

# Antrag zur Durchführung einer Projektgruppe

- 1 Thema** TabScript - Handschrifterkennung auf Android-basierten Tablets
- 2 Zeitraum** Wintersemester 2012/13 - Sommersemester 2013
- 3 Umfang** Je 8 SWS
- 4 Veranstalter** Dipl.-Inf. Leonard Rothacker und Prof. Dr.-Ing. Gernot A. Fink  
Informatik XII, Arbeitsgruppe *Mustererkennung in eingebetteten Systemen*  
{leonard.rothacker | gernot.fink}@tu-dortmund.de

## 5 Aufgabe

Die Aufgabe der Projektgruppe soll darin bestehen, eine Software zu entwickeln, mit der natürliche menschliche Handschrift zu einer möglichen Eingabemethode auf einem Android [1] basierten Tablet-Computer wird. Die tatsächliche Anwendung auf dem Tablet soll durch die Teilnehmer mitbestimmt werden. Die Veranstalter werden jedoch Vorschläge machen (Notiz-Applikation, Chat-Applikation, PDF-Annotation, Kalender/TODO-Applikation) und auf die Realisierbarkeit im Rahmen der Projektgruppe achten.

### 5.1 Motivation

Als konsequente Weiterentwicklung sogenannter Smartphones gewinnen Tablet-Computer mehr und mehr Marktanteile. Mit ihren berührungssensitiven Bildschirmen eignen sie sich im privaten Bereich vor allem als mobile Entertainment Geräte. Sie werden jedoch vielfach auch als elektronischer Notizblock, zum Beispiel in Besprechungen, eingesetzt. Im Gegensatz zu einem Laptop gestaltet sich die Interaktion dabei eher wie mit einem klassischen Notizblock. Das Gerät liegt vor dem Nutzer auf dem Tisch und verdeckt ihn nicht, so dass eine direktere Kommunikation mit anderen Gesprächsteilnehmern möglich wird. An dieser Stelle endet jedoch die Analogie zu dem Erstellen von Notizen mit Papier und Stift: Der Nutzer muss sich durch die hohe Zahl verschiedener Applikationen und Anwendung verstärkt konzentrieren. Eine wichtige Komponente dabei ist die Eingabe von Text. Tablet-Computer bieten in der Regel ein sogenanntes Soft-Keyboard, ein Ersatz für eine herkömmliche Tastatur. Die Tastatur wird dabei durch eine Software Applikation auf dem Bildschirm dargestellt. Ein wesentlicher Nachteil ist, dass so die Tastatur den Bildschirm mindestens teilweise verdeckt. Je weniger der Bildschirm verdeckt wird, desto kleiner müssen aber auch die einzelnen Tasten dargestellt werden, was die Eingabe zunehmend erschwert. An dieser Stelle soll die Projektgruppe ansetzen. Durch die Entwicklung und Integration einer Applikation zur Handschrifterkennung soll die Interaktion mit dem Tablet intuitiver werden. Nutzer sollen so in die Lage versetzt werden, ihre Eingabe direkt in die Applikation zu schreiben, also nicht mehr an ein Soft-Keyboard gebunden sein. Die Eingabe von Text wird damit schneller, einfacher und ist ohne den Wechsel von Fenstern in der Applikation möglich. Die Anwendung ist zudem höchst aktuell. Nur wenige kommerzielle Applikationen existieren, wie zum Beispiel die Microsoft Handschrifterkennung [6] oder die in das Betriebssystem integrierte Handschrifterkennung des Samsung Galaxy Note [5]. Insbesondere ist den Veranstaltern keine frei verfügbare Applikation bekannt, die die in der Projektgruppe angestrebten Funktionen umsetzt.

### 5.2 Zielsetzung

Die Aufgabe dieser Projektgruppe ist es, eine Applikation zur Handschrifterkennung auf Tablet-Computern zu entwickeln. In der Anwendung tätigt der Nutzer somit seine Eingabe, indem er Wörter und Sätze mit einem (kapazitiven) Spezialstift direkt auf den Bildschirm schreibt. Eine Einschränkung auf bestimmte Schriftstile,

wie zum Beispiel Druckschrift, soll nicht notwendig sein. Je nach Typ der entwickelten Applikation soll auch möglichst auf eine separate Schreibregion verzichtet werden. Vielmehr soll die Stelle, an der geschrieben wurde, sinnvoll für die Eingabe einbezogen werden. Zum Beispiel ist es sinnvoll, den erkannten Text an der Stelle einzufügen, an der er in einer Notiz-Applikation geschrieben wurde. Der Prozess der Handschrifterkennung auf dem Tablet-Computer gliedert sich dann grob in folgende Abschnitte (vgl. [3, 7]), die auch für die thematische Einteilung der Projektgruppe sinnvoll erscheinen:

- **Schriftmodell:** Da die Erkennung auf Basis der Trajektorie eines Schriftzugs erfolgt, kommen Verfahren aus dem Bereich der *Online* Handschrifterkennung zum Einsatz. Zunächst wird die Trajektorie gefiltert, um die für die Erkennung irrelevante Information zu entfernen (Vorverarbeitung). Dann werden die Informationen extrahiert, die für die Erkennung bedeutungsunterscheidend sind (Merkmalsberechnung). In einem statistischen Modell erfolgt abschließend die Generierung von verschiedenen Hypothesen für die Transkription des Schriftzugs (vgl. [3, 7]). Für das statistische Schätzen der Modellparameter steht eine Stichprobe von Online Handschriftdaten zur Verfügung [4].
- **Sprachmodell:** Ausgabe des Schriftmodells sind verschiedene Hypothesen für einzelne Buchstaben oder Wörter. An dieser Stelle ist es möglich den Kontext der Eingabe einzubeziehen und die verschiedenen Hypothesen neu zu gewichten. Sprachmodelle spiegeln dazu im Allgemeinen die Wahrscheinlichkeit bestimmter Buchstaben oder Wortsequenzen wider (vgl. [7]). Eine zentrale Frage ist wie diese Wahrscheinlichkeiten anwendungsbezogen geschätzt werden können. In der Projektgruppe sollen hierbei Methoden des Web-Mining einen Schwerpunkt darstellen. Dabei werden Wahrscheinlichkeiten dieser Sequenzen mit Hilfe von dynamisch aus dem Internet abgerufenen Texten geschätzt [8]. Eine zusätzliche Möglichkeit besteht in der Einbindung der auf dem Tablet-Computer gespeicherten Wörterbücher, auf die über entsprechende APIs zugegriffen werden kann [1]. Diese beinhalten insbesondere auch die speziell durch den Nutzer verwendeten Wörter.
- **Anwendungsintegration:** Die Handschrifterkennung muss so umgesetzt werden, dass eine sinnvolle Anwendung auf dem Tablet-Computer resultiert. Die wahrscheinlichsten Erkennungsergebnisse sollen dem Nutzer angezeigt werden. Die Herausforderungen in der Projektgruppe liegen insbesondere in dem Verständnis und richtigen Einsatz von Hilfsbibliotheken. Diese dienen dazu bestimmte Informationen abzufragen (z.B. ermitteln der Trajektorie des Schriftzugs) und das Erkennungsergebnis im Kontext der Anwendung zu verwenden. Je nach Vorlieben der Projektgruppe und Möglichkeiten der eingesetzten Bibliotheken kann dies sowohl die Modifikation bestehender als auch das Erstellen neuer Applikationen betreffen. Einige Vorschläge zu möglichen Anwendungen finden sich weiter unten.

Da für die Anwendungsintegration die Handschrifterkennung direkt auf dem Tablet-Computer implementiert werden muss, ist die effiziente Umsetzung geeigneter Algorithmen unerlässlich. Andernfalls ist keine reaktive Interaktion mit dem Nutzer möglich. Die gegebenen Hardware- und Software-Voraussetzungen auf dem Tablet machen diese Integration somit zu einer der wichtigsten Herausforderungen der Projektgruppe. Das Erstellen der statistischen Schrift- und Sprachmodelle braucht, im Gegensatz dazu, aber nicht auf dem Tablet selbst zu erfolgen.

Zur Steigerung der Identifikation der Teilnehmer mit dem Projekt wird die konkret zu entwickelnde Anwendung nicht vorab festgelegt, sondern gemeinsam in der Diskussion zwischen Veranstaltern und Teilnehmern definiert. Folgende Beispiele dienen somit zur Orientierung und sollen die verschiedenen Aspekte der Eingabe durch Handschrift verdeutlichen:

- **Notizblock:** Der Tablet-Bildschirm zeigt einen linierten Notiz-Zettel auf den der Nutzer seine Eingabe direkt schreiben kann. Um die Eingabe besser zu lokalisieren bietet sich eine Zoom-Funktion an.
- **Chat:** Der Tablet-Bildschirm zeigt den Gesprächsverlauf zwischen einem oder mehreren Gesprächsteilnehmern. Zur Text-Eingabe steht ein Schreibfeld/-region zur Verfügung. Die in Chats üblichen *emotiv*en Äußerungen, zum Beispiel Smileys, sollen durch einfache Zeichnungen (Gesten) eingefügt werden können.

- (PDF-) Dokumenten-Annotation: In Besprechungen, Vorlesungen, etc. dienen den Teilnehmern vorliegende Dokumente häufig als Grundlage der Diskussion. In der Anwendung soll es möglich sein Anmerkungen direkt in das Dokument zu schreiben. Die Eingabe soll erkannt und an entsprechender Stelle als Annotation eingefügt werden.
- Kalender / TODO Anwendung: Das Erstellen neuer Einträge für den Kalender / TODO Liste soll durch Handschrifterkennung erfolgen. Durch die potentiell hohe Zahl wiederkehrender Ereignisse / Aufgaben, bietet sich die dynamische Erstellung eines Lexikons aufgrund der zuvor erfassten Daten an.

Die Arbeit der Projektgruppe umfasst die Einarbeitung in die nötigen Techniken und Werkzeuge, die methodische Konzeption der entsprechenden Funktionalitäten sowie deren prototypische Umsetzung in einem integrierten Softwaresystem. Erklärtes Ziel ist ein funktionsfähiger Demonstrator, gegebenenfalls in einem vereinfachten Szenario.

### 5.3 Grundlagen

Für die Projektgruppe stehen aktuelle Linux-PCs mit modernen Entwicklungswerkzeugen für die Softwareentwicklung bereit. Außerdem stehen zwei aktuelle Tablet-Computer der Firma Samsung sowie kapazitive Spezialstifte zur Verfügung. Für eine einfache und schnelle Evaluation von relevanten Methoden und Verfahren kann auch auf die *rapid prototyping* Umgebung MatLab zurückgegriffen werden. Darüber hinaus wird das Entwicklungs Tool-Kit ESMERALDA [2] durch den Veranstalter betreut. Es beinhaltet Implementierungen der gängigen Verfahren, die auch zur Handschrifterkennung verwendet werden. Das in Abschnitt 5.2 angesprochene Training der Modellparameter soll mit diesen bereits vorhandenen Implementierungen durchgeführt werden. Die Implementierung des Dekodierers kann bei Schwierigkeiten mit der Tablet-Umsetzung als Vorlage dienen. Das Tablet-Betriebssystem ist Android [1]. Daher müssen die Algorithmen auf dem Tablet in Java umgesetzt werden, wobei nur ein eingeschränktes SDK zur Verfügung steht.

### 5.4 Perspektiven

Die Projektgruppe bietet den Teilnehmern die Möglichkeit, praktische Erfahrungen in einem dynamischen Forschungsumfeld mit unmittelbarem Anwendungsbezug und hoher praktischer Relevanz zu sammeln. Inhaltliche Schwerpunkte dabei sind die folgenden Arbeitsgebiete, die durch die Gruppe bearbeitet werden können:

- Erarbeiten der Methodik zur Handschrifterkennung.
- Konzeption einer Mensch-Maschine-Schnittstelle für Handschrift.
- Verwendung eines mobilen Endgerätes und dessen Funktionalitäten in einem Softwareframework.
- Automatischer Zugriff auf verschiedene Web-Services als Informationsquelle.

Außer den o.g. fachspezifischen Schwerpunkten lassen sich

- Erfahrungen mit teamorientierter Softwareentwicklung (insb. Aufgaben- und Zeitplanung) erlangen,
- Erfahrungen mit dem Werkzeugeinsatz bei der Entwicklung komplexer Softwaresysteme sammeln.

## 6 Teilnahmevoraussetzungen

Folgende Kenntnisse sind für die Teilnahme an der Projektgruppe *notwendig*:

- Programmierkenntnisse (C/C++ oder Java)

Weiterhin sind Kenntnisse in folgenden Bereichen *hilfreich*:

- Mensch-Maschine-Interaktion, Eingebettete Systeme, Mustererkennung
- Erfahrung in der Softwareentwicklung unter UNIX/Linux und Android.

## 7 Minimalziele

1. Erstellen eines statistischen Schriftmodells mit Hilfe des ESMERALDA Tool-Kit und der IAM-On Datenbank [4].
2. Erstellen eines statistischen Sprachmodells mit Hilfe des ESMERALDA Tool-Kit und eines bezüglich der Anwendung geeigneten Lexikons.
3. Automatische Erstellung eines Lexikons durch geschicktes Abrufen der Informationen aus dem Internet und / oder entsprechend geeigneter Quellen.
4. Effizientes Umsetzen des Erkenners auf der Android-Plattform unter Java. Die Erkennungsergebnisse müssen dem Nutzer schnell und deutlich angezeigt werden.
5. Erarbeitung eines Konzepts für die Integration der Handschrifterkennung in eine Anwendung.
6. Umsetzung dieser Anwendung auf der Android-Plattform unter Java.
7. Vorführung eines Prototypen auf dem Tablet-Computer.

## 8 Literatur

- [1] Android Open Source Project. Homepage. <http://source.android.com/>, April 2012.
- [2] Gernot A. Fink and Thomas Plötz. Developing pattern recognition systems based on Markov models: The ESMERALDA framework. 18(2):207–215, 2008.
- [3] S. Jaeger, S. Manke, and A. Waibel. Npen++: An on-line handwriting recognition system. In *Proc. 7th Int. Workshop on Frontiers in Handwriting Recognition*, pages 249–260, Amsterdam, September 2000.
- [4] Markus Liwicki and Horst Bunke. Iam-ondb – An on-line English sentence database acquired from handwritten text on a whiteboard. In *Proc. Int. Conf. on Document Analysis and Recognition*, volume 2, pages 956–961, Seoul, Korea, 2005.
- [5] Samsung Galaxy Note. Technische Daten, April 2012. [http://de.samsung.com/de/microsites/galaxynote/eigenschaften\\_specs.aspx](http://de.samsung.com/de/microsites/galaxynote/eigenschaften_specs.aspx).
- [6] James A. Pittman. Handwriting recognition: Tablet PC text input. *IEEE Computer*, 40(9):49–54, 2007.
- [7] Thomas Plötz and Gernot A. Fink. *Markov Models for Handwriting Recognition*. SpringerBriefs in Computer Science. Springer, 2011.
- [8] Christoph Ringlstetter, Max Hadersbeck, Klaus U. Schulz, and Stoyan Mihov. Text correction using domain dependent bigram models from web crawls. In *Proceedings of the International Joint Conference on Artificial Intelligence*, Hyderabad, India, Januar 2007.

## 9 Rechtliche Hinweise

Die Ergebnisse der Projektarbeit, inklusive der dabei erstellten Software, sollen der Fakultät für Informatik uneingeschränkt zur freien Forschung zur Verfügung stehen. Darüber hinaus sind keine Einschränkungen der Verwertungsrechte und keine Vertraulichkeitsvereinbarungen vorgesehen.

## 10 Realisierung

### Phase 1 (Bis Ende 2012): Seminarbetrieb

Sammeln von Erfahrungen mit dem ESMERALDA Took-Kit und der Android-Plattform.

Einarbeitung in das Themengebiet und Sichtung der relevanten Literatur.

Vorstellung vielversprechender bekannter Ansätze zur Lösung der Teilprobleme in Vorträgen.

Erarbeiten der Konzeption für eine Anwendung auf dem Tablet-Computer.

Planung der Arbeitspakete und Arbeitsgruppen für die weiteren Entwicklungsarbeiten.

- Ergebnis:*
- Konzepte und Auswahl geeigneter Methoden zur Realisierung der Handschrifterkennung und deren Anwendung.
  - Statistische Modelle für Schrift und Sprache.
  - Aufnahme erster Trajektorien auf dem Tablet-Computer und Erkennung dieser mit (ausschließlich) den in ESMERALDA zur Verfügung stehenden Programmen.

### Phase 2 (Bis Ende WS 2012/13): Gruppenarbeit mit Diskussionen / Besprechungen im Team

Entwurf einer Softwarearchitektur für das Gesamtsystem und Realisierung eines Systemrahmens.

Realisierung eines Prototypen für die Handschrifterkennung auf dem Tablet-Computer.

Realisierung eines Prototypen für die dynamische Erstellung eines Lexikons (Sprachmodell).

Konzept für die Texteingabe durch Handschrift.

Realisierung eines Prototypen für die Anwendung auf dem Tablet-Computer.

Erstellung des Zwischenberichts.

- Ergebnis:*
- Definition der Softwarearchitektur und interner Datenrepräsentation.
  - Lauffähiger Handschrifterkennner in Java (Android kompatibel).
  - Dynamische Erstellung des anwendungsabhängigen Lexikons.
  - Evaluierung der Erkennungsleistung mit dem ESMERALDA Tool-Kit (Testdaten).
  - Applikation auf dem Tablet-Computer (ohne Handschrifterkennung).
  - Zwischenbericht.

### Meilenstein 1 (Ende WS 2012/13)

Einzelne Module wurden separat mit Erfolg getestet (Testumgebung).

### Phase 3 (Bis Ende Mai 2013) Gruppenarbeit mit Diskussionen/Besprechungen im Team

Individuelle Weiterentwicklung der Module mit regelmäßiger Integration (kurze Zyklen, z.B. wochenweise) zu einem vollständigen Prototypen des Gesamtsystems.

Synchronisation der Anwendungs-Daten (in der Cloud, z.B. Google Konto, etc.).

Planung des Abschlussberichts (Struktur, Zuständigkeiten etc.).

### Meilenstein 2 (Ende Mai 2013)

Vollständig lauffähiges Demonstrationssystem (gesamte Funktionalität realisiert).

### Phase 4 (Bis Ende SS 2013)

Evaluation des Systems mit realen "Benutzern". Hierfür ist die Durchführung einer Nutzerstudie angedacht. Es soll sowohl die Zuverlässigkeit der entwickelten Methoden anhand quantitativer Kriterien bewertet als auch subjektive Nutzerbewertungen ausgewertet werden.

Erstellung des Abschlussberichts.

## 11 Erweiterungsmöglichkeiten

Über die Minimalziele hinaus eröffnen sich einige Erweiterungsmöglichkeiten, die im Rahmen der Projektgruppe bearbeitet werden können.

Der wichtigste Aspekt bezieht sich auf die automatische Adaption der statistischen Modelle an bestimmte Nutzer. Wie im Antrag beschrieben, werden diese für Schrift und Sprache vorab erstellt und auf dem Tablet-Computer für die Erkennung verwendet. Solche Modelle für verschiedene Schreiber zu erstellen ist schwierig, da menschliche Handschrift stark variiert. Ein einzelner Tablet-Computer wird aber in der Regel nur von sehr wenigen Nutzern verwendet, so dass eine stärkere Adaption an die entsprechenden Schreibstile die Erkennungsleistung verbessern kann. In diesem Sinne wäre auch die Unterstützung von Profilen für die verschiedenen Nutzer eines Tablet-Computers wünschenswert. Um solch eine automatische Adaption zu realisieren, müssten die statistischen Modelle aktualisiert werden, sobald der Nutzer das Erkennungsergebnis manuell korrigiert.

Ein weiterer wesentlicher Aspekt ist die effiziente und ressourcenschonende Umsetzung der Handschrifterkennung auf dem Tablet-Computer. Nach der Implementierung der Standardverfahren ergeben sich fortgeschrittene Möglichkeiten zur Effizienzsteigerung. Im Allgemeinen zielen diese darauf ab den Suchraum für die Erkennungshypothesen geschickt einzuschränken. Letztendlich kann es auch sinnvoll sein Teile des Erkenners in der Sprache *C* zu programmieren. Mit der Android *native* API ist es dann möglich diesen Programmcode aus einer Java Applikation heraus zu verwenden. Dieses Auslagern von Programmteilen steigert die Effizienz insbesondere bei Operationen die verstärkt auf die CPU zugreifen. Solche Operationen werden bei der Erkennung von Handschrift oft benötigt.

Abschließend ergeben sich eine Fülle von Erweiterungsmöglichkeiten bei der Umsetzung der Applikation auf dem Tablet. Diese variieren natürlich je nach Anwendung, aber typischer Weise lassen sich Verbesserungen der Nutzerinteraktion, wie die grafische Programmgestaltung oder die intuitive Menüführung verbessern. Ein ähnlicher Punkt ist die Vernetzung von Informationen. Tablet-Computer sind mobile Geräte, die auf die Synchronisation und den Abruf von Daten aus dem Internet und lokalen Netzwerken angewiesen sind. So könnte zum Beispiel auch die Entwicklung eines eigenen Web-Service für den Datenaustausch interessant sein. Dieser Aspekt zielt insbesondere auf Anwendungen ab, bei denen kein entsprechender Service zur (kostenlosen) Verfügung steht (zum Beispiel bestimmte kostenpflichtige, beziehungsweise eingeschränkt kostenlose, Google-Dienste).

## 12 Beantragung von Ressourcen

Für die Softwareentwicklungsarbeiten sind Arbeitsplätze am Fachbereich (inkl. durch die IRB bereitgestellte Rechnerzugänge und Projektgruppeninfrastruktur) notwendig sowie ein Besprechungsbereich für regelmäßige Teamsitzungen der Projektgruppenmitglieder.

## 13 Anrechnung zur Lehrverpflichtung der Antragsteller

Bei Zustandekommen der Projektgruppe soll die Betreuung durch die Antragsteller wie folgt auf deren Lehrverpflichtung angerechnet werden:

### Wintersemester 2012/13

Gernot A. Fink: 4 Stunden  
Leonard Rothacker: 4 Stunden

### Sommersemester 2013

Gernot A. Fink: 4 Stunden  
Leonard Rothacker: 4 Stunden