this irregular response, where in practice that type of error may not in fact be significant.

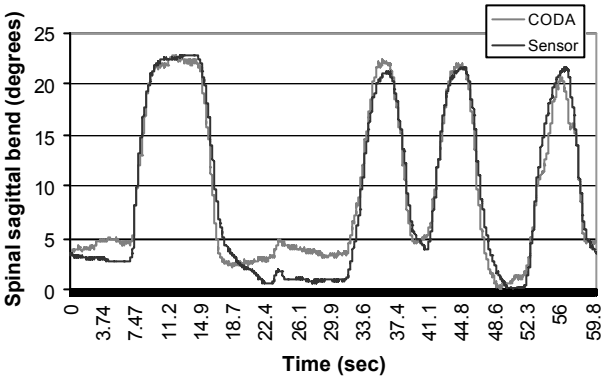


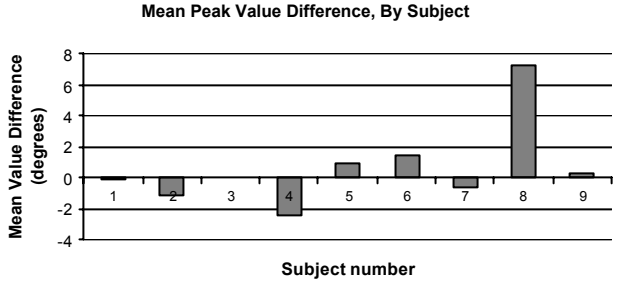
Fig. 2 Sample CODA and sensor response

In place of Bland-Altman analysis, we performed a peak analysis which allowed the sensor response to be evaluated on both value and timing (amplitude and phase) independently of one another.

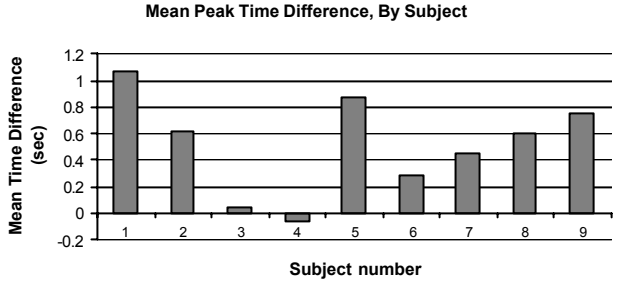
IV. Peak analysis

Three peaks and two troughs from each of the nine datasets were evaluated. For each peak/trough, the minimum or maximum value was extracted for both the sensor and the CODA measurement. The timestamp for this value was also extracted. The difference in magnitude between the peak/trough values and in time between the peak/trough times were recorded for each, and the resulting datasets analyzed for mean and standard deviation.

This analysis resulted in quantified error for scale and timing. The mean value difference for the entire set of peaks/troughs was 0.64±3.1degrees (mean±sd), and the mean time difference was 0.53±0.8 seconds (mean±sd). These data are depicted by subject in Figure 3(a) and (b). As seen, the high standard deviation in value difference was strongly influenced by one significant outlier. Without the outlier, the mean value difference is 0.19, std. dev. 1.74. Reasons for the extremity of this outlier are unclear: one possibility is the subject's anatomical irregularity of a fairly recessed spine due to prominent development of the adjacent latissimus dorsi muscles.



(a)



(b)

Fig. 3 Mean value (a) and time (b) differences, by subject

From the peak analysis, we can see that the two signals are well matched for the purposes of posture measurement. Posture is a very low frequency signal, and thus time differ-

ences of a even a few seconds are tolerable. In our examination of current methods of posture analysis [12,18], we have discovered that experts display coefficients of variation of 30 to 60 percent, indicating a large amount of intra-expert variation. When asked to repeatedly identify the transition point between good and bad posture, experts gave responses that covered an average of 20.29% of the range of motion. While the comparisons are not identical (peak analysis vs. identification of a more nebulous threshold), the sensor's accuracy with comparison to CODA is within 3.67 degrees (without the outlier) and 2.13 seconds of the actual value. Our subjects had an average range of motion of 25.32 degrees, thus 3.67 degrees represents only 14.5% of the range of motion.

The error in the wearable sensor is most likely due to the slippage of the sensor and textile over the skin. This is minimized by close fit and extensibility of the garment itself. While other textile-based means of body motion capture [20] have utilized redundancy as a means of minimizing sensor error, we believe redundant sensors would not improve our data: redundancy minimizes the error inherent in a sensor, while our error arises from movement of the garment housing, which would be experienced by all sensors.

Based on this additional peak analysis, we have reconfirmed that the wearable sensor, while slightly less accurate than the gold-standard motion capture system, is more accurate than traditional visual analysis and thus easily accurate enough for our purposes.

V. Desktop Interface

The next step in the development of a field-deployable wearable posture monitor is the design and implementation of a user interface. This initial development has been undertaken in two phases: hardware development and software development.

A. Sensor Hardware

Our sensor hardware is based around a PIC 16F76 microcontroller. The voltage fluctuations over the light-detector part of the bend sensor (a light-sensitive resistor) are first amplified and then passed through the PIC's ADC converter. The digital values are then sent over a serial Bluetooth connection to the host computer.

B. Sensor Software

The software interface consists of a calibration mode and an operation mode. The calibration mode asks the user to record their best posture and their threshold posture. In a clinical setting, this could be a trained posture or could be set by the clinician. In a more casual setting, this could be entirely user-determined. The software then uses the range to calculate a degree of "padding" for the threshold value: the top 2/3 of the range is the "safe" zone. The last third is the "warning" zone, and below the threshold is the "unhealthy" zone.

Following calibration, the software enters operation mode. The calibration window is replaced by a system tray icon. The icon is a colored circle, with a white outline of a seated stick figure. As the user passes from safe to warning to unhealthy posture, the icon changes from green to yellow to red, and the posture of the stick figure changes (Figure 4). If an unhealthy posture is maintained for too long, a popup warning message (Figure 4(c)) is displayed to attract the user's attention.

The system tray icon interface allows the user to monitor his/her own progress at will, without active interruption, until such point as an unhealthy position is maintained for too long. Additionally, while the system is in operation mode, the data are logged to a log file for future analysis.

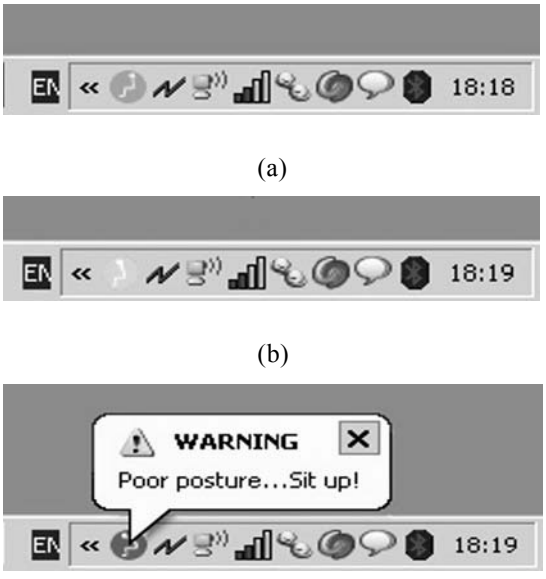


Fig. 4 Interface system tray icon: safe (a), warning (b), and unhealthy (c) modes.

VI. Conclusion

The sensor garment described presents a truly wearable solution for the problem of field monitoring of seated pos-

ture. With very minimal compromise of the user's comfort and status quo, it is possible to monitor his/her seated posture with greater accuracy than expert visual analysis.

Although the system is highly functional in its current form, it is not yet finalized. Current work is underway to optimize the physical size and housing for the wearable hardware component. This is necessary both to optimize the sensor's performance, which is affected to some degree by movement of the light-sensing component, and to maximize the wearability and comfort of the garment-integrated device. Additional studies are also being conducted to evaluate the relationship between garment ease (tightness of the garment itself) and sensor performance, and to evaluate the impact of donning and doffing the garment. Once the system is finalized, it will be deployed in long-term field evaluations.

Acknowledgment

This work is supported in part by University College Dublin under grant number SF-020, and in part by Science Foundation Ireland under grant number 03/IN.3/I361.

References

1. J. Dennerlein and P. Johnson, "Different computer tasks affect the exposure of the upper extremity to biomechanical risk factors". *Ergonomics* 2006; 49(1): 45-61
2. U. Bergqvist, E. Wolgast, B. Nilsson and M. Voss, "Musculoskeletal disorders among visual display terminal workers: individual, ergonomic, and work organizational factors". *Ergonomics* 1995; 38: 763-776
3. J. Faucett and D. Rempel, "VDT-related musculoskeletal symptoms: interactions between work posture and psychosocial work factors". *American Journal of Industrial Medicine* 1994; 26: 597-612
4. A. Aaras, G. Horgen, H. Bjorset, O. Ro and M. Thoresen, "Musculoskeletal, visual and psychosocial stress in VDU operators before and after multidisciplinary ergonomic interventions". *Applied Ergonomics* 1998; 29(5): 335-354
5. M. Fogleman and G. Brogmus, "Computer mouse use and cumulative trauma disorders of the upper extremities". *Ergonomics* 1995; 38: 2465-2475
6. F. Gerr, M. Marcus, C. Ensor, et al. "A prospective study of computer users: 1. Study design and incidence of musculoskeletal symptoms and disorders". *American Journal of Industrial Medicine* 2002; 41: 221-235
7. G. Szeto, L. Straker, and P. O'Sullivan, "A comparison of symptomatic and asymptomatic office workers performing monotonous keyboard work: 2. Neck and shoulder kinematics". *Manual Therapy* 2005; 10: 281-291
8. US Bureau of Labour Statistics 2003 http://www.bls.gov/iif/oswhc/osh/os/osch0024.pdf
9. A. Lindegard, C. Karlberg, E. Wigaeus Tornqvist, A. Toomingas and M. Hagberg, "Concordance between VDU-users' ratings of comfort and perceived exertion with experts' observations of workplace layout and working postures". *Applied Ergonomics* 2005; 36: 319-325
10. F. Gerr, M. Marcus, C. Monteilh, L. Hannan, D. Oritz and D. Kleinbaum, "A randomised controlled trial of postural interventions for prevention of musculoskeletal symptoms among computer users". *Occupational and Environmental Medicine* 2005; 62: 478-487
11. G. Li and P. Buckle, "Current techniques for assessing physical exposure to work-related musculoskeletal risks, with emphasis on posture-based methods". *Ergonomics* 1999; 42(5): 674-695
12. P. Walsh, L.E. Dunne, B. Caulfield, and B. Smyth, "Marker-Based Monitoring of Seated Spinal Posture Using a Calibrated Single-Variable Threshold Model", *Proc of EMBC2006*.
13. L.K. Simone and D.G. Kamper, "Design considerations for a wearable monitor to measure finger posture". *Journal of NeuroEngineering and Rehabilitation* 2005; 2(5)
14. T. Martin, M. Jones, J. Edmison, T. Sheikh, Z. Nakad, "Modeling and simulating electronic textile applications". *Proc of **LCTES 2004*** (2004)
15. L.E. Dunne, S. Brady, R. Tynan, et. al, "Garment-based body sensing using foam sensors". *Proc of AUIC 2005*
16. K.S.C. Kuang, W.J. Cantwell, and P.J. Scully, "An evaluation of a novel plastic optical fibre sensor for axial strain and bend measurements". *Measurement Science and Technolog* 2002; 13: 1523-1534.
17. L.E. Dunne, P. Walsh, B. Smyth, and B. Caulfield. "Design and evaluation of a wearable optical sensor for monitoring seated spinal posture". *Proc of ISWC 2006.*
18. P. Walsh. "Measuring Seated Spinal Posture". Master's Thesis, University College Dublin. 2006.
19. J. M. Bland and D. Altman, "Statistical methods for assessing agreement between two methods of clinical measurement," Lancet, vol. 1, pp. 307-310, 1986.
20. F. Lorussi, W. Rocchia, E. Scilingo, A. Tognetti, and D. De Rossi. "Wearable, redundant fabric-based sensor arrays for reconstruction of body segment posture", *IEEE Sensors Journal,* 2004; 4(6):807-818.

Address of the corresponding author:

Author: Lucy Dunne
Institute: University College Dublin
Street: CSI, Block 8 Belfield Office Park
City: Clonskeagh, Dublin 4
Country: Ireland
Email: lucy.dunne@gmail.com

Ambient and Wearable Sensor Fusion for Activity Recognition in Healthcare Monitoring Systems

Julien Pansiot[1,2], Danail Stoyanov[1,2], Douglas McIlwraith[1], Benny P.L. Lo[1,2] and G. Z. Yang[1,2]

[1] Royal Society/Wolfson MIC Lab, Department of Computing, Imperial College London, United Kingdom
[2] Institute of Biomedical Engineering, Imperial College London, United Kingdom

Abstract— **The use of wearable sensors for home monitoring provides an effective means of inferring a patient's level of activity. However, wearable sensors have intrinsic ambiguities that prevent certain activities to be recognized accurately. The purpose of this paper is to introduce a robust framework for enhanced activity recognition by integrating an ear-worn activity recognition (e-AR) sensor with ambient blob-based vision sensors. Accelerometer information from the e-AR is fused with features extracted from the vision sensor by using a Gaussian Mixture Model Bayes classifier. The experimental results showed a significant improvement of the classification accuracy compared to the use of the e-AR sensor alone.**

Keywords— **blob sensor, wearable sensor, sensor fusion, activity recognition**

I. Introduction

Monitoring the status of the elderly or chronically ill patients in their own homes is an essential requirement for delivering more effective pervasive healthcare. By continuous monitoring of key physiological parameters of the patients, wearable sensors can provide a rich source of information about their current status of health [1]. There is also increasing evidence to suggest that major episodes are often preceded by changes in behaviour and domicile activity, which may be detected from detailed information about the posture, gait and general activity of the patient based on ambient sensing [2]. To achieve truly pervasive health monitoring, it is necessary to integrate wearable/ implantable sensors provided by a body sensor network (BSN) with data acquired from ambient environment sensors.

Recently, various systems for home monitoring have been developed based on either wearable or ambient sensors. Frameworks using wearable sensors are typically based on accelerometers [15,16], ECG sensors [18], pulse oximeters ($Sp0_2$) [16], temperature [15], and bend sensors [12,13,18]. Other wearable sensors include humidity, acoustic and light sensors [14]. By the use of wearable sensors, it provides an effective means of monitoring the biophysical status of the patient. Due to the lack of a global reference, however, it can be difficult to use this data to infer certain physical activities. For example, a wearable accelerometer positioned on the head can detect local motion but not whether the subject is standing or sitting. Either ambient or additional wearable sensors (typically positioned on the joints) need to be used for achieving the required body posture differentiation.

For monitoring environments based on ambient sensors, current systems include the use of cameras [4,11,21,24], IR sensors, ambient sound [21], heat, as well as contact sensors mounted on furniture [19]. These systems can provide information about the spatial location and general activity of the subject within the environment. The weakness of ambient sensing is that it is often too ambiguous to differentiate detailed information about the subject, which in many cases can only be derived from a wearable system.

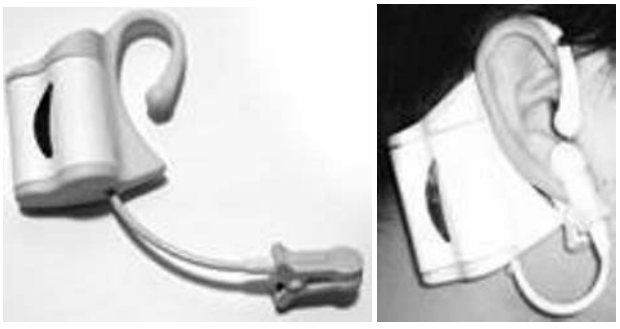
Existing research has shown that there is a complementary relationship between the two sensing paradigms. Effective sensor fusion can be used to combine the strengths of ambient and wearable sensors by fusing sensory data at hardware, raw data, feature, or decision levels [22]. At the hardware level, it can be achieved by using simple thresholds [15]. At the data level, dimensionality reduction such as Principal Component Analysis (PCA) is often deployed before further pattern classification techniques are applied. At this level of sensor fusion, modelling methods such as Gaussian mixtures [14], Bayes networks [21] or Hidden Markov Models (HMM) methods are common.

The purpose of this paper is to develop a framework for improved activity recognition by integrating an ear-worn activity recognition (e-AR) sensor with ambient blob sensors. Data independently obtained by each sensor is preprocessed for dimensionality reduction before the application of a Bayesian classifier. To assess the improved accuracy of the proposed method, we evaluated the classifier against a lab-based home monitoring scenario. Significant improvements in the recognition rates of all activities have been achieved when compared to using wearable or ambient sensors alone.

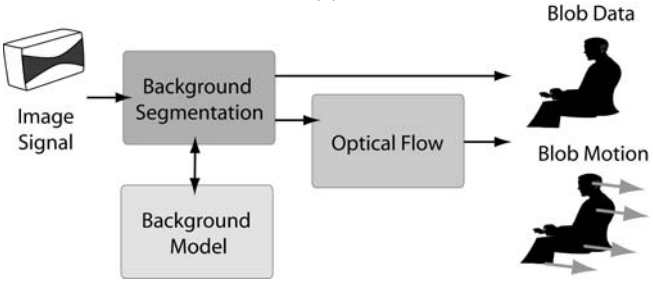
II. System Design

A. Wearable e-AR Sensor

The e-AR sensor [16] is based on the BSN platform [20] that consists a Texas Instruments MSP430 processor, Chipcon CC2420 radio transceiver, Atmel 512KB EEPROM, MCC ChipOX SpO_2 module and a 3-axis accelerometer. The integrated e-AR sensor used for this study is shown in Fig. 1(a). For activity recognition in this study, the main information used was derived from the 3-axis accelerometer whereas SpO_2 signals of the e-AR sensor were not used.



(a)



(b)

Fig. 1 (a) The e-AR sensor used in this study [16] and (b) the data processing pipeline for the ambient blob sensor.

B. Ambient Blob Sensors

The ambient sensor proposed in this study is a self-contained module consisting of a video sensor, on-board processor, wireless communication and battery [4]. It has a wall mount design and can be integrated into the home environment similar to a PIR security device. Video data observed by the device is processed on board in real-time and the sensor communicates only derived signal metrics such as the silhouette of a moving object and its local motion in the form of optical flows. Communication between ambient nodes is used to provide large scale tracking and improves the overall system robustness. The ambient sensor being under development, CCTV cameras were used in this experiment. Under this sensing paradigm, blobs representing the monitored subject are first extracted from the video signal using a background statistical model, where every pixel is represented as a Gaussian mixture distribution maintained over time as proposed by Lee [10]. Incoming signals are compared with the existing background model and segmented into a binary map of foreground and background. The use of normalized RGB colour space reduces the sensitivity of the algorithm to shadows. Post-processing of the foreground object based on mathematical morphology is used for noise removal.

In addition to the extraction of blob profiles, the optical flow within the blob is also extracted, which is based on the classical technique proposed by Horn and Schunck [8]. Optical flow can be considered as a natural extension of the blob silhouette as it also captures the motion of the limbs. This information has been used previously for gesture recognition [25] and activity recognition in a multi-resolution framework [9]. The complete data processing work flow is summarized in Fig 1(b).

It is important to note that the silhouette and optical flow extracted by each ambient sensor do not carry any appearance information and no image data is transmitted to other devices. This is important for home care environments where privacy is of high priority.

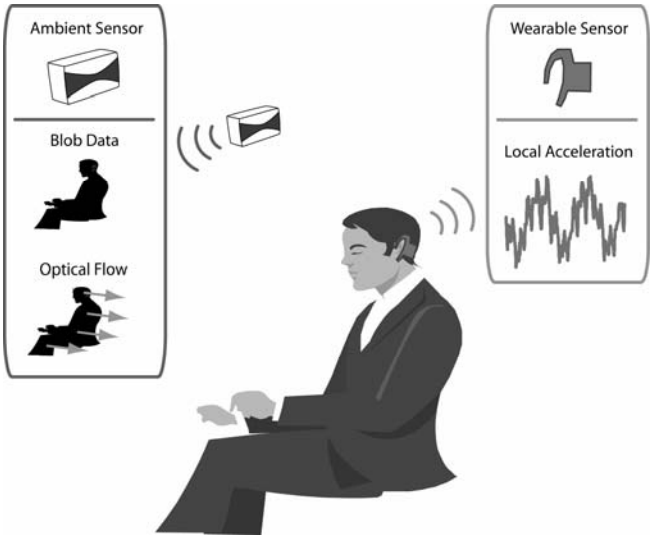


Fig. 2 A schematic diagram of the proposed ambient and wearable monitoring system.

C. Overall System Integration

For the motion data provided by the e-AR sensor, the accelerometer signal is intrinsically linked to the patient's movement. On its own, it is capable of differentiating activities such as walking, standing, and sleeping. For certain activities, however, the readings from the e-AR sensor are ambiguous, *e.g,* standing and sitting still. In these cases, the e-AR sensor cannot correctly classify the patient's activity as it cannot obtain a global perspective of the body's position just from the head motion alone. It is expected that by fusing the e-AR data with the ambient sensors at the data level, it is a possible to obtain a much more reliable activity classification result for a wider range of activities. Fig. 2 illustrates a schematic diagram outlining the proposed system.

III. Ambient and Wearable Sensor Fusion

For effective sensor fusion, two types of features are extracted from the e-AR accelerometers: tilt and movement frequency spectrum. To derive the tilt information, the accelerometers are pre-calibrated such that the acceleration due to gravity can be evaluated. A record of the total acceleration in the three axes allows the calculation of the gravity constant component on each of the accelerometers. It is therefore possible to separate the relative head acceleration and gravity acceleration to the tilt with respect to the vertical axis. Moving window Fast Fourier Transform (FFT) was also computed on the acceleration data to the intrinsic frequency of the movement.

Table 1 Features used in classification

Sensor	Feature	Size
e-AR	Acceleration X axis	1
e-AR	Acceleration Y axis	1
e-AR	Acceleration Z axis	1
e-AR	FFT acceleration X axis	11
e-AR	FFT acceleration Y axis	11
e-AR	FFT acceleration Z axis	11
e-AR	Head tilt X	1
e-AR	Head tilt Z	1
Blob	Blob speed estimation	1
Blob	Blob aspect ratio	1
Blob	Subject height estimation	1
Blob	Subject optical flow intensity	1
Blob	Subject optical flow correlation	1
Blob	Subject optical flow aspect ratio	1

From the ambient sensor, the derived features used for sensor fusion include the aspect ratio and mean velocity of the blob. The calculation of the optical flow is based on the iterative application of the following equation:

$$V^{k+1} = \overline{V^k} - I\frac{I_x\overline{V_x^k} + I_y\overline{V_y^k} + I_t}{\alpha^2 + I_x^2 + I_y^2}$$

for iteration k for a given pixel. In this equation V^k is the optical flow vector, I , I_x , I_y and I_t the image intensity and its partial derivatives and $\overline{V^k}$ the average of V^k in its neighbourhood. In the above equation, α is a regularization constant to ensure the smoothness of the result derived. After noise filtering, the main moving elements of the field-of-view are extracted and the bounding box is calculated from the the eigenvectors of the covariance matrix of the blobs as proposed by Lahanas *et al.* [23]. A complete list of the features used for the classifier is summarised in Table 1.

Sensor fusion is performed based on a Gaussian Bayes EM classifier based on the e-AR and the blob sensor data. A Gaussian Mixture Model (GMM) is used to model each activity class. For the implementation of the classifier, we used the Bayes Net Toolkit (BNT) [5] and a total of nine classes were modelled to describe different activities. The activities used for classification in this study include walking, standing still, standing with head tilted on the side, sitting at the dining table, reading at the table, eating, sitting on the sofa, lounging on the sofa and eventually lying down.

For each of the activities considered, three quarters of the data were used for training of the inference system and the rest for validation. For training, an Expectation-Maximisation (EM) iterative method was used to compute the maximum likelihood fit [6]. Given a dataset $\{x_{1..R}\}$ to be classified in c classes, and assuming that the conditional probability density function (PDF) *P(X=x)* for each of these classes is Gaussian, we try to find the best fit of:

$$\lambda_t = \{\mu_{1..c}(t), \sum_{1..c}(t), p_{1..c}(t)\}$$

where $\mu_i(t)$, $\sum_i(t)$ and $p_i(t) = P(w_i)(t)$ are the mean, the covariance and the estimates of the weights of the mixtures at the iteration *t*, respectively. The expectation and maximization steps are performed iteratively until convergence. The expectation step for each class *i*, based on Bayes' law can be represented as the following:

$$P(w_i|x_k,\lambda_t) = \frac{p(x_k|w_i,\lambda_t)P(w_i|\lambda_t)}{p(x_k|\lambda_t)} = \frac{p(x_k|w_i,\mu_i(t),\sum_i(t))p_i(t)}{\sum_{j=1}^{c} p(x_k|w_j,\mu_j(t),\sum_j(t))p_j(t)}$$

Because the Gaussian PDF is differentiable, the maximization of the likelihood step for each class *i* can be expressed as:

$$\mu_i(t+1) = \frac{\sum_{k=1}^{R} P(w_i \mid x_k, \lambda_t) x_k}{\sum_{k=1}^{R} P(w_i \mid x_k, \lambda_t)}$$

$$\sum_i (t+1) = \frac{\sum_{k=1}^{R} P(w_i \mid x_k, \lambda_t)(x_k - \mu_i(t+1))(x_k - \mu_i(t+1))^T}{\sum_{k=1}^{R} P(w_i \mid x_k, \lambda_t)}$$

$$p_i(t+1) = \frac{1}{R} \sum_{k=1}^{R} P(w_i \mid x_k, \lambda_t) x_k$$

Once the model is computed through EM, the remaining data is used to evaluate the accuracy of the classifier based on the marginal probability of every activity. The highest probability was chosen for the final classification.

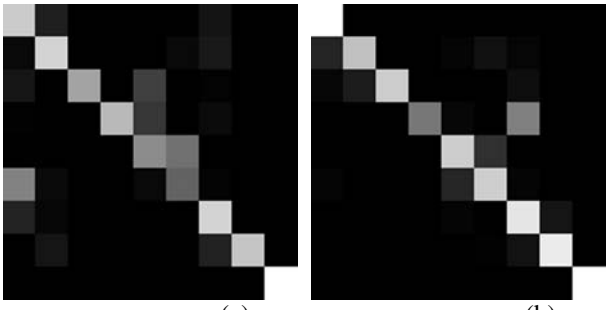
IV. Experiments and Results

To evaluate the proposed technique, a purpose built simulated home environment was used. Data was recorded for two actors wearing an e-AR sensor. A total of 9 activities were performed by the actors with each activity lasted for approximately one minute. The classification results by using the proposed method as compared those with the e-AR sensor alone are presented in Table 2. A detailed analysis of inter-class misclassification as illustrated as confusion matrices with and without sensor fusion is provided in Figure 3.

From the results shown in Table 2, it is evident that by incorporating ambient sensing with the e-AR sensor, activity classification is improved significantly. This is particularly obvious for classes where the e-AR sensor was ambiguous due to a lack of global information. For example, reading is not easily classified with the e-AR sensor only, as very little temporal and global orientation information is available. With the use of the blob sensor, this improves significantly because of the strong difference of the appearance cue. The same effects are visible in differentiating classes such as standing with the head tilted and sitting activities such as eating and reading, which are not well classified by the e-AR sensor alone. In these cases, the optical flow features provide a good clue about the type of activity, which significantly improves the sensitivity and specificity of the system. In the current implementation, however, the differentiation between sitting on the chair and on the sofa is relatively low, as evident from the confusion matrix shown in Figure 3.

Table 2 Comparison of activity classification rates between using a wearable sensor alone and with the proposed combined system

Class	Activity	e-AR sensor alone	e-AR + ambient sensing
1	Walking	79%	100%
2	Standing	83%	75%
3	Standing (head tilted)	65%	80%
4	Sitting	73%	47%
5	Reading	55%	80%
6	Eating	39%	81%
7	Sitting (sofa)	84%	90%
8	Lounging	77%	92%
9	Lying down	100%	100%



(a) (b)

Fig. 3 The confusion matrices showing how the algorithm differentiates eating and reading activities with (a) the e-AR sensor only and (b) after sensor fusion. The main non-diagonal element in (b) is the confusion between sitting on a chair and on the sofa.

V. Conclusions and Future Work

In this paper, we have proposed a sensor fusion framework for integrating ambient and wearable e-AR sensors. Our experiments illustrate the practical value of the method by improving classification rates for most activities investigated in this study. This clearly demonstrates the fact that ambient environment sensors can be used to overcome some of the ambiguities in activity recognition by using wearable sensing alone. This is a desirable feature for the effective deployment of future pervasive patient monitoring systems.

In the current system, we did not explicitly handle the spatial dependency between the ambient sensor and the patient. This projective relationship can influence the ambi-

ent sensor readings. We are currently looking into resolving this issue by using multi-view geometry to derive pose invariant 3D representations. Other areas for further improvement include the development of robust learning capabilities of the ambient sensors and real-time implementation of the proposed sensor fusion paradigm directly on the sensor nodes.

References

1. Yang G-Z (2006) Body Sensor Networks
2. Verghese J, Lipton R B, Hall C B, Kuslansky G, Katz M J, Buschke H (2002) Abnormality of gait as a predictor for non-Alzheimer's dementia. N Engl J Med, 347:1761-1768
3. Stauffer C, Grimson W E L (2000) Learning patterns of activity using real-time tracking. IEEE Transactions on Pattern Analysis and Machine Intelligence, 22:747-757
4. Lo B P L, Wang J L, Yang G Z (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. Murphy K. Bayes Net Toolkit. http://bnt.sourceforge.net
6. Dempster A P, Laird N M, Rubin D B (1977) Maximum likelihood from incomplete data via the EM algorithm, by Journal of the Royal Statistical Society. Series B (Methodological) Royal Statistical Society
7. McKenna S J, Nait-Charif H (2004) Summarising contextual activity and detecting unusual inactivity in a supportive home environment. Pattern Analysis and Applications 7(4), 386-401, Springer-Verlag, December 2004
8. Horn B K P, Schunck B G (1981) Determining optical flow. Artificial Intelligence, vol 17, pp 185-203.
9. Wactlar H, Yang, Jie, Chen, D, Hauptmann A, Christel M (2004) Infrastructure for Machine Understanding of Video Observations in Skilled Care Facilities: Implications of Early Results from CareMedia Case Studies. UbiComp 2004: The 2nd International Workshop on Ubiquitous Computing for Pervasive Healthcare, Applications, Nottingham, U.K., September 6-7, 2004
10. Lee D-S (2005) Effective Gaussian Mixture Learning for Video Background Subtraction. IEEE Transactions on Pattern Analysis and Machine Intelligence ,vol. 27, no. 5, pp. 827-832, May, 2005.
11. Pansiot J, Stoyanov D, Lo B P L, Yang G-Z (2006) Towards image-based modeling for ambient sensing. . International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2006), 3-5 April 2006
12. Amft O, Junker H, Lukowicz P, Troster G, Schuster C (2006) Sensing muscle activities with body-worn sensors. International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2006), 3-5 April 2006
13. Krebs D E, Huddleston J I, Goldvasser D, Scarborough D M, Harris W H, Malchau H (2006) Biomotion community-wearable human activity monitor: total knee replacement and healthy control subjects. International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2006), 3-5 April 2006
14. Grajales L, Nicolaescu I V (2006) Wearable multisensor heart rate monitor. International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2006), 3-5 April 2006
15. Laerhoven K V, Gellersen H W, Malliaris Y G (2006) Long term activity monitoring with a wearable sensor node. International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2006), 3-5 April 2006
16. Aziz O, Lo B, King R, Darzi A, Yang G-Z (2006) Pervasive body sensor network: an approach to monitoring the post-operative surgical patient. International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2006), 3-5 April 2006
17. Giorgino T, Quaglini S, Lorassi F, De Rossi D (2006) Experiments in the detection of upper limb posture through kinestetic strain sensors. International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2006), 3-5 April 2006
18. Park C, Pai H, Chou, Bai Y, Matthews R, Hibbs A (2006) An Ultra-Wearable, Wireless, Low Power ECG Monitoring System, to appear in Proc. IEEE BioCAS, Nov 29 - Dec 1, 2006. The British Library, London.
19. Reeves A A, Ng J W P, Brown S J, Barnes N M (2006) Remotely supporting care provision for older adults. International Workshop on Wearable and Implantable Body Sensor Networks (BSN 2006), 3-5 April 2006
20. Lo B, Yang G-Z (2005) Key Technical Challenges and Current Implementations of Body Sensor Networks. IEE Proceedings of the 2nd International Workshop on Body Sensor Networks (BSN 2005), 1-5, April 2005
21. Chen D, Malkin R, Yang J (2004) Multimodal detection of human interaction events in a nursing home environment. Proceedings of the 6th international Conference on Multimodal interfaces (State College, PA, USA, October 13 - 15, 2004). ICMI '04. ACM Press, New York, NY, 82-89.
22. McCullough C L, Dasarathy B V, Lindberg P C (1996) Multi-level sensor fusion for improved target discrimination, Decision and Control. Proceedings of the 35th IEEE , vol.4, no.pp.3674-3675 vol.4, 11-13 Dec 1996
23. Lahanas M, Kemmerer T, Milickovic N, Karouzakis K, Baltas D, Zamboglou N (2000) Optimized bounding boxes for three-dimensional treatment planning in brachytherapy. Medical Physics 27 (10): 2333-2342 October 2000
24. Wang J, Lo B P L, Yang G-Z (2005) Ubiquitous Sensing For Posture/Behavior Analysis. IEE Proceedings of the 2nd International Workshop on Body Sensor Networks (BSN 2005), pp.112-115, April 2005.
25. Cutler R, Davis L (1998) View-based detection and analysis of periodic motion. International Conference on Pattern Recognition, Brisbane, Australia, August 1998

Address of the corresponding author:

Author:	Julien Pansiot
Institute:	Department of Computing, Imperial College London
Street:	180 Queen's Gate, South Kensington
City:	London
Country:	United Kingdom
Email:	julien.pansiot@imperial.ac.uk

Physical Activity Monitoring for Assisted Living at Home

Roozbeh Jafari[1], Wenchao Li[1], Ruzena Bajcsy[1], Steven Glaser[2], Shankar Sastry[1]

[1] Department of Electrical Engineering and Computer Science
[2] Center for Information Technology Research in the Interest of Society
University of California, Berkeley, California, USA

Abstract— **We propose a methodology to determine the occurrence of falls from among other common human movements. The source data is collected by wearable and mobile platforms based on three-axis accelerometers to measure subject kinematics. Our signal processing consists of preprocessing, pattern recognition and classification. One problem with data acquisition is the extensive variation in the morphology of acceleration signals of different patients and under various conditions. We explore several effective key features that can be used for classification of physical movements. Our objective is to enhance the accuracy of movement recognition. We employ classifiers based on neural networks and k-nearest neighbors. Our experimental results exhibit an average of 84% accuracy in movement tracking for four distinct activities over several test subjects.**

Keywords— **Fall detection, movement monitoring, wearable and ubiquitous computing, signal processing.**

I. Introduction

Health care costs in developed countries are rapidly increasing due to a substantial increase the elderly population. Monitoring of daily physical activities can be a key to evaluating the actual quality of life among the elderly. We believe that the overall health and wellness of elderly sectors of the population can greatly benefit from the use of information and communication technology (ICT), especially for the homebound [1, 2]. New technology allows the creation of small sensor "Motes" which combine a variety of micro-machined transducers, a micro-controller to reduce data into information, and a wireless link to the outside world. Privacy is greatly increased by a decentralized system, where distributed data cannot be easily corrupted. Sensor platforms integrated into clothing provide the possibility of enhanced reliability of accident reporting and health monitoring. Such devices improve the independence of people needing living assistance. In this paper we present data and analyses that show differences in movement parameters between young and aged control groups. A new classification scheme is introduced that allows learning the idiosyncrasies of the individual subject (hence group).

In order for ICT-based systems to gain widespread acceptance and use, important social concerns must be addressed and many technical challenges overcome. Some of these concerns involve wearability and ease of use, cost, maintenance and the effectiveness of privacy. One of the technical challenges posed by such systems is the fusion and analysis of the many streams of data provided by numerous sensing elements. In this paper we present a prototype of a Mote-based system that is capable of predicting the need for medical attention, and notifying emergency services of an acute illness or accident. In particular, we want to accurately announce falling of the elderly. This work is a part of the information technology for assisted living at home (ITALH) project at Berkeley.

II. Related work

Various motion sensors can be adapted to monitor daily physical activity, ranging from mechanical pedometers [3], actometers [4] to accelerometers [5].

Accelerometers are the most commonly used motion sensors for physical activity assessments. These sensors respond to both the frequency and intensity of a movement, and are superior to pedometers and actometers, which are attenuated by impact or tilt and can only measure body movement over a certain threshold. Current MEMS technology enables us to build very small and lightweight accelerometer-based Motes that can be worn for days or even weeks.

Anliker et al. proposed a portable telemedical monitor (AMON) [6] for high-risk cardiac/respiratory patients. This system includes continuous collection and assessment of multiple vital signs, intelligent multi-parameter medical emergency detection, and a cellular link to a medical center. By integrating the whole system in a low profile, wrist-worn enclosure, continuous long-term monitoring can be performed without interfering with the patients' everyday activities and restricting their mobility. Specific movement patterns, however, may not be recognized using AMON.

Najafi et al. suggested a method of physical activity monitoring which is able to detect body postures (sitting, standing, and lying) and periods of walking in elderly persons using only one kinematic sensor attached to the chest [7]. The wavelet transform, in conjunction with a simple kinematics model, was used to detect different postural transitions (PTs) and walking periods during daily physical activity. This approach may not be applicable to light-weight and wearable processing units due to its computational complexity.

The most recent work presents the implementation of a real-time classification system for the types of human movement associated with the data acquired from a single, waist-mounted triaxial accelerometer unit [5]. The decision making algorithm is based on the detection of the angular orientation of the accelerometer. The decision algorithm does not account for individual differences, and signal processing is always performed on a fixed interval of data (one second).

Aminian outlined the advantage of new technologies based on body-fixed sensors and particularly the possibility to perform field measurement [8]. The system integrates outputs from accelerometers and gyroscopes for human movement tracking.

III. System architecture

We have constructed a prototype of a decentralized sensor Mote that is designed to eventually fit into an integrated communication scheme such as illustrated in Figure 1. The current device uses Bluetooth communication to facilitate connection to mobile phones and laptop computers.

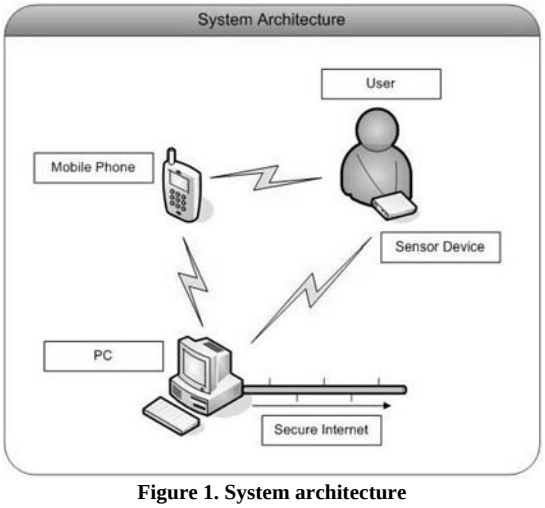


Figure 1. System architecture

The prototype device utilizes three-axis accelerometers for fall detection, and GPS to measure out-of-building mobility, as shown in Figure 2.

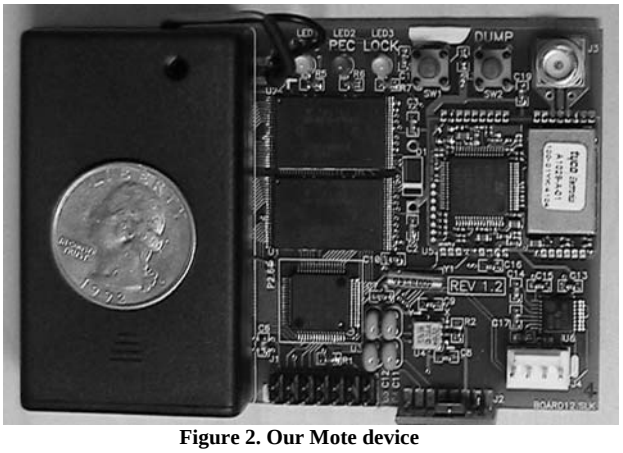


Figure 2. Our Mote device

Laptop PCs were used as the central home server, and Nokia 6680 mobile phones were used as the mobile servers. In addition, the mobile phones are used off the telephone network and communicate with the laptops via Bluetooth, which emulates a remote service, for example, a medical assistance's connection.

IV. Data acquisition

We propose the following framework for simultaneously detecting falls and classifying physical activities. In particular, we are interested in classifying transition movements including sit-to-stand, stand-to-sit, lie-to-stand and stand-to-lie. These four transition movements all involve rapid changes of acceleration in some or all axes. Rapid falls can be detected by simple thresholding when there is a sharp change of acceleration on the z-axis, but slow falls, such as collapsing after a heart attack, are much harder to discern due to a smaller change in acceleration. In addition normal activates of the subject must be discernible from emergency conditions to prevent false positive alarms. We propose a strategy of detecting falls by integrating three-axis accelerations with a fall-resembling activity classifier. With this approach, we can increase successful fall detection without increasing the rate of false positives. In the future, we can integrate rate gyroscopes and a biological sensor to further differentiate slow falls from a normal stand-to-lie movement (going to bed for example). The following flow chart illustrates our methodology where the board is attached to the body for our experiments.

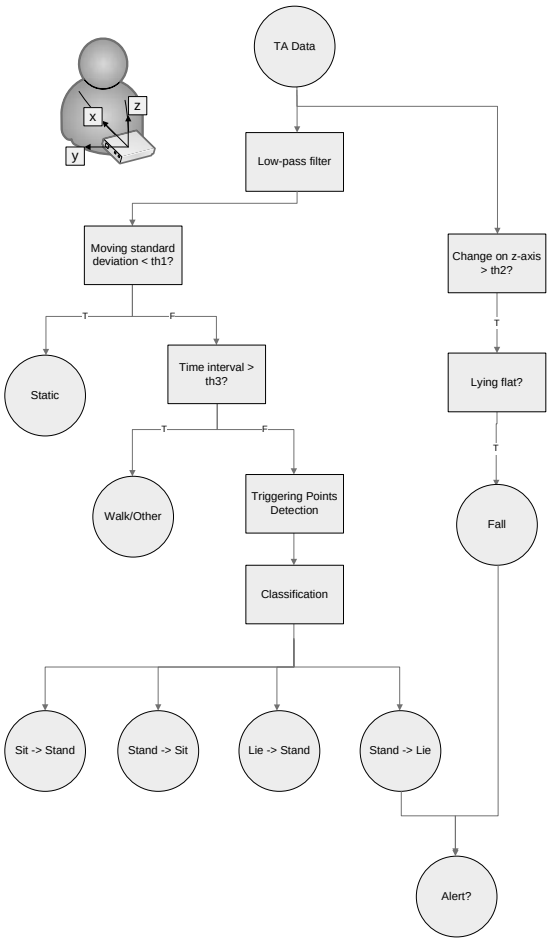


Figure 3. Signal processing flow-chart (*th1*, *th2*, *th3*: thresholds)

Given the goal of classifying movements based on subject motion, the functionality of our automated pattern recognition system is divided into two basic tasks: the description task which generates attributes of a movement using feature extraction techniques, and the classification task which classify this movement based on its attributes.

A. Preprocessing

The filtering is performed by a sixteen coefficient smoothing filter. The intuition behind using 2^n filter coefficients is that the division can be efficiently performed by the 16-bit MCU with a right-shift register. Unfiltered data is fed into the normal fall detection detector to avoid filtering out potential falls.

For classifying transition movements, hard thresholds (*th1* and *th2*, as shown in Figure 3) are first used to differentiate the activities into static, walk/other and transitions. To distinguish activity vs. rest, a majority vote from all three axes is used to make a decision. A triggering point detection mechanism is then used to further narrow down the regions of interesting activity. We define triggering points as places where the mean of the next *k* samples is greater or less than the mean of the previous *k* samples by a given threshold. This threshold is obtained by multiplying the maximum amplitude change over the transition interval by a given ratio (*r*). The choice of this ratio is only dependent on the window size *k*. We used *r=0.25* in our experiments.

B. Feature extraction

- Postural orientation: We employ the concept of postural orientation as indicated in [5]. We do this by tracking the angle between all sensors and the gravity (Constant inclination feature) from the beginning to the end of a movement. In fact, the absolute value of the z-axis itself is a fairly reliable indicator of the lying posture, since the axis would be usually pointing in a direction parallel to the floor, and does not change much even the person rolls from side to side.
- Singular value decomposition (SVD): One of the challenges of bio-signal analysis is to develop efficient methods to perform structural pattern recognition. SVD can be a valuable measure in obtaining such a characterization. SVD is a common technique for analysis of multivariate data, and therefore, data with unknown morphology might be well suited for analysis using this metric [9, 10]. Daily physical activities are composed of basic individual movements. For example, lying down can be composed of first sitting down and then lowering the upper part of body. Moreover, sitting and lying movements themselves may be decomposed into more fundamental motions. Such compositions are typically seen as several data segments with local maximums and minimums, as shown in Figures 4 and 5. We applied SVD analysis to these local extremums, weighing each extremum by its distance from the preceding extremum.

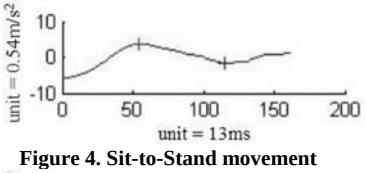


Figure 4. Sit-to-Stand movement

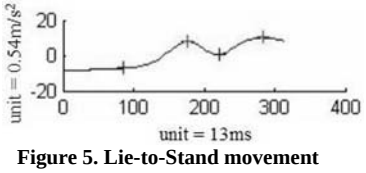


Figure 5. Lie-to-Stand movement

- Skewness: Skewness is computed on a region of interest by treating the signal as a distribution plot rather than a point plot. This is done by generating a sample space in which every acceleration value *v* at sample *t* on the original graph is converted to *v* points with value *t* in the new sample space. Skewness is then calculated based on this sample space. This approach differs from previous approaches because it can estimate the asymmetry of the profile of the signal. This measure is applied to both the x and y axes. The principle drawback of using this feature is that it is computationally expensive.
- Maximum amplitude change: Maximum amplitude change of the signal gives us an idea of how abrupt the transition is. Moreover, there is a direct correlation of this metric with the energy level of an individual. This is one of the causes of classification inaccuracy across different age groups.

C. Classification

We utilize both a neural network [11] and k^{th} nearest neighbor (k-NN) algorithm [12] to classify movements. Neural network classifiers are data-driven, which may better adapt to an idiosyncratic motions over time, whereas k-NN provides scalability for distributed sensing platforms.

V. Experimental results

Two young test subjects aged twenty and twenty-one imitated 68 types of falls. Our fall detection approach (Figure 3), accurately identified 118 out of 132 falls. The subjects also imitated 36 fall-resembling movements. Our normal fall detection approach based on hard thresholds resulted in 23 false-positives out of 168 fall/non-fall movements. The raw data from z-axis of an accelerometer is shown in Figure 6.

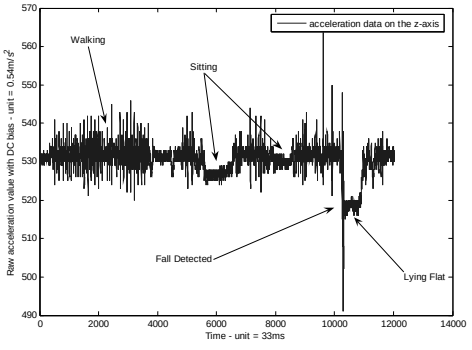


Figure 6. Raw acceleration data from z-axis

Furthermore, we carried out another set of experiments on four young subjects with an average age of twenty and seven elderly subjects with an averaged age of sixty four. They were asked to perform the following tasks:

- Walking on a straight line over a span of twenty feet and six times (as a control).
- Sit down and stand up three times each, from three different seats. The seats were a hardtop and a cushioned chair, as well as a couch. The test subjects were allowed to pause or walk around in between these actions.
- Lying down and getting up three times from the same bed.

Corresponding video sequences of the experiments were captured for comparing to and justifying our classification results.

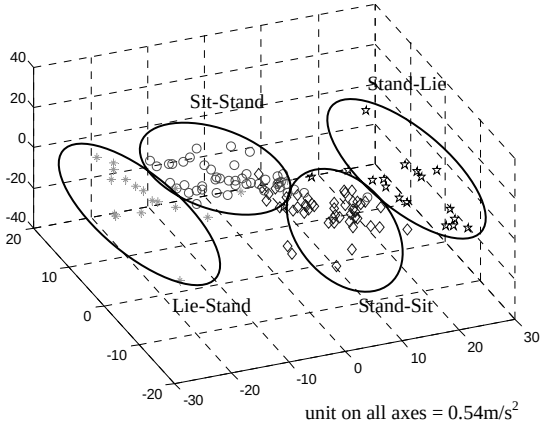


Figure 7. Constant inclination feature

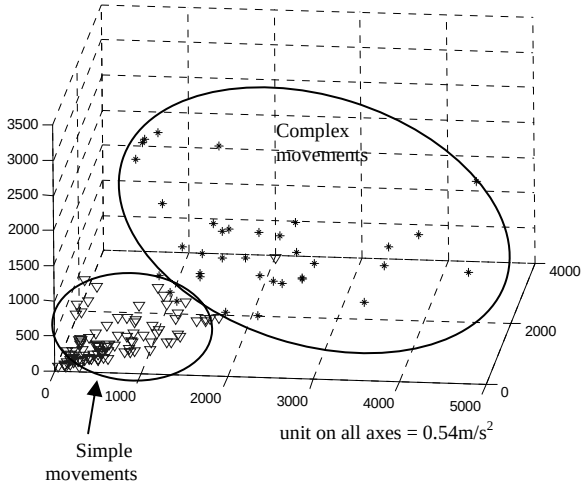


Figure 8. SVD feature

Figure 7 illustrates the clustering performed using the constant inclination features. The four movements, that include sit-to-stand, stand-to-sit, lie-to-stand and stand-to-lie, have been presented with several legends. Figure 8 exhibits the clustering performed with SVD feature. This figure demonstrates SVD can effectively discriminate simple movements from complex movements.

For the neural network classifier, we used a feed-forward design with eight first hidden layer nodes, four second hidden layer nodes and four output nodes for the four activities. The network was trained using back-propagation [13] with 100 epochs and an error margin of 0.01. We also used k-NN with *k* values of 5, 9 and 10.

Tables 1 to 4 illustrate the accuracy of classification using the constant inclination features. In this set of analyses, each classifier is trained and tested on the two age groups separately.

	Sit-Stand	Stand-Sit	Lie-Stand	Stand-Lie	% of success
Sit-Stand	26	4	4	2	**72.2**
Stand-Sit	0	29	0	8	**78.4**
Lie-Stand	0	0	5	3	**62.5**
Stand-Lie	3	0	0	5	**62.5**

Table 1. Results of using neural network trained with data from elderly subjects and tested on elderly subjects, training set size = 10.

	Sit-Stand	Stand-Sit	Lie-Stand	Stand-Lie	% of success
Sit-Stand	21	0	3	0	**87.5**
Stand-Sit	0	19	1	2	**86.4**
Lie-Stand	0	0	6	0	**100**
Stand-Lie	0	0	0	6	**100**

Table 2. Results of using neural network trained with data from young subjects and tested on young subjects, training set size = 5.

	Sit-Stand	Stand-Sit	Lie-Stand	Stand-Lie	% of success
Sit-Stand	31	4	1	0	**86.1**
Stand-Sit	1	35	0	1	**94.6**
Lie-Stand	2	0	5	1	**62.5**
Stand-Lie	0	3	0	5	**62.5**

Table 3. Results of using k-NN trained with the same set of data from elderly subjects and tested on elderly subjects, training set size = 10, k = 10.

	Sit-Stand	Stand-Sit	Lie-Stand	Stand-Lie	% of success
Sit-Stand	22	0	2	0	**91.7**
Stand-Sit	0	22	0	0	**100**
Lie-Stand	0	0	6	0	**100**
Stand-Lie	0	0	0	6	**100**

Table 4. Results of using k-NN trained with the same set of data from young subjects and tested on young subjects, training set size = 5, k = 5.

Next, we use the same constant inclination feature but train our classifiers with all the data from one age group and test them on the other group, as illustrated in Tables 5 to 8.

	Sit-Stand	Stand-Sit	Lie-Stand	Stand-Lie	% of success
Sit-Stand	41	5	0	0	**89.1**
Stand-Sit	4	41	0	1	**89.1**
Lie-Stand	8	0	7	3	**38.9**
Stand-Lie	0	7	3	8	**44.4**

Table 5. Trained neural network with constant inclination feature from young subjects and tested on elderly subjects.

	Sit-Stand	Stand-Sit	Lie-Stand	Stand-Lie	% of success
Sit-Stand	43	3	0	0	**93.5**
Stand-Sit	5	42	0	0	**89.4**
Lie-Stand	2	2	13	1	**72.2**
Stand-Lie	0	2	2	14	**77.7**

Table 6. Trained k-NN classifier with constant inclination feature from young subjects and tested on elderly subjects – k = 9.

	Sit-Stand	Stand-Sit	Lie-Stand	Stand-Lie	% of success
Sit-Stand	27	0	2	0	**93.1**
Stand-Sit	0	26	0	1	**96.3**
Lie-Stand	4	0	7	0	**63.6**
Stand-Lie	1	2	0	8	**72.7**

Table 7. Trained neural network with constant inclination feature from old subjects and tested on young subjects.

	Sit-Stand	Stand-Sit	Lie-Stand	Stand-Lie	% of success
Sit-Stand	25	0	4	0	**86.2**
Stand-Sit	0	25	0	2	**92.6**
Lie-Stand	1	0	10	0	**90.9**
Stand-Lie	0	1	0	10	**90.9**

Table 8. Trained k-NN classifier with constant inclination feature from old subjects and tested on young subjects, k = 9.

The results indicate that training classifiers with the data from elderly subjects and testing them on the young subjects yields higher accuracies than the reverse. We propose that this discrepancy is due to the fact that movements are more pronounced for the young subjects, thereby resulting in better clustering.

Next, the analysis was performed on classifying simple vs. complex movements. We use the SVD feature and the results are shown in Tables 9 to 12.

	Sit-Stand + Stand-Sit	Lie-Stand + Stand-Lie	% of success
Sit-Stand + Stand-Lie	59	14	**80.8**
Lie-Stand + Stand-Lie	8	8	**50.0**

Table 9. Results of using neural network trained with data from old subjects and tested on old subjects, training set size = 20.

	Sit-Stand + Stand-Sit	Lie-Stand + Stand-Lie	% of success
Sit-Stand + Stand-Lie	43	3	**93.5**
Lie-Stand + Stand-Lie	0	12	**100**

Table 10. Results of using neural network trained data from young subjects and tested on young subjects, training set size = 10.

	Sit-Stand + Stand-Sit	Lie-Stand + Stand-Lie	% of success
Sit-Stand + Stand-Lie	70	3	**95.9**
Lie-Stand + Stand-Lie	2	14	**87.5**

Table 11. Results of using k-NN trained with the same set of data from old subjects and tested on old subjects, training set size = 20, k = 10.

	Sit-Stand + Stand-Sit	Lie-Stand + Stand-Lie	% of success
Sit-Stand + Stand-Lie	40	6	**87.0**
Lie-Stand + Stand-Lie	1	11	**91.7**

Table 12. Results of using k-NN trained with the same set of data from young subjects and tested on young subjects, training set size = 10, k = 5.

For brevity, the classifications results using other features are omitted. On average, the classifications using maximum amplitude change and skewness achieved 76% and 80% success rate respectively.

VI. Conclusion

We have devised four novel features and implemented them through two classification schemes. The features include constant inclination, SVD, skewness and maximum amplitude change. Overall, we can classify four transition movements that include sit-to-stand, stand-to-sit, lie-to-stand and stand-to-lie, with an accuracy of 84%. Both constant inclination and skewness are effective features for classification in the elderly group, while maximum amplitude change is the most effective classification feature for the young group. SVD is useful in distinguishing complex movements from simple movements. The results indicate that training classifiers with the data from elderly subjects and testing them on the young subjects yields higher accuracies than the reverse. We propose that this discrepancy is due to the fact that movements are more pronounced for the young subjects, thereby yielding in better clustering. This is an important issue that needs to be addressed in this domain where the system must be adaptively customized for individuals. This issue was not explored in the previous work.

Acknowledgment

The authors would like to acknowledge Dr. Mikael Eklund for his valuable efforts integrating the research platform. The authors also would like to acknowledge NSF for grants CCR-0325274 and CCF-0424422.

References

[1] J. Porteus and S. Brownsell, "Exploring Technologies for Independent Living for Older People, A report on the Anchor Trust/BT Telecare Research Project," Anchor Trust, Fountain Court, Oxford Spires Business Park, Kidlington, Oxon 2000.

[2] P. E. Ross, "Managing CARE Through the Air," *IEEE Spectrum*, pp. 26-31, December 2004.

[3] H. G. C. Kemper and R. Verschuur, "Validity and reliability of pedometers in habitual activity research," *Eur. J. Appl. Physiol*, vol. 37, pp. 71-82, 1977.

[4] P. Avons, P. Garthwaite, H. L. Davies, P. R. Myrgatroyd, and W. P. T. James, "Approaches to estimating physical activity in the community: Calorimetric validation of actometers and heart rate monitoring," *Eur. J. Clin. Nutr.*, vol. 42, pp. 185-196, 1988.

[5] D. M. Karantonis, M. R. Narayanan, M. Mathie, N. H. Lovell, and B. G. Celler, "Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 10, pp. 156-167, 2006.

[6] J. A. Anliker U Fau - Ward, P. Ward Ja Fau - Lukowicz, G. Lukowicz P Fau - Troster, F. Troster G Fau - Dolveck, M. Dolveck F Fau - Baer, F. Baer M Fau - Keita, E. B. Keita F Fau - Schenker, F. Schenker Eb Fau - Catarsi, L. Catarsi F Fau - Coluccini, A. Coluccini L Fau - Belardinelli, D. Belardinelli A Fau - Shklarski, M. Shklarski D Fau - Alon, E. Alon M Fau - Hirt, R. Hirt E Fau - Schmid, M. Schmid R Fau - Vuskovic, and M. Vuskovic, "AMON: a wearable multiparameter medical monitoring and alert system."

[7] K. Najafi B Fau - Aminian, A. Aminian K Fau - Paraschiv-Ionescu, F. Paraschiv-Ionescu A Fau - Loew, C. J. Loew F Fau - Bula, P. Bula Cj Fau - Robert, and P. Robert, "Ambulatory system for human motion analysis using a kinematic sensor: monitoring of daily physical activity in the elderly."

[8] K. Aminian, "Monitoring Human Movement with Body-Fixed Sensors and its Clinical Applications," in *Computational intelligence for movement sciences : neural networks and other emerging techniques*, Rezaul Begg and M. Palaniswami, Eds.: Hershey, PA : Idea Group Pub., c2006, 2006.

[9] M. P. Fargues, R. Cristi, and M. M. Vanderkamp, "Modeling and classification of biological signals using least squares Prony-SVD AR modeling [underwater acoustics]," 1993, pp. 445-448 vol.1.

[10] L. Chuanjun, K. Latifur, and P. Balakrishnan, "Real-time classification of variable length multi-attribute motions." vol. 10: Springer-Verlag New York, Inc., 2006, pp. 163-183.

[11] G. L. Jonghun Baek, Wonbae Park, Byoung-Ju Yun, "Accelerometer Signal Processing for User Activity Detection," in *Knowledge-Based Intelligent Information and Engineering Systems*. vol. 3215: Springer Berlin / Heidelberg, 2004.

[12] T. Cover and P. Hart, "Nearest neighbor pattern classification," *Information Theory, IEEE Transactions on*, vol. 13, pp. 21-27, 1967.

[13] E. Wilson, "Backpropagation Learning for Systems with Discrete-Valued Functions," in *World Congress on Neural Networks* San Diego, 1994.

Address of the corresponding author:

Author: Roozbeh Jafari
Institute: Electrical Engineering and Computer Science Department, University of California, Berkeley
Street: 475 Hearst Memorial Mining Building
City: Berkeley
Country: USA
Email: rjafari@eecs.berkeley.edu

6th Session
Wearable Computing

BLIG: A New Approach for Sensor Identification, Grouping, and Authorisation in Body Sensor Networks

Jacob Andersen[1] and Jakob E. Bardram[2]

[1] Centre for Pervasive Healthcare, Department of Computer Science, University of Aarhus, Aarhus, Denmark
[2] IT University of Copenhagen, Copenhagen, Denmark

Abstract— **Using body sensor networks (BSN) in critical clinical settings like emergency units in hospitals or in accidents requires that such a network can be deployed, configured, and started in a fast and easy way, while maintaining trust in the network. In this paper we present a novel approach called *BLIG* (Blinking Led Indicated Grouping) for easy deployment of BSNs on patients in critical situations, including mechanisms for uniquely identifying and grouping sensor nodes belonging to a patient in a secure and trusted way. This approach has been designed in close cooperation with users, and easy deployment and ease of use are top priorities. We present an initial implementation and evaluation of the presented technology.**

Keywords— **body sensor network, WPAN, healthcare**

I. Introduction

The goal of wireless body sensor networks (BSNs) is to avoid cables and the drawbacks that come with them; cables are inherently unstable as cables and plugs tend to loosen or break under moderate stress, and cables tend to get tangled into each other. Patients with wired monitoring equipment are fixed to a bed or a chair and cannot easily move, or be moved, around. This has made the use of sensor technology impracticable in critical medical situation where sensors must be deployed in seconds and where patients need to be moved frequently. In our study of medical work in settings like emergency units in hospitals and larger accidents, we have discovered that sensors are not used at all, even though the same users say that sensors in many situations would be very useful [1]. Cables, however, have two very important features: first, the grouping of sensors and peripheral units like displays is *palpable* because you can see which nodes are attached to each other. Second, the sensor network is to a large degree *secure* – it is very difficult for an adversary to gain access to the medical data.

The overall objective of our work is to create a wireless BSN which is easy to deploy and use, while maintaining an appropriate level of security and privacy. In particular, we want a wireless BSN to include the two important features of *palpability* and *security* that a cabled BSN possesses. Such a BSN is especially required in critical medical circumstances like large accidents. Based on our studies of monitoring in e.g. the home of patients, there is clear evidence that these requirements are also valid in such more mundane areas of medical monitoring

This paper presents BLIG (short for: Blinking Led Indicated Grouping), which is an approach to deployment of wireless body sensor networks (BSN) on patients in critical situations, like accidents. We present the requirements for BLIG and explain how the design fulfils these requirements for fast, easy, and secure deployment. We then present the current implementation of BLIG and report from a user evaluation session with professional paramedics trying out the technology. We then conclude the paper and shortly discuss our plans for further work.

A. Problems

Figure 1 illustrates the problems BSNs inherently are dealing with. Throughout healthcare, usability issues are always important, demanded by the critical situations and the quick work pace. Security issues, such as privacy of medical information and logging of events are required by law and ethics. Furthermore, as battery technology is developing at a very slow rate, and energy storage per unit of volume or mass is lower than we would like it to be, we want to conserve power as much as possible and to avoid large and bulky equipment.

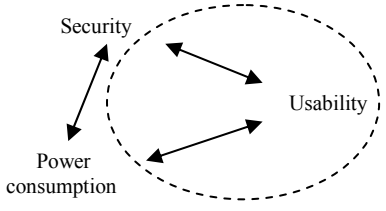


Fig. 1 The BSN problems and their relations. The dashed ellipse encloses the issues addressed by BLIG.

Solving the three problems tend to be problematic, as they conflict with each other, as illustrated by the arrows on the figure: good security properties requires lots of memory and processing power – conflicting with power conservation

– and keys may need to be distributed and users must log in and out, conflicting with simple and easy use. Also when designing the user interfaces, power saving requirements can have a huge impact on what is possible and what is not.

As illustrated by the dashed ellipse on the figure, BLIG was designed to be extremely simple to use, and at the same time have some reasonable security and power properties.

B. Requirements

Our design is grounded in various user-centred design processes ranging from supporting emergency workers involved in large accidents [1], to supporting mobility in hospitals [2], to supporting home care monitoring of patients and elderly people [3,4]. Based on this work we have distilled the following set of requirements:

- **Fast:** Most medical work is time-critical and often done on the move, so deployment of BSNs has to be swift. Data from a train collision emergency exercise showed that only a few seconds are spent with each victim during triage (sorting and prioritising victims). Hence, our goal is that deployment must take *only a few seconds*.
- **Easy:** Deployment takes place in a hectic environment where the users have to attend to many parallel tasks. Hence, the deployment must be very easy. Our goal is to enable users to deploy sensors *using only one hand*, while using the other hand for other things, such as supporting the patient.
- **Resilient and scalable:** The BSN must work reliably in a rugged environment like an outdoor accident scene or in the home, with little or no infrastructure available. Hence, deployment operations must only rely on communication between the actual units involved – i.e. no server or similar network resources should participate. This requirement is to ensure that the basic functionality can be accessed even during a major server or network breakdown. This requirement also ensures a scalable network since network latency will not slow down the process of adding or removing sensors.
- **Secure:** Medical data should be protected and access to it authorised. The goal is that the wireless BSN should not be inferior compared to existing (cabled) technology. The requirements are that communication should be encrypted, only authorised parties can gain access to the network and its data, sensors should be able to provide a security proof, and no unauthorised node should be able to deduce anything about the sensor (apart from its presence). [1]
- **Palpable:** Users need to be able to understand the BSN and investigate its configuration. The goal similarly is that investigating the wireless BSN should be just as easy as investigating how a cabled BSN is set up. The requirement is that users within a few seconds can tell which nodes belong to the same group, i.e. patient.

II. Related work

A method proposed by Baldus et. al. in [5] gives each clinician a personal pen with an infrared transmitter. This pen is then used to arrange sensor nodes on a single patient together in a group. The main problem remains: how can users quickly investigate these groups to inspect which nodes are members? No answer is given to this question.

The CodeBlue project [6] uses service discovery to establish links between sensors and displays. As new sensors are attached to patients, they just pop up at the displays. However, apparently there is no restriction limiting which displays are allowed to show the data, which means that everyone can get access to the information and no logging of who views the data takes place. Furthermore, only a numeric sensor id identifies the data source. This makes it hard to figure out which patient the data originates from. This implies a greater risk of mixing up patients and therefore more time spent checking and double-checking the associations between patient and sensor id.

III. BLIG design issues

It is important to realise that the requirements presented above are of a *technical* as well as a *usability* nature. We need a method for BSN identification, pairing, and authorisation which is technically adequate, while ensuring that users can use it quickly, with a high level of confidence, and understand the logical connections.

The BLIG method offers a solution to the usability issues: how to set up sensors in a fast, easy, and palpable way. Along with BLIG, we are developing a corresponding protocol with good security and power consumption properties. Although this protocol is far beyond the scope of this paper, we will give a few hints towards the relation between these properties and BLIG.

An example scenario, using the BLIG at an emergency could be this: an emergency worker wants to add 3 sensors to a patient. He takes the first sensor, turns it on, brings it close to his id-tag and then places it on the patient. He then

[1] It is beyond the limits of this paper to engage in a thorough threat analysis. But a few examples include: Adversaries gaining access to medical information about a patient, adversaries counterfeiting sensor readings and thus causing ill-treatment of the patient, and adversaries jamming the radio channels or launching a Denial-of-Service attack.

takes the next sensor, turns it on, brings it close to his id-tag and brings it close to the first sensor (in arbitrary order), and finally places it on the patient. Likewise for the third sensor (which is brought close to either one of the existing sensors as well as the id-tag). Later when he needs to reuse a sensor, he detaches it from the first patient, turns it off and on (re-set) and repeats the above procedure with another patient.

A. Identity of the patient

In most – if not all – existing patient monitoring equipment (such as [7,8]), and many related research projects (such as [6]), the sensors have no information about the identity of the patient. Instead, the sensors simply deliver the data to some (perhaps predefined) destination. It is then the job of the clinician to associate the incoming data at this destination with the identity of the patient. This involves two steps: attach the physical sensor to the patient, and associate the data with the identity of the patient (e.g. by typing in the patient's Social Security Number (SSN) at the monitoring central). If the second step does not take place at the patient's bed, the result is an increased risk of mistakes.

Data produced by the sensors are logically tied to the identity of the patient. However, in many cases the identity of the patient is unknown (e.g. in the accident scenario) and even if known, it is impractical to enter patient names or SSNs into sensors in a BSN. Therefore, in line with the method presented by Baldus et al. [5] we propose to group sensors on a patient together. This way, the identity of the sensor group is not depending on whether the true identity of the patient is known or not. When the identity of the patient is established, a mapping between the patient's "true id" and the sensors' group-id (the shared identity of all sensors on a single patient) can be made, e.g. by adding a PDA with proper communication capabilities to the sensor group and type in the "true id", such as the name or SSN. This procedure is carried out only once for each hospitalised patient – in contrast to once for each time a sensor is attached, changed or the patient is moved, which is the case for current sensors. Furthermore, the procedure would normally be carried out at the bedside, reducing the risk of errors such as patient mix-ups.

In contrast to the method suggested by Baldus et al. [5], however, we do not require the use of a special patient identification node or tag. We are proposing to let the first sensor attached to the patient generate a unique group-id. This id will be replicated to all sensors subsequently added to the patient. This way, not one specific sensor has a special role. This will make the sensor network more robust with respect to single point of failure.

B. Grouping

Once the group-id has been established (using the first sensor node) other sensor nodes can be attached to the BSN group. Hence, we need a method of grouping and un-grouping sensors. This is done by handing over the group-id to the node joining the group. Each sensor has a button for grouping (which could also serve as power button). When a patient has no sensors attached, the first sensor will generate its own unique group-id. Successive sensors are added to the group by pressing the grouping button while holding the new sensor close to any one of the already attached sensors. Note that we can use any node since the group-id is replicated amongst all participating nodes. Short range communication hardware (to be explained later) will be used to add the new sensor to the group of the nearby sensor. By 'short range' we mean no more than about half a meter, i.e. not a radio transceiver. For security reasons, which will be explained below, the user (clinician or emergency worker) doing the grouping must carry an authorisation-node (e.g. part of an id-tag), which shall also participate in the attachment procedure. Removing a sensor from a group is simply a matter of powering this sensor off. When a sensor is removed from the patient the remaining sensors will keep the original shared group-id.

C. Security

The short range communication channel is used to add new sensors to a group. During this procedure the new node must use the short range communication hardware to communicate with both an existing sensor on the patient and an authorisation-node worn by the user. These three nodes (or two, if the new node is the first to be attached to this patient) will cooperate to generate an unforgeable certificate for the new sensor. This certificate will be used by the new sensor to prove to others in or outside the group (e.g. another sensor or a display) that it truly belongs to the group and is in fact placed on the corresponding patient. Assuming "short range"' means less than about half a meter, the procedure places all nodes at the same place at the same time. An audible sound from the user's node during this procedure ensures that we have the user's attention; therefore the certificate is a proof that this user has approved the grouping.

If all users and their authorisation-nodes are trusted, this argument (by transitivity) states that all sensors belonging to the same group (with valid certificates) will in fact be placed on the same patient. Each time the authorisation node participates in an action, it provides a clear audible sound. This will alert the user if an adversary should try to abuse his or her node for an unauthorised action.

Due to the authorisation certificates a user can (perhaps remotely) inspect a group of nodes using e.g. a PDA to investigate who initially deployed the nodes. Furthermore, by (Danish) law [9] clinicians are required to document their actions. Automatically collecting and logging these certificates, the sensor network can now provide this documentation, saving the clinicians a lot of work.

D. Group palpability

In contrast to e.g. Baldus et al. [5], the grouping can be performed using only one hand and the user should be able to deploy a sensor within a few seconds. Our usability goals, however, were more ambitious: the user must also be able to quickly inspect and understand the grouping.

For this purpose, the presented design incorporates a mechanism which allows a sensor to reveal which group it belongs to. Each sensor node has a number of light-emitting diodes (LEDs) in different colours. All nodes in a group can present a synchronised blinking behaviour – i.e. the LEDs on all nodes belonging to one group will slowly and synchronously blink in matching colours. Each group will have its own unique blinking pattern (very long sequence of colours), which helps the user(s) to quickly verify whether all nodes on one patient are correctly grouped and whether the sensors on two (adjacent) patients belong to different groups (hence the name BLIG: Blinking Led Indicated Grouping). Sensors that fall off a patient are easy to re-establish on the correct patient.

To the user, the colour sequence will appear to be random. Technically, the sequence will be a repetitive pattern with a very long period, determined uniquely as a function of the group id and controlled by a common clock maintained between the nodes belonging to a group. This scheme is similar to the Frequency Hop Spread Spectrum (FHSS) radio modulation method used by e.g. Bluetooth (IEEE802.15.1) [10]. Instead of hopping among a set of radio channels, this scheme will hop among a set of colours. Clock drift is compensated by exchanging clock information between nodes. Since clock drifts in the order of up to tens of milliseconds are not a problem and the oscillators are rather accurate, this synchronisation can be a very rare event (i.e. not a scalability issue).

E. Hardware and power saving

The sensors' primary communication uses radio links. Furthermore, the sensors should have hardware for short range communication to be used in grouping and authorisation. Also, the sensors will need at least one button and a group of LEDs in different colours.

In order to save power, the LEDs should not be blinking all the time – only when a new sensor is added and up to a maximum of a few minutes after that. On the scene of a major emergency this behaviour might be unwanted, but then a command could be broadcast in the sensor networks at the emergency site, forcing all sensors to keep blinking until they leave the site. Furthermore, in the development of the protocol care has to be taken towards conserving power as well. Generally, we want the sensors to consume as little power as possible. Authorisation nodes, on the other hand, are allowed to spend more power, since they should only last for a single shift before being recharged. Therefore, radio and short range communication receivers can be switched on all the time and heavy computations can be performed here. This is exploited by the underlying protocol.

Short range communication can be achieved using a number of technologies. The key requirement is that communication can only take place if the devices are in a close proximity of each other – in the order of about half a meter. A number of technologies could be used, including light, sound, infrared light, ultrasound, and electromagnetic induction. For our prototypes, we have chosen induction, and we are not going to concern ourselves further with this issue. However, if BLIG is brought beyond the prototype implementation, further investigations will be necessary. In particular the security properties of each type should be compared: we assume that the local communication is truly *local*, but would it be possible for a technically skilled adversary to establish a "local" communication link at a distance? For instance, if infrared was used instead, could the adversary use an infrared laser from an adjacent room or through a window? Would visible light be better? – A visible light laser would certainly be easier to spot than an infrared laser. Other properties, like power consumption of different technologies, should also be examined.

IV. Prototype

A prototype intended for workshop based proof-of-concept testing has been developed. This prototype implementation was developed only to demonstrate the BLIG grouping mechanism, thus the secure protocol mentioned earlier, which we are also developing, has not been implemented in this prototype yet. The node hardware platform we use is the Berkeley mote of Telos revision B design ("Tmote sky" from Mote*iv*) extended with a coil and a couple of amplifier circuits for the short range communication channel. Figure 2 shows two motes. Mote 7 is a regular sensor node, which will be attached to patients, and mote

200 is a "clinician mote", i.e. it plays the role of the clinician's authorisation-node. This mote also has a piezoelectric buzzer. Figure 3 shows a rough schematic of the extension: two amplifiers built from a few transistors and an op-amp IC takes care of bridging the digital input and output from the mote's microcontroller with the coil. Both amplifiers can independently be turned off and on in order to save power. As of now, we use a very simple frequency modulation technique to transmit a number between 0 and 31 over the communication channel in less than 100 milliseconds. This is more than sufficient for the current prototype, but in the future, as the secure protocol will be implemented, we will have to improve this, in order to get a fast general data packet transfer.

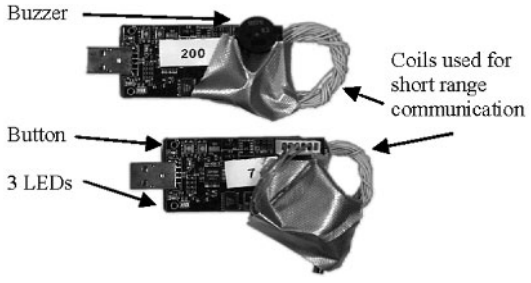


Fig. 2 A clinician mote and a regular sensor mote

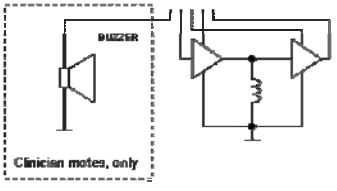


Fig. 3 Schematics of the short range communication extension.

Regular sensor nodes have 4 different states: ungrouped, searching, grouped, and error. When the node is turned on or reset, it enters the ungrouped state. While in this state, all LEDs are off. Pressing the grouping button, the node will enter the searching state for a maximum of 10 seconds, during which the LEDs will be glowing (with smooth intensity variations). If the node meets an authorisation-node on the short range channel within the time limit, it will generate a new group-id and enter the grouped state using this group-id. If the node meets both an authorisation-node and another regular node, which is already in the grouped state, within the time limit, the node will receive the group-id and current time from the other regular node. The node will then enter the grouped state using this group-id and synchronise its clock. If the node does not meet an authorisation-node within the time limit, it will enter the error state. In this state the red LED will be blinking aggressively. As a node enters the grouped state, it will begin blinking slowly. Having 3 coloured LEDs (red, green, and blue), there are 8 different combinations available. The combination chosen at any given time is a function of the group-id and a sequence number, which is incremented at a fixed interval.

Authorisation-nodes work in a straight-forward way: they simply provide a short beep sound from the buzzer each time they encounter a regular node in the searching state on the short range channel.

A PC running a small terminal-based Java program can be used to monitor the network, listing the sensors attached to each "patient" (group-id), and the identity of the user who attached it to the patient. The Java program is also capable of showing sensor readings.

V. Preliminary evaluation

Earlier prototype implementations without the short range communication hardware[2] were demonstrated at two workshops arranged by the PalCom project [11] in the fall of 2005 and the spring of 2006, involving participants from the local fire brigade and police force, emergency workers and trauma centre physicians. The reactions were very positive.

The participating clinicians expressed their appreciation with the easy grouping of the sensors and the tangible way groups can be recognised. Also, it was appreciated that the blinking LEDs are easy to see in all weather conditions, bright daylight and at night (something we never even considered), and that no small displays are involved, making it easy to inspect connections at a distance and without glasses. A few simple scenarios were played around the table, and the clinicians demonstrated that they could easily use and understand this method. They also clearly encouraged us to continue the development.

VI. Future work

At the time of this writing, we are planning to begin involving participants from a hospital ward. This hospital ward has both operating theatres and intensive care.

We are currently planning a new workshop where we are going to test the current version of the prototype described above. At this workshop we will be using people acting as patients and "faked" sensor readings on the displays. We

[2] "Proximity" between two units was emulated by pressing a button on both units

will be testing how quickly the clinicians are able to perform the different tasks, establish an overview of the local sensors and identify errors.

Later on, prototypes equipped with real sensors (ECG, pulse oximeter, blood pressure meter etc.) will be constructed and used in workshops, and perhaps as one part of a larger pilot test planned in the near future at the above mentioned hospital.

The pros and cons of different kinds of local communication hardware should be examined in regard to power consumption and how to ensure true local communication. As mentioned earlier, we have chosen not to look further into this subject at the moment.

The idea of using the blink pattern of LEDs to convey the notion of grouping of devices to an observing human is new and needs to be tested further, involving experts in the human visual perception capabilities. How fast should the colours change? Are some blinking patterns better than others? What if the clinician / emergency worker is colour-blind? Would different rhythms be helpful? Could blinking patterns and rhythms for instance be used to convey information about the condition of the patient? In the case of a major accident (like a train derailing) where triage is necessary, could the blinking pattern carry a message about the victim's condition (on a scale from "OK" to "critical care needed")?

More work need to be put into the protocol. In particular, we need to formalise and prove the security properties. We also need to do some more extensive scalability analysis and tests – the technology should be capable of handling rather large-scale emergencies with high densities of victims (think of train collisions or terror attacks).

VII. Conclusion

The idea of grouping all sensors on a patient together is not revolutionising. Others have gone down this road before us [5,12]. Our contribution is to the way this group identity will be made tangible for the user. The user will need a fast, one-button method of adding (and removing) a sensor device to a group, and a way to quickly inspect the correctness of the grouping (all sensors on the same patient are in the same group and the sensors on two different patient do not belong to the same group).

We have presented a new method of setting up networks of medical sensors on and around patients in hospitals or at emergency scenes. Besides easing the monitoring of the patient, this method offers a simple and intuitive way of identifying the patient, Furthermore, we are confident that the underlying protocol, which we are currently developing, will prove to be as secure against hardware (including network) failures and most adversary attacks, as the current (cabled) technology.

Acknowledgements

The research presented in this paper is done as part of the Activity-Based Computing Project (http://activity-based-computing.org/) funded by the Danish Research Council under the NABIIT programme, and as part of the PalCom project under the 6th Framework Programme, Information Society Technologies, Disappearing Computer II (http://www.ist-palcom.org).

References

1. Kristensen M, Kyng M, Nielsen E. T. (2005) IT support for healthcare professionals acting in major incidents, SHI 2005, 3rd Scandinavian conference on Health Informatics, pp 37–41
2. Bardram J E, Bossen C (2005) Mobility Work – The Spatial Dimension of Collaboration at a Hospital, Computer Supported Cooperative Work, 14(2), pp. 131–160 DOI 10.1007/s10606-005-0989-y
3. Bardram J E, Bossen C, Thomsen A (2005) Designing for transformations in collaboration: a study of the deployment of homecare technology, GROUP'05: Proceedings of the 2005 international ACM SIGGROUP conference on supporting group work, pp. 294–303 DOI 10.1145/1099203.1099254
4. Aaløkke S, Bunde-Pedersen J, Bardram J E (2006) Where to, Roberta?: Reflecting on the role of technology in assisted living, NordiCHI'06: Proceedings of the 4th Nordic conference on Human-computer interaction, pp 373–376 DOI 10.1145/1182475.1182515
5. Baldus H, Klabunde K, Müsch G (2004) Reliable Set-Up of Medical Body-Sensor Networks, EWSN'04 1st European Workshop on Wireless Sensor Networks, Springer LNCS 2920, pp. 353–363
6. Shnayder V, Chen B, Lorincz K, Fulford-Jones T R F, Welsh M (2005) Sensor Networks for Medical Care, Technical Report TR-08-05, Division of Engineering and Applied Sciences, Harvard University
7. GE Healthcare: Patient monitoring solutions, www.gehealthcare.com/usen/patient_mon_sys/wireless_and_telemetry/products/telemetry_sys/
8. Welch Allyn telemetry monitoring, monitoring.welchallyn.com/products/wireless/
9. Danish Ministry of Health, guidelines on nursing records (in Danish only), www.retsinfo.dk/DELFIN/HTML/C2005/0922960.htm
10. Haartsen J C, Zürbes S (2002) Frequency hop selection in the Bluetooth radio system, IEEE 7th International Symposium on Spread Spectrum Techniques and Applications, vol 1, pp. 83–87 DOI 10.1109/ISSSTA.2002.1049291
11. PalCom "major incidents" web site, www.ist-palcom.org/examplesOfWork/accidents.php
12. Ng J W P, Lo B P L, Wells O, Sloman M, Peters N, Darzi A, Toumazou C, Yang G-Z (2004) Ubiquitous Monitoring Environment for Wearable and Implantable Sensors (UbiMon), UbiComp'04: The 6th International Conference on Ubiquitous Computing, Poster Proceedings.

Address of the corresponding author:

Author: Jacob Andersen
Institute: Department of Computer Science, University of Aarhus
Street: Aabogade 34
City: 8200 Aarhus N.
Country: Denmark
Email: andersen@daimi.au.dk

Approximate Data Collection using Resolution Control based on Context

David Jea[1], Winston Wu[1], William J. Kaiser[1], and Mani B. Srivastava[1]

[1] Electrical Engineering Department, UCLA, Los Angeles, USA

Abstract— **Approximate data collection is an important mechanism for real-time and high sampling rate monitoring applications in body sensor networks, especially when there are multiple sensor sources. Unlike traditional approaches that utilize temporal or spatio-temporal correlations among the measurements of the multiple sensors observing a physical process to reduce the communication cost, in this paper we explore the idea of assigning different context-dependent priorities to the various sensors, and allocating communication resources according to data from a sensor according to its priorities. Specifically, a higher number of bits per sample is allocated to sensors that are of higher priority in the current context. We demonstrate that the proposed approach provides accurate inference results while effectively reducing the communication load.**

Keywords— **Approximate Data Collection, Data Aggregation, Context-aware systems, Body Sensor Networks.**

INTRODUCTION

Increasingly richer sensing sources have been incorporated in body sensor networks (BSN) to collect more accurate and detailed physiological information of a person for medical industry. However, the bandwidth of the shared wireless medium is limited and moreover higher data volume have a negative impact on battery life as the system energy consumption is dominated by wireless communication. This has hindered the progress of real-time high data rate monitoring applications in BSN. The situation gets worse when the sensor nodes communicate with each other through a low-power radio, which usually provides lower data rate levels. In-network aggregation is often used to address this issue by reducing communication load through distributing sensor data query operators down to the sensor nodes. The partial results from the different sources are then combined to reduce data redundancy at the intermediate nodes along the path to the base station. Similarly, instead of sending raw sensing data, sensor nodes can perform feature extraction on local data and send feature vectors only. However, in the light of deployment experiences with various sensor network applications, the above approaches are unattractive to domain experts [1], who often want to be able to reconstruct detailed sensor waveforms and not just events and features. Users in BSN research community also often need to communicate back complete sensed information.

To carry detailed sensing information over the resource-limited wireless medium, lossless and lossy compression algorithms may be employed to reduce communication load at the cost of increased encoding complexity. For a highly resource-constrained device, such as a sensor node, a lossless encoding scheme is usually too computationally intensive to be practical. A reasonable compromise to achieve the application fidelity requirements and bandwidth demands is to use lossy data compression schemes that trade information quality for lower computational complexity and higher compression ratio.

Approximate data collection [3] is an in-network lossy compression scheme for collecting data from sensor nodes. Current techniques mostly exploit temporal or spatio-temporal correlations among the measurements of the physical process made by different sensors. It requires domain experts to define the degree of precision (error bound ε) required for their queries. The returned answer is guaranteed to be ε-approximation of the real value with probability ρ. In this paper, the approximate data collection is addressed from a different angle. We embrace the idea of "*Based on the current context, the system dynamically adjusts the degree of precision for each sensor. Those sensors of interest provide more detailed information while the rest maintain just enough confidences in a inferred context.*" Indeed, from the viewpoint of a user, the information at different sensors has different significance levels (and hence should be given different priorities) in different contexts.

We believe that the "sensors of different priorities" concept is especially suitable for BSN. For example, say, when early signs of heart-attack (context) are being detected, the electrocardiogram (ECG) signals will become the highest priority information for diagnosis and should be delivered immediately and with high fidelity. Meanwhile, accelerometers used for classifying the physical context of a user (walking, running, limping, etc.) now become lower priority information sources that can be turned off. In this scenario, however, it is possible to further enhance our confidence in detecting a heart-attack event if the system allows partial accelerometer information to be involved. For example, running could trigger angina while walking might not [16].

We argue that an energy or bandwidth constrained BSN should not blindly collect maximum fidelity data from all the sensors since different sensors have different significance levels for different inferred contexts. Given a context,

the system shall deliver data with high fidelity or fine granularity for sensors of interest (high priority), and reduce information collected from the remaining sensors (low priority). However, it is domain experts who would have the knowledge to assign the significance level of each sensor under different contexts. Therefore, this paper does not address how to make such decision. Rather, this paper explores the above idea and presents an approach that enables **data resolution** (number of bits per sample) used to represent information from different sensors to be adapted to their priorities, while maintaining desired overall confidence in the inference result.

We propose in section I a framework that uses feedback paths between the BSN and the inference engine (a Bayesian classifier in our design) to dynamically control the required resolution based on the inferred contexts. However, the system shall only reduce the resolution of low-priority sensors if it can still provide users certain levels of confidence (threshold) in the inferred context. We then look into the details of such decision. The principles in part B of section I allows our system to determine that how the reduced information will affect user confidence in the inferred context. Section II presents a preliminary implementation and evaluation of our framework. We investigate the case in which the system equally reduces the resolution for all sensors and compare the results to the case where the system operates at different resolutions for different sensors based on their information priorities. Section III summarizes the related work. Section IV describes our conclusions.

I. Proposed Approach

A. Framework

We consider a general one-hop, star topology BSN scenario that consists of one base station and several sensor nodes. The physiological information is constantly collected by sensor nodes and delivered to a base station. An inference engine at the base station takes the features of sensor data as inputs and generates user context as output. Sensor nodes are assigned priorities in each context. The system attempts to collect detailed information from high priority sensor nodes and reduce fidelity information from low priority nodes. A user can setup the lowest acceptable resolution of a sensor to avoid data being degraded too much, yielding an unacceptable distortion level. The features extracted from these samples still allow the inference engine to have desirable confidence level in the current context. In this manner, approximate data collection is achieved based on the context.

Figure 1 shows the block diagram of the proposed approximation data collection on the base station. The initial setup is to collect data samples at the highest resolution from all the sensor nodes. The base station extracts features from these samples, and the inference engine classifies the context according to these features. The system imposes data rate constraints to the sensor nodes when it detects a network congestion event, reserves bandwidth for higher priority sensor nodes or simply wants to save energy. Given this data rate constraint, a naive solution is to equally reduce data resolutions and hence the data rate requirements of all nodes. In contrast, in our proposed approach, the base station selects low priority nodes based on the inferred context, and reduces data resolution of these nodes. The base station then uses the collected data samples to classify current context and possibly adjusts the resolution if it is not confident in the inferred context.

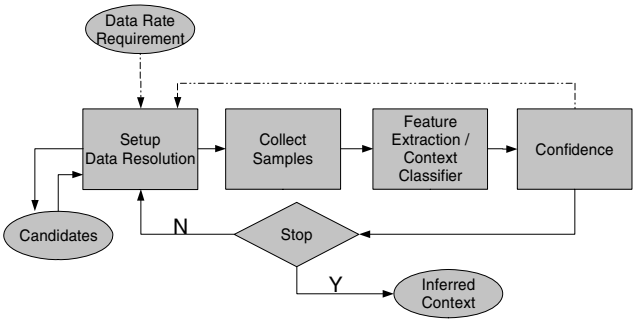


Figure 1: Block Diagram of the Base Station

Figure 2 depicts that each sensor node controls resolution of the collected samples based on the resolution requirement prescribed by the base station. Note that the finest resolution that a BSN node can provide is the original data without any resolution reduction.

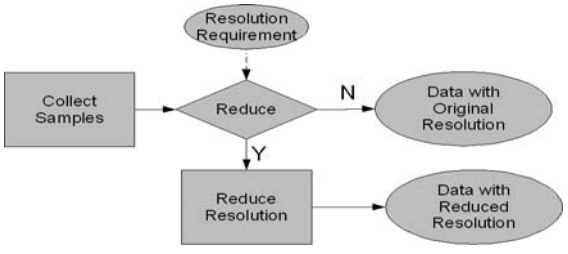


Figure 2: Block diagram of the BSN node

B. Principles

Our inference engine employs a Bayesian network [11] that combines multiple sensor observations for the purpose of inferring the context of a subject. The assumption is that we are confident in an inferred context if its occurrence probability is greater than T_H and the occurrence probabili-

ties of all the other contexts are less than T_L. The goal is to determine the current context using a reduced resolution version of sensor data.

Denote the n possible observations by O_1, O_2, …, O_n and the observation space by O_s, $O_s=\{O_1 \cup O_2 \ldots \cup O_n\}$.

The probability of a context C given an arbitrary observation O of a sensor is represented by the conditional probability $P(C|O)$.

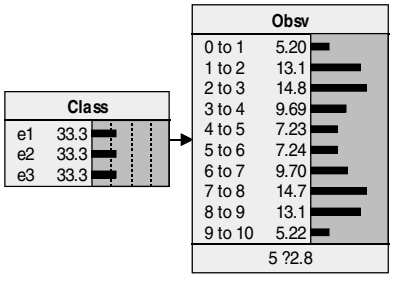
We note that for a partial (incomplete) observation O, knowing only the fact that O belongs to the set O_s is sufficient to determine a probability range for context C given O. To be specific, $P(C|O)$ has a lower bound and upper bound given by Eq. (1):

$$\forall O \in O_s$$
$$\exists i, j: \; P(C/O_i) \leqslant P(C/O) \leqslant P(C/O_j), \quad \cdots\cdots(1)$$
$$where\ i=\underset{m=1\ldots n}{argmin}\, P(C/O_m),$$
$$j=\underset{m=1\ldots n}{argmax}\, P(C/O_m)$$

Moreover, it is also sufficient to upper-bound the occurrence probability of any context other than C (denoted by C') given O, obtained in Eq. (2).

$$\forall C' \neq C,$$
$$\exists k: \; P(C'/O) \leqslant P(C'/O_k), \quad \cdots\cdots\cdots(2)$$
$$where\ k=\underset{m=1\ldots n}{argmax}\; P(C'/O_m)$$

If we know for a fact that an incomplete observation Θ belongs to O_s, we can use the $P(C|O_i)$ in Eq. (1) to lower-bound $P(C|\Theta)$. Similarly, $P(C'|O_k)$ in Eq. (2) is used to upper-bound $P(C'|\Theta)$. If the lower bound of $P(C|\Theta)$ is greater than T_H and the upper bound of $P(C'|\Theta)$ is less than T_L, then we are confident that, based on the assumption, the current context is C. Note that a low resolution observation is an incomplete observation, whose rough range instead of real (high-resolution) value is known. By replacing a complete observation by an incomplete (lower resolution) one (Θ), Eqs. (1) and (2) can serve to decide the corresponding confidence level in the inferred context. In the case that we are not confident about the inference results, higher resolution or other observations would be required.



	P(c1|O)	P(c2|O)	P(c3|O)
0-1	89	11	0
1-2	88.9	11.1	0
2-3	79.1	20.9	0
3-4	47.7	52.3	0.01
4-5	9.83	89.5	0.65
5-6	0.63	89.7	9.62
6-7	0.01	52.7	47.3
7-8	0	21.2	78.8
8-9	0	11.3	88.7
9-10	0	10.5	89.5

Figure 3: An example Bayesian classifier

In the following we present a simple classifier example for illustration purposes. The example is extended based on a Bayesian belief networks software, Netica [14]. Assume a sensor's observation (ranging from 0 to 10) has a resolution of unit 1. The sensor observations are characterized by a normal distribution, whose mean and standard deviation values are, respectively, equal to 2 and 1 when context one (c1) occurs, 5 and 2 when context two (c2) occurs, and 8 and 1 when context three (c3) occurs. Assume these three contexts take place with equal probabilities as shown to the left in Fig. 3. After we train the classifier with some random simulated samples, the table to the right in Fig. 3 shows the probability of each context given an observation. Fig. 4 shows the resulting confidence levels in a context based on reduced resolution observation.

Low Resolution Obsv	Lower bound		
	P(c1|O)	P(c2|O)	p(c3|O)
0 – 2	88.90%	11.00%	0.00%
2 – 4	47.70%	20.90%	0.00%
4 – 6	0.63%	89.50%	0.65%
6 – 8	0.00%	21.20%	47.30%
8 – 10	0.00%	10.50%	88.70%

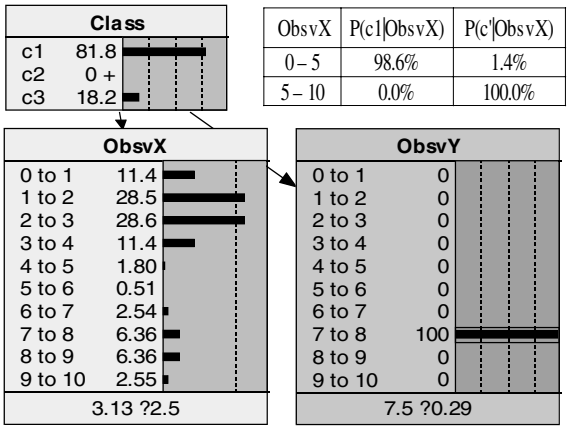
Figure 4: Results of resolution reduction to example classifier

We now extend the discussion to a BSN scenario that contains multiple sensors. In such a scenario, knowing the observation of a sensor will "pull" other sensors to be inclined towards its inference. This means that given that the observation from sensor one is x and the most likely context is C, the distribution of observations on sensor two is now in favor of context C. Note that observations from two sensors (x and y) are conditionally independent given context C in a naive Bayesian network. In this manner, the probability of context C for an observation y at sensor two, conditioned on that x is observed by sensor one, is given by Eq. (3):

$$P(C/y,x)=\frac{P(y/C)P(C/x)}{P(y/x)} \quad \cdots\cdots \quad (3)$$

Similar to the earlier discussion in the context of single sensor case, we may provide a lower bound of the probability of context C for a low resolution (incomplete) observation Y, $P(C|Y, x)$, $Y=\{y_i \leq y \leq y_k\}$, and also an upper bound of the probability of context C' ($C' \neq C$) given Y, $P(C'|Y, x)$. After replacing an complete observation y by a lower resolution representation Y, the lower bound of $P(C|Y, x)$ and the upper bound of $P(C'|Y, x)$ may determine our confidence level in the inferred context C.

We now extend the example by adding a second sensor, whose observations are also characterized by a normal distribution, with mean and standard deviation values equal to 7 and 2 when context one ($c1$) occurs, 2 and 1 when the second context ($c2$) occurs, and 4 and 2 when context three ($c3$) occurs. Figure 5 shows that if an observation from sensor two (indicated by ObsvY in Fig. 5) ranges from 7 to 8, then the observations of sensor one is pulled towards the in-



ObsvX	P(c1|ObsvX)	P(c'|ObsvX)
0 – 5	98.6%	1.4%
5 – 10	0.0%	100.0%

Figure 5: An example with two sensors and shows that the observation from a sensor pulls the other sensor towards its inference. A low resolution of ObsvX can greatly enhance the confidence to inferred context (c1).

ferred context, *c1*. In that case, given ObsvY = 7 ~ 8, a highly reduced resolution version (0 to 5) of sensor one observation (ObsvX) is able to increase our confidence level in context *c1* to *98.6%* and decrease to *1.35%* in other contexts $C' \neq c1$, as shown in the table to the top right of Fig. 5.

Note that most practical Bayesian inference engines take data features to classify context rather than plain observations. The paper does not address this gap in depth.

II. Implementation & Evaluation

We use real data that are collected from testbed experiments for evaluation. The system uses two three-axis accelerometers to collect physiological information of the subject and send back to the base station. We divide the collected data into two sets, training set (75%) and testing set (25%). When running the training set data, conditional probability tables are updated by a supervised learning setting in order to train the classifier to learn the characteristics of each context. The inference engine then extracts features from the testing data set and classifies the subject's context based on learned information. We evaluate the confidence levels in all contexts with different resolution reduction versions of the testing data set. The degree of reduction depends on the number of least significant bits (LSB) being reduced per sample, in which we normalize to ten bits.

We modified a system originally developed for Medical Embedded Device for Individualized Care (MEDIC) in [7]. The system now contains a base station (Nokia 770 [12]) and two sensor nodes (BlueSentry [13]). Each sensor node is equipped with one three-axis accelerometer and uses Bluetooth radio to communicate with the base station. The sensor nodes continuously relay sampled information (a total of six sensing channels) back to base station, and the base station saves received information into files with timestamps. In this experiment, for evaluation purpose, we use an off-line inference engine to perform classification. The prototype system is shown in Figure 6.

Figure 6: The prototype hardware system includes one base station and two sensor nodes. A knee brace is used for emulating limping context.

We investigated four contexts for this evaluation: walking, left leg limping, right leg limping, and stationary. We performed the experiments in the sixth floor hallway of UCLA Engineering IV building. The subject wore two sensor nodes (each with an accelerometer) around his waist, one on the left-hand side, and the other one on the right-hand side. Both accelerometers had the sampling rate set at 100 Hz on each channel (x, y, and z axis). We collected data under each context for approximately 15 to 30 minutes. To emulate limping context, we use a knee brace to enforce a fixed leg. The experiments collects a total of one and a half million samples for these four contexts (walk: ~27.4%, left leg limping: ~26.9%, right leg limping: ~26.6%, stationary: ~19.0%).

To extract features, the base station first converts data from time domain to frequency domain through the fast Fourier Transform (FFT) algorithm. Assume X(n) is the transformed version of the input time-series sensor data and N is its length. It then picks the amplitude of dominant frequency and calculates the integrated spectral energy as shown in Eqs. (4) and (5) as the two features considered in this evaluation ([10]). The inference engine takes these features as inputs and estimates the most probable context.

$$F_{amp} = max\left\|X(n)\right\| \quad \cdots\cdots\cdots(4)$$

$$F_{energy} = \sum_{1}^{N} X(n) * X(n) \quad \cdots\cdots\cdots(5)$$

We choose a naive Bayesian classifier as the inference engine in the system. A naive Bayesian classifier often works well in real world and assumes that features are conditionally independent. The classifier is a Bayesian network that contains two types of node: feature nodes and context nodes. Each feature node represents one of the features that

are extracted from collected data. Our inference engine employs continuous domain feature nodes and then quantizes the real value into a discrete level state. There are five data levels per "amplitude of dominant frequency" (ADF) feature node and ten data levels per "spectral energy" (SE) feature node. Uniform step-size quantization is applied, with the quantization dynamic range determined by the maximum and minimum value of its training data set. A better setup is possible but beyond the scope of this work. Note that the current setup is shown to be quite sufficient to classify the contexts and to fulfill our evaluation purpose. With the feature nodes being setup, the context node shows the inference result and has the four discrete context states that we consider (walking, left leg limping, right leg limping, and still). Due to space limitation, the result of the trained Bayesian network is presented in [15].

In this classifier, the inference of stationary context, "still", always has an accuracy of 100%. In addition, irrespective of how much information is being reduced in testing data set, its accuracy still remains at 100%. This is due to its unique characteristics (near zero) in all features, and with reduced data resolution, we actually push the data towards the stationary context. In the highly resolution reduction case, all the inferred results become "still" context. We thus omit its accuracy in following discussion.

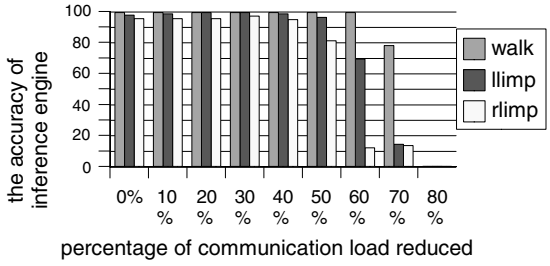


Figure 7: The effects of reducing data resolution in time domain on test data set of every sensing channel.

We refer the feature space to the union of all level states across all feature nodes in the rest of this paper. For example, in Fig. 5, ObsvX node has 10 discrete states and ObsvY node also has 10 discrete states, the feature space is therefore the 20 states from both nodes. After having learned from the training data set, we find that nearly **50%** of the feature space suggests a confidence level greater than 80% for the "**walking**" context or a confidence level less than 20% for any of the rest contexts. On the contrary, only 3.3% and 7.8% of the feature space indicates the same confidence level (80%, 20%) for "left leg limping" and "right leg limping" contexts, respectively. In other words, the inference engine can reduce the data resolution of a feature for "walk" context, by combining 50% states of the feature space as discussed in section 3, without losing noticeable

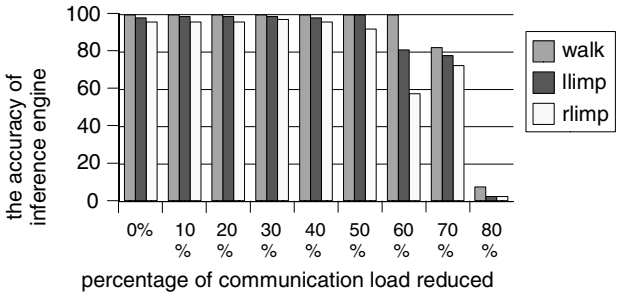


Figure 8: Keeping original resolution on the sensor of interest (the y-axis of accelerometer on left leg). The effects of reducing data resolution in time domain on test data set of the rest sensing channels.

confidence, while the "left leg limping" and "right leg limping" contexts are more sensitive to such resolution reduction. In Fig. 7, we show the resulting inference accuracy of the naive solution that equally reduces resolution on all sensing channels. It is noted that "walking" context manages to maintain high accuracy level even with largely reduced data resolution, while the accuracy levels for the other two contexts are significantly degraded as the resolution is reduced beyond a certain level.

In Figure 8, in contrast to the naive approach used in Fig. 7, the sensing information collected by the y-axis of the accelerometer on the left waist (ACCY_L) is set to high priority and assigned to transmit with complete information. In doing so, we believe that further resolution reduction on other sensors can be performed to provide the same inference accuracy. As described earlier, based on the assumption that ACCY_L is of more importance for the considered context, the other sensing channels are for the large part playing the role of assisting and refining the inference result of ACCY_L. We verified this by comparing the results of Figure 7 to Figure 8 for the cases where sixty or seventy percent of information are discarded on the testing data set. Figure 8 intelligibly achieves more accurate inferences using these highly degrade samples. Figure 9 shows the lowest resolution allowed in both approaches to yield 80% confidence in all inferred contexts. It is thus clearly desirable and necessary to use a more intelligent solution, such as the proposed approach, for adapting the contexts that are sensitive to data resolution reductions to collect data.

III. Related work

Research communities in different fields have long investigated in monitoring applications. Directed Diffusion [1] is a data-centric protocol for general data collection in sensor networks. Cougar [8] and TinyDB [9] is a database abstraction of sensor networks and provide declarative in-

	walk	left leg limping	right leg limping
Figure 7	5	5	5
Figure 8	3	4	5

Figure 9: The lowest resolution (number of bits) out of 10 bits for a sensing channel that results 80% confidence to the inferred context.

terfaces. Skordylis et al. conduct a complete survey [2] in approximate data management for sensor networks.

In [3], Chu et al., motivated by existing deployment, attack the "SELECT *" problem for sensor networks. The proposed approach compresses data using replicated dynamic probabilistic models to guarantee a fixed error bound from the measurement readings. The work is similar to BBQ approach proposed in [4] but pull-based. Both work establish statistic models of real worlds to reduce sensing and communication overheads. Jain et al. [5] explore temporal correlation of a data source and introduce a dual Kalman filter architecture to conserve network bandwidth. The server and the source maintain same copies of a Kalman filter. The server uses the filter to predict data from the source, and the source updates the corresponding filter on server if the prediction is outside of given precision.

Lazaridis and Mehrotra [6] propose the piecewise constant approximation (PCA) approach for data producer to transmit time series to data archiver. The approach represents time series with a sequence of segments (c_i, e_i), where c_i is a constant value for time in $[e_{i-1}+1, e_i]$.

The Bayesian classifier has been proved to be effective for motion recognition applications. In [17], twelve three-axis accelerometers and a naive Bayesian classifier are used to capture postures and activities of a user. Korpip et al. [18] apply naive Bayesian networks to classify user daily activities based on audio features. Du et al. [19] divide features into global and local classes, and present a dynamic Bayesian network model to recognize interacting activities.

IV. Conclusions

In this paper, we have presented a framework and its underlying principles that enable context-based resolution controls to collect approximate data samples. We adapt a two-tier network model and aim non-aggregate data collection for users. The system provides deterministic error bounds through resolution reduction. Our work explores reducing communication load based on context and using rough data resolution for low priority nodes. We demonstrate the implementation and conduct evaluations using real data collected from the experiments. The proposed approach yields accurate inferred contexts even when the resolution of low priority sensors has been greatly degraded.

Acknowledgment

The research described in this paper was funded in part by the NSF award # CNS-0306408, NSF grant # ANI-00331481, NIH grant # NLM T15LM07356, and by UCLA Center for Embedded Networked Sensing. Any opinions, findings and conclusions or recommendations expressed in this paper are those of the author(s) and do not necessarily reflect the views of the NSF or CENS.

References

1. C. Intanagonwiwat, R. Govindan, and D. Estrin., "Directed diffusion: A scalable and robust communication paradigm for sensor networks," *Proc of ACM MOBICOM*, 2000.
2. A. Skordylis, N. Trigoni and A. Guitton, "A Study of Approximate Data Management Techniques for Sensor Networks," *Proc of the 4th Workshop on Intelligent Solutions in Embedded Systems*, 2006.
3. D. Chu, A. Deshpande, J. M. Hellerstein, and W. Hong, "Approximate Data Collection in Sensor Networks using Probabilistic Models," *Proceedings of the 22nd International Conf. on Data Engineering* (ICDE), 2006.
4. A. Deshpande, C. Guestrin, S. Madden, J. Hellerstein, and W. Hong. "Model-driven data acquisition in sensor networks," *In the 30th International Conf on Very Large Data Bases* (VLDB), 2004.
5. A. Jain, E. Chang, and Y.-F.Wang. "Adaptive stream resource management using Kalman Filters," *Proceedings of the 2004 ACM SIGMOD international conference on Management of data*, 2004.
6. I. Lazaridis and S. Mehrotra. "Capturing sensor-generated time series with quality guarantees," *Proceedings of the 19th International Conf. on Data Engineering (ICDE)*, 2003.
7. W.H. Wu, A.A.T. Bui, M.A. Batalin, L.K. Au, J.D. Binney, and W.J. Kaiser, "MEDIC: Medical Embedded Device for Individualized Care," *Artificial Intelligence in Medicine* (2006 submitted).
8. Y. Yao and J. Gehrke. "The Cougar approach to in-network query processing in sensor networks," *ACM SIGMOD record,* 2002.
9. S. Madden, M. J. Franklin, J. M. Hellerstein, W. Hong. "TinyDB: an acquisitional query processing system for sensor networks," *ACM Transactions on Database Systems (TODS),* 2005.
10. W.H. Wu, A.A.T. Bui, M.A. Batalin, D. Liu, and W.J. Kaiser, "Incremental Diagnosis Method for Intelligent Wearable Sensor Systems," *IEEE Trans Info Technol Biomed* (2006 accepted).
11. Richard E. Neapolitan, "*Learning Bayesian Networks*", Prentice Hall, 2004.
12. Nokia 770, http://www.nokiausa.com/770
13. BlueSentry, http://www.gridenabled.com/bluesentry.html
14. Netica, http://www.norsys.com/
15. Trained Bayesian Network, http://www.ee.ucla.edu/~dcjea/researches/drc/TrainedBayesNet.jpg
16. American Heart Association, http://www.americanheart.org/presenter.jhtml?identifier=4472
17. N. Kern, B. Schiele, and A. Schmidt, "Multi-sensor activity context detection for wearable computing," *Proc. of Eur. Symp. on Ambient Intelligence (EUSAI)*, pp. 220-232, November 2003.
18. P. Korpip, M. Koskinen, J. Peltola, S. Mkel, and T. Seppnen, "Bayesian approach to sensor-based context awareness," Personal and Ubiquitous Comp. J, vol. 7, pp. 113-124, February 2003.
19. Y. Du, F. Chen, W. Xu, and Y. Li, "Recognizing Interaction Activities using Dynamic Bayesian Network," *Proceedings of the 18th International Conference on Pattern Recognition (ICPR'06)*, 2006.

Recognizing Soldier Activities in the Field

David Minnen[1], Tracy Westeyn[1], Daniel Ashbrook[1], Peter Presti[2], and Thad Starner[1]

[1]Georgia Institute of Technology, College of Computing, GVU, Atlanta, Georgia, USA. {dminn, turtle, anjiro, thad}@cc.gatech.edu

[2]Georgia Institute of Technology, Interactive Media Technology Center, Atlanta, Georgia, USA. peter.presti@imtc.gatech.edu

***Abstract*— We describe the activity recognition component of the Soldier Assist System (SAS), which was built to meet the goals of DARPA's Advanced Soldier Sensor Information System and Technology (ASSIST) program. As a whole, SAS provides an integrated solution that includes on-body data capture, automatic recognition of soldier activity, and a multimedia interface that combines data search and exploration. The recognition component analyzes readings from six on-body accelerometers to identify activity. The activities are modeled by boosted 1D classifiers, which allows efficient selection of the most useful features within the learning algorithm. We present empirical results based on data collected at Georgia Tech and at the Army's Aberdeen Proving Grounds during official testing by a DARPA appointed NIST evaluation team. Our approach achieves 78.7% for continuous event recognition and 70.3% frame level accuracy. The accuracy increases to 90.3% and 90.3% respectively when considering only the modeled activities. In addition to standard error metrics, we discuss error division diagrams (EDDs) for several Aberdeen data sequences to provide a rich visual representation of the performance of our system.**

***Keywords*— activity recognition, machine learning, distributed sensors**

I. Introduction

The Soldier Assist System (SAS) was created for DARPA's Advanced Soldier Sensor Information System and Technology (ASSIST) program. The stated goal of the ASSIST program is to "enhance battlefield awareness via exploitation of information collected by soldier-worn sensors." The task was to develop an "integrated system and advanced technologies for processing, digitizing and reporting key observational and experiential data captured by warfighters" [1]. SAS allows soldiers to utilize on-body sensors with an intelligent desktop interface to augment their post-patrol recall and reporting capability (see Figure 1). These post-patrol reports, known as after action reports (AARs) are essential for communicating information to superiors and future patrols. For example, a soldier may return from a five hour patrol and his superiors may ask for a report on all suspicious situations encountered on patrol. In generating his report, the soldier can use the SAS desktop interface to view all of the images taken around the time of a particular activity such as taking a knee or raising his field weapon. These situations are likely to contain faces, places, or other events of interest to the soldier. These images can be used to jog the soldier's memory, provide details that may have gone unnoticed during patrol, and provide relevant data to be directly included in the soldier's AAR for further analysis by superiors.

Fig. 1. Screenshot of the SAS desktop interface

The automatic recognition of salient activities may be a key part of reducing the amount of time spent manually searching through data. Working with NIST, DARPA, and subject matter experts, we identified 14 activities important for generating AARs: running, crawling, lying on the ground, a weapon up state, kneeling, walking, driving, sitting, walking up stairs, walking down stairs, situation assessment, shaking hands, opening a door, and standing still. During interviews, soldiers identified the first five activities as the most important.

Given these specific requirements, our sensing hardware was tailored to these activities and for the collection of rich multimedia data. The on-body sensors include six three-axis accelerometers, two microphones, a high resolution still camera, a video camera, a GPS, an altimeter, and a digital compass (see Figure 2). The soldier activities are recognized by analyzing the accelerometer readings, while the other sensors provide content for the desktop interface.

Given the number of sensors and the magnitude of the data, we used boosting to efficiently select features and learn accurate classifiers. Our empirical results show that many

of the automatically selected features are intuitive and help justify the use of distributed sensors. For example, driving can be detected by identifying engine vibration detected by the hip sensor while in a moving vehicle.

In Section V, we present cross-validated accuracy measurements for several system variations. Standard accuracy metrics do not always account for the impact that different error types have on continuous activity recognition applications, however. For example, a trade-off could exist between identifying all instances of a raised weapon (potentially with imprecise boundaries) and obtaining precise boundaries for each identified event but failing to find all of the instances. To better characterize the performance of our approach, we discuss error division diagrams (EDDs), a recently introduced performance visualization that allows quick, visual comparisons between different continuous recognition systems [2]. The SAS system favors higher recall rates rather than higher precision rates because the activity recognition is used to highlight salient information for a soldier. Failing to identify events can cause important information to be overlooked, while a soldier can easily cope with potential boundary errors by scanning the surrounding media.

II. Related Work

Many applications in ubiquitous and wearable computing support the collection and automatic identification of daily activities using on-body sensing [3], [4], [5]. Westeyn and Vadas *et al.* proposed the use of three on-body accelerometers to support the identification and recording of autistic self-stimulatory behaviors for later evaluation by autism specialists [6]. In 2004, Bao and Intille showed that the use of two accelerometers positioned at the waist and upper arm were sufficient to recognize 20 distinct activities using the C4.5 decision tree learning algorithm with overall accuracy rates of 84% [7]. Lester *et al.* reduced the number of sensor locations to one position by incorporating multiple sensor modalities into a single device [8]. Using a hybrid of both discriminative and generative machine learning methods (modified AdaBoost and HMMs) they recognized 10 activities with an overall accuracy of 95%. Finally, Raj *et al.* simultaneously estimated location and activity using a dynamic Bayesian network to model sensor activity and location dependencies while relying on a particle filter for efficient inference [9]. Their approach performs similarly to that of Lester *et al.* and elegantly combines activity recognition and location estimation.

III. Soldier Assist System On-Body Sensors

The SAS on-body system is designed to be unobtrusive and allow for all-day collection of rich data. The low power, lightweight system consists of sensors spanning several modalities along with a wireless on-body data collection device. Importantly, the sensors are distributed at key locations on the body, and all but two are affixed to standard issue equipment already worn by a soldier.

Fig. 2. Left: Soldier wearing the SAS sensor rig at the Aberdeen Proving Grounds. Right: Open Camelbak showing OQO computer and connectors.

The major sensing components used for soldier activity recognition are the six three-axis bluetooth accelerometers positioned on the right thigh sidearm holster, in the left vest chest pocket, on the barrel of the field weapon, on a belt over the right hip, and on each wrist via velcro straps. Each accelerometer component consists of two perpendicularly mounted dual-axis Analog Devices ADXL202JE accelerometers capable of sensing $\pm$2G static and dynamic acceleration. Each accelerometer is sampled at 60Hz by a PIC microcontroller (PIC 16F876). The Bluetooth radio is a Taiyo Yuden serial port replacement module that handles the transport and protocol details of the Bluetooth connection. Each sensor is powered by a single 3.6V 700mA lithium ion battery providing 15-20 hours of run-time.

Data from the wireless accelerometers (and other streaming sensors) are collected by an OQO Model 01+ palmtop computer. The OQO and supporting equipment is carried inside the soldier's Camelbak water backpack. The OQO sits against the Camelbak water reservoir to both cushion the OQO against sudden shock and to to help avoid overheating. Two 40 watt-hour camcorder batteries provide additional power to the OQO and supply main power to other components in the system. These batteries provide the system with a run time of over 4 hours.

Our system also contains off-the-shelf sensing components for location, video, and audio data. Affixed to the shoulder of the vest is a Garmin GPSMAP 60CSx. The unit also houses a barometric altimeter and a digital compass.

Two off-the-shelf cameras are mounted on the soldier's helmet. The first is a five megapixel Canon S500 attached to the front of the helmet on an adjustable mount. A power cable and USB cable run from the helmet to the soldier's

backpack. The OQO automatically triggers the camera's shutter approximately every 5 seconds and stores the images on its internal hard drive. A Fisher FVD-C1 mounted on the side of the helmet. This camera records MPEG-4 video at 640x480 to an internal SD memory card.

Our system has two microphones for recording the soldier's speech and ambient audio in the environment. One microphone is attached to the helmet and positioned in front of the soldier's mouth by a boom. The second microphone is attached to the arm strap of the Camelbak near the chest of the soldier. Both microphones have foam wind shields and are powered by a USB hub attached to the OQO.

The various sensors are not hardware synchronized. To account for asynchronous sensor readings, we ran experiments to estimate the latency associated with each sensor. The collection and recording software running on the OQO is responsible for storing a timestamp with each sensor reading and applying the appropriate temporal offset to account for the hardware latency. This provides rough synchronization across the different sensors and frees later processes, such as the data query interface and the recognition system, from needing to adjust the timestamps individually.

IV. Learning Activity Models

Two fundamental issues must be addressed in order to automatically distinguish between the 14 soldier activities that our system must identify. The first is how to incorporate sensor readings through time, and the second concerns choosing a sufficiently rich hypothesis space that allows accurate discrimination and efficient learning. Although several models have been proposed that explicitly deal with temporal data (*e.g.,* hidden Markov models (HMMs) and switching linear dynamic systems (SLDSs)), these models can be very complex and typically expend a great deal of representational power modeling aspects of a signal that are not important to our system. Instead, we have adopted an approach that classifies activities based on aggregate features computed over a short temporal window similar to the approach of Lester *et al.* [8].

A. Data Preprocessing

Several steps are needed to prepare the raw accelerometer readings for analysis. First, we resample each sensor stream at 60Hz to estimate the instantaneous reading of every sensor at the same fixed intervals. Next, we slide a three second window along the 18D synchronized time series in one second steps. For each window, we compute 378 features based on the 18 sensor readings including mean, variance, RMS, energy in various frequency bands, and differential descriptors for each dimension. We also compute aggregate features based on each three-axis accelerometer. Originally our feature list was augmented with GPS derivatives, heading, and barometer readings. However, preliminary experiments showed that this did not significantly improve performance and these features were not included in the experiments discussed in this paper. This finding is in line with that of Raj *et al.* who also determined that GPS-derived features did not aid activity recognition in a similar task using a different sensor package [9].

Effectively, we have transformed a difficult temporal pattern recognition problem into a simpler spatial classification task by computing aggregate features. One potential difficulty is determining which features are important for distinguishing each activity. Rather than explicitly choosing relevant features as a preprocessing step (either manually or through an automatic feature selection algorithm), we build models using an AdaBoost framework [10] by selecting the dimension and corresponding 1D classifier that minimize classification error at each iteration. This framework automatically selects the best features for discrimination and is robust to unimportant or otherwise distracting features.

B. Boosted 1D Classifiers

AdaBoost is an iterative framework for combining binary classifiers such that a more accurate ensemble classifier results [10]. We use a variant of the original formulation that includes feature selection and support for unequal class sizes [11]. The ensemble classifier is based on the sign of a value derived from a weighted combination of the weak classifiers called the margin. The magnitude of the margin gives an indication of the confidence of the classifier in the result. Margins may not be comparable across different classifiers however, but we can use a method developed by Platt to convert the margin into a pseudo-probability [12]. This method works by learning the parameters of a sigmoid function ($f(x) = 1/(1+e^{A(x+B)})$) that directly maps a margin to a probability. We fit the sigmoid to margin/probability pairs derived from the training points and only save the sigmoid parameters (A and B) for use during inference.

C. 1D Classifiers

During each round of boosting, the algorithm selects the feature and 1D classifier that minimizes the weighted training error. Typically, decision stumps are used as the 1D classifier due to their simplicity and efficient, globally optimal learning algorithm. Decision stumps divide the feature range into two regions, one labeled as the positive class and the other as the negative class. In our experiments,

we supplement decision stumps with a Gaussian classifier that models each class with a Gaussian distribution and allows for one, two, or three decision regions depending on the model parameters.

Learning the Gaussian classifier is straightforward. The parameters for the two Gaussians ((μ_1, σ_1) and (μ_2, σ_2)) can be estimated directly from the weighted data. The decision boundaries are then found by equating the two Gaussian formulas and solving the resulting quadratic equation for x: $(\sigma_1^2 - \sigma_2^2)x^2 - 2(\sigma_1^2\mu_2 - \sigma_2^2\mu_1)x + [(\sigma_1^2\mu_2^2 - \sigma_2^2\mu_1^2) - 2\sigma_1^2\sigma_2^2 log(\frac{\sigma_1}{\sigma_2})] = 0$.

D. Multiclass Classi cation

The boosting framework can be used to learn accurate binary classifiers but direct generalization to the multiclass case is difficult. Two common methods for combining multiple binary classifiers into a single multiclass classifier are the one-vs-all (OVA) and the one-vs-one (OVO) approaches. In the one-vs-all case, for C classes, C different binary classifiers are learned. For each classifier, one of the classes, call it ω_i, is taken as the positive class while the $(C-1)$ others are combined to form the negative class. When a new feature vector must be classified, each of the C classifiers is applied and the one with the largest margin (corresponding to the most confident positive classifier) is taken as the final classification. Alternatively, we can base the decision on the pseudo-probability of each class computed from the learned sigmoid functions.

In the one-vs-one approach, $C(C-1)/2$ classifiers are learned, one for each pair of (distinct) classes. To classify a new feature vector, the margin is calculated for each class pair (m_{ij} is the margin for the classifier trained for ω_i vs. ω_j). Within this framework many methods may be used to combine the individual classification results:

- *vote[ms,ps,pp]* - each classifier votes for the class with the largest margin; ties are broken by using *msum*, *psum*, or *pprod*.
- *msum* - each class is scored as the sum of the individual class margins; the final classification is based on the highest score: $\underset{i}{argmax} \ \sum_{j=1}^{C} m_{ij}(x)$
- *psum* - each class is scored as the sum of the individual class probabilities; the final classification is based on the highest score: $\underset{i}{argmax} \ \sum_{j=1}^{C} p(\omega_i|m_{ij}(x))$
- *pprod* - the classification is based on the most likely class: $\underset{i}{argmax} \ \prod_{j=1}^{C} p(\omega_i|m_{ij}(x))$

E. Combining Classi cations from Overlapping Windows

As discussed in Section IV-A, we transform a temporal recognition problem into a spatial classification task by computing features over three second windows spaced at one second intervals. This means that typically there are three different windows that overlap for each second of data. To produce a single classification for each one second interval, we tested three methods:

- *vote* - each window votes for a single class and the class with the most votes is selected
- *prob* - each window submits a probability for each class, and the most likely class is selected: $\underset{i}{argmax} \ \prod_{j=1}^{3} p(\omega_i|window_j)$
- *probsum* - each window submits a probability for each class, and the class with the highest probability sum is selected: $\underset{i}{argmax} \ \sum_{j=1}^{3} p(\omega_i|window_j)$

V. Experimental Results

To evaluate our approach to recognizing soldier activity, we collected data in three situations. In each case, the subject wore the full SAS rig and completed an obstacle course designed to exercise all of the activities. We collected 14 sequences totaling roughly 3.3 hours of activity data and then manually labeled each sequence based on the video and audio captured automatically by the system. The first round of data collection took place in a parking deck near our lab. We created an obstacle course based on feedback from the NIST evaluation team in order to collect an initial data set for development and testing. The second data set was collected at the Aberdeen Proving Ground under more realistic conditions to supplement our initial round of training data. Finally, we have data collected while the subject shadowed an experienced soldier as he traversed an obstacle course set up by the NIST evaluation team.

Figure 3 shows a graph of performance versus the number of rounds of boosting used to build the ensemble classifier. We learned the ensemble classifier with 100 rounds and used a leave-one-out cross validation scheme to evaluate performance. In this case, we learned a classifier based on 13 of the 14 sequences and evaluated it on the 14^{th} sequence. As you can see in the graph, performance does vary with

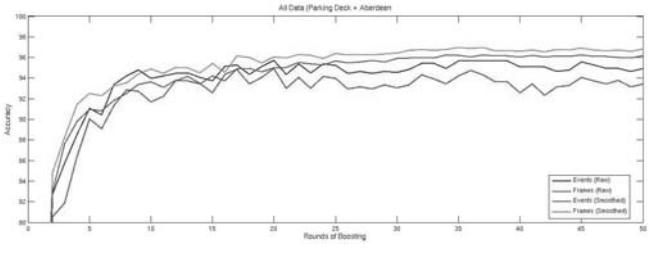


Fig. 3. Event and frame-based accuracy vs. number of rounds of boosting

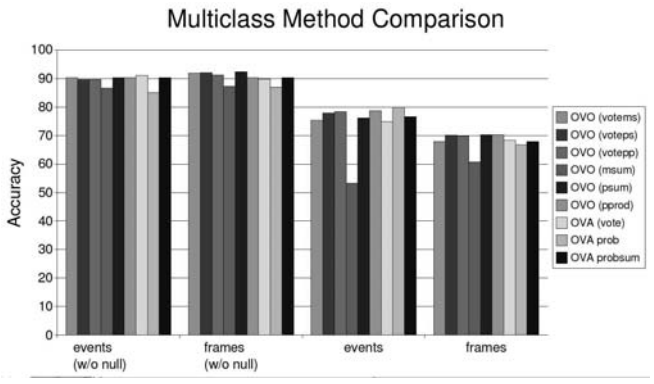


Fig. 4. Comparison of different binary classifier merge methods to handle multiclass discrimination

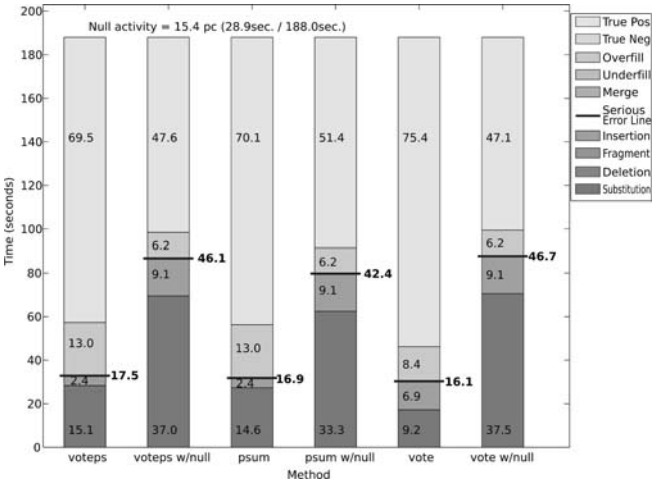


Fig. 5. Error division diagrams comparing three multiclass methods both with and without the null class. The bold line separates "serious errors" from boundary and thresholding errors.

additional rounds of boosting but basically levels off after 20 rounds. Thus we only used 20 rounds of boosting in subsequent experiments.

Although all of the features used by the activity recognition system are selected automatically, manual inspection reveals that they are often easily interpreted. For example, one of the primary features used to classify "shaking hands" is the frequency of the accelerometer mounted on the right wrist. Interestingly, the learning algorithm selected high frequency energy in the chest and thigh sensors as useful features for detecting the "driving" activity. We believe this is due to engine vibration detected by the sensor when the soldier was seated in the humvee. This cue highlights the strength of automatic feature selection during learning since high-frequency vibration is quite discriminative but might not have been obvious to a system engineer if features had to be manually specified.

The official performance metric established by the NIST evaluation team measured recognition performance during those times when one of the 14 identified soldier activities occurred. We call the unlabeled time the "null activity" to indicate that it constitutes a mix of behaviors related only by being outside the set of pre-identified activities. Importantly, the null activity may constitute a relatively large percentage of time and contain activity that is similar to the identified activities. It is therefore very difficult to model the null class or to detect its presence implicitly as low confidence classifications for the identified activities. Performance measures for our approach in both cases (*i.e.,* when the null activity is included and when it is ignored) are given in figure 4.

As discussed in Section IV-D, there are two common approaches for combining binary classifiers into a multiclass discrimination framework. Figure 4 displays the performance using each of the methods. The performance is fairly stable across the different methods, though the inferior accuracy achieved using the *msum* method validates the idea that while the margin does indicate class membership, its value is not generally commensurate across different classifiers. Interestingly, although the recognition performance is comparable, the average OVO training time was less than half that of the OVA methods (11.9 vs. 29.9 minutes without a null model and 23.9 vs. 47.3 minutes including the null activity). Thus, selecting an OVO method can either significantly reduce training time or allow the use of more sophisticated weak learners or additional rounds of boosting without increasing the run time for model learning.

We can visualize and compare the trade-offs between various event level errors and frame level errors using Error Division Diagrams (EDDs). EDDs allow us to view the nuances of the system's performance not completely described by the standard accuracy metric involving substitution, deletions, and insertions. For example, EDDs allow comparisons of errors caused by boundary discrepancies. These discrepancies include both underfilling (falling short of the boundary) and overflowing (surpassing the boundary). EDDs also simultaniously compare the fragmentation of an event into multiple smaller events and the merging of multiple events into a single event. The different error types and a more detailed explanation of EDDs can be found in Minnen *et al.* [13]. Figure 5 shows EDDs for three different multiclass methods. The EDDs show that the majority of the errors are substitution or insertion errors.

Finally, Figure 6 shows a color-coded activity timeline

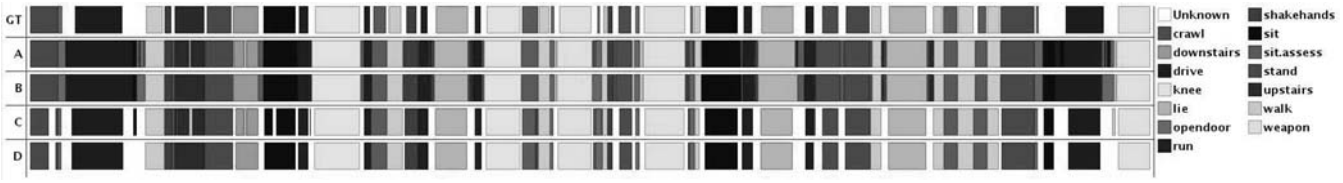


Fig. 6. Activity timeline showing the results of four system variations on one of the parking deck sequences. The top row (*GT*) shows the manually specified ground truth labels. System *A* was trained without any knowledge of the null class, while System *C* learned an explicit model of the null activity. Systems *B* and *D* are temporally smoothed versions of Systems *A* and *C*, respectively. In the smoothed systems, there is less fluctuation in the class labels. Also note that Systems *A* and *B* never detect the null activity, and so their performance would be quite low if the null class was included in the evaluation. Systems *C* and *D*, however, perform well over the entire sequence.

that provides a high-level overview of the results of four system variations. Each row, labeled *A*-*D*, represents the results of a different system. These results can be compared to the top row, which shows the manually specified ground truth activities. The diagram depicts the effects of two system choices: explicit modeling of the null activity (Systems *C* and *D*) and temporal smoothing (Systems *B* and *D*).

As expected, Systems *A* and *B* perform well over the identified activities but erroneously label all null activity as a known class. Many of the mistakes seem reasonable and potentially indicative of ground truth labeling inconsistencies. For example, walking before raising a weapon, crawling before and after lying, and opening a door between driving and walking are all quite plausible. Systems *C* and *D*, however, provide much better overall performance by explicitly modeling and labeling null activity. In particular, System *D*, which also includes temporal smoothing, performs very well and fixes several mistakes in System *C* by removing short insertion and deletion errors.

VI. Conclusion

The activity recognition component of the Soldier Assist System can accurately detect and recognize many low-level activities considered important by soldiers and other subject matter experts. By computing aggregate features over a window of sensor readings, we can transform a difficult temporal recognition problem into a simpler spatial modeling task. Using a boosting framework over 1D classifiers provides an efficient and robust method for selecting discriminating features. This is very important for an on-body, distributed sensor network in which some sensors may be uninformative or even distracting for detecting certain activities. Finally, we show that performance is relatively unaffected by the choice of multiclass aggregation method, though one-vs-one methods lead to a considerable reduction in training time.

VII. Acknowledgments

This research was supported and funded by DARPA's ASSIST project, contract number NBCH-C-05-0097.

References

[1] DARPA, "ASSIST BAA #04-38, http://www.darpa.mil/ipto/solicitations/open/04-38_PIP.htm," November 2006.

[2] J. Ward, P. Lukowicz, and G. Tröster, "Evaluating performance in continuous context recognition using event-driven error characterisation," in *Proceedings of LoCA 2006*, May 2006, pp. 239–255.

[3] N. Kern, S. Antifakos, B. Schiele, and A. Schwaninger, "A model for human interruptability: Experimental evaluation and automatic estimation from wearable sensors." in *ISWC*. IEEE Computer Society, 2004, pp. 158–165.

[4] M. Stäger, P. Lukowicz, N. Perera, T. von Büren, G. Tröster, and T. Starner, "Soundbutton: Design of a low power wearable audio classification system." in *ISWC*. IEEE Computer Society, 2003, pp. 12–17.

[5] P. Lukowicz, J. Ward, H. Junker, M. Stäger, G. Tröster, A. Atrash, and T. Starner, "Recognizing workshop activity using body worn microphones and accelerometers," in *Pervasive Computing*, 2004, pp. 18–22.

[6] T. Westeyn, K. Vadas, X. Bian, T. Starner, and G. D. Abowd, "Recognizing mimicked autistic self-stimulatory behaviors using hmms." in *Ninth IEEE Int. Symp. on Wearable Computers*, October 18-21 2005, pp. 164–169.

[7] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive*. Berlin Heidelberg: Springer-Verlag, 2004, pp. 1–17.

[8] J. Lester, T. Choudhury, N. Kern, G. Borriello, and B. Hannaford, "A hybrid discriminative/generative approach for modeling human activities," in *Nineteenth International Joint Conference on Artificial Intelligence*, July 30 - August 5 2005, pp. 766–772.

[9] A. Raj, A. Subramanya, D. Fox, and J. Bilmes, "Rao-blackwellized particle filters for recognizing activities and spatial context from wearable sensors," in *Experimental Robotics: The 10th International Symposium*, 2006.

[10] R. E. Schapire, "A brief introduction to boosting," in *IJCAI*, 1999, pp. 1401–1406.

[11] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Computer Vision and Pattern Recognition*, 2001.

[12] J. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," in *Advances in Large Margin Classifiers*. MIT Press, 1999.

[13] D. Minnen, T. Westeyn, T. Starner, J. Ward, and P. Lukowicz, "Performance metrics and evaluation issues for continuous activity recognition," in *Performance Metrics for Intelligent Systems*, 2006.

Probabilistic parsing of dietary activity events

Oliver Amft, Martin Kusserow and Gerhard Tröster
Wearable Computing Lab, ETH Zurich, Switzerland
email: amft@ife.ee.ethz.ch

Abstract—**Dietary behaviour is an important lifestyle aspect and directly related to long-term health. We present an approach to detect eating and drinking intake cycles from body-worn sensors. Information derived from the sensors are considered as abstract activity events and a sequence modelling is applied utilising probabilistic context-free grammars. Different grammar models are discussed and applied to dietary intake evaluation data. The detection performance for different foods and food categories is reported. We show that the approach is a feasible strategy to segment dietary intake cycles and identify the food category.**

Keywords—**Automatic dietary monitoring, nutrition, intake cycle, eating detection, PCFG, Earley-Stolcke parsing.**

I. Introduction

Nutrition is a key aspect of our everyday life and health. While pure over-consumption in time frames of months and years leads to the predominant overweight and obesity, many other forms of malnutrition exist. Often malnutrition is a confounding factor for developing chronic illnesses. Since nutrition is related to daily living behaviour, modifying eating behaviour requires changing lifestyle.

Besides caloric value, nutrition behaviour includes a variety of aspects such as duration and frequency of eating and drinking activities, rate of intake as well as the type of foodstuff itself. Information about these parameters on a daily basis provide insight into the dietary activities and can be integrated in lifestyle coaching, e.g. reminders to maintain a lunch duration of at least 15 minutes.

Our work aims at developing methods to monitor dietary behaviour automatically using wearable systems. In this paper we present an approach to infer eating and drinking activity as well as food categories from activity events derived in three on-body sensing domains.

A. Automatic dietary monitoring

We expect that by utilising wearable systems useful assistive systems for dietary monitoring are conceivable. Such systems could provide a rough estimate on the food consumption and could provide valuable insight into daily eating behaviour. This includes a rough estimation of food type, e.g. ratio of fluid and solid nutrient combined with the timing information, e.g. event schedule and meal durations over the day.

We target non-invasive wearable systems relying on information from the following three sensing domains: (1) the identification of characteristic arm and trunk movements associated with food intake using inertial sensors [1], (2) the analysis of food chewing sounds from an ear microphone [2] and (3) the detection of swallowing from collar-worn sensors [3]. These sensing domains are modelled as activity event sources by appropriate continuous pattern detectors. These events constitute the input for the event sequence detection presented in this work.

B. Decomposition of hierarchical activity

While many human actions may not be feasibly sensed and modelled as a whole, they can be described as a hierarchical activity process. Consequently, such an activity process is composed of separate sub-activities, often aligned in a sequential order. Given that a sufficient abstraction was chosen, patterns of identified sub-activities can be recognised from sensor data. An example for such an activity consisting of a sub-activity event sequence are dietary intake cycles. These cycles consist of movements to prepare a food piece and manoeuvre it to the mouth, e.g. using fork and knife, chewing the food with multiple closing and opening cycles of the jaw and eventually swallowing the food bolus. Usually several intake cycles are used to consume a food product or meal. The combination of these sub-activities in their correct order forms the superior activity *eating*.

Sequences of sub-activities that are linked to form a meaningful action suggest the analogy to linguistic terms, e.g. words (=sub-activity events) and sentences (=action, consisting of sub-activity event sequences). Following the example of an intake cycle described above, a syntax is given by the fact that foods may be chewed and swallowed only after they have been prepared and moved to the mouth. We hypothesise that sub-activities follow a grammatical structure and henceforth can be interpreted as computable language. Given that this hypothesis holds, the high-level segmentation of intake cycles can be achieved and moreover, structure parameter such as number of chews and food category estimates per intake cycle become available.

The detection of the event sequences, in linguistic terms the parsing of symbols, has to deal with the following main problems: (1) the input sequence may not follow the assumed language syntax in all situations and (2) the input sequence may be partially incorrectly detected by the event pattern detectors. Both problems violate the applied grammar and a standard language parser would simply give up. Obviously the applied parsing method and grammar has to cope with such situations, however accounting for the violations. As a solution a probabilistic context-free grammar (PCFG) parser is used in this work.

C. Probabilistic parsing of activity events

A grammar G can generally be described by $G = (T, NT, P, S)$. Here T is a set of terminal symbols, NT is a set of non-terminal helper symbols, P is a set of production rules of the grammar and S is the start symbol.

The prototype production rule of a context-free grammar is described in Eq. 1. These production rules require that the left hand side corresponds to a non-terminal symbol X that is expanded by the set of terminal and non-terminal symbols $(NT \cup T)^*$ at the right hand side when required. This concept of production rules permits the modelling of embedded symbol sequences by parsing from outside to inside instead of left to right.

$$X \rightarrow \lambda,\ with\ X \in NT,\ \lambda \in (NT \cup T)^* \qquad (1)$$

A context-free grammar is extended to a PCFG by assigning a probability P to each production rule. This principle is shown in Eq. 2.

$$X \rightarrow \lambda\ \ [P] \qquad (2)$$

Conceptually, this probability is conditional on the selection of the non-terminal symbol X for derivation. The aspect of "contextual freeness" is reflected by the independence of the production rules X_i in a complete PCFG, Eq. 3.

$$\forall i: \quad \sum_j P\left(X_i \rightarrow \lambda_j\right) = 1 \qquad (3)$$

While several problems can be tackled with this approach, we concentrate on the scoring task: we intend to estimate the probability that a symbol sequence was generated by a certain grammar. For this task J. Earley developed a parsing algorithm [4]. This algorithm was extended to probabilistic processing by Stolcke [5].

Further in this section, related works for activity sequence modelling and activity parsing are discussed. Section II describes our detection approach in the three on-body sensing domains and introduces the activity event parsing method. In Section III a experimental procedure is sketched to acquire and analyse evaluation data. Section IV reports the achieved performance of the event parsing approach. Finally Section V summarises the work and Section VI provides an outlook on future research.

D. Related works

Many attempts have been made to decompose activities into individual events of varying granularity and apply learning machines to identify the events individually. However the combined detection of the activity events sequences is favourable to reason about the superior activities. The methods applied at this level include Hidden Markov Models (HMMs), Bayesian networks, PCFGs and combinations thereof.

For HMMs different solutions have been proposed to model higher-level temporal structures including hierarchical HMMs and layered HMMs. Generally these HMM-based solutions require high training efforts, e.g. the availability of a large training corpus and extensive parameter search in order to tune the large amount of model parameters. Layered HMMs attempt to reduce this complexity by training layers independently [6]. Bayesian networks are by far the most flexible framework for reasoning and have been applied to the recognition of human activities, e.g. [7]. Moreover combinations with other reasoning approaches, such as PCFGs have been attempted, e.g. [8].

Research work relying on PCFGs for activity recognition have been presented in the domain of image recognition mainly, e.g. Moore and Essa [9], Ivanov and Bobick [10] and Yamamoto et al. [11]. Moore and Essa used the Earley-Stolcke parsing algorithm to detect activities in the card game Black Jack. The identification of player strategies were targeted. The authors proposed a complex error handling concept. Ivanov and Bobick demonstrated single recognition results for music conduction and activity surveillance at a parking lot. A simpler error handling was used in this work. Yamamoto et al. applied PCFGs to the Japanese tea ceremony and tracked the correct activity execution.

II. Sensing and detection principle

The sensing domains used for our approach in dietary behaviour monitoring are further detailed in this section. Moreover the event detection method using PCFGs is introduced and basic dietary behaviour models are presented.

A. On-body sensing domains

To analyse dietary behaviour, we evaluated three sensing domains that are obviously related to dietary intake activities and provide insight into the eating micro-structure. For each sensing domain different sensing modalities are used to detect activity events in continuous data. The following event types are derived:

1) Movement events from inertial sensors, e.g. gestures of the arms during drinking or eating with specific tools.
2) Chew events from an ear-worn microphone sensor.
3) Swallow events from neck muscle contraction (Electromyography electrodes) and a stethoscope microphone integrated into a sensor-collar, worn at the neck.

Fig. 1 illustrates the applied sensing modalities and their respective positioning. Inertial sensors have been attached to the lower and upper arms and the back, a microphone is worn at the ear and the sensor-collar at the neck. Data acquisition using this setup is further described in Section III.

Events for each sensing domain are discriminated into different categories: for motion events we distinguished gestures using fork and knife, using a spoon, drinking from a glass or bottle and simple hand-only gestures. A similar approach was taken for the chew and swallow events to discriminate the food texture and food bolus consistency respectively.

The search for individual event types is regarded as independent pattern detection problem. This detection was discussed in previous works [1]–[3]. For the remainder of

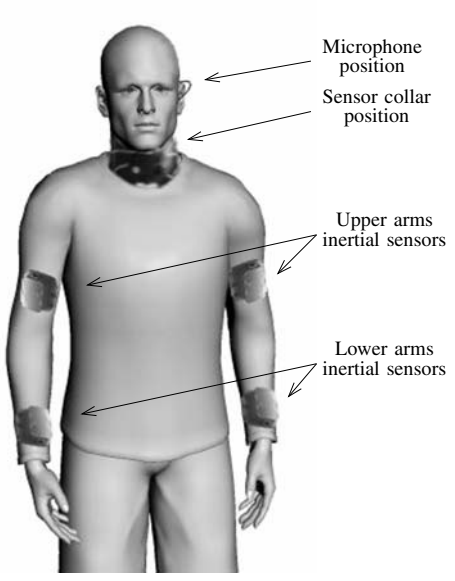


Fig. 1. Schematic sensor positioning at the body.

this paper a correct detection of these events was assumed in order to analyse the grammar modelling feasibility.

We refer to a sequence of the events containing motion E_M, chew E_C and swallow events E_S as an intake cycle with the number of occurrences N_M, N_C, N_S, Eq. 4.

$$E_{Cycle} = (E_M^{N_M}, E_C^{N_C}, E_S^{N_S}) \quad with\ N_M = 1, N_C \geq 1, N_S \geq 1 \tag{4}$$

We restricted our intake cycle model to consist of one initial movement event only, $N_M = 1$. This is useful in order to segment individual intake cycles and analyse the natural processing of these single "bites" in isolation. The food type estimation is facilitated by the abstraction, since the food item will not change during a cycle. Certain cycles event types may not be available in all intake cycles, e.g. there are usually no chew events for drinking activities.

B. Earley-Stolcke parsing algorithm

The aim of our event sequences analysis is to derive an event level segmentation that resembles the intake cycles and classify the food type in parallel. For this goal events are interpreted as terminal symbols of an Earley-Stolcke parser.

The parser processes symbols of the input stream sequentially by applying the defined PCFG. While processing, the parser keeps track of all possible derivations of the symbol sequence. With every new input symbol the number of possible derivations is increased as new alternatives appear or decreased when multiple solutions are resolved. For this purpose the parser keeps a set of states for each position in the input stream. A state is described by the notation shown in Eq. 5.

$$i\ :\ {}_kX \rightarrow \lambda.\mu\ [\alpha, \gamma],\ with\ \ \lambda, \mu \in (NT \cup T)^* \tag{5}$$

The index i, ($i \geq 0$) and the dot "." refers to the current position in the input stream, index k, ($k \leq i$) indicates the begin of a sub-string given by the non-terminal X. The variables α and γ refer to forward- and inner probability respectively. The forward probability $\alpha_i({}_kX \rightarrow \lambda.\mu)$ is the summarised probability of all paths of length i that end at ${}_kX \rightarrow \lambda.\mu$. The inner probability γ is the summarised probability of all paths of length $i-k$ starting at $k :_k X \rightarrow .\lambda\mu$ and ending at $i :_k X \rightarrow \lambda.\mu$.

The Earley-Stolcke parsing algorithm consists of the states *Prediction*, *Scanning* and *Completion*. A brief summary of the algorithm operation is provided below, a more in-depth elaboration can be found in [5]. For every input symbol the states are processed and the probabilities α, γ are updated. In the prediction step all non-terminals are expanded as long as non-terminal symbols are available. In the scanning step a new input symbol is read and matched to a terminal. When a match was found, the position index i is incremented. All expansions that are not matched in this step are omitted from the current set of states. The completion step is the finalisation of the non-terminal derivation. All fully expanded non-terminals are added to the set of states. Prediction and completion steps can have loops due to cyclic expansions. These are resolved by the parsing concepts *left corner relation* and *unit production relation* [5].

A vital aspect for the PCFG-application in activity parsing is the handling of errors in the symbol sequence. Many related works expect that the input has a low error rate, e.g. Yamamoto et al. [11] and Moore et al. [9]. However the latter work provides a full framework to cope with multiple insertions, substitutions and deletions by hypothetically continuing parsing paths. It can be assumed that the complexity of the parsing algorithm increases significantly due to this complex error handling. Ivanov and Bobick [10] utilised grammar modifications and multivalued input vectors to address insertion and substitution errors. In this paper, we followed the approach of Ivanov and Bobick.

C. Parsing of dietary activities

Since the relevant activities are very different regarding their activity event structure, e.g. eating and drinking consist of different events, each type of intake cycle was modelled with a dedicated PCFG. For each such PCFG we were interested in solving the scoring problem and determine, how well the current event sequence match the specific grammar.

Fig. 2 shows the parsing concept and the parser instantiations used in the evaluation. Events generated from one or multiple sensor pattern detectors are parsed by N parsers, where N equals the number of different PCFGs. Using a dedicated parsing instance for each grammar, provides a scalable solution that can tolerate multiple differently structured nutrition activities. Eventually all parsing results are combined to a final decision based on the best matching sequence indicated by the parser forward probability.

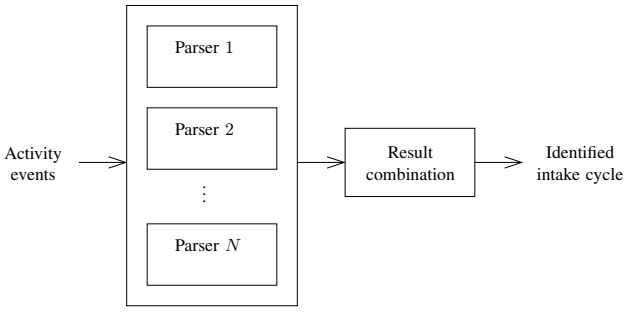


Fig. 2. Concept of parsing using different PCFGs.

D. Probabilistic models for eating and drinking

For eating cycles we exploited the freedom of the parsing concept by modelling different foods and food categories by an individual grammar. Eq. 6 describes a generic rule for eating based on the intake cycle specification provided above (Eq. 4). Every cycle is described by an initiating movement symbol followed by non-terminal chew and swallow symbols[1]. The non-terminal symbols are expanded to a sequence of chew and swallow terminals based on the received chew and swallow events. The model is restricted to swallow terminals from chewed foods only, hence $E_{S,Chewed}$.

$$\begin{aligned} FOOD &\rightarrow E_M\ CHEW\ SWALLOW\ [1.0] \\ CHEW &\rightarrow E_C\ [0.1] \\ &\mid E_C\ CHEW\ [0.9] \\ SWALLOW &\rightarrow E_{S,Chewed}\ [0.5] \\ &\mid E_{S,Chewed}\ SWALLOW\ [0.5] \end{aligned} \tag{6}$$

The eating grammar shown above accounts for the number of occurrences of chew and swallow events (N_C, N_S) by the probabilities associated to each production rule. Typical food intake cycles contain multiple chew events, described by a high probability of one chew event followed by further chew events (0.9), while the derivation of a single chew event indicates the end of a chew sequence. These probabilities have been chosen manually. Swallow events are modelled in this grammar as finalisation of the intake cycle occurring as one or multiple events.

Contrary to the eating cycle grammar, drinking requires less event types. Here, chewing is not involved in the cycle. Similar to the eating grammar, multiple swallowing events may occur. The movement is restricted to the drink gesture, $E_{M,Drink}$. For drinking a swallow terminal $E_{S,Fluid}$ (fluid bolus item) is required. The grammar is shown in Eq. 7.

$$\begin{aligned} DRINK &\rightarrow E_{M,Drink}\ SWALLOW\ [1.0] \\ SWALLOW &\rightarrow E_{S,Fluid}\ [0.5] \\ &\mid E_{S,Fluid}\ SWALLOW\ [0.5] \end{aligned} \tag{7}$$

[1]Following the nomenclature in related works, non-terminal symbols are printed in upper case letters.

These grammar rules are applied and further discussed in the evaluation described in Section IV.

III. Evaluation procedure

A. Evaluation data set

In order to analyse the performance of our parsing approach we recorded a data set of eating and drinking activities using the sensors as described in Section II above. The sensors were positioned as shown in Fig. 1. While the test user was eating different food products an observer annotated the recordings online. In a post-processing step this annotation was manually refined by reviewing the waveforms to obtain sections that reflect the boundaries of the described event types. The annotation information for every event was then used as input for the parsing evaluation.

Tab. I summarises the recorded foods. In total 3799 events were recorded and annotated from eating and drinking of one test user consuming 11 foodstuffs in 162 intake cycles.

TABLE I
Description of the recorded food data set.

Food item	Description
Drink	Drinking from a glass. Drinking does not involve chewing.
Cornbar, Biscuit, Peanuts, Potato chips	Eating the food items using the hand to bring the food to the mouth. The foods are of dry texture during chewing.
Lasagne	Eating lasagne using fork and knife. The cooked food is of soft texture. The swallow bolus is of variable consistency.
Lettuce	Eating using fork and knife. The food is of wet texture. The swallow bolus is of variable consistency.
Bread	Eating bread using the hand to bring the food to the mouth. The food is of soft texture during chewing.
Soup	Eating a soup from a bowl using a spoon. This food item provides no chewing events.
Apple	Eating an apple using the hand to bring the food to the mouth. The food is of wet texture. The swallow bolus is of variable consistency.
Yoghurt	Eating from a mug using a spoon. This food item provides no chewing events. The swallow bolus is of variable consistency.

B. Performance analysis

Since there is no automatic algorithm training step involved in the applied parsing approach, we did not partition the data into training and testing set. Instead, we used the entire data set to test the parsing and the grammars.

To analyse performance, we utilised the metrics *Precision* and *Recall*, commonly used for algorithm evaluation in information retrieval applications. These metrics are derived as follows:

$$Recall = \frac{\text{Recognised intake cycles}}{\text{Relevant intake cycles}} \quad (8)$$

$$Precision = \frac{\text{Recognised intake cycles}}{\text{Retrieved intake cycles}} \quad (9)$$

Relevant intake cycles corresponds to the annotated number of actually conducted intake cycle instances. *Retrieved intake cycles* represents the number of cycles returned by the parsing algorithm. Finally, *Recognised intake cycles* refers to the correctly returned number of cycles. Both metrics are defined for the value range $[0, 1]$. A recall value close to one indicates a good sensitivity of a method to return relevant intake cycles, while a precision value close to one indicates that the method does return few insertion errors.

IV. Results

In the first analysis step we aimed at exploring the sequential properties of the intake cycles and feasibility of the grammar models. For this purpose we applied the simple eating and drinking grammars as defined in Eq. 6, 7 to the individual foods. For movement and swallow events the abstract event instances were used as described in Tab. I. For chew events we assumed in this step that every food can be modelled by a food-specific symbol. Fig. 3 shows the achieved parsing performances using the metrics precision and recall.

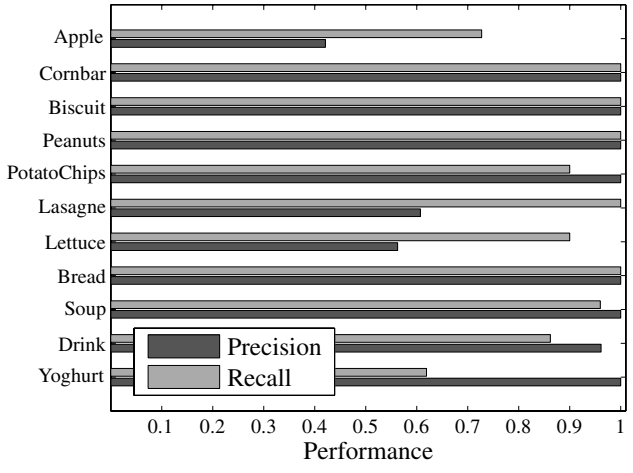


Fig. 3. Performance chart for the intake cycle detection of the simple food grammars shown in Eq. 6, 7. For precision and recall, best performance is found towards high values.

These performance values show that the simple model is not a feasible solution for all food types. For several food items many insertion errors were retrieved, indicated by the low precision value at ∼0.6 or below. The used grammar requires a strict sequence of chew and swallow events while many foods contain alternating chew and swallow events, e.g. apple and lasagne. These food items contain more fluid than dry foods, e.g. peanuts, that lead to increased swallow rates. Moreover, the intermediate swallows are an additional food-specific feature that could be explored.

In the following step the food model was refined for non-dry foods to incorporate the typical intermediate swallowing activity. Eq. 10 shows the updated grammar.

$$\begin{aligned} FOOD &\rightarrow E_M\ MAST^{+}\ [1.0] \\ MAST &\rightarrow CHEW\ SWALLOW\ CHEW [0.2] \\ &\quad |\ CHEW\ SWALLOW\ MAST\ [0.8] \\ CHEW &\rightarrow E_C\ [0.1] \\ &\quad |\ E_C\ CHEW\ [0.9] \\ SWALLOW &\rightarrow E_{S,Chewed}\ [0.5] \\ &\quad |\ E_{S,Chewed}\ SWALLOW\ [0.5] \end{aligned} \quad (10)$$

$^{+}$*MAST.* = Mastication

Using this model, we repeated the analysis of step 1. Fig. 4 shows the parsing performances for this analysis using precision and recall. A clear improvement for food items containing fluid was achieved, e.g. for lasagne the precision increased from ∼0.6 to 1 indicating that no insertion errors were retrieved when parsing the data set with this grammar.

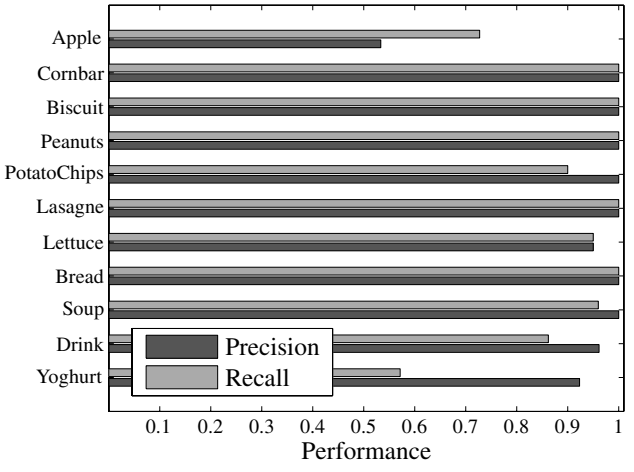


Fig. 4. Performance chart for the intake cycle detection of the refined non-dry food grammars shown in Eq. 10. For precision and recall, best performance is found towards high values.

In a further step we analysed the performances of intake cycles grouped into food categories. We defined the groups based on the similarity of food texture, movement and swallow type. The group "Dry" contained bar, biscuit, peanuts and chips. Yoghurt and soup were grouped into "Spoon" since no difference in the activity event sequences was expected for the food items: movement and swallows are similar for both foods and both are not chewed. Fig. 5 shows the precision and recall results for the "Dry" and "Spoon" food groups in comparison with the remaining foods.

The very good performance for the group "Dry" indicates that all food items in this group are similar in their event

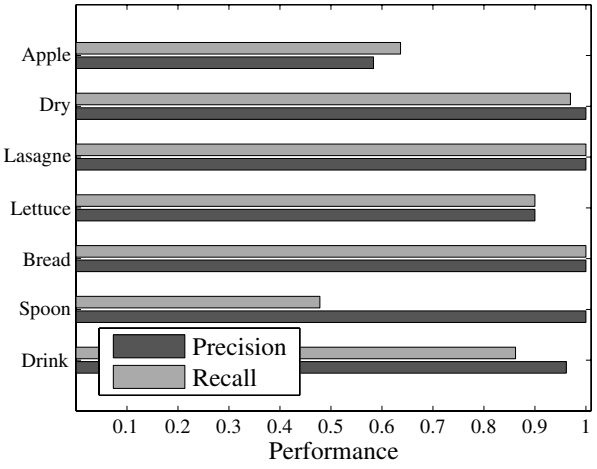


Fig. 5. Performance chart for detection of intake cycles of "Dry" and "Spoon" groups in comparison with the remaining foods. For precision and recall, best performance is found towards high values.

sequence structure. However the new "Spoon" group suffered from high deletion errors, indicated by the low recall. This is mainly due to the weak matching of the grammar on yoghurt intake, since yoghurt consisted of highly fluctuating number of swallows.

V. Conclusion

We presented an approach to detect dietary intake cycles from on-body activity event sequences. The event sequences were modelled using probabilistic grammars. The approach was evaluated with sensor data annotations and the algorithm performance was derived for detecting intake cycles.

We analysed different variants of the grammars, starting with simple and strict sequencing rules. The analysis however showed, that these rules were not capable to catch intermediate swallows in certain food cycles. Hence, we adapted the grammars to better accommodate the observed sequences. With the refined rules the detection rates of non-dry foods improved clearly. This analysis addressed the basic intake cycle modelling on individual foods. In order to handle multiple food items a further abstraction from individual foodstuffs was needed. For this purpose the food items were grouped by similar texture and intake characteristics. We analysed the feasibility of using one grammar for the detection in each food group.

Overall detection rates of $\sim$80% were achieved for precision and recall in the food category analysis. This indicates that the intake cycle modelling using probabilistic grammars is a feasible solution. The evaluation was performed with event data acquired from one subject only. However we expect that the approach is scalable to multiple users since no automatic model training was used that would fit the model to the event data. Hence the grammar models applied in this work were rather tuned for food features than for the test user.

VI. Future work

We plan to further analyse the PCFG approach for detecting dietary intake activities. While the general feasibility of probabilistic grammars was shown in this work, the interconnection with event detection methods will be investigated. Moreover we intend to evaluate the method on further food items and test users.

Acknowledgement

The authors are grateful to Dr. Yuri Ivanov for providing his Earley-Stolcke parser implementation for reference. The parser was re-implemented for this project. The work was supported by the Swiss State Secretariat for Education and Research (SER).

References

[1] O. Amft, H. Junker, and G. Tröster, "Detection of eating and drinking arm gestures using inertial body-worn sensors," in *ISWC2005: IEEE Proceedings of the Ninth International Symposium on Wearable Computers.*, October 2005, pp. 160–163.

[2] O. Amft, M. Stäger, P. Lukowicz, and G. Tröster, "Analysis of chewing sounds for dietary monitoring," in *UbiComp 2005: Proceedings of the 7th International Conference on Ubiquitous Computing.*, M. Beigl, S. Intille, and J. Rekimoto, Eds. Springer: Lecture Notes in Computer Science, September 2005, pp. 56–72.

[3] O. Amft and G. Tröster, "Methods for detection and classification of normal swallowing from muscle activation and sound," in *Proceedings of the First International Conference on Pervasive Computing Technologies for Healthcare*. ICST, November 2006.

[4] J. C. Earley, "An efficient context-free parsing algorithm," *Communications of the ACM*, vol. 13, no. 2, pp. 94–102, Feb. 1970.

[5] A. Stolcke, "An efficient probabilistic context-free parsing algorithm that computes prefix probabilities," *Computational Linguistics*, vol. 21, no. 2, pp. 165–201, 1995.

[6] N. Oliver, A. Garg, and E. Horvitz, "Layered representations for learning and inferring office activity from multiple sensory channels," *Computer Vision and Image Understanding*, vol. 96, no. 2, pp. 163–180, Nov. 2004, special issue on event detection in video.

[7] Y. Du, F. Chen, W. Xu, and Y. Li, "Recognizing interaction activities using dynamic bayesian network," in *ICPR 2006: Proceedings of the 18th International Conference on Pattern Recognition*, vol. 1, Aug 2006, pp. 618–621.

[8] K. Kitani, Y. Sato, and A. Sugimoto, "Deleted interpolation using a hierarchical bayesian grammar network for recognizing human activity," in *Proceedings of the 2nd Joint IEEE International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance*, Oct 2005, pp. 239–246.

[9] D. Moore and I. Essa, "Recognizing multitasked activities from video using stochastic context-free grammar," in *AAAI 2002: Proceedings of the American Association for Artificial Intelligence*, 2002, pp. 770–776.

[10] Y. Ivanov and A. Bobick, "Recognition of visual activities and interactions by stochastic parsing," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 852–872, Aug 2000.

[11] M. Yamamoto, H. Mitomi, F. Fujiwara, and T. Sato, "Bayesian classification of task-oriented actions based on stochastic context-free grammar," in *Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition*. IEEE, April 2006, pp. 317 – 323.

7th Session
Communication II

An All-Detailed Architecture of a RF Wireless Transmitter

Mourad MELLITI[1] Salem HASNAOUI[1]
[1]Department of Computer and Communication Technologies - National School of Engineering of Tunis (ENIT)
AL-Manar University, BP37, 1002, Tunis - TUNISIA
email : mourad.melliti@laposte.net

***Abstract*–It is widely known that (FHSS) frequency-hopped spread spectrum is the mostly used spread technique which uses multilevel modulation and cover both infrastructure field and Ad-Hoc networks. This paper presents the architecture and implementation of a FHSS RF radio transmitter. The complete radio transmitter chain was implemented in which a direct digital frequency synthesizer (DDFS) was used for fast frequency hopping and FSK modulation. The radio transmits data in the 900 MHz ISM band at a rate of up to 160 kbps using frequency shift keying (FSK) modulation with frequency hopping at a rate of up to 80 khops/sec. The transmitter is also designed to take part of a detailed transceiver chain, to be modular and reusable in other wireless communication systems.**

***Keywords*—Transmitter, FHSS, DDFS, FSK, Reusable.**

I. INTRODUCTION

Needs for wireless communications grow in domains of our daily life and particularly in industrial domain. The near future shows new industrial equipments, connected together, within wireless networks that permit to have an increased versatility in the management of disparate equipments. As a consequence, researchers are developing technologies that will eventually lead to a fully integrated single chip transmitter based on CMOS technology. A noncoherent FSK based FHSS system has been chosen due to its power saving potential and reduced hardware complexity. The original proposed system is capable of sending data at rates up to 160 kbps, and hopping over the entire 902-928 MHz ISM band at 20 khops/sec [1]. The transmitter combines a direct digital frequency synthesizer (DDFS), a digital-to-analog converter (ADC), a double quadrature upconverter, an integrated oscillator, and power amplifier with variable output.

II. TRANSMITTER ARCHITECTURE

The transmitter performs modulation, upconversion, and power amplification. Two major transmitter architectures exist [4]. The first one as shown in figure 1. (a) is called direct-conversion because the carrier frequency is equal to the local oscillator frequency. There are several drawbacks of this simple architecture. First, the quadrature upconverter cannot suppress the LO signal completely at its output. Often referred to as 'LO leakage', this residue LO appears at the transmitted spectrum and acts as interference to other receivers. Second, since the shielding from the power amplifier to the local oscillator is finite, the strong signal of the power amplifier will corrupt the local oscillator spectrum. To avoid this phenomenon (referred to as 'LO pulling'), a two-step transmitter architecture is proposed. As shown in figure 1. (b), the baseband signal is upconverted twice such that the power amplifier output spectrum is far from the frequencies of the local oscillator. The first BPF suppresses the harmonics of the IF signal and the second BPF removes the unwanted sideband due to the simple upconversion.

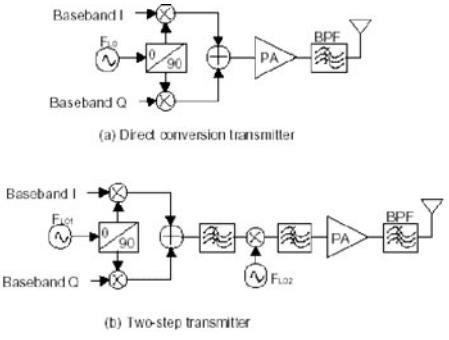


Fig. 1. Transmitter architecture

We notice that in both the direct-conversion and the two-step transmitters, quadrature upconversion is used. This is also referred to as the Weaver's single sideband (SSB) modulation technique, it partially suppresses the unwanted image such that the requirement for the BPF following the upconversion mixer is relaxed. A comparison of the simple mixing and the SSB modulation is shown in figure 2.

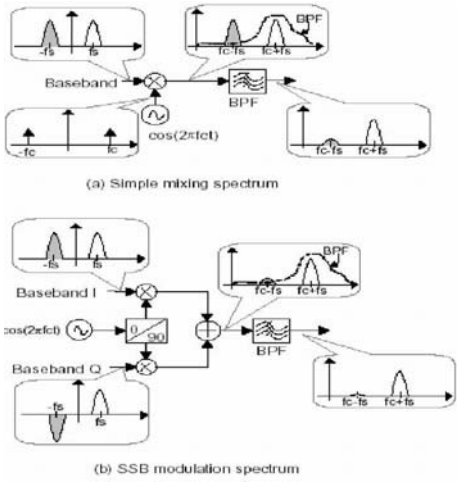


Fig. 2. Simple mixing and quadrature mixing spectra

Taking into consideration the issue of image suppression, hardware complexity and the requirement for fast frequency hopping [3], a variation of the direct-conversion transmitter architecture is adopted. Instead of modulating the carrier with the baseband signals, a low IF baseband signal is used to modulate a local oscillator with frequency outside the transmit band. To transmit the frequency-hopped FSK signals in the 902-928 MHz ISM band, a fixed frequency local oscillator at 880 MHz is modulated with a hopping FSK signal ranging from 22 to 48 MHz. The chosen structure avoids the problem of LO pulling in a conventional direct-conversion transmitter. Furthermore, the LO leakage issue is alleviated because now the residue LO is out of the transmit band and thus could be further suppressed by the BPF following the upconverter. More importantly, the residue image due to imperfect quadrature mixing is at least 44 MHz away from the signal and could be further suppressed by using a relatively high Q BPF.

This transmitter structure puts a more stringent requirement on the baseband signal generation. Because now the baseband circuitry needs to generate accurate quadrature hopping FSK signals from 22 MHz to 48 MHz. To tackle this problem, a high-speed direct digital frequency synthesizer (DDFS) is used. The operation principle is described in the next section.

II. 1. HOPPING SYNTHESIZER

We know that the final transmitted tone has a frequency witch is determined by the following equation:

$$F_{tone} = F_{LO} + F_{hop} + F_{data} \qquad (1)$$

where F_{LO} is the fixed local oscillator frequency of the transmitter, F_{hop} is the hopping frequency determined by the PN sequence, and F_{data} is one of the M data bearing frequencies of an ordinary M-FSK system. One way to accomplish the modulation process is to generate a hopping baseband signal with frequency equal to the sum of F_{hop} and F_{data}, and use this to modulate a fixed local oscillator with a frequency of F_{LO}. The other method is to generate a hopping local oscillator with frequency of $F_{LO}+F_{hop}$, then modulate it with data dependent tone F_{data}. The second method requires a fast hopping synthesizer while the first method only requires fast hopping baseband signal generation. Figure 3 shows a traditional phase locked loop (PLL) based synthesizer, which uses programmable frequency dividers in the loop to generate different frequencies [8]. Though it has some advantages such as wide tuning range and good spectral purity, the fundamental problem of the analog PLL-based synthesizer is its low settling speed and limited frequency resolution [4]. Although the resolution of an analog PLL can be improved by means of a fractional-N synthesizer, the hopping speed of an analog synthesizer is limited by the settling time of the loop, which is inversely proportional to the loop bandwidth.

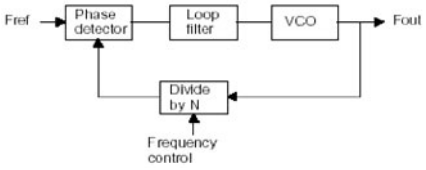


Fig. 3. Analog PLL based frequency synthesizer

To provide fast frequency hopping, a direct digital frequency synthesizer (DDFS) combined with a digital-to-analog converter (DAC) is an attractive alternative. The DDFS allows fast and precise manipulation of its output frequency under full digital control. It also enables very high resolution in the incremental selection of output frequency. The basic functional block diagram and signal flow of a generic DDFS is shown in figure 4.

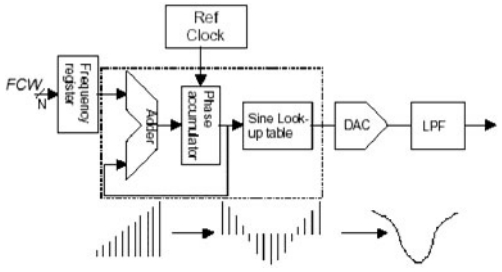


Fig. 4. Direct Digital Frequency Synthesizer (DDFS)

The phase accumulator is a variable-modulus counter that increments the number stored in it each time it receives a clock pulse. When the counter overflows it wraps around, making the phase accumulator's output continuous. The frequency tuning word sets the modulus of the counter, which effectively determines the size of the increment that gets added to the value in the phase accumulator on the next clock pulse. The output of the phase accumulator is used to address a ROM in which the samples of a sine wave are stored. Thus the larger the added increment, the faster the accumulator overflows and wraps around, which results in a higher output frequency. The relationship of the output frequency F_{out}, reference clock F_{CLK} and the frequency tuning word F_{CW} (which stands for frequency control word) is determined by the formula:

$$F_{out} = \frac{FCW \quad F_{CLK}}{2^N} \qquad (2)$$

where N is the number of bits in the accumulator.

The digital sine wave output of the DDFS drives a DAC that reconstructs the sine wave in analog form. A generic output spectrum of a DAC is shown in figure 5. If a reference clock of 100 MHz is used to generate an output sine wave of 20 MHz, the output spectrum contains the fundamental plus aliased signals that occur at multiples of the reference clock frequency ± the desired output frequency [6].

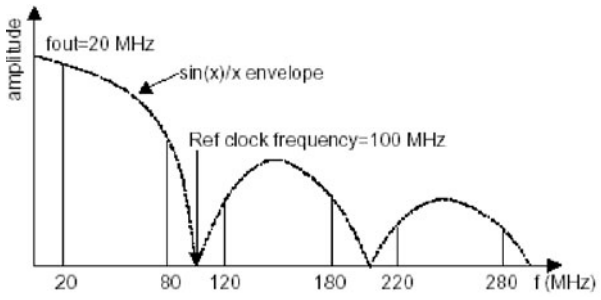


Fig. 5. DAC Output spectrum

As shown in the figure, the aliases have relatively strong energy and follow a sin(x)/x envelope roll-off. To suppress the images, an anti-alias LPF is needed after the DAC. To relax the requirement of the anti-alias LPF, the highest frequency to be generated should be lower than 40% of the reference clock frequency.

The specification of the DAC, LPF and the DDFS are chosen such that 50 dB spectral purity can be achieved with the hopping synthesizer.

III. TRANSMITTER IMPLEMENTATION

The block diagram of the transmitter's structure is shown in figure 6. The DDFS-DAC is used to generate a tone whose frequency is the sum of the proper hopping frequency and the FSK modulated frequency. After the LPF, it is upconverted by a quadrature mixer to the 902-928 MHz ISM band. After a BPF, a pre-amplifier amplifies the signal. Then a digitally controlled variable attenuator attenuates the signal before passing it to the final power amplifier and antenna.

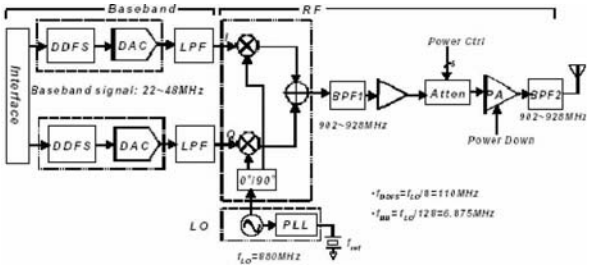


Fig. 6. Block diagram of the transmitter

The following sections describe some of the key features and the detailed design of the transmitter.

III. 1. LOCAL OSCILLATOR AND CLOCKING STRATEGY

Since all the frequencies are generated from the Local Oscillator (LO) and especially when the spectral purity of the tones generated by the DDFS is directly related to the phase noise performance of the LO. So it is essential that we specify the noise performance according to the system requirement.

III. 1. 1. LOCAL OSCILLATOR SPECIFICATION

The noise of the LO might limit the possible receiver performance if it is not properly designed. All oscillators have sideband noise and it is important to determine the amount of sideband noise that is acceptable. A generic power spectrum of an LO is given in figure 7.

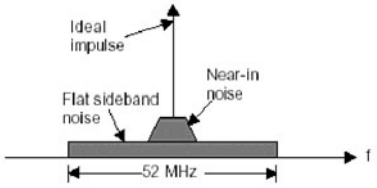


Fig. 7. Generic spectrum of an LO

The spectrums consists of an impulse at the desired frequency of operation, a flat noise sideband across the whole band of interest, and near in noise close to the impulse. The noise sideband beyond ± 26 MHz will not contribute noise because the received signal is band limited. Closer to the impulse, the noise floor is amplified by the oscillator and causes the flat noise to rise toward the carrier which gives rise to the near-in noise [5].

III. 1. 2. DESIGN OF THE LOCAL OSCILLATOR

The block diagram of the local oscillator sub-system is illustrated in figure 8.

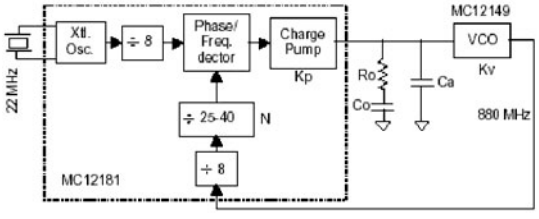


Fig. 8 Local oscilator sub-system

The local oscillators of the transmitter is designed using Motorola's low power VCO chip MC12149 and a PLL chip MC12181. The two chips plus a loop filter and a crystal reference form a complete PLL sub-system. With the crystal reference oscillator frequency chosen to be 22 MHz, and the programmable divide ratio N equal to 40, an output frequency of 880 MHz is obtained for the transmitter. Ro and Co plus the charge pump form a basic second order loop filter. An optimum loop bandwidth of 17.5 kHz is chosen to obtain both minimum phase noise and maximum reference sideband suppression. For a second order system with a damping factor of ζ=1, the loop natural frequency is 17.5kHz/2.5 or 7 kHz. Then the loop transfer function is given by

$$T(s) = \frac{R_0 C_0 s + 1}{(N_t C_0 / K_p K_v) s^2 + R_0 C_0 s + 1} \qquad (3)$$

$$T(s) = \frac{(2 \quad / \quad _0)s \quad 1}{(1/ \quad _0^2)s^2 + (2 \quad / \quad _0)s + 1} \qquad (4)$$

where N_t is the total divide ratio which is 8×40=320, K_v is

the VCO gain which is measured to be 23MHz/Volt, K_p is the phase detector/charge pump gain which is 2.3 mA/radian. R_o and C_o are the major loop filter component as shown in figure 8. Equating the coefficients and given the fact the natural frequency ω_0 is 7 kHz, we can calculate the value for R_o and C_o. The capacitor C_a is used to provide additional reference sideband suppression and its value should be less than one tenth of C_o.

III. 1. 3. CLOCKING STRATEGY

There are three major clocks used in the system. They are the DDFS clock, baseband circuitry clock, and the symbol clock. To simplify the generation of the clocks, all of them are derived from the LO. The use of one crystal oscillator source reduces harmonic coupling and other distortions possible with a multi-oscillator system. Another advantage of this is that all clocks are synchronized. This is proven to be essential in the design of the DDFS interface. The clocking scheme is illustrated in figure 9.

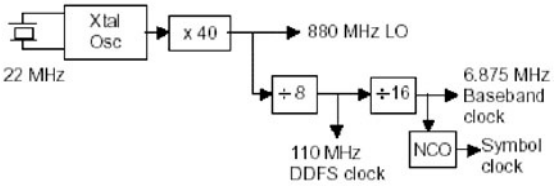


Fig. 9. Transmitter clocking strategy

The crystal used in the PLL is 22 MHz. Then all frequencies of the transmitter are derived from the LO, whose frequency is 40 times that of the crystal oscillator. The DDFS clock frequency is chosen to be 110 MHz such that the highest frequency it can generate (48 MHz) covers the upper limit of the hopping bandwidth. Also since the reference clock of the DDFS is the divided-by-eight version of the LO, its phase noise is also eight times samller than that of the LO [5]. Thus a high quality agile frequency synthesizer can be built upon the DDFS. The baseband frequency is also synchronized with the DDFS clock frequency so fast frequency and phase update of the DDFS is possible. In the baseband circuit, an NCO further generates the symbol clock which ranges from 2 kBaud to 80 kBaud.

III. 2. HOPPING SYNTHESIZER

As described in the last sections, the hopping synthesizer is based on a DDFS/DAC combination and a LPF is needed after the DDFS/DAC to suppress the aliased images due to the sampling operation. Because the highest frequency generated from the DDFS/DAC is 48 MHz, the nearest image is at 110-48=62 MHz. A LPF with a sharp transition edge is needed. A seventh order elliptic LC type LPF filter with a cutoff frequency of 50 MHz is designed for this purpose.

Although the passband ripple is large in the simulation due to use of ideal inductors, the measured frequency response of the implemented LPF (figure 10) has a relatively flat passband while providing more than 50 dB of suppression for frequencies higher than 62 MHz.

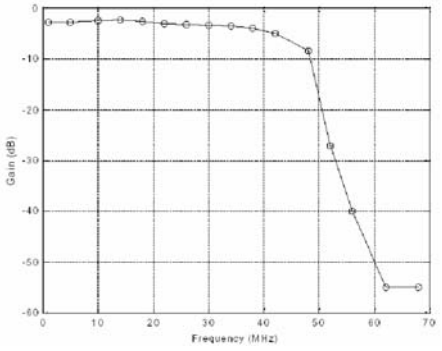


Fig. 10. Measured frequency response of the DAC LPF

III. 3. QUADRATURE UPCONVERTER

Two DDFS/DAC chips generate quadrature hopping signals from 22 MHz to 48 MHz. These signals are used to modulate a quadrature upconverter mixer to produce hopping signals in the 902-926 MHz ISM band. A RF Micro Device RF2402 UHF quadrature modulator chip is used here (Figure 11). The chip contains differential amplifiers for the baseband inputs, a 90° hybrid phase splitter for the LO, a limiting LO amplifier, two balanced mixers, a combining amplifier and an output RF amplifier which drives a 50Ω load [9].

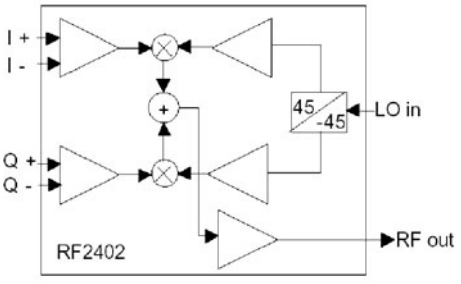


Fig. 11. Block diagram of RF2402 quadrature upconverter

This chip was chosen rather than the traditional ring diode modulator [7] for two reasons. One reason is that better sideband suppression can be achieved with quadrature modulation. Due to its monolithic construction, the chip exhibits excellent amplitude balances and high phase accuracy. Sideband suppression is 25 dB without any adjustment, and carrier suppression can be adjusted to be more than 40 dB. Another advantage is the use of active components for mixing and the inclusion of an on-chip RF amplifier which compensates for losses in the modulator [9].

Careful design and matching of the I-Q branches are important to ensure adequate sideband and carrier suppression. In the design, several adjusting points are inserted in the I-Q branches. These adjusting points include the current of the DAC on the DDFS/DAC chip, the gain of the I-Q buffers, and the biasing voltage levels

for the upconverter. Care is also taken to match the components for the two LPFs. With such careful design, an image suppression of 35 dB and carrier suppression of more than 40 dB are achieved. By adding a BPF after the mixer, we can further reduce the sideband and carrier by another 5-10 dB.

III. 4. POWER AMPLIFIER

The upconverted signal is passed to a digitally controlled programmable attenuator and then to the final power amplifier (PA). The attenuator chip, RF2420 from RF Micro Devices is a multistage GaAs monolithic programmable attenuator. The attenuation is set by five bits of digital data and covers a 44 dB range in 2 dB steps. This attenuator in conjunction with the PA realizes a digitally controlled RF power transmitter system. A RF2103P medium power linear amplifier is chosen as the final PA. The modulation used is phase continuous FSK which can be efficiently amplified by non-linear amplifiers because of its constant envelope nature. The PA chip delivers a maximum power of 24 dBm to a 50Ω load via a matching network at its output. The details of the matching network can be found in the schematic in Appendix.

III. 5. DETAILED ARCHITECTURE

The detailed architecture of the transmitter is shown in figure 12.

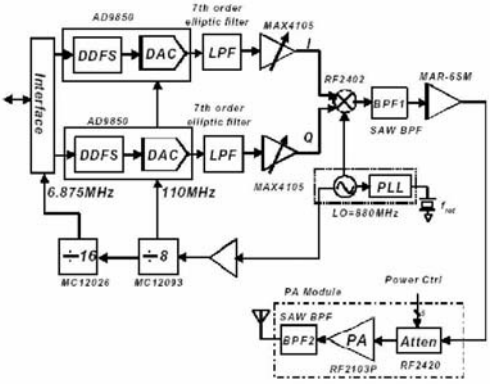


Fig. 12 Detailed transmitter architecture

A PLL sub-system generates the LO at 880 MHz. The LO is divided by 8 to generate the DDFS/DAC clock at 110 MHz. Then the 110 MHz clock is further divided down by 16 to 6.875 MHz and used to clock the baseband circuits. Two DDFS/DAC chips are controlled by the baseband frequency control word (FCW) to synthesize quadrature hopping tones from 22 to 44 MHz. Two 7th-order elliptic LPFs after the DDFS/DAC suppress the aliased images. Two wide band buffers (MAX4105) bring the I-Q signals close to 5 Volts peak-to-peak which is required to drive the mixer at the highest efficiency. The LO is also buffered to about 1 dBm to drive the mixer. The amplified I-Q signals are mixed with the LO and brought to the 902-928 MHz ISM band. A SAW filter centered in the ISM band rejects the residue sideband and carrier while also providing good 50Ω matching [2].

A monolithic MMIC amplifier MAR-6SM from Mini Circuit is used to compensate the loss due to the SAW filter. Then the signal is attenuated according to transmit power control word and finally amplified by the PA. There is also another BPF after the PA to reduce spurious response out of the ISM band. The power amplifier drives a planar inverted F antenna. The inverted F antenna shows good 50Ω matching and transmission characteristics in the 902-928 MHz ISM band [9]. The measured results of the transmitter are discussed in the next sections.

IV. TRANSMITTER MEASUREMENT RESULTS

In this section, the output spectrums of the transmitter are presented. Notice that the spectrum shown here is taken before the power amplifier. Figure 13 is the binary FSK spectrum of the transmitter sending pseudo-random data.

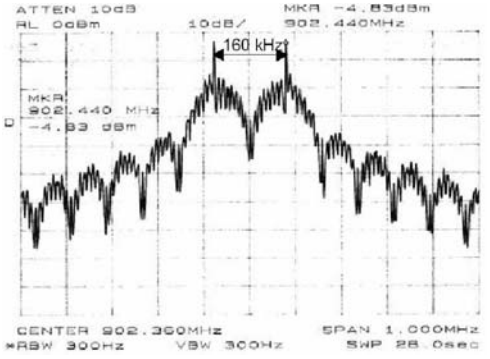


Fig. 13. Binary FSK spectrum

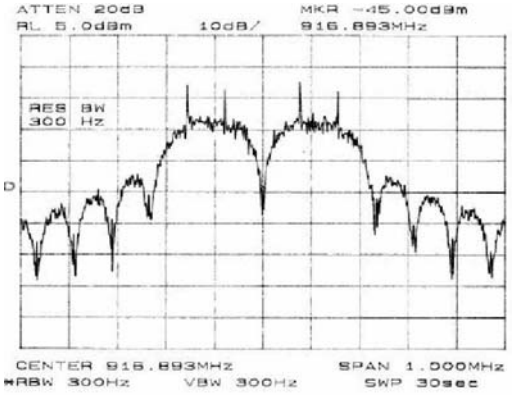


Fig. 14. 4FSK spectrum

Figure 14 is the output spectrum of the transmitter sending pseudo-random data with 4FSK modulation. The symbol rate is 74 kBaud in this test. The two plots clearly demonstrate the superiority of using DDFS/DAC to synthesize FSK signals. Due to the phase continuity of the DDFS synthesizer, the side lobes of the FSK spectrum are much lower. Altogether these two plots verify the basic FSK modulation functionality of the transmitter.

The next figure illustrates the hopping spectrum of the transmitter.

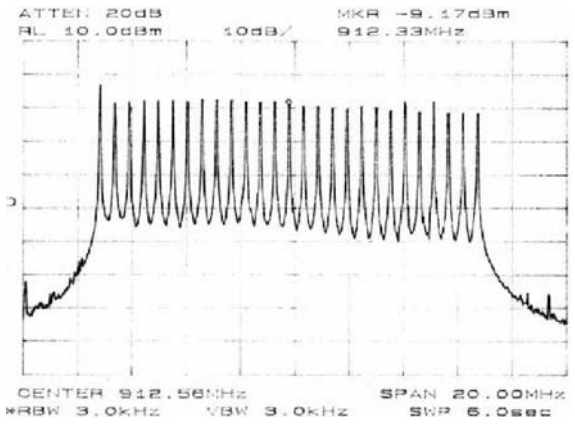


Figure 15 Transmitter hopping spectrum

In this test, the transmitter is configured to hop every 4 symbols (20 khops/sec) within a total bandwidth of 14 MHz starting from 902 MHz. A total of 27 discrete hopping channels are clearly shown in the plot. However, the amplitude of the hopping tones are not the same. The fluctuation is mainly due to the frequency dependence of the DAC/LPF section and the limited bandwidth of the buffers following the DAC/LPF. Also the hopping rate is relatively slow (20 khop/sec) thus the averaging effect of the spectrum analyzer makes it less accurate in reading the instantaneous power of each hopping tone.

V. CONCLUSION

In this paper, the design and implementation of a FHSS transmitter has been described. Discussions on access techniques, radio architecture, design methodology, detailed circuit structure and measurement results are also presented.

A direct digital frequency synthesizer was used in the transmitter for fast frequency hopping and FSK modulation. The radio transmits data in the 900 MHz ISM band at a rate of up to 160 kbps using 4FSK modulation with frequency hopping at the rate of up to 80 khops/sec. The transmitter is designed to be modular and reusable in other wireless communication systems. If modulation schemes other than FSK are utilized, the basic structure of the transmitter can still be used with minimal hardware modifications.

REFERENCES

[1] J. Min, A. Rofougaran, H. Samueli, A. A. Abidi. An all-CMOS architecture for a low-power frequency-hopped 900 MHz spread spectrum transceiver. *Proc. of IEEE Custom Integrated Circuits Conference*, pp. 379-382, 1994.

[2] P. R. Gray and R. G. Meyer. Future directions in silicon ICs for RF. *Proc. of IEEE Custom Integrated Circuits Conference* pp. 83-90, May 1995.

[3] A. A. Abidi. CMOS wireless transceivers: the new wave. *IEEE Communications Magazine.* pp. 119-124, Aug. 1999.

[4] B. Razavi. RF Microelectronics. *Prentice Hall PTR, Upper Saddle River*. 87 NJ 1998.

[5] D. B. Leeson. A simple model of feedback oscillator noise spectrum. *Proc. of the IEEE*. pp. 329-330, Feb 1966.

[6] A. Rofougaran, G. Chang, J. J. Rael, J. Y.-C. Chang, M. Rofougaran, P. Chang, M. Djafari, J. S. Min, E. W. Roth, A. A. Abidi, , H. Samueli. A single-chip 900 MHz spread-spectrum wireless transceiver in 1µm CMOS-part II: receiver design. *IEEE J. Solid-State Circuits* Vol. 33, No. 4, pp. 535-547, April 1998.

[7] J. J. Rael. A 915 MHz CMOS frequency synthesizer. *Master thesis*, Dept. of Electrical Engineering. UCLA 1995.

[8] Sumit Mohan. Implementation of a versatile baseband FSK processor for fast frequency-hopped spread-spectrum communications *Master thesis* Department of Electrical Engineering, UCLA, 2000.

[9] H. Zou. Design and Implementation of a Radio Transceiver for a Fast Frequency Hopped Spread Spectrum testbed. *Master thesis*, Dept. of Electrical Engineering University of California, 2000.

Towards Plug-and-Play Interoperability for Wireless Personal Telehealth Systems

L. Schmitt, T. Falck, F. Wartena, and D. Simons

Philips Research Europe

Abstract— **This survey paper serves as an introduction to the challenges and needs related to wireless personal telehealth systems and provides an overview of ongoing activities in industry and various communication standards, aiming to enable plug-and-play interoperability. Specifically, we address a recently founded industry consortium, the Continua Health Alliance, and ongoing standardization efforts within the family of ISO 11073/IEEE 1073 standards, the Bluetooth SIG, and the ZigBee Alliance.**

Keywords— **Personal Telehealth, Wireless Communication, Interoperability, Standardization, Continua Health Alliance**

I. Introduction

In 2004 the costs for healthcare in the United States have increased above 16 percent of the gross national product (GNP) [1]. This correlates with the fact of an aging population and an increasing number of chronically ill people. While in 2000, the proportion of the population in the United States aged ≥65 was 12.4 percent, this is expected to increase to 19.6 percent in the year 2030 [2]. Considering diabetes as an example, the proportion of persons of age ≥65 with a chronic diabetes condition is approximately one in five (18.7%) [2]. Even assuming that this ratio remains constant in the future and does not get worse, the expected overall increase in percentage of persons with chronic diabetes conditions due to the aging of society is alarming. This development is not restricted to the US but it is a worldwide problem that particularly developed countries are facing. Hence, a major challenge in healthcare is to improve the quality of care for an increasing number of patients using limited financial and human resources.

Personal telehealth systems, including remote patient monitoring and management, are increasingly recognized as having the potential to help overcoming that challenge. Per definition, in a personal telehealth system the caregiver is geographically separated from the care consumer with the care plan being individually tailored to the patient's needs. This patient-centered concept of bringing the care from the hospital or the doctors office to the patient at home results in cost-reduction and improved quality of care. Being able to more frequently observe the patient's state of health by performing remote measurements of the patient's vital signs enables optimizing the patient's medication and treatment. This results in longer independent living for older patients and lower mortality rates. Through increased frequency of daily automated, but personalized, patient intervention, the care providers can manage a broader range of chronic disease patients, improving efficiency. However, the biggest opportunity for reduction in costs is not in lower costs for nurse visits but rather in a reduced need for high-cost chronic care and hospitalization.

In addition to the care being provided in a remote and personalized way, an important factor for enabling the success of future telehealth systems is to make the last hop to the patient wireless. By introducing wireless technology cumbersome cables can be eliminated, enabling greater physical mobility and making the system more unobtrusive and ubiquitous for the patient. For the wireless link at the patient side the technologies being considered range from enabling a simple cable-replacement to allowing real networking of vital sign measurement devices, as for example in the context of body sensor networks (BSN).

An example for a sophisticated system solution incorporating the aspects mentioned above is Motiva [3], a TV-based remote patient monitoring platform from Philips. The platform comprises daily, personalized patient interactions, delivered via broadband connection to the home television. Patients receive reminders and messages, educational videos, and feedback on their vital signs comprising weight, blood pressure, and blood glucose levels, based on a tailored care plan defined by their caregiver at enrollment.

However, besides the benefits of wireless personal telehealth systems mentioned above, there exist also several challenges and issues still unresolved or with tentative answers so far. Just to name some examples: An inherent problem of telehealth is the lack of context information of the raw medical data that is transmitted to the remote caregiver, not knowing for example the patient's posture or current activity level when measuring blood pressure. In contrast to lack of information, the amount of data being transmitted has to be carefully chosen in order to prevent cognitive overload of caregivers or providers. Besides having an overall safe system, it should be reliable. This is especially an issue if wireless connections are employed, as these are prone to interference and variations in quality of the propagation channel. Furthermore, the system should be secure in terms of being protected against unwanted access, and assure privacy of the patient's data. From a legal per-

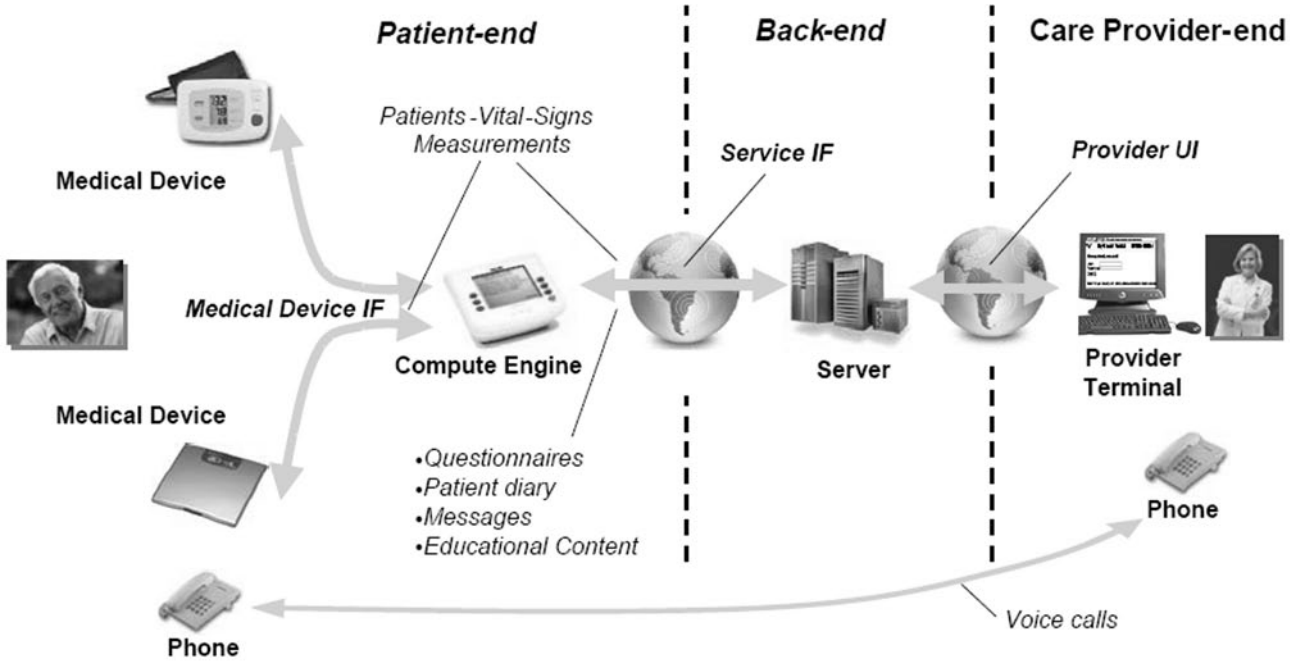


Figure 1: Model of a typical personal telehealth system for chronic disease management

spective there is the question of liability for remote health systems. Another important topic relates to the business aspects for the system implementers: Who are the payers of telehealth systems and what are the strategies for reimbursement? Besides health insurance companies, a new class of payers might be the patients themselves, who actively manage their health by consuming health care rather than being provided with it.

All these issues have to be considered in designing solutions to the overall complex problem of personal telehealth care. However, a lot of isolated and proprietary solutions exist today. Indeed there is an enormous conglomeration of personal health devices and services lacking interoperability, and hence preventing that the issues mentioned above are solved in a unified and standardized way. Thus, it is exactly the approach of enabling plug-and-play interoperability and connectivity within the context of wireless personal telehealth, which is necessary for the success of future telehealth systems.

In the remainder of this paper we elaborate in more detail the problem of interoperability in the personal telehealth domain (Section II) and give an overview of current ongoing standardization activities. After presenting an overview of a recently formed industry consortium addressing the overall problem of interoperability in Section III, we address current standardization activities regarding interoperability on upper-layers and the application layer of the OSI stack in Section IV and V, respectively. Finally concluding remarks are given in Section VI.

II. The interoperability problem

An example of a personal telehealth system for chronic disease management is shown in Figure 1. Typically it consists of three parts. A patient-end with personal health devices in the patient's home or on the patient's body, a back-end part for storing the data and a care provider-end, where the personal health consultant has access to the patient's data.

The devices at home are medical devices (e.g. blood pressure meters, weight scales, pulse rate meters, or thermometers) for measuring the patient's vital signs and usually some kind of compute engine (e.g. a PC or a mobile phone) for gathering the data and providing a user interface for the patient. The back-end part contains a server component, which is connected via a network, e.g. the Internet, to the patient-end part. Similarly, at the care provider-end access to the patient's data on the server is enabled via some network, which could also be e.g. the Internet.

Obviously there is a need for intercommunication among the various components within the system. Focusing on the patient-end part there are several alternative wireless tech-

nologies, as e.g. IEEE 802.15.1 (Bluetooth), IEEE 802.15.4, or IEEE 802.11 (WLAN), to establish a wireless link between medical devices and compute engines. Indeed this has been accomplished with a number of personal health devices. However, the standardization process with regard to medical device interoperability lacks behind technical possibilities. Virtually all of these solutions are specialized applications with proprietary interfaces unique to the two devices being linked. That means if for example the blood pressure meter from company *X* can communicate with the same companies compute engine, almost certainly a blood pressure meter from vendor *Y* can not interoperate with the compute engine from company *X*. Ensuring compliance on the physical layer between two devices does not ensure interoperability, as there are many different way to transmit the same information over a physical layer interface. In order to ensure plug-and-play interoperability between multi-vendor devices, the devices must be able to *understand* the format and the content of the messages they communicate to each other.

Hence, on the patient-end side the problem of device interoperability has to be solved on three principle levels: On lower-layers a standardized transport technology enabling basic connectivity has to be developed. On upper-layers application profiles have to be developed, which define what capabilities of the transport technology have to be used to best support the application requirements. Finally on application level standardized data models and formats have to be developed, which represent an abstract unique mapping of the real world entities. While a significant amount of problems on the lower layers has been solved already and mature standards are available, more work at levels closer to the application is needed. For a discussion of medical device interoperability refer also to [4].

Concerning the back-end part at some point in the system the data has to be translated into HL7 (Health Level 7) [5], which is usually employed by archival repositories.

III. The Continua Health Alliance

In June 2006, 22 industry-leading technology and health companies joined forces and formed an open nonprofit industry alliance [6]. Since its inception the alliance has grown continuously, now comprising 36 promoter companies (see Figure 2 for an overview of the Continua promoters), and another 32 contributor companies.

According to its mission *"establish an ecosystem of interoperable personal health systems that empower people & organizations to better manage their health and wellness"* the alliance plans to select connectivity standards and set out guidelines for interoperability. The objective is not to develop new standards, but to leverage existing ones as much as possible, and to close recognized interoperability gaps by means of interoperability guidelines.

Figure 2: The 36 promoter companies of the Continua Health Alliance [6] in November 2006.

Regarding the architectural structure of personal telehealth systems as indicated in Figure 1, Continua addresses the whole range from the medical device at the patient's home to the back-end services by defining interoperable interfaces. Currently, several wired and wireless standards are under investigation for selection to establish connectivity at the various interfaces. For the connection of medical devices to the system these include Bluetooth, USB, ZigBee, Wi-Fi, amongst others. For connections regarding the in-home network the list includes wireless and wired Ethernet, and power line communications. Last but not least, regarding the connectivity from the patient's home to the back-end services some candidates amongst others are Cable, DSL, Cellular (e.g. GPRS or CDMA), WiMax, and POTS[1]. For a very good online resource of wireless technologies please refer to [7].

Besides technical aspects, an objective of the alliance is to establish a certification program with a consumer recognizable logo for the devices. Furthermore, the alliance plans to collaborate with government regulatory agencies regarding consistent policies for the use of hi-tech personal health devices at home, and to develop new ways for reimbursement of personal health systems. These different areas of

[1] Plain Old Telephone Service

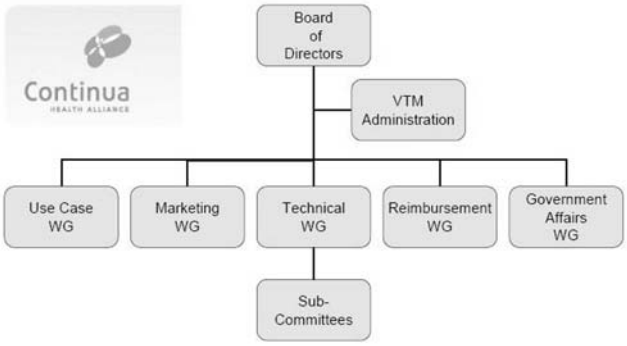


Figure 3: The Continua Health Alliance organizational structure.

involvement are also reflected by the organizational structure of the alliance, as shown in Figure 3.

Regarding the scope of application domains addressed, Continua does not focus exclusively on the chronic disease management domain, but also extends to elderly monitoring, i.e. monitoring the health and healthcare needs of aging people, and proactive health and fitness.

According to the timeline of Continua [8] the first version of the certification program development is expected to be completed at the end of 2007, such that the first certified products might enter the market in 2008. While the first version products are expected to cover the most basic devices and use cases in the envisioned application domains, extended versions are planned in the following years.

IV. Data models and formats

As pointed out in Section II, to solve the interoperability problem on application level it is necessary that devices speak a common language by means of a common nomenclature, data types, message syntax, and encoding rules. Many national and international organizations work on standards that enable upper-layer medical information exchange. The most important standards include DICOM [9], HL7 [5], and the ISO 11073/IEEE 1073 [10] family of standards - often also referred to as *Medical Information Bus* (MIB) or x73 standards. DICOM is rather a standard for transmitting medical imaging data, including also handling, storing and printing, and HL7 is a comprehensive set of standards for the exchange of healthcare information between computer applications. Whereas the ISO 11073/IEEE 1073 standard is a family of standards intended to enable medical devices to interconnect and interoperate with other medical devices. See Table 1 for an overview of some important parts of the x73 standard.

The standard is based on an object-oriented system management paradigm. An object oriented data model, the domain information model (DIM), defined in ISO 1173-10201, is used to specify objects, attributes, attribute groups, event reports, and communication services, that may be used to communicate device data and to configure medical devices and functionalities. The standardized nomenclature (ISO 11073-10101) comprises a set of numeric codes that identify every item that is communicated between systems. Related to the general domain information model, there exist device specializations for several medical devices (see Table 1), which provide guidelines for how the DIM should be constrained for application to specific devices. Application profiles according to the 2yyzz-series provide specific sets of restrictions for the use of the object model and service model tailored for a class of communication needs. They are independent of specific device types or specific lower communication layers. For example, the Polling Mode application profile (ISO 11073-20201) defines a context free polling mode to access the device medical data and is especially intended for small-scale devices. For an overview of the x73 documents and the x73 concepts for medical device communication refer also to [11].

The x73 standards for point-of-care medical device communication are mainly designed for acute monitoring and treatment applications in a particular diagnostic, bed or treatment area in the hospital domain [11]. Besides general requirements like patient and user safety of medical devices, minimal user interaction and unambiguous association, the key objectives for clinical domain applications addressed by the standard are real-time plug-and-play interoperability and frequent network reconfiguration. According to the employment of bedside devices some attention has also been paid to the reduction of implementation complexity and computational burden at the devices. For example, the message overhead is moderate and the encoding and parsing of protocol data units (PDUs) is very efficient due to the concept of *canned* messages (message templates can be filled in memory in which only the actual updated values must be copied [11]).

However, these design objectives align only partly with the requirements for personal telehealth systems, where especially sensor and battery powered devices demand for very low computational complexity and low power consumption. For wireless devices the latter requirement not only implies to minimize transmit power, but also to reduce transmission time by means of minimizing protocol overhead. On the other hand, in personal telehealthcare the network configuration and user association is rather static than dynamic and there is no distinct requirement for real-time streaming and real-time alarms today, as the setting is usually not acute.

In view of the rising activities in the personal telehealth care domain and as a consequence of the diverging require-

Table 1: The ISO 11073-xyyzz series of standards (status: s = published standard, d = draft, new = new project authorization request).

OSI Layer	*Part*	*Title / Content*	*Stat.*
	00000	Health Informatics – Point-of-care medical device communication – Framework and overview	D
	00103	Health Informatics – Personal health device communication – Technical report – Overview	new
7	1yyzz	Data & Information Definitions	
	10101	Health Informatics – Point-of-care medical device communication – Nomenclature	S
	10201	… - Domain information model (DIM)	S
	10300	… - Device specialization – Framework and overview	D
	10301	… - Infusion device	D
	10302	… - Vital signs monitor	D
	10303	… - Ventialtor	D
	10304	… - Pulse oximeter	D
	10305	… - Defibrillator	D
	10306	… - ECG monitoring	D
	10307	… - Blood pressure	D
	10308	… - Temperature	D
	10309	… - Airway meter	D
	10310	… - Cardiac output	D
	10311	… - Airway gas analyzer	D
	10312	… - Hemodynamic calculator	D
	10313	… - Pulmonary calculator	D
	10314	… - Respirator	D
	10315	… - Weighting scale	D
	10316	… - Dialysis device	D
	10400	Health Informatics – Personal health device communication – Device specialization – Common framework	new
	10404	… - Pulse oximeter	new
	10406	… - Heart rate monitor	new
	10407	… - Blood pressure monitor	new
	10408	… - Thermometer	new
	10415	… - Weighting scale	new
	10417	… - Glucose meter	new
7-5	2yyzz	Application profiles	
	20101	Health Informatics – Point-of-care medical device communication – Application profiles – Base standard	S
	20102	… - MIB elements	D
	20201	… - Polling mode profile	D
	20202	… - Baseline profile	D
	20601	Health Informatics – Personal health device communication – Application profile – Optimized exchange protocol	new
4-1	3yyzz	Transport and physical profiles	
1	4yyzz	Physical layer interface	
3	5yyzz	Internetworking support	
4	6yyzz	Application gateways	
	9yyzz	Related concepts	

ments mentioned above several new project authorization requests have been submitted recently to the IEEE standards association and the new ISO/IEEE 11073 *Personal Health Data* (PHD) working group has been established. The standards that this group is going to specify are the parts indicated as "new" in Table 1. They include specializations of 6 personal health devices and a new application profile aimed to address the needs and requirements of personal health devices mentioned above. Related to this activity it is also likely that the DIM will be extended to include additional items being unique to the personal health domain as e.g. patient context information describing the posture of a patient when measuring blood pressure. It is also expected, that the list of devices in the 104zz-series is further extended in future.

It should be noted that the issue of upper layer interoperability as addressed by the x73 standards is an integral part of the agenda of the Continua alliance, which makes the results of the new ISO/IEEE 11073 PHD working group likely to be adopted later on by Continua.

V. Wireless transports

Among the many wireless short-range communication standard activities worldwide, the IEEE 802.11 (WLAN) and the IEEE 802.15 (WPAN) family of standards have the largest impact on wireless today. Although providing lower data rates than IEEE 802.11, it is mainly the IEEE 802.15 family that fits best the requirements of small-scale personal health devices in terms of low power consumption and low complexity. After the activities in IEEE 802.15.3 (high-rate WPAN) have been abandoned the two major WPAN technologies today are IEEE 802.15.1 (better known as Bluetooth[2] [12]) and IEEE 802.15.4 [13] in combination with ZigBee [14] building on top of it, with IEEE 802.15.4 defining only the physical and MAC layer of the OSI stack. Operating in the license-free worldwide 2.4GHz ISM (Industrial, Scientific and Medical) band, IEEE 802.15.4 operates additionally in two less-crowded sub-GHz bands. See Table 2 for more details and an overview of other important system parameters of Bluetooth and ZigBee[3].

Let us briefly compare both technologies regarding personal telehealth applications, which are mainly characterized by episodic measurements. The data rate of ZigBee with a maximum rate of 250 Kbit/s is less than the theoretical maximum rate of 3 Mbit/s for Bluetooth, however still fulfilling the requirements of virtually all personal telehealth applications.

Both technologies use spread-spectrum (SS) techniques for increased interference robustness. Due to the direct-sequence approach of ZigBee in one of 16 non-overlapping channels it is possible to run non-interfering networks si-

[2] The IEEE adopted the V1.1 version of the Bluetooth specification.

[3] From now we refer to the combination of IEEE 802.15.4 and ZigBee simply as ZigBee.

Table 2: Some Bluetooth and ZigBee System parameters

	Bluetooth	ZigBee
Frequency Band	2.4 GHz	2.4 GHz (868 / 915 MHz bands)
Channel access	Slotted TDMA (FHSS)	CSMA/CA (DSSS)
# simultaneous networks	~ 10 (PER > 10 %)	16
Data Rate Per Channel	1-3 Mbit/s (theoretical)	250 kbit/s
Network Topology	Star	Star, P2P, Cluster Tree, Mesh
Multihop Capability	No	Yes
Network Size	8	65,535
Power consumption	low	very low

multaneously. According to the frequency-hopping (FH) approach comprising all of the 79 channels this is not strictly possible with Bluetooth as indicated in the table.

Although the transceiver physical characteristics are very similar in terms of transmit power, the power consumption for ZigBee is considered to be less than for Bluetooth in personal telehealth applications. This is mainly due to the FHSS approach of Bluetooth, where synchronization to the FH-sequence of the master (e.g. the compute engine) is required by a medical device. When entering a power saving mode either a timer running is required to stay in sync or synchronization is required when reconnecting to the network which can takes in average little less than one second. However, for both methods the resulting additional power consumption, and hence the effect on battery lifetime, is significant if frequent episodic measurements are employed. In contrast, ZigBee devices do not require synchronization with the network as they use a carrier sense multiple access (CSMA) approach.

Another main difference of both technologies is the supported network topology. While Bluetooth can be considered mainly as a single-hop wireless cable-replacement technology ZigBee supports a variety of different network technologies including most general mesh networking with multi-hop routing. A good comparison of both technologies is also given in [15], though with the focus on body-worn sensor networks. To summarize, due to their different technological characteristics it is expected that ZigBee and Bluetooth will take on complementary roles in the broad domain of personal telehealth.

An important advantage of Bluetooth, which makes it a prime candidate for wireless telehealth applications today, is its standard maturity and availability for end user products. The latest version of Bluetooth [12] is version 2.0 with the next release being expected early 2007. Furthermore, end of 2006 the Bluetooth SIG claimed that already more than 1 billion Bluetooth devices have been shipped so far.

However, both Bluetooth and ZigBee basically cover only the lower OSI layers up to the transport layer. In order to establish interoperability from the physical layer to the application layer for personal health devices, they can be used as a transport under the data model and format specified in the upper layers by x73. There are in principle two different ways to achieve this. Either certain transport profiles for the various wireless transport technologies are defined in the 11073-3yyzz series of standards, or certain application profiles are defined on top of the wireless transports. The latter option is currently pursued for Bluetooth. In May 2006 the Bluetooth SIG announced the formation of a *Medical Devices* working group. This group initially comprised 19 industry companies and aims at defining an application profile for personal health and fitness devices. As the group desires to leverage existing standards where possible, it can be expected that it will adopt the data models and format of the x73 PHD working group. Also, the group's profile itself might be adopted by the Continua alliance.

It is expected that the formation of a similar working group will be pursued in the ZigBee alliance as well. The ZigBee Application Framework group has already issued a call for participation in a *Personal / Home Health Care* study group in September 2006.

VI. Conclusions

Personal telehealth systems have the potential to cope with today's major healthcare challenge of improving the quality of care for an increasing number of chronically ill patients using limited financial and human resources. Today, there exist a lot of technologies and technology standards, but still being isolated solutions to the overall complex problem. However, there is significant activity in industry and in standardization organizations aiming at enabling real plug-and-play multi-vendor interoperability in the personal telehealth ecosystem. An industry consortium as the Continua Health Alliance is a step towards the right direction. However, one of the next steps should also be building a large-scale real-world demonstration that shows the realization of the vision.

Acknowledgment

We would like to express our gratitude to the *Bundesministerium für Wirtschaft und Technologie (BMWI)* for partially funding this work under contract *01MT305*.

References

1. Centers for Medicare and Medicaid Services at http://www.cms.hhs.gov/apps/media/press/release.asp?Counter=1750
2. Centers for Disease Control and Prevention. Public health and aging: trends in aging—United States and worldwide. *MMWR* 2003; 52(06):101-106.
3. D. Simons, T. Egami, and J. Perry, *Remote Patient Monitoring Solution,* in Advances in Healthcare Technology (Eds.: G. Spekowius and T. Wendler), Springer, 2006.
4. R. Schrenker, and T. Cooper, „Building the Foundation for Medical Device Plug-and-Play Interoperability," Medical Electronics Manufacturing, Spring 2001.
5. HL7 – Health Level 7 at http://www.hl7.org/
6. The Continua Health Alliance at http://www.continuaalliance.org
7. Palowirless Wireless Resource Center. Palowireless Pty Ltd. at http://www.palowireless.com/
8. http://www.continuaalliance.org/news_events/white_papers/
9. DICOM – Digital Imaging and Communications in Medicine at http://medical.nema.org/
10. IEEE 1073, Standard for Medical Device Communications at http://www.1073.org/
11. IEEE 1073, Standard for Medical Device Communication – Part 00000: Framework and overview.
12. Bluetooth Special Interest Group, Specification of the Bluetooth System, Version 2.0 + EDR , November 2004.
13. IEEE Std. 802.15.4™-2003, Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Low-Rate Wireless Personal Area Networks (LR-WPANs).
14. ZigBee Alliance, ZigBee Specification, Version 1.0, December 2004.
15. J. Espina, T. Falck, and O. Mühlens, *Network Topologies, Communication Protocols, and Standards*, in Body Sensor Networks (Ed.: G.-Z. Yang), Springer, London, 2006.

Address of the corresponding author:

Author: Lars Schmitt
Institute: Philips Research Europe
Street: Weisshausstrasse 2
City: 52066 Aachen
Country: Germany
Email: lars.schmitt@philips.com

On-Body Measurements and Characterization of Wireless Communication Channel for Arm and Torso of Human

E. Reusens[1], W. Joseph[1], G. Vermeeren[1] and L. Martens[1]

[1] Department of Information Technology, Ghent University / IBBT, Gaston Crommenlaan 8 box 201, B-9050 Ghent, Belgium
E-mail: {Elisabeth.Reusens, Wout.Joseph, Gunter.Vermeeren, Luc.Martens}@intec.Ugent.be

Abstract—**This paper discusses the propagation channel between two half-wavelength dipoles placed near a human body. Different parts of the body are investigated separately. Statistical properties of the wireless on-body channel have been investigated. Path loss parameters and time domain channel characteristics are extracted from the measurement data. Path loss models for the arm and torso have been derived. A comparison with a path loss model near a flat, homogeneous medium has been made.**

Keywords— **Propagation channel, human body, wireless body area network (WBAN), path loss, delay spread**

I. Introduction

A Wireless Body Area Network (WBAN) connects independent nodes (e.g. sensors and actuators) that are situated in the clothes, on the body or under the skin of a person. The network typically expands over the whole human body and the nodes are connected through a wireless communication channel. According to the implementation, these nodes are placed in a star or multihop topology [1]. A WBAN offers many promising new applications in home/health care, medicine, multimedia, sports and many other areas, all of which make advantage of the unconstrained freedom of movement a WBAN offers.

An important step in the development of a WBAN is the characterization of the physical layer of the network, including an estimation of the delay spread and the path loss between two nodes on the body. This requires a detailed characterization of the electromagnetic wave propagation and antenna behavior near the human body. Propagation near flat, homogeneous and layered phantoms has been investigated in [2]-[3]. Papers [4]-[5]-[6] characterize the wireless on-body channel for some specific configurations of the transmitter and receiver.

In this paper, measurements are performed on a real human using two half-wavelength dipoles, considering different parts of the human body separately. A large number of transmitter and receiver positions are considered. The on-body channel is characterized for the arm and torso. Path loss models have been determined and the lognormal behavior has been validated. Also models for the delay spread and mean excess delay have been developed and the statistical behavior has been analyzed.

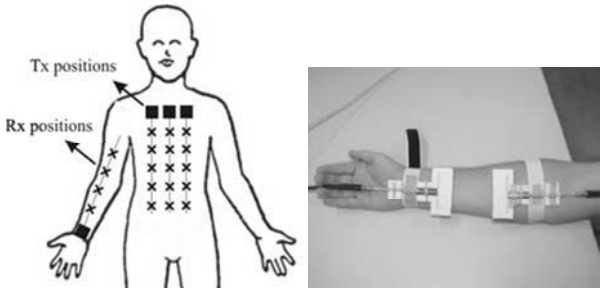
In Section II the measurement setup and the different configurations that were investigated, are described. Then the on-body channel characteristics are examined. Section III presents path loss models for both the arm and torso. In Section IV a time domain analysis is performed and the mean excess delay and delay spread are discussed. Finally, Section V summarizes the conclusions of this paper.

II. Measurement setup

A. Configuration

Two identical half-wavelength dipoles at 2.45 GHz, with a length of 5.7 cm and a diameter of 1 mm, are placed at various positions on a human body. The dipoles are balanced for radiation near the body using a λ/4-bazooka balun. In all measurements the dipoles are placed parallel to each other and lined up for maximal power transfer. The propagation channel characteristics depend strongly on the height of the antenna above the body [2]. In this paper we investigate the wireless channel for a separation of 5 mm between body and antennas.

Models are developed for the channel characteristics along the human arm and torso. Fig. 1 shows the measurement setup and the antenna positions on the arm and torso. Tx and Rx represent the transmitting and receiving antenna, respectively.



(a) (b)

Fig. 1 Measurement setup: (a) antenna positions on the body (■ = Tx and × = Rx position), (b) picture of the setup for measurements on the arm.

First, measurements are performed on a stretched arm. The transmitting antenna Tx is placed at a fixed position on the wrist and the receiving antenna Rx is moved along various positions towards the shoulder (see Fig. 1a). The distance between the antennas varies from 5 cm up to 40 cm in steps of 1 cm. These measurements are performed on two persons (male and female, age 23). A total of 214 measurements are carried out to extract propagation statistics.

The measurements on the torso are performed on a male person lying on a table with the arms along the body. The transmitter is placed at approximately shoulder height at one of three different positions (left, middle, or right, see Fig. 1a). The receiver is placed directly below the transmitter and is moved along a straight line in steps of 2 cm. The antenna separation varies from 5 cm up to 30 cm. A total of 102 measurements are performed on the torso.

B. Measurements

The measurements are performed in a modern laboratory/ office. A vector network analyzer (Rohde & Schwarz ZVR) is used to determine the S_{21}(f)-parameter between Tx and Rx for the different positions (see Fig. 1) in the frequency range from 300 kHz to 4 GHz. This frequency range is necessary to be able to distinguish the direct and the reflected waves in the time domain. We obtain a resolution of 0.25 ns.

Time domain analysis is performed by calculating the channel impulse responses $s_{21}(t)$ using the Inverse Fast Fourier Transform (IFFT) of the measured S_{21}-parameters. Fig. 2 shows the normalized impulse response from 0 ns to 40 ns for a measurement on the arm and an antenna separation of 20 cm. It can be seen that most energy is received via the direct path with different multipath reflections after some time.

III. PATH LOSS MODEL

To model the path loss between the transmitting and the receiving antenna as a function of the distance *d*, we use the following semi-empirical formula, expressed in dB and based on the Friis formula in free space:

$$P_{dB}(d) = P_{0,dB} + 10n\log(d/d_0) = -|S_{21}|_{dB} \quad (1)$$

where $P_{0,dB}$ is the path loss at a reference distance d_0 (10 cm in this paper), and n is the path loss exponent, which equals 2 in free space. The path loss in this paper is defined as $-|S_{21}|_{dB}$, which allows us to regard the setup as a two-port network for which we determine S_{21}. In the following we present path loss models for the human arm and torso.

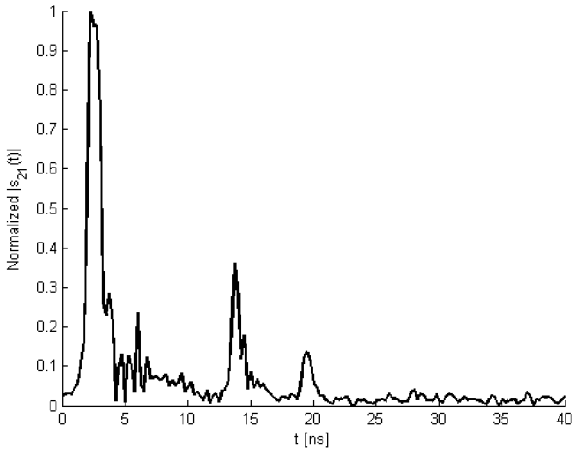


Fig. 2 Measured normalized impulse response on the arm for an antenna separation of 20 cm.

A. Arm and torso

Fig. 3 shows the measured path loss versus Tx-Rx separation for the arm and torso. The circles indicate the individual measurements taken along the stretched arm and the full line represents the path loss model obtained through fitting of the arm measurement data. The crosses indicate the individual measurements taken along the torso and the dotted line represents the path loss model obtained through fitting of the torso measurement data. Table 1 shows the parameter values of the fitted path loss models for the arm and torso according to equation (1), and the variation σ of the measurement results around the model.

The path loss increases with antenna separation as expected. The path loss along the torso and along the arm follow the same course, but the path loss along the torso is higher than the path loss along the arm. This is probably due to the higher absorption in the larger volume of the trunk, and because the surface of the trunk is less flat than the surface of the stretched arm. The path loss exponent of both models is almost the same ($n \approx 3.3$). For the measured path loss along the torso, we observe a slightly higher variation around the path loss model (standard deviation $\sigma = 6.1$ dB) compared with the measurements along the arm ($\sigma = 4.1$ dB). This is because the measurements along the torso were performed on three different lines: left, middle, and right.

The reference path loss $P_{0,dB}$ and the path loss exponent n obtained in this paper, are consistent with previous results: in [7], a path loss exponent of $n = 3.1$ and a path loss value of $P_{0,dB} = 44.6$ dB at a reference distance $d_0 = 10$ cm were measured in a large empty room for waves traveling along the front of the torso.

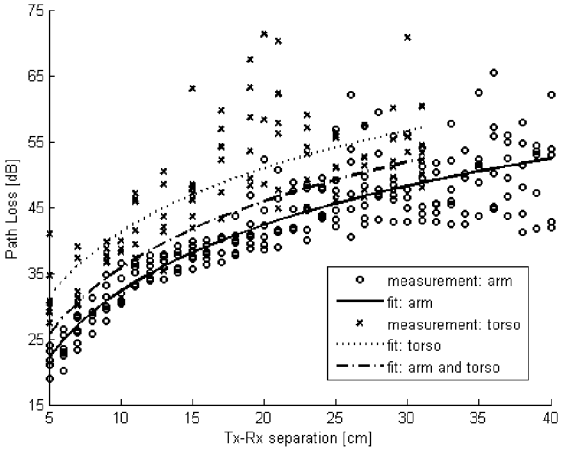


Fig. 3 Measured path loss and fitted models versus antenna separation.

Table 1 Parameter values of the path loss models for the arm and torso.

parameter	arm	torso	arm + torso
d_0 [cm]	10	10	10
$P_{0,dB}$ [dB]	32.2	41.2	35.7
n [-]	3.35	3.23	3.38
σ [dB]	4.1	6.1	6.2

Instead of investigating different parts of the body separately, we can also define an averaged path loss model for the whole human body. For this purpose we consider a path loss model obtained through fitting of all measurement data. The path loss model for the arm and torso measurements altogether is also shown in Fig. 3 (dashed line). This curve lies between the previously derived path loss models. The parameter values of this model are shown in Table 1.

B. Cumulative distribution function

Figs. 4 and 5 show the cumulative distribution function (CDF) of the deviation of measured path loss and model for the arm and torso. It was shown in [8] that the variation around the mean path loss is well described by a lognormal distribution (see lognormal fits in Figs. 4 and 5).

The mean values μ and the standard deviations σ of the lognormal distributions for the arm and torso, fitted using a least-square error method, are provided in Table 2. For waves traveling along the torso, we again observe a slightly higher variation (σ = 5.5 dB) compared with transmission along the arm (σ = 3.4 dB). The values of the CDF parameters in Table 2 indicate that the path loss models show very good correspondence with the measurement results. The mean values of the fitted CDFs of the deviation is close to 0 dB and the standard deviations differ less than 1 dB from the standard deviation of the measured values (see Table 1: 0.7 dB and 0.6 dB difference for arm and torso, respectively).

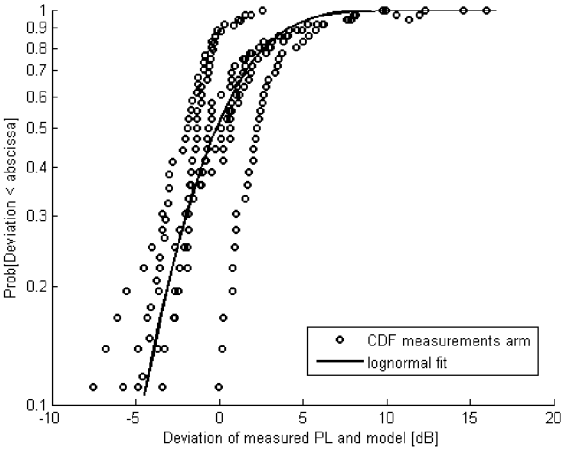


Fig. 4 Cumulative distribution function of the deviation of measured path loss and model along the arm and lognormal distribution fit.

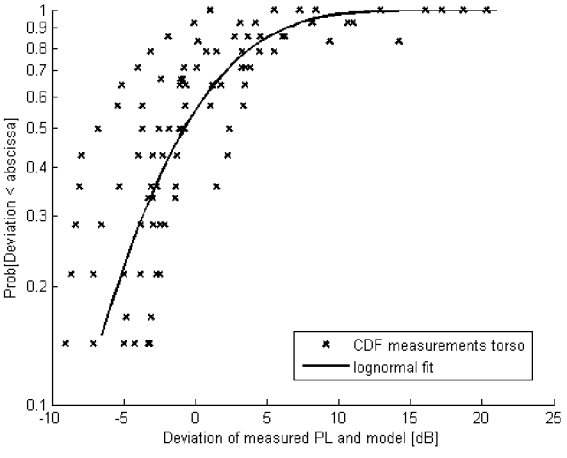


Fig. 5 Cumulative distribution function of the deviation of measured path loss and model along the torso and lognormal distribution fit.

Table 2 CDF of the deviation of measured path loss and model for the arm and torso: lognormal distribution fit.

lognormal fit	arm	torso	arm + torso
σ [dB]	3.4	5.5	4.1
μ [dB]	-0.2	-0.7	-0.3

C. Comparison with flat, homogeneous medium

In Fig. 6 we compare the path loss models obtained using a least-square error fit of the measurement data (for the arm and torso) with a model derived from simulation results near a flat, homogeneous medium representing muscle tissue (specific relative permittivity $\varepsilon_r = 53.57$ and conductivity $\sigma = 1.81$ S/m at 2.45 GHz) [3]. In this path loss model, $P_{0,dB}$ and n are defined by a formula depending on the height h above the medium (see equation (3) in [3]).

Fig. 6 shows that the models derived from the on-body measurements deliver a higher path loss than the model for the flat, homogeneous medium. Path loss models for flat, homogeneous tissues may thus underestimate the actual path loss near a human body. The path loss exponent of the flat, homogeneous muscle simulating tissue ($n = 3.87$) is somewhat higher than those of the arm and torso.

IV. Delay spread model

Time domain analysis is performed by analyzing channel impulse responses $h(t) = s_{21}(t)$, that are calculated from the measured frequency transfer functions from 300 kHz to 4 GHz (see Section II. B) and performing the IFFT.

Power delay profiles (PDP) were calculated for all the measurement positions. For the following analysis only peaks (see Fig. 2) of less than 35 dB below the maximum value of the PDP are taken under consideration. The PDP is characterized by the first central moment (mean excess delay τ_0), and the square root of the second moment (RMS delay spread τ_{rms}). The RMS delay spread provides a figure of merit for estimating data rates for multipath channels [9]. The mean excess delay τ_0 is defined as

$$\tau_0 = \frac{\sum_{i=1}^{N} \tau_i \left|h(\tau_i)\right|^2}{\sum_{i=1}^{N} \left|h(\tau_i)\right|^2} \tag{2}$$

and the delay spread τ_{rms} as

$$\tau_{rms} = \sqrt{\frac{\sum_{i=1}^{N} \tau_i^{\,2} \left|h(\tau_i)\right|^2}{\sum_{i=1}^{N} \left|h(\tau_i)\right|^2} - \tau_0^{\,2}} \tag{3}$$

where $h(\tau_i)$ is the time domain impulse response at time τ_i, obtained from the measurement data.

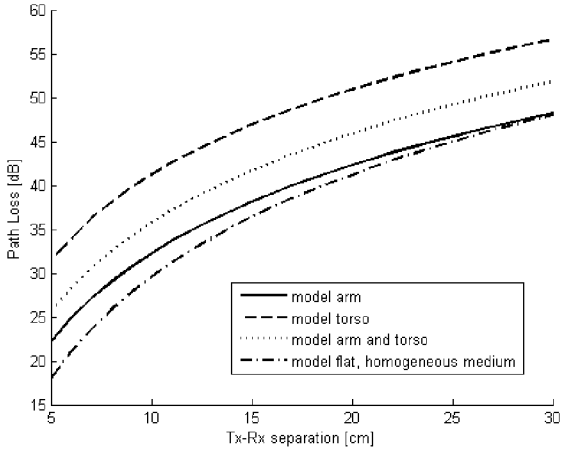


Fig. 6 Comparison of the path loss models derived from the measurements along the arm and torso, with a model near a flat, homogeneous medium.

A. Arm and torso

Figs. 7 and 8 show the mean excess delay τ_0 and the RMS delay spread τ_{rms} versus antenna separation for both measurement cases (arm and torso). The values of τ_0 and τ_{rms} increase with antenna separation. A linear fit for τ_0 and an exponential and logarithmic fit (with breakpoint d_{bp}) for τ_{rms} are performed, using a least-square error method. The fits have the following equations (with distance d in cm):

$$\tau_0(d) = A \cdot d + B \;[ns] \tag{4}$$

$$\tau_{rms}(d) = C\left(e^{D \cdot d} - 1\right) \;[ns] \text{ for } d \le d_{bp} \tag{5}$$

$$\tau_{rms}(d) = E + F \ln\left(d / d_{bp}\right) \;[ns] \text{ for } d > d_{bp} \tag{6}$$

The parameter values of the fits are shown in Table 3. For the measurements along the torso, the variation around the fitted models is slightly higher ($\sigma = 1.9$ ns for τ_0 and $\sigma = 1.4$ ns for τ_{rms}) compared with the measurements along the arm ($\sigma = 1.1$ ns for τ_0 and $\sigma = 1.1$ ns for τ_{rms}).

Table 3 Parameter values of the models for the mean excess delay τ_0 and the RMS delay spread τ_{rms}, for the arm and torso.

parameter	arm	torso
A [1/cm]	0.35	0.76
B [ns]	-1.11	-2.71
C [ns]	1.53	0.58
D [1/cm]	0.09	0.23
E [ns]	10.07	12.38
F [ns]	6.17	6.13
d_{bp} [cm]	23.0	13.6

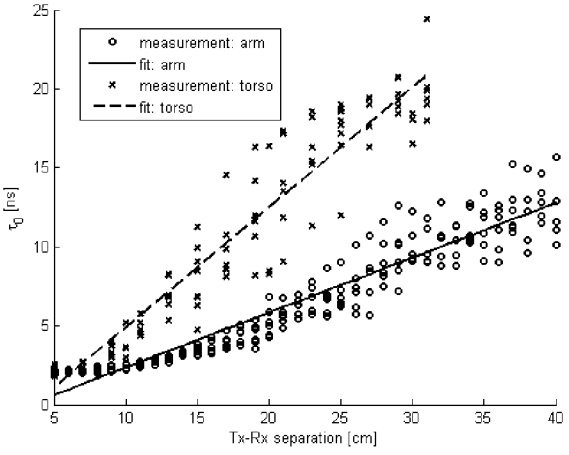


Fig. 7 Mean excess delay τ_0 and linear fit versus antenna separation.

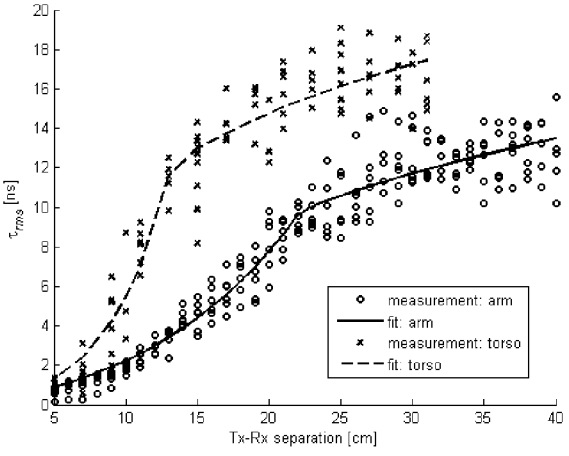


Fig. 8 RMS delay spread τ_{rms} and exponential/logarithmic fit versus antenna separation.

B. Cumulative distribution function

In this section we verify whether τ_0 and τ_{rms} follow a lognormal distribution. Other different empirical distributions were applied in previous research [6], however, lognormal proved to be the best fit. We investigate the deviation of the measured values and the models, see equations (4)-(6). Figs. 9 and 10 show the CDFs of these deviations and a lognormal distribution fit for τ_0 and τ_{rms}, respectively, for the measurements along the torso. For the measurements along the arm, we obtain similar CDF distributions.

The mean values and the standard deviations of the CDFs for the arm and torso, fitted using a least-square error method, are provided in Table 4. These parameters indicate that both models show very good correspondence with the measurement results. The mean values are close to 0 dB and the standard deviations differ less than 0.1 dB from the variation of the measured values discussed in Section IV. A.

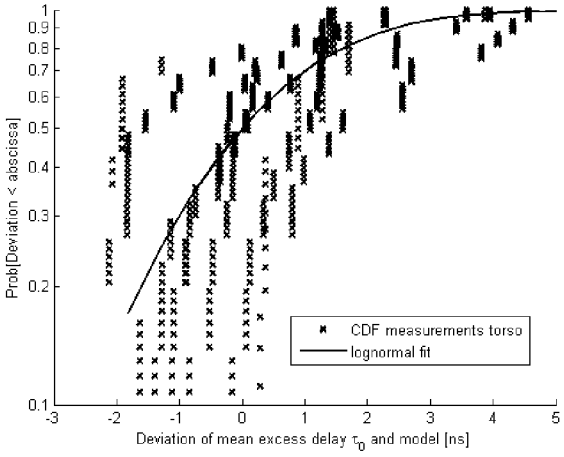


Fig. 9 CDF of the deviation of mean excess delay τ_0 and model along the torso and lognormal fit.

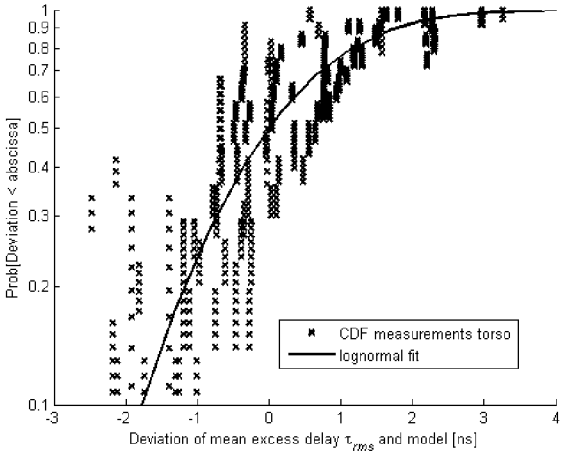


Fig. 10 CDF of the deviation of RMS delay spread τ_{rms} and model along the torso and lognormal fit.

Table 4 CDF of the deviation of mean excess delay τ_0 and RMS delay spread τ_{rms} and models for the arm and torso: lognormal distribution fit.

lognormal fit	**τ_0**		**τ_{rms}**	
	arm	**torso**	**arm**	**torso**
σ [ns]	1.1	1.9	1.0	1.4
μ [ns]	-0.05	0.00	-0.05	-0.005

V. Conclusions

Path loss models for the human arm and torso have been derived. It is found that the path loss along the torso and along the arm follow the same course, but the path loss along the torso is higher than the path loss along the arm. This is probably due to the higher absorption in the larger volume of the trunk, and because the surface of the trunk is less flat than the surface of the stretched arm. The cumulative distribution functions of the deviation of measured path loss and models are well described by a lognormal distribution, for both the arm and torso. The mean values and the standard deviations of the lognormal fits indicate that the path loss models show excellent correspondence with the measurement results.

The models derived from the on-body measurements deliver a higher path loss than the model for a flat, homogeneous muscle simulating medium. Path loss models for flat, homogeneous tissues may thus underestimate the actual path loss near a human body.

Finally, the mean excess delay τ_0 and the RMS delay spread τ_{rms} were studied. These parameters increase with antenna separation. Models have been determined for τ_0 and τ_{rms} and show excellent correspondence with the measurement results. The cumulative distribution functions of the deviation of τ_0 and τ_{rms} and the derived models are well described by a lognormal distribution, for both the arm and torso.

Acknowledgment

This work was supported by the Interdisciplinary institute for BroadBand Technology (IBBT) and the Research Foundation - Flanders (FWO). The authors would like to thank Toon Depessemier for helping us during the measurements.

References

1. Latré B, Vermeeren G, Moerman I, Martens L, Louagie F, Donnay S, Demeester P, Networking and propagation issues in body area networks, 11th Symposium on Communications and Vehicular Technology in the Benelux 2004, SCVT 2004, Ghent, Belgium, Nov. 2004.
2. Roelens L, Joseph W, Martens L, Characterization of the path loss near flat and layered biological tissue for narrowband wireless body area networks, International Workshop on Wearable and Implantable Body Sensor Networks (BSN2006), MIT Media Lab, Boston, USA, Apr. 2006, CD-ROM.
3. Roelens L, Van den Bulcke S, Joseph W, Vermeeren G, Martens L, Path loss model for wireless narrowband communication above a flat phantom, IEE Electronics Letters, vol. 42, no. 1, pp. 14–15, Jan. 2006.
4. Fort A, Desset C, De Doncker P, Wambacq P, Van Biesen L, An ultra-wideband body area propagation channel model – from statistics to implementation, IEEE Trans. Microwave Theory and Tech., vol. 54, no. 4, pp. 1820-1826, Apr. 2006.
5. Ryckaert J, De Doncker P, Meys R, de Le Hoye A, Donnay S, Channel model for wireless communication around human body, IEE Electronics Letters, vol. 40, no. 9, pp. 543-544, Apr. 2004.
6. Alomainy A, Hao Y, Hu X, Parini C G, Hall P S, UWB on-body radio propagation and system modeling for wireless body-centric networks, IEE Proceedings-Communications, Feb. 2006, vol. 153, pp. 107-114.
7. Fort A, Desset C, Ryckaert J, De Doncker P, Van Biesen L, Wambacq P, Characterization of the ultra wideband body area propagation channel, IEEE International Conference on Ultra-Wideband (ICU), Zurich, Switzerland, Sept. 2005, CD-ROM.
8. Fort A, Desset C, Ryckaert J, De Doncker P, Van Biesen L, Donnay S, Ultra wide-band body area channel model, IEEE Conference on Communications (ICC), Seoul, Korea, May 2005, vol. 4, pp. 2840-2844.
9. Saunders S R (1999) Antennas and propagation for wireless communication systems. Wiley, West Sussex.

Address of the corresponding author:

Author: Elisabeth Reusens
Institute: Department of Information Technology, Ghent University
Street: Gaston Crommenlaan 8 bus 201
City: B-9050 Gent
Country: Belgium
Email: Elisabeth.Reusens@intec.ugent.be

8th Session
In-Body Sensors

Healthy Aims Overview

Dr Diana Hodgins MBE DSc (Honorary)

Managing Director, European Technology for Business Ltd, Codicote, England

Abstract-

The Healthy Aims project, www.healthyaims.org, is a 23M€ four year project funded under the IST FP6 programme, with 16M€ combined funding from the EU and Switzerland. The project goal is to develop a number of intelligent medical implants and diagnostic systems integrating a range of underpinning Micro and Nano technologies. After three years a number of results are already available, disseminated to over 400,000 people worldwide and summarised in this paper.

The project has 25 partners from ten EU countries, including seven SMEs, six clinical partners, five LEs and seven academics and research groups. This combination of disciplines has enabled the Consortium to design, fabricate and carry out laboratory and clinical trials on the range of medical products under test. It has also enabled the project to set realistic specifications for core components that form an integral part of the implants, for example the implantable battery and the wireless data communications system. A further valuable benefit is that the micro and nano structures are designed specifically for use in an application, so again the specification requirements are clearly defined.

Keywords- **implants, electrodes, wireless, sensors**

I. MEDICAL PRODUCTS UNDER DEVELOPMENT

The products under development are:

- Retina, cochlear, and intra-cranial pressure (ICP) sensor implants
- Functional stimulation systems for bladder and bowel control and restoring arm movement
- Glaucoma and sphincter sensors and sensor systems for human body motion

The active implants all provide electrical stimulation directly to nerves or muscles in the body. The sensor systems make a direct measurement of parameters that can be used as indicators for particular aspects of health and well-being.

The technological challenges that have been addressed within the project in order to produce the products are described in the following sections.

II. WIRELESS COMMUNICATIONS

In order to minimise infections implanted devices must be completely self contained and therefore means of transmitting data from in the body to on the body must be wireless. The carrier frequency for data transfer is defined by the application requirements and medical safety standards. In some cases it is also combined with inductive power transmission. For the devices being developed within Healthy Aims the following choices have been made.

Table 1: Carrier frequencies for data transfer

Product	Carrier freq for data
Glaucoma sensor	27.3 MHz
Retina implant	IR
Functional Electrical Stimulation (FES)	403 MHz (MICS)
Cochlear implant	5MHz
ICP sensor	13.56MHz

The integration of a Medical Implant Communication Service (MICS) chip into two FES implants has been developed by Zarlink. A demonstrator is now available and Zarlink, UCL and FTM are now working on the integration of the chip and the antenna with the implant enclosure. The goal is to demonstrate the complete implant system in a pseudo body environment by the end of the project.

III. NOVEL ELECTRODES

Electrical stimulation is applied via electrodes and these have unique requirements for each of the product applications. The retina and cochlear implants face the greatest challenge in terms of the required number of electrodes/mm^2, the preferred topology and interconnects.

The team of electrode partners, INEX, EPFL, IMEC and ITE have worked with the product partners to produce prototype electrodes for the retina and cochlear implants and these are now being tested by the two product partners.

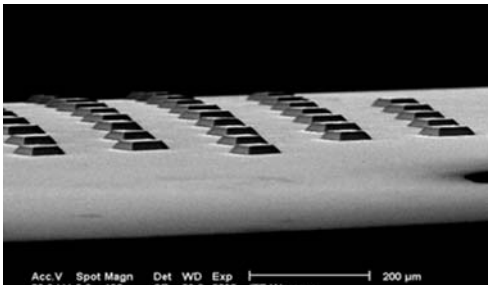


Fig 1
Truncated prism electrodes for the retina implant

Fig 2
Modiolus electrodes for the cochlear implant

IV. BIOMATERIALS FOR COATING THE IMPLANTS

All implant components need to either be fabricated using a biocompatible material or be completely encapsulated with a biocompatible material in order to prevent the release of leachables from the implant components to the body and to stop water ingress to the components. These biocompatible coatings must remain stable for the lifetime of the implant, which in many cases can be the same as the lifetime of a person. In this project QMUL has been working with the implant product partners to develop a biocompatible coating that is thin, flexible and rugged enough for handling during surgery. Preliminary results from silicone coated with diamond like coating (DLC) and IPM modified silicone indicate that these could meet the physical requirements and these could improve the water resistance compared to silicone alone. Further studies are ongoing and results will be available next year.

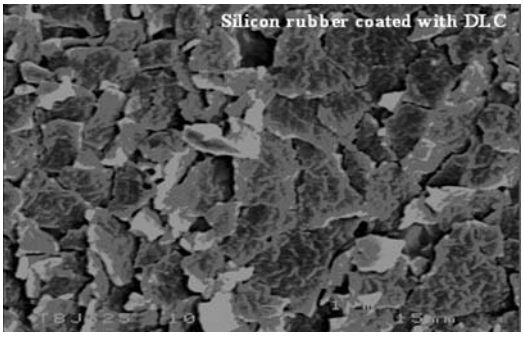


Fig 3
Silicone rubber coated with diamond-like carbon (DLC)

V. ENHANCING THE ELECTRODE - NERVE INTERFACE

The electrode - nerve interface has a significant effect on the power requirement. INEX has been working on ways of optimising this interface by:

- Discouraging growth of scar tissue on electrodes which increase the impedance at the electrode interface
- Encouraging the growth of neuronal cells on the electrodes which encourages electron flow at the interface

Physically adapting the electrode surface with 1µm grooves has been shown from in–vitro testing to align neuronal cells and will next be validated using real electrodes for the retina and cochlear implants.

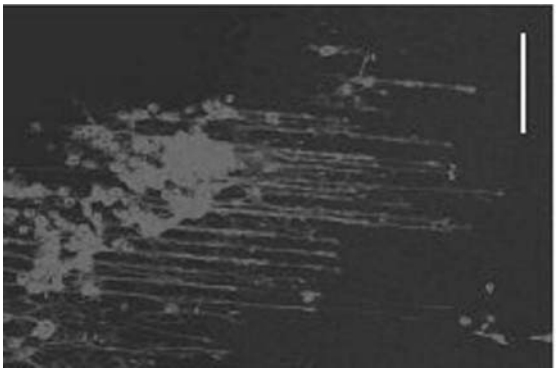


Fig 4
Alignment of electrically active neurons to the electrode surface

VI. INTEGRATION OF ELECTRODE ARRAYS WITH THE ELECTRONICS

The connection between electrode arrays and the electronic circuit which enables each electrode to be individually addressed is essential for the retina and cochlear implants. This has been jointly addressed by the product partners and technology partners developing the electrodes. A considerable amount of effort has also been put into developing methods for thinning ASICs and flip chip techniques to mount the ASICs onto flexible membranes. These techniques have already been applied to the prototype glaucoma sensor, the implants and the intra-cranial pressure sensor (ICP) system. ASIC thinning to 50µm and subsequent flip chip bonding has been successfully demonstrated and provides adequate compliance for these applications.

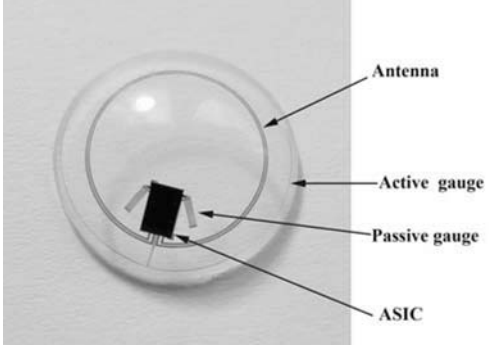
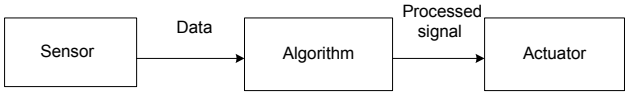


Fig 5
Mounting of the thinned ASIC in the glaucoma contact lens sensor

VIII. EMBEDDED SOFTWARE AND ALGORITHMS

Some form of intelligence needs to be integrated into the product hardware in order to provide the patient with a system solution. Within this project a number of algorithms have been developed which interpret sensor signals and produce an output. In many cases that output is then used as the input to the actuator.



Specific examples that are being developed in the project are:

- Processing arm motion data and providing a single trigger for the FES system only when it is recognised that the patient intends to reach for an object
- Processing image data from glasses worn by user and providing a pattern of electrode array stimulation in real time for the retina implant
- Processing human body motion data and providing a classification of activity

For the sensor - actuator systems the algorithms are embedded into processors, which form part of the on-the-body unit. It was therefore important that the systems were designed to minimise power to enable the systems to work whilst the user is awake and recharge whilst asleep.

IX. POWER SOURCES

Power sources are a critical component with any of the medical devices or diagnostic systems. For the in-the-body applications the requirements are extremely challenging. Saft and CEA have produced a prototype secondary cell that can be used for the FES and Cochlear applications and laboratory trials have already been successfully completed.

Parameter	Specification	Unit
No: of charges	4000	-
Life time	10	Years
Dimensions	5x10x22	Mm
Minimum voltage	3	Volts
Average current	5	mA
Min charge time	1	hour

Table 2 Secondary cell specification

Fig 6
Prototype Secondary cell

Work is now underway on the integration of the battery and charging circuit into the implant enclosure. A key requirement is battery charging must not significantly raise the body tissue temperature.

For lower power implants, for example a Pacemaker, a direct Biofuel cell is being developed by Dinamic and IMTEK. The goal is to produce a Biofuel cell that can provide adequate power for a Pacemaker or equivalent for up to 10 years without any intervention.

Already a prototype has been demonstrated that produce $4\mu W/cm^2$ using biocompatible materials.

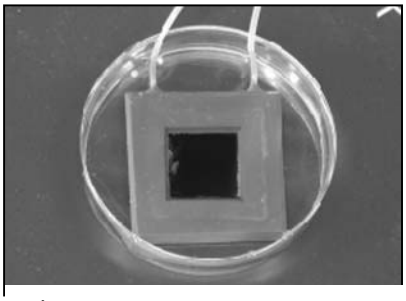


Fig 7
Prototype direct biofuel cell

X. THE PRODUCTS

The products have provided the necessary platform for technology integration. At the end of Year 3 many of the technologies have already been integrated into prototype systems, and in some cases clinical trials have started. This has been possible by including clinical experts/surgeons into the Consortium together with end manufacturers with knowledge and experience in obtaining approvals for clinical trials for active implants.

For the hand and wrist extension FES some clinical trials have already been completed with a 2 channel surface stimulator. Clinical trials are planned for next year with 2 channel passive implantable system (electrodes implanted, power and active electronics external). By the end of the project a 6 channel demonstrator fully implantable (with the battery and MICS system in the implant) will be available. This will be triggered using the sensor system and algorithm also developed in the project.

Work on two other FES applications for bladder and bowel stimulation started in Year 3. Animal trials have already started with one clinical team and will be continued in Year 4.

Prototype retina implants have been produced by IIP and are just starting clinical trials. Preliminary clinical results from Medical University of Graz indicate that some vision will be given to the patient.

Fig 8
Prototype retina implant

The glaucoma sensor has already completed initial clinical trials with a wired unit. The first wireless contact lens version has now been produced combining the 50µm thick ASIC and flip chip attachment method and this will undergo clinical trials next year.

ICP prototypes are also available from CMT and lab trials have shown excellent stability over a 30 day period, which is the critical performance requirement and some animal trials are planned for next year at NBNHST.

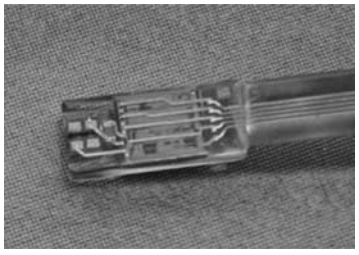


Fig 9
Sensor head for the ICP sensor system

Fig 10
Schematic showing the positioning of the ICP sensor system

The cochlear implant was available at the start of the HA project. However, advancements in electrode fabrication within the project have enabled CTC to look at product enhancement. In the final year of the project CTC will undertake laboratory trials on electrode systems and if these are successful will then prepare for clinical trials.

The activity monitor hardware is now available and has already started laboratory trials. In the final year of the project this product will be extended to provide full motion data, simulating the 'Gait Lab' environment.

XI. SUMMARY

The goal of the project to develop new medical implants and diagnostic systems using micro and nano-technology is well on its way to being met. Some products are now entering clinical trials, some are just commencing animal trials and others will have laboratory demonstrators at the end of the project. This considerable achievement is considered to be attributable to the fact that the clinical teams have led the product partners, who themselves have defined to the technology partners their exact requirements. Spending time on the specifications early on in the project has meant that there have been few design iterations. The result is that the project has managed to progress all the way to clinical trials in some cases. The final year will undoubtedly bring some fascinating results from these trials and will lead the way for future medical implant development.

ACKNOWLEDGMENT

This research project is funded by the European Information Societies Technology Programme (IST-2002-1-001837).

Address of the corresponding author:

Diana Hodgins
Project Co-ordinator Healthy Aims
European Technology for Business (ETB) Ltd
Codicote Innovation Centre
St. Albans Rd
Codicote
Herts
SG4 8WH
UK
Tel: +44-(0)1438 822822
Fax: +49-(0)1438-822811
Email: diana.hodgins@etb.co.uk

Communication and Control for Reanimating Paralyzed Limbs via a Network of Wireless Micro-Implants

Nuria Rodríguez[1,2], Jack Weissberg[1], Gerald E. Loeb[1]

[1] Alfred. E. Mann Institute for BioMedical Engineering, University of Southern California. Los Angeles, California 90089
[2] TECNUN-CEIT, P. Manuel de Lardizábal 15. San Sebastián, Spain 20018

Abstract— **Functional reanimation of a paralyzed limb requires stimulation and sensing in multiple sites distributed throughout the limb. This paper describes a communication and control system for BION2™, a biomimetic system based on wireless, injectable neuromuscular implants that can detect the user's intention, activate muscles and monitor the ensuing movements to support distributed feedback control with delays comparable to those achieved by the spinal cord. Various design features optimize the safety and efficacy of a system that must be reconfigurable to achieve many different clinical requirements despite highly constrained power and bandwidth.**

Keywords— **Communications protocol, Sensor control protocol, RF battery-less sensors, implantable sensors, implantable microstimulators**

I. INTRODUCTION

Neuromuscular electrical stimulation has been used as a rehabilitation therapy to prevent and treat disuse atrophy in patients with upper motor lesions such as stroke and spinal cord injury. BIONs™ (BIOnic Neurons) are relatively new, injectable, inductively powered micro-implants (2mm diameter x 16mm long) that provide selective and repeatable exercise of individual muscles without requiring surgery [1, 2]. The individually implants addressed receive power and command signals from a control unit through an external coil (see Fig. 1).

In principle, similar electrical stimulation can be used to create functional movement of a limb, but this requires interfaces to sense the motor intentions of the user and the ongoing limb movements for real-time, closed-loop control. BION2, the next generation BION implants now under development, will incorporate several sensing modalities. This paper describes the design of the bidirectional telemetry system that will control and synchronize the function of 1-20 such implants that must be configured dynamically to support a wide variety of sensorimotor behaviors

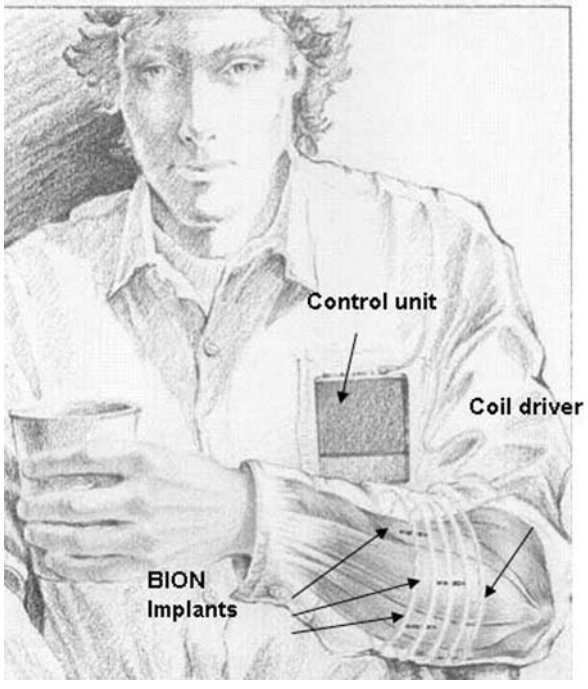


Fig. 1- BION System Overview

II. MUSCLE CONTROL

Muscle force depends on the number of motor units that have been recruited and their frequencies of firing, as well as nonlinear effects of the length and velocity of motion (shortening or stretching) of the muscle fibers. BIONs are usually implanted near the entry zone of the nerve containing the motor axons, where increasing the intensity of pulsatile stimuli can recruit an increasing percentage of the motor units. Stimulus efficacy is generally proportional to the charge of each pulse, the product of pulse current and pulse

duration that are specifiable dynamically as part of the command structure for each BION.

The usual firing rates for motor units are relatively low (20-30pps). If the stimulus rate is set too high, the recruited muscle fibers will fatigue quickly. If it is set too low, the force produced by the muscle may have a substantial ripple at the stimulus rate and it will be difficult to ramp up force quickly such as for responding to perturbations. When controlling a limb, it is desirable to provide bursts at high frequency in order to achieve strong contractions and to reduce force ripple at low stimulation rates by staggering the stimulus pulses from synergistic sites. Both features are accommodated by the frame architecture described below, which provides mechanisms to generate duplicate stimulus pulses with identical parameters within a given frame and to contract the muscles asynchronously as required for smooth movements.

III. NATURAL FEEDBACK SENSORS AND BION SENSORS

Biological muscles are endowed with a variety of sense organs, including spindle stretch receptors, whose sensitivity can be adjusted dynamically by the fusimotor gain control system. These signals are combined with those from cutaneous sensors and with the descending command signals from the brain in a sophisticated set of excitatory and inhibitory interneurons in the spinal cord that provide much of the input to the motor neurons.

The BION2 system has to provide stimulation commands and sensory feedback with response times similar to those of the biological reflex responses to perturbations. This time is limited by physiological constraints like the transit delay to and from the spinal cord (20-30ms). Muscle fibers respond sluggishly to changes in their neural activation with time constants of about 50ms (rising phase) and 100ms (falling phase) to step changes in that activation. These delays pose serious challenges for the use of closed-loop control in both biological and prosthetic systems but they provide useful hints about specifications of a biomimetic communication and control scheme (see below) for the following modalities of sensing:

A. Patient intention

The electromyogram (EMG) is a stochastic pattern of electrical potentials that arises from the temporospatial overlap of asynchronously firing motor units. EMG can be recorded to detect residual voluntary control from which to infer the user's intentions. The EMG sensing scheme included in BION2 is based on integrating samples for a period that can be determined dynamically by the external controller (see frame architecture description below).

B. Muscle response

The recruitment of the muscle in response to each stimulus pulse is used to map a range of stimulus intensities to a percentage activation of the muscle and to adjust stimulation parameters on-line to cope with shifts of this recruitment curve due to mechanical deformation of the contracting muscle. The EMG recording and integration subsystem can be used for this task but the sampling time must be controlled and synchronized with the stimulation pulses of the same and other implants.

C. Posture Sensing

In order to coordinate movements and compensate for perturbations, the controller needs information about the starting posture and ongoing trajectory of these movements. Three complementary sensing modalities will be incorporated:

- Orientation with respect to gravity and translational acceleration determined by a MEMS accelerometer that has 2 axes of piezoresistive bridge elements [3].
- Absolute position through the detection of reference magnetic fields created outside the body [4].
- Posture of distal joints of the hand and fingers, inferred from changes in the relative position of implants in the muscles that operate those joints [5].

IV. PROTOCOL REQUIREMENTS AND BION IMPLEMENTATION

The BION2 control and communication protocol was designed to meet the following critical requirements:

- Efficient use of forward and reverse data rate
- Predictable intervals for reflex adjustment
- Tolerance for transmission errors
- Dynamic reinitialization of individual implants

Most of these aspects either mandate or reflect the selection of a communication protocol based on a sequence of frames, each of which consists of the same number and length of messages and the same order of actions.

A. Efficient use of forward & reverse data rate

BIONs are intended to be used for a wide range of clinical applications with different numbers of implants and

varied requirements for sensing, stimulation and reaction speed. The main objective of the communication protocol is to allow the limited bit rate in each direction to be reconfigured as needed for such applications. This configuration occurs during an initialization transmission to each implant, which sets the number of bits and the data that they represent for both forward and reverse telemetry during a given operational session.

- *Duplex communication*: The physical layer is designed to allow full duplex communication. Forward telemetry provides 120kb/s via a frequency-shift keyed (FSK), Manchester-encoded signal over a 480kHz carrier frequency that provides the clock and power for all implants. Reverse telemetry provides 480kb/s via on-off keyed (OOK) bursts of a crystal-stabilized 400 MHz carrier limited by the bandwidth of a single MICS channel.
- *Prefixed time slots*: The time slots for the forward and reverse transmissions for each implant are fixed by the external controller when the system is turned on and the implants are initialized (see below). The use of these predefined channels avoids collisions between transmissions from several BION implants sending data to the controller. In addition, the controller can identify data from each BION without including headers in back telemetry and each BION can identify the presence of incoming data from the controller also without headers in forward telemetry. All devices and actions are synchronized by counting clock cycles based on the inward telemetry and power carrier transmitted by the external controller.

B. Predictable frame intervals for reflex adjustment

Frame in this context is a programmed sequence of events performed by all implants that repeats at a rate consistent with the default stimulation frequency for each muscle (typically 20-30pps). The duration of a frame and the exact timing of events within the frame are controlled by two unique Manchester code violations that cannot arise from the data sequences themselves.

- A *Frame Sync* is a signal that triggers a new frame or sequence of events. The frame time is set by the external controller in real time.
- *Internal Syncs* are signals that can be received at any time inside the frame. They are responsible for triggering each event in the frame (stimulation, starting and ending points of sensing operations or reinitialization if necessary) in an order that is programmed as part of the initialization command.

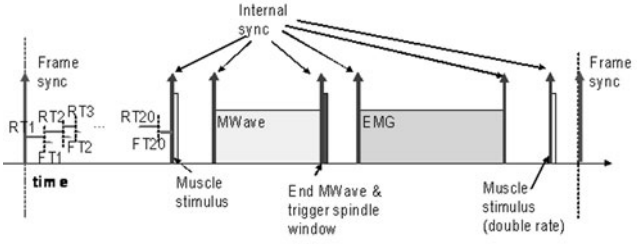


Fig. 2- Control signals in a BION frame. (RTi: Reverse Telemetry and FTi, the Forward Telemetry, BION i.)

Frame Sync and Internal Sync concepts are illustrated in Fig. 2 by the sequence of events included in a frame and their respective timing. The first event after a Frame Sync is the communication stage: reverse and forward telemetry from and to each implant, respectively, according to bit counts and time slots that are programmed as part of the initialization of each implant. After the communication stage, the BIONs wait for the Internal Syncs that are used to trigger various events during the frame. The BION example included in Fig. 2 stimulates the muscle, measures the muscle response by integrating the M-wave for a variable period, then detects the joint position by measuring the amplitude of brief stimulus artefacts generated by other implants, and finally integrates the background EMG activity over a variable period to provide voluntary command data. All sensor data gathered during these functions are held in a LIFO register until the next frame, when its reverse telemetry slot arrives and it sends the data back to the controller. Thus the delay for responding to any sensory feedback or command signal is two frames.

Inside the implant one of the initialization registers is called "Sync Mask" because it determines which BION action (stimulation or various sensing modes) is triggered by each successive Internal Sync signal. The combination of Frame Syncs, Internal Syncs and Sync Masks is one of the key points that make the control inside a frame very flexible while permitting tight synchronization of events between implants:

- *Improving sensing modality accuracy and dynamic range*: the integration time of two sensing modalities (EMG and M-wave) is controlled from the outside with Internal Syncs. The external controller can adjust them dynamically without requiring transmission and storage of these timing parameters.
- *Increasing the strength of muscle contraction*: as mentioned in section II, one powerful and rapid way to increase the effectiveness of muscle stimulation is by increasing the stimuli rate. This can be done using Internal Syncs to trigger this "extra stimulus pulse" in

some selected BION implants. As illustrated in Fig. 2, the last Internal Sync is sent by the controller only when an additional stimulation pulse is required.

- *Smoothing limb movements*: as mentioned in section II, in order to produce smooth joint torques, the physiological activity of motor units should be asynchronous. This can be achieved by strategically phasing Sync Mask values for stimulation as shown in Fig. 3.

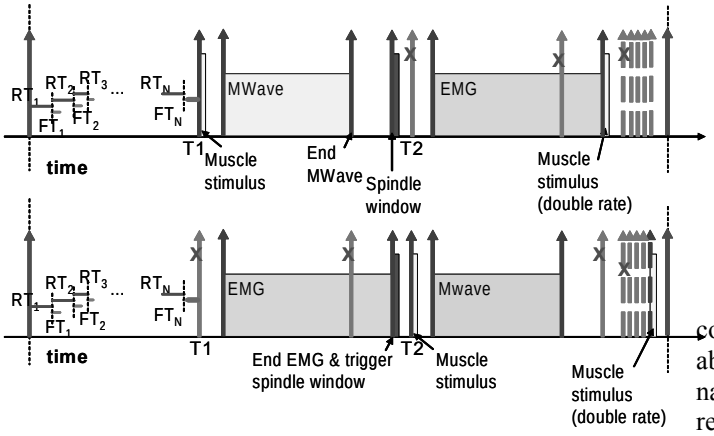


Fig. 3- Control signals in a BION frame for two different sets of BION to obtain a smooth joint torque

In Fig. 3 the Internal Sync signals that are not used for the corresponding BION are marked with an X and the optional ones are represented as broken lines. In this particular case, some of the BION implants in the sytem are stimulating the muscles at time T1 and some of them at T2. After stimulation, each implant records the M-wave response. In the case shown, to obtain double rate stimulation additional Internal Syncs can be sent, and the double stimulation rate is achieved with the 8th Internal Sync for the first BION and with the 13th for the second one.

C. Error tolerance by limiting consequences

Communication errors can be classified into critical errors and bearable errors depending on the potential consequences. Bearable errors are errors in dynamic data transmission whose possible range is limited to "safe values" (see below). Critical errors in communication that must be avoided are those whose consequences are indeterminate. The initialization of each implant is verified by a complete echo of these data (see section □.D) because it configures the system to be tolerant of transmission errors in either direction that may occur during actual operation.

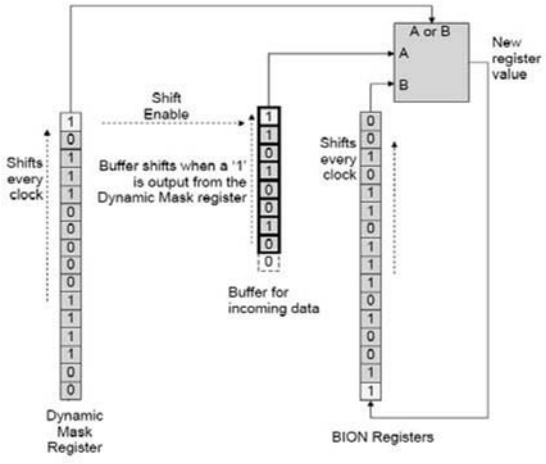


Fig. 4- Dynamic Mask operation

To ensure that the maximal stimulation error caused by communication errors in the dynamic parameters are bearable and will never be dangerous for the patient, the "Dynamic Mask" mechanism is used. The Dynamic Mask is a register included in each BION whose bits correspond to each bit stored in the parameter registers for stimulation and sensing modalities. If a bit in the parameter registers is to be affected by dynamic data on a frame-by-frame basis, its corresponding bit in the dynamic mask register is set to a "1"; if that value is to remain unchanged, its corresponding bit in the dynamic mask register is set to a "0", protecting that bit to stay at the value initially set in the parameter registers (Fig. 4). Dynamic mask register has two purposes: to avoid dangerous errors and to allow frequently changing parameters to be adjusted while minimizing the number of bits sent in each frame. Because the possible errors have been limited to a safe range by the Dynamic Mask, it is better for the BION to act with the wrong parameters than to skip stimulation and allow the muscle to relax.

Forward telemetry data specifying dynamic parameters is sent with a parity bit so that transmission errors can be detected. Error conditions are sent as status bit with the reverse telemetry so that the external controller can take appropriate action

D. Dynamic reinitialization of individual implants

All of the programmable parameters of an implant (including the critical communication timing and mask registers described above) are held in volatile RAM that depends on power received by inductive coupling from the external coil-driver. The very movement created by the muscle stimulation can cause shifts in the relative alignment be-

tween the coil-driver and one or more implants. If the received voltage in a given implant drops under a critical level, that implant goes through a reset operation, clears all internal registers, and does not participate in the next communication phase. The absence of reverse telemetry carrier during the preassigned time slot for this implant informs the external controller which implant needs reinitialization. The possibility of such occasional, isolated drop-outs makes it important to enable dynamic reinitialization of individual implants without affecting the ongoing performance of the other implants.

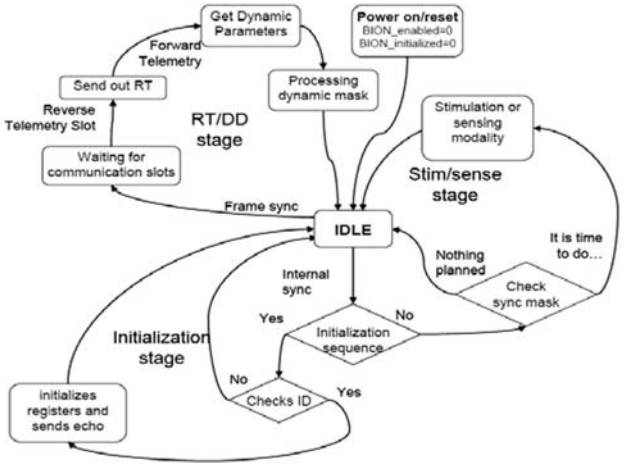


Fig. 5- State Diagram

Any implant that is in an uninitialized state will respond only to a special initialization code consisting of an Internal Sync followed by an initialization sequence (see Fig. 5). The first part of an initialization command is the complete, unique 32-bit identification code. If this ID code matches the ID code that is hard-wired into the ROM of the implant's IC, the subsequent data are used to initialize all programmable registers and are echoed via reverse telemetry to the external controller for complete verification. The implant is now initialized but not yet enabled, which occurs only in the next frame, if and when the external controller detects its reverse telemetry signal in the correct time-slot and synchronizes this implant with the whole system.

Thus the state diagram in Fig. can be divided into three stages: initialization, communication and stimulation/sensing. The controller can reinitialize any device at any time that the device is in the "Idle" state, i.e. not transmitting or receiving data. This option can be useful not only to recover a BION that lost power but also to allow changes in parameters that are outside the previous dynamic range.

V. Conclusions

The BION2 communication protocol incorporates several strategies useful in systems in which multiple devices with a range of possible functions must be configured dynamically to work with command and data channels that have limited bit rates and non-zero error rates. The global signals Frame Sync and Internal Sync and the run-time assignment of bit slots for inward and outward full-duplex telemetry for each device greatly reduces time normally allocated to headers and device addresses in reconfigurable systems. In addition, the Dynamic Mask described in this paper allows the controller to select specific parts of sensing and stimulating parameters for dynamic adjustment, while protecting the system from gross error. This enables aggressive use of the available carrier bandwidth to achieve high data rates without requiring complex and time-consuming error correction. Fig 6 summarizes the influence of design features on the performance characteristics of the communication protocol.

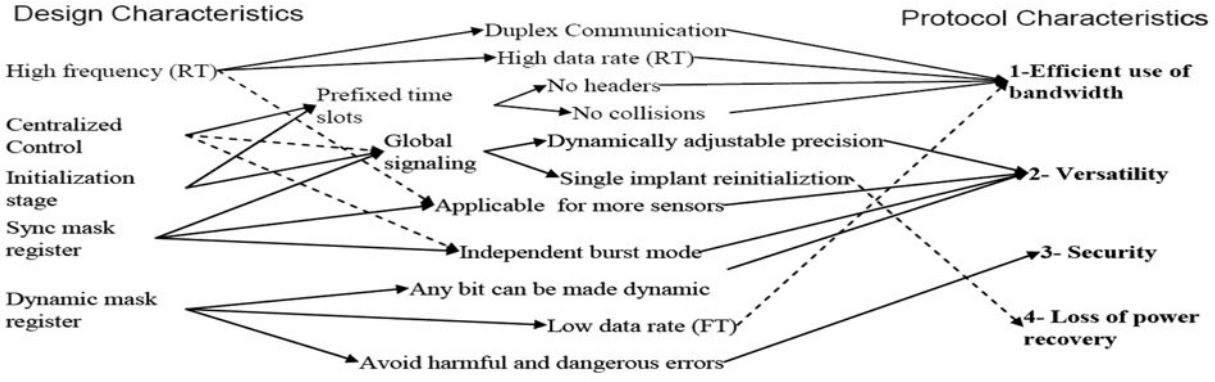


Fig 6- Summary of protocol characteristics

ACKNOWLEDGEMENT

Supported by NIH BRP Grant #R01EB002094 and the NSF Engineering Research Center for Biomimetic Micro-Electronic Systems. The authors would like to thank Centro de Estudios e Investigaciones Tecnicas de Gipuzkoa (CEIT) and Tecnun (Universidad de Navarra) for their funding and support to N.Rodriguez.

REFERENCES

1. Loeb GE, Richmond FJR, Baker LL, *"BION Injectable Interfaces with Peripheral Nerves and Muscles"*, Neurosurgical Focus, vol 20 May 2006
2. Cameron T et al, "*Micromodular implants to provide electrical stimulation of paralyzed muscles and limbs"*, IEEE Trans Biomed Eng 1997; 44: 781-790
3. Q. Zou, W. Tan et all, "Single-axis and Tri-axis Piezoelectric Bimorph Accelerometer," IEEE/ASME Journal of Microelectromechanical Systems, in press
4. Tan, Wei and Loeb, G.E., "Feasibility of prosthetic posture sensing via injectable electronic modules," IEEE Trans. Neural Systems & Rehab. Engng., in press
5. Sachs, N.A. and Loeb, G.E. "Development of a BIONic muscle spindle for prosthetic proprioception," IEEE Trans. Biomedical Engineering, in press

Address of the corresponding author:

Author: Nuria Rodriguez
Institute: Alfred E. Mann Institute for Biomedical Engineering
phone: 213-821-2462; fax: 213-821-1120
Street: 1042 Downey Way
City: Los Angeles, California 90024
Country: USA
Email: nuriar@usc.edu

Minimizing Thermal Effects of In Vivo Body Sensors

Daniel Garrison

Virginia Polytechnic Institute and State University, College of Engineering, Virginia, USA

Abstract— **In sensor networks, energy availability is often viewed as a constraint in the long-term sustainability of the network. Duplicate message suppression and message aggregation are high-level layer approaches to reduce power consumption. Body sensor networks have similar constraints however the use of aggregator nodes may increase the heat dissipation in the tissue surrounding these nodes. Due to additional processing, this additional heat may raise the tissue temperature to a level damaging to the individual. We propose a model of data reporting that balances power consumption with body temperature thresholds. We present a generic algorithm along with implementation analysis via simulation.**

Keywords— **Data reporting, aggregation, thermal effects**

I. Introduction

Data aggregation has been studied as a way to reduce the density of data in sensor networks [3]. This reduction takes place at two levels. First, individual sensors (source nodes) reduce the transmission range required to report their data to a base station (sink). Data is rather reported to a cluster node in close proximity with that cluster node acting as an intermediate sink where data aggregation takes place. Second, through aggregation, this cluster head then acts on the behalf of the source nodes and reports abbreviated values to the base station.

It has been shown that transmission power on average uses more energy than an equivalent level of processing [5]. While this suggests that data aggregation is preferred, exceptions to this heuristic exist. One is when applications cannot tolerate a loss in data details often found through aggregation. Due to tissue sensitivity, another exception is in certain applications of body sensor networks that utilize implanted nodes. An explanation follows.

Body sensor networks use nodes to detect, record, and transmit data pertaining to physiological changes in an individual. Depending on the localization requirements, tissue that surrounds in vivo sensors may be sensitive to certain temperature increases [4]. Increases result from conduction and radiation emitted during sensor operations [1]. As a result, certain medical applications cannot utilize sensors.

Using data aggregation to increase a system's lifetime can further restrict the range of medical applications of biosensor networks. This is due to additional power density dissipation found in cluster heads when performing aggregation. The result of this added processing is a further increase in temperature over normal node operations. The additional increase makes more types of tissue susceptible to damage, thereby discounting the use of data aggregation.

In regards to an implanted node, tissue surrounding a node is generally heated in one of three ways: sensor circuitry power dissipation, sensor antenna radiation, and/or radio frequency (RF) charging radiation. All contribute to tissue heating irrespective of node aggregation. While we discuss each further, it is worth noting that RF charging is a method of energy replenishment. As such, energy conservation, while still important to the longevity between charges, becomes secondary to tissue preservation.

The motivation for our work is simple. We seek a method that prolongs intervals between the recharging of implanted sensors while preventing tissue damage that may result from power conservative aggregation. We propose a dual-mode method of data reporting that allows for data aggregation in body sensor networks. We suggest switching between data aggregation and independent reporting to prevent tissue damage in the individual. Ceasing aggregation for a period of time allows heat dissipation to occur more effectively in the tissue, thereby promoting cooling.

Our contributions are as follows:

An energy efficient method of data reporting. Through data aggregation, we allow for energy conservation while preventing tissue damage from occurring.

A scalable method of aggregate data reporting. We do not require a minimum number of nodes to prevent tissue damage, thereby affording aggregation in more applications.

A method that allows for optimal cluster head placement. We do not require cluster head rotation to dissipate heat from tissue, therefore allowing for optimal cluster head placement. This optimum can increase the quality of transmissions and thereby reduce temperature increases which would result from retransmission processing.

A method that affords better source node placement. As we do not require all nodes to serve as cluster heads, certain nodes may be able to be surgically implanted in more appropriate locations.

A generic algorithm that allows bi-level data reporting. We allow this to continue indefinitely while preventing tissue damage. An analysis is provided to determine specific values for switching between reporting methods.

The balance of the paper is organized as follows. Section 2 gives a description of related work. Section 3 identifies the system model. Section 4 introduces our algorithm along with the simulation setup environment. Section 5 provides a set of notable results. Section 6 concludes.

II. Related work

Of interest to our work in body sensor networks is work that addresses thermal effects of active sensors in the body. Gosalia, et. al [6] looked at the thermal effects of external radiation on the eye. While measuring temperature rises, possible sources were ignored as only one tissue heating factor was considered. We however consider multiple tissue heating contributors. Additionally, they use a continuous source of heat whereas we switch processing methods, thereby varying the levels and locations of heating.

Demarco, Lazzi, et. al [7,8] further investigated thermal elevations in retinal tissues, however their work considered only two factors leading to temperature increase. These include radiation caused by telemetry and also power dissipation from the circuitry. Unlike our work, they neglected a large contributor, namely RF charging.

Tang, et. al [1] present a model closest to our work. While mentioning the retina, they generalize their simulation to address homogeneously composed tissue. Our work makes a similar generalization. Their work addresses thermal effects of aggregation by rotating cluster heads. Similar to our proposal, they consider three factors in heating. They do not however accurately consider blood perfusion, a natural method of cooling. Additionally, their proposal does not allow sensor placement by temperature sensitive heterogeneous tissue, as all nodes must serve as cluster heads. Our implementation does not require rotation, thereby allowing for optimal placements of all sensors. In addition, for cooling purposes, they require more than the minimum number of nodes needed to form an aggregating group. Our solution allows for aggregation scaling down to minimal levels.

Concerning radio frequency radiation absorption, a few notable works focus on human tissue. Rui and Foster [11], and Moneda, et. al [10] used a electric dipole to measure the absorption levels of RF waves to tissue. Scott [12] looked at the temperature rises in the human eye by infrared (IR) radiation, an alternate form of sensor charging [5]. Each work again however considers only one method of tissue heating. While we do not consider (IR) charging as it is bound by a line of sight, we do use the Specific Absorption Rate (SAR) of radiation into tissues, as found in [9].

Finally, in regards to switching between modes of reporting data back to the base station, as we know, we are the first to investigate this form of tissue cooling.

III. System model

While our model allows sensor placement near heterogeneous tissue, we assume smooth tissue homogeneity for radiation measurement purposes. Power dissipation density is a measure of the power consumption of the sensor over the footprint of the sensor [2]. As power consumption varies between sensor hardware, we also assume an average power consumption level. We differentiate between power consumed by a cluster head node and a source biosensor.

While the temperature of the tissue could be measured by optional sensors added to a body sensor, we have chosen not to consider this for two reasons. First, is the added unit cost of these sensing capabilities. Second, is to keep the sensor footprint small to contribute less to tissue damage. Therefore, we use SAR to calculate levels of radiation received into the tissue per unit weight. SAR is measured in W/kg and is determined from the induced electric field by:

$$SAR = \frac{k\left|E_{rms}\right|^2}{\rho_t} \tag{1}$$

where E_{rms} is the rms amplitude of the induced electric field, k is the electrical conductivity of tissue, and ρ_t is the mass density of the tissue. This spatial distribution of deposited power measurement rises predominately during periods of RF charging. As many countries, including the United States, derive regulations on exposure to thermal radiation using this equation, it is pertinent for our work.

Radiation from antenna communication, either cluster head or base station, is assumed to use similar configurations as RF charging. A perpendicular dipole antenna is assumed with its effects on thermal heating being characterized through near and far field. Tissue close to the antenna is considered the near field, whereas outlying regions are considered the far field. While we assume consistent degradation of radiation absorption at increasing distances from the antenna, further derivation can be found in [17,18,19].

Most work assumes frequencies used for RF charging usually occur between 2MHz and 40MHz, with absorption rates varying as the square of frequency in this range [14,15]. Power is modeled as an incident plane wave with a percentage of the wave transmitted into tissue and the balance reflected off the body surface. The boundary between air and body is characterized by a difference in conductivity, density, specific heat, perfusion, and heat flux or power density, which we consider. In addition, we note that the power density degrades exponentially the deeper body penetration goes. Further derivation of the relationship between the incident and transmitted power density while considering the induced electric field is found in [16].

As the human body affords many body sensor network configurations, we limit our analysis to a two-dimensional control volume. This control volume is a small group of sensors which report to a single cluster head during aggregation. The cluster head is responsible for collecting this data, aggregating it, and transmitting it to the base station. We assume that nodes have been configured using a cluster-based approach; an approach that has demonstrated merit over tree-based configurations [13]. In addition, as we allow for optimal placement of nodes, this is also assumed.

Nodes are charged through RF waves directed toward the control volume region. While IR charging allows focused charging, we chose RF charging for its flexibility as waves can be received even when the biosensors are out of the line of sight of the emitter. Other forms of charging [5] do not produce enough energy to sustain biosensor operation.

An increase in the control volume temperature is the result of sensor circuitry powering, sensor antenna radiation, and RF charging radiation. To determine the distribution of this tissue heat, we consider Penne's bio-heat equation [20]:

$$\rho_t c \frac{\partial T}{\partial t} = \nabla \cdot k \nabla T + JE - Q_h \qquad (2)$$

where ρ_t is again the tissue mass density, c is the tissue specific heat, and $\partial T/\partial t$ is the partial derivative of temperature over time. Better said, it is the rate at which temperature increases in the control volume. Through $\nabla \cdot k \nabla T$ and its Laplacian operator, the thermal conductivity gradient of the tissue is introduced. J is the density of the current, while E refers to the electric field as a measure of its intensity. Finally, Q_h is the loss due to blood perfusion taking heat away from the tissue. Our work and analyses assume that sensors are implanted in both cartilage and muscle.

Algorithm 1 Source Node

```
1:  INPUT: aggregation_max (period until aggregated transmission); temperature_max (aggregation threshold*); temperature_min (cooling threshold*); time_max (temperature_max period from equation (3) using time and temperature_min*); time_min (temperature_min period from equation (3) using time and temperature_max*)
2:  interval = 0
3:  time_temperature_max = time_max + time
4:  do until time >= time_temperature_max
5:      Transmit individual sensed values to the cluster head
6:  loop
7:  time_temperature_min = time_min + time
8:  do until time >= time_temperature_min
        Transmit individual sensed values to base station
9:  loop
```

*values are application dependent

IV. Methodologies

Our algorithm for dual-mode reporting works as follows. For a period *time_max,* aggregation of data occurs as source sensor nodes send their data to the cluster head. At (or before) the end of period *time_max*, the cluster head transmits aggregated data to the base station. This occurs in intervals of *aggregation_max.* Once period *time_max*, has ended, all nodes enter time period *time_min.* During *time_min*, each source node transmits directly to the base station. The cluster head node acts as an ordinary source node. As nodes only need to transmit when a reportable event has occurred, thermal effects are minimized. This contrasts with the continuous processing that occurs at a cluster head node where the probability of receiving reportable data from at least one of its cluster nodes is higher. These processes are further detailed in Algorithms 1 and 2. Our algorithms are purposely generalized as particular implementations vary, as do their details. It should be noted that sensor node charging occurs during both modes at unknown intervals as this action is external to the system.

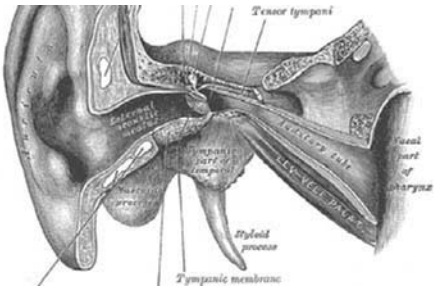
To solve the steady-state and transient two-dimensional heat transfer problem in biological tissue, we extend SAR and Penne's bio-heat equation. To measure distribution and rates of change in tissue temperature throughout the control volume, we discretized these measurements to the Cartesian coordinate system. These manipulations produced the following equation which is also applicable to our simulation:

$$\frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y}\right)=wc_b(T-T_b)-\dot{q}+\rho_t c\frac{\partial T}{\partial t} \qquad (3)$$

Algorithm 2 Cluster Head Node

```
1:  INPUT: aggregation_max`; temperature_max`; temperature_min`; time_max`; time_min`
2:  interval = 0
3:  time_temperature_max = time_max + time
4:  do until time >= time_temperature_max
5:      do until interval >= aggregation_max
6:          Receive sensed values from source nodes
            if time >= time_temperature_max then Exit Loop
            'Premature exit if aggregation time exceeds
            increment interval
7:      loop
8:      Aggregate values
9:      Transmit aggregate values to base station
10: loop
11: time_temperature_min = time_min + time
12: do until time >= time_temperature_min
13:     Transmit individual sensed values to base station
14: loop
```

`value descriptor found in Algorithm 1



[28]

Fig 1. Physical representation of the experimental area.

where T is the tissue temperature while T_b is the blood temperature. Through $\dot{q}$ we represent the local metabolic heat generation rate per unit volume. Again we denote the tissue mass density by ρ_t, and c as the tissue specific heat. As one experimental set considers the thermal effects of capillary blood through sensor implantation in muscle, we also account for this. We define the variable Q_h in Penne's bioheat equation by representing perfusion as w with the specific heat of blood as c_b. Thermal conductivity is denoted as k while t represents time in seconds. Cartesian coordinates are naturally denoted as x and y with partial derivatives being performed throughout the equation. Finally, to this equation we introduce the SAR value of RF charging.

While our solution is applicable beyond our simulation case, for analysis sake, we limit our work to the use of sensors implanted in the outer and middle human ear (Figure 1). Limited work has been done to address non-functioning parts of a human's hearing [22]. Our simulation discretized a test bed to sixty cells, measuring temperature at 39 points in the tissue and at sensors (Figure 2). Sensors are implanted subcutaneously in the dorsal wall of the external acoustic meatus near the tympanic membrane, also known as the ear canal. As the absence of blood vessels in the cartilage of the canal prevents blood perfusion, a secondary implantation was also simulated. This implantation occurs in the middle ear and attaches to the tensor tympani muscle. Muscle naturally allows for blood perfusion, thereby giving sets of simulations that can be contrasted. Besides blood perfusion, we assume similar tissue makeup for our experiments and use the values in Table 1. Of note is that due to the location of our simulation we do not model metabolic activity as it is almost non-existent in the region of the ear.

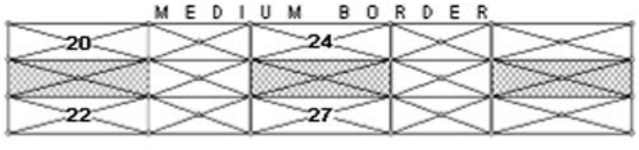


Fig 2. Sensor placement within the subcutaneous tissue with temperature discretization matrix and various temperature measurement points.

Table 1. Experimental values for inner ear and outer ear tissue.

Variable	Value (outer ear)	Value (inner ear)	Units
k	0.54 [24,25]	0.54 [24,25]	W/mK
ρ_t	1050 [24,25]	1050 [24,25]	kg/m^3
T	37.0	37.0	˚C
c	3700 [24,25]	3700 [24,25]	J/kgK
w	1.0	8.3 [23]	kg/m^{3}s
$\dot{q}$	0.0	0.0	W/m^3
T_b	37.0	37.0	˚C
c_b	3800 [26]	3800 [26]	J/kgK

The ear canal is roughly 26mm in length whereas the tensor tympani muscle averages 20mm. As the implant preference is near the tympanic membrane, we focus our simulation test bed on the inner 20mm of the canal. Past work has involved a piezoelectro transducer attached to the middle ear ossicle where sensors interface directly with the brain while recharging though induction. We align three sensors linearly with the middle node periodically acting as cluster head and the top medium being the skin surface (Figure 2).

We chose this experimental setting due to its appropriateness and to extend work previously done. However direct application to the work done in this area is difficult as attachment may still be required from the outer ear to the inner ear ossicle, possibly leading to irreversible damage. Our secondary approach remedies this by moving the complete application to the middle ear. This location bypasses dealing with the tympanic membrane which separates these regions. In addition, by using this experimental approach we can introduce blood cooling as we model tensor tympani muscle attachment. Further, RF charging is shown by placing the charging probe in the canal next to the sensors.

Our simulation test bed is realistic while novel. It reduces the invasiveness of current solutions and limits transmission power due to aggregation. Additionally, our simulation addresses the criticality of temperature increases due to sensors. This criticality comes as sensors placed in this area are close to the brain and more importantly the hypothalamus. As the hypothalamus regulates body temperature, heating of the hypothalamus can result in an individual becoming drowsy and ultimately experiencing hypothermia [21]. As part of our analysis of temperature rise and fall, we set the initial temperatures of the tissue and blood to 37˚C.

We model nodes placed 1mm below the surface of the tissue in both experimental test beds. While sensor material makeup varies, for our simulation we use sensors housed in an aluminum casing [1,2,27]. These sensors, capsule in shape, extend 4.5mm in length [22] and 1mm in height. Due to differences between tissue makeup and foreign materials like sensors, we adapted the earlier equation (3) to apply to the heat transfer within a sensor, again discretized to the Cartesian coordinates:

$$\frac{\partial}{\partial x}\left(k\frac{\partial T}{\partial x}\right)+\frac{\partial}{\partial y}\left(k\frac{\partial T}{\partial y}\right)=\rho c\frac{\partial T}{\partial t}-\dot{q} \qquad (4)$$

where T is the temperature of the sensor; assumed to be the same as the surrounding tissue of implantation. Time and the thermal conductivity are once again depicted by t and k, respectively. The density of the aluminum is represented by ρ, and the metal's specific heat by c. Partial differentiation is performed on different values including the x and y coordinates to determine sensor temperature change rates.

These rates of change are determined while considering the sensor boundaries with either the muscle or cartilage. We assume that non-cluster heads in this small configuration consume roughly one-fifth the power needed by cluster heads in both circuitry powering and communication. Except for the boundary between the air and tissue, we assume there is no heat flux between the test bed and the surrounding tissue. This is logical as the borders of our test bed do not signify the border of the tissue, therefore are adiabatic.

V. Analysis of Results

As noticed in our results, the application of sensors with RF charging has the potential of substantially heating its surrounding tissue in the outer ear. We saw that subcutaneously implanting the sensors in cartilage causes temperature to rise up to 1.58˚C. Contributing to this increase is aggregation and the continuous RF charging during the 6000 seconds needed for temperature leveling. By turning off aggregation, and relying on individual reporting during continuous powering, we found a maximum temperature rise of 1.33˚C with leveling again at 6000 seconds.

When implanting the sensors in the middle ear, namely the tensor tympani muscle where blood perfusion occurs, the maximum temperature was first seen 610 seconds into the heating cycle. There we found a 0.21˚C increase in the tissue temperature. These measurements again model consistent RF charging and aggregation. Removing aggregation, resulted in a decrease in temperature by 0.04˚C to 0.17˚C. So, from these results we see that our alternate placement of hearing assistance sensors in the middle ear results in a dramatic decrease in tissue temperature rise.

It should be noted that 39 reading points were chosen to determine the effects of RF charging over the medium border while considering the role that the sensors played both in their physical obstruction and circuitry power dissipation and antenna radiation. Therefore, temperature differences observed correspond to their locations within the test bed with reading points closest to the cluster head and RF charging source receiving the highest increase. The lowest temperature points were those located farther from the RF charging source where they were shielded by other sensors.

From our calculations we determined *time_max* based upon *temperature_max*. While temperature thresholds are widely varied, we chose the point of leveling to prevent prolonged exposure. We determined the *time_max* for the worse case scenario which corresponds to when aggregation and RF charging occurred simultaneously. This scenario was used as the prediction of RF charging is impossible.

For the outer ear simulation, threshold temperature occurred 5500 seconds into the clustering cycle at node 24, located on the medium border above the cluster head (Figure 2). Therefore we set *time_max* for a *temperature_max* of 1.58˚C to 5500 seconds for ear cartilage implementations. For the middle ear test bed, this again occurred at node 24, so for our algorithm we set *time_max* to 660 seconds when *temperature_max* is set to 0.21˚C.

According to our algorithm, when we reach *time_max,* we then switch to independent reporting mode. This allows the tissue to cool. During this new time period, while non-aggregating nodes now emit more radiation due to transmissions with the base station, cooling of the tissue was still observed. As expected, cooling of the tissue took longer in cartilage than muscle as no blood perfusion exists.

As data reporting is still occurring during this phase, complete cooling of the tissue to initialization values is not expected nor was observed. In the ear cartilage, normal independent reporting temperatures level at 37.20˚C while in muscle with blood perfusion, leveling is at 37.02˚C. As these are the lowest temperatures expected, we set the *temperature_min* to their corresponding leveling temperatures. With these values set in our algorithm, it is important to determine *time_min* for each of these implantation locations. In order to prevent incremental temperature increases, we choose this when all reading points reached their corresponding *temperature_min.* For the outer ear environment, this occurred 5400 seconds into this cycle of non-aggregate reporting. Therefore for ear cartilage implementations, *time_min* is set to 5400 seconds. The 5400 seconds began after the 5500 seconds of aggregation. In inner ear environments when attaching to muscle, *time_min* was calculated and set to 700 seconds; the *time_min* value.

From our simulation, we saw that our dual mode of reporting works well by allowing energy savings during aggregation while maintaining tissue temperature in a safe range. From the determined and calculated temperature threshold values and calculated time threshold values, our algorithm can be implemented on sensors within the ears.

Finally, when RF charging was absent in cartilage during aggregation, we saw temperature leveling at 37.45˚C in most regions of the tissue. From this, we conclude that aggregation contributed to a 0.45˚C increase in ear cartilage

sensor implantations. We found similar trends but to a lesser extent in the muscle, having aggregative processes contribute 0.04˚C to the muscle tissue temperature increase.

VI. Conclusion

From our work we have shown our dual-mode of data reporting as a viable option in the design of a biosensor system. Through aggregation and its energy saving nature, we have been able to increase the time between node charges and thereby increase the application space for biosensor deployments. Further, we have demonstrated worse-case temperature heating scenarios in the use of a biosensor, a factor critical to adequately addressing tissue damage. We demonstrated applicability and savings for networks of small size, thereby demonstrating a solution, unlike others, that scales well. In addition, because of the nature of our generic algorithm, we allow for the optimal placement of sensors and cluster heads while achieving these savings. Finally, through a series of experiments we demonstrated the usefulness, benefit and gentleness of our generic algorithm to many applications in body sensor networks.

References

1. Tang Q, Tummala N, Gupta S.K.S, Schwiebert S (2005) Communication Scheduling to Minimize Thermal Effects of Implanted Biosensor Networks in Homogeneous Tissue, IEEE Transactions Biomedical Engineering
2. Tang Q, Tummala N, Gupta S.K.S, Schwiebert S (2005) TARA: Thermal-Aware Routing Algorithm for Implanted Sensor Networks, International Conference on Distributed Computing in Sensor Systems (DCOSS)
3. Heidemann J, Silva F, Intanagonwiwat C, Govindan R, Estrin D, Ganesan, D (2001) Building Efficient Wireless Sensor Networks with Low-Level Naming, 18th ACM Symposium on Operating Systems Principles
4. Hirata A, Matsuyama S.I, Shiozawa T (2000) Temperature rises in the human eye exposed to EM waves in the frequency range 0.6-6 GHz, IEEE Transactions on Electromagnetic Compatibility, (42)4:386
5. Schwiebert L, Gupta S.K.S, Weinmann J (2001) Research Challenges in Wireless Networks of Biomedical Sensors, In Proceedings of the 7th ACM MOBICOM
6. Gosalia K, Weiland J, Humayun M, Lazzi G (2004) Thermal Elevation in the Human Eye and Head Due to the Operation of a Retinal Prosthesis, IEEE Transactions on Biomedical Engineering, 51(8): 1469
7. DeMarco S.C, Lazzi G, Wentai L, Weiland J.D, Humayun M.S (2003) Computed SAR and Thermal Elevation in a 0.25-mm 2-D Model of the Human Eye and Head in Response to an Implanted Retinal Stimulator - Part I: Models and Methods, IEEE Transactions on Antennas and Propagation, 51(9):2274
8. Lazzi G, DeMarco S.C, Wentai L, Weiland J.D, and Humayun M.S (2003) Computed SAR and Thermal Elevation in a 0.25-mm 2-D Model of the Human Eye and Head in Response to an Implanted Retinal Stimulator - Part II: Results and Methods, IEEE Transactions on Antennas and Propagation, 51(9):2286
9. International Electrotechnical Commission (IEC): Medical Electrical Equipment, Part 2-33: Particular Requirement for the Safety of Magnetic Resonance Systems for Medical Diagnosis. IEC 60601-2-33 Ed.2, 2002.
10. Moneda A.P, Ioannidou M.P, and Chrissoulidis D.P (2003) Radiowave Exposure of the Human Head: Analytical Study Based on a Versatile Eccentric Spheres Model Including a Brain Core and a Pair of Eyeballs, IEEE Transactions on Biomedical Engineering, 50(6):667
11. Riu P.J, Foster K.R (1999) Heating of Tissue by Near-field Exposure to a Dipole: A Model Analysis, IEEE Transactions on Biomedical Engineering, 46(8):911
12. Scott J.A (1988) The Computation of Temperature Rises in the Human Eye Induced by Infrared Radiation, Physics in Medicine and Biology, 33:243
13. Shankar V, Natarajan A, Gupta S.K.S, Schwiebert L (2001) Energy-efficient Protocols for Wireless Communication in Biosensor Networks, In Proceedings of the 12th IEEE International Symposium on Personal, Indoor and Mobile Radio Communication, October
14. Heetderks W.J (1998) RF powering of millimeter- and submillimeter-sized neural prosthetic implants, IEEE Transactions on Biomedical Engineering, 35(4):323
15. Mokwa W, Schnakenberg U (2001) Micro-transponder systems for medical applications, IEEE Transactions on Instrumentation and Measurement, 50(6):1551
16. Ulaby F.T (1999) Fundamentals of Applied Electromagnetics. Prentice Hall
17. "A Practical Guide to the Determination of Human Exposure to Radiofrequency Fields", NCRP Report No. 119, (1993).
18. Office of Engineering Technology (1993) Understanding The FCC Regulations for Low-Power, Non-Licensed Transmitters In OET BULLETIN NO. 63, October
19. Kuster N, Balzano Q (1992) Energy Absorption Mechanism by Biological Bodies in the Near Field of Dipole Antennas Above 300 MHz, IEEE Transactions on Vehicular Technology, 41(1)
20. Pennes H.H (1948) Analysis of Tissue and Arterial Blood Temperatures in the Resting Forearm, Journal of Applied Physiology, 1:93
21. Rechtschaffen A, Siegel J.M (2000) Sleep and Dreaming, In Principles of Neuroscience: Fourth Edition, Kandel, E.R., Schwartz, J.R., and Jessel, T.M., eds., 936-947, McGraw-Hill, New York
22. Zenner H, Leysieffer H (1988) Totally Implantable Hearing Device for Sensorineural Hearing Loss, The Lancet, 352(9142):1751
23. Cain C.A, Umemura S (1986) Concentric-Ring and Sector-Vortex Phased-Array Applicators for Ultrasound Hyperthermia, IEEE Transactions on Microwave Theory and Techniques, 34(5):542
24. Sekins K.M, Emery A.F (1982) Thermal Science for Physical Medicine, Therapeutic Heat and Cold, Lehmann, J.F., ed., 81-85, Williams & Wilkuns
25. Giering K, Minet O, Lamprecht I, Müller G (1995) Review of Thermal Properties of Biological Tissues, Laser-induced Interstitial Thermotherapy Müller, G.J., Roggan, A., eds., 45-65, SPIE-The International Society for Optical Engineering
26. Hodson D.A, Barbenel J.C, and Eason G (1989) Modelling Transient Heat Transfer Through the Skin and a Contact Material, Physics in Medicine and Biology, 34:1493
27. Leland J, Recktenwald G (2003) Optimization of a Phase Change Sink for Extreme Environments, In Proceedings of the Nineteenth Annual IEEE Semiconductor Thermal Measurement and Management Symposium
28. Gray H (1918) Anatomy of the Human Body. Lea & Febiger

Subject Index

G

H

I

L

M

N

O

P

Q

R

S

T

W

Z

Authors Index

G

H

I

J

K

L

M

N

O

P

R

S

T