

OPTICAL CHARACTER RECOGNITION TECHNIQUES IN MAIL SORTING:

A REVIEW OF ALGORITHMS

by

Jonathan J. Hull, Ganapathy Krishnan,
Paul Palumbo and Sargur N. Srihari

Department of Computer Science
University at Buffalo, State University of New York
Buffalo, New York 14260

Interim Technical Report No. 1
Advanced Character Recognition Project
United States Postal Service

JUNE 1984

TECHNICAL REPORT NUMBER 214

Abstract

A study of Optical Character Recognition (OCR) techniques employed in automatic mail sorting equipment is presented. Methods and algorithms for image preprocessing, character recognition, and contextual postprocessing are discussed and compared. The objective of the study is to provide a background in the state-of-the-art of this equipment as the first element in a search for techniques to significantly improve the capabilities of postal address recognition.

1. INTRODUCTION

2. METHODS USED

- 2.1 Published Literature Search
- 2.2 Patent Literature Search
- 2.3 Experimental Investigation

3. USPS-OCR: PAST AND PRESENT

- 3.1 First and Second Generation OCR
- 3.2 Scanning Hardware
- 3.3 Electronic Mail
- 3.4 Performance Evaluation of OCR

4. OCR ALGORITHMS FOR MAIL SORTING

- 4.1 Pitney Bowes - Elsag

- [4.1.1 Elsag Technical Specifications](#)
- [4.1.2 Elsag Preprocessing](#)
- [4.1.3 Elsag Character Recognition](#)
- [4.1.4 Elsag Post Processing](#)
- [4.1.5 Elsag Architecture](#)

4.2 [Burroughs - Nippon Electric](#)

- [4.2.1 NEC Technical Specifications](#)
- [4.2.2 NEC Preprocessing](#)
- [4.2.3 NEC Character Recognition](#)
- [4.2.4 NEC Post Processing](#)

4.3 [AEG-Telefunken](#)

- [4.3.1 Telefunken Technical Specifications](#)
- [4.3.2 Telefunken Preprocessing](#)
- [4.3.3 Telefunken Character Recognition](#)
- [4.3.4 Telefunken Post Processing](#)
- [4.3.5 Telefunken Conclusions](#)

4.4 [Toshiba](#)

- [4.4.1 Toshiba Technical Specifications](#)

4.5 [Recognition Equipment, Inc. \(REI\)](#)

- [4.5.1 REI Technical Specifications](#)

5. [CONCLUSIONS](#)

ACKNOWLEDGEMENTS

6. [BIBLIOGRAPHY](#)

- [6.1 Background](#)
- [6.2 Elsag](#)
- [6.3 Nippon Electronic](#)
- [6.4 Recognition Equipment Inc.](#)
- [6.5 Telefunken](#)

1. INTRODUCTION

This is the first interim technical report of the Advanced Character Recognition Project

conducted at the State University of New York at Buffalo. The project, supported by the United States Postal Service Technology Resource Department, has three tasks. This report covers task 1 (performed during the period of January 10, 1984 - April 15, 1984).

The main objective of task 1 was to evaluate optical character recognition (OCR) techniques employed in commercial mail sorting equipment and currently available to the United States Postal Service (USPS). This was to be done primarily by describing the methodology and algorithms commercial equipment employ for image processing and character recognition and by providing a technical evaluation where possible. Task 1 was to be performed to serve as a frame of reference for tasks 2 and 3. The goal of Task 2 is to assess the state-of-the-art in character recognition and image processing in the context of applicability to USPS mail sorting. The objective of Task 3 is to develop a plan for future research in these areas that will substantially advance the state-of-the-art in mail address recognition.

In order to accomplish the goal of Task 1 it was decided to review and report on documentation concerning the USPS OCR experimentally evaluate current commercial techniques and to project future research was also envisaged.

This document contains a report of activities that describes the methods used for locating published and patent literature, and the experimental software developed. A background report is also given on the past and present USPS OCR environment and research. A description is provided of methods, algorithms and technical specifications of character recognition equipment manufactured by three firms: Pitney Bowes - Elsag, Burroughs Nippon Electric (NEC), and Telefunken. Only technical specifications are provided for two others: Toshiba and Recognition Equipment, Inc. A bibliography of relevant literature is also presented that includes papers in the open technical literature, U.S. patents and USPS-sponsored technical reports.

2. METHODS USED

Technical work on the Advanced Character Recognition project commenced after a meeting of representatives of the USPS Technology Resource Department, their technical consultant Arthur D. Little Inc., and contractors in the Advanced Character Recognition project, viz., SUNY at Buffalo, University of Maryland, and Recognition Equipment, Inc. on January 12, 1984 in Washington, D.C.

It was decided to limit the review of algorithms to relatively current OCR technology reported by firms that manufacture mail sorting equipment. These firms were determined to be Pitney Bowes-Elsag, Burroughs-NEC, AEG-Telefunken, Recognition Equipment (REI), Hotchkiss-Brandt, and Toshiba. Some of the manufacturers were contacted to locate relevant material. Due to the proprietary nature of most of the material, it was decided that a search of the published and patent literature was the most feasible approach to this review. An implementation of some of the algorithms was considered necessary to gain insight into some of the techniques.

2.1 Published Literature Search

The published scientific literature was searched using a combination of manual and computer-assisted means to construct a bibliography on character recognition techniques developed or implemented by these firms. These efforts led to locating significant material for Telefunken and Elsag and to a lesser extent for NEC and REI. Nothing of substance was located for

Hotchkiss-Brandt and Toshiba. Literature published in reputed scientific journals can provide objective technical information, as demonstrated by the discussions on Elsas character recognition (section [4.1.3](#)) and Telefunken character recognition (section [4.3.2](#)). Published literature may however describe techniques not yet commercially viable.

2.2 Patent Literature Search

Due to the unavailability of substantial published literature on several of the manufacturers, a manual and computer- assisted search of the patent literature was essential. A patent search was conducted using the facilities of the Buffalo and Erie County Public Library and the Cassis and BRS database services; the search was based on keyword indices of manufacturers and subject classifications. A relatively large number of patents for NEC and REI, and to a lesser extent Elsas, were located and determined pertinent; see the bibliography for a list. These have been reviewed as described in Section 4. Patent literature provides more technical details, but fewer objective evaluations, than published literature. However, these techniques are more likely to be the ones implemented in current equipment.

2.3 Experimental Investigation

Current commercial OCR techniques involve three successive steps. These are: image processing, character classification and contextual post-processing. Supporting experimental work on each of these areas was deemed useful and is described next.

2.3.1 Image Processing

A) To develop image processing capabilities necessary to this project, a software package, called HIPS (Human Information Processing System), was brought up on a VAX-11/780 and interfaced to a Grinnell display. One of the image data tapes received from REI were decoded and processed using HIPS.

B) A sample of two trays of third class dead-letter mail was received from the Buffalo USPS Sectional Center. To provide a demonstration database of high resolution envelope images, a carefully selected sample of eight envelopes was digitized with a laser scanner located at Kenmore Colorplate in Kenmore, New York. The envelopes were scanned at a resolution of 500 pixels per inch at three different gain settings. A tape of the images was then decoded for processing and display on the VAX. This data will be used to study the effect of high resolution scanning on recognition algorithms in Tasks 2 and 3.

C) The Elsas adaptive image thresholding technique described in section [4.1.1](#) was implemented (in the "C" language) to study its effectiveness.

D) A synthetic address generation system was developed. This allows for the creation and display on the Grinnell screen of images of city/state/zip lines from the DOPO (Directory of Post Offices) database. Parameters can be controlled to alter the typefont and to introduce varying levels of skew and noise into these images. The objective of this effort is to develop the capability to generate a controlled set of data so that large scale experimentation could be conducted to test the effect of alteration of the aforementioned parameters.

2.3.2 Character Recognition

As part of an experimental investigation of recognition algorithms, the Telefunken polynomial discriminant approach to character recognition (see section [4.3.2](#)) was implemented in APL. This has provided an understanding of the strengths and weaknesses of this method.

An experimental comparison of the Telefunken method with other recognition methods would be valuable. Such an investigation requires a large, labelled set of isolated, normalized, binary character images. A request for such a data set was turned down by REI due to its proprietary nature. Presently, the binary video data of address lines provided by REI is being used to generate such a database.

2.3.3 Contextual Post-processing

The Telefunken postprocessing approach, based on hashing functions (see section [4.3.3](#)), was implemented to accept input from the Telefunken character recognition simulation. To provide a database of city/state/zip combinations to test this algorithm, the USPS Address Information Systems Support Center provided the DOPO database. The zip-a-list database was also obtained so that this information along with a good implementation of the post-processing algorithm would allow a useful simulation of the Telefunken approach to be constructed. This would provide a valuable empirical base for comparison with other character, word, and address recognition approaches.

3. USPS-OCR: PAST AND PRESENT

OCR techniques have a long history in mail sorting applications. In this section we report on first and second generation OCR, scanning hardware, the relationship of OCR to electronic mail, performance evaluations of presently installed OCR equipment, and an on-site evaluation of an operational OCR. The report is primarily in the form of reviews of several documents that provide background information on the past and present USPS OCR environment and on USPS-OCR research. Complete citations for the documents are given in section 6.1.

3.1. First and Second Generation OCR

3.1.1 "Optical character readers in the USPS," circa 1977.

A short history of OCR in the USPS is given in this document. It does not indicate firms that manufactured the equipment. It discusses the first and second generation OCRs and projects the needs of the third generation.

First generation systems use flying spot scanners to locate the line to be read. These systems also use two MTUs (Mail Transport Unit) and two scanners that time-share a recognition unit. Recognition is character by character followed by a lookup in a comparison unit.

Second generation systems are described in detail. These systems extract parameters that describe each line of data. These parameters are passed to the character recognition processor that recognizes individual characters. This information is used to format character data for context analysis.

The second generation indexing OCR of REI is then described vaguely. It differs from earlier systems in that it has one MTU. The A field of the address is defined to be City, State, and Zip. The B field contains other parts of the address. The A and B field information is verified, coded,

and sprayed on the envelope in the form of a bar code.

3.1.2 "Zipping the mail," Optical Spectra, August, 1973.

This document describes the OCR II system developed by Philco-Ford and installed in Boston's south postal annex. This apparently was the only installation of this device. This machine was notable for reading the entire address, including house number or box number and patron's name. The author of this article was in charge of the OCR project at Philco-Ford.

3.1.3 "Foundry Font reading tests," USPS Technical Notes, 1976.

These are four documents detailing the procedures and results of tests of the ability of the three experimental second generation OCRs: Philco-Ford OCRII, IBM A-OCR, and REI CSOCR-II, to read foundry fonts. Results indicate that "none of the three machines did very well from the standpoint of character recognition performance". However, among its competitors, the REI machine produced the best overall performance. All machines were about equal for upper case numerics, but the REI and IBM machines were noticeably better than the Philco-Ford machine in performance on lower case alphas and numerics.

3.2 Scanning Hardware

3.2.1 G.U. Merkson, "Scanner Review," Technical Note PTR-02-80, October, 1979.

Laser and semiconductor array scanners are reviewed for application to postal information processing. These technologies were chosen for study because of their ability to provide the required resolution and throughput specifications demanded by page scanning for electronic mail systems (EMS) and envelope scanning. The resolution needed in these areas is from 200 to 300 pixels per inch with a throughput rate of 16 to 38 megapixels per second for EMS and 23 megapixels per second for OCR. This corresponds to processing twenty pages per second for EMS and twelve envelopes per second for OCR.

The deflection systems used in laser scanners, which direct the scan of the light beam, are reviewed to determine which are appropriate for postal needs. Five types of electromechanical deflection techniques are considered. These are torsion rod, taut band, galvanometer, piezoelectric, and multifaceted mirror deflectors. All these methods are determined to be unsuitable for postal needs except the torsional mirror which is deferred for study to a future time.

Acousto-optic deflection is then considered under which Bragg deflection is discussed at length. Due to the necessity of focusing the light beam as well as deflecting it, acoustic traveling wave lenses and chirp lenses are also considered.

Solid state semiconductor arrays are devices that convert reflected light into electrical energy. Both linear and two dimensional versions are available. However, only linear versions are considered here. Of these, MOS (standard self scanned photodiode arrays), CCD (charge coupled devices), CCPD (charge coupled photodiode arrays), and CID (charge injection devices), are discussed.

MOS devices are reported on favorably with it being pointed out that they have a broad spectral response, including the blue portion of the spectrum. This is a significant problem for

at least one of the current postal OCRs.

CCD devices are then discussed. Even though their spectral response is not as good as that of MOS devices, they are considered a competitive technology.

CCPD devices are briefly mentioned. They are not considered further due to their maximum sample rate of two megapixels per second, too slow to meet the aforementioned specifications.

CID devices are similar to CCD sensors except that the internal method for reading a signal from a scanning cell is different. These devices are favorable for postal applications due to their low cost and potentially high yield. However, they are presently only available in area array format and will not be investigated further at this time.

It is mentioned that the time delay and integration method of scanning will be investigated further by the Naval Ocean Systems Center (NOSC).

This report concludes with the recommendation that two types of laser scanners and two types of solid state imaging devices be investigated further for their application to the postal domain. These are a laser scanner with acousto-optic beam deflection followed by an acoustic traveling wave lens or a chirp lens. The solid state devices to be investigated are CCD and MOS devices. It appears as though the author of this report contemplated building a complete version of each system. The following section discusses the experimental CCD implementation mentioned above.

3.2.2 G.U. Merkson, "Characteristics of Solid-State photo diode array scanners for Postal Service applications," USPS Technical Note 08-83, May 1983.

Solid-state photo diode array scanners are investigated for application to electronic mail (FAX), and OCR. FAX was the domain of interest for this report since its requirements are more stringent than OCR. However, the results also apply to OCR. A demonstration model of a CCD page scanner was constructed and tested in a USPS laboratory. This report details this device and offers suggestions for its use in postal applications. Most of this report has been deleted. However it is obvious that the deleted portions and much of the included material deal with detailed technical questions not pertinent to our investigation.

"Seventh Annual Report, Advanced Mail Systems Scanner Technology," Naval Ocean Systems Center, May, 1982.

This is one in a long series of reports by NOSC on their investigation of scanning technologies for application to the postal domain. The major thrust of this report appears directed toward the electronic mail domain and the investigation of scanning and memory technologies therefore. This report is primarily an update of the activities of the contractor and will not be discussed further at this time.

3.3 Electronic Mail

3.3.1 R.O. Sheppard, Jr., "Feasibility and Implementation of an Adaptive Recognition Technique," Postal Technology Research Report, January, 1982.

An adaptive character recognition technique using a flying spot scanner interfaced to a PDP-11/20 minicomputer is described. The application domain is the recognition of a line of billing information from the header of a letter in an electronic mail system. This technique uses the first ten digits in the line as prototypes in the recognition of the digits in the rest of the line.

Each test document contained a line of 30 digits, the first ten were "0" through "9", the remaining 20 were randomly chosen. Each test document was photographed and the resultant negative was used for digitization with the flying spot scanner. Segmentation is done by fetching data from the input line until it is determined that a complete character has been found. A character thus found is centered in a 20x16-bit field and passed to a storage buffer.

The first ten characters of a data line were used to generate "reference" and "don't care" templates. The reference template is a smoothed version of an ideal character, the don't care template is a mask used to eliminate black-white transition areas in the reference from consideration during matching.

Recognition is performed by comparing each unknown character to each template using a Hamming distance where the "on~ bits of the don't care mask are not counted. To account for misalignment, the unknown is shifted to five different horizontal and five different vertical positions (a total of 25 different positions) and the distance is calculated at each position. The lowest Hamming distance for the 25 calculations is stored as the score for the character, the other 24 are discarded. This procedure is repeated for each unknown character using the ten reference templates. The template with the lowest score provides the recognition.

The performance of the system is quite good with one reject error and no substitution errors when tested on a sample of 1120 characters from 19 different fonts.

A section of this report is then devoted to a description of a prototype hardware implementation of this method. Most of this material is not germane to our investigation. It is concluded that this technique provides high recognition performance at low cost with the capability of processing five documents per second. In contrast to a multi-font system, a user would not be limited to a predefined set of fonts but would rather be constrained with respect to format.

Overall, this is a very interesting and well-written report. The technique is described clearly and succinctly enough to be implemented directly.

3.4 Performance Evaluation of OCR

3.4.1 "Report on the field testing of Commercial OCR's".

Results of tests of five commercially available OCR's during 1979 and 1980 are reported. Each OCR was installed in a different post office and extensive testing with "live" mail was done. This report is a summary of the conduct and results of those tests. The tests appear to have been conducted using a plan similar to that discussed in [3.4.2](#) below. The average percentage of mail that was unreadable due to script, misfaced, no ZIP, partial ZIP, obscured ZIP, illegal ZIP, degraded print, out of scheme, foreign, and return to sender were categorized according to mail type as below:

mail type	% unreadable
------------------	---------------------

CIRCS	3
Collection	40
LVMM	5
Metered	12
Managed	26

mail type	C	D
CIRCS	65.9	73.3
Collection	39.1	34.0
LVMM	76.3	84.0
MB	61.6	58.7
MMP	52.8	49.0

One of the OCR's was designed to also read script addresses. It achieved 11% to 13% recognition on such addresses. As for pertinence to this investigation, this report presents only data on nine reasons for rejection attributable to faulty OCR.

3.4.2. R.O. Sheppard, Jr., "Effects of high density laser printing on OCR performance," USPS Research and Development Laboratories, September, 1980.

The results of tests on the ability of six OCRs (Elsag, NEC, Telefunken, Toshiba, REI, and the IBM AOCR) to read laser-printed addresses is reported. These tests were carried out in conjunction with those described in [3.4.1](#) above.

A test deck of addresses was prepared using the IBM 3800 and Xerox 9700 laser printers. Three parameters (space between characters, space between lines, and size of characters) were varied and addresses were generated to represent a sample of address throughout the United States. This test data was used to determine the ability of the machines to read such images. Based on the results of these tests, possible areas of improvement in the algorithms of each machine are suggested.

Based on the minimum dimensions determined readable by humans and producible by a laser printer, the following address format requirements are recommended for incorporation in postal OCRs:

- Character size: height \geq 50 mils
 - - width \geq 35 mils

Character separation:

- - at least 10 mils

line spacing: 13 mils

The machines tested did not possess the capability to read such addresses with the required accuracy. This was mainly because of a general requirement for characters greater than 65

mils high. The report suggests that it would be worthwhile to increase the capabilities of these machines to the above specifications.

It should be noted that this document was not supplied by USPS but was located by independent work of the staff of this project. It supplied the technical specifications presented in Chapter 4.

3.4.3 "OCR/CS Operational Test Plan," Usps Research and Development Laboratories, April 15, 1983.

This is a plan for testing one Burroughs and one Pitney Bowes OCR. Multiple tests were run with each of three categories of mail: managed mail, large volume metered mail, and meter belt mail. Data was gathered on the mechanical and recognition performance, the characteristics of mail that was rejected or incorrectly accepted, maintenance history, and the accept rate of five and nine digit zip codes.

The purpose of these tests was to provide data for operational and economic evaluations of OCR/CS and the zip+4 automation program as well as what could be done to improve the readability of mail. A major portion of this document describes the test procedure in detail giving data forms, etc.

3.4.4 USPS memo, 12/28/83.

A limited test of the read reject errors of the Burroughs and Pitney Bowes OCRs is summarized. During a five day period, sample envelopes were extracted from the normal mail stream presented to each OCR. These were run through each machine twice after which those output in the read-reject stacker were categorized according to 16 reasons for rejection. The largest causes for rejection were handwritten address and extraneous data.

3.4.5 On-site OCR Evaluation

During a visit to the Washington D.C. post office on January 12, 1984, an on-site evaluation of OCR equipment installed there was conducted. This included observations of its performance during operational runs on two categories of live mail. During the first run the machine reported 52% correct performance, during the second, 92% correct performance. This reflects the spectrum of mail differences: the first run contained handwritten, typed, and foundry font addresses and the second had well-prepared mail from a few sources. During this visit mail output in the reject stacker was inspected and an experienced operator pointed out reasons for rejection. These reasons for rejection are summarized below:

1. skewed address
2. touching characters
3. portion of address obscured in a window envelope
4. smeared characters caused by ink-jet or dot matrix printing
5. small printing
6. extraneous matter printed on the face of the envelope which caused confusion or failure in address block determination
7. city/state/zip did not appear on the same line
8. hand-generated address
9. pairs of characters confused (e.g. "3"s and "8"s, "G"s and "6"s, and "O"s and "0"s, were

frequently observed confusion pairs)

10. foundry fonts or spectral confusion caused a large position of business reply mail to be rejected.

From this it was apparent that much of the rejected mail was machine printed and could be read easily by humans. The stated reasons for rejection were marginal and point out the necessity for improvements in performance.

4. OCR ALGORITHMS FOR MAIL SORTING

In this section OCR algorithms implemented in commercial mail sorting equipment are described. The discussion is organized by manufacturers, viz., Pitney Bowes-Elsag, Burroughs-NEC, AEG-Telefunken, Toshiba, and Recognition Equipment, Inc. A summary of the technical specifications for each machine are presented (extracted from [Sheppard, 9/80]). Specific algorithms are then discussed that pertain to preprocessing operations such as thresholding, address block segmentation and detecting envelope windows. Also discussed are character segmentation, character recognition and contextual post-processing methods.

4.1 Pitney Bowes - Elsag

After the technical specifications, an adaptive thresholding technique, a character recognition algorithm based on a syntactic approach and a post-processing technique are described. This is followed by a discussion of an architecture for a postal address reading machine.

4.1.1 Elsag Technical Specifications

Scanning:

- 512 bit self-scanned array, each pixel covers 0.125 mm by 0.125 mm.

Character Separation:

- Either fixed pitch or more than 4.92 mils between each character.

Character Dimensions:

- 1 mm \leq height \leq "8mm;" width $<$ 5mm.

Line Spacing:

- should be 0.125mm.

4.1.2 Elsag Preprocessing

An adaptive thresholding technique based on local intensities rather than a fixed, predetermined value is described in the patent [Guilianio, et. al., 1977].

Intensity data (high intensity = "white" and low intensity = "black") is output by a linear photodiode array as it scans a continuously moving document. The intensity $S(x,y)$ of a point P

(x,y) in this image is functionally compared to the intensities of pixels in a 9×9 neighborhood about this point and a binary value $Z(x,y)$ is output. The neighborhood is constructed as shown in **Figure 1** where area A_1 contains the pixels in the 3×3 neighborhood about $P(x,y)$, A_2 contains the four 3×3 neighborhoods diagonally adjacent to A_1 and A_3 contains everything else. Pixels in A_1 are most significant in the comparison to determine an output value, those in A_2 are less significant and those in A_3 are not considered at all.

Figure 1

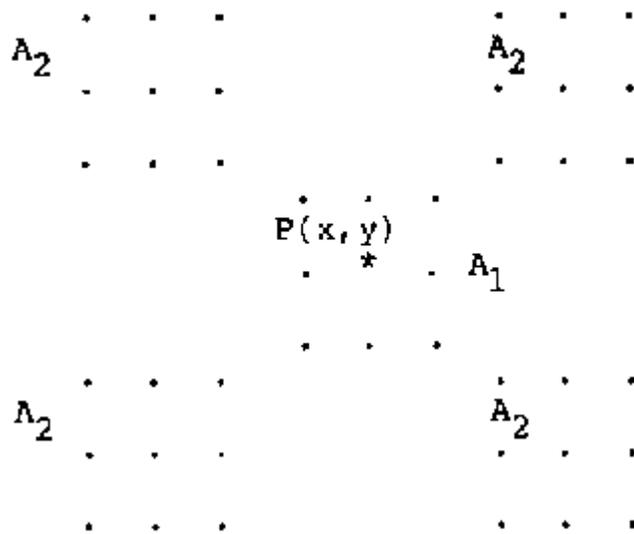


Figure 1. Image point $P(x,y)$, nearby 3×3 neighborhood A_1 , and four diagonally adjacent neighborhoods A_2 .

The relative importance of pixels in A_1 and A_2 is determined by the weighting function $w(g,h)$ defined below: $w(g,h) = K$ if $P(x+g, y+h)$ is in A_1 ,

$w(g,h) = 0$ if $S(x+g, y+h) \geq K'$

$w(g,h) = -K'/mes A_2$ if $S(x+g, y+h) < K'$ and

$P(x+g, y+h)$ is in A_2 ,

$w(g,h) = 0$ if $P(x+g, y+h)$ is in A_3 .

$w(g,h)$ is used in $R(x,y)$ to determine a weighted difference of average intensity values in A_1 and A_2 by multiplying $w(g,h)$ and $S(x+g, y+h)$, double integrating the result with respect to g and h and evaluating the integral over the area $(A_1 \cup A_2 \cup A_3)$.

$$\begin{aligned}
 R(x, y) &= \iint_{A_1 \cup A_2 \cup A_3} S(x+g, y+h) w(x+g, y+h) dg dh \\
 &= K \iint_{A_1} S(x+g, y+h) dg dh - K' / \text{mes } A_2 \iint_{A_2} S(x+g, y+h) dg dh
 \end{aligned}$$

This can be further simplified because of the use of discrete data as:

$$R(x, y) = K \left(\frac{1}{\text{mes } A_1} \left\{ \sum_y \sum_x S(x, y) \right\} \right) - K_3 / \text{mes } A_2 \left(\sum_y \sum_x S(x, y) \right)$$

which is the weighted difference between the mean intensity value in A_1 and the weighted mean of the intensity values in A .

$R(x, y)$ is used as shown below to determine the final binary output value $Z(x, y)$.

- $Z(x, y) = 1$ if $R(x, y) < 0$ and $S(x, y) \geq K''''$
- $Z(x, y) = 0$ if $R(x, y) \geq 0$ and $S(x, y) \geq K''''$
- $Z(x, y) = 0$ if $S(x, y) \leq K''''$

Thus Z effectively smoothes the image. If the intensity value of $P(x, y)$ is greater than a threshold but the average intensity in A_1 is less than that in A_2 , the pixel is set to high intensity (value = 1). If the pixel is above a threshold but the average intensity in A_2 is less than that in A_1 , it is set to low intensity (value = 0). When the original intensity is below a threshold it is set to low intensity in all cases.

This appears to be a good technique that takes into account local intensity variation, however, it still requires several apriori-determined values as does a less sophisticated method. It might also be mentioned that the presentation of this method in [Guiliano, et. al., 1977] is marred by many typographical mistakes

4.1.3 Elzag Character Recognition

A syntactic technique for character recognition based on a context sensitive grammar with four levels of productions is discussed in [Stringa, 1978]. Operations at the image level associate an input pattern with a predefined set of pattern primitives, one per line of a character raster image. First level or "orthographic rank" productions map an input string of such primitives into another string of primitives that represent a standardized version of the original character. Productions at the second level or "lexical rank" in the hierarchy associate a string of pattern primitives with a string of English words. This string is checked for syntactic correctness as an English sentence by productions at the next level, known as the "syntactic rank". Productions at the highest level, known as the "semantic rank", are responsible for the overall description of a character and its recognition.

The primitives associated with each row of an input image fall into four classes according to how the raster intersects horizontal or vertical lines in the character image. **Figure 2** defines each of these classes and presents prototype raster lines in each class. This example, which continues in later figures,

Figure 2

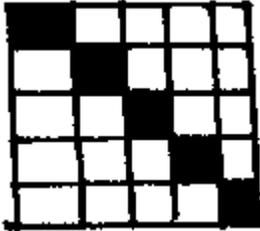
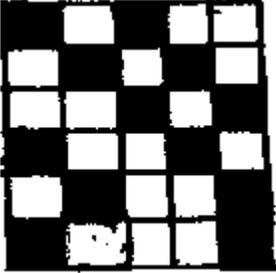
type	definition	patterns	primitive
A	Intersection with horizontal line		c1 c2 c3 c4
B	Intersection with one unique vertical, oblique or curved line		w1 w2 w3 w4
C	Intersection with two lines of type B		b1 b2 b3 b4
D	Intersection with three lines of type B		d

Figure 2: Definitions of the four types of pattern primitives and their corresponding prototypes.

is excerpted from [Stringa, 1978]. The terminal symbol that each prototype is associated with is also shown. Each row of an input character image is matched with these prototypes according to a similarity measure. The resultant string of terminal symbols serves as input to the orthographic level of the grammar. An example of an input character, the matching prototypes, and the associated string of terminals is presented in [Figure 3](#).

Orthographic rank productions are applied to a string of terminal symbols to remove high frequency noise and standardize minute features such as serifs, etc. The result is another string of terminals that represents a normalized version of an input character. An example of the effect of applying these productions to the string of [figure 3](#) is presented in [figure 4](#).

Lexical rank productions associate a string of English words with a string of terminal symbols output by the orthographic rank productions. An example of the application of lexical-rank productions to the output of the orthographic rank productions in

Figure 3

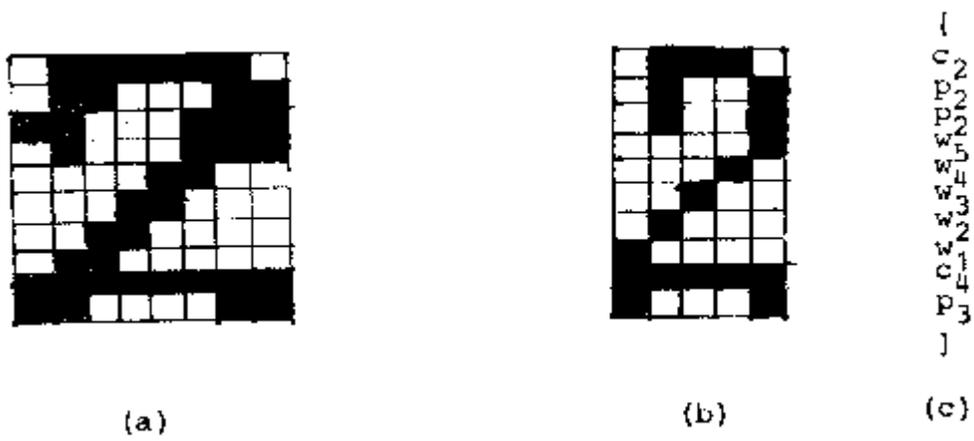


Figure 3: An input character (a), its matching pattern prototypes (b), and string of terminal symbols (c).

Figure 4

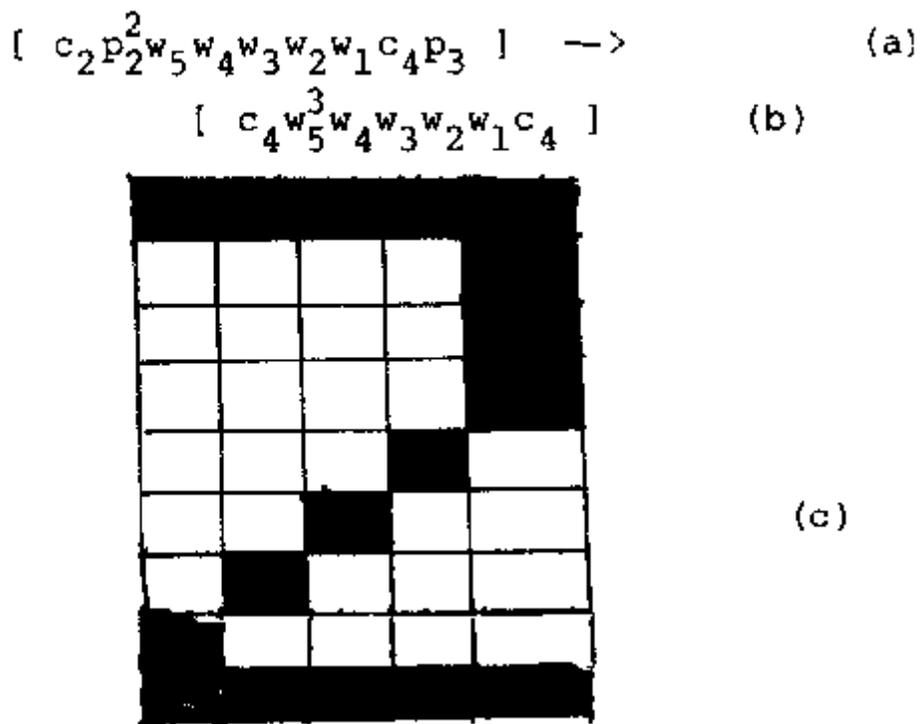


Figure 4: a string of terminals: (a), output of orthographic productions: (b), and (c): pattern corresponding to (b).

Figure 4 is presented in **Figure 5**.

Syntactic rank productions temporarily rewrite a string of English words as a sequence of syntactic class indicators. This string is then verified for correctness. The syntactic rank productions and the form of a correct phrase are given in **Figure 6**. An input phrase is allowable if it matches this phrase or can be formed by a concatenation of such phrases with interposing conjunctions.

Semantic rank productions classify a syntactically correct string by a similarity measure determined during a learning phase. The exact way in which this is done is not clear from the

Figure 5

$[c_4 w_3^3 w_4 w_3 w_2 w_1 c_4] \rightarrow$
 $[c_4 [\text{LONG OBLIQUE LINE FROM RIGHT TO LEFT}] [c_4] \rightarrow$
 $[\text{LONG HORIZONTAL LINE FROM LEFT TO RIGHT}]$
 $[\text{LONG OBLIQUE LINE FROM RIGHT TO LEFT}]$
 $[\text{LONG HORIZONTAL LINE FROM LEFT TO RIGHT}]$
 $[\text{LONG HORIZONTAL LINE FROM LEFT TO RIGHT THEN}$
 $\text{LONG OBLIQUE LINE FROM RIGHT TO LEFT THEN}$
 $\text{LONG HORIZONTAL LINE FROM LEFT TO RIGHT}]$

Figure 5: An example of parsing at the lexical level

Figure6

HORIZONTAL | OBLIQUE | VERTICAL | LONG| MEDIUM | SHORT --> (adjective)
 LINE --> (noun)
 LEFT | MIDDLE | RIGHT --> (adverb)
 FROM | TO --> (preposition)
 AND | THEN --> (conjunction) (adjective) (adjective) (noun) (preposition) (noun) (preposition) (adverb)

Figure 6: Syntactic rank productions and the syntactic form of all legal strings of English words

discussion in the paper.

A character is recognized by this system by first associating its raster image with a set of primitives of the grammar. This set of primitives is considered as a one dimensional string at the first level in the grammar that is parsed into a different representation at each level of the grammar with recognition determined by the closest matching semantic production. Experimental results reported in this paper indicate a 96% to 100% recognition rate with 100 to 150 semantic productions per class. This approach is interesting as a commercial application of syntactic pattern recognition, a basic method that lacks a large number of practical applications.

4.1.4 Elsag Post Processing

The post processing techniques of the Elsag equipment are briefly and vaguely discussed in [[Stringa, 1974](#)] and

[Stringa, 1980]. They differ slightly and will both be described here.

In the first method, "weighted distances" are computed between input characters and each character class. The distance between an input word and a dictionary word is computed as the sum of the distances between corresponding characters. The dictionary word with minimum distance provides word level recognition.

This technique is extended to consider the complete address with individual words processed simultaneously. The final decision on routing a mailpiece is done by minimization of an unspecified distance function.

A similar but even more vague postprocessing strategy is discussed in [Stringa, 1980]. Here the system is based on a hierarchy of Resemblance Measures (RM). The lowest level in the hierarchy is concerned with character recognition and produces an intershape RM between an input character and several character classes. The lexical level produces an interword RM between input characters and characters of a dictionary word. The phrase level produces an interlanguage RM between the output of the word level and the syntactic classes "city", "state", and "zip". This is used to verify syntactic correctness of output words. The semantic level is responsible for final recognition. It produces an interconcept RM over a sentence (city/state/zip combination) with routing presumably determined by the city/state/zip combination with the minimum interconcept RM.

Regardless of the terminology used to describe these techniques, they both appear to be standard postprocessing methods. There does not seem to be anything to distinguish them from other approaches that match a set of word hypotheses to a dictionary of address words and determine routing based on a distance function over the scores of word matches. However, this conclusion is based on the vague information provided in the references cited above. A more quantitative set of conclusions is not possible without further information.

4.1.5. Elsag Architecture

The ELSAG postal address reading machine is realized by the SARI implementation (Automatic Address Recognition system). This is a multiprocessor architecture that can process over 40-50,000 mail items per hour and has been operational in the French and Italian Postal Administrations as well as in the USPS. Each SARI consists of a relatively universal parallel processing system called EMMA (Associative Multi Mini Processor). EMMA was designed for pattern recognition applications and has been used for microfilm and credit certificate reading, speech recognition and for traffic control as well as for the postal OCR environment of SARI [Manara and Stringa, 1981], [Stringa, 1983].

Two major design objectives were realized in EMMA. The first was the utilization of fault tolerance, i.e. the decrease of critical through-points by using few problem-specific hardware. This allows graceful degradation of the system so that when a hardware fault occurs, recognition performance and not throughput is affected. The second objective was the implementation of numerous parallel processors to enhance the computing power of the system. The number of processors (called Associative Units [AU]) is variable between installation and application ranging from 10 to 110 (the USPS EMMA system contains 46 AU).

These objectives result in a modular system suitable for cheap mass production that is able to perform fast, efficient, and reliable calculations at high speed. EMMA is also a highly flexible real time data processing system with a wide range of applications.

An EMMA configuration consists of a central monitor and "families" of interconnected AU's. The monitor, a DEC PDP-11 minicomputer, supervises the activity of the entire system, including the peripherals common to all families, and provides communication, fault detection and reconfiguration functions.

A family is formed by up to 128 AU's, connected by a family bus. Communication between AU's in the same family is controlled by a Data Exchange Coordinator (DEC). The DEC also controls communication to the other family subsystems through the transmitting (TRA) and receiving (REC) interfaces on the family bus. The family bus is designed for direct data access (DMA) message transmission where each AU is polled in daisy-chain fashion until the requesting AU is located. Each message sent between AUs of the same family can be up to 1023 words long and is transmitted either word by word (with a possible delay between words) or as a continuous stream.

Message transfers take place in two steps. The first one transfers the message from the originating AU to the

DEC for storage in a data storage register(SD). The message is then transferred directly from the SD to the receiving AU. Other information used by the DEC includes:

1. address of the AU requesting the transfer (AU_0);
2. the addresses of the transmitting and receiving AUs;
3. the number of blocks to transfer;
4. starting locations in the transmitting and receiving AUs;
5. exchange frequency.

If no message exchange is in progress or an exchange has been suspended due to the unavailability of an AU, the control unit of the DEC waits. When a transfer is initialized, AU_0 coordinates the transfer with the DEC by transfer of the information in 1-5 above. If either the transmitting or the receiving AU are active and the new request is of lower priority, the request is suspended. If it is of higher priority, the active message transfer is suspended and the new transfer initiated. If neither the transmitting or receiving AU are active the message is transmitted in a time slice manner without suspension. At the completion of a message transfer, a termination code is sent to the requesting AU. Each AU contains a processor, a dual-port memory and an I/O port for communication to family and AU specific peripheral devices. The AU can also be connected to an external memory if more storage is needed. It has 32 normal instructions associated with most other processors using a program counter (PC), an accumulator (A), and an index and count register. These instructions include arithmetic logic operations, I/O, branching, shifting, and incrementing. Two special instructions are included for associative searches over memory words, even if the memory is not an associative memory. These instructions provide a fast means for performing a comparison between bit patterns needed in the SARI application. These distance measures can be implemented in Parallel using a separate AU for each measure. Each distance measuring AU is controlled by a supervisor AU which collects and evaluates the final results, providing an efficient means of template matching.

EMMA's operating system (EMOS) manages communication between AU's (including the monitor; the AUs perceive the monitor as another AU) and is hierarchically partitioned into the AU, DEC and the monitor. To facilitate this communication, two table data structures are present in every AU. The MAPPING TABLE acts as a symbol table that translates the symbolic names of routines to addresses within the host AU allowing easy reconfiguration and limiting the accessibility of a routine to those locations stored in the table. The second table, a FLAG TABLE, is used for storage of synchronization flags between routines (and AU).

Fault diagnostics are done by using intrinsic hardware redundancies to limit the influence of faults. EMMA contains both hardware and software tools for fault identification and for system reconfiguration. These include hardware error detection as well as both distributive and centralized diagnostic procedures. The diagnostic procedures monitor the state of either the AU (for distributive diagnostics) or the monitor (for centralized diagnostics). The distributive diagnostics are resident in each AU and include diagnostic tools for testing each AU, software error detection and network error detection. Centralized diagnostics regulate the performance of each of the AU diagnostics that are sent to the monitor. For the SARI application, diagnostic mechanisms include time-out of an AU not returning a result to the controlling unit and table verification.

In current applications, EMMA obtained an availability rate of 99.76% (in an EMMA configuration of 106 AU).

In the standard SARI configuration, three families are used. After an envelope is transported past the scanning unit, the image is sent through a high speed bus to the first family where it is filtered and segmented. After leaving this family, the image is transported to a family where similarity measures are calculated for each character. The output of the measurement family then enters the contextual analysis family where the final recognized characters are sent to the monitor. They are then redirected to the transport control system where routing is performed based on the results of the context analysis stage.

4.2. Burroughs - Nippon Electric

The technical specifications of the NEC equipment, several stages of preprocessing, its character recognition component, and its postprocessing system are discussed in this section.

4.2.1 NEC Technical Specifications

Scanning:

prescanning is by a 512 bit self-scanned array with a pixel size of 0.125mm. Primary scanning is by flying spot that has a minimum pixel size of 0.09mm.

Character Separation:

Character Dimensions:

2mm <= height <="6mm.</P">

Line Spacing:

should be 0.5 mm or more.

Skew:

up to 5 degrees is tolerable.

4.2.2 NEC Preprocessing

Several preprocessing techniques have been patented by NEC. Methods for address block location on windowed envelopes, prescanning for Japanese mail postal code location and character segmentation are discussed here.

4.2.1.1 Prescanning for Address Block Location

A method of address block location on windowed envelopes is described in the patent [Miura,1979]. This method is based on the increased reflectivity of transparent window material as opposed to the regular background of an envelope and the hypothesis that if this area of increased reflectance is detected then the address block will be contained in it.

When the leading edge of a mail item passes a control point in the device, a beam of light from the scanning unit successively sweeps the item in a top to bottom manner as the item moves through the transport unit from left to right. This beam of light can be received by one of two optical receiving units. The first of these units is responsive to the window portion of the item and is positioned to receive the regularly reflected light beam produced by the paraffin or cellophane film covering the window. The second optical receiver will catch the irregularly reflected light produced when the window portion is not present. The first receiving unit will then record the height of the reflected light with respect to the leading edge position received from the point signal processing unit. This is done with a linear photo-diode array aligned parallel to direction of the transport unit that is responsive to the beam of light reflected from of a window bearing mail item. The array elements will be "off" if the window beam does not reach it and will be "on" otherwise. This information is used to compute the coordinates of the window. If a window is detected on the item then its coordinates are sent to a character recognition scanner, otherwise, another method of address block location, such as that discussed next, is used.

4.2.1.2 Prescanning for Postal Code Location

A pre-scanning technique is described in [Isono,1977] that locates the area of an envelope image containing a postal code using information from the prescanner. The coordinates of the postal code number are then sent to a high resolution scanner that gathers data for the character recognition circuit. This enables a smaller area of the envelope to be processed using high resolution image data, thus saving time and capacity.

The pre-scanning section comprises a lamp, a 128 bit photodiode array and some circuitry. A clock pulse activates the photodiode array. The resultant stream of information (in bits) is fed to a pattern shift register which consists of 27 X 132 bits. In addition to the 128 bits from the photodiode array, 4 blanking bits are used, presumably as a delimiter. The photodiode array output is binary. A logical 1 signifies a dark spot and a logical 0 signifies a white spot. The shift register is connected to word end detectors, a height detector, and a mask detector. Each of the word end detectors is a logic circuit with inputs from selected stages of the shift register. Each word end detector acts as a mask of fixed shape. As the stream of bits moves through these masks, recognition is triggered when a match is found. These masks are basically of two types--a left mask and a right mask of varying heights. The purpose of these detectors is, as the name implies, to detect the beginning and ends of words. The height detection circuit outputs a signal depending upon the height of the field of postal code

numbers:

- S(mall) 2.5 ~ 3.43 mm.
- M(edium) 3.43 ~ 4.68 mm.
- L(arge) 4.68 ~ 5.93 mm.
- E(xtra Large) 5.93 ~ 6.86 mm.

Counters (one for the X location and one for the Y location) record the position of words. A line is assumed to be a postal code when only one word is found on that line. The postal code is corrected for skew by plotting the word ends in the X and Y directions. Transformations are then performed in increments of +1, -1, +2, -2 degrees, according to the equations:

- - $X = x \cdot \cos(\theta) - y \cdot \sin(\theta)$
 - $Y = y \cdot \cos(\theta) + x \cdot \sin(\theta)$

to bring the postal code into horizontal alignment. The character recognition circuitry receives X,Y coordinates for the postal code from the prescanner. This area is then scanned with higher resolution, and it segmented into individual characters as described below.

4.2.1.3 Character Segmentation

A technique for segmenting a frame containing three letters of approximately equal height and width is described in [Miyazaki,1980]. This technique uses both a grey level and a binary representation for the frame.

The binary image is first narrowed down on both the left and right extremes until the first and last column are no longer entirely white. The resulting frame is then divided into three approximately equal areas where each division line would be an approximate segmentation of the three letters. Then a range of C columns on either side of the two divisions are searched in the binary image. If the range contains an entirely white column followed (not necessarily immediately) by a non-white column, then a proper segmentation has been made. If an entirely white column is not located in this range then segmentation is done at the column in the multi-level image whose sum of grey-level values is minimum over a similar range of columns.

If one segmentation has been made, a new initial segmentation is calculated for the remaining two-letter block as the difference of its left and right extents and the procedure described above is run on this image. After each letter has been segmented, the upper and lower portions of its frame are collapsed until a row of the letter contains a non-white element. The segmented character that results is then directed to the recognition unit for processing.

This technique can be implemented easily and efficiently in contrast to a feature extraction method. The segmentation, however, in the multi-level domain is purely artificial and can result in a one-letter segment containing portions of a neighboring letter.

4.2.3 NEC Character Recognition

The character recognition method referenced in [Hoshino,1977] as being used in NEC mail processing equipment is discussed here. Input to this approach is from a vertical raster containing 68 elements, where a document travels 10 mils in the period of a vertical scan. The analog signals are reduced to binary signals that correspond to the black/white regions of the document. The next step provides for the consolidation of the black/white signals into a relatively small number of binary bits. The black/white signals obtained in groups of four adjacent scans, referred to as cells, are consolidated into a single binary bit for that cell. The decision regarding the black/white nature of a given cell depends on five basic factors. These are :

- - - The occurrence of a short vertical line element in the cell.
 - The occurrence of a horizontal line element in the cell.
 - The total black signal in the cell.
 - The black/white decision for the preceding horizontal adjacent cell.

- The black/white decision for the preceding vertical adjacent cell.

When these data reduction rules are applied, the area scanned is reduced to a 17 X 10 array of cells.

The binary bits corresponding to each cell are entered into a 170 bit shift register. The occurrence of each succeeding cell causes all of the bits in the shift register to be shifted one position. In this manner, each character is shifted into every possible position in the register. This makes it possible to accommodate all character registrations by implementing the recognition criteria for one registration and testing them after every shift of the register. The recognition criteria define acceptable patterns for each character by explicit and-or logic connected to appropriate positions of the shift register. All character recognition criteria operate in parallel. As the logic is developed, conditions are introduced to make specific discriminations between characters. Whenever a common feature is required in several characters, it can be defined as a zone condition implemented independently, and then used as a separate entity in the logic of each of the characters. These common features consist of line elements, white areas, or other configurations of discrete black and white cells.

The performance of such a character recognition unit is limited by the font used. Typically, this technique cannot be used in multiple font recognition which is essential for address recognition. The hardware design is ad hoc and allows no flexibility. It should be noted that the reference to this technique as being part of the NEC machine is unclear. It is the opinion of the members of this project that a more sophisticated method would be needed.

4.2.4 NEC Post Processing

A contextual post-processing system for word recognition is described in [Hoshino,1977]. Word comparison is performed directly in hardware and a reliability measure is computed based on the extent of the match.

The block number of every word is computed based on the following coding scheme :

A, B	0
C, D	1
E, F, G	2
H, I	3
J, K, L	4
M, N, O	5
P, Q	6
R, S	7
T, U, V	8
W, X, Y, Z	9

An index table address is formed by concatenating the codes for the first three characters of an input word. The contents of that location in the index table is the address of a block of dictionary words with the first three characters corresponding to those codes.

An input word is compared with the words stored in these blocks. A distance is computed by incrementing a counter to determine the degree of match. If a threshold is crossed, then the word is assumed to be recognized correctly.

The number of different addresses which may be read by the OCR system from the various postal matters handled is 10,000. The reliability of single character recognition is 90% . Hence the probability of two or more characters being recognized correctly is 97.2 % . This justifies the use of basing the hash code on the first three characters in a word.

4.3 AEG-Telefunken

The recognition process of the Telefunken equipment is documented in several papers in the open literature. No patents were located for this equipment. A multifont word recognition system developed for postal address reading is described in [Schurmann, 1978] and [Doster, 1978].

4.3.1 Telefunken Technical Specifications

Scanning:

512 bit self scanned array. Pixel size is 0.12mm by 0.12mm.

Character Separation:

Either by match to fixed pitch information, or by vertical column of white space.

Character Dimensions:

height \geq 1.8mm.

Line Spacing:

should be 0.23mm.

up to 25 degrees tolerable for lines with less than 29 characters.

4.3.2. Telefunken Preprocessing

For outgoing mail, the bottom line of the address is read. For incoming mail, the second to last line containing street and house name is also read. An input image is a 512 x 1024 b/w raster. Preprocessing consists of suppression of underscoring lines, line-skew correction by shearing, and segmentation by a hierarchy of different segmentation procedures especially tailored for fixed and variable character spacing. Centering to the center of gravity, and normalizing of various stroke widths and sizes are also done.

4.3.3 Telefunken Character Recognition

The input to the Single Character Recognizer (SCR) is a stream of 16 X 16 rasters each of which represents a character. This raster is converted into a 256 element measurement vector (v). There are 'K' classes to be discriminated. In this particular example, $K = 32$ for the alphabetic channels, and $K = 16$ for the numeric channel. Each of these channels is an independent Single Character Recognition Unit.

This SCR system uses a least mean-square classifier system with a quadratic discriminant function. The binary measurement vector ' v ' described above is mapped onto a feature vector ' x '. This mapping is defined by a combination of heuristic and mathematical methods. If a truly quadratic discriminant function were used, the number of terms in the ' x ' vector would be extremely large. Typically, around 1200 elements are chosen.

- $x(v) = (v_0, v_1, v_2, v_3, v_4 \dots v_1v_2, \dots v_nv_n)$
 - $(x_0, x_1, x_2, x_3, x_4 \dots x_n \dots \dots x_m)$

Based on these terms, a set of discriminant functions can be designed that are suitably combined to form the K dimensional discriminant vector d .

- - $d = (d_1, d_2, d_3, \dots d_K)$

Each of these K discriminant functions d_i is defined to be a linear expression in x

- - $d_i = a_{0i} + a_{1i}x_1 + a_{2i}x_2 + a_{3i}x_3 + \dots a_{mi}x_m$ ($i = 1, \dots K$)

In expanded form, the discriminant function can be written as

$$\begin{array}{c}
 | a_{01} a_{02} \dots a_{0m} | \\
 | a_{11} a_{12} \dots a_{1m} | \\
 | \dots | \\
 | a_{K1} a_{K2} \dots a_{Km} |
 \end{array}
 \begin{array}{c}
 | x_1 | \\
 | x_2 | \\
 | \dots | \\
 | x_m |
 \end{array}
 =
 \begin{array}{c}
 | d_1 | \\
 | \dots | \\
 | d_m |
 \end{array}$$

The matrix of coefficients is termed the A matrix and is optimized using the least mean-square technique. The coefficients are calculated based on several hundred thousand characters extracted from live mail. The number of coefficients to be calculated is roughly 80,000 for all the channels. Once the A matrix has been computed, calculating the discriminant function is a matter of matrix multiplication. The maximum value in the discriminant vector is taken to be class of the feature vector. Typically, reliable vectors are accompanied by small distances between d , the discriminant vector and y , the objective vector. The objective vector is a vector with K elements, with zeros in all positions except that position corresponding to the decision class. The square of the error vector

$$r = y - d$$

decision

is used as a figure of reliability. This measure is used to control the number of choices that are transmitted from the single character recognition system to the contextual post-processing system. Three thresholds are determined empirically to control the number of choices to be presented to the contextual postprocessing unit.

The BOND procedure outlined in the summary of the paper "Multifont OCR Post-processing system", by Rosenbaum and Hillard is used to distinguish between numeric and alphabetic subfields. For each character position, the Single Character Recognition System offers a set of alternatives. The JOKER is offered if the characters discriminated are unreliable. Using these positional alternatives, we can form word alternatives. For each word alternative, the probability of the word being correct is computed

- - - $p(w) = \text{Product}(P_i)$
 - $i = 1$ to word length.

Taking the natural logarithm of this product, a confidence measurement is determined for each alternative. The set of alternative words is ranked according to confidence values.

4.3.4 Telefunken Post Processing

The set of words so determined is now looked up in a dictionary of postal words. Two hashing functions, HF1 and HF2 are used to access the dictionary. HF1 is used for the left half of the word, and HF2 is used for the right half of the word. The reasoning here is that it is unlikely that both the left and right halves of the word are incorrect. The hashing functions are a function of the word length and the characters in the word. In the first pass, word length is assumed to be preserved. The first hashing function HF1 produces an index i into an index table IT1. The address stored in cell i of IT1 points to an entry within a separate storage area. Each entry of this storage area consists of two link fields, a field for the place name and a field for additional postal information. The lengths of the fields for the place name and the postal information are variable. The entry accessed by the hashing function is the word sought or the first word of a sequence of collisions. All collisions of one IT1 place are linearly linked in the data-storage section.

If a match is not found, segmentation errors are taken care of by assuming that the actual word length is one more than the calculated value. If this fails to produce a match, a word length of one less is assumed. Different word lengths result in different hashing indexes. If a match is found, then the city and state are verified with the

Zip code. In case the word has any Jokers, letters are substituted based upon the frequency of occurrence of the letter in the dictionary. The number of different letters substituted is controlled. In precisely the same manner, the number of words that are looked up in the dictionary is pre-specified to optimize system performance.

4.3.5 Telefunken Conclusions

The main drawback of the Single Character Recognition technique is the large amount of computation time required to perform matrix multiplication. The order of matrix multiplication varies between 2.5 and 3 depending upon the method used. However, it should be noted that the multiplication could be reduced to addition in an implementation. Storage, to a lesser extent, is not so much a problem because a sum total of 80000 coefficients have to be stored.

The algorithm assumes strictly numeric or strictly alphabetic subfields. However, when dealing with Canadian Zip codes, mixed subfields must also be taken into account.

This technique borrows many ideas from the IBM technique for multifont character recognition. Notably, the idea of directing the input raster to three different channels is an extension of the IBM two channel technique.

4.4. Toshiba

Only the technical specifications could be located for this machine. They are reproduced below.

4.4.1 Toshiba Technical Specifications

Scanning:

128 bit self scanned array for three data lines. Pixel size is 0.15mm by 0.125mm.

Character Separation:

Touching digits are tolerated but touching letters are not.

Character Dimensions:

height \geq 2.0mm; width \geq 1.0mm.

Line Separation:

at least 0.10mm.

Up to seven degrees tolerated.

4.5 Recognition Equipment, Inc.

Several patents were located for REI equipment. Because of a difficulty in determining the pertinence of this material to Task 1, it is not evaluated here. Only the technical specifications are presented.

4.5.1. REI Technical Specifications

Scanning:

512 bit self scanning Reticon array.

Character Separation:

Characters may touch if font is fixed pitch.

Character Dimensions:

height \geq 81 mils.

Line Separation:
approximately 19 mils.

Line Skew:
Up to six degrees is tolerable.

CONCLUSIONS AND FUTURE WORK

Several techniques for image processing, character recognition, and contextual postprocessing employed in commercial mail sorting equipment were reviewed. This was based on USPS-supplied background documentation, published literature and patent literature. The background material provided the history of postal OCR and USPS-sponsored OCR research efforts. Published and patent literature review provided insight into the specifications and techniques of the Elmag, Nippon Electric, Telefunken, Toshiba, and Recognition Equipment, Inc. address reading equipment. These efforts have provided a solid base of information for the next phase of this investigation in which the state-of-the-art in technology applicable to OCR will be studied.

The preprocessing component of present machines involves scanning, address block location, and character segmentation. Among the techniques studied, most rely on heuristic decisions. Most notably, address block location is based on primitive image characteristics such as the reflectivity of a window region rather than higher level information such as a preclassification of the components of a probable address block.

Character recognition techniques employed in present machines vary widely. The Elmag approach relies on a syntactic description of an input pattern and a multi-level parsing procedure to provide recognition. The Telefunken method is based on a quadratic discriminant function approach that requires training on a large set of data. The NEC method is supposedly based on a template matching procedure. These methods span a wide range of character recognition algorithms from the most-powerful to the most-basic.

The postprocessing methods of postal equipment include a basic attempt at substitution error correction by Hamming distance by the NEC equipment as well as a more powerful consideration of single substitutions, insertions, and deletions by the Telefunken algorithm. The use of context in these algorithms is limited to string error correction. More sophisticated approaches to utilizing context are now available but are not implemented in these machines.

Generally the preprocessing, character recognition, and postprocessing methodologies of the Elmag, NEC, and Telefunken equipment seem simple compared to what is known to be available. This is most likely due to real-time constraints engineered into these machines, i.e., when processing 12 envelopes per second a sophisticated recognition scheme is probably too costly. However, with projected improvements in hardware and software technology, the use of artificial intelligence techniques and multispectral (color) information should allow more sophisticated OCR algorithms to be used in mail sorting equipment of the 1990's and beyond.

ACKNOWLEDGMENTS

The authors would like to thank Jeff Wiedl and Roy Wass of Kenmore Colorplate, Inc. for their assistance in providing the time and facilities needed for laser scanning.

Don D'amato and Sarah Long of Arthur D. Little provided useful references to the published literature.

6. BIBLIOGRAPHY

6.1. Background

1. H.C. Glass and J.S. Gibbons, "Foundry font reading tests and addendum #1 to foundry font reading tests," *Technical Note PTR-6-76*, Electronic Sciences Division, Pattern Recognition and Communications Branch USPS Research and Development Department (January, 1976).
2. H.C. Glass and J.S. Gibbons, "Foundry font reading tests Final Report, APPENDIX A - IBM AOCR character recognition data," *Technical note PTR-6A-76*, Electronic Sciences Division, Pattern Recognition and

Communications Branch, USPS Research and Development Department (March, 1976).

3. H.C. Glass and J.S. Gibbons, "Foundry font reading tests Final Report, APPENDIX B - REI CSOCR II Character recognition data," *Technical Note PTR-6A-76*, Electronic Sciences Division, Pattern Recognition and Communications Branch, USPS Research and Development Department (March, 1976).

4. H.C. Glass and J.S. Gibbons, "Foundry font reading tests final report," *Technical Note PTR-6A-76*, Electronic Sciences Division, Pattern Recognition and Communications Branch, USPS Research and Development Department (March, 1976). Executive Summary.

5. W.J. Matthews, "Zipping the Mail," *optical Spectra*, (August, 1973).

6. B. Levine, "OCR Reject Analysis," *Memo to Martin Sack* (December 28, 1983).

7. G.U. Merkson, "Scanner Review," *Technical Note PTR-02-80*, Electronic Systems Division, Pattern Recognition and Communications Branch USPS Research and Development Laboratories (October, 1979).

8. G.U. Merkson, "Characterization of solid-state photodiode array scanners for postal service applications," *Technical Note PTR-08-83*, Physical Sciences Branch, Engineering Sciences Division, USPS Research and Development Laboratories (May, 1983).

9. R.O. Sheppard, Jr., "Feasibility and implementation of an adaptive recognition technique," *PTR research report*, USPS Research and Development Laboratories (January, 1978).

10. R.O. Sheppard, Jr., "Effects of high density laser printing on OCR performance," *Test Report no. 7ECE01T744*, Pattern Recognition and Communications Branch, Electronic Sciences Division, USPS Research and Development Laboratories (September, 1980).

11. Citations from the Inspec Database, *Automation of the Postal Service (January 1975 - October 1981)*, NTIS Published Search (1982).

12. OCR/CS Operational Test Plan, *USPS Research and Development Laboratories* (April 15, 1983).

13. Optical character Readers in the United States Postal Service, *USPS-Supplied about 1977*.

14. Report on the field testing of commercial OCR's, *USPS supplied approx. 1980*.

15. "Seventh annual report advanced mail systems scanner technology," *NOSC TR 812*, Naval Ocean Systems Center, San Diego, California (May, 1982).

6.2. Elsag

16. G. Barbagelata, B. Contera, M. Giraudi, and L. Stringa, "Electronic Data Processing System with data transfer between independently operating miniprocessors," *U.S. Patent 4099233*, Elsag (July 4, 1978).

17. E. Guiliano, O. Parta, and L. Stringa, "Electronic Character Reading System," *U.S. Patent 4047152*, Elsag (September 6, 1977).

18. R. Manara and L. Stringa, "The EMMA System: An industrial experience on a multiprocessor," pp. 215-227 in *Language and Architectures for Image Processing*, ed. S. Levialdi, Academic Press (1981).

19. L. Stringa, "MARS: A hand and machine-written address reader implemented on a multimini based associative processor," *Proceedings of the Second International Joint Conference on Pattern Recognition*, pp. 443-447 (1974).

20. L. Stringa, "LCD: A formal language for constraint-free hand-printed character recognition," *Proceedings of the Fourth International Conference on Pattern Recognition*, pp. 354-358 (November, 1978).

21. L. Stringa, "State of the art and perspectives of the linguistic-semantic approach to industrial OCR," pp. 325-335 in *Pattern Recognition in Practice*, ed. L.N. Kanal, North Holland Publishing Company (1980).

22. L. Stringa, "EMMA: An industrial experience on large multiprocessing architectures," *The 10th Annual International symposium on Computer Architecture*, (June 13-17, 1983).

6.3 Nippon Electronic

23. Y. Hoshino, "Word recognition apparatus," *U.S. Patent 4010445*, Nippon Electric Company, Ltd. (March 1, 1977).

24. T. Isono and T. Nakada, "Automatic postal-code-number reading system," *U.S. Patent 4034341*, Nippon Electric Company, Ltd. (July 5, 1977).

25. T. Miura and Y. Nishijima, "Arrangement for detecting a window area of a window-having mail item," *U.S. Patent 4158835*, Nippon Electric Co., Ltd. (June 19, 1979).

26. T. Miyazaki and Yukio Hoshino, "Letter Segmenting Apparatus for OCR Comprising Multi-level Segmentor operating when binary segmentor fails," *U.S. Patent 4206442*, Nippon Electric Co., Ltd. (June 3, 1980).

27. K. Sugita, T. Miura, M. Iida, H. Ueda, N. Tsukakoshi, Y. Nishijima, J. Tsukumo, and Y. Hoshino, "Development of Bi-Functional OCR," *NEC Research and Development*, (60) pp. 64-71 (January, 1981).

6.4 Recognition Equipment Incorporated

28. C.H. Carlson, "Character Presence Detector," *U.S. Patent 4135148*, Recognition Equipment Incorporated (January 16, 1979).

29. D.R. DuVall, "Independent Channel automatic gain control for self-scanning photocell array," *U.S. Patent 4157533*, Recognition Equipment Incorporated (June 5, 1979).

30. Dale R. DuVall, "Adaptive Correlator for Video processing," *U.S. Patent 4162481*, Recognition Equipment Incorporated (July 24, 1979).

31. J.D. Erwin, D.R. Duvall, and R.K. Habitzreiter, "Single read station acquisition for character recognition," *U.S. Patent 4013999*, Recognition Equipment Incorporated (March 22, 1977).

32. L.L. Hilley and M.W. Noff, "Character Recognition Unit," *U.S. Patent 4075605*, Recognition Equipment Incorporated (February 21, 1978).

6.5. Telefunken

33. W. Doster, "Contextual postprocessing system for cooperation with a multiple-choice character-recognition system," *IEEE Transactions on Computers*, Vol. C-26(11) (November, 1977).

34. W. Doster and J. Schurmann, "An application of the modified viterbi algorithm used in text recognition," *Proceedings of the Fifth International Conference on Pattern Recognition*, pp. 853 - 855 (December, 1980).

35. J. Schurmann, "A Multifont word recognition system for postal address reading," *IEEE Transactions on Computers*, Vol. C-27(8) pp. 721-732 (August, 1978).

36. J. Schurmann, "Reading Machines," *Proceedings of the sixth International Conference on Pattern Recognition*, (October, 1982), pp. 1031-1044.