

A knowledge-based approach for recognition of handwritten Pitman shorthand language strokes

P NAGABHUSHAN and BASAVARAJ S ANAMI[†]

Department of Studies in Computer Science, University of Mysore,
Manasgangotri, Mysore 570 006, India

[†] Present address: Department of Computer Science & Engineering, Basaveshwar
Engineering College, Bagalkot 587 102, India
e-mail: pnagabhushan@hotmail.com

MS received 6 November 2000; revised 20 November 2001

Abstract. The Pitman shorthand language (PSL) is a recording medium practised in all organizations, where English is the transaction medium. It has the practical advantage of high speed of recording, more than 120–200 words per minute, because of which it is universally acknowledged. This recording medium has its continued existence in spite of considerable developments in speech processing systems, which are not universally established yet. In order to exploit the vast transcribing potential of PSL a new area of research on automation of PSL processing is conceived. It has three major steps, namely, shape recognition of PSL strokes, their validation and English text production from these strokes.

The paper describes a knowledge-based approach for the recognition of PSL strokes. Information about location and the direction of the starting point and final point of strokes are considered the knowledge base for recognition of strokes. The work comprises preprocessing, determination of starting and final points, acquisition of quadrant knowledge, graph-based traversal and finally a rule-based inference process for generating phonetic equivalent of English language characters for the strokes. The proposed work is thoroughly tested for a large number of handwritten strokes.

Keywords. Pitman shorthand language; character recognition; English text production; primitives; knowledge base.

1. Introduction

The system of shorthand writing was invented by Sir Issac Pitman (Pitman 1976) and is used to note down dictated/spoken text. It is widely practised in all organizations where English is the transaction medium. Advantages of the practical ability of PSL are universally acknowledged and this recording medium continues to exist in spite of considerable developments in Speech Processing Systems (SPS), which are not universally established yet. Further, PSL could also find global applications such as in the areas of private and secret communication, compact

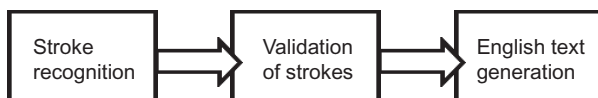


Figure 1. Three phases in automatic text generation from PSL.

storage of documents, speech to text conversion, adaptation of PSL to other languages, e-mail compression etc. The fast developments in both computer and communication technologies promise development of the above global applications. Hence, the problem of automatic generation of complete English text, in the form of a printed document, from spoken/dictated matter is conceived and addressed as a new research area.

Leedham & Downtown (1984, 1986, 1987, 1990) and Hemanth Kumar (1998) have addressed the problem of automatic recognition of PSL strokes and techniques like Hough transformation and regional decomposition have been discussed. Kuroda (1999) has discussed the successor method based on stochastic regular grammar but makes use of the knowledge of starting point in recognition of handwritten Chinese characters. Samouelian *et al* (1994) have presented a knowledge-based approach for recognition of English language consonants. However, no work, based on knowledge, on recognition of PSL strokes/consonants is reported in the literature.

This paper proposes a knowledge-based approach for recognition of PSL strokes that form the main constituents of any PSL document. The work is part of an attempt made to automate the process of text generation from PSL documents, which is viewed as a three-phase pattern recognition problem. The phases are: (i) shape recognition, (ii) stroke validation, and (iii) English text production, which are as shown in figure 1. The last phase consists of the generation of correct and equivalent English text from the PSL document. The efficiency of this phase depends on the accuracy of the shape recognition phase and hence it is considered a challenging task, especially for handwritten strokes. The strokes recognized are converted to their equivalent English characters. The authors have worked on the composition of English text from phonetic text documented through PSL, which is an initial work in this direction and a paper has also been published (Nagabhushan & Anami 1999).

The paper is organised into seven sections. Section 2 deals with PSL/basic stroke set and their description. The preprocessing of strokes and knowledge-based approach are given in § 3 and 4 respectively. Section 5 describes the overall algorithm adopted in this approach. The results are discussed in § 6 and the work is concluded in § 7. The bibliography is given at the end of the paper.

2. PSL character set

A practitioner of PSL is called a stenographer and the character set of PSL consists of 24 consonants, 12 vowels and other basic punctuation symbols. Consonants are the basic strokes in PSL. In general, a stroke in PSL represents a character or a word in English at the simplest level and a phrase or a sentence at the complex level. The stroke recognition phase addresses the recognition of the shapes of strokes. The basic strokes and their equivalent characters in English are given in table 1. A separate group is working on the development of robust techniques for recognition of handwritten PSL stroke shapes (Nagabhushan & Murali 2000). Normally, the stenographer writes the strokes with respect to a base line, as shown in figure 2. The writing of strokes above or on or through this base line depends on the first sounding vowel in the word. The writing of each stroke is practised in a definite manner and hence knowledge of writing in PSL is considered for recognition of these strokes in this approach.

Table 1. PSL strokes and English consonants.

Character	Phonetic name	Stroke nature	English consonant
<i>Explodents</i>			
\	Pee	Thin	P
\	Bee	Thick	B
	Tee	Thin	T
	Dee	Thick	D
/	Chay	Thin	CH
/	Jay	Thick	J
—	Kay	Thin	K
—	Gay	Thick	G
\	Ef	Thin	F
<i>Continuants</i>			
\	Vee	Thick	V
(Ith	Thin	TH
(Thee	Thick	TH
)	Ess	Thin	S
)	Zee	Thick	Z
)	Ish	Thin	SH
)	Zhee	Thick	ZH
<i>Nasals</i>			
∩	Em	Thin	M
∪	En	Thin	N
∪	Ing	Thick	NG
<i>Liquids</i>			
(Ar, ray	Thin	R
) /	El	Thin	L
<i>Coalescents</i>			
✓	Way	Thin	W
✓	Yay	Thin	Y
<i>Aspirates</i>			
∩	Hay	Thin	H
∩	Hay	Thin	H

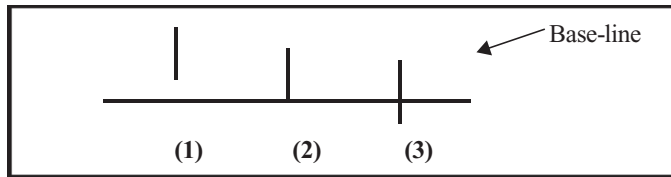


Figure 2. Three positions for strokes.

The vowels are classified into six long vowels as heard in the words like wah!, ale, each, all, oak, ooz etc. and six more short vowels as heard in the words like at, etch, it, odd, pub, cook etc. These vowels have six positions, namely first, second and third on each side of the stroke. The vowels are represented by the symbols, dash (–) and dot (·) and can be thin or thick based on the intensity of sound. The vowels are written before or after the stroke depending upon whether the vowels precede or succeed the consonant sound during pronunciation (Pitman 1976). Thus, automatic English text generation from a PSL document requires efficient recognition of PSL characters/strokes. The work is limited to the recognition of the basic set of consonants, namely, explodents, continuants, nasals, liquids, coalescent and aspirates.

The first eight consonants, represented by straight-line strokes, are called “explodents” because in pronouncing them, outgoing breath is forced in a sudden gust through the barriers that were previously closed. The next eight, represented by upright or sloping curves, are called “continuants”, because in uttering these outgoing breath is allowed to escape in a continuous stream through a similar barrier partially open, instead of being expelled suddenly. Closing successive barriers in the mouth against the outgoing air-stream produces the “nasals”, so that air has to escape through the nose, and these represented by horizontal curves.

The “liquids” flow into union with other consonants and thus make double consonants as in the words cliff, dry etc. where the consonants ‘l’ and ‘r’ blend with the preceding consonants. The “coalescents” precede vowels and coalesce or unite with them. The “aspirate” breathes upon a following vowel. Thus, by breathing upon the vowel ‘a’ in the word “at” the word changes to “hat”. Other than these, PSL has many advanced features normally practised by an expert stenographer that are not considered in this paper. Keeping in mind the manner in which the strokes are written, a knowledge-based approach is developed for the recognition of PSL consonants/strokes (Pitman 1976).

3. Preprocessing of the strokes

The handwritten PSL strokes are scanned and the binary image obtained is used for further recognition. The preprocessing phase comprises thinning and trimming of the stroke images, followed by filling up of discontinuities present in the images (Gose *et al* 1999). This phase is a prerequisite for the traversal of strokes.

(a) *Thinning*: The binary image is converted to a single pixel width image by removing unwanted information based on the 8-neighbours technique, where the neighbouring 8 pixels for a given pixel are checked for the nonzero value and zero to nonzero transitions. The algorithm uses the rule, depicted in figure 3a, to either delete or retain the pixel in question. The number of neighbouring 1’s is counted for the pixel in question and a threshold value

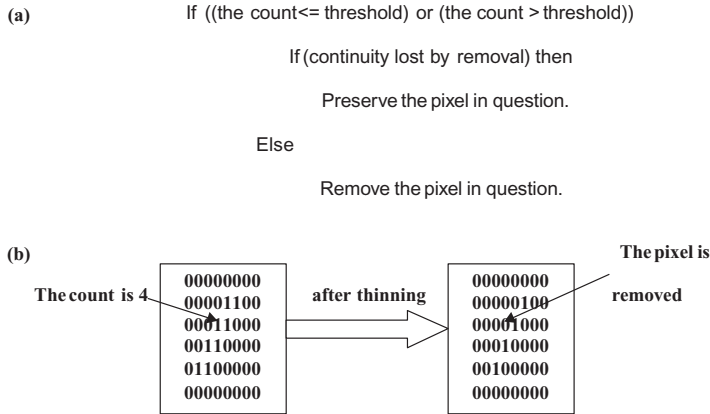


Figure 3. (a) Rule for the retention/elimination of pixels. (b) Illustration of thinning process.

is assumed for either deletion or retention. The “threshold” is defined as the number of neighbouring pixels ON for a given pixel. The threshold value of 3 has given good results and is discussed in section. The thinning algorithm also eliminates isolated pixels, which are pixels that have no neighbours. An example illustrating the thinning process is given in figure 3b.

(b) *Trimming and filling:* The binary image of the given stroke may have a large number of leading and trailing zeros making image handling unwieldy. Trimming is used to remove these leading and trailing zeros without disturbing the image. Thus, the actual image can be used for further recognition. There can be discontinuities of one to two pixel width due to noise in the scanned image. In order to render a continuous image for traversal of a stroke during recognition it is necessary to fill in the discontinuities, if any, in the image. This is again based on the 8-neighbours technique. Figure 4 illustrates the process of trimming and filling.

The trimmed image is either a square or a rectangle. The property of the rectangle is that the intersection of the diagonals passes through the centre of the rectangle and it is used in this approach. The gap in the image expressed in terms of pixels is called the limit for filling the gap.

4. Knowledge based approach

Consonants are the result of audible friction or stopping of breath in some part of the mouth or throat. Simple geometric shapes like straight lines, curves and hooks repre-

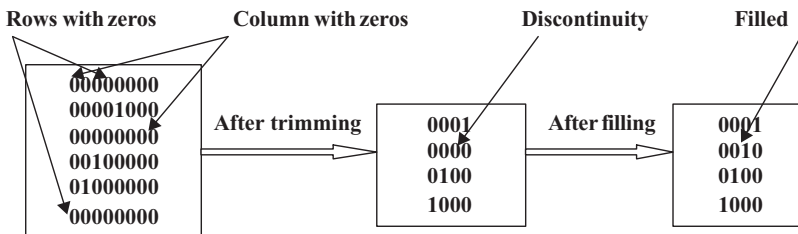


Figure 4. Illustration of trimming and filling process.

sent all consonant sounds. These three shapes form the basic elements and are called strokes/characters in PSL. These consonants are listed in phonographic order in table 1. The strokes are written thin or thick depending upon light and heavy sounds respectively, viz thin vertical stroke represents “T” and the thick vertical stroke represents “D”. Every stroke can represent a simple character or a complex word or a sentence in PSL. Words, phrases and sentences are represented by a concatenation of strokes. However, the strokes are written in a definite manner and have two extreme points called pen-down and pen-up points, which are considered the extreme points of the stroke. The pen-down and pen-up points are also referred to as starting point and final point respectively (Patterson, 1999).

In this approach, the basic strokes are considered for recognition because the concatenated strokes are segmented to the basic strokes for their recognition by detecting the corner points or the point joining the strokes. The knowledge of extreme points is used in this method for stroke recognition because it can be distinguished from other points by the property that it has only one neighbour. A careful study of handwritten PSL documents is carried out and it is observed that the above mentioned points lie in one of the four quadrants, when the stroke is enclosed in a frame and is true for every stroke other than hooks. This information is used as the heuristic knowledge in the recognition process. In the case of hooks, it is observed that at one of the extreme points there is a cluster of points and this knowledge is used for the recognition of hooks.

The extreme points of the strokes are found by traversing the stroke using nearest neighbour technique. The following piece of knowledge is used in determining the neighbouring pixels during stroke traversal.

- Corner pixels have utmost three neighbours.
- Pixels on the frame boundary have utmost five neighbours.
- All other pixels have utmost eight neighbours.

Figure 5 gives a few strokes as illustration. The four quadrants considered become inadequate for recognizing the strokes representing the characters with phonetic names ‘Tee’, ‘Dee’, ‘Kay’, and ‘Gay’. These are the vertical and horizontal strokes coinciding with the lines dividing the frame into four quadrants. Hence, two more regions are defined, one along the X -axis and the other along the Y -axis, in conjunction with the already defined four quad-

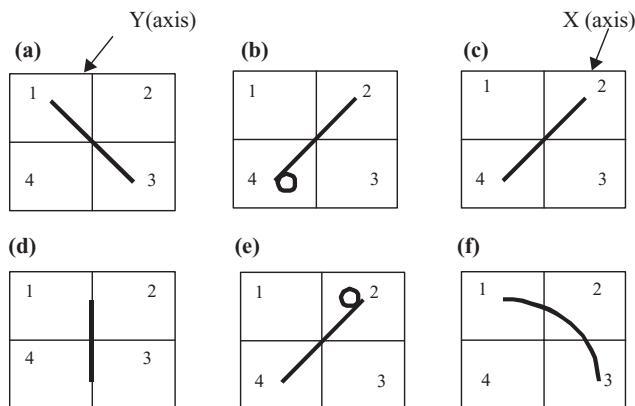


Figure 5. Strokes and their starting points. (a) ‘Bee’, (b) ‘hay’, (c) ‘jay’, (d) ‘tee’, (e) ‘hay’, and (f) ‘ar’.

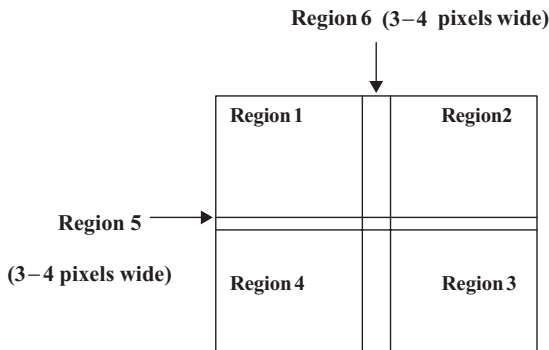


Figure 6. Different regions of the frame.

rants. For the sake of uniformity, the frame is assumed to have totally six regions, as depicted in figure 6. Information about regions in which the given stroke starts, passes and finally ends is used in stroke recognition along with the heuristic knowledge.

The following are the heuristics adopted during the recognition of PSL strokes/character.

- To distinguish between thin and thick strokes, a concept called ‘threshold’, which represents the width of the stroke in pixels, is introduced. The pixel in question is preserved when the number of neighbouring 1’s is less than or equal to the threshold or the number is more than the threshold but continuity may be lost by its removal.
- A stroke is a straight line if it passes through the centre of the trimmed image. This heuristic coupled with the knowledge of the regions containing the starting and final points is used to recognize the strokes.
- A stroke is said to be a hook if it is a straight line with a cluster of pixels at one of the extreme points.

Further, the combination of heuristics and the region to classify the strokes is given in the form of a complete decision tree and is given in figure 7.

5. Algorithm description

The overall approach is summarized as a sequence of steps. Accept the binary image and subject it to preprocessing. Assume suitable values for parameters like limit and threshold. Extreme points of the stroke and region information are determined. The heuristic knowledge is applied and the decision tree is used for recognition of the character. These steps are implemented as separate sub-algorithms and are described in this section.

(a) *Binary image generation:* This algorithm converts the scanned images of strokes to their equivalent binary images.

Begin

Step 1. BMP file of the image is accessed and the starting point, height and width of the given image is obtained.

Step 2. For each pixel in the BMP file of the image


```

If the pixel is ON then
    1 is stored in the binary file.
Else
    0 is stored in the binary file.
    
```

End.

(b) Thinning of stroke: This algorithm thins the given stroke for further recognition and the knowledge about the nature of the stroke, i.e. thick or thin, is obtained. The algorithm “Can-thin” decides whether the image requires thinning or not.

Begin

Begin can-thin.

Step 1. For the pixel under test the number of 1’s in the neighbouring pixels is counted.
 Step 2. If (the count \leq the threshold) then

```

        Return (0)
    Else
        Return(1)
    
```

End can-thin

Step 1. Raster-scan the image from left to right

Step 2. If (can-thin) then

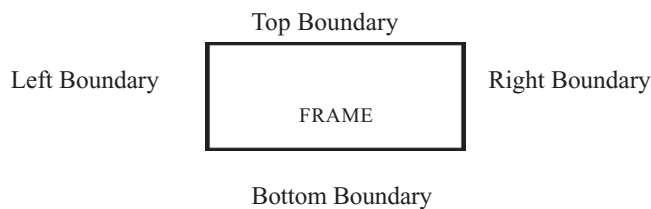
```

        Set test pixel to ‘o’
    Else
        Preserve the test pixel.
    
```

Step 3. The steps 1 and 2 are repeated for the image from right to left, top to bottom and bottom to top.

End.

(c) Trimming algorithm: This algorithm decides the boundary for the strokes and fixes the entire image into a frame of size MxN as shown below.



Begin

- Step 1. The binary image is scanned from the top and the top boundary is fixed.
- Step 2. The binary image is scanned from the bottom and the bottom boundary is fixed.
- Step 3. The binary image is scanned from the left and the left boundary is fixed.
- Step 4. The binary image is scanned from the right and the right boundary is fixed.
- Step 5. The coordinates of extreme points are stored.

End.

(d) *Knowledge about extreme point:* This algorithm finds the extreme or end point on the stroke. This point has exactly one neighbour.

Begin

1. The image is raster scanned.
2. The pixel that has only one neighbour is found, which is one extreme point.

End

(e) *Tracing the stroke:* This algorithm stores the coordinates of the pixels, which are ON along the stroke.

Begin

1. The neighbour of the extreme point is determined.
2. If this neighbour pixel is ON and already visited then it is marked as traced.
3. Repeat step 2 until the end is reached.
4. The final point is noted.

End

6. Results and discussions

The PSL strokes are of size $1/6^{\text{th}}$ of an inch and the images of size 64×64 are considered for the recognition. The proposed knowledge based approach works with 100% accuracy for the ideal inputs. The large number of handwritten strokes, a subset of the sample set is given in figure 8, are tested with this approach and the results are tabulated as in table 2. The efficiency is reduced for the handwritten inputs. The input set for each character is made up of 250 handwritten strokes. Thus, the size of the input data set is 250×24 strokes. The two major parameters that govern the accuracy of recognition are the threshold and the limit values.

6.1 *Effect of threshold values on the success rate of strokes.*

The effect of threshold values on the accuracy of recognition of a stroke is a matter of interest in this approach. The success rate for all characters for different threshold values at different limit values is summarized in table 2. The limit values of 4,6,8 and 10 and threshold values of 1,2 and 3 are considered.

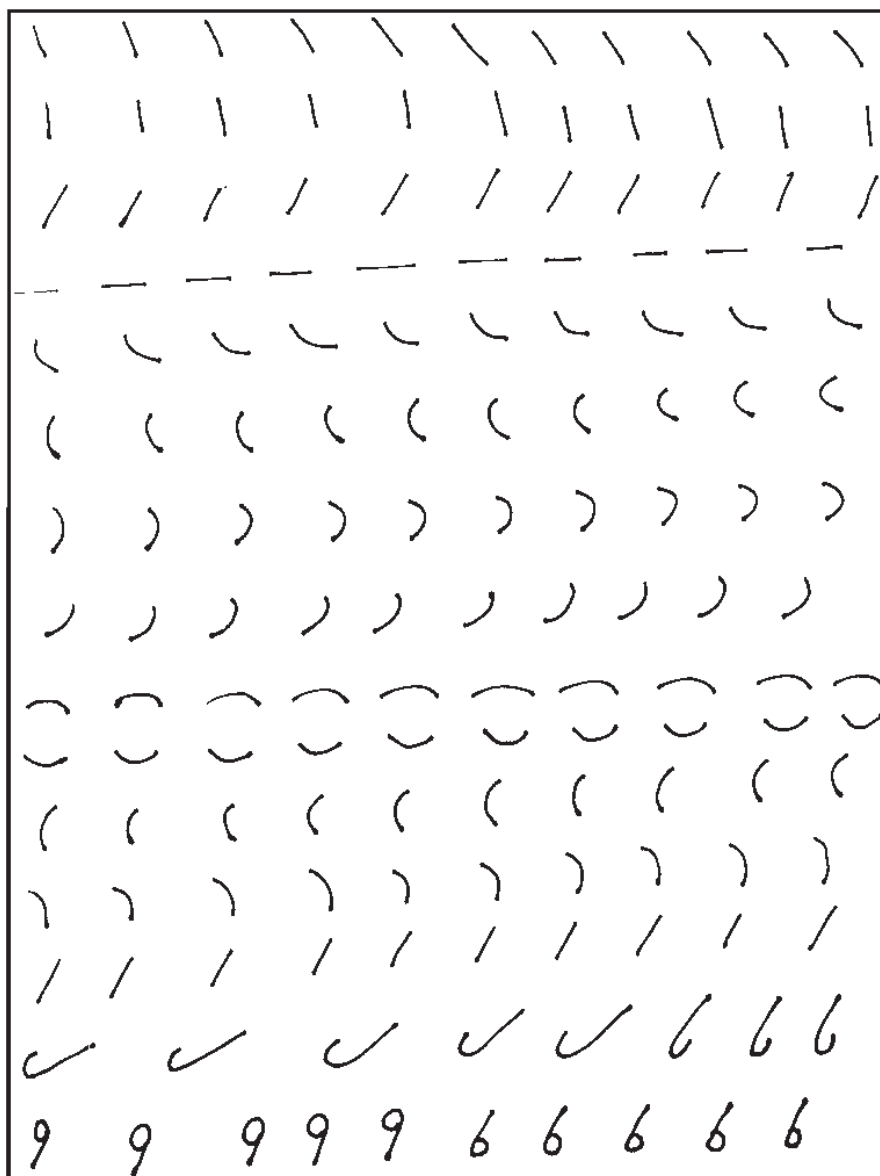


Figure 8. Samples of scanned handwritten PSL strokes.

It is observed that the success rate for the thick strokes like B, D, Gay, etc. is less. This is attributed to the resolution of the scanner. The success rate also depends on the clarity of writing and thickness of writing is considered as a relative term. The success rate increases noticeably, when the thick characters are written with more intensity. Otherwise, the characters will be recognized as thin strokes, leading to reduction in the success rate. For example, the thick stroke representing 'B' is interpreted as the thin stroke 'P'. Any increase in the threshold value leads to partial thinning of the strokes, i.e. thinning does not reduce the stroke to single pixel width. Higher threshold values (> 4) affect adversely the recognition of strokes. It is

Table 2. Success rate for limit = 4, 6, 8, 10 and threshold = 1, 2, 3..

Characters	Limit = 4			Limit = 6			Limit = 8			Limit = 10		
	Threshold			Threshold			Threshold			Threshold		
	1	2	3	1	2	3	1	2	3	1	2	3
P	92	90	88	93	92	88	96	94	93	96	94	92
B	80	78	76	81	78	75	80	81	80	80	81	78
T	92	90	88	90	88	78	92	91	90	92	90	89
D	84	84	80	85	86	79	84	82	79	84	80	81
CH	88	86	80	89	84	76	88	85	79	88	79	75
J	76	77	68	75	76	70	80	78	79	80	78	72
K	92	90	81	90	82	73	92	88	78	92	89	81
G	80	78	76	81	78	71	80	79	74	80	76	70
F	80	81	77	78	76	73	80	79	77	76	74	69
V	76	77	76	77	68	69	76	73	69	72	68	62
TH	84	80	75	85	78	69	84	79	76	72	69	70
S	88	86	79	85	78	71	88	79	77	84	79	74
Z	76	75	68	77	76	77	76	70	72	76	78	73
M	68	66	60	66	67	64	72	68	66	60	62	64
N	68	65	58	68	64	61	72	68	69	68	70	71
NG	68	64	69	65	63	60	68	66	60	64	62	59
SH	84	80	81	80	78	70	92	89	80	84	80	77
ZH	76	72	69	65	64	61	84	79	77	80	78	74
R	88	80	78	90	85	79	96	87	80	74	75	74
L	84	78	76	86	83	80	92	79	74	84	80	78
W	68	69	62	65	67	65	72	66	64	60	61	58
Y	68	62	60	63	66	60	68	65	60	52	50	51
H	52	50	50	60	60	61	56	52	50	48	49	46

observed that the optimum results are obtained when threshold is 1 and the bar chart depicting this is given in §6.2.

6.2 Effect of variations in the limit on the success rate

Another reason for low success rates is the discontinuity present in the scanned images as shown in the figure 9. As explained in the §3, the filling algorithm is used to fill the gaps

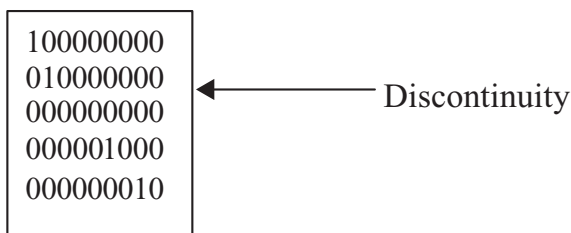


Figure 9. Discontinuities in scanned image.

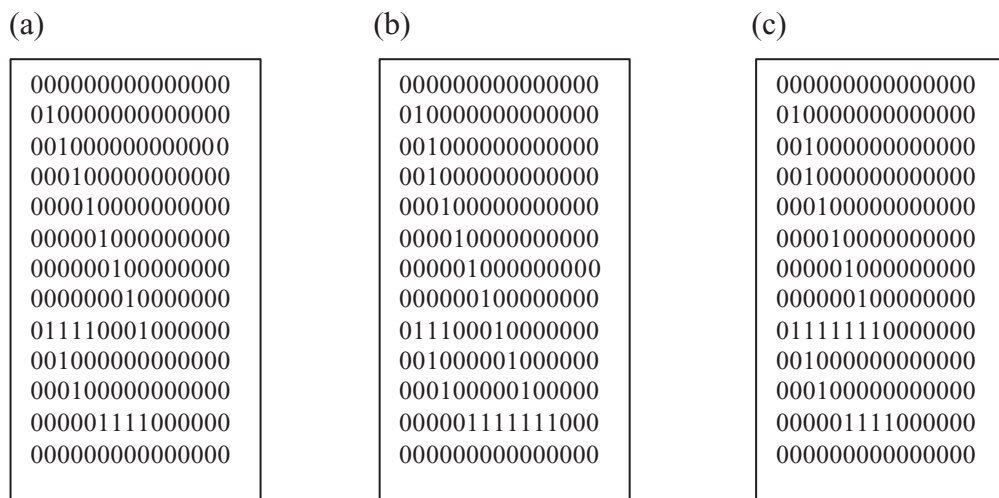


Figure 10. Effect of varying the limit. (a) Scanned image, (b) intended output of filling, (c) obtained output.

between the stroke pixels. The gap expressed in terms of pixels is called the limit and is used in filling the gap. Normally, a gap of 1 to 8 pixels is observed in the scanned images. Hence it is appropriate to choose a limit value less than 9.

Variation in the limit value has direct effect on the rate of recognition of strokes. The results obtained for limit = 4, 6, 8, and 10 are summarized in table 2. Any increase in limit value beyond a certain number has an adverse effect on filling the stroke discontinuities, as is illustrated in the figure 10. If the limit value is less than say a value = 3, then discontinuous ends are not connected. On the other hand, if the limit value is beyond a certain number then erroneous result is obtained, shown in figure 10c. The success rate of stroke recognition for different values of limit at different threshold values is summarized in table 2. It can be observed that the success rate for limit = 10 is less than the success rate obtained for limit = 8. The optimal results are obtained when threshold = 1 and limit = 8. The PSL strokes are categorized into 5 types namely, explodents (1), continuants (2), nasals (3), liquids (4), and coalescents (5). The results obtained are quite satisfactory and are accurate for good handwritten PSL strokes.

7. Conclusion

Results of the experiments conducted to bring out the efficacy of the approach presented are encouraging. The success rate is high for explodents, continuants, nasals, and liquids, which is in the range 87–94%. However, for coalescents it is somewhat lower, in the range 65–70%, which is attributed to the small circles/hooks present at the end of the strokes. The heuristic approach has a recognition rate of 99 to 100% for ideal strokes. The approach performs poorly for largely distorted handwritten strokes. In general, stenographers take down dictation neatly because later generation of the English text depends on the readability of handwritten strokes. The work has global application in the areas of secret communication, legal transcription, compression etc. The other features of PSL like vowels, diphthongs, word segmentation etc. is under investigation.

The authors acknowledge the help extended by Babu Kiranagi in the preparation of this paper.

References

- Gose E, Johnsonbaugh R, Jost S 1999 *Pattern recognition and image processing* (New Delhi: Prentice Hall of India)
- Harowitz E, Sahani S 1976 *Fundamentals of data structures* (Woodland Hills, CA: Computer Science Press)
- Hemanth Kumar G 1998 *Automation of text production from Pitman shorthand notes*. Ph D thesis, University of Mysore, Mysore
- Kuroda K, Harada K, Hagiwara M 1999 Large scale online handwritten Chinese character recognition using successor method based on stochastic regular grammar. *Pattern Recogn.* 32: 1307–1315
- Leedham C G, Downtown A C 1984 On-line recognition of short forms in Pitmans' handwritten shorthand. *Proc. 7th Int. Conf. on Pattern Recognition*, Montreal, pp 2.1058–2.1060
- Leedham C G, Downtown A C 1986 On-line recognition of Pitmans' handwritten shorthand - An evaluation potential. *Int. J. Man-Machine Studies* 24: 375–393
- Leedham C G, Downtown A C 1987 Automatic recognition and transcription of Pitman's handwritten shorthand - An approach to short forms. *Pattern Recogn.* 20: 341–348
- Leedham C G, Downtown A C 1990 Automatic recognition and transcription of Pitman's handwritten shorthand. *Computer processing of handwriting* (eds) R Plamondon, C G Leedham (Singapore: World Scientific)
- Leedham C G, Downtown A C, Brooks C P, Newell A F 1984 On-line acquisition of Pitman's handwritten shorthand as a means of rapid data entry. *Proc. Int. Conf. on Human-Computer Interaction*, London, pp 2.86–2.91
- Nagabhushan P, Anami B S 1999 A knowledge based approach for composing English text from phonetic text documented through Pitman shorthand language. *Int. Conf. On Computer Science (ICCS-99)*, New Delhi, pp 318–327
- Nagabhushan P, Anami B S 2000 A knowledge based approach for recognition of grammalogues and punctuation symbols useful in automatic English text generation from Pitman shorthand language documents. *Proc. Natl. Conf. on Recent Trends in Advanced Computing (NCRAC-2000)*, Thirunelveli, pp 175–183
- Nagabhushan P, Murli 2000 Tangent feature values and cornarity index to recognise handwritten PSL words. *Proc. Natl. Conf. on Document Analysis and Recognition (NCDAR)*, Mandya, India, pp 49–56
- Patterson D W 1999 *Introduction to artificial intelligence and expert systems* (New Delhi: Prentice Hall of India)
- Pitman I 1976 *Pitman shorthand instructor and key* (Wheeler Publisher)
- Samouelian et al 1994 Knowledge based approach to English consonant recognition. *Proc. Int. Conf. On Acoust. Speech & Signal Process.*, pp 77–80, Piscataway, NJ