

Kapitel 1

Einleitung

Als Kommunikationsform und Erweiterung des menschlichen Gedächtnisses hat die geschriebene Sprache in all ihren Formen über viele Jahrhunderte hinweg die Entwicklung der Kulturen in entscheidender Weise beeinflusst. Seit der Erfindung des Buchdrucks verlagerte sich die Erstellung von Schriften zunehmend vom manuellen in den maschinellen Bereich. Dieser Prozeß wird scheinbar durch die starke Verbreitung leistungsfähiger Computer und deren Vernetzung zunehmend beschleunigt. Aufgrund dieser Entwicklungen wird nun - an der Schwelle zum dritten Jahrtausend - zunehmend die Frage nach der zukünftigen Bedeutung der Handschrift aufgeworfen.

Die skizzierten Entwicklungen mögen einerseits den Schluß nahelegen, dass die Nutzung von Handschrift als Kommunikationsmittel in Zukunft nicht mehr relevant sein wird. Andererseits, ebenfalls bedingt durch die technologischen Entwicklungen, scheint sich aber eine weitere Nutzungsform der Handschrift zu erschließen. Insbesondere die zunehmende Miniaturisierung von Computern, bei gleichzeitig zunehmender Rechenleistung, wecken den Bedarf nach fortschrittlichen Mensch-Maschine-Schnittstellen. Die Nutzung einer Tastatur erscheint bei Miniaturisierungsgraden heutiger PDA (personal digital assistants) oder Mobiltelefonen nicht mehr zweckmäßig. Darum rücken natürlichere Kommunikationsformen zwischen Mensch und Maschine stärker ins Blickfeld. Zu solchen Kommunikationsformen zählt zweifellos die gesprochene Sprache, wie auch die Handschrift. Dabei stellt sich offensichtlich weniger die Frage ob besser Sprache *oder* Handschrift verwendet werden sollte, sondern wie sich gesprochene Sprache *und* Handschrift ergänzend kombinieren lassen.

Während sich die gesprochene Sprache als Eingabemodus größerer Textmengen anbietet, scheint die Handschrift z. B. besonders geeignet für die Eingabe persönlicher Notizen.

Weiterhin stellt sich die Realisierung einer robusten Spracherkennung in Umgebungen mit starken Hintergrundgeräuschen als problematisch dar. Auch hier gilt die Handschrift als ergänzender Eingabemodus zur gesprochenen Sprache. Darüber hinaus sind einige Dokumentbestandteile, wie z. B. Skizzen oder Formeln, unter Verzicht auf Tastatur und Maus

sinnvoller Weise nur mittels eines Stiftes einzugeben.

Wie auch immer die speziellen Anwendungen für eine Handschrift-basierte Mensch-Maschine-Schnittstelle aussehen, langfristig werden die mit der Schnittstelle verbundenen Erkennungsalgorithmen drei wesentliche Forderungen erfüllen müssen:

- Die Erkennungssysteme sollten schreiberunabhängig arbeiten, oder aber im schreiberabhängigen Modus mit begrenztem Datenmaterial trainierbar sein.
- Das Erkennungssystem sollte über einen möglichst großen Wortschatz verfügen.
- Es sollten keine besonderen Bedingungen an den Schreibstil gestellt werden. Das System sollte durch den Benutzer mit seinem persönlichen Schreibstil bedienbar sein. Derzeitige verfügbare Produkte, mit einem fest definierten, speziellen Zeichensatz, können lediglich als Übergangslösung angesehen werden.

Wenngleich das äußerst facettenreiche Problem der Handschrifterkennung nicht in allen Details und mit absoluter Erkennungssicherheit gelöst werden kann, so sollen doch im Rahmen dieser Arbeit einige Lösungsvorschläge erarbeitet und präsentiert werden, die in die Richtung der oben genannten Anforderungen nach robuster und schreibstilunabhängiger Erkennung zielen.

Die für den Aufbau der Systeme akquirierten Trainings- und Testbeispiele orientieren sich konsequent an diesem Ziel. Hierbei werden die Testpersonen vor der Datenaufnahme nicht instruiert in besonderer Weise zu schreiben. Es findet keine Unterscheidung zwischen gedruckter Handschrift, vollständig verbundener (kursiver) Schrift oder diversen Mischtypen statt.

Wie in Abb. 1.1 dargestellt, bildet die Verarbeitungskette bestehend aus Vorverarbeitung, Merkmalsextraktion, Modellebene und Erkennung die Grundstruktur des Gesamtsystems. Die vorliegende Arbeit mit der Aufteilung der Kapitel bildet diese Struktur nach.

In Kapitel 2 werden somit zunächst die wesentlichen Vorverarbeitungsmethoden erklärt. Dies umfasst im wesentlichen die Glättung der Rohdaten, sowie die Normalisierung der Schrift. Die Normalisierung wird durchgeführt, um vorab die Variabilität der Handschrift bezüglich der Schrift- und Zeilenneigung, sowie der Schriftgröße einzuschränken.

In Kapitel 3 werden einige Verfahren zur Merkmalsentnahme untersucht. Um später die Modellparameter zuverlässig schätzen zu können, ist es das Ziel der Merkmalsextraktion, aus der Handschrift mit möglichst wenig Parametern die relevante Information zu extrahieren.

Kapitel 4 und 5 betreffen die Modellierungsaspekte. Während in Kapitel 4 die Frage nach der optimalen Modellierungsform beantwortet wird, werden in Kapitel 5 Vorschläge zur Strukturverknüpfung kontextabhängiger Modelle präsentiert. Bei kontextabhängigen Modellen handelt es sich um multiple Modelle für ein Zeichen. In Abhängigkeit der möglichen Kombinationen der Nachbargrapheme entsteht eine große Anzahl solcher kontextabhängigen Mo-

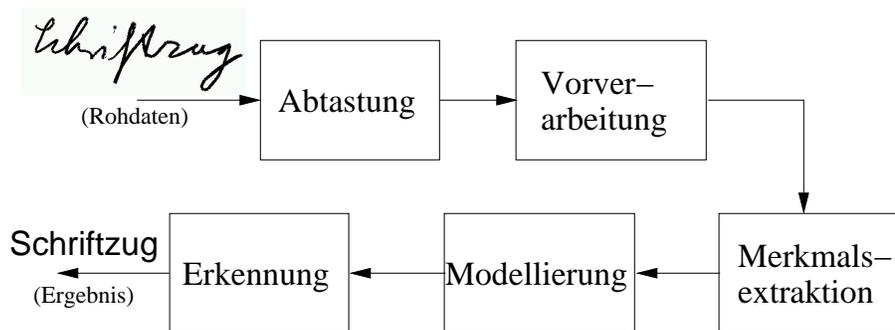


Abbildung 1.1: Systemübersicht

delle. Die Trainierbarkeit dieser Modelle wird schließlich durch die geschickte Verknüpfung von Modellparametern ermöglicht.

Bei verbundener, oder zumindest stückweise verbundener Handschrift, ist für die Erkennung die Angabe der erlaubten Wörter und der darin enthaltenen Buchstabenfolge in Form eines Lexikons notwendig. Mit dem in Kapitel 6 vorgestellten Ansatz ist es möglich diese Restriktion weiter zu reduzieren. Das feste Lexikon und die darin starr vorgegebenen Buchstabenfolgen werden durch eine statistische Beschreibung ersetzt. Die Ausdehnung der statistischen Beschreibung auf einen hinreichend großen Bereich erfordert dabei allerdings spezielle Dekodierungsalgorithmen.

Die Kapitel 2 und 3 gelten als allgemeingültig und sind als weitgehend unabhängig von der Zielanwendung und von der Funktionsweise des verwendeten Erkennungsansatzes zu verstehen. Kapitel 4 und insbesondere Kapitel 5 beziehen sich auf die Hidden Markov Modell basierte Texterkennung mit großem Vokabular, wobei in Kapitel 6 die Restriktion des festen Lexikons weiter reduziert wird. Kapitel 7 schließlich zeigt ein weiteres interessantes Anwendungsszenario für die Online-Handschrifterkennung. Hierbei handelt es sich um die Erkennung und Verarbeitung handgeschriebener Formeln. Die in Kapitel 7 beschriebenen Verfahren setzen auf den zuvor beschriebenen Algorithmen auf. Darüber hinaus bedarf es für die Analyse der zweidimensionalen Struktur von Formeln weiterer Schritte, die in diesem Kapitel beschrieben werden. Das System wird schließlich zu einem interaktiven, handschriftlichen Formeleditor erweitert.

Die Ergebnisse werden gesondert zu den einzelnen Kapiteln diskutiert. Eine Zusammenfassung schließt die Arbeit ab.