

Touching Syllable Segmentation using Split Profile Algorithm

L.Pratap Reddy[1], T.Ranga Babu[2], N.Venkata Rao[3] and B.Raveendra Babu [4]

**[1] Department of Electronics and Communication Engineering,
JNTUH College of Engineering, Hyderabad - 500085, Andhra Pradesh, India**

**[2] Department of Electronics and Communication Engineering,
RVR & JC College of Engineering, Guntur - 522019, Andhra Pradesh, India**

**[3] Department of Electronics and Communication Engineering,
Sri Vasavi Engineering College, Tadepalligudem - 534101, Andhra Pradesh, India**

**[4] Department of Computer Science and Engineering,
RVR & JC College of Engineering, Guntur - 522019, Andhra Pradesh, India**

Abstract

The most challenging task of a character recognition system is associated with segmentation of individual components of the script with maximum efficiency. This process is relatively easy with regard to stroke based and standard scripts. Cursive scripts are more complex possessing a large number of overlapping and touching objects, where in the statistical behavior of the topological properties are to be studied extensively for achieving highest accuracy. Certain amount of similarity exists between unconstrained hand written text as well as South Indian scripts in terms of topology, component combinations, overlapping and merging characteristics. The concept of syllables and their formulations is an additive complexity with regard to Indian scripts. In this paper the statistical behavior of the cursive script, Telugu, is presented. The topological properties in terms of zones, component combinations, behavioural aspects of syllables are studied and adopted in the segmentation process. The statistical behaviour of cursive components are evaluated. Split Profile Algorithm is proposed while handling touching components. The proposed algorithm is evaluated on different fonts and sizes. The performance of the proposed algorithm is compared with two approaches methods viz aspect ratio and syllable width approaches.

Keywords: *Segmentation, Connected Components, Syllable Segmentation, Touching Syllable, Split Profile Algorithm.*

1. Introduction

Document image analysis and Optical Character Recognition (OCR) systems are under continuous research for decades. The transformation of paper media into the searchable and revisable text format gives a great boost in the field of language technology based research. Automated content creation from printed or written form of ancient and later versions of documents is the major area of OCR research. Achieving accurate results under all possible conditions remains as a challenging task. The first step in this process is to achieve maximum efficiency in character segmentation, which inturn reflects in OCR accuracy.

Script topology plays a dominant role in the segmentation process. Structural features describe the patterns of topology and geometry while exploring global and local properties. White spaces and pitch information are the useful primitive parametric data of any segmentation system. The notion of detecting vertical white space between successive characters is an important concept while dissecting images of machine print and even in hand written documents. Apart from this, other topological features like height, width and orientation etc., are useful parameters. In case of fixed width characters, pitch information provides effective segmentation. However, variable width characters are found in almost all scripts due to large number of font designers. As a result, various segmentation approaches are proposed [1] in literature to handle this complexity.

Structural properties of natural language script are another useful piece of information, to be adopted while choosing the segmentation approach. Scripts can be further classified as stroke based, cursive and hybrid. Segmentation of stroke based scripts can be performed by

making use of properties like horizontal, vertical and slant line information.
Segmentation approaches that are to be adopted for cursive scripts are complex when compared with stroke based scripts. Character shapes of these scripts posses variable widths and sizes, originated from a combination of more than one component. The topological or structural properties of individual components and their associative nature with other components transform the final shape, occasionally leading to overlapping and touching phenomena. Segmentation issues of these scripts are to be addressed by considering common statistical properties along with specificities of the respective class of formulations. On the other hand, hybrid model is associated with a set of strokes as well as curves. The primitive shape (glyph) is to be treated as the main focus of this model.

2. Review

Various segmentation methods adapted in document image processing are described [1] by Richard G.Casey and Eric Lecolinet. Profile based approach proposed [2] by K.Ohta which is considered as a simple and effective method for segmenting a print line. These approaches are reported to be effective for non-cursive writing systems and still found their applications even in handwritten recognition systems. Detection of white spaces can be effectively carried out on a structured text image. Identification and extraction of vertical strokes is made simple using this method. Analysis of peaks and valleys of profile patterns extended the scope of profile method for partial adaptation into touching character segmentation. In [1], the profile was first obtained and then the ratio of second derivative of this pattern to its height is used as a criterion for identifying segment separation. The peak of the derivative, which is associated with projection minima converge the splitting points along the thin horizontal lines.

A peak-to-valley function is proposed in [3] by Y.Lu with further improvements. Spatial domain characteristics based on the topological features of the script are explored. Valleys between successive peaks are extracted from the profile function. An average function is used to identify the extract segment point with a specific reference to touching characters. A selection criterion of the segmentation boundary is associated with discriminating function of topological features of individual characteristics.

Bounding box approach [4] is proposed by M.Cesar and R.Shinghal as an alternative to profiling method. They reported that the method is effective on stroke based script. Splitting and merging of character component is reported in [4]. The connected component approach proposed in this work is mainly concentrated on defining specific rules using height and width parameters of bounding box.. The adjacency relationships between bounding boxes, their size and aspect ratios are explored for splitting mechanisms. Segmentation effectiveness is reported with high degree of accuracies even at low computational requirements. Extension of connected component approaches is proposed in [5] by G.Seni and E.Cohen for segmentation of hand written words in a document image.

The CJK script models are more predominant in strokes and relationships among the strokes are well structured [6]. Latin text, European language models describe the dominance of strokes. The linear property of strokes is explored in using profile based approach while segmenting characters of all the above scripts.

North Indian scripts are hybrid in nature, combining strokes and curves, where strokes are dominated by curves. Linear spatial relationship in the form of shirorekha (a straight line combining components) can be found within the topological structure. The resultant form of this linear relation is treated as zone, which is used to establish correlation among strokes and curves within the syllables. The top zone of the character resembles stroke like geometry. The positional information of zone is identified by finding the linear region from the profile function of script line. Segmentation is achieved by exploring the statistical properties [7,8,9,10,14] of zone information using profile based methods.

Arabic and South Indian scripts are dominated by curve like components. In Arabic scripts, the formation of character is nonlinear and base line is identified by the peak of vertical profile function[11]. South Indian scripts are derived from the writing style on palm leaves, resembling cursive nature in machine print as well as hand written. The process of character formulations resulting from component combinations with zero width joiner and some times with non joiner leading to overlapping phenomena. Character formation in these scripts (also known as syllable) deals with two part glyphs in certain cases, deviating from the linear process. Notable number of non-linear combinations exists in these scripts. Segmentation is to be addressed by taking into consideration of all processes, linear and non-linear combinations. In this context, the statistical behavior of component shapes within the boundaries of text line, any word and even a syllable, influences the segmentation strategies.
In South Indian scripts, curves are more predominant and extraction of zone information is complex. Syllable is formed by a set of curves with high degree of similarity among them. The individual components in the syllable are

extracted and associated relationship is studied using zone information. The extraction of zone information is complex because of the non-linear distribution of glyphs in the upper and bottom regions. Common properties are reported in [12,13] by extensive statistical evaluation of the profile function.

Profile method [11] found its use not only in printed text but also in hand written text. Handwritten profile information of the script line is used to identify the linear portions of the script characters. The profile information of a word differs from that of a line. At the same time, multiple word combinations of a script line posses linear behavior in the profile patterns. Extensive statistical evaluation of various script lines is necessary while formulating rules in the zone identification process. Similar studies are extended to syllable segmentation in the present paper. A detailed description of segmentation model is presented in the following section.

3. Segmentation Model

The segmentation model in this paper explored the statistical patterns of profile vector which signifies the topology and geometry of printed text with specific reference to cursive script of Telugu. Preprocessing steps like binarisation, skew correction are carried out on the document image before proceeding to segmentation algorithm.

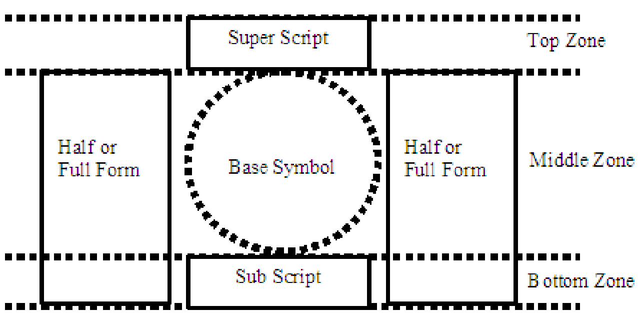


Fig. 1 Canonical Syllable Model.

Four phases are proposed in the segmentation process. First phase deals with line segmentation and extraction of zone information, second phase deals with syllable segmentation, third phase addressed the classification of segmented syllables into touching and non-touching objects and fourth phase is emphasised on segmentation of touching objects using Split Profile Algorithm (SPA).

In the first two phases, connected component approach is adopted for segmentation of syllable. Syllable model proposed [13] by Pratap et al, presented in Fig. 1 and is adopted in the present phase.

Topology of the syllable can be decomposed into component like glyph objects. One base symbol object, also treated as essential component, is the minimum topology. In a complex conjunct syllable a maximum of four other components will be positioned, as in the above figure. The number of components may vary in between 1 and 5. However the topological features in the form of zones is difficult to extract due to the inherent nature of zero width joiner and non joiners between components. This phenomenon reflected in the form of touching syllables that are predominant in various font sizes.

Using the above model, syllable segmentation and classification of touching syllable is carried out in phase 2 and 3 respectively. In the last phase, segmentation of touching syllable is addressed with the help of SPA. Topology of various syllable components is studied after splitting the profile. Prediction of segmentation threshold is carried out in the separation process of touching syllables.

3.1 Line and Zone Separation

Different scripts posses varied structural features. However machine printed document images are structured in nature with a similarity around script line distribution. The linear property from pixel distribution of Horizontal projection Profile (HPP) is adopted for line segmentation. HPP is obtained using Eq.(1)

$$H_{PP}(i) = \sum_{j=1}^{N} f(i,j) \qquad (1)$$

where i = 1,2,3,.. M (Height of the object)
j = 1,2,3….N (Width of the object)

White spaces between text lines are treated as delimiters in ideal case. However under the influence of noise the profile distribution between lines reflects the random nature of noise information. In the present case, we considered ideal scenario, where the noise component is negligible. Starting point and ending point of script line is found with certain amount of black pixel distribution, using which the lines are segmented.
Pixel distribution of script line is studied on various document images. Certain amount of linear behavior is found in the form of peaks and valleys, reflecting the zone information. The geometry of individual syllable does not match with zones, which is also the case with certain words where as multiple combinations of words found to

be linear. One peak in the first half of profile distribution is observed. This peak matched with zone separation line between top and middle zones. However zone separation line between middle and bottom zone is reflected in the form of maximum slope in the later half of the profile formation.

The detailed algorithm for Line and Zone separation is listed below

Step 1 Extract horizontal profile vector for the whole document image.
Step 2 Divide the Lines profile vector which consists of starting and ending of the line in the image and mark them as top line and bottom line respectively.
Step 3 Extract first line from the document image
Step 4 Divide the script line profile vector into two halves along the length.
Step 5 Find the row with peak value of black pixel density in the first half of the profile vector and mark it as Head Line.
Step 6 Find the rows with peak values of black pixel density in the later half of the profile vector.
Step 7 Find the slope of the valley from the row in Step 6 with respect to the successive row.
Step 8 The row with the maximum slope of the valley is identified as the base Line.
Step 9 Extract the Top-Middle-Bottom zones of the script line. Get the zone information for different script lines.

Original document image is presented in Fig.2 Segmentation of lines and extraction of zone information, as defined in steps 1-5, is presented in Fig.3 and Fig.4.

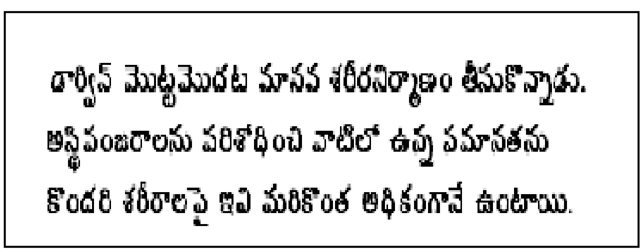


Fig. 2 Original Document Image.

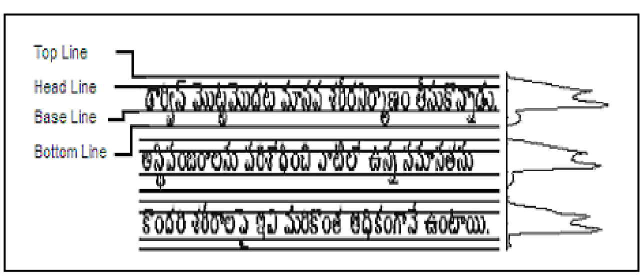


Fig. 3 Line & Zone separation fro horizontal profile.

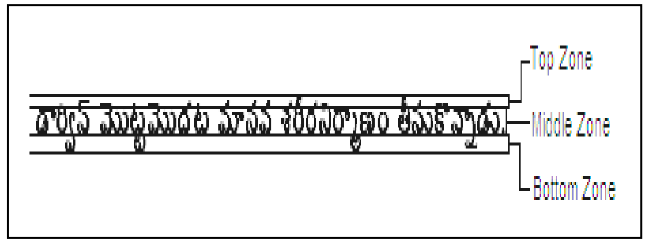


Fig. 4 Zone Information of text sample.

3.2 Syllable Segmentation

In an ideal scenario, individual glyph components (Fig.1) can be decomposed using zone information. The canonical space is extracted from the text document using connected component approach with a reported [13] segmentation efficiency of 95.72%, without addressing the touching syllables. Similar approach is adopted in the present phase. The component objects that are separable, are identified with the help of labeling approach. Grouping of core and non-core components are carried out while segmenting syllable objects. These objects may include touching syllables also. The syllable segmentation is given below

The detailed algorithm is as follows:

Step 1 Extract the zone information for different script lines.
Step 2 Label the pixels by scanning them from left to right in each row and row after row.
Step 3 Find adjacent labels and connect the labels to form basic components.
Step 4 Find unique labels and extract component with unique labels
Step 5 Identify whether it is core component or non-core component using zone information
Step 6 Merge the non-core components with core components using zone information.
Step 7 Extract the Syllable from the input text image and draw the bounding box.

The concept of core component and other components in a syllable as proposed by Pratap Reddy et al. [13] is a reflective mechanism of the topology and geometry with regard to Telugu script. This property is explored in steps 1-7 and the result is presented in Fig.5.

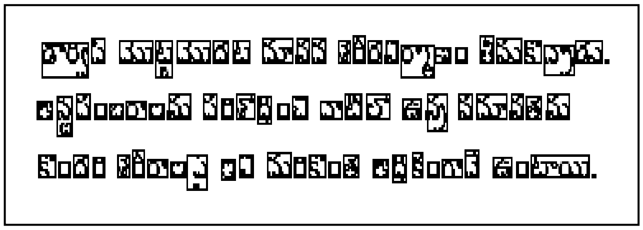


Fig. 5 Segmented Syllables of document image.

3.3 Classification of Correctly Segmented Syllables

In the process of improving segmentation efficiency, it is required to classify the correctly segmented syllables against the others. The syllable objects, extracted from the previous stage are a combination of touching and non-touching syllable objects. Aspect ratio (the relation between component height and width) is a simple approach adopted for this purpose which is defined in Eq.(2).

$$\text{Aspect Ratio}(A) = \frac{\text{Component Width}}{\text{Component Height}} \qquad (2)$$

The optimum threshold Th_{ASP} is defined in Eq.(3)

$$Th_{ASP} = \left[\frac{\sum_{i=1}^{N} A(i)}{N} \right] \qquad (3)$$

Th_{ASP} is an averaging function of the relationship between component objects, which is used as a threshold value. The syllable object with $A \geq Th_{ASP}$ is treated as touching syllable. Segmented output after adaptation of aspect ratio is presented in Fig.6 and Fig.7.

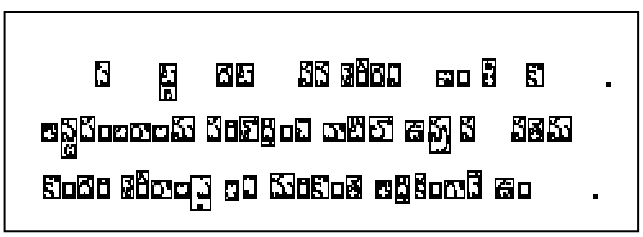


Fig. 6 Syllables classified as Single Syllable using Aspect Ratio.

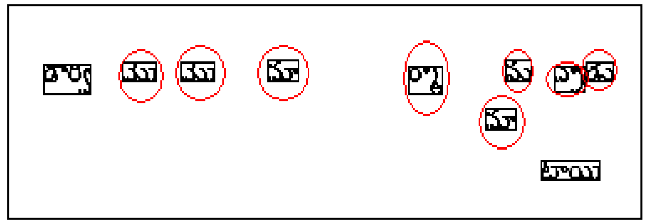


Fig. 7 Syllables classified as Touching syllable using Aspect Ratio

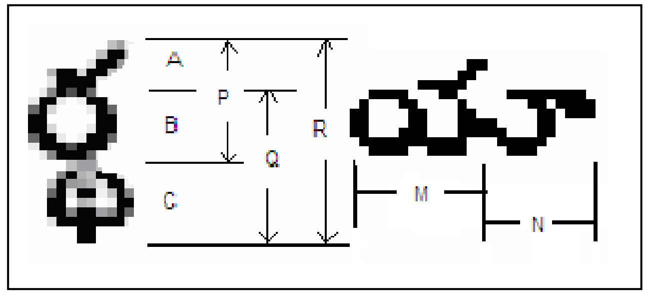


Fig. 8 Height and width variation of Syllable.

Topology of the script is so complex, at times height and width of a syllable posses large variations as described in Fig.8

The variables A,B,C define the height of top zone, middle zone and bottom zone layers. P,Q,R variables are computed as A+B, B+C, A+B+C respectively. It is quite interesting to see that the syllable height may have varying combination among B, P, Q, R. Similarly two different specifications can be made with regard to syllable width M, defined as core width and N, as modifier width. Two possibilities always exist. The width of the syllable may be M or some times extends to M + N depending on the type of modifiers. The value of M and N differs for each core component as well as modifiers. The possible combination of aspect ratio finally leads to 8 different types for each type of core component. There are 36 different types of core components, where the width and sizes differs slightly. In addition to that another 13 modifiers exists. The entire complexity can be visualised in the form of classification error leading to isolated syllables pushed into touching syllables. This error influences segmentation efficiency up to a large extent.

While resolving the segmentation problem and to reduce the classification error, it is necessary to identify effective classification mechanism. As per the evaluation in [15] the syllable groups and their frequency statistics reveal the fact that modifier combinations are only 16% of the text information. Large numbers of syllable groups are focused around core width. Average syllable width is one of the parameter for better judgment.

The optimum threshold for classification of syllable objects is redefined as Th_{SYW} (Average Syllable Width), and is given in Eq.4

$$Th_{SYW} = \left[\frac{\sum_{i=1}^{N} W(i)}{N} \right] \tag{4}$$

The outcome of the classifier using Th_{SYW} is presented in Fig.9 & Fig.10. It is observed that the error is now reduced further, while classifying touching syllable only.

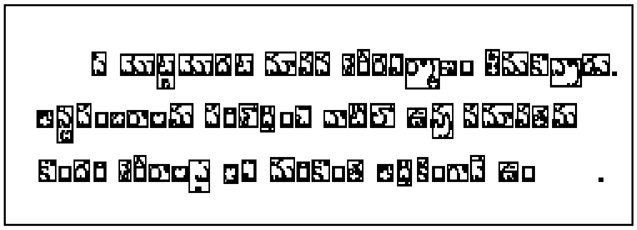


Fig. 9 Syllables classified as Single Syllable using Syllable Width.

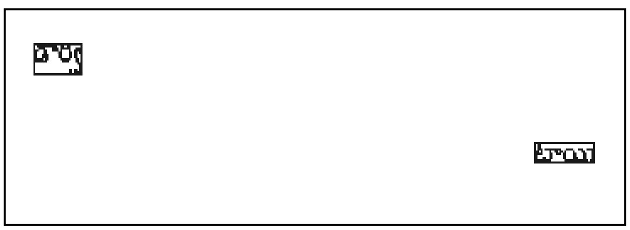


Fig. 10 Syllables classified as touching syllable using Syllable Width.

After careful analysis on 38,419 syllables (12,14,16 &19 font sizes) of Anupama font, it is found that segmentation efficiency is 91.39% for Aspect ratio, where as 94.48% for Syllable width. An improvement of 3 % is observed. For Gowthami font (30,871 syllables) segmentation efficiency is observed as 86.63% and 89.20% for aspect ratio and syllable width respectively. In the case of Priyanka font average efficiency is observed as 81.73% and 87.32% respectively.

3.4 Split Profile Algorithm

The topology of the syllable, where cursive nature is more predominant, provides information about the placement and their relationship among the glyphs. The syllables are formed with the placement of glyph(s) with varying style and sizes. In certain occasions one syllable may overlap with the next syllable. This situation may appear at different zones of the syllable, resulting in complex behavior. The statistical analysis on different fonts (Anupama, Gowthami and Priyanka) reveals the fact that large number of overlapped characters is present in small sized font structures. Some syllables with large width may be classified as touching syllable. They will play a major role on segmentation efficiency. Connected component approach fails in segmenting the touching characters for the reasons explained above.

To handle this problem we propose an alternative method, Split Profile Algorithm, where the profile information of the touching syllable will be split into two parts and the geometric features are studied for identification of touching labels.

Two parameters are introduced in the algorithm. The first one 'Horizontal Split Threshold' (ST_H) which is used to identify the splitting index of the text line segment. The second one 'Vertical Split Threshold' (ST_V), to identify the segment boundary of touching syllable.

Horizontal split threshold is defined as

ST_H = Head Line + ((Bottom Line – Head Line)/2)

Vertical split threshold is defined as

$ST_V = 2^{nd}$ transition point from low to high in the lower part of the profile

Either the top segment or bottom segment can be considered for extraction of syllable index. The topology of the touching characters is mostly associated with positional structure of glyph like components. In this context one of the line segments will provide positional information of the above components. Vertical profile of the respective portions of line segments is plotted for this purpose. The profile part with multiple segments is used for defining segmentation index. From the statistical study it is observed that the starting point of the second component is the most likely predicted segment boundary from ST_V. The split profile algorithm which was applied is as follows

The algorithm for SPA

Step 1 Get the horizontal profile of the syllable
Step 2 Split the horizontal profile using 'ST_H'
Step 3 Get the vertical profile for the lower part of the image.
Step 4 Split the touching syllable by using Vertical Split Threshold 'ST_V' index from the scaled profile obtained from step 3.

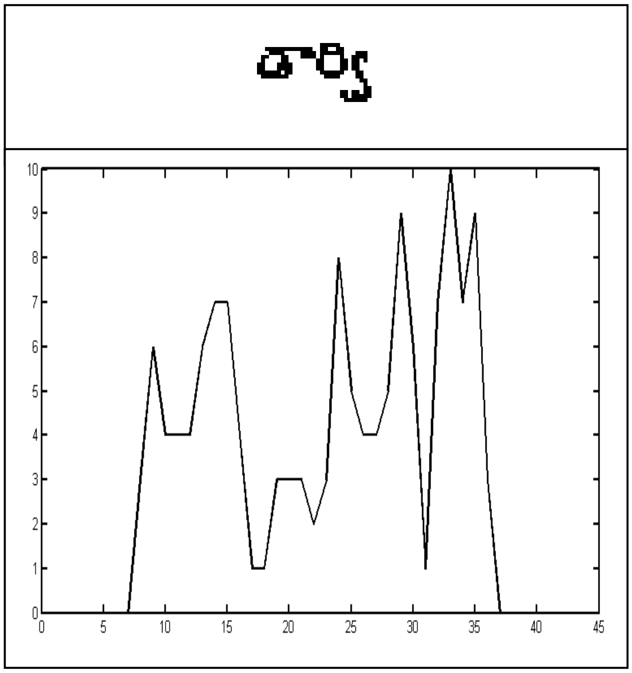


Fig. 11 Vertical Profile of the First Syllable

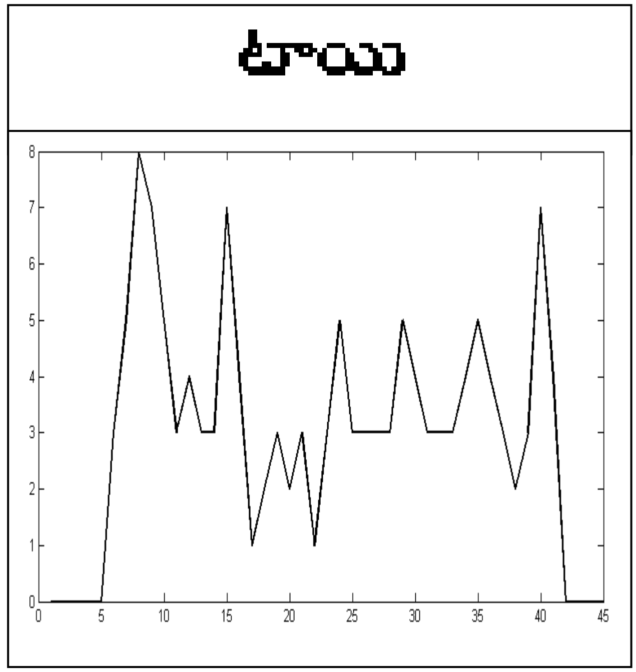


Fig. 13 Vertical Profile of the Second Syllable.

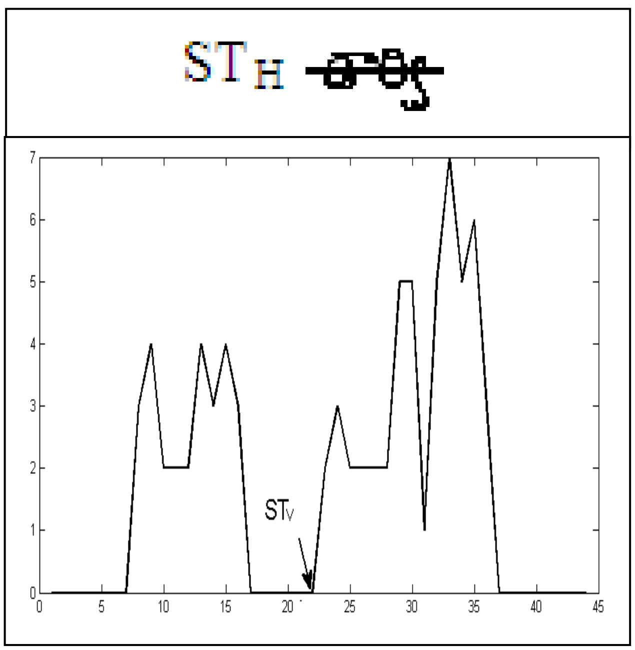


Fig. 12 Scaled Vertical Profile for the lower part of the First Syllable.

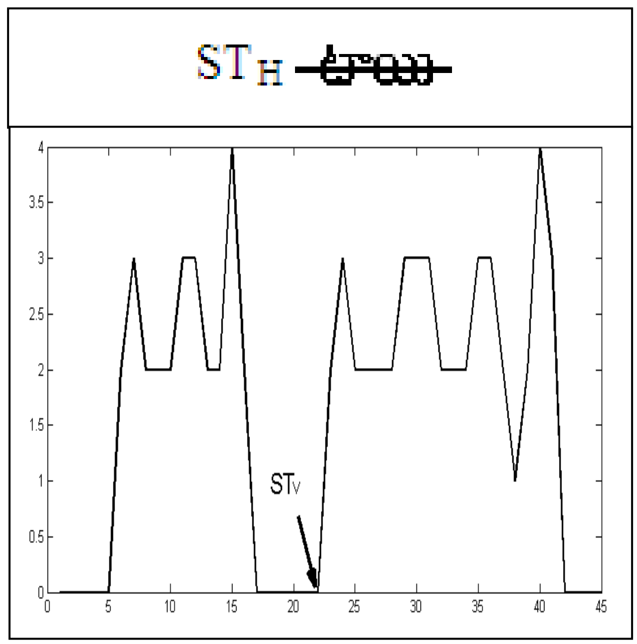


Figure 14. Scaled Vertical Profile for the lower part of the Second Syllable

Table 1: Output for the first syllable after SPA

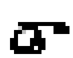
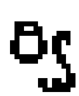
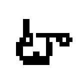
	

Table 2: Output for the second syllable after SPA

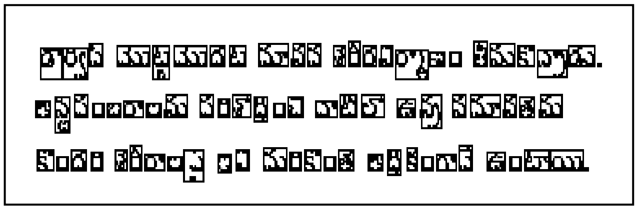


Fig. 15 Segmented Syllables of text sample after SPA

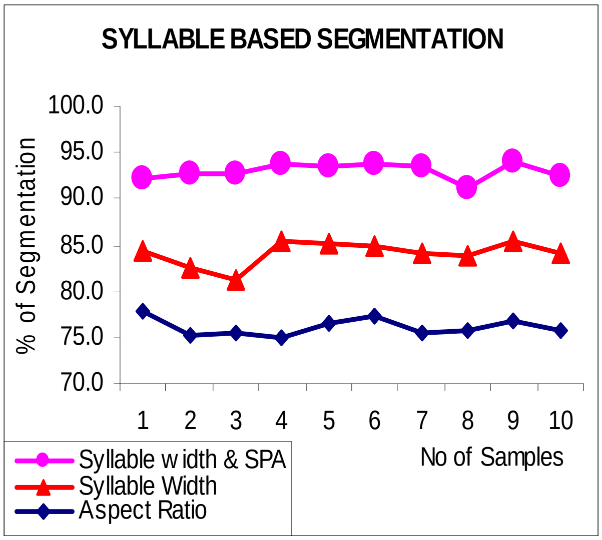


Fig. 16 ANUPAMA for size 12.

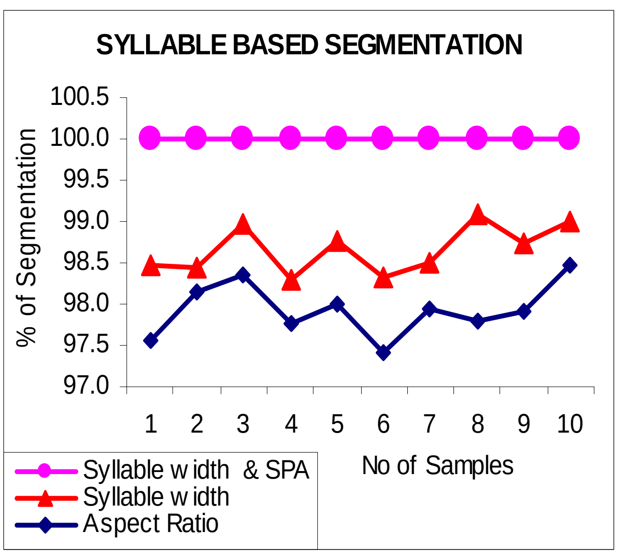


Fig. 17 ANUPAMA for size 14.

The first bounding box of the touching syllable in the earlier Fig.10 is considered for evaluation. Fig.11 and Fig.13 depict the complete profile of the two syllables, where as Fig.12 and Fig.14 is presented with scaled vertical profile of the linear segment after applying the ST_H and the split profile ST_V. Segmented syllables after adaptation of ST_V are presented in Table 1 and 2. Fully segmented syllables after the Split Profile Algorithm is presented in Fig.15. Various combinations of touching syllables, which lie in various zone combinations are studied with the help of above algorithm.

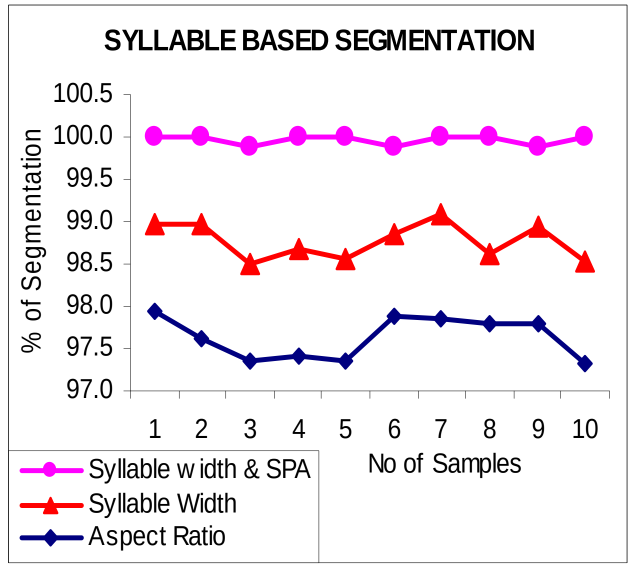


Fig. 18 ANUPAMA for size 16.

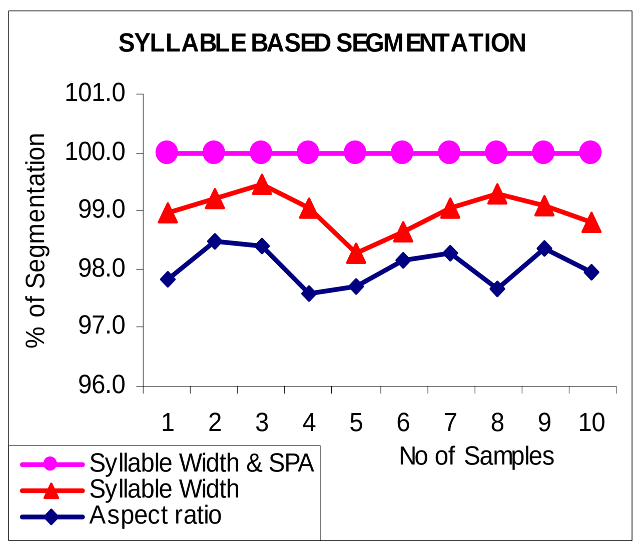


Fig. 19 ANUPAMA for size 19.

Efficiency comparisons between Split Profile Algorithm, Syllable width and Aspect Ratio are presented in Fig.16 to Fig.19

4. Results

The proposed algorithm is evaluated on 1,11,582 syllables of Anupama, Gowthami, and Priyanka font. Segmentation is carried out on font sizes of 12,14,16 and 19. Syllable segmentation efficiencies of aspect ratio, syllable width and Split Profile Algorithm for Anupama font is presented in figures 15 to figures 18. SPA outperform with 100% segmentation efficiency on the sample set of size 14, 16 and 19. The syllable width based approach is observed with average segmentation efficiency of 98%. The aspect ratio based approach is observed with segmentation efficiency ranging from 97.93% to 98.04%. For font size 12, SPA is observed with maximum segmentation efficiency of 92.98% against 84.15% and 76.12% with syllable width and aspect ratio respectively. However, when evaluated on samples of font sizes 12,14,16 and 19, the average segmentation efficiency of SPA is observed as 92.98%, 100%, 99.96% and 100% where as syllable width approach is observed as 84.15%, 98.65%, 98.77% and 98.98% and aspect ratio is found to be 76.12%, 97.93%, 97.63% and 98.04% respectively. Comparison of segmentation efficiencies for different fonts and sizes presented in Table 3, 4 and 5

Table 3: Aspect Ratio

Font Size	12	14	16	19
Font Type	Segmentation efficiency			
Anupama	76.12	97.93	97.63	98.04
Gowthami	73.75	95.00	92.95	95.42
Priyanka	50.44	85.13	97.72	98.98

Table 4: Syllable Width

Font Size	12	14	16	19
Font Type	Segmentation efficiency			
Anupama	84.15	98.65	98.77	98.98
Gowthami	76.89	97.01	96.68	97.14
Priyanka	62.21	91.41	98.61	99.52

Table 5: Split Profile Algorithm

Font Size	12	14	16	19
Font Type	Segmentation efficiency			
Anupama	92.98	100.00	99.96	100.00
Gowthami	88.95	99.77	99.26	99.40
Priyanka	76.68	97.55	99.67	99.96

5. Conclusions

Topology and geometry is observed to be one of the important information of any script. Extensive study of the statistical properties with regard to topology is crucial while improving segmentation accuracy. In this paper an attempt is made towards this direction on popular cursive script Telugu. Profile function is considered for separating the linear region over non linear portions in the script line as well as touching syllables. A general approach (connected component approach) on these scripts is compared with the proposed Split Profile Algorithm. The highest performance of average segmentation efficiency with SPA is observed as 99.98%, 99.47% and 99.05% on ANUPAMA, GOWTHAMI and PRIYANKA fonts respectively. Experimental evaluation of the proposed algorithm on small font sizes is in progress. Extension of the proposed algorithm at the level of segmentation and classification with apreori knowledge is in progress.

References

[1] Richard G.Casey and Eric Lecolinet, "A survey of methods and strategies in character segmentation" IEEE Transctions on Pattern Analysis and machine Intelligence, vol. 18, No. 7, pp. 690–706, July 1996.

[2] K. Ohta, I. Kaneko, Y. Itamoto, and Y. Nishijima, "Character Segmentation of Address Reading/Letter Sorting Machine" for the Ministry of Posts and Telecommunications of Japan, NEC Research and Development, vol. 34, no. 2, pp. 248-256, Apr. 1993

[3] Y. Lu, "On the Segmentation of Touching Characters," lnt'l Conf. Document Analysis and Recognition, Tsukuba, Japan, pp. 440-443, Oct. 1993.

[4] M. Cesar and R. Shinghal, "Algorithm for Segmenting Handwritten Postal Codes," Int'l J. Man Machine Studies, vol. 33, no. 1, pp. 63-80, July 1990.

[5] G. Seni and E. Cohen, "External Word Segmentation of Off- Line Handwritten Text Lines," Pattern Recognition, vol. 27, no. 1, pp. 41-52, Jan. 1994.

[6] K.W. Gan, K.T. Lua, "A new approach to stroke and feature point extraction in Chinese character recognition". Pattern Recognition Letters, Vol. 12 , no. 6, pp 381-386 , June 1991

[7] B. B. Chaudhuri, U. Pal, "A complete printed Bangla OCR system" Pattern Recognition, vol.31, No. 5, pp. 531- 549, March 1998.

[8] Veena Bansal, R. M. K. Sinha, "Integrating knowledge sources in Devanagari text recognition system", IEEE Transactions on Systems, Man, and Cybernetics, Part A : Systems and Human, vol. 30, no. 4, pp 500-505, July 2000.

[9] M. K. Jindal, G. S. Lehal, R. K. Sharma, "A Study of Touching Characters in degraded Gurmukhi Script", World Academy of Science, Engineering and Technology, vol.4,pp 121-124, 2005

[10] U. Pal and Sagarika Datta, "Segmentation of Bangla Unconstrained Handwritten Text", Proceedings of the Seventh International Conference on Document Analysis and Recognition (ICDAR 2003),August 2003.

[11] Liana M. Lorigo, Venu Govindaraju, "Offline Arabic Handwriting Recognition: A Survey" IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 28, no. 5, pp. 712-724, May 2006

[12] Pratap Reddy L, Satyaprasad L, ASCS Sastry, "Middle Zone Component Extraction and Recognition of Telugu Document Image", Ninth International Conference on Document Analysis and Recognition,(ICDAR 2007), Vol 2, pp 584 – 588, September 2007.

[13] Pratap Reddy, L. Sastry, A.S.C. Rao, A.V.S. Venkat Rao, N., "Canonical syllable segmentation of Telugu document images", TENCON 2008 - IEEE Region 10 Conference , pp 1-5, Nov-2008.

[14] M. K. Jindal, R. K. Sharma, G. S. Lehal "Segmentation of touching characters in upper zone in printed Gurmukhi script", ACM Annual Bangalore Compute Conference, Article No.: 9, 2009 .

[15] L. Pratap Reddy, "A New Scheme for Information Interchange in Telugu through Computer Networks", Ph.D. thesis, Department Electronics and Communication, JNTUniversity, Hyderabad, INDIA, May 2001.

L.Pratap Reddy received his B.E degree in Electronics and Communication Engineering from Andhra University, India in 1985, M.Tech. degree in Electronic Instrumentation from Regional Engineering College, Warangal, India in 1988 and Ph.D. degree from Jawaharlal Nehru Technological University, Hyderabad, India in 2001. From 1988 to 1990 he was lecturer in Department of Electronics and Communication Engineering at Bangalore Institute of Technology, Bangalore, India, from 1991 to 2005 he was faculty member at JNTU College of Engineering, Kakinada, India. Since 2006 he is with Department of Electronics and Communication Engineering and presently he is Professor and Head of the department at JNTUH College of Engineering, Hyderabad, India. His current activity in research and development includes, apart from telecommunication engineering subjects, Image Processing, Pattern Recognition and Linguistic processing of Telugu language. He published more than 50 research publications in various National, Inter National conferences and Journals. He is active member in professional bodies like ISTE, IE, IETE, and CSI.

T.Ranga Babu received his B.E. degree in Electronics and Communication Engineering from University of Madras, India in 1992, M.S. degree in Electronics & Control Engineering from Birla Institute of Technology and Science, Pilani, India in 1999 and M.Tech. degree in Electronics and Communication Engineering with specialization in Digital Electronics and Communication Systems from Jawaharlal Nehru Technological University College of Engineering (Autonomous), Anantapur, India in 2004. From 1992 to 2003 he was worked as faculty member in the Department of Electronics and Communication Engineering at different educational institutions and from 2003 he is working as Assistant Professor in the Department of Electronics and Communication Engineering at R.V.R & J.C. College of Engineering, Guntur, India. Currently he is working in the area of Image Processing, Natural Language Processing, Pattern Recognition and pursuing his Ph.D. in Jawaharlal Nehru Technological University Hyderabad, India. He is a Life member of ISTE, IETE, CSI and IACSIT.

N.Venkata Rao received his B.E degree in Electronics and Communication Engineering from Bangalore University, India in 1985 and M.Tech. degree in Electronics and Communication Engineering with specialization in Instrumentation and Control Systems from Jawaharlal Nehru Technological University, Kakinada, India in . He has 22 years of teaching experience as Assistant Professor, Associate Professor, Professor and he is currently working as Professor in the Department of Electronics and Communication Engineering, Sri Vasavi Engineering College, Tadepalligudem, India. He has published 7 research papers in various National, Inter National conferences. Currently he is working in the area of Natural Language Processing, Image Processing, Pattern Recognition and pursuing his Ph.D. He is a Life member of ISTE and IETE.

B.Raveendra Babu received his M.S degree in Software Systems from Birla Institute of Technology and Science, Pilani, IN 1997, M.Tech degree in Computer Science and Engineering from Anna University, Chennai IN 2000, Ph.D degree from S.V.University, India in 1992. He has 26 years of teaching experience as Assistant Professor, Associate Professor, Professor and he is currently heading the Department of Computer Science and Engineering at RVR & JC College of Engineering, Guntur. India. He is presently Chairman, Board of studies for Computer Science and Engineering, Acharya Nagarjuna University, India. He has published more than 20 research publications in various National, Inter National conferences, proceedings and Journals. His research areas of interest include VLDB, Image Processing, Pattern analysis and Wavelets. He is a life member for ISTE and CSI, member for ACM and IEEE Computer society.