Technical Analysis

Alternatives to Optical Character Recognition

As a starting point, OTA examined possible alternatives to optical character recognition technology for postal automation.

OTA concluded that the strongest competition to postal automation is likely to come from electronic mail. If a significant portion of the paper-based mainstream were to divert to either Generation II (electronic input-hardcopy output) or Generation [11 (electronic input and output) electronic mail, then the need for optical character recognition technology would be reduced. However, in a previous (1982) study of electronic mail and message systems, OTA found that, even under very optimistic assumptions about growth of electronic mail, there most likely will be a significant residual volume of paper mail at least through the year 2000.

Therefore, while electronic mail is a strong competitor of postal automation, the major effects of electronic competition are likely to be delayed for at least 10 to 15 years. From this vantage point, there is a window of opportunity for further use of paper-based automation technology.

Some leaders in the optical character recognition industry already recognize that "the current information revolution promises to replace the traditional media on which information has been carried (paper) with electronic media.... Thus, OCR provides a bridge between the Paper Age and. the [formation Age. It is a transitional system which aids users who have one foot in each era. But as the (electronic) Information Age matures, the role of OCR promises to diminish.... The irony is that OCR will be faced with increased opportunities before the electronic axe falls" (Schantz, 1983, p.?).

Optical character recognition technology reads printed alphanumeric characters (letters and numbers) and recodes these characters into machine-readable forms such as a barcode. There are other recognition technologies like magnetic or mechanical, but these seem clearly impractical for conventional mail.

For example, rather than showing addresses in black and white alphanumeric printing on envelopes, the address information could be stored in bits of information in magnetic stripes on the envelope, similar to credit cards and farecards. However, magnetic stripes are not readable by the human eye and would be difficult for senders to encode and receivers to decode. Another form of magnetic code is MICR (magnetic ink character recognition) used on bank checks. This code is readable but lacks alphabetic characters and requires special equipment. Alternatively, a mechanical code could be used. Thus, address information could be stored as punched holes in cards or tape attached to the envelope. These would be difficult to read and would require special punches. ‘

In sum, while there are electronic, magnetic, and mechanical alternatives to optical character recognition, none are both readable and readily and cheaply available as a substitute at the present time. In the final analysis, OTA found that, at least for the U.S. mail, it is as yet difficult to improve on the information carrying ability, readability, and cost effectiveness of printed characters on paper. As long as this is the case, then optical character recognition technology is the technology of choice.

Alternatives to a 9-Digit ZIP Code

OTA also examined whether there are viable alternatives to the 9-digit ZIP or ZIP+4 code.

OTA concluded that there are alternative codes. and, indeed, some are used today by other countries. For example, both Canada and Britain use alphanumeric zip codes, that is, a combination of letters and numbers. Other code schemes have been suggested, for example, using individual telephone numbers as zip codes. Telephone numbers would permit sorting down to the level of each individual street address.

In 1976, USPS considered a wide range of alternative ZIP schemes, including scrapping the 5-digit ZIP, using an alphanumeric code, and adding a check digit (e.g., a tenth digit to the 9-digit code). USPS ruled out any change in the basic 5-digit ZIP, since almost all mail (about 94 percent as of 1976) used a ZIP code. A change in the 5-digit code was judged by USPS to be unfair and excessively burdensome to mailers. This left the alternative of adding 3, 4 or 5 digits to the existing 5-digit codes. USPS elected to add 4 digits. Three was ruled out since this would have required an alpha or alphanumeric add-on. Five was likewise ruled out, since the additional digit, while helping to detect code errors and preventing letters from sorting to the wrong destination, would have increased mailing list information and maintenance cost. (The USPS barcode does contain a correction character.)

At the March 5 OTA workshop, several participants expressed the view that the current 9-digit ZIP was not the best code, but that it was too late to make any major changes. The 5-digit ZIP is almost universally accepted and used (98 percent usage) and the 9-digit ZIP directory is now completed. ZIP+4 codes are being distributed to and beginning to be used by large business mailers.

At this juncture, OTA concluded that there is no realistic alternative. If the ZIP+4 becomes widely used, USPS could consider adding a tenth digit (for error checking purposes) at some future time. If ZIP+4 does not become widely used, alternative codes could be considered.

Performance of Single-line Optical Character Readers (OCRs)

OTA reviewed available data on performance of the single-line OCRs now being installed by USPS. Single-line OCRs read only the last line of an address -- usually containing the city, State, and 5-or 9-digit ZIP code.

[n the 1976-80 period, when the basic USPS automation program was developed, the single-line optical character reader was, in the opinion of USPS, the only proven equipment. Even so, in 1980 when USPS issued specifications for single-line OCRs, no U.S. manufacturer had OCRs meeting USPS specifications. As a result, and to meet USPS domestic content requirements, four foreign companies teamed with U.S. manufacturers who were licensed to produce single-line OCRs.

Two U.S. OCR manufacturers (under foreign license) were selected -- Burroughs (under license to NEC) and Pitney Bowes (under license to ELSAG) -- along with one bar code sorter (BCS) manufacturer, Bell and Howell.

Both OCR manufacturers experienced start-up problems in meeting USPS performance specifications. However, based on review of current performance data and on-site inspection, OTA concluded that the already installed OCRs now essentially meet USPS performance specifications. [illustrative OCR performance data are presented in figure 1.

Figure 1

Illustrative Optical Character Reader (OCR) Performance Data

Burroughs/NEC

Actual Test Results	Accept Rate	Throughput Rate	Error Rate
Meter Belt Mail	80.0%	28,500 pieces/hour	1.50%
Managed Mail	60.0%	31,300 pieces/hour	2.40%
Performance Specs			
Meter Belt Mail	62.1%	28,908 pieces/hour	2.37%
Managed Mail	53.4%	30,083 pieces/hour	2.00%

Pitney Bowes/ELSAG

Actual Test Results	Accept Rate	Throughput Rate	Error Rate
Meter Belt Mail	72.0%	28,500 pieces/hour	1.10%
Managed Mail	62.0%	30,500 pieces/hour	1.10%
Performance Specs			
Meter Belt Mail	67.1%	26,224 pieces/hour	1.50%
Managed Mail	57.8%	26,730 pieces/hour	1.50%

Definitions:

Accept Rate	Percent of letters read by the OCR as a portion of the total fed into the OCR.
Throughput	Total mail pieces fed through the OCR per hour.
Error Rate	Percent of letters sent to the wrong pocket.
Meter Belt Mail -	High quality, generally OCR readable mail- from large mailers.
Managed Mail -	Mail from other Post Offices containing widely variable levels of OCR readable mail.

Source: United States Postal Service: August 1983 2-week tests.

The 252 OCRs already purchased by USPS averaged $645,000 per unit in capital cost, although the Burroughs price was apparently significantly lower (approximately $300,000 per unit) than the Pitney Bowes price.

These prices include all OCR equipment (e.g., mail transport, optical character recognition unit, computer directory, ink jet printer, sorting stackers) plus installation, acceptance (equipment must perform to USPS specifications prior to acceptance by USPS), and 2 years worth of spare parts. There is no obvious explanation for the significant price difference between Burroughs and Pitney Bowes, since both manufacturers bid on the same number of units meeting the same technical performance specifications. USPS apparently split the procurement between two vendors in order to encourage future competition for subsequent procurements, but at an additional cost of about $37.8 million (126 units times the estimated $300,000 price differential between Pitney Bowes and Burroughs).

In the so-called Phase II procurement, USPS intends to purchase an additional 403 OCRs. Competitive bids have been solicited from four qualified U.S. manufacturers (all under license to foreign companies):

U.S. Manufacturers	Foreign Licenser
Burroughs	NEC (Japan)
Pitney Bow-es	ELSAG (Italy)
Recognition Equipment Inc.	Toshiba (Japan)
ElectroCom	Telefunken (Germany)

USPS has budgeted for a capital cost of $660,000 per unit for the 403 OCRs. Bids have been received by USPS, but a selection decision and contract award(s) will not be made before late June 1984.

Performance of Bar Code Sorters (BCSs)

Performance of the BCSs has not been in dispute. USPS has already procured 248 units from Bell and Howell at a capital cost of approximately $129,000 each. The BCSs sort 24,000 to 28,000 letters per hour depending on the type of sort, with an accept rate of 96 percent. USPS plans to procure an additional 452 BCSs as part of the Phase II procurement, and has budgeted approximately $154,000 per unit. USPS has signed agreements with five BCS manufacturers to conduct release-loan tests during the summer of 1984. The manufacturers are:

Bell and Howell (U. S.)
Hotchkiss-Brandt Sogeme (France)
Leigh Instruments (Canada)
National Presort (U. S.)
Telefunken (Germany)

USPS will require that at [east 75 percent of the machine cost be of domestic manufacture. Therefore, any foreign manufacturer whose equipment tests satisfactorily will have to license a U.S. company to produce all or most of the machines in order to qualify for the competitive procurement.

Performance of Multi-Line OCRs

Over the last few years, multi-line OCR technology has emerged from the laboratory and prototype stage to operational units. OTA examined available research and data on multi-line OCR performance. The major difference between single- and multi-line OCRs is that the multi-line machines read up to four lines of the address while the single-line machines read only the bottom line (with city, state, and ZIP).

At present, USPS has no definite plans for use of multi-line OCRs. However, USPS has largely funded the development of a multi-line OCR by Recognition Equipment Inc. (REI), a U.S. company based in Dallas, Texas. REI has a proven track record in optical character recognition technology and is a leading U.S. corporation in high performance OCRs.

There are a total of five prototype REI multi-line OCRs operating at postal installations (two in Chicago; one each in New York, Philadelphia, and Dallas) The REI OCRs are known as RCS/OCR for Read Code Sort/Optical Character Readers.

The computer software of the RCS/OCR is programed so that the address search is "bottom up." That is, the bottom line containing city, State, and ZIP code (5- or 9-digit) is read first, followed by the second line containing the street number and name, followed by the third (and, if necessary, fourth) line containing company name, office building, etc. The address information on the envelope is compared with information maintained in a computerized ZIP+4 address directory. Once a match between the address information on the envelope and in the directory is obtained, a bar code is applied to the envelope, which from then on is sorted automatically down to the carrier level. The multi-line provides additional redundancy since, for example, the street number and name as well as city and State can be cross-checked against the ZIP code.

A direct comparison between single- and multi-line machine performance is difficult, since the USPS has not subjected both machines to equivalent acceptance testing on a comparable mail base. REI had proposed comparative testing, but this suggestion was declined by USPS on the grounds that it would be unfair to other potential multi-line OCR manufacturers and would violate the ongoing competitive procurement process for single-line OCRs. USPS also asserts that the REI multi-line OCR did not

meet USPS performance specifications in 1980, when USPS initially decided to use single-line OCRs, and that test results on the prototype multi-line OCRs were not available until April 1983, after USPS had decided to purchase single-line OCRs.

OTA did not itself investigate and has reached no conclusions on the OCR procurement history. However, OTA did conclude that, as of May 1984, the preponderance of evidence indicates that multi-line OCR performance is essentially equivalent to that of single-line OCR performance for processing 9-digit ZIP mail, and that multi-line performance is substantially better for processing 5-digit ZIP mail to the 9-digit level.

For 9-digit ZIP (ZIP+4) mail, USPS performance data indicate that the Burroughs and Pitney Bowes single-line OCRs and the REI multi-line OCRs all correctly read, code, and sort 98 to 99 percent of OCR-readable ZIP+4 mail. For purposes of mail flow analysis, USPS assumes 100 percent correct reading of ZIP+4 mail.

For 5-digit ZIP mail, both single-line and multi-line OCRs correctly read, code, and sort 98 to 99 percent correctly to 5 digits. However, only the multi-line OCR can read, code, and sort 5-digit ZIP mail to 9 digits.

USPS has estimated that the multi-line OCR can read, code, and sort 60 percent of 5-digit ZIP mail to 9 digits. USPS believes that this 60 percent estimate may be high, since USPS assumed that the total local metropolitan area would be included in the OCR computer directory. If the local directory has less than total coverage, the read-code-sort rate would be reduced. Also, USPS notes that adequate test data are not available on how the multi-line OCR performs two-stage encoding (e.g., placing a 5-digit bar code on a non-local letter at an originating post office and subsequently placing the

4-digit add-on bar code -- or the entire 9-digit bar code -- on the letter at the destination post office).

In contrast, OTA has concluded that the 60 percent USPS estimate may be low for the following reasons. First, full coverage computerized local address directories appear to be technically and economically feasible. USPS already has partial local directories in several metropolitan areas. Second, there is no evidence that two-stage encoding will pose a significant problem for multi-line OCRs. Reasserts that two-stage encoding can be accomplished with no significant degradation in performance. Third, it is reasonable to expect that production model multi-line OCRs would have improved performance compared to the prototype RSC/OCRs. USPS has found that single-line OCR performance improved 5 to 10 percent between prototype and production.

With respect to overall productivity, USPS has concluded that the single-line and multi-line OCRs are roughly equal. Average data from USPS performance reports are shown in figure 2.

In addition to REI, OTA has identified two other companies that manufacture multi-line OCRs: Telefunken (Germany) and ELSAG (Italy). Japanese firms may have the capability and interest, judging from their activity in the single-line OCR market. Other than REI, no U.S. companies are known to currently have multi-line OCR capability. At one time, Control Data Corporation, IBM Corporation, and Ford-Aerospace (Philco-Ford) all had single-line OCR products, and might have been able to develop multi-line OCRs, but left the business in the mid-1970's. Burroughs, Pitney Bowes, and ElectroCom have acquired single-line OCR capability under licenses to foreign manufacturers (NEC, ELSAG, and Telefunken, respectively).

Figure 2

Productivity of Single- and Multi-Line
Optical Character Readers

	Gross Accept Rate (%)	Throughput (pieces/hour)	Productivity (pieces/hour)
Burroughs Single-line OCR	52.8	22,324	8,527
Pitney Bowes Single-line OCR	51.8	19,305	9,127
REI Multi-line OCR	51.3	22,095	10,397

Definitions:

Gross Accept Rate - Pieces of mail accepted by the machine per hour divided by pieces of mail fed to the machine per hour.

Throughput - Pieces of mail fed through the machine per hour.

Productivity - Throughput divided by workhours to arrive at pieces of mail processed per work hour.

Source: United States Postal Service

While it is difficult to extrapolate from foreign experience to U.S. postal needs, due in part to major differences in the mail make-up and use of postal codes, multi-line OCRs appear to be performing well in other countries. ELSAG has 20 two-line OCRs operating in French post offices with an apparently very low error rate (0.1 to 0.5 percent). Telefunken has two- or three-line OCRs operating in Norway, the Netherlands, and Britain. The British Post Office reports that its one Telefunken three-line OCR is undergoing a field trial to be completed by December 1984, and is handling United Kingdom mail at rates between 28,000 and 30,000 letters per hour. Productivity and error rates are not known.

Feasibility of Local and National Directories

[n order to read, code, and sort 5-digit ZIP mail to the 9-digit level, multi-line OCRs require a computerized address directory against which the address information can be compared to ascertain the correct 9-digit ZIP code, then apply the corresponding bar code, and finally sort the letter.

OTA has reviewed the current state-of-the-art in computerized directories to determine if such directories for postal purposes are technically and economically feasible.

Accordingly to USPS, about 40 percent of mail is local and 60 percent non-local. [f a multi-line OCR is to process local mail to the 9-digit level at the originating post office, a local directory is needed. If a multi-line OCR is to process all mail (local and non-local) to the 9-digit level at the originating post office, a national directory is needed.

Until recently, the absence of a local or national directory was a limiting factor for use of multi-line OCRs. However, in the 1981-83 period, USPS completed a national ZIP+4 directory and local ZIP+4 directories for major metropolitan areas. The national directory is stored on a computer in San Francisco and is essentially the sum of all local directories.

In the four cities where multi-line OCRs are already operating (Chicago, New York, Philadelphia, Dallas), the local directories have been partially converted to a format usable by the multi-line OCRs. For example, the Philadelphia multi-line OCR uses a converted local ZIP+4 directory containing about 185,000 local ZIP+4 codes covering more than one-half of the addresses in the Philadelphia metropolitan area.

USPS and OTA agree that the conversion of existing local ZIP+4 directories to a multi-line OCR format is technically feasible. This would be simply an extension of the partial conversions already accomplished in the four cities noted above.

Whereas local directories clearly would be necessary for multi-line OCR operation, OTA was not able to determine whether national directories would offer any significant advantage. With local directories only, non-local mail would have to be processed by multi-line OCRs twice, once at the originating post office to the 5-digit level and a second time at the destination post office to the 9-digit level. A national directory would, in theory, eliminate the need for two-stage OCR processing. After the initial OCR processing, the ZIP+4 bar coded mail could bypass subsequent OCR processing and be handled entirely by the less expensive bar code sorters.

USPS argues that this would not result in a reduction in the number of OCRs, since these machines would still be needed for processing outgoing mail at the destination post

off ice. In addition, USPS points out that when taken together, local directories in all major metropolitan areas would amount to a de facto national directory. Local mail (about 40 percent of the total) would be processed using a local directory in the originating post office, and non-local mail (the other 60 percent) would be processed using local directories in the various destination post offices.

Nonetheless, OTA estimated the technical and cost requirements of a national directory, even though the need for such a directory has not been firmly established.

The memory size of a national directory has been grossly estimated at 20 gigabits or 20 billion bits. A directory of this size would contain all 20 million ZIP+4 codes plus address information including street number and address, city, State, and, where necessary, building floor and suite numbers. The size could be reduced to include only the most frequently used ZIP+4 codes and related address information. For example, if 15 percent of ZIP+4 codes account for 75 percent of ZIP+4 code use, then a memory size of 3 billion bits might suffice.

Currently available magnetic disc memory technology can provide a 3 billion bit capacity at a cost of about $30,000. But the average access time appears to be too long. Optical disc and magnetic bubble memories have similar limitations. Large random access memories (RAMs) have fast access times (microseconds as compared to milliseconds) and may be the best approach. A 3 billion bit directory using 256K RAMs might cost on the order of $300,000 in 3 years (1987 dollars). If five OCRs shared each directory, then the cost per OCR would be about $60,000.

[n sum, a national directory would be technically feasible with memory technology now coming on the market. A full national directory of 20 billion bits of information

would be very expensive (about $4 million each in 1987), even if shared among five OCRs ($800,000 per OCR). If the memory size is reduced to 3 billion bits, the cost would be about $300,000 per memory, or $60,000 per OCR if used on a shared basis (as above).

In contrast, a local directory would require much smaller memory size (e.g., about 72 million bits for the Philadelphia metropolitan area). OTA estimated that the cost of a typical local directory would be about $20,000 (roughly $300,000 times 72 million/3 billion times a multiplier of 3x). At this low cost, sharing a directory among several OCRs may not be necessary. But if shared among five OCRs, the cost per OCR would be further reduced to about $4,000. In all likelihood, the capital cost of local directories would be small compared to either the total cost of multi-line OCRs or to the cost of directory data conversion and maintenance.

Feasibility of Single-line to Multi-line Conversion

OTA reviewed the technical feasibility and cost of converting single-line OCRs to multi-line. OTA concluded that conversion would be technically feasible and that the USPS estimate of conversion cost -- $200,000 per machine -- is as good as can be developed from available information.

Conversion is relatively simple because a large part of the single-line OCR could be retained almost as is, as discussed below.

Letter sorter -- no change.

Letter transport -- no change. The current mechanical transport is designed to move faced and bottom justified letters at a constant speed past an OCR window. This function is common to single- and multi-line OCRs and would not change.

"Prelook" window -- probably no change. The prelook is used to find the address location, and probably could be used as is.

Lenses and scanners -- probably no change if the scan height and pixel resolution are adequate. Otherwise, the OCR hardware from the lens system back (including "prelook") would need to be replaced.

OCR electronics -- probably some change needed to upgrade the image registers, recognition logic, etc. Since the multi-line OCR must process much more address information than the single-line OCR, the basic scanning and clocking rates may have to be increased. If the single-line OCR already captures three or four lines of the address at sufficient speed and resolution, then little or no change in OCR electronics would be needed -- only a change in computer software and directory.

Computer software -- change needed to upgrade the software so that multiple address lines could be processed in a "bottom-up" fashion, and the resulting data properly formatted and queued into the directory.

Computerized directory -- change needed to expand the directory from city, State, and ZIP+4 to include street number and name and, as appropriate, building floor and suite.

Ink jet printer and verifier -- probably no change needed with a local directory. [f a national directory is used, the access and processing time may lengthen to the point where the ink jet printer and verifier (sprays the bar code on the bottom right-hand edge of each envelope) would need to be moved farther downstream in the transport path.

The actual cost of single- to multi-line OCR upgrade can only be determined by detailed engineering analysis and a competitive procurement process. It is possible that the conversion could be accomplished by an OCR manufacturer other than the original source, although this would require a high degree of technical cooperation between the two OCR vendors.

Technical Opportunities for Improved Performance

OTA identified several areas where technical performance of postal automation might be improved in the future. These include bar-coded reply envelopes, mailer printing of bar codes, character recognition upgrades, address format standards, and increased postal research and development.

Bar-coded reply envelopes. Mailers-- and especially large business mailers -- already frequently provide reply envelopes to customers, presumably to facilitate payment of bills such as those mailed out monthly by utility, telephone, and gas companies. Preprinting of the bar code along with the return address on the reply envelope would appear to be cost-effective, and could permit processing of business reply mail by the less expensive BCSs rather than OCRs.

Some business mailers are already preprinting bar codes. But in order for this to be successful, the bar codes must be readable by the Bell and Howell barcode sorters. That is, the color spectrum of the ink and location of the bar code on the envelope must match the capability and location of the photo-detector in the BCS. Also, in order to achieve savings by bypassing some or all OCR processing, the bar-coded reply envelopes would need a unique facing indicia mark (FIM) so that these envelopes could be detected (perhaps by the facer/canceler) early in the mainstream and diverted to BCSs.

Mailer printing of outgoing bar codes.' Most major business mailers already use computers to store, update, and print out addresses on outgoing mail. Addresses are printed directly on the envelopes, on mailing labels, or on letters that show through window envelopes. Almost all business mailers have already entered the 5-digit ZIP codes into their address data bases. USPS would, of course, like them to convert the addresses from 5-digit to 9-digit ZIP codes.

At present, few mailers have actually converted their address files. partly because of concern over the cost of conversion. Some participants at the March 5 OTA workshop pointed out that, if and when mailers convert, consideration should be given to building in a capability to print outgoing bar codes in addition to or as a substitute for the ZIP+4 numeric codes. In principal, outgoing bar-coded mail could, with proper FIMs on the envelopes, bypass the OCRs completely and go directly to BCSs.

In practice, the technical and economic feasibility is unclear. For example, high speed non-impact printers (such as laser printers) could be programed to print the bar code immediately under the last address line on an envelope or label. However, if the barcode location is too far up from the bottom of the envelope, then the Bell and Howell bar code sorters used by USPS would not be able to read the bar code unless a second scanning channel was added. Ink jet printers or special photo offset printing devices could also be used, but this would involve significant equipment procurement and processing changes on the part of mailers. And proper location of the bar code could be difficult when labels or window envelopes are used. Special bar code window envelopes or barcode labels might be needed. Bell and Howell indicated to OTA that addressing and/or inserting machines could be easily used to print bar codes on outside envelopes of outgoing mail, and that 30 to 35 billion pieces of mail annually could be processed in this way. A clear understanding of the potential and pitfalls must await further study.

Character recognition upgrades. The percentage of OCR-readable mail actually read by the single-line OCRs used by USPS, although meeting USPS performance specifications, is still far less than 100 percent. USPS has assumed an average read rate of 70 percent. One way to improve performance is to upgrade the character recognition technology.

Current character recognition technology uses mask matching, whereby character patterns are stored in electronic memory and matched against the actual characters in the address. But because there are so many different sizes, shapes, and forms of alphanumeric characters in U.S. mail addresses, not all characters are stored in electronic memory. If the computer cannot make a match within 100 milliseconds, the address is not read.

The overall read rate could be improved by, first, studying the rejected addresses and determining what types of characters are not being read. Then, OCR manufacturers could be solicited to develop improved character recognition technology that would read some or all of the characters rejected most frequently.

Address format standards. Some OCR rejects are due to problems with the location and format of addresses on envelopes. To receive a ZIP+4 discount, mailers must meet several mandatory requirements. These include: a barcode clear zone (at the bottom right-hand portion of the envelope) in which no printing or markings whatsoever are permitted; an OCR read area in which the city, State, and ZIP+4 code line must be visible and unobstructed (no extraneous printing or markings); machine-printed address with uniform character single and line spacing; and a reasonable degree of color contrast between the address and mail piece.

However, there are several voluntary guidelines with respect to margins; State abbreviations; character fonts; character pitch, height, and height to width ratio; character and line skew; upper case characters; and line and character spacing. Also, while black ink on white paper is preferred, other color combinations are permitted (except for brilliant colors and reverse printing and any others that do not meet minimum reflectance standards).

The overall read rate could be improved by studying the rejected address and determining the cause(s). The most common causal factors could be mitigated by:

1. better enforcement of mandatory format requirements;

2. improved compliance with voluntary format requirements, possibly through incentives, and if necessary, by making some voluntary requirements mandatory; and/or

3* technical upgrades as discussed previously, so that the OCRs can read a wider variety of address characters and formats.

Postal research and development. The USPS record on postal automation R&D is mixed. On the negative side, USPS continues to underspend on R&D, despite repeated recommendations from congressional oversight committees and the Commission on Postal Service to raise postal R&D closer to private sector levels. Postal R&D was about $24 million in 1983, or about one-tenth of one percent (0.1 percent) of revenue. This compares to a U.S. industry average of perhaps 3 percent. Second, postal R&D organization and management appear to have lacked stability, clear direction, and, at times, top-level commitment. Third, despite 20 years of USPS investment in optical character recognition R&D, when USPS solicited manufacturers in 1980 for single-line OCRs, no U.S. manufacturer was judged to be qualified. All U.S. manufacturers previously receiving USPS support for single-line OCR R&D had withdrawn from the market by 1980, in part due to USPS indecision on an automation strategy.

For example, USPS awarded development contracts for first generation OCRs to Philco Corporation in 1960-65, and for second generation OCRs to IBM and Philco-Ford in 1968 and 1969, respectively. Both companies developed successful OCR designs and were subsequently awarded prototype contracts. The Philco-Ford second generation OCRs were installed in Boston in November 1971 and successfully tested in early 1972; and IBM OCRs were installed in New York in June 1972 and tested in late 1972. By the time USPS had settled on a postal automation strategy and solicited manufacturers to provide OCRs on a release-loan basis in 1979-80, the only companies left in the single-line OCR business were foreign manufacturers. This was despite the fact that the Philco-Ford second generation OCR units in Boston remained operational until 1982, and the IBM OCR units in New York are still in service.

Thus, it is at least arguable that USPS could have reasonably opted for wide deployment of single-line OCRs in the early 1970's, perhaps using a 5-digit bar code (5-digit ZIP code use had reached 84 percent by 1972). Had USPS opted for this strategy, some U.S. manufacturers of single-line OCRs might well have stayed in the business. And it would be reasonable to expect that OCR technology would be further advanced than it is today.

On the positive side, USPS has established a good track record in narrowly focused R&D on improvements to upgrade existing equipment. For example, the multi-position letter-sorting machine (MPLSM), in wide use since the late 1960's, has been upgraded several times, most recently by a not yet fully implemented electronic ZIP retrofit (known as EZR) that allows four-digit keying of ZIP+4 codes. Facer cancellers, single position letter sorting machines, and flat sorting machines also have been, or will be, upgraded. USPS equipment upgrades are highlighted in figure 3.

Figure 3

Summary of USPS Equipment Upgrades

Multi-position Letter Sorting Machine (MPLSM)

1969 - ZIP Mail Translator (ZMT) - Converted MPLSMs from simultaneous keying (cordal) to sequential keying. ZIP Codes keyed by operators then could be translated by the ZMT into BIN assignments.

1973 - Engineering Data Isolation Technique (EDIT) - An electronic modification to the ZMT which allowed keyed data to be monitored for accuracy.

1974 - Automatic Density Analysis Profile Technique (ADPT) -- An Upgrade that enabled automatic tabulation of MPLSM sweepside BIN densities.

1976 - Electronic Sort Processor (ESP) - A modification to MPLSMs which replaced mechanical code setting with electronic code setting to provide more accuracy, reduce maintenance costs, and reduce noise.

1978 - Micro-Key - An upgrade to the ZMT which allowed the first digit of a carrier route number to be locked in each time an operator keyed an incoming secondary distribution.

1981 - ZIP Data Logger (ZDL)

1982 - Electronic ZIP Retrofit (EZR)-- A modification to MPLSMs to allow four digit keying of ZIP+4 Codes.

Facer Cancelers

1982 - Micro Mark - A modification to Mark II Facer Cancelers to upgrade the electronics to Solid State circuitry.

Single Position Letter Sorting Machine (SPLSM)

1972 - Automated Business Mail Processing System (ABMPS) - A modification to the Universal Business Machine (UBM) SPLSM to allow automated distribution of destinating Bar Coded Business Mail.

Flat Sorting Machine (FSM 775)

1984 - Software Modification - An upgrade to provide Micro key capability on the FSM 775.

Source: United States Postal Service

Also on the positive side, despite some variability in funding and commitment, USPS has provided enough support over the last 14 years to Recognition Equipment,Inc. (of Dallas, Texas) such that REI has developed one of the leading multi-line OCRs on the world market. Ironically, at various times from the late 1960's to the late 1970's, USPS appeared to actually favor the multi-line over the single-line OCR. [n the late 1970's, USPS procured one multi-line OCR from REI, and as insurance solicited every known OCR manufacturer in the world to provide a single-line OCR on a release-loan basis.

As it turned out, the REI multi-line OCR did not satisfy USPS performance requirements, but the single-line OCRs of five foreign manufacturers did (ELSAG, NEC, Telefunken, Toshiba, and ITT Belgium). USPS decided to deploy single-line OCRs and awarded production contracts to Pitney Bowes (under license to ELSAG) and Burroughs (under license to NEC) in early 1981. According to USPS, for insurance purposes an additional contract was awarded to REI for five multi-line OCRs. These were installed and tested between June 1982 and April 1983. As discussed earlier, USPS test data indicate that the multi-line OCR performance is now fully competitive with single-line OCR performance.

Finally, USPS may wish to consider: (1) new approaches to R&D and procurement (including the release-loan testing process) with a view towards speeding up the time delay from R&D to installation of new equipment; and (2) new ways to organize mail processing in order to achieve faster and more reliable delivery.