NEW MULTIMODAL INTERACTION TECHNIQUES FOR MOBILE DEVICES

(PROPOSAL/EXPOSÉ)

TOBIAS HIRSCH

tobias.hirsch@mailbox.tu-berlin.de, 214675

Exposé zur
Magisterabschlussarbeit
im Fach
Kommunikationswissenschaft

Fachgebiet Audiokommunikation
Institut für Sprache und Kommunikation
Fakultät 1 – Geisteswissenschaften

Quality and Usability Lab
Institut für Softwaretechnik und Theoretische Informatik
Fakultät 4 – Elektrotechnik und Informatik

Technische Universität Berlin

Februar 2010

Gutachter

Prof. Dr. Stefan Weinzierl
Fachgebiet Audiokommunikation
Technische Universität Berlin

Hamed Ketabdar, Ph. D. (Betreuer)
Quality and Usability Lab
Deutsche Telekom Laboratories/Technische Universität Berlin

CONTENTS

ABSTRACT

Recently a new interaction technique, suitable for gesture input, was developed by researchers of the Quality and Usability Lab (Q&U Lab) respectively Deutsche Telekom Laboratories (T-LABS). With focus on this new method, a part of the thesis consists of analysing this technique in direction of gesture based interaction. This involves the evaluation of machine learning techniques and developing of a classifier that models hand gestures. After being able to identify different gestures, another task will be to find application areas for mobile devices, which ideally results in the implementation of a demonstrator on a mobile phone, e. g., iPhone.

ZUSAMMENFASSUNG

Kürzlich wurde von Forschern des Quality and Usabilty Labs bzw. der Deutsche Telekom Laboratories eine neue, für den Einsatz von Gesten geeignete, Interaktionstechnik entwickelt. Mit Schwerpunkt auf dieser wird ein Teil der Abschlussarbeit aus der Analyse der Technik in Richtung handgestenbasierter Interaktion bestehen. Dies beinhaltet die Evaluation verschiedener Methoden des machinellen Lernens und die Entwicklung eines Klassifikators zur Gestenmodellierung. Nachdem verschiedene Gesten korrekt identifiziert werden können, besteht eine weitere Aufgabe im Auffinden von Einsatzgebieten für mobile Endgeräte, was im Idealfall mit der Umsetzung eines Demonstrators auf einem Mobiltelefon abgeschlossen wird.

1 INTRODUCTION

The thesis is mainly concerned with the study of a new interaction technique. Particular interest lies in mobile use cases, such as authentication, security, gaming and music. Therefore we will use approaches from different areas like human computer interaction (HCI), ubiquitous computing (UC), machine learning (ML) or sound design. For legal reasons the new method, developed by the Q&U Lab, can only be described closer within the thesis.

Mobile devices suffer from problems of limited screen space and paucity of buttons or other interaction techniques (Gaunet and Briffault, 2009, p. 373), all in all "awkward input" (Nielsen, 2009), and thus have a reduced usability compared to more traditional user interfaces of laptop or desktop systems (pointing device/keyboard/screen) (Brewster et al., 2003, p. 1).

Unlike the notebook or PC, they give us new opportunities for ubiquitous computing (Essl and Rohs, 2009, p. 197). Especially their size enables them to being used nearly everywhere and yet stay in the background, without demanding our focus of attention, like normal PCs (Weiser, 1991).

According to a recent survey of comScore and Compete (2009, taken from emarketer.com) among US mobile phone users, Smartphones are mainly used for mobile media, E-mail, News, Applications, Social Networking and instant messaging (IM). But they hold a lot more potential than this, which is shortly described in the following.

With the arise and through advancements of mobile/smart phones tools with sufficient computing power and several sensoring capabilities (touch, position, range sensing, motion, orientation) are widely available. This allows us to explore new ways of interaction and application possibilities. By using the input of different sensors and applying methods of machine learning, we can deploy new or enhance existing interaction techniques. The lately developed one is here of special interest. Afterwards we use defined classes (special gestures) and the according classifier (responsible for matching a gesture to one of the defined ones) to control a to be developed tool on a mobile smart phone.

Thereby we think that it is very practicable and desirable to bring gesture based input to the mobile platform, as "sensed gesture [...] may be less demanding of visual attention" (Hinckley et al., 2000) and "when mobile [...] visually demanding interface designs [are] hard to operate" (Brewster et al., 2003).

2 STATE OF THE ART

2.1 GESTURES IN HCI AND UC

"The methods, theories, and models in human information processing are currently well developed." (Proctor and Vu, 2008, p. 58) and in addition "researchers in academia, commercial laboratories, and industrial development have been producing ubicomp prototypes for more than 15 years." (Quigley, 2010, p. 244).

Research on or implementation of gesture based input mainly concentrates on vision-based recognition, pen, touch, mouse gestures and accelerometers/gyros capturing motion information (Hinckley et al. (2004, p. 533), Quigley (2010, pp. 271–274), Hinckley (2008, pp. 169–170) and Wilson (2008, pp. 131–132)).

Because of similarities in the frequency domain of the sensed data in activity/gesture recognition with accelerometers (see Bao and Intille (2004), Ravi et al. (2005)) and the new input technique (again vague for legal reasons), approaches used in this discipline should fit to our needs. Therefore we can draw on data recently determined for a thesis conducted at the `T-Labs` about gesture-based authentication on mobile devices using accelerometers and gyroscopes.

2.2 SIGNAL PROCESSING AND MACHINE LEARNING

One part of machine learning is finding useful features to use in a classifier. Because the to be analyzed interaction technique is new, there is no literature about it. But as stated above there are chances, that we can use procedures from gestures with accelerometers.

Many classifiers or models exists and partly are able "to recognize which of several known gestures the user has performed" Wilson (2008, p. 185). To name a few: Fisher's linear discriminant, likelihood, neural network, decision trees, Support Vector Machine, Bayesian network, dynamic Bayesian network and Hidden Markov Model. All these models are well described in machine learning literature (Heijden et al., 2004; Bishop, 2006) and implemented in data mining software like `Weka`.

2.3 GESTURES IN APPLICATIONS

Exemplary only one potential case example is described.

MUSIC Although "Turning mobile devices into musical instruments has already been explored by a number of researchers", "Mobile music performance is only rarely discussed conceptually or analytically." (Essl and Rohs, 2009, pp. 197, 198)

It is useful to analyse current widespread tools (*not only* instruments) for music-making and performance ranging from acoustic to electronic ones. What makes them playable and enjoyable? What are commonly used standards in music production. By finding similarities in usage (e. g., special gestures, movements) and transferring them to our mobile application, it should enhance joy of use and usability. Especially papers published for the International Conference on New Interfaces for Musical Expression (http://www.nime.org/) and the Organised Sound journal offer clues for further research (e. g., Levitin et al. (2002), Traube et al. (2003), Paine (2004)).

3 APPROACHES AND METHODS

CLASSIFY INPUT TECHNIQUE According to Hinckley (2008, pp. 162–167) you can describe input devices with a set of attributes:

- input device properties (number of dimensions, indirect/direct, device acquisition time, accuracy, learning time, comfort, cost)
- input device states (e. g., out of range, tracking, dragging)
- elemental tasks (e. g., select, position)
- transfer function (mapping input to desired function)
- feedback

The new input technique will be classified by these properties.

FEATURE SELECTION AND CLASSIFICATION Feature extraction will be done with MATLAB. For classification (training and testing) either data mining software (`weka`, `RapidMiner`) or MATLAB with toolboxes, including PRTools (pattern recognition, `http://www.prtools.org`), are available. Figure 1 on page 6 shows for our needs a very simplified work flow of machine learning. Whether machine learning techniques such as Hidden Markov Models (HMM) and artificial neural networks (NN) are more efficient for classification than discriminant analysis techniques must be evaluated within the thesis.

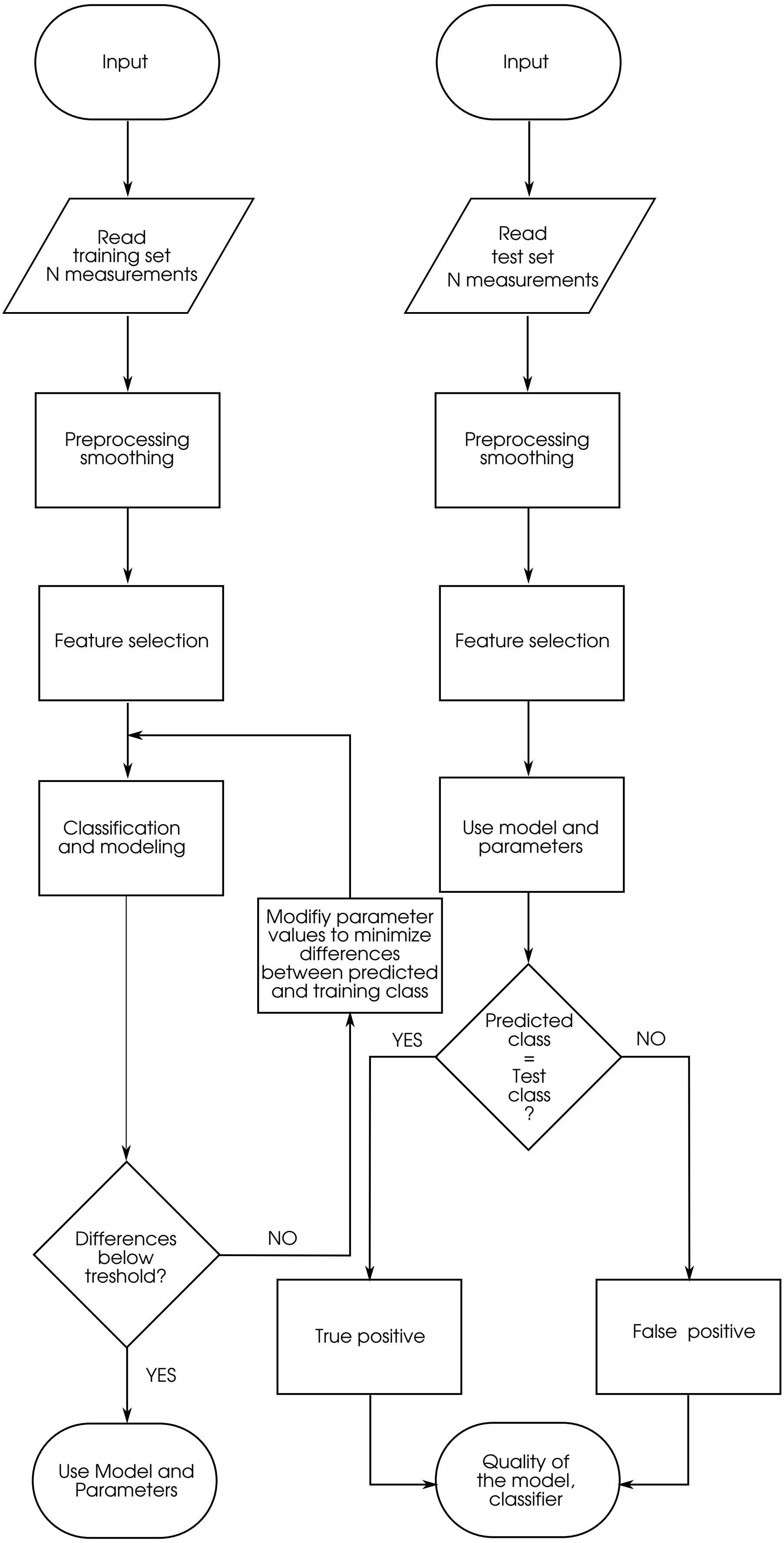

Figure 1: Flow Chart—Machine Learning, loosely based on information by Wilson (2008, pp. 183–187)

4 TIME SCHEDULE

STEP	CONTENT	TIME
Preperation	Investigation of literature, review of previous work done at T-Labs, machine learning, human computer interaction, gesture based authentication, mobile music-making	4 weeks
Signal processing	Preprocessing, feature selection, classification and modelling	4 weeks
Application	Develop a small application using the new interaction technique	4 weeks
Writing	Take part in preparing papers/publications and writing the thesis	6 weeks

Table 1: Time Schedule

Ling Bao and Stephen S. Intille. Activity recognition from user-annotated acceleration data. In A. Ferscha and F. Mattern, editors, *Pervasive*, volume 3001, pages 1–17. Springer-Verlag Berlin Heidelberg, 2004. (Cited on page 3.)

Christopher M. Bishop. *Pattern Recognition and Machine Learning*. Springer Science+Business Media, LLC, 2006. (Cited on page 3.)

Stephen Brewster, Joanna Lumsden, Marek Bell, Malcolm Hall, and Stuart Tasker. Multimodal 'eyes-free' interaction techniques for wearable devices. In *CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 473–480, New York, NY, USA, April 2003. ACM. (Cited on page 2.)

comScore and Compete. Android: Crashing the smartphone party, December 2009.
URL http://www.emarketer.com/Article.aspx?R=1007439, 2010-02-10. (Cited on page 2.)

Georg Essl and Michael Rohs. Interactivity for mobile music-making. *Organised Sound*, 14 (2):197–207, 2009.
URL http://www.eecs.umich.edu/~gessl/georg_papers/os09-mobileinteractivity.pdf, 2010-02-08. (Cited on pages 2 and 4.)

Florence Gaunet and Xavier Briffault. Designing a ubiquitous audio-based memory aid. In Judith Symonds, editor, *Ubiquitous and Pervasive Computing: Concepts, Methodologies, Tools, and Applications*, volume 1, chapter 2.11, pages 370–388. Information Science Publishing, September 2009. (Cited on page 2.)

F. van der Heijden, R.P.W. Duin, D. de Ridder, and D.M.J. Tax. *Classification, Parameter Estimation and State Estimation: An Engineering Approach using MATLAB*. John Wiley & Sons Ltd, 2004. (Cited on page 3.)

Ken Hinckley. Input technologies and techniques. In Sears and Jacko (2008), chapter 9, pages 161–177. (Cited on pages 3 and 5.)

Ken Hinckley, Jeff Pierce, Mike Sinclair, and Eric Horvitz. Sensing techniques for mobile interaction. In *UIST '00: Proceedings of the 13th annual ACM symposium on User interface software and technology*, pages 91–100. ACM, 2000. (Cited on page 2.)

Ken Hinckley, Robert J. K. Jacob, and Colin Ware. Input/output devices and interaction techniques. In Allen B. Tucker, editor, *Com-*

puter Science Handbook, chapter 20, pages 523–554. Chapman & Hall/CRC, 2. edition, 2004. (Cited on page 3.)

Daniel J. Levitin, Stephen McAdams, and Robert L. Adams. Control parameters for musical instruments: a foundation for new mappings of gesture to sound. In *Organised Sound*, volume 7 (2), pages 171–189, 2002. (Cited on page 4.)

Jakob Nielsen. Mobile usability, July 2009.
URL http://www.useit.com/alertbox/mobile-usability.html, 2010-02-10. (Cited on page 2.)

Garth Paine. Gesture and musical interaction: Interactive engagement through dynamic morphology. In *NIME '04: Proceedings of the 2004 conference on New interfaces for musical expression*, pages 80–86, Singapore, Singapore, 2004. National University of Singapore. (Cited on page 4.)

Robert W. Proctor and Kim-Phuong L. Vu. Human information processing: An overview for human-computer interaction. In Sears and Jacko (2008), chapter 2, pages 44–62. (Cited on page 3.)

Aaron Quigley. From gui to uui: Interfaces for ubiquitous computing. In John Krumm, editor, *Ubiquitous Computing Fundamentals*, chapter 6, pages 238–284. Chapman & Hall, 1. edition, 2010. (Cited on page 3.)

Nishkam Ravi, Nikhil Dandekar, Preetham Mysore, and Michael L. Littman. Activity recognition from accelerometer data. In *In Proceedings of the Seventeenth Conference on Innovative Applications of Artificial Intelligence (IAAI)*, pages 1541–1546. AAAI Press, 2005. (Cited on page 3.)

Andrew Sears and Julie A. Jacko, editors. *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies, and Emerging Applications*. Human Factors and Ergonomics. Lawrence Erlbaum Associates, 2. edition, 2008. (Cited on pages 8 and 9.)

Caroline Traube, Philippe Depalle, and Marcelo Wanderley. Indirect acquisition of instrumental gesture based on signal, physical and perceptual information. In *NIME '03: Proceedings of the 2003 conference on New interfaces for musical expression*, pages 42–47, Singapore, Singapore, 2003. National University of Singapore. (Cited on page 4.)

Mark Weiser. The computer for the 21st century. *Scientific American: Special Issue on Communications, Computers, and Networks*, September 1991.
URL http://www.ubiq.com/hypertext/weiser/SciAmDraft3.html, 2010-01-17. (Cited on page 2.)

Andrew D. Wilson. Sensor- and recogntion-based input for interaction. In Sears and Jacko (2008), chapter 10, pages 177–199. (Cited on pages 3 and 6.)