International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 10 Issue: 03 | March - 2026 SJIF Rating: 8.659 ISSN: 2582-3930

Analysis of Kannada Handwritten Vowels

PRASHANTH. M C

ABSTRACT

Language identification for handwritten document images is an open document analysis problem. Handwritten character recognition has received extensive attention in academic and production fields. The recognition system can either online or offline. There is a large demand for handwritten character recognition and handwritten documents. India is a multi-lingual and multi-script country, where eighteen official scripts are accepted and have over hundred regional languages. Handwritten character recognition for Indian language is an important problem where there is relatively little work has been done. Particularly difficult is the problem of recognition of kagunita- the compound characters resulting from the consonant and the vowel combination. To recognize a kagunita, we need to identify the vowel and the consonant present in the kagunita character image. Unlike the Latin script used for the English language, it does not have upper case or lowercase. It has only one case of writing. Moreover, each alphabet contains more curves than straight lines. Hence handwritten Kannada character recognition is a challenging task. We had taken hundred handwritten datasets of different users. Handwritten Kannada vowels are scan and converted into binary image and normalized into a size of 64 x 64 pixels. Scanned image is segmented, we get a extracted of single Kannada vowels and then stored in a database. By using median filter we removed noise from the image. Using morphological thinning function, we get thinned Kannada vowels of different handwritten data sets. Overlapping the similar characters we get a overlapped image. By using morphological erosion and dilation function we get a standard representative image. Using this representative image, we can test different samples to get recognition accuracy. The performance of the proposed method is experimentally evaluated and the promising results and findings are presented.

I.INTRODUCTION

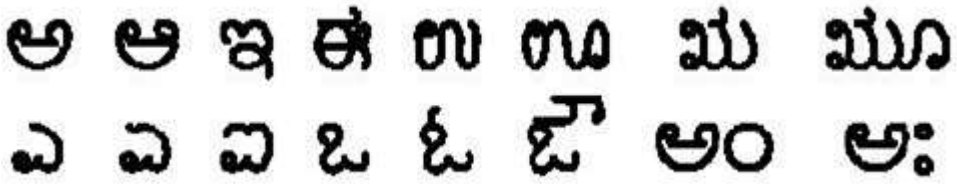
OCR systems are now available commercially at affordable cost and can be used to recognize many printed fonts. Even so, it is important to note that in some situations this commercial software is not always satisfactory and problems still exist with unusual character sets, fonts and with documents of poor quality. Unfortunately, the success of OCR could not extend to handwriting recognition due to large variability in people's handwriting styles. Handwritten Kannada characters are more complex for recognition than English characters due to many possible variations in order, number, direction and shape of the constituent strokes. The number of authors is attempts to make for developments of OCR system for Kannada characters. From the literature survey, it reveals that, handwritten character recognition of foreign languages like English, Chinese, Japanese, and Arabic are reaches to saturation point, but there is room for Indian languages like Kannada script. The Kannada character is complicated to segmentation and reorganization compare to English languages, because of Kannada character complex in nature. This has motivated us to design a recognition system for Kannada character recognition.

Handwriting processing is a domain in great expansion. The interest devoted to this field is not explained only by the exciting changes involved, but also the huge benefits that a system, designed in the context of a commercial application, could bring. Two classes of recognition systems are usually distinguished: Online systems for which handwriting data are captured during the writing process, which makes available the information on the ordering of the strokes, and offline systems for which recognition takes place on a static image captured once the writing process is over. In this work, we have chosen Offline handwritten character recognition for Kannada.

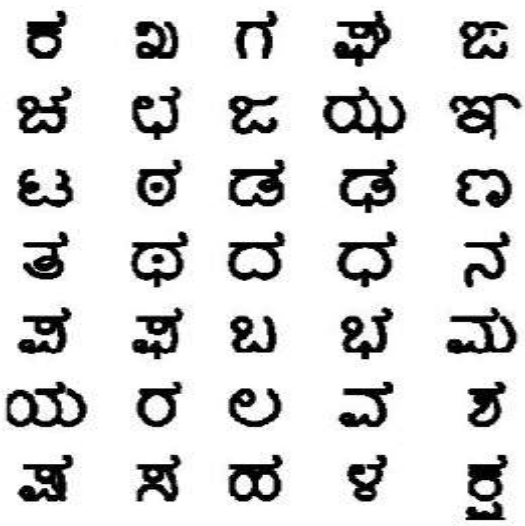
II.LITERATURE SURVEY

Kannada is the official language of the southern Indian state of Karnataka. Kannada is a Dravidian language spoken by about 44 million people in the Indian states of Karnataka, Andhra Pradesh, Tamil Nadu and Maharashtra. The Kannada alphabets were developed from the Kadamba and Chalukya scripts, descendents of Brahmi which were used between the 5th and 7th centuries AD. There are 13 Vowels (Swara), 2 Yogavaha and 34 Consonants (Vangana) in modern Kannada script. In this paper we constrain ourselves to recognition of handwritten Kannada vowels. Printed

Kannada vowels and their corresponding handwritten vowel samples. Fig 1 shows handwritten vowel samples and Fig.2 shows handwritten kagunita samples respectively,



Kannada vowels

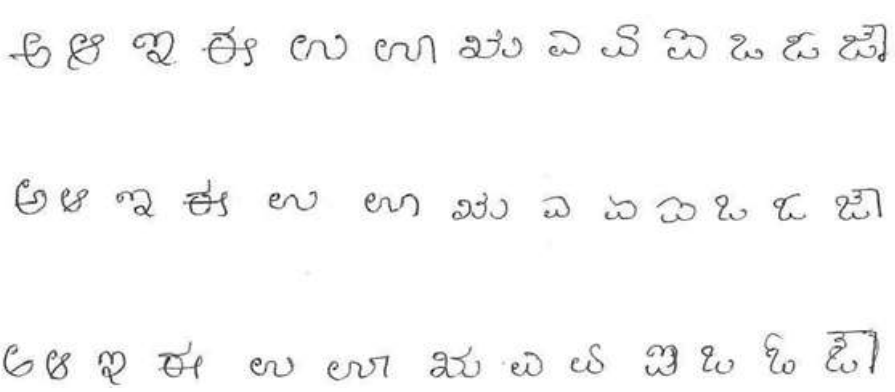


Kannada kagunita characters

The standard database for Kannada handwritten vowels character is not available; therefore, our own database created. Data collected from different professionals belonging to schools, colleges, and commercial sectors. We collected 1625 images from 125 writers are considered for the experimentation purpose. A flat bed scanner was used for digitization. Digitized images are in gray tone with 300 dpi and stored as BMP format. We have used global threshold binarizing algorithm to convert them to two-tone (0 and 1) images (Here '1'represents object point and'0'represents background point). Scanned isolated Vowel images often contain noise that arises due to printer, scanner, print quality, etc. therefore, it is necessary to filter this noise before we process the recognition of Kannada vowels. The noise removed by using median filter and scanning artifacts are removed by using morphological opening operation.

III. PROPOSED SYSTEM

The page of text is scanned through a flat bed scanner at 300dpi resolution and binarized using a global threshold computed automatically based on a specific image. A binary image is obtained by considering the character as ON pixels and the background as OFF pixels. The binarized image is processed to remove any skew so that text lines are aligned horizontally in the image. The standardized database for Kannada handwritten characters is not available therefore; our own database is created. We have collected data from different persons. Each person has been asked to write each sample of Kannada vowels. Noise removal is performed by employing morphological area opening operation. The samples collected are as shown below:



3. 1.1SEGMENTATION

Segmentation step contains line segmentation, word segmentation and character segmentation. Methods for character segmentations are based on i) white space and pitch ii) projection analysis and iii) connected component labeling.

The segmented image is divided into a single labeled character and stored in a particular database. The labeled characters are as shown below:

Char_1.bmp Char_2.bmp Char_3.bmp Char_4.bmp Char_5.bmp

Char_6.bmp Char_7.bmp Char_8.bmp Char_9.bmp Char_10.bmp

Char_11.bmp Char_12.bmp Char_13.bmp

3.1.2. THINNING:

Thinning is the process to extract and apply additional constraints on the pixel elements that need to be preserved such that a linear structure of the input image will be recaptured without destroying its connectivity. In the context of image processing, outline from a pattern until all the lines or curves are of unit width or of single pixel wide. Thus, the reduced pattern is known as the *skeleton*. Skeletonization or thinning is a very important preprocessing step in pattern Analysis. One major advantage of thinning is to reduction memory space required for storing the essential structural information presented in a pattern moreover it simplifies the data structure required in pattern analysis.

3.1.3 OVERLAPPING:

The overlapped images are used for finding the standard representative for the collection of different handwritten Kannada vowels of different users. We have taken 5 datasets of similar character of different users. The collected images are then overlapped. The images that are collected will not be in a correct manner as they are of different users. On further process we get a single representative for different similar vowels of different data sets.

Morphological operation:

Morphology is a broad set of image processing operations that process images based on shapes. Morphological operations apply a structuring element to an input image, creating an output image of the same size. In a morphological operation, the value of each pixel in the output image is based on a comparison of the corresponding pixel in the input image with its neighbors. By choosing the size and shape of the neighborhood, you can construct a morphological operation that is sensitive to specific shapes in the input image.

The most basic morphological operations are dilation and erosion. Dilation adds pixels to the boundaries of objects in an image, while erosion removes pixels on object boundaries. The number of pixels added or removed from the objects in an image depends on the size and shape of the structuring element used to process the image. In the morphological dilation and erosion operations, the state of any given pixel in the output image is determined by applying a rule to the corresponding pixel and its neighbors in the input image.

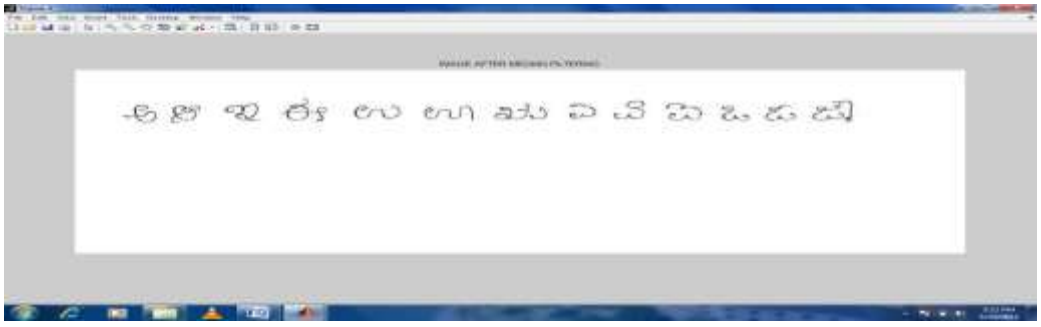
The rule used to process the pixels defines the operation as dilation or an erosion. This table lists the rules for both dilation and erosion. In dilation the value of the output pixel is the *maximum* value of all the pixels in the input pixel's neighborhood. In a binary image, if any of the pixels is set to the value 1, the output pixel is set to 1. In erosion the value of the output pixel is the *minimum* value of all the pixels in the input pixel's neighborhood. In a binary image, if any of the pixels is set to 0, the output pixel is set to 0. Opening operation is the process of dilation followed by erosion and closing operation is the process of erosion followed by dilation.

3.1.4 REPRASENTATIVE OF IMAGES:

After applying all the morphological operations, we get a single standard image, which is considered as representative character. This representative character is further used for matching purpose.

IV. EXPERIMENT AND RESULT

Input images:



Segmented images:

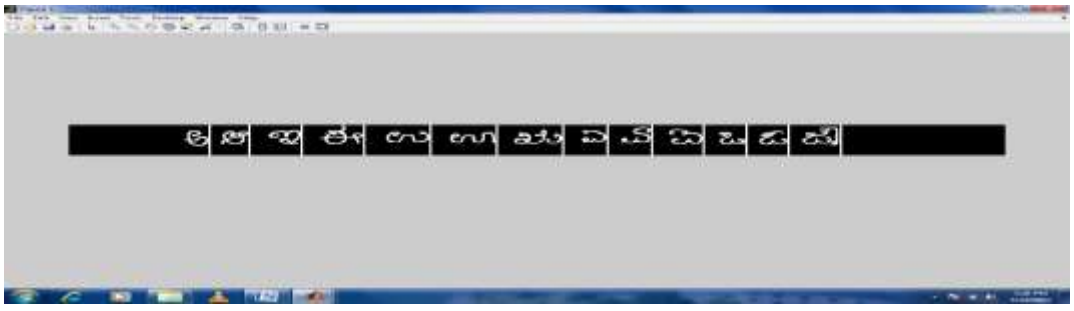


Image after thinning

Overlapped image

Dilated image

Representative image

CONCLUSION

For independent recognition of handwritten Kannada vowels only 13 global spatial features are considered. The proposed directional spatial features showed quite encouraging performance with respect to handwritten Kannada vowels. The aim of this work is to build the standard representatives for handwritten Kannada vowels which is later used for matching purpose. The work can be extended for recognition of handwritten Kannada characters, using these representative images. By using these representative images, we can match the different handwritten Kannada characters to get recognition accuracy.

REFERENCES

1. N. Sharma, U. Pal, F. Kimura, "Recognition of Handwritten Kannada Numerals", 9th International Conference on Information Technology (ICIT'06), ICIT, pp. 133-136.
2. Rajput, G.G., Mallikarjun Hangarge, "Recognition ofIsolated Handwritten Kannada Numeral based on Imagefusion method", PREMI07, LNCS, Vol. 4815, SpringerKolkata, pp153-160, 2007.
3. B.V. Dhandra, R.G.Benne and Mallikargun Hangargi,"Isolated Handwritten Kannada Numeral recognition basedon Template matching", IEEE-ACVIT -07, pp.1276-1282,Dec-2007.
4. Dinesh Acharya U, N. V. Subba Reddy, Krishna Moorth Makkithaja, "Multilevel Classifiers in the Recognition of Handwritten Kannada Numerals", Proceedings of World Academy of Science Eng and Technology, vol. 32, Aug 2008.
5. Jomy John, Pramod K. V, Kannan Balakrishnan , "Handwritten Character Recognition of South IndianScripts: A Review",National Conference on Indian Language Computing, Kochi, Feb 19-20, 2011
6. S. K. Sangame, R.J. Ramteke and Yogesh V. G, "Recognition Of Isolated Handwritten Kannada Characters Using Invariant Moments and Chain Code", World Journal of Science and Technology 2011
7. Sangame S.K., Ramteke R.J., Rajkumar Benne, "Recognition of isolated handwritten Kannada vowels", Advances in Computational Research, ISSN: 0975–3273, Volume 1, Issue 2, 2009, pp-52-55
8. Anil K. Jain, Douglass Zonker, "Representation and Recognitionof handwritten Digits using Deformable Templates", IEEE,Pattern analysis and machine intelligence, vol.19, no-12, 1997.
9. T. V. Aswin and P. S. Sastry (2002). "A font and size independent OCR system for printed Kannada documents using Support Vector Machines", Sadhana Vol.27 part 1. pp.35 – 58.
10. P.Nagabhushan, S.A.Angadi, B.S.Anami, "A fuzzy statistical approach of Kannada Vowel Recognition based on Invariant Moments", Proceedings of NCDAR-2003, Mandy, Karnataka, India, pp275-285, 2003.