Zeitschrift: Bulletin des Schweizerischen Elektrotechnischen Vereins, des Verbandes Schweizerischer Elektrizitätsunternehmen = Bulletin de l'Association suisse des électriciens, de l'Association des entreprises électriques suisses

Herausgeber: Schweizerischer Elektrotechnischer Verein ; Verband Schweizerischer Elektrizitätsunternehmen

Band: 69 (1978)

Heft: 22: Sondernummer Elektrotechnik 1978 = Edition spéciale Electrotechnique 1978

Artikel: Elektronische Lesegeräte

Autor: Kis, F.

DOI: https://doi.org/10.5169/seals-914959

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Mehr erfahren

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. En savoir plus

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. Find out more

Download PDF: 09.12.2025

ETH-Bibliothek Zürich, E-Periodica, https://www.e-periodica.ch

Elektronische Lesegeräte

Von F. Kis

681.327.1;

1. Einführung

Elektronische Lesegeräte sind Datenerfassungsgeräte, denen heute immer mehr auch Datenverarbeitungsfunktionen aufgetragen werden, um ihre Wirtschaftlichkeit und Einsatzfähigkeit zu erhöhen. Die Information soll nicht nur gelesen, sondern in zweckmässiger Form weitergegeben werden und gegebenenfalls gewisse Vorgänge auslösen. Die zum Lesen unerlässlichen Erkennungsvorgänge verleihen diesen Geräten eine Art Intelligenz, die Beachtung verdient, insbesondere auch deshalb, weil in der Datenerfassung und -verarbeitung hierdurch neue Möglichkeiten geboten werden, die ständig wachsende Datenflut zu bewältigen.

Die Schriftzeichenerkennung stellt einen kleinen Bereich der *Zeichenerkennung* im allgemeinen dar, deren Spektrum noch andere aktuelle Gebiete umfasst, wie Spracherkennung, Identifikation von Partikeln oder Blutkörperchen, Qualitätsprüfung, Auswertung von Flug- und Satellitenaufnahmen, Stellen von Diagnosen anhand von vorgegebenen Krankheitssymptomen, usw. Es handelt sich im allgemeinen um die Identifikation von strukturierten Signalen oder Gestalten, welche in bestimmte Gruppen oder Klassen eingeteilt werden können. Wesentlich dabei ist, dass die Zeichen der gleichen Klasse nicht identische, sondern ähnliche Erscheinungen beinhalten (Fig. 1). Die Identifikation solcher Zeichen war bis vor kurzem nur Menschen möglich. Sie wird in Zukunft immer mehr Erkennungsmaschinen übertragen. Die komplizierte Informationsverarbeitung bei Erkennungsvorgängen in Lebewesen kann zwar noch nicht verstanden werden. Man baut Maschinen nach dem letzten Erkenntnisstand. Dabei spielen die Intuition und die moderne Technologie eine grosse Rolle.

2. Über Lesegeräte allgemein

Die konventionellen Dateneingabeeinheiten bilden bekanntlich einen Engpass in der Datenverarbeitung. Die technische Entwicklung auf diesem Gebiet ist infolge der aufwendigen Realisierung gegenüber den heutigen schnellen EDV-Anlagen bedeutend zurückgeblieben. An moderne Dateneingabeeinheiten werden zwei wichtige Forderungen gestellt: Schnelligkeit und Sicherheit, also weniger Fehler bei grösserer Geschwindigkeit.

Ein weiterer Grund zur Einsetzung von Lesegeräten ist die Vermeidung von Datenzwischenträgern wie Lochkarten und Magnetbänder. Das meist unvermeidbare Kopieren der Daten bildet eine potentielle Quelle von Fehlern, was zu Unterschieden zwischen den primären und sekundären Datenträgern führen kann. Die von Mensch und Maschine lesbaren Urbelege können jederzeit kontrolliert und archiviert werden.

Die ersten elektronischen Lesegeräte für praktische Anwendungen kamen in den 50er Jahren auf den Markt. Seither ist eine eindeutige Aufwärtsbewegung auf diesem Gebiet festzustellen, insbesondere in den USA, aber auch in Europa und Japan.

Das Einsatzgebiet von optischen Lesegeräten, in Fachkreisen *OCR-Geräte* (Optical Character Recognition) genannt, erstreckt sich auf Banken, Versicherungen, Industrieunternehmen, Rechenzentren, Handelsbetriebe, Postwesen und auch auf kleinere Betriebe. Untersuchungen zeigen, dass die Kosten der Datenerfassung oft 30...50 % der gesamten Datenverarbeitung ausmachen. Es ist heute möglich, durch richtiges Einsetzen eines geeigneten OCR-Lesers drei konventionelle Datenerfassungsplätze einzusparen.

Die optisch-elektronischen Lesegeräte können wie folgt gruppiert werden:

Belegleser ist oft ein Sammelname für alle beleglesenden Maschinen. Er bezeichnet aber meistens Geräte, die Checkformate mit oder ohne Sortierfunktionen bearbeiten. Man spricht von Belegsortierung, wenn die gelesenen Dokumente in mehr als zwei Ablagefächer verteilt werden. Belegleser zeichnen sich durch hohe Durchlaufgeschwindigkeiten aus, wobei in den meisten Fällen nur eine Datenzeile auf dem Beleg gelesen werden muss. Sie werden hauptsächlich in Geldinstituten eingesetzt und erreichen eine Leseleistung von bis zu 140 000 Belegen pro Stunde.

Seiten- und Mehrfunktionsleser (Allround-Leser) bearbeiten Belege von Checkformat bis zu A4-Grösse und darüber. Ihre Verarbeitungsgeschwindigkeit erreicht etwa 400...1000 Belege/h. Seitenleser werden im allgemeinen als Multi-Font-Leser eingesetzt, d. h. sie lesen die für maschinelles Lesen geeigneten Schriftarten (Fontarten) wie OCR-A, OCR-B, IBM-1428, NCR-NOF, welche in Lesefelder aufgeteilt, aber auf dem gleichen Beleg auftreten können. Daneben sind auch Handschriftzeichen (Blockschrift) sowie Strichmarkierungen möglich.

Journalstreifenleser lesen Streifen, die von Buchungsmaschinen, Registrierkassen und Additionsmaschinen mit einer geeigneten Schriftart bedruckt werden. Seitenleser sind oft auch für Journalstreifenlesung ausgerüstet (Fig. 2).

Markierungsleser sind, verglichen mit den oben genannten, einfachere Geräte, welche Belege mit Strichmarkierungen lesen. Diese werden im allgemeinen dort benützt, wo eine Auswahl von vorgegebenen Möglichkeiten getroffen werden soll, z. B. auf Lottoscheinen, Prüfbogen und Formu-

2 2 2 2 2 3 3 3 3 3 3

Fig. 1 Verschiedene Zeichen der Klasse 2 und 3

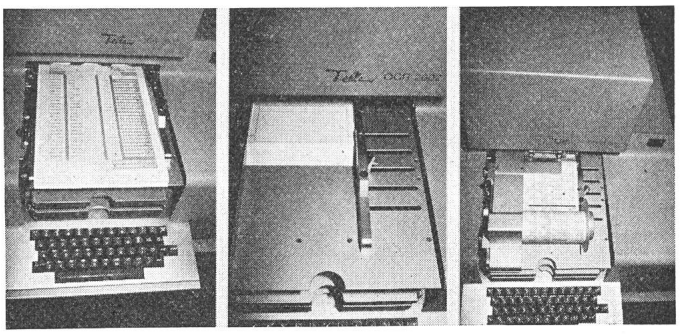

Fig. 2 Wahlweiser Einsatz eines Seitenlesers zur Lesung von A4-Dokumenten, A6-Dokumenten und Journalstreifen

laren für statistische Auswertungen. Markierungslesen benötigt den einfachsten Erkennungsvorgang, nämlich eine Ja- oder Nein-Entscheidung, abhängig von der Schwärzung eines Zeichenfeldes.

Handleser haben in den letzten Jahren immer mehr Verbreitung gefunden. Sie lesen OCR-A oder OCR-B und arbeiten z. B. in Verbindung mit Kassenterminals. Charakteristisch ist für diese Leser, dass sie einen beweglichen Abtastkopf besitzen, mit welchem die Informationszeile (z. B. einer Etikette) durch eine einzige Handbewegung abgetastet wird.

Die aufgeführte Gruppierung der Lesegeräte ist mehrheitlich anwendungsorientiert und entspricht der heute üblichen Einteilung. Da eine Beschreibung der Geräte bezüglich Schriftart oder Belegart wenig sinnvoll erscheint, sollen zunächst die typischen Funktionen eines Mehrfunktionslesers behandelt werden.

3. Schriftzeichenerkennung

Ein System für Schriftzeichenerkennung kann aufgeteilt werden in die Abtasteinheit, die Vorverarbeitung der abgetasteten Zeichen, die Merkmalextraktion, die Erkennung, die Aufbereitung der Zeileninformation (Datenverarbeitung), den Belegtransport und die Ausgabeeinheit. Damit sind auch die einzelnen Problemkreise eines Erkennungssystems gegeben. In einem modernen elektronischen Lesegerät stehen diese Systemteile unter Kontrolle eines schnellen Mikroprozessors oder Minicomputers, der gleichzeitig auch eine komfortable Bedienung und vielseitige Anwendung des Lesers ermöglicht.

3.1 Abtasteinheit

Das zweidimensionale Bild jedes einzelnen Zeichens muss in elektrische Signale umgewandelt werden, und zwar nach einem vorgegebenen Abtastprogramm, z. B. rasterförmig oder konturfolgend. Dabei begnügt man sich in der Schriftzeichenerkennung mit einer Schwarz-Weiss-Quantisierung der Information, welche sich durch eine Abtastschwelle zwischen dem Weissbereich des Datenträgers und dem Schwarzbereich des Zeichendruckes festlegen lässt.

Ein konventioneller und bewährter Abtaster ist die Flying-Spot-Röhre (lichtintensive Elektronenstrahlröhre), kombiniert mit einem Photomultiplier oder mit einem Halbleiter-Lichtsensor. Durch die einfache Steuerbarkeit des Abtaststrahles ergibt sich eine grosse Flexibilität bezüglich des Abtastprogrammes. Deshalb eignet sich diese Röhre heute noch vorzüglich als Konturabtaster, in welchem der Lichtpunkt so gesteuert wird, dass die Konturen des Zeichens von ihm sozusagen abgefühlt werden. Aus dem zeitlichen Verlauf der Steuerspannung kann dabei auf die Topologie und Geometrie des vorliegenden Zeichens geschlossen werden. Nachteile dieser Methode sind grosse Abmessungen, hohe Beschleunigungsspannung und beschränkte Abtastgeschwindigkeit als Konturabtaster.

Grundsätzlich ist es gleichgültig, ob ein Zeichen rasterförmig oder nach einem anderen Programm abgetastet wird, solange relevante Information dabei nicht verlorengeht. Die Art der Abtastung kann aber die darauffolgende Verarbeitung der Zeicheninformation wesentlich beeinflussen.

Dank der heutigen Halbleitertechnologie sind sog. Solid-State-Line-Scanners mit grosser Sensorendichte erhältlich. Sie ermöglichen, die ganze A4-Breite auf 1024 Fotodioden (ca. 2,5 cm) linear abzubilden. Jede einzelne Fotodiode besitzt eine zugeordnete Kapazität, welche sich durch einen zur Belichtung proportionalen Strom auflädt. Die Abfrage der Zeile erfolgt sequenziell mit Hilfe eines im Line-Scanner integrierten Schieberegisters. Die Amplitude des Videoimpulses entspricht der Aufladung der Diodenkapazität. Die Binärquantisierung der Videoinformation kann also während der Abfrage mit einer einzigen Schwellwertschaltung erfolgen.

Ein Abtaster muss in der Lage sein, die Schwarz-Weiss-Verteilung des Zeichens ohne Verzerrung in einem zweidimensionalen Speicher wiederzugeben. Zur Erfüllung dieser Aufgabe werden an die Papierqualität des Beleges bestimmte Anforderungen gestellt: Die Papieroberfläche muss glatt, matt, weiss und ohne Fluoreszenzwirkung sein. Das Papier muss ausserdem undurchsichtig sein und darf keine schwarzen Einschlüsse aufweisen. Für diese Eigenschaften existieren Test- und Messmethoden, welche die Qualität des sog. OCR-Papiers vorschreiben. Die diesbezüglichen Qualitätsempfehlungen der Gerätehersteller zur Erreichung optimaler Leseresultate sind heute ohne Probleme erfüllbar. Es sind auch sog. Blindfarben auf dem Papier zulässig, damit für das Ausfüllen des Beleges Hilfsfelder geschaffen werden können. Diese Blindfarben sind ebenfalls vom Gerätehersteller definiert und hängen von der Spektralempfindlichkeit des verwendeten Fotosensors ab.

Um gute Kontrastunterschiede auf dem Dokument zu erzielen, sollten in der Schreibmaschine Einmalbänder verwendet werden, wobei die Schwärzung keine Reflexion des Lichtes verursachen darf. Für Handschrifteintragungen eignen sich am besten spezielle OCR-Stifte (z. B. Pentel-Schreiber) oder spitze Bleistifte (Härte H oder HB).

3.2 Vorverarbeitung der Zeicheninformation

Diese Verarbeitungsstufe besteht aus mehreren Schritten und bezweckt die Vorbereitung der abgetasteten Zeicheninformation für die Merkmalextraktion. Dabei soll die nicht relevante Information des Zeichens reduziert werden.

Jedes Zeichen ist mit Hintergrundrauschen infolge kleiner Druck- und Schreibfehler behaftet. In leistungsfähigen Geräten ist während der Abtastung ein zweiter Schwellwert für Grauwerte in Funktion, durch welche neben der Schwarz-Weiss-Verteilung auch eine Grau-Weiss-Verteilung der Zeichen entsteht. Mit Hilfe eines Algorithmus kann aus den zwei quantisierten Feldern auf die wahrscheinliche Verteilung geschlossen werden. Der Algorithmus definiert einen zweidimensionalen Operator, der als lokaler «Tiefpassfilter» (z. B. 5×5 Rasterpunkte) in der x- und y-Richtung über alle Positionen des Rasters geschoben wird. Man vergleicht also damit die Graupunkte mit den Nachbarpunkten, um kleine Kontinuitätsstörungen in den Konturen auszugleichen und kleine Störflecken zu eliminieren.

Das Resultat dieses Vorganges ist das Rasterbild des Zeichens, das in ein zweidimensionales Schieberegister (Abtastspeicher) geleitet wird. Dieses ist so dimensioniert, dass in ihm alle Zeichen einer Zeile (ca. Belegbreite) leerstellengetreu abgebildet werden können. Dieser Speicher bildet den Ausgang für die weitere Verarbeitung der Zeileninformation.

Das Segmentieren (Herauslesen der Zeichen und Leerstellenstücke) der Zeileninformation vollzieht sich dadurch, dass der Inhalt des Abtastspeichers unter Kontrolle mehrerer

Mess- und Prüfschaltungen sowie des Mikroprozessors oder Minicomputers aus dem Schieberegister geschoben wird. Damit sind Messungen, wie Höhe und Breite der einzelnen Zeichen sowie Anzahl der Leerstellen zwischen ihnen, verbunden. Bereits an dieser Stelle werden zwei infolge Verschmutzung einander berührende Zeichen und vom Benützer durchgestrichene Zeilen (Line Delete) erfasst. Da in einer Zeile bei Multi-Font-Lesern verschiedene Schriftarten möglich sind (siehe Fig. 6), müssen alle Messungen unter Programmkontrolle des Mikroprozessors ausgeführt werden. D. h., dass im ganzen Erkennungsvorgang (inklusive Vorverarbeitung) die Messungen von der im Lesefeld befindlichen Schriftart abhängig sind (Zeichen und Leerstellen sind z. B. bei Handschrift grösser). Die Aufteilung des Beleges in Lesefelder erfolgt durch den Benützer mit Hilfe eines Anwenderprogrammes.

Die Zeichen einer zur optischen Lesung geeigneten Maschinenschrift weisen eine regelmässige Grösse auf. Anders verhält es sich mit den Handschriftzeichen (Blockschrift). Sie sind, trotz Schreibempfehlungen, nicht nur in der Form verschieden, sondern auch in der Grösse und Strichdicke. Handschriftzeichen müssen deshalb auf eine vom System vorgesehene Grösse transformiert werden. Die Grundlage zur Grössentransformation bilden die gemessene Höhe und Breite des Zeichens. Die während der Transformation entstehende Verformung verursacht eine Verzerrung des Zeichens. Um diesen unvermeidbaren Informationsverlust möglichst klein zu halten, müssen bestimmte Algorithmen als Verknüpfung zwischen benachbarten Zeilen und Kolonnen des gerasterten Zeichens angewandt werden.

Eine weitere Vorverarbeitung ist das Zentrieren des in der Grösse normierten Zeichens, und zwar in die Mitte oder Ecke eines hiefür vorgesehenen Zeichenfeldes. Durch diesen Vorgang werden auch die Bezugspunkte der Merkmalmessungen festgelegt. Die Rastergrösse des Zeichens muss etwa 12 × 20 Rasterpunkte betragen. Experimente bei Menschen mit alphanumerischen Zeichen haben gezeigt, dass die Erkennungssicherheit gut und schnell ist, solange das Zeichen auf der Netzhaut etwa die Grösse von 20 × 20 Rezeptoren bedeckt. Darunter fällt die Sicherheit und Schnelligkeit rapid ab, unter einer Feldgrösse von 10 × 10 Rezeptoren hört die Erkennung ganz auf.

3.3 Merkmalextraktion

Das Zeichenfeld dient nun zur Messung von charakteristischen Merkmalen der Zeichen. Dabei ist wesentlich, ob es sich um Handschriftzeichen oder um die übliche optische Maschinenschrift handelt. Abhängig von der durch den Benützer programmierten Fontart auf dem Dokument müssen die entsprechenden Meßstufen, welche im Zeichenfeld wirken, in Funktion gesetzt werden.

Maschinenschrift hat den Vorteil, dass die Zeichen der gleichen Klasse praktisch die gleiche Form und Grösse aufweisen. Für Maschinenschrift genügt es in vielen Fällen, dass jeder Rasterpunkt des Zeichenfeldes als ein Binärmerkmal aufgefasst wird. Im Falle der Feldgrösse von 12 × 20 ergeben sich dabei pro Zeichen 240 Binärmerkmale, die abhängig von der Schwarzverteilung des Zeichens den Wert 1 oder 0 annehmen.

Bei der *Handschrift* ist eine Lösung an Hand der Rastermerkmale illusorisch. Arbeiten in der Neurophysiologie bestätigen, dass bei Lebewesen bereits unmittelbar hinter der Retina eine Reduktion der ankommenden Information auf sog. Eigenschaften des abgebildeten Zeichens stattfindet, welche in unserem Sinne einer Merkmalextraktion gleichkommt. Ein Rezept daraus für technische Systeme abzuleiten, ist bisher nicht gelungen, und man ist bei der Realisierung von Erkennungssystemen weitgehend auf Intuition und Experimente angewiesen. Ideale Merkmale sollten folgende Eigenschaften aufweisen:

– Um nicht-korrelierte Messwerte über das Zeichen zu erhalten, sollten die angewandten Messungen voneinander linear unabhängig sein (eine Messung sollte keine Teilmessung einer anderen sein).

– Die Beschreibung der Zeichen durch Merkmale sollte die Zeichen derselben Klasse einander näherbringen (kleine Hammingdistanz). Unterschiede zwischen klassenfremden Zeichen sollten dagegen durch sie hervorgehoben werden (grosse Hammingdistanz).

Hieraus resultiert auch die Definition der charakteristischen Merkmale der Zeichen. Solche sind eine möglichst kleine Anzahl physikalischer oder mathematischer Grössen, welche die Zeichen einer Klasse (z. B. der Klasse 5) eindeutig absondernd von klassenfremden Zeichen beschreiben können. Merkmale für Handschriftzeichen können von verschiedenen Arten sein, wie topologische (z. B. Endpunkte und Verzweigungen), geometrische (z. B. Krümmungsart und Form der Konturen) oder irgendwelche abstrakte (z. B. mathematische Projektionen). Sie können Formelemente (Bestandteile) oder Masken der Zeichen darstellen oder aber als reine Zahlen von Distanzmessungen (innerhalb der Zeichenkonturen) auftreten. Da heute noch nicht gesagt werden kann, welche Merkmale die beste Lösung liefern, wird jede für ein gegebenes System geeignete Art angewandt.

Die Meßschaltungen der Merkmale, die in verschiedenen Teilen des Zeichenfeldes wirken, sind recht kompliziert und aufwendig. Der Grund dafür liegt in der hohen Zahl der

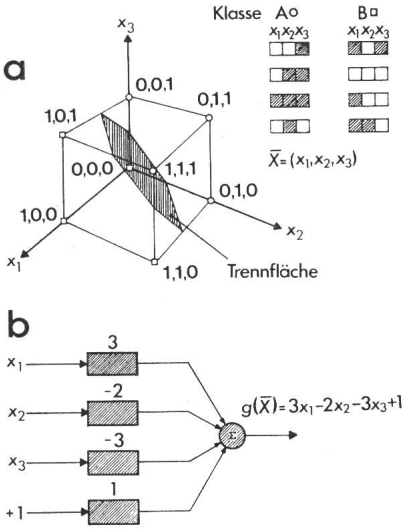

Fig. 3 Unterscheidung von zwei Zeichenklassen A und B mittels drei binärer Merkmale x_1, x_2, x_3 (Zweiklassenproblem im dreidimensionalen Binärraum)

a räumliche Darstellung
b gewichtete Korrelation

$\bar{X}$ gehört zur Klasse A, falls $g(\bar{X}) < 0$
$\bar{X}$ gehört zur Klasse B, falls $g(\bar{X}) > 0$

Rasterpunkte und darin, dass die Messungen möglichst invariant sein müssen gegen Störungen und Änderungen in der Strichdicke sowie gegen kleine Drehungen und Verschiebungen des Zeichens. Ausserdem müssen die Meßschaltungen im ganzen Zeichenbereich «verschiebbar» und eventuell «drehbar» sein, damit Messungen verschieden orientiert an verschiedenen Zeichenteilen ausgeführt werden können.

3.4 Erkennung

Die 240 Binärvariablen eines Zeichenfeldes stellen einen 240dimensionalen Binärraum dar, in welchem sich den Zeichen entsprechende Punkte (insgesamt 2^{240}) in den Eckpunkten eines Würfels von derselben Dimension Platz einnehmen. In diesem Raum gelten analoge Gesetzmässigkeiten wie in einem dreidimensionalen Raum (Fig. 3). Die Punkte einer Zeichenklasse befinden sich mehr oder weniger in einem Haufen, d. h. im Bereich kleiner Hammingdistanzen zueinander. Die Erkennung bei Maschinenschriftzeichen besteht demzufolge darin, dass die Punkthaufen durch Trennflächen, sog. Hyperebenen, voneinander separiert werden. Die Neigung dieser Trennebenen wird durch entsprechende «Gewichtung» der Eingangsvariablen zu jeder Klasse definiert. Damit werden bestimmte Punkte des Zeichens höher und andere niedriger bewertet, wodurch deren Summationsbeitrag (Korrelationsbeitrag) entsprechend grösser oder kleiner wird. Fig. 3 und 4 zeigen Entscheidungsnetzwerke zu 2 bzw. R Klassen mit gewichteten Eingangsvariablen des gerasterten Zeichens. Zur Bestimmung der Gewichte können rechnerische Methoden nur beschränkt angewandt werden. Eine für den praktischen Einsatz brauchbare Lösung wird in den meisten Fällen durch intuitive Gewichtsveränderung, verbunden mit einer Trainingsphase anhand von in der Wirklichkeit vorkommenden Zeichen, erarbeitet. Im Falle schwer separierbarer Klassen besteht die Möglichkeit, sog. Unterklassen der Zeichen zu bilden, um bestimmte Punkthaufen doch separieren zu können. Eine derartige Lösung entspricht einer stückweise-linearen Separierung im Raum. Fig. 5 zeigt die Arbeitsweise einer stückweise-linearen Separierung in einem zweidimensionalen «Raum», in welchem die eingezeichneten Klassenpunkte eine willkürliche Verteilung haben und voneinander linear nicht trennbar sind [1; 2; 3] [1]).

Im Prinzip können für die Erkennung von Handschriftzeichen die gleichen Überlegungen gemacht werden; den

[1]) Siehe Literatur am Schluss des Aufsatzes.

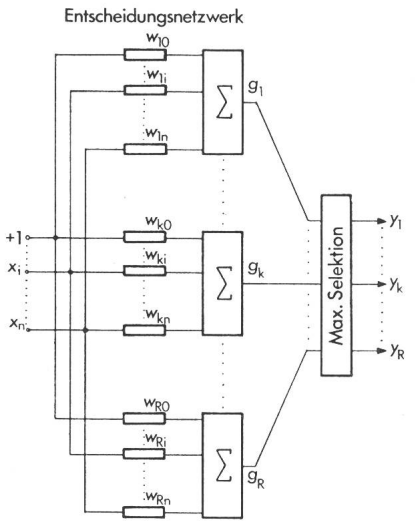

Fig. 4 Entscheidungsnetzwerk für R Zeichenklassen bei n binären, gewichteten Merkmalen

$\overline{X} = (+1, x_1, \ldots x_n)$ Binärvektor eines unbekannten Zeichens

$\overline{W}_k = (w_{k0}, w_{k1}, \ldots w_{kn})$ Gewichtsvektor der Klasse k

$g_k = \overline{W}_k \cdot \overline{X} \quad k = 1, \ldots R$ Vergleichswert der Klasse k

S_0 Sicherheitsdistanz

$\overline{X}$ ist ein Zeichen der Klasse l, falls
$g_l > S_0 + g_k \quad k = 1, \ldots R; \; l \neq k$

Dabei ergibt sich am Ausgang
$y_l = 1$ und $y_k = 0 \quad k = 1, \ldots R; \; l \neq k$

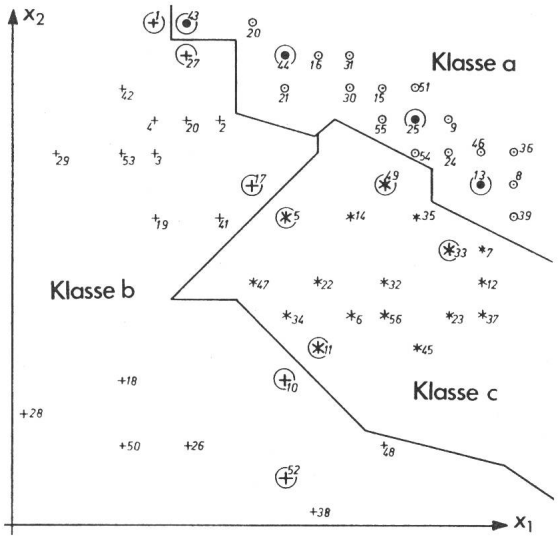

Fig. 5 Stückweise lineare Separierung im zweidimensionalen «Raum»

Die Punktverteilung kann dabei irgendein physikalisches Problem darstellen (z.B. x_1 = Temperatur, x_2 = Gasdruck). Die dargestellten Zustände (Punkte) sollen mit Hilfe der für das Problem repräsentativen Fälle (Referenzmengen) in drei Klassen eingeteilt werden.

$\overline{R}_a = \left\{\overline{A}_{13}, \overline{A}_{25}, \overline{A}_{43}, \overline{A}_{44}\right\}$ Referenzmenge der Klasse a

$\overline{R}_b = \left\{\overline{B}_1, \overline{B}_{10}, \overline{B}_{17}, \overline{B}_{27}, \overline{B}_{52}\right\}$ Referenzmenge der Klasse b

$\overline{R}_c = \left\{\overline{C}_5, \overline{C}_{11}, \overline{C}_{33}, \overline{C}_{49}\right\}$ Referenzmenge der Klasse c

$\overline{X} = (x_1, x_2)$ ein beliebiges Zeichen

$\overline{R}_k = (r_1, r_2)$ ein Referenzzeichen der Klasse k

g_k = Vergleichswert der Klasse k

$$g_k = \overline{R}_k \cdot \overline{X} - \frac{\overline{R}_k^2}{2}, \; k = \text{a, b, c}$$

$\overline{X}$ ist ein Zeichen der Klasse k, falls
$g_k > g_l \quad k, l = \text{a, b, c} \quad k \neq l$

Die Gleichung der Trennstücke ist gegeben durch

$$(\overline{R}_k - \overline{R}_l) \cdot \overline{X} - \frac{\overline{R}_k^2 - \overline{R}_l^2}{2} = 0 \qquad k, l = \text{a, b, c} \qquad k \neq l$$

Entscheidungsraum bilden dabei die gemessenen Zeichenmerkmale. Man spricht hier also über einen Merkmalraum, in dem die Zeichen abhängig davon, ob die Merkmale binär (0, 1), ternär (– 1, 0, + 1) oder mehrwertig quantisiert sind, in Punkte abgebildet werden. Infolgedessen könnte der in Fig. 3 gezeichnete Raum als Merkmalraum aufgefasst werden und dabei die Binärvariablen durch gemessene zweiwertige Merkmale (1 Merkmal vorhanden, 0 Merkmal fehlt) ersetzt werden. Die Separierbarkeit der Klassen im Merkmalraum ist auch ein Mass für die Leistungsfähigkeit der Merkmale.

Es kommt oft vor, dass man für Handschriftzeichen einen sog. logischen Entscheidungsbaum (logische Verknüpfung der Merkmale) für die Erkennung einsetzt. Der Grund hiefür liegt in der Natur der Merkmale. Ihre Messung liefert meistens logische Grössen über verschiedene Eigenschaften der Zeichen, die in der Logik dreiwertig wirken, nämlich als ja-, nein- oder «don't care»-Aussage. Ausserdem kann die Sequenz (Erscheinungsfolge) der Merkmale, welche vom Abtastprogramm abhängt, als eine weitere «Dimension» in der Logik berücksichtigt werden. Der Entscheidungsbaum muss selbstverständlich, ähnlich wie das Summationsnetzwerk in Fig. 4, so viel Ausgänge wie Anzahl zugelassener Klassen aufweisen.

Zur leistungsfähigen Handschriftlesung müssen eine relativ grosse Anzahl Merkmale (50...100) angewandt werden, für die sich ein logisches Netzwerk schlecht eignet. Statt dessen stellt man die gemessenen Merkmale als einen Binärvektor der Dimension n auf (n Anzahl Merkmale). Dieser kann dann mit in Festwertspeichern (ROM) gespeicherten Referenzvektoren der Klassen verglichen werden, um seine Klassenzugehörigkeit festzustellen. Solche Speicher sind heute bis zu 16 kbit erhältlich und ermöglichen eine schnelle Zugriffzeit. Bei dieser Methode muss die Dreiwertigkeit der Referenzvektoren berücksichtigt werden. Dies bedeutet praktisch die Verdoppelung der Speicherkapazität. Der Erkennungsmechanismus kann hier auch wie eine stückweise-lineare Separierung im Raum dargestellt werden. Die Klassenzugehörigkeit eines Zeichens ergibt sich, indem eine Übereinstimmung zwischen Merkmalvektor und einem Referenzvektor gefunden wird. Keine Übereinstimmung oder Übereinstimmung in mehreren Klassen müssen als nicht erkannt gewertet werden. Bei dieser Methode ist die Erkennung auf die Merkmale abgestützt, d. h., die erlaubten Abweichungen in der Grösse und Form der Zeichen sind durch die Toleranz dieser Messungen gegeben.

3.5 Aufbereitung und Übertragung der Zeileninformation

Die Zeichen jeder abgetasteten Belegzeile werden nach der Segmentierung einzeln klassifiziert, und das entstehende Resultat wird einem temporären Speicher (Ausgabespeicher) zugeführt. In ihm reihen sich die Zeichen, inklusiv Leerstellen, zu einer Datenzeile zusammen, die, wenn nötig, weiter verarbeitet und in ein mit dem Leser on-line geschaltetes Gerät übertragen werden kann.

In der ersten Zeile in Fig. 6 ist eine Belegzeile in Originalform dargestellt. Die Lesezonen sind dabei der Reihe nach Artikelnummern in OCR-B-Schrift, Warenbezeichnung im nicht zu lesenden (non-scan) Bereich, Einheitspreis in OCR-B-Schrift und Mengenangabe als Handschrifteintragung. Die zweite Zeile stellt die aufbereiteten Daten des

```
   (OCR-B)     (Non-Scan)        (Preis)      (Menge)

  00213701     Milchnuss          18.20        150
  00213701                        18.20     150
00213701  18.20    150
          ¦
0020012000¦00213701  18.20    150
          ¦
```

Fig. 6 Formatierung für die Datenübertragung

Ausgabespeichers dar. Dabei ist zu beachten, dass der Benützer im Beispiel durch Anwenderprogramm einen Non-Scan-Bereich definiert hat und sich im Handschriftbereich infolge breiterer Zeichen weniger Leerstellen ergeben haben. Die dritte Zeile zeigt die vom Benützer gewünschte Formatierung der übertragenen Daten, die für die Weiterverarbeitung der Daten in einer EDV-Anlage vorgesehen sind. Dabei ist in diesem Fall das erste Datenfeld 8stellig links, das zweite 7stellig rechts und das letzte 6stellig rechts justiert. Die Formatierungsangaben sind ebenfalls im Anwenderprogramm durch Programmparameter definiert. Im Programm kann man in der Regel wählen, ob die Daten von einem sogenannten Line-Header begleitet werden sollen oder nicht. Die vierte Zeile in Fig. 6 zeigt diesen Fall. Das zehnstellige Feld enthält dabei die Information für das On-Line-Gerät bezüglich Lage und Verarbeitung der betreffenden Belegzeile. Im dargestellten Fall handelt es sich um die 12. Zeile auf dem Beleg, welche der Formatierungsart (Zeilenart) Nr. 2 entsprechend gelesen und verarbeitet worden ist. Der Line-Header gibt gleichzeitig auch darüber Auskunft, ob in der betreffenden Zeile nichterkannte Zeichen aufgetreten und ob diese korrigiert sind. Dies ermöglicht eine zusätzliche Kontrolle der gelesenen Daten in der EDV-Anlage.

3.6 Belegtransport

Die zu lesenden Belege sind in der Regel gestapelt. Von den Stapeln werden sie einzeln abgezogen und oft mit hoher Geschwindigkeit, genau ausgerichtet unter der Lesestation vorbeigezogen. Dabei dürfen keine Doppelabzüge und keinerlei Beschädigungen der Belege entstehen. Ferner muss verhindert werden, dass sich die Dokumente infolge der Reibung an Transportrollen und -flächen elektrisch aufladen. Der Transport der Belege erfolgt entweder durch Reibung an Friktionsrädern oder durch Saugluft. Durch letztere kann extrem hohe Durchlaufgeschwindigkeit erreicht werden. Die heute realisierte maximale Geschwindigkeit bei Belegsortierern beträgt 6 m/s mit 6 Zoll hohen Belegen, wobei die Dicke des Papiers in den erlaubten Grenzen (0,06...0,18 mm) keine grosse Rolle spielt.

In einem universellen Seitenleser reichen die Dokumentformate z. B. von Lochkartengrösse bis zu A4-Grösse und variieren auch in der Papierdicke (z. B. 60...170 g/m²). Weiter ist es heute fast selbstverständlich, dass eine Mehrfunktionsanlage auch Journalstreifen entweder in einer festen Station oder mit Hilfe eines entsprechenden Zusatzgerätes zur Lesung aufnehmen kann.

3.7 Ausgabefächer

Abhängig von der Art des elektronischen Lesers sind zwei oder mehrere Ausgabefächer vorgesehen. Im letzteren Fall spricht man von einer Sortieranlage. Das Sortieren erfolgt in Abhängigkeit der im Leser erfassten Daten.

Seitenleser besitzen im allgemeinen zwei Ablagefächer. In dem einen kommen Belege an, welche fehlerfrei gelesen werden konnten, im zweiten solche, die aus irgendeinem Grund nochmals bearbeitet oder überprüft werden müssen.

Nach jedem gelesenen Beleg wird ein sog. Document-Header übertragen, welcher Angaben über die Lesung, das Ablagefach und die Verarbeitung des Beleges enthält. Solche sind z. B. Auftrags-Nr., Belegnummer, Ablagefach, Grund des Ablegens, Anzahl Zeilen des Beleges, Anzahl und evtl. Statistik der nicht erkannten Zeichen.

4. Besondere Eigenschaften der Lesegeräte

4.1 Datenverarbeitung im Leser

Die bisherige Ausführung zeigt, welch komplexe Aufgabe ein Leser zu erfüllen hat. Dies ist mit Hilfe eines schnellen Mikroprozessors heute möglich, bei dem die Mikroinstruktionszeit z. B. etwa 160 ns beträgt. Auf diese Schnelligkeit ist ein Universal-Leser angewiesen, insbesondere während der Merkmalextraktion der Zeichen.

Ausser der kontrollierten Abtastung, Erkennung und formatierten Übertragung muss der Leser auch spezifische Wünsche der Benützer erfüllen. Solche sind z. B.

- Prüfen eines Feldes auf numerischen Inhalt
- Addieren oder Vergleichen von Dezimalzahlen
- Durchführung einer Prüfzifferrechnung
- Laden eines Anwenderprogrammes abhängig von der Belegnummer (falls verschiedene Belege in einem Stapel vorkommen)
- Unterdrücken oder Verschieben von Leerstellen zwischen Daten
- Blocken der Daten für die Übertragung
- Zweimaliges Klassifizieren von Handschriftzeichen
- Aussteuerung des Beleges nach Auffinden eines vorgegebenen Zeichens
- Stempelung der fehlerfrei gelesenen Belege für die Archivierung
- Programmiermöglichkeit der EDV-Anlage durch die Tastatur des Lesers
- Fehlermeldungen auf Monitor des Lesers

Alle diese Vorgänge sollten im Leser durch Parameter im Anwenderprogramm abgerufen werden können.

4.2 Korrektur von nicht erkannten Zeichen

Durch Programmparameter besteht die Wahl, nicht erkannte Zeichen als Spezialzeichen, z. B. als ausgefüllte Vierecke zu übertragen, sie für die Korrektur sichtbar zu machen (Fig. 7) oder die Lesung des Beleges abzubrechen. Im zweiten Fall erscheint das nicht erkannte Originalzeichen (unter der Punktmarkierung) vergrössert auf dem Monitor zusammen mit seinen benachbarten Zeichen. Durch die Tastatur ist es dann möglich, eine beliebige Stelle zu berichtigen oder die ganze Zeile zu ignorieren. Die ausgeführte Manipulation wird dabei unabänderbar im Line-Header oder Document-Header protokolliert.

4.3 Programmiertes Lesen

Die formatierte Übertragung der Daten oder das Lesen verschiedener Schriftarten benötigt das Einlesen eines Anwenderprogrammes vor dem Lesen der Belege. Diese Absicht kann dem Leser z. B. durch das Drücken einer bestimmten Taste signalisiert werden. Die Speicherung des Anwenderprogrammes für die Dauer eines Arbeitsganges wird im RAM (Random Access Memory) des Mikroprozessors vorgenommen. Falscher Aufbau oder nicht erlaubte Programmparameter werden während des Speichervorganges auf dem Monitor als Fehler gemeldet. Es sind im allgemeinen drei Arten von Programmparameter anwendbar:

Jobspezifische Parameter:

- Festlegen einer Auftragsnummer
- Datenlänge für die Übertragung der Zeileninformation
- Anzahl Datenzeilen in einem Block
- Wahl des Ausgabe-Code ASCII oder EBCDIC
- Art der Korrektur (Monitoranzeige oder Übertragung)
- Definieren eines «End of Page»-Zeichens
- Unterdrückung von Einer-Leerstellen zwischen Daten

Feldspezifische Parameter:

- Nummer der Zeilenart (eine bestimmte Fontverteilung)
- Fontdefinition (OCR-A, OCR-B, IBM 1428, NCR-NOF, Strichmarkierung, Handschrift oder Non-Scan)
- Anzahl Zeichen oder Wahl der «nicht Formatierung» (Image-Mode)
- Horizontalbegrenzung der Felder entsprechend der Fontverteilung (Lesezonen)
- Justierung der Daten (links oder rechts)
- Spezielle Behandlung des Feldes

Parameter zur vertikalen Verteilung der Zeilenarten:

- Vertikale Begrenzung jeder spezifizierten Zeilenart

4.4 On-Line-Betrieb mit elektronischem Leser

Eine Möglichkeit für die Erfassung der Ausgabedaten eines Lesers ist der Anschluss an eine Magnetbandstation (Fig. 8), wozu normalerweise kein Datenumsetzer benötigt wird. Dabei werden die Daten in geblockter oder ungeblockter Form auf Magnetband übertragen. Im Falle eines On-Line-Anschlusses an eine Rechenanlage, ein Modem oder Peripheriegerät benötigt man dagegen ein Datenumsetzungsgerät. Vom Gerätehersteller werden sog. *P*rogrammable *C*ommunication *C*ontroller (PCC) angeboten, die mit einem Mikroprozessor ausgerüstet sind und für den richtigen

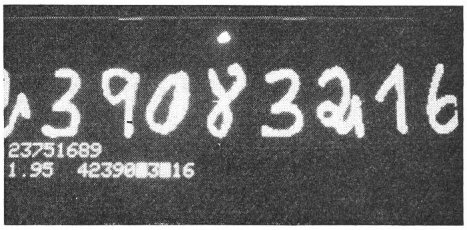

Fig. 7 Anzeige zur Korrektur nicht erkannter Zeichen
Die zwei Rückweisungszeichen können über die Eingabetastatur korrigiert werden

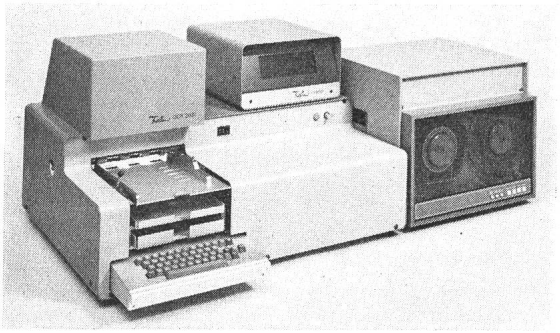

Fig. 8 Universeller Seitenleser mit Eingabetastatur, Kontrollschirm und Magnetbandstation

Anschluss an eine praktisch beliebige Anlage sorgen. Die Datenumsetzung kann dabei sowohl die Datenstruktur wie auch die erforderlichen Steuer- und Kommunikationssignale vom Eingang wie vom Ausgang her erfassen. Der PCC ermöglicht ausserdem die Vorprogrammierung einer EDV-Anlage durch die Eingabetastatur des Lesers. Der Monitor des Lesers kann dabei für Rückmeldungen von der angeschlossenen Anlage als Kontrolle benützt werden.

5. Lesesicherheit elektronischer Leser

Absolute Maßstäbe für die Reject- und Substitutionsrate eines Lesers gibt es nicht. Beide sind stark von der Druckqualität der zum Lesen angebotenen Schrift abhängig. Die erwähnten Schriftarten sind die am meisten zum optischen Lesen verwendeten, wobei OCR-A eine amerikanische und OCR-B eine europäische Normschrift ist. Sie beinhalten den ganzen alphanumerischen Satz inklusive Hilfszeichen und eignen sich infolge grosser Unterschiede in der Zeichenform vorzüglich zur maschinellen Verarbeitung. Da die Druckqualität der Zeichen nicht eindeutig definiert werden kann, formulieren Gerätehersteller die Reject- und Substitutionsrate bezogen auf ihre Testdokumente. Im Falle der OCR-Schrift ist eine in der Praxis bestätigte Angabe: 1 zurückgewiesenes Zeichen auf 10 000 bis 50 000 gelesener Testzeichen. Die besonders gefürchtete Substitution sollte dabei zehnmal kleiner sein. Abweichungen in der Druckqualität oder im Datenträger können das Resultat verschlechtern.

In der Handschrifterkennung sind meistens nur die zehn Ziffern und einige Hilfszeichen aus der Buchstabenreihe, wie *E* und *X*, im Gebrauch. Diese müssen in der Regel zudem nach den vom Gerätehersteller empfohlenen Vorschriften geschrieben werden. Zur Übung der restriktiven Form werden oft auch Kurse durchgeführt. Die in der Praxis erreichbare Rejectrate beträgt heute ca. 0,5...2 % der gelesenen Handschriftzeichen. Auch hier sollten Substitutionen zehnmal seltener auftreten als zurückgewiesene Zeichen.

Literatur

[1] *F. Kis:* Automatische Zeichenerkennung. Techn. Rdsch. 63(1971)8, S. 17...21, Nr. 9, S. 27...31.

[2] *N. J. Nilsson:* Learning machines. New York, McGraw-Hill, 1965.

[3] *L. Uhr:* Pattern recognition: Theory, experiment, computer simulations and dynamic models of form perception and discovery. New York a. o., John Wiley, 1966.

Adresse des Autors

Ferenc Kis, Dr. sc. techn., Feller AG, Forschungslabor, Geissacherstrasse 6, 8126 Zumikon.