

Probleme der Mustererkennung bei der automatischen Briefverteilung = Problèmes de la reconnaissance de critères de référence dans le tri automatique des lettres

Autor(en): **Meyer-Brötz, Günter / Schürmann, Jürgen**

Objektyp: **Article**

Zeitschrift: **Technische Mitteilungen / Schweizerische Post-, Telefon- und Telegrafienbetriebe = Bulletin technique / Entreprise des postes, téléphones et télégraphes suisses = Bollettino tecnico / Azienda delle poste, dei telefoni e dei telegrafi svizzeri**

Band (Jahr): **53 (1975)**

Heft 2

PDF erstellt am: **22.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-875586>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern.

Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden.

Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Probleme der Mustererkennung bei der automatischen Briefverteilung¹

Problèmes de la reconnaissance de critères de référence dans le tri automatique des lettres¹

Günter MEYER-BRÖTZ und Jürgen SCHÜRMAN, Ulm/Donau

656.816.31 : 656.851.433.5 : 681.327.12

Zusammenfassung. In grossen Postämtern werden halbautomatische Briefverteilanlagen benutzt, bei denen Operateure (innen) bei der abgehenden Post die Postleitzahl, bei der ankommenden Post den Strassennamen in Codeform aufbringen. Ausgehend von den Anforderungen automatischer Codierplätze beschreiben die Autoren ein von AEG-Telefunken für die Deutsche Bundespost entwickeltes Funktionsmuster zum Lesen der handelsüblichen Druckschriften (oder etwa 70% der Briefsendungen). Sie treten auf die dabei zu lösenden Probleme der Mustererkennung und deren Lösung ein. Die Klassifizierung der Einzelzeichen wird nach einem Prognoseverfahren vorgenommen, für Adresseninterpretation wird ein «Orts- und Strassen-Wörterbuch» mit Ähnlichkeitsvergleich eingesetzt, bei Mehrdeutigkeiten dient die Postleitzahl zur Unterscheidung.

Résumé. Dans les grands offices postaux, on utilise souvent des installations automatiques de tri des lettres où des opérateurs (opératrices) appliquent sous forme de code le numéro postal d'acheminement sur le courrier partant et le nom des rues sur le courrier arrivant. Se fondant sur les exigences liées aux places de codage automatique, les auteurs décrivent un prototype mis au point par AEG-Telefunken pour les Postes fédérales allemandes, qui sert à lire des caractères d'imprimerie usuels (ce qui correspond à environ 70% des envois). Ils exposent les problèmes touchant à la reconnaissance de critères de référence et aux solutions apportées. La classification des caractères se fait selon une méthode de pronostics, tandis qu'un lexique des localités et des rues sert à interpréter les adresses d'après un processus de comparaison quant à la similitude. En cas d'ambiguïté, le numéro postal d'acheminement sert de critère de différenciation.

Problemi riguardanti il riconoscimento di campioni nella spartizione automatica delle lettere

Riassunto. Negli importanti centri postali si impiegano impianti semiautomatici per la spartizione delle lettere; sulle lettere in partenza operatori (-trici) registrano numeri postali d'avviamento, su quelle in arrivo codici corrispondenti alle denominazioni stradali. Gli autori, prendendo come spunto le esigenze d'un posto di codificazione automatico, descrivono un prototipo funzionale sviluppato dalla AEG-Telefunken per la posta della Repubblica Federale Tedesca, in grado di leggere caratteri ordinari stampati (ossia circa il 70% della corrispondenza). Essi espongono i problemi e le soluzioni adottate in merito al riconoscimento di campioni. La classificazione dei singoli caratteri si basa sul procedimento a diagnosi; per l'interpretazione degli indirizzi si impiega un «vocabolario delle località e delle strade con comparazione di rassomiglianza», e, in caso di possibile equivoco, è il numero postale d'avviamento che serve quale criterio di discriminazione.

1 Stand der Technik bei der halbautomatischen Briefverteilung

Bei der rein manuellen Verteilung kann eine Arbeitskraft maximal 1200 Sendungen in der Stunde in mehreren Arbeitsgängen sortieren. In grossen Postämtern werden deshalb zunehmend maschinelle Briefverteilanlagen benutzt, die im wesentlichen die folgenden Arbeitsabläufe umfassen (Fig. 1):

- Formattrennen
- Aufstellen und Stempeln
- Codieren
- Verteilen

Die Formattrennanlage scheidet zu grosse, zu kleine, zu dicke und zu lange Sendungen aus. Die verbleibenden Sendungen werden aufgestellt, indem die Lage der Briefmarke ermittelt wird, dann gestempelt und einem Handcodierplatz entweder direkt oder indirekt (über einen Fernsehmonitor) zugeführt. Hier lesen Operateure bei der abgehenden Post die Postleitzahl, bei der ankommenden Post den Strassennamen und geben über eine Tastatur die Postleitzahl oder einen Extraktionscode – es kann beispielsweise der erste, dritte, vierte und letzte Buchstabe des Strassennamens sein – in einen Rechner ein. Dieser liefert die Druckinformation für einen Codedruker. Der Code wird in Phosphoreszenz- oder Magnetfarbe in Form

¹ Leicht gekürzte Fassung eines anlässlich des 10. Technischen Presse-Colloquiums von AEG-Telefunken (17./18. Oktober 1974) gehaltenen Vortrages.

1 Etat de la technique dans le tri semi-automatique des lettres

Par la méthode de tri manuel pur, un agent peut trier 1200 envois à l'heure au maximum, en plusieurs opérations. De ce fait, les grands offices postaux sont de plus en plus souvent équipés d'installations de tri mécanique assurant pour l'essentiel les opérations suivantes (fig. 1):

- séparer les formats
- redresser et oblitérer
- coder
- trier

La machine à séparer les formats élimine les envois trop grands, trop petits, trop épais et trop longs. Les lettres restantes sont redressées – par détection de l'emplacement du timbre – puis oblitérées et dirigées sur une place de codage, soit directement, soit indirectement (sous surveillance d'un moniteur de télévision). A ce niveau, les opérateurs lisent le numéro postal d'acheminement pour le courrier partant et le nom des rues pour celui qui arrive et introduisent dans une calculatrice, par le biais d'un clavier, le numéro postal d'acheminement ou un code d'extraction représentant, par exemple, la première, la deuxième, la troisième, la quatrième et la dernière lettre du nom de la rue. La calculatrice fournit l'instruction nécessaire à l'imprimeur de code. Les signes-codes sont imprimés sous forme

¹ Version légèrement abrégée d'un exposé présenté lors du 10^e colloque de la presse technique organisé par AEG-Telefunken (17/18 octobre 1974)

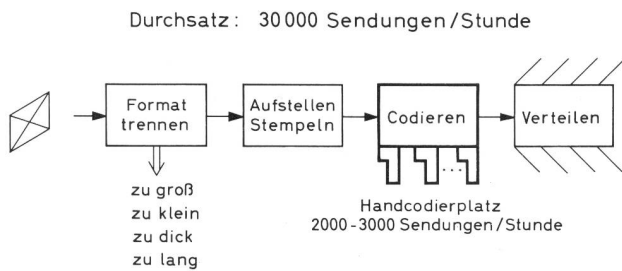


Fig. 1
Schema einer halbautomatischen Briefverteilanlage – Schéma d'une installation semi-automatique de tri des lettres

Durchsatz: 30 000 Sendungen/Stunde – Capacité de traitement: 30 000 envois à l'heure
 Format trennen – Séparer les formats
 Zu groß – Trop grand
 Zu klein – Trop petit
 Zu dick – Trop épais
 Zu lang – Trop long
 Aufstellen – Redresser
 Stempeln – Oblitérer
 Codieren – Coder
 Verteilen – Distribuer
 Handcodierplatz 2000...3000 Sendungen/Stunde – Place de codage manuel 2000...3000 lettres/heure

eines Strichcodes aufgedruckt, der maschinell leicht lesbar ist und die abschliessenden Verteilvorgänge steuert. Die grösste von AEG-Telefunken gelieferte und installierte Feinverteilmaschine ermöglicht eine Verteilung in bis zu 480 Fächer. Moderne Verteilmaschinen können etwa 30 000 Sendungen/Stunde verarbeiten. Da an einem Handcodierplatz dagegen nur 2000 bis 3000 Sendungen/Stunde codiert werden können, müssen je Verteilmaschine 10...15 Codierplätze vorgesehen werden.

2 Anforderungen an den automatischen Codierplatz

Es geht um die Aufgabe, die lange Reihe von Handcodierplätzen soweit wie möglich durch *einen* automatischen Anschriftenleser zu ersetzen, der sich in ein historisch gewachsenes Organisationsschema einfügen muss. Die vielfältigen Probleme und Schwierigkeiten dieser Aufgabe sind dem Nichtfachmann kaum bewusst. Im Laufe der Jahrzehnte haben sich Gewohnheiten eingebürgert, nach denen Postkunden ihre Sendungen beschriften. Diese Routine ist Allgemeingut aller Postbenützer und auch den menschlichen Arbeitskräften unmittelbar vertraut, die bisher von Hand sortierten oder codierten. Hierbei war es kein Nachteil, wenn die bestehenden Vorschriften für das Beschriften nicht ausdrücklich und verbindlich formuliert waren und beim Sortieren unreflektiert beachtet wurden.

Zum Zeitpunkt jedoch, an dem die Funktion des Briefsortierens vollständig einer Maschine übertragen werden soll, wird es unerlässlich, diese Vorschriften klar zu fassen. Sie bilden einen wesentlichen Bestandteil der Aufgabenbeschreibung, nach der diese Maschine zu bauen ist. Bevor sie gebaut wird, sollte genau der Zweck bekannt sein, dem sie dienen soll. Es ist höchst unbefriedigend, eine Maschine zunächst zu entwickeln und erst hinterher ihren Zweck nach den Leistungen zu definieren, die sie erbringt. Beim Bearbeiten des Projekts eines automatischen Anschriftenlesers ergeben sich eine Reihe von Punkten, an denen zur Abgrenzung der Aufgabenstellung strikte Festlegungen getroffen werden müssen. Maschinen, die lesen, sind im Prinzip nicht neu. Die Belege aber, die in Beleglesemaschinen verarbeitet werden, sind durch Vorgabe

de bâtonnets phosphorescents ou magnétiques pouvant être lus facilement par voie automatique et utilisés pour la commande des processus de tri. La plus grande machine de tri détaillé, fournie et installée jusqu'ici par AEG-Telefunken permet un tri aboutissant à 480 cases. Les machines de tri modernes sont capables de traiter quelque 30 000 envois à l'heure. Or, une place de codage manuel ne permettant de coder que 2000 à 3000 lettres à l'heure, il faut en prévoir 10 à 15 par installation de tri.

2 Exigences auxquelles doit satisfaire une place de codage automatique

Le problème à résoudre est de remplacer si possible par *un seul* lecteur automatique d'adresses la longue rangée de places de codage manuel, ensemble qui devra s'intégrer dans un schéma d'organisation dicté par la routine. Un profane a peine à imaginer la multiplicité et la difficulté des problèmes que pose cette tâche. Au cours des décennies, les clients de la poste ont pris l'habitude de libeller les adresses des envois d'une certaine manière. Cette routine est profondément ancrée dans le public et les agents qui trient et codent le courrier à la main en connaissent toutes les particularités. De ce fait, aucun inconvénient n'était à craindre du fait que les prescriptions sur le libellé des adresses ne sont ni précises ni de caractère obligatoire, puisque les agents de tri étaient capables d'interpréter l'information par réflexe.

Toutefois, dès l'instant où le tri des lettres doit être entièrement confié à une machine, il devient indispensable de formuler ces prescriptions de manière précise. Elles constituent une partie intégrante du cahier des charges de la machine. Avant de la construire, il faut définir clairement le but auquel elle servira. Vouloir développer une installation et ne préciser son but qu'après, au vu de ses performances, est très peu satisfaisant. Lors de l'élaboration du projet d'un lecteur automatique d'adresses, il faut donc définir tout d'abord strictement une série de points, en vue de circonscrire la tâche. Les systèmes de lecture optique ne sont pas nouveaux quant à leur principe. Toutefois, les documents que traitent les lecteurs ad hoc sont soumis à de telles limitations – format, emplacement de l'information à lire, type d'écriture, qualité de l'impression, qualité du papier, etc. – qu'on ne saurait imposer leur emploi à la plupart des clients de la poste.

Certaines Administrations postales, notamment celles de l'Union soviétique et du Japon, ont introduit des restrictions sévères. Dans ces régions, l'utilisateur doit inscrire sur l'enveloppe un code purement numérique, en un endroit, exactement déterminé, selon des règles d'écriture précises, de sorte que les enveloppes à panneau transparent, par exemple, ne peuvent pas être utilisées. L'autre extrême est le système «Advanced Optical Character Reader», développé aux Etats-Unis, où l'intéressé est pratiquement libre de libeller l'adresse où bon lui semble. L'Administration postale des Etats-Unis ayant investi près de 50 millions de dollars de 1955...1972 pour le développement de systèmes de reconnaissance de caractères capables de lire et de trier les envois postaux, on assiste actuellement à l'essai de la deuxième génération de lecteurs d'adresses.

Un prototype de lecteur d'adresses développé par les établissements AEG-Telefunken sera mis à l'essai par les Postes fédérales allemandes à Berlin; toutefois, la machine ne travaillera pas encore à son rendement définitif. Elle

des Formats, des Ortes der zu lesenden Information, der Schriftart, der Druckqualität, der Papierqualität usw. so starken Einschränkungen unterworfen, wie sie meistens Postkunden nicht zugemutet werden können. Einzelne Postdienste, etwa in der Sowjetunion oder Japan, haben derartige extreme Einschränkungen eingeführt. Dort muss der Postbenützer einen rein numerischen Code auf dem Briefumschlag an genau bestimmter Stelle nach genau vorgeschriebenen Schreibregeln aufbringen, so dass zum Beispiel Fensterkuverts nicht verwendet werden können. Das andere Extrem, bei dem dem Postkunden fast keine Restriktionen in der Anordnung der Anschrift auferlegt werden, wird mit dem Advanced Optical Character Reader System in den USA verfolgt. Dort geht jetzt die zweite Generation von Anschriftenlesemaschinen in die Erprobung, nachdem die amerikanische Postverwaltung in den Jahren 1955 bis 1972 nahezu 50 Millionen Dollar in die Entwicklung von optischen Zeichenerkennungssystemen für das Lesen und Sortieren von Postsendungen investiert hat.

Ein noch nicht mit der vollen, endgültigen Leistung arbeitendes Funktionsmuster des von AEG-Telefunken entwickelten Anschriftenlesers wird Anfang nächsten Jahres bei der Deutschen Bundespost in Berlin in Erprobung gehen. Es verarbeitet alle Sendungen im Standardformat – Briefe, Fensterkuverts, Postkarten – deren Adresse in einer der handelsüblichen Druckschriften – ausser Kursiv- und Frakturschriften – geschrieben ist. Dadurch sollen etwa 70% aller Sendungen erfasst werden. Solche mit nicht maschinell lesbaren Beschriftungen, wie etwa Handschrift, werden automatisch ausgesondert. Gelesen werden muss der volle alphanumerische Text der Abgangs- oder Eingangsinformation. Die Leistungsfähigkeit des Geräts wird 60 000 Sendungen/Stunde betragen, so dass zwei Verteilanlagen bedient werden können. Es erschien der Bundespost nach vielen Diskussionen aber zweckmässig, aus Gründen der Wirtschaftlichkeit des automatischen Anschriftenlesers einige Vorschriften für die Postbenützer ins Auge zu fassen, die sich auch an internationale Normen anlehnen:

- Der Anschriftenblock muss sich innerhalb einer Erwartungszone befinden (Fig. 2)
- Zusatzinformationen, wie Reklameaufdrucke, sind nur links oder oberhalb des Anschriftenblockes zugelassen
- Die Anschrift muss linksbündig geschrieben sein mit der Anordnung von oben nach unten:
 - Empfänger
 - Eingangsinformation (Strassenname, Hausnummer) mit Zustell- oder Abholungsangaben
 - Abgangsinformation (Bestimmungsort), einzeilig mit postamtlichen Leitangaben (Fig. 2).

Vor allem die letzte Vorschrift begrenzt den Aufwand für den Anschriftenleser stark, da eindeutig festgelegt ist, dass für die Abgangssortierung nur die unterste Zeile, für die Eingangssortierung die zweitunterste Zeile der Anschrift gelesen werden muss. Trotz dieser Einschränkungen liegt aber eine komplexe Aufgabenstellung der Mustererkennung vor, die erst jetzt nach den Fortschritten der Computertechnik und -technologie in Angriff genommen werden konnte.

3 Ausgewählte Probleme der Mustererkennung

Der maschinelle Anschriftenleser, der den Handcodierplatz ersetzen soll und sich daher organisch in die gesamte



Fig. 2
Briefvorderseite mit eingeschränkter Anordnung der Anschrift – Recto d'une lettre avec possibilité restreinte de disposition de l'adresse

Erwartungszone – Zone supputée
Anschriftenblock – Bloc de l'adresse

traite tous les envois de format standard – lettres, enveloppes à fenêtre, cartes postales – dont l'adresse est imprimée en caractères d'imprimerie courants, sauf les écritures cursive et gothique. Ce faisant, elle sera capable d'interpréter quelque 70% des adresses; celles qui ne sont pas lisibles mécaniquement, par exemple les adresses écrites à la main, sont éliminées automatiquement.

La machine doit lire le texte alphanumérique complet de l'information d'arrivée et de départ. Avec une capacité de lecture de 60 000 envois à l'heure, elle peut desservir deux installations de tri. Après de nombreuses discussions, les Postes fédérales ont jugé rationnel, pour des raisons de rendement du lecteur automatique d'adresses, de prévoir un certain nombre d'instructions à l'intention des usagers, instructions qui se fondent d'ailleurs sur des normes internationales:

- le bloc de l'adresse doit se trouver à l'intérieur d'une zone supputée (fig. 2)
- les adjonctions, telles que les empreintes-réclame, ne sont admises qu'à gauche ou au-dessus du bloc de l'adresse
- il faut aligner le premier mot de chaque ligne de l'adresse sur la gauche et la libeller de haut en bas dans l'ordre suivant:
 - destinataire
 - information d'arrivée (nom de la rue, numéro de la maison) avec indications relatives à la distribution ou au retrait
 - information de départ (lieu de destination), en une ligne, avec indications d'acheminement de la poste (fig. 2).

Le lecteur d'adresses peut surtout être simplifié en raison de la dernière disposition qui fixe clairement que seule la dernière ligne de l'adresse doit être lue pour le tri au départ et l'avant-dernière pour le tri à l'arrivée. Malgré cette limitation, la reconnaissance de critères de référence représente un problème complexe que seuls les progrès réalisés dans la technologie et les techniques des ordinateurs ont permis d'attaquer.

3 Problèmes typiques de la reconnaissance de critères de référence

Le lecteur d'adresses mécanique appelé à remplacer la place de codage manuel et à s'intégrer harmonieusement dans l'ensemble de l'organisation de tri assure les fonctions suivantes:

- lecture optique
- recherche de la ligne

Verteilerorganisation einfügen muss, führt folgende wesentlichen Funktionen aus (Fig. 3):

- Abtasten
- Zeilen suchen
- Bildvorbereitungs-Prozeduren durchführen
- Vereinzelte Schriftzeichen erkennen
- Adresse interpretieren
- Codieren

Jede einzelne Funktion stellte eine schwierige Aufgabe dar. Wie die Probleme gelöst wurden, wird im folgenden skizziert.

31 Abtasten des Anschriftenblocks und Suchen der relevanten Anschriftenzeile

Die Postsendung bewegt sich kontinuierlich mit einer Geschwindigkeit von etwa 4 m/s von links nach rechts unter einer vertikalen Spalte von 512 integrierten Photodioden vorbei und wird dabei also scheinbar von rechts nach links abgetastet. Die elektrischen Ausgänge der Photodioden werden zu diskreten Zeitpunkten abgefragt, so dass insgesamt eine flächenhafte Quantisierung, das heisst eine Rasterung des Vorlagenfeldes in $120 \mu \times 120 \mu$ grosse Bildpunkte, entsteht. Das entspricht einer Rasterpunktdichte von knapp 70 Punkten/mm². Die analogen elektrischen Signale der Photodioden werden mit einer durch die Helligkeit des Untergrundes und den Kontrast des Aufdruckes gesteuerten Quantisierungsschwelle binär quantisiert, wobei einem hellen Bildpunkt der Zahlenwert 0, einem dunklen der Wert 1 zugeordnet wird. Damit ist eine zweistufige Amplitudenquantisierung durchgeführt. Die Abtastpunkte der Erwartungszone werden in 512 Schieberegister von je 1024 bit Länge solange fortlaufend eingelesen, bis die linke Vorderkante des Anschriftenblocks dadurch gefunden ist, dass die Zahl der schwarzen Bildpunkte in einer Spalte abrupt abnimmt (linksbündige Anschrift!). Damit steht die Anschrift in einem Speicher von 512 kbit zur weiteren Verarbeitung zur Verfügung.

Die Zeilenstruktur der Anschrift spiegelt sich in Schwärzungsbalken wieder, deren horizontale Ausdehnung sehr

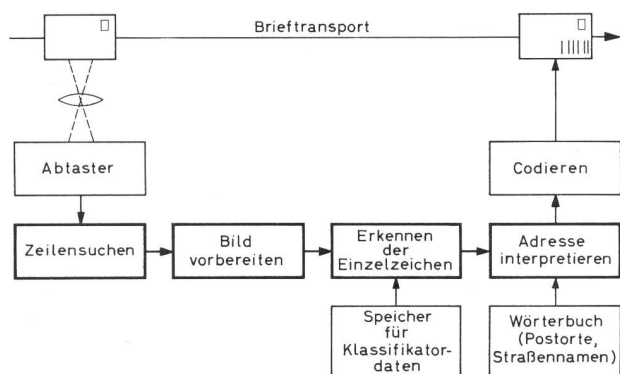


Fig. 3 Funktionsgruppen des automatischen Codierplatzes - Unités de la place de codage automatique

- Brieftransport - Transport des lettres
- Abtaster - Lecteur optique
- Zeilensuchen - Recherche de la ligne
- Bild vorbereiten - Préparation de l'image
- Erkennen der Einzelzeichen - Reconnaissance des caractères
- Speicher für Klassifikatordaten - Mémoire des données du classeur
- Adresse interpretieren - Interprétation de l'adresse
- Wörterbuch (Postorte, Strassennamen) - Lexique (localités postales, noms des rues)
- Codieren - Coder

- préparation de l'image
- reconnaissance de certains caractères
- interprétation de l'adresse
- codage.

Chacune de ces opérations représentait un problème difficile à résoudre. Les solutions apportées seront décrites dans les pages qui suivent.

31 Lecture optique du bloc de l'adresse et recherche de sa ligne déterminante

L'envoi postal défile continuellement de gauche à droite à une vitesse de 4 m/s devant une rangée latérale de 512 photodiodes intégrées et semble donc exploré de droite à gauche. Les sorties électriques des photodiodes sont interrogées à des moments discrets, si bien qu'il en résulte une quantification de la surface, c'est-à-dire une décomposition du sujet exploré en une trame formée de points de $120 \mu \times 120 \mu$, répartis à une densité de quelque 70 points par mm². Les signaux électriques analogiques des photodiodes sont quantifiés en mode binaire selon un seuil de quantification déterminé par la luminosité du support et le contraste de l'impression, la valeur «0» étant assignée à un point clair et la valeur «1» à un point sombre. Il en résulte une quantification d'amplitude à deux niveaux. Une fois lus, les points de la zone supputée sont introduits sériellement dans 512 registres à décalage d'une capacité de 1024 bits chacun, jusqu'à ce que la marge de gauche du bloc de l'adresse soit trouvée, ce qui se traduit par une brusque diminution des points noirs explorés (mots de chaque ligne de l'adresse alignés sur la gauche). Ainsi, l'adresse est inscrite dans une mémoire de 512 kbits où l'on peut en disposer pour le traitement ultérieur.

La structure de la ligne est restituée électriquement par des plages de noir dont l'étendue horizontale est nettement plus grande que l'étendue verticale. Si les lignes étaient toujours placées horizontalement, il serait relativement facile d'en déterminer les coordonnées verticales par projection horizontale de la répartition de noir dans le bloc de l'adresse (fig. 4). Toutefois, comme il faut s'attendre que les lignes soient souvent disposées passablement en biais, ce principe doit être amené à un plus haut degré de perfection. Le résultat de cette technique que nous nous abstiendrons de décrire en détail est illustré par la figure 5:



Fig. 4 Suchen der relevanten Anschriftzeile durch Horizontalprojektion und Maximumsuche in der Randverteilung - Recherche de la ligne déterminante de l'adresse par projection horizontale et détection du maximum dans la distribution marginale de l'information

Randverteilung - Distribution marginale de l'information

gross ist im Verhältnis zu ihrer vertikalen. Lagen die Zeilen immer genau horizontal, könnte man die vertikalen Ortskoordinaten der Zeilen verhältnismässig einfach durch eine Horizontalprojektion der Schwärzungsverteilung des Anschriftenblocks ermitteln (Fig. 4). Da mit einer erheblichen Schräglage der Zeilen gerechnet werden muss, muss dieses Prinzip sehr verfeinert werden. Das Resultat dieser Technik, die hier nicht näher beschrieben werden kann, zeigt an einem Beispiel Fig. 5: die unterste Zeile aus dem anormal schräg liegenden Anschriftenblock wurde aufgefunden und durch Scheren gerade gerichtet. Als Ergebnis der Zeilensuchoperation steht die für die Verteilung massgebende Anschriftenzeile in einem Speicher von $64 \cdot 1024 = 64 \text{ kbit}$.

32 Bildvorbereitung für die automatische Zeichenerkennung

Aufgabe des Bildvorbereitungssystems ist es, aus der Anschriftenzeile die Einzelzeichen herauszupräparieren und von dabei anfallenden Störungen zu befreien. Darüber hinaus ist es Aufgabe dieses Arbeitsgangs, solche Zeichenvariationen, die verabredungsgemäss keinen Einfluss auf die Bedeutung der Zeichen haben sollen, etwa Lagevariationen oder auch Grössenvariationen, auszuschalten.

321 Segmentierung

Aufgabe der Segmentierung ist es, aus dem fortlaufenden Rasterbild der Anschriftenzeile die Rasterbilder der Einzelschriftzeichen herauszuschneiden. Wenn die Zeichenbilder in sich zusammenhängen und vollständig von weiss umgeben sind, ist das trivial. In der Praxis muss aber mit verklebten und zerrissenen Zeichen gerechnet werden. In diesen Fällen muss der Segmentieralgorithmus so weit wie möglich von Vorkenntnissen über die Zeichenteilung und die Zeichenabmessungen Gebrauch machen.

Bei Schriften mit variabler Teilung – vor allem Buchdruck – liegt im allgemeinen ein sauberer Druck vor. Sie lassen sich im wesentlichen durch Benutzung von Weisswegtechniken segmentieren. Hierbei wird ein Weisskanal definierter Breite eingeführt, beispielsweise von 5 Bildelementen (Fig. 6). Ein Weissweg gilt als gefunden, wenn innerhalb des Weisswegkanals ein ausschliesslich über weisse Bildelemente führender Weg vom oberen zum unteren Bildfeldrand existiert. Als Schnittlinie wird in Fig. 6 die linke Kante des innerhalb des Weisswegkanals am weitesten links verlaufenden Weissweges verwendet. Dadurch wird das Zeichenschwarz rechts herausmodelliert.

Schriften mit fester Teilung haben eine weitaus grössere Bedeutung als solche mit variabler Teilung. Während bei Schreibmaschinenschriften mit den Teilungen 10, 11 und 12 Zeichen/Zoll gerechnet wird, haben Schnelldrucker grundsätzlich 10 Zeichen/Zoll. Zum Segmentieren dieser

Helmut Schmider
Hauptstraße 26
4901 Hiddenhausen-Schweichel-Bermbeck

ZEILE 1 19.06.1974
4901 Hiddenhausen-Schweichel-Bermbeck

Fig. 5 Auffinden und Scheren schrägliegender Zeilen – Détection et redressement des lignes placées en biais

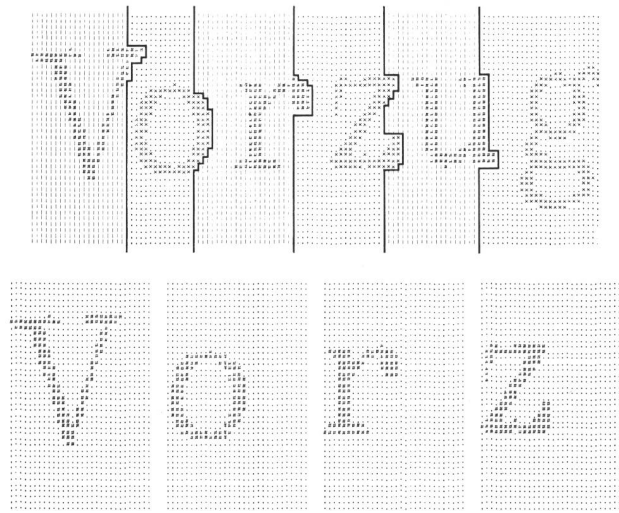


Fig. 6 Weissweg-Segmentierung – Segmentation des zones claires

La ligne inférieure du bloc de l'adresse, anormalement disposée en biais, a été détectée et redressée. Après l'opération de détection de ligne, les données servant au tri, contenues dans la ligne déterminante de l'adresse, sont inscrites dans une mémoire de $64 \cdot 1025 = 64 \text{ kbits}$.

32 Préparation de l'image pour la reconnaissance automatique de caractères

La tâche du système de préparation de l'image est de préparer chaque caractère, transformé en signal, de la ligne de l'adresse et de le débarrasser des perturbations qui pourraient l'affecter. En plus de cela, ce processus élimine les paramètres qui, par définition, ne doivent pas avoir d'influence sur la signification du caractère, à savoir les variations de position et de dimension.

321 Segmentation

Le processus de segmentation consiste à découper l'image tramée continue de la ligne de l'adresse de manière à isoler l'image tramée de chaque caractère. Lorsque les caractères se suivent régulièrement et sont entièrement entourés d'une zone blanche, cette opération est très simple. Or, en pratique, les caractères sont souvent accolés ou disjoints. En pareils cas, l'algorithme de segmentation doit tenir compte dans une mesure aussi large que possible de toutes les règles connues sur la séparation et les dimensions des caractères.

En typographie surtout, où l'espacement entre caractères est variable, l'impression est généralement irréprochable. Les caractères peuvent être séparés par la méthode des zones claires. Ce système repose sur le principe de canaux d'éléments blancs d'une largeur définie – par exemple 5 éléments de trame (fig. 6). Par définition, un canal d'éléments blancs est trouvé lorsqu'un chemin ne comportant que des éléments blancs conduit du haut en bas de l'image à l'intérieur d'une zone claire. Comme le montre la figure 6, la coupure se fait le long de la limite extrême-gauche du canal d'éléments blancs situé le plus à gauche de la zone claire. Cette césure est marquée par un trait qui épouse les contours des éléments noirs du caractère.

Les caractères à espacement fixe revêtent une importance beaucoup plus grande que ceux dont l'espacement est variable. En principe, les imprimantes rapides alignent dix caractères par pouce, alors que les machines à écrire en impriment 10, 11 ou 12 sur la même longueur. La seg-

Schriften macht man davon Gebrauch, dass eine feste Teilung vorliegt, die durch Messung bestimmt werden muss. Die feste Teilung prägt dem Bild der Zeile eine periodische Struktur auf, die sich weniger in den Zeichen als in den Zeichenlücken bemerkbar macht. Ein Zeilenausschnitt wird also daraufhin untersucht, wieweit eine Zeichenlückenstruktur gegebener Periodizität Bestandteil des Bildausschnittes ist. Die gesuchte Struktur spiegelt sich in der Randverteilung $r(i)$ wider, die man durch Vertikalprojektion der Anschriftenzeile erhöht (Fig. 7). Zum Segmentieren wird die so erhaltene Zeilenlückenstruktur mit einer idealen periodischen Struktur, einer Kammfunktion $k(i)$, verglichen. Dazu wird die Korrelationsfunktion

$$\Phi(j) = \sum k(i) r(j+i)$$

gebildet, die ihre lokalen Maxima an der Stelle der Zeichenlücken hat (Fig. 7). Da das Segmentiersystem feste und variable Zeichenteilungen verarbeiten können soll, müssen die beiden beschriebenen Segmentiertechniken miteinander kombiniert und einander untergeordnet werden. Durch die Weisswegsuche wird auch Anfang und Ende einer Zeile gefunden.

322 Bildreinigung

Für die nachfolgenden Operationen ist es oft zweckmässig, die Zeichen vom Hintergrundrauschen zu befreien, das bei kontrastarmen Zeichen auftritt. Ein Beispiel hierfür ist Teil (a) von Figur 7. Verwendet werden dafür lokale Operationen, die mehr oder minder isoliert stehende schwarze Bildpunkte, die von weissen Bildpunkten umgeben sind, eliminieren. Je nachdem, eine wie grosse weisse Umgebung erforderlich ist, um einen schwarzen Bildpunkt als Störung zu definieren und damit auszulöschen, wirken die Bildreinigungsprozeduren unterschiedlich aggressiv. Sie können schliesslich dazu führen, dass nicht nur störende Punkte im Hintergrund des Zeichens, sondern bei dünnem Druck auch die Zeichen selbst «ausradiert» werden, wie die zweite Zeile in Figur 8 zeigt. Es ist deshalb erforderlich, die Stärke der Bildreinigung in Abhängigkeit von der Strichdicke der Zeichen zu steuern, die aus der gesamten Zeichenfläche und dem Umfang der Zeichen näherungsweise ermittelt werden kann. Den Erfolg dieser

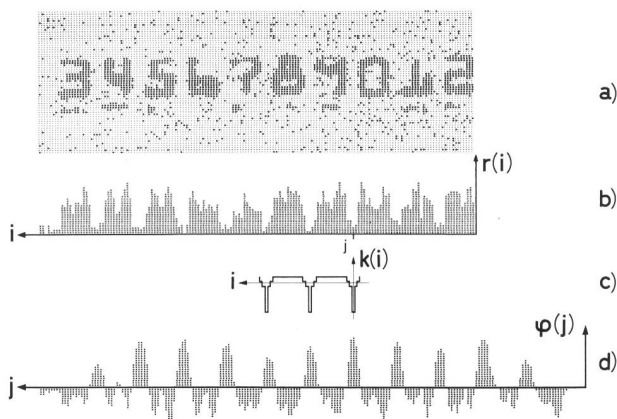


Fig. 7
Kamm-Segmentierung – Segmentierung selon une fonction en dents de peigne

- a) Rasterbild – Image tramée
- b) Randverteilung – Distribution marginale de l'information
- c) Kammfunktion – Fonction en dents de peigne
- d) Korrelationsfunktion – Fonction de corrélation

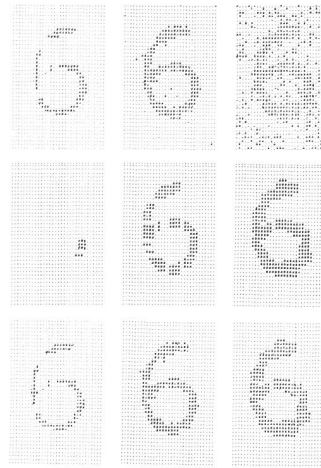


Fig. 8
Bildreinigung – Eputation de l'image

segmentation de ces caractères est possible du fait qu'un espacement fixe pouvant être mesuré les sépare. Cet écart fixe confère à la ligne une structure périodique qui ressort plus de l'espacement que des caractères eux-mêmes. On recherche Jonc, pour un fragment donné de ligne, dans quelle mesure une structure d'espacement entre caractères de périodicité définie fait partie du secteur d'image analysé. La structure recherchée apparaît dans la distribution marginale de l'information $r(i)$ que l'on accentue par projection verticale de la ligne de l'adresse (fig. 7). La segmentation se fait par comparaison entre la structure d'intervalles de ligne ainsi obtenue et une structure périodique idéale, soit une fonction en dents de peigne $k(i)$. En définitive, on obtient une fonction de corrélation

$$\Phi(j) = \sum k(i) r(j+i)$$

ayant des maximums locaux aux endroits où se trouvent les espacements entre caractères (fig. 7). Le système de segmentation devant pouvoir traiter des caractères à espacement fixe et variable, il est nécessaire de combiner les deux techniques de segmentation décrites et de leur assigner une hiérarchie. Par recherche de canaux d'éléments blancs, on détermine aussi le début et la fin d'une ligne.

322 Eputation de l'image

Pour les opérations qui suivent, il est souvent rationnel de débarrasser les signaux obtenus du bruit de fond qui les affecte lorsque le contraste est faible. La partie a) de la figure 7 illustre ce phénomène. Le processus consiste à éliminer les points noirs plus ou moins isolés dans leur environnement de points blancs. Suivant l'étendue de l'environnement blanc d'un point noir – critère servant à définir ce dernier en tant que perturbation et à motiver son effacement – l'effet d'épuration est plus ou moins prononcé. Il peut finalement en résulter non seulement un effacement des points noirs gênants de l'environnement du signal, mais aussi la suppression du signal lui-même, s'il résulte d'une impression trop fine, comme on le voit à la deuxième ligne de la figure 8. De ce fait, il est nécessaire de commander le processus d'épuration en fonction de l'épaisseur des contours des caractères, critère qui peut être approximativement défini par la surface totale qu'ils occupent et par leur nombre. La dernière ligne de la figure 8 montre dans quelle mesure l'épuration de l'image en fonction de l'épaisseur des traits peut être menée à bien.

Massnahme, die Bildreinigung durch die Strichdicke zu steuern, zeigt die unterste Zeile von Figur 8.

323 Zentrierung

Da Verschiebungen des Zeichens innerhalb des Rasterfeldes keinen Einfluss auf die Bedeutung der Zeichen haben, ist es zweckmässig, dafür zu sorgen, dass ein- und dieselbe Konfiguration von schwarzen Bildelementen vor weissem Hintergrund nicht in mehreren gegeneinander verschobenen Positionen im Rasterbild vorkommt. Die wichtigsten Verfahren, eine Lagenormierung zu erreichen, sind die Zentrierung auf eine horizontale und vertikale Zeichenkarte oder auf den Flächenschwerpunkt des Zeichens. Es ist wünschenswert, dass nicht schon geringfügige Änderungen der Schwarzweiss-Konfiguration zu einer deutlichen Änderung der Zentrierposition führen. In dieser Hinsicht sind die Schwerpunkt-Zentrierverfahren den Kanten-Zentrierverfahren überlegen.

324 Grössennormierung

Zu den Zeichenvariationen, die verabredungsgemäss keinen Einfluss auf die Bedeutung haben sollen, gehört die Zeichengrösse. Das Prinzip der Grössennormierung ist es, aus einer Menge von vorher vereinbarten Abbildungsmassstäben einen geeigneten auszuwählen und das Rasterbild des vorliegenden Zeichens damit aus dem Eingaberaster in ein Ausgaberraster so abzubilden, dass es die bestimmte Grösse erhält. Das Ausgaberraster, das die grössennormierten Zeichen enthält, ist zweckmässigerweise gröber, das heisst es enthält weniger Bildpunkte.

Der Abbildungsmassstab

$$m = \frac{h}{h'}$$

ist das Verhältnis der Höhe h des Zeichens im Eingaberaster zur angestrebten, normierten Höhe h' im Ausgaberraster, jeweils gemessen in Bildpunkten. Für die Grössennormierung ist es also notwendig, die Zeichenhöhe im Eingaberaster zu messen. Dies wird am Zeichenprofil, das heisst der Randverteilung in vertikaler Richtung, vorgenommen. Im einfachsten Fall wird nur der Zeichenschatten verwendet und von oben und unten sein Beginn und Ende ermittelt. Am Beispiel einiger Schriftzeichen mit stark variierender Höhe zeigt *Figur 9* die ausgleichende Wirkung der Grössennormierung. Die Bildvorbereitungs-Operationen der Segmentierung, Zentrierung und Grössennormierung für eine Anschriftzeile zeigt *Figur 10* noch einmal im Zusammenhang. Im übrigen ist es angesichts der zahlreichen Zeichenvariationen, mit denen der nachfolgende Klassifikator fertig werden muss, unangemessen, bei den Prozeduren der Bildvorbereitung Perfektionismus zu betreiben.

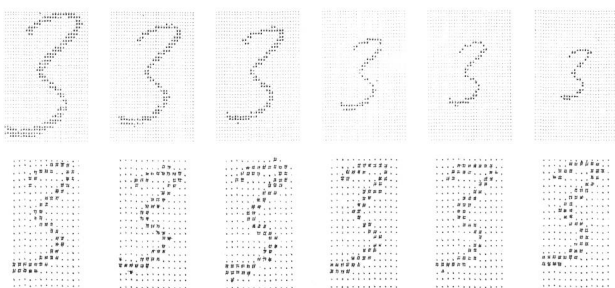


Fig. 9
Grössennormierung - Calibrage

323 Centrage

Les décalages de caractères à l'intérieur de la zone tramée n'ayant pas d'influence sur leur signification, il est rationnel de veiller à ce qu'une configuration donnée de points noirs dans un environnement blanc ne se produise pas, dans la zone tramée, pour plusieurs positions décalées les unes par rapport aux autres. Les procédés de normalisation de position les plus importants sont le centrage en fonction d'une carte des positions types horizontales et verticales des caractères ou d'une surface de référence assignée à chaque caractère. Il est souhaitable que d'infimes modifications de la configuration du noir et du blanc ne conduisent pas déjà à un décentrage marqué. A cet égard, les procédés de centrage par surface de référence sont plus efficaces que ceux qui reposent sur le suivi des contours.

324 Calibrage

Parmi les variables qui, par définition, ne doivent pas avoir d'influence sur la signification du texte lu, il faut mentionner la grandeur des caractères. Le principe du calibrage consiste à choisir, parmi un grand nombre d'échelles fixées à l'avance, celle qui convient à la reproduction du caractère et à lui conférer une grandeur définie, de sorte que la trame d'entrée soit convertie en une trame de sortie normalisée. La structure de la trame de sortie - contenant les caractères calibrés - doit être de préférence plus grossière, c'est-à-dire contenir moins de points significatifs.

$$\text{L'échelle } m = \frac{h}{h'}$$

est le rapport entre la hauteur h du caractère au niveau de la trame d'entrée et la hauteur normalisée h' à celui de la trame de sortie, mesuré en points d'image. Il est donc nécessaire de mesurer la hauteur des caractères, lorsqu'ils sont tramés initialement, pour obtenir un calibrage. On y parvient par la mesure soit du profil du caractère, soit de la distribution marginale de l'information projetée en direction verticale. Dans le cas le plus simple, seule l'ombre du caractère est significative et sa limite supérieure et inférieure permet de déterminer le début et la fin d'un caractère. A l'exemple d'une série de caractères à hauteurs très différentes la *figure 9* illustre l'effet égalisateur du calibrage. La *figure 10* montre l'ensemble des opérations de préparation de l'image d'une ligne de l'adresse, à savoir la segmentation, le centrage et le calibrage. Par ailleurs, il serait peu indiqué de faire preuve de perfectionnisme lors de la préparation de l'image, vu les nombreuses variantes de caractères dont le classeur qui suit doit s'accommoder.

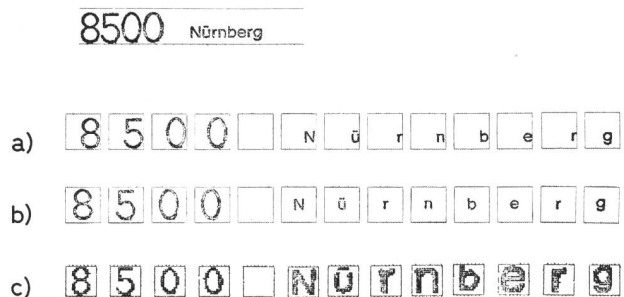


Fig. 10
Bildvorbereitung - Préparation de l'image

Nach Abschluss der Bildvorbereitung liegt das gerasterte Einzelzeichen in einem 16 Rasterelemente breiten und 16 Rasterelemente hohen Bildfeld vor. Numeriert man die Elemente zum Beispiel von links oben nach rechts unten durch, so wird die Schwärzungsverteilung jedes Zeichens durch eine Folge von 256 Zahlenwerten $x_1, x_2, \dots, x_i, \dots, x_{256}$ repräsentiert, die jeweils die Zahlenwerte $x_i = 1$ (entsprechend einem schwarzen Bildelement) oder $x_i = 0$ (für ein weisses Bildelement) annehmen können. Aufgabe des Klassifikators ist es, nur aus der Kenntnis der Schwärzungsverteilung, also aus der Zahlenfolge x_1, x_2, \dots, x_{256} , die mutmassliche Klassenzugehörigkeit anzugeben. Zur Lösung dieser Aufgabe gibt es eine unübersehbare Fülle von Vorschlägen. AEG-Telefunken hat sich nach jahrelangen Forschungsarbeiten entschlossen, einen Weg zu gehen, der aus der Ideenwelt der *Prognosestechnik* stammt. Die Aufgabe der Klassifizierung kann dann so beschrieben werden: Aus der gemessenen Schwärzungsverteilung soll eine Prognose über die vermutete Klassenzugehörigkeit abgegeben werden, oder genauer, es soll eine Schätzung d_k für die Wahrscheinlichkeit berechnet werden, mit der das Zeichen jeder der verabredeten Klassen zugehört. Der Index k des grössten Schätzwertes d_k weist auf die vermutete Klasse hin.

Das mathematische Handwerkzeug der Prognosestechnik ist die Regressionsanalyse. Es wird ausgegangen von einem in den x_i quadratischen Ansatz für die Schätzgleichung

$$d_k = a_{0,k} + a_{1,k}x_1 + \dots + a_{256,k}x_{256} + a_{257,k}x_1 \cdot x_2 + \dots + a_{32896,k}x_{255} \cdot x_{256}$$

für jede der zu unterscheidenden Klassen. Der vollständige quadratische Ansatz hat bei 256 Bildpunkten eine Länge von 32 897 Termen, eine auch für moderne Grossrechenanlagen völlig undiskutable Grösse. Die Ansatzlänge wird daher zunächst nach heuristischen Gesichtspunkten verkürzt, indem eine «Reichweitengrenze» angegeben wird, das heisst ein maximaler Abstand zwischen zwei Bildpunkten x_i und x_j , den diese haben dürfen, um als quadratisches Glied $x_i \cdot x_j$ in den Ansatz aufgenommen zu werden. Eine weitere Verkürzung der Ansatzlänge ergibt sich während des mathematischen Optimierungsprozesses zur Berechnung der zunächst noch unbekanntenen Koeffizienten $a_{i,j,k}$. Diese Optimierung stützt sich auf eine «Lernstichprobe», in der viele verschiedene Zeichenformen aus jeder Bedeutungsklasse vorkommen, gekennzeichnet mit ihrer wirklichen Bedeutung. Übrig bleibt schliesslich eine Schätzgleichungslänge von 1024 Gliedern für die Klassifizierung von 26 Buchstaben oder 512 Gliedern für die Klassifizierung der 10 Ziffern. Die Koeffizienten $a_{i,j,k}$ werden offline berechnet und in den Speicher für die Klassifikatordaten (vgl. Fig. 3) übernommen.

Das Ergebnis der Klassifikatoradaption zeigt *Figur 11* an einem Beispiel. Für die linke 5 hat d_5 , für die rechte 6 d_6 eindeutig den grössten Wert. Bei den mittleren beiden Zeichen haben d_5 und d_6 etwa die gleiche Grösse: der Klassifikator kann sich – in Übereinstimmung mit der menschlichen Anschauung – in diesen Fällen nicht eindeutig entscheiden. Derartige Fälle werden auch an die nachfolgende Adressinterpretation als unsicher weitergegeben. Bei einer vorläufigen Klassifikatoradaption auf die 26 Grossbuchstaben der verschiedensten Schriftarten wurde eine Fehlerrate von 10^{-4} bei 10^{-2} als unsicher ge-

Une fois l'image préparée, on dispose du caractère tramé qui se présente sous forme d'une trame ayant 16 éléments de large et 16 éléments de haut. Si on les numérote, par exemple de gauche en haut à droite en bas, la répartition de noir de chaque caractère peut être exprimée par une série de 256 valeurs booléennes $x_1, x_2, \dots, x_i, \dots, x_{256}$, pouvant prendre soit la valeur $x_i = 1$ (pour un élément noir) ou $x_i = 0$ (pour un élément blanc). La tâche du classeur est de déterminer l'appartenance probable à une classe, en vertu de la répartition du noir, soit en fonction des variables booléennes x_1, x_2, \dots, x_{256} . Un nombre immense de propositions répondent à la solution de ce problème. Après des années de recherches, les établissements AEG-Telefunken ont décidé d'opter pour une méthode reposant sur la *technique des pronostics*. Dans cette optique, la classification peut être définie ainsi qu'il suit: Un pronostic concernant l'appartenance probable à une classe doit être élaboré à partir de la distribution de noir mesurée, ou, plus précisément, par estimation de d_k pour la probabilité où le caractère considéré ferait partie de l'une des classes définies.

L'outil mathématique de la technique des pronostics est l'analyse régressive. En ce qui concerne l'équation d'estimation, on part d'une expression bilinéaire en x_i

$$d_k = a_{0,k} + a_{1,k}x_1 + \dots + a_{256,k}x_{256} + a_{257,k}x_1 \cdot x_2 + \dots + a_{32896,k}x_{255} \cdot x_{256}$$

pour chacune des classes à différencier.

L'expression bilinéaire complète possède, compte tenu de 256 points d'image, une longueur de 32 897 termes, volume d'information que même de grands centres de calcul modernes ne sauraient raisonnablement traiter. De ce fait, on raccourcit tout d'abord la longueur de l'expression selon des considérations heuristiques, ce qui revient à fixer une «limite de portée». Cela consiste à choisir un espacement maximal entre deux points d'image x_i et x_j , afin de pouvoir introduire $x_i \cdot x_j$ en tant que terme bilinéaire dans l'expression. Un autre raccourcissement de l'expression se produit durant le processus mathématique d'optimisation visant à calculer le coefficient $a_{i,j,k}$ encore inconnu. Cette optimisation se fonde sur un «échantillon» comprenant plusieurs formes de caractères de chacune des classes significatives affectées de leur signification véritable. Il reste, pour finir, une équation d'estimation de 1024 termes

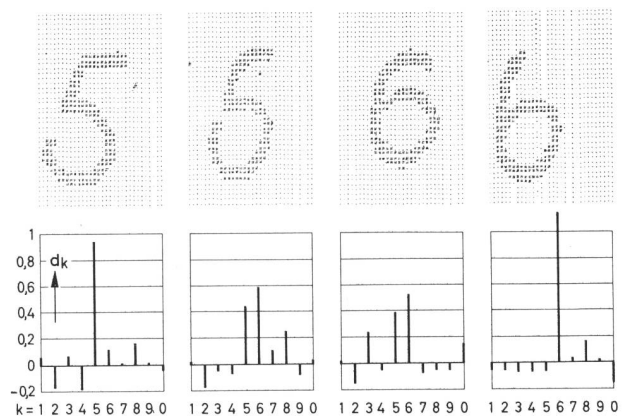


Fig. 11 Schätzung der Klassenzugehörigkeit – Estimation de l'appartenance à une classe

kennzeichneten Zeichen erreicht. Die endgültige Klassifikatoradaption muss mit Hilfe einer sehr grossen Stichprobe von wirklichen Anschriften vorgenommen werden, an der auch die vorher beschriebenen Techniken des Zeilensuchens und der Bildvorbereitung getestet werden müssen. Die Deutsche Bundespost nimmt dazu eine Million abgetastete und gerasterte Anschriften mit etwa 30 Millionen Schriftzeichen auf Magnetband auf. Da im Funktionsmuster des Anschriftenlesers viele Funktionen in Software verwirklicht sind, können Detailkorrekturen einfach durchgeführt werden.

34 Adresseninterpretation und Codierung

Je nach Betriebsart muss der automatische Anschriftenleser die Abgangs- oder die Eingangsinformation erfassen und interpretieren. Die Abgangsinformation besteht aus Postleitzahl, Ortsnamen und allenfalls der Nummer eines Zustellpostamtes, während sich die Eingangsinformation aus Strassennamen und Hausnummer, häufig auch aus der Angabe eines Postfaches zusammensetzt. In beiden Fällen zerfällt die zu lesende Information in mehrere Bestandteile, deren wichtigste «Wörter» wie in der Umgangssprache Buchstabenfolgen variabler Länge sind. Einschliesslich der Umlaute und des Zeichens β stehen im Deutschen 30 verschiedene Buchstabenklassen zur Verfügung. Es ist leicht auszurechnen, wie viele unterschiedliche Zeichenfolgen sich bei gegebener Wortlänge im Prinzip bilden lassen. Bei einer Wortlänge 10 ergeben sich beispielsweise $30^{10} \approx 10^{14}$ verschiedene konstruierbare Wörter.

Charakteristisch für das Anschriftenlesen ist nun, dass von der fast unermesslich grossen Anzahl konstruierbarer Wörter nur ein nahezu verschwindend kleiner Anteil tatsächlich vorkommt. So enthält das amtliche Postortverzeichnis der Deutschen Bundespost insgesamt nur etwa 23 000 gültige Ortsnamen und davon nur etwa 3300 mit der Wortlänge 10. In der Sprache der Informationstheorie bezeichnet man diesen Sachverhalt als Redundanz. Die ausserordentlich redundante Schreibweise der zu lesenden Information gestattet es im Prinzip, mit Störungen fertig zu werden, die zwischen Absenden und Empfangen der Information auftreten können. Dabei kann es sich um falsche Schreibweise, Druckfehler und Druckmängel, Verschmutzungen der Adresszone, Funktionsmängel des Abtast- und Bildvorbereitungssystems oder auch um Mängel des Einzelzeichen-Erkennungssystems handeln.

Die Korrektur solcher Übertragungsfehler erweist sich als ein Mustererkennungsproblem, das gewissermassen eine Hierarchiestufe über dem Problem der Erkennung isolierter Schriftzeichen angesiedelt ist. Die Entscheidung zielt hier nicht mehr auf die Bedeutungsklasse des Einzelzeichens, sondern auf die im Wörterbuch enthaltenen gültigen Wörter. Die Messungen, die zur Entscheidungsfindung herangezogen werden, sind nicht mehr die elementaren Schwarzweissmessungen am Rasterbild eines Einzelschriftzeichens, sondern es sind hier die Schätzungen des Einzelzeichen-Erkennungssystems für sämtliche Buchstabenpositionen des zuzuordnenden Wortes.

Das Einzelzeichen-Erkennungssystem erzeugt Schätzungen für die mutmassliche Bedeutung des betrachteten Schriftzeichenbildes. Häufig lässt sich daraus unmittelbar eine Entscheidung ableiten, die mit hoher Wahrscheinlichkeit richtig ist. In unsicheren Fällen wird zusätzlich eine «zweite Wahl» angeboten, in besonders unsicheren Fällen darüber hinaus auch noch eine dritte, in ganz unsicheren

pour la classification de 26 lettres ou de 512 termes pour la classification de 10 chiffres. Les coefficients $a_{i,k}$ sont calculés off line et introduits dans la mémoire des données de classification (voir fig. 3). L'exemple de la figure 5 montre comment le classeur interprète (adapte) l'appartenance d'un caractère à une classe. Pour le chiffre 5, à gauche, d_5 a clairement la plus grande valeur tandis que pour le chiffre 6, à droite, la plus grande valeur est exprimée par d_6 . Pour les deux signes du milieu, d_5 et d_6 ont à peu près la même grandeur: le classeur – conformément à la logique humaine – ne peut se décider. De tels cas, s'ils surviennent dans l'interprétation d'adresses, sont acheminés dans la catégorie «peu sûr». Capable provisoirement d'interpréter les 26 majuscules de divers types de caractères, le classeur fonctionne à un taux d'erreur de 10^{-4} pour 10^{-2} signes reconnus peu sûrs. L'interprétation du classeur, au stade final, devra être mise au point à l'aide d'échantillons très nombreux d'adresses véritables, auxquels seront également appliquées les techniques décrites de la recherche de ligne et de la préparation de l'image. A cet effet, les Postes fédérales allemandes enregistrent sur bandes magnétiques un million d'adresses lues électroniquement et tramées, ce qui représente quelque 30 millions de caractères. De nombreuses fonctions du prototype de lecteur d'adresses étant réalisées en software, il est possible de procéder de manière simple à des corrections de détails.

34 Interprétation d'adresses et codage

Suivant le mode d'exploitation, le lecteur automatique d'adresses doit saisir et interpréter les informations de départ ou les informations d'arrivée. Les informations de départ comprennent le numéro postal d'acheminement, le nom de la localité et éventuellement le numéro de l'office postal de distribution. L'information d'arrivée, en revanche, se compose du nom de la rue et du numéro de la maison ainsi que, très souvent, d'une indication de la case postale. Dans les deux cas, l'information à lire se subdivise en plusieurs parties, dont les «mots» significatifs sont, comme dans le langage courant, des séries de caractères de longueur variable. Compte tenu des signes particuliers de la langue allemande, il faut disposer de 30 classes de caractères. Il est facile de calculer combien de combinaisons de signes peuvent être formées pour une longueur de mot donnée. Un mot de 10 lettres permet théoriquement de construire $30^{10} \approx 10^{14}$ combinaisons de mots.

Ce qui caractérise la lecture d'adresses est le fait qu'un nombre relativement restreint de mots doivent pratiquement être pris en considération parmi le nombre presque incommensurable de combinaisons possibles. Ainsi, la liste officielle des localités des Postes fédérales allemandes ne contient qu'environ 23 000 noms de localités valables, dont 3300 seulement ont une longueur de 10 caractères. En informatique, ce phénomène est appelé redondance. Cette écriture très redondante de l'information à lire permet en principe de venir à bout des erreurs pouvant survenir en chemin. Il peut s'agir, en l'occurrence, de fautes d'orthographe, de fautes et de défauts d'impression, de zones d'adresses maculées, de défauts de fonctionnement des systèmes de lecture et de préparation de l'image ou de dérangements à l'ensemble de reconnaissance de caractères.

La correction de telles erreurs de transmission est en fait un problème de reconnaissance de critères de référence, problème situé hiérarchiquement un échelon plus

Fällen kann die Entscheidung vollständig verweigert werden. Die Fallunterscheidung ist so konzipiert, dass mit hinreichend hoher Wahrscheinlichkeit die richtige Bedeutung des Einzelschriftzeichens in den erzeugten Alternativen enthalten ist. Aus den wählbaren Lösungsvorschlägen für die einzelnen Buchstabenpositionen eines ganzen Wortes kann dann eine Menge von alternativen Lösungsworten gebildet werden. Mit hinreichend hoher Wahrscheinlichkeit ist das korrekte Wort in der Menge der Alternativen enthalten. Mit sehr geringer Wahrscheinlichkeit enthält die Menge von Möglichkeiten gleichzeitig auch falsche, aber wiederum gültige Wörter. Durch Nachschlagen im Wörterbuch findet man die gültigen Wörter in der Menge der angebotenen Alternativen.

Der einfachste Fall ist die Wörterbuchsuche mit Hilfe eines unmittelbaren Deckungsvergleiches. Das ist der Spezialfall eines allgemein konzipierten Ähnlichkeitsvergleiches. Indem Ähnlichkeitsmasse eingeführt werden, die auf die Besonderheiten der beim Anschriftenlesen zu erwartenden Störungen abgestimmt sind, wird das Adressen-Interpretationssystem in die Lage versetzt, auch mit Fehlern fertig zu werden, die über fehlerhafte Entscheidungen beim Einzelzeichenerkennen hinausgehen. Darunter haben die Segmentierfehler des Bildvorbereitungssystems, durch die die Zeichenanzahl im gelesenen Wort verändert werden kann, die grösste Bedeutung.

Nach dem Identifizieren von Orts- oder Strassennamen werden die weiteren Bestandteile der Abgangs- oder Eingangsinformation in die Adresseninterpretation einbezogen. Die Postleitzahl hat dabei die Aufgabe, die bei zahlreichen Ortsnamen vorkommenden Mehrdeutigkeiten aufzulösen. Im übrigen dient sie zur Absicherung der über den Ortsnamen getroffenen Entscheidung. Bei unleserlichen Ortsnamen kann sie diesen ersetzen.

Im Fluss der Informationsverarbeitung eines Anschriftenlesers stellt die Adresseninterpretation die höchste Hierarchiestufe dar. Dementsprechend benötigt sie für ihre Aufgabe auch den grössten Vorrat an Hintergrundkenntnissen. Den meisten Speicherplatz beanspruchen dabei die Wörterbücher mit den Ortsnamen für die Abgangscodierung und mit den Strassennamen für die Eingangscodierung. Insgesamt sind etwa 4000 kbit erforderlich. Bestandteil dieser für die Adresseninterpretation bereitzustellenden Wörterbücher sind über die zulässigen Schreibweisen der Wörter hinaus statistische Daten, die Angaben über die Häufigkeit enthalten, mit denen die Wörter des Wörterbuches am Aufstellungsort des Anschriftenlesers auftreten. Als letzter Schritt im Arbeitsablauf des Anschriftenlesers wird die aus dem Anschriftenblock entnommene Verteilinformation auf die Sendung (Brief oder Postkarte) in Form des Codeaufdruckes wie beim manuellen Codierplatz rückübertragen.

Aus dem Umfang der Probleme und Lösungswege, die in diesem Beitrag skizziert wurden, ist unmittelbar klar, dass die Autoren nicht allein über eigene Forschungsarbeiten berichten, sondern dass hier vielmehr eine Konzeption dargestellt wurde, hinter der die Leistung einer grossen Forschungs- und Entwicklungsmannschaft steht.

haut que la reconnaissance de caractères isolés. La décision ne dépend plus, dans le cas particulier, de la classe de signification du caractère, mais des mots valables contenus dans le lexique. Les mesures servant à prendre les décisions ne sont plus les mesures élémentaires de noir et de blanc faites au niveau de l'image tramée d'un caractère; il s'agit ici d'estimations du système de reconnaissance de caractères portant sur toutes les positions des lettres du mot à classer.

Le système de reconnaissance de caractères fournit des estimations concernant la signification supputée du caractère imprimé à lire. Très souvent, il en résulte sans délai une décision d'un taux de vraisemblance très élevé. Dans les cas peu sûrs, un «deuxième choix» est considéré ou un troisième lorsque l'incertitude est particulièrement grande. Dans les cas douteux, il y a refus de décision et, partant, rejet.

Le système de décision reconnaît avec une probabilité suffisamment élevée la signification véritable des caractères parmi le choix des possibilités offertes. Parmi les solutions proposées, qui sont des alternatives quant à la position des caractères à l'intérieur d'un mot complet, il est possible de retenir une multitude de mots-solutions pouvant convenir. Toutefois, le mot correct est contenu avec un degré de probabilité suffisamment élevé dans le choix des alternatives. Il est, de plus, fort improbable que ce choix renferme aussi des mots faux pouvant néanmoins être reconnus exacts. En compulsant le lexique, on trouve les mots valables parmi le choix des alternatives proposées. Le cas le plus simple est la recherche dans un lexique par la méthode de coïncidence directe, ce qui constitue un cas particulier du test de similitude pris dans un sens très général. Par l'introduction de critères de similitude spécialement adaptés aux particularités des perturbations à prévoir dans la lecture d'adresses, le système d'interprétation des adresses est mis en mesure de venir également à bout d'erreurs situées à un niveau plus élevé que les décisions erronées touchant la reconnaissance de caractères isolés. A cet égard, les erreurs de segmentation du système de préparation de l'image revêtent la plus grande importance, car elles peuvent modifier le nombre de caractères du mot lu.

Une fois les noms des localités et des rues identifiés, les autres parties constituantes de l'information de départ ou de celle d'arrivée sont retenues pour l'interprétation de l'adresse. Dans ce contexte, le numéro postal d'acheminement sert à lever le doute lors d'ambiguïtés dans les noms de localités. Il confirme aussi l'exactitude des décisions prises au sujet desdits noms et peut même les remplacer s'ils sont illisibles.

Parmi toutes les informations que traite un lecteur d'adresses, le processus d'interprétation de l'adresse est placé au niveau hiérarchique le plus élevé. De ce fait, l'ensemble qui résout cette tâche doit avoir la plus grande capacité de connaissances fondamentales. Ce qui, quant à la mémorisation, prend le plus de place, sont les lexiques

(Suite et fin page 61)