

# The Role of Holistic Paradigms in Handwritten Word Recognition

Sriganesh Madhvanath, *Member, IEEE*, and Venu Govindaraju, *Senior Member, IEEE*

**Abstract**—The Holistic paradigm in handwritten word recognition treats the word as a single, indivisible entity and attempts to recognize words from their overall shape, as opposed to their character contents. In this survey, we have attempted to take a fresh look at the potential role of the Holistic paradigm in handwritten word recognition. The survey begins with an overview of studies of reading which provide evidence for the existence of a parallel holistic reading process in both developing and skilled readers. In what we believe is a fresh perspective on handwriting recognition, approaches to recognition are characterized as forming a continuous spectrum based on the visual complexity of the unit of recognition employed and an attempt is made to interpret well-known paradigms of word recognition in this framework. An overview of features, methodologies, representations, and matching techniques employed by holistic approaches is presented.

**Index Terms**—Handwriting recognition, holistic paradigms, analytical methods, reading theory, pattern recognition.

## 1 INTRODUCTION

**H**ANDWRITTEN Word Recognition (HWR), also called *Isolated* Handwritten Word Recognition, deals with the problem of machine reading handwritten words. There are two different problems that fall under the purview of handwritten word recognition: Offline HWR and Online HWR.

Offline HWR deals with the problem of reading a handwritten word *offline*, that is, at some point in time (minutes, months, years) after it was written. A handwritten word is typically scanned in from a paper document and made available in the form of a binary or gray-scale image to the recognition algorithm.

The problem differs from *online* HWR where the writing is with a special pen on an electronic notepad or a tablet and where temporal information, such as the position and velocity of the pen along its trajectory, is available to the recognition algorithm. Since most algorithms for online HWR attempt to recognize the writing as it is being written, online HWR is also sometimes referred to as “real-time” HWR.

This survey focuses on the task of offline HWR. However, the discussion is pertinent to the online problem as well.

### 1.1 The Offline HWR Task

Some applications of offline HWR today are recognition of handwritten check amounts, interpretation of handwritten addresses on pieces of mail, reading handwritten responses

on forms, and automatic filing of faxes. The handwritten text must be located, extracted, made free of artifacts stemming from the medium (underlines and background from the check leaf, boxes from forms, postal marks from the piece of mail), separated into lines if necessary, and, finally, into individual words before it can be recognized. These steps are generally nontrivial and research issues in their own right. We assume in this survey that the complex task of segmentation of the image of the handwritten word or phrase of interest from its surroundings has already been accomplished by prior processes. The tasks of segmentation and recognition of words are generally accomplished sequentially based upon different features of the image. They are consequently difficult to combine, except superficially in the sense that word recognition is used to choose from multiple word segmentation hypotheses. We will focus in this survey on the task of *recognition* of the isolated word or phrase using the appropriate lexicon (Fig. 1).

The handwritten word or phrase may be constrained by the application to be in a particular style. For example, forms often request that the responses be handprinted. In general, however, handwritten words may be cursive, purely discrete, touching discrete, or a mixture of these styles (Fig. 2). While for some applications of online HWR, a single author assumption can be made and the algorithms tuned to a particular style of writing, this assumption cannot generally be made for the offline problem. Consequently, the recognition algorithm must deal with a variety of author-specific idiosyncrasies.

Moreover, there is little or no control in most offline scenarios on the type of medium and instrument used. The artifacts of the complex interactions between medium, instrument, and subsequent operations such as scanning and binarizations present additional challenges to algorithms for offline HWR. Offline HWR is, therefore, generally regarded as much more difficult than its online counterpart.

• S. Madhvanath is with the IBM Almaden Research Center, San Jose, CA 95120. E-mail: srig@almaden.ibm.com.

• V. Govindaraju is with the Center of Excellence for Document Analysis and Recognition (CEDAR), Department of Computer Science and Engineering, State University of New York at Buffalo, 226 Bell Hall Amherst, NY 14260. E-mail: govind@cedar.buffalo.edu.

Manuscript received 18 Nov. 1999; revised 18 Apr. 2000; accepted 6 Nov. 2000.

Recommended for acceptance by R. Plamondon.

For information on obtaining reprints of this article, please send e-mail to: tpami@computer.org, and reference IEEECS Log Number 110966.

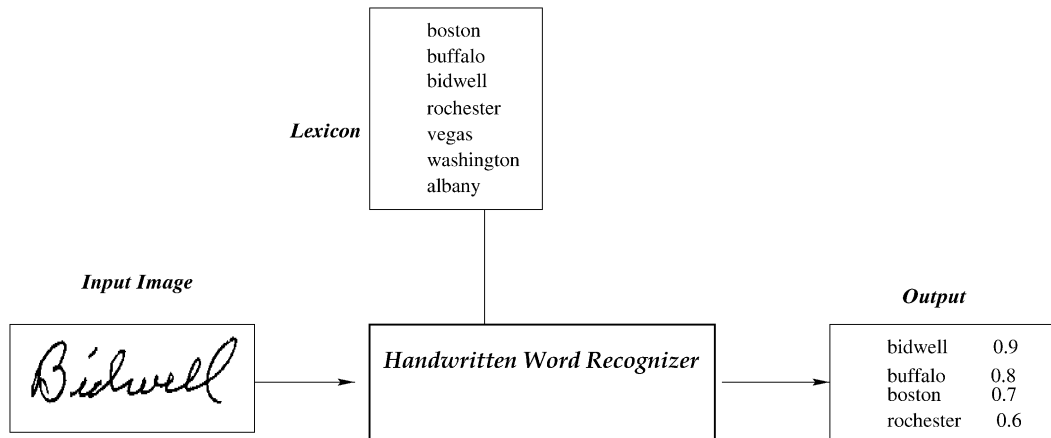


Fig. 1. I/O behavior of offline HWR. Input is the word image and a lexicon of possible choices. Output is the lexicon sorted by some confidence measure.

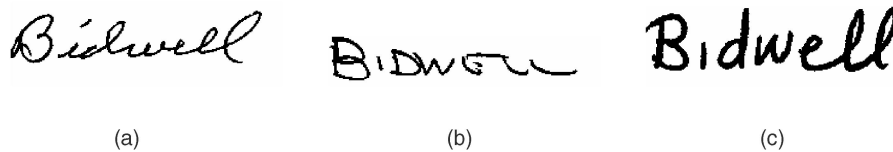


Fig. 2. Examples of handwriting styles. (a) Cursive, (b) discrete touching, and (c) mixed.

Since words are fairly complex patterns and owing to the great variability in handwriting style, the HWR task is a difficult one. In fact, it is only made tractable when a lexicon of valid words is provided. The lexicon is usually determined by the application domain. For example, there are only 33 different words that may appear in the so-called legal amounts on handwritten checks. The lexicon for this application, is hence, both small and *static*, i.e., constant across all recognition instances (Fig. 3).

The lexicon used for street name recognition in Handwritten Address Interpretation (HWAI) is generally comprised of street name candidates generated from knowledge of the zip code and the street number. This is an example of an HWR application where the lexicon is *dynamic*, i.e., varying from one instance to the next (Fig. 4). Some

applications, such as the reading of handwritten prose, may involve very large lexicons of over 20,000 words. The nature of the lexicon is crucial to the design of HWR algorithms for a particular application.

## 1.2 Holistic Approaches

From the earliest days of research in HWR, two approaches to the problem have been identified. The first approach, often called the *analytical* approach, treats a word as a collection of simpler subunits such as characters and proceeds by segmenting the word into these units, identifying the units and building a word-level interpretation using the lexicon. The other approach treats the word as a single, indivisible entity and attempts to recognize it using features of the word as whole. The latter approach is referred to as the *word-based* or *holistic* approach and is

Two thousand and forty five only.

|         |           |          |          |         |
|---------|-----------|----------|----------|---------|
| one     | two       | three    | four     | five    |
| six     | seven     | eight    | nine     | ten     |
| eleven  | twelve    | thirteen | fourteen | fifteen |
| sixteen | seventeen | eighteen | nineteen | twenty  |
| thirty  | forty     | fifty    | sixty    | seventy |
| eighty  | ninety    | hundred  | thousand | dollars |
| dollar  | and       | only     |          |         |

Fig. 3. Handwritten legal amount recognition involves the recognition of each word in the phrase matched against a static lexicon of about 33 words.

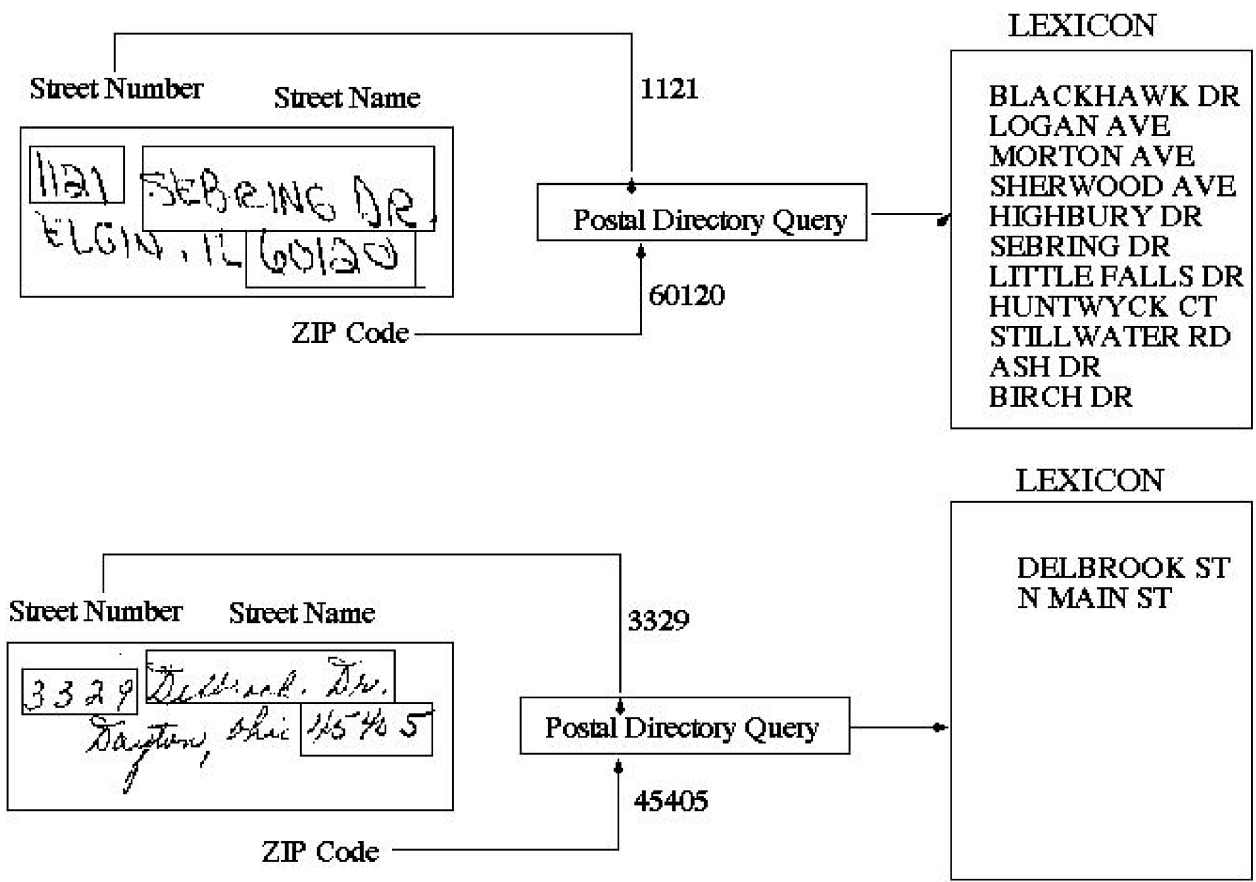


Fig. 4. A lexicon is dynamically created using the zip code and the street number. Note that the street name images are matched with different lexicons generated. (a) In zip code 60120 there are 11 streets which have the street number 1121. (b) In zip code 45405 there are two streets with the street number 3329.

inspired in part by psychological studies of human reading, which indicate that humans use features of word shape such as *length*, *ascenders*, and *descenders* in reading (Fig. 5).

Because analytical approaches decompose HWR into the problem of identifying a sequence of smaller subunits, the chief problems they face are 1) *segmentation ambiguity*: deciding where to segment the word image (Fig. 6) and 2) *variability of segment shape*: determining the identity of each segment (Fig. 7), [13].

Holistic approaches circumvent these problems because they make no attempt to segment the word into subunits.

Instead, they rely on features and matching at the word-level to determine the identity of the word.

### 1.3 Relevance of the Holistic Paradigm

Analytical approaches that decompose handwritten words into characters or other subunits derived from characters do not generally distinguish between static and dynamic lexicons; random strings of characters are recognized as effectively as valid words.

For holistic approaches, on the other hand, every word is a different class. The holistic features and matching scheme

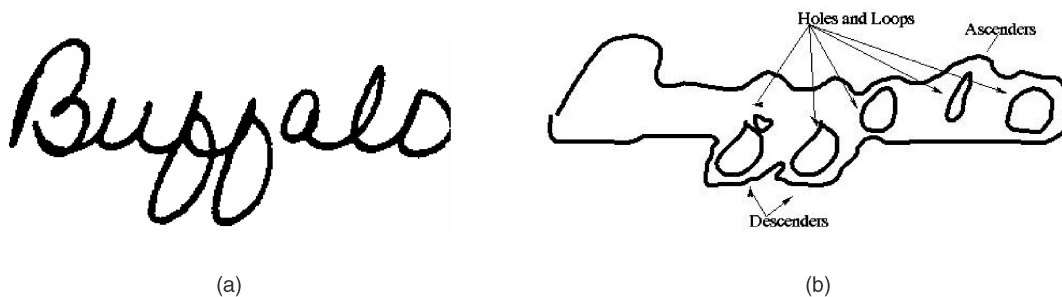


Fig. 5. (a) Word image. (b) Word-shape features do not refer to individual characters and include length, ascenders, descenders, loops, etc.

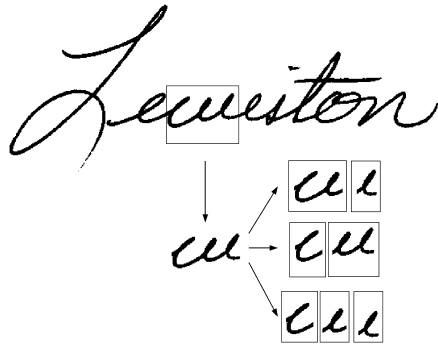


Fig. 6. Ambiguities in segmentation: The letter(s) following the “e” can be “w,” or “ui” or “iu” or “iii.”

used must be coarse enough to be stable across exemplars of the same class(word), i.e., across a variety of writing styles, but fine enough to be able to distinguish exemplars of different classes. Given that words are complex 2D patterns and given the large variety of writing styles, these are difficult criteria to satisfy when the number of classes is large or unknown. Hence, holistic approaches have been used traditionally in application scenarios wherein the classes are few and fixed. For example, the check amount recognition task (Fig. 3). Moreover, when the lexicon is small and static, it becomes possible to collect a large number of training samples of each class. Training may then be performed in the traditional sense of estimating class-conditional densities of features from the training samples or storing prototypical feature vector exemplars for each class.

When the lexicon is large or dynamic [11], [22], [23], [34] (handwritten address interpretation example in Fig. 4), the ability of any given set of holistic features to distinguish between word classes is diminished. In addition, it is difficult or impossible from a practical standpoint to obtain representative samples of all word classes for training a holistic classifier. For these reasons, there is consensus among researchers in the field of HWR that the utility of the holistic approach is either in the small, static lexicon scenario or in the filtering of large lexicons. For example, a survey of the state of the art in online HWR [52] concludes that

While the [whole-word] approach can be useful for small vocabularies, current thinking is that it is not viable for the general problem [of classification of handwritten words].

There are two issues that must be emphasized in this context.

First, classification is only part of the problem of recognition of offline handwritten words. Given the

difficulty of the task, practical recognition engines must employ multiple classification algorithms and complex strategies for combining classifier decisions [35]. Fig. 8 shows the role of a holistic recognizer in the complex combination of recognizers used by a handwritten address interpretation system [36].

Second, the merit of a particular paradigm is best judged by its cost/accuracy benefits, rather than by accuracy alone. An algorithm that is highly accurate at classifying words is not viable in practice if the computational cost involved is unreasonable. Conversely, an algorithm, such as the holistic recognizer, with relatively low accuracy may prove beneficial if used in conjunction with more accurate algorithms and if the additional computational burden is relatively small.

This investigation into holistic approaches is further motivated by the following observations:

- **Intrinsic advantages of the holistic paradigm.** By circumventing segmentation issues and treating each word as a class unto itself, holistic approaches have the potential to model effects that are unique to the class. For example, they can model *coarticulation effects*, i.e., the changes in the appearance of a character as a function of the shapes of neighboring characters. Fig. 5 shows the two “f”s written have different shapes depending on what precedes and follows them. Generally speaking, algorithms based on the holistic paradigm are computationally efficient.
- **Orthogonality of holistic features.** Holistic features provide information about the word that is clearly orthogonal to the knowledge of characters in it and it stands to reason that the introduction of this knowledge should improve recognition. For example, a Holistic approach may succeed when the writing is so poor that the individual characters cannot be distinguished but the overall shape of the word is preserved (Fig. 9).
- **Evidence from psychological studies.** A large body of evidence from psychological studies of reading (Section 2) points towards the use of a holistic approach in conjunction with analysis of letter identities—humans do not, in general, read words letter by letter. A computational theory of reading should include the holistic paradigm.
- **Potential benefits for HWR engines.** The recognition of unconstrained handwritten words is a challenging problem that may be addressed only when a lexicon is available. Existing recognition algorithms show a decline in both accuracy as well

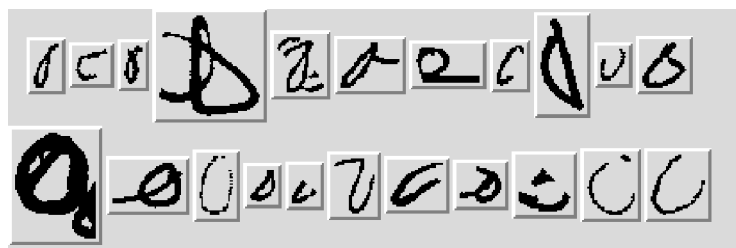


Fig. 7. Wide variability in shapes of characters (“o” in this example) even when taken from the writing of the same writer.

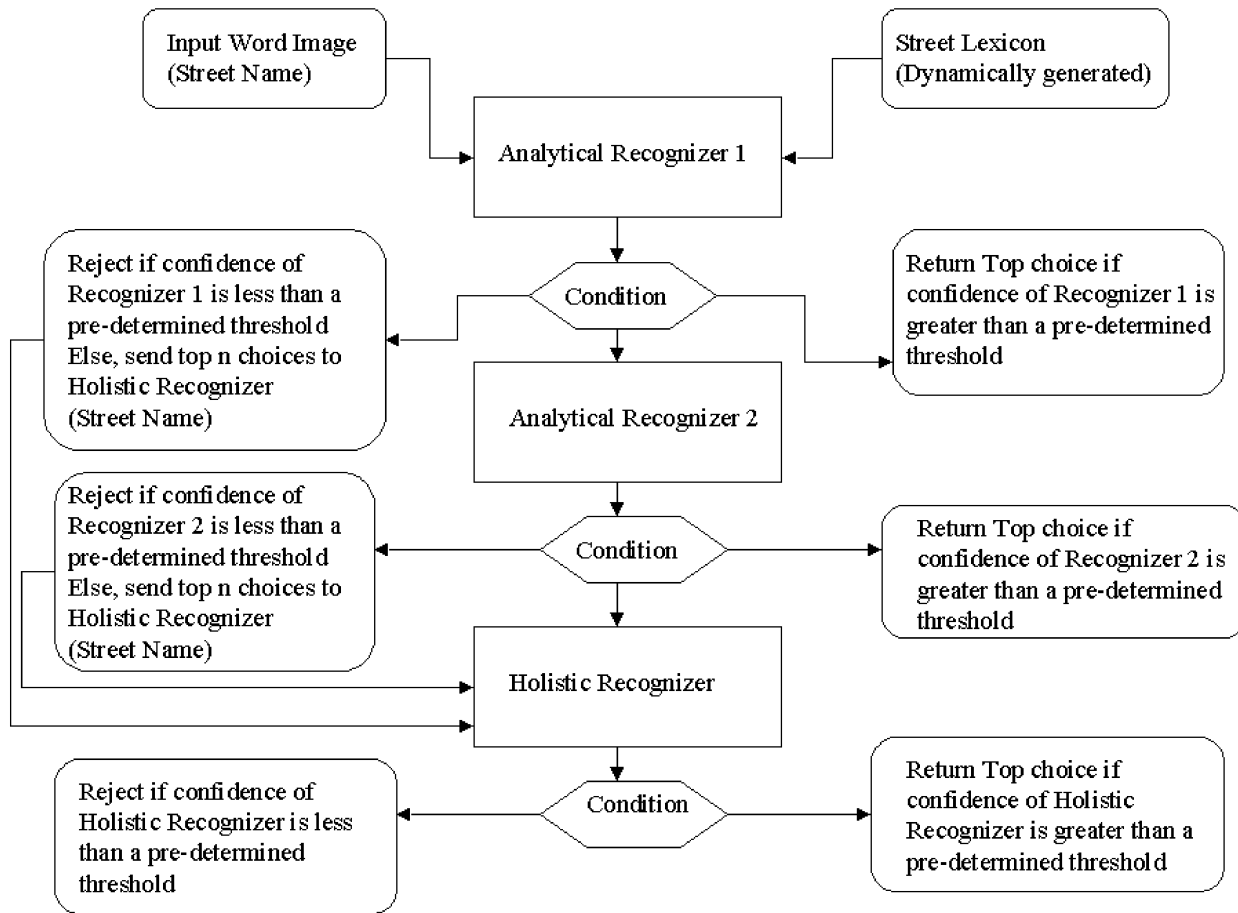


Fig. 8. Combination of three word recognizers, two analytical and a holistic in the context of handwritten address interpretation. The holistic recognizer serves as a “tie-breaker” between the top choices of the two analytical recognizers.

as computational efficiency when confronted with real-world recognition scenarios involving noisy images and large lexicons. The use of multiple classifiers with substantially different features and approaches, decision combination methods, and complex strategies for thresholding are some ways of combating the decline in performance (Fig. 8).

We hope that this survey will encourage the reader to reexamine the consensus about the role of the holistic paradigm in offline handwritten word recognition.

#### 1.4 Organization of Survey

In Section 2, we discuss some of the findings from experiments with human readers. An important motivation for investigating the holistic paradigm comes from the fact that humans use holistic features in reading and tend to read whole words at a time. Psychological studies have demonstrated the robustness of human reading skills in the presence of large distortion or incomplete information at

lower levels of the text hierarchy. Fluent reading appears to involve the recognition of word patterns rather than individual letter patterns. In Section 3, we attempt to refine the distinction between holistic and nonholistic approaches in order to better comprehend the methods proposed in the literature. We discuss various broad classifications of the holistic methods surveyed in Section 4 and survey holistic features, representations, and matching methodologies. The survey is summarized in the concluding section.

## 2 THE HOLISTIC PARADIGM AND THE PSYCHOLOGY OF READING

It is no surprise that a dichotomy analogous to holistic/analytical approaches to machine recognition of words is also the center of a long-standing debate in reading studies. An excellent survey of this debate is presented by Soltysiak [51] and forms the basis for this section.

Visual recognition of words has been widely investigated by psychologists during the past century (for example, [8], [44], [56], [37], [27]) and has produced two very different interpretations. *Holistic* theories suggest that words are identified directly from their global shape; the opposing view of *hierarchical* theories is that recognition results from identifications of component letters. These theories do not hypothesize different mechanisms for

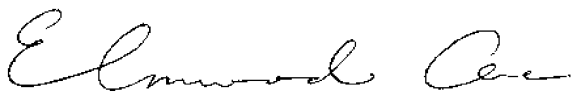


Fig. 9. Images with low character information cause problems for analytical approaches.

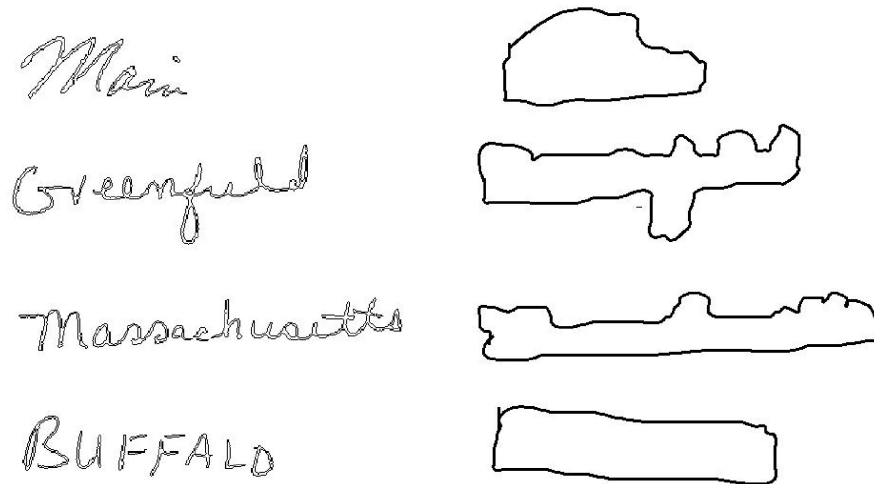


Fig. 10. The word shape of cursive words alone contains sufficient information to classify the image as one of the lexicon words. Words written completely in upper case, such as BUFFALO in the example do not possess such information.

printed and handwritten words; the studies supporting them deal primarily with printed words.

Holistic theories of reading propose that reading is accomplished using stored encodings of shapes of words. They predict that lowercase words are easier to read than uppercase words and that familiar words such as function words are easier to read than unfamiliar words. They also predict degraded recognition performance when word shape is disrupted, with this degradation being more pronounced for words compared to nonwords, and familiar words compared to unfamiliar ones.

Fig. 10 illustrates how word shape contains sufficient information to classify words in certain small lexicons. The perceptual features that are being invoked are perhaps the length of the word, the relative positions of ascenders and descenders, and other cues. It is also to be noted that if a word is written entirely in uppercase, there are no shape features present.

Hierarchical theories, on the other hand, hypothesize that words are recognized from letters and letters from features detected in the stimulus. Letter detectors are thought to contain information solely about letter identities and not their visual form and letters are thought to be processed in parallel. The role of a hierarchical mechanism in reading is widely accepted. As argued by Coltheart in 1981, abstract letter identification enables reading of words in typeface never before encountered. McClelland (1977) argued that it is identification of letters that allows words to be recognized as they are "the only invariant cues to the identity of words." There is, however, evidence to suggest that this need not be the sole means by which words are recognized. Recently, models that combine these conflicting interpretations have been proposed based on evidence from studies of individuals with acquired dyslexia (especially [25]) and studies of reading development (see, for example, [48]). These models propose that holistic and hierarchical processes operate in parallel in both the developing and the skilled reader.

Different holistic theories define word shape in different terms—word envelopes, shapes and sizes of individual

letters, arrangements of ascenders, descenders and neutrals, digrams, and spelling units. Early evidence for holistic theories was provided by studies that showed that the time required to initiate naming of a word was less than that of a single letter—the well-known *word superiority effect* [8].

Moreover, word regularity effects (regular words such as MINT read aloud faster than irregular words such as PINT) and semantic priming effects (context facilitating word recognition) have been found to be more pronounced for upper than lowercase words [55] and have argued to indicate holistic recognition of words when word shape is available and more detailed analytical recognition process in the absence of such information. In another study, subjects were asked to guess the next word in a sentence given varying amounts of information about the next word [30]. It was found that guessing accuracy was enhanced when word length information was provided and further improved when word shape information was made available.

Studies involving proofreading tasks [39], [19] provide further evidence for word shape in word recognition. These tasks involved recognition of words in text passages, the words mutilated by substituting or deleting letters. Certain mutilations involved deletion of a perceptual feature such as an ascender or descender, or substitution of a perceptual feature by a neutral character, causing a large change in word shape (e.g., "fastest" became "fascest" or "faset"). Others involved deletion of a neutral character or substitution of a neutral character by another, and caused only a small change (e.g., "fastest" became "fastect" or "fastet"). It was observed that misspellings that preserved word shape were less noticeable than those that disrupted word shape. This has been argued to support a two-stage model of visual analysis [2], [3] involving the cyclical interaction of a passive global process which selects a set of words matching the shape of the stimulus and an active local process which "fills-in" details to permit full identification of the stimulus.

The evidence for holistic recognition of words is commonly criticized as being inconclusive because of the failure of studies to independently manipulate confusion of

letters from that of word shape; apparent word shape effects may be explained at the letter level using the argument that lowercase letters are more distinct than their uppercase counterparts [21]. However, there is evidence for a parallel reading mechanism in humans that is based solely on word shape. Studies of individuals with acquired reading disorders or dyslexia suggest that different forms of dyslexia result from impairment at different levels of the human visual word recognition system. Surface dyslexia and letter-by-letter reading [48] may be explained in terms of damage to the word representation or its connections with the letter detectors.

However, there is evidence from studies of individuals with deep acquired dyslexia, especially "TM," described by Howard in [25]. TM had great difficulty matching words across case. TM was much worse at reading words with letters separated by "+" (as in "w+o+r+d+s"), while words with letters separated by spaces (w o r d s) were read as well as when not so separated. In addition, TM was significantly worse at reading case alternated (WoRdS) and diagonally written words; and unable to understand abbreviations when presented in inappropriate case (e.g., E.G.).

TM is unable to either extract or use the abstract identities of the letters that constitute the word. This directly contradicts the prediction of hierarchical theories that no word can be recognized if letter identity information is unavailable. These results are seen as evidence for a reading mechanism that is completely independent of letter identities. A visual word recognition system with two available routes has been proposed by a number of researchers. There are also a number of theories about how the two kinds of information are combined for fluent reading, which are beyond the scope of this review. It is clear from these studies, however, that word shape plays a significant role in visual word recognition both in conjunction with character identities, as well as in situations wherein component letters cannot be discerned.

This review would not be complete without some mention of the research that suggests that the mechanisms used for recognition of cursive script may differ in fundamental ways from those used for printed words [56]. Most hierarchical models of reading assume a model of parallel processing wherein features of individual letters simultaneously activate words in the mental lexicon [53]. The fact that individual letters are easy to segment from the background would suggest that word shape features such as ascenders and descenders provide little additional information to aid in the recognition of letters and, ultimately, the word. Clearly, this is not true of cursive script. In one study conducted at the Nimjen University in the Netherlands, the presence of ascenders and descenders was found to have an impact on both reading speed and error rate [47]. In particular, reading speed was seen to decrease for cursively written words which have no ascenders or descenders.

### 3 PARADIGMS FOR VISUAL WORD RECOGNITION

In the Section 1, we presented two paradigms for word recognition: analytical and holistic. In this section, we attempt to refine the distinction between these paradigms

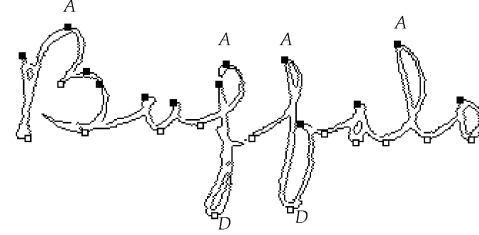


Fig. 11. Ascenders and descenders are perceptual features. Features such as the proportion of pixels in the left and right segments, number of extrema, the perimeter of the word, etc., are examples of holistic features.

and clarify what a holistic approach is and what it is not. A review of the literature reveals a variety of interesting methods which are difficult to classify as one or the other. In addition, there appear to be at least two different senses in which the term "holistic approach" has been used in the literature: 1) an approach that matches words as a whole and 2) an approach that uses word shape features. It is important to distinguish holistic features from holistic approaches. A holistic approach may or may not use word shape features. For example, it may use pixel direction distribution features [54]. Conversely, a classifier may use word shape features in an approach that is not holistic to perform segmentation [15], [20] and/or character recognition.

The term "global features" has been used by some researchers to refer to "simpler aspects of word shape that can be easily and reliably measured" [54]. Often, this refers to estimates of word length and counts of perceptual features such as ascenders and descenders (Fig. 11).

Hierarchical theories of reading postulate the use of letter models as part of the recognition process, whereas holistic theories of reading suggest that the word identity is determined directly from word shape features extracted from the stimulus (Fig. 10). The holistic/analytical distinction differs from this holistic/hierarchical dichotomy encountered in reading studies in significant ways. Analytical approaches for handwritten word recognition are not limited to the use of letters as models; conversely, an approach that uses nonletter models would be considered analytical rather than holistic. In fact, the holistic/analytical classification is a continuous spectrum rather than a dichotomy, as is evidenced from the methods surveyed. Fig. 12 illustrates this particular point. At the one end, we have a word represented as an array of pixels and on the other as an ASCII string. The features move from the fine-grained pixels to the holistic shape of the word. In the middle, we have features ranging from the purely analytical, such as strokes, loops, and characters, to the holistic, such as histogram profiles of words and pixel density distributions. The exact line where we depart from the analytical and move to holistic is in fact a gray band.

#### 3.1 Features and Models

Features may be used directly to determine the identity of a word image or they may be used to determine the identity of intermediate entities which constitute the word image. We will refer to these intermediate entities as submodels. Features such as strokes of characters, loops, t-crossings,

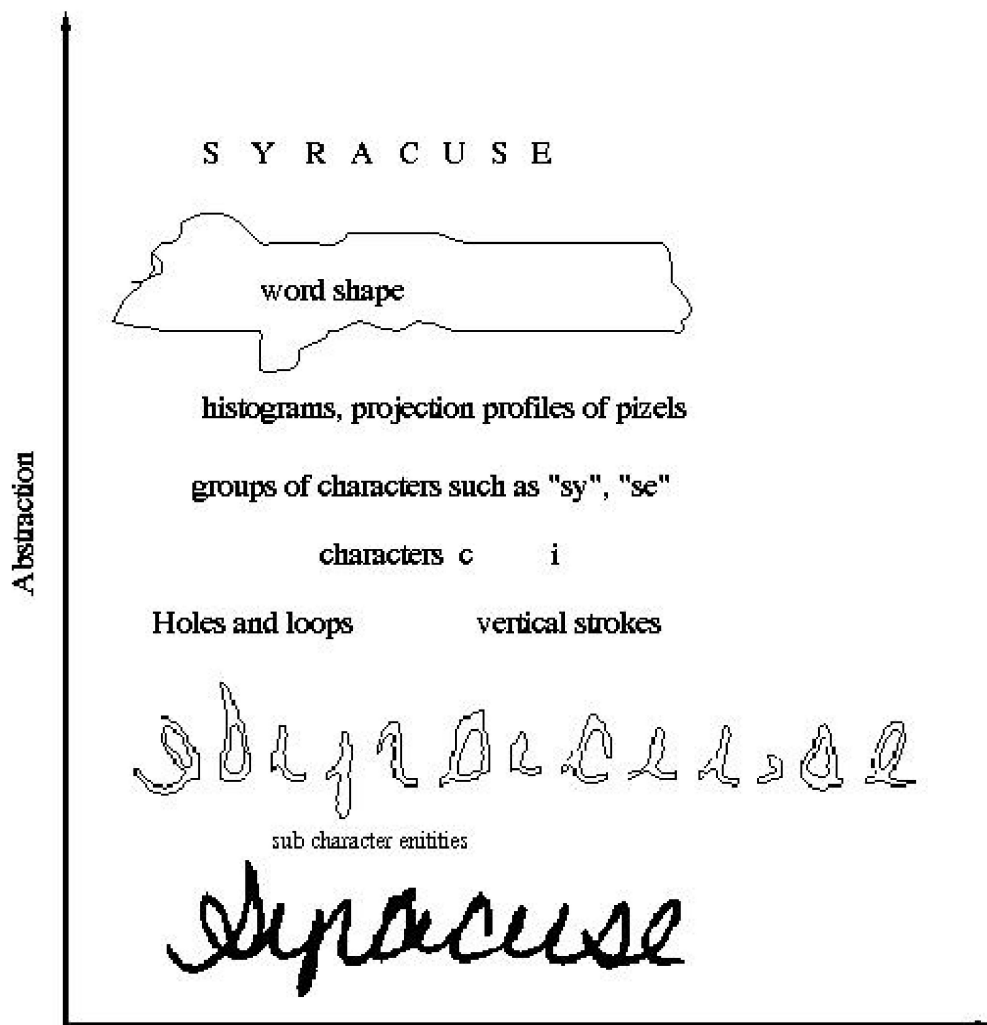


Fig. 12. The continuum of features moving from the fine-grained pixel level features to the coarser features of strokes, then characters, groups of characters and, finally, the ASCII representation of the word. The line between holistic features and analytical features is a actually subjective.

i-dots, and individual characters themselves are all regarded as submodels. The use of submodels is motivated by the realization that exemplars of different classes are not unrelated, rather, they are composed of common subentities. The burden of dealing with variations in the input may effectively be shifted to the level of submodels, where they are more constrained and easier to characterize and compensate for. The assumption implicit in the use of submodels is that they are consistent when they appear in different positions and in exemplars of different classes.

We would like to formally define a holistic approach as one that does not use submodels as part of its classification strategy. We refer to nonholistic approaches as model-based approaches. These approaches have traditionally been called "analytical," but holistic approaches may involve detailed analysis of the word as well. In fact, in an early survey of HWR [29], Frishkopf's approach of encoding the trace of the word as a sequence of extrema points and matching the entire sequence against a lexicon of similarly encoded words—essentially a holistic approach—is classified under analytical approaches. Therefore, we prefer the term "model-based" to describe approaches that employ

submodels. However, we will continue to use the familiar term "analytical" to refer to such approaches.

Previous reviews of the literature in HWR have proposed similar taxonomies for methods [7]. These taxonomies have been based largely on the scheme used to segment the word image, and holistic approaches have been described as those which use no segmentation or implicit segmentation. The uniqueness of the taxonomy presented here, in our opinion, lies in the fact that it is based on submodels used in the recognition process, rather than the segmentation scheme used.

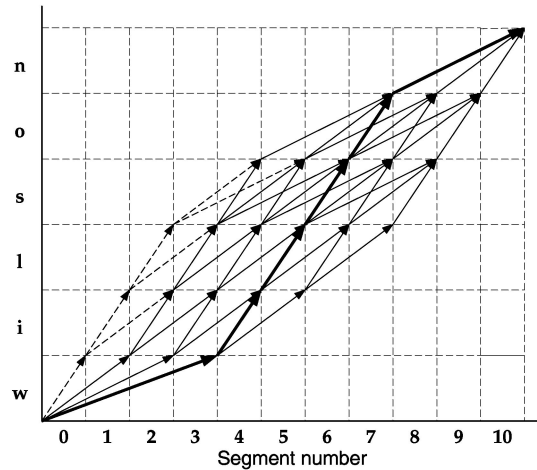
### 3.2 Analytical (Model-Based) Approaches

The central idea in a model-based approach is to identify parts of the word image as one of a predefined set of models known to the classifier [46]. A "circular" situation arises here: in order for a piece of the image to be identified as a model, it must first be segmented; but in order for it to be correctly segmented, it must be identified as a valid model first. Casey and Lecolinet [7] introduce the term "dissection" to refer to a partitioning of the image based on image features alone (i.e., without involving recognition). Different model-based





(a)



(b)

Fig. 13. (a) Image features based dissection of a word image. (b) Dynamic matching of the segments (intermediate entities) with the lexicon word WILSON.

methods place different emphasis on the subprocesses of recognition and dissection to arrive at a final segmentation such that each identified segment corresponds to one of the models. At one end of the spectrum are methods which perform an image-feature-based dissection and use recognition of segments to detect and correct segmentation errors (Fig. 13). At the other end of the spectrum are methods which scan the image looking for models. This scanning may be of either the image or some representation thereof. Here, segmentation is driven by recognition and some researchers have used the confusing term “segmentation-free” to describe such approaches. Fig. 14 illustrates this process. The recognition of different characters “peaks” as a window (of the size of a character) slides along the image.

### 3.3 Holistic Approaches

Holistic approaches do not attempt to label parts of the image using sets of models; instead they extract holistic features from the word image and use the features directly to arrive at the word identity. In order for this feature-level matching to be possible, every candidate from the lexicon must have a feature representation similar to that used to represent the image features. The process of constructing a lexicon in which each lexicon entry is represented by its holistic features, or statistics about holistic features in the case of probabilistic methods, is sometimes referred to as “inverting the lexicon.” Holistic methods described in the literature have used a variety of holistic features, representations, and matching methodologies.

Fig. 10 can be used to illustrate this point. Let us assume that our holistic features are {length, number of ascenders, number of descenders}. The lexicon can be inverted as follows:

```

MAIN [4 0 0]
GREENFIELD [10 3 1]
MASSACHUSETTS [13 3 0]
BUFFALO [7 3 2]
    
```

Note that the word “BUFFALO,” written in uppercase, does not lend itself to holistic features. Assuming that we can derive the features from the shape of the words, the task of recognition, given the inverted lexicon, is quite trivial.

### 3.4 Remarks

Model-based approaches transform the problem of modeling the variability in the signal at the word-level, to modeling it at the level of submodels. The choice of submodels is critical, because of the assertion that these are identical irrespective of where they occur in words. The choice of characters as submodels may be the most obvious, but it is not the only one. It is difficult to capture all of the variability of handwritten words in terms of 26 submodels, especially when the shape of a character is a function of its neighbors (coarticulation effects).

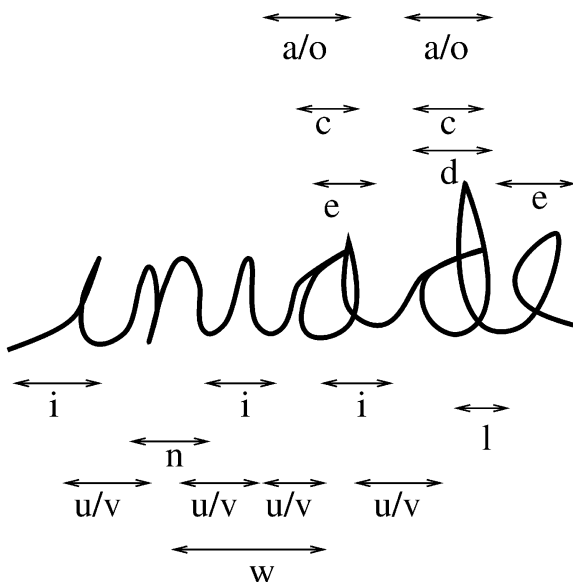


Fig. 14. Looking for instances of each character in a lexicon word (INVADE in this example) in the word image.

Clearly, if the number of classes were finite and large amounts of training data were available, building a separate model for each word directly from the features would yield the best classification, since it would not be constrained by submodels. In practice, these conditions are met only when the lexicon is small and static.

So far, we have skirted the issue of how features differ from submodels. At the lower levels of the hierarchy are simple structures such as edges and other features of the pattern which are grouped into increasingly complex visual entities which approach characters in visual complexity at higher levels of the hierarchy. A variety of these visual entities, both ones that appear in the human visual system as well as arbitrary others, of increasing visual complexity may be imagined that span the spectrum from image pixels at one end to the word identity at the other.

We draw a line somewhere in the middle of this spectrum and call the visual entities preceding it "holistic features" and the ones following it "subword models." This distinction is necessarily a subjective one and is difficult to make unambiguously in some cases. Examples of word features are pixel density distributions, vertical and horizontal strokes, and perceptual features. Examples of subword models are the letters themselves, *graphemes* and other pseudoletter segments that result from an image-based segmentation scheme, and *n-grams* (letter combinations).

Although holistic and analytical approaches are commonly distinguished by the observation that the latter are "segmentation-based," the fact is that holistic and analytical paradigms comprise a continuum of approaches to word recognition. As noted by Casey and Lecolinet [7] in their survey of cursive word recognition, some form of segmentation is involved in all pattern recognition methods, which for holistic methods is their feature extraction phase.

The main difference lies in the level of abstraction of the segmented elements: features (that is to say low-level elements) in the case of holistic methods, versus pseudoletters in the case of analytical methods.

We have described the distinction between static and dynamic lexicon scenarios in Section 1.2, Figs. 3, and 4. Several examples of holistic recognizers that work with dynamic lexicons have been developed for use in the postal handwritten address interpretation context and described in the literature [11], [22], [23], [34]. Most systems used in the check recognition application [28] work with static lexicons.

## 4 OVERVIEW OF HOLISTIC METHODS

The potential role of a holistic approach in a particular application scenario is inevitably linked to the size and static/dynamic nature of the lexicon. Examples of systems for classification, lexicon reduction, and verification are presented in this section.

Holistic methods can be categorized as follows:

1. *Domain*: online and offline.
2. *Lexicon use*: static and dynamic with applications in lexicon reduction and verification.
3. *Features*: low-level, intermediate-level, and high-level.

4. *Feature representation*: vectors, assertions, sequences, and graphs.
5. *Hybrid Methods*: Methods that explicitly use a combination of methodologies.

We will survey several methods that qualify as holistic methodologies by our definition (Section 3) and group them in the above categories. Clearly, methods will belong to several categories among the items listed above. We have attempted to highlight certain aspects of these methods under the various categories.

### 4.1 Domain

The earliest applications of the holistic paradigm were developed in the 1960s and 1970s for online HWR where the word was written on an electronic tablet, or on a screen with a lightpen. In this section, we use these efforts as starting points to cover the landscape of holistic approaches to offline HWR and we will refer to them throughout this section.

The whole word matching approach seems to have been first explored by Frishkopf and Harmon at Bell Labs in 1961 [15]. Words written on an electronic tablet are represented by the sequence of local  $x$  and  $y$  extrema along the trace of the word and compared with similarly encoded lexicon entries by looking for contiguous subsequences of similar extremes.

Around the same time, Earnest [12] at the Mitre Corporation designed a lexicon filter which used just the counts of ascenders and descenders and the presence or absence of a t-bar (i.e., global features) to identify similar words in a 10K lexicon.

Frag [14] represented the entire trace of the word by its chain coded representation. The approach involved extracting an 8-directional code sequence from the cursive input and using a first or second order Markov chain for recognition from a small, static lexicon of words.

Brown and Ganapathy [5] developed a system wherein a fixed 2D grid is imposed on the word image and the number of features (cusps, extensions, etc.) in each grid element counted. The resulting feature vector was compared with stored exemplars obtained from training using a Euclidean metric and  $k$ -NN is used to determine the final class.

The method presented by Miller [38] involved segmenting the online trace of a cursive word and classifying each of the stroke segments as one of a set of segments (codebook), obtained from unsupervised clustering of training words. An angular metric was used to rank a small static lexicon of three-character assembly language opcodes.

More recently, a highly accurate word recognition system that uses noncharacter specific features has been described [10]. The authors talk of the intent of the writer as one of conveying a complete message and not necessarily being careful about the individual characters. They describe the use of features such as cusps, crosses, dots, and breaks.

#### 4.1.1 Remarks

Whatever may be the arguments put forth by developers of purely analytical word recognizers [13], holistic methods have been implemented and successfully used in practical systems [10].

## 4.2 Lexicon Use

The work of Bertille et al. [31] for recognition of unconstrained offline words in a small *dynamic* lexicon context is inspired by the earlier work of Salome et al. [32] in recognizing check amounts with a small static lexicon. Following segmentation of the word, loops, ascenders, and descenders are extracted and quantified into a small number of levels. Different symbols are assigned to each possible combination of loops and extensions that may be discovered within a segment, to yield a set of 27 symbols. A symbol descriptor for the word is thus obtained. Letters and common pseudoletters produced by segmentation (such as the first half of “n”) are represented by a total of 65 three-state models. Given a training word and its corresponding descriptor, a word model for the word is constructed by concatenating letter and pseudoletter models and trained on a set of training strings (descriptors). The system achieves top choice accuracies of 91.2 percent, 77.6 percent, and 42.5 percent with dynamic lexicons of size 10, 100, and 1,000, respectively.

Holistic methods have been used for reduction of large lexicons, mainly in the online handwriting and printed domains. Perceptual features, especially, have been used internally by many analytical classifiers to rapidly discard dissimilar lexicon entries.

Verification of handwritten phrases may be thought of as the task of verifying that a given image of a word or phrase is that of a given ASCII string (or one of a given set of ASCII strings), frequently the result of another recognition algorithm.

The online system of Bramall and Higgins [1] is one of many that employ global features implicitly for lexicon reduction. The “candidate word hypothesization” phase of the system involves the use of features such as length, counts of features, and relative positions of ascenders and descenders to reduce a large static lexicon of 20,000 words to 184 on the average with 92 percent accuracy.

Global features of the word shape such as length and the presence of ascenders and descenders are useful for detecting unlikely matches in the lexicon, either explicitly as a lexicon filter, or implicitly as part of a word classifier. Earnest’s lexicon filter for online script [12] described earlier, for example, uses just the counts of ascenders and descenders and the presence or absence of a t-bar to identify similar words in a 10K lexicon.

### 4.2.1 Remarks

For a holistic classifier, lexicons and training are tightly interwoven, since the lexicon entries are exactly the classes to be distinguished. Most of the literature deals only with small, fixed lexicons; in these cases, enough samples of each class are available to train the classifier in the conventional sense [5], [14]. This is clearly impractical in the case of large or dynamic lexicons.

In the latter case, it becomes necessary to obtain feature vectors for the lexicon entries via other means. In the machine print domain, it is possible to synthetically generate training samples of various fonts and sizes and even model forms of distortion [24]. Unfortunately, this method cannot be applied to unconstrained handwriting,

owing primarily to the wider variety of handwriting styles and defects (breaks, fragmented strokes, skew, slant, open, and filled up loops) which are beyond the scope of existing models of handwriting. When the style of handwriting is constrained (to be online cursive, for example) and the features extracted are coarse, it may be possible to define production rules to determine whether an image descriptor derived from the image can be generated from a given lexicon entry [16]. For unconstrained handwriting, coarse holistic features such as ascenders, descenders, and length of a lexicon entry can be predicted from the features of the constituent characters using heuristic rules [33]. In these cases, training is in the form of heuristics or production rules being used to synthesize feature vectors corresponding to lexicon entries.

## 4.3 Features

There has been extensive research in the design of features for the recognition of isolated characters, which may be in theory applied to the recognition of entire words. Pixel-based features such as template correlation, transformations, and series expansions; features based on distribution of pixels derived from zoning, moments, n-tuples, characteristic loci, crossings, and distances; and low-level geometrical and topological features, such as strokes and curves in various directions, end points and intersections, and properties of the contour have been studied and extensively reviewed [47].

Although easy to extract and fairly insensitive to noise, features based on pixels or their distributions tend to be dependent upon position alignment and highly sensitive to distortion and style variations.

The last category of geometrical and topological features is by far the most popular for isolated handprinted characters, owing to their higher tolerance for distortion and stylistic variations and certain affine transformations. They form the lower tiers of a continuum of *structural* features (so named because they describe the characteristic geometry and topology of the word) that have been used for holistic recognition of words.

Fig. 12 can be referred to once again to illustrate the gradations of features from the fine (low-level) to the coarse (high-level).

### 4.3.1 Low-Level

Highly local, low-level structural features such as stroke direction distributions [45] have been applied successfully for holistic recognition of machine printed words. Hull et al. [54] experimented with both stroke direction distributions as well as *local shape templates* detected by convolution and thresholding [26]. In fact, they were found to perform better than either pixel-based features or higher level structural features such as perceptual features, whose detection is often unreliable [24]. Structural features at this level, however, are generally unsuitable for offline HWR, on account of wide variation in style.

Farag’s method for online HWR [14] may also be thought of as using low-level features since the entire trace of the word was represented as an 8-directional chain code.

### 4.3.2 Intermediate-Level

Structural features at the intermediate-level include edges, end-points, concavities, diagonal and horizontal strokes, and exhibit a greater abstraction from the image (pixel or trace) level. The cusps and extensions extracted by the method of Brown and Ganapathy [5] and the local extrema extracted by Frishkopf and Harmon [15] may also be classified as intermediate-level structural features.

Dzuba et al. [10] describe a holistic word recognizer that works with a whole word or a phrase. They use features that reflect the importance of vertical extremas.

Guillevic and Suen [18] describe a feature-based holistic method for check recognition. For a training word of length  $n$ , a grid with  $n$  equal columns is used to capture ascender, descender and midzone loop positions (extracted from the contour) in the form of an  $n$ -bit vector. Strokes in the vertical, horizontal, and diagonal directions are extracted using morphological operators. The final feature vector for each training word is the concatenation of these binary vectors, along with counts of ascenders, descenders, loops, and length measured as the number of center-line crossings. Given a test word, the features are extracted using positions relative to the horizontal extent of the image. The authors report a top choice accuracy of 72 percent with a static check amount lexicon of 32 words.

Olivier et al. [42] take a structural description approach for holistic recognition of words from a static check amount lexicon. The center line intersects the thinned representation of the word at anchor points and divides it into structural primitives such as *upper loop* and *lower connection* (eight in all). The authors refer to these primitives as "strokes." Strokes sharing an anchor point taken together constitute a "grapheme." A set of 42 graphemes is obtained from all the graphemes found in a training set of words by an unsupervised clustering procedure. An image may now be represented as a sequence of either strokes or graphemes. The top choice of the stroke-based classifier, grapheme-based classifier, and their combination is reported to be 34 percent, 70 percent, and 72 percent, respectively.

### 4.3.3 High-Level

Perceptual features such as ascenders, descenders, loops, and length are easily perceived by the human eye, and we have reviewed evidence for their use in human reading. They are by far the most popular for holistic recognition of handwritten words.

Ascenders and descenders, while of uniform height and relatively easy to detect in machine print, are heavily subject to vagaries of style in handwriting, making their accurate detection a challenge. In theory, ascenders and descenders may be extracted by looking for parts of the word in the upper and lower zones, respectively. This, in turn, entails accurate reference line determination, which often fails in the presence of large skew, uneven writing, curved baseline, and for "top-heavy" images (e.g., "Falls"). Ascenders and descenders may also be detected directly from a run-length or contour representation.

Dots and holes may be computed by connected component analysis or alternatively by chain code analysis.

Some features such as diagonal strokes and arcs may be easier to extract from the skeletonized image [24].

Word length is a particularly important perceptual feature [30] and may be estimated in the online case from the number of times the script traverses the "center line" as the ratio of this number to a statistic representing the number of traverses of the center line per letter of the average English word [4]. This method extends itself readily to offline script, but, in practice, the accuracy of the estimate is not satisfactory. Of course, the number of center-line crossings may be used in its raw form as a measure of length and compared with the estimated number of crossings for a given lexicon word. Other notions of length include the number of lower contour minima, the number of vertical strokes, and the number of possible segmentation points (ligatures and breaks).

Earnest's lexicon filter [12] which used counts of ascenders and descenders and the presence or absence of a t-bar is an early example of the use of perceptual features in a holistic HWR method. Simon and Baret [49] describe an approach to cursive script recognition that involves decomposing a cursive word into a pseudoperiodic signal (regular features) modulated by nonperiodic signals (irregular features). The irregularities are in essence perceptual features.

O'Hair and Kabrisky [41] describe the use of two-dimensional low frequency Fourier coefficients as features for holistic recognition of printed text. The low frequency components contain enough general information to uniquely identify the word from a fixed lexicon of possible words, but not the specific details of font and style. The latter are encapsulated by the high frequency components, which are ignored in the match. It is not clear that this approach will succeed with handwriting, given the large scale variations in writing style.

Miller's approach [38] of segmenting the online trace of a cursive word and classifying each of the stroke segments using a codebook is an example of an approach that uses a segmentation scheme in conjunction with a simple set of segment categories. These models are often derived from or are compositions of medium or high-level structural features. These methods may be classified as being holistic or analytical depending on the subjective decision as to whether the segments are complex enough to be called models.

### 4.3.4 Remarks

To summarize, the features best suited for holistic recognition of handwriting, as apparent from these studies, are higher level structural features, such as edges and end points, and perceptual features, such as dots, holes, ascenders, descenders, and t-bars. A particularly important perceptual feature is word length and many measures of length may be envisaged. The algorithmic accuracy of detection of perceptual features depends on the style and neatness of writing.

## 4.4 Feature Representation

The scheme used to represent holistic features is clearly a function of the features themselves and whether they are

low-level, medium-level, or high-level. Here, we review the predominant representation schemes.

#### 4.4.1 Feature-Vectors and Matrices

Feature-vector representations are commonly used to represent low-level or intermediate-level features. The image is divided into sections using a fixed or variable grid, features are extracted from the sections, and the counts of different features in different sections are represented as a Boolean, integer or real-valued vector, or a matrix. Representing high-level features as a feature vector is less common.

The sectioning scheme is a form of implicit segmentation of the image or its representation and is often as important as the features themselves. The feature extraction of Brown and Ganapathy [5] divided the image into  $n$  equal sections and resulted in a 138-dimensional feature vector. Hull et al. [54] used a variable grid based on reference lines detected in the image.

Pixel-level features are uniquely identified by the  $x$ - $y$  coordinates of the pixel. Low-level structural features are typically extracted by superimposing a rectangular grid on the image. This grid may either be fixed [5], or it may be variable (image-dependent) [24], [33].

For higher level structural features such as edges, end points, and perceptual features, the presence or absence of each feature is important and, consequently, the representation should allow matching of corresponding features in *nearby* cells as well. These features are generally represented more robustly by a graph or a string of codes, each code referring to a different feature or combination of features, although there are instances in the literature of binary feature vectors being used to represent the presence of higher level structural features in different sections of the word [18], [50].

#### 4.4.2 Counts and Assertions

These are the simplest representation of high-level features. For example, Earnest's lexicon filter [12] used just the counts of ascenders and descenders and the presence or absence of a t-bar. Such simple features (sometimes called "global features") are often used to discard dissimilar word candidates from the lexicon.

#### 4.4.3 Sequences

The word is represented as a sequence of symbols representing a set of structural primitives, which correspond to intermediate or high-level features or combinations of such features. This constitutes an implicit segmentation of the image into the structural primitives. Some hybrid methods explicitly segment the image and extract holistic features from the segments.

A *location coded string representation* tags each code with the "positions" in which it occurs, again with reference to a fixed or variable global reference frame. For instance, "O:256" may indicate that there are three holes in the image, located at the second, fifth, and sixth positions along the length of the image [24].

Since a word is approximately a one-dimensional signal that flows from left to right, a sequential representation of such codes may suffice as a description of shape.

Accordingly, a *symbol string representation* denotes the image as a sequence of codes. Adjacency and relative locations of structural features of different types are readily captured by these descriptors [16]. Features that may be located above or below others are better described in separate strings.

Moreau [40] extracted vertical and horizontal strokes, loops, i-dots, and t-bars from the offline image to obtain a string descriptor and compared the descriptor with unique prototypes of words found in French check amounts and their more common orthographic deviations. The unique prototype for each class was obtained as the mode of the descriptors obtained from training samples of that class.

Salome et al. [32] extracted ascenders, descenders, loops, i-dots, and unattached t-bars from the contours of connected components and obtained a string descriptor. Word length was estimated as the number of letter segments obtained as a by-product of a separate analytical subsystem. The Levenshtein metric was used to compare the test string with reference strings obtained from training corresponding to a small lexicon of check amounts.

#### 4.4.4 Graph Structures

The whole image may be represented by a graph with features as nodes and spatial relationships between them as the edges [17]. Graph representations are powerful in that they can represent both positions of features as well as relationships between them. Fig. 15 is an illustration of such a graph structure.

Paquet and Lecourtier [43] describe a check amount recognition method where the intersections of the middle line with the word ("guiding points") are first determined. Stroke following is initiated at each guiding point and each point is coded by features of the primitive stroke segments starting and ending at the point. The "graph" (essentially the thinned image) obtained from stroke tracing is analyzed into seven types of primitives (upper strokes, lower strokes, upper connection, etc.) and a symbol string describing the structure of the graph is achieved. The test string is compared with empirically obtained reference strings using the Levenshtein metric.

Camillerapp et al. [6] labeled singular vertices (end-points, crossings, and points of local curvature) in the skeletonized gray-level image and obtained a tree of stroke primitives. Each tree node was described by the type of primitive, vertical word zone position, and its relative horizontal position within the word. Each lexicon word was coded as a similar tree of primitives, except that each node could describe a set of primitives covering variations that may be expected at that point.

#### 4.4.5 Remarks

The choice of a representation scheme depends on the implementation constraints and on the eventual matching strategy used. The types of features seem to be common in that they are not specific to individual characters. Statistical classification techniques use feature vectors, heuristic matching techniques predominantly use counts and assertions, symbolic matching methods primarily use lexicon coded strings, and graph representation methods naturally favor graph matching algorithms.

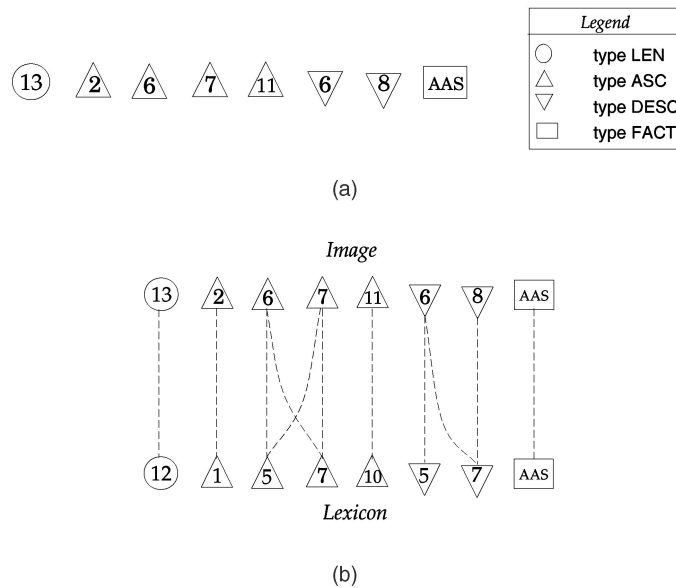


Fig. 15. (a) Wordgraph for the image of "Buffalo." (b) Possible node associations between the image and the lexicon entry *Buffalo*. A possible node association is denoted by a dotted arc between a node of the first graph and one of the second.

#### 4.5 Hybrid Methods

Some methods adopt both analytical and holistic features. The commercial French check processing system of Simon et al. [50] uses seven operators to estimate the probability of each of the 25 lexicon classes. The first of these is based on comparing a structural description known as the "holography" of the test word with prototypes of each class and is based on Gorsky's earlier work [17]. The other six operators are based on different perceptual features:

1. length (number of segments obtained from the analytical method),
2. positions of ascenders and descenders,
3. positions of sets of overlapping horizontal segments,
4. mid-zone loops,
5. dots above the half-line, and
6. mid-zone crosses.

Dodel and Shinghal [9] describe a hybrid analytical-holistic method for offline words to identify the correct class from a static lexicon of 31 words. Aspect ratio (horizontal extent/midzone width) and relative positions of ascenders and descenders are used to achieve direct recognition of some words such as "eight" and partial recognition of others.

Hull et al. [54] estimate length of printed words from character segmentation and word case from reference lines and use these global features to filter the lexicon. Their "segmentation-based" approach is actually holistic since it involves segmenting the word image into characters and concatenating the pixels corresponding to each segmented character (normalized to a  $24 \times 24$  grid) into a  $24 \times 24 \times N$  vector, where  $N$  is the length of the word. They use the Baird templates [26] and stroke direction distribution features.

##### 4.5.1 Remarks

It is natural for word recognition engines to consider a hybrid of recognizers for best performance. Analytical and holistic methods can complement each other's strengths and provide for a robust system.

## 5 SUMMARY

The Holistic paradigm in handwritten word recognition is one that treats the word as a single, indivisible entity and attempts to recognize it using features of the word as whole, and is inspired by psychological studies of human reading, which indicate that humans use features of word shape such as *length*, *ascenders*, and *descenders* (see Fig. 5) in reading.

Holistic approaches circumvent the issues of segmentation ambiguity and character shape variability that are primary concerns for analytical approaches and may succeed on poorly written words where analytical methods fail to identify character content. However, their treatment of lexicon words as distinct pattern classes has traditionally limited their application to recognition scenarios involving small, static lexicons.

Given the difficulty of the task of reading handwriting, practical recognition engines must use multiple classification algorithms and complex strategies for combining classifier decisions and thresholding based on classification confidences for rejection of classification errors. In this survey, we have attempted to take a fresh look at the potential role of the Holistic paradigm in handwritten word recognition, in the light of this observation.

The Holistic paradigm draws inspiration from studies of individuals with acquired dyslexia, studies of reading development, and studies involving proofreading tasks which provide evidence for the existence of a parallel holistic reading process in both developing and skilled

readers; however, there appears to be no consensus on how word shape information is combined with letter identities.

We have attempted to characterize approaches to recognition as a continuous spectrum based on the visual complexity of the unit of recognition employed. Holistic features may be distinguished from subword models in the visual processing hierarchy by their relatively lower visual complexity; however, this distinction is subjective. A holistic approach may be defined as one which does not search for subword models. Analytical approaches are more accurately called model-based approaches.

Holistic systems generally adopt either a *feature-extraction* or a *structural description* approach to the problem of representing word shape. The features themselves may be classified broadly as being pixel-based or structural. Higher level structural features appear to be best-suited for holistic recognition of handwriting and are represented as feature vectors, location-coded and symbol strings, and graphs, to name a few common ones. The matching methodology adopted is related closely to the representation of features.

Holistic recognition systems are characterized by an integration of training and the lexicon, whose presence is often an implicit assumption in the design of holistic word recognition algorithms. Most implementations of holistic approaches in the offline HWR domain have been used for the classification of small, static lexicons. Lexicon reduction and verification of recognition results have recently emerged as other applications of the holistic paradigm.

Given the evidence from reading studies, the intrinsic advantage of computational economy, and orthogonality with respect to analytical approaches, we believe that the holistic paradigm holds immense promise for realizing near-human performance.

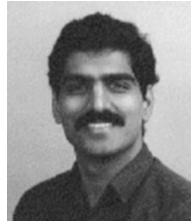
## REFERENCES

- [1] P.E. Bramall and C.A. Higgins, "A Blackboard Approach to On-Line Cursive Handwriting Recognition for Pen-Based Computing," *Proc. Third Int'l Workshop Frontiers in Handwriting Recognition (IWFHR-III)*, pp. 295-304, 1993.
- [2] D.E. Broadbent and M.H.P. Broadbent, "General Shape and Local Detail in Word Perception," *Attention and Performance VI*, S. Dornic, ed., Erlbaum, 1977.
- [3] D.E. Broadbent and M.H.P. Broadbent, "Priming and the Passive/Active Model of Word Recognition," *Attention and Performance VIII*, R.S. Nickerson, ed., Lawrence Erlbaum, 1980.
- [4] E.R. Brocklehurst and P.D. Kenward, "Preprocessing for Cursive Script Recognition," *NPL Report*, Nov. 1988.
- [5] M.K. Brown and S. Ganapathy, "Cursive Script Recognition," *Proc. Int'l Conf. Cybernetics and Soc.*, pp. 47-51, Oct. 1980.
- [6] J. Camillerapp, G. Lorette, G. Menier, H. Oulhadj, and J.C. Pettier, "Off-Line and On-Line Methods for Cursive Handwriting Recognition," *From Pixels to Features III: Frontiers in Handwriting Recognition*, S. Impedovo and J.C. Simon, eds., pp. 273-288, North-Holland, 1992.
- [7] R.G. Casey and E. Lecolinet, "A Survey of Methods and Strategies in Character Segmentation," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, no. 7, pp. 690-706, July 1996.
- [8] J.M. Cattell, "The Time Taken Up by Cerebral Operations," *Mind*, vol. 11, pp. 220-242, 1886.
- [9] J.-P. Dodel and R. Shinghal, "Symbolic/Neural Recognition of Cursive Amounts on Bank Cheques," *Proc. Third Int'l Conf. Document Analysis and Recognition (ICDAR)*, pp. 15-18, Aug. 1995.
- [10] G. Dzuba, A. Filatov, D. Gershuny, and I. Kil, "Handwritten Word Recognition—The Approach Proved by Practice," *Proc. Fifth Int'l Workshop Frontiers in Handwriting Recognition*, pp. 99-112, 1998.
- [11] A.C. Downton, E. Kabir, and R. Birch, "Recognition and Verification of Postcodes in Handwritten and Hand-Printed Addresses," *Proc. 10th Int'l Conf. Pattern Recognition*, pp. 469-473, 1990.
- [12] L.D. Earnest, "Machine Recognition of Cursive Writing," *Proc. IFIP Congress*, pp. 462-466, 1962.
- [13] S. Edelman, T. Flash, and S. Ullman, "Reading Cursive Handwriting by Alignment of Letter Prototypes," *Int'l J. Computer Vision*, vol. 5, no. 3, pp. 303-331, 1990.
- [14] R. Farag, "Word Level Recognition of Cursive Script," *IEEE Trans. Computers*, vol. 28, no. 2, Feb. 1979.
- [15] L.S. Frishkopf and L.D. Harmon, "Machine Reading of Cursive Script," *Proc. Fourth London Symp. Information Theory*, 1961.
- [16] N. Nasrabadi, G. Seni, and R. Srihari, "An Online Cursive Word Recognition System," *Proc. IEEE Computer Vision and Pattern Recognition*, June 1994.
- [17] N.D. Gorsky, "Off-Line Recognition of Bad Quality Handwritten Words Using Prototypes," *Fundamentals of Handwriting Recognition*, S. Impedovo, ed., pp. 199-217, Springer-Verlag, 1993.
- [18] D. Guillevic and C.Y. Suen, "Cursive Script Recognition Applied to the Processing of Bank Cheques," *Proc. Third Int'l Conf. Document Analysis and Recognition (ICDAR)*, pp. 11-14, Aug. 1995.
- [19] R.N. Haber and R.M. Schindler, "Errors in Proofreading: Evidence for Syntactic Control of Letter Processing?" *J. Experimental Psychology: Human Perception and Performance*, vol. 7, pp. 573-579, 1981.
- [20] L. Harmon, "Automatic Recognition of Print and Script," *Proc. IEEE*, vol. 60, no. 10, pp. 1165-1176, 1972.
- [21] L. Henderson, *Orthography and Word Recognition in Reading*. Academic Press, 1982.
- [22] A. Hendrawan and A.C. Downton, "Verification of Handwritten British Postcodes Using Address Features," *Fundamentals of Handwriting Recognition*, S. Impedovo, ed., pp. 313-317, Springer-Verlag, 1993.
- [23] A. Hendrawan, A.C. Downton, and C.G. Leedham, "A Fuzzy Approach to Handwritten Address Verification," *Proc. Third Int'l Workshop Frontiers in Handwriting Recognition (IWFHR-III)*, pp. 207-216, 1993.
- [24] T.K. Ho, "A Theory of Multiple Classifier Systems and Its Application to Visual Word Recognition," PhD thesis, State Univ. of New York at Buffalo, 1992.
- [25] D. Howard, "Reading Without Letters?" *The Cognitive Neuropsychology of Language*, M. Coltheart, G. Sartori, and R. Job, eds., Lawrence Erlbaum, 1987.
- [26] L.D. Jackel, H.S. Baird, H.P. Graf, and W.E. Hubbard, *A VLSI Architecture for Binary Image Classification*, pp. 275-186, North Holland, 1989.
- [27] G.W. Humphreys, L.J. Evett, and P.T. Quinlan, "Orthographic Processing in Visual Word Recognition," *Cognitive Psychology*, vol. 22, pp. 517-560, 1990.
- [28] G. Kim and V. Govindaraju, "Bank Check Recognition Using Cross Validation between Legal and Courtesy Amounts," *Int'l J. Pattern Recognition and Artificial Intelligence*, vol. 11, no. 4, pp. 657-674, 1997.
- [29] N. Lindgren, "Cursive Script Recognition," *IEEE Spectrum*, May 1965.
- [30] R. N. Haber, L.R. Haber, and K.R. Furlin, "Word Length and Word Shape as Sources of Information in Reading," *Reading Research Quarterly*, vol. 18, pp. 165-189, 1983.
- [31] J. Bertille, M. Gilloux, and M. Leroux, "Recognition of Handwritten Words in a Limited Dynamic Vocabulary," *Proc. Third Int'l Workshop Frontiers in Handwriting Recognition*, pp. 417-422, 1993.
- [32] J.C. Salome, M. Leroux, and J. Badard, "Recognition of Cursive Script Words in a Small Lexicon," *Proc. First Int'l Conf. Document Analysis and Recognition (ICDAR '91)*, pp. 774-782, Sept./Oct. 1991.
- [33] S. Madhvanath and V. Govindaraju, "Holistic Lexicon Reduction," *Proc. Third Int'l Workshop Frontiers in Handwriting Recognition (IWFHR III)*, pp. 71-81, 1993.
- [34] S. Madhvanath and V. Govindaraju, "Holistic Lexicon Reduction for Handwritten Word Recognition," *Proc. Eighth SPIE Int'l Symp. Electronic Imaging: Science and Technology*, Jan./Feb. 1996.
- [35] S. Madhvanath, G. Kim, and V. Govindaraju, "Chaincode Processing for Handwritten Word Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 21, no. 9, Sept. 1999.
- [36] S. Madhvanath, E. Kleinberg, and V. Govindaraju, "Empirical Design of a Multiclassifier Thresholding/Control Strategy for Recognition of Handwritten Street Names," *Int'l J. Pattern Recognition and Artificial Intelligence*, vol. 11, no. 6, pp. 933-946, 1997.

- [37] J.L. McClelland, "Preliminary Letter Identification in the Presentation of Words and Nonwords," *J. Experimental Psychology: Human Perception and Performance*, vol. 2, pp. 80-91, 1976.
- [38] G.M. Miller, "Real-Time Classification of Handwritten Script Words," *Proc. IFIP Congress*, pp. 218-223, Aug. 1971.
- [39] A.F. Monk and C. Hulme, "Errors in Proofreading: Evidence for The Use of Word Shape in Word Recognition," *Memory and Cognition*, vol. 11, pp. 16-23, 1983.
- [40] J. Moreau, "A New System for Automatic Reading of Postal Checks," *Int'l Workshop Frontiers in Handwriting Recognition (IWFHR-2)*, pp. 121-132, Sept. 1991.
- [41] M.A. O'Hair and M. Kabrisky, "Recognizing Whole Words as Single Symbols," *Proc. First Int'l Conf. Document Analysis and Recognition (ICDAR '91)*, pp. 350-358, Sept./Oct. 1991.
- [42] C. Olivier, T. Paquet, M. Avila, and Y. Lecourtier, "Recognition of Handwritten Words Using Stochastic Models," *Proc. Third Int'l Conf. Document Analysis and Recognition (ICDAR)*, pp. 19-23, Aug. 1995.
- [43] T. Paquet and Y. Lecourtier, "Handwriting Recognition: Application to Bank Cheques," *Proc. First Int'l Conf. Document Analysis and Recognition (ICDAR '91)*, pp. 749-757, Sept./Oct. 1991.
- [44] G.M. Reicher, "Perceptual Recognition as a Function of Meaningfulness of Stimulus Materials," *J. Experimental Psychology*, vol. 81, pp. 274-280, 1969.
- [45] K. Yamamoto, S. Mori, and M. Yasuda, "Research on Machine Recognition of Handprinted Characters," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 6, no. 4, pp. 386-405, July 1984.
- [46] K.M. Sayre, "Machine Recognition of Handwritten Words: A Project Report," *Pattern Recognition*, vol. 5, pp. 213-228, 1973.
- [47] L. Schomaker and E. Segers, *Advances in Handwriting Recognition*, S.-W. Lee, ed., vol. 34, World Scientific, 1999.
- [48] P.H.K. Seymour, "Developmental Dyslexia," *Cognitive Psychology: An International Review*, M.W. Eysenck, ed., John Wiley and Sons, 1990.
- [49] J. Simon and O. Baret, "Formes Regulieres et Singulares; Application a la Reconnaissance de L'écriture Manuscrite," *C.R. Acad. Scr. Paris*, vol. 309, pp. 1901-1906, 1989.
- [50] J.C. Simon, O. Baret, and N.D. Gorski, "A System for the Recognition of Handwritten Literal Amounts of Checks," *Proc. IAPR Workshop Document Analysis Systems*, A.L. Spitz and A. Dengel, eds., pp. 265-287, 1994.
- [51] S.J. Soltysiak, "Visual Information in Word Recognition: Word Shape or Letter Identities?" *Proc. Workshop Integration of Natural Language and Vision Processing*, 1994.
- [52] C.C. Tappert, C. Suen, and T. Wakahara, "The State of the Art in On-Line Handwriting Recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 12, no. 8, Aug. 1990.
- [53] J. Theios and J.G. Muise, *The Word Identification Process in Reading*, vol. 2. Hillsdale, N.J.: John Wiley & Sons, 1977.
- [54] J.J. Hull, T.K. Ho, and S.N. Srihari, "Word Recognition with Multilevel Contextual Knowledge," *Proc. First Int'l Conf. Document Analysis and Recognition (ICDAR '91)*, pp. 905-915, Sept./Oct. 1991.
- [55] G. Underwood and K. Bargh, "Word Shape, Orthographic Regularity, and Contextual Interactions in a Reading Task," *Cognition*, vol. 12, pp. 197-209, 1982.
- [56] D.D. Wheeler, "Word Recognition Processes," *Cognitive Psychology*, vol. 1, pp. 59-85, 1970.



**Sriganesh Madhvanath** received the BTech degree in computer science from the Indian Institute of Technology, Bombay, in 1990 and the MS and PhD degrees from the State University of New York at Buffalo, in 1993 and 1997, respectively. He worked as a research assistant at the Center for Excellence in Document Analysis and Recognition (CEDAR) from 1991 to 1996. He then joined the Image and Multimedia Systems Group at the IBM Almaden Research Center, San Jose, California as a member of the research staff. His research interests include classifier decision combination, genetic algorithms, and document analysis and recognition. He is a member of the IEEE.



**Venu Govindaraju** received his PhD in computer science from the State University of New York at Buffalo in 1992. He has coauthored more than 100 technical papers in various international journals and conferences and has one US patent. He is currently the associate director of CEDAR as well as an associate professor in the Department of Computer Science and Engineering at the University of New York. He is the associate editor of the *Journal of Pattern Recognition* and the area chair of the IEEE SMC technical committee for pattern recognition. Dr. Govindaraju has been a coprincipal investigator on several federally sponsored and industry sponsored projects. He is presently leading multiple projects on postal applications. He is the program cochair of the upcoming International Workshop on Frontiers in Handwriting Recognition in 2002. He is a senior member of the IEEE.