Tagungsband der

Magdeburger-Informatik-Tage 2. Doktorandentagung 2013 (MIT 2013)

Herausgeber:

Robert Buchholz
Georg Krempl
Claudia Krull
Eike Schallehn
Sebastian Stober
Frank Ortmeier
Sebastian Zug

16. Juli 2013

Impressum:

Verlag: Otto-von-Guericke-Universität Magdeburg

Verlagsnummer: 85720

Otto-von-Guericke-Universität Magdeburg
Fakultät für Informatik
Postfach 41 20
39016 Magdeburg

Herausgeber:

Robert Buchholz
Georg Krempl
Claudia Krull
Eike Schallehn
Sebastian Stober
Frank Ortmeier
Sebastian Zug

Redaktionsschluß: 16. Juli 2013

Redaktion/Gestaltung:
Eike Schallehn

Herstellung:
Magdeburger DigitalDruckerei GmbH

ISBN 978-3-940961-96-9

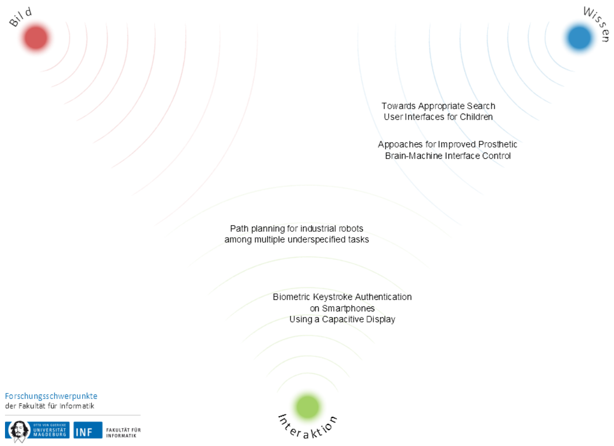
`http://www.cs.uni-magdeburg.de/`

Vorwort

Dieser Band ist eine Zusammenfassung der Beiträge zum 2. Magdeburger-Informatik-Tag (MIT). Hier präsentierten ausgewählte junge Wissenschaftler der Fakultät für Informatik der Otto-von-Guericke-Universität ihre fortgeschrittenen Promotionsprojekte. Ziel der Tagung ist die Vorstellung anerkannter Forschungsergebnisse unserer Fakultät über Fachgebiets- und Universitätsgrenzen hinweg.

Das Forschungsprofil an der Fakultät Informatik richtet sich an den drei Schwerpunkten Bild, Wissen und Interaktion aus. Der Schwerpunkt „Bild“ beschäftigt sich mit der Repräsentation, Analyse und Vermittlung bildhafter Information. Dies beinhaltet speziell die Bereiche Bildverstehen, Modellierung, Bilderzeugung und Visualisierung. Forschungsarbeiten im Schwerpunkt „Wissen“ beschäftigen sich mit den methodischen und technologischen Grundlagen des Erwerbs, der Modellierung und Repräsentation, der Verwaltung und der Verarbeitung von Daten, Informationen und Wissen. Der Schwerpunkt „Interaktion“ adressiert mit Forschungsarbeiten zu Multimodalität, Zuverlässigkeit, Sicherheit und Technologie wichtige Herausforderungen moderner Mensch-Technik-Interaktion sowie der Interaktion technischer Geräte untereinander.

Die folgende Grafik ordnet die vorgestellten Arbeiten in Bezug zu diesen drei Leitmotiven ein.



Der vorliegende Tagungsband dokumentiert sowohl die Vielseitigkeit als auch die Konvergenz der Forschungsaktivitäten von Nachwuchswissenschaftlern an der Fakultät für Informatik.

Magdeburg, den 16. Juli 2013

Robert Buchholz
Georg Krempl
Claudia Krull
Eike Schallehn
Sebastian Stober
Frank Ortmeier
Sebastian Zug

Inhaltsverzeichnis

Path planning for industrial robots among multiple under-specified tasks

Sergey Alatartsev
Computer Systems in Engineering
Faculty of Computer Science, Otto-von-Guericke University
Magdeburg, Germany
E-mail: sergey.alatartsev@ovgu.de
Supervisor: Prof. Frank Ortmeier

***Abstract*— This paper is an overview of the currently going PhD project. The main goal of the project is to optimize a sequence of industrial robotic tasks. We want to make use of the fact that robotic tasks often allow some freedom during the execution. In contrast, this extra freedom is ignored by state-of-the-art approaches. This work refers to extra freedom as under-specification. This paper presents a comprehensive overview of the state-of-the-art approaches in task sequencing. We state the thesis goals and make a short introduction to the methods developed within the PhD project. Work packages are formed based on the list of unaccomplished goals.**

I. INTRODUCTION

Nowadays industrialized countries with their high labor costs have to rely on production automation to keep their competitive advantage. One of the most flexible and powerful automation technologies that are available today is industrial robotics. Since acquisition and programming of an industrial robot are very expensive, the feasibility of using robots in production facilities depends on the efficiency with which the robot can perform its task: the more production steps a robot can perform in a given time interval, the higher the production rates are; as a consequence, the faster the robot can compensate for its initial acquisition and programming costs, the higher is the competitive advantage it provides to the company.

Virtually all robotics scenarios consist of two different types of robotic movements. The first category includes movements that are specifically required for the job, e.g., welding a seam, deburring a sharp edge or cutting a shape. During these movements the tools (e.g., a welding torch) are switched on. We call this category *effective movements* or *effective tasks*. Another category is *supporting movements* or *supporting tasks*. Supporting movement are between two effective movements. Supporting movements are not directly needed for a given job. However, they are necessary to sequence one effective movement after the other. This concept is illustrated in Fig. 1 on a simple welding example. The two welding seams – (2) and (4) – are effective movements, whereas (1), (3) and (5) are supporting movements, which are only necessary to execute the effective movements.

S. Alatartsev and F. Ortmeier are affiliated with the working group Computer Systems in Engineering, Otto-von-Guericke University of Magdeburg, Germany. {sergey.alatartsev, frank.ortmeier}@ovgu.de

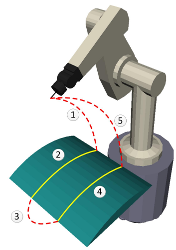


Fig. 1. Example of the alternating stages

There are two major characteristics that affect the efficiency of a given robot. The first is how fast the robot can perform its tasks (effective movements). Another characteristic is how long it takes the robot to move into a new position to perform the task after the previous task was completed (supporting movements). While effective movements are usually understood as rigid paths (defined by the task/application), supporting movements are open to computational optimization.

The current industrial practice is to let an engineer program both effective and supporting movements. In contrast, state-of-the-art research already provides solutions for computing collision-free supporting movements automatically and/or optimizing the schedule of effective tasks.

In the currently going PhD project, we intent to go one fundamental step further: we observe that not only supporting movements but also many effective movements are under-specified. For example, many robot tasks (e.g., cutting, welding, deburring) have to be performed along a given open-end curve or a given closed contour. But the actual starting point for the task is usually irrelevant. This last fact is typically neglected in existing approaches for computing optimal paths. This significantly limits the potential for optimization of the robot movements. Therefore, we propose to under-specify effective movements, by omitting precise definition of the starting/ending points and corresponding end-effector orientation. For example, each point of a closed contour could be a starting point.

Although under-specification in effective tasks greatly increase the effectiveness of optimization, another important flexibility comes from robot kinematics (i.e., multiplicity of inverse kinematics solutions). For example, the costs of supporting movements usually depend heavily on the axis configuration[1] that a robot uses for reaching any given starting/ending point of two subsequent effective tasks. In addition, real-life applications impose collision constraints that have to be considered.

The planned result of this PhD project is a set of algorithms to compute (near-) optimal robot paths by making use of under-specification (closed and open-end contours and end-effector orientation freedom), involving robot flexibility (multiple inverse kinematic solutions) and collision constraints. Optimality can be defined with respect to different metrics such as distance, time or energy consumption.

The remainder of the paper is organized as follows. Background is presented in Section II. We state the goals of the thesis in Section III. Section IV presents results of the accomplished work. Future work packages are listed in Section VI. We conclude in Section VII.

II. BACKGROUND

This section presents a comprehensive overview of the task sequencing methods in industrial robotics.

A. State of the Art of Industrial Robot programming

Today, there are two different ways to program robots in an industrial setting: either it is done *online* or *offline*. In online programming, the robot is directly taught a movement which it has to replay later in production mode [1]. The offline approach is typically based on simulation. The trajectory that has to be executed by the real robot later in production is calculated or taught in a simulated environment. Regardless of programming method, robot programming is still based on the knowledge and skills of the programmer. It requires intuition if good/optimal movements are desired. The quality of a path depends only on the experience of the programmer, and almost no optimization approaches are involved [2], [3].

B. Task sequencing in the offline environments

1) Industrial offline environments: Task sequence optimization is partially included in large simulation and programming environments, but has very limited functionality. For example, a customization for DELMIA[2] adopted for drilling applications[3] allows automatic sequencing of drilling tasks, but uses only a simple greedy algorithm[4]. Sequence optimization is not presented in RobotWorks[5]. The tasks are specified precisely with one entrance configuration, i.e., none of the additional flexibility of effective movements is included.

2) Scientific offline environments: There exist several frameworks in robotics, mostly applied in research or service robotics. Mainly logic of execution is strictly defined by a programming language (e.g., VLP\C# for MRDS [4], or Python\C++ for OpenRAVE [5]) or by a state machine (e.g., SMACH in ROS [6]). Additional information could be found in the good survey [7].

C. Existing algorithmic approaches

In general, calculating an optimal movement for multiple effective tasks depends on solving two sub-problems: (1) calculating collision-free movements between each pair of effective tasks (local planning) and (2) determining a sequence in which the tasks should be performed (global sequencing). Though both problems are well-known separately, in robotics they have to be solved in conjunction. A global sequencer cannot calculate the distance between two particular tasks without a local planner, and – vice versa – a local planner cannot lead to a global optimal path without a global sequencer.

1) Local planning: In recent years a significant amount of research has been focused on intelligent planning for supporting movements. The goal is to calculate an optimal path between two robot configurations. Normally the planning is done in C-space[6]. The two most well-known approaches are Probabilistic RoadMaps (PRM) [9] and Rapidly-growing Random Trees (RRT) [10]. These approaches typically do neither consider the optimization of a sequence of tasks nor can they deal with under-specification. However, they are very effective for planning optimal paths/trajectories between two effective tasks.

2) Global sequencing: Normally the problem of searching for an optimal sequence can be reduced to the Traveling Salesman Problem (TSP). This problem deals with the search for an optimal sequence of points with given distances between them and a constraint that every point should be visited once [11]. One of the first who discovered the application of the TSP in industrial robotics for point-to-point movements was Dubowsky et al. [12]. Later the TSP application was extended with other features like involving multiple inverse kinematics solutions (solved with Genetic Algorithm (GA) [13]), robot base layout optimization (solved with GA [14]) or involving extra constraints (e.g., maximum velocity and acceleration, solved with Ant colony optimization algorithm [15]). Mentioned approaches do not involve collision constraint into the planning. No under-specification in task descriptions is considered.

[1]In general, every 6D point (position and orientation) can be reached by a robot with eight different robotic configurations (e.g., elbow up vs. elbow down).

[2]See http://www.3ds.com, accessed on June 25, 2013.

[3]DELMIA V5 Robotic Drilling Application, see http://www.delfoi.com/web/products/delfoi_products/en_GB/drilling-app/, accessed on June 25, 2013.

[4]See http://en.wikipedia.org/wiki/Greedy_algorithm, accessed on June 25, 2013.

[5]Compucraft Ltd, RobotWorks, see http://www.compucraftltd.com, accessed on June 25, 2013.

[6]C-space is the space of all possible angle values for every robot joint, also called joint- or axis space [8].

3) Combined approaches: The combination of a global sequencer and local planning got a lot of attention within the last 10 years and different researchers refer to it with different names. Wurll et al. [16] was one of the first who observed this problem and referred to it as *Multi-Goal Path Planning (MTP)*. Here, a goal is usually understood as a point in T-space[7] or in C-space. Later other researchers proposed approaches to solve MTP (e.g., [17], [18], [19]), however, all of them assumed that tasks are uniquely specified.

4) Under-specification: All of the above mentioned algorithms require an *exact* specification of the target points in T-space or in C-space. No optimization process was involved in allocation of target points.

Under-specification of tasks was investigated in the collision-free path planning domain [20]. This approach addresses under-specification by using rectangular boxes instead of single target points. The planning can only be performed from the current position to one of many goals. The algorithm chooses the optimal goal only from the point of view of a local planner.

Gentilini et al. [21] used some under-specification for the robot with mounted camera on its end-effector. They model this problem as a Traveling Salesman Problem with Neighborhoods (TSPN). TSPN is an extension to standard TSP, which allows each point to be moved within a given area [22]. They represented TSPN as Mixed-Integer Non-Linear Program and implemented a heuristic to speed up the exact solver.

D. Summary

Industrial robot programming is still intuition-based approach. Existing industrial and research programming environments require a very precise specification of robot movements. Optimization of task sequence is ignored or presented with very primitive methods.

Existing scientific approaches model sequencing problem mostly as TSP. Very few approaches consider extra-freedom of task execution, but they are either too application-specific or do not involve robot kinematics or collision constraints or sequencing in general.

We want to fill in this gap and to develop a general and flexible format of task description as well as sequencing algorithms that will consider robot flexibility and allow constructing near-optimal collision-free paths.

III. GOALS OF THE THESIS

The **main goal** of this doctoral project is to develop methods for computing near-optimal robotic movements by making use of task under-specification. In particular, this leads to the following intermediate subgoals:

1. Develop algorithms to solve sequencing of tasks that are represented as:
   1.1. 1D closed contours.
   1.2. 1D open-end curves.
   1.3. 3D volumes
2. Allow selecting "one from many" entrance/ending specifications.
3. Extend the task description with under-specification of the robot end-effector orientation.
4. Involve robot kinematics into the planning.
5. Involve collision constraints into the planning.

We point out three types of possible tasks geometries: closed contours (e.g., cutting a hole or gluing), open-end curves (e.g., welding a seam or deburring an edge) and 3D volumes (e.g., remote laser welding or object inspection). These types of under-specification cover a large variety of domains where the tasks have to be optimized. Efficient and effective method has to be developed that is capable to solve sequencing of such types of tasks.

IV. ACCOMPLISHED WORK

A. Analysis

1) Existing sequence optimization problems: Previously we have already stated some modeling problems that were applied by other researchers. Further, a short outlook on TSP-like problems is presented. These modeling problems are applied or planned to be applied in the current PhD project.

The most known modeling problem is Traveling Salesman Problem (TSP) [11]. The goal is to find the minimal-cost circle tour that visits all the points. Although, this problem can efficiently model the sequencing problem it gives no freedom for complex tasks.

There are other TSP-like problems, that allow some freedom. One of them is Generalized TSP (GTSP), which is an extension of TSP, where the minimal-cost circle tour should visit sets of points. However it leads to discretization of the tasks and causes unwanted errors.

Close-Enough TSP (CETSP) [23] is another modeling problem. CETSP goal is to find the minimal-cost circle tour such that every point is visited within a certain radius. The main drawback is that only a minor part of robotics tasks could be efficiently modeled with CETSP.

Another related problem that takes flexibility into account is the Touring-a-sequence-of-Polygons Problem (TPP) [24]. The goal of the TPP is to find the minimal-cost circle tour that visits a sequence of regions. Although the problem require a predefined sequence, it is capable to find the minimal-cost tour point positions in the regions.

The most general among all discussed problems is TSP with Neighborhoods (TSPN) [22]. The goal of the TSPN is to find the minimal-cost circle tour through a set of regions such that every region is visited once. It has no discretization error (in contrast to GTSP), it supports areas of different shapes (in contrast to CETSP) and it is capable of calculating both a sequence and point positions (in contrast to TPP).

2) Existing methods for solving sequencing problems: We suggest to make use of TSPN for modeling under-specified robot tasks are that represented as closed contours. This type of under-specification is very common in many industrial applications, e.g., cutting, deburring, welding.

[7]T-space is a 6D space where 3 dimensions specify the robot's end-effector position and 3 dimensions specify its orientation.

Previously, TSPN received significant attention in the domain of approximation algorithms [25]. One can find comprehensive surveys in [26] and [27].

As the TSP got larger attention from the researchers than TSPN, a natural idea could be to adapt existing successful methods. TSP adaptations were already successfully done for GTSP (e.g., Lin-Kernighan [28]) or for CETSP (e.g., splitting the problem to the combination of TPP and TSP [23]). Another approach is to optimize points allocation inside the areas with a greedy algorithm (i.e., the nearest point from the next largest area is added to the set) and then calculate the TSP tour [29]. One more way to apply algorithms from the TSP domain is to represent TSPN as GTSP, since the transformation from GTSP to TSP is known (e.g., [30]). However in practice it greatly increases the search space [31]. TSPN was also solved with Mixed-Integer Non-Linear Programming solver [21]. Proposed method is capable of solving TSPN close to optimum, however it was tested on the benchmarks with up to 16 regions though some industrial applications consist of much larger number of tasks[8].

Heuristics for TSP could be split into two categories: tour-construction and tour-improvement heuristics [33]. After a tour is obtained, it can be improved by applying tour-improvement heuristics, e.g., 2-Opt or 3-Opt [34].

3) Suitable simulation environment choice: OpenRAVE [5] was chosen for prototypical implementation. It is an open source and easy to use tool. It has predefined library IKFast for calculating inverse kinematics (in contrast to MRDS [4]). OpenRAVE supports COLLADA, and programming is done in C++/Python that simplifies the comparison with state-of-the-art approaches.

B. Developed algorithms

Similar to other approaches, we apply known algorithms from the TSP domain to solve TSPN. We propose two heuristics: Constricting Insertion Heuristic (CIH) [35] and Constricting 3-Opt (C3-Opt).

1) Constricting Insertion Heuristic: We split the TSPN into two problems TSP and TPP. We solve these problems not one after another (in contrast to [23]), but rather simultaneously. Insertion Heuristic (IH) [36] was chosen for solving TSP. As a TPP solver, so called Rubber-band algorithm (RBA), proposed by Pan et al. [37], is applied.

Basic principle of IH is to incrementally construct a tour by adding new points into specific positions. There are multiple modifications of the IH, but commonly they are different in three ways: strategy 1 of initializing a tour (e.g., tour with points from the convex hull), strategy 2 of choosing a point to insert (e.g., nearest or farthest), strategy 3 of choosing where in the tour to insert these point (e.g., minimizes the tour distance).

We involve RBA into strategy 3. It means that whenever a new point should be added to the tour, the tour is constricted with RBA (constricted means that point locations on the borders of the contours are optimized). Multiple calls of RBA leads to tours of quality better than [23], and the heuristic is much faster than [21].

To search for the point location on the contour any one-dimensional optimization technique could be applied, e.g., Golden Section Search optimization method [38] or Bisection method.

2) Constricting 3-Opt: 3-Opt heuristic [39] is a certain case of the K-exchange algorithm developed by Lin [40]. Evaluation [33] shows that if K is greater than three, the effectiveness of the approach decreases. Therefore, normally the case K=3 is used.

The basic idea is to sequentially select K edges from the tour and reconnect them in all possible ways. If the reconnection causes a decrease of the tour cost, the algorithm starts from the beginning; otherwise, the next K edges are selected. The algorithm stops when there are no more exchanges that could improve the tour. There are seven possible ways of reconnecting three edges in 3-Opt. It was shown in [41] that only two ways of reconnecting are required to cover all possible combinations.

We proposed the adaptation of the 3-Opt for TSPN. We refer to this approach as Constricting 3-Opt. The core idea is to constrict the tour (i.e., to perform a local search) after every reconnection. If this local search gives near-optimal tour improvement, it could lead to the local optimum fast. Therefore, we applied a simplified RBA algorithm for constricting. Position of every point, newly added to the tour, is optimized. We also check that at some point of time, cost of the newly created tour could exceed the cost of the original tour. In that case it make no sense to continue the swapping.

After the new tour is constructed (i.e., after the swapping is done) and the points are optimized (i.e., simplified RBA is executed) then its cost is compared with the original tour. In case if swapping and constricting bring an improvement, the new tour replaces original tour.

C. Evaluation

Proposed algorithms were evaluated on two different sets of benchmarks. The first benchmark set with known optimal values was proposed by Gentilini et al. [21] with up to 16 contours. It is available online[9]. The second benchmark set was proposed by Alatartsev et al. [35] with up to 70 contours. It is also available online [42]. It is hard to calculate the optimal values for large test instances, therefore we make the comparison with best obtained values by different heuristics.

Summary of the evaluation is presented in Table I. CIH appeared to be a fast heuristic with a small average error. The drawback is a high maximum error that appears on some of the instances. One possible improvement is to simply apply 3-Opt heuristic to optimize the tour and then apply RBA to optimize the point positions on the contours. This leads to a slight decrease of the errors, although it requires more computational time. We refer to this approach as CIH (3-$Impr.$). As C3-Opt is an improvement heuristic, we evaluate it on three different input tours: generated randomly

[8]There are welding applications with 250 goals [32].

[9]TSPN Instances: http://wpweb2.tepper.cmu.edu/fmargot/ampl.html

TABLE I

EVALUATION OF THE PROPOSED METHODS (EXCEPT HIS) ON TWO SETS OF TEST INSTANCES

		Heuristics					
		HIS	CIH	CIH (3-$Impr.$)	NN→ C3-Opt	Rand→C3-Opt	CIH→C3-Opt
Test instances	average error	0.148	0.319	0.285	0.190	0.003	0.001
of	max error	1.090	6.260	5.440	2.390	0.070	0.020
Gentilini et al. [21]	average $t(ms)$	650.417	11.972	14.351	38.206	62.645	28.728
Test instances	average error	-	5.067	3.366	1.552	0.262	0.320
of	max error	-	13.980	13.950	5.170	1.500	2.080
Alatartsev et al. [42]	average $t(s)$	-	1.652	3.992	146.896	244.805	68.924

(Rand), by Nearest Neighborhood heuristic (NN) and by CIH. Random tour allows obtaining better improvements, but it is the most time consuming approach. The best choice is to use a combination of CIH and C3-Opt, as it produces small errors in a feasible computational time.

V. WORK IN PROGRESS

Above presented algorithms (CIH and C3-Opt) are effective methods that are able to produce near-optimal tours. However, these methods use Cartesian space metrics and do not involve robot kinematics into the solving process. Current work is concentrated on an adaptation of the developed techniques to the robot model.

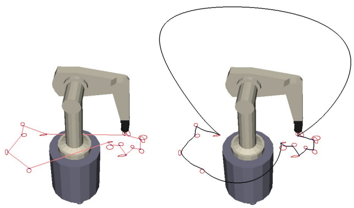


Fig. 2. Difference between the shortest paths in Cartesian space (left) and configuration space (right)

We use Puma 560 robot that has to cut out several contours. The current approach is to first construct the tour and then obtain the robot motion in the configuration space. Often there could be a huge gap between distance estimation in Cartesian and configuration spaces, see Fig. 2. The current goal is to resolve this problem and involve it into the optimization process.

VI. FUTURE WORK

Above we described the results of the accomplished work. In order to fulfill the stated thesis goals, the following work packages have to be done:

A. Work Packages

1) WP1: In this work package, an algorithm for the sequencing of open-end curve and 3D volume tasks has to be developed. These problem were already solved by researchers (e.g., [43], [21], [44])). Although solutions for particular cases exist, a combination of these types of extra-freedom with previously stated under-specification of task is the core research part of this work package. One possible idea is to extend previously developed algorithms with some other local search methods. For 3D volumes local search methods should work in multi-dimensional spaces, e.g., Gauss-Seidel or Gradient descent methods [45].

2) WP2: In this work package, we are going to extend the created algorithmic approaches to allow optimization of the choice of one entrance/ending among possible regions. In particular, it could lead to solving of GTSP for regions, but not for configurations.

3) WP3: Many tasks require specific orientation of the end-effector while being performed, e.g., keeping torch in the normal direction to the welding seam. Nevertheless, often there is a certain freedom, e.g., in remote laser welding maximum deviation from the normal vector could be 15 degrees [44]. Often such freedom also appears in grasping tasks, spot-welding, etc. In this work package we are going to make use of that type of under-specification. One possible way is to make planning not in Cartesian space but in T-space. Another way could be to make planning first in Cartesian space and when the tour is obtain, make constricting already in T-space.

4) WP4: The process of choosing a right sequence is closely related to the distance metrics. Algorithms consider this information as reliable. However, it is computationally expensive to call collision-free planner every time, when a distance information is required. There are some solutions for sequencing problems modeled as TSP (see Section II-C.3). But collision-free path planning in combination with under-specified sequencing is still an open problem.

5) WP5: In this work package we will evaluate the developed methods on the real world cases and compare the obtained results with currently applied methods such as greedy planning. Heuristics often depend on different parameters (e.g., desired precision or convergence speed). Tuning of these coefficients on the real scenarios should be conducted.

VII. CONCLUSION

In this paper we presented the state of the currently going PhD project. The overall aim of the doctoral project is to develop new methods for automatic generation of near-optimal robotic movements among multiple under-specified tasks. We restrict ourselves to scenarios where the specification of effective tasks is available in computer-readable format and where no online planning (e.g., sensor integration) is used.

The core idea is to make use of under-specification of effective tasks to open up potentials for optimization. Two efficient heuristics were developed that are capable of solving the under-specification of the closed contour tasks. Current work is concentrated on involving robot kinematics into the planning. The next steps are to extend variety of supported under-specified tasks and develop missing algorithmic approaches.

REFERENCES

[1] B. Siciliano and O. Khatib, Eds., *Springer Handbook of Robotics*. Springer Verlag, Berlin, 2008.

[2] G. Biggs and B. MacDonald, "A survey of robot programming systems," in *Proceedings of the Australasian Conference on Robotics and Automation*, 2003.

[3] Z. Pan, J. Polden, N. Larkin, S. V. Duin, and J. Norrish, "Recent progress on programming methods for industrial robots," in *Proceedings for the joint conference of 41st Internationel Symposium on Robotics and 6th German Conference on Robotics*, 2010.

[4] J. Jackson, "Microsoft robotics studio: A technical introduction," *Robotics & Automation Magazine, IEEE*, vol. 14, pp. 82–87, 2007.

[5] R. Diankov and J. Kuffner, "OpenRAVE: a planning architecture for autonomous robotics," Robotics Institute, Pittsburgh, PA, Tech. Rep. CMU-RI-TR-08-34, July 2008.

[6] M. Quigley, B. Gerkey, K. Conley, J. Faust, T. Foote, J. Leibs, R. W. E. Berger, and A. Ng, "ROS: an open-source Robot Operating System." in *Open-Source Software workshop of the International Conference on Robotics and Automation (ICRA)*, 2009.

[7] A. Elkady and T. Sobh, "Robotics middleware: A comprehensive literature survey and attribute-based bibliography," *Journal of Robotics*, vol. 2012, p. 15, 2012.

[8] S. LaValle, *Planning algorithms*. Cambridge University Press,, 2006.

[9] L. Kavraki, P. Svestka, J. Latombe, and M. Overmars, "Probabilistic roadmaps for path planning in high-dimensional configuration spaces," in *IEEE International Conference On Robotics and Automation*, 1996, pp. 566–580.

[10] J. J. Kuffner and S. M. Lavalle, "RRT-Connect: An efficient approach to single-query path planning," in *IEEE International Conference On Robotics and Automation*, 2000, pp. 995–1001.

[11] D. L. Applegate, R. E. Bixby, V. Chvatal, and W. J. Cook, *The traveling salesman problem: a computational study*. Princeton University Press, 2007.

[12] S. Dubowsky and T. Blubaugh, "Planning time-optimal robotic manipulator motions and work places for point-to-point tasks," *IEEE Transactions on Robotics and Automation*, vol. 5, pp. 377–381, 1989.

[13] P. T. Zacharia and N. A. Aspragathos, "Optimal robot task scheduling based on genetic algorithms," *Robotics and Computer-Integrated Manufacturing*, vol. 21, pp. 67–79, 2005.

[14] K. Baizid, R. Chellali, A. Yousnadj, A. Meddahi, and T. Bentaleb, "Genetic algorithms based method for time optimization in robotized site." in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2010, pp. 1359–1364.

[15] W. Bu, Z. Liu, and J. Tan, "Industrial robot layout based on operation sequence optimisation," *International Journal of Production Research*, vol. 41, pp. 4125–4145, 2009.

[16] C. Wurll, D. Henrich, and H. Wörn, "Multi-Goal Path Planning for Industrial Robots," *International Conference on Robotics and Application (RA99)*, 1999.

[17] S. N. Spitz and A. A. G. Requicha, "Multiple-goals path planning for coordinate measuring machines," in *IEEE International Conference on Robotics and Automation*, 2000, pp. 2322–2327.

[18] L. B. Gueta, R. Chiba, J. Ota, T. Ueyama, and T. Arai, "Coordinated motion control of a robot arm and a positioning table with arrangement of multiple goals," in *IEEE International Conference on Robotics and Automation, ICRA*, 2008, pp. 2252–2258.

[19] M. Saha, T. Roughgarden, J.-C. Latombe, and G. Sánchez-Ante, "Planning tours of robotic arms among partitioned goals," *Robotics Research*, vol. 25, pp. 207 – 223, 2006.

[20] D. Berenson, S. Srinivasa, and J. Kuffner, "Task space regions: A framework for pose-constrained manipulation planning," *International Journal of Robotics Research (IJRR)*, vol. 30, no. 12, pp. 1435–1460, October 2011.

[21] I. Gentilini, F. Margot, and K. Shimada, "The travelling salesman problem with neighbourhoods: MINLP solution," *Optimization Methods and Software*, vol. 0, pp. 1–15, 2011.

[22] E. M. Arkin and R. Hassin, "Approximation algorithms for the geometric covering salesman problem," *Discrete Applied Mathematics*, vol. 55, pp. 197–218, 1995.

[23] W. Mennell, "Heuristics for solving three routing problems: close-enough traveling salesman problem, close-enough vehicle routing problem, sequence-dependent team orienteering problem," Ph.D. dissertation, University of Maryland, 2009.

[24] M. Dror, A. Efrat, A. Lubiw, and J. S. B. Mitchell, "Touring a sequence of polygons," in *35th annual ACM symposium on Theory of Computing*. ACM Press, 2003, pp. 473–482.

[25] J. S. Mitchell, "A constant-factor approximation algorithm for TSP with pairwise-disjoint connected neighborhoods in the plane," in *Proceedings of the 2010 annual symposium on Computational geometry*, ser. SoCG '10, 2010, pp. 183–191.

[26] S. Arora, "Approximation schemes for NP-hard geometric optimization problems: A survey," *Mathematical Programming*, vol. 97, pp. 43–69, 2003.

[27] J. S. Mitchell, *Shortest paths and networks*. Chapman & Hall/CRC, 2004, ch. 27, pp. 607–641.

[28] D. Karapetyan and G. Gutin, "Lin–Kernighan heuristic adaptations for the generalized traveling salesman problem," *European Journal of Operational Research*, vol. 208, no. 3, pp. 221–232, 2011.

[29] K. M. Elbassioni, A. V. Fishkin, and R. Sitters, "Approximation algorithms for the euclidean traveling salesman problem with discrete and continuous neighborhoods," *International Journal of Computational Geometry and Applications*, pp. 173–193, 2009.

[30] P. Oberlin, S. Rathinam, and S. Darbha, "Today's traveling salesman problem," *IEEE Robotics & Automation Magazine*, vol. 17, no. 4, pp. 70–77, 2010.

[31] X. H. Shi, Y. C. Liang, H. P. Lee, C. Lu, and Q. X. Wang, "Particle swarm optimization-based algorithms for TSP and generalized TSP." *Information Processing Letters*, vol. 103, no. 5, pp. 169–176, 2007.

[32] Z. Pan, J. Polden, N. Larkin, S. Duin, and J. Norrish, "Automated offline programming for robotic welding system with high degree of freedoms," in *Advances in Computer, Communication, Control and Automation*, ser. Lecture Notes in Electrical Engineering, Y. Wu, Ed. Springer Berlin Heidelberg, 2012, vol. 121, pp. 685–692.

[33] D. S. Johnson and L. A. McGeoch, *Local search in combinatorial optimization*, E. Aarts and J. K. Lenstra, Eds. John Wiley and Sons, London, 1997.

[34] K. Helsgaun, "An effective implementation of the Lin–Kernighan traveling salesman heuristic," *European Journal of Operational Research*, vol. 126, pp. 106–130, 2000.

[35] S. Alatartsev, M. Augustine, and F. Ortmeier, "Constricting Insertion Heuristic for Traveling Salesman Problem with Neighborhoods," in *23rd International Conference on Automated Planning and Scheduling (ICAPS)*, 2013.

[36] R. Hassin and A. Keinan, "Greedy heuristics with regret, with application to the cheapest insertion algorithm for the TSP," *Operations Research Letters*, vol. 36, no. 2, pp. 243–246, Mar. 2008.

[37] X. Pan, F. Li, and R. Klette, "Approximate shortest path algorithms for sequences of pairwise disjoint simple polygons." in *Canadian Conference on Computational Geometry*, 2010, pp. 175–178.

[38] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical recipes – the art of scientific computing*, 3rd ed. Cambridge: Cambridge University Press, 2007.

[39] F. Bock, "An algorithm for solving traveling-salesman and related network optimization problems," in *Unpublished manuscript associated with talk presented at the 14th ORSA National Meeting*, 1958.

[40] S. Lin, "Computer solutions of the traveling salesman problem," *The Bell System Technical Journal*, 1965.

[41] C. Rego and F. Glover, *The Traveling Salesman Problem And Its Variations*. Kluwer Academic Publishers, 2002, ch. Local search and metaheuristics, pp. 309–368.

[42] S. Alatartsev, M. Augustine, and F. Ortmeier. (2012) http://euromover.cs.uni-magdeburg.de/cse/robotics/tspn/.

[43] G. Reinhart, U. Munzert, and W. Vogl, "A programming system for robot-based remote-laser-welding with conventional optics," {*CIRP*} *Annals - Manufacturing Technology*, vol. 57, pp. 37 – 40, 2008.

[44] A. Kovács, "Task sequencing for remote laserwelding in the automotive industry," in *23rd International Conference on Automated Planning and Scheduling (ICAPS)*, 2013.

[45] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*, 1st ed. Springer Netherlands, 2004.

Towards Appropriate Search User Interfaces for Children

Tatiana Gossen, M.Sc.
Data and Knowledge Engineering Group, Faculty of Computer Science,
Otto von Guericke University Magdeburg, Germany
tatiana.gossen@ovgu.de
Supervisor: Prof. Dr.-Ing. Andreas Nürnberger

***Abstract*—The Internet is an integral part of the lives of our children nowadays. Using the Internet, in particular search engines, children search for information for school, for their individual interests or simply for entertainment. Unfortunately, research shows that children face many difficulties when searching and browsing for information, even though there are several search engines that are designed especially for children. In our work we emphasize the importance of usability of search user interfaces (SUIs) for children. We aim to develop search user interfaces that are conform with children's skills and competencies. In this paper, we provide an overview of the undergoing PhD thesis, its objectives, achieved results and work planned in the future. Therefore, a detailed plan of research is given. It consists of four parts, i.e. a survey in the field of information retrieval for children, evaluation of current children' web search engines, analysis of children's information-seeking behavior and design of web search user interfaces for children.**

Keywords—Information Retrieval, Children, Search User Interface

I. Introduction

The time of the digital natives has come. Nowadays, computer and Internet access is available in almost every household and the number of children who use modern technologies and Web services from early age grows every year. According to recent report [12], children ages five to nine spend about 28 minutes daily online and this time continuously grows. The German 2010 KIM[1] study [13] reports that about 60% of the German children of ages six to thirteen use the Internet and 70% of them use search engines. Children exploit the Internet for entertainment, e.g. online games and to communicate or chat with others. Furthermore, they also use the Internet for research related to their school activities or for other information like news, weather and sports [14]. More than half of the children search the Internet predominantly alone [13].

In order to support children in their search, special search engines for children have been launched. Currently, their main purpose is helping children to find child appropriate content in the WWW. Unfortunately, still not all children succeed in information inquiry and especially younger children experience difficulties. For example, almost 50% of six and seven years old children claim to have problems in information acquisition [13]. Another important aspect, that is often forgotten, is the usability of those search engines. It is significant that these search engines match the particular skills of children to increase their usability for this specific user group. Furthermore, these search engines are designed for children in general, i.e. the target group consists of users of a very wide age range. That is why, research on usability of search systems for children is of importance.

When designing tools for children, there is a need to target very narrow age groups [15]. However, we use age groups in the sense of a more abstract category defining a set of specific capabilities when growing up[2]. Cognitive abilities, emotional maturity and knowledge of a fourteen years old and a seven years old child strongly differ. In this work we concentrate on children of ages seven to eleven[3]. We chose this age range using the knowledge about human cognitive and psychosocial development [17], which explains that human development occurs in stages. Human abilities are different at each stage and new knowledge, abilities and skills are built upon the acquired ones. The age boundaries for each development stage are approximate and may vary from child to child. More specifically, the age range we chose falls into the "industry versus inferiority" period of child's psychosocial development, age 6–12 [18]. In this period it is important that a child succeeds in finding information. In this way, he or she feels competent and develops self-confidence. In contrast, if a child is not able to find good results, he or she may develop a feeling of incompetence [18]. Besides the immaturity in the emotional domain, children's cognitive abilities are also not fully formed [19]. According to Piaget [19] children of age seven to eleven are in the *concrete operational stage* of development with its unique cognitive characteristics. The special characteristics of children are challenging and should be considered in the development of web search engines, including the design of web search user interfaces (UIs).

The aim of this work is to develop a web search engine that meets the needs of children, i.e. fits their cognitive abilities, knowledge and provides the necessary emotional support. We focus on the front end. Our user interface should support the children in their search in a web document collection. Furthermore, we narrow down the target group and consider children of primary school age as potential users. We mainly

[1]KIM is a German acronym for Children and Media ("Kinder + Medien, Computer + Internet"). It is a German study which is regularly conducted in the form of interviews.

[2]Child's age provides only an approximation of his or her capabilities. For an individual child an appropriate mapping to the age group has to be found, e.g. using psychological tests.

[3]As we are interested in textual information search, our user group should at least be able to read. Children can read simple texts usually by the age of nine [16]. But in case an information retrieval system has an integrated text to speech reader this restriction can be relaxed

concentrate on information search in a web document collection. During searching children intend to find information relevant to their information needs. Though web documents mostly have a textual form and are written in natural language, they can also contain pictures and other media. However, our primary focus is on textual information retrieval. In order to search successfully, children require special search engines that should be designed considering the specific requirements and needs of children. Additionally, an appropriate (web) document collection, i.e. documents for a specific age group, is an essential component in order for children to succeed in search. However, in this paper, we are not further discussing the current usability, quality or amount of websites meant for children.

This paper is organized as following. We first describe our plan of research in Sect. II. Then, in Sect. III we describe relevant research in the field of human computer interaction in information retrieval for children. In Sect. IV we describe challenges when designing web SUIs for children and propose feasible solutions that are demonstrated with a new designed SUI for children. We conclude and give directions for future work in Sect. V.

II. Plan of Research

In order to achieve our goal, the following work packages (WP) are considered to be necessary:

WP1: Survey in the field of information retrieval for children

- Aspects of children development important for information retrieval tasks
- Children's information seeking behavior
- Existing algorithms and user interface concepts
- Existing information retrieval systems for children

WP1 is completed. The results are published in [1]. The research on information retrieval for young users is still in its infancy. Currently, there are no solid and approved solutions for a child-specific IR system. Information retrieval for young users is a complex topic. It is strongly related to the cognitive science on human development and sociological studies on information-seeking behaviour of children. Much knowledge is already gained in those two fields that can be transferred to information retrieval for young users.

In our opinion, the current problem in research of information retrieval for young users is that researchers view children between pre-school age and teenagers as a consistent user group. They often do not consider that children of different age require different solutions that influence the design of information retrieval systems. Theories of human development confirm, that children in different development stages do differ in cognitive abilities and motor skills. That is why, when designing UI concept and algorithms for children, there is a need to target very narrow age groups. Current research also suffers from the lack of evaluation of recently proposed algorithms and user interface concepts. Children's information-seeking behaviour was studied mostly on keyword oriented IR systems. It would be beneficial to apply relatively new technologies like *eye tracking* to study children's usage of IR systems. Furthermore, it is still unclear how to deal with children's *loopy* browsing style. Loopy means that children click, repeat searches and revisit the same result web page more often than adults. In fact, no solution was proposed to solve this problem. This type of browsing behaviour can be a sign of children's cognitive overload. In the future, mechanism to prevent this overload should be developed. There is also no study of mechanisms for emotional support of children during the search, which is also a potential future direction.

There is also much potential in the development of new ranking algorithms for children. Until now, only one algorithm, AgeRank [20], was proposed and evaluated, but unfortunately not with children. There are some conceptual suggestions what elements should effect the ranking for children, e.g. complexity, interestingness, and affective value. Ranking algorithms based on these suggestions should be implemented in the future. An open research question here would be how much influence each of the mentioned components has on the target ranking value.

WP2: Evaluation of current children' web search engines

- *Methodology development for usability assessment of web search engines and its application on current search engines*
  We review recent work in this field and evaluate to what extent current search engines for children are appropriately supporting the motor and cognitive skills of primary school-age children. We base our study on findings of previous research and derive criteria to assess existing search engines. This work is important to estimate the current state of industrial search engines for children.

WP2 is completed. The details can be found in [3]. We suggest using the size of buttons and length of the home page as criteria for assessment how well a web search engine matches the motor skills of children. In order to evaluate to what extend current search engines for children are appropriately supporting the cognitive skills of children, we advise using such criteria as type of the search tool, support of backtracking and presentation of search results. A good presentation of search results, in turn, depends on the number of results, font size, multimedia usage, highlighting of keywords and spell checking. Our conducted study offers an overview of the quality of current search engines based on a quantitative analysis. The results of our usability assessment show that current search engines for children not always match the skills and abilities of children. We also found that most search engines for children do not offer observable advantages over the common search engine Google. This lack of adaption can lead to childrens frustration during the search. In order to avoid these problems, it is important not only to take child-friendly content into account, but also the search interface itself has to be child friendly, so that children are able to use it without problems. In order to design such child-friendly interfaces, more research is needed.

WP3: Analysis of children's information-seeking behavior

Analysis of children's information-seeking behavior is important as it provides information about how children use information retrieval systems and what difficulties they face.

The results can be further applied to improve IR systems for children.

- *Logfile analysis of targeted web search engines*
  We analyze logfiles of German web search engines for children. The aim of this research is to analyze fundamental facts about how children's web search behavior differs from that of adults. We show differences to previous results, which are often based on small lab experiments.
- *Analysis of the search behavior of young users with the help of eye tracking*
  We carry out a user study with the help of eye tracking device to uncover the differences in information-seeking behavior and perception of SUI elements between grown ups and children. Comparative user studies of this kind have not been done before. The results of this research would provide us with further knowledge how to design better SUI for children.

W3 is in progress. To our knowledge, this is the first large-scale log study of search behavior on children web search engines. Our large-scale analysis suggests that children search queries are more information-oriented and shorter on average. The purpose of *informational queries* is to find information about a topic assumed to be available on the web, in order to read about it. Meanwhile, the adults most frequent queries are *navigational* or *transactional*, with the immediate intent to reach a particular website that the user has in mind, or even further carry out some transactions, e.g. purchasing a product. Children indeed make a lot of spelling errors and often repeat searches and revisit web pages. More details can be found in [10]. Eye tracking analysis is in progress. Using eye tracking we did a first user study with primary school children in Biederitz and grown ups.

WP4: Design of web search user interfaces for children

- *Deriving of challenges in interface design and proposal of solutions*
  We derive the main challenges in designing search user interfaces for children and discuss possible solutions for each challenge.
- *Concept development*
  We design a novel web search user interface for children to demonstrate these solutions.
- *Prototyping & Evaluation*
  We do prototypes of proposed web search user interface. In order to support our work different evaluations in form of user studies, among others with eye tracking and Wizard of Oz experiment, are conducted.

W4 is in progress. We will present our first results in Sect. IV. This section is based on our publications [4], [6].

Our current research is focused on making the SUI adaptable. Search user interfaces are usually designed and optimized for a certain user group. Users within the group are similar, for example, in cognitive, fine motor and other abilities. These abilities influence the decisions made in the user interface (UI) design process of SUIs. However, especially for young and elderly users, the design requirements change rapidly due to fast changes in users' abilities, so that a flexible modification of the SUI is needed. In order to overcome this issue, we suggest to develop an *evolving search user interface* (ESUI). It adapts the UI dynamically based on the derived capabilities of the user interacting with it. The focus of a current work is a prototype study towards the concept of an ESUI. We aim to develop an adaptable SUI first to examine the user acceptance and different ways to adapt the UI elements of a SUI. This adaptable SUI should provide means for personalization, i.e. it allow changes not only in properties of user interface elements like color but also adapts the UI elements themselves and there placement to user preferences.

III. Related Work

Our survey about information retrieval (IR) for children in general where all the components of an IR systems (both front end and back end) are covered is published in [1]. In this section we describe relevant research in the field of human computer interaction in information retrieval (HCIR) for children. The related work can be subdivided into four parts: research on children's information seeking behavior, studies about children's preferences in web search interfaces, existing web search systems for children, and concepts of search user interface (SUI) for children.

A. Information Seeking Behavior

Catalog Oriented Search: In general there are two interface types for search engines, that are currently used: catalog and query oriented search engines. In query oriented search engines the user needs to input some keywords, whereas in catalog oriented search he browses/navigates[4] through predefined categories. Search engines that integrate both interface types are also common [22] and combine advantages of both. Researchers found that the browsing performance of children is better and that children prefer browsing [23]. Nevertheless, the results of a recent study [24] suggest that children prefer typing keywords rather than "browsing the main categories". This can be explained with the fact that participants have grown more accustomed to keyword-oriented search UIs (had already experience with *Google*). However the search UIs used in the study does not seem to have good navigation capabilities, e.g. categories are hidden within the interface. One reason for children to prefer navigation is that navigation imposes less cognitive load. Less user knowledge is required to recognize and react to offered terms than to recall concepts from the memory. Borgman et al. [23] explain that navigation fits to children's "natural tendency to explore". It also better fits to the fine motor skills of children. Whereas keyword oriented search engines require correct spelling and typing, navigation is possible with simple point-and-click interaction. Nevertheless there are potential problems in category navigation. As children have only little domain knowledge and a smaller vocabulary than adults do, they may have problems finding the right category. ***Catalog oriented search is easier for children than a keyword oriented search.***

[4] Although the term "browsing" is heavily used in the literature about children's information seeking behavior in the context of "category browsing", it is more accurately to apply the term "navigation" (e.g. see [21]).

Keyword Oriented Search: The limited domain knowledge of children is also a problem in keyword oriented search engines. In order to formulate a search query, the user needs sufficient domain knowledge to think about useful keywords [25]. Selecting keywords is difficult for children, because it requires the ability of thinking in abstract categories [26]. Therefore, children tend to input full natural language queries [27]. Children do not use advanced search syntax like boolean operators [28]. Furthermore, children often use too vague or too specific keywords in queries [29], [28]. This makes it more difficult for children to get relevant results. Furthermore, most children have difficulties with typing. They are not able to type commands without looking at the keyboard (touch-typing). Instead they typically hunt-and-peck on the keyboard for correct keys. By looking at the keyboard while typing, children often do not spot spelling mistakes. Utilizing keyword oriented search, which requires correct spelling, is difficult for children [30], [31]. ***Children have difficulties to find the right keywords for a query and problems with spelling.***

Further Aspects of Search: Compared to adults, children have a different browsing style in Web documents. Children's search behavior can be described by many looping (repeating) and backtracking (clicking the "Back"-button of a browser) actions, with fast reading of the retrieved documents and little focus on the search goal, while adults' browsing style is *linear* or *systematic* [28]. Children click, repeat searches and revisit the same result web page more often than adults [10]. This characteristic agrees with children's lower cognitive recall, i.e. children forget about an already visited page or are lost. Children also have difficulties to evaluate the relevance of retrieved documents to their information need [24]. Children are frustrated by too many results and do not have the ability to quickly determine the most relevant and "best" documents [32]. In task-oriented search, children look for the final "concrete" answer in documents, without trying to read and understand the content [28]. Most children visit only the first result page and click on the first item in the result list [33]. ***Children have a loopy browsing style and difficulties to determine the most relevant result.***

Findings about Children's Preferences: Naidu [34] found that children in general prefer websites with many pictures. It is consistent with Large et al. [35] whose user study results suggest using attractive screen designs, based especially on effective use of color, graphics, animation. They also suggest allowing for individual user personalization in areas such as color and graphics. Budiu and Nielsen [31] also came to the conclusion that children like movement, graphics, funny sounds, and colors. But designers should not exaggerate. Too much multimedia can overwhelm children [31]. Budiu and Nielsen [31] also found that metaphors, especially spatial navigation, work very well with children. But if the behavior of metaphorical UI elements is inappropriate usability problems appear. Furthermore, there is also evidence that children can experience difficulties with too advanced metaphorical navigation interfaces whose meaning they do not understand [24]. With regard to graphical objects, children understand icons better that text because it does not require good reading skills. But icons should also represent real-world concepts children are familiar with. ***Children like attractive colorful UI design.***

Straightforward text fonts (14 pt for young children and 12 pt for older children) and simple text layouts make reading easier for children. Both adults and children avoid reading long texts on the Web [31]. Interface elements should be large enough as fine motor skills of children are still developing and are not as good as by adults. The time for selecting a mouse target gets larger, the smaller the target object is. This means, that larger target sizes allow children to make selections more quickly [36]. Therefore, Budiu and Nielsen [31] suggest to make the clickable targets big. During the web search children tend to formulate natural queries instead of using keywords for search [31]. Thus, a large search box should be used in keyword search interfaces. Researchers also suggest that search interfaces should address both educational and entertainment needs of children [35]. Certain mouse interactions are very difficult for children. For example, they have difficulties with drag-and-drop interactions, because they can not coordinate dragging and holding at the same time [37]. However, better design decisions might help to decrease the errors by drag-and-drop interactions [38]. Children often do not use complex interactions like scrolling a web page [34]. This mostly applies to younger children. Children older than nine are fairly comfortable with scrolling [31]. ***Clickable UI elements and text fonts should be large. SUI should enable simple interactions.***

B. Existing Web Search Systems & Interface Concepts

Existing Web Search Systems for Children: An overview of existing web search engines for children and a methodology for the usability assessment of web search engines for young users can be found in our paper [3]. ***Unfortunately, current search engines for children not always match the skills and abilities of children [3].*** Existing web search engines for children provide first of all a child-safe content. Here we list several English and German search engines: *quinturakids.com*, *kidrex.org*, *onekey.com*, *askkids.com*, *kidsclick.org*, *blinde-kuh.de*, *helles-koepfchen.de*, *fragfinn.de*, *dipty.com* etc. For example, Google-based search engine *Kidrex* supports a safe search and has a funny start page. In other ways it is the same as Google, including additional information such as advertisement. Most of these web search engines (e.g. *blinde-kuh.de*, *askkids.com*, *kidsclick.org*) also care about text complexity and provide those web pages as search results, that are easy to read for children. *kidsclick.org* also provides information about the reading level of the retrieved web pages. The majority of the engines have a colorful design of a start page which should attract children attention (e.g. *quinturakids.com*). In every other aspect, existing web search engines for children have the same design as common search engines do. They have a keyword-based interface. Thus, a child should input a text query to an input field in the first place to initiate the search. A few web search engines also have flat categories. For example, *quinturakids.com* has five categories (music, history, animal, computer games, sport and recreation) which are shown as moving kites.

The visualization of search results is also very similar to standard search engines (like Google), i.e. a vertical results

Fig. 1. Simple Search mode in International Children's Digital Library, childrenslibrary.org.

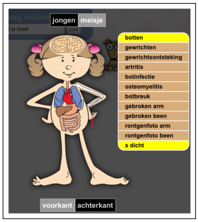


Fig. 2. *EmSe* search service: the *Body Browser* with query suggestions [44].

list of text snippets. The positive fact is, that some search engine also provide pictures along with text summaries to support the relevance filtering process of children (e.g. *helles-koepfchen.de*). Usually, there are ten results per page (as by Google). But, for example, search engine *kidsclick.org* places all the results on one page. Thus, it can be forty-fifty results on the page. The query input field moves to the bottom of the page, after results. This makes scrolling unavoidable which may be difficult for children.

One drawback of many search engines for children is the absence of spellchecking and query suggestion mechanisms which are especially important for children. For example, *askkids.com* does not process misspelled queries. Another drawback is small text fonts which make it hard to read for children. Thus, current web search engines for children not always match the skills and abilities of children. Because of that, using them might frustrate children. In order to avoid these problems, it is important not only to take child-friendly content into account. The search interface has to be child-friendly (usable for children), so that children are able to use it without problems.

Besides web search engines, there is research in digital libraries for children. Hutchinson et al. [25] developed a searching and browsing tool suitable for children for the International Childrens Digital Library. They considered children's differences in motor skills designing the system and provided large icons and simple point-and-click actions to interact with the system. Besides searching, the system supports browsing where child appropriate categories are used. These categories are represented by icons to support children with undeveloped reading abilities. Search results can be filtered by different parameters using the ca-tegory buttons. Sequential clicking on the categories leads to Boolean conjunctive operations which is also represented in the user interface. The International Childrens Digital Library is publicly available[5].

Search Interface Concepts: TeddIR [39], *CollAge* [40], *JuSe* [41] and *Imagepile* [42] are search user interfaces that are designed for young children, but mostly for preschoolers. In *TeddIR* children search books by putting tangible objects, which represent the search terms, on the screen. This helps children to overcome difficulties in spelling and deriving query terms. The *CollAge* web information retrieval system integrates search results for children's web queries with child-oriented multimedia results. For an user query, the system runs a Google image search and returns images as results in addition to existing search results. *ImagePile* displays the results as a pile of images where the user navigates horizontally to support children in determining the relevance of results. *Junior Search (JuSe)* is an interface that enables searching through adaptable picture dictionaries. Children can construct queries using these pictures. *JuSe* uses categories derived from children's vocabulary lists and parents can adjust the list, e.g. add new words.

Lingnau et al. [43] presented the *Show and Tell (SAT)* system for children's interactive search. The goal of the system is to train children's ability to identify interesting material and to connect information through telling stories. *SAT* uses a book metaphor. Children have to select an image (e.g. from a Website) using a drag-and-drop operation in order to start a search. This image is placed on the book cover. The book, its left side, is then filled with search results associated with the search image. Children can create their own stories by selecting items from search results and placing them on the right book side.

EmSe [44] is a search service for children in a hospital environment. It was designed for children of ages 8–12. The authors share the idea of a guidance avatar which is also supported by the SUI we propose in this paper. In order to overcome terminology difficulties a novel visual querying interface *Body Browser* (Figure 2) is offered which lets children explore medical information. Furthermore, *EmSe* provides children with relevant documents where medical terms are annotated with explanations.

From our point of view, current research suffers from the lack of evaluation of recently proposed SUI concepts. Children's information-seeking behavior was studied mostly on classic keyword oriented IR systems. Their searching behavior and acceptance of "new" user interfaces should be examined in comparative user studies to compare them with existing alternative UI concepts. Children's perception of user interface elements, e.g., different forms of results visualization, should also be compared in the future.

[5]http://www.childrenslibrary.org/icdl/SimpleSearchCategory?ilang=English, accessed on 13.07.2012.

IV. Knowledge Journey Design

In this section we first conduct a requirement analysis and derive challenges when designing web SUIs for children. We also discuss feasible solutions. After that, we present our search user interface *Knowledge Journey (KJ)* and describe its main components and features.

A. Design Challenges & Solutions

Emotional Support: Based on Erickson's theory of psychosocial development [18] children require emotional support and a feeling of success. So far, this problem was not covered in HCIR for young users. In the case of an ideal search engine children would always be satisfied with search results and would not get frustrated if their search does not end successfully. Until the ideal search engine is designed we suggest supporting children by their search which can be achieved for example by proper guidance. The idea here is to provide children with enough help to support their search process to avoid frustration. We propose building a guidance figure that captures children's failures, e.g. getting no results or spelling mistakes, and explain how to do better. In contrast to adults, less experienced young users (and thus those who especially require support) are willing to read instructions and thus would pay attention to well-designed help instructions [15]. Furthermore, spoken instructions would be appreciated by children whose reading skills are not well developed.

Language Support: Children, especially in the primary school age, read slowly and are still learning to write [16]. In addition, children have a limited domain knowledge [26] and difficulties with typing using a keyboard [30]. This results in problems with query formulation and spelling errors [28], [45]. Thus, spelling correction and query suggestion mechanisms in keyword based search tools are important. Furthermore, a search UI for children should provide different possibilities for children to formulate their information need. Previous research addresses this problem by suggesting alternative ways for query formulation like using a predefined term dictionary in *JuSe* or a set of tangible objects which represent the search terms in *TeddIR* (see Sect. III-B). We suggest using a menu with various categories that correspond to children's information needs. This menu should be image based and audio supported to support dual information coding [46] and therefore to allow ergonomic and fast navigation within it. Besides navigating using the menu, we also suggest to provide the opportunity of keyword-oriented search supported by spelling correction mechanisms. Children can choose the way they want to start searching. With an increasing domain knowledge (possibly gained from navigation in categories) children can employ keyword-oriented search more efficiently.

Cognitive Support: According to theories of human cognitive development, human development occurs in a sequential order in which new knowledge, abilities and skills build upon the previously acquired ones [47]. Piaget [19] describes four development stages. Children in the concrete operational stage of development learn to reason logically and have difficulties with thinking abstractly. Their understanding is limited to concrete and physical concepts. It is important to design categories which match the cognitive abilities of children. Therefore, categories used in the menu should not be abstract and the menu should have a flat hierarchical structure. Metaphors used in the user interface should be familiar to children and have a connection to the physical world (this is also advised in [31]).

Support Type	*SUIs*
Emotional Support	EmSe
Language Support	EmSe, JuSe, ICDL
Cognitive Support	ICDL, EmSe
Memory Support	SAT
Interaction Support	EmSe, ICDL, Imagepile
Relevance Support	CollAge, Imagepile

TABLE I. Correspondence between existing web SUIs and types of support they offer.

Memory Support: According to the information processing theory [17], information processing of children differs from the adults' in terms of how they apply information and what memory limits they have, i.e. children can represent and process less information. Information retrieval processes may cause children's memory to overload. This explains children's "looping" behavior during the information seeking process. Children click, repeat searches and revisit the same result web page more often than adults do [28], [45]. Thus, it is also important to show a clear back-button or just present the search result in the same window (e.g. using frames) to prevent children from getting lost. In our opinion, the aspect of memory support is not covered by the current research and researchers should pay more attention to it. Research would benefit from new approaches in *personal information management* for children. To support children's cognitive recall we suggest providing a result storage functionality.

Interaction Support: The information processing rate influences the fine motor skills of children [48], [36]. Children's performance in pointing movements, e.g. using a mouse, are lower than that of adults. Therefore, the search user interface should prefer simple point-and-click interactions and clickable interface elements should be large enough to be easily hit (this is also consistent with [31]).

Relevance Support: Children also have difficulties to judge the relevance of the retrieved documents to their information need [24]. Children are frustrated by too many results and do not have the ability to determine the most relevant and "best" documents [32]. A child-suitable form of results presentation can support children's judgment of results' relevance and provide relevance clues. Each result item should at least have a website image and its description. Akkersdijk et al. [42] also suggest displaying the results using a *Coverflow* technique where the user navigates horizontally. Coverflow allows users to concentrate on one item at a time. It also does not require complex interactions like scrolling as a vertical results list used in common search engines.

B. Knowledge Journey

Current research in SUIs for children (Sect. III-B) do not cover all the mentioned aspects. Especially cognitive support

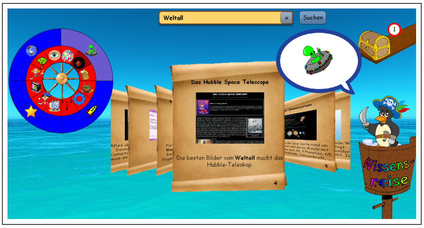


Fig. 3. Screenshot of the *Knowledge Journey* user interface: a guidance figure and a treasure chest on the right hand side, query input elements on the top, a navigation menu on the left hand side and a coverflow with search results in the middle.

is not studied well (Table I). The proposed SUIs only partially support children by their search. In opposite to previous research we aim to develop SUI that covers all the aspects.

We considered the requirements for UI design and developed a SUI for children called *Knowledge Journey (KJ). KJ* is a SUI designed for primary school children. It aims to support children in different ways, providing emotional, language, cognitive, memory, interaction and relevance support. KJ is audio supported, contains possibilities for both searching through text input and navigating using menu categories, has a guidance figure for emotional support and a result storage functionality to support cognitive recall. Knowledge Journey uses the metaphor of a treasure hunt where a user takes a journey to gather relevant search results. The interface of KJ is shown in Fig. 3. More details can be found in [4].

We evaluated our SUI in a user study with children. We emphasize that in general it is not easy to conduct user studies where children are involved. We have to collaborate with primary schools in order to organize user studies. For some schools, e.g. the mentioned school in Biederitz, it is necessary to obtain a permit from the local authorities (ministry of education). All the relevant agreements from the caretakers have to be obtained. Parents sign a release in advance agreeing to allow their children participate in the study. Therefore, studies with ten to thirty participants are common in the field. Some researchers even try to omit direct evaluations with children and conduct only usability expert revisions.

In order to evaluate our SUI we did a comparative user study with 28 participants of age seven to twelve. We evaluated our search interface against a classic keyword-based SUI with a vertical result listing. To our knowledge we are the first who did a comparative user study of search UIs for children with a fixed backend. We used a latin square design and let each participant compare between the two UIs. We mainly concentrated on the user satisfaction. We consider this factor to be more relevant usability factor for children as positive attitudes towards the system keeps children motivated. This is also consistent with other research, e.g. Mohd et al. [49] suggest that *fun* should be measured when evaluating online systems designed for children. The results of our study indicate that our SUI has much potential. 61% of participants preferred our SUI over a Google-like UI and 18% liked both. They liked new features of knowledge journey, particularly the treasure chest. More details about the study can be found in [4].

V. Conclusion

We aim to develop an IR system that is conform with children's skills and competencies. In specific, we emphasize the importance of usability of search user interfaces for children. Our goal is to develop an ergonomic search user interface for children of primary school age as potential users. In order to achieve this goal several steps are necessary. We already did a survey in the field of information retrieval for children and analyzed the specifics of IR for children. Based on findings of previous research we derived criteria to assess the usability of existing search engines and performed a usability case study. We did a logfile analysis of several German search engines for children in order to study different aspects of children's information-seeking behavior. We derived challenges in SUI design for children and proposed several solutions. We demonstrated these solutions in a first prototype of a SUI for kids and evaluated it in a user study. In the future, we are going to work on making the SUI adaptable. We also continue the development of novel SUI concepts for children.

List of Publications

[1] T. Gossen and A. Nürnberger. Specifics of Information Retrieval for Young Users: A Survey. *Information Processing & Management*, vol. 49, no. 4, pages 739–756, 2013.

[2] T. Gossen, M. Kotzyba, and A. Nürnberger. Graph Clusterings with Overlaps: Adapted Quality Indices and a Generation Model. *Neurocomputing*, 2013. in press. DOI: 10.1016/j.neucom.2012.09.046.

[3] T. Gossen, J. Hempel, and A. Nürnberger. Find it if you can: usability case study of search engines for young users. *Personal and Ubiquitous Computing*, pages 1–11, 2012. DOI: 10.1007/s00779-012-0523-4.

[4] T. Gossen, M. Nitsche, and A. Nürnberger. Knowledge journey: A web search interface for young users. In *Proceedings of the Sixth Symposium on Human-Computer Interaction and Information Retrieval (HCIR 2012)*. ACM, 2012.

[5] T. Gossen, M. Kotzyba, and A. Nürnberger. Adaptation and enhancement of evaluation measures to overlapping graph clusterings. In *Proceedings of the 1st International Conference on Pattern Recognition Applications and Methods*, 2012.

[6] T. Gossen, M. Nitsche, and A. Nürnberger. Search User Interface Design for Children: Challenges and Solutions. In *Proceedings of euroHCIR 2012 Workshop*, 2012.

[7] S. Haun, T. Gossen, A. Nürnberger, T. Kötter, K. Thiel, and M. Berthold. On the Integration of Graph Exploration and Data Analysis: The Creative Exploration Toolkit. *Bisociative Knowledge Discovery*, pages 301–312, 2012.

[8] T. Gossen, M. Nitsche, S. Haun, and A. Nürnberger. Data Exploration for Bisociative Knowledge Discovery: A Brief Overview of Tools and Evaluation Methods. *Bisociative Knowledge Discovery*, pages 287–300, 2012.

[9] T. Gossen, S. Haun, and A. Nürnberger. How to evaluate exploratory user interfaces? In *Proceedings of the SIGIR 2011 Workshop on "entertain me": Supporting Complex Search Tasks*, pages 23–24, 2011.

[10] T. Gossen, T. Low, and A. Nürnberger. What are the real differences of children's and adults' web search? In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information*, pages 1115–1116. ACM, 2011.

[11] T. Gossen, K. Bade, and A. Nürnberger. A comparative study of collaborative and individual web search for a social planning task. In *Proceedings of the LWA 2011 Workshop*, 2011.

REFERENCES

[12] A. Gutnick, M. Robb, L. Takeuchi, J. Kotler, L. Bernstein, and M. Levine, "Always connected: The new digital media habits of young children." The Joan Ganz Cooney Center at Sesame Workshop, 2011.

[13] Medienpädagogischer Forschungsverbund Südwest, "KIM-Studie 2010. Kinder+ Medien," *Computer+ Internet. Stuttgart*, 2011.

[14] E. Newburger and U. C. Bureau, *Home Computers and Internet Use in the United States, August 2000*. Bureau of the Census, 2001.

[15] J. Nielsen, "Children's Websites: Usability Issues in Designing for Kids," *Jakob Nielsen's Alertbox*, 2010.

[16] A. Stuart, "When Should Kids Learn to Read, Write, and Do Math?" WebMD, 2007, online at http://children.webmd.com/features/when-should-kids-learn-read-write-math, accessed 18.07.2012.

[17] R. Kail, *Children and their development*. Prentice Hall Upper Saddle River, NJ, 2001.

[18] E. Erikson, "Children and society," *New York: Narton*, 1963.

[19] J. Piaget, B. Inhelder, and B. Inhelder, *The psychology of the child*. Basic Books, 1969, vol. 5001.

[20] K. Gyllstrom and M.-F. Moens, "Wisdom of the Ages: Toward Delivering the Children's Web with the Link-based AgeRank Algorithm." Proc. of the Int. Conf. in Information and Knowledge Management (CIKM), 2010.

[21] R. McAleese, *Navigation and browsing in hypertext*. Intellect Books, Oxford, 1989.

[22] M. Hearst, *Search user interfaces*. Cambridge University Press, 2009.

[23] C. Borgman, S. Hirsh, V. Walter, and A. Gallagher, "Children's searching behavior on browsing and keyword online catalogs: the Science Library Catalog project," *Journal of the American Society for Information Science*, vol. 46, no. 9, pp. 663–684, 1995.

[24] H. Jochmann-Mannak, T. Huibers, L. Lentz, and T. Sanders, "Children searching information on the Internet: Performance on children's interfaces compared to Google," *Workshop of the 33rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 27–35, July 2010.

[25] H. Hutchinson, B. Bederson, and A. Druin, "The evolution of the international children's digital library searching and browsing interface," in *Proc. of the 2006 conf. on Interaction design and children*. ACM, 2006, pp. 105–112.

[26] H. Hutchinson, A. Druin, B. Bederson, K. Reuter, A. Rose, and A. Weeks, "How do I find blue books about dogs? The errors and frustrations of young digital library users," *Proc. of HCII 2005*, 2005.

[27] G. Marchionini, "Information-seeking strategies of novices using a full-text electronic encyclopedia," *Journal of the American Society for Information Science*, vol. 40, no. 1, pp. 54–66, 1989.

[28] D. Bilal and J. Kirby, "Differences and similarities in information seeking: children and adults as Web users," *Information Processing & Management*, vol. 38, no. 5, pp. 649–670, 2002.

[29] Y. Kammerer and M. Bohnacker, "Children's web search with Google: the effectiveness of natural language queries," in *Proc. of the 11th International Conf. on Interaction Design and Children*, ser. IDC '12. New York, NY, USA: ACM, 2012, pp. 184–187.

[30] P. Solomon, "Children's information retrieval behavior: A case analysis of an OPAC," *Journal of the American Society for Information Science*, vol. 44, no. 5, pp. 245–264, 1993.

[31] R. Budiu and J. Nielsen, *Usability of Websites for Children: Design Guidelines for Targeting Users Aged 3-12 Years, 2nd edition*. Nielsen Norman Group Report, 2010.

[32] A. Large and J. Beheshti, "The Web as a classroom resource: Reactions from the users," *Journal of the American Society for Information Science*, vol. 51, no. 12, pp. 1069–1080, 2000.

[33] A. Druin, E. Foss, L. Hatley, E. Golub, M. Guha, J. Fails, and H. Hutchinson, "How children search the internet with keyword interfaces," in *Proceedings of the 8th International Conference on Interaction Design and Children*. ACM, 2009, pp. 89–96.

[34] S. Naidu, "Evaluating the usability of educational websites for children," *Usability News*, vol. 7, no. 2, 2005.

[35] A. Large, J. Beheshti, and T. Rahman, "Design criteria for children's Web portals: The users speak out," *JASIST*, vol. 53, no. 2, pp. 79–94, 2002.

[36] J. Hourcade, B. Bederson, A. Druin, and F. Guimbretière, "Differences in pointing task performance between preschool children and adults using mice," *ACM Transactions on Computer-Human Interaction (TOCHI)*, vol. 11, no. 4, pp. 357–386, 2004.

[37] E. Strommen, "Children's use of mouse-based interfaces to control virtual travel," in *Proc. of the SIGCHI conf. on Human factors in computing systems: celebrating interdependence*. New York, NY, USA: ACM, 1994, pp. 405–410.

[38] A. Donker and P. Reitsma, "Drag-and-drop errors in young children's use of the mouse," *Interacting with computers*, vol. 19, no. 2, pp. 257–266, 2007.

[39] M. Jansen, W. Bos, P. van der Vet, T. Huibers, and D. Hiemstra, "TeddIR: tangible information retrieval for children," in *Proc. of the 9th Int. Conf. on Interaction Design and Children*. ACM, 2010, pp. 282–285.

[40] K. Gyllstrom and M. Moens, "A picture is worth a thousand search results: finding child-oriented multimedia results with collAge," in *Proc. of the 33rd int. ACM SIGIR conf. on Research and development in information retrieval*. ACM, 2010, pp. 731–732.

[41] T. Polajnar, R. Glassey, K. Gyllstrom, and L. Azzopardi, "Enabling Picture-based Querying and Learning with the JuSe Interface," in *Proc. of the 2nd Child Computer Interaction: Workshop on UI Technologies and Educational Pedagogy at CHI*, 2011.

[42] S. Akkersdijk, M. Brandon, H. Jochmann-Mannak, D. Hiemstra, and T. Huibers, "ImagePile: an Alternative for Vertical Results Lists of IR-Systems," *Tech. Report TR-CTIT-11-11, Centre for Telematics and Information Technology, Univ. of Twente*, no. ISSN 1381-3625, 2011.

[43] A. Lingnau, I. Ruthven, and M. Landoni, "Show and Tell: Supporting Childrens Search by Interactively Creating Stories." in *In N. J. Belkin, C. L. A. Clarke, N. Gao, J. Kamps, & J. Karlgren (Eds.), Proc. of the SIGIR Workshop on "entertain me": Supporting Complex Search Tasks*. ACM, 2011, pp. 11–12.

[44] C. Eickhoff, L. Azzopardi, D. Hiemstra, F. de Jong, A. de Vries, D. Dowie, S. Duarte, R. Glassey, K. Gyllstrom, F. Kruisinga *et al.*, "Emse: Initial evaluation of a child-friendly medical search system," in *IIiX Symposium*, 2012.

[45] T. Gossen, T. Low, and A. Nürnberger, "What are the real differences of children's and adults' web search," in *Proc. of the 34th int. ACM SIGIR conf. on Research and development in Information*. ACM, 2011, pp. 1115–1116.

[46] A. Paivio, *Mental representations: A dual coding approach*. Oxford University Press, 1990, no. 9.

[47] J. Ormrod and K. Davis, *Human learning*. Merrill, 1999.

[48] S. Card, T. Moran, and A. Newell, "The model human processor- An engineering model of human performance," *Handbook of perception and human performance.*, vol. 2, pp. 45–1, 1986.

[49] Y. Y. Mohd, M. Landoni, and I. Ruthven, "Assessing fun: young children as evaluators of interactive systems," in *Proceedings of the Workshop on Accessible Search Systems held at the 33st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, New York, USA*, 2010.

Approaches for Improved Prosthetic Brain–Machine Interface Control

Christoph Reichert
Department of Neurology
University Medical Center A.ö.R.
D-39120 Magdeburg, Germany
Email: christoph.reichert@med.ovgu.de
Supervisor: Prof. Dr. Rudolf Kruse

***Abstract*—In the present study, two essentially different approaches are presented, aiming to contribute progress in prosthetic control driven by human brain activity. The first approach addresses the limitations of noninvasive brain wave recording techniques providing a low information transfer rate and, consequently, very limited prosthetic control in terms of the number of degrees of freedom. The complex process of grasping an object is intended to be executed by an intelligent, autonomous robotic system. Therefore, the experiments presented here investigate how accurate one of several virtual reality objects can be selected for grasping only by directing attention to the target object. Two established brain responses were studied to achieve brain controlled object selection, both resulting in reliable detection rates within a closed-loop brain–machine interface (BMI). The second approach of this study operates at a higher level of neural interfacing and concentrates on the decoding of actual arm movements. For this purpose, invasively recorded electrocorticographic (ECoG) signals were decoded to predict velocity components of one hand. In parallel, a classifier serves as a movement detector discriminating movement related brain activity from resting state. The combination of both predictors represents a hybrid brain decoder that makes use of conceptually different decoding techniques. As a result, the accuracy of velocity predictions can significantly be improved with the proposed methods by cancellation of noise in nonmovement intervals resulting in prevention of involuntary movement.**

I. Introduction

Persons who lost their ability to move due to an impairment of their central or peripheral nervous system suffer from enormous restrictions in their everyday life and permanently rely on the assistance of others. A brain–machine interface (BMI) could provide an opportunity to partially restore their mobility. The BMI circumvents peripheral nerves by connecting the central nervous system with a computer that translates brain activity into commands or actions and opens the feasibility to make use of intact brain function to control a prosthetic device for grasping or even more complex arm movements.

In humans, the use of noninvasive techniques, like the electroencephalogram (EEG), is preferable over invasive recordings. However, only a small number of commands can be discriminated with noninvasively assessed motor imagery and, as a consequence, these systems do not allow for full control of complex manipulators with many degrees of freedom. The low information transfer rate that is currently achievable with those systems suggests the requirement of new strategies for noninvasive prosthesis control. In this work, a strategy is presented that enables the selection of several alternative grasping targets by attentional tasks of comparatively low cognitive load. The approach aimes to combine both development of efficient brain decoding techniques and the use of autonomous actuator control to overcome the limited information transfer from noninvasive BMIs. Two selection paradigms were applied in virtual reality to demonstrate the feasibility to initiate grasps to natural objects. Here, the implementation of BMIs relies on magnetoencephalogram (MEG) recordings.

Otherwise, if invasive techniques will become established for BMIs, higher signal quality would be available permitting higher information transfer rates. In the second part of this work this possipility is focussed and an approach to decode end point velocity of upper limbs is presented. This is realized by constructing an asynchronous movement detector being able to discriminate arm movement and nonmovement and working parallel to a velocity component decoder. This method opens up the possibility to prevent the generally oscillating output of the velocity decoder when no output is necessary. The aim of this approach is to improve noise-prone velocity predictions for potential BMI use at nonmovement intervals.

The paper is organized as follows. In the next section, related work is reported and the link to the presented approaches is stated. The third section describes the methods of the already realized tasks followed by the obtained results in section four. In section five the results are discussed. Finally, in section six objectives for future work are described.

II. Related Work

In the recent years, highly invasive techniques were tested to control prosthetic devices by voluntary modulation of brain activity [1], [2]. Commonly, movement commands are generated by motor imagery tasks aiming to decode the μ-rhythm [3]. However, a considerable percentage of people are unable to control motor imagery BMIs [4], [5]. In contrast, it was shown that a larger fraction of people is able to select items in speller paradigms using an oddball task [6] or a paradigm based on steady-state visual evoked potentials (SSVEP) [7]. In the oddball task a P300, a positive EEG deflection, is evoked when a rare target stimulus appears in a series of irrelevant stimuli. While it is often assumed that the accuracy of visually stimulated P300 spellers is independent of the direction of the eye gaze it has recently been shown that the performance of the matrix speller drops significantly if the eyes are not moved toward the target [8]. The reason is that two EEG components, the P300 and the N200, contribute information when the centre of regard is moved to the target [9], whereas only the P300 is present if the eyes do not move. This could render the P300 paradigm less useful for patients who cannot move their eyes. In this work, it will be demonstrated that the paradigms relying on attention to visual stimulations can successfully be applied to initiate targeted grasps in a visually complex virtual environment with multiple realistic objects. Importantly, it will be shown that one of the paradigms developed here works independently of user's ability to direct gaze towards the target object. This is of high relevance for the targeted user group.

The second part of the study acts on the assumption that noninvasiveness is not an essential requirement for a user. Therefore, in this part the feasibility to improve the output of predicted hand movement reconstructed from invasive recordings is investigated. Hence, brain data of higher quality compared to MEG is used here. In the BMI literature two different approaches are commonly used to reconstruct trajectories from brain activity. The first one is reconstructing a computer cursor trajectory, generally resulting from wrist movements of a joystick [10], [11] or from categorical imagination [12]. The second way is to reconstruct actual end point kinematics which are based on arm movements [13], [14]. Here, corresponding to the latter approach to reconstruct trajectories, decoding of end point velocity components based on arm movements performed during a center-out-reaching task is focused.

In contrast to the reconstruction of continuous trajectories, classification of discrete states offers an alternative view on brain patterns derived from movement. The classified states can be used as control signals as well. For instance, ECoG signals were classified to discriminate arm movement directions [15] and to discriminate categorical actual and imagined movements [16], [17]. Recently, two different gesture types were successfully discriminated [18].

For neuro-prosthetic control it is of high importance that the system can be controlled spontaneously. Recent work reported progress in onset detection of hand movements [19] and detection of reaching movement initiation [20]. However, in many of the proposed paradigms so far motor actions are event correlated, i.e. coupled to a predetermined point in time [21], but asynchronism, the independence of external events, is a requirement to a BMI that enormously increases the comfort of application but constitutes an insufficiently solved challenge. Several studies which predict limb movements not even consider a rest condition. Especially copy and tracking tasks [13], [14], [22] solely decode movement. The question how the decoder deals with the typically noisy data during resting state remains unclear. In the present study, a hybrid approach is proposed which asynchronously decodes from the ECoG directed velocity components and simultaneously detects idle phases to prevent involuntary movement. The innovation in this work is the combination of two conceptually different methods that induce an improvement of an asynchronously predicted continuous control signal in terms of reduced noise in rest intervals.

III. Materials and Methods

A. *Object selection for grasping*

Two paradigms were implemented in a BMI fashion relying on different brain wave characteristics and providing different advantages for potential users. The first paradigm is based on the ability to measure SSVEPs when subjects direct their gaze to one of several stimuli flickering at predefined frequencies. Since the SSVEP is most prominent when the stimulus falls in the fovea centralis, participants have to fixate the target. Consequently, potential users would be required to be able to move their eyes (dependent BMI) even though the main BMI target group is not able to do so. This drawback is compensated in the second paradigm employed in this study. The so called oddball paradigm relies on the P300 potential which is evoked approximately 300 ms after a rare stimulus was perceived in a series of irrelevant stimuli. The expected brain response is even present under covert attention constraints and therefore permits the implementation of an independent BMI.

The experiments were employed in six to nine runs. The runs were performed in three different modes that served different purposes. In instructed selection mode (ISM) subjects were instructed to attend a predetermined target object. In this mode true classifier labels are available which are required to train the classifier and to reliably determine the decoding accuracy. In the first two ISM runs random feedback was provided since in these initial runs no classifier was available. The second selection mode was the free selection mode (FSM) where subjects were free to choose the target object. Finally, the grasp selection mode (GSM) differed from the FSM only in terms of the feedback which was a grasp of a virtual robot to the decoded target object followed by lifting and presenting the object rather than a short static feedback. The labels of FSM and GSM trials were applied to calculate the decoding accuracy but not to train the classifier.

1) Stimulus presentation: In order to simulate a realistic environment, we presented objects in a virtual reality (VR) scenario (Figure 1). The VR scene was projected on a screen located 1 m in front of the subject at a refresh rate of 60 fps. We defined circular regions on a table which were used i) to stimulate the subjects with visual events by lighting up the object background and ii) to provide cues and feedback by colouring the region's shape. For static feedback the decoded target object was marked by a green circle immediately after classification. All other objects were marked by red circles. To provide realistic feedback the model of a robot (Mitsubishi RV E2), equipped with a three finger gripper (Schunk SDH), was part of the scene. The virtual robot is designed to mimic actual movements of the real robot. Specifically, an autonomously calculated grasp to the selected object is visualized.

2) Recordings: As a noninvasive technique, the MEG was recorded with a whole-head BTi Magnes 248-sensors system at a sampling rate of 678.17 Hz. Simultaneously, the electrooculogram (EOG) was recorded for subsequent inspection of eye movements. MEG data and event channels were instantaneously forwarded to a second workstation capable of processing the data in real-time. In two experiments the intention of volunteers to select one of several selectable realistic objects for grasping was determined. The decoding results were intended to initiate a grasp of a robotic gripper.

Fig. 1. VR scenario used for visual stimulation. This snapshot shows a flash event of an object.

In the experiments, 19 subjects participated in the SSVEP based paradigm and 17 subjects participated in the P300 based paradigm whereas seven of them participated in both. All of the participants gave written informed consent. The studies were approved by the ethics committee of the Medical Faculty of the Otto-von-Guericke University of Magdeburg.

3) Object selection by visually evoked potentials:

a) Paradigm: In the paradigm described here, object selection was carried out by recognizing SSVEPs. Four different objects were presented which were positioned at equidistant points with a view angle of 8.5° between adjacent objects and 12° between diagonal opposed objects. Moreover, the VR scenario applied here differs from the depicted scene in Figure 1 by absence of the robot model and by a rotated view angle directed to the longitudinal side of the table. The target object was predefined by the experimenter before each ISM run. A trial started by changing the positions of the objects randomly. It followed a short orientation interval to permit the fixation on the target object at its new position. Subsequently, all object backgrounds were flickered at frequencies 6.67 Hz, 8.57 Hz, 10.0 Hz and 15 Hz, respectively while frequencies were locked to positions upper left, lower right, lower left and upper right, respectively. The stimulation interval length was 5 s immediately followed by feedback presentation. The subjects were instructed to subsequently signalize the correctness of the feedback by thumb button press and erroneous feedback by index finger button press. In total, 32 trials per run were employed. In this experiment no GSM trials were performed.

b) Data processing: An interval of length 4.5 s, starting from flicker onset was considered the data segment to be classified. Frequency decomposition was performed by calculating the discrete fourier transform at the four alternative destination frequencies. Since visual components were expected to contribute to the SSVEP detection, only 59 occipital located sensors were involved in the analysis to empirically reduce the feature space. A regularized logistic regression classifier was applied to discriminate which of the four frequencies were attended by the subject. After each ISM run, the classifier was trained with the recently acquired data.

4) Object selection by covert attention:

a) Paradigm: In the oddball paradigm employed here, six objects were aligned as shown in Figure 1 while the visual angle between outmost left and right objects was 8.5°. The objects were marked by flashing their background for 100 ms. These intensifications were displayed in random order with an inter-stimulus-interval of 300 ms. Each object was marked five times per selection trial resulting in a stimulation interval length of 10 s.

Subjects were instructed to fixate the black cross centred to the objects and to count how often the target object was marked. The counting ensured that attention was maintained on the stimulus stream. In addition, subjects were instructed to avoid eye movements and blinking during the stimulation interval.

One run consisted of 18 selection trials. The experiment started with ISM, whereas the target object was cued by a light grey circle and subjects were instructed to attend this target object. In the following FSM and GSM runs, an erroneous detection of the attended object was signalled by the subject saying the word “no” after the feedback was presented and silence otherwise.

b) Data processing: The online data stream was cut into intervals including only the stimulation sequence, immediately after the respective data were available. The MEG data were then band-pass filtered between 1 Hz and 12 Hz and down sampled to 32 Hz sampling rate. Then, the stimulation interval was cut in overlapping 1000 ms segments starting at each flash event. In ISM, the segments were labelled as target or nontarget segments depending on whether the target or a nontarget object was marked.

As a suitable classifier for high dimensional feature spaces in binary discrimination tasks and a proven technique in single trial MEG discrimination [23], [24] the linear support vector machine (SVM) algorithm was chosen. In contrast to traditional P300 analysis, epochs were not averaged but the P300 presence or absence was classified in single epochs, while the selection of appropriate features was left to the pattern recognition algorithm. Classification was performed in the time domain, meaning that the magnetic flux at 32 time steps per sensor was involved as classifier input. To reduce the number of features, 104 sensors located farthest from the vertex (the midline sensor at the position halfway between inion and nasion) were empirically excluded since this region is the expected site of the P300 response. The number of sensors was further reduced by selecting the 64 sensors providing the highest weights in an initial SVM training. The selected feature set was then used to retrain the classifier after each run conducted in ISM.

B. Reconstruction of arm movement

In this section, data accompanying actual arm movements and recorded invasively from ECoG electrodes are analysed. The analysis comprises two prediction methods, linear regression and binary classification. Electrodes were implanted to locate the focus of intractable epilepsy disease and were exclusively placed on clinical grounds. Four patients (age range 18–35 years) participated in the experiment, all having a different number of runs recorded (2 to 16 runs), depending on recording time. The experimental recordings were conducted in parallel to the clinical monitoring and presented minimal risk to the participants as approved by the UC San Francisco and UC Berkeley Committees on Human Research. All participants provided informed consent.

1) Task: Patients performed a center-out-reaching-task sitting upright in a hospital bed. The targets were presented on a screen that was horizontally located in front of the patient. The participants were instructed to use a stylus to perform the movements. The movement performing arm was chosen contralateral to the grid location, i.e. opposed to the hemisphere reflecting activity of this limb, in 21 of 23 runs. Arm movements were cued by the appearance of target objects on the touch screen. The number of targets varied with six or eight (in 4 runs 8 targets were presented). Each run started with the appearance of a rectangular target in the center of the screen. After a delay of 600 ± 200 ms one of the peripheral targets appeared. The target positions were located equally distributed on an imaginary circle around the center (radius of 15 cm) and remained fixed through the session. After another delay of 100-500 ms the center field disappeared, which was the cue to execute the movement to the target location. When the stylus reached the target location, the central field reappeared and the participant moved the stylus back to the center, which initiated the next trial. Target positions were determined randomly, but the number of trials per target was equal within one block. The total number of trials per block ranged from 120 to 320 (avg: 200.4, std: 52.3).

2) Recordings: Brain signals were recorded intracranial from subdural implanted electrode arrays. Except patient P2 who had six strips of six electrodes implanted, the electrodes used to record ECoG signals were 64-channel 8 x 8 grids with 4 mm electrode diameter and 10 mm center-to-center spacing. The stylus x and y position on the screen as well as event triggers were synchronously recorded with the brain signals. A 256-Channel pre amplifier (PZ2-256, Tucker Davis Technologies) was used at a sample frequency of 3 kHz.

3) Data Processing: The recorded ECoG signal time series data were re-referenced by removing the common average reference (CAR) by subtraction of the average signal over electrodes, whereas electrodes with poor quality were rejected from the CAR calculation. Slow signal drifts were removed by high pass filtering the data at 0.5 Hz. Position parameter x and y were low pass filtered at a cutoff frequency of 5 Hz to reduce noise.

From the preprocessed ECoG data a spatio-temporal feature space was constructed, involving low frequency potentials (LP) as well as spectral features of the high gamma band. For LP extraction, the ECoG recordings were smoothed with a 5 Hz lowpass 4^{th} order butterworth filter and subsequently down sampled to 40 Hz for classification and to 10 Hz for regression. High gamma band features were extracted by multitaper decomposition [25], using a time window of 500 ms and five tapers which corresponds to a bandwidth of 6 Hz. The taper coefficients between 65 Hz and 118 Hz were accumulated representing the high gamma activity magnitude. The time window was shifted in steps that corresponded to the respective LP sampling rate. In order to involve temporal features, sampling points within a prediction window of width 800 ms were concatenated. This procedure was applied to each channel. No feature selection was applied to the data. Rather, algorithms that internally weight relevant features and produce sparse solutions of predictor coefficients were applied.

In order to achieve brain controlled reconstructions of continuous velocity components of the hand, a linear regression was calculated between the LP features and one of the velocity components v_x, v_y and the magnitude $\|\vec{v}\|$. In matrix form, the equation to be solved can be written as

$$\vec{m} = \vec{w}\mathbf{E} \tag{1}$$

where $\vec{m}$ is a vector of the output signal to be estimated, i.e. the elements in $\vec{m}$ represent time series of v_x, v_y or $\|\vec{v}\|$. The ECoG signals are represented by the matrix $\mathbf{E}$ which essentially is a train set in the spatio-temporal feature space. The vector of weights is denoted as $\vec{w}$ and solves the equation with minimal error. Each column in the signal matrix $\mathbf{E}$ represents one of the shifted prediction windows at time step t and affects the corresponding element in the output signal $\vec{m}$. To solve the regression problem, the ridge regression algorithm was applied. This method shrinks the coefficients in $\vec{w}$ by imposing a constraint to their size [26], leading to a sparse solution. In order to determine the complexity parameter, we applied five exponentially growing, empirically chosen values to three different subsets of the current train set and chose the parameter revealing the highest average correlation for the final regression.

In parallel to the regression, a SVM was trained to discriminate movement from nonmovement intervals. The classification was done independent of the direction of the arm movement, thus serving as an asynchronous movement detector. Here, asynchronous means that the classified time intervals are not locked to any known event. This is realized by continuous temporal shifting of the prediction window analogous to the regression. The feature space was also represented by LP features and additional high gamma features. Another difference to the feature space used for regression is the higher temporal resolution, but the prediction window width and spacing was identical. Labels for movement and nonmovement were determined by thresholding the velocity $\|\vec{v}\|$ of the stylus. Two thresholds were calculated for the labeling, one threshold defined the upper limit for nonmovement, the second defined the lower limit for movement. In order to obtain an equal number of samples per class, samples of the larger class closest to the respective threshold were rejected. Time intervals involving velocities in between the thresholds were excluded from the training phase but classified in the test phase. Although they are classified, it is not possible to evaluate the correctness in the validation approach, but it ensures the time-continuity in terms of a detection task.

Finally, both predictors were combined to achieve an improved control of decoded hand velocity. The binary classifier output was transformed to a movement probability time series. Specifically, each prediction was weighted with surrounding predictions to estimate a movement probability vector $\vec{p}$ by convolving the classifier output $\vec{c}(c_t \in \{0, 1\})$ with a $L1$

normalized Bartlett window of size 5 sample points (corresponding to a window size of 500 ms). This smoothing step reveals a time series of movement probabilities caused by the classifier. Assuming that this estimation is highly correlated with the actual hand velocity, movement probability was adopted to choke the noisy reconstruction. Since velocity components are directly related to movement, the probabilities in $\vec{p}$ were multiplied element wise with the predictions obtained by regression analysis:

$$\hat{m}_t = p_t m_t \quad (2)$$

This procedure suppresses the reconstructed velocity parameter at moments when the detection algorithm assesses a movement of the patient's arm unlikely, and it retains the velocity when the classifier detected a movement.

4) Validation: The evaluation of both predictors and their combination was achieved in a cross validation fashion, where one run was equally split in ten intervals and one of the intervals was left out for validation. The remaining task relevant brain signals (avg: 375.6 s, std: 74.4 s) were used to estimate the regression coefficients $\vec{w}$ and train the classifier. By repeating this procedure for each validation interval, predictions were determined over the whole recording block except at interval limits where the prediction window takes 800 ms. All processing steps involving information of hand movement were performed separately in each cross validation step. Reconstructed velocity components were evaluated by determining the Pearson correlation coefficient r between recorded and reconstructed velocity components. Detection accuracy was determined as percentage of correct predictions involving only time points where the current velocity clearly defines movement or nonmovement.

IV. Results

A. Accuracy of the object selection BMI

Decoding accuracy was determined as the ratio of correctly selected objects divided by the total number of selections. All subjects performed the task reliably above guessing level which was 25 % and 16.7 %, respectively.

In the SSVEP based selection experiment, on average 74.4 % of trials were correctly classified by the BMI while single subject performance ranged from 43.2 % to 91.7 % correct selections. In ISM, average decoding accuracy was 72.0 %, in FSM 81.3 %. Single run decoding accuracies are depicted in Figure 2(a). The histograms suggest that decoding accuracy increases with the number of runs performed. Furthermore, the figure shows that in FSM subjects tend to achieve higher decoding accuracies whereas the increase is not statistically significant.

The experiment relying on P300 based object selection permitted correct recognitions of 77.7 % on average of all trials performed while single subject accuracies ranged from 55.6 % to 92.1 %. In the ISM, the average accuracy was 73.9 % and 85.9 % in the FSM. Here, this performance difference is statistically significant (Wilcoxon rank sum test: p=0.03). When subjects received feedback by moving the virtual robot to the grasp target the average accuracy was even higher with 91.2 %. Figure 2(b) depicts the evolution of decoding accuracies over runs. The average decoding accuracy increases gradually over time. The histograms also demonstrate the higher accuracy achievement in FSM runs. Note that the system achieved perfect detection in eight of the twelve subjects who received virtual grasp feedback. However, only six selections were performed by each subject in these grasp selection runs.

An established measure for comparison of BMIs is the information transfer rate (ITR) which combines decoding accuracy and number of alternatives to a unique measure. Here, the ITR was calculated according to the method of Wolpaw et al. [27]. This makes the selection paradigms better comparable due the different number of alternatives and stimulus duration. The SSVEP based object selection revealed bit rates of 1.4 to 17.4 bit/min (avg: 10.0 bit/min). A comparable bit rate was achieved in the P300 based selection with 3.4 to 12.0 bit/min for single subjects and 8.1 bit/min on average. A paired t-test on the ITRs of subjects who participated both experiments revealed no significant difference in bit rates per minute. However, bit rates per trial were significantly higher with the P300 based selection ($t_6 = 3.3, p < 0.05$). This indicates comparable performances regarding speed but a higher hit rate in the P300 task. Note that the maximum achievable bit rate with the applied stimulation schemes is 24.0 bit/min and 15.5 bit/min, respectively.

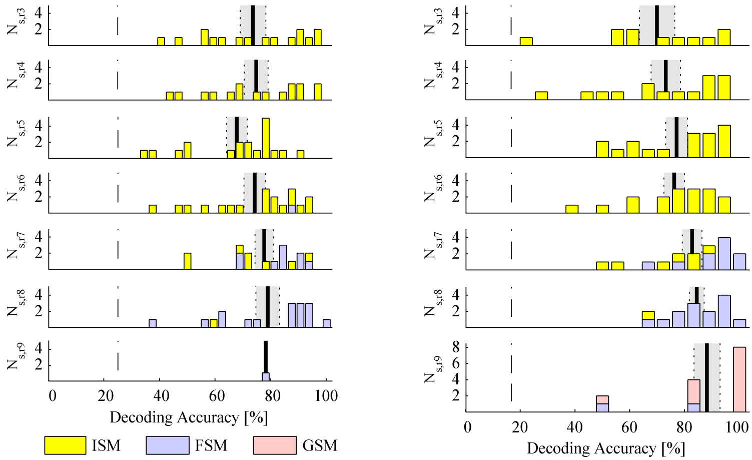
The covert attention requirement was controlled online by observation of the subjects' eyes on a video screen and offline by inspection of EOG measurements . All subjects followed the instruction to keep fixation in the P300 based selection experiment.

B. Arm movement predictions

The velocity components v_x, v_y and $\|\vec{v}\|$, considered for brain activity decoding and suited to move a prosthetic device, could all be predicted with significant correlations using the proposed regression approach. Correlation coefficients ranged from 0.567 to 0.869 for velocity magnitude and from 0.100 to 0.827 for v_x and v_y components. It can be observed that reconstructions of $\|\vec{v}\|$ are substantially more accurate (avg: $r = 0.739$, std: 0.077) than those of directional velocity components ($p < 0.001$). Prediction accuracies of components v_x (avg: $r = 0.542$, std: 0.219) and v_y (avg: $r = 0.512$, std: 0.227) reveal no significant difference.

Decoding accuracy of the binary classifer ranged from 78.5 % to 97.4 % correct (avg: 91.5 %, std: 4.3 %). Averaged decoding accuracy for single subjects were P1: 92.5 % (std: 4.6 %), P2: 90.2 % (std: 1.9 %), P3: 89.8 % (std: 0.9 %) and P4: 88.2 % (std: 5.3 %). Note that the samples which are involved in the calculation of the decoding accuracy are clearly separated concerning the absolute velocity of the moved stylus. A conclusion about unspecified samples in terms of movement is not given with the discrimination rate, because even though there is a classifier output for those samples, true labels were not available.

The final step was to merge the outputs of the movement detecting classifier and the velocity reconstruction aiming an improvement of the unstable output. When weighting velocity parameter reconstruction with the probability of movement p_t correlation increases on average by 0.135–0.144 (std: 0.043–0.062) depending on the predicted component (Figure 3(b)).



(a) SSVEP (b) P300

Fig. 2. Performance histograms. The ordinate indicates the number of subjects N_s who achieved the respective decoding performance out of 33 possible percentage bins (0/32–32/32) in the SSVEP based selection (a) and 19 bins (0/18–18/18) in the P300 based selection (b) outlined on abscissa. Each histogram shows data from a different, chronologically numbered run (r3–r9). Yellow bars indicate ISM runs, blue bars FSM runs and light red bars indicate runs in which grasp feedback was presented. Bars falling in the same accuracy bin are stacked. Vertical dashed lines indicate the guessing level and thick lines indicate the average decoding accuracies over subjects (standard error marked grey).

A paired t-test over all performed runs revealed significantly higher correlation coefficients for each of the three reconstructed velocity components ($10.2 \leq t_{22} \leq 15.9, p < 0.001$). To demonstrate the reduction of undesired movement, test data were split in fast movement and slow-movement by using the median velocity magnitude as threshold. Afterwards, the correlation of measured and reconstructed velocity magnitude was calculated separately. In nonmovement intervals the correlation increase is considerably higher than in the movement intervals. Figure 3(c) demonstrates these improvements. In Figure 3(a) it is shown how the movement probability determined by classification helps to reduce noise from velocity parameter reconstructions based on regression. In this example, the discrimination rate for detection of movement is 90.5 % which enables an enhancement of the correlation from 0.686 to 0.888.

V. Discussion

In the present work, two approaches for improved control in neural interfaces were investigated, a noninvasive BMI to initiate a grasp to one of several objects and a feasibility study for decoding invasively recorded cortical motor activity to drive a prosthetic device.

The first approach demonstrated that well-known brain responses are suited to reliably select one of several realistic objects for grasping. Two selection paradigms, based on SSVEPs and on the P300, were implemented to asses the rate of successful selections. It turned out that both paradigms reveal a comparable ITR, measured in units of time. However, the P300 paradigm allows for more accurate object selection and generally allows for a greater number of objects to choose from. Furthermore, in contrast to the SSVEP experiment the performance in the P300 experiment was obtained without orienting gaze to the target, despite earlier reports, indicating that eye movements greatly improve performance in a P300 speller [8], [28], [9]. Thus, a P300 based gripper might be useful for patients who cannot easily orient gaze to the target object. Importantly, in both experiments user training and simultaneous decoder calibration were kept below 10 to 20 minutes which is short compared to the amount of user training needed in invasive prosthetic control based on motor activity [1]. In the experiment employed with the oddball paradigm, it was demonstrated that the performance of the system improves with training and with permitted spontaneity, while this trend also was present in the SSVEP experiment. Furthermore, the results suggest that performance improves even further when realistic visual feedback is provided by moving a robotic gripper that performs an online calculated grasp. Regarding the length of the chosen stimulation intervals one can expect that the reliability of the system can be further increased by extending the stimulation interval [29]. Note that system reliability is often more important for the user than a rapid but error prone detection of intention.

The first part of the study has the advantage that non-invasive recording techniques can be used to drive a BMI. Furthermore, autonomous robotic grasping permits security features to prevent accidental actions. However, this approach

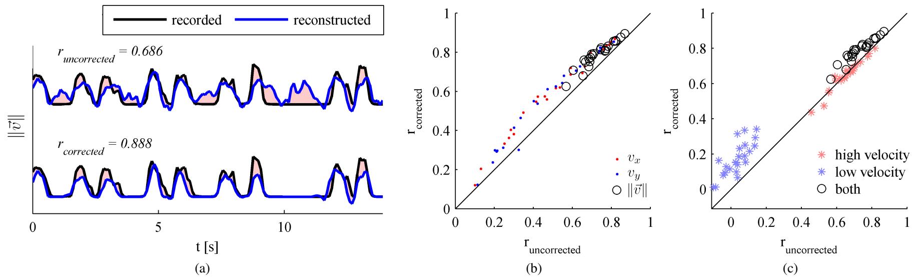


Fig. 3. Comparison of corrected and uncorrected velocity predictions. (a) Noise reduction on velocity magnitude. The upper time series shows the recorded velocity magnitude compared with the reconstruction for an exemplary interval. The light red surface demonstrates the error of the reconstruction. In the lower time series the probability of movement determined by classifier predictions was multiplied with the reconstruction. This reduces the error in nonmovement intervals and consequently increases the correlation between recorded and reconstructed signal. (b) Correlation coefficients of single blocks for all reconstructed velocity components before and after correction. (c) Correlation improvements for the velocity magnitude separately for high and low velocities.

currently provides no asynchronous control and is limited to the implemented gestures. Therefore, the second approach of this study addresses the question how reliable end point movement of one hand can be reconstructed from brain activity asynchronously. A linear regression was used to predict velocity components using a spatio-temporal feature space of low-frequent potentials revealing significantly correlated reconstructions. This initially demonstrates the ability to reconstruct directed arm movements from cortical activity. In a parallel analysis, a movement detection was implemented. The binary output was smoothed to get a probability of movement and weighted with the linear regression output. This approach causes a suppression of the directed movement decoding if the classifier detects nonmovement and consequently increases the correlation of recorded and reconstructed velocity components. The benefit solely occurs in actual nonmovement intervals. This step improved the correlation of recorded and reconstructed hand movement highly significant for all reconstructed velocity components. It is important to note that a high discrimination accuracy of the classifier is essential to improve the velocity reconstruction. Otherwise, a contrary effect could be achieved. The applied algorithms are potentially applicable in a BMI and, specifically, they are running asynchronously, i.e. no information about any event onset is necessary. This is an important requirement for autonomous controllable prosthetic devices.

VI. Outlook

Although suitable decoding accuracies were achieved with object selection paradigms, it is preferable to further improve decoding accuracies. Potential increase of decoding performance might occur with powerful feature extraction which have to be tested offline with the recorded MEG data. One established approach to be focused is spatial filtering. The main goal is to reduce the number of sensors and simultaneously enhance the signal-to-noise ratio. As a further advance, in future projects the paradigms might be extended to operate asynchronously.

Similar objectives are focussed to improve the invasively induced motor control. The enhancement of the signal-to-noise ratio and extraction of informative and discriminative features is supposed to improve predictions. Particularly, in ECoG measurements the detection of epileptic activity and noisy channels is a central challenge to be solved.

VII. Conclusion

As an important step towards the development of assistive systems for severely impaired patients, the results reported here indicate that noninavsive BMI in combination with an intelligent actuator can be used in real world settings to grasp and manipulate objects. Brain-machine interfacing based on covert attention is a fundamental requirement for patients who cannot easily orient gaze to the target object. The oddball paradigm presented here simulates a realistic setting for impaired people using only attention-based brain activation for control. From the second part of this study, it can be concluded that reconstructed velocity components of real endpoint arm movement decoded from cortical activity can significantly be improved by training a binary classifier in parallel to suppress noise in nonmovement intervals. The implementation of asynchronous movement detection is an important step to a more reliable BMI control of a prosthetic limb. One can argue that the suggested methods considerably can improve the controllability of a prosthetic device controlled by cortical activity.

Acknowledgment

This work has been supported by the EU project ECHORD number 231143 from the 7th Framework Programme and by Land-Sachsen-Anhalt Grant MK48-2009/003. The author would like to thank H. Hinrichs and J.W. Rieger for appropriation of funds and their qualified comments; M. Kennel and U. Schmucker from the Fraunhofer Institute for Factory Operation and Automation IFF, Magdeburg, for providing the VR application and L. Secundo, E. F. Chang, R. Ivry, J.W. Rieger, R.T. Knight from the Helen Wills Neuroscience Institute, UC Berkeley, for providing the ECoG data.

REFERENCES

[1] L. R. Hochberg, D. Bacher, B. Jarosiewicz, N. Y. Masse, J. D. Simeral, J. Vogel, S. Haddadin, J. Liu, S. S. Cash, P. van der Smagt, and J. P. Donoghue, "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm." *Nature*, vol. 485, no. 7398, pp. 372–375, May 2012.

[2] M. Velliste, S. Perel, M. C. Spalding, A. S. Whitford, and A. B. Schwartz, "Cortical control of a prosthetic arm for self-feeding." *Nature*, vol. 453, no. 7198, pp. 1098–1101, Jun 2008.

[3] G. Pfurtscheller, C. Neuper, C. Guger, W. Harkam, H. Ramoser, A. Schlögl, B. Obermaier, and M. Pregenzer, "Current trends in Graz Brain-Computer Interface (BCI) research." *IEEE Trans Rehabil Eng*, vol. 8, no. 2, pp. 216–219, Jun 2000.

[4] C. Vidaurre and B. Blankertz, "Towards a cure for BCI illiteracy." *Brain Topogr*, vol. 23, no. 2, pp. 194–198, Jun 2010.

[5] C. Guger, G. Edlinger, W. Harkam, I. Niedermayer, and G. Pfurtscheller, "How many people are able to operate an EEG-based brain-computer interface (BCI)?" *IEEE Trans Neural Syst Rehabil Eng*, vol. 11, no. 2, pp. 145–147, Jun 2003.

[6] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger, "How many people are able to control a P300-based brain-computer interface (BCI)?" *Neurosci Lett*, vol. 462, no. 1, pp. 94–98, Oct 2009.

[7] I. Volosyak, D. Valbuena, T. Lth, T. Malechka, and A. Grser, "BCI demographics II: how many (and what kinds of) people can use a high-frequency SSVEP BCI?" *IEEE Trans Neural Syst Rehabil Eng*, vol. 19, no. 3, pp. 232–239, Jun 2011.

[8] P. Brunner, S. Joshi, S. Briskin, J. R. Wolpaw, H. Bischof, and G. Schalk, "Does the 'P300' speller depend on eye gaze?" *J Neural Eng*, vol. 7, no. 5, p. 056013, Oct 2010.

[9] S. Frenzel, E. Neubert, and C. Bandt, "Two communication lines in a 3 3 matrix speller." *J Neural Eng*, vol. 8, no. 3, p. 036021, Jun 2011.

[10] A. P. Georgopoulos, F. J. P. Langheim, A. C. Leuthold, and A. N. Merkle, "Magnetoencephalographic signals predict movement trajectory in space." *Exp Brain Res*, vol. 167, no. 1, pp. 132–135, Nov 2005.

[11] E. C. Leuthardt, G. Schalk, J. R. Wolpaw, J. G. Ojemann, and D. W. Moran, "A brain-computer interface using electrocorticographic signals in humans." *J Neural Eng*, vol. 1, no. 2, pp. 63–71, Jun 2004.

[12] G. Schalk, K. J. Miller, N. R. Anderson, J. A. Wilson, M. D. Smyth, J. G. Ojemann, D. W. Moran, J. R. Wolpaw, and E. C. Leuthardt, "Two-dimensional movement control using electrocorticographic signals in humans." *J Neural Eng*, vol. 5, no. 1, pp. 75–84, Mar 2008.

[13] D. W. Moran and A. B. Schwartz, "Motor cortical activity during drawing movements: population representation during spiral tracing." *J Neurophysiol*, vol. 82, no. 5, pp. 2693–2704, Nov 1999.

[14] A. B. Schwartz and D. W. Moran, "Motor cortical activity during drawing movements: population representation during lemniscate tracing." *J Neurophysiol*, vol. 82, no. 5, pp. 2705–2718, Nov 1999.

[15] T. Ball, A. Schulze-Bonhage, A. Aertsen, and C. Mehring, "Differential representation of arm movement direction in relation to cortical anatomy and function." *J Neural Eng*, vol. 6, no. 1, p. 016006, Feb 2009.

[16] P. Shenoy, K. J. Miller, J. G. Ojemann, and R. P. N. Rao, "Generalized features for electrocorticographic BCIs." *IEEE Trans Biomed Eng*, vol. 55, no. 1, pp. 273–280, Jan 2008.

[17] R. M. Demirer, M. S. Ozerdem, and C. Bayrak, "Classification of imaginary movements in ECoG with a hybrid approach based on multi-dimensional Hilbert-SVM solution." *J Neurosci Methods*, vol. 178, no. 1, pp. 214–218, Mar 2009.

[18] T. Pistohl, T. Ball, A. Schulze-Bonhage, A. Aertsen, and C. Mehring, "Prediction of arm movement trajectories from ECoG-recordings in humans." *J Neurosci Methods*, vol. 167, no. 1, pp. 105–114, Jan 2008.

[19] Z. Wang, A. Gunduz, P. Brunner, A. L. Ritaccio, Q. Ji, and G. Schalk, "Decoding onset and direction of movements using Electrocorticographic (ECoG) signals in humans." *Front Neuroeng*, vol. 5, p. 15, 2012.

[20] E. Lew, R. Chavarriaga, H. Zhang, M. Seeck, and J. d. R. Millan, "Self-paced movement intention detection from human brain signals: Invasive and non-invasive EEG." *Conf Proc IEEE Eng Med Biol Soc*, vol. 2012, pp. 3280–3283, 2012.

[21] T. Milekovic, J. Fischer, T. Pistohl, J. Ruescher, A. Schulze-Bonhage, A. Aertsen, J. Rickert, T. Ball, and C. Mehring, "An online brain-machine interface using decoding of movement direction from the human electrocorticogram." *J Neural Eng*, vol. 9, no. 4, p. 046003, Aug 2012.

[22] G. Schalk, J. Kubánek, K. J. Miller, N. R. Anderson, E. C. Leuthardt, J. G. Ojemann, D. Limbrick, D. Moran, L. A. Gerhardt, and J. R. Wolpaw, "Decoding two-dimensional movement trajectories using electrocorticographic signals in humans." *J Neural Eng*, vol. 4, no. 3, pp. 264–275, Sep 2007.

[23] J. W. Rieger, C. Reichert, K. R. Gegenfurtner, T. Noesselt, C. Braun, H.-J. Heinze, R. Kruse, and H. Hinrichs, "Predicting the recognition of natural scenes from single trial MEG recordings of brain activity." *Neuroimage*, vol. 42, no. 3, pp. 1056–1068, Sep 2008.

[24] F. Quandt, C. Reichert, H. Hinrichs, H. J. Heinze, R. T. Knight, and J. W. Rieger, "Single trial discrimination of individual finger movements on one hand: A combined MEG and EEG study." *Neuroimage*, vol. 59, no. 4, pp. 3316–3324, Feb 2012.

[25] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proceedings of the IEEE*, vol. 70, no. 9, pp. 1055–1096, 1982.

[26] T. Hastie, R. Tibshirani, J. Friedman, and J. Franklin, "The elements of statistical learning: data mining, inference and prediction," *The Mathematical Intelligencer*, vol. 27, no. 2, pp. 83–85, 2005.

[27] J. R. Wolpaw, N. Birbaumer, W. J. Heetderks, D. J. McFarland, P. H. Peckham, G. Schalk, E. Donchin, L. A. Quatrano, C. J. Robinson, and T. M. Vaughan, "Brain-computer interface technology: a review of the first international meeting." *IEEE Trans Rehabil Eng*, vol. 8, no. 2, pp. 164–173, Jun 2000.

[28] M. S. Treder and B. Blankertz, "(C)overt attention and visual speller design in an ERP-based brain-computer interface." *Behav Brain Funct*, vol. 6, p. 28, 2010.

[29] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, "An efficient P300-based brain-computer interface for disabled subjects." *J Neurosci Methods*, vol. 167, no. 1, pp. 115–125, Jan 2008.

Biometric Keystroke Authentication on Smartphones Using a Capacitive Display

Matthias Trojahn
Volkswagen AG
Wolfsburg, Germany
Email: matthias.trojahn@volkswagen.de
Supervisor:
Jun.-Prof. Dr. Frank Ortmeier
Otto-von-Guericke University
Magdeburg, Germany

***Abstract*—Security of mobile devices is becoming an increasingly important topic in recent years. Different authentication solutions to secure the devices exist using biometric methods. This paper presents the work of a PhD project for authentication by exploiting keystroke dynamics on a device with a capacitive display. The major objective is to use the keystroke methods on a touchscreen and to improve the error rates compared to studies with a normal hardware keyboard (old mobile phones with 12 keys or computers). For this, new data of the capacitive and inertial sensors of smartphones is extracted and evaluated. Furthermore, different scenarios are tested to evaluate whether this method is suitable for daily usage. Effects of various devices are analysed and evaluated just as learning effects of this method. The results for the different studies show that a keystroke authentication can be used on capacitive displays with good error rates. This method improves the security of mobile devices by doing a two-factor authenticatiom (password and keystroke dynamics).**

Keywords*—smartphone security; keystroke authentication; capacitive display; inertial sensor; continuous authentication; context independence

I. Introduction

So far, more than one billion smartphones have ever been sold [1]. They are so small that they fit in each pocket and can be carried around everywhere. Thus, they are considered as a user-friendly device compared to normal computer regarding their size. However, these devices can easily be lost or stolen. In cabs in London alone, more than 55,000 mobile devices were lost in half a year [2].

Due to the number of devices and their flexibility they are becoming a security problem [3]: A lot of devices store crucial data or are able to access sensitive information via apps installed on the device.

At the same time, if an authentication method is used to restrict the access of the system in most cases this is only realised by passwords or *personal identification numbers* (PINs). But these are not sufficient. Social engineering (negotiation technique to aquire passwords) or shoulder surfing (looking on the device during typing of a different user) are only two methods to retrieve the password [4]. This means additional security mechanisms are required.

Keystroke authentication can be used during typing the password. It extracts the rhythm of the typing behaviour of one user which represents a biometric authentication method. For this rhythm, special features are extracted and used for authentication.

Keystroke authentication is not known very well to users. Face recognition is one of the most accepted biometric authentication methods which works on smartphones [5].

The goal of this work is to show that keystroke dynamics on mobile devices with a capacitive display can be efficiently applied for authentication.

In this paper, the background for keystroke dynamics will be explained in Section II together with the related work. Section III explains the existing challenges for keystroke authentication. The challenges will be solved with the concept for a keystroke authentication described in Section IV. Section V will focus on the construction of the realised case studies to analyse individual challenges. Seven different studies are presented. These are used to generate results using different algorithms. The results show, for example, how to reduce the learning effect or how to do a device comprehensive keystroke authentication where enrolment and verification is not on the same device. Then, the open points will be explained. The last section will conclude this paper and will present some information about the future work.

II. Background and Related Research

Typical biometric authentication methods contain two phases (enrolment and verification/identification) (see [6]). Figure 1 represents both phases with the included steps. Data acquisition, pre-processing and feature extraction are part of each phase. Extracted features during enrolment are stored in the database. In contrast, features extracted during verification or identification are used for classification or comparison.

Existing research studies used different approaches during this authentication process. These will be described in the following.

The type of data acquisition depends on the used sensors. Current keystroke publications use the keyboard of a computer

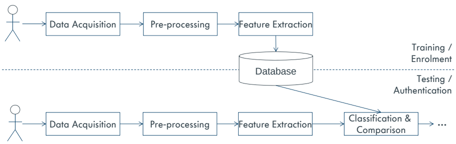


Fig. 1. Two steps and their phases during the authentication process (adapted from [7])

or of a 12-key mobile phone. With these hardware keys it can be recorded which key was pressed. Furthermore, the time for pressing and releasing can be extracted.

In combination with the pre-processing, features will be extracted. First, the duration is used which is the time between pressing and releasing a key. Second, digraph describes the time between pressing one and the next key. This can be generalised to *n* keys (n-graph). But normally only digraphs or trigraphs (time difference between three keys) were extracted [8]. If a specific password is required for authentication (instead of just the correct timing for arbitrary letter combinations), wrong key presses may occur and have to be made undone. Another feature can then be exported which represents how often a wrong key was pressed.

Classification can be done with different existing classifiers (e.g. *support vector machines* (SVM), neuronal networks or statistical classifiers). Together with the previously described features it can be shown how good an authentication system is. Three different error rates can be calculated: In general, the *false acceptance rate* (FAR - which represents the chance for an intruder to access the system) and the *false rejection rate* (FRR - the chance of an authentic user to be rejected) are used to compare different studies. The third rate is the *equal error rate* (EER) which represents the point where FAR is equal to FRR. The EER is also used in some research paper to compare their results.

The survey by Banerjee and Woodard (2012) presents an overview about the different studies in the field of keystroke authentication [9]. The survey shows that most studies were done with less than 50 subjects.

First keystroke dynamic studies arose in the desktop computer field between 1985 and 1990 ([10], [11]). They both used only the digraph as a feature and a statistical classifier. In this time, the error rates were around 10 %. Over the years they were improved by selecting more features and using other classifiers (e.g. neuronal networks). An EER of 3.8 % was reached by Napier et al. [12].

In addition to the text-dependent keystroke authentication, Monrose and Rubin [13] were the first who did a text-independent keystroke identification (successful in over 90 %). Today, keystroke dynamic on computer keyboards is already well known in the industry, like Psylock [14] or ID Control [15], too.

The use of keystrokes on mobile phones started in the last years. In 2006, Clarke and Furnell [16] adapted keystroke authentication on the mobile 12-key layout of the phones. In addition to duration and digraph, they used the error rates during typing. These features were extracted on a hardware keyboard. The generated EER in this study was between 13.3 % and 17 % for 32 subjects for duration and digraphs, respectively.

Two years later, Buchoux et al. [17] showed that a four-digit long numerical password is too short to authenticate a person. Then, several studies focused on the improvement of error rates ([18], [19]).

A few approaches exist where a touchscreen is used for authentication. For example, De Luca et al. [20] presented an authentication using a 3x3 matrix. Here, a pattern had to be drawn by connecting points. In this experiment, in addition to the time, the features pressure, size and exact coordinates were used. They reached an accuracy rate of 77 % and an EER of around 20 %.

Besides the multiple studies, Banerjee and Woodard created various public databases storing keystroke behaviour. One disadvantage with behaviour based on biometric authentication is that in most cases a new study has to be done to analyse new features or scenarios. Furthermore, these databases do not contain data information about size or X- and Y-coordinates which can be exported with the capacitive display. Due to these disadvantages, new studies had to be conducted to solve the challenges described in the next section.

III. Research Challenges

With the growing market of smartphones the focus is changing to devices with capacitive displays. Touchscreens do not have physical keys and thus no physical feedback (e.g. cannot feel whether two keys are pressed at the same time). Only vibrations can be used as kind of physical feedback. Moreover, the size of the keys is smaller using a QWERTY keyboard layout in comparison to the 12-key layout.

The main limitation of the existing studies are that they only use hardware keyboards and no capacitive displays. Furthermore, the most known studies for keystroke dynamics are using a low number of different subjects in an experiment, especially, for mobile devices. The average number of subjects is between 15 and 35 people [9]. This is not enough to see the scalability of this authentication model. In addition, some studies have error rates which are higher than 10 % for FAR and FRR. In this case, every tenth person who knows the password is able to unlock the system (security) and only in nine of ten authentication attempts the real user gets accepted (usability). Both is not sufficient.

Another limitation is that the studies focus only on an idealized setting (typing during sitting). But this is not the usual case of usage for mobile devices. We can distinguish between four main groups with different scenarios. The first group describes scenarios that focus the change of the physical condition of a person (e.g. movements like walking, tiredness or stress). The next group is defined by the influence of the environment. This group contains, for example, mobile movements (bus or train), temperature and light. The third group describes the physical condition of the device like the size or alignment of the device. The last one contains, for example, which hand is used and how many fingers.

The main objective of this PhD thesis is to generate a keystroke authentication system for smartphones on touchscreen displays and analyse how it can be used in practice (FRR smaller than 2 %). From the previously described problems different challenges and questions can be derived which have to be solved for a keystroke authentication system on touchscreen displays:

a) New sensors: In general, it should be checked if it is possible to authenticate via keystroke on a touchscreen device. Furthermore, the influence of smaller keys on the error rates has to be analysed. With higher error rates this method cannot be used for an initial authentication (authentication if the device is locked) on these devices.

b) Quality of new features: Are pressure, size, exact coordinates and values of the inertial sensors (gyroscope and accelerometer) improving the authentication rate? The single features and feature groups have to be evaluated to show which features should be added to an authentication system. This will improve the error rates for a keystroke authentication.

c) Comparison of different classifiers: A good classifier for keystroke dynamics is able to differ between intra-class differences of one user and inter-class differences between multiple users. Previous studies use different classifiers, but they cannot be compared. They are thus unsuitable for evaluating classifiers. It has to be evaluated which classifier should be used for keystroke authentication. Is it possible to use them also for the new features? The different classifiers have to be tested in combination with the features and multiple fusion strategies.

d) Improvement of FAR and FRR: Reliability of the system (FAR) and the usability (FRR) should be decreased to under 2 % if the person is focusing on the typing and does not do anything else. In this case the keystroke dynamics could be used as an initial authentication system.

e) Influence of context changes: Like other biometric methods keystroke authentication cannot be used in every situation [21]. Scenarios are, for example, walking, sitting or running during typing. For different scenarios it has to be evaluated which influence they have on our typing behavior. Previous studies focus on the normal keystroke authentication scenario where all participants are sitting and typing the password. But this is not the general situation during authentication. Furthermore, it has to be shown which effect the changes have on error rates and if it is possible to recognize the scenario changes. Moreover, if the scenario is recognised, it is possible to adapt the authentication model for this scenario.

f) Device independent and comprehensive authentication: It has to be analysed which influence different devices have on the error rates during keystroke authentication. Moreover, it has to be examined whether enrolment and verification can be done on separate devices.

g) Learning effect: It has to be analysed which influence the effect of learning has on the authentication. If a person is training to type on a device, the behaviour is changing so the following questions have to be solved. Is an authentication possible without changing the enrolment model? Can methods be created to reduce this effect?

h) Re-Authentication: A system that authenticates the person not only during initial authentication (re-authentication system) could improve the security. Continuously checking of the users credentials with keystroke dynamics is possible and should be evaluated. Therefore, a text-independent keystroke authentication has to be analysed on the capacitive display. It has to be tested which features can be used for authentication. Moreover, are other authentication methods possible?

IV. Design of a Concept for Biometric Authentication Using Keystroke

In this section the concept and methods of keystroke authentication will be explained. This will be based on the traditional process for biometric authentication (see Section II). The concepts to reduce the learning effect and to generate the re-authentication model are not complete yet.

A. Data Acquisition

New sensors in smartphones are generating additional data. Useful sensors for the keystroke authentication are the capacitive display and the inertial sensors (gyroscope and accelerometer).

With the capacitive display more than the time and which key was pressed can be extracted. The X- and Y-coordinates where the screen was touched as well as the physical pressure on the device and the fingertip size can be extracted. All the data can be extracted during pressing and releasing the finger on the screen. In addition, new actions could be recorded like movements with the finger on the screen or multi-touch events with more than one finger.

The other sensors, accelerometer and gyroscope, are extracting the movements of the device during typing. While the accelerometer measures accelerations, the gyroscope is measuring changes of the orientation. Both have to be recorded regularly in a predefined interval or in combination with the timestamps of the events of the capacitive display.

B. Pre-Processing

The new information of the sensors which has been exported requires new pre-processing steps. All the data is exported at the timestamp they occur. The focus using multi-touch events is to extract which action is referring to which key (key up, key down and key move events).

In addition, we do not want only to evaluate the clinical process (scenario *sitting* during authentication) so some transformations have to be done like device transformation and the context transformation.

1) Context analysis and transformation: Sitting, walking or using the non-dominant hand influence our typing behaviour. The impact of these scenarios has to be analysed and a detection of different scenarios has to be implemented to improve the error rates for these scenarios. During walking, the inertial sensors could be used to recognise if a person is moving. In addition, an algorithm to detect the dominant hand improves the error rates if different hands are used. This algorithm is based on the fact that pressure and size are distributed differently on the device using the left or the right hand. The study has to be designed that an enrolment is done in general during sitting. After this, the evaluation for each scenario is tested against the model during sitting. For this, multiple algorithms had to be designed to predict the model for other scenarios. For example, the time values (di- and trigraph) are bigger during walking because the person has to focus on walking, too.

2) Device transformation: Not only the context could change. There are some more device dependencies which have to be accounted for to enable device-independent authentication. The device is tracking a number of different values for each feature. In addition, the values are not normalized. Especially, this has been seen for the values of pressure and size.

The following steps have to be focused, if a device independency is requested:

- The time-based feature showed constant changes between the devices in a pre-test. Di- and trigraph changes can be explained by the different size of the devices.
- The X/Y-coordinates depend on resolution and dimension of the device. They could be transformed to the coordinate system of another model.
- The discretisation levels for pressure and size differ between models and have to be converted. The problem is that converting the values of a device with a lower quality to another with a higher quality increases the error rates.
- The gyroscope data showed no specific similarities which could be calculated. These have to be tested whether they can be used for different devices.

C. Feature Extraction

Section II describes the features which have been extracted during typing on a hardware keyboard. With the capacitive display and the inertial sensors more features could be extracted which are based on the data described in Subsection IV-A.

Values for pressure, size, X- and Y-coordinates (horizontal and vertical position on the screen) as well as the gyroscope (X, Y and Z) and the accelerometer (pitch and roll) have been exported during pressing and releasing one key. The accelerometer exports additional values (like azimuth) but these depend on the direction the user is facing in the moment of authentication (geographic direction). So they cannot be used for keystroke authentication because each time a person is standing facing a different direction.

Altogether the amount of features *nof(k)* which can be exported during authentication (here only the key press events) depends on the number of letters *k* in the password which can be seen in Equation 1.

$$nof(k) = \begin{cases} 9 \times k + (3 \times k - 3), & if\ (k > 2) \\ 9 \times k + (2 \times k - 1), & if\ (k = 2) \\ 10 \times k, & if\ (k = 1) \end{cases} \quad (1)$$

For a password with five letters 57 features could be extracted which are 45 more than for the traditional keyboards (duration, digraph and trigraph). If releasing events are captured as well, $9 \times k$ more features could be extracted. In general, more adequate features (these show good results if only they are used for classification) mean better error rates.

D. Classification & Comparison

To compare different classifiers a framework has to be generated. Basically, data mining tools like Weka [22] can be used for identifying or verifying a person. But for each different test case (e.g. selecting different features or different thresholds) new Weka files have to be created. The following aspects have to be considered for flexible abstract classification:

- selection of the classifier
- authentication mode: verification or identification (for mobile devices verification is sufficient, because normally only one user uses the device). During an identification the current sample is compared with each user model in the database. The best par will be used. During verification the current sample will only be compared with a preselected one and it will be decided whether they are equal enough.
- set of features
- weight for the features
- which fraction of recorded samples should be used for training and evaluation

V. Findings

The studies which are described in the first subsection are used to describe the current state of the multiple challenges from Section III.

A. Case Studies

In total, seven different case studies were conducted to answer the research questions. A keyboard layout for Android OS was developed for this to extract all the necessary information during typing.

In addition, individual applications were designed for each study. All studies started with the input of an anonymised identifier to identify persons who take part in multiple studies. Then, some descriptive data is requested to give an overview about the representativeness of each study. In addition, some knowledge and opinion to some keystroke related observations (e.g. if they have problems to type with touchscreen displays) are asked to answer. In each study some person did not have previous experience with touchscreens. That is why in the beginning the person had the possibility to train. After this, the different studies started which can be distinguished with the following facts:

1) The first study investigated the usage of the new display type (touchscreen), in addition, to the new features of the capacitive display ([23], [24]).
2) A study was done to show the differences between the 12-keys layout and the QWERTY-layout. Here, numerical and alphabetic passwords were distinguished [25].

3) A big study with more than 150 subjects was done to show the influence of stress during typing ([26], supervised master thesis [27]). In addition, the scalability was tested due to different numbers of subjects. This is also the first study to investigate gyroscope data as an authentication feature.
4) Another input form (Swype) where words are formed by sliding the finger on the screen to connect letters was used (pre-study with 16 [28] – supervised bachelor thesis – and the actual study with 42 subjects). After a training, the subjects had to swype five preselected words 10 times in a row.
5) In a pre-study, 65 different devices (in total 91 devices) were analysed to see how the data is influenced by the different sensors of the devices. Three main groups exist which differ in the number of different pressure values, as well as in screen size. For each group a representative was selected for the main study. The 80 subjects had to enter three different passwords 20 times with the same length on three smartphones and one tablet.
6) The influence of five different scenarios was investigated. The subjects (80 in total) had to stand, to walk, to sit, use their non-dominant hand, and listen to music while they are typing multiple passwords 20 times. At the same time the importance of the length of passwords was analysed.
7) To test the influence of learning a study was started with 400 repetitions of one word for each subject.

For all studies the obtained data from the capacitive display are the time values (duration, di- and trigraph), pressure, size and the exact coordinates. The inertial sensor data is data of the gyroscope and the accelerometer. Different devices were selected and used (e.g. Galaxy Nexus, Samsung S2 and S3, HTC Desire). Even though they are from the same model range, the sensors are different. For passwords different German words (e.g. *anna*, *sommer* or *donnerwetter*) were selected to focus in the studies only on the typing behaviour.

B. Results for Time Features on a Capacitive Display

In one of the first studies (passphrase with 17 letters) it was analysed which error rates could be obtained by using the old features (duration, digraph and trigraph). Table I shows the results for each feature group and the fusion of all three.

TABLE I. RESULTS FOR DIFFERENT FEATURE COMBINATIONS

FAR (in %)	FRR (in %)	Selection of features
8.03	12.3	residence time
12.66	11.64	digraph (two different keys)
13.63	33.33	trigraph (three different keys)
6.61	8.03	residence time + digraph + trigraph

It can be seen that with a passphrase of 17 letters good error rates have been generated if the basic features are fused. These error rates are better than the proposed ones by Clarke and Furnell [16] presented in Section II. This means that capacitive displays can be used as well as for keystroke authentication.

C. Quality of New Features and Device Dependencies

As already stated the quality of features regarding the error rates depends on the sensor used in the smartphone. In Table II the error rates for all evaluated features are presented in comparison to the basic features. In this study, passwords consisting of six characters were used.

TABLE II. DIFFERENT ERRS (IN %) FOR INDIVIDUAL FEATURE GROUPS ON DIFFERENT DEVICES FOR DIFFERENT PASSWORDS

		Nexus	S2	S3
duration	treter	20.91	18.81	20.04
	module	19.64	19.16	19.05
	sommer	18.21	19.46	18.57
digraph	treter	23.02	21.73	18.17
	module	17.77	17.02	14.47
	sommer	14.59	15.83	14.67
trigraph	treter	24.32	22.95	21.18
	module	20.43	19.44	16.27
	sommer	16.1	18.28	15.67
pressure	treter	18.84	39.65	50
	module	17.76	30.85	50
	sommer	15.52	26.86	50
size	treter	29.34	17.44	24.75
	module	28.62	16.06	24.31
	sommer	26.26	15.97	22.44
XY	treter	25.77	24.94	23.2
	module	27.22	26.95	22.86
	sommer	20.89	20.67	22.41
gyroscope	treter	36.27	37.39	37.1
	module	38.85	39.76	35.17
	sommer	36.94	41.02	37.46
accelerometer	treter	31.49	19.77	22.56
	module	32.91	19.93	25.53
	sommer	36.28	18.89	21.38

The first fact which could be seen is that the error rates even for the basic features are worse than seen in Table I. That can be explained by the length of the password because shorter passwords are generating less features which can be used for classification. The EER for the new features which were extracted by the capacity display have nearly the same accuracy for authentication. The inertial sensors have not that good error rates (most have EERs bigger than 30 %). But between the devices differences can be seen, especially, for the feature *pressure*. An EER of 50 % is not useful because it has the same quality as guessing. The reason for bad results is that on the S3, *pressure* is only a binary measure (pressure/no pressure).

Table II shows that some feature groups are more suitable for authentication than others. This means a weighted fusion should be done to improve the error rates. But for different devices best results are obtained with different configurations.

D. Comparison of Different Classifiers

As part of a supervised bachelor thesis different classifiers (support vector machine, neuronal network and statistical) were compared [29]. The computation time for training and evaluation for SVM and neuronal network is too high to generate and evaluate the model on the device. Only the statistical classifier could be used completely on the mobile device. SVM and neuronal network can only be evaluated on the device.

For neuronal networks and SVM the models have to be generated on a server. If the system is working in real time, it has to be calculated how many distributed servers are required to compute the model in less than one minute. It is time and computational power consuming to do it directly on the device for most classification algorithms. This can be done for example by using a Markov model [30].

In the study different error rates could recorded using different number of subjects. The neuronal network was more suitable for datasets containing less than ten subjects (EER < 1 %). For bigger datasets with 152 subjects the statistical classifier showed better results.

An important fact is that the error rates are depending on the size of the dataset. The reason is that typing behaviour is not unique [31]. But it provides enough information to confirm the identity of a person ([31], [32]). As a result of this experiment the statistical classifier should be used because of the better results and the better performance.

E. Improvement of FAR and FRR

If the old (duration, di- and trigraph) and new features are combined, better results can be reached. The FAR of 6.61 % and the FRR of 8.03 % with the basic features have been improved to a FAR of 1.3 % and FRR of 0.79 % (with all features). For this, the word *donnerwetter* in study 6 was analysed during sitting. The 20 repetitions were splitted into 13 training and seven evaluation datasets and were classified using a k-Nearest-Neighbor classifier. In addition, the statistical classifier used weights based on the quality of the feature groups (see Table II). Compared to the existing error rates on mobile devices with hardware keys (see Section II) EERs of less than 10 % could be reached which are shown in the next two subsections.

F. Influence of Context Changes

These results are based on the supervised bachelor thesis [33]. Five different scenarios were analysed (standing, walking, sitting, using the non-dominant hand and during listening to music). Table III shows the error rates when enrolment and authentication are done in the same situation.

TABLE III. COMPARISON OF THE ERROR RATES (IN %) WITH AN ENROLMENT FOR EACH SCENARIO. THE BIGGEST ERROR RATES ARE MARKED WITH A DARK RED; THE BEST ERROR RATES ARE DARK GREEN

scenario	**anna**		**sommer**		**donnerwetter**	
	FAR	FRR	FAR	FRR	FAR	FRR
standing	2.77	1.56	1.64	0.95	0.94	1.12
walking	4.84	6.69	4.01	3.37	1.72	2.31
sitting	2.20	2.84	1.38	0.63	0.69	1.27
hand	3.56	4.57	3.18	2.77	2.98	1.79
music	3.60	1.42	0.89	0.47	0.86	0.95

With more letters in the password the average of the error rates improve. The best error rates have been achieved if the subject is sitting and using the dominant hand (with or without music). The highest error rates are caused by walking or using the non-dominant hand, which can be explained by high intra-person differences (differences between multiple attempts by one person).

In addition, it has been compared how the error rates are if the model is generated during sitting and used during other scenarios. This can be seen in Table IV.

The error rates during sitting are the lowest and equal to Table III. The reason is that in both cases the same training and evaluation data are used. In all other cases, error rates are much higher. Even in the best case FAR and FRR are six percentage points. Especially with the non-dominant hand, no keystroke authentication can be done because the error rates are too high for verification. The typing behaviour is different between the two scenarios. Without a context transformation the keystroke can only be used if the same scenario is already known and has been recognised.

TABLE IV. COMPARISON OF THE ERROR RATES (IN %) IF AN ENROLMENT IS DONE DURING SITTING AND USED FOR OTHER SCENARIOS.

scenario	**anna**		**sommer**		**donnerwetter**	
	FAR	FRR	FAR	FRR	FAR	FRR
standing	13.13	15.46	10.72	9.32	7.43	10.96
walking	16.55	19.71	12.32	12.76	12.19	8.87
sitting	2.20	2.84	1.38	0.63	0.69	1.27
hand	19.72	41.09	21.18	41.68	21.42	38.20
music	8.50	15.87	6.49	8.07	6.19	7.30

G. Device Transformation

Table V shows the results of the study 5 if enrolment and verfication is done on the same device.

TABLE V. TOTAL ERROR RATES (IN %) VARIOUS PASSWORDS AND DEVICES. THE MOST SUITABLE COMBINATION OF PASSWORD AND DEVICE IS MARKED BOLD.

	Nexus		S2		S3	
	FAR	FRR	FAR	FRR	FAR	FRR
treter	5.72	5.81	6.09	5.43	6.76	10.19
module	**4.09**	**3.12**	4.9	5.19	5.60	8.87
sommer	3.27	4.97	5.87	4.65	7.69	5.42

It can be seen that keystroke authentication is device dependent. The Galaxy Nexus showed the best results for all passwords using the k-Nearest-Neighbor classifier.

If the same model is used on multiple devices, the error rates increase to a value between 27 % and 36 %.

If, on the other hand, the algorithm described in Section IV-B2 is used for device transformation, the error rates improve. This can be seen in Table VI.

The table shows that converting from or to the Samsung S2 shows the highest error rates in our study. The error rates are higher than 10 % on average. This means an improvement compared to the results if no converting was used. With a more complex password the error rates are better. The word *treter* represents only a small area of the keyboard. In almost every test case the error rates of the words *module* and *sommer* show better results. Furthermore, converting between Samsung S3 and the Galaxy Nexus showed the best results. This can be explained by nearly the same display size and the same dpi values (308 and 315). The feeling for the user is nearly equal between both devices (e.g. weight or size).

VI. OPEN ISSUES

Two of the challenges are not completely solved at this point. These points and what is planned will be described in this section.

A. Learning Effect

The existing challenge of the learning effect can be solved by implementing an updateable model. It depends on the previous experience how steep the graph (learning curve) is. For inexperienced persons the model has to be updated on a regular basis. In a history the changes need to be saved to use this data for a better authentication.

TABLE VI. ERROR RATES (IN %) IF DIFFERENT DEVICES ARE USED FOR ENROLMENT AND VERIFICATION

enrolment	**Nexus**				**S2**				**S3**			
verification	S2		S3		Nexus		S3		Nexus		S2	
	FAR	FRR	FAR	FRR	FAR	FRR	FAR	FRR	FAR	FRR	FAR	FRR
treter	12.2	18.2	10.4	8.6	12.7	12.4	13.6	13.7	7.8	5.4	12.9	10.7
module	12.8	11.8	7.6	6.1	12.8	9.6	13.2	10.7	6.0	5.9	13.3	10.2
sommer	13.4	12.2	7.6	6.2	12.6	9.1	13.3	9.6	7.9	3.7	14.1	8.0

The acquisition of the data to evaluate the learning effect is not yet finished. Further subjects have to take part in this study. After the study the data have to be evaluated with respect to how much the typing behaviour changes and which impact it has for authentication. Based on this information and results from the already completed studies decisions about possible update circles for each feature group can be made.

B. Re-Authentication

The first step is to generate a text-independent keystroke authentication, i.e. an authentication system based exclusively on the typing behavior irrespective of the words actually typed [34]. In addition, gait recognition (walking behaviour) methods for smartphones have to be adapted by other researcher [35]. The gait recognition could be used, especially, if no interaction is done with the device. Both authentication forms have been fused and used to determine the trust of device if it is still the same person as before. The basis is a defined trust model which calculates the certainty in a time box and overall. If the trust is too low, the device gets locked.

VII. SUMMARY

In this paper, we first identified several security problems with passwords on mobile devices. Other authentication methods than only passwords are required to increase the security. We proposed a system using keystroke behaviour as a biometric authentication method on smartphones with the capacitive display and inertial sensors.

We presented different research challenges to keystroke authentication. Different case studies were conducted to analyse these challenges. The new keyboard layout (QWERTY in combination with soft keys) and new features were considered and showed an improvement of the error rates (FAR of 1.3 % and FRR of 0.79 %). Moreover, a concept for an abstract classifier was generated to test different scenarios with multiple configurations. Other studies only used simplistic scenarios to analyse the keystroke behaviour. That is why the influence of different situation on the error rates was observed. Some situations are not applicable without big decreasing of the error rates (e.g. during walking or usage of the non-dominant hand). In addition, we focused on the device comprehensive and independent authentication. As a result, not all devices turned out to be as good as others. This depends, among others on the sensors in the device. Furthermore, learning effect and text-independent keystroke authentication were discussed.

Overall, the proposed techniques and studies showed an important step towards more effective and robust implementations of keystroke authentication systems. Using the presented results, an authentication system could be implemented and used which improves the security and is user-friendly at the same time. It was shown that good results depend on the right devices and that they have to be tested before using them. Moreover, like other biometric authentication methods some limitations exist regarding their usability in each scenario. It is impossible to generate templates for all scenarios, especially, for those which cannot be recorded (e.g. stress). In these situations, the thresholds have to be adjusted regarding the focus of usability or security. This could mean if the focus is on security that the user can only access the system in a special situation.

However, further work is planned in order to evaluate the effect of learning and how this effect can be reduced. Furthermore, in the next stage it is planned to convert the text-independent authentication to a re-authentication system which analyses the person during usage of the device. This would increase the security level for the devices because intruders can be found at several points during usage.

REFERENCES

[1] S. Bicheno, "Global smartphone installed base by operating system for 88 countries: 2007 to 2017," 09.10.2012.

[2] J. Twentyman, "Lost smartphones pose significant corporate risk," 2009. [Online]. Available: http://www.scmagazineuk.com/lost-smartphones-pose-significant-corporate-risk/article/126759/

[3] M. Trojahn and F. Ortmeier, "Designing an enterprise security strategy for mobile intranet access," in *International Workshop on Trustworthy Computing*, 2013.

[4] L. O'Gorman, "Comparing passwords, tokens, and biometrics for user authentication," *Proceedings of the IEEE - Engineering our future*, vol. 91, no. 12, pp. 2019–2020, 2003.

[5] A. K. Jain, L. Hong, and S. Pankanti, "Biometric identification," *Communications of the ACM*, vol. 43, no. 2, pp. 90–98, 2000.

[6] C. Vielhauer, *Biometric User Authentication for IT Security: From Fundamentals to Handwriting*, ser. Advances in information security. Springer-Verlag, 2006.

[7] M. Karnan, M. Akila, and N. Krishnaraj, "Biometric personal authentication using keystroke dynamics: A review," *Applied Soft Computing*, vol. 11, no. 2, pp. 1565–1573, 2011.

[8] L. F. Cranor and S. Garfinkel, *Security and Usability: Designing Secure Systems That People Can Use*, illustrated ed. O'Reilly Media, 2005.

[9] S. Banerjee and D. Woodard, "Biometric authentication and identification using keystroke dynamics: A survey," *Journal of Pattern Recognition Research*, vol. 7, pp. 116–139, 2012.

[10] D. Umphress and G. Williams, "Identity verification through keyboard characteristics," in *International Journal of Man-Machine Studies*, vol. 23, 1985, pp. 263–273.

[11] S. Bleha, C. Slivinsky, and B. Hussien, "Computer-access security systems using keystroke dynamics," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, 1990, pp. 1217–1222.

[12] R. Napier, W. Laverty, D. Mahar, R. Henderson, M. Hiron, and M. Wagner, "Keyboard user verification: toward an accurate, efficient, and ecologically valid algorithm," *International Journal of Human-Computer Studies*, vol. 43, no. 2, pp. 213–222, 1995.

[13] F. Monrose and A. D. Rubin, "Authentication via keystroke dynamics," in *Proceedings of the 4th ACM conference on Computer and communications security*, ser. CCS '97. New York, NY, USA: ACM, 1997, pp. 48–56.

[14] Psylock GmbH, "Psylock demo," 2009. [Online]. Available: http://www.psylock.com

[15] IDControl B.V., “Keystrokeid,” 2012. [Online]. Available: www.idcontrol.com

[16] N. L. Clarke and S. M. Furnell, “Authenticating mobile phone users using keystroke analysis,” in *International Journal of Information Security*, vol. 6. Berlin, Heidelberg: Springer-Verlag, 2006, pp. 1–14.

[17] A. Buchoux and N. L. Clarke, “Deployment of keystroke analysis on a smartphone,” in *Proceedings of the 6th Australian Information Security & Management Conference*, 2008.

[18] P. Campisi, E. Maiorana, and A. Neri, “User authentication using keystroke dynamics for cellular phones,” in *IET Signal Processing*, vol. 3, Issue: 4, 2009, pp. 333–341.

[19] S. Zahid, M. Shahzad, S. A. Khayam, and M. Farooq, “Keystroke-based user identification on smart phones,” in *Proceedings of the 12th International Symposium on Recent Advances in Intrusion Detection*, ser. RAID ’09. Berlin, Heidelberg: Springer-Verlag, 2009, pp. 224–243.

[20] A. d. Luca, A. Hang, F. Brudy, C. Lindner, and H. Hussmann, “Touch me once and i know it’s you!: implicit authentication based on touch screen patterns,” in *SIGCHI Conference on Human Factors in Computing Systems (CHI12)*, 2012, pp. 987–996.

[21] M. Trojahn and P. Marcus, “Towards coupling user and device locations using biometrical authentication on smartphones,” in *Internet Technology And Secured Transactions, 2012 International Conference For*, 2012, pp. 736–741.

[22] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, “The weka data mining software: An update,” in *SIGKDD Explorations*, 2009, vol. 11.

[23] M. Trojahn and F. Ortmeier, “Toward mobile authentication with keystroke dynamics on mobile phones and tablets,” in *7th International Symposium on Security and Multimodality in Pervasive Environment (SMPE)*, 2013.

[24] M. Trojahn and F. Ortmeier, “Keystroke dynamics authentication on mobile devices with a capacitive display,” in *Biometrics: Technologies, Systems and Applications - BTSA 2013*, 2013.

[25] M. Trojahn and F. Ortmeier, “Biometric authentication through a virtual keyboard for smartphones,” in *International Journal of Computer Science & Information Technology (IJCSIT)*, 2012.

[26] M. Trojahn, F. Arndt, M. Weinmann, and F. Ortmeier, “Emotion recognition through keystroke dynamics on touchscreen keyboards,” in *15th International Conference on Enterprise Information Systems*, 2013.

[27] F. Arndt, “Erkennung von Emotionen anhand des Tippverhaltens auf Touchscreen-Tastaturen,” Master’s thesis, TU Braunschweig, Braunschweig, 2012.

[28] M. Graumann, “Extracting biometrical features via swype,” Bachelor Thesis, Otto-von-Guericke University, Magdeburg, 2012.

[29] B. Behrendt, “Evaluierung von Klassifikatoren zur Authentifizierung mithilfe des Tippverhaltens,” Bachelor Thesis, Otto-von-Guericke University, Magdeburg, 2013.

[30] M. Trojahn, L. Pan, and F. Schmidt, “Developing a cloud computing based approach for forensic analysis using ocr,” in *7th International Conference on IT Security Incident Management & IT Forensics*, 2013.

[31] A. A. Ross, K. Nandakumar, and A. K. Jain, *Handbook of Multibiometrics (International Series on Biometrics)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc, 2006, vol. 6.

[32] H. Crawford, “Keystroke dynamics: Characteristics and opportunities,” in *Privacy Security and Trust (PST), 2010 Eighth Annual International Conference on*, 2010, pp. 205–212.

[33] T. Meier, “Ubiquitäre biometrische Authentifizierung mithilfe des Tippverhaltens,” Bachelor Thesis, Otto-von-Guericke University, Magdeburg, 2013.

[34] M. Trojahn and F. Ortmeier, “Re-authentication model for mobile devices,” *8. GI FG SIDAR Graduierten-Workshop über Reaktive Sicherheit*, p. 9, 2013.

[35] M. O. Derawi, C. Nickel, P. Bours, and C. Busch, “Unobtrusive user-authentication on mobile phones using biometric gait recognition,” in *Proceedings of the 2010 Sixth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, ser. IIH-MSP ’10. Washington, DC, USA: IEEE Computer Society, 2010, pp. 306–311.