Advances in Computer Science and Information Technology (ACSIT)
p-ISSN: 2393-9907; e-ISSN: 2393-9915; Volume 3, Issue 5; July-September, 2016, pp. 436-440

http://www.krishisanskriti.org/Publication.html

Mobile Phone Based Malayalam Text Detection and Localization from Natural Scene Images

Anit V. Manjaly[1], Benson C. C[2] and B. Shanmuga Priya[3]

[1,3]CMS College of Science and Commerce, Coimbatore, Tamilnadu
[2]University of Calicut, Kerala
E-mail: [1]manjaly.anit@gmail.com, [2]bensonuoc@gmail.com, [3]shanmugapriyaa.b@gmail.com

Abstract—*Text information detection and extraction from the scene images is a prominent research problem in computer vision. The process of Text Information Extraction (TIE) involves key phases including pre-processing, text detection, text localization, and segmentation. Variations in text may occur because of differences in size, style, orientation, alignment of text, low image contrast and complex backgrounds. It creates difficulties during the extraction of the text. In this paper we propose a very simple and efficient method for the detection and localization of the Malayalam text from the natural scene images captured by using mobile phone camera. The proposed work is based on the technology of mathematical morphology. Performance of the algorithm is evaluated using the precision rate and recall rate .Experimental results show that the proposed algorithm efficiently extracts the Malayalam texts from the natural scene images.*

Keywords*: Text localization, Text extraction, Mathematical Morphology, Precision rate, Recall rate.*

1. INTRODUCTION

In all areas of life the visual information plays a vital role. Today, much of this information is represented and processed digitally. Digital image processing (DIP) is ubiquitous with application ranging from photography to printing, television to tomography and from robotics to remote sensing. Recent studies in the field of pattern recognition and computer vision shows an immense amount of interest in content retrieval from images and videos. This content can be in the form of shape, objects, texture, color and the relationships between them. The semantic information in an image can be useful for content based image retrieval, indexing and classification purposes [1, 2].

Text Information Extraction (TIE) has become a major research area over the last few years due to its applications in text-based image indexing, content based image indexing, text-based image search and text translation within the signboard images [3]. Text extraction involves detection, localization, tracking, extraction and enhancement of text in an image.

Three major categories of images with text information are Document images, Scene text images and Caption text images. Document images are in the form of scanned image or photographs of any type of documents, book covers, CD covers etc. Text images are further classified as scene text images and caption text images. Natural images that contain text are called scene text images. The caption text image is the one in which text is inserted or superimposed in the image. The Fig. 1 (a), (b) and (c) shows the samples of these three types of images.

(a) Document Image b) Caption Image c) Scene Image

Fig. 1: Different types of text images

Text contained in an image is very difficult to extract due to its variable size, alignment, inter-character distance, color, motion, edge and compression [1]. The algorithm which successfully works on one language may not be suitable for another language. Localization of texts in some languages such as English is comparatively easier than the Indian languages as almost all English characters contains vertical or horizontal lines.

There are lots of research works reported in the text localization of English and other languages [4-9].No promising research result is reported for Malayalam text localization and is still an open problem. It is due to the curved nature and other complexities of the Malayalam language. In this paper, we attempt to implement Malayalam text localization from scene images through a multistage process. The operations performed in this paper are based on the mathematical morphology. Mathematical morphology is a shape based tool for extracting information from an image [10].

This paper is ordered as follows. Next section discusses the related works in text localization from scene images. Third section explains the materials and methods used. Results and evaluations are included in the fourth section. Finally, conclusions are described in the last section.

2. RELATED WORKS

Automatic identification and extraction of text regions from images is regarded as an important and challenging research area for the scientists and engineers. Many promising research results have been published for identification, extraction and recognition of text information present in the images and videos. Large amount of information are embedded in natural scenes and are often required to be automatically recognized and processed. This requires automatic detection, segmentation and recognition of visual text entities. The system should be capable of capturing images, automatic detection and recognition of texts.

Many methods are reported in localization of text from scene images in English and other languages. They are region-based methods [11-14], connected component based methods [15-17], edge based methods [18-20], and texture-based methods [21-23].Here we adopted one of the powerful tools of image processing called mathematical morphology for the text localization. Malayalam text localization can be efficiently done with morphology based methods.

Wonder Alexander Luz Alves and Ronaldo Fumio Hashimoto proposed a classical method for extracting text from scene images [4]. Their classification contains three steps including text region candidate extraction, text region candidate selection and text region localisation. Thomas Retornaz and Beatriz Marcotegui proposed another method for scene text localization based on the ultimate opening [5].Rodolfo P. dos Santos et al., proposed another method for text line segmentation based on Morphology and Histogram production [6]. It is a multistage process with eight stages such as feature extraction, Y histogram projection, text line separation, false line exclusion, line region recovery, X histogram projection, false word exclusion and text selection.

Lee and Kankanhalli applied a CC-based method to the detection and recognition of text on cargo containers, which can have irregular lighting conditions and characters with different shapes and sizes [24]. Jui-Chen Wuet.al proposed another classical method for text line extraction based on morphology [7]. They applied opening and closing operator separately on the image and find the difference between them. A closing operator is applied to the resultant image. After applying one of the thresholding methods, they found the connected components from the binary image. Zhong et al. used a Connected Component based method, which uses color reduction. The peaks in a color histogram in the RGB color space is used by them for quantizing the color space [25].

3. MATERIALS AND METHODS

The input is in the form of an image for the text information extraction system. The images can be in color or gray scale, un-compressed or compressed. Fig. 2 shows the stages in text information extraction.

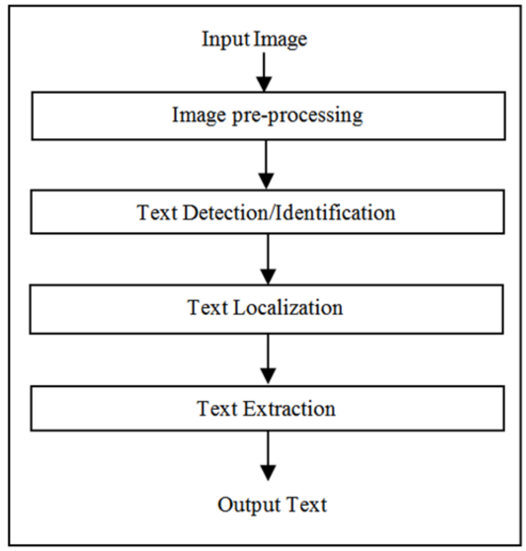

Fig. 2: Stages in Text Information Extraction

3.1. Data Acquisition and Image Database Creation

The availability of a dataset that contains adequate number of appropriate samples is a critical part of any pattern recognition research. The accuracy of these types of research highly depends on the presence of moderately large and representative samples of the database used. In the present work, a large number of sample images with text information are highly essential for the conduct of experiments to analyse the performance of proposed methods.The database used for the study consists of 100 scene images of name boards, posters, book/CD covers, traffic sign boards and documents. All the images are captured using Micromax Canvas A1 Android phone.

3. 2. Mathematical Morphology

Mathematical morphology tool for extracts image components that are useful in the representation and description of regions or shapes such as boundaries, skeletons, and the convex hull. Erosion and dilation are the two fundamental operations in mathematical morphology. Many other operations can be derived from the erosion and dilation. Opening and closing are the two such derived operations. Erosion is used to shrink the components of an image and dilation is used to expand the components.

Let A is a set (binary image) and S be the structuring element in Z^2. Then the erosion of A by the structuring element B, denoted by AΘS is the set of all pixels in which S placed at that pixel is contained within A. In another way we can say that the Erosion of A by S is the set of all points z such that S, translated by z, is contained in A.

$$A\Theta S = \{z \mid (S)\, z \subseteq A\} \quad (1)$$

Erosion shrinks or thins the object. It is also used to exclude the small unwanted objects from the given image.

Dilation of A by S in Z^2, denoted by A⊕S

$$A\oplus S= \{z \mid (\hat{S})\, z \cap A \neq \emptyset\} \quad (2)$$

Dilation of A by the structuring element S is the set of all displacements z such that A overlap by at least one element. Dilation is used for growing or thickening of an image. It is used to highlight the small objects in an image.

Opening and closing are the other two operations in the mathematical morphology. Opening operation makes the contour of an object smoothened and break the narrow lines. Like that closing operation also smoothens the contours but it usually eliminates discontinuity and small gaps between the objects. Opening operation is denoted by A○S and closing operation is denoted by A●S, where A is an image and S is the structuring element. In opening erosion is performed followed by dilation and in closing dilation is performed followed by erosion.

$$\text{Opening, } A\circ S = (A\Theta S) \oplus S \quad (3)$$

$$\text{Closing, } A\bullet S = (A\oplus S)\, \Theta S \quad (4)$$

Another important idea used in the mathematical morphology is the morphological reconstruction. Using morphological reconstruction we can extract some connected components of an image, which are marked in the image. Morphological reconstruction can be implemented by applying series of dilation operations, until the edge of the marker image fits under a second image, which is called the mask image.

3.3. Steps involved in the Text Localization

Localization of Malayalam texts from the scene images involved the following steps.

Step 1: Convert the color image, *C* into Gray scale image, *G* as a pre-processing stage.

$$G = \text{Gray}\,(C) \quad (5)$$

Step 2: Apply morphological opening and closing to the image using Disc structuring element $SE_1$and store the results in the separate variables, *OP* and *CL*.

$$OP=G\circ SE_1 \quad (6)$$

$$CL=G\bullet SE_1 \quad (7)$$

Step 3: Calculate the morphological gradient by subtracting opening image from the closed image.

$$GR=CL-OP \quad (8)$$

Step 4: Binarize the resultant image using Otsu's binarization method.

$$BIN=\text{Otsu}\,(GR) \quad (9)$$

Step 5: Apply morphological dilation to the resultant image using 3x3 structuring element SE_2.

$$DIL=BIN \oplus SE_2 \quad (10)$$

Step 6: Avoid non-text regions by calculating the aspect ratio and the volume of the connected components.

Step 7: Compute the boundary around each word and draw the bounding box.

Step 8: Superimpose the result to the input image.

Automatic extraction of Malayalam texts from scene images is a complicated task due to its particular characteristics. Here we used morphological operations for the text extraction. As a pre-processing stage we converted the image into gray scale image. Then morphological opening and closing are separately applied to the gray scale image using disc structuring element. Here we used disc structuring element since the repeated experiments shows that it provides more accurate results than the other structuring elements. Then we found the morphological gradient of the image by subtracting opened image from the closed image. After that binarize the resultant image using Otsu's method [26].Otsu's method is a global thresholding technique that works on the assumption that the grey level histogram of the image has a bimodal distribution and it iteratively selects an optimal threshold such that it minimizes the within-class variance of the two modes, or equivalently maximizes the between-class variance. The resultant image is dilated using 3x3 structuring element. Then avoid the non-text regions by using some common properties of the Malayalam text. For that we used aspect ratio and volume. Repeated experiments enabled us to fix the value of aspect ratio and volume as a constant. Finally the result is superimposed into the input image for getting the actual text.

4. EXPERIMENTAL RESULTS

In the experiment part, algorithm explained in the previous section is applied on the images in the dataset. The output of the proposed algorithm is represented in image 3.

(a) **(b)**

Fig. 3. (a) Input Image, (b) Output image

The results of the proposed algorithm are evaluated using precision and recall rate. These two standard measures are used to evaluate the performance of the text localization algorithms. Precision and recall rate are calculated using the following equations. Where NC, NM, NFp are the number of correctly retrieved text regions, missed ones, and false positive respectively.

$$Recall = \frac{NC}{NC + NM}$$

$$Precision = \frac{NC}{NC+NFp} \qquad (11)$$

Performance of the proposed algorithm is given in the table 1 and the graphical representation of the performance is represented in Fig. 4 and 5. 100 images are divided into 10 sets with 10 images in each set. Average recall rate and precision rate of each set is separately calculated and then overall average is identified. Overall average recall rate of the proposed algorithm is 93.597 and overall average precision rate is 95.523.

Table 1: Performance analysis of the proposed algorithm

	Average Recall Rate	Average Precision rate
Set 1	92.35	91.01
Set 2	92.98	92.55
Set 3	94.97	90.78
Set 4	92.45	89.86
Set 5	93.56	93.82
Set 6	96.15	94.58
Set 7	94.37	93.68
Set 8	91.09	90.20
Set 9	95.25	95.11
Set 10	92.80	93.64
Overall average	**93.597**	**95.523**

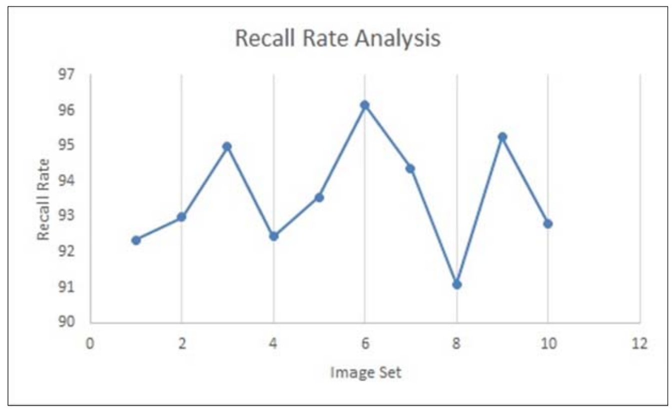

Fig. 4. Recall Rate Analysis

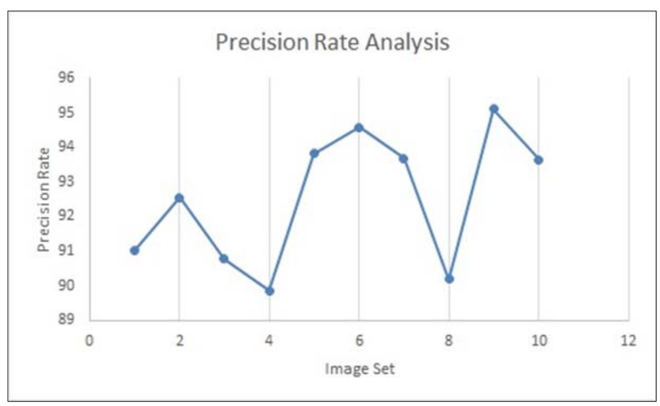

Fig. 5. Precision Rate Analysis

5. CONCLUSIONS AND FUTURE ENHANCEMENTS

In this paper we proposed an efficient and easy method for the localization of the Malayalam texts from the scene images. We got the overall average recall rate as 93.597 and overall average precision rate as 95.523. The proposed algorithm has lots of applications in various areas. Although there exists many OCR for the Malayalam text recognition, none of them gives the desired output. The proposed work is the first stage of creating a successful OCR which recognizes Malayalam characters. In future, we can improve the performance of the algorithm by localizing texts in complex backgrounds, different lighting conditions etc.

REFERENCES

Keechul Jung, Kwangin Kim and Anil K. Jain, "Text information extraction in images and video: a survey", 5, The journal of the Pattern Recognition society, 2004, 977-997.

Rainer Lienhart and Axel Wernicke, "Localizing and Segmenting Text in Images and Videos", IEEE Transactions on Circuits and Systems for Video Technology, 12 (4), 2002.

[1] C. P. Sumathi, T. Santhanamand G. Gayathri Devi, "A Survey on Various Approaches of Text Extraction in Images," International Journal of Computer Science & Engineering Survey (IJCSES), 3(4), 2012.

[2] Wonder Alexander Luz Alves and Ronaldo Fumio Hashimoto, "Text Localization in scene Images by Morphological Filters", SIBGRAPI – Conference on Graphics, Patterns and Images, 2009.

[3] Thomas Retornaz and Beatriz Marcotegui, "Scene Text Localization based on the Ultimate Opening," Proceedings of the 8th International Symposium on Mathematical Morphology, Rio de Janeiro, Brazil, 2007.

[4] Rodolfo P. dos Santos, Gabriela S Clemente, Tsang IngRen and George D. C. Calvalcanti, "Text Line Segmentation Based on Morphology and Histogram Projection," 10th International Conference on Document Analysis and Recognition, 2009.

[5] Jui-Chen Wu,Jun-Wei Hsieh and Yung-Sheng Chen, "Morphology-based text line extraction," Springer, 3 August 2007.

[6] S. T. Deepa and S.P. Victor, "A Novel Method for Text Extraction," International Journal of Engineering Science & Advanced Technology,2 (4), 2012, 961 – 964.

[7] Datong Chen, Jean-Marc Odobez and HerveBourlard, "Text Detection and Recognition in Images and Video Frames," Elsevier, Pattern Recognition 37, 2004, 595-608.

[8] R. C. Gonzales and R. E. Woods, Digital Image processing, Third Edition, Prentice Hall, 2008.

[9] J. Ohya, A. Shio and S. Akamatsu, "Recognizing characters in scene images", IEEE Transactions on Pattern Analysis and Machine Intelligence, 16 (2), 1994, 214–224.

[10] C.M. Lee and A. Kankanhalli, "Automatic extraction of characters in complex images", International Journal of Pattern Recognition and Artificial Intelligence, 9 (1), 1995, 67–82.

[11] Y. Zhong, K. Karu and A.K. Jain, "Locating text in complex color images", Pattern Recognition 28 (10), 1995, 1523–1535.

[12] R. Lienhart and W. Effelsberg, "Automatic text segmentation and text recognition for video indexing", Technical Report TR-98-009, PraktischeInformatik IV, University of Mannheim, 1998.

[13] A.K. Jain and B. Yu, "Automatic text location in images and video frames", Pattern Recognition 31 (12), 1998, 2055–2076.

[14] S. Messelodi and C.M. Modena, "Automatic Identification and skew estimation of text lines in real scene images", Pattern Recognition. 32, 1992, 791–810.

[15] E.Y. Kim, K. Jung, K.Y. Jeong and H.J. Kim, "Automatic text region extraction using cluster-based templates", Proceedings of International Conference on Advances in Pattern Recognition andDigital Techniques, Calcutta, India, 1999, 418–421.

[16] S.-W. Lee, D.-J. Lee and H.-S. Park, "A new methodology for gray-scale character segmentation andrecognition", IEEE Transactions on Pattern Recognition and Machine. Intelligence. 18 (10), 1996, 1045–1050.

[17] Y.M.Y. Hasan and L.J. Karam, "Morphological text extraction from images", IEEE Transactions on Image Processing 9 (11), 2000, 1978–1983.

[18] D. Chen, K. Shearer and H. Bourlard, "Text enhancement with asymmetric Alter for video OCR", Proceedings of International Conference on Image Analysis andProcessing, Palermo, Italy, 2001, 192–197.

[19] B. Sin, S. Kim, B. Cho, Locating characters in scene images using frequency features, Proceedings of International Conference on Pattern Recognition, Vol. 3, Quebec, Canada, 2002, pp. 489–492.

[20] W. Mao, F. Chung, K. Lanm and W. Siu, "Hybrid Chinese/English text detection in images and video frames", Proceedings of International Conference on Pattern Recognition, Vol. 3, Quebec, Canada, 2002, 1015–1018.

[21] K. Jung, K. Kim, T. Kurata, M. Kourogi and J. Han, "Text scanner with text detection technology on image sequence", Proceedings of International Conference on Pattern Recognition, Vol. 3, Quebec, Canada, 2002, 473–476.

[22] C.M. Lee, and A. Kankanhalli, "Automatic Extraction of Characters in Complex Images", International Journal of Pattern Recognition and Artificial Intelligence, 9 (1), 1995, 67-82.

Yu Zhong, KalleKaru, and Anil K. Jain, Locating Text In Complex Color Images, Pattern Recognition, 28 (10), 1995, 1523-1535.

[23] N. Otsu's, "A threshold selection method from gray level histograms," IEEE Transactions on systems, MAN and Cybernetics, 9 (1), 1979.