COGNITUS – Intelligente Formularerfassung

Die maschinelle Erkennung hand- und maschinengeschriebener Dokumente hat in den letzten Jahren zu großen Kostensenkungen bei vielen industriellen Anwendungen geführt, z.B. bei der Erfassung von Überweisungsformularen, Kreditkartenquittungen und Schecks sowie beim automatischen Sortieren von Postsendungen. Dabei wird meist eine hohe Verarbeitungsgeschwindigkeit (mehrere Formulare pro Sekunde) bei gleichzeitig sehr niedrigen Fehlerraten verlangt, was sowohl den Einsatz leistungsfähiger Rechner als auch Entwicklung zuverlässigerer und effizienter Schrifterkennungsverfahren erfordert. Am "Zentrum für Paralleles Rechnen" wurde hierzu ein umfassendes und flexibles System zur Formularerfassung entwickelt, das neuartige Zeichenerkennungsmethoden eng mit einer dynamischen Segmentierung und einer Kontextanalyse koppelt, um so die erforderliche Zuverlässigkeit zu gewährleisten.

Ansatz

Bei den meisten im Einsatz befindlichen Erfassungssystemen wird die Bearbeitung in eine Folge unabhängiger Bearbeitungsschritte unterteilt. In einer Vorverarbeitungsstufe wird aus dem Farbbild die binäre Schriftinformation herausgefiltert und das Formular grob in Felder zerlegt. Eine Segmentierungseinheit zerlegt die Schrift in einzelne Zeichen, die die Erkennungseinheit einzeln klassifiziert. Auftretende Fehler werden meist in einer aufwendigen manuellen Nachbearbeitung korrigiert. Solche "Optical Character Recognition" (OCR)-Techniken haben sich für maschinelle Schriften als ausreichend zuverlässig erwiesen und sind schon lange in industriellem Einsatz.

Handschriften hingegen erfordern aufgrund ihrer Variabilität und auftretender Mehrdeutigkeiten aufwendigere Techniken, die auch Kontextinformationen ausnutzen. Schon eine korrekte Segmentierung in Einzelzeichen stellt sich als sehr schwierig heraus, weil viele Zeichen zerrissen oder durch Ligaturen verbunden sind. Oft ist erst nach der Erkennung eine Entscheidung über die korrekte Zerlegung möglich. In vielen Fällen können einzelne Zeichen selbst vom Menschen nicht einer eindeutigen Klasse zugeordnet werden, so daß eine Entscheidung im Zusammenhang getroffen werden muß, z.B. durch Abgleich mit einem Wörterbuch oder anderen Datenbanken.

Das System "COGNITUS", das am ZPR entwickelt wurde, basiert auf leistungsfähigen hierarchischen Nächste-Nachbar-Klassifikationsverfahren, die erlauben, eine große Vielfalt von Schriftarten abzudecken und damit sehr niedrige Fehlerraten zu erzielen. In enger Zusammenarbeit zwischen der Segmentierungs-, Klassifizierungs- und Kontextkorrektureinheit wird die global wahrscheinlichste Interpretation ermittelt, wodurch eine sehr hohe Zuverlässigleit erreicht wird. Das Programm wurde zunächst zur Erfassung hand- und maschinengeschriebener Überweisungsformulare entwickelt, ist jedoch leicht vom Benutzer auf andere Anwendungen anpaßbar. Durch eine effiziente Implementierung wird eine hohe Verarbeitungsleistung von etwa einem Formular pro Sekunde schon auf einem Pentium-200 Prozessor erreicht. Für noch höhere Leistungsanforderungen ist der Einsatz von Parallelrechnern leicht möglich. Im folgenden gehen wir auf die einzelnen Komponenten des Systems näher ein.

Zentrum f"ur Paralleles Rechnen — Arbeitsgruppe Prof. Dr. A. Bachem

Telefon: (0221) 470–6030

Hausanschrift: Weyertal 80, D-50931 Köln — Telefax: (0221) 470–5160

Postanschrift: ZPR, Universität zu Köln, Albertus-Magnus-Platz, D-50923 Köln — Email: zpr@zpr.uni-koeln.de

Klassifikation durch binäre Clusterverfahren

Zur Zeichenerkennung werden "Nächste-Nachbar-Verfahren" eingesetzt, die eine interessante Alternative zu sonst üblichen neuronalen Netzwerken darstellen. Nach Normierung der Zeichen erfolgt die Klassifizierung durch einen *direkten Vergleich* der binären Muster mit einem großen Satz von Prototypen. Traditionelle Clusterverfahren werden benutzt, um in einer "Lernphase" vorab aus einer sehr großen Menge von Beispielmustern charakteristische Referenzmuster zu ermitteln. Am ZPR wurden binäre Clusteralgorithmen entwickelt, die eine erhebliche Beschleunigung der Schrifterkennung erreichen und die erlauben, auch mit extrem großen Datensätzen zu arbeiten.

Nächste-Nachbar-Verfahren erzielen sehr niedrige Fehlerraten, wenn ausreichend viele Prototypen (mehrere 100000) zum Einsatz kommen, die eine große Vielfalt von Formen abdecken können. Um den dann anfallenden Rechenaufwand zu reduzieren, wurden am ZPR *hierarchische Verfahren* entwickelt, die in einfachen Fällen eine Klassifikation mit wenigen Prototypen vornehmen während in Zweifelsfällen sehr aufwendige Opertationen eingesetzt werden, um ein höchstes Maß an Sicherheit zu garantieren. Auf diese Weise können etwa 1000 Ziffern pro Sekunde bei einer Fehlerrate von nur 0.5% erkannt werden.

Durch Kombination dieser Methode mit neuronalen Netzwerken, die anders arbeiten und eine andere "Sichtweise" einbringen, kann die Zuverlässigkeit der Erkennung weiter gesteigert werden. Die Untersuchung solcher "Multi-Agent-Klassifizierer" ist Thema einer Zusammenarbeit mit anderen Hochschulen.

Dynamische Segmentierung

Es wurden effiziente binäre Verfahren entwickelt, die ein Formular in etwa 0.1 Sekunden in Felder und Einzelzeichen zerlegen, wobei bei Mehrdeutigkeiten eine Vielzahl von Möglichkeiten ausgegeben wird. Die Erkennungseinheit bewertet all diese Segmente, und die wahrscheinlichste Folge wird erst danach durch Optimierungstechniken ermittelt.

Datenbankabgleich

In Zusammenarbeit mit der Parsytec GmbH, Aachen wurde eine Kontextkorrektur durch Datenbankabgleich für Überweisungsformulare entwickelt. Dabei wird der Zusammenhang zwischen Kontonummer und Konteninhaber sowie zwischen der Bankleitzahl und den Banknamen ausgenutzt, um die geforderte Sicherheit zu erreichen. Techniken der *dynamischen Programmierung* werden eingesetzt, um zu einer erkannten Symbolfolge den nächsten Eintrag in einer Kundendatenbank zu ermitteln. Durch eine stark optimierte Implementierung können pro Sekunde 100 Formulare abgeglichen werden. Bei Zurückweisung von 17% der Fälle wird beispielsweise die Fehlerrate für die Kontennummern auf weniger als 0.1% reduziert.

Implementierung

Das System wurde bislang mit etwa 200 000 maschinengeschriebenen und 60 000 handgeschriebenen Zeichen trainiert. Es unterscheidet zuverlässig zwischen den beiden Schifttypen und verwendet entsprechend unterschiedliche Referenzmuster. Auf einem Pentium 200 MHz erfordert die komplette Bearbeitung eines Überweisungsformulares nur 0.2 Sekunden. Das Programm liegt portabel als ANSI-C-Bibliothek vor. Der Benutzer kann alle wesentlichen Parameter zur Ablaufsteuerung und zum Formularlayout modifizieren und damit leicht unterschiedliche Anwendungen realisieren.

Ansprechpartner

Barthel Steckemetz
Zentrum für Paralleles Rechnen (ZPR)
Universität zu Köln
Weyertal 80
D-50931 Köln

Tel.: 0221/470-6039
Fax.: 0221/470-5160
email: `<steckemetz@zpr.uni-koeln.de>`

Stand: 1.3.98