Sensing Opportunities for Physical Interaction

Florian Michahelles and Bernt Schiele
Perceptual Computing & Computer Vision Group,
ETH Zurich, Switzerland
{michahelles, schiele}@inf.ethz.ch

ABSTRACT

Today's technology offers a wide variety of sensors. Although many sensing applications have been produced, there is no support for the design of applications offering physical interaction. In order to make a step towards such a design framework this paper analyzes different means of sensing of humans and human activity. In particular we identify six sensing goals, referred to dimensions of sensing: ID (1), Object Use (2), Location (3), Bio Signs/Emotions (4), Activity (5) and Interaction among humans (6). Those dimensions together with different sensor placements are used to review and analyze ubiquitous computing research related to physical interaction and sensing. The final discussion draws conclusions from this analysis with respect to appropriateness of sensors and sensor placement for different sensing dimensions.

1. INTRODUCTION 1PAGE

Mobile devices give access to computing services without the constraint of sitting in front of a desktop computer. This poses new challenges for human-computer interaction: mobile users can be busy with real-world activities at any time while using mobile devices, such as crossing a busy street, discussing in a meeting or riding a bicycle. Previous work [15, 29, 24, 13] proposes physical interaction, e.g. tilting a device for configuring a device's functionality, as new and convenient forms of interaction for mobile user scenarios. The notion of implicit interaction takes this a step further and suggests to sense "an action, performed by the user that is not primarily aimed to interact with a computerized system but which such a system understands as input." [28]. That means, the user interacts with physical objects in a natural way, but a computer system also can extract inputs from these actions for the use in meaningful applications. System inputs generated from interaction with physical objects already have been used for coupling physical objects with computer applications as tangible user interfaces [18], computer-assisted furniture assembly [1], future restaurant scenarios [17], tracking a patient's medicine cabinet [31] or work-flow monitoring in a chemical lab [3]. Empowering a computer system to process physical user inputs requires augmentation of today's computer nerve-endings, such as mouse and keyboard, by sensors: perception and interpretation of real world phenomena enables a computer system to participate in the user's physical environment and serve the user in an appropriate way. Today's technology offers an astounding variety of sensors more or less suited for different applications: accelerometer, oximeter, microphone, gyroscope, temperature, skin resistance, etc. However, this variety of sensors makes it difficult for an application designer to choose the most appropriate subset which depends on the application. Although quite a variety of applications have been produced, there is no support for the design of applications offering physical interaction, such as toolkits or style-guides available for GUI development.

This paper was presented at "Physical Interaction (PI03) Workshop on Real World User Interfaces", a workshop at the Mobile HCI Conference 2003 in Udine (Italy). September 8, 2003. The copyright remains with the authors. Further information and online proceedings are available at http://www.medien.informatik.uni-muenchen.de/en/events/pi03/

This paper is a first step to develop a conceptual framework, that allows to categorize existing sensors and evaluates their utility in various applications. Eventually, this framework shall guide application designers to choose meaningful sensor subsets, inspire new systems and evaluate existing applications. The paper is structured as follows: Section 2 briefly summarizes related work, section 3 describes a conceptual categorization framework of sensors and reviews existing ubicomp applications by means of the framework. Section 4 presents an evaluation of sensing technology in respect to the framework. Section 5 concludes the paper.

2. RELATED WORK

Related research attempts have been made to develop frameworks and infrastructure for reusable sensing mechanisms. The context toolkit [26] supports the development of context-aware applications with useful abstractions from the actual sensors. However, it mostly deals with the context recognition on an abstract level decoupled from the variety of sensor technology. Furthermore, it limits applications to single sensor usage as only one context abstraction can be mapped to one physical sensor. In contrast to that, the TEA architecture [32] focuses on low-level abstractions for simple sensors, which depends to much on the used sensors and, as such, does not provide reusable perception mechanisms either. The sensor classification scheme [37] facilitates the comparison and classification of sensors.

3. FINDING THE APPROPRIATE SENSORS

Advances in sensor technology such as form-factor, power consumption, processing requirements and cost-effective fabrication offer a wide variety of integration into devices and appliances. An application that enables implicit interaction uses sensors as nerve-endings to perceive the environment. But what are the appropriate sensors? Instead of an technology-oriented view we take the perspectives of a designer and an engineer: sensing goals, referred to as dimensions of sensing, and actual placement of sensors.

3.1 Logical View: Dimensions of Sensing

Typically, application designers are more interested in the opportunities sensors can offer, than in the actual technology itself. As physical interaction shall happen between human and machine, all characteristics that describe the user's situation are of interest to the application. For that, in the last years a very general definition has been established [9]: "Context is any information [...] to characterize the situation of an entity". Unfortunately this definition is too general and does not really help for application design. Thus, we identify six sensing goals, referred to as dimensions of sensing, that give a more precise description of user context.

The first dimension is a user's *ID* - this has been widely used already, e.g. for customizing and personalizing services without requiring explicit user inputs [25, 4]. In fact, we use a more general definition of *ID* ranging from differentiating people to actual identification. The second dimension is *Location*; it has been the dominant implicit input used in ubiquitous computing applications [35, 8]. It does include 3D coordinates but also semantic location descriptions. The third dimension, *Activity*, describes the task the user is performing which ranges from simple moving patterns [33] to precise job descriptions. The fourth dimension, *Object Use*, comprises collocation of a user to an object [25], carrying an object [20] and the actual use [1]. The fifth dimension, *Bio Signs/Emotions*, describes the internal state of the user. Research in this area is still in its infancy. First results could be obtained with heart-rate and skin-resistance, for reasoning about a user's affects [23]. The sixth dimension, *Human Interaction*, characterizes the relationship between humans including simple collocation, listening to a speaker, gaze, and actual interaction as discussion. In section 3.3 we will use these six dimensions of sensing together with choices of sensor placement to categorize current sensing technology.

3.2 Physical View: Placement of Sensors

An application engineer has to consider the possibilities of sensor placement in the physical world: e.g. a traffic jam can be remotely detected by a camera or locally at each car through mutually exchanged distance and speed information. Both choices are appropriate for the intended purpose but have different side-effects: the camera has to be mounted once and works for all cars, but only at one location. As a side-effect its use could be extended to other applications, e.g. criminal search. The local set-up instead requires individual effort at each car but users have the choice to participate in the system or not and it works everywhere.

We identify four different categories of sensor placement. *In Environment* refers to stationary installed sensors, e.g. in the floor, walls, where placement can only be changed with effort. Whereas *In Environment* installations work with all users at the stationary location *on Human* has the opposite characteristics: only users wearing the sensors can participate, but therefore they are not bound to a location. *On Object* is in between the two previous categories, as objects can be personal and can be carried with a human (e.g. key), but also stay at a certain location (e.g. chair). This distinction depends on the object. Additionally, *mutual collaboration* defines sensing system that always require more than one unit in order to operate properly, e.g. triangulation of signal strength for localization.

3.3 Review: Sensors in Ubicomp Research

Based on own experience with sensors and literature review we compiled a table (Fig. 1) characterizing sensor technology in respect to the six sensing dimensions and the four sensor placement possibilities. This table should be used as reference for application developers during the process of finding the appropriate sensor for their application.

In each table cell sensors are aligned due to bandwidth consumption and quality of perception in respect to the dimension. The alignment due to precision and bandwidth should be seen as rough estimation for relative comparison between sensors occurring in the same table field. For recognizing a person's *ID* the table shows four choices of sensors for installation *in environment* in the upper left cell. Obviously, the best results can be achieved with biometric sensors [36], such as finger print or iris scan, as represented by vertical alignment in the cell. Methods based on vision [10], audio or load-cells embedded into the floor [6] deliver less quality. Horizontal alignment in the cell shows, that data generated by load-cells and finger print sensors consumes lower bandwidth than methods based on vision or audio. Inertial sensors placed *on object* and *on human* can be used to sense typical movements, e.g. perceiving the signature at a pen, for identification. [27] reports about using vision, [19] about using audio worn *on human* for people identification. Location systems [16] can also be used for identifying people at different locations. As these systems require both sensors worn by human and installed base stations those system appear in the *mutual collaboration* column. For detecting *object use* load-cells [30] have been proven useful installed both *in environment* and *on object*. Object classification with vision is well established in static settings, occlusion during dynamic use can be challenging. Audio is another option, if the object use generates characteristic sounds. The use of inertial force sensors placed *on object* has been successfully used [15, 29, 24, 13] for *object use*. Obviously, motion during *object use* can be also sensed *on human* but with less quality. Audio *on human* is also possible [21] but is an indirect measurement compared to *on object* placement. Location systems can give hints as well for *object use*, e.g. teleporting X Windows to user's current location [25]. *Location* is the most explored sensing dimension in ubiquitous computing. Load-cells [30], vision [5] and audio [7] have been explored in different project. Coarse *location* can be also gained through passive-infra-red sensors, mechanical switches or IR-barriers.

Placement	Quality of Sensing	Installed in Environment	On Object	On Human	Mutual Collaboration
ID	high low	biometric sensing vision audio load-cell	inertial sensors	audio inertial sensors	location systems
Obj. Use	high low	load-cell audio vision switch/lightbarriers	inertial sensors load-cell force/distance/capacity light	audio inertial sensors	location systems
Location	high low	load-cell radar, laser Vision PIR audio switches, IR-barrier	GPS pressure, humidity	GPS pressure, humidity	location systems diff. GPS
Bio/Emots	high low	vision audio	force/load touch	GSR oximetry inertial sensors temperature	-----
Activity	high low	Smart Board Load-cell vision PIR, pressure, capacity	-----	inertial sensors strain strips GPS	location systems
Interaction (humans)	high low	Load-cells vision audio	-----	inertial sensors vision micro GPS	-----
		Bandwidth low high	Bandwidth low high	Bandwidth low high	Bandwidth low high

Table 1: Placement vs. Dimension

On object and *on human* the primary outdoors is GPS[1], more low-level information deliver humidity, inertial [34] or pressure sensors. The variety of location systems based on *mutual collaboration* is huge: differential GPS, ultra-sound, radio etc. Sensing *bio signs/emotions* with *in environment* sensor-settings is difficult: [10] and [12] report vision and audio for reasoning on user's *bio signs/emotions*. Augmented objects measuring force and touch [2] can give some hints about *bio signs/emotions*. However, most promising are *on human* measurements such as [14, 22]. *Activity* can be well sensed with special purpose system, such as commercially available smart white boards. Load-cells, passive infra-red, pressure and capacity sensors can be used for low-level detection only. *On human* sensing has been well explored for motion activity [11]. Location system can give hints reasoned from semantical location descriptions. *Interaction* among humans has not been explored very well. *In environment* sensing systems based on vision, load-cells and audio could help to perceive characteristics of interaction, such as collocation, gestures or speech. The *on object* field is blank as objects are not involved here. *On human* the same sensors can be used as for *activity* if measurements are correlated among interactors. Location system do not really help here, as collocation is not significant for interaction.

4. DISCUSSION

As a result of the review presented in the previous section, this section discusses the appropriateness of different sensor placement for the six sensing dimensions. Table 2 evaluates our classification of sensing technologies due to the dimensions of sensing and sensor placement. In particular we differentiate between *not applicable*, if a combination does not make sense, *possible* for instances with very low quality of perception, and *good* and *very good* for more promising solutions.

[1]Actually GPS is a mutual technology requiring a receiver in collaboration with satellites in space. However, as satellites are so ubiquitous and invisible anyway we consider them as a "natural" resource and view the receivers only.

It points out, that *in environment* placement is the primary choice for *ID* sensing. As our focus is on human sensing it is not surprising that *on object* is well suited for *object use*. *On human* is suited for direct measurements of human-centric sensing aspects, such as *bio signs/emotions* and *activity*. *Mutual collaboration* sensors, such as the location systems perform best *location* sensing, but also can give hints for other dimension. Quite interestingly, each sensor placement is meaningful for it least one sensing dimension.

Placement	In Environment	On Object	On Human	Mutual
ID	++	---	+	+
Obj. Use	o	++	+	o
Location	+	o / +	o / +	++
Bio/Emots	+	o	++	---
Activity	+	---	++	o
Interaction	+ / ++	---	+	o

-- not applicable, o possible, + good, ++ very good

Table 2: Evaluation

Looking placements more globally, table 2 depicts that *in environment* and *on human* offer the best sensing results. Analyzing the dominant factors for each placement, it points out that video and audio are most prominent for *in environment* sensing. However, the perception quality depends on computational expensive recognition, as video and audio per se can only provide indirect measurements which are less power than e.g. direct *activity* with inertial sensors. Nevertheless, once an environment has been augmented with sensors, e.g. Smart Rooms, applications work without additional instrumentation of users or objects. It also can give hints for human-human interaction from an outer view.

As physical interaction with everyday object mostly involves movements, such as grasping, moving or turning, the dominant sensor technology for *object use* are inertial sensors. *On human* placement is suited for various sensors such as inertial sensors, audio, bio sensors and also video to a certain extend. In regards of human sensing *on human* also represents the closest to phenomena placement. Due to the high relevance of location in the real world *mutual collaboration* sensors can provide coarse information about *object use*, *activity* and *in environment*. This also explains why in the first years of context-aware computing mostly location was regarded as context. It can do a lot but in direct comparison with *on object* and *on human* sensing location system are in an inferior position.

5. CONCLUSION

This paper is a first step to systematize the use of sensor technology. Therefore, six dimension of sensing have been identified representing the sensing goals for physical interaction. We reviewed existing ubiquitous computing research for an evaluation of sensing technology with respect to the dimensions of sensing and physical sensor placement opportunities. This enables to support application designers finding appropriate sensors during system design.

Acknowledgments

The Smart-Its project is funded in part by the Commission of the European Union under contract IST-2000-25428, and by the Swiss Federal Office for Education and Science (BBW 00.0281).

6. REFERENCES

[1] S. Antifakos, F. Michahelles, and B. Schiele. Proactive Instructions for Furniture Assembly. In *UbiComp*, Gothenburg, Sweden, 2002.

[2] W. Ark, DC. Dryer, and DJ. Lu. The Emotion Mouse. In *Proc. of HCI International*, 1999.

[3] L. Arnstein, G. Borriello, S. Consolvo, C. Hung, and J. Su. Labscape: A Smart Environment for the Cell Biology Laboratory. *IEEE Perv. Comp.*, 1(3), 2002.

[4] T. Bohnenberger, A. Jameson, A. Krüger, and A. Butz. User acceptance of a decision-theoretic, location-aware shopping guide. In *IUI 2002*. New York, 2002.

[5] B. Brumitt, B. Meyers, J. Krumm, J. Kern, and S. A. Shafer. Easyliving: Technologies for intelligent environments. In *HUC*, pages 12–29, 2000.

[6] Philippe Cattin. *Biometric Authentication System Using Human Gait*. PhD thesis, ETH Zurich, 2001.

[7] T. Darrell, D. Demirdjian, N. Checka, and P. Felzenswalb. Plan-view Trajectory Estimation with Dense Stereo Background Models. In *ICCV*, 2001.

[8] N. Davies, K. Mitchell, K. Cheverst, and G. Blair. Developing a Context Sensitive Tourist Guide. In *MobileHCI*, 1998.

[9] A. Dey and GD. Abowd. Towards a Better Understanding of Context and Contex-Awareness. Technical Report 22, GeorgiaTech, 1999.

[10] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski. Classifying facial actions. *IEEE T. Patt. Analy. and Mach. Intell.*, 21:974 – 989, 1999.

[11] J. Farringdon, AJ. Moore, N. Tilbury, J. Church, and PD. Biemond. Wearable sensor badge and sensor jacket for context awareness. In *ISWC*, 1999.

[12] R. Fernandez and R.W. Picard. Modeling Drivers' Speech under Stress. *Speech Comm.*, 40:145–159, 2003.

[13] BL. Harrison, KP. Fishkin, A. Gujar, C. Mochon, and R. Want. Squeeze me, hold me, tilt me. In *CHI*, 1998.

[14] J. Healey and RW. Picard. Startlecam: A cybernetic wearable camera. In *ISWC*, pages 42–49, 1998.

[15] K. Hickley, J. Pierce, M. Sinclair, and E. Horvitz. Sensing Techniques for Mobile Interaction. In *UIST*, pages 91–100, 2000.

[16] J. Hightower and G. Borriello. Location systems for ubiquitous computing. *Computer*, 34(8):57–66, 2001.

[17] LE. Holmquist, R. Maz, and S. Ljungblad. Designing Tomorrow's Smart Products - Experience with the Smart-Its Platform. In *DUX*, San Francisco, 2003.

[18] H. Ishii and B. Ullmer. Tangible bits: Towards seamless interfaces between people, bits and atoms. In *CHI*, pages 234–241, 1997.

[19] N. Kern, B. Schiele, H. Junker, P. Lukowicz, and G. Troster. Wearable Sensing to Annotate Meeting Recording. In *ISWC*, Seattle, 2002.

[20] M. Langheinrich, F. Mattern, K. Romer, and H. Vogt. First steps towards an event-based infrastructure for smart things. In *Ubiquitous Computing Workshop at PACT*, 2000.

[21] P. Lukowicz, H. Junker, M. Staeger, T. v. Bueren, and G. Troester. WearNET: A Distributed Multi-sensor System for Context Aware Wearables. In *UbiComp 2002*, 2002.

[22] F. Michahelles, P. Matter, A. Schmidt, and B. Schiele. Applying Wearable Sensors to Avalanche Rescue: First Experiences with a Novel Wearable Avalanche Beacon. In *IMC Workshop*, Rostock, Germany, 2003.

[23] RW. Picard and J. Klein. Computers that Recognise and Respond to User Emotion: Theoretical and Practical Implications. *Interacting with Computers*, 14(2):141–169, 2002.

[24] J. Rekimoto. Tilting operations for small screen interfaces. In *UIST*, pages 167–168, 1996.

[25] T. Richardson, F. Bennett, G. Mapp, and A. Hopper. Teleporting in an X Window System Environment. *IEEE Personal Comm.*, 1994.

[26] D. Salber, AK. Dey, and GD. Abowd. The context toolkit: Aiding the development of context-enabled applications. In *CHI*, pages 434–441, 1999.

[27] J. Scheirer, R. Fernandez, and RW. Picard. Expression Glasses: A Wearable Device for Facial Expression Recognition. In *CHI Short Papers*, Pittsburgh, 1999.

[28] A. Schmidt. Implicit human-computer interaction through context. In *MobileHCI*, Edinburgh, 1999.

[29] A. Schmidt, K. A. Aidoo, A. Takaluoma, U. Tuomela, K. Van Laerhoven, and W. Van de Velde. Advanced interaction in context. *LNCS*, 1999.

[30] A. Schmidt, M. Strohbach, K. Van Laerhoven, A. Friday, and HW. Gellersen. Context acquisition based on load sensing. In *Ubicomp 2002*, 2002.

[31] F. Siegemund and C. Floerkemeier. Interaction in Pervasive Computing Settings using Bluetooth enabled Active Tags and Passive RFID Technology together with Mobile Phones. In *PerCom*, Fort Worth, USA, 2003.

[32] TEA - Technology for Enabling Awareness. http://www.teco.edu/tea/.

[33] K. Van Laerhoven, K. Aidoo, and S. Lowette. Real-time analysis of Data from Many Sensors with Neural Networks. In *ISWC*, Zurich, October 2001.

[34] E. Vildjiounaite, E. Malm, J. Kaartinen, and P. Alahuhta. Location Estimation Indoors by Means of Small Computing Power Devices, Accelerometers, Magnetic Sensors and Map Knowledge. In *Pervasive*, pages 211–224, Zurich, 2002.

[35] R. Want, A. Hopper, V. Falcao, and J. Gibbons. The active badge location system. Technical report, 1992.

[36] J. Wayman, A. K. Jain, D. Maltoni, and D. Maio. *Biometric Systems: Technology, Design and Performance Evaluation*. Springer, 2003.

[37] R.M. White. A Sensor Classification Scheme. *IEEE Trans. on Ultrasonics, Ferroelectrics, and Frequency Control*, pages 124–126, 1987.