Offline Tamil Handwritten Character Recognition Using Zone Based Feature Extraction Technique

K.Mohamed Sayeed Khan[1], Dr.S.Shajun Nisha [2], Dr.M.Mohamed Sathik[3]
[1,]Dept of Comp Science
[2]Asst Prof &Head, Dept of Comp Science
[3]Princpal
[1, 2, 3] Sadakathullah Appa College ,Rahmath Nagar,Tiruneveli.

Abstract- *The Tamil characters in the ancient Stone Inscriptions are used for the recognition and extraction. Text extraction from images is an active research area as it is contributing more to Content based image indexing and Text & database retrieval. Text embedded / inserted in images is used to describe the contents of an image & also it can be easily extracted compared to other semantic contents. This paper proposes a new technique for the extraction of Tamil characters from the image. The proposed system consists of four steps: pre processing, detection, extraction and recognition. While pre processing, the noise is removed from the image and the image is enhanced for further processing. We apply edge detection technique to detect the characters from the image. For extraction of the characters, we do line determination and character separation using horizontal and vertical projections. Finally, we apply Discrete wavelet transformation to recognize the characters from the image. For classification, we use KNN classifier. The proposed system is tested with various test images and the experimental results prove that the proposed system outstands the other state of art methods.*

Keywords- Line Detection, Extraction, Character, Detection, Separation, Tamil, Inscriptions.

I. INTRODUCTION

With the recent advances in digital technology, more and more databases are multimedia in nature, containing images and video in addition to the textual information. Text extraction from images is an active research area as it is contributing more to Content based image indexing and Text & database retrieval. Text embedded / inserted in images is used to describe the contents of an image & also it can be easily extracted compared to other semantic contents and it facilitates applications such as keyword-based image search, automated processing & reading of documents, text-based image indexing, Pre processing for OCR technique and multimedia processing. Variation in Font style, size, Orientation, alignment & complexity of background makes the text extraction as a challenging task. Text in images/video images are classified into overlay text and Scene text.

Caption text is the one which is inserted text & otherwise called as superimposed/artificial text. Natural images /embedded are called as Scene text/graphics text. In this scenario, there is a need for unified methodology to extract text from various kinds of images so as to suit users' information needs. There are eight kinds of text characteristics that are frequently used in video text detection, localization, and extraction: contrast, color, orientation, stationary location, stroke density, font size, aspect ratio, and stroke statistics. From the point of view of multilingual text processing, they are classified into language-independent characteristics and language-dependent characteristics.

Language-Independent Characteristics

Contrast:

Some methods assume that a text has high contrast against its background; however, this cannot be guaranteed due to complex background in video scenes. Actually, against a clear background, even a low-contrast text can be easily recognized, while in a complex background, only the high-contrast text can be discerned. Thus, we set an adaptive contrast threshold according to the background complexity.

Color:

We believe that most text strings originally have uniform color, but the lossy compression causes the color bleeding effect at the edge of text. Thus, the color of a text string is in a range whose breath is related to the local background color.

Orientation:

No matter whether it is an artificial text or a scene text, if it is closely related to the video content, it must have been presented in a way easy to read, i.e., with a certain

contrast, a reasonable size, and an approximately horizontal orientation. Thus, the constraint of horizontal text orientation is adequate for the video indexing purpose.

Stationary Location:

Most texts of interest are stationary over time. Scrolling texts are only used in credit titles and special effects. To keep the efficiency of video indexing system, stationary texts only handled.

Language-Dependent Characteristics

According to the linguistic classification, English, French, and Spanish belong to alphabetic literal, whereas Chinese, Japanese, and Korean belong to ideograph. Their differences in the following four aspects affect the video text processing.

Stroke Density:

The stroke density of an English text string is roughly uniform but that of a Chinese text string is not. Because the stroke number of a Chinese character may vary from 1 to more than 20 and each character occupies the same-sized block space, the stroke density varies significantly .This poses an obstacle to employ the region-growing techniques. However, the stroke density of a whole string is still high. *Font Size:*

To display all strokes clearly, an English string requires at least 8-pixel-high font size, while a Chinese string requires at least 20-pixel-high font size due to the large stroke number. This is the basis of size constraint to discard regions that are too small.

Aspect Ratio:

In English, characters (letters) must be combined to form meaningful words. In Chinese, each individual character has its own meaning, and it can also be combined with other characters to express more meaning in terms of phrases. Sometimes, a text string may contain some characters with long spaces, but they are not separable .Thus, the aspect ratio constraint should be adapted to Chinese characters

.

Stroke Statistics:

Since English characters consist of mainly vertical strokes, many methods only consider vertical edges to detect or extract English texts. However, Chinese characters consist of four directional strokes. Moreover, those filters to extract strokes do enhance the stripe-like strokes but suppress the intersections of strokes. Since English characters do not contain many intersections, this will not affect the character recognition too much. However, Chinese characters contain many intersections and these intersections are important clues for character recognition. Therefore, a multilingual method has to handle both strokes of all directions and intersections well. The two types of texts are described below.

Scene Text

Scene text occurs naturally in the background as a part of the scene, such as the advertising boards, banners, and so on.

Overlay Text

Rather than that, overlay content is superimposed on the video scene and used to help watchers' understanding. Since the overlay content is profoundly conservative and organized, it can be utilized for video ordering and recovery. The accompanying are the issues in content discovery and extraction.

Raw sensor image and sensor noise:

In low-priced Handheld Image Devices (HID), pixels of a raw sensor are interpolated to produce real colors, which can induce degradations. Demosacing techniques, viewed more as complex interpolation techniques, are sometimes required. Moreover, sensor noise of an HID is usually higher than that of a scanner.

Viewing angle:

Scene text and HIDs are not necessarily parallel creating perspective to correct.

Motion Blur:

During acquisition, some motion blur can appear or be created by a moving object. All other kinds of blur are included in some other imaging conditions such as focus, for example.

Focus:

HIDs are not really outfitted with auto-center and focal point deviation can firmly obscure the picture.

Lighting:

In real images, real (uneven) lighting, shadowing, reflections onto objects, inter-reflections between objects may make colors vary drastically and decrease analysis performance.

Resolution and Aliasing:

From webcam to professional cameras, resolution range is large and images with low resolution must also be taken into account.

Outdoor/non-paper objects:

Different materials cause different surface reflections leading to various degradations and creating inter-reflections between objects.

Scene text:

Backgrounds are not necessarily clean and white, and more complex ones make text extraction from background difficult. Moreover scene text such as that seen in advertisements may include artistic fonts. Due to the large diversity of backgrounds and text to handle, it is rather difficult to detect and extract text in the image.

Non-planar objects:

Text embedded in bottles or cans suffer from deformation.

Unknown layout:

There is no a priori information on structure of text to detect it efficiently.

Objects in distance:

Distance between text and HIDs can vary, and character sizes may vary in a wide range, leading to a wide range of character sizes in a same scene.

Moving objects:

A moving camera (by its mobile context) or moving objects may induce unknown motion blur, which is difficult to model, leading to degraded imaging conditions, previously detailed.

The following paragraphs explain the related works of the proposed system.

In text extraction algorithm using Dual Tree Complex transform, the images with Tamil text are considered for extraction. The text extracted can be binarized and sent to text to audio software which can be used for visually challenged persons. The Methods are Applying the Dual tree complex wavelet into the input text image, we get the higher and lower sub bands, in all the sub bands detecting the edges. Edges are detected using canny edge detection method as there are many curvature in the Tamil characters. Farras filters are used for orthogonal2D channel perfect reconstruction. The advantages are approximate shift invariance. Good selectivity and directionality in 2-dimensions with Gabor-like filters. Perfect reconstruction (PR) using short linear phase filters. Limited redundancy, independent of the number of scales efficient order-N computation- only times the simple DWT for m-

D. The disadvantages are lack of shift invariance.

A morphological binary map is generated by calculating difference between the closing image and the opening image. Then candidate regions are connected by using a morphological dilation operation and the text regions are determined based on the occurrence of text in each candidate. The detected text regions are localized accurately using the projection of text pixels in the morphological binary map and the text extraction is finally conducted. The methods are Morphological binary map is generated using closing and opening image. Connected components for each candidate region are generated. The advantages are to reduce the processing time. Texts are extracted based on color polarity computation method. The disadvantages are to detect and extract the text with different orientations

A system for automatically detecting signboards and extracting text from them is presented. First it locates the signboard's position on the image by the Hough transform, edge density checking and quadrilateral finding. Then, with a geometric transformation, the system changes the signboard from an arbitrary quadrilateral to a rectangular shape. Finally, it performs segmentation to extract the characters on the signboard. In our system, the signboards can have large variations such as rotation and scaling and are not limited to be rectangular in the images. The methods are All the color information in the signboard are turned into grayscale and applying Hough transform to finding edges and noise reduction. Type segment means a segment is part of the side of the quadrilateral and its end point is not a corner. Type corner means that the end point of a segment is a corner. Type alternative is used when there is more than one segment with the same intersection point as the previous end point.

Different viewing angles of images, a rectangular signboard are usually distorted to a quadrilateral shape. To transform the signboard back to an undistorted rectangle the geometric transformation are applied. Apply the vertical projection to detect a text. The advantages are to guide the navigation of an autonomous robot. Performance is good. The disadvantages are more robust detection of edges and character or word recognitions

In Overlay Text Detection in Complex Video Background, the characteristics that there are transient color changes existing between text and its adjacent background, to generate transition maps are explained. And this work applied the connected components to smooth the connected rectangular shapes. The study results demonstrated that the text detection method based on the transition map can effectively detect the text regions under different text contrasts, fonts and background complexities. The methods are Transition Map generation, Search for candidate text regions, Determination of overlay text regions. The Advantages are accuracy generate smooth text regions. The Disadvantages are only binary text can be extracted in overlay text regions.

The rest of the paper is organized as follows. Section II describes the proposed methodology in detail. Section III deals with the experimental results obtained. Section IV covers the conclusion and future works of the paper.

II. PROPOSED METHODOLOGY

The proposed system of text detection and recognition can be divided into the following modules.

i) Pre Processing
ii) Text Detection
iii) Text Extraction
iv) Text Recognition

A). Pre Processing:

As the input images obtained are often noisy in nature, we have to pre process the images before doing further process of extraction. The pre processing step consists of two steps: Filtering and Contrast Enhancement;

Filtering:

The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing. Median filtering is very widely used in digital image processing because, under certain conditions, it preserves edges while removing noise. The advantages are No reduction in contrast across steps, since output values available consist only of those present in the neighborhood. Median filtering does not shift boundaries, as can happen with conventional smoothing filters. Since the median is less sensitive than the mean to extreme values, those extreme values are more effectively removed.

Contrast Enhancement:

Histogram equalization is a technique for recovering some of apparently lost contrast in an image by remapping the brightness values in such a way as to equalize, or more evenly distribute, its brightness values. As a side effect, the histogram of its brightness values becomes flatter. A key advantage of the method is that it is a fairly straightforward technique and an invertible operator. So in theory, if the histogram equalization function is known, then the original histogram can be recovered. The calculation is not computationally intensive

B) Text Detection and Extraction

The second and third module can be combined as text detection and extraction. In this module, we have three steps: Edge detection, Line Determination and Character Separation.

Edge Detection:

We used canny edge detection technique in the proposed method. Canny has shown that the first derivative of the Gaussian closely approximates the operator that optimizes the product of *signal-to-noise ratio* and localization Filter image with derivative of Gaussian. Find magnitude and orientation of gradient. One of the advantages of the canny edge detection is the simplicity, and the fact that all edges detected are closed loops.

Line Determination:

We detect a line, using horizontal projection to find the line. We find the lines from the Inscriptions. Horizontal projection angles to the screen, allowing users to place the projector anywhere in the room and still display a perfect image.

Character Separation:

In this step, we extract or separate a single character from the given inscription or image after the previous steps of

processing. We use the vertical projection technique for separating a character from the inscription.

C) Text Recognition:

This is the final module of the proposed system. We use optical character recognition to train the data and each character should be matched for the trained data. Using discrete wavelet transform (DWT) we can find the result of Tamil font from the Inscriptions. A discrete wavelet transform (DWT) is any wavelet transform for which the wavelets are discretely sampled. A key advantage of this technique over Fourier transforms is that it has temporal resolution and it captures both frequency and location information

III. EXPERIMENTAL RESULTS

The proposed system is implemented using a Mat lab program where it is evaluated for recognize the image. The performance of the algorithm is evaluated on several real images. These pictures are the most widely used standard test images used for image recognition algorithms. The image contains a nice mixture of detail, flat regions, shading and texture that do a good job of testing various image processing algorithms. These are still in the industry standard for tests. It is a good test images. These images are used for many image processing researches.

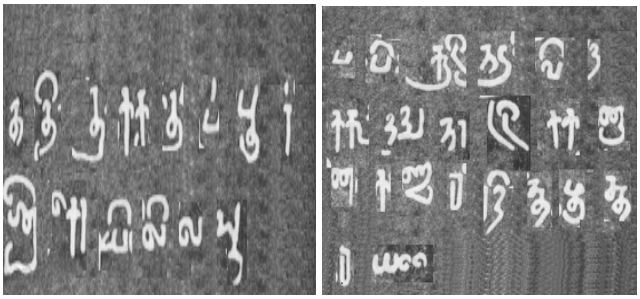

Figure 1 Experimental Images

The above figures are used as input images for analyzing the performance of this approach. This image is the gray scale image. The size of image is 300 x 171. The image type is BMP. The input images are subjected to pre processing and the result image after pre processing is shown in figure 3
.

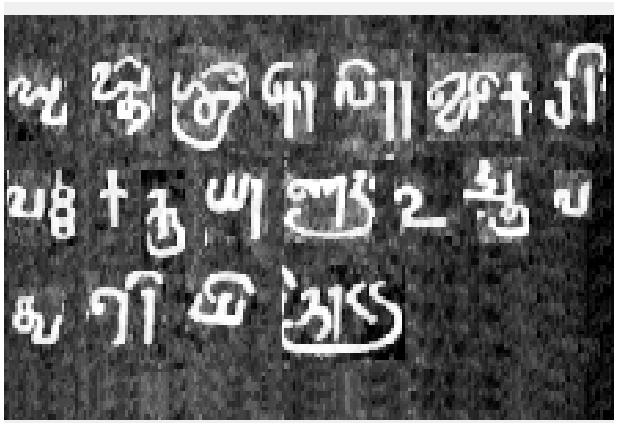

Figure 2 Input image after Pre processing

The canny edge detection is applied on this image, which is followed by the line detection. The resultant image after line detection is shown below.

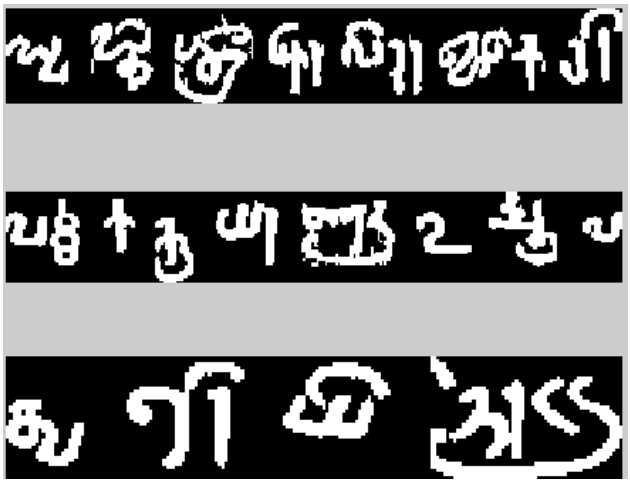

Figure 3 After Line Detection

The individual characters are separated after line detection. They are shown in the following figures. In this performance analysis of text detection and extraction methods we have to analyze the above three text detection and extraction techniques and find out which one is the best. We carried out computational experiments to conform the effectiveness of the above techniques.

There are various tools are available to evaluate the performance of the text detection and extraction techniques. They are as follows:

Precision Rate (PR) - Precision rate (P) is defined as the ratio of correctly detected characters to sum of correctly detected characters plus false positives.

$$\text{Precision rate} = \frac{\text{Correctly detected characters}}{\text{Correctly detected characters} + \text{False positives}} \times 100\%$$

Figure 4 Character Separation

Recall Rate (RR) – The Recall rate (R) is defined as the ratio of the correctly detected characters to sum of correctly detected characters plus false negatives.

$$\text{Recall rate} = \frac{\text{Correctly detected characters}}{\text{Correctly detected characters} + \text{False negatives}} \times 100\%$$

TABLE I
Comparison of Differential Approach and DWT

Images	Differential Approach	DWT
Image1	81%	88%
Image2	70%	75%
Image3	83%	89%
Image4	81%	87%
Image5	70%	79%

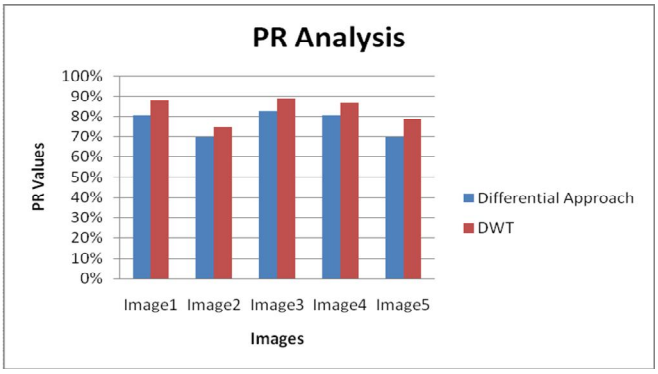

Figure 5Comparison between Differential Approach and DWT

IV. CONCLUSION AND FUTURE WORK

In this thesis a new approach is used for text detection and extraction. The tamil characters in the ancient Stone Inscriptions are used for the recognition and extraction. To detect, extract and recognize the Tamil character from the image this thesis performed three steps. They are preprocess, detection and extraction and recognition. In the above three process the first stage is the preprocess. In preprocess stage the noise data can be removed and contrast of an image can be increased. In the second that is character detection and extraction stage, edge detection technique is performed to detect the characters. The canny edge detection operator is used for edge. And then line determination and character separation is done using horizontal and vertical projections. They are applied to extract the line and characters. Then extracted characters are stored in the separate folder. The Final stage is the recognition. In this stage Discrete Wavelet Transform (DWT) is applied to the characters in the training folders and then DWT is also applied in the extracted characters from the input image to find the mean value. These mean value are considered as the features. The find the difference value between the trained character's mean value and the input image's characters mean value by using the Euclidean distance. Finally these distance value can be classified using KNN classifier.

In future, the above technique can be applied on the text with very complex background. It can also be extended to the moving videos.

REFERENCES

[1] "Automatic Urban Road Extraction using Airborne Laser Scanning/ Altimetry and High Resolution Satellite Data" Poonam S. Tiwari . H. Pande . Ashwini Kumar Pandey , springer, J. Indian Soc. Remote Sens. (June 2009) 37:223–231

[2] "Automated updating of road databases from aerial images", International Journal of Applied Earth Observation and Geo information" Emmanuel Baltsavias*, Chunsun Zhang1, 6 (2005) 199–213)

[3] "Comparison of pixel-based and object-Oriented classification approaches using landsat-7 ETM spectral bands" M. Oruc a, *, A. M. Marangoz a, G. Buyuksalih a ZKU, Engineering Faculty, 67100 Zonguldak, Turkey

[4] "Hamprecht1 Quantitative Assessment of Image Segmentation Quality Random Walk Relaxation Times", Bjorn Andres1, Ullrich Koethe1, Andreea Bonea1, Boaz Nadler2, and Fred A.,

[5] "Performance evaluation of atomi system for road database updating from aerial film,ADS40, , IKONOS AND QUICKBIRD orthoimagery"

[6] C. Zhanga, E. Baltsaviasb, L. O'Sullivanc: Stilla U, Rottensteiner F, Hinz S (Eds) CMRT05. IAPRS, Vol. XXXVI, Part 3/W24 --- Vienna, Austria, August 29-30, 2005

[7] "Multi-sensor imagefusion in remote sensing: concepts, methods and applications," Pohl C and Genderen JLV (1998). International Journal of Remote Sensing19(5):823–854

[8] "Automatic Feature Recognition and Extraction from Remote Sensing Imagery." Granzow Ed (2001), National Consortia on Remote Sensing in Transportation. U.S Department of Transportation

[9] "Automatic linear feature extraction by dynamic programming and LBS-snakes.", Gruen A and Li H (1995) , Photogrammetric Engineering and Remote Sensing 63(8): 985–995

[10] "Road Extraction in rural and urban areas. Semantic Modeling for the Acquisition of Topographic Information from Images and Maps", Hinz S, Baumgartner A, Steger C, Mayer H, Eckstein W, Ebner H and Radig B (1999) pp 133–153

[11] "Emperical Evaluation of Automatic Extracted Road Axes. In Bowyer K and Phillips P (Eds), Emperical Wiedemann C, Heipke C, Mayer H and Jameet O (1998)

[12] "Hierarchical image object-base structural analysis toward urban land use classification using high-resolution imagery and Airborne LIDAR data. In: the 3rd international symposium on remote sensing of urban areas," Zhan Q, Molenaar M and Tempfli K (2002) vol.1. Istanbul, Turkey, pp 251–258

[13] "A hierarchical object-based approach for urban land use classification from Remote Sensing data". PhD thesis, Zhan Q (2003), ITC