Albert-Ludwigs-Universität Freiburg
Embedded System Engineering

Wearable Computing Systems

Summer-Term 2015

Seminar Course

Prof. Dr. Kristof Van Laerhoven
Dipl. Inf. Philipp Marcel Scholl

Preface

This document presents the summary of the Wearable Computing Systems Seminar held in the summer term 2015 at the Albert-Ludwigs-University Freiburg. The seminal papers created by the participating students summarize the topics of each group.

An important issue for wearable computing systems is their input modality. The first topic therefore treats possible sensors for Human Activity Recognition, looking at their measurement principles, as well as their wearability and how these can be objectively evaluated. Processing the acquired sensor values with different feature set was the next topic, which was found to be very application-specific and a number of applications are accordingly presented. The detection of Human Activities can be enhanced when additional model knowledge is brought into the system. Another topic treats different ways of creating computable human body models to achieve this.

The fundamentals of wearable computing systems was another topic for this seminar. Most influential papers from the last twenty years for wearable computing were summarized. The application of computing systems as support tools in wetlabs was selected as one topic, since it presents an application area with stringent requirements. Wearable Computing Systems can also stretch into computing system for animals, which is why two groups looked into the application of monitoring birds in the field.

We would like to thank all the participanting students for their highly motivated work during the seminar course, and we are proud of the generated results.

- The Editors, 15. Oct 2015

Table of Contents

Marc Pfeifer und Benjamin Völker
Sensors in Human Activity Recognition 1

Sebastian Böttcher und Mireille Rehse
Feature Set in Activity Recognition 9

Ilia Dobrusin, Paul Gavrikov und Murat Killi
Human Body Models 16

Tobias Paxian und Martin Dold
Wearable Computing System Fundamentals 24

Jan Donath, Johannes Güttler und Florian Wolling
Wetlab Support Systems 31

Andreea Oltean und Cosmin Zvariste
An Analysis of Wireless Sensor Networks for Bird Monitoring 36

Alexander Richter, Philip Klein und Florian Tautz
Bird Monitoring by Wireless Networks 42

Sensors in Human Activity Recognition

Marc Pfeifer
Albert Ludwigs University
Freiburg, Department of Computer Science
Email: pfeiferm@informatik.uni-freiburg.de

Benjamin Voelker
Albert Ludwigs University
Freiburg, Department of Computer Science
Email: voelkerb@informatik.uni-freiburg.de

Abstract*—Human Activity Recognition (HAR) describes the automatic recognition of gestures, movements and activities of a human with an embedded or non-embedded computer system. Research in this area started in the middle of the 1990's with the introduction of MEMS sensors which made it possible to directly equip a human body with the needed sensor-system. The sensors can be seen as the main interface between the human body and wearable computers and are therefore the first stage in the Human Activity Recognition flow. The selection of the appropriate sensors directly influence the accuracy of the whole system as well as its size, obtrusiveness and energy-consumption. Due to this importance, this paper mainly focuses on the advantages and disadvantages of different sensors used in HAR, how they evolve over time and what needs to be considered to implement a practical HAR system in a wearable context. The most used sensors and their working principles are briefly explained in the context of HAR-systems. The evolution of sensors in HAR-systems from the first simple and self-made sensors-system to modern smartphones with high accurate sensors is discussed as well as the wearability of a sensor- or HAR-system. Appropriate measures for obtrusiveness are presented and the latest topic of social wearability is discussed. Different body positions to place sensors are closely related to obtrusiveness and social wearability. Even if it is not possibly to find a single perfectly matching body position different results of current research work is shown. Another very common but not well defined term with respect to sensors and HAR-systems is modality. Therefore the most reasonable definition and a corresponding work in which a gyroscope is replaced by a magnetometer is presented. Finally different sensor sampling rates and signal resolutions are analysed due to their influence on the recognition accuracy and an optimal value for each parameter is proposed. With the given content the paper provides an introduction in the field of sensors so the fellow reader will be able to choose the appropriate sensors for a given application.

I. Introduction

The recognition of human activities has become a task of high interest especially for medical-, military-, security- or sport applications. For example people with dementia or diabetes need to follow a predefined daily routine to get used to their medical condition. Consequently recognizing activities like sitting, walking or sleeping can give useful feedback about the person's behaviour by spotting abnormal activities.

The activity recognition process can be performed either with external or wearable sensing [1]. External sensing was state of the art in the middle of the 1990's. One example for such an external sensor platform is a smart house that is capable of recognizing a lot of complex activities with external sensors like visual cameras or proximity sensors. The drawbacks of external sensors are that they are usually expensive and require a lot of maintenance. Furthermore the recognition is restricted to the area where the sensor is attached and thus require the user to stay in this area to recognize his or her activities. Wearable sensing was mainly initiated by the improvements of MEMS (Micro-Electrical-Mechanical-System) sensors at the end of the 1990's. These small sensors made it possible to turn the task of activity recognition into a wearable computing problem by equipping the human body with a sensor system. Figure 1 shows the flow traversed to recognize an activity with a wearable sensing platform. At first raw data needs to be collected from the available sensors. Relevant information is highlighted in the raw data set and extracted into a feature set. This feature set is evaluated online or offline in a classification algorithm that was trained supervised or semi-supervised. The classification algorithm finally generates a label representing the currently performed activity.

Figure 1 shows the importance of sensors for the task of activity recognition since data collection is the initial stage in the recognition flow. The selection and placement of the sensors influence many important parameters like the quality and quantity of the collected data, the prediction accuracy, the energy consumption and finally the cost and wearability of the whole system. To pick up this problem, three key questions were previously defined and are addressed in this paper. Which sensor is suitable for a given application? What are the drawbacks of a given sensor and how to handle them? Up to which limit are already available sensors (e.g. in smartphones) sufficient?

To address these questions a lot of research was done by looking at many different papers concerning the general definition of human activity recognition. Since only a few papers focus solely on the topic "sensors", a lot of papers where reviewed in parts and selected papers were examined for interesting fields. As a starting point [1] was selected because it provides a good overview of the whole topic. From the citations in this work other suitable works were selected and discovered afterwards. The search-keyword "activity recognition" was used and combined with specific key words like "sampling rate". The conference *International Symposium on Wearable Computers* was mainly used as a paper-source but other conferences were also included if matching papers were found. From the large amount of information the above stated key questions are answered by the following structure: An overview of sensors that are already used in HAR-systems is given together with their working principle in Section II. Section III explains how sensors and their application changed in the field of activity recognition during the past two decades. Afterwards in Section IV the term obtrusiveness is defined and its importance for a practical HAR-system is discussed.

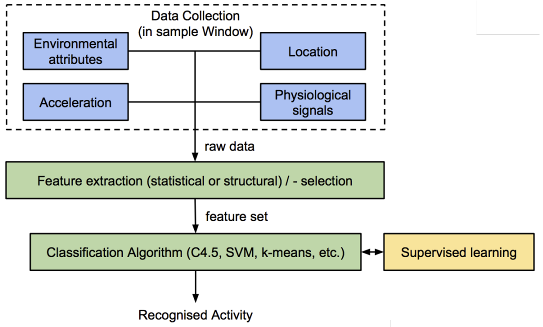


Fig. 1. Overview of the Human Activity Recognition (HAR) Flow, reprinted from [1].

Section V deals with different possibilities to measure the same physical quantity. The influence of different sampling rates and signal resolutions on the recognition accuracy is finally discussed in Section VI followed by a conclusion and summary of the work.

II. Attribute Groups and Sensors in HAR-Systems

There is a wide range of different sensors which are used in Human Activity Recognition Systems. A few examples are accelerometers, photodiodes, pressure-sensors, GPS-receiver and heart rate sensors. All these sensors measure different physical effects, but with respect to HAR they could be combined into groups, which describe a certain measured attribute that is used for the recognition. Different suggestions for the definition of these groups exist but in general these attribute groups could be defined as follows (cf. [1]): inertial sensors, environmental attributes, location and physiological signals. In the following the four groups are explained in detail and the most important sensors of each group are introduced.

A. Inertial sensors

The first and most relevant attribute group are inertial sensors which describe movements, twists and turns of the human body. The most used sensors in this attribute group are the accelerometers. They are used in nearly every HAR-system and the measured acceleration is usually the key-data for the recognition. Examples can be found in [2], [3] and [4]. Reasons for this high popularity are the low price, the relative low energy consumption compared to other sensors and the already high availability for instance in modern smartphones [1].

Figure 2 shows the working principle of a general accelerometer. The sensor consists of a seismic mass m (red) which is hold at its neutral position e.g. with springs. During an external acceleration the mass is deflected from its neutral position. The amount of deflection can be determined for example capacitively and is proportional to the force F. Finally the acceleration can be calculated using the *equivalence principle*:

$$F = m \cdot a$$

To determine the true acceleration of the sensor, the earth acceleration has to be considered or removed in the result. However in HAR-systems this is often omitted, since the earth acceleration can be used to get information about the sensor

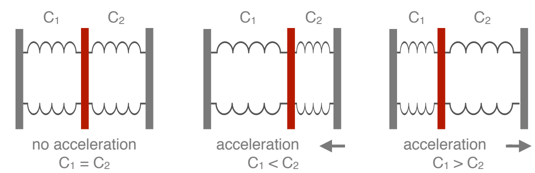


Fig. 2. Working principle of an accelerometer.

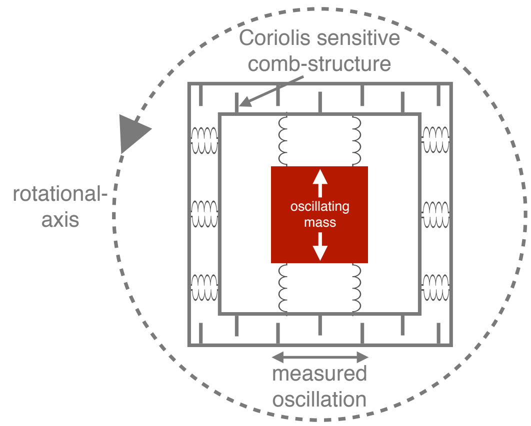


Fig. 3. Working principle of a gyroscope.

position, which is one of the key-data for the recognition procedure. Today's accelerometers are usually produced in MEMS technology and offer three sensitive axis and a digital output. Beside the already mentioned advantages accelerometers are small and have a simple working principle which make them very reliable. The main disadvantage is that the sensor has to be calibrated to its local area since the earth acceleration is not homogeneous and can change with the global location.

An accelerometer is often combined with a gyroscope, which represents the second most used sensor of the inertial sensors attribute group. [3]. Gyroscopes measure the angular velocity of the local coordinate system. Angular velocities hold much information about the currently performed activity.

Figure 3 shows the working principle of a gyroscope. It measures the deflection of an oscillating seismic mass due to the Coriolis force. These sensors are typically build in MEMS technology and the deflection is often measured capacitively. The precision of the MEMS process leads to small form factors and sensors that measure the angular velocity with a precision of one tenth of a degree [5]. A drawback of gyroscopes is, that they only detect angular changes and must be calibrated with a reference point. To avoid drift, they also needs to be constantly corrected or recalibrated.

A third sensor-type which fits in the inertial sensors attribute group is the magnetometer. Magnetometers are used in HAR-Systems to support other sensors e.g. accelerometers [3] or to replace other sensors completely [6].

A magnetometer measures vector components of the magnetic field in its surrounding. The working principle is shown in Figure 4. The deflection of a beam due to the Lorenz force is usually measured piezoresistive or capacitive. Magnetometer

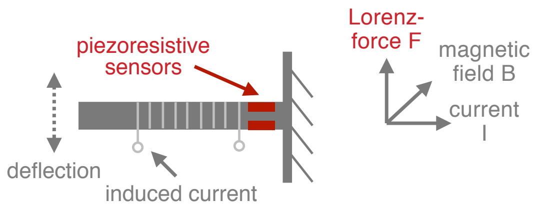


Fig. 4. Working principle of a magnetometer.

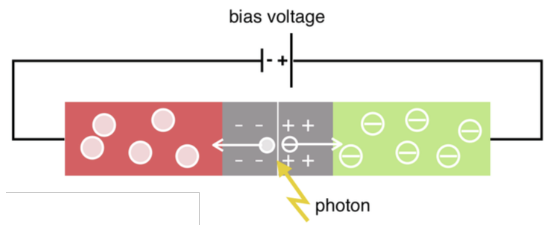


Fig. 5. Working principle of a photodiode.

can be used to measure the earth magnetic field and thus to obtain the current azimuth heading. The azimuth heading is absolute in the earth coordinate system and thus could be used as a reference for other sensors. Unfortunately magnetometers are prone to environmental magnetic fields and need to be calibrated to the local magnetic field.

B. Environmental attributes

The second attribute group describes information about the surrounding of a person. These information are typically combined with acceleration data to further distinguish activities. If the surrounding is for example dark and the accelerometer indicate no movement, the person could be sleeping. If the surrounding is bright and no movement is detected, the person could also rest in the park.

To get the mentioned information about the brightness of the surrounding, light sensors can be used which exist in different types like photodiodes, photoresistors, phototransistors or solar cells. One system that includes a light sensor can e.g. be found in [4]. The working principle of a photodiode is shown in Figure 5. Incoming photons create electron-hole pairs in the depletion zone of the diode due to the photoelectric effect. These pairs are instantly separated because of the charged areas which cause a flowing current. Although light sensors are small and inexpensive, they can easily produce misleading data if they are covered.

A second example for environmental sensors are pressure sensors which can be used to get relative altitude information. Figure 6 shows the working principle of a piezoresistive air-pressure sensor. Strain in the membrane, which is related to the pressure difference of the current air-pressure and the pressure p_0 inside the cavity, is measured either piezoresistive, piezoelectric or capacitive. This simple concept has the drawback that only large changes in height (more than $1m$) can be measured. However this information is often enough to recognize ascending, descending or elevator movements [4].

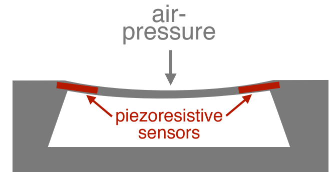


Fig. 6. Working principle of a pressure sensor.

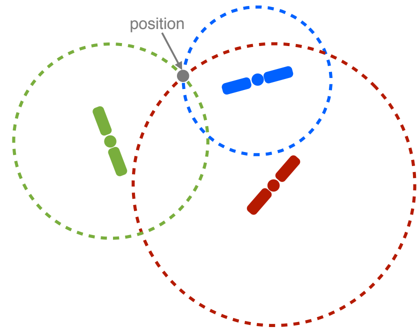


Fig. 7. Working principle of the Global Positioning Sytem.

C. Location

The third attribute group gives information about the location of the person. Location data can also support the activity recognition process since it is for example unlikely for a person to brush his teeth on the street. Travel distances and speed can be calculated as well if multiple location data with timestamps are available.

The most prominent sensors in this group are GPS receivers, which provide the global position. Therefore GPS satellites consistently send their positions and the current time. If the GPS receiver receives at least data of three different satellites, it can calculate the absolute distances to each of the satellites with a technique called Time Of Arrival (TOA). Triangulation as shown in Figure 7 is finally used to calculate the global position of the GPS receiver. The calculated position data is absolute and could be used as reference for other sensors. The major drawbacks of GPS devices are their energy consumption, the high price and that they only work outdoor.

D. Physiological signals

The last of the four attribute groups describes the vital and body parameters of a human. They are also often used to support other sensors and to improve the recognition accuracy. One example for vital parameters is a persons heart rate. If the person is moving a lot with a high heart rate the probability that the person is doing sports is very high. If on the opposite the heart rate is low while another sensor senses much motion the person could be for example standing on a boat.

The working principle of a heart rate sensor could be seen in Figure 8. The sensor employs the so called electrocardiography. Small electrical changes on the skin, which are produced by the heart muscle, are measured with special electrodes. Afterwards the current heart rate is derived from these electrical changes. The biggest advantage of this method is, that it results raw body data which is very accurate.

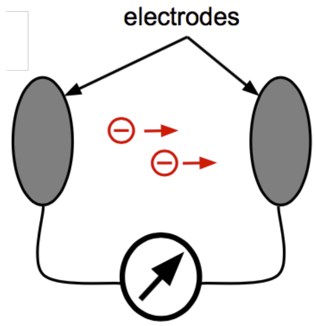


Fig. 8. Working principle of a heart rate sensor.

Fig. 9. On the left: MEMS-Accelerometer *Analog Devices ADXL202* [16]. On the right: MEMS-Gyroscope *Murata ENC-03-J* [17].

However the heart rate is only slowly adjusting to the currently performed activity. This means that it can for example still be high even if the person is not doing sports any more. Nevertheless heart rate sensors are used in some works like [7] and [8].

The above explained sensors are the most relevant ones in each group. A lot of research in Human Activity Recognition also focus on other sensor types or even self-made sensors. To mention only a few: vibration sensors are used in [9], microphones are used in [10] and force sensors are used in [11].

III. Evolution of sensors in HAR-Systems

First works in the research area of Human Activity Recognition came up in the middle of the 1980's. Since there were no small MEMS sensors available in the first years, HAR-systems often consist of stationary, non-wearable components like cameras or motion capture systems that only cover a certain area. Examples for such systems are the tennis action recognition done in [12] or the activity recognition done in [13].

The introduction of small MEMS accelerometers and gyroscopes (like shown in Figure 9) in middle of the 1990's made it possible to directly equip a human body with different sensors and thus to build wearable HAR-systems. Because of this new opportunity and its many advantages, the research area of HAR had a first hype during this time period. Many new HAR-systems were introduced, which were at this time often self-build and consisted of early MEMS accelerometers and gyroscopes with only one sensing axis and an analogue output. Examples for such systems could be found in [14] and [15]. As time elapsed the development of sensors continued along with HAR-systems. While accelerometers and gyroscopes were often equipped with two sensing axis and analogue outputs in the middle of the 2000's, nearly every system since 2010 uses sensors with at least three sensing axis, a digital output and high accuracy.

Considering the whole research area of HAR some change over time could also be observed. From the beginning of the hype in the middle/end of the 1990's until around 2004 many works focus around newly, completely self-build HAR-systems. After 2004 the number of this works decreases. Research focused more and more on the feature extraction and classifiers. To test the results of this works already existing sensors-platforms or primarily recorded data-sets were often used. An example for such an of-the-shelf sensor-platform is the *XSens* [5] which consists among other sensor of an accelerometer, a gyroscope and a magnetometer. One example where this platform is successfully used is [18]. In this work the authors attached respectively six *XSens* modules on eight different test persons and tried to recognize 19 different activities. With this setup and a Naive Bayes Classifier a recognition accuracy of around 99.2% was achieved. Since the year 2010 the HAR research area had a second big hype. This was caused by the starting currency of modern smartphones which are already equipped with accelerometers and often many other sensors. The fact that no extra and often obtrusive hardware must be carried to get data for HAR caused the development of many new systems which use smartphones as their primary sensors. One example is [19]. Here the authors tried to recognize exercise activities with a smartphone worn in an armholster on the upper arm. In a person independent classification with a k-Nearest-Neighbour Classifier they achieved an accuracy up to 93.6%, which clearly shows that there is much potential with this new technology. Underlying this trend also many new interesting topics arose like the development of systems which don't require the smartphone to be worn at a specific position on the body. Research in this topic is still ongoing and new technologies like smartwatches bring new possibilities and problems to solve.

IV. Obtrusiveness

While the recognition accuracy of HAR-Systems gets more and more improved with time, further research is also going towards the feasibility of HAR-Systems. A feasible HAR-System should not require its user to wear many uncomfortable sensors, nor should it require to interact too often with the system [1]. If these aspects are not fulfilled it is not or only hardly possible to make people use the system outside a lab-environment. The term "obtrusiveness" combines the different dimensions of discomfort and invasiveness with respect to a wearable HAR-system. The miniaturisation caused by technical progress has a huge impact on the invasiveness of HAR-systems. For instance, in the beginning of the 21st century people had to carry huge notebooks in backpacks to have enough mobile computational power for the recognition process ([14], [3], [20]). With the invention of smartphones and smartwatches backpacks could be replaced by trouser pockets [21] or armholster [19] resulting in more practicable HAR-systems that are less invasive.

6	No exertion at all
7	Extremely light
8	
9	Very light
10	
11	Light
12	
13	Somewhat hard
14	
15	Hard (heavy)
16	
17	Very hard
18	
19	Extremely hard
20	Maximal exertion

Fig. 10. The scale which is used to determine the exertion via the *Borg Relative Perceived Exertion* method [22].

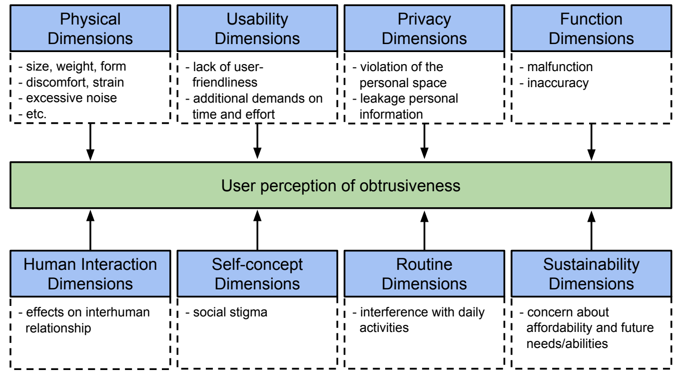


Fig. 11. The different dimensions of obtrusiveness in our society, reprinted from [23].

A. Obtrusiveness measures

To compare different HAR-systems also in terms of obtrusiveness a standardized measure would be useful. Unfortunately there is not much research on this topic. One reason for this is probably that the perceived obtrusiveness is always very subjective and defining an objective measure is difficult. However the authors in [22] tried to come up with such a measure. They state three main effects which influence the obtrusiveness: The physiological effects which describe the energy that is needed to e.g. carry the HAR-system on the body, the biomechanical effects which describes the direct influence on body like musculoskeletal loading and the comfort effects which describe the feeling of well-being while using a HAR-system. Each of this effect could be measured precisely with complex and expensive methods but instead of using these, the authors introduced very cheap and simple, still standardized methods. This was done to provide a measure that could be used in field without much effort. The physiological effect could be for example simply measured with a method called *Borg Relative Perceived Exertion (RPE)*. In this method test persons wearing the HAR-system rate their exertion according to a standardized scale shown in Figure 10. Similar scale-methods are introduced to measure the biomechanical- and comfort-effects, whereas the biomechanical effects are further divided into “posture and movement” and “musculoskeletal loading”. With the received score for each effect and a provided table, a so called “Level of Effect” could be obtained. This could afterwards be translated into an level of obtrusiveness with a second table. A first field test of the authors on an existing wearable HAR-systems showed good results for measure.

B. Dimensions of obtrusiveness

Like already mentioned above the term of obtrusiveness does not only describe the level of discomfort of a HAR-systems. Instead it has many different dimensions which are summarized in [23] and shown in figure 11. The physical dimension of obtrusiveness describes the physical discomfort due to its size, weight, form or noiselevel. The usability dimension depict the user-friendliness and the additional time effort needed to handle the system. An increasingly important point is the privacy dimension which describes how personal information is handled. The functional dimension describes the accuracy of a system and how faultless it behaves. The dimensions shown in the lower part of Figure 11 can be pooled to the term “social wearability”, which is further described in the next section.

C. Social Wearability

According to [24] Maslow’s hierarchy of human needs can be applied to the context of wearable devices. “If basic needs are met, attention shifts to higher order needs” [25]. Basic needs are summarized in the upper part of Figure 11 while higher order needs involve the social acceptance of a device or the self-actualisation by wearing the device. The social wearability of a device is represented by it’s “expressive” (color, texture, form), “referential” (brand, trend, social role) and “interacting” (passive or active gestures) characteristic [24]. According to a case study, most people are concerned by negative visual properties while wearing or interacting with devices that are attached on the body. The preferred body position is the wrist or the forearm, whereas the torso and hip are considered as awkward and sexual suggestive. For interacting with clothing, garment features or edges afford interactions and support “natural gestures”. A body-map of such interactions is displayed in Figure 12 separately for men and woman. Lines show garment features or edges while light gray fields indicate body positions that provide easy access. Dark gray fields highlight social problematic locations which differ in gender and depending on the cultural and locational affiliation.

D. Body Positions

Since obtrusiveness is closely related to the body position where the device is attached, one may ask if the body position impair the recognition accuracy of a HAR-System. As stated in [21] the recognition of basic activities like sitting, standing, walking or running is not affected by the body position of the sensor. Furthermore a social acceptable location like the wrist performed best in an experiment where six different body locations were compared with the associated recognition accuracy. However according to [1] there are many works which all came to different result for the perfect body positions for sensors. To conclude this topic no clear statement could be made.

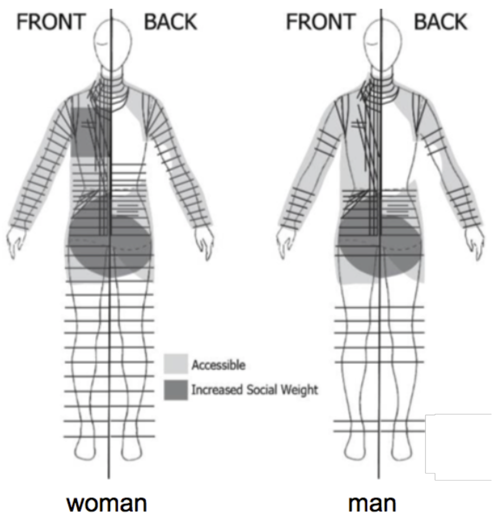


Fig. 12. A body map for possible locations to interact with clothing [24].

V. Modalities

The term "modalities" is very interesting with respect to sensors in Human Activity Recognition, but different definitions exist. In general modalities describe different ways to measure the same physical quantity. One example would be to replace an accelerometer by another sensor but keeping the acceleration as the sensor-output. Unfortunately the word modalities is often used to describe the use of another physical quantity to solve the same problem with for instance a new sensor. Even if it would be possible to measure the same physical effect with the new sensing method, the data is not transformed to the previously used quantity. Instead the new data is used independently and a classifier is trained directly on this data.

But some works exists that fit in the general definition. The authors in [6] tried to replace a gyroscope in a HAR-system with a magnetic-field-sensor, while keeping the angular velocity as the measured quantity. One motivation for this is to replace a gyroscope completely by an already existing magnetometer which can save cost, space and energy. As the result the authors came up with a formula to calculate the 3D angular velocity from data of the 3D magnetic field sensor. Since magnetometers could be highly disturbed by environmental magnetic fields or ferromagnetic material the authors also investigate the sources of error for their method. Fortunately the disturbance of the sensor output is not that critical in this context because the derivation is considered in the formula. Only high inconsistent fields or magnetic fields with strong curvature in its field lines disturb the calculation of the angular velocity. To test the influence of this possible error-sources the authors also did some evaluation. Therefore primarily recorded HAR-data-sets were used that include magnetometer and gyroscope data. By using the provided formula they tried to reconstruct the angular velocities measured by the gyroscope with the magnetometer data. From the results which are shown in Table I it can be concluded that the accuracy depends much on the body position where the sensor is placed. The authors claimed that sensors worn on the wrist or leg are more often in contact with a disturbing magnetic field than for example a sensor that is attached on the head. Because the accuracy was

TABLE I. Accuracy of reproduced angular velocity data (from [6]).

Placement	Mean Error	Median Error	Stand. Dev.
Head	17%	11%	90%
Torso	21.1%	12.5%	110%
Back	23.5%	19.5%	150%
Wrist	53.2%	44.2%	214%
Lower leg	34%	23.4%	162%
Upper arm	32.0%	25%	173%

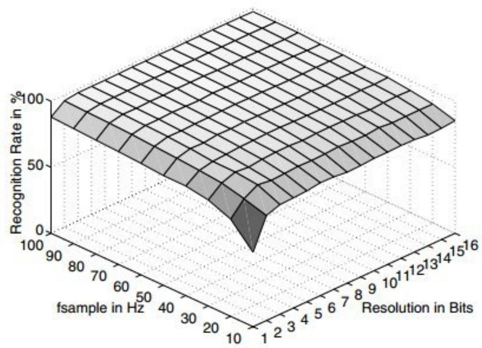
not good in general, the authors did a second evaluation where they used their method directly in a HAR-system. But the previously accomplished results, especially for sensors worn the arm, were justified. The recognition accuracy was always worse when the calculated angular velocity was used compared to the use of the direct output from the gyroscope. So even if the authors came to the conclusion that it is possible to replace a gyroscope with a magnetometer by using their method one may conclude that this is not yet possible. But using the proposed method magnetometers could be used to compensate e.g. shift of other sensors employing the same techniques as developed for gyroscopes.

VI. Sampling Frequency and Signal Resolution

In each digital HAR-system a appropriate sampling frequency and signal resolution must be chosen for each sensor. The signal resolution as well as the sampling frequency should thereby be as small as possible to save e.g. energy but it should not affect the recognition in a negative sense on the other side. Many research works focus on the influence of these parameters on the accuracy of the whole HAR-systems. One of this works is [26]. Here the authors equipped test-persons with 3D accelerometers and recorded acceleration data while the persons performed different activities. The sampling frequency at which the data was recorded was $100Hz$ and the resolution was $16Bit$. Afterwards the authors downsampled and quantized the data-sets onto different quality-levels and fed each set into three different classifiers (Decision-Tree, k-Nearest-Neighbour and Naive Bayes). Since the results are nearly the same for each classifier only the outcome of the Decision-Tree is shown in Figure 13. One may see, that if the sampling frequency is above or equal $20Hz$ and the signal resolution is above or equal $2Bit$, the recognition accuracy stays constant and the parameters are thus optimal. Fortunately other works came to the same result as well like e.g. [21]. Furthermore they showed that the optimal value for the sampling frequency of $20Hz$ even holds for light sensors.

VII. Conclusion

This paper highlights the importance of sensors in HAR-systems as they provide the core data for the recognition process and mainly influence the recognition accuracy, cost and wearability of any HAR-system. While taking a look back on the key-questions stated in Section I all of them were answered in this work. By introducing different relevant sensors groups with their most prominent sensors a good overview was given to choose a sensor for a specific application. The shown advantages and drawbacks for each sensor are also a first part of the answer to the second question, which covers the handling of such drawbacks. This question was further investigated in the section dealing with obtrusiveness



(a) C4.5

Fig. 13. The recognition accuracy of a Decision-Tree classifier in dependence of the sampling rate and the signal resolution [26].

where the importance of an unobtrusive system in HAR was highlighted. Many aspects for building such a system were given beside a standardized measure of obtrusiveness. The third question which asks for a limit up to which already available sensors are sufficient was mainly answered in the last two parts. In the section covering modalities, an adequate definition of the term with respect to the question was given and one example was presented in which the authors tried to replace a gyroscope with a magnetometer. Unfortunately this was not really a success and is representing one limitation of sensors (in this case for magnetometers). In the last part a second obvious limit for sensors was given by identifying the optimal sampling frequency for sensor with $20Hz$ and the optimal signal resolution with $2Bit$. Since nearly all modern sensors fulfill this limit they are appropriate for a use in HAR-systems in general. This assumption was further proven in Section evolution of HAR-system. In this section one work was introduced which used an of-the-shelf sensors platform and compared to another work which used an smartphone placed somewhere on the upper arm. In both cases the system achieved a recognition accuracy of more than 90%, which shows that sensors in smartphones are totally suitable for the application of Human Activity Recognition as well.

Through the research for this work we learned the fundamentals of Human Activity Recognition and discovered the steps needed to set up a HAR-system from scratch. We think that more accurate and smaller sensors could help to improve the recognition accuracy of existing HAR-systems. In addition we see a lot of research going in the direction of activity recognition with already available devices like smartphones or smartwatches. The variety of sensors inside these devices and their social acceptance will establish practicable HAR-systems. The fact that these devices are available to nearly everyone supports the ongoing research which will further improve the recognition accuracy and the simplicity in usage.

References

[1] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *Communications Surveys & Tutorials, IEEE*, vol. 15, no. 3, pp. 1192–1209, 2013.

[2] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *ACM SigKDD Explorations Newsletter*, vol. 12, no. 2, pp. 74–82, 2011.

[3] S.-W. Lee and K. Mase, "Activity and location recognition using wearable sensors," *IEEE pervasive computing*, vol. 1, no. 3, pp. 24–32, 2002.

[4] T. Choudhury, S. Consolvo, B. Harrison, J. Hightower, A. Lamarca, L. LeGrand, A. Rahimi, A. Rea, G. Bordello, B. Hemingway *et al.*, "The mobile sensing platform: An embedded activity recognition system," *Pervasive Computing, IEEE*, vol. 7, no. 2, pp. 32–41, 2008.

[5] Xsens the leading innovator in 3d motion tracking technology. [Online]. Available: https://www.xsens.com

[6] K. Kunze, G. Bahle, P. Lukowicz, and K. Partridge, "Can magnetic field sensors replace gyroscopes in wearable sensing applications?" in *Wearable Computers (ISWC), 2010 International Symposium on*. IEEE, 2010, pp. 1–4.

[7] Ó. D. Lara, A. J. Pérez, M. A. Labrador, and J. D. Posada, "Centinela: A human activity recognition system based on acceleration and vital sign data," *Pervasive and mobile computing*, vol. 8, no. 5, pp. 717–729, 2012.

[8] E. M. Tapia, S. S. Intille, W. Haskell, K. Larson, J. Wright, A. King, and R. Friedman, "Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor," in *Wearable Computers, 2007 11th IEEE International Symposium on*. IEEE, 2007, pp. 37–40.

[9] D. Gordon, H. R. Schmidtke, M. Beigl, and G. Von Zengen, "A novel micro-vibration sensor for activity recognition: Potential and limitations," in *Wearable Computers (ISWC), 2010 International Symposium on*. IEEE, 2010, pp. 1–8.

[10] M. Stäger, P. Lukowicz, and G. Tröster, "Implementation and evaluation of a low-power sound-based user activity recognition system," in *Wearable Computers, 2004. ISWC 2004. Eighth International Symposium on*, vol. 1. IEEE, 2004, pp. 138–141.

[11] G. Ogris, T. Stiefmeier, P. Lukowicz, and G. Tröster, "Using a complex multi-modal on-body sensor system for activity spotting," in *Wearable Computers, 2008. ISWC 2008. 12th IEEE International Symposium on*. IEEE, 2008, pp. 55–62.

[12] J. Yamato, J. Ohya, and K. Ishii, "Recognizing human action in time-sequential images using hidden markov model," in *Computer Vision and Pattern Recognition, 1992. Proceedings CVPR'92., 1992 IEEE Computer Society Conference on*. IEEE, 1992, pp. 379–385.

[13] R. Polana and R. Nelson, "Low level recognition of human motion (or how to get your man without finding his body parts)," in *Motion of Non-Rigid and Articulated Objects, 1994., Proceedings of the 1994 IEEE Workshop on*. IEEE, 1994, pp. 77–82.

[14] K. Van Laerhoven and O. Cakmakci, "What shall we teach our pants?" in *Wearable Computers, The Fourth International Symposium on*. IEEE, 2000, pp. 77–83.

[15] A. Y. Benbasat and J. A. Paradiso, "Compact, configurable inertial gesture recognition," in *CHI'01 Extended Abstracts on Human Factors in Computing Systems*. ACM, 2001, pp. 183–184.

[16] Analog devices adxl202. [Online]. Available: http://www.roboeq.ir/uploads/images/shop/products/L1339363496274030.jpg

[17] Murata enc-03-j. [Online]. Available: http://cache.amobbs.com/bbs_upload782111/files_9/ourdev_245063.gif

[18] K. Altun and B. Barshan, "Human activity recognition using inertial/magnetic sensor units," in *Human behavior understanding*. Springer, 2010, pp. 38–51.

[19] M. Muehlbauer, G. Bahle, and P. Lukowicz, "What can an arm holster worn smart phone do for activity recognition?" in *Wearable Computers (ISWC), 2011 15th Annual International Symposium on*. IEEE, 2011, pp. 79–82.

[20] N. Kern, B. Schiele, and A. Schmidt, "Multi-sensor activity context detection for wearable computing," in *Ambient Intelligence*. Springer, 2003, pp. 220–232.

[21] U. Maurer, A. Smailagic, D. P. Siewiorek, and M. Deisher, "Activity recognition and monitoring using multiple sensors on different body positions," in *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*. IEEE, 2006, pp. 4–pp.

[22] J. F. Knight, D. Deen-Williams, T. N. Arvanitis, C. Baber, S. Sotiriou, S. Anastopoulou, and M. Gargalakos, "Assessing the wearability of wearable computers," in *Wearable Computers, 2006 10th IEEE International Symposium on*. IEEE, 2006, pp. 75–82.

[23] B. K. Hensel, G. Demiris, and K. L. Courtney, “Defining obtrusiveness in home telehealth technologies,” *Journal of the American Medical Informatics Association*, vol. 13, no. 4, pp. 428–431, 2006.

[24] L. E. Dunne, H. Profita, C. Zeagler, J. Clawson, S. Gilliland, E. Y.-L. Do, and J. Budd, “The social comfort of wearable technology and gestural interaction,” in *Engineering in Medicine and Biology Society (EMBC), 2014 36th Annual International Conference of the IEEE*. IEEE, 2014, pp. 4159–4162.

[25] A. H. Maslow, “A theory of human motivation.” *Psychological review*, vol. 50, no. 4, p. 370, 1943.

[26] H. Junker, P. Lukowicz, and G. Troster, “Sampling frequency, signal resolution and the accuracy of wearable context recognition systems,” in *Wearable Computers, 2004. ISWC 2004. Eighth International Symposium on*, vol. 1. IEEE, 2004, pp. 176–177.

Feature Sets for Activity Recognition

Wearable and Network Embedded Systems Seminar 2015

Sebastian Böttcher
Faculty of Engineering
University of Freiburg
Georges-Köhler-Allee 101
79110 Freiburg, Germany
Email: boettchs@informatik.uni-freiburg.de

Mireille Rehse
Faculty of Engineering
University of Freiburg
Georges-Köhler-Allee 101
79110 Freiburg, Germany
Email: rehsem@informatik.uni-freiburg.de

***Abstract*—During the *wearable and network embedded systems* seminar, students discussed current topics in the fields of wearable tech, sensor networks, embedded systems and activity recognition. In the course of the semester, each group gives a number of talks to present their intermediary results, concluding in a survey paper reporting their final findings for the topic.**

In this report we summarize and review recent advancements in HAR by specifically looking at the development of feature sets used in research. Features are essential in reducing the large data sets that are generally used in HAR to their characteristic traits, which then train classifiers to recognize the activities labeled in the training data sets.

I. Introduction

Human activity recognition is a relatively new field of study, with some early works of research dating back to the 1990's. Since then numerous advances in the fields of machine learning, computer vision, but also sensor technology and general processing power have continually driven HAR to a point of relative certainty in recognition of basic human movement. Recent research papers report accuracy rates of 90 % and more on activities like *walking*, *running*, *sitting*, *ascending* and *descending stairs*.

In the summer semester of 2015 the embedded systems group of the faculty of engineering at the University of Freiburg offered the *wearable and network embedded systems* seminar, during which 7 groups of students researched the topics *Feature Sets for AR*, *Habitat Monitoring*, *Fundamentals*, *Sensors*, *Body Models*, and *Wetlab Support*. All groups report on their preliminary findings in five presentations during the course, and write a survey paper summarizing their respective topics as well as their research techniques.

In this paper we report on the seminar topic of *Feature Sets for Activity Recognition*. We selected 46 papers and two Ph.D. theses for review and pay regard to types and number of features selected by researchers, activities that are classified, influence of the classification scope on feature selection, sensors that are used to collect data, and the classification accuracy achieved by the respective systems. Since one of the goals of the seminar is to teach students to extract knowledge from a large number of research papers, we also summarize the methodology we applied during the seminar.

In Sec. II, we give a brief overview of the topic of human activity recognition and outline the general recognition pipeline applied in most of the relevant research work. Sec. III lists different types of features and feature sets, and discusses their relevance for activity recognition. In our exploration of the topic at hand, we came upon several cases of use of comparably unconventional sensors utilized for the task of activity recognition. In Sec. IV we describe some of these systems and their sometimes special use of features. Furthermore, Sec. V describes the effect of sensor displacement on AR systems and the features used in them. In Sec. VI we summarize the methodology we applied during our research and give some simple statistics on the papers we studied. Finally, Sec. VII concludes our research.

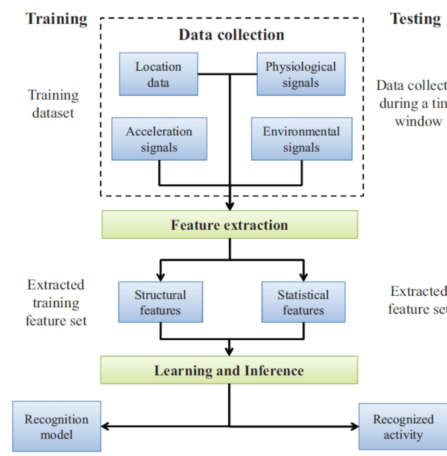


Fig. 1. **Activity recognition pipeline [1]:** Data flow in human activity recognition systems, with training and test data based on wearable sensors.

II. Activity Recognition Overview

In general, most activity recognition systems follow a specific sequence of processes to identify the activities performed by humans. In principle (see also Fig. 1), first the relevant data is collected by a sensor system, usually worn by test subjects. The data might then be segmented, after which features are selected and extracted for an accurate and efficient representation of the data. In the last step, these features are used to train a classifier, which is then able to detemine activities from non-training test data.

The following descriptions of the consecutive steps in the activity recognition chain are based on the *'Tutorial on Human Activity Recognition'* by Bulling et al. [2]:

Data Collection

The first step in any HAR system is always data acquisition. Test subjects timeframes to label the data. In standard research, acceleration data is widely used in a multitude of systems, and is generally accepted as the most efficient form of motion data in human activity recognition. However, there are many other types of sensors that can be used, ranging from standard IMU data like gyroscope and magnetometer data, to more exotic types like EEG signals, sound or GPS coordinates. Relevant for data collection is also the sampling rate, which typically is in the range of 5 Hz to 100 Hz, depending on intended accuracy and sensor system architecture limitations.

Data Segmentation

After collecting the raw data, it is often segmented into the relevant portions containing the activities. The raw data oftentimes has long stretches of useless data, e.g. when the person is standing still or just generally not performing any actions. These portions are usually not interesting for the application and can be discarded for less computation time. There are several approaches to segmentation: The sliding window, that consecutively extracts a fixed-size portion of the data for processing. Energy-based approaches segment the data according to the intensity of the movement. Finally, additional sensors may be used to gather information about the current activity. For example, GPS data might be used to segment IMU data to specific locations.

Feature Extraction

To be able to properly separate activities, features are selected and extracted from the data. These describe the activities in the data as effectively as possible and discrimate between the different activities. All features that are extracted form the feature space, in which feature points close together should correspond to the same activities. There are several types of features (see also Sec. III): Signal-based features are statistical time or frequency domain features and the most common found in research. Body model features are based on a 3D representation of the human body and describe properties like trajectories. Finally, event-based features, e.g. for eye movements or muscle activity. In general, features lessen the computational load on the recognition system, so selecting the minimum best set of features is an important part in the activity recognition chain. Features can either be selected manually, or automatically via ranking algorithms. These methods identify the most important features, lower training times in the next step, and generalize the feature set. They can be grouped into filter methods, which individually score single features, wrapper methods, which apply search algorithms, and hybrid approaches which use machine learning techniques to select features.

Training and Classification

In the last step of the activity recognition chain, machine learning algorithms are used to classify activities in non-labeled test data. Beforehand, they are trained with the features previously extracted and the labels noted by researchers. Popular methods include *dynamic time warping*, *hidden Markov models*, *support vector machines*, and *C4.5 decision trees*.

III. Feature Types

The choice of appropriate features is an important task to achieve a high classification accuracy. There are several feature types and methods which can be recognized to facilitate this procedere. In addition to the commonly used statistical features (Bao and Intille [3]), Zhang and Sawchuk [4] introduced physical features, which are optimized for human movements in activity recognition. The feature selection depends on several factors: The activities to be recognized, which types of sensors are used, and how much of them. Is it possible to wear them on different body parts? A straightforward approach would be to use all features and thereby improve the accuracy. But several studies have shown, that the feature space should be as small as possible. One reason is the reduction of the computational cost, especially with regard to wearable devices. Furthermore there are features which can falsify the results and the classification will get unclear.

In this section we introduce features for activity recognition with accelerometers, gyroscopes and inertial sensors since these are the most widely used sensor types in this field.

A. Statistical Features

Statistical features are basic features, which are used in most activity recognition procedures. These features are extracted from each sensor axis individually. Tab. I shows statistical features commonly used for AR with accelerometers and gyroscopes.

Feature	Description
mean	DC component of the signal over the window
median	median signal value over the window
standard deviation	variation of the signal over the window
variance	square of standard deviation
root mean square	quadratic mean value of the signal over the window
averaged derivatives	mean value of the first order derivatives of the signal over the window
pairwise correlation	correlation between two axes
spectral entropy	distribution of frequency components

TABLE I. Common statistical features and brief descriptions [4]

By taking the mean acceleration it is possible to recognize poses without movement like sitting, standing still and lying. The variance distinguishes activities by the different acceleration values e.g. walking from jogging. The correlation can differentiate activities that involve translation in single or multi-dimension, such as walking from stair climbing. These correlation features can also recognize activities that involve several body parts. By means of the frequency domain entropy, activities with similar energy values can be distinguished. For instance, the discrete FFT hip acceleration while biking only shows one component at 1 Hz. Running result in a complex hip acceleration and many frequency components.

B. Physical Features

Physical features are extracted from multiple sensor channels, with sensor location and orientation known a priori.

Movement Intensity:
Euclidean norm of the total acceleration vector, for which the mean and the variance over the window is computed.
This measures the immediate intensity of human movements.

Normalized Signal Magnitude Area:
Acceleration magnitude summed over three axes and normalized by the window length.
This feature is an indirect estimation of energy expenditure.

Eigenvalues of Dominant Directions:
Covariance matrix of the acceleration data along the axis in each window. The eigenvectors correspond to the dominant directions, the eigenvalues to the corresponding relative motion magnitude.
This computes the acceleration along the axis, e. g. for jumping or running forward.

Correlation between Acceleration along Gravity and Heading Directions:
Euclidean norm of the total acceleration vector along the heading direction. Then calculate correlation coefficient between heading direction and gravity.

Averaged Velocity along Heading and Gravity:
Euclidean norm of the averaged velocities along y and z axes over the window.

Dominant Frequency:
Maximum of the squared discrete FFT component magnitude from each axis.

Energy:
Normalized sum of the squared discrete FFT component magnitudes from each axis.

IV. UNCONVENTIONAL SENSORS

During our research, we found that the majority of authors use traditional sensors and means to record data, like accelerometer readings over a few hours. However, in some papers more unconventional sensors were used to classify activities diverging from the usual movements, like reading, swallowing, finger taps on skin, etc. In this chapter we present some of these papers and the special features used in recognizing some of these activities.

A. Wearable Acoustics

In [5] acoustic sensors are introduced to record and classify sounds which are produced in the mouth and throat area.

For the prototype a microphone is embedded into a bluetooth headset and covered with a chestpiece of a stethoscope. A microphone in the earpiece of a headset is very sensitive to sounds emerging in the mouth region, e.g. caused by eating, drinking, speaking, whispering, whistling, laughing, sighing, coughing.

To recognize the activities a sample length of 5 s and a sampling rate of 22 050 Hz were chosen. The three relevant domain features are time, frequency and cepstral. To calculate the frequency and cepstral features, the data was pre-processed with FFT. The average and standard deviation was calculated across all frames for each feature.

Time-domain feature	Zero-crossing rate: rate of sign changes along a signal
Frequency-domain features	Total spectrum power: logarithm of summed spectrum power
	Subband Powers: Frequency spectrum divided into logarithmic subbands
	Brightness: frequency centroid
	Spectral rolloff: frequency below which 93 % of the distribution is concentrated
	Specral flux: average variation value of spectrum between to adjacent frames
Cepstral feature	Mel-frequency cepstral coefficients

TABLE II. SUMMARY OF THE DOMAIN FEATURES AND BRIEF DESCRIPTIONS

For training and testing two protocols are used: *Leave-one-participant-out* and *Leave-one-sample-per-participant-out*. The results of the laboratory evaluation show that it is possible to achieve an average accuracy of 49 % over all activities by using the classification with the *Leave-one-participant-out* protocol. With the *Leave-one-sample-per-participant-out* protocol the accuracy was improved by 30 %. Activities like taking a deep breath, drinking and sighing are most difficult to detect, while whistling and speaking achieve high accuracies.

B. Eye Movement

Eye-based activity recognition (EAR) is used to detect human activity by analyzing the eye movements like saccades, blinks and eye movement patterns. In [6] electrooculography is applied to detect the signals produced by the eye movements to recognize activities like copying a text, reading a printed paper, taking handwritten notes, watching a video and browsing the web.

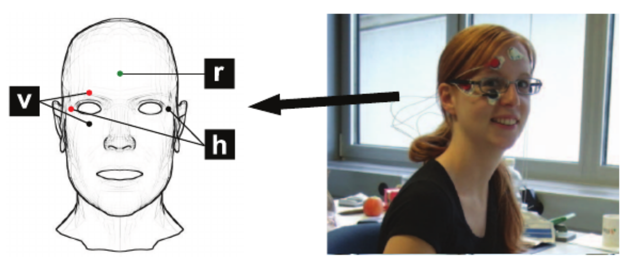


Fig. 2. Electrooculography (EOG): Electrode placement for detecting the eye movements [6]

Electrooculography enables to model the eye as a dipole by placing electrodes around the eye. If the eye moves from center position toward one of the two electrodes, this electrode ”sees” the positive side of the retina while the opposite electrode ”sees” the negative side of the retina. Consequently, a potential difference occurs between the electrodes. Assuming that the resting potential is constant, the recorded potential is a measure of the eye’s position.

The feature selection is based on the different movements like saccades, fixations, blinks and the wordbook of eye movement patterns. Eye movement patterns are strings of a certain length of eye movement sequences. The movements are differentiated by the direction the person is looking at.

Movement	Features
Saccade (S)	mean (mean), variance (var), maximum (max) of the EOG signal amplitudes (Amp), rates (rate) of small (S), large (L), positive (P), negative (N) saccades in horizontal (Hor), vertical (Ver) direction
Fixation (F)	mean (mean), variance (var) of the horizontal (Hor), vertical (Ver) EOG signal amplitudes (Amp) within or duration (Dur) of fixation or rate of fixation
Blinks (B)	mean (mean), variance (var) of the blink duration or blink rate (rate)
Wordbook (W)	wordbook size (size), maximum (max) difference (diff) between maximum and minimum mean (mean), variance (var) of all occurrence counts (Count) in the wordbook of length (-lx)

TABLE III. SUMMARY OF THE USED FEATURES [6]

For classification and feature selection the *Leave-one-person-out* scheme was used. In the following the top three features are listed which were selected by the mRMR algorithm.

read: W-maxCount-l2, W-meanCount-l4, W-varCount-l2
browse: S-rateSPHor, W-varCount-l4, W-varCount-l3
write: W-carCount-l4, F-meanVarVertAmp, F-varDuration
video: F-meanVarVertAmp, F-meanVarHorAmp, B-rate
copy: S-varAmp, S-meanAmpSNHor, S-meanAmpLPHor

The experiments reached an accuracy of more then 62 % to detect the activities.

C. Wearable Face Recognition System

A wearable system based on Google Glass is presented by Mandal et al. [7]. The system includes several levels like face detection, eye localization, face recognition and a user interface to display the information. The challenge is to handle different lighting conditions, various face poses and faces of various scaling.

In the first step the system detects a face in a FPV image by using the OpenCV face detector. OpenCV is only adequate for frontal views, so additionaly an algorithm based on Haar features and an Adaboost classifier is used to detect non-frontal view faces. OpenCV is also applied to localize the eyes. In this case it performs well for front-view faces, even with closed eyes. For oblique view and faces with various scale the ISG eye detector algortihm is used. The next step normalizes the image by the following steps: integer to float conversion, geometric normalization, masking, histogram equalization, pixel normalization and scaling all on the same picture size. To be able to recognize a person's face and provide the related personal informations, eigenfeature regularization and extraction is performed.

In the training stage the normalized and preprocessed images are clustered into subclasses where each person is one class. Then the within-subclass scatter matrix is computed and the eigenfeature regularization scheme is applied to get the regularized features. In the next step the total-subclass and between-subclass scatter matrices are computed and the features are chosen for which the corresponding eigenvalues are largest.

In the recognition stage every incoming image vector is transformed into a feature vector by using the feature regularization and extraction matrix and after this a classifier can be applied.

In the experiments for the eye localization an accuracy of 70 % for non-frontal view faces and of 90 % for frontal view faces is achieved. For the face recognition, the system achieves over 90 % accuracy rate when using the PCA method and more than 100 features.

D. Worn Capacitive

Cheng et al. [8] present a system of capacitive sensors made from conductive textile electrodes to observe changes of capacitance in several places of the human body. They design front-end boards with attached conductive textile material cut into shape to record data from several locations: Placement on the chest, wrist and neck (Fig. 3) provide the ability to recognize activities like breathing, head motions, speaking, swallowing, heartbeat, etc.

For data analysis, 45 different features are extracted from 1.5 s sliding windows, including signal mean, variance, and maximum. The recorded data sets include activities performed while sitting and walking, among them swallowing, nodding, speaking, and head movements. A linear discriminant classifier is employed and shows an accuracy of detecting the activities while walking and sitting, and only while sitting, of 69 % and 77 % respectively.

Additionally, the detection of swallowing and the amount swallowed, and the detection of respiration rate is analyzed. To detect swallowing, a feature similarity search is applied to the features previously mentioned, with a resulting performance of 80 % recall and 60 % precision. The amount is recognized using a linear discriminant classification using the features variance, minimum, and maximum, with the results being "comparable [...] to previous investigations using audio and Electromyography [9]". To find the rate of breathing of a test person, data from the neck sensor is filtered and subjected to a hill-climbing peak detection algorithm. The method shows a detection rate of 80 % to 90 %, however the authors note that it may fail when breathing is irregular, in which case the heartbeat could be picked up by the sensor.

The researchers conclude that the method "is a viable and highly interesting new sensing concept for wearable monitoring." They furthermore note its use in biomedical and healthcare investigation, and pose that "further work should address the integration and optimization for individual applications".

E. Muscle Activity

With their wearable muscle activity sensing system, Amft et al. [10] show the validity of force sensitive resistors in monitoring individual muscle activity. They conduct experiments

Fig. 3. **Placement of capacitive sensors [8]:** The textile based sensors are placed on the chest, wrist, and neck, with integration in a pullover collar.

with test subjects performing four distinct arm movements and recording the signals from the FSR and an additional fabric stretch sensor, both attached to the subjects lower arm. The recorded data from both devices is thresholded and indicates that both methods can be used to detect arm movements. However, only the FSR method can discriminate between single muscles, whereas the FSS reacts to any change in arm circumference, thus the arm motions are indistinguishable.

Lukowicz et al. [11] use force sensitive resistors attached to the upper leg to interpret muscle activity to recognize modes of locomotion. A sensor is attached on the upper leg of a test subject and fastened with an elastic band. The subject performs a number of walking actions and the data is recorded an labeled. Some physiological facts are used to formulate features that best describe the modes of locomotion:

1. No signals during any leg swing phase.
2. For normal walking, the front and back leg muscles alternate their activity phases between pushing off and putting down the feet.
3. An increase in muscle activity and a decrease in delays between peaks indicates faster walking.
4. For both walking downstairs and upstairs, the front leg muscle is dominant, however during the former activity, all leg muscles are active, while during the latter the back muscles are mostly inactive.

From these, the swing phase, the ratio and the delay between front and back muscle activity can be derived as features. In Fig. 4, the feature space for one experiment clearly shows the separability of the four considered modes of locomotion. The paper does not include classification of the activities, however mentions it as possible future work.

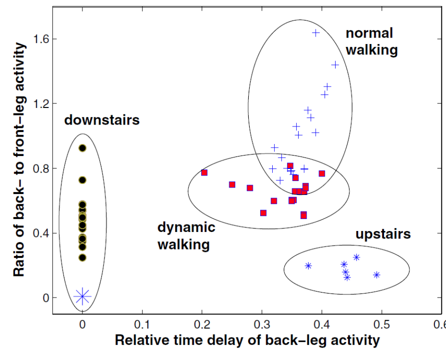
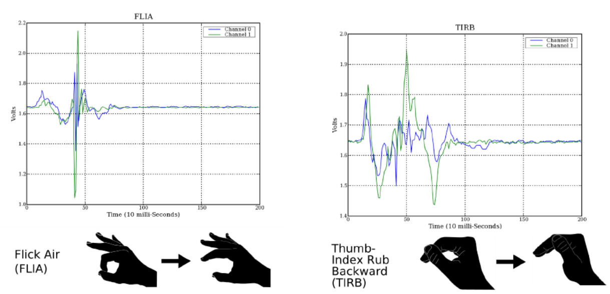


Fig. 4. **Muscle activity features [11]:** The four modes of locomotion can be separated using the described features.



(a) (b)

Fig. 5. **Hand gestures with data [12]:** Two example hand gestures with start and end positions, and the recorded two second data window.

F. Bio Acoustics

The Hambone system developed by Deyle et al. [12] uses two piezoelectric sensors to record acoustic data transmitted via bone to the wrist-attached sensor platform. Whenever the subjects hand pose changes or certain gestures are performed, acoustic waves are generated by the skin, which are transmitted on the surface of the skin or through the bone. Piezoelectric sensors can observe these waves while not being affected by other external sound waves. From the recorded data, a hidden Markov gesture model is generated by the HMM classifier implemented in the Georgia Tech Gesture Toolkit [13]. Additionally, the same method can be applied to the feet, with the sensor device attached to the ankles. The experiments include seven distinct hand and foot gestures (see also Fig. 5).

The models, trained with 20 sample data windows for each gesture, have an overall correctness of 81 %, 82 %, and 100 % with data from one wrist sensor, one ankle sensor, and two wrist sensors respectively. Additionally, the models are tested across different users, which results in accuracies of 63 % and 69 % for wrist and ankle sensors respectively. The authors conclude that the system "is a viable approach to mobile device interaction."

Harrison et al. [14] proposed the Skinput system, another device that uses acoustic signals to recognize fingertaps and their locations on a users arm. Devices like these are often applied to UI problems and can even be combined with some kind of projection on the skin to make it more convenient to use.

V. Sensor Displacement

In real-life applications it is not possible to have an exact fixed position for the sensors which have to record the data. For AR often the sensors implemented in smartphones are used and for this several positions are conceivable, like for example the trouser pocket, the bag or grasped in the hand. These positions affect the sensor signal in different ways and thus make it more difficult to get a clear signal which is neccessary to characterize the different activities.

Sensor displacement can be decomposed into the following sub-problems: On-body placement, sensor oriention, and displacement within a body part [15].

On-body placement describes the different possible body locations where the device can be carried. Some options can be the trouser pocket, jacket pocket, bag, at the belt or in a holster at the shoulder. There are some approaches to handle these placement variations like methods to recognize the body part location where the sensor is. For this we analyze the body constraints and how body parts behave, e.g. a sensor at the shoulder has a different movement characteristic than at the wrist. Another approach is to train your classifier for different body locations. This is realizable if there are not so many activities and different body locations, otherwise the state space is getting very large. The recognition of the body part can also be reached by taking more than one sensor into account or to choose features which are placement-invariant, but there are only a few activities which can be recognized in this way.

The sensor orientation is important for detection of activities which involve translation in different dimensions, like taking a stairway, climbing or jumping. This problem could be solved if the subject performs some easy calibration gestures. It is also conceivable to measure the acceleration over all axes to recognize the gravity.

Even at a fixed body position a displacement within this body part can be possible, like e.g. shifting around while jogging or movement on a belt. To focus on this displacement occurence, comprehension of the human movements is required. One approximation is to assume a rigid body model. By means of this idealization we can describe all movements through translation and rotation and derive the appropriate features. During translation all points of the model have the same speed and thus the accelerometer is location invariant while the gyroscopes produce no signals. During rotation the signal of the gyroscope is the same at all points, while the acceleration differs and is placement dependent. If the motions are dominated by translations or changes in orientation, the acceleration features have to be chosen. If they are dominated by rotation or rotation and translation, the gyroscope provides stable features.

VI. Research Summary

Researching the given topic, finding papers, analyzing and distilling them are significant parts of the wearables seminar. In this chapter, we give a brief overview of our methodology, and present some statistics extracted from the papers we read.

A. Methodology

The seminars key point is teaching students to research a certain topic, find scientific papers on it, and compile them into a survey. To help with structuring the process, each group gives biweekly short presentations on their progress. We chose the following structure for our presentations:

1. **Topic Overview:** We explain the general activity recognition pipeline and pose some key questions we feel are important for our remaining research.
2. **Feature Types:** After briefly touching on features in the first part, we now give an in-depth look at certain types and selection methods.
3. **Unconventional Sensors:** We have noticed a lot of researchers applying IMUs in their work, so we specifically try to find papers that don't.
4. **Specialization Part 1:** To adress some of our previously posed key questions, we analyze sensor displacement as a problem, and present a thesis on motif discovery.
5. **Part 2 + Summary:** We take a look at wearable face recognition and present some statistics and a short summary of all talks.

To find relevant papers, we primarily use https://scholar.google.de/ and http://ieeexplore.ieee.org/. Keywords like *human activity recognition*, *features*, and *wearable* are enough to yield a relevant set of papers. For more specialized topics, additional terms like e.g. *sensor displacement* and *face recognition* help with filtering. In general, we saved all papers we analyzed for later use. A bibliography database was kept using *JabRef* [16].

Some key questions we posed during our research:

- Which features are selected for specific activities?
- Which classifiers are selected for a recognition task? *(considered as off-topic later on)*
- How precise is recognition of everyday activities?
- Which challenges are addressed by specific features?
- Are there feature sets for specific sensing modalities?

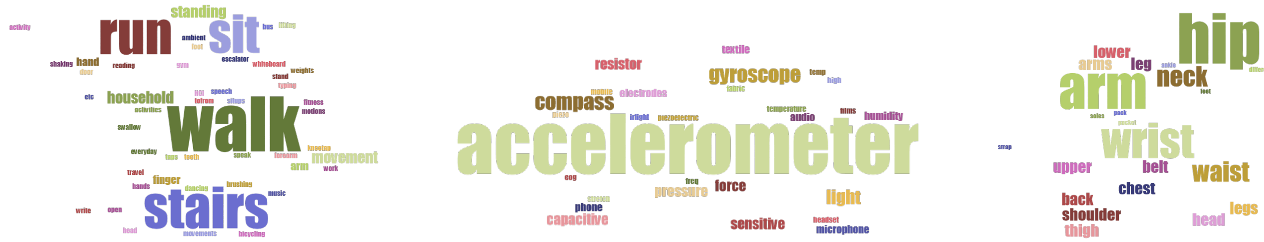
B. Statistics

In the course of the seminar, we analyzed 50 works of research, of which 46 were scientific papers, two dissertations and two books. For the purpose of this statistics summary, we only regard the scientific papers, with the oldest one published in 1995 (second oldest in 2002).

31 papers presented extractable data, with the points of interest being types and number of features, types and number of activities classified, sensor types used and their position, and any notes on the overall accuracy of the presented recognition system. Some key points from the data:

- 20 researchers chose an accelerometer or IMU as their main source of activity data.
- 23 systems use one or more statistical features in time domain.
- The average number of features used is 127, with 7 feature sets larger than 100.
- The average number of activities recognized is 7.5, overall ranging from 1 to 20.
- 19 papers report an overall accuracy of at least one part of their system of greater than 90 %.

Overall, from the data we gathered we can conclude that the number of features used is increasing, which seems to have an effect of increased classification accuracy.



(a) (b) (c)

Fig. 6. **Word clouds:** (a) The activities classified by the various recognition system; (b) The sensors used for recording data; (c) The body positions the sensors are attached to or are positioned at in general.

We summarized parts of the extracted statistics into word clouds (Fig. 6), clearly showing some of the preferences in state of the art research in the field. Almost all researchers used some form of IMU, either attached on the arm, or positioned at the hip, indicating the use of a smartphone or similar device. Among the most popular activities that are classified by the systems are different modes of locomotion, as well as more recently various household activities.

VII. Conclusion

Our research showed that ever since human activity recognition became a field of study, statistical and FFT features are very prevalent in recognition systems. They constitute most of the feature sets used when an IMU or similar device is the main source of activity data and seem to produce the best results even today. Recently, mobile phones have only added to this. They are very accessible, integrate all the sensors needed, and have become powerful enough to even do online feature extraction and classification in some cases.

References

[1] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *Communications Surveys & Tutorials, IEEE*, vol. 15, no. 3, pp. 1192–1209, 2013.

[2] A. Bulling, U. Blanke, and B. Schiele, "A tutorial on human activity recognition using body-worn inertial sensors," *ACM Computing Surveys (CSUR)*, vol. 46, no. 3, p. 33, 2014.

[3] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Pervasive computing*. Springer, 2004, pp. 1–17.

[4] M. Zhang and A. A. Sawchuk, "A feature selection-based framework for human activity recognition using wearable multimodal sensors," in *Proceedings of the 6th International Conference on Body Area Networks*. ICST, 2011, pp. 92–98.

[5] K. Yatani and K. N. Truong, "Bodyscope: a wearable acoustic sensor for activity recognition," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 2012, pp. 341–350.

[6] A. Bulling, J. Ward, H. Gellersen, G. Tröster *et al.*, "Eye movement analysis for activity recognition using electrooculography," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 33, no. 4, pp. 741–753, 2011.

[7] B. Mandal, S.-C. Chia, L. Li, V. Chandrasekhar, C. Tan, and J.-H. Lim, "A wearable face recognition system on google glass for assisting social interactions," in *Computer Vision-ACCV 2014 Workshops*. Springer, 2014, pp. 419–433.

[8] J. Cheng, O. Amft, and P. Lukowicz, "Active capacitive sensing: Exploring a new wearable sensing modality for activity recognition," in *Pervasive Computing*. Springer, 2010, pp. 319–336.

[9] O. Amft and G. Tröster, "Methods for detection and classification of normal swallowing from muscle activation and sound," in *Pervasive Health Conference and Workshops, 2006*. IEEE, 2006, pp. 1–10.

[10] O. Amft, H. Junker, P. Lukowicz, G. Tröster, and C. Schuster, "Sensing muscle activities with body-worn sensors," in *Wearable and Implantable Body Sensor Networks, 2006. BSN 2006. International Workshop on*. IEEE, 2006, pp. 4–pp.

[11] P. Lukowicz, F. Hanser, C. Szubski, and W. Schobersberger, "Detecting and interpreting muscle activity with wearable force sensors," *Lecture Notes in Computer Science*, vol. 3968, pp. 101–116, 2006.

[12] T. Deyle, S. Palinko, E. S. Poole, and T. Starner, "Hambone: A bio-acoustic gesture interface," in *Wearable Computers, 2007 11th IEEE International Symposium on*. IEEE, 2007, pp. 3–10.

[13] T. Westeyn, H. Brashear, A. Atrash, and T. Starner, "Georgia tech gesture toolkit: supporting experiments in gesture recognition," in *Proceedings of the 5th international conference on Multimodal interfaces*. ACM, 2003, pp. 85–92.

[14] C. Harrison, D. Tan, and D. Morris, "Skinput: appropriating the body as an input surface," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2010, pp. 453–462.

[15] K. Kunze and P. Lukowicz, "Dealing with sensor displacement in motion-based onbody activity recognition systems," in *Proceedings of the 10th international conference on Ubiquitous computing*. ACM, 2008, pp. 20–29.

[16] M. O. Alver and N. N. Batada, "Jabref," http://jabref.sourceforge.net/.

Human Body Models

Ilia Dobrusin
Technische Fakultät
Albert-Ludwigs-Universität Freiburg
Georges-Köhler-Allee 101
79110 Freiburg, Germany
Email: ilia.dobrusin@mars.uni-freiburg.de

Paul Gavrikov
Technische Fakultät
Albert-Ludwigs-Universität Freiburg
Georges-Köhler-Allee 101
79110 Freiburg, Germany
Email: pg74@mars.uni-freiburg.de

Murat Killi
Technische Fakultät
Albert-Ludwigs-Universität Freiburg
Georges-Köhler-Allee 101
79110 Freiburg, Germany
Email: mk497@venus.uni-freiburg.de

***Abstract*—This review describes different systems for capturing and modeling the human body posture and body motion. We present different papers on retrieving body model data from both, body worn and stationary sensors and from captured image data. Additionally some recently developed approaches are discussed. For the image based methods the HumanEVA data set as a standardized comparison method is discussed.**

***Keywords*—Body Model, Gesture Recognition, Activity Recognition, Hand Tracking, 3D Interaction, Image Processing**

I. Introduction

Reconstructing the human body from body data has many applications in science, professional use and consumer oriented products. These applications include a variety of fields such as gesture recognition in human machine interaction, motion capturing for games and movie productions, smart environments, sport motion analysis and therapy, surveillance and automotive actions in the context of machine learning [1]. To be able to reconstruct a body model data has to be captured, evaluated and translated in a mathematical model of the human body. In this review we will introduce both image and body-worn approaches and the underlying mathematical models. We discuss both full body as well as partial body models, which focus on specific parts of the body. In chapter X we will show some recent approaches by Google [2] [3] and Microsoft Research [4] using radar technology and the Doppler effect respectively.

A. Rigid Body

Most models introduced and discussed in this review use an idealization of the human body in terms of a simplified mathematical construct consisting of rigid body parts with joints with different degrees of freedom to connect the parts. By attaching sensor to those body parts or by extracting information from image data it is possible to construct mathematical representations of the human body. As part of this work several different approaches are introduced and discussed (Figure 1).

II. Body worn Systems

In this chapter we describe systems which utilize inertial measurement units (IMUs) or similar sensor based approaches to retrieve spatial information of the human posture. An IMU is an electronic device consisting of three accelerometers, three gyroscopes and in some cases three additional magnetometers and is used to calculate changes in the six degrees of freedom (translation and rotation in the three dimensions). X-Sense [6] is a system using a bodysuit with attached sensors is used to capture the body motion and can be seen as a state-of-the-art system for capturing body movements and the conversion to a 3D model of the body. The ability to use the sensor data for modeling the body depends on a calibration step, which often consists of predefined poses and gestures, allowing the systems to adjust lengths and distances of different body parts. X-Sense provides a robust infrastructure for capturing and processing sensor data and is was the starting point for further research on systems relying on sensor data.

A. Digits - Wrist-worn gloveless sensor [7]

Our hands and fingers provide a sophisticated interface for interactions with the physical world. For interactions with computer systems we mostly still rely on physical contact when dealing with inputs or instructions towards computer systems. Digits provides a system, which could reduce the necessity of physical contact and improve the way people interact with computer systems. Digits is a wrist-worn system that senses the full 3D hand pose without off-body sensors or full instrumentation of the hand. In contrast to data gloves, which cover the hands completely, Digits does not restrict natural movement.

1) Hardware setup and data acquisition : The device is attached to the wrist and is a self contained system with different sensors providing the needed information (Figure 2). The infrared laser line projector projects a thin line across the hand which intersects with the fingers as they are moved, providing 3D information about the finger position. In addition to the laser line data an array of infrared LEDs is used to

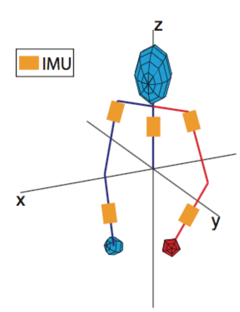


Fig. 1. [5] Simple body model retrieved from five body worn sensors.

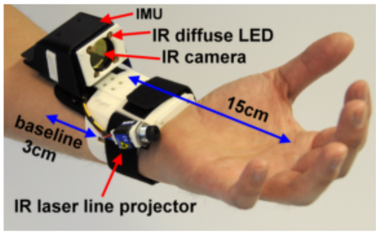


Fig. 2. [7] Digits - system overview

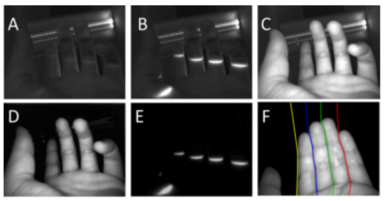


Fig. 3. [7] Background subtraction. A) Active illumination off. B) IR laser and background IR. C) IR LEDs and background IR. D) Background subtracted IR LEDs. E) Background subtracted IR laser. F) Finger separation.

illuminate the user's hand and retrieve the positions of the fingertips by subtracting the background from the images from the infrared camera (Figure 3). To increase the accuracy an IMU is added to the setup and provides additional data for possible gesture and motion recognition.

2) Kinematic model: As described in the previous section, Digits has two main sensor systems: laser line infrared sensing and an array of infrared LEDs. Both system allow separate models to be constructed. The laser line method utilizes a forward kinematic model, where the three joints of each finger are measured and combined with additional information about the natural constraints of finger movements (Figure 4).

By using the images from the infrared camera Digits is capable to perform fingertip detection by extracting depth information from the 2D images. The researchers described both methods as separate approaches and finally combined the information from those methods to a new inverse kinematic model, by building the model from the fingertips.

With a combination of the two methods to a new model the researchers were able to increase the overall accuracy of the system and create a more stable and failure proof solution compared to the systems used in a separate manner. The main advantage of the whole system is the cost efficiency and a low algorithmic complexity. Some limitations of the system occur due to being an mostly image based system (although it relies on body-worn sensors). Since images can only retrieve 2D information, some gestures could be misinterpreted when the fingers are crossed or the hand is holding an object which covers parts of palm and fingers.

B. Wrist contour sensor [8]

One of the main problems in recognition of large body movements is, that they are not good in recognition of small body movements (e.g. hand motion). So they developed a wrist worn device [8] which enable to realize some applications

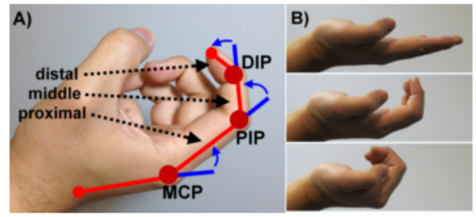


Fig. 4. [7] Illustration of index finger joints and bones.

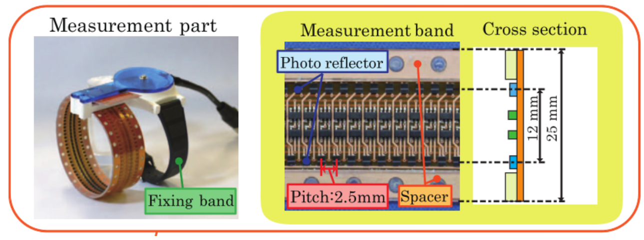


Fig. 5. [8] Measurement band with 185mm area, 2.5mm pitch and 75 photo reflectors. The fixing band assists the attachment of the device.

like remote control of some devices (e.g. home electronics).

1) Components: The device is separated in two parts: the wrist-watch-type measurement part, and the battery and control part.

- measurement band (Figure 5)
  - measurement area: 185mm
  - measurement pitch: 2.5mm
  - distance resolution: 0.1mm (3.5mm range)
  - sampling rate: 10Hz
- photo reflectors (infrared-light distance sensor)
  - 75 photo reflectors
- wireless module (communication with PC)
- battery

2) Hand recognition process: The measurement band is divided in two arrays. Each array has 75 photo reflectors. With the help of the photo reflectors the measurement band can measure the distance between the band and the surface of the wrist. Therefore the measurement band measure the changes of muscle and tendon of the human hand.
Because of small differences among the raw data of hand classes, feature extraction is essential. So we need two potential feature types: Normalized contour data, because each muscle and tendon is different in thickness and each sensor element has different variation range of distance. The feature process samples the maximum and minimum distance and normalized distance data to 0 to 1 (Figure 6).
Contour statistics, are the statistics from the data, such as sum of distances, maximum distance and so on. With the calibration data the statistics are normalized to overcome slippage or personal differences.

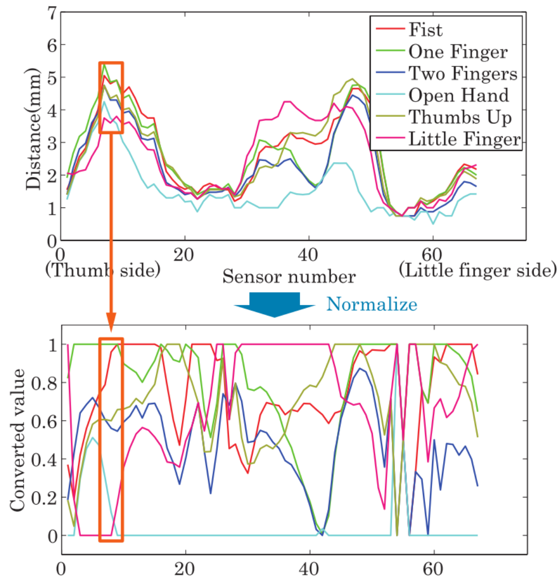


Fig. 6. [8] An example of normalized data

For the classification the k-Nearest Neighbor method and AdaBoost method are used.

C. Electromyography

Electromyography is able to measure the potential that is generated when muscles are activated electrically or neurotically.
Whilst EMG was used for quite some years in medicine to analyse disabilities and to track rehab progress, researches found more ways to integrate EMG. Another UI interface control Thalmics Myo [9] is based on EMG Measurements. But for a broader use we will look at EMGs for exoskeletons. Since muscle structures differ from muscle to muscle and muscles themselves differ from human to human TU-Berlin [10] researches focused on building a leg orthosis by just using EMG on a few muscles in combination with accelerometers and hall-effect sensors.
A complete model of the knee torque would have had involved an interplay of 13 muscles, and even more for the complete leg. For a start researchers picked the two most important muscles for flexing and extending the knee.

1) Working Principle: EMG measurements are made by a small linux machine and are being transformed using EMG-to-force functions, inverse and forward dynamics into a knee model that is then passed to a motion controller that moves the actuator accordingly (Figure 7).
The system is even capable to detect intended motion i.e.. a person might not be able to move his leg on his own, but unless he is paraplegic, his muscles will still try to move the leg. This signal can be picked up, boosted and passed. The actuator will then be able to do, what the leg cant do by its own.

2) Limitations: The body model created by this procedure contains legs with feet, shanks, tights and the torso. However

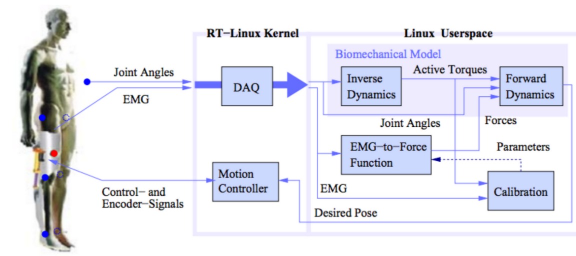


Fig. 7. [10] Electromyography - System overview

body mass, dimension and determination of the measurement point require input for each individual. Additionally each orthosis needs to be calibrated for a user.
There is a shift in time for motions, that could be disturbing.

3) Accuracy: The paper does mention an accuracy rate or 4.9 (with an standard-deviation of 5.9 and a maximum error of 15.4).
However KIT researchers [11] were able to increase the accuracy of the detection in general by using arrays of EMG electrodes and Hidden Markov Models. Additionally they were able to achieve session- and person-independence. EMG standalone recognition was improved by +25% to 33% accuracy and EMG measurements combined with IMUs were improved to 97.8%.

D. Pedalvatar

For a more realistic motion of human action Pedalvatar [12] is a human skeleton representation with a foot-rooted kinematic model using IMU sensors.

1) Kinematic model: For the reconstruction of the human pose they used a forward kinematic model in the tree structure. The condition to build up a kinematic model is, that at least one foot must be static. Therefore three cases exists to build up the model (left foot, right foot or both) (Figure 8).
With the help of orientation matrices, which are given from the IMU sensors, the model can build up step by step. The procedure is to consider a component as a parent of another component of body parts, starting with the foot as the first parent.

2) State Machine - dynamic root switch: However with this representation there exists one important problem. While the human body is in motion the static foot is always changing.
Using the angular velocity of the right / left foot (with the help of the IMU sensors) the state machine compare the value with a given threshold and can decide which foot is currently the root foot of the kinematic model (Figure 9).

E. Limitations of sensor based systems

Sensor based systems provide a variety of methods to retrieve spatial information about the human body and parts of it. Since the measured data is almost always tied to specific parts of the body, the discussed methods fail to provide information about the position of the examined body in space. One

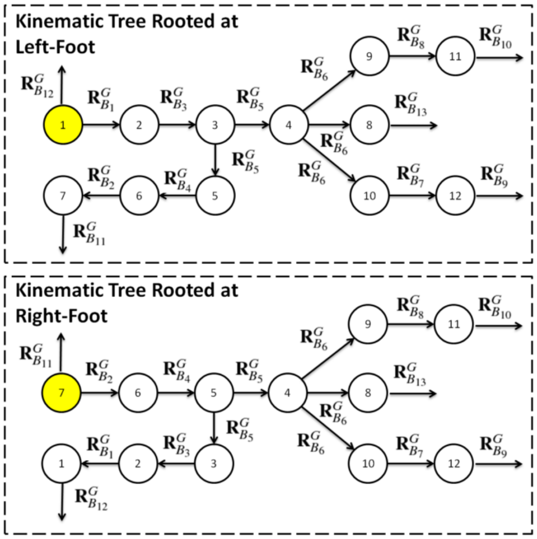


Fig. 8. [12] Forward kinematic trees rooted at different feet

$$\begin{aligned}
1: &\quad \text{if } \|\omega_L\| \geq \tau \ \& \ \|\omega_R\| < \tau \\
2: &\quad \text{if } \|\omega_L\| < \tau \ \& \ \|\omega_R\| \geq \tau \\
3: &\quad \text{if } \|\omega_L\| < \tau \ \& \ \|\omega_R\| < \tau \\
4: &\quad \text{if } \|\omega_L\| \geq \tau \ \& \ \|\omega_R\| \geq \tau
\end{aligned}$$

Fig. 9. [12] Four cases to decide which foot is the static foot of the model.

limitation is the absence of information about the clearance of the feet and whether the body has contact with the ground. In addition, body worn systems can limit the moving ability of the body (or parts of it) leading to compromised data. Since sensors and systems build with sensors get smaller with technical improvements, this factor becomes smaller over time.

III. Image based systems

In contrast to the sensor based approaches this chapter will cover image based systems. We will show a generic approach for generating human body models from image data, both static images as well as image sequences.
Additionally we will discuss two papers focusing on data sets for the usage in gesture and activity recognition.

1) Generic Pose Estimation: When creating body models from images most approaches use a variety of the generic approach which we will cover in this section.
For image based reconstruction there are different sources which can be used for construction a model. The data is either 2D, which is the most common approach, as well as 3D data (added depth information). Furthermore this data can consist of single (discrete) images for reconstructing a single body configuration, but also can be a sequence of images with a fixed or variable frame rate (typically around 30 frames per second for a off-the-shelf camera). Depending on the usage of the system, it can either directly extract configuration information as discussed in the next section, or it can use given information to predict possible subsequent positions using stochastic methods, as described later in this chapter.

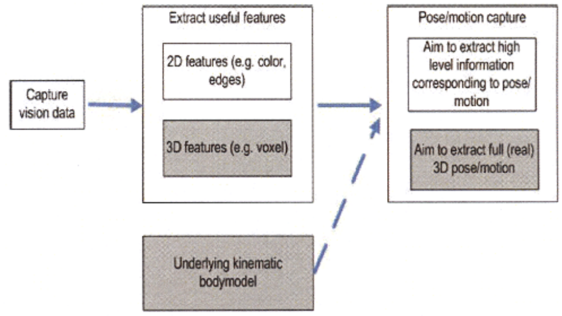


Fig. 10. [13] The generic approach for reconstructing a body configuration from image data.

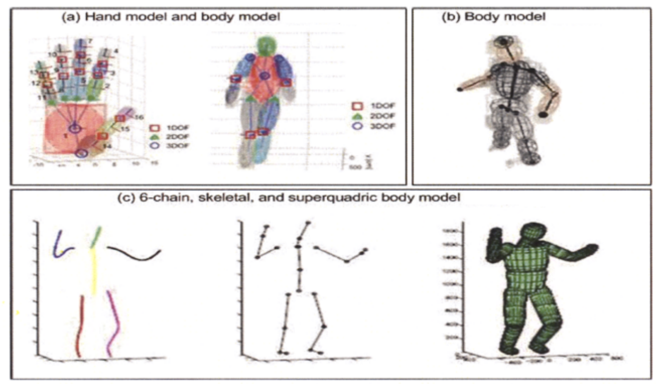


Fig. 11. [13] A set of different body models

The generic approach in Figure 10 uses 2D or 3D vision data and combines this information with a kinematic body model, resulting in a 3D pose or motion. It is necessary to reconstruct voxel data from the images. One possibility is a shape-from-silhouette reconstruction. To achieve good results a multi camera setup in combination with background subtraction could be used. Of course, this proposed method can be exchanged by any method returning similar results. After the image data is processed, we need to map the voxel data and the kinematic model to the body model configuration. This can be described by the following formula:

$$M : (Y, C) \rightarrow \Theta \tag{1}$$

Y is the voxel data, C the cinematic constraints and Theta the resulting body configuration.

There are several different body models, displayed in Figure 11. The top left models are based on a Kinematically Constrained Gaussian Mixture Model with 27 degrees of freedom (DOF) for the hand model and 19 DOF for the body model [14]. To achieve good results, the researchers use a probabilistic approach using the Expectationmaximization algorithm [15].
The top right picture [16] shows a body model using Gaussian clusters to represent the body parts. A skeletal model with Gaussian Blobs attached to the bones with additional color

information is used to break down complex movements into basic motion, enabling a real time rendering of the model. The method is limited to specified set of trained movement sequences.
The models on the bottom of Figure 11 show a probabilistic approach using the laplatian eigenspace to perform an estimation of the skeletal and superqaudric parameters. [17]

A. 2D Human Pose Estimation

In the last years the progress of human pose estimation increased, but current data sets are limited. So the researchers created a novel benchmark MPII Human Pose [18] and analysed three approaches for human pose estimation.

1) Data collection: For a wider pallet of challenges they collected images of human activities with a wider range of viewpoints (human poses, clothing types, interactions with various objects, etc).
Also consider YouTube as a data source and collect frames of human pose in videos. In the data sets are annotated body joints, 3D viewpoint of head and torso and the position of eye and nose.

2) Evaluation metrics: After the data collection they used some common evaluation metrics like the PCP metric, the self modified PCPm metric and PCKh metric for the state of art. In combination with the benchmark and the evaluation metrics they analysed the state of art of four human pose estimation. For the analysis are two full body approaches and two upper body approaches considered.

3) Goal: The performance of the human pose approaches are divided in sub-performance like body pose performance, activity performance, viewpoint performance, etc. The goal in this paper was to evaluate the robustness of the four approaches and get the existing limitations.

B. Real Time 3D Body tracking using VLMMs

Variable Length Markov Models are used to predict and / or detect body configurations in frames where body parts are hidden or not detectable e.g. because of a fast movements, whilst even saving storage in contrast to traditional Markov Models. That saving allows the system to operate in real time.

1) Working Principle: The algorithm is based on learned configurations that are extended by newly recognised features over time to enhance the detection rate. Once a frame with a complete body configuration is found, next configuration candidates (so called motion prior) are predicted and tested against the next received frame. This configuration will be learned the algorithm will continue to loop (Figure 12).
Motion prior detection is being accelerated by breaking down complex movements into elementary motions that are detected individually. Additionally, applying joint constraints on rotations can help.

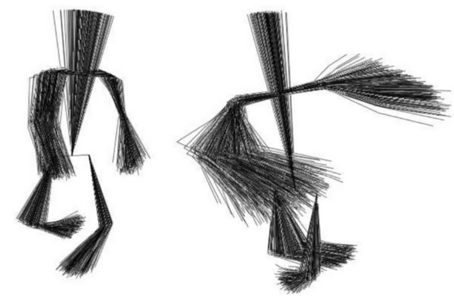


Fig. 12. [16] Visualisation of motion prior candidates

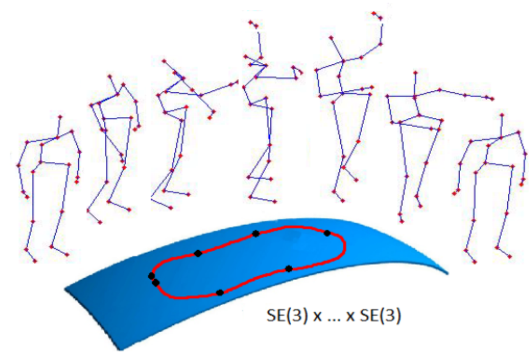


Fig. 13. [19] Human actions as curved in Lie group

2) Results: Although the paper did not mention an accuracy rate, it did state that occasional failures could be detected and be recovered from. A single 2 GHz computer was able to provide up to 10 fps with a pool of 1000 particles. The body model was represented by a 25 dimensional vector containing information about 14 body segments.

C. 3D Skeleton Representation with Lie Groups

For the past several decades human action recognition has been an important area for application like video games, robotics, etc. In this paper they developed a representation with the help of Lie groups [19].

1) Lie group: The new representation models the 3D geometric relationships between various body parts. For the modelling of the parts we need rotations and translations in 3D space. These motions are members of the special Euclidean group SE(3), therefore the representation lies in the Lie group SE(3) x x SE(3), a curved manifold. Human actions can be modeled as curves in this Lie group (Figure 13).

2) Skeleton representation: The representation of the human body is image based by using depth sensor. These sensors provide 3D depth data of the scene, which offers useful information to recover 3D human skeletons.
For the representation of the skeleton as a kinematic model we need the relative 3D geometry between different body parts. The relative geometry can obtained by using the rotation and translation of one body part to the position of the other.

Mathematically the rotation and translation in a 3D space are member of the special Euclidean group SE(3), which is a matrix Lie group. So we can model each human action as curves in the Lie group.
A Lie group is a curved manifold because of this property its difficult to classify curves, which is necessary for the action recognition. To overcome this situation we map the action curves of the Lie group to Lie algebra and build up the kinematic models of the human pose.

D. Evaluation and Comparison

When comparing different pose estimation and tracking methods, we need a method to evaluate the results and compare the outcomes. One solution could be a projection of the estimated 3D body pose to the original image, comparing the result manually. For motion tracking and motion capturing methods, it is possible to apply the body model to a virtual character model and see if the movements appear to be natural. The problem with those solutions is the lack of direct comparison to other approaches.
Every methods is evaluated by the researchers using different types of error measures, e.g. average RMS angular error on joint angles, normalized error in joint angles, silhouette overlap to compare the overall accuracy and many more. The problem with those evaluation measures is the fact that most of the data is not made public or is very specific to the method and its application.

1) Human Eva I / II: For this specific set of problems the Human Eva [20] evaluation system was created. Human Eva consists of different data sets containing image sequences from predefined actions in an idealized environment. Human Eva includes three different sets: training, validation and test, consisting of different types of actions performed by different people. It also includes a set of error measures and a base-line algorithm for comparing algorithms. This common data set is used to determine which system or method provides the best results and helps researchers to find better solutions without the necessity to build complex testing scenarios by themselves. The captured activities are available as raw image data as well as motion capturing data.

The setup of Human Eva uses four human subjects in a defined area with a 4 (Human Eva II) or 7 (Human Eva I) camera setup (Figure 14). The actions performed by subject include walking, jogging, throwing, catching, a simple gesture, a boxing gesture and a combination of all those activities. All in all, the data set contains 80000 frames and 15 on body data points. The main focus is to have natural appearance in the motions, as well as having a fully clothed subject, in contrast to other evaluations, where an idealized body with tight clothing is preferred.
The difference between the Human Eva I and II is a simplified setup with less cameras and a simplified set of actions. The Human Eva data sets can be used for a variety of evaluations between different systems and is freely available.

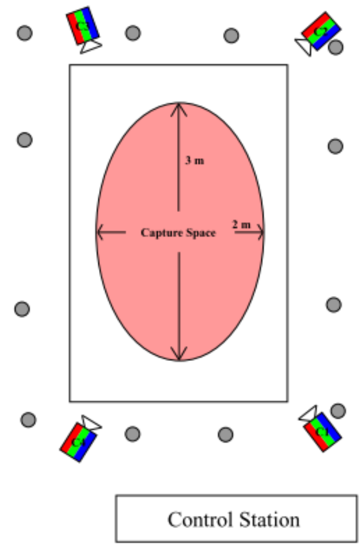


Fig. 14. [20] Human Eva II setup

Fig. 15. [21] Usage of Soli to control volume

IV. Recent Developments

A. Google ATAP Project Soli

Google ATAP [2][3][21] has revealed a hand tracking system based on radar waves on Google I/O 2015. Since there are no papers or technical documents on Google Soli, all information is taken from Googles keynote and press statements. This information might be different by now (Figure 15).

1) Working principle: A small (smaller than $1cm^2$, incl. antennas) System-on-Chip (SoC) emits 60 GHz radar waves which are then reflected from the user's hand and registered by that same SoC. Using machine learning a hand model is being calculated and passed to any third party connected via IO pins. In particular Soli uses Doppler images, IQs and spectrograms to recognise movement, velocity and distance.
Soli speed ranges somewhere between 2000 - 16000 measurements per second. The accuracy is so high, that even the natural hand shaking that every human does, is registered.

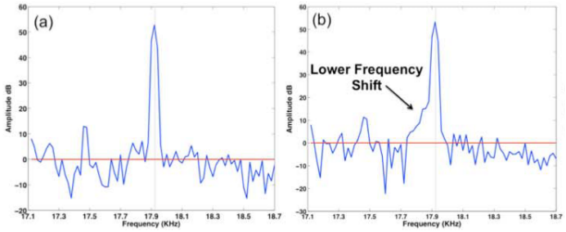


Fig. 16. [4] (a) emitted tone (b) measured frequency shift

2) Usage: Since Soli wants to be understand as a virtual context control, Google has only revealed information about gestures that can be detected. These are: dragging, rotating or clicking a virtual surface, but it is also capable of recognising hovering, finger crossing, fists and more gestures which traditional systems typically weren't able to detect. It is also notable that Soli can sense through certain materials and thus could even be placed inside a smartwatch or smartphone.

B. Microsoft Research Soundwave

Microsoft Research [4] has shown a system that is able to detect in-air gestures using nothing else than a simple notebook or PC with speakers and a microphone. Although it is not really able to reconstruct a body model, it can be used to detecting user presence and small gestures. Microsoft recommends to combine it with image recognition to maximise usefulness.

1) Working Principle: The system makes use of the Doppler effect to detect whether an object (in this case a human or a hand) is moving towards or away from the PC. Since the frequency in front of a moving object is higher than the frequency behind it, the PC is able to detect a frequency shift. This is done by periodically emitting a high frequency tone that is reflected on the moving object and received by the microphone (Figure 16).
Afterwards the software creates a spectrogram by comparing the emitted tone to the received one. A movement away from the PC will cause a shift in lower frequencies, whilst a movement towards it will cause a shift in higher frequencies. No frequency shift should be measured when no motion is present.

Combining multiple measurements allows the system to recognise small gestures like: Two Handed Pull/Push, Pull Back, Flicks, Quick Taps, Slow Taps.

2) Limitations: The frequency shift would be bigger the higher the emitted tone is - however, since Microsoft wants to use normal notebooks, they have to pick a frequency of 18 - 22 kHz to support as many devices as possible. This frequency is high and infrasonic for most adults, however children and pets might still be able to hear and be disturbed by the emitted tone. The system needs a 500ms setup, that calibrates the emitted tone frequency. The effective range is limited to approx. 1m.
Using the Doppler recognition while typing might disturb it and deliver false positives. So the system should be shut down, while a user is interacting with a keyboard or touchpad.

3) Accuracy: A three user test in two sessions - at home (approx. 45 dB SPL) and in a cafe (approx. 72 dB SPL) - with a total of 600 performed gestures, showed a average of 86 - 100% correctly detected gestures. False-positives were not included in that measurement and ranged between 2.5 to 6 per minute.

V. Conclusion

We discussed different systems and methods to retrieve body models and body configurations from both sensor and image data. Both types enable us to capture human postures, gestures and activities performed by subjects. The more recent developments show that activity and gesture recognition can simplify our methods when interacting with computers.
One common problem is the discrepancy between the ideal environment used by the researchers and the real world. Also the latency between the action of the subject and their output of method is a limiting factor for some approaches, since some applications require low latency. For systems with no such restrictions, such as sports analysis and motion capturing for games and films, higher latencies can be tolerated.
With faster and better algorithms, increasing computational power in embedded devices and new mathematical models, the human machine interaction can be improved and also lead to new input paradigms aside from classical hardware such as keyboards, mice and (touch-)screens.

References

[1] M. B. Holte, C. Tran, M. M. Trivedi, and T. B. Moeslund, "Human 3D Pose Estimation and Activity Recognition from Multi-View Videos : Comparative Explorations of Recent Developments," no. c, pp. 1–16, 2011.

[2] "http://de.engadget.com/2015/06/02/google-project-soli-neues-magisches-interface-mit-viel-prazisio/."

[3] "http://thenextweb.com/google/2015/05/29/google-unveils-project-soli-a-radar-based-wearable-to-control-anything/ ."

[4] Microsoft Research, "SoundWave - Using the Doppler Effect to Sensce Gestures," pp. 4–7, 2012. [Online]. Available: http://youtu.be/rFM59B3tYI4

[5] U. Blanke and B. Schiele, "Towards human motion capturing using gyroscopeless orientation estimation," *Proceedings - International Symposium on Wearable Computers, ISWC*, 2010.

[6] D. Roetenberg, H. Luinge, and P. Slycke, "Xsens MVN : Full 6DOF Human Motion Tracking Using Miniature Inertial Sensors," *Hand The*, pp. 1–7, 2009. [Online]. Available: http://en.souvr.com/product/pdf/MVN_white_paper.pdf

[7] Kim, O. Hilliges, S. Izadi, Butler, J. Chen, I. Oikonomidis, and P. Olivier, "Digits: freehand 3D interactions anywhere using a wrist-worn gloveless sensor," *Proceedings of UIST 2012*, pp. 167–176, 2012. [Online]. Available: http://dl.acm.org/citation.cfm?id=2380139\npapers://c80d98e4-9a96-4487-8d06-8e1acc780d86/Paper/p9825

[8] R. Fukui, M. Watanabe, M. Shimosaka, and T. Sato, "Hand Shape Classification with a Wrist Contour Sensor (Comparison of Feature Types and Observation of Resemblance among Subjects)," pp. 1–11.

[9] "https://www.thalmic.com/myo/ ."

[10] C. Fleischer and G. Hommel, “EMG-Driven Human Model for Orthosis Control,” *proceedings of the 6th annual Human Interaction with Machines workshop*, pp. 69–76, 2006.

[11] M. Georgi, C. Amma, and T. Schultz, “Recognizing Hand and Finger Gestures with IMU based Motion and EMG based Muscle Activity Sensing,” *Proceedings of the International Conference on Bio-inspired Systems and Signal Processing*, pp. 99–108, 2015.

[12] K.-c. Chan and C. C. L. Wang, “Pedalvatar : An IMU-Based Real-Time Body Motion Capture System Using Foot Rooted Kinematic Model.”

[13] C. Tran, M. M. Trivedi, C. Vision, and S. Diego, “Human body modeling and tracking using volumetric representation: selected recent studies and possibilities for extensions,” 2008.

[14] S. Y. Cheng and M. M. Trivedi, “Articulated Human Body Pose Inference from Voxel Data Using a Kinematically Constrained Gaussian Mixture Model.”

[15] A. P. D. N. M. L. D. B. Rubin, “Maximum Likelihood from Incomplete Data via the EM Algorithm,” vol. 39, no. 1, pp. 1–38, 1977.

[16] F. Caillette, A. Galata, and T. Howard, “Real-time 3-D human body tracking using variable length Markov models,” *British Machine Vision Conference 2005, BMVC 2005, Oxford, UK, September 2005. Proceedings.*, pp. 469–478, 2005. [Online]. Available: http://cms.brookes.ac.uk/staff/PhilipTorr/BMVC2005/table_of_contents.html

[17] A. Sundaresan and R. Chellappa, “Model driven segmentation of articulating humans in Laplacian Eigenspace,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 30, no. 10, pp. 1771–1785, 2008.

[18] M. Andriluka, L. Pishchulin, P. Gehler, and B. Schiele, “2d human pose estimation: New benchmark and state of the art analysis,” *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014.

[19] R. Vemulapalli, F. Arrate, and R. Chellappa, “Human Action Recognition by Representing 3D Skeletons as Points in a Lie Group,” *2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014.

[20] L. Sigal and M. J. Black, “HumanEva: Synchronized Video and Motion Capture Dataset for Evaluation of Articulated Human Motion Leonid Sigal and Michael J. Black,” no. September, pp. 1–30, 2006.

[21] “https://www.youtube.com/watch?v=0QNiZfSsPc0 .”

ESE Seminar Topic IV - Fundamentals

Tobias Paxian
Albert-Ludwigs-Universität Freiburg
Institut für Informatik
Georges-Köhler-Allee 106
79110 Freiburg
Email: paxiant@informatik.uni-freiburg.de

Martin Dold
Albert-Ludwigs-Universität Freiburg
Institut für Informatik
Georges-Köhler-Allee 106
79110 Freiburg
Email: martin.dold@saturn.uni-freiburg.de

***Abstract*—This paper examines the most influential papers in the field of wearables and ubiquitous computing of the last 18 years. It discusses the question how this can be measured, and shows some possible methods to find influential papers in this field. Furthermore the progresses in this fields are reported. Mainly discussed papers are focused on privacy, security, context and context awareness, activity recognition and location detection.**

I. Introduction

The seminar wearables of the new professorial chair embedded systems at the university of Freiburg tries to give an deeper knowledge in the areas of wearables and ubiquitous computing. Therefore the papers of the main conferences in this field shall be searched for interesting papers in different areas. This paper is the final report of two participants, Martin Dold and Tobias Paxian, of the seminar. It is the summary due to the investigation in the progress of the most influential papers of the last 18 years.

In the section II it is discussed how to find a good measurement for influential papers. What are common indices for papers having an impact to the field of wearable computing and what can be used due to time limits. The subsection II-A gives an overview how commonalities of all the selected papers can be found. To show the progress in the field, the following four sections are divided into smaller time periods. Each of this time periods presents papers of two influential themes at that time. The last section VII gives a conclusion to all the seen fields.

tb, md
Juli, 2015

II. Selection Criteria

The authors main task is to analyze the progress in wearable computing of the time frame 1998 until now by looking at the following conferences:

1) The International Symposium on Wearable Computers (ISWC)
2) Ubiquitous Computing (UbiComp)
3) International Conference on Pervasive Computing
4) IEEE International Conference on Pervasive Computing and Communications (PerCom)

The first three conferences are predefined by project specification, whereas the last one is added because of interesting papers found during research.

It is hard to give an exact number of published papers of all these conferences during the last 18 years, but there are roughly 1700 papers published during that time frame. Because of this huge paper flood we are looking for metrics to select the most influential papers over time.

Here fore, we consider the following metrics at the very beginning:

1) Best Paper Awards (by ISWC, UbiComp and PerCom):
   Elected by a committee of the conference between 2004 and 2014.
2) 10-years impact award (by UbiComp):
   Elected by a committee of the conference in 2012 and 2014 taking into account the papers of past 10 years.
3) Citation count of several search engines:
   Used as indicator as it tells how many other researchers referenced to the authors work.
   a) Google Scholar
   b) CiteSeerX
   c) Microsoft Academic Search
   d) ACM
   e) Web of Knowledge
4) h-index:
   Common used metric to rate authors.
5) Impact Factor:
   Common used metrics to rate journals.
6) Published by Prof. Dr. Laerhoven:
   As founder of the wearable systems seminar this author very likely writes interesting papers.
7) Own Interests:
   With respect to our personal opinions and interests we add some key words to the search.

Finally, the h-index and the impact factor are excluded from the selection criteria metrics as these do not fit well into our topic. As the h-index is calculated per author it is not possible for us to search for the full list of authors. On the other hand the impact factor is intended to rate journals, but not papers.

Additionally, the two authors choose an additional theme as part of personal interest “Security and Privacy” and tries to follow its evolution over time within this paper. Applying all these criteria leads to a list of 74 papers to be considered as influential.

As a summary on this chapter, the following figure 1 provides a graphical overview of the selection criteria.

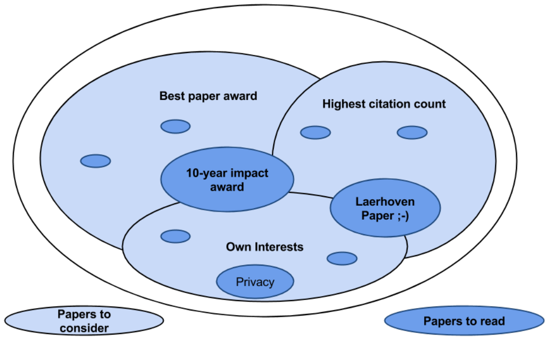


Fig. 1. Graphical overview of the selection criteria

A. Detecting commonalities

Once the papers are selected, the question arises of how to detect commonalities between the papers. The authors attempt to answer this question is stated in the following chapter.

Initially, due to the structure of the course this paper is attended to and to handle the large number of 74 papers, the time frame from 1998 to 2014 is divided into four smaller time periods:

1) year 1998 to 2003
2) year 2004 to 2006
3) year 2007 to 2011
4) year 2012 to 2014

The time periods are chosen such that each of them includes roughly 18 papers to be considered and to read. Within each time frame, the abstracts and conclusion chapters of the papers are read and summarized by key words. On the one hand, the key words given by the authors themselves are chosen, on the other hand further key words are added by the authors of this paper. As a result, each time period is covered by a list of summarizing key words. Counting the appearance of specific key words then enables us to give an estimation on what major topics and technologies discussed in the papers of each period. Based on this knowledge and aligned with the authors personal interests, two to four papers of each period are chosen to be read completely and be presented in more detail.

Accordingly, this paper is divided into the four time frames too and each of the following chapters covers one time frame. Analogously to the course, the papers presented in the weekly sessions are described in more detail in this paper.

III. Time frame 1998 until 2003

Within this chapter the time period from 1998 to 2003 is analysed for the topics of context-awareness and privacy. It is shown that two fundamental papers are published that try to provide defintions and their general position in the field of ubiquitous computing for both topics.

It can be seen that at this time frame ubiquitous computing became more and more popular. In one popular paper augmented reality came in discussion. Additionally it becomes obvious that the papers are yet 12-17 years old. The papers are about mobile phones, PDAs and Handheld computers. Nothing yet about smartphones.

A. Context and Context Awareness

The paper with the most citations, within all time frames is published at this time period. “Towards a better Understanding of Context and Context-Awareness” [1] has a total of 4116 citations at the Google Scholar search engine. This paper claims about the quantity of definitions of *context* and *context-awareness*. To get a better basement and a better comparability of papers using this terms, the authors give an overview of what context meant till that time and give a fundamental definition of *context* and *context-awareness* for future papers.

Previous definitions of context were done by enumeration of examples or by choosing synonyms for context. At that paper they do it descriptionally:

> *Definition: context* Context is any information that can be used to characterize the situation of an entity. An entity is a person, place or object that is considered relevant to the interaction between a user and an application, including the user and applications themselves. [1]

There are four primary categories of context:

- where - location
- who - identity
- what - activity
- when - time

All others are secondary, e.g. phone number or wheather forecast.

> *Definition: context-awareness* A system is context aware if it uses context to provide relevant information and/or services to the user, where relevancy depends on the users task.

The authors categorized context-aware applications into three categories.

- Presentation of information and services to a user.
- Automatic execution of a service.
- Tagging of context to information for later retrieval.

These definitions were the basis for many other papers writing about context and context-aware applications.

B. Principles of Privacy-Aware Ubiquitous Systems

In the very first time period examined within this paper there were two papers published by Langheinrich, M. that are considered to be key milestones in the topics of security and privacy of ubiquitous computing. This consideration is acknowledged by the facts that the first one published in 2001 [2] has one of the highest citation counts (744 at Google Scholar) within the given time frame. Furthermore, it was awarded with the 10-years impact award by UbiComp in 2011.

The paper "tries to serve as an introductory reading for the interested computer science researcher, especially in the field of ubiquitous computing. It gives a brief background on privacy - its history and the issues surrounding it, touches on various legal implications, and tries to develop a comprehensive set of guidelines for designing privacy-aware ubiquitous systems" [2].

With regard to the history of privacy in legislation, the author identifies the US privacy act of 1974 as one of the most influential pieces of early privacy legislation by creating "the notion of fair information practices, a significant policy development that influenced privacy policies worldwide" [2]. According to the author, a similar influential law was introduced in Europe in 1995 by "Directive 95/46/EC on the protection of individuals with regard to the processing of personal data and on the free movement of such data".

"Following the fair information practices and their recent enhancements through the enactment of the European Directive, we can identify seven main areas of innovation and system design that future research in ubiquitous computing will need to focus on" [2]. In order to discuss approaches of the privacy principles over time from 1998 to 2014, but to not mislead the reader, the following paragraph lists the principles that are discovered, improved or implemented in further papers only.

The first principle that is named *Notice* technically deals with the system of announcement and therefore tries to answer the questions of what data is collected and for which use case.

In the second section the principle of *Anonymity and Pseudonymity* is described. The authors notice that the task of anonymity is not always feasible in ubiquitous computing with existing technologies, e.g. the uniqueness of MAC addresses allows identification of devices through wireless and wired networks. Furthermore, anonymity (wherever feasible) restricts applications in some way, e.g. as no authentication is possible. For such situations, the principle of Pseudonymity is claimed to be a key concept.

The next principle is summarized as *Adequate security* whereas the meaning of "adequate" is most important key word here. According to the author we must be looking for secure and authenticated communication channels in general, but must be aware of situations where this goal is limited by technical constraints. This applies true in case of small embedded devices that are typically limited in computation and battery power. In the time frame of 2004 to 2006 this topic is recaptured in a later chapter within this paper.

In his second paper [3] published in 2002, Mr. Langheinrich introduces a proposal to a privacy aware system that underlines and implements the principles and guidelines presented in the first paper. One key element in this architecture is the *Personal Privacy Proxy (PPP)* that is a cloud application storing the user defined privacy profile. This profile holds information about which ubiquitous services a user would like to enable if offered by a ubiquitous environment. As a counter-part, a ubiquitous environment holds privacy proxies for each provided service. In case a user enters the ubiquitous environment, RFID tags announce *Privacy Beacons (PB)* that are received by the users *Privacy Assistant (PA)*, e.g. a PDA. The PA uses the information given in the PB and triggers the PPP to automatically negotiate the terms of service in the background. As a final result, a user may use the ubiquitous printer offered by the environment, but a ubiquitous camera or display service is disabled due to profile of the users PPP.

As the principles and guidelines of [2] and [3] are considered to be fundamental steps in the field of privacy and security, the authors try to follow their history over the time period from 1998 to 2014 within the following chapters of this paper.

IV. Time frame 2004 until 2006

This chapter discusses the papers published in the time frame of 2004 to 2006 and thereby provides a closer look on the aspects of activity recognition and privacy.

At this period it becomes visible that the hardware foorprint decreases over time. Now the body worn sensors are discussed as well as the activity recognition which is only possible because of increasing CPU-speed.

A. Activity Recognition

At this time period activity recognition is mentioned most often. The paper "Activity recognition from user-annotated acceleration data" [4] is with its 1679 citations in Google Scholar the most cited paper at that time frame. Additionally it won the 10-years impact award in 2014 at the UbiComp conference.

At this paper it is tried to recognize activities under semi-naturalistic conditions. Only semi-naturalistic because 20 everyday activities are given in different categories but the recognition weren't supervised and not in a laboratory. Furthermore it is not asked directly like *use the web to find out what the world's largest city in terms of population is* instead of *work on a computer*. To finally know which activities were done at which time frame, the algorithms are trained only by user labeled data.

Till that paper, the most approaches to recognize activities were done in a laboratory. Normally supervised and done with a very small activity set.

The collected data is based on acceleration data, collected with five biaxial accelerometers. The 120g sensor boards are attached to hip, wrist, arm ankle and tight. With 20 test persons and two 90 minute training sessions they used three different protocols. The first one trained the algorithm only on the first session data and tested the algorithm with the second session of the same person. Here the overall results were about 74% accuracy. The second one was trained with all subjects except one and tested at the one person left. Here the overall accuracy was about 85%. In addition to that they tried out which sensor-positions are the most important ones. There they showed that only with the two sensors tigh and wrist recognition rates 3.3% less than with all sensors are possible.

A totally different approach in recognizing activities presents: Activity Recognition in the Home using simple and ubiquitous sensors [5]. It is 910 times cited at Google Scholar and won the 10-years impact award at the UbiComp conference 2014 as well.

In that paper the authors try to recognize activities due to devices used at home. For example using the dish washer means that probably before someone has eaten something and is now cleaning the kitchen. Or by flushing the toilet it is obvious that someone has used it before.

B. Security and privacy aspects of RFID and embedded internet

Two technologies supporting the ubiquitous computing are analysed and improved regarding security concerns within this time frame.

Firstly, the paper "Security and Privacy Aspects of Low-Cost Radio Frequency Identification Systems" [6] analyses the RFID technology for technical, economical and security-relevant aspects. Being cited by 1588 other papers (according to Google Scholar), it has the second highest citation count within this time period.

At the time of publication in 2004, RFID tag costs are in the US0.50 $- US1.00$ range [6]. "To achieve significant consumer market penetration RF tags will need to be priced in the US0.05 $- US0.10$ range [...]. At this price range, providing strong cryptographic primitives is currently not a realistic option." [6]. This argument is underlined by the fact, that "Hardware implementations of symmetric encryption algorithms like AES typically have on the order of 20,000-30,000 gates" [6] whereas "a practical US$0.05 design [...] may be limited to hundreds of bits of storage, roughly 500-5,000 gates" [6].

The authors try to overcome this economical gap by introducing more lightweight cryptographic solutions that fit to RFID with regard to the wireless technology as such and to hardware footprint as well. The proposed solution is a basic security scheme based on one-way hash functions. Furthermore, two enhancements on the basic algorithm are provided, whereas one of them even serves as improved anti-collision scheme. From a technological point of view, this paper enhances the proposed privacy-aware solution [3] of using RFID tags for the principle of *Notice*. Furthermore, it addresses the principle of *Adequate security* by showing the technological upper bound of 2004, but also fulfilling it by combining feasible algorithms.

Secondly, the paper "Sizzle: A standards-based end-to-end security architecture for the embedded Internet" [7] published in 2005 presents a technological milestone in the field of embedded internet in security aspects. The committee of UbiComp shared the same understanding as it was awarded with Mark Weiser Best Paper Award in in the same year.

The researchers from MIT successfully ported a fully-featured secure webserver on embedded platforms like the Berkeley/Crossbow family (e.g. Mica2dot, Mica2 and MicaZ) that include Atmel 8-bit micrcontrollers with down to 4 MHz processing power. The key leading achievement is that the webserver implements the TLSv1.0 protocol and thus offers HTTPS service. Therefore, this paper proofs the feasibility of embedding strong cryptographic algorithms and protocols on hardware and processing power limited platforms, e.g. with less than 50 kByte of flash memory and 4 kByte of RAM. This opens a new perspective in the field of embedded internet as it will allow further ubiquitous services running even on tiny microcontrollers but still allowing strong encryption and authentication methods. With respect to Mr. Langheinrichs principles of privacy [2], this paper addresses the principle of *Adequate security* such that it enlarges the set of feasible security algorithms for ubiquitous devices.

V. TIME FRAME 2007 UNTIL 2011

This chapter summarizes the influential papers published within the time frame of 2007 to 2011 whereas the topics of event recognition and privacy are emphasized.

A. Electrical Event Detection

How to sense through an existing infrastructure in a home. "At the flick of a switch: Detecting and classifying unique electrical events on the residential power line" [8] tries to answer this question. It has a high Google Scholar citation count for this time period of 295. Prior works tried to use the plumbing infrastructure to infer basic activities via microphones. Others tried to localize subjects indoor via the residential power line. This work not only tries to localize persons indoor, but also tries to sense which electrical powerline device is used actually.

Applications for this can be seen in the section of energy monitoring, by logging which devices are switched on for how long. Other applications could be home automation, if someone switches the light on in a room, the settings of the room can be adjusted, such as music or temperature. Even healthcare for elder people can be a point, by studying the everyday activities of a person. This can be used to recognize variation in the activities and could be a good indicator for diseases like Alzheimer.

For sensing through existing infrastructure, for each home only one single powerline plug-in-sensor is needed. The sensor consists of an custom powerline interface, an USB Data Acquisition Oscilloscope and a Laptop to record and evaluate all the events. The Dual Input Oscilloscope is bandpass filtered and has a sampling rate of 100 million samples per second.

The task then was to classify electrical noise. It can be differentiated between resistive loads and inductive laods. Resistive loads are a transient noise pulse when for example a light is turned on or off. Inductive loads on the other hand are from either mechanical switching like from a motor which produces continuous noise signal until it is turned off, or from solid state switching like from a power supply for Personal

Computers. It is extremely difficult to predict the transient noise - therefore the algorithm is trained with about 80 events in each of the six homes.

The overall results show that between 85% and 95% of all events can be classified correctly. But it is important to train the algorithm several times during the testing period of several weeks. Several limits for the event detection are finally discussed. Firstly it is difficult to detect all events in industrial sizes. If the building becomes too big, more precisely if the power line length becomes to long, the signals are too weak to detect them all with only one sensor. On the other hand compounded events are difficult to learn and similar events which are close to each other in the sense of same power line length are difficult to differentiate as well.

B. Inference attacks on GPS location tracks

With regards to the topics of security and privacy an interesting and influencing paper [9] was published by Krumm, J. in 2007 called "Inference attacks on location tracks". At the time of writing this paper its citation count at Google Scholar search engine is 331 and is thereby the highest citation count within the given time frame from 2007 to 2011. Target of the paper was to quantify the risk of such attacks based on measured GPS data points. In order to do so, the authors did a user study including 172 individuals that provided their home location and identity at start up to compare the later results. GPS receivers were installed in the cars of the volunteers, recording GPS data points in a median interval of 6 seconds and 64.4 meters. After two weeks of recording the authors then analysed the data set using different techniques in a stepwise approach. Furthermore, an investigation on countermeasures to the inference attacks is done and several countermeasures are analysed too.
The first task of the stepwise approach was determining the home location out of the timestamped GPS data traces. Four different inference algorithms were used on the data set where each of them returns a single GPS coordinate as a guess of the subjects home. These are namely:

1) *Last destination:* This algorithm is based on the heuristic that the last destination of the day is often a subjects home.
2) *Weighted median:* Here, it is assumed that the subject spends more time at home than at any other location and therefore each coordinate in the survey is weighted by the dwell time at that point.
3) *Largest cluster:* This heuristic assumes that most of a subjects coordinates will be at home.
4) *Best time:* It learns a distribution over time giving the probability that the subject is at home.

As an intermediate result, the authors showed that the first three algorithms guess roughly 11 percent of the subjects homes correctly, whereas only 3.5 percent of the returned data by the *Best time* algorithm is correct.

In a second step, the home locations (returned as GPS data points in the first step) are passed to the Windows Live Search engine in order to identify the volunteers name and address.

As an overall result of both steps, a total of roughly 5 percent of the subjects identities are analysed correctly using the *Last destination*, *Weighted median* and *Largest cluster* algorithm in the first step. In case of *Best time* algorithm a total percentage of only 1,2 percent of the subjects are identified correctly.

The authors analyse three key problems as reason for the comparably small number of correct results. Firstly, the data given by the GPS receivers are claimed to be inaccurate such that the inference algorithms would perform better in case of more data sets with higher resolution. Secondly, the database that is used in the second step of reverse geocoding might be outdated and/or inaccurate as well. The third argument is the behaviour of the subjects itself as the parking locations of the cars distant from home locations. This argument becomes even stronger in case a car is parked to multiunit buildings that enlarges the set of possible identities living in this area.

Based on the experience gained in the observation, the following three countermeasures are presented:

1) *Spatial cloaking:* A circle of radius r is chosen around the subjects home location and all GPS data points within this circle are deleted.
2) *Noise:* Adding Gaussian noise with 50 meter standard deviation.
3) *Rounding:* Each point is snapped to nearest point on a 50x50 meter grid.

The authors "are trying to find how much we have to corrupt the GPS data for the three countermeasures to significantly reduce the number of correct address inferences" [9]. As a final result it is shown that the radius r of the *Spatial cloaking* must be chosen to 2000m in order to drop the number of correct addresses to zero. For the *Noise* and the *Rounding*, a standard deviation and a discretization of 5000m is required to get the same result.

Linking this paper to the principles of privacy [2] it can be seen as a further improvement and field study regarding the principle of *Anonymity and Pseudonymity*.

VI. Time frame 20012 until 2014

This section analyses the last time period of 2012 to 2014. It is shown that the topic of security is still of note by presenting a solution on secure bootstrapping of ubiquitous displays. On the other hand, a hardware and prototyping relevant papers is shown that discusses the new research field of circuit printing.

A. Instant Inkjet Circuits

2013 UbiComp honours the paper "Instant Inkjet Circuits: Lab-based Inkjet Printing to Support Rapid Prototyping of UbiComp Devices" [10] with a best paper award. It was 42 times cited according to Google Scholar, which is quite high for a paper less than two years old. For that time frame it was additionally the second highest citation count within that time frame. The paper shows a cheap way to print circuits with a cheap inkjet printer. The same professorship even opened an actual kick-starter project[1] which is funded with more than 80.000$ yet.

Until that approach inkjet circuit printing was only possible with expensive printers priced several tens of thousands of dollars. Furthermore a sintering process at more than 150 degrees for several hours is necessary. Another way to get a circuit relatively fast and convenient are Printed Circuit Boards (PCB) milling machines. They need about one hour, but have a noisy and messy production process. Milling flexible substrates was quite difficult. Vinyl cutting machines are a third possibility. They have a cheap initial investment of round about 200$ plus 10$ per meter for the film. But it is relatively time consuming and thin traces can easily break.

The scientists used an printer manufactured by Brother Co because they have nozzles which eject higher volumes of ink and leave therefore a greater amount of conductive ink on the paper. They used the printer Brother DCP-J140w, for 77$, because it was the least expensive model available on amazon at that time. In addition they bought empty cartridges for 9$ and silver nanoparticles ink. The nanoparticles were smaller than 0.1 micrometer, throughout experimentation they found out that the Mitsubishi Paper Mill ink were the best for their purposes. It costs 50$ per m^2, or 5 Cent per meter for a 1mm wide trace.

The circuits are printed best to chemically coated paper to absorb ink effectively and prevent smearing. The surface has to be rough to establish nano-scale conductive structures. Glossy photo paper worked fine for that. Transparent or white PET (polyester) film from Mitsubishi Paper Mill is used as well. To get the most possible ink on the surface the ink is loaded in all of the CMYK cartridge positions. Printer settings are adjusted to best print quality, color mode vivid and color density is set to +2. Line thickness down to 0.25 mm are still evaluate good results with these settings.

Connecting components with a soldering process is unsatisfactory due to temperature. Due to that two solutions are presented. Conductive tapes allow interconnection between substrates through the adhesive thickness (z-axis). It is strong enough for most prototyping issues, but if a stronger connection is needed, the use of silver epoxy is suggested.

It looks very promising for developers to print their own circuits at home easy, cheap and fast. To get this work even more easy the Kickstarter project were launched.

B. Secure Bootstrap of Ubiquitous displays

In the last time period from 2012 to 2014 the field of security is addressed by the paper "Secure Bootstrapping of Cloud-Managed Ubiquitous Displays" published by Sethi, M. et al. published in 2014 [11]. At the time of writing this paper, the citation count at Google Scholar was still 0, but the potential influence is shown by the Best Paper Award given in 2014. The paper can be grouped into four different items that are stated in the following section.

[1]https://www.kickstarter.com/projects/1597902824/agic-print-printing-circuit-boards-with-home-print

Firstly, the authors provide an analysis of the current technological status on the initial setup of an ubiquitous display and thereby define the key problems to deal with. In particular, the problem is described as a two-phase configuration of the device: In a first step, the device must be configured to access a network that provides access to the internet. E.g. in case of Wireless-LAN (that is expected to be one of the most common used networks) the device must be configured to join a specific WLAN SSID and the matching password must be entered. In a second step, once access to the internet is granted, authentication to a cloud service is required that is typically done by some user name and password as well. Generally speaking, an user input is required whereas an user may be any technician or non-technician that shall power up the device for the very first time. The problem here results in the fact that many of the displays have very limited or even no input capabilities. Furthermore, when considering the large number of expected ubiquitous displays, it appears reasonable to look for a fast bootstrap solution that requires less user-interaction.

Therefore, the authors offer a solution architecture on secure bootstrapping that requires no input capabilities on the devices at all. Key element of the proposed solution is to display a QR-code on the device screen after power up, that can be scanned by the user. The QR holds default information about the device e.g. a (manufacturer) specific ID, an URL to automatically connect to and a random number that is generated new after each power up. From a security point of view, a standard Diffie-Hellmann key exchange between the cloud service and the device is performed whereas the random number is used as out-of-band channel authentication within this process. The authors successfully managed to combine different state-of-the-art technologies into their solution such as WPA-EAP and RADIUS with respect to security and HTML5/CSS3 to bring any content to the display.

In a third step, the authors setup a prototype of their proposed solution and even include a user study using this prototype. Main output of this study is that the users (a mix of technicians and non-technicians/inexperienced) felt like they actually did "nothing" to securely boot up the device. One may follow from this fact, that the proposed solution is somewhat bullet-proof as even unexperienced user can handle it. Moreover, the authors emphasize that the solution is even working for other devices like printers, RFID-tags, speakers etc. Generally speaking, it works for any device type that can display the content of the QR code to the user.

Mapping this paper to any principle of privacy-aware systems [2] appears to be impossible as it addresses a topic that is actually not covered by [2] or [3]. Nevertheless, it can be seen as an approach in supporting the migration of large ubiquitous systems that is not considered by Mr. Langheinrich yet.

VII. Conclusion

Target of this paper is to examine the most influential papers in wearable and ubiquitous computing over the last 18 years from 1998 up to now. To do so, we evaluated different types of metrics on quantifying the potential influence of papers.

Additionally, a reasonable set of selection criteria was chosen to handle the paper flood of roughly 1700 papers published within the past years. With respect to the personal interests, privacy and security was added to the key words of the selection criteria.

In the research field of privacy and security, it is shown that the papers of Mr. Langheinrich ([2] and [3]) defined principles and guidelines of privacy aware systems for the first time. Therefore, the influence of the papers to this research area is considered very high. Apparently, this understanding is not only true for the authors of this paper, but also for the committee of UbiComp conference, that recognized Mr. Langheinrichs studies with the 10-years impact award. In the following, the authors study and present further papers of the privacy topic. Each of these papers is mapped to the privacy principles and their improvement for the specific field is determined. In detail, it is shown that major steps are done in 2004 ([6]) and in 2005 ([7]) to improve the meaning of *Adequate security* principle and to lift the upper bound limits of security implementations in hardware and software. In 2007, [9] analyzed and showed the danger of inference attacks on GPS location tracks and provides countermeasures to protect the users identity and thereby its privacy. In the last time period, the paper [11] even goes beyond the scope of the privacy principles by presenting a fast and secure bootstrap mechanism that improves the migration process such that migrating a large number ubiquitous devices appears to be realizable.

The terms “context” and “context awareness” needed a general definition other papers could use. That’s the reason why “Towards a better Understanding of Context and Context-Awareness” [1] got the highest citation count at all.

Other topics like activity recognition became more relevant due to decreased hardware footprint around the year 2004. Furthermore “Activity recognition from user-annotated acceleration data” [4] tried out a new approach how to record data in circumstances as naturalistic as possible.

A new idea how to detect events was reported in 2007. “At the flick of a switch: Detecting and classifying unique electrical events on the residential power line” [8] tried out to detect events using the yet implemented infrastructure of a house.

“Instant Inkjet Circuits: Lab-based Inkjet Printing to Support Rapid Prototyping of UbiComp Devices” [10] won a best paper award with a cheap approach to print circuits.

In general it can be said, papers with new ideas are cited most. If they are the first ones to write down the approach in an “easy to read” paper. Let’s be curious where the research field will develop in the next decades.

REFERENCES

[1] A. K. Dey and G. D. Abowd, “Towards a Better Understanding of Context and Context-Awareness,” *Computing Systems*, vol. 40, no. 3, pp. 304–307, 1999.

[2] M. Langheinrich, “Privacy by designprinciples of privacy-aware ubiquitous systems,” *Ubicomp 2001: Ubiquitous Computing*, pp. 273–291, 2001. [Online]. Available: http://link.springer.com/chapter/10.1007/3-540-45427-6_23

[3] ——, “A privacy awareness system for ubiquitous computing environments,” *UbiComp 2002: Ubiquitous Computing*, pp. 237–245, 2002.

[4] L. Bao and S. S. Intille, “Activity Recognition from User-Annotated Acceleration Data,” *Pervasive Computing*, pp. 1 – 17, 2004. [Online]. Available: http://www.springerlink.com/content/9aqflyk4f47khyjd

[5] E. M. Tapia, S. S. Intille, and K. Larson, “Activity Recognition in the Home Using Simple and Ubiquitous Sensors,” *Pervasive Computing*, vol. 3001, pp. 158–175, 2004.

[6] S. Weis, S. Sarma, and R. Rivest, “Security and privacy aspects of low-cost radio frequency identification systems,” *Security in pervasive computing*, pp. 201–212, 2004. [Online]. Available: http://www.springerlink.com/index/YVMFPKWC9NQ6HQDW.pdf

[7] V. Gupta, M. Wurm, Y. Zhu, M. Millard, S. Fung, N. Gura, H. Eberle, and S. Chang Shantz, “Sizzle: A standards-based end-to-end security architecture for the embedded Internet,” in *Pervasive and Mobile Computing*, vol. 1, no. 4, 2005, pp. 425–445.

[8] S. N. Patel, T. Robertson, J. A. Kientz, M. S. Reynolds, and G. D. Abowd, “At the Flick of a Switch : Detecting and Classifying Unique Electrical Events on the Residential Power Line,” *UbiComp*, pp. 271–288, 2007.

[9] J. Krumm, “Inference Attacks on Location Tracks,” *Pervasive Computing*, vol. 10, no. Pervasive, pp. 127–143, 2007. [Online]. Available: http://www.springerlink.com/index/TG64551RW2716103.pdf http://research.microsoft.com/en-us/um/people/jckrumm/publications 2007/inference attack refined02 distribute.pdf

[10] Y. Kawahara, S. Hodges, B. S. Cook, C. Zhang, G. D. Abowd, and G. D. A. Yoshihiro Kawahara, Steve Hodges , Benjamin S. Cook , Cheng Zhang, “Instant Inkjet Circuits: Lab-based Inkjet Printing to Support,” *Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing*, pp. 363–372, 2013. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2493432.2493486

[11] M. Sethi, E. Oat, M. D. Francesco, and T. Aura, “Secure Bootstrapping of Cloud-Managed Ubiquitous Displays,” 2014.

Wetlab Support Systems

Florian Wolling
Faculty of Engineering
University of Freiburg
Georges-Köhler-Allee 101
79110 Freiburg, Germany
wollingf@tf.uni-freiburg.de

Jan Donath
Faculty of Engineering
University of Freiburg
Georges-Köhler-Allee 101
79110 Freiburg, Germany
donathj@informatik.uni-freiburg.de

Johannes Güttler
Faculty of Engineering
University of Freiburg
Georges-Köhler-Allee 101
79110 Freiburg, Germany
guettlerj@informatik.uni-freiburg.de

***Abstract*—Wetlabs in general are not known to be a place of using supportive systems so far. Having a look at a wetlab environment may reveal that most of the work is done the same way for years. Using notebooks to document every step by hand, taking materials to evaluate them later and so on.**
This report shows how we found papers dealing with different stages of experiments in wetlabs and whole laboratory setups. Furthermore we have a closer look to what a wetlab environment consists of, what kind of supportive systems have been developed within the last years, why they are used or why not and which general problems occur.
To evaluate this topic we used the most common search engines that are addressed to fields where wetlabs are established.
We present sample implementations of supportive systems developed for the use in wetlabs and biological laboratories, state their advantages and disadvantages and how the users accept them.
***Index Terms*—Wetlab support, biology, laboratory, bench, hardware, preparation, documentation, collaboration, automation**

Introduction

Initially, no member of our group had any or only few experience of working in a wetlab environment. We had a rough idea of how wearable and embedded devices could be of use within this setting. So our first step was to inform ourselves about what wetlabs exactly are. This research included what materials are used, the workflow of experiments and which problems may be caused by using devices and especially computers with user interfaces in a laboratory environment. Here we sketched our first questions what problems and conditions have to be considered.
Afterwards we had another closer look on the initially suggested papers about the eLabBench [1] and the Labscape [2] where two exemplary environments where presented. We also recognized that these papers reference papers describing fundamentals of wetlab work and additional information. So we decided to have a look on the references whether they could be of use to get information about more supportive systems.

Research

Finding work about supportive systems to be used and being tested in a wetlab environment was a challenge. The first source for our research was Google Scholar [3]. The main reason for this choice was the amount of available papers in the database and the availability of not only one certain category. So we were able to collect papers from chemical conferences as well as from those specialized in wearable computing.
We also considered search engines bound to a certain field where wetlabs or wearable devices are likely to be used. Those engines consisted of:

- IEEE Xplore [4]
- PubMed / NCBI [5]
- ScienceDirect [6]
- Berkeley OskiCat [7]

But searching for work using the keyword 'wetlab' mostly presented results about work that has been done within wetlabs or descriptions about new approaches on how to do a certain experiment. Even the number of hits when searching for wetlabs was way too high to be a use for further research. Google Scholar had a result count of about 15,000. And even with a refinement using keywords like 'smart' and 'support' the number of hits only decreased down to about 9,000.
Also referencing other engines gave us a big amount of hits but with different emphases. So we checked different papers from all the engines but recognized that topics concerning wetlab supportive systems were not common.
But after all, this gave us a broad view about all the different sections concerning work in wetlabs. This resulted in a classification of subtopics where we thought it might be a benefit to have a closer look onto small areas instead of wetlab setups as a whole. We identified the following items within a wetlab:

- Preparation
- Hardware
- Analysis
- Documentation
- Collaboration

Using a separation like this resulted in more specific results with a manageable amount of papers per topic. Including all items we started with a selection of 19 papers from various conferences. Here, the topics about whole setups and hardware were the ones with the highest count of matching papers.
Finding papers about documentation and analysis was a way more complicated because most of the papers we found dealed with documentation or analysis in general and were not referenced to the use especially in wetlabs.
The most difficult parts to find any sources for were collaboration and preparation. The main reason is that these topics are seen as standalone systems not directly referenced to wetlab usage or being part of a whole setup.

Figure 1. Example of existing support system: Pipettor robot 'Andrew' [8].

HARDWARE

There are many suggestions for supporting hardware in wet-labs. Mostly it is part of a bigger support system. This chapter deals with existing and commonly used supportive systems as well as with approaches using stationary, portable or wearable devices.

Existing Hardware

Big labs often use general support systems like conveyor belts which we will not describe here because they were not designed for this purpose. Other systems for big labs are decappers, centrifuges with automated loading, analysers, pipettors and so on.
A very common system for biology and medical labs is an automated PCR (polymerase-chain-reaction) which is used to duplicate gene-fragments. The System is very simple to use and saves the laboratory worker from losing a lot of time by doing the PCR manually. Other important devices are real-time-analysers, for example for blood-gas analysis, which are not only used in labs but also in operation rooms.

Portable and Wearable Devices

The will to document research during the work in laboratories in contrast to doing all the documentation afterwards in the office leads to using not only lab-notebooks but also laptops and tablets if applicable. The big disadvantage of these devices is that one not only could contaminate the device and take the contamination out of the lab. But as the device often does not stay in the lab, it might bring contamination into the lab and messes up the experiment [1]. Thus all portable devices should never leave the laboratory.
To reduce physical interaction with devices it might be reasonable to use wearables for tasks where the user does not want to use non-stationary device [9]. In [10] the use of Google Glass in biological laboratories is studied, setting focus on the support for novice researchers. The new scientists could display experiment instructions and partially use hands-free interaction with Glass. Nevertheless for some tasks touch-input is necessary which reduces the field of application for these features to low-risk laboratories. A combination of wearable devices is suggested in [9]. Here the authors present a combination of a head mounted display, a Smart Watch to measure the acceleration and a RFID reader worn under a glove to support not only the documentation process but also to enable the researcher to access previously recorded data.

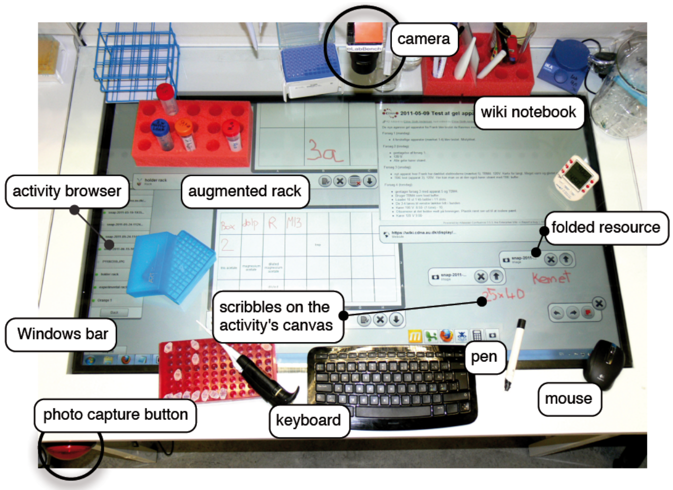


Figure 2. Laboratory setup of the eLabBench [1].

Stationary Devices

We found two main approaches to support laboratory workers with stationary systems. One is to rarely modify the working environment and to coexist beside the existing workflow. The other one is to integrate the whole workflow into the system. An example for the first is the Labscape environment [2]. In the described setting a workplace of a researcher is equipped with a touchscreen along with a mouse and a keyboard. For some tasks it could be useful to have a barcode- or RFID-scanner. Because the work often takes place in different places within the lab, several workstations are distributed over the whole possible working area.
Another approach for a stationary system is the eLabBench [1], a big touch-display replacing the labbench and supplying some features to the user. Running a Microsoft Windows system, the user can use all kinds of Windows applications besides the eLabBench software. For the input it is equipped with a pen for the touch-screen besides the common mouse and keyboard-input. In this way the user has the possibility to scribble notes directly on the screen. Another feature of the system is a camera targeting the bench's surface. A big photocapture button allows the user to use the camera hands-free if necessary.

DOCUMENTATION

Every experiment, no matter if it has a biological topic or not, is in vain without a proper documentation. The effort to produce meaningful results is for nothing if the documentation can not show the consequential aspects and the reviewers can not follow and reproduce the experiment.

In a classic experimental workflow, documentation is the final stage after the actual experiment which is introduced by a preparation. A traditional documentation is manually written in natural language. This leads to a high influence of subjectivity and uncertainties. Usually the final documentation is transcribed from a bunch of working documents and notes to a lab book several days after the actual experiment took place. To publish the argumentation, these documents are extracted again. Thereby in every stage details are omitted according to the transcribing process and the repeatability and confirmability gets worse. [2, p. 18-19]

According to a "shift from an analytical natural science to a design or engineering science" [1, p. 1], biological science now handles with projects that mainly consist of a loop where a product is improved until it meets the predefined requirements. In each round of such a loop and also in each similar type of experiment, standard procedures are executed that do not differ and do not need to be described repeatedly in every documentation. Only the differences between the common standard and the chosen way need to be described. In some cases scientists also tend to produce too much text and overhead for only small improvement which would simply require an annotation. Thus a reduction of information helps to focus on the important adjustments.

In all mentioned cases a supported workflow in combination with a supported, formalized documentation can help to extract, to manage and to provide the essential information.

Supported Workflow

The principle of a supported workflow integrates the documentation into the preparation and execution stages of a traditional workflow. It intends to lower the barrier between office and laboratory area and to make the documentation process more incidental and parallel to procedural work.

Therefore the laboratory needs to be equipped and upgraded, as stated in the previous section. The experimental procedures and the documentation as well need to be at least partly automated so that the scientist is not distracted of the scientific process and is allowed to concentrate on scheduling and the information management. While the automation of standard procedures is already common, even in smaller laboratories, an appropriately automated documentation needs libraries of definitions and documentation fragments. On the one hand these information can be provided by the manufacturer of a device. The hardware then represents a closed procedure with certain parameters. On the other hand modified, self-developed or manual procedures can be documented and shared by scientists in a social network. These networks are one kind of collaboration that is stated later on.

Formalized Documentation

For formalized documentation standard interchange languages and data formats are used to archive information. It is not needed to manually code, the files are automatically generated by appropriate devices that are connected in a network.

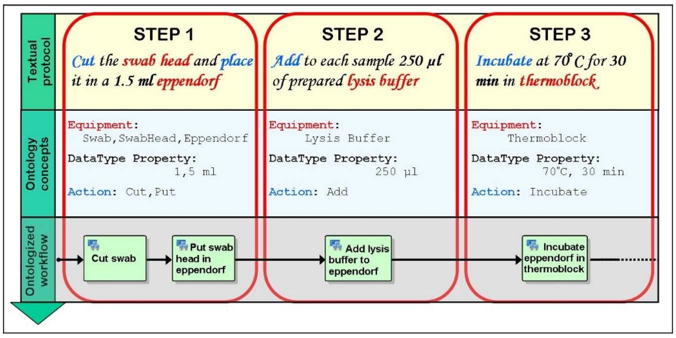


Figure 3. From textual protocols to COW formalized documentation [11].

The documentation consists of two levels. The first level contains all collected information, even procedural data like ambient temperature, in a form that allows machines to automatically process it or just to repeat the experiment. These information are not of direct use for scientists but allow e.g. to reconstruct the ambient conditions and to find causes of failures. The second level above packs the directly relevant information in a human-readable form and is used for reviewing and as a source of inspiration for further research. If possible the information is illustrated in diagrams, graphics or photos but if this is not possible or reasonable it bases on written language, comparable to natural text. Nevertheless it reduces the overhead compared to manually written reports.

There are two types of work that are formalized using two different strategies. On the one hand there are individual operations, called entities, that are defined in an ontology. On the other hand there are manual or automated procedures that are combined in a workflow. A third and new approach combines these strategies to design formal protocols, called Combining Ontologies with Workflows (COW). [11]

Ontology Strategy

Ontologies describe single entities, individual operations. They are hierarchically set up and based on actions that lead to defined goals. Their simplified structure automatically reduces ambiguity and redundancy of human-written reports and allows computers to detect structural or logical errors and also allows automated reasoning. [11]

Workflow Strategy

Workflows are a flows of instructions in a receipe style. They contain a set of activities that are combined by an execution controller. The controller allows sequences, parallelism, choices and synchronization of its entities. [11]

Combining Ontologies with Workflows

Figure 3 shows an illustration of how the combination of ontology and workflow strategy (COW) works. A textual protocol is filled with connecting words that do not contain any important information but allow a fluently reading by

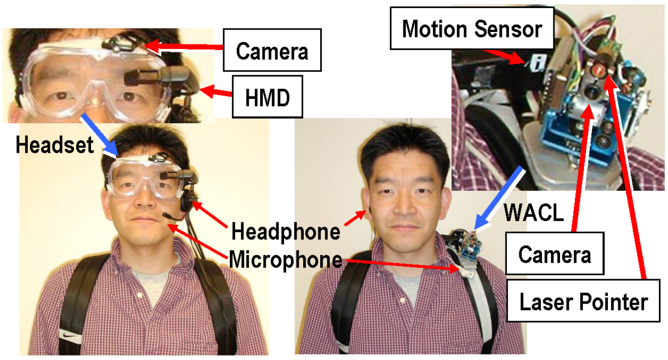


Figure 4. Human-mounted hardware for User-Expert Collaboration [12].

humans. Ontologies break these connections up and just define objects, quantities and relations or actions that lead to a certain goal. COW then uses the resulting ontologies as entities of a workflow to provide a step-by-step schedule.
COW is a logical consequence regarding to both strategies in global and not only each one out of its point of view. It allows the reuse of ontologies or even workflow bricks or templates that can be shared over social media and other collaborative networks. This minimizes the preparation and reduces the time to start with the actual experiments. The formalization and additional highlighting of critical steps helps the user to focus on important information and helps to avoid reinventions by just automatically checking equalities of projects. But to develope the ontologies for a futur use in workflows is labour intensive and together with an upgrade of the laboratory equipment it is expensive. The generated reports are uninspiring and beside a high page amount they contain a lot of white space. Especially loops or more complex constructs are not readable by humans anymore. [11]

COLLABORATION

Finding references to collaboration tools or hardware was not that successful. It is common to use systems like Skype, Dropbox or versioning systems to communicate and share files with each other.
But also these approaches have some withdraws. Talk to each other, getting help or guidance while doing an experiment is not that comfortable.
Using laptops to be guided or to demonstrate is circuitous because it may be rearranged when changing the view or the working place.

User-Expert Collaboration

That is a circumstance which Kurata, Takeshi, et al. wanted to improve by introducing their shoulder-worn active camera/laser [12]. It is worn on the body and project the view of the experimentalist to another person. The setup consists of a camera, a freely rotatable laser mounted to the shoulder of the user in the lab, a microphone and headphone to communicate with each other. The remote person is able to rotate the laser and use it to point to things or places. Even the camera is rotatable so that the remote user is able even to look elsewhere than the experimentalist.
Another implementation to antagonize the lack of collaboration within a laboratory environment consists of a head-mounted display, camera, microphone and headphones. Even here the remote person is able to get the view of the experimentalist and both are able to speak to each other. The display is used to show images to the experimentalist. These can be images captured by the camera and anotated by the remote person or images showing additional information or sketches. Both implementations are meant to be used productively in lab environments. But both need time to get used to it or are extinguished as too uncomfortable.
By introducing new hardware or processes they should be established without additional effort towards the users.

Group Collaboration

Another kind of collaboration we found is the sharing of information and data with a larger amount of people. Formerly it was usual to inform yourself about new methods and results by reading the published papers of others. But these only contain the final results and data and is written to understand the principle behind the work but not the exact process. It is also hard to reconstruct the results because the raw data is not freely available.
Therefore it is necessary to establish a system to document every single step that was made, saving all data and having the property to share certain or all details with others.
This is not only useful for others to reconstruct but also for the publisher to get additional ideas and feedback.
One system having these abilities is called *Prism* [13]. Unlikely other systems where people implement something thinking someone could use it somehow, the authors of *Prism* first run several studies to have an idea what elements could be useful and are necessary.
After observing biologists at work they started implementing an online environment which allows to link different streams to each other. These streams contain e.g. online content, personal notes, emails, binary files or calendars.
While developing this environment it was desired that the users send feedback actively and all the time. During the main

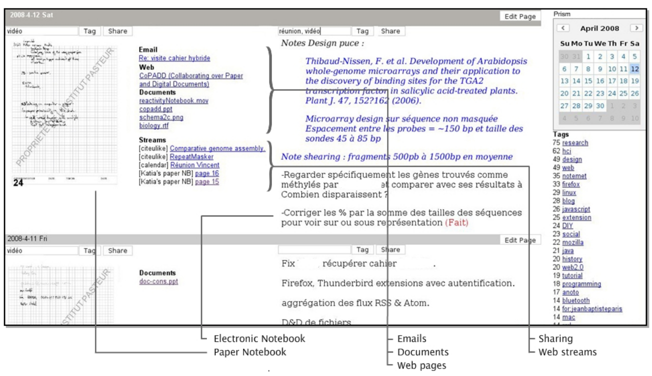


Figure 5. Sample view of the Prism environment [13].

developing phase all interactions were logged and evaluated to improve the system with respect to the users' wishes.

This concept emphasizes the whole project because the environment was build around the users' workflow. This allowed to establish the system without additional introduction.

But not only people using this system benefited from it. Using standards also others are able to access the shared contents. This allowed other groups and researchers to retrace every step at every point of time, to get the state of the processed data after every step and to use data to do their own researches or just to repeat and to review the results.

Conclusion

After examining several approaches withing the different subcategories it figures out one basic problem: Comfort. All solutions need time to get used to it, are unintuitive, too complex or just obtrusive. But a benefit for a scientist is a system that does not consum additional time and money without a quick success.

Solutions for a whole laboratory setup, e.g. Labscape or eLabBench, are stationary and the scientist is always confronted with its presence and can not move it away. However wearable support devices allow the use but do not interfere with the daily routine in a laboratory. The main problem of these devices may be the benefit in a real use case. At the beginning it might be fun to use it but after a while they are getting inconvenient.

Another big issue is the documentation of projects. Most systems seem to be too complex for the use in daily research. While the supported documentation instead of a laboratory notebook or even sheets of paper seems to yield benefits, the automated documentation is not ready for use in real science. A positive prospective view offers Prism where documentation, information management and social networking are combined to reduce time-to-publish and time-to-review. The main difference between the developement approach of Prism in comparison to others is the study in the background to discover the real needs of its users, the employees of wetlabs. Cencepts like Prism will be the next step of collaborative research and lead to more transparency and repeatability.

With respect to all possible fields withing wetlabs, it can be said that supporting systems, in whatever kind they appear, have a high potential to improve this conservative market. But the benefits have to be high enough to exceed the inconveniences that e.g. wearable devices imply. While a wetlab allows small-field solutions and hinders the one big solution, there will always be improvement with respect to individual needs for special cases.

References

[1] A. Tabard, J. D. Hincapié Ramos, and J. Bardram, "The elabbench in the wild: supporting exploration in a molecular biology lab," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2012, pp. 3051–3060.

[2] L. Arnstein, C.-Y. Hung, R. Franza, Q. H. Zhou, G. Borriello, S. Consolvo, and J. Su, "Labscape: A smart environment for the cell biology laboratory," *Pervasive Computing, IEEE*, vol. 1, no. 3, pp. 13–21, 2002.

[3] Search mask of google scholar. [Online]. Available: https://scholar.google.de/

[4] Search mask of the ieee xplore digital library. [Online]. Available: http://ieeexplore.ieee.org/Xplore/

[5] Search mask of pubmed. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/

[6] Search mask of sciencedirect. [Online]. Available: http://www.sciencedirect.com/

[7] Search mask of oskicat ucb library catalog. [Online]. Available: http://oskicat.berkeley.edu/

[8] Andrew, the pipetting robot. [Online]. Available: http://www.andrewalliance.com/wp-content/uploads/2015/02/andrewalliance-pipettingperformances.pdf

[9] P. M. Scholl and K. Van Laerhoven, "Wearable digitization of life science experiments," in *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM, 2014, pp. 1381–1388.

[10] G. Hu, L. Chen, J. Okerlund, and O. Shaer, "Exploring the use of google glass in wet laboratories," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*. ACM, 2015, pp. 2103–2108.

[11] A. Maccagnan *et al.*, "Combining ontologies and workflows to design formal protocols for biological laboratories," in *Automated Experimentation Journal*, 2010.

[12] T. Kurata, N. Sakata, M. Kourogi, H. Kuzuoka, and M. Billinghurst, "Remote collaboration using a shoulder-worn active camera/laser," in *Wearable Computers, 2004. ISWC 2004. Eighth International Symposium on*, vol. 1. IEEE, 2004, pp. 62–69.

[13] A. Tabard, W. E. Mackay, and E. Eastmond, "From individual to collaborative: the evolution of prism, a hybrid laboratory notebook," in *Proceedings of the 2008 ACM conference on Computer supported cooperative work*. ACM, 2008, pp. 569–578.

An Analysis of Wireless Sensor Networks for Bird Monitoring

Andreea Sorina Oltean
Faculty of Engineering
University of Freiburg
Freiburg, Germany
Email: olteana@informatik.uni-freiburg.de

Cosmin Gabriel Zvariste
Faculty of Engineering
University of Freiburg
Freiburg, Germany
Email: zvaristc@informatik.uni-freiburg.de

***Abstract*—Wireless sensor networks are of great use when talking about monitoring environmental areas without the need of a human intervention. Therefore, they are well suited for bird monitoring, including collecting reliable data directly from the source while having a long lifetime with low energy consumption. In this paper we will analyze the important aspects when considering designing a wireless sensor network for this purpose, considering problems regarding energy efficiency and harvesting, and self-healing networks. We will debate about protocols used in the wireless network, together with data approaches for a better management in what concerns energy saving. Towards the end, we will present two well-documented experiments of great relevance for short and long range communication.**

***Index Terms*—Wireless Sensor Network, bird monitoring, energy efficiency, energy harvesting, self-healing.**

I. Introduction

Biologists and scientists are interested in obtaining reliable data about environment with high fidelity. The usage of traditional data loggers for habitat monitoring proved to be too intrusive and affect the wildlife in a bad manner. Scientists showed that even a 15 minutes visit to a bird colony can cause up to 20% mortality among eggs and chicks in a breeding year. Therefore a new way of studying the wildlife was required. The need of an inconspicuous sensing method of the environment that can provide real-time data was fulfilled by WSNs (Wireless Sensor Networks). Wireless sensor networks could work unattended for long periods of time (up to years), being able to self-organize, self-heal and harvest their energy from the environment. Another advantage is the small size of the nodes, making them suitable to be attached directly to the birds without affecting their normal activities, offering therefore great tracking possibilities in the case of migrations.

II. Energy efficiency in Wireless Sensor Networks

Most of the time the WSNs used for habitat monitoring are long term networks that should work unattended from a few months up to several years. For covering the energy restrictions of a long term WSN, in this paper we will present two main approaches that can help save energy: duty-cycling and data-driven approach.

A. Duty-cycling

In order to make a better distinguish of how one can control the energy consumption, one can divide the duty-cycling approach in two complementary methods: Topology Control and Power Management. The first one tries to exploit the nodes' redundancy, making sure that just a minimal subset of nodes, that can fulfill connectivity, is active at the same time. The rest of the nodes is put in a sleeping state, waking them up in order to replace an active node that will run out of power or die from external causes. The second approach, the power management, relies on the fact that an active node does not need to keep its radio up all the time, and since the radio module is the most energy consuming element (Figure 2), switching it off when there is no network activity will considerably increase the WSN's lifetime.

1) Power Management: A radio module from a WSN node has four states: reception, transmission, idle and sleep. The problem is that not only in transmission and reception mode there is high power consumption, but it also consumes a lot of energy in idle mode, energy that is wasted since the network is actually not communicating. Therefore, in order to prolong the life of the WSN one can put the radio in a sleep state, a much lower power consuming state. Alternating these sleep and wake-up modes is known as duty cycling and it can be implemented on top of the MAC(Medium Access Control) layer, on application or network layer, or implemented directly in the MAC layer. Therefore, the Power Management approach can be further divided in two subcategories: Sleep/Wakeup protocols and MAC protocols.

*a) **Sleep/Wakeup protocols**:* The advantage of Sleep/Wakeup protocols is that they are flexible and can be written depending on the application's needs, in principle, on any preexisting MAC protocol. Further we will classify these protocols into three categories: on-demand protocols, rendezvous protocols and asynchronous schemes.

On-demand protocols rely on the idea that a node will wake up only when another one wants to communicate with it. The problem is in how to acknowledge the node that someone wants to send data to. To solve this, beside the main radio

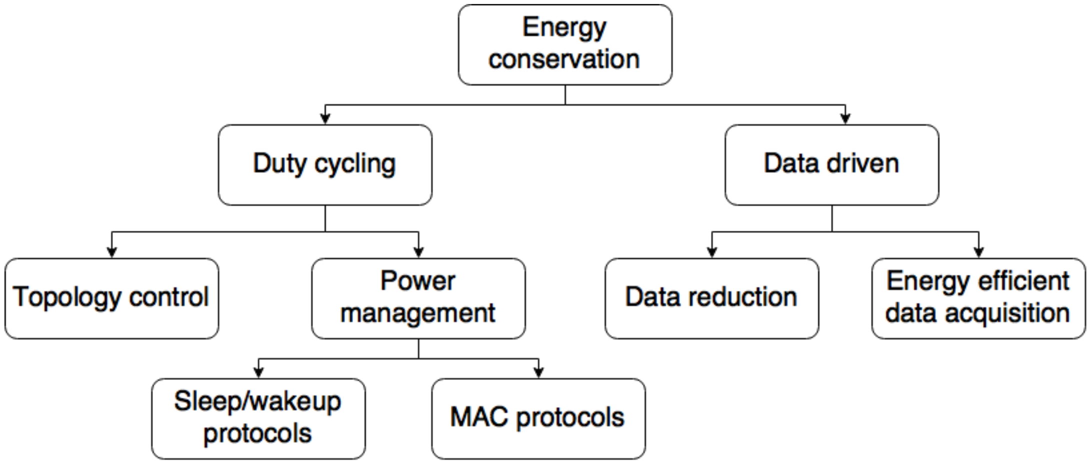


Fig. 1. Energy saving approaches

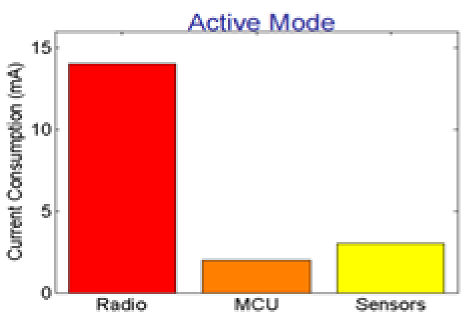


Fig. 2. Current consumption of different modules [6]

module that is used for data transfer, another low consumption radio is used just to receive the request for communication start up. In this way the power hungry radio used for data transfer stays most of the time in sleep mode and only the low power consumption radio stays idle waiting for signals.

Rendezvous protocols are based on the idea that all the nodes from a neighborhood wake up simultaneously, at a given time, and communicate with each other for a short time interval. After this, they all go back into sleep mode. The advantage of this method is that the broadcast from a node is guaranteed to reach all its neighbors. On the downside, having all the nodes communicating at the same time will cause a high number of collisions, and thus energy waste due to the need of retransmitting the packages. Also, there might be problems with synchronizing the clocks of all nodes.

Asynchronous schemes implies that a node can wake up whenever it has any data to be sent, transmit it and then go back to sleep. Considering that the receiving node must be also awake and ready for communication, one can say that this method can very well be used in single hop networks, were a gateway that is up all the time is present.

b) ***MAC protocols:*** The second approach for having a good power management regarding the communication is with the help of MAC protocols with low duty cycle such as TDMA(Time Division Multiple Access) based protocols. TDMA protocols provide unique time slots for each node for receiving/transmitting data, removing therefore the interference among nodes. Even though the energy efficiency is very good, there are problems when it comes up to scalability and flexibility of this method.

B. Data Driven

Data driven approach refers to reducing the amount of collected data from the nodes, without modifying the performance and accuracy of the intended application. Data driven approaches could save up even more energy by two means that will be further discussed.

1) Unneeded samples: When collecting information for a long period of time, there is no need to communicate all the data. It is know that a sudden change in the environment cannot occur instantly, but it will require intermediary states (e.g. the temperature cannot jump with 10C in one second). Keeping the spatial-temporal correlation in mind, the redundant information will not be sent to the sink, causing more energy saving on the sensing subsystem due to less communication [5].

2) Data reduction: In essence, this technique also aims at reducing the amount of transmissions or the length of the transmitted data packets by three different approaches.

In-network processing: data aggregation. This mechanism's main idea is to reduce the number of bytes sent and received by using compression and computation techniques. The wireless network implies a large amount of data, and sometimes could be aggregated at various levels in order to save energy for a transmission action. The information is collected from more nodes (e.g. from nodes that are very close to each other, as the data collected might be very similar). The idea is to perform data aggregation at the nodes between the sink and the source (e.g. compute the average value), but this may lead to an application specific problem.

Data compression involves encoding the data at the nodes and decoding it at the sinks. Information will be compressed before sending, therefore the size of the data will be reduced and energy will be saved. The decompression will be made at the base station, after the data will be acknowledged.

Data prediction consists in building an abstraction of a sensed phenomenon as a model describing the data evolution. This model can predict the values sensed by sensor nodes within certain error bounds and reside both at the sensors and at the sink [5]. Sometimes, the queries could be answered using the model and not the sensed data, since computing a new value is much less energy consuming. This model can be self-reconfigurable and adapt to new exception values. The model found at the sources ensures the paradigm's correctness by comparing the acquired data with the predicted value from the model; the one at the sink answers queries without requiring access to any data from the sensor. Different types of data prediction have been tackled. One can begin using the probabilities or the statistical properties of the event to build up a probabilistic model (e.g. can obtain a probability density function after a training phase). Another solution can use time series forecasting: prediction is made using historical values by periodical sampling; this explicitly considers the internal structure of the data, in comparison to the statistical approach. This model can generally be consisted as made of a classic pattern and a random error (e.g. auto-regressive models).

Energy efficient data acquisition. Sometimes, the reduced number of times the nodes communicate may not be enough. The nodes could have power hungry transducers, A/D converters or active sensors, which cannot be directly controlled. Also, depending on the device or application, acquisition of data may take a lot of time. Therefore, the number of data samples given by the sensors must also be reduced. One first solution could be an adaptive sampling approach. As said before [5], the idea behind is that data may slowly change in space and time and that the subsequent samples do not differ so much from one another. This can be exploited to reduce the number of pieces of information. Secondly, the hierarchical sampling idea represents a trade-off between accuracy and energy conservation and it is based on the idea that a node has more types of sensors, and that not all the sensors must be active in the same time.

III. Energy Harvesting

Even though a WSN could use some smart way of saving energy, like presented in chapter 2, the lifetime of the nodes will still be limited to the capacity of its battery. This is the greatest problem a WSN will face, especially in habitat monitoring, where changing the battery of a mote is sometimes impossible due to the interaction with wildlife, interaction unwanted by scientists. Therefore to improve even more the duration that a WSN can operate, a way to provide power to the nodes is needed. One can do this with the help of energy harvesting, converting ambient energy into electrical energy. Most common sources of harnessing power, which can be used in habitat monitoring WSNs, are: solar power, mechanical energy (piezoelectric or vibrations) and thermal energy (body heat).

Here we can distinguish two kinds of architecture: Harvest-Use and Harvest-Store-Use. The difference between the two models is the presence of the battery. The first architecture, Harvest-Use, relies just on the energy that the node can harvest at the moment, having no ways to store it for further usage. In this case the node needs continuously harvested energy above the operating point, otherwise the mote will shut down. On the other side, the Harvest-Store-Use architecture is equipped with a battery, making it able to work in an environment with abrupt variances in harvesting, due to the stored energy.

A. WSN-HEAP

A particular type of WSN that uses harvested energy is the WSN- HEAP (Wireless Sensor Networks Powered by Ambient Energy) [2]. The main characteristic of those networks is that they rely solely on the harvested energy stored in super-capacitors. Beside the disadvantage of a limited amount of charge/deplete cycles that batteries present, they also are not environmentally friendly since they are prone to leakage. With the usage of super-capacitors, the WSN-HEAP takes care of both of the problems, providing a suitable solution for long term deployment of habitat monitoring networks.

A problem that occurs in WSN-HEAP is that one cannot foresee when a node will harvest enough energy to send its data to the gateway. Therefore implementing any TDMA protocols for communication is difficult, asynchronous schemes being more suitable. Also multi-hop networks are hard to implement since we can have no neighbors awake to route the package further. In the following graph we present a comparison between a WSN-HEAP and a battery based WSN.

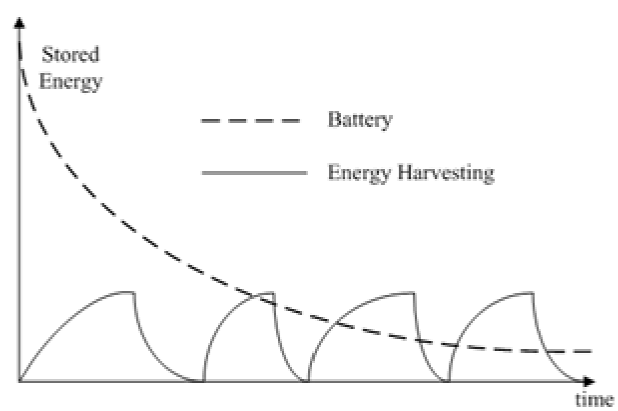


Fig. 3. Types of energy storing over time [2]

IV. Self-Healing

A successful application based on a wireless sensor network should be, in essence, reliable, especially the ones placed in a natural environment, such as the ones for monitoring wildlife. In ordinary laboratory research, when a node fails, that node is discarded and the network is reorganized. But this may cause, besides the trouble of reorganizing the network or replacing the node, unexpected maintenance costs and also could decrease the function of the WSN. In order to guarantee a reliable, robust and a low cost operating WSN, the concept of self-healing has been introduced. This implies using a system with reconfigurable hardware and in the following, we will present two main paradigms that use FPAAs and FPGAs [3]. The main concept for both paradigms is based on the fact that if part of a node fails, the whole node is not abandoned, but it will reconfigure itself in order to continue working.

A. Redundancy-based self-healing

In this paradigm, in the circuit built with the help of FPGA or FPAA redundant modules of important circuits are built; they are connected with the FPGA/FPAA in order to form self-healing modules. Besides this, fault diagnostic circuits are also built, and with their help, when a part fails and they detect the fault, the FPGA/FPAA dynamically reconfigures such as it disconnects the failed part and switches to the redundant part of the module. Therefore, the faulty part is abandoned and the node continues to work normally. The downside is that this requires complicated hardware design, bigger node dimension, so it gets more expensive.

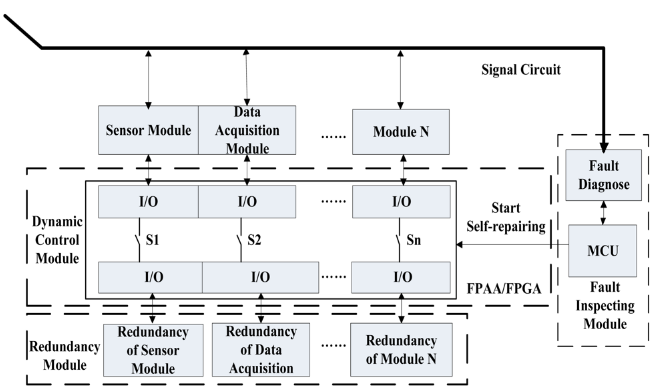


Fig. 4. Redundancy-based self-healing WSN node paradigm [3]

B. FPGA/FPAA-based self-healing

Here, the FPGA/FPAA's internal modules substitute the main analog or digital circuits of the system. When a failure is encountered, the FPGA/FPAA dynamically reconfigures in order to use one of its own internal module to replace the defected part. The advantage is that there is no need for extra redundant modules and this simplifies the design and cost of the system. On the negative side, the FPGA/FPAA cannot reproduce a sensor's accuracy perfectly, hence this solution is not applicable for high precision tasks.

The results from the experiments made in [3] show that the recovery speed of the FPGA/FPAA based was faster due to the fact that the circuit switching speed between the internal modules of the FPGA/FPAA is faster. Also, it was shown that when using a redundancy base plan, the energy consumption was higher because it costs more energy when redundancy modules are adopted.

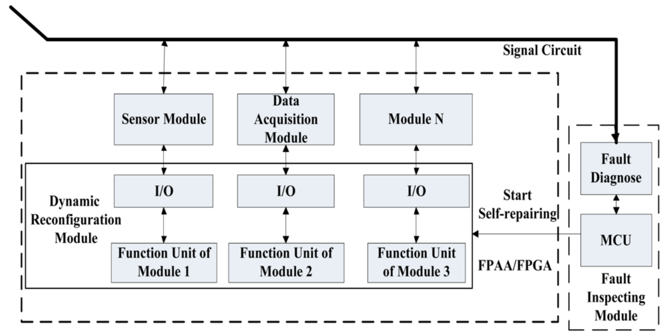


Fig. 5. FPGA/FPAA-based self-healing node paradigm [3]

V. Communication

In the following chapter we will present two different ways of communication inside a WSN, providing advantages and disadvantages for each and also examples of how each were used in field experiments. We will distinguish between short ranged radio communication used for nest monitoring and cellular based communication best used for tracking birds during their migrations.

A. Short range radio communication

Short range radio communication is a low cost, low power method to send data between motes in a WSN. For stationary bird monitoring, like nest monitoring, this is one of the best ways to establish a connection between the nodes. Since the breeding areas are usually relatively small, nodes placed on birds or nests are close to each other, making them able to route the packets towards the gateway via a multi-hop system, or even send the data directly to a gateway if they are in its range. This is the case of the next experiment we will present which took place on Great Duck Island in Canada and had the goal of studying Storm Petrels during their breeding period.

In order to obtain reliable data about the microclimate the birds prefer for their incubation and hatching, motes were developed small enough not to interfere with the daily life of the birds, when placed inside the nests. Since the nests were actually underground burrows, another node was placed above the burrow, at the ground surface, so the scientists could examine the differences between inside and outside the burrow. The nodes were based on a Mica board developed by UC Berkeley, and equipped with sensors for temperature, humidity, light, pressure and infrared radiation(used to see if the burrow is populated or not).

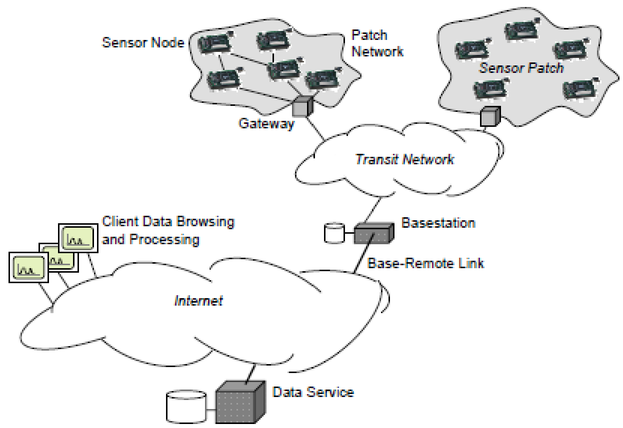


Fig. 6. The architecture of the system [1]

As shown in Figure 6 the sensors were organized in patches (clusters), each cluster equipped with a gateway that was receiving the reading from the nests and forwarding them via a transit network to the base station. Afterwards the data is uploaded in a database over internet where scientists have access to it via user interfaces. In order to save energy, the nodes were in sleep mode primarily, waking up every 70 seconds to send data as 36 byte data packages. The gateway on the other hand is working at 100% duty-cycle being able to coordinate the activity inside the patch and help with additional computations and storage. For this, the gateway is equipped with additional solar panels to harvest energy and a rechargeable battery.

B. Cellular based communication

As birds come in different size and shapes, the wireless sensor network that monitors their activity must adapt to that specifically type of bird. For migratory birds that travel hundreds of kilometers per day though continents, a classic wireless sensor network is not suitable. For this large scale connectivity network, a cellular based communication has been chosen [4]. Cellular technology has the benefit that it already covers high areas due to the preexisting coverage of cellular networks, but the drawback is the network's holes in the wildlife: only the populated areas have GSM coverage. This disadvantage can be overcome by adding a small memory module to keep the data until the birds reach an area with GSM coverage.

The CraneTracker experiment [4] monitors the endangered species of Whooping Cranes that do an annual migration of 4000km between Texas and Canada, travelling up to 950km every day. The challenges faced were to cover this huge area, but to keep as shorter delays as possible (less than 24 hours) to establish eventual cause of disconnection from the network, which may show cause of death of the cranes. The intended network lifetime should be around 5 to 7 years and the weight of the system that what attached to the bird as a backpack should be less than 110g (which represents 2% of the bird's body weight). Besides the classical modules (GPS, temperature sensor, compass, radio module), the crane tracker mounted on the birds was equipped with a very low-weight GSM module GE865 from Telit, which supported UART communication for easily interfacing with the microcontroller; the module is working on international bands, meaning it can be deployed in many places around the world and with its help, a coarse localization using cellular tower information could be made. The tracker also holds a lithium-polymer battery, a flexible solar panel from PowerFilm for energy harvesting and a 512kB memory for storing the information when out of GSM coverage. Beside the moving tracker, the backend component gathered the information from the gateways and used a web service to give the results to the user. Once the birds were reaching the breeding area, the network would switch on short range radio communication since the nests were in fixed locations relatively close to each other (Fig. 7).

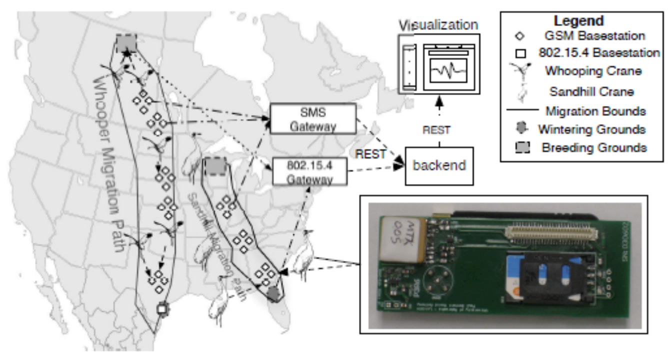


Fig. 7. The architecture of the system [4]

Due to the endangerment of the species, some incremental deployments have been made, starting with field experiments in enclosures on wild turkeys to deploying the tracker on five wild sandhill cranes; this last part of the experiment is the one most close to its intended purpose, as the sandhill cranes have been monitored for over a year migrating along the American continent. The results show that two of the five subjects completed the migration to south, and the other three's tracker either fell off the bird or failed to recharge. During the migration, they recorded around 330 GPS locations, the flying periods of the birds, the speed and the altitude. Also different charge/recharge cycles have been recorded due to the difference of location; most of the time, they stay in the breeding grounds, more under the vegetation, while in the air, they are in direct sunlight. The overall results of the experiment confirmed the viability of the cellular network in these extreme geographic distances, giving enough insight of the cranes' habits.

VI. Conclusion

The paper showed why wireless sensor networks represent a suitable solution for bird monitoring. We discussed some important aspects of a WSN providing advantages and disadvantages as well as where the different solutions should be used. An important characteristic of these networks is the low energy consumption, and for this to be achieved we presented multiple approaches both from communication point of view as well as from data management part. For prolonging even more the lifetime of a WSN, one can use energy harvesting methods like solar panels or mechanical energy from birds movement. Since the contact between humans and the wildlife should be as little as possible, if we want unaltered data, we pointed out how one can equip a WSN with self-healing attributes. In the end we discussed two experiments that showed how we can adapt a WSN for our needs, proving we can use WSNs for monitoring birds both in a small area (like nesting zones), or a large migration path over more countries.

Acknowledgment

The authors would like to thank Kristof Van Laerhoven and Philipp M. Scholl from University of Freiburg for offering the opportunity to write this paper.

References

[1] R. Szewczyk, J.Polastre, et al. *An Analysis of a Large Scale Habitat Monitoring Application* 2004.

[2] Winston K.G. Seah et al. *Wireless Sensor Networks Powered by Ambient Energy Harvesting (WSN-HEAP) Survey and Challenges* 2009.

[3] Shenfang Yuan et al. *Providing Self-Healing Ability for Wireless Sensor Node by Using Reconfigurable Hardware* 2012.

[4] David Anthony et al. *Sensing Through the Continent: Towards Monitoring Migratory Birds Using Cellular Sensor Networks* 2012.

[5] Giuseppe Anastasi et al. *Energy conservation in wireless sensor networks: A survey* 2008.

[6] Abdelmalik Bachir et al. *MAC Essentials for Wireless Sensor Networks* 2010.

[7] Sujesha Sudevalayam and Purushottam Kulkarni *Energy Harvesting Sensor Nodes: Survey and Implications* 2011.

[8] David Anthony et al. *CraneTracker100 Platform Description* 2012.

[9] Zahra Rezaei et al. *Energy Saving in Wireless Sensor Networks* 2012.

[10] Sumit Chaudhar et al. *Energy Efficient Techniques for Data aggregation and collection in WSN* 2012.

[11] Joseph Robert Polastre *Design and Implementation of Wireless Sensor Networks for Habitat Monitoring* 2003.

[12] Alan Mainwaring, Joseph Polastre, Robert Szewczyk, David Culler, John Anderson *Wireless Sensor Networks for Habitat Monitoring* 2002.

[13] Hongwei Zhang, Anish Arora *GS3: scalable self-configuration and self-healing in wireless sensor networks* 2002.

Bird monitoring by wireless sensor networks

Alexander Richter
Department of Computer Science
University of Freiburg
Freiburg, Germany 79085
Email: ar.richter@gmail.com

Philip Klein
Department of Computer Science
University of Freiburg
Freiburg, Germany 79085
Email: philipkln.91@gmail.com

Florian Tautz
Department of Computer Science
University of Freiburg
Freiburg, Germany 79085
Email: florian@nyronet.de

***Abstract*—The behavior of birds is astonishing and many scientists are hoping to find out more about how birds can do extraordinary things such as flying for multiple days in a row. There are many challenges when it comes to bird monitoring, and many of them are more and more conquered by modern technology. Wireless sensor networks can be used to listen for bird sounds and thereby identify certain species. Another application for wireless sensor networks is monitoring of nesting behaviour by placing sensors directly in or close to bird nests. These application present interesting challenges for computer scientists and biologists and necessitate collaboration between these fields. Habitat monitoring has been a practiced for many years and was continuesly improved with additional sensors and features like sound detection or images. Wearable sensors for birds are especially challenging due to the fact, that a bird's flight behavior is influenced if the carried sensors are heavier than 3% of the bird's body weight. Renowned researchers all around the world are pushing the limits of technology to create smaller and lighter tracking devices, that eventually will make the recapturing of birds obsolete.**

***Index Terms*—bird monitoring, habitat monitoring, nest observation, wireless sensor networks, song recognition, bird backpack, lab-on-a-bird.**

1. Introduction

The continuously advancing technology regularly allows for new ways of bird monitoring. In this paper we give an overview about various approaches. We will start with wireless sensor networks in general as they form the foundation for the systems specific to bird monitoring. The core issues we focus on are error handling and reprogrammability. From there on we will take a look at various approaches to bird monitoring. We will go over simple nodes with sensors, to image based monitoring and detecting the birds via their songs.

Thereafter we take a look at systems that are partly attached to the bird and finally reach the point where the whole nodes are attached to the birds. The research and information was gained from papers which stated how modern technology is used in the different types of fields.

August 17, 2015

2. Paper selection criteria

Papers were collected by using the search terms: “sensor networks”, “habitat monitoring”, and “bird sound detection”. When reading papers, related articles on Google Scholar and alike were taken into consideration as well. Articles from the conferences IWSC, UbiComp, Sensys, and EWSN each with high citation counts were preferred in the selection process. Papers which were hardly related to the keywords or had a very low citation count in relation to their age were rejected. The same applies for work with a publication date older than 10-15 years or when the corresponding references were inexplicable.

3. Wireless sensor networks

A Wireless sensor network (WSN) is a collection of microcontroller-based sensor nodes (also referred to as “motes”) that communicate via some wireless network technology. This approach is useful because such small nodes only have a small impact on the monitored animals and advances in technology have made the hardware relatively cheap.

Especially in animal monitoring, the size of the node package is usually constrained and this, together with the requirement for long time operation, makes power supply – and thus power consumption of the nodes – the main challenge for designing the nodes. This problem can be approached from different directions, including node design, network layout, and software.

3.1. Network layout

The types of network layout can generally be divided in singlehop and multihop networks. In singlehop networks, the nodes do not perform any routing and network packets are transmitted directly from source to destination. In multihop networks, packets can be routed from source to destination via intermediate nodes. Multihop networks can span a larger area without having to increase transmission power but routing also has an impact on power consumption for the intermediate nodes.

A typical network layout for animal monitoring usually consist of multiple sensor nodes in a singlehop or multihop structure and a gateway that is connected (sometimes over the internet) to computers that permanently store the collected data or issue commands to the network. The network layout has to be chosen carefully to avoid inequal load and consequential node failures. Network simulation software and smaller scale test runs can help mitigating these issues.

3.2. Handling node failures

Even in a carefully crafted network with very sturdy nodes, software bugs and unforeseeable environmental factors make it next to impossible to avoid node failures completely. Wireless sensor networks should be designed in a way that allows recovery from such failures, or at the very least detection of failures, to allow timely human intervention. An easy way to detect and predict node failures are regular health reports from all nodes, containing information on the state of the power source and network routing graph.

Detecting and handling failures of multiple nodes presents a special challenge [1]. If a part of the network gets cut off from the gateway it is impossible to tell the difference between just the intermediate nodes failing and the whole part of the network that is no longer reporting failing, requiring manual intervention. This can be avoided to some degree by ensuring there always exist at least a certain number of alternative routes between nodes.

3.3. Reprogrammability

Due to changing requirements for the WSN, changing environmental conditions, or node failures it is useful to have a way of reprogramming some or all nodes of a network. The goal is transmitting the required data power efficiently but as quickly as possible.

Maté [2] is a bytecode interpreter built on top of TinyOS which aims to make programs using a common set of operations smaller. It provides abstractions for many complex tasks that can be invoked with small 1-byte instructions. Programs are split up into capsules of 24 such instructions each. Maté provides a routing algorithm and everything required for capsules to forward themselves to other nodes with a single instruction.

3.3.1. Data dissemination. Due to the memory, processing power, and electric power constraints, specialized techniques for propagation of data in wireless sensor networks in a fast, yet resource-efficient way were developed. In the following we take a closer look at two of them.

Deluge [3] is an epidemic protocol for disseminating large data objects such as full binary images for node operating systems. It splits the object to be disseminated into pages which are split into packets for network transmission. The node maintains an object description with the age of the stored pages and utilizes 16 bit CRC for integrity checks. The protocol describes three states a node can be in: In MAINTAIN state, a node advertises its version of the object's pages to neighbouring nodes (if there are not too many other nodes broadcasting at the same time), switches to RX state when receiving an advertisement for a newer page or switches to TX when receiving a request for one of the advertised pages. In RX mode, the node requests a page advertised by an other node and then receives the page's packets. In TX mode, the node transmits the requested page's packets. The authors claim that Deluge achieves a transmission rate of 90 bytes/second under real-world conditions.

Wireless sensor networks usually operate at very low duty cycles (frequently under 1%). This results in traditional flooding algorithms being rather inefficient because most of the time only very few neighbours are awake at the same time to receive a transmission. Opportunistic flooding [4] seeks to reduce redundancy in transmissions while also achieving fast data dissemination. To achieve this, the protocol determines the routing tree with the lowest power consumption and the expected transmission delay. The expected delay is shared with previous-hop nodes which can then, based on this information, decide to transfer outside the energy-optimal tree if this is likely to result in a significantly earlier arrival of a packet for the following node. This structure is updated dynamically and decisions are made for each packet individually. Collisions are handled by delaying transmission of packets depending on link quality and only sending when the channel is clear. Opportunistic flooding achieves a significant improvement when compared to traditional flooding methods and achieves flooding delay and energy cost values close to the optimal case.

4. Bird nest monitoring

In "An analysis of a large scale habitat monitoring application" [5] the authors describe an experiment run over the course of a few months in 2003. They deployed wireless sensor networks to monitor nesting behaviour of birds on Great Duck Island, Maine. The applied network layouts include singlehop and multihop networks and two different kinds of nodes for monitoring weather conditions and bird nests respectively.

The key challenges identified in the paper are node lifetime, obtrusiveness, reliability, and ease of deployment. The system used in the experiment achieved satisfactory results with respect to most of these requirements. Most nodes lasted multiple months and provided useful data. Nodes in the multihop network exhausted their power supply earlier than nodes in the singlehop network but the multihop network covered a significantly larger area. Ease of deployment and diagnostic tools were identified as areas with room for improvement in later projects.

4.1. Sound based monitoring

This method has the great advantage that one doesn't need to reach the nests, which can be a challenge in difficult terrain like rain forests. Biologists are very interested in a system that monitors when each bird species sings because this would allow them to learn a lot about bird communication.

To construct such a system, the nodes need an unidirectional microphone to listen in every direction and a mechanism to start and stop recording. The mechanism to record can't just be activated through noise due to the high background noise in the birds habitat. But luckily most birds sing in a frequency that differs from the background noise so one can set a range of frequencies that will activate the recording. Analyzing the recorded sounds is still difficult due to the high noise to signal ratio.

The next step is it to match the recorded song to a certain species or even individual. Many classifiers and learning algorithms like neural networks or Bayesian models come to mind for this task but so far the research is dominated by hidden Markov models (HMM) and Gaussian mixture models.

Hidden Markov models (HMM) are a great fit for bird song recognition for multiple reasons. Many recordings of bird songs miss the beginning or end of the song. For HMM this doesn't matter. They are also easy to implement and don't require much memory or processing power. Trifa et al had great success with this approach which they detailed in their paper "Automated species recognition of antbirds in a Mexican rainforest using hidden Markov models" [6]. They mention that going forward, it would be important to increase the signal to noise ratio due to the fact that HMM are susceptible to noise. They reached an accuracy of 99.5% with clean signals but only 90% in a realistic setting.

Gaussian mixture models are already widely used for human speech recognition, which allows reusing a lot of that research for bird song recognition. Another advantage is that not many learning samples, which are usually hard to acquire, are required for satisfying results. Like hidden Markov models, this approach is susceptible to noise.

This setup uses more power than the previous one, where only simple nodes were used. To deal with this, one can decide to add an external energy source, like solar power, to the node or design it in such a way that changing the battery is a quick an easy process.

4.2. Image based monitoring

A system which sends images of bird nests to a central server would be great for biologists. It would spare them a lot of time currently spent on checking the nests for any changes and would capture important events which could easily be missed with the classic approach. This system would of course create some new challenges. It needs to send a lot of data through the network which results in high power consumption which, in combination with the camera, would definitely require a power source.

Sending large amounts of data through a network requires a lot of energy compared to doing some local processing on the node. This means that it can be advantageous to compress images before sending them. Conventional compression algorithms like JPEG aren't a good choice because they need more computing power and memory than usually available. Paek and his fellow researchers decided to use an algorithm called PackBits which was originally developed by Apple [7]. The advantages of PackBits include that it is a simple algorithm and works with limited memory as it doesn't need to buffer the full image. The algorithm works on gray scale images. When it runs it checks how many consecutive values are within a certain threshold of each other. These values are then averaged and compressed to [length][average value]. An example of how the algorithm works taken from the paper by Paek [7]:

> If the threshold is 10 and the original image data are 1, 2, 1, 3, 2, 3, 2, 2, 15, 20, 15, 20, 100, 110, 105, 105, 100, 110, then this image is compressed into [8][2],[4][18],[6][105] and will later be decompressed to image data 2, 2, 2, 2, 2, 2, 2, 2, 18, 18, 18, 18, 105, 105, 105, 105, 105, 105.

As one can easily see, choosing a high threshold results in high compression but also causes a higher loss of information.

One downside of this algorithm is that the compressed image requires 100% reliable data recovery which in turn means that data would sometimes have to be sent multiple times. In the paper by Paek [7] the full image recovery was especially crucial because the server ran automatic checks on the images to see if a bird is on it. Another issue identified in the experiment is that the compression ratio isn't as good as anticipated when tuned for acceptable image quality.

4.3. RFID

With the help of Radio Frequency Identification one can monitor when the bird leaves or enters its burrow. In the paper "Wireless Sensor Network for habitat monitoring on Skomer Island." [8] the authors combined this approach with a weighing scale at the burrow entrance to monitor the weight of birds entering or leaving. With this information, they wanted to draw conclusions about how much food the birds bring back home for their offspring and how frequently such events occur. To set this up they had to catch the birds first to put and RFID chip on them.

It is easy to add sensors in and near the nest to get more information like the temperature and humidity which then can be transmitted when birds enter or leave. With this information one can look for interesting patterns, for example how hunting times are influenced by time, weather and other factors. The multihop network for such an experiment doesn't need to be very complex due the fact that the collected information is very small data-wise.

An interesting addition to this project would be to combine it with a camera in the nest because the information already gathered would facilitate taking pictures at the right moment.

5. Wearable bird monitoring

When it comes to monitoring individual birds, the most reliable way is to attach sensors directly to the bird itself. Unfortunately, this method entails many challenges due to the fact that birds can only carry about 3% of their own body weight without changing their flight behavior. That makes wearable bird monitoring especially difficult for smaller sized birds.

Bird ringing is one the oldest methods for the tracking of individual birds. But in order to attach a ring to a bird it has to be captured. In addition, the bird has to be recaptured every time so that it can be identified. Though this method has proven to be quite effective it only allows limited insight into the birds movement since it only provides the bird's current location once captured.

Instead, many scientists use energy efficient microcontrollers and small GPS-receivers in order to store the bird's location in regular intervals and, afterwards, retrieve this information by recapturing the bird and downloading the information via USB. A more advanced approach is to utilize wireless communication so that the recorded information can be downloaded as the bird flies by without recapturing it. In order to fulfill the high energy demands of wireless communication and the requirement of low weight for the bird to carry, batteries are often not an option. Alternatively, energy harvesters, such as solar panels, are used to generate the energy necessary to power the device, along with a small backup battery. The more sensors are in use and the more data is stored, the higher the energy consumption. Therefore, the use of a combination of sensors to record activity, mortality, acceleration, temperature, light conditions, location, etc. must be thought out. The bird's location, for example, doesn't need to be saved every second in order to understand its route. Saving one location every few hours could be sufficient depending on the application. Other sensors such as temperature sensors could be obsolete if the weather conditions in the area of interest are known ahead of time. Additionally, the use of accelerometers can also be reduced to a minimum by only recording a few seconds of data when trying to recognize the bird's activities such as flying or sitting.

This approach is described in more detail in the paper: "A flexible GPS tracking system for studying bird behavior at multiple scales" [9]. In the paper, a device with the dimensions 62x30x12 mm, a weight of about 12 g, a microprocessor, a built in storage, a GPS receiver, an accelerometer, a radio transceiver, a solar panel, and a small backup battery with a capacity of 65 mAh, are used to monitor bird movements without the need of recapturing them by receiving recorded data wirelessly at a breeding area.

The use of solar energy is very effective, but at the same time, with birds living in areas without abundant sunshine one cannot rely on solar power. The amount of energy needed to bridge between solar charges can be too large to be fulfilled by enlarging the battery size. Therefore, alternative energy harvesters such as piezo electric energy harvesters, that harvest energy from the birds movements, are being tested in order to become independent of such restrictions. In theory, the amount of energy that can be harvested from a flying bird is about 0.01-10 mW which can be enough to power an electric circuit like the one mentioned above. At Cornell University in New York there has been a lot of interesting research in this field and also a paper called "Testing of Vibrational Energy Harvesting on Flying Birds" [10].

Birds are capable of astonishing things and biologists are burning to learn more about various aspects of bird behavior. The perfect bird monitoring system has yet to be built, but an ideal wearable bird monitoring device would enable insights that, so far, were never possible. A combination of multiple energy harvesters and a small battery could create a self-sustaining power supply. By using ARGOS, which allows to send information directly to a satellite, the recapturing or proximity of the bird for downloading data is no longer necessary. The use of an inertial measuring unit can tell which state the bird is currently in, coordinates are updated multiple times per second, spherical images from the birds surroundings are streamed live and a on-chip lab continuously monitors the birds metabolism. But until then, researchers keep fighting the 10 g barrier to monitor small birds with as many sensors as possible.

6. Conclusion

The main challenge identified in the papers we reviewed was supplying the systems with sufficient power for long time operation. This problem can be approached from different directions, namely making the nodes use less power or supplying them with more power via different forms of energy harvesting. For ground based projects, solar energy is a useful power source and for on-bird systems piezoelectric energy harvesters may be used. Ideally a combination of power efficiency and energy harvesting is employed to ensure reliable operation over a long period of time. The strict power constraints play into most areas of system design, from design of the node hardware over software implentation to network layout.

Other challenges that are frequently mentioned in the papers are obtrusiveness (animal behaviour should not be influenced by the monitoring technology) and usablity. Tackling these issues required careful design of the nodes and collaboration with the scientists that are going to use the system in the field. The system should be small, easy to deploy, have some form of reprogrammability or reconfigurability, and be as long-lived as possible.

In addition to these general challenges, each project presents its own unique problems. While ground based applications monitoring by sound detection or imagery require a lot of processing power but transmit data only over a short distance, air based approaches require a way to store data and potentially transmit the collected data over a longer distance.

References

[1] Younis, Mohamed, et al. *Topology management techniques for tolerating node failures in wireless sensor networks: A survey.* Computer Networks 58 (2014): 254-283.

[2] Levis, Philip, and David Culler. *Maté: A tiny virtual machine for sensor networks.* ACM Sigplan Notices. Vol. 37. No. 10. ACM, 2002.

[3] Hui, Jonathan W., and David Culler. *The dynamic behavior of a data dissemination protocol for network programming at scale.* Proceedings of the 2nd international conference on Embedded networked sensor systems. ACM, 2004.

[4] Guo, Shuo, et al. *Opportunistic flooding in low-duty-cycle wireless sensor networks with unreliable links.* Computers, IEEE Transactions on 63.11 (2014): 2787-2802.

[5] Szewczyk, Robert, et al. *An analysis of a large scale habitat monitoring application.* Proceedings of the 2nd international conference on Embedded networked sensor systems. ACM, 2004.

[6] Trifa et al. *Automated species recognition of antbirds in a Mexican rainforest using hidden Markov models*

[7] Jeongyeup Paek, John Hicks, Sharon Coe and Ramesh Govindan, *Image-Based Environmental Monitoring Sensor Application Using an Embedded Wireless Sensor Network*

[8] Freeman, Robin, at al. *Wireless Sensor Network for habitat monitoring on Skomer Island*

[9] Bouten, Willem and Baaij, EdwinW. and Shamoun-Baranes, Judy and Camphuysen, KeesC.J., *A flexible GPS tracking system for studying bird behaviour at multiple scales*, Springer-Verlag, 2013.

[10] Michael W. Shafer, Robert MacCurdy, Ephrahim Garcia Cornell University, *Testing of Vibrational Energy Harvesting on Flying Birds*, ASME 2013 Conference on Smart Materials, Adaptive Structures and Intelligent Systems, Snowbird, Utah, USA, September 1618, 2013.