Extraction of Kannada text from different sets of Images

[1]Leena. A, [2]Pooja. A.P
[1]PG student, [2]Assistant professor
[1]Department of Electronics and Communication
[1]VVCE, Mysore, India

***Abstract* - Extraction of text content has raised immense thought from the Computer vision group as of late. Extraction process is said to be tough in images that is taken by camera. Extraction of content is a vital assignment because of it wide mixed bag of uses, for example, route, human Computer association, substance primarily based data indexing and recovery, image comprehension, image pursuit and indexing, programmed annotation and organizing of images. Significantly extraction of Kannada content from pictures is tough because of geometry and between characters separating which is non-uniform in Kannada script. The primary point of the work is to develop a system for extraction of Kannada text and to expel the non message area from shaded complex foundation pictures. So content can be seen clearly, that is helpful for a few applications. Therefore concentrating on the Kannada text from images possible with complicated foundation here a basic and effective Curvelet Transform is utilized. It changes an image by reducing it into an arrangement of directional sub-groups with details of surface captured in distinctive orientations and scales. The algorithms are executed in matlab and applied to an image set with distinctive text size, text designs and thickness.**

***IndexTerms*-Curvelettransform,canny,dilation,clustering.**

I. INTRODUCTION

With the rapid growth in large image collections, efficient ways of extracting advantageous information from them have gained immense research attention lately. Extracting content in images has turned into a potential application in abundant fields like intelligent carriage frameworks, robotics and so on. Text Extraction from image is anxious with extracting the accordant text data from a accumulating of images. Rapid development of digital technology has resulted in digitization of all categories of materials. There is a amount of advantageous application can be found in extraction of text which appears in testing of articles, archival manuscripts, journals and magazines, scanned article, maps, newspapers, microfilms, currency notes, identification of vehicle plates, stamps, Tele vision programs, commercial card, web pages, different argument account cartoon regions, and all these can be converted to images which is accustom with a lot of arduous analysis issues in extraction of text and detection.

Images are classified into document images, caption text images and scene text images. A document usually contains text and few graphic cartoon elements. Document images are acquired via scanning journal, written articles, degraded article pictures, handwritten chronological article along with book cover and so on, Caption text is referred in the direction of Overlay text or else Cut line text and unusually superimposed on top of the video/image by the side of cutting along with it constantly describes or identifies the area of the image/video content and Appearance of text after video or else image shot via recording devices or camera is termed as Scene images It can be seen easily while a portion of the scene along with text contain required linguistics data like in advertisements which is having artistic fonts, names of streets along with institutes names, shops, highway signs, information of traffic control, groceries containers, price boards, vehicle authorization detection/recognition and so on. Example of Document, Caption and Scene images are shown below in Figure 1, Figure 2 and Figure 3.

Fig 1: Document image

Fig 2: Caption image

Fig 3: Scene image

Largely built digital article analysis is gaining popularity which is designed for Indian scripts and speedy increment in computational approach which enables a major role in methodologies for extracting the information. Text appearances in the

greater part of the article are commonly skew and compact in nature. For the reason of overlapping and touching fonts the instability of the image is extreme. Kannada is a prehistoric language along with extra widespread letters appearing in order range than English language. The Kannada letter set is classed into two fundamental classes that are sixteen Vowels and thirty four consonants. Here introducing the pattern of direction of pixel gradient along with features of magnitude pixel in order to point out the dominant text pixels from the canny edge map in the given input frame. Text representatives are obtained during mapping out the dominant text pixel to the edge map of canny. In order to find out candidate text representatives the broken down segments can be eliminated by using connected component analysis. To seek out direction of text, Curvelet transformation is designed which group out the word patches in the direction of text to get the complete text line from the image. Matlab software is used in order to develop and test the algorithm. The execution of the methodology is conferred by exhibiting promising results of text for an array of images taken from distinctive types of book and magazine covers, Street banners and caption pictures with complicated foundation. The experimental results are found to be exceptionally satisfactory.

II. LITERATURE SURVEY

Some of the related works on text extraction are summarized below:

Nagendra D R, Nandish G R, Mahendra B V, Bhagya H K [1] present an effective and basic strategy to extract Kannada text in images most likely with complicated background. Curvelet is used which transform an image to decay it into directional sub-groups along with texture information taken in diverse orientations and scales. It has the fundamental components of wavelets, particularly, multi scale and time frequency localization, also offers a high level of directionality and anisotropy. As a replacement of detecting the edges only in horizontal, vertical & diagonal directions, this approach is intended to detect the edges appearing in multi directions to deal with text oriented designed for curved/ bended letters.

Sumathi, C.P. and G. Gayathri Devi [2] present Gamma modification to find out a gamma value which is used for enhancing the forefront information in an image. Also it uses gray balanced co-occurrence matrices, measurement of textures, thresholding methods. In order to eliminate non text section and to show the text section in the image proposed preprocessing approach is very helpful. To find out a correct gamma value, a range of gamma values from 0.1 to 10, with period of 0.1 is used for the image resulting in 100 pictures. After conversion of RGB to gray image, Gray Level co-occurrence matrix for each image is computed to extract the textural features contrast and energy. Each image is handled designed for 4 combinations of radius along with angle to facilitate in 1 and 0°, 1 and 45°, 1 and 90° and lastly 1 and 135°. Contrast along with Energy measures of matrices of the 4 orientations is averaged and Otsu's threshold algorithm is used in order to calculate threshold value for all images.

M. M. Kodabagi, Hemavati C. Purad [3] used a novel technique for recognizing fundamental Kannada characters in common scene images is proposed. Zonal wise horizontal and vertical profile based components of character images are used. The work has two stages training and testing. In training, zone wise vertical and horizontal profile based components are extracted from training samples and neural system is trained. In testing, the test image is prepared to obtain features and it will be recognized using neural system classifier, also it makes use of DCT based texture feature in order to identify the script of low resolution display board images. Because of minimal information context, noise and bigger skew misclassification will appear which affects the execution of the texture based work.

Deepak Kumar, A G Ramakrishnan [4] proposed a strategy by introducing DCT and ART in order to obtain feature vectors in the role of transformation module as soon as binarizing. The MAPS tactic is used as an amount of binarization step. DCT along with angular radial transform are utilized in order to extract the features of characters once the standardization for scale and translation is done. Here nearest neighbor classifier is designed for classification purpose and the novelty of this system is appearing in binarization and reduces the dimension of the extracted feature vectors.

Keshava Prasanna, Ramakhanth Kumar P, Thungamani. M, Monohar Koli [5] proposed a method which contains 3 basic stages i.e., content extraction, text detection and vocalization synthesis. In order to extract the text from images/videos an effective algorithm is used which is able to detect easily, localize and extract Kannada text from images which contains convoluted background. This method is based on the application of a color reduction technique, a SD based technique for detection of edges, and the localization of text regions using new CC properties. Output will be the text boxes along with clear-cut background, and then it can be fed into an optical character recognition for character recognition. This process is very effective along with respect to distinctive text dimensions, orientation, arrangement along with background complexities and textual style hues.

Kohei Arai1, Herman Tolle [6] introduced a methodology to detect and extract the content from commercial snapshot images. Their methodology utilizes boundary-based system and cc group approach known as blob extraction method. By using constant blob extraction method content blobs is obtained furthermore non content blobs. In order to separate the text from the image blend of homogeneity edge detection filter along with proper threshold number is used. Every associated pixel will be distinguished with the same shading list as in different blob object. To assemble a blob like a content blob or else non content blob, a small amount of fundamentals from content blob candidate will be extracted.

III. BASIC BLOCKS

Here presented a direct tactic to extract Kannada text from images which is having complicated background. Using a Curvelet which transform an image to decay it into set of directional sub-groups with texture points of interest captured in many orientations and scales. The proposed method is demonstrated as follows

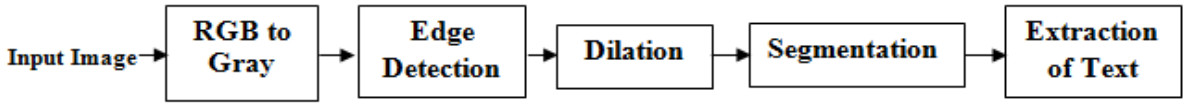

Fig 4: Block diagram

Input image:

Three different types of image sets are used namely document image, caption text image, scene text image. Algorithm will be evaluated on images which suffer from lack of determination and low characteristic, angle distortion, complicated background, font style and thickness, background moreover as foreground texture, camera position which can introduce geometric distortions, image resolution, shadows, irregular lightening, low contrast and large signal dependent noise, slant and tilt.

RGB to gray:

In the preprocessing step, first the input RGB image is changed over to gray scale image because if the image information is in RGB it can't be used for segmentation straight forwardly.

Edge detection:

The name itself says it is used for detection of edges, which find outs the sharp discontinuities appearing in an image. Canny will smoothes the image in order to wipe out the noise, then image gradient will be detected to highlight regions along high spatial derivatives. Canny find outs the regions and eliminate the pixel which is not at the greatest and gradient array will be decreased by hysteresis.

Dilation:

Morphological dilation operators are applied to join separated text edges of every point of interest component sub-band in a transformed binary image. Dilation is a process which grows or else improves the region of interest with the help of structuring element of obliged shape and size.

Segmentation:

It isolates an image into a various discrete regions such that the pixels have high similarity in each region and high contrast between regions. With a specific end goal to cluster out the text pixels k means clustering is utilized. The K-means clustering algorithm is a basic system for evaluating the mean of an arrangement of K-groups.

Extraction of text:

Curvelets allow time frequency localization of wavelets although shows a very high level of directionality and anisotropy, and its singularities be able to approximate along with a small amount of coefficients. It is a multiscale directional transform that allows a practically ideal nonadaptive sparse representation of objects with edges. It transforms an image to decompose it into a set of directional sub-groups along with surface subtle elements taken in distinctive orientations and scales. In order to extract the text line word patches will be gathered based on direction of these word patches. And it offers a high level of directionality.

Fast discrete Curvelet transform is implemented using Wrapping function which is based on Fourier samples that takes a 2D input image in the form of Cartesian array

$$f[m, n], 0 \leq m \leq M , 0 \leq n \leq N \quad (1)$$

Where, M and N= Dimensions of the array

The result will be a collection of Curvelet coefficients $c^D(j, l, k_1 k_2)$sorted out by a,

- scale j
- orientation l
- spatial location parameters k_1 and k_2

$$c^D(j, l, k_1 k_2) = \sum_{0 \leq n \leq N}^{0 \leq m \leq M} f[m, n] \, \emptyset^D_{j,l,k_1 k_2}[m, n] \quad (2)$$

Where, $\emptyset^D_{j,l,k_1 k_2}$ = Digital Curvelet waveform and superscript

D = digital.

Wrapping based CT is a multiscale pyramid which comprises of a number of Subbands by distinctive scales comprising of diverse orientations and positions in the frequency domain. At a high frequency level, Curvelets are consequently adequate and it is a needle shaped element and at low frequency level it is a non-directional coarse element. The steps for wrapping based FDCT algorithm shown below,

- In order to obtain Fourier samples two dimensional FFT is

$$\hat{f}[m, n], -\frac{n}{2} \leq m, n < \frac{n}{2} \quad (3)$$

- For every scale j and angle l , the product will be formed

$$\tilde{U}_{j,l}[m, n]\hat{f}[m, n] \quad (4)$$

- Product will be wrapped around the origin by obtaining

$$\hat{f}_{j,l}[m, n] = W\left(\tilde{U}_{j,l}\hat{f}\right)[m, n] \quad (5)$$

Where m, n, $\theta = 0 \leq m < 2^{j}, 0 \leq n < 2^{j/2}$ and $-\frac{\pi}{4} \leq \theta < \frac{\pi}{4}$

- IFFT is applied to obtain $\hat{f}_{j,l}$, Therefore collecting the discrete coefficients $c^{D}(j, l, k_1 k_2)$

Finally the properties of the text region in an image will be measured and lastly the detected word is segmented from the input image.

IV. RESULT

STEP 1: Input image

Fig 5: Input Image

The image will be read first whose content must be segmented.

STEP 2: Conversion of Gray image

Fig 6: Gray Image

In order to extract the text, segmentation as to be done but if the input image contains RGB value then it can't be utilized for segmentation straightforwardly, so it has to convert it into gray image.

STEP 3: Selection of ROI

Fig 7: Region of interest selected from gray image

Region of interest will be selected in order to eliminate the possible background which suffers from blurring effects of assorted lighting, sophisticated movement and transformation, unknown layout and shadowing

STEP 4: Detection of edges

Fig 8: Edge detected input image

Next edges will be extracted by using the canny edge detection algorithm. Canny edge accepts a grayscale as its input, and precedes a 0's and 1's image of a similar size as output image.

STEP 5: Performance of dilation.

Fig 9: Dilution of text

Text pixels are clustered together by using a morphological dilation operation and remove the pixels that are far from the candidate text regions.

STEP 6: Extraction of text

Fig 10: Display of the Extracted Text

After the dilation, the non text region will be eliminated and the properties of the text region in an image will be measured and finally the detected word is extracted from the input image.

V. CONCLUSION

The work of text extraction from images is proposed utilizing the canny edge detector, morphological dilation operator, and the Curvelet transform. The goal of the segmentation process is to divide the image into disjoint areas, in such a direction to facilitate the text is removed from the background, and fonts don't seem to be combined at the same time. The evaluation of the proposed algorithm was done on an information set consisting of diverse images. For scene images ROI selection is made in order to view the text properly. The capability of the text extraction tactic is to know whether an image contains text or not. The Algorithm not only constrained to Kannada text, it can be utilized for different languages also.

VI. ACKNOWLEDGMENT

Author owes deep gratitude to the ones who have contributed greatly in completion of this manuscript. Foremost, would like to express sincere gratitude to the guide Ms. Pooja A.P, Assistant Professor, Department of Electronics and Communication, VVCE, Mysuru, for providing her profound insights and attention to details have been true inspirations to this script.

REFERENCES

[1]. Nagendra d r, Nandish g r, Mahendra b v, Bhagya h k, "a novel approach for extracting Kannada text in image", proceedings of 5th sarc-irf international conference, Bangalore, India, 04th may, 2014. ISBN: 978-93-84209-13-1

[2]. Sumathi, C.P. and G. Gayathri Devi, "Automatic Text Extraction from Complex Colored Images Using Gamma Correction Method", Journal of Computer Science, ISSN: 1549-3636, Pp no: 705-715**,** 2014.

[3] M. M. Kodabagi1, Hemavati C. Purad**, "**Script identification Of text in low resolution images of display boards Using DCT coefficients" , international journal of graphics and multimedia (IJGM), Volume 4, Issue 1, January - April 2013, pp. 31-40

[4] Deepak Kumar, A G Ramakrishnan, "Recognition of Kannada characters extracted from scene images", DAR '12, December 16 2012, Mumbai, IN, India

[5] Keshava Prasanna, Ramakhanth Kumar P, Thungamani.M, Manohar Koli, "Kannada Text Extraction From Images And Videos For vision Impaired Persons",International Journal of Advances in Engineering & Technology© IJAET, ISSN: 2231-1963,Volume No**:** 1,Issue no: 5 Pp no: 189-196, Nov 2011.

[6] Kohei Arai1, Herman Tolle (2011)," Text Extraction From Tv Commercial Using Blob Extraction Method", International Journal Of Research And Reviews In Computer Science Vol.2, No. 3