OPTICAL IMAGE SCANNERS AND CHARACTER RECOGNITION DEVICES: A SURVEY AND NEW TAXONOMY

Amar Gupta
Sanjay Hazarika
Maher Kallel
Pankaj Srivastava

Working Paper #3081-89

Massachusetts Institute of Technology
Cambridge, MA 02139

ABSTRACT

Image scanning and character recognition technologies have matured to the point where these technologies deserve serious consideration for significant improvements in a diverse range of traditionally paper-oriented applications, in areas ranging from banking and insurance to engineering and manufacturing. Because of the rapid evolution of various underlying technologies, existing techniques for classifying and evaluating alternative concepts and products have become largely irrelevant. A new taxonomy for classifying image scanners and optical recognition devices is presented in this paper. This taxonomy is based on the characteristics of the input material, rather than on speed, technology or application domain.

1. INTRODUCTION

The concept of automated transfer of information from paper documents to computer-accessible media dates back to 1954 when the first Optical Character Recognition (OCR) device was introduced by Intelligent Machines Research Corporation [1]. By 1970, approximately 1000 readers were in use and the volume of sales had grown to one hundred million dollars per annum [3]. In spite of these early developments, through the seventies and early eighties scanning technology was utilized only in highly specialized applications.

The lack of popularity of automated reading systems stemmed from the fact that commercially available systems were unable to handle documents as prepared for human use. The constraints placed by such systems served as barriers, severely limiting their applicability. In 1982, Ullmann [2] observed:

> "A more plausible view is that in the area of character recognition some vital computational principles have not yet been discovered or at least have not been fully mastered. If this view is correct, then research into the basic principles is still needed in this area."

However, ongoing research in the area of automated reading (conducted as part of the wider field of pattern recognition) is leading to new products, and the market for scanning devices has expanded dramatically. Concomitantly, there has been tremendous growth in the market for third-party scanning software, as well as software and peripherals for specialized tasks such as handwriting recognition and data indexing, storage and retrieval [3].

This paper analyzes the technology used in the image scanning and character recognition industry and presents a taxonomy for evaluating the capabilities of current systems. Although the main focus is on optical scanning and character recognition technologies, an attempt is made to identify auxiliary technologies and products as well. In addition, the paper discusses future trends and characteristics of systems which are likely to become available in the next ten years.

Inclusive of this introductory section, there are nine sections in the paper. Section Two presents a discussion of the principal recognition approaches and technologies. The third and fourth sections classify the scanning industry by market segment and function respectively. Section Five discusses ancillary technologies such as scanning software, data storage, indexing and retrieval products and handwriting recognition software. In Section Six, future trends in products and technologies are discussed. In Section Seven, a new taxonomy is developed for evaluating the capabilities of current scanning systems. Section Eight describes a set of representative scanners and applies the taxonomy to judge their performance. Finally, the conclusions of the paper are presented in Section Nine.

2. RECOGNITION TECHNIQUES

The recognition of text, the scanning of images, and the raster to vector conversion of technical drawings have usually been considered independently of each other. The technologies corresponding to these three areas must be integrated in order to accurately scan and process complex technical documentation. In the case of most non - technical applications, only the first two areas need to be considered. A framework that allows recognition of both text and images is presented in this section.

The three major stages in the processing of a document are preprocessing, recognition, and post-processing. These are discussed in the following subsections.

2.1 Preprocessing

Preprocessing is the conversion of the optical image of characters, pictures, and graphs from the document into an analog or digital form that can be analyzed by the recognition unit. This preparation of the document for analysis consists of two parts: image analysis and filtering [4]. The significance of these parts is described in the following paragraphs.

(a) **Image Analysis**

The first stage of image analysis is scanning [5,9,25]. Scanning provides a raster image of the document with sufficient spatial resolution and gray scale level for subsequent processing. In the case of a picture or a graph, the issue of gray scale level is more important that in the case of text. For text, this phase consists of locating character images. With the exception of high end scanners, which employ contextual analysis as well, reading machines are character-oriented. Each character is treated as a unique event and is recognized independently of other characters. This implies that the document must first be segmented into distinct characters, and then the identity of each character recognized separately.

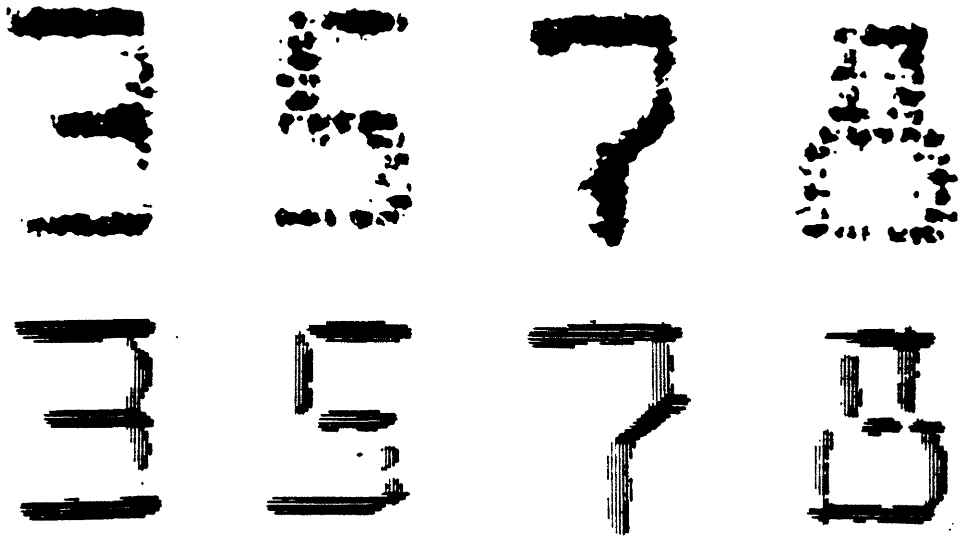
The process begins with the optical system taking a raster image of the area that is supposed to enclose the character. Alternatively, the raster image representing the character is "cut out" of the image of the document. In either case, the image is transmitted sequentially to a single-character recognition subsystem. If the image, or the information on the features of the character constituting the image, possesses characteristics which are significantly different from the characteristics maintained by the character recognition subsystem, then the particular area is deemed to be either an unrecognized character or noise. Depending on the system, the output is expressed as a flag or a blank. In some systems, any graphic or character input which is outside the size limitations is not flagged but skipped. The formatted output, in such a case, contains a blank zone corresponding to the input of the improper size.

(b) **Filtering**

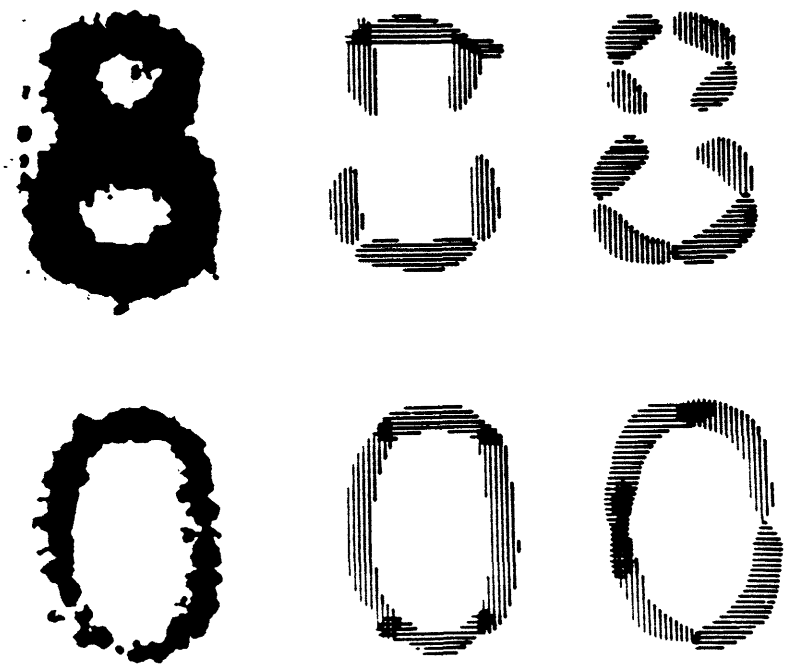
After the image is analyzed, filtering takes place. Filtering minimizes the level of noise. The latter may be caused either in the source document or by the opto-electrical transformation mechanism. The process of filtering also enhances the image for easier recognition. One filtering process that eases the extraction of the features of the character in the recognition phase was proposed independently by Lashas [5] and Baird, et al. [6]. They presented two OCR readers in which black marks constituting the character are transformed into quantized strokes. Their approach is depicted in Figure 1.

Preprocessing: New Approaches

The preprocessing phase consists of deriving a high level representation of the contents of the image. The scanned document is seen as a set of blocks corresponding to independent ways of representing information such as text, line drawings, graphs, tables, and photographs.



CCR-A font numerals and horizontal and vertical quantized stroke segments



CCR-B font numerals and four detected directions of quantized strokes

FIGURE 1: QUANTIZED STROKES

Lashas, A., et. al. "Optical Recognition Based On Analog Preprocessing and Automatic Feature Extraction". Computer Vision, Graphics and Image Processing.vol. 32. pp. 191 - 207.

Understanding of this document involves the following:

(a) Identification of the major blocks of information.

(b) Identification of the spatial relationships between the different blocks (for example, the reading order must be recognized so that the logical connections between different blocks of text or graphics can be easily derived).

(c) Identification of the layout features such as the number of columns, margins and justification.

(d) In the case of text, further identification of headlines, footnotes and other similar characteristics.

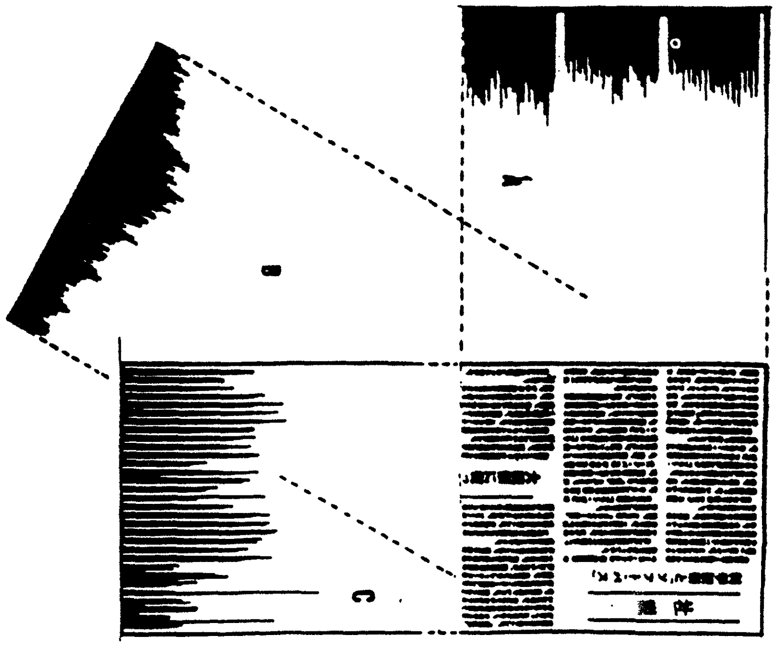
Typically, the textual portion is first distinguished from other information, and then columns of text are recognized. Next, these columns are split into lines of text which are in turn segmented into single character images.

A reader which is able to accept a free format document was described by Masuda et al. [7]. Their scheme of area-segmentation uses projection profiles obtained by projecting a document image on specific axes. Each profile shows the structure of the document image from a particular angle. The projection profile is very sensitive to the direction of projection. In order to check for skew normalization, the image of the document is incrementally rotated and the horizontal and vertical projection values are noted. Through an analysis of the intensity of these values and an examination of different areas, general text and headlines are separated (Fig. 2). Another reading system based on this approach is described by Kida, et al. [8].

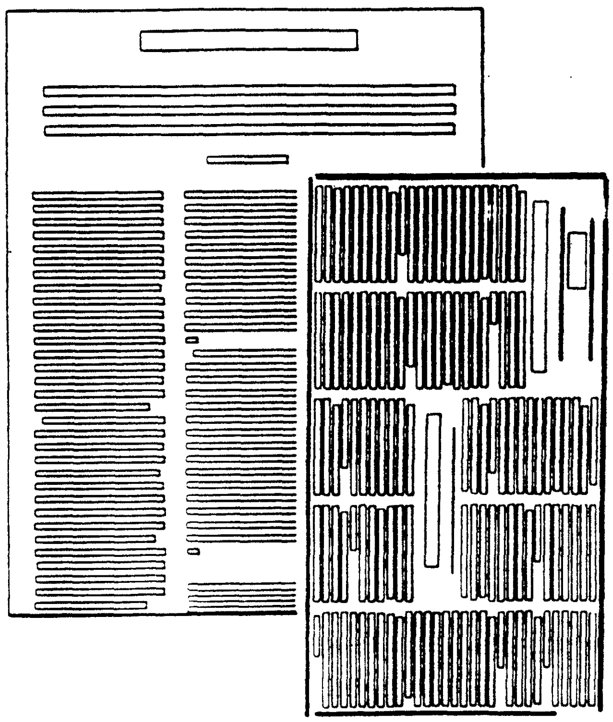
Commercial system today require areas to be manually segmented [2,5,7]. While segmentation methods are fairly rudimentary in low-end scanners, methods that permit the operator to define several text and graphics "windows" within a page are available on high-end products. These methods include the use of a mouse or a light pen to define text, graphic, or numerical zones.

2.2 Recognition

Recognition occurs at the level of characters in most commercial page readers. However, high-end scanners are now complementing character recognition with sophisticated contextual analysis. Techniques used for character recognition and contextual analysis are discussed in the following paragraphs.



Projection Profiles



Text line areas extraction

FIGURE 2: SEPARATION OF TEXT

Masuda, I. et. al. "Approach to a Smart Document Reader". IEEE Conference on Computer Vision and Pattern Recognition. 1985. pp. 550 - 557.

Character Recognition

The character recognition problem is essentially one of defining and encoding a sequence of primitives that can represent a character as accurately as possible. The most common approaches to character recognition are:
(a) template matching, and (b) feature extraction [9].

(a) **Template Matching Technique**

Among the oldest techniques for character recognition, template matching involves comparing the bitmap that constitutes the image of the character with a stored template [9]. The amounts of overlap between the unknown shape and the various stored templates are computed, and the input with the highest degree of overlap is assigned the label of the template.

The primitive in this case is a very simple function that assigns one value to a point of the bitmap if it is black, and another if it is white [5,9,25]. The performance of different readers using this technique depends on the decision algorithms. This method offers high-speed processing especially for monofont pages. Even in the case of documents containing several fonts which are extracted from a limited set of fonts of a given size, this technique can be very effective. Moreover, this method is relatively immune to the noise generated during the opto-electrical transformation.

However, this method is less effective in the case of unknown fonts, multiple fonts and multisize characters. Further, it imposes constraints on the format of the page in areas such as the spacing of the letters and the position of the text. If these constraints are relaxed, template matching becomes slow and costly, because of the need to translate, scale, and rotate input images.

(b) **Feature Extraction**

In contrast to the template matching technique which emphasizes the importance of the overall shape, the feature extraction approach focuses on the detection of specific components of the character [5]. This approach assumes that local properties and partial features are sufficient to define every character.

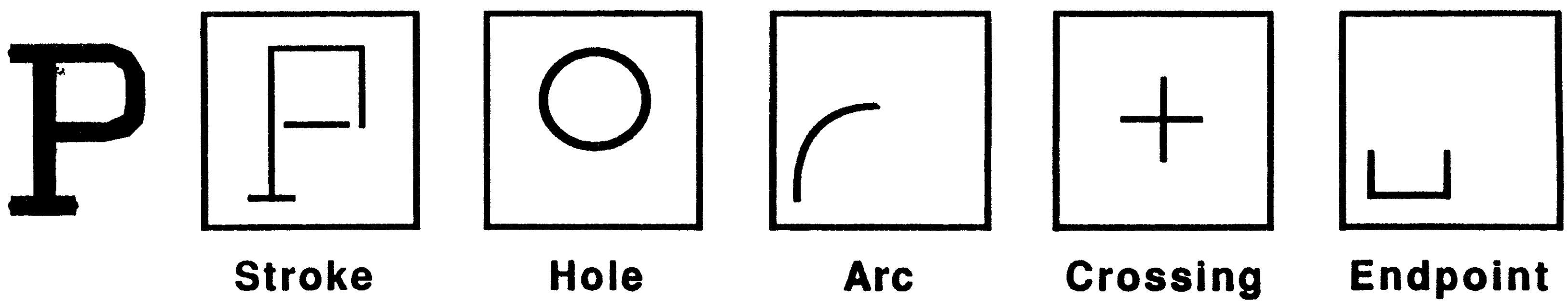
Feature extraction techniques identify local aspects such as pronounced angles, junctions and crossing, and define properties such as slope and inflection points. In one method of recognition, a boolean or a numerical function that characterizes each feature is calculated and then applied to the given image [5,7]. Another method involves defining a partial mask that can be systematically positioned on the pattern [7]. In a more recent strategy, recognition is based on the analysis of the direction and connective relationships of the strokes of the character [7].

Structural Analysis Methods

The use of structural analysis methods is a recurrent theme in the literature [1,4,5,6,7,10]. These methods are frequently utilized in commercial reading machines. Each character is defined as a set of topological primitives such as strokes, segments, holes, and arcs, and these machines analyze the characteristics of the scanned inputs to detect such primitives.

Isolated, unbroken characters are first recognized using a structural description [9]. These descriptions are independent of the position and the size of the character. The shape is then parameterized so that the results of the structural analysis can be compared with a stored list of shapes. Next, the shapes are clustered to discriminate between characters and to classify them. The power of structural analysis depends of the number of features of each character used by the system. An example of the structural analysis of a character is shown in Figure 3.

FIGURE 3: EXAMPLE OF STRUCTURAL ANALYSIS OF A "P"



Character Classification

So far, the discussion has centered on single characters in a single font. Most documents contain multiple characters with considerable variation across fonts. A statistical approach is used for dealing with these variations. In reading machines equipped to handle multiple fonts, tests of character recognition are based on statistical data constituted by a training set and a test set. When the term "omnifont" is applied to a reader, the implication is that the reader has been trained on a large set of fonts [6].

Feature extraction is generally complemented by classification. In order to be classified, characters have to be discriminated from each other. The classifier is designed to construct a discriminant function underlying a complex probability distribution, taking into account the non-linear variations in the characters [26].

The two essential steps in recognition - feature extraction and classification - have different optimization criteria. It is felt that feature extraction should be optimized with respect to the reconstruction of the character, while classification should be optimized with respect to recognition [4]. Consequently, feature extraction should not be directed towards recognition through definition of features that minimize the classification process.

Classifiers are usually adapted to different fonts by the manufacturer. However, some machines provide a classifier which allows for the adaptation of new fonts by the operator. An on-line training capability for multifont recognition can also be provided [4]. This concept of trainability is similar to that employed in modern speech recognition systems, which overcome the variability of voice between speakers by requiring that adaptation be performed during a training phase prior to regular operation. The main disadvantage of trainable systems lies in their slow training speed, as the training set contains several thousands of characters.

Limitations of Current Recognition Techniques

Ideally, feature extraction techniques should be able to generate a set of characters that are independent of font and size. However, the wide variety of fonts encountered in office environments results in a huge number of possible characteristics. Furthermore, ambiguities occur due to the similarity in the features of different characters across fonts. For example, the distinction between 1 (one) , I (capital I) and l (el) across and even within fonts is not obvious. Moreover, there are a number of characters which cannot be correctly identified without further information about their size or position. For example, the character "O" (capital o) in one size corresponds to the lower case "o" in a larger size; further, this character is the same as 0 (zero) in several fonts [11].

The user can control the accuracy of the system by adjusting the error and reject rates [12,25]. The error rate is the percentage of the characters in a document that are incorrectly identified by the system. The reject rate is the probability that the system cannot identify a character at the desired level of confidence (i.e., if the reject rate is 5%, the system will flag 5% of all characters, on the average, as being unrecognizable). Clearly, there is a tradeoff between the error rate and the reject rate - if a lower reject rate is selected, the error rate will increase and vice versa [12,25]. Consequently, a reader may possess a very low error rate, but it may flag or reject every character that does not offer a high probability of correctness as defined by the decision algorithm [13].

There is naother aspect to accuracy. For example, a reader may recognize the characters in "F16" correctly, but then flag them as rejects because the word is not in its built-in dictionary. Also, the inability of readers to distinguish between ambiguous characters such as "I" (capital ai), "1" (el) and "1" may or may not be considered to be an error.

The error rate of the entire recognition process is dependent not only on the functioning of the single character recognition subsystem but also on the level of preprocessing [12]. For example, different kinds of segmentation errors may occur, and lines of text may be missed or misaligned in the case of a document containing several columns. Situations involving such missed or misaligned lines can be minimized by preprocessing (described in Section 2.1).

With respect to performance evaluation, character recognition is sometimes deemed to be a branch of empirical statistics [4]. There is no reliable way of modelling the accuracy of a reading machine except by comparison with a standard set of norms. The impracticability of statistical modelling is due to the fact that the pattern generating process and its multivariate statistics are influenced by a number of barely controllable, application-dependent parameters.

Recognition techniques are prone to ambiguities, apart from involving heavy computation. The sole reliance on the physical features of the characters results in a high rate of errors and rejects. To combat this problem, recognition methods are complemented by contextual and syntactical analysis aided by customizable lexicons, especially in high-end systems [9,11,13].

In the case of low quality documents with different kinds of noise and broken and touching characters, the error rate is high. Contextual analysis becomes a necessary condition for reducing the rate of errors and rejects in such situations.

Separation of Merged Characters

The breaking of images into character blocks is based on the assumption that each character is separated by horizontal or sloped line of blank spaces. However, in the case of tight kerning, inadequate resolution of the scanner, poor quality of the document, or high brightness threshold, adjacent characters spread into each other. Separating merged characters involves the following three processes:

(a) Discriminating multiple characters from single character "blobs";

(b) Breaking the blobs into components that are identified as single character blobs;

(c) Classifying the different characters and deciding whether to accept or reject the classification.

Defining a set of criteria for distinguishing between adjacent characters is difficult because of the many ways in which characters merge together and the fact that merged characters contain misleading strokes. Separation of characters is a computationally intensive task that significantly slows down the overall recognition process [2,4,8,12]. Separation is usually possible in a limited number of cases and for pairs of characters only.

Contextual Analysis Techniques

Contextual analysis is of two types: layout context analysis and linguistic content analysis [9]. The former covers baseline information on the location of one character with respect to its neighbors. For example, it generates formatting information and is usually language-independent. Linguistic analysis on the other hand involves the use of spelling, grammar and punctuation rules. For example, a capital letter in the middle of a word (with lower case neighbors) is not accepted. Layout context analysis capabilities are available on several systems available today. However, only the high-end systems offer some degree of linguistic analysis capabilities.

Dictionary Lookup Methods

Several contextual methods for word recognition are based on the analysis or comparison of a word or string of characters with stored data for type and error correction. Spelling checkers are commonly used to complement character recognition devices [13]. Such checkers may be part of the recognition software, or they may be part of the text composition software that the operator uses for correcting the processed document. In either scenario, when the size of the dictionary is large, the search time can be very long. Furthermore, if the contextual information is not considered, the scanned word may be incorrectly converted into some other word [13]. Several high speed correction methods use the concept of "similarity measures", which are based on the fact that most errors are due to character substitution, insertion, or deletion [22]. The similarity between two words (the correct word and the garbled word) of equal length is measured using the Hamming distance. The Modified Levenstein distance [9] generalizes the similarity measure for substitution, insertion and deletion. The minimization of this distance forms the optimization criterion. Tanaka [14] proposed two methods that yield 10 - 35% and 35 - 40% higher correction rates than a typical dictionary method and reduce computing time by factors of 45 and 50 respectively. One of these methods is based on the assumption that different characters can be classified into classes or groups that are independent of each other, so that a character in one class is never misrecognized as a character in another class. This categorization helps to significantly reduce the probability of errors.

Use of Statistical Information

The frequency of occurrence of different characters varies widely. For example, the letter "e" has the highest frequency of appearance in an English document. Also, the letter "t" is the most frequent first letter of a word, and the letter "q" is always followed by a "u." Further, several sequences of character combinations are more likely to appear than others. The frequency of occurrence of a character within a given string of text can be efficiently modeled by a finite-state, discrete-time Markov random process [22]. The use of statistical distributions of different combinations of N characters (N-grams) allows the implementation of error correction algorithms such as the Viterbi algorithm. According to Hou [9], this algorithm can reduce the error rate by one half.

Linguistic Context Analysis

Characters are primitives of strings constrained by grammatical rules. These rules define legitimate and non-legitimate strings of characters. Based on the recognition of some words, the class (noun, verb, etc.) to which they belong can be identified along with the applicable syntactic analysis to identify misspellings, misplacements, and other syntactic errors. A string of words or a sentence can be decomposed using a parsing tree. There are several efficient parsing trees for different types of grammatical structures [9,12,15].

Syntactic analysis assumes that the text is constrained by one type of grammar. This assumption need not hold in all cases. Several technical documents contain uncommon words: serial numbers, technical words, and abbreviations. Such situations require interaction between the technical operator and the system, or the use of dictionary lookup methods. Except in such special situations, linguistic context analysis techniques are well suited to the task of identifying and correcting reading errors [15].

Future of Recognition Technology

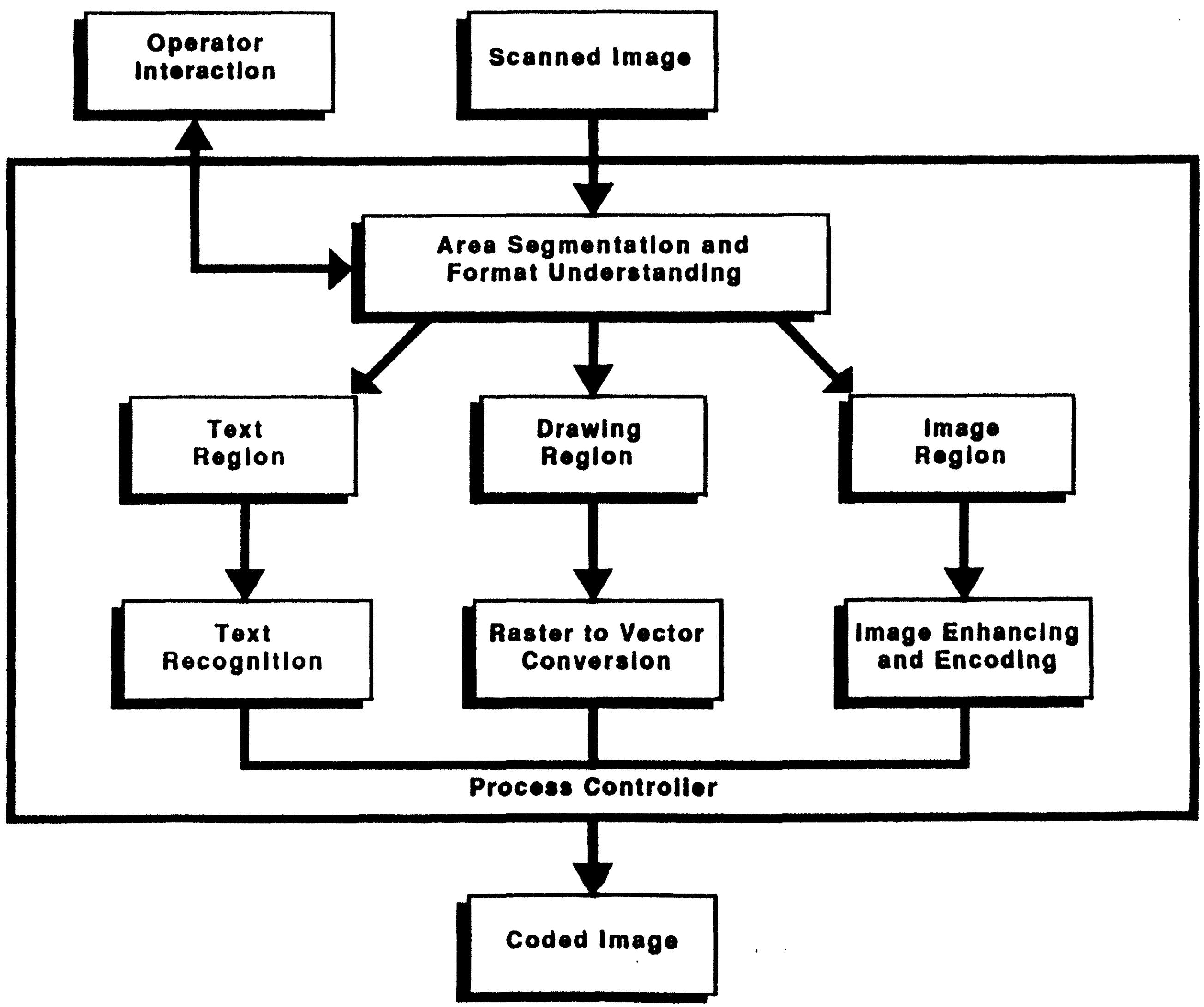
Conventional techniques for character recognition based solely on geometric or analytical properties are not sufficient to accurately process complex documents at high speed [22,25]. While the use of contextual analysis improves accuracy, it does not increase processing speed. To improve speed, it becomes necessary to simultaneously analyze the document from several points of view. For example, one unit can analyze the image, another can recognize characters, and a third can perform contextual analysis of the text.

The use of a single recognition technique is insufficient to solve all ambiguities encountered during the recognition process. As such, general algorithms must be complemented with customized rules for special cases, such as distinguishing letters that are easily confused (e.g., "o" and "O" and "6" and "b"), or recognizing characters printed in pieces (such as "i", "j", and ";"). Ideally, a single process control unit should collect all the pieces of information from the subsystems, order them, and then make appropriate final decisions [4]. Such a strategy (functionally depicted in Figure 4) allows the recognition of different parts to be done in parallel.

2.3 Postprocessing

After the text and graphics are recognized, they must be encoded for transmission to, and processing by, a word processor, a graphic editor or a desktop publishing program [5]. In the case of text, the image is most commonly converted into the ASCII format. In the case of graphics, the storage requirements for the document vary greatly, depending on the extent of compression. The storage capacity may pose a major constraint in some cases. For example, in systems which scan at 300 dots per inch (dpi), one single 8.5" x 11" page requires over one megabyte of

FIGURE 4: SCHEMATIC REPRESENTATION OF A SINGLE-PROCESSOR IMAGING SYSTEM



memory if stored in raw bitmap format. While the advent of optical disks with storage capacities in gigabytes tends to mitigate the storage problem to some extent, it is usually desirable to use more efficient storage strategies.

3. CLASSIFICATION OF SYSTEMS

For analytical purposes, the scanning industry can be divided into four segments: (a) Low-end scanners; (b) High-end scanners; (c) Integrated systems; and (d) Software. The characteristics of each segment are explained in the following paragraphs.

(a) Low-end scanners

These are small scale systems, usually desktop-based with relatively low sophistication in dealing with page scanning and image processing applications [1]. They are dependent on a host computer system for processing and storage (commonly an IBM personal computer or a Macintosh). The capabilities of these systems are enhanced by third party image and text processing software that is generally executed on the host computer.

Low-end scanners are either automatic feed or flatbed, and can handle between 20-100 documents per minute (or 20-50 pages per minute). The recognition technology tends to be simple - generally the matrix matching technique is used. Costs are low, ranging from $1,800 to $7,000. The software cost averages around $2,000 [3]. These scanners are geared towards low volumes (10 - 100 pages or documents at a time), and relatively low accuracy (1 - 10 letters per page misread or flagged for error on the average). The Scan-Optics Easy-Reader 1720 and the Datacopy 730 are examples of popular low-end scanners.

Most of these machines can scan a page or a portion of a page as an image or as characters of text, depending on the intent of the user. In some cases, the user can specify which zones within the input page contain text and which zones include graphic images. Some of the systems use separate phases for scanning text and images respectively.

Hand held scanners have entered as the lowest end products in the scanner market [3]. These usually offer limited page and text reading capabilities without requiring any additional software. They are either plugged directly into the personal computer like a keyboard or linked by a simple interface (such as a parallel card). They are transparent to the host computer, its operating system, and its application software. The volume handled is very low, because of the need to physically move the scanner with the hand. The scanning speed ranges from about 50 to 200 characters per second, and the number of fonts recognized is limited to 5 or 10. Their costs are quite low ($200 - $1000) [16]. The Caere Corp OCReader is an example of a widely used hand-held scanner.

(b) **High-end scanners**

These are large, sophisticated image/page/document processors with stand-alone or host based capabilities [1]. They offer mature hardware and software devoted to the tasks of text recognition and image scanning. These systems include sophisticated dedicated processors, editing and correction workstations, as well as customized recognition software that uses sophisticated techniques like feature extraction [30]. They support their own storage systems and can be readily integrated with other common systems.

High-end scanners feature multi-font and multi-size flexibility to handle multiple fonts and multiple sizes along with automatic feeding capability that can cope with high volumes. Large numbers of documents (1,000-10,000) can be accurately processed at high speeds (200 or more pages per minute or 750 or more documents per minute). Costs range from $15,000 upwards with the median falling between $25,000 - $30,000 [16]. Examples of high-end scanners include the Calera CDP Series and the Kurzweil family of scanners.

(c) Integrated Systems

These systems serve as automated image processing facilities for the office, and are designed to network with the existing computer and communications facilities. A typical system, controlled by a common operating system such as UNIX, consists of an enhanced office network that provides multi-media, multi-user, and multi-tasking access to its diverse resources [17].

An integrated image system offers various data entry facilities such as optical character recognition devices (both low and high end), image scanners, keyboard workstations and facsimile equipment. The system is usually managed by a dedicated processor (a mainframe or a minicomputer), although LAN-based systems are becoming increasingly popular. The processor is procured from a vendor, or the system is configured to operate within the existing computational facilities. Data storage is centralized and is accessible from all the workstations. Software is executed either on the processor or on a decentralized basis [17].

Typically, an integrated system will support a variety of software packages, including "off the shelf" word processors, database and statistics programs, and other professional software tools. Similarly, at the output level there is support for a variety of printers, monitors, and facsimile equipment.

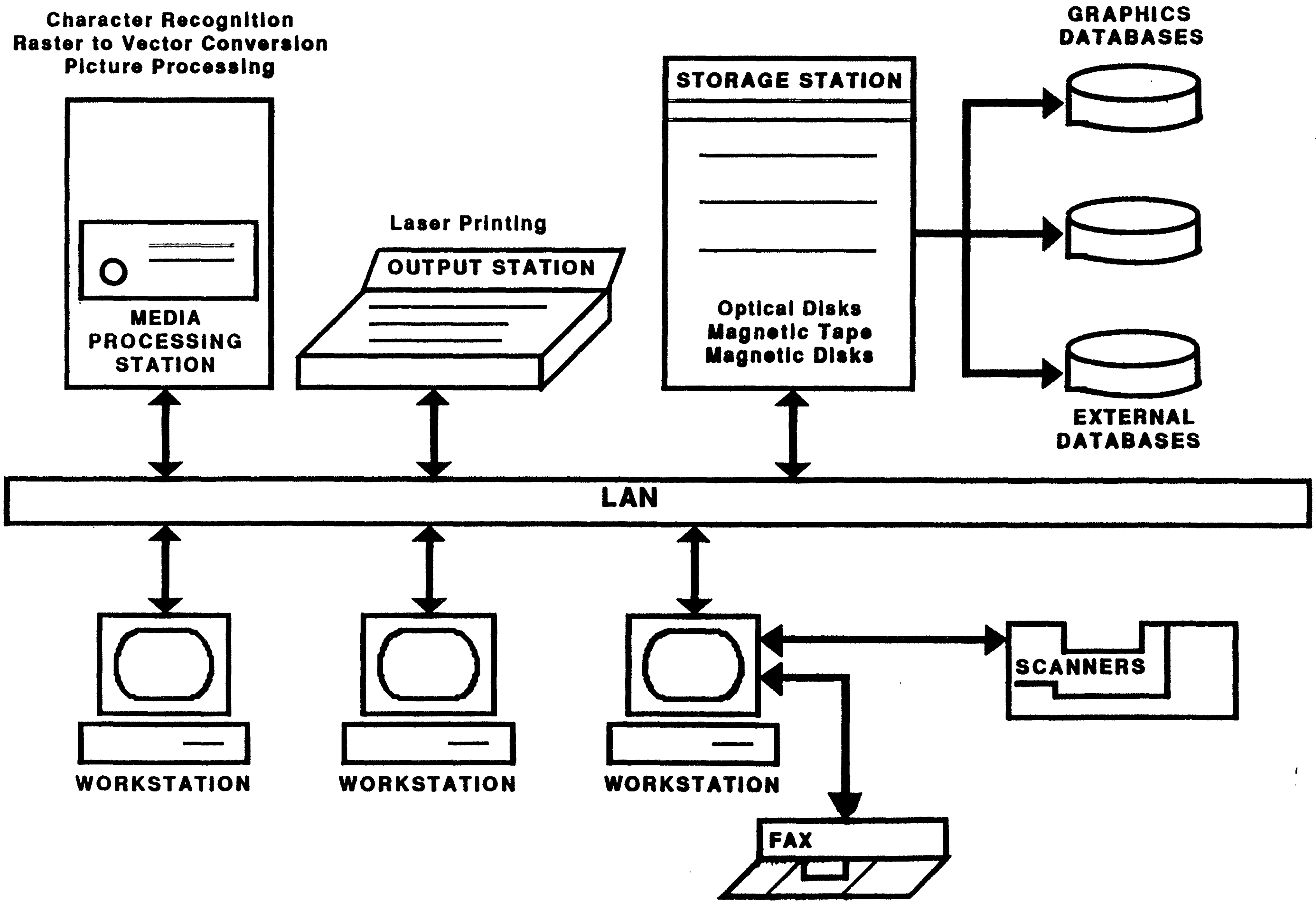
Several integrated imaging systems are available, including the Scan Optics ScanEdit 3200, the Wang Integrated Image Processing System, and the DRS 4000. IBM has announced its own system, the Image Plus, which is expected to be released in 1990. Figure 5 provides a schematic representation of a typical integrated imaging system. This paper does not offer a detailed treatment of integrated systems, as they represent an amalgamation of new imaging technology with conventional computing and communications power, rather than a new technological innovation.

(d) **Software**

Third party software tools offer powerful scanning and processing capabilities for both images and text. As mentioned earlier, they are substantially enhancing the capabilities of low-end scanners. In addition, software packages utilizing new approaches from areas such as database management, expert systems and neural networks are being developed for specialized applications [1]. Most packages are compatible with the popular desktop scanners as well as with Macintosh and IBM personal computers.

The orientation of this paper precludes a detailed analysis of software products. However, Section Seven provides a brief treatment of currently available products.

FIGURE 5: AN INTEGRATED IMAGING SYSTEM



4. FUNCTIONAL CLASSIFICATION OF MACHINES

Apart from classifying products by market segment, it is feasible to classify the image and character recognition industry by function. Thus, "reading machines" can be divided into four major categories: (a) Document Readers; (b) "Process Automation" Readers; (c) Page Readers; and (d) Image Readers. These groups are described in the following paragraphs.

(a) Document Readers

This category of readers was the first to be introduced. Developed during the sixties and seventies, these machines are oriented towards transaction processing applications such as billing and form processing [4,8,25]. In earlier machines, the source document was prepared in a rigid format using a stylized type font and the character set was limited to numeric characters plus a few special symbols and sometimes alphabetic letters. Contemporary machines have greater flexibility in font recognition and in coping with material of poor quality generated by high speed printers. The speed of processing is very high, typically between 400 and 4000 characters per second. An important feature is on-line correction with the help of bitmaps of unrecognized characters.

(b) **"Process Automation" Readers**

The main goal of these readers is to control a particular physical process [4,8,9,25], such as automatic sorting of postal letters. Since the objective is to direct each piece of mail into the appropriate sorting bin, whether or not the recognition process results in the correct decision for every single character is not critical in this case. Since a reader of this type is designed with specific applications in view, it is not explicitly considered in this paper. However, the class of integrated systems discussed in Section Three above is capable of performing "process automation" functions.

(c) **Page Readers**

These reading machines were originally developed to handle documents containing normal typewritten fonts and sizes [4,8,9,25]. Initially intended for use in the newspaper and publishing industries, these machines were designed with the capability to read all alphanumeric characters. Until the late seventies, devices of this type were quite restrictive in terms of their requirements for large margins, constrained spacing, and very high quality of printing. Further, they accepted only a small number of specially designed fonts such as OCR-A and OCR-B. The reading speed for a monofont page was several hundred characters per second and the price of the machine itself was around one hundred thousand dollars per unit. The above situation was significantly altered by the introduction of the Kurzweil Omnifont Reader in 1978. As its name implies, this machine was able to read virtually any font. Since then, several other machines belonging to this category have been introduced, and their capabilities have been substantially enhanced. For example, in many cases the user can specify the particular zones on the page to be read, and the format of each zone. On-line error correction is another common feature. In addition, prices have fallen steeply to under twenty-five thousand per system.

(d) **Image Readers**

These machines were originally designed to meet Computer Aided Design (CAD) needs, and subsequently came into more generalized use. They capture items such as drawings in the form of their mapped image and then reprocess that image to generate an equivalent vector graphics file.

The functional classification described above is fast becoming obsolete. As the matrix in Figure 6 indicates, one of the main trends in the scanning industry is the convergence of separate functions within a single system [1,3]. One machine can now perform functions that previously required multiple machines. An equally important

FIGURE 6: THE SCANNING INDUSTRY - A FUNCTIONAL VS. MARKET CLASSIFICATION MATRIX

MARKET →

FUNCTIONAL ↓

	LOW-END SCANNERS	HIGH-END SCANNERS	INTEGRATED SYSTEMS
DOCUMENT READER	X	X	X
PROCESS AUTOMATION READERS			X
PAGE READER	X	X	X
IMAGE READER		X	X

trend is the growth in the market for ancillary products such as third party scanning software, storage-end devices and data indexing and retrieval software [3,17,18]. The next section explores this trend.

5. ANCILLARY PRODUCTS

The rapid growth of the scanning market has been accompanied by a proliferation of ancillary products. These fall into two broad groups:

(a) Low-End Enhancement Products

This group consists of low-cost products which are designed to augment the capabilities of low-end scanning systems. They interface with the scanner to perform one or more functions of a high-end system. Third party scanner enhancement software and low-end interfacing devices fall into this group.

(b) Complementary Products

This group consists of products which complement and support the capabilities of scanning systems. Storage-end devices and software for specialized scanning applications are examples of complementary products.

The following paragraphs briefly describe the trends and products within each group.

5.1 Low-End Enhancement Products

These products emerged in response to the substantial costs of high-end systems. By performing one or more functions of a high-end system, they enable relatively basic systems to emulate high-end systems at lower costs. Scanner enhancement software and peripheral interfacing devices are typical examples of this phenomenon, and are discussed in the following paragraphs.

(a) **Third Party Scanner Enhancement Software**

These are page recognition packages like OmniPage, Advantex, OCR ReadRight and SPOT. They perform tasks such as creating scanning templates, specifying key fields, spelling and grammar checking, and separating columns, graphics and text. Ranging in price from $600 to $4000 [3], these relatively user friendly packages are compatible with most popular scanners and support the more widely used word processing, spreadsheet and image compression formats.

Although scanner enhancement programs enable users to perform functions beyond the capability of low-end systems, the claims of some vendors that these programs can replace high-end scanners are only partially true. In comparison to high-end systems, such programs are slower and less versatile in the number of fonts they can recognize. They are also less accurate, and are not very good at separating columns, graphics and text. However, they represent a serious alternative for many users who do not need or cannot afford a high-end system [16]. A good example is the desktop publishing industry, where user friendliness is a primary concern. High-end system vendors have had relatively little success in accessing this market, but third party packages used with low-end scanners have been extremely popular. According to International Data Corporation [18], overall sales for these products have grown by 215% in the period between 1986 and mid 1989.

(b) **Low-End Interfacing Devices**

This group of scanner ancillaries is the hardware equivalent of scanner enhancement software. These products are usually plug-in boards for personal computers, and they interface with the host system and its low-end scanner to perform high-end scanning functions. Like scanner enhancement programs, these devices enable users to recognize text and graphics and "read" them into word processing and other formats. Specification of key fields and template creation are also supported. Ranging in price from $1000 to $5000 [16], low-end interfacing devices suffer from the same drawbacks as their software counterparts, such as slower speeds and limited recognition of fonts.

5.2 Complementary Products

This group of products complement and support the overall capabilities of scanning systems. Storage-end hardware and data indexing and retrieval software are included in this group. In addition, there are a number of products which combine scanning technology with technology from other areas to perform specialized scanning applications. One example of this trend is the advent of neural network based programs for handwritten character recognition [19].

(a) **Storage-End Devices**

Scanning systems require much more storage space than is available with traditional magnetic storage media [28]. Using an optical disk system with several gigabytes of storage space is one way of meeting this need. A survey by Frost and Sullivan in 1989 [3] found that between 1986 and 1988, sales of optical disk systems grew by over 150%. These systems consist of one or more optical disk drives (usually Write Once Read Many or WORM drives) and sometimes a "jukebox" to perform disk storage, retrieval, loading and operating tasks [21]. However, these systems are very expensive and a typical system with two disk drives and a five platter jukebox costs between $27,000 and $33,000 [18]. In addition, one must pay for sophisticated software packages which store, index and retrieve data on CD-ROM's.

One alternative to optical disk technology is being marketed by Scan Optics, Inc. This is the Image EasyFile System, which uses 8mm cassette tapes (on up to 28 concurrent drives) as the storage medium. Each tape has a capacity of 2.3 gigabytes. As data on cassette tapes have to be accessed in a serial manner, this system is much slower than an optical disk system. However, at $ 17,000 for a 28 drive unit, Image EasyFile is considerably cheaper than an equivalent optical disk storage system. A cassette based system is a viable alternative to an optical disk system in situations where the access frequency per data point is low.

(b) **Software for Specialized Applications**

An interesting development in the scanning industry has been the joint application of scanning technology and technology from other fields to handle specialized scanning tasks. A good example is the development of neural network based scanning software for handwritten character recognition. Neural network computer models are based on models of the functioning of the human brain. They emulate thought processes such as learning, and can be taught to perform complex tasks [19]. The handwriting recognition application emerged as a result of the use of neural networks in the wider area of pattern recognition.

A number of packages are currently available, either off-the-shelf or as part of a customized service that configures handwriting recognition solutions to customer needs. NestorWriter of Nestor Inc., Providence, Rhode Island is an example of the former, and Inscript of Neurogen, Inc., Brookline, Massachusetts is an example of the latter. The main attraction of these packages is that they can be trained to recognize a much wider range of handwritten characters than is possible with conventional scanning software.

Most products in this area are designed for DOS, OS/2 and Macintosh environments. Data are entered through a scanner or a digitizing tablet, and recognition speeds range from 1 to 3 characters per second. As in the case of conventional scanning software, there is a tradeoff between the accuracy rate and the rate of rejection. The accuracy rate can be made arbitrarily high if tolerance for error is low, and vice versa.

The main problem with these packages is that it takes a very long time to train them. In addition, they scan at much slower speeds than conventional systems do. Until these problems are resolved, they are unlikely to be widely used. However, they have excellent long term potential, and should eventually come into widespread use by the scanning industry. Short term prospects are brightest in areas such as scanning numerical amounts on checks and signature verification.

The next section discusses broad trends in the optical image scanning and character recognition industry.

6. TRENDS AND PROJECTIONS

Rapid advances in hardware and software have created a major demand for scanning technology in diverse industries ranging from banking to defense and from engineering to medicine [17]. Instead of focusing on each application separately, this section focuses on broad trends. Figure 7 presents a schematic representation of the evolution of the capabilities of scanning technology.

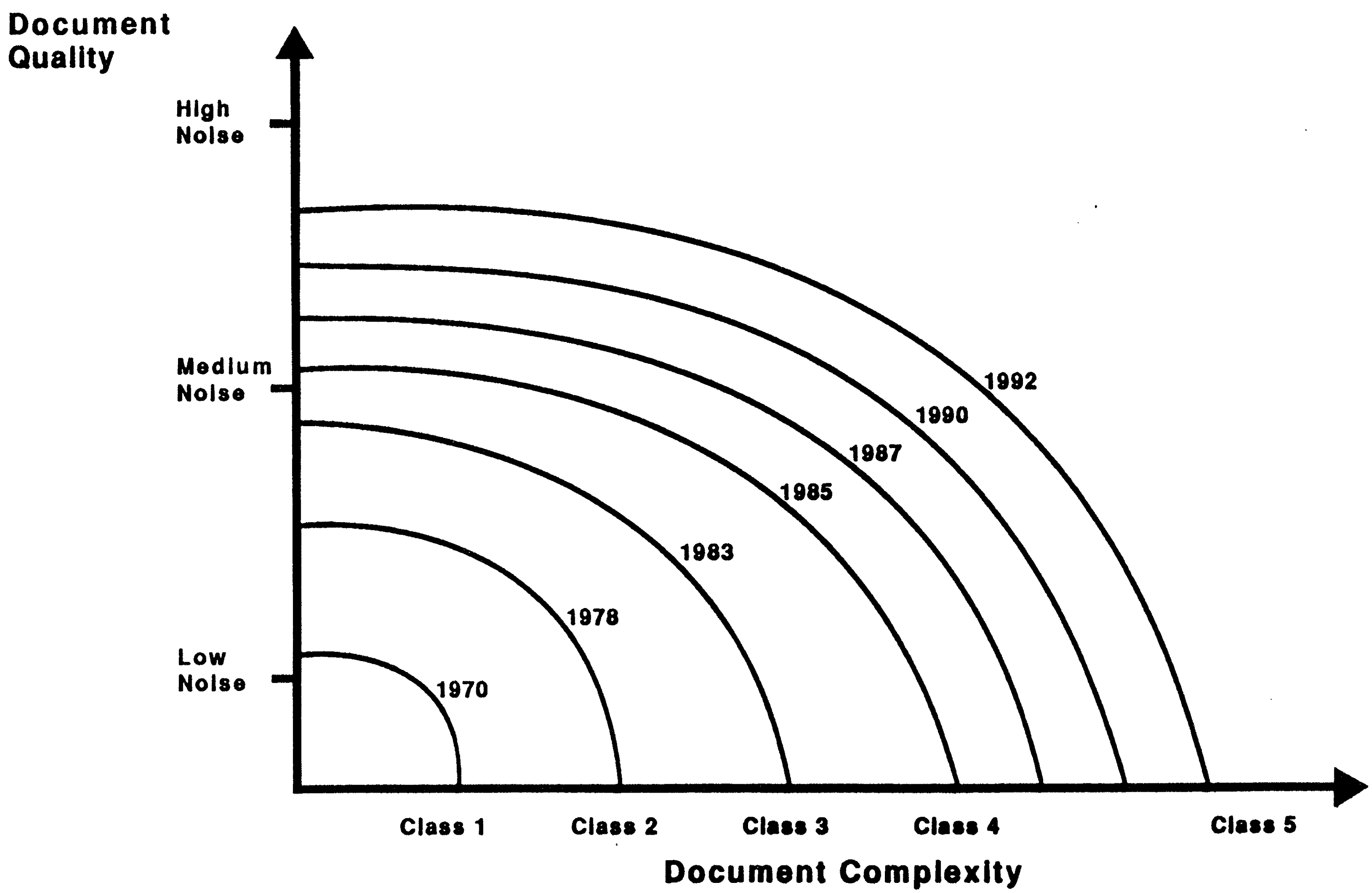
6.1 Falling Costs

One of the primary reasons for the growing demand for scanning technology is the fall in the cost of products. As technology improves, costs will continue to decline. Figure 8 depicts this trend of falling costs in terms of the system costs per font.

6.2 Coalescence of Text, Image and Graphics Processing

Graphical information can be stored either as simple bitmaps or as a set of standard geometric entities. The latter option allows for easy editing and storage in a library of symbols. Currently available scanners do not offer these capabilities. However, developments are taking place in the area of Computer Aided Design (CAD) for converting raster images of line drawings into vector graphics files, which can be easily modified with graphic

FIGURE 7: EVOLUTION OF SCANNING CAPABILITIES



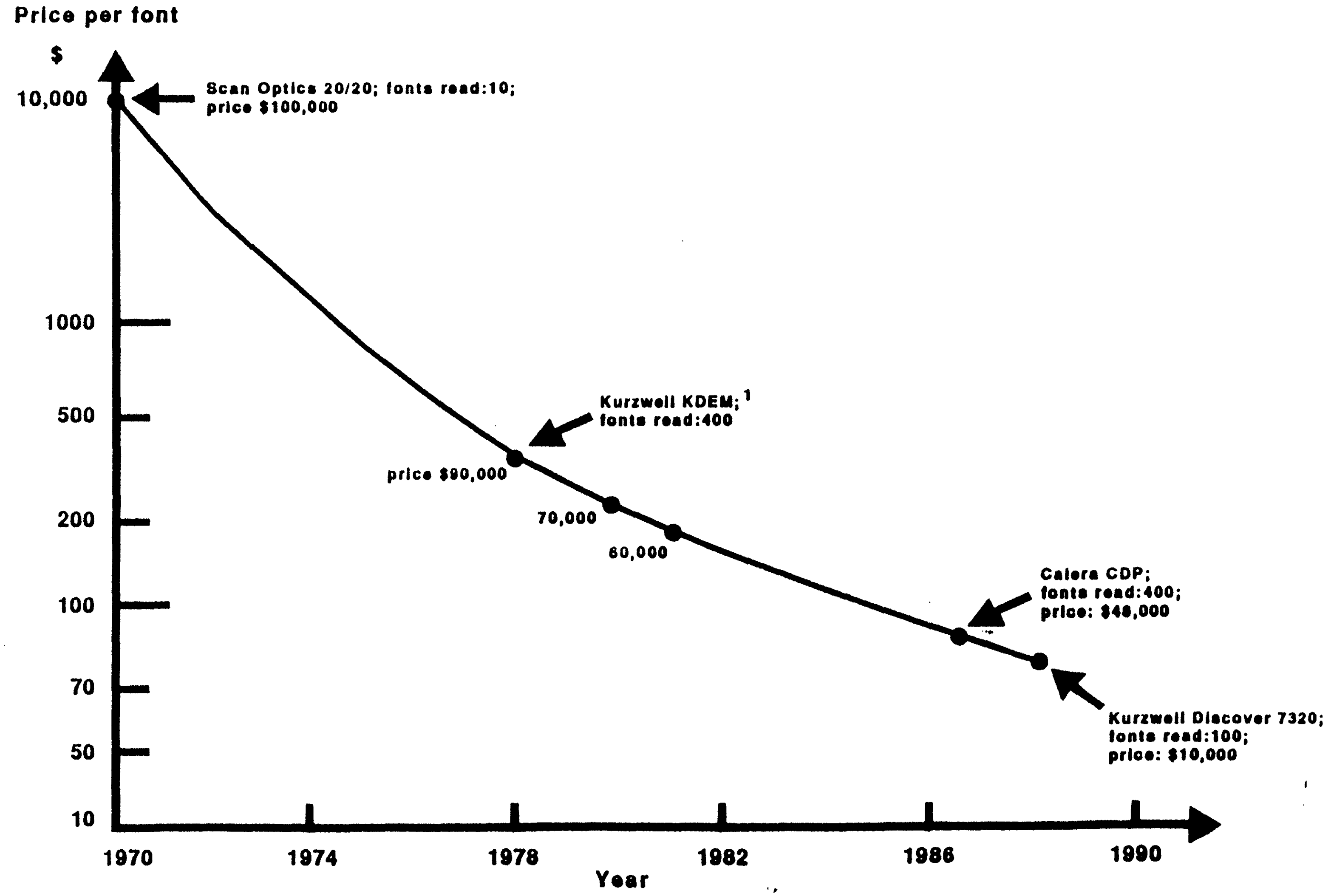
DOCUMENT COMPLEXITY

Class 1: Single column, monospaced, single pitch
Class 2: Single column, multifont, mixed spacing
Class 3: Single column, some images, any formatted text
Class 4: Multicolumn document, tables
Class 5: Mulitcolumn, mixed text and images

DOCUMENT QUALITY

Low Noise: Original typewritten or typeset document, clearly separated characters, no skewing
Medium Noise: Easily readable photocopy or original laser print, characters not touching
High Noise: Broken and touching characters, fading ink, skewed text

FIGURE 8: DECLINING COSTS OF READING MACHINES PER FONT RECOGNIZED (1970-89)



[1] Kurzweil KDEM price decreased over time

NOTE: This figure is not drawn to scale

line editors. Raster-to-vector conversion systems are now available from several companies including Houston Instruments and Autodesk Inc. However, at present only approximate solutions are supported. For example, all curves are decomposed into line segments. It is expected that ideas from the area of raster to vector conversion will be combined with scanning technologies to enable text, images, and line graphics to be edited and processed through a composite package.

6.3 Integration and Networking

One of the most significant trends in the information processing sector as a whole is the increasing integration of disparate data processing equipment into a single composite system [17]. This trend is reflected in the scanning industry by the emergence of the integrated imaging system (described in Section Three).

The trend towards integration can be observed in other areas of the scanning industry. For example, optical character recognition capabilities are being increasingly integrated into facsimile equipment [29]. At the same time, facsimile capabilities are also becoming available on document readers such as the Datacopy 730.

6.4 Impact of Artificial Intelligence Techniques

Advances in Artificial Intelligence techniques in areas related to character recognition and image analysis will lead to faster, more accurate and versatile reading machines [23]. Semantic analysis and natural language parsing aids for contextual analysis will improve the process of identifying letters and words, reducing the error rates, the reject rates and the need for operator intervention [24]. Eventually, reading machines will be able to handle virtually all printed material [18,27].

6.5 Use of Special Purpose Hardware

A highly accurate omnifont reader requires sophisticated algorithms for vectorization and classification, separation of merged characters, contextual analysis, and syntactical analysis. These algorithms can benefit from microprogramming and special purpose hardware which is optimized for specific applications. Sophisticated readers such as the Calera CDP 9000 and Kurzweil 5100 use special purpose hardware.

6.6 Limiting Factors

Although it appears that ongoing research will enhance the capabilities of scanning systems, several factors continue to impede a wider acceptance of the technology. These are as follows:

(a) The accuracy provided by current systems is still inadequate for several applications. The editing of errors and the presence of unrecognized characters continue to be major bottlenecks. The time taken by the operator to overcome these bottlenecks limits the throughput of the overall system.

(b) Wide acceptance of the new technology will occur only after it becomes possible to automatically handle complex documents containing both text and graphics [13,17,20]. The need to hire and train expert operators continues to inhibit the widespread adoption of scanning systems.

(c) Broken strings and touching characters in a document still constitute a major hurdle for even the most sophisticated machines. Scanning systems remain highly sensitive to variables such as paper quality and thickness, clarity and font. Major developments in syntactical and semantic analysis are needed before reading machines realize their full potential.

The major trends and projections are summarized in Table 1.

TABLE 1: SCANNERS - TRENDS AND PROJECTIONS

	1975	1982	1989	PROJECTED (within 1-3 years)	PROJECTED (within 8 years)
PROCESSING CAPABILITY	Text (characters)	Text plus images (ASCII text and bit maps)	Text plus images (ASCII text and bit maps)	Text and images; Separate abilities to deal with vector graphics	Text and images plus vector graphics (vector graphics files)
FONTS	Typewritten text only	Most printed as well as typewritten fonts	Virtually all printed material	All printed material	All printed fonts including some handwritten text
IMAGE ANALYSIS	Matrix matching for character recognition of selected fonts	Feature analysis of letters within the text, and matrix matching techniques for identifying characters	Context analysis spelling checking (lexicons), feature extraction for identifying characters	Natural language syntax analysis based techniques for text recognition; aids for distinguishing text and graphics	Semantic analysis, natural language parsing aids for context analysis and correct interpretation of letters and automatic identification of text and image areas
EDITING	Simple editors for correcting correcting errors within data scanned	Text editors and word processors for editing pages	Separate editing facilities for text (text) editors/word processors) and images (through graphic editors)	Limited degree of integrated editing	Integrated editor for text, images, and vector graphics with interface for other packages like DBMSs
UTILIZATION AREAS	Document processing with no editing requirements	Documents with text processed through text editors	Pages scanned and processed through word processors, and graphic editors	Documents with text and graphics processed	Customization of printed page processing software
ERROR RATES	2-3 percent for a good quality document	2-3 percent for a typewritten page	1 percent for a typewritten page	0.5 percent for a typewritten page; 1 percent for a typeset page	0 percent for a typewritten page; 0.1-0.4 percent for a typeset page

7. A TAXONOMY FOR PERFORMANCE EVALUATION

The conventional criteria for evaluation of a reader are speed and accuracy. However, by expressing the speed simply in characters per second and the accuracy in a single figure, one is overlooking the fact that the performance of a scanning system is heavily dependent on the characteristics of the input. For example, it takes longer to process a document with multiple fonts and multiple character sizes than it does to process a monofont, monosize document [9]. The presence of graphics and the formatting of the document also affects the reading speed. Further, the quality of the print is a major factor affecting the accuracy of the system; broken and touching characters, low contrast, and skewed text result in high error rates and reject rates as well as in a significant reduction in the speed of reading. The speed and the error rate presented in the technical documentation supplied by the vendors consider the characteristics in only one case - usually the perfect one.

In order to assess the capabilities of a system, one must consider not only the scanning speed but also factors such as time spent in editing and correcting, time spent in training the operators and time spent in training the system itself where applicable [13]. In the case of a document containing several graphs and multiple columns, the time spent in editing the document can be significantly greater than the scanning time [12,13]. Consequently, it is difficult to obtain an accurate estimate of the overall speed by simply observing the elapsed time for the scan operation.

With rapid evolution in the technologies used in the new generation of products and their broad functionality, it becomes necessary to use a larger repertoire of evaluation criteria that gives appropriate weightage to the major factors that determine the efficiency of the scanning process, such as: complexity of document; quality of document; recognition technology used; and man - machine interface. Unfortunately, there is no framework that currently meets this need. In order to fill this void, a new taxonomy for the evaluation of scanning systems is developed in this section.

7.1 Document Complexity

The above facts highlight the need for a measure of document complexity that takes into account the diversity of fonts utilized, the size of characters, and the proportion of images in the document. Since no measure exists, a five-part document complexity classification scheme is proposed to facilitate systematic analysis. This classification groups documents into the following classes (in increasing order of complexity):

(a) Class 1: Basic Text - Only Documents: All material is in a single font, with a single pitch and uniform spacing. An example of this class is a typewritten document, as shown in Figure 9 (a).

(b) Class 2: Single Column Documents with Multiple Fonts and Mixed Spacing: This covers text-only documents with proportional spacing, as well as typeset and laser printed documents with multiple formats (such as bold or hyphenated). A sample Class 2 document is shown in Figure 9 (b).

(c) Class 3: Single Column Documents with Segregated Text and Images: Such documents contain all material in a single column format. The text is justified or hyphenated and there are some images. These images can be easily separated from the text (separate zones for text and images), as in the case of the example in Figure 9 (c).

(d) Class 4: Multicolumn Documents: Such documents contain two or more columns on a page. Although they have mostly text, there are some images and tabular material. A printed page from a newspaper will fall under this category. A

- Figure references are not consistent (some with, some without parentheses).
- The references quoted on page 6 are nowhere indicated in the text, except for the * on page 5, line 14.
- The figures are of poor quality. Figs. 2 and 3 are unnecessarily complicated, Fig. 4 lacks one legend in the ordinate.

All the above could be corrected by a complete rewrite. The idea is interesting, having a minimum configuration microprocessor display current, power and maximum demand by a TLU method on current values obtained safely by a non-invasive current transformer. I would still question the following:

- Only current is measured and the data that are displayed are based upon, previous, actual current and power measurements, stored in the table. As power is still a function of current <u>and</u> voltage (and phase angle if the load is not purely resistive), the results will be in error if the line voltage changes. In addition, unless it is the intention to use this only on one particular load, the type of load will influence the results, making this system not quite general purpose.
- No mention is made of the actual hardware used. From the references I presume it is an Intel machine. The type of ADC (speed, number of bits) and the method of display (printer, 7-segment displays) are also not mentioned.
- Page 3: "for every possible digital value, the current and power are determined and stored ...". In what format? The description on page 4 and the flowcharts of Figs. 2 and 3, leave much to the imagination, unless the reader knows exactly the configuration of their microcomputer. The references to the "status", "addresses" and "data" field suggest to me that some type of single board computer, that has these fields as displays, was used. Can it be assumed that all readers are familiar with this?
- If I follow the flowchart in Fig. 3 (maximum demand display) correctly, when a subsequent current value is determined to be less than the previous one, a search through the complete table is made, using the same, old, input data before displaying the, same, results. I don't know how much memory is included in their system, but surely the previous value could be stored somewhere to avoid this unneeded search.
- Similarly, if the next value is determined to be higher than the previous one, again a complete table search is made instead of starting at the address of the last (lower) value.
- In the three phase measurement system, a simple two comparison schedule should be added to indicate the highest load, rather than leaving this up to the human operator.

...3

<u>FIGURE 9 (a)</u>: SAMPLE DOCUMENT – CLASS 1

HIGHLIGHTS OF
KNOWLEDGE-BASED INTEGRATED INFORMATION SYSTEMS ENGINEERING

AMAR GUPTA AND STUART E. MADNICK

Large organizations must necessarily rely on multiple computer systems for a number of reasons, such as the increasing size of the organization and the growing reliance on computerized data. In virtually all cases, dissimilar and incompatible hardware and software systems are operating on a concurrent basis. While these systems may meet the objectives for which each was designed, their heterogeneity presents a major obstacle to ready access and assimilation of the information they contain.

The objective of Integrated Information Systems, or Composite Information Systems (CIS), is to mitigate the problem described above. Such systems must be geared to span applications, functional areas, organizational boundaries, and geographic separations in order to present a unified picture to the user. While designing such systems, it is necessary to look at a number of inter-related strategic, technical, and organizational issues. The goal of the Knowledge-Based Integrated Information Systems Engineering (KBIISE) effort, highlighted in this report, is to survey the state-of-the-art of methodologies for addressing these needs and identify areas needing increased research focus.

Strategic issues include motivating cooperation between multiple organizations, each with its own goals, priorities, and security needs. One critical success factor for such cooperation is participant consensus on the issue of access to each others' technical and non-technical information. There is an urgent need to clearly define the domains of shared information, the potential benefit to each group that participates, and the role and the responsibility of each constituent.

Under technical issues, the evolution of distributed heterogeneous information systems is studied. Being inherently more complex than conventional databases, such systems require powerful semantics and update capabilities, as well as sophisticated concurrency control and recovery mechanisms. Both the semantics and the syntax of individual queries and updates must be mapped across systems. In addition to physical connectivity issues, it becomes essential to develop new techniques for incorporating logical connectivity across systems. Such techniques combine ideas from the fields of database technology, communication technology and expert systems technology.

Organizational issues cover the process of making controlled changes in complex organizational environments. The prevailing theory of inter-organizational networks explains the multiple forces that modulate behavior of individuals, groups, and organizations. There is a need to develop focused standards to serve as the foundations for neutral representation as well as for the development of more elaborate standards.

The problem of distributed heterogeneous information systems occurs in many disciplines ranging from manufacturing to banking and from maintenance to logistics. A number of government organizations including NASA, NBS, and different agencies of the Department of Defense are faced with this problem. Within the U.S. Air Force, several efforts are directed at finding solutions in this area. In order to attain success, it is necessary that government, industry, and academia work together to develop common standards and to fill major voids that exist. A plan for concerted action is developed in this report.

FIGURE 9 (b): SAMPLE DOCUMENT – CLASS 2

ATTACHMENT C

Materials for IES AdCom Meeting (May24,1985,Toronto) Material—A

Suggestion For IE Society Enhancement

Objective

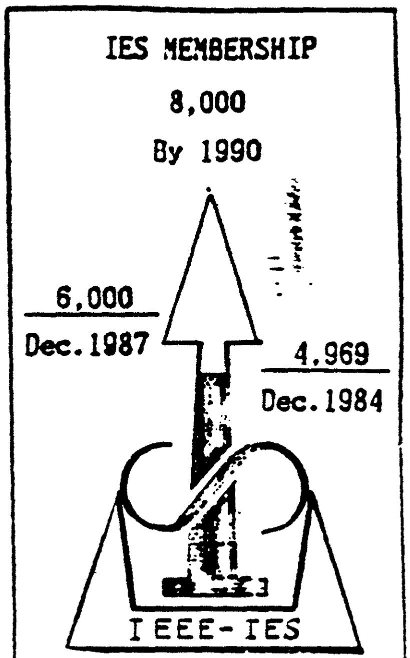
To promote enhancement activities for the purposes of increasing IES membership and improving and strengthening IES functions, image, and effectiveness.

Overall Plan

The overall plan to enhance our society involves addressing four areas as follows;

1. Method of Promotion and Scope
2. Goal
3. Enhancement Activities
4. Assessment of IES (status of IES)

Each of the above areas will be addressed in the following sections.



The IES's GOAL is a membership of 8,000 by 1990. This barometer will track progress toward that goal.

1.Method of Promotion and Scope

The enhancement activities are aimed at invigorating our society. These activities encompass all aspects which are likely to improve the management efficiency of the IES AdCom and they extend over a wide range.

To promote enhancement activities, there are 3 promotional methods to be used singly or in combination, depending on the nature of what we are doing;

(1) Top—down managerial activities from IES AdCom

(2) Technical Committee activities which extend horizontally

(3) Section/Chapter activities which have a bottom—up effect

§ Society activities with team work and cooperation; see Fig.1—1

FIGURE 9 (c): SAMPLE DOCUMENT - CLASS 3

TABLE VI
MICROPROCESSOR OPERATING SYSTEM SUPPORT BENCHMARKS—STACK EXERCISER [8]

			8086		Z 8000		68000	
S. No.	Table size	No. of tables	Path length	Time (clocks)	Path length	Time (clocks)	Path length	Time (clocks)
1	100	2	1838	18399	44	2220	436	5326
2	100	3	2751	27621	60	3332	648	7938
3	200	2	3638	36407	44	4236	836	10328
	Total (1)+(2)+(3)			82427		9788		23592
	Performance relative to 8086			1.00		8.42		3.49

TABLE VII
GENERAL CHARACTERISTICS OF 32-BIT MICROPROCESSORS [8]

	BELLMAC-32A	HP 32-BIT CPU	INTEL iAPX 432
YEAR OF COMMERCIAL INTRODUCTION	1982*	1982*	1981
TECHNOLOGY	2.5-μm DOMINO CMOS	1.5/1.0-μm NMOS	NMOS
NO. OF TRANSISTORS	146,000	450,000	219,000 ON 3 CHIPS
SIZE OF CHIP	160,000 MIL²	48,400 MIL²	100,000 MIL² EACH
POWER DISSIPATION	0.7 WATT AT 8 MHz	4 WATTS	2.5 WATTS/CHIP
PIN COUNT	63 ACTIVE 64 TOTAL	83	64 PER CHIP
BASIC CLOCK FREQUENCY	10 MHz	18 MHz	8 MHz
DIRECT ADDRESS RANGE (BYTES)	2^{32}	2^{29} REAL; 2^{41} VIRTUAL	2^{24} REAL; 2^{40} VIRTUAL
NO. OF GENERAL-PURPOSE REGISTERS	16 USER-VISIBLE	28 (NOT ALL GENERAL-PURPOSE)	NO REGISTERS VISIBLE TO USER
NO. OF BASIC INSTRUCTIONS	169	230	221
NO. OF ADDRESSING MODES	18	10	5

*Currently for internal use only.

TABLE VIII
EXECUTION TIMES [2], [16], [17]

MACHINE	LANGUAGE	WORD SIZE	TIME (MILLISECONDS)			
			SEARCH	SIEVE	PUZZLE	ACKER
VAX-11/780	C	32	1.4	250	9400	4600
	PASCAL (UNIX)	32	1.6	220	11 900	7800
	PASCAL (VMS)	32	1.4	259	11.530	9850
68000 (8 MHz)	C	32	4.7	740	37.100	7800
	PASCAL	16	5.3	810	32.470	11 400
	PASCAL	32	5.8	960	32.520	12.320
68000 (16 MHz)	PASCAL	16	1.3	196	9180	2750
	PASCAL	32	1.5	246	9200	3080
8086 (5 MHz)	PASCAL	16	7.3	764	44 000	11,100
432/REL 3 (8 MHz)	ADA	16	4.4	978	45.700	47 800
80286 (8 MHz)	PASCAL	16	1.4	168	9138	2218
80286 (10 MHz)	PASCAL	16	1.1	135	7311	1774
HP 32-BIT CPU* (18 MHz)	PASCAL	32	NA	NA	7450	2590
NS 16032* (7 MHz)	PASCAL	32	NA	NA	24 000	9900

*Indicates experimental prototype.
**Indicates vendor-provided information.

be monitored. Consider, for example, the machine tool application shown in Fig. 1. There are several independent variables to be controlled and monitored as follows:

a) rotation speed of parts,
b) lateral movement of parts,
c) lateral position and angle of each of the cutting tools.

To increase productivity, it is essential that multiple tools must operate concurrently; also, at no instant should any tool interfere with another tool. Each of the independent variables can be monitored by a separate microprocessor, but the overall concurrency of tool operation will be determined by the ability to maintain noninterfering profiles. The latter aspect is dependent on the communication bandwidth between the computing elements. The overall performance of such multiprocessor configurations is affected by the number of processors, the communication mechanism among the computing resources, the characteristics of the computational workload, and the control program. Whereas the major constraint in single processor systems is the speed of the processor itself, the major constraint in multiprocessor systems is usually the speed of the interconnection mechanism used for communicating between the computing elements.

Fig. 2 shows an example of a multiple microprocessor-

FIGURE 9 (d): SAMPLE DOCUMENT – CLASS 4

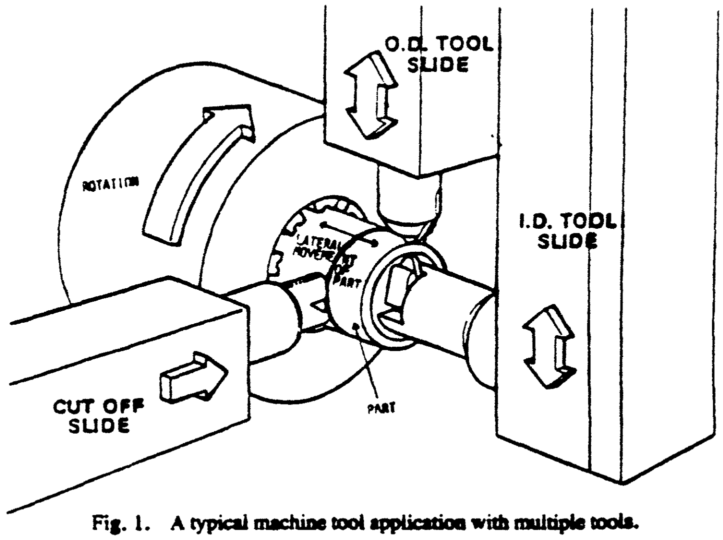


Fig. 1. A typical machine tool application with multiple tools.

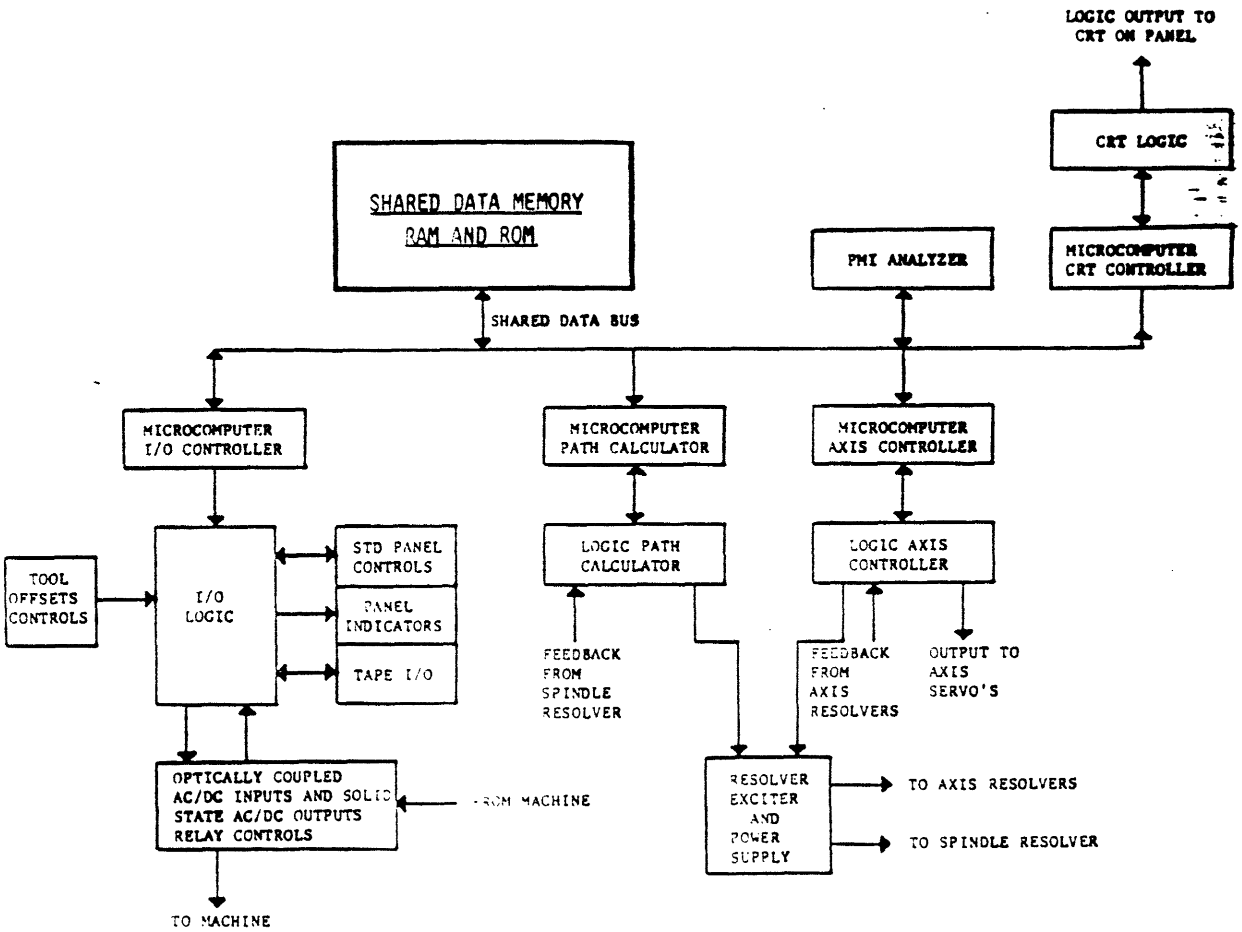


Fig. 2. Computer system configuration for machine tool application.

based multi-axis machine tool control that uses a shared common data bus. Machine tool throughput is limited not so much by the number of microprocessors attachable to the shared bus, but more by the utilization under load of the shared data bus and data memory. The memory utilization for two typical functions executed under the multi-axis machine tool control environment is summarized in Table IX. The total utilization figures have been broken up by principal microprocessor tasks. In each case, the program code resident in each microprocessor subsystem has been optimized to make minimum use of shared memory. For nonoptimized programs in which flags and status registers are maintained in shared memory, frequent accesses to shared memory are essential. This often has the undesirable result of driving up memory and bus utilizations to 70-90 percent. At this level, queuing contention will have a severe adverse impact upon system sample Class 4 document is shown in Figure 9 (d).

(e) Class 5: Integrated Documents: Such documents contain both text and images. A typical document of this class contains multiple columns, with several charts or illustrations within each column, as shown in the example in Figure 9 (e).

The above five - tier complexity classification scheme is utilized to evaluate a broad range of products in Section Six. The five sample documents shown in Figure 9 were selected after extensive study; they are used for the purposes of evaluating products and validating the taxonomy.

7.2 Document Quality

Since the performance of a scanner is highly dependent on the quality of the input documents, it became necessary to carefully control the variations in the quality of the documents used in the benchmark tests. Although it is difficult to establish rigorous measures, documents can be broadly grouped into three classes, based on their quality:

(a) Low noise documents: This category comprises original typewritten and typeset documents, with normal leading and clearly separated characters. Skewing is absent or negligible. In addition, these documents have no hyphenation or kerning.

(b) Medium noise documents: This category comprises original laser printed documents or high quality dot matrix printed documents, as well as good photocopies of such documents. The contrast is good and skewing is low (under 2%). Further, the characters do not touch each other. There may, however, be some instances of kerning, hyphenation, and uneven leading.

(c) High noise documents: This category comprises second or later generation photocopies with broken segments of text and characters touching each other. Usually, there is low contrast and skewed text.

The above three - tier scale represents one measure for defining quality. Strictly speaking, quality is a continuous variable with multiple dimensions (e.g., quality of characters, quality of background, contrast ratio and amount of skewing). Consequently, this scale represents a first level quantization of this variable.

7.3 Recognition Techniques

Recognition technique is a qualitative variable that tries to capture the sophistication of the technique used in the recognition process. Various recognition techniques such as matrix-matching and feature extraction offer different capabilities for reading. The implications of using different recognition techniques were examined in Section Two.

7.4 Man - Machine Interface

In order to minimize the total cost of scanning and editing documents, one important factor to consider is the interface with the reading machine. A higher-speed system that requires special skills and training of dedicated operators may, at times, be less desirable than a lower-speed system with a very user friendly interface. Two of the evaluation criteria that represent this variable are trainability and document handling. In addition, it is also important to examine the interface between the reader and other computational equipment.

The concepts outlined in subsections 7.1 and 7.2 above are independent of the equipment utilized. Also, they do not require an understanding of technical terms. Document complexity and document quality are therefore good variables for an evaluation exercise. Such an evaluation is documented in the next section.

8. EVALUATION OF PRODUCTS

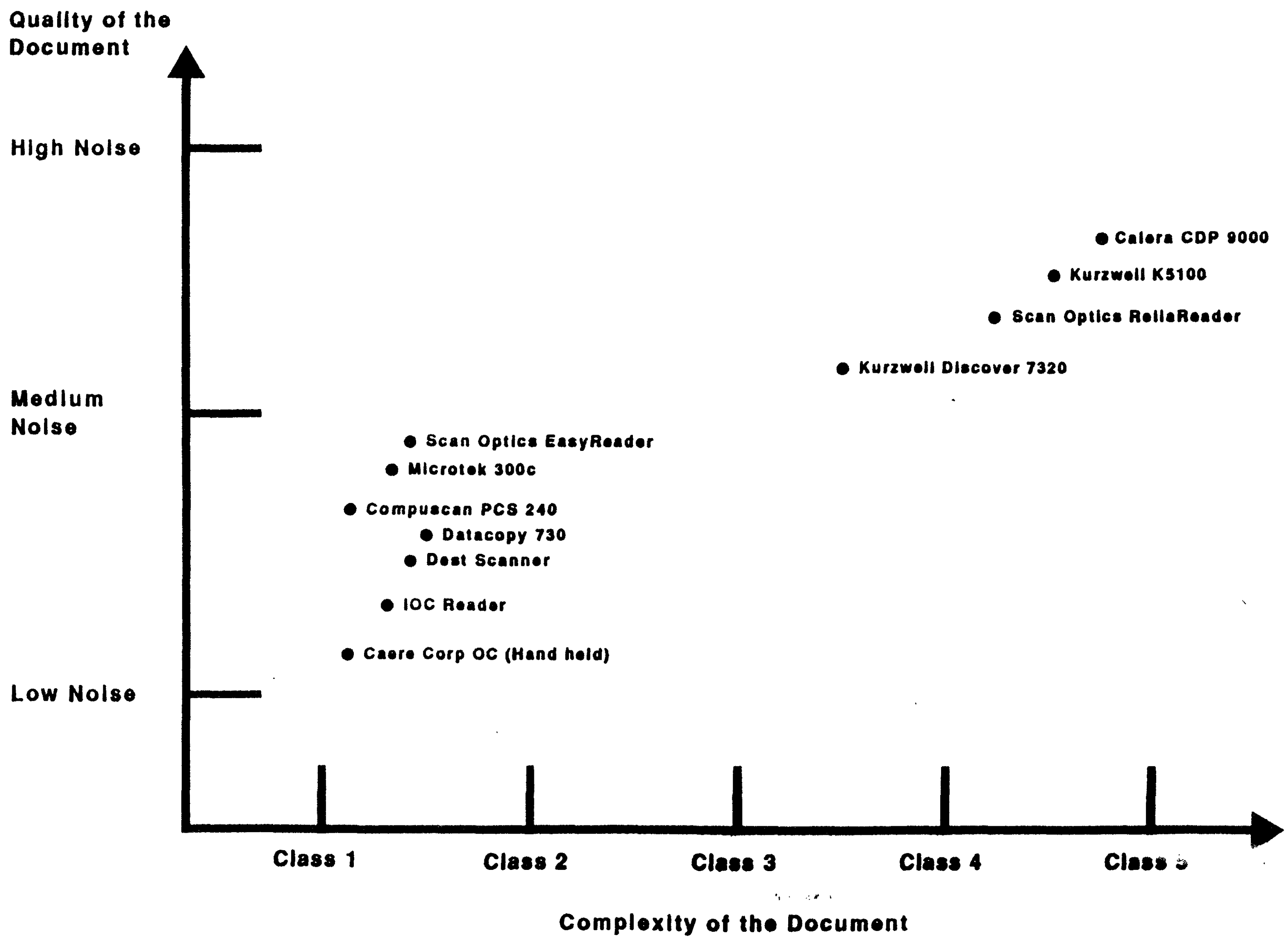
The number of scanners available today runs into the hundreds, and it is very difficult to generate an accurate and exhaustive comparison; even if it could be generated, it would soon become obsolete. Consequently, in this paper only a few representative products have been analyzed. The list is as follows:

Product	Family
(a) Datacopy 730	Low-end scanner
(b) Caere Corp OCReader 500 Series	Hand-held scanner
(c) Scan Optics EasyReader 1720	Low-end scanner
(d) DestScanners PCScan 1000/2000	Low-end scanner
(e) Sharp JX-450 Color Scanner	Low-end scanner
(f) Truvel Truscan TZ-3BWC	Low-end scanner
(g) Kurzweil Discover Model 30	Low-end scanner
(h) Calera CDP 9000	High-end scanner
(i) Scan Optics ReliaReader	High-end scanner
(j) Kurzweil 5100	High-end scanner

8.1 Results of the Evaluation

The results of the tests on these products are presented in Figures 10-12. Figure 10 depicts the relative performance of the systems in terms of a document complexity

FIGURE 10: CAPABILITIES OF CURRENT PRODUCTS ON DOCUMENT COMPLEXITY/QUALITY MATRIX



versus document quality graph. In Figure 11, relative performance is measured by graphing document complexity against scanning speed. Finally, Figure 12 maps the performance of the systems on a column and graphics - handling matrix. A more comprehensive comparison appears in the Appendix.

At the low end, all the scanners were able to read Class 1 and Class 2 documents without much difficulty. Class 3 documents were read most successfully by the Scan Optics EasyReader and the Kurzweil Discover Model 30. None of the low-end scanners was very successful with Class 4 and Class 5 documents.

All the high-end scanners were able to handle Class 1, 2 and 3 documents. The Scan Optics ReliaReader was not as accurate as the Calera CDP 9000 and the Kurzweil 5100 in the case of the Class 4 document. The Class 5 document, with no clear separation between text and graphics, was handled most effectively by the Calera CDP 9000 system. In the case of such complex documents, the process of editing and reconstituting the original format is time consuming, exceeding ten minutes for the sample used in the benchmark study.

Characters that touch each other and those with broken

FIGURE 11: CAPABILITIES OF CURRENT PRODUCTS ON DOCUMENT COMPLEXITY/SPEED MATRIX

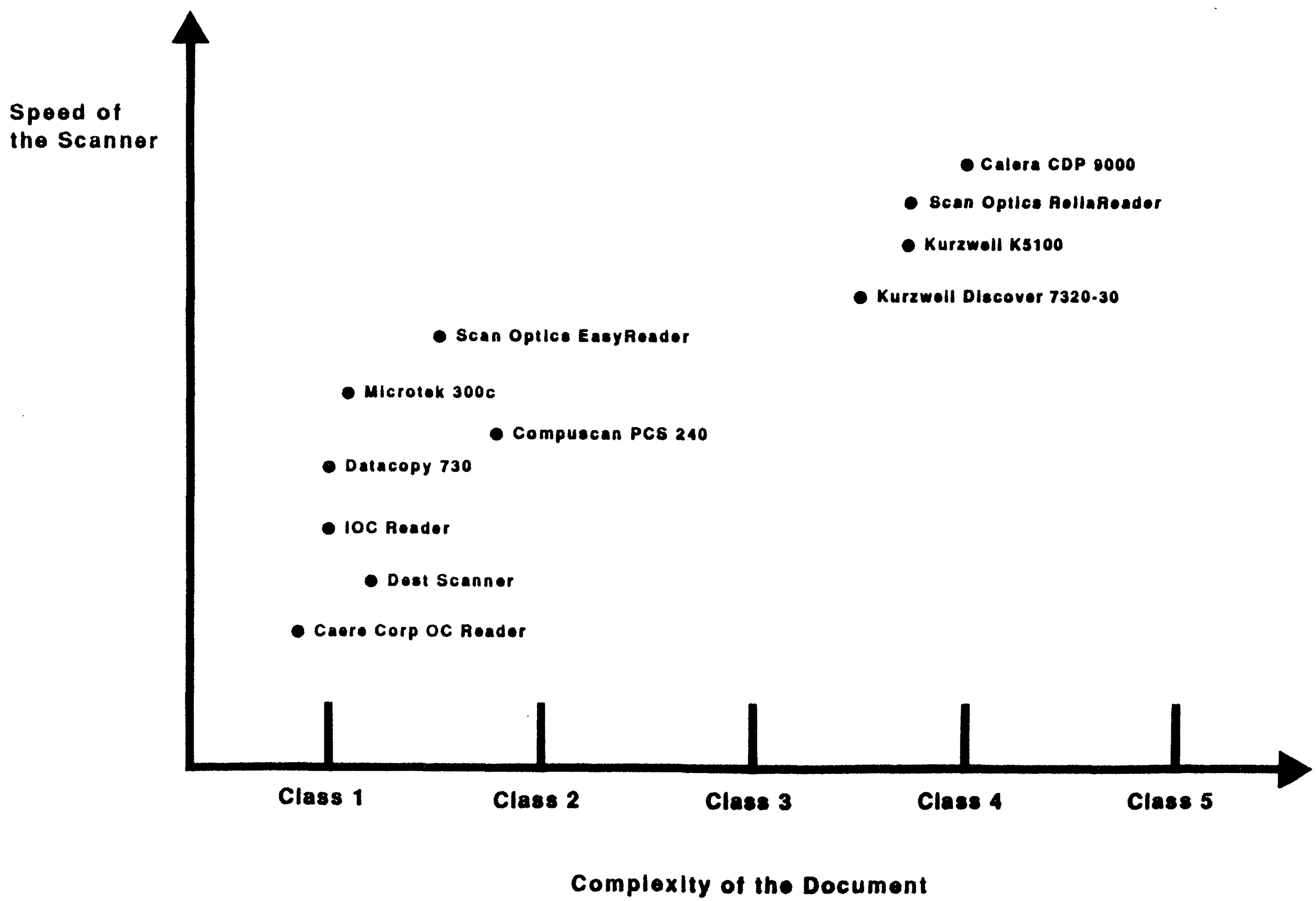
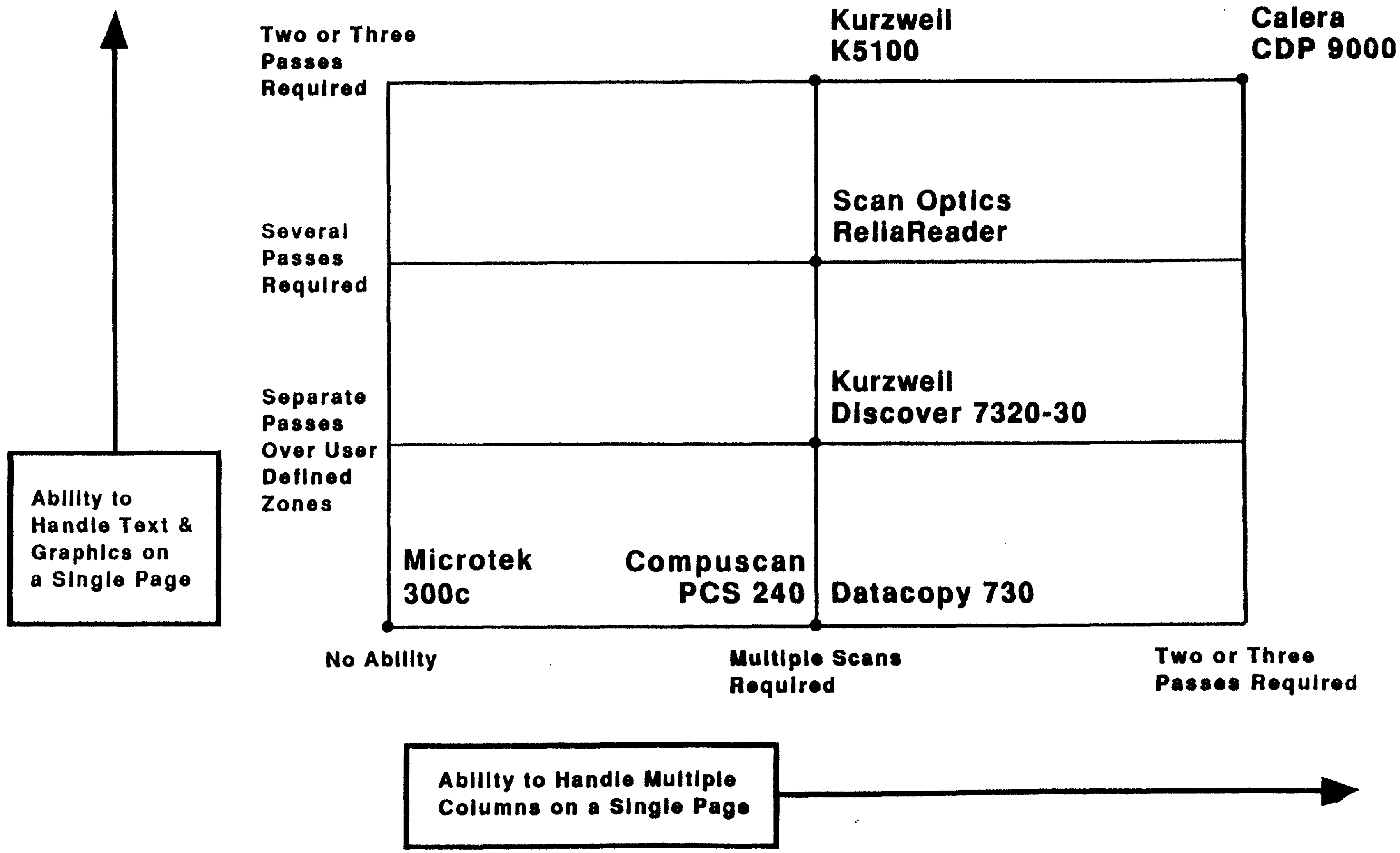


FIGURE 12: CAPABILITIES OF CURRENT PRODUCTS ON COLUMN & GRAPHICS HANDLING MATRIX



strokes were the major sources of error for all the systems tested. The complexity of the document being scanned severely impacted the scanning speed and the accuracy. The accuracy of all the scanners was found to be highly sensitive to the quality of the document. Even a typewritten document caused a significant number of errors in cases where the quality of the documents was low. A single handwritten mark or even a speck of dirt is a potential source of error for the reading mechanisms employed in most systems.

9. CONCLUSION

The image scanning and character recognition industry is undergoing a period of rapid growth in which different technologies and concepts are being amalgamated. The traditional frameworks for classifying various approaches in this discipline have lost their relevance. In this paper, a new taxonomy was proposed. This taxonomy is based on the characteristics of the input, rather than on the application or the price of the product. Based on the quality and complexity of the input material, it is feasible to predict when off-the-shelf technology will attain the maturity threshold needed to motivate the adoption of these automated technologies in a particular industry.

REFERENCES

[1] All About Optical Readers, Datapro Reports, Datapro Research Corporation, Delran, NJ, 1989.

[2] Ullman, J. R., "Advances in Character Recognition", in Applications of Pattern Recognition, Ed. K. S. Fu, C. R. C. Press, Inc., Boca Raton Florida, 1982, p.230.

[3] Market Report on Imaging, Frost & Sullivan, Inc., 1989.

[4] Shurman, J., "Reading Machines", Proceedings of the International Conference on Pattern Recognition, 1982, pp. 1031-1044.

[5] Lashas, A., et. al., "Optical Character Recognition Based on Analog Preprocessing and Automatic Feature Extraction", Computer Vision, Graphics and Image Processing, Vol. 32, May 1985, pp.191-207.

[6] Baird, H., Simon, K., and Pavlidas, T., "Components of an Omnifont Reader", International Symposium on Pattern Recognition, John Wiley & Sons, 1986, pp. 471-493.

[7] Masuda, I., et al., "Approach to a Smart Document Reader", Conference on Computer Vision and Pattern Recognition, McGraw Hill, 1985, pp. 550-557.

[8] Kida, H., et al., "Document Recognition Systems for Office Automation", Proceedings of the International Symposium on Pattern Recognition, John Wiley & Sons, 1986, pp. 343-362.

[9] Hou, H. S., Digital Document Processing, John Wiley & Sons, 1983.

[10] Suen, C. Y., Berthod, M. and Mori, S., "Automatic Recognition of Handprinted Characters - the State of the Art", Proceedings of the IEEE, Vol. 68, No. 4, April 1987, pp. 469-487.

[11] Harris, B., "Characteristics of Imaging Systems", Byte, March 1988, pp. 76-87.

[12] Mason, J., "Error Rates of Imaging Systems", Imaging, February 1987, pp. 42-46.

[13] Johnson, R., "What High-End Scanners Offer You", Computer World, April 1989, pp. 17-29.

[14] Tanaka, Eiichi, Kohasigushi, Takahiro and Kunihoka, Shiamura, "High Speed String Correction for OCR", Proceedings of the International Conference on Pattern Recognition, John Wiley & Sons, 1986, pp. 340-343.

[15] Wilson, P., "Error Correction in Sophisticated Scanners", Imaging, May 1988, pp. 78-89.

[16] "Survey of Scanning Products", PCWorld, May 1989, pp. 43-49.

[17] "Networked Systems in the Automated Office", Business Week, August 1989, pp. 45-56.

[18] Report on Scanners, International Data Corporation, Framingham, Ma., 1989.

[19] Buffa, Michael, G., "Neural Network technology Comes to Imaging", Advanced Imaging, November/December 1988, pp. 42-50.

[20] Kapoor, Ajit, "Electronic Imaging and Information Processing", The Office, December 1988, p. 56.

[21] Dollard, Tony and Odorisio, Linda, "State of the Art - State of the Industry: A Review of Commercially Available Optical Disk-based Engineering Data Systems", Proceedings of the ATI 87 Conference, Monterey, California, February 1987, pp.456-478.

[22] K. S. Fu, Syntactic Pattern Recognition and Applications, Prentice - Hall, Englewood Cliffs, NJ, 1982.

[23] Nandhakumar, N. and Aggarwal, J., "The Artificial Intelligence Approach to Pattern recognition - A Perspective and Overview", Pattern Recognition, Vol. 18, No. 6, 1985, pp 383-389.

[24] Rich, E., Artificial Intelligence, McGraw - Hill, New York, 1983.

[25] Mantas, J., "An Overview of Character Recognition Methodologies", Pattern Recognition, Vol. 19, No. 6, 1986, pp. 122-134.

[26] Ward, Jean Renard and Blesser, Barry, "Interactive Recognition of Handprinted Characters for Computer Input" IEEE Computer Graphics and Applications, Vol. 23, September 1985, pp. 231-240.

[27] Berman, Ari P., "Image Processing With Embedded Expert Systems", Advanced Imaging, No. 15, November/December 1988, pp. 51-60.

[28] Jones, Tony and Safdie, Elias, "Scanner Selection - More Important Than You Think", Proceedings of the International Electronic Imaging Exposition and Conference, Prentice - Hall, 1987, pp. 341-351.

[29] Brobst, Paul L., "Fax Networking", Proceedings of the International Electronic Imaging Exposition and Conference, Prentice - Hall, 1987, pp. 512-522.

[30] Stern, Randall, "From Intelligent Character Recognition to Intelligent Document Processing", Proceedings of the International Electronic Imaging Exposition and Conference, Prentice - Hall, 1987, pp. 236-245.

APPENDIX

CHARACTERISTICS AND CAPABILITIES OF THE

PRODUCTS EVALUATED IN THE STUDY

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE SET OF CURRENT PRODUCTS

EVALUATION	DestScanners PC 2000	Kurzweil K5100	SO ReliaReader
DOCUMENT PROCESSING Scanning Speed [1]	2.5 pages/min	5 pages/min	1000 chars/sec
Formatting (Alignment, tabs)	Need to specify formats	Preserves tabs, justified text, indents	Need to specify formats
Multiple Column Handling	Multiple scans required	Single pass scan required for multiple columns; no intervention or pretraining required for new formats	Multiple scans required
Handling of graphics/text on same document (Single pass, User-definable zones)	Ability to integrate images into text/ document	One pass text/graphics scanning	Only text handling device
Capability to overlap user-defined zones (Text and graphics)	No	More than one text/graphics zone on page with overlap distinguished by document separator	No
Non-scan colors	Pastels	Red	Light pastels
ERROR HANDLING Accuracy of recognition	Typewritten, good; typeset low	Very high for original quality material	Very high for originals or photocopies

[1] Scanning speed refers to speed taken to scan a high-quality document in a recognizable font.

CHARACTERISTICS AND CURRENT PRODUCTS REPRESENTATIVE

EVAL. CRITERIA	DestScanners PC 2000	Kurzweil K5100	SO ReliaReader
Error flagging Capability	Flagging of unrecognized characters	Interactive batch correction/ on-line; flags unrecognized characters	Error correction either online or batch
GRAPHICS HANDLING Resolution (in d.p.i.)	30 - 300 dpi, software selectable	Upto 400 dpi	300 dpi
Extent of Compression	EPSF(Encapsulated Post Script Format)	CCITT Group 3 or 4 standard	CCITT standard
Raster to Vector Conversion on Line drawings	No	No	No
TECHNOLOGY Recognition Technology	Matrix matching	Feature extraction aided by intelligent character recognition to join broken characters	Matrix matching and feature analysis
Actual Documents Scanned	Yes	Yes	Yes
PHYSICAL DOCUMENT HANDLING (Text/Graphics)	Text or image	Text and graphics	Text (documents)
Document Input Procedure	Automatic feeder optional	Automatic or manual feeder	Automatic feeder
Maximum Hopper Capacity	35 pages with optional sheet feeder	Upto 50 sheets	1000 sheets extendable
Maximum Size Handled	6"x6" to 8.5"x14"	Max. 8.5"x 14" Min. 4"x 5"	3"x4" to 8.5"x14"

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE

SET OF CURRENT PRODUCTS

EVAL. CRITERIA	DestScanners PC 2000	Kurzweil K5100	SO ReliaReader
RECOGNITION OF TEXT Fonts or Styles Read (#) Quality of document (Photocopy, original, dot matrix)	Selected typefonts from typewriters, impact and laser printers; can process good quality first or second generation material	Omnifont, omnisize, mixed typeface; normal and proportional (typewritten, typeset, laser print, dot-matrix, and other computer generated print, and photocopies)	Text and alphanumeric handprints; most fonts can be processed
Point-Size Limitation	8 - 12	6 - 24	6 - 24
Multiple Fonts in Document	No	Automatic omnifont recognition	Automatic omnifont recognition
Font Selection (User specified or automatic)	Through software	Omnifont, omnisize; automatic font selection	Automatic or preset font selection
Character spacing (Mono/ Proportional)	10, 12 pitch and proportionally spaced typefonts	Monospace and Proportional	Monospace or proportional
Character & Layout Attributes (Bold, Italics, Underlining)	Bold and italics	Bold, italics, underscores and super/subscript preserved	Recognized but not preserved
Contrast Flexibility	3 levels under software control	Automatic adjusting for background contrast	Self-adjusting opacity guage
Trainability	Non-trainable	Omnifont recognition, no pretraining; learns from errors	Needs presetting for formats of documents to be scanned

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	DestScanners PC 2000	Kurzweil K5100	SO ReliaReader
OUTPUT HANDLING Editing Program	TIFF (Tagged Image File Format)	ASCII text format and TIFF, PCX graphics format	ASCII text format and TIFF graphics format
Output interface (Ethernet, RS232C, Hosts)	IBM PC, RS232, MAC compatible through SCSI interface card	RS232, Ethernet, SCSI, PC-AT's	Standalone or customized storage devices
Output Format (Wordprocessor, database,wksheet	MacWrite, MSW, or unformatted	16 word processing conversion formats; Lotus format, and databases	Popular word processing formats and databases
Output medium (Host-dependent or standalone)	Host dependent	Host dependent: IBM PC, AT, XT; DOS 3.0 or more; 10 Mb hard disk	Standalone/integrated in Scan Optics office systems
QUOTED PRICE Including software	Under $5,000	$17,950 including co-processor	Over $100,000
REMARKS Major Advantages	High accuracy image processing with 8 bit gray scale(256 levels) and bi-level and half-tone models; sizing and scaling capabilities included	Excellent font reading capabilitiy and supports most conversion formats accurately reducing post-processing	High volume accuracy and speed document reader; can read up to 10,000 documents per hour

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	DestScanners PC 2000	Kurzweil K5100	SO ReliaReader
Major Disadvantages	Needs optional hardware for text processing	No handwritten character recognition capability	Not a very powerful page reader; needs dedicated operator
DEVELOPMENTS EXPECTED Near term, < 1 year	Expected to preserve formatting (e.g. tabs)	Increased processing speed	Not known
Longer term, 1-5 years	Multiple column handling; ability to process all quality material	Limited handwritten character recognition	Increased power in page reading capabilities and less dependent on dedicated operator
LOCAL CONTACT	No contact made	Joan Ennis, Spaulding Inc., Waltham, MA	Gail Graham Scan-Optics, Hartford, CT

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE

SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	CompuScan PCS 240	Data Copy 730	Kurzweil Discover 7320-30
DOCUMENT PROCESSING Scanning Speed [1]	110 char/sec or 3 - 4 pages/min	2 - 3 pages/min	30 - 60 char/sec
Formatting (Alignment, tabs)	User definable	Only alignment	User definable
Multiple Column Handling	Multiple scans required	Multiple scans required	Operator definable zones using a light pen
Handling of graphics/text on same document (Single pass, User-definable zones)	User specified zones	Only graphics or text on a page	Several passes with operator specification of zones at each pass
Capability to overlap user-defined zones (Text and graphics)	No	No	No
Non-scan colors	Red and light pastels	Warm colors	All except light pastels
ERROR HANDLING Accuracy of recognition	99% for typewritten matter	80% and above on original material	Depends on the extent of training

[1] Scanning speed refers to speed taken to scan a high-quality document in a recognizable font.

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	CompuScan PCS 240	Data Copy 730	Kurzweil Discover 7320-30
Error flagging capability	Inserts non-recognized chars.to flag errors	On-line correction and training	3 levels of flagging
GRAPHICS HANDLING			
Resolution (in dpi)	Upto 300 dpi	Upto 300 dpi	Upto 300 dpi
Extent of Compression	30 to 75%	PC software dependent	Depends on the PC software used
Raster to Vector Conversion Capability for Line Drawings	Not available	Not available	Not available
TECHNOLOGY			
Recognition Technology	Matrix matching	Matrix matching	Feature extraction
Actual Documents Scanned	Yes	Yes	Yes
PHYSICAL DOCUMENT HANDLING			
Data Form	Text/graphics	Text/graphics	Both
Document Input Procedure	Automatic	Flat-bed	Flat-bed 30 page hopper optimal

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE

SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	Compuscan PCS 240	Data Copy 730	Kurzweil Discover 7320-30
Maximum Hopper Capacity	50 sheets	1	30
Size Handled	8.5" x 11"	8.5" x 14"	Max. 8.5" x 14"
RECOGNITION OF TEXT			
Fonts or Styles Read	12 fonts; cannot read typeset matter	Cannot read type-set matter	Virtually any font
Quality of Documents (Photo copies, original or dot matrix)	Any quality can be scanned	Any quality can be scanned	Quality dot matrix and good photocopies
Point-size limit	8 - 12	8 - 12	8 - 24
Multiple fonts in document	Can be processed	Can be processed	Can be processed
Font Selection	Automatic	Automatic	Not applicable
Character spacing (Mono/proportional)	Mono spacing	Mono spacing	Any line spacing
Character & Layout Attributes (Bold, italics, Underlining)	Can process underlined text only	Can process underlined text only	Distinguishes only underlined and nonunderlined text
Contrast flexibility	None	None	15 levels of contrast

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE

SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	Compuscan PCS 240	Data Copy 730	Kurzweil Discover 7320-30
Trainability	Not trainable	User trainable for new fonts	No operator intervention, AI software
OUTPUT HANDLING Output Form	ASCII file for text & bit mapped image	ASCII for text and bit mapped image	ASCII text, bit mapped images
Output interface (Ethernet, RS232C or hosts)	RS232C, IBM PC compatible hosts	RS 232C, IBM PC compatible hosts	IBM PC/AT or compatibles
Output format (Wordprocessor, database, worksheet)	ASCII text file compatible with popular word processors	ASCII text file compatible with popular word processors	ASCII file and common PC word processors
Output Medium (Standalone or host-dependent)	Host dependent	Host dependent	Host dependent
QUOTED PRICE Incl. Software	$6,000	$3,000 - $4,000	$6,995
REMARKS Major advantages	Special provision for security of data	Trainable for new fonts	Omnifont reader; able to read several kinds of documents; user friendly interface; user definable lexicon of 10,000 words

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	Compuscan PCS 240	Data Copy 730	Kurzweil Dis- cover 7320-30
Major Disadvantages	Requires a lot of operator intervention	Can process one sheet at a time	Needs training
DEVELOPMENTS EXPECTED Near term, < 1 yr	Not known	New version to pro- cess typeset matter under development	Improvement in the software
Longer term, 1 - 5 years	Omnifont processing capability	Omnifont processing capability	Significant im- in the accuracy
Other Comments	One of the most versatile low-end products available	Regarded as bench- mark product in the low-end area	Hardware for this product is the same as that for Abaton and Microtek scan- ner
LOCAL CONTACT	D.C. Victor Yuan SYNTECH, Inc. Hauppage, NY	D. C. Victor Yuan SYNTECH, Inc. Hauppage, NY	Joan Ennis, Spaulding Co. Waltham, MA

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE

SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	SCAN OPTICS EASYREADER	HAND HELD OCREADER 500 SERIES	TRUSCAN TZ-3BWC
DOCUMENT PROCESSING Scanning Speed [1]	120 documents/min	200 characters/sec	Color: 0.141"/sec B/W: 0.83"/sec
Formatting (Alignment, tabs)	User specified justification; 0 or blank fill	Not available; needs post-processing	Not available; image scanner only
Multiple Column Handling	Must preset page format for >1 columns	Handheld, i.e. user dependent	Not available
Handling of graphics/text on same document (Single pass, user-definable zones)	Text only	Text only (alphanumerics)	Image scanner only
Capability to overlap user-defined zones (Text and graphics)	Must preset page format to overlap zones	Text reading only, handwaving	Image zones only
Non-scan colors	Drop-out color is red with green & black options	All colors except grayscale	Scans all common colors and many more on palette
ERROR HANDLING Accuracy of recognition	High accuracy	60 - 70% on original quality material	High accuracy

[1] Scanning speed refers to speed taken to scan a high-quality document in a recognizable font.

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	SCAN-OPTICS EASYREADER	HAND HELD OCREADER 500 SERIES	TruScan TZ-3BWC
Error flagging capability	Online/batch correction using flagging of errors	Audible and visual feedback confirms data validation	Scanning error reported with operations halted
GRAPHICS HANDLING			
Resolution (in dpi)	Upto 300 dpi	Upto 300 dpi	300 pixels/scanline 75 - 900 dpi
Extent of Compression	Standard CCITT	Not available	Standard CCITT
Raster to Vector Conversion Capability for Line Drawings	Not available	Not available	Not available
TECHNOLOGY			
Recognition Technology	Feature extraction	Matrix matching	Not relevant; scanner not reader
Actual Documents Scanned	Yes	Yes	Yes
PHYSICAL DOCUMENT HANDLING			
Data Form	Text/forms	Text	Images, 3D objects
Document Input Procedure	Automatic feeder	Hand Held Scanner	Open scan table with overhead lens tower

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	SCAN-OPTICS EASYREADER	Hand Held OCReader 500 Series	TruScan TZ-3BWC
Trainability	User programmability for different page formats	User programmability with full flexibility in designing applications	Not possible
OUTPUT HANDLING Output Form	ASCII or EBCDIC	ASCII data transmission emulating keyboard	TIFF and PCX
Output interface (Ethernet, RS232C or hosts)	IBM PC/AT and compatibles	IBM PC, MAC	Apollo, IBM PC, MAC
Output format (Wordprocessor, database, worksheet)	Most PC-based editing software	ASCII text file compatible with popular word processors	Industry-standard format compatible with most graphics editing programs
Output Medium (Standalone or host-dependent)	Host dependent	Host dependent	Host dependent
QUOTED PRICE Incl. Software	Around $10,000	Less than $1,000	$3,000 - $4,000
REMARKS Major advantages	Fast text/forms processor at desktop level with good error checking and editing	Interfaces like a modular I/O peripheral device	High resolution upto 900 dpi; flexible scanning flatbed allowing page, bound document, and 3D object scanning

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	SCAN-OPTICS EASYREADER	Hand Held OCReader 500 Series	TruScan TZ-3BWC
Maximum Hopper Capacity	500 documents	1 page at a time	One page
Size Handled	11.7" X 14.3 "	any size scannable	14.5" x 24"
RECOGNITION OF TEXT			
Fonts or Styles Read	Text,alphanumeric (handwritten)	6 different fonts can be processed	Graphics only
Quality of Documents (Photo copies, original or dot matrix)	Any quality can be scanned	Only originals can be scanned	Any generation quality can be scanned
Point-size limit	6 - 24	8 - 14	Any size scanned
Multiple fonts in document	Yes	No	Any fonts scanned
Font Selection	Automatic	Manual	Scans all fonts
Character spacing (Mono/proportional)	Proportional	Monospace	Scans all spacing
Character & Layout Attributes (Bold, Italics, Underlining)	Bold, italics, and underlining	Bold, italics, and underlining	Scans in attributes
Contrast flexibility	Automatically adjusting opacity guage	Cannot handle changes in contrast	Contrast levels for red, green,and blue under operator control

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE

SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	SCAN-OPTICS EASYREADER	Handheld OCReader 500 Series	TruScan TZ-3BWC
Major Disadvantages	No image scanning, and less powerful as a page scanner than as a document scanner.	No image scanning	Scanning speed is slow; no text processing
DEVELOPMENTS EXPECTED Near term, < 1 yr	Not known	Read multiple formats	Not Known
Longer term, 1 - 5 years	Automated multiple column handling	Omnifont recognition	Text processing capabilities
Other Comments	Local area networking feature through Scan-Optics 3200 systems; through a control unit, it can communicate with a variety of host systems	Completely transparent to host's hardware, operating system, and application programs	Can scan 3D objects because of flexible scanner bed
LOCAL CONTACT	Gail Graham, Hartford, Connecticut	No contact made	No contact made

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE

SET OF CURRENT PRODUCTS

EVALUATION	Microtek 300 C	Calera CDP 9000	IOC Reader
DOCUMENT PROCESSING Scanning Speed [1]	Host dependent: 80 char/sec with IBM AT	250 char/sec 6 - 10 pages/min	100 char/sec 3 pages/min
Formatting (Alignment,tabs)	Stores the formatting	Preserves tabs, justified, and paragraphs	Maintains formatting
Multiple Column Handling	No	Single pass for multiple columns, charts, tables, and photographs in document; 256 user definable zones	Needs 2 passes; Vertical length of graphics/text area can be specified
Handling of graphics/text on same document (Single pass, User-definable zones)	No	Intermixed text and graphics; template-based user definable zones; registration zones to adjust for page skew	No
Capability to overlap user-defined zones (Text and graphics)	No	Can have more than one zone or page for text/graphics with overlap	No
Non-scan colors	No	Light pastels	Light green, yellow
ERROR HANDLING Accuracy of recognition	Good on typewritten, low on typeset	99.9% recognition rate for optimal documents	Very good for specific fonts

[1] Scanning speed refers to speed taken to scan a high-quality document in a recognizable font.

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE SET OF CURRENT PRODUCTS

EVAL. CRITERIA	Microtek 300 C	Calera CDP 9000	IOC Reader
Error flagging capability	Yes	Online/batch correction; flagging of indeterminate characters	One level of flagging
GRAPHICS HANDLING Resolution (in d.p.i.)	300 dpi	300 dpi from CDP scanner & 1200 dpi from others	200-400 dpi
Extent of Compression	Depends on formatting	CCITT Group 3 or 4 standard	CCITT standard
Raster to Vector Conversion on Line drawings	No	No	No
TECHNOLOGY Recognition Technology	Feature extraction (structural analysis)	Proprietary recognition algorithms aided by contextual analysis using 3 built-in dictionaries	Matrix matching
Actual Documents Scanned	Yes	Yes	Yes
PHYSICAL DOCUMENT HANDLING (Text/Graphics)	Both	Same page text, graphics	Text, graphics
Document Input Procedure	Automatic feeder	Automatic sheet feeder	Automatic sheet feeder
Maximum Hopper Capacity	50 pages with optional sheet feeder	Up to 50 sheets	30 sheets
Maximum Size Handled	8.5"x 14"	Max. 8.5"x 14" Min. 3"x 5"	Max. 8.5"x 14" Min.5"x 5"

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE SET OF CURRENT PRODUCTS

EVAL. CRITERIA	Microtek 300 C	Calera CDP 9000	IOC Reader
RECOGNITION OF TEXT			
Fonts or Styles Read (#)	Limited number of fonts	Omnifont, omnistyle, omnisize; handles normal and proportional (typewritten, typeset, laser print, dot-matrix, and other computer generated print);20,000 fonts.	10 fonts read including Pica, Prestige, and Courier. Can read dot matrix but has trouble with photocopies.
Quality of document (Photocopy, original, dot matrix)			
Point-Size Limitation	6 - 12	6 - 28	10 - 12
Multiple Fonts in Document	Yes	Yes, without pretraining or setting	Yes
Font Selection (User specified or automatic)	Automatic	Automatic font selection	User specification possible. Automatic mode takes longer.
Character spacing (Mono/ Proportional)	Typewritten material	Mono & proportional	Mono only
Character & Layout Attributes (Bold, Italics, Underlining)	Reads but does not restitute the attributes.	Bold, italics and underlining read and preserved to exact (x,y) text location.	Yes, but only underlining coded material.
Contrast Flexibility	3 levels for text, 8 levels for graphics	Auto-thresholding to adjust background	Not adjustable
Trainability	Non-trainable	Capable of omnifont recognition without machine training.	Not trainable

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	Microtek 300 C	Calera CDP 9000	IOC Reader
OUTPUT HANDLING Editing Program	Editing with common PC word processors	ASCII or ANSI text and bitmapped images	ASCII text and bit-mapped images
Output interface (Ethernet, RS232C, Hosts)	IBM & compatibles, Macintosh	RS232, Ethernet, PC-AT's, Sun workstations	IBM PC/XT/AT and other MSDOS compatibles
Output Format (Wordprocessor, database,wksheet	Common wordprocessor formats	Most popular word processors, desktop publishing tools, databases	Wordprocessor and graphics editing
Output medium (Host-dependent or standalone)	Host dependent	Host dependent/ stand alone; stand alone needs host to set parameters	Host dependent
QUOTED PRICE (Including software)	$2,500	$29,000-$30,000 with basic imaging and reading software	$3,500
REMARKS Major Advantages	Very good graphic capabilities; good performance of the software considering the price; good user interface	Exceptional font reading capability; capability to process tables and forms; suitable for intensive production use	Specific font reading capabilities can be developed by the company to order; excellent for half-tones; 128 shades of grey available

CHARACTERISTICS AND CAPABILITIES OF REPRESENTATIVE

SET OF CURRENT PRODUCTS

EVALUATION CRITERIA	Microtek 300 C	Calera CDP 9000	IOC Reader
Major Disadvantages	Limited capabilities for typeset and proportionally spaced material; poor ability to handle multiple column and mixed text and graphics documents	No capability for handwritten character recognition	Cannot read typeset material
DEVELOPMENTS EXPECTED Near term, < 1 years	Improved reading of typeset material	Networking to other scanners and hosts	Ability to read photocopies
Longer term, 1-5 years	Improved accuracy; could become an excellent tool for desktop publishing & the low to medium range market	Raster to vector conversion for line drawings; advanced AI techniques for natural language parsing leading to even higher accuracy	Typeset reading capability
LOCAL CONTACT	Ken Barber Provue, Inc. 145 South St. Boston. MA	Robert Warner Calera Corp. 8 New England Exec Park, Burlington, MA	Bob Haskell System Automation, Inc. 15 Lakefield, Office Park Wakefield, MA