KANNADA TEXT EXTRACTION FROM IMAGES AND VIDEOS FORVISION IMPAIRED PERSONS

Keshava Prasanna[1], Ramakhanth Kumar P[2], Thungamani.M[3], Manohar Koli[4]
[1, 3] Research Assistant, Tumkur University,Tumkur, India.
[2]Professor and HOD, R.V. College of Engineering,Bangalore, India.
[4] Research Scholar, Tumkur University,Tumkur, India.

ABSTRACT

We propose a system that reads the Kannada text encountered in natural scenes with the aim to provide assistance to the visually impaired persons of Karnataka state. This paper describes the system design and standard deviation based Kannada text extraction method. The proposed system contain three main stages text extraction, text recognition and speech synthesis. This paper concentrated on text extraction from images/videos. In this paper: an efficient algorithm which can automatically detect, localize and extract Kannada text from images (and digital videos) with complex backgrounds is presented. The proposed approach is based on the application of a color reduction technique, a standard deviation base method for edge detection, and the localization of text regions using new connected component properties. The outputs of the algorithm are text boxes with a simple background, ready tobe fed into an OCR engine for subsequent character recognition. Our proposal is robust with respect to different font sizes, font colors, orientation, alignment and background complexities. The performance of the approach is demonstrated by presenting promising experimental results for a set of images taken from different types of video sequences.

KEYWORDS: *SVM, OCR, AMA, CCD Camera, Speech synthesis.*

I. INTRODUCTION

Recent studies in the field of computer vision and pattern recognition show a greatamount of interest in content retrieval from images and videos. Text embedded in images contains large quantities of useful semantic information, which can be used to fully understand images. In this world maximum objects can be analyzed and identified by reading the text information present on that object

Automatic detection and extraction of text in images have been used in many applications such as document retrieving; a document image analysis system is one that can handle text documents in Kannada, which is the official language of the south Indian state of Karnataka. The input to the system is the scanned image of a page of Kannada text. The output is an editable computer file containing the information in the page. The system is designed to be independent of the size of characters in the document and hence can be used with any kind of document in Kannada. The task of separating lines and words in the document is fairly independent of the script and hence can be achieved with standard techniques. However, due to the peculiarities of the Kannada script, we make use of a novel segmentation scheme whereby words are first segmented to a sub-character level, the individual pieces are recognized and these are then put together to effect recognition of individual aksharas or characters. The Kannada alphabet (50) is classified into two main categories 16 Vowels and 34 consonants as shown in figure 1 and figure 2 words in Kannada are composed of aksharas[13] which are analogues to characters in English words. We use a novel feature vector to characterize each segment and employ a classifier based on the recently developed concept of Support Vector Machines (SVM)[14], address block location, content based image/video indexing, mobile robot navigation to detect text based landmarks, vehicle license detection / recognition, object identification, etc. The blind peoples are almost dependent on others. They cannot read and analyze objects their own. In making blind peoples readable extraction of textual information plays very vital role. Textual information extraction helps blind peoples in various aspects such as identifying the objects,identifying and self-reading of the text books, newspapers, current and electric bills, sign boards, personal letters etc.

OCR systems available for handling English documents, with reasonable levels of accuracy. (Such systems are also available for many European languages as well as some of the Asian languages such as Japanese, Chinese etc.) However, there are not many reported efforts at developing OCR systems for Indian languages. The work reported in this project is motivated by the fact that there are no reported efforts at developing document analysis systems for the south Indian language, Kannada. In most OCR [13] systems the final recognition accuracy is always higher than the raw character recognition accuracy. For obtaining higher recognition accuracy, language-specific information such as co-occurrence frequencies of letters, a word corpus [14], a rudimentary model of the grammar etc. are used. This allows the system to automatically correct many of the errors made by the OCR subsystem. In our current implementation, we have not incorporated any such post-processing. The main reason is that, at present we do not have a word corpus for Kannada. Even with a word corpus the task is still difficult because of the highly in flexional nature of Kannada grammar. The grammar also allows for combinations of two or more words. Even though these follow well-defined rules of grammar, the number of rules is large and incorporating them into a good spell-checking application for Kannada is a challenging task.

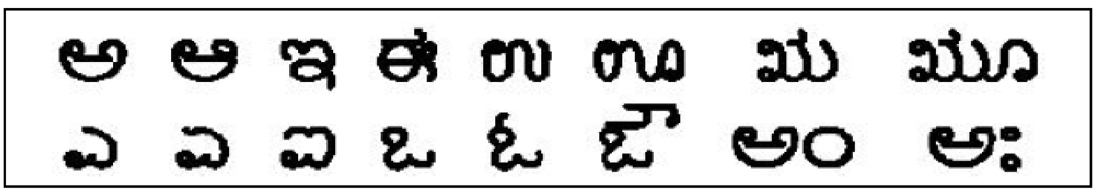

Figure 1: Vowels in Kannada [13]

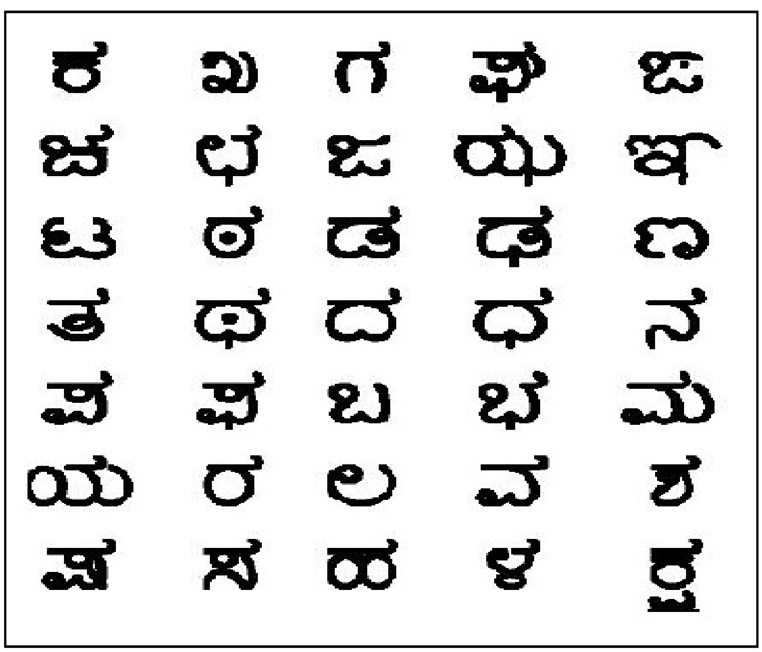

Figure 2: Consonants in Kannada [13]

II. RELATED WORK

Due to the variety of font size, style, orientation, and alignment as well as the complexity of the background, designing a robust general algorithm, which can effectively detect and extract text from both types of images, which is full of challenges. Various methods have been proposed in the past for detection and localization of text in images and videos. These approaches take into consideration different properties related to text in an image such as color, intensity, connected – components, edges etc. These properties are used to distinguish text regions from their background and / or other regions within the image.

[1]. **Xiaoqing Liu et al [1, 2]:**The algorithm proposed is based on edge density, strength and orientation. The input image is first pre-processed to remove any noise if present. Then horizontal, vertical and diagonal edges are identified with the help of Gaussian kernels and based on edge density, strength and orientation text regions are identified. This approach is based on the fact that edges are most reliable features of text.

[2]. **JulindaGllavata et al [3]:**The algorithm proposed is based on connected component based method. This approach is based on the fact that text is collection of characters usually comes in a group. The input image is first pre-processed to remove any noise if present. Then an input image is converted from RGB to YUV model and Y-channel is processed, horizontal and vertical projections are calculated. Then with the help of horizontal and vertical threshold text regions are identified.

[3]. **Wang and Kangas et al [4]:**The algorithm proposed is based on color clustering. The input image is first pre-processed to remove any noise if present. Then the image is grouped into different color layers and gray component. This approach utilities the fact that usually the color data in text characters is different from the color data in the background. The potential text regions are localized using connected component based heuristics from these layers. Also an aligning and merging analysis (AMA) method is used in which each row and column value is analyzed. The experiments conducted show that the algorithm is robust in locating mostly Chinese and English characters in images; sometimes false alarms occurred due to uneven lighting or reflection in the test images.

[4]. **K.C. Kim et al [5]:**The text detection algorithm is also based on color continuity. In addition it also uses multi-resolution wavelet transforms and combines low as well as high level image feature for text region extraction, which is a hierarchical feature combination method to implement text extraction in natural scenes. However, authors admit that this method could not handle large text very well due to the use of local features that represents only local variations of images blocks.

[5]. **Victor Wu et al [6]:**The text finder algorithm proposed is based on the frequency, orientation and spacing of text within an image. Texture based segmentation is used to distinguish text from its background. Further a bottom – up 'chip generation' process is carried out which uses the spatial cohesion property of text strokes and edges. The results show that the algorithm is robust in most of the cases, expect for every small text characters that are not properly detected. Also in case of low contrast in image, misclassifications occur in the texture segmentation.

[6].**Qixiang Ye et al[7,8]:**The approach used in [7, 8] utilizes a support vector machines (SVM) classifier to segment text from non – text in an image or video frame. Initially text is detected in multi scale images using non edge based techniques, morphological operations and projection profiles of the image. These detected text region are then verified using wavelet features and SVM. The algorithm is robust with respect to variance in color and size of font as well as language.

[7].**SanjeevKunteet al [11]:**The Kannada character detection algorithm is based on Neural Network concept. The input image is first pre-processed to remove any noise if present. Neural classifiers are effectively used for the classification of characters based on moment features.

[8]. **Te´ofilo E. de Campos et al [12]:**The character detection algorithm is based on SVM. It evaluate six different shape and edge based features, such as Shape Context, Geometric Blur and SIFT, but also features used for representing texture, such as filter responses, patches and Spin Images.

III. PROPOSED WORK

In this Proposed Work, a robust system for automatically extracting Kannada text appearing in images and videos with complex background is presented. Standard deviation based edge detection is performed to detect edges present in all directions.
The identification of the used script can help in improving the segmentation results and in increasing the accuracy of OCR by choosing the appropriate algorithms. Thus, a novel technique for Kannada script recognition in complex images will be presented. Figure 3 shows the general configuration of proposed system. The building elements are the TIE, the CCD-camera and the voice synthesizer.

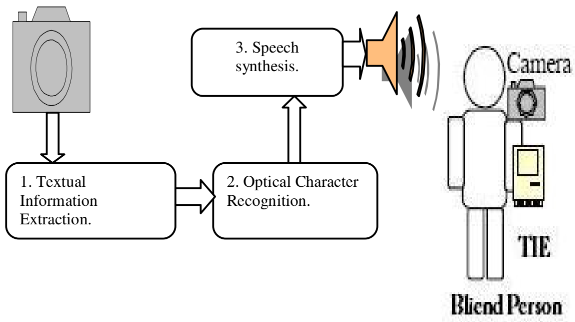

Figure3. System configuration (walk-around mode)

Proposed system contains three main steps after acquiring image with the help of CC-camera.

1. Textual information Extraction.
2. Optical character Recognition.
3. Speech Synthesis.

As the first step in the development of this system, simple standard deviation based method for Kannada text detection method is proposed.
The different steps of our approaches are asfollows.

1. Image preprocessing.
2. Calculate Standard Deviation of Image.
3. Detection of Text Regions.

Step 1: Image Preprocessing. If the image data is not represented in HSV color space, it is converted to this color space by means of appropriate transformations. Our system only uses the intensity dataFigure 5 (V channel of HSV) during further processing. A median filtering operation is then applied on theV (intensity) band to reduce noise before a contrast-limited Adaptive Histogram Equalization is applied for contrast enhancement.

Figure4.Original Image

Figure5. V channel

Step 2: **Edge Detection.** This step focuses the attention to areas where text may occur. We employ a simple method for converting the gray-level image into an edge image.
Our algorithm is based on the fact that the characters processes high standard deviation compared to their local neighbors.

$$\mathbf{Std(x)=1/\,(N\text{-}1)\ \sum(V\ (i)\text{-}\mu(x))\ ^{2}}\ \text{……………}\ \mathbf{(1)}$$

i€W(x)

Where x is a set of all pixels in a sub-window W(x), N is a number of pixels in W(x), μ(x)is mean value of V(i)and i €W(x). A window size of 3X7 pixels was used in this step.

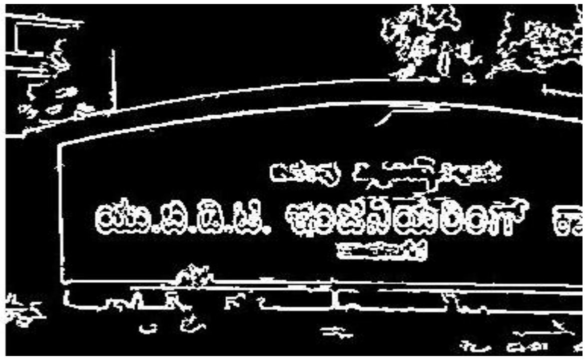

Figure6. Standard Deviation Image

Step 3:Detection of Text Regions. Steps used in Kannada Text location are different compared to English text localizationbecause features of both texts are different. Height and width ratio, Centroid difference and orientation calculations used in English text extraction are not suitable for Kannada text extraction.

Normally, text embedded in an image appears in clusters, i.e., it is arranged compactly. Thus, characteristics of clustering can be used to localize text regions. Since the intensity of the feature map represents the possibility of text, a simple global thresholding can be employed to highlight those with high text possibility regions resulting in a binary image. A morphological dilation operator can easily connect the very close regions together while leaving those whose positions are far away to each other isolated. In our proposed method, we use a morphological dilation operator with a 7×7 square structuring element to the previous obtained binary image to get joint areas referred to as text blobs. Two constraints are used to filter out those blobs which do not contain text [1 ,2], where the first constraint is used to filter out all the very small isolated blobs whereas the second constraint filters out those blobs whose widths are much smaller than corresponding heights. The retaining blobs are enclosed in boundary boxes. Four pairs of coordinates of the boundary boxes are determined by the maximum and minimum coordinates of the top, bottom, left and right points of the corresponding blobs. In order to avoid missing those character pixels which lie near or outside of the initial boundary, width and height of the boundary box are padded by small amounts as in Figure 7.

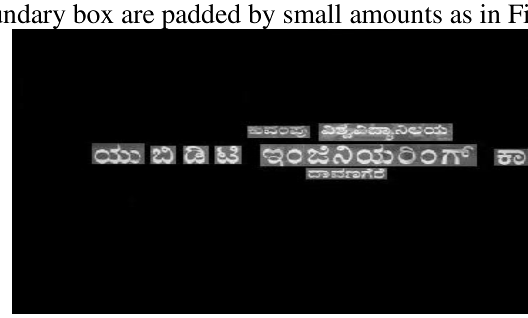

Figure 7.Final results for the example given in Figure. 5

IV. EXPERIMENTAL EVALUATION

The proposed approach has been evaluated using datasets containing different types of images Figure 8,9,10. The whole test data consists of 300images where 100 of them were extractedfrom various MPEG videos

Figure 8. Results of House Boards

Figure 9. Results of Wall Boards

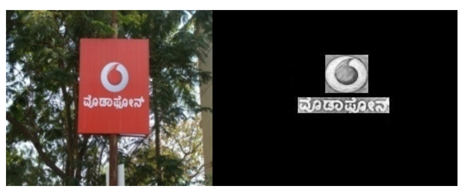

Figure 10. Results of Banners.

The precision and recall rates (Equations (2) and (3)), have been computed based on the number of correctly detected words in an image in order to further evaluated the efficiency and robustness. The precision rate is defined as the ration of correctly detected words to the sum of correctly detected words plus false positive. False positive are those regions in the image, which are actually not characters of text, but have detected by the algorithm as text regions.

$$\text{Precision Rate} = \frac{\text{Correctly detected words}}{\text{Correctly detected words + False Positives}} * 100\% \qquad \text{............ (2)}$$

The Recall rate is defined as the ratio of correctly detected Words to the sum of correctly detected words plus false negatives. False negatives are those regions in the image, which are actually text characters, but have been not detected by the algorithm.

$$\text{RecallRate} = \frac{\text{Correctly detected words}}{\text{Correctly detected words + False Negatives}} * 100\% \quad \text{...... (3)}$$

Table 1. Analysis of precession rate and recall rate

TEST DATA	NO OF IMAGES	PRECISSION RATE	RECALL RATE
FROM IMAGES	200	92.2	88.6
FROM VIDEOS	100	78.8	80.2
TOTAL	300	80.5	84.4

V. CONCLUSION

In this paper, Text extraction is a critical step as it sets up the quality of the final recognition result. Itaims at segmenting text from background, meaning isolating text pixels from those ofbackground. we presented the design of a Kannada scene-text detection module for visually impaired persons. As the first step in the development of this system, simple standard deviation based method for Kannada text detection have been implemented and evaluated.

VI. FUTURE WORK

The main challenge is to design a system as versatile as possible to handle all variability in daily life, meaning variable targets with unknown layout, scene text, several characterfonts and sizes and variability in imaging conditions with uneven lighting, shadowing and aliasing. Variation in Font style, size, Orientation, alignment & complexity ofbackground makes the text segmentation as a challenging task in text extraction.

We plan to employ an OCR system to check the recognition performance for the text images produced by the proposed algorithm andalso employ a Speech Synthesizer to spell the recognized text to vision impaired persons. Finally, work will focus on new methods for extracting Kannada text characters with higher accuracy.

REFERENCES

[1].Xiaoqing Liu and JagathSamarabandu , An Edge-based text region extraction algorithm for Indoor mobile robot navigation, Proceedings of the IEEE, July 2005.

[2].Xiaoqing Liu and JagathSamarabandu, Multiscale edge-based Text extraction from Complex images, IEEE, 2006.

[3].JulindaGllavata, Ralph Ewerth and Bernd Freisleben, A Robust algorithm for Text detection in images , Proceedings of the 3 international symposium on Image and Signal Processing and Analysis, 2003.

[4].Kongqiao Wang and Jari A. Kangas, Character location in scene images from digital camera, the journal of the Pattern Recognition society, March 2003.

[5]K.C. Kim, H.R. Byun, Y.J. Song, Y.W. Choi, S.Y. Chi, K.K. Kim and Y.K Chung, Scene Text Extraction in Natural Scene Images using Hierarchical FeatureCombining and verification , Proceedings of the 17International Conference on Pattern Recognition (ICPR ’04), IEEE.

[6] Victor Wu, RaghavanManmatha, and Edward M.Riseman,Text Finder: An Automatic System to Detect and Recognize Text in Images, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 21, No. 11, November 1999.

[7]Qixiang Ye, Qingming Huang, Wen Gao and DebinZhao,Fast and Robust text detection in images and video frames, Image and Vision Computing 23, 2005.

[8]Qixiang Ye, Wen Gao, Weiqiang Wang and Wei Zeng,A Robust Text Detection Algorithm in Images and Video Frames, IEEE, 2003.

[9]Rainer Lienhart and Axel Wernicke, Localizing and Segmenting Text in Images and Videos, IEEE Transactions on Circuits and Systems for Video Technology, Vol.12,No.4, April 2002.

[10]Keechul Jung, Kwang in Kim and Anil K. Jain, Text information extraction in images and video: a survey, the journal of the Pattern Recognition society, 2004.

[11]SanjeevKunte and R D Sudhaker Samuel, A simple and efficient optical character recognition systemfor basic symbols in printed Kannada text.

[12]Nobuo Ezaki, Marius Bulacu, Lambert Schomaker, Text Detection from Natural Scene Images: Towards a System for Visually Impaired Persons, Proc. of 17th Int. Conf. on Pattern Recognition (ICPR 2004), IEEE Computer Society, 2004, pp. 683-686, vol. II, 23-26 August, Cambridge, UK.

[13]T V Ashwin and P S Sastry, "A font and size-independent OCR system for printed Kannada documents using support vector machines", S¯adhan¯a Vol. 27, Part 1, February 2002, pp. 35–58. © Printed in India
[14] Department of Computer Sciences, University of Texas at Austin, Support Vector Machines, www.cs.utexas.edu/~mooney/cs391L/svm.ppt,The VC/SRM/SVM Bible:

Keshava Prasanna received B.E from Bangalore University and M.Tech in Information and Technology in the year 2005.He has experience of around 13 years in academics. Currently pursuing Ph.D. and working as Research Assistant in Tumkur University, Tumkur. Life membership in Indian Society for Technical Education (ISTE).

Ramakanth Kumar P completed his Ph.D. from Mangalore University in the area of Pattern Recognition. He has experience of around 16 years in Academics and Industry. His areas of interest are Image Processing, Pattern Recognition and Natural Language Processing. He has to his credits 03 National Journals, 15 International Journals, and 20 Conferences. He is a member of the Computer Society of India (CSI) and a life memember of Indian Society for Technical Education (ISTE). He has completed number of research and consultancy projects for DRDO.

Thungamani. M received B.E from Visvesvaraya Technological University and M.Tech in Computer Science and Engineering in the year 2007.She has experience of around 08 years in academics. Currently pursuing Ph.D. and working as Research Assistant in Tumkur University, Tumkur. Life membership in Indian Society for Technical Education (MISTE) The Institution of Electronics and Telecommunication Engineers (IETE).

Manhoar Koli received B.E from Visvesvaraya Technological University and M.Tech in Computer Science and Engineering.He has experience of around 08 years in academics. Currently pursuing Ph.D. as Research Scholar in Tumkur University, Tumkur.