

# Scene Text Extraction from Invariant Complex Image by Using Gaussian Filter and Fuzzy Logic

Nouri Ali<sup>1</sup>, Husayn Abo showfa<sup>2</sup>, Osama Almabrouk<sup>3</sup>

Asst. Lecturer, Dept. of General, Higher Institute of Medical Sciences, Bani Walid, Libya<sup>1</sup>

Asst. Lecturer, Dept of CS, College of Electronics Technology, Bani Walid, Libya<sup>2</sup>

Asst. Lecturer, Dept of CS, College of Electronics Technology, Bani Walid, Libya<sup>3</sup>

**ABSTRACT:** Many images is rapidly expanding area that daily use to posts and pages design on web pages as well as videos contain visible text .Usually, we can read characters printed on sheets of papers according to some rigid format restriction. Scene text is important feature to be extracted , especially in scene which vary along its size, scale, orientation and colour intensity. This paper addresses the major challenges associated with an edge-based text region detection ,extraction and recognition algorithm. Based on this aim is to provided solid solution with respect to font size, colour intensity and orientation which implemented with various kernels in whole directions with 45<sup>0</sup> angle increased respectively in kernel directions which has been started in 0<sup>0</sup> angle completely directed until 315<sup>0</sup> angle into its position and angle respectively and due to effectively localize which involves further enhancing regions by eliminating non-text regions and retain candidate text regions, but some merged characters need morphology to separate characters. And it can be used fuzzy logic to recognize the character based on it, that is a logical system which is an extension of multi-value logic. While fuzzy logic technique was used to identify the character after extracting text from image. We were testing when got robust result . The proposed method with various kinds of natural images and confirmed which extraction rate are high.

**KEYWORDS:**Edge-Detection, Gaussian-Edge detector, Optical Character Recognition(OCR), Text Localization, Text Extraction, Morphology, Fuzzy Logic.

## I. INTRODUCTION

The demands to providing a developed candidate text that contains in image and video usually provides brief and important information about it. A sign is an object that suggests the presence of a fact. It can be a displayed structure bearing letters or advertise a place of business. Signs are everywhere in our lives. They make our lives easier when are familiar with them. Edge detection from a natural scene is useful in many applications. A well-know example is vehicle license or post detection. Edge detection that is the process of finding edges in images. Edge detection an image significantly reduces the amount of data and filters out useless information, while preserving the important structural properties in an image that signs containing text from a natural scene[1][2]. In recent years, great progress has been made in Optical Character recognition(OCR) technique that can only handle an text against plain monochrome background, and extract text from a complex background commercially ,OCR engines cannot yet detect and recognize text embedded in complex background directly[4] .Candidate regions of digital are extracted from an image through edge extraction ,enhancement and labeling. Since the digits in the image often have skew and slant correction[3]. Extraction text from image has been relied upon mainly on the properties of text .In the past few years. It was witnessed rapid growth in a number and variety of application using fuzzy logic that is logical system which is an extension of multi-value logic. It was used to identify the character after extracting text from an image[5].

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

Although these systems can be very useful, but cannot commentary useful in many situations, because the characters have many deficiencies that could be improved in larger skew and slant than ones in the document images scanned using scanner[6]. However, the objective of this research is to develop a robust system that can automatically detect text by using edge based various kernels can be used for this operation, the whole set of eight kernels are produced by taking each kernel with specific direction as rotating its coefficient circularly. Edge detection has been well worked for digital image corrupted and sharp intensity edge-based with multi-scale and multi-orientation whereas this operator can be used in practical image which generally suitable performance. The text data is particularly interesting because text data can contain image of varying text due to differences in size, orientation and alignment as well as complex background that make the problem of automatic text extraction extremely challenging. This method has been applied in a prototype English sign translation system which normally only has a horizontal, vertical and diagonal alignment. In order to be useful candidate text which be localized/extracted algorithm should be robust with respect to font size, styles, orientation, alignment and effects of lighting and the computation time must be small enough to be used in real-time.

In present paper, we introduce a fast and robust algorithm of text identification based on an edge-based text extraction which is robust with respect to font size, styles, color/intensity, orientation, effects of illumination, and complexity image background, and can quickly and effectively localize and extract text from real scenes.

## II. RELATED WORK

This section discusses issues related to the study, which describes the-state-of-art of categories segmentation and focus on the recognition text. They also describe analytic segmentation, run-length coding algorithm, and properties of text and identifying character. The fuzzy logic is used to determine the character.[13] it has introduced categories of segmentation which division of images into meaningful structure i.e. image segmentation is often an essential step in image analysis, object representation, visualization and many other image processing tasks. Here, there are many methods that find particular pixels that makeup an object. With this technique, detected edge in image is assumed to represent object boundaries, and used to identify these objects and from through known the boundaries of the object we can recognition the object. [14] The edge detection is fundamental tool used in most image processing applications to obtain information from the frames as precursor steps to feature extraction and object segmentation.[4] They employed the image is varied with distance from the camera. When pictures are taken with different distance, it affects the resolution of image and text. Acquisition text via scanner device would display image containing texts as seen in figure 2.2. Meanwhile, detection text is suitable form that text area identification. Since limited text detection that possibly localization of the text. [4] Algorithm based on Sober edge operation in four directions and by Sober operator are dense that precise edge localization. It gets gradient map in three component RGB. Meanwhile, morphology operation "close" and "open" operation the edge-pixels adjacent to each other. The "open" operation is used to disconnect edge map where it is too narrow to contain text. Afterwards, use projection profile of image block that compact representation of spatial pixel. Then, bounding edge-dense blocks in edge map after profile projection and compare the performance with it. [6] They used the method is based on finding text edge using information content of sub image coefficients of discrete Wavelet transformed input image, the most text used image are well characterized by their contained edge and "dense edge" are distinct characteristics of text blocks which can be used in detecting possible text region and using Sobel detect is effective in extracting strong edge in image. [3] Generate the edge map for text detection, they mainly depend on the edge information to detect text. Then, two edge maps are generated for detection of scrolling text and edge map is generated by performing Sobel detection to entire input image. Ideally, text area identification that whose task is to sure from that detection is text. [29] Which three rules of confirm from candidate text blocks. These rules are divided on. First, text block height and text block width are that text line will contain at least two word. Second, is used to limit the size of text T1 and T2 are set as 8 pixels and 32 pixels respectively. Text blocks whose height is small than 8 pixels or large than 32 pixels can be found in zoom image. Third, is to eliminate blocks that contain too few edge pixels, but with this, sometimes remain noise similar of text in the size, that is non-text. So use Wavelet feature and SVM classifies of verifying from candidate text. Although text has text property, this text is quite weak and irregular, only that they include some strokes i.e. horizontal, vertical, up-right-slanting and up-left-slanting stroke. These strokes

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

are regular to some extent when we consider them as one block, but never regular as to each pixel. Furthermore, [3] which false alarm due to complex background could be filtered in horizontal edge map are used to identify the text region and to get the refined text region. So calculate the pixels with horizontal edge in detected text region to identify whether it is true text region.[4] 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 threshold 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 position is far away to each other isolated. And use a morphological dilation operator that 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, 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.

Considering the Otsu method is often used to calculate the threshold to segment text from the background. For the vertical motion text, a vertical adaptive threshold is applied meanwhile horizontal adaptive threshold for the horizontal motion text. In corresponding, [7] Which use Strong & boosted edges after dilation are combined with addition followed by AND operation which forms the text region. Results of Dilation & logical operation and mapping to the original image to get text regions Remaining Non text regions are identified & eliminated by removing from a binary image .Meanwhile, [4] Their method that cannot handle character embedded in shade, the text used or complex background .The stage to extract accurate binary character from localized text regions. So that can the existing OCR directly for recognition, they use uniform white character pure black ground.[7] They have been shown pattern recognition (PR) is a subject researching object description and classification method, it is also a collection of mathematical, statistical, heuristic and inductive techniques of fundamental role in executing the tasks like human beings on computers .Pattern recognition includes a lot of methods which implies the development of numerous applications in different fields. The practicability of these methods is intelligent emulation.

### III.PROPOSED METHOD

The design of this system rely entirely on image processing algorithm and techniques. that described functional diagram of proposed algorithm is presented in figure 1.As can be seen in the flowchart ,a colour image is entered to the system as the input data and the segmented text on a clear black background will be the output. In the following, we will describe each stage in detail. The basic in previous knowledge of edge-based text extraction should be comply of following points.

1. Input image or real image with original colors.
2. Create a Gaussian pyramid by convolving the input image with Gaussian kernel and successively down-sample each direction by half.
3. Create directional kernels to detect edge  $0^{\circ}$ ,  $180^{\circ}$ ,  $45^{\circ}$ ,  $225^{\circ}$ ,  $90^{\circ}$ ,  $270^{\circ}$ ,  $135^{\circ}$  and  $315^{\circ}$  orientation.
4. Convolve each image in the Gaussian pyramid with each filter
5. Collect kernels to entire edges.
6. Dilate the resultant image using sufficiently structuring element (3x3) to cluster candidate text together.
7. Create final output image in white pixels against a plain black background.
8. Candidate text based on its properties like width and height
9. Use fuzzy logic to recognize the character.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

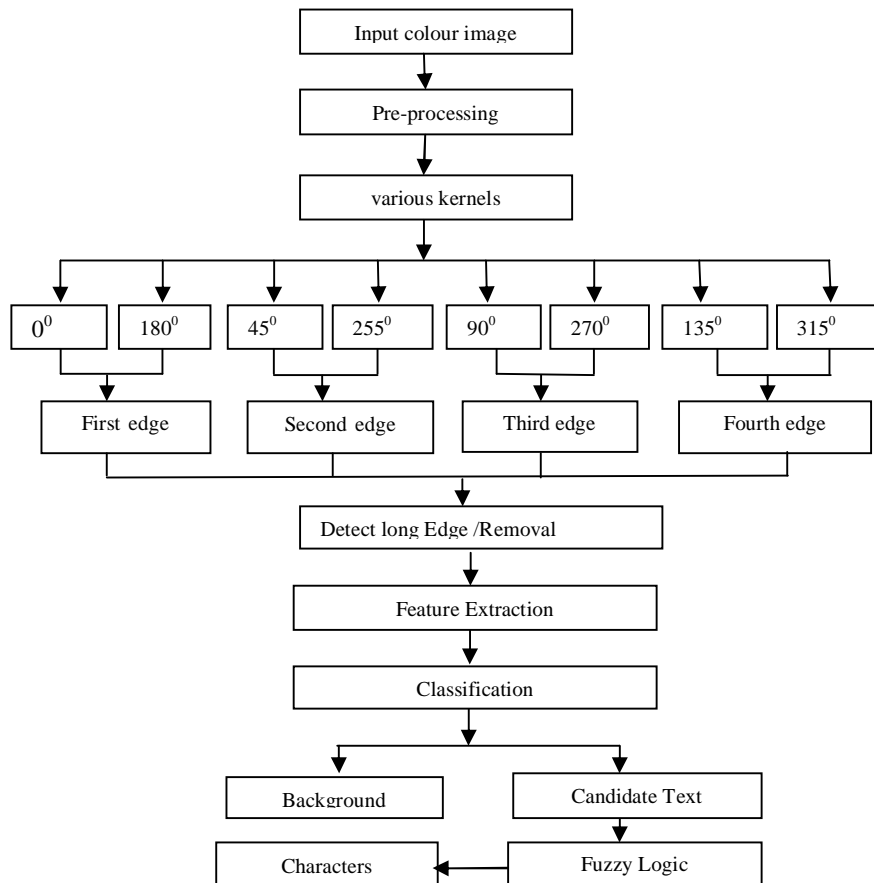


Fig. 1 Proposed method

## 1. IMAGE PRE-PROCESSING

The pre-processing stage is performed on the input images commonly captured by digital cameras (as such images are usually contaminated with noise). Firstly, as in fig 2 shows picture is often in the RGB colour space we first convert the 3 colour components to intensity component (grey images). Then employ a weighted median filter to reduce noise, after filtering step, a great part of noise will be have removed while the edges in the image are still preserved. Also we attempt to expose the picture by manipulating the brightness and contrast of the image. Therefore, we apply histogram equalization to distribute light to get suitable image.



Fig.2 Original image

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

## 2. CANDIDATE TEXT

This stage aims to build a feature map by using three important properties of edges: edge strength, density and variance of orientation. The resulting feature map is a gray-scale image represents the possibility of text.

### 2.1. OPERATIONAL FILTERING.

In our proposed method, in fig 3 we use magnitude of the second derivative of intensity as a measurement of edge strength as this allows better detection of intensity peaks that normally characterize text in images. The edge density is calculated based on the average edge strength within a window. Considering effectiveness and efficiency, eight orientation  $0^{\circ}, 180^{\circ}, 45^{\circ}, 225^{\circ}, 90^{\circ}, 270^{\circ}, 135^{\circ}$  and  $315^{\circ}$  are used to evaluate the variance of orientations, where  $0^{\circ}, 180^{\circ}$  denote horizontal directions,  $90^{\circ}, 270^{\circ}$  denote vertical directions, and  $45^{\circ}, 135^{\circ}, 225^{\circ}$  and  $315^{\circ}$  are the four diagonal directions, respectively. 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  $3 \times 3$  square structuring element to the binary image obtained from the previous step to get joint area referred to as text blobs. A convolution operation with compass operator as shown in figure 2 result in eight oriented edge images.

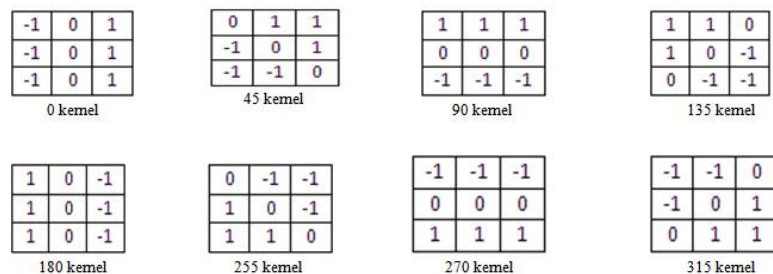


Fig.3 Various kernels of edge detection

After convolving the image with the orientation kernels, we collected the edges detection with them due to the final edges detection as shown in question below.

Edge first= Edge"0"+ Edge "180"

Edge second= Edge"45"+ Edge "225"

Edge third= Edge"90"+ Edge "90"

Edge fourth = Edge"135"+ Edge "315"

Edgetotal= Edge first+ Edge second+ Edge third+ Edge fourth

Feature map that has been created. A weighting factor is associated with each pixel to classify it as candidate or non-candidate for text region. A pixel is a candidate for text if it is highlighted in all of the edge maps created by the directional filters. Thus, the feature map is a combination of all edge maps at different scales and orientations with the highest weighted pixels present in the resultant map. Actually, after edges detected and convolve each edge with Gaussian pyramid. We combined eight regions obtained in each kernel with them, to obtain entire edges detection with whole angles. As below in figure 4 shown entire edges .

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

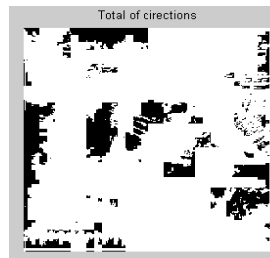


Fig.4 The Total Edge

## 2.2. EDGE SELECTION

Vertical edges form the most important strokes of characters and their lengths also reflect the heights of corresponding characters. By extracting and grouping these strokes, we can locate text with different heights (sizes). However, in real scene under an indoor environment many other objects, such as windows, doors, walls, etc ..., also produce strong vertical edges. Thus, not all the vertical edges can be used to locate text. However, vertical edges produced by such non-character objects normally have very large lengths. Therefore, by grouping vertical edges into long and short edges, we can eliminate those vertical edges with extremely long length and retain short edges for further processing. After, threshold these long vertical edges may become broken short edges which may cause false alarms (positives). So the proposed method uses a two stage edge generation method. The first stage is used to get strong vertical edges by collect edge “0” and edge “180” who described in questions below.

$$\text{Edge } v = \left| \begin{matrix} \text{strong} \\ \text{Edge } v = \left| \text{Ev} \right| z \\ \text{Edge "0" bw} + \text{Edge "180" bw} \end{matrix} \right.$$

Where Ev is “0”+”180” intensity edges image which is the 2D convolution result of the original image with the “0” and “180” kernels,  $\left| \cdot \right| z$  is a threshold operator to get a binary result of vertical edges, it is not very sensitive to the threshold value. The second stage is used to obtain weak vertical edges described in below.

$$\begin{aligned} \text{dilated} &= \text{Dilation}(\text{Edge } v) \ 3 \times 3 \\ \text{closed} &= \text{closing}(\text{dilated}) \ m \times 3 \end{aligned}$$

$$\text{Edge } v = \left| \begin{matrix} \text{weak} \\ \text{Ev} \times (\text{closed} - \text{dilated}) \end{matrix} \right| z$$

Where the morphological dilation with rectangular structuring element of size  $3 \times 3$  is used to eliminate the effects of slightly slanted edges and a vertical linear structuring element  $m \times m$  is then employed in closing operator to force the strong vertical edges closed. The resultant vertical edges are a combination of strong and weak edges as described in question below.

$$\text{Edge } v = \left| \begin{matrix} \text{strong} & \text{weak} \\ \text{Edge "v" bw} + \text{Edge "v" bw} \end{matrix} \right.$$

The results of two stages edges generation method the resultant vertical edge images are already done after use the question in the above. Morphological thinning operator followed by connected component labelling and analysis algorithm are then applied on the resultant vertical edges as described question in the below

$$\begin{aligned} \text{Thinned} &= \text{thinning}(\text{Edge } v) \\ \text{Labeled} &= \text{BWlabell}(\text{Thinned}, 8) \end{aligned}$$

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

Where the morphological thinning operator makes the widths of the resultant vertical edges on one pixel thick. Since high value in the length labelled image represents a long edge, a simple threshold described question the below that used to separate short edges.

$$\text{Short v bw} = \left\lfloor \frac{\text{lengthlabeled}}{E_v} \right\rfloor z$$

### 3. FEATURE EXTRACTION

The "dense edges" are the distinct characteristics of the text blocks with detect possible text regions. Here by finding the edges in the mentioned mages and fusing the edges contained in each image, the candidate text regions can be found, as the result.

#### 3.1. TEXT REGION LOCALIZATION

The localization text regions through image acquainted comes the role of pre-processing as image complement. In general, detecting a target from a color image is a complex process. In order to identify the outline of a target easily, the input image converted into a grayscale format. Grayscale image conversion operations are useful for enhancing the visibility of brightness variations among gray levels in regions of a digital image fine details are not better. Filtering is used to soften the effect of noise on the image. Binarization is the conversion process of grayscale image into black and white image[8].The conversion of gray image into black and white is done based on threshold value, as shown in the fig 4.

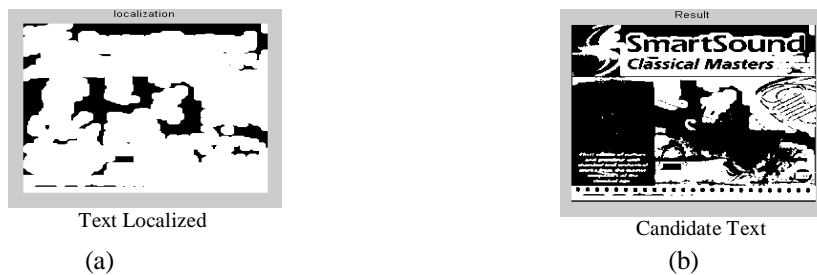


Fig.4 Candidate operation (a) Text Localized (b) Candidate Text

As a result, the area as well as the width and height ratio of each text blob cannot too small. In our proposed which do not contain text. First constraint is used to filter out all the very small isolated blobs described in next equation, where it uses a relative area value instead of the absolute value for each blob. Therefore, there is no specific limitation for the font size which makes more sense for the real world scenes.

$$\text{Area}_{\text{regions}} \geq (1/25) * \text{Area}_{\text{max}}$$

The second constraint described in next equation filters out those blobs whose widths are much smaller than corresponding height. A threshold of "0.25" is considered for the worst case, such as single western letter "1" or English digit "1". Experiment results show that this value(0.25) is not sensitive to different situation. The detection filter image is shown in

$$\text{Ratio}_{\text{win}} = (\text{width}_{\text{region}} / \text{height}_{\text{region}}) \geq 0.25$$

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

## 3.2. BOUNDARY BOXES GENERATION

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. figure 5 shows the boundary boxes of text regions. In order to avoid missing those character pixels which are lying near or outside of the initial boundary, width and height of the boundary box. As known, regions with text in them will have significantly higher values of average edge density, strength and variance of orientation than those of non-text regions. We exploit these three characteristics to refine the candidate regions by generation feature which suppresses the false regions and enhances true candidate text regions. This procedure is described in below.

Candidate=Dilation (Short v bw)  $m \times m$

Where, the morphological dilation with a  $m \times m$  structuring element employed in the selected short vertical edge image and used to get potential candidate text regions

## 3.3. DETERMINE BORDERS

The borders should be removed; this will reduce the image size. Only the rectangular part of the image which contains text or texts in the image that will remain as shown in fig 5. That means, there are many connected components which assigned by pixel "1" while the background assigned by "0". To remove the borders, there four stages that are involved. Firstly, the start from up-bottom that from first row, if the row does not contain the pixel "1", it will be removed and this will continue until there exists pixel "1" in a row. In this case, this row containing the pixel "1" is the border of image in above part. Furthermore, the same case is applicable to bottom-up until there exists a row containing the pixel "1"- this is in the horizontal stage. Afterwards, the same operation is made in the vertical, but the differences in column instead from row, which start from first column then continues until there exists the pixel "1" in any column may be first column, that means from left-to-right. Moreover, the reverse make with right-to-left.



Fig.5 Box Text

## 3.4. DIVIDE TEXT INTO ROWS

Now, after determine the region of text we are going to divide text into rows and also we used dilation and erosion of separate reaches characters as found in word "masters" in fig 5, the result is shown in Figure 6. After moving the borders, the area will now be divided into rows. [12] They makes every text to have two lines i.e. every text with start line in the first row and end line in the end row in the text. Furthermore, the start line and end line are different from text to text because of multi-scale variances in texts. So every text depends on the size of the text which has properties assigned for it. That means, length and width of connected components exist side by side i.e. horizontally in many rows. We can say this represents the text whose properties and size are different from the properties and size of another text, or multi-scale. In this base, we can know start line and end line of each text. After that, can deal with it as shown in fig 6



# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

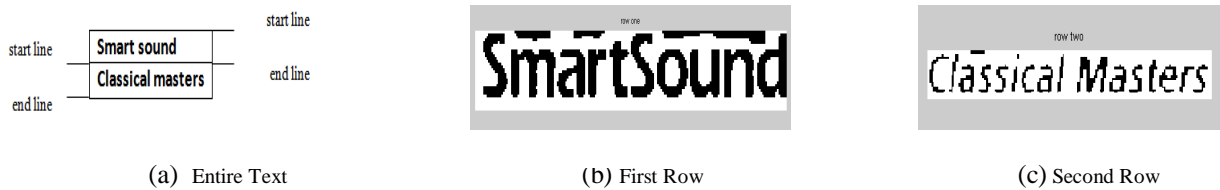


Fig.6 Divide Text

### 3.5. DIVIDE ROWS "LINES INTO WORDS"

The single line is then divided into words .Because that. [12]. They employed the empty area before and after the text are removed, but there main question must answer on it "here is size or scale differences from text to another also space between the words depend on the scale .So, in this problem can answer on this question, we can depend on length and width of connected components if is character and on this basic we know the space between words and between characters in the word based on ratio the length of the width. The word may be a single character or more than that, the size may be different in the same word and it is not necessary that the word has a meaning. As shown in fig 7.

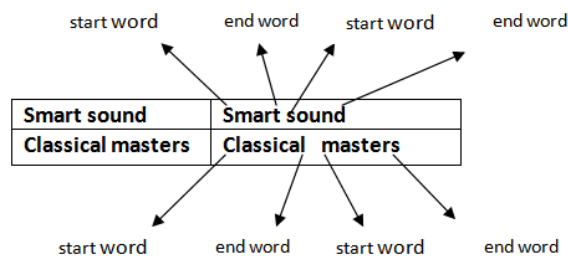


Fig.7 Divide the Rows into Words

### 3.6. DIVIDE WORD INTO CHARACTERS

Each word is then divided into characters and saved in the array, Again on this basic , we have been known the character in each word and each word in text .Afterward , each connected component known then used the fuzzy logic of recognize the character and any connected component not recognize is noise or distortion in the remain image. As shown in fig 8 below

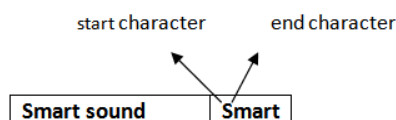


Fig.8 Divide Word into Characters

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

## 4. CHARACTER EXTRACTION

Existing OCR(Optical Character Recognition) engines can only accept simple inputs, i.e., it only deals with printed characters against clean background and cannot handle characters embedded in shaded, textured or complex background. The purpose of this stage is to extract accurate binary characters from the localized text regions so that we can use the existing OCR directly for recognition. In our proposed, we use uniform white character pixels in pure black ground.OCR system available require the input image to be such that the characters can be easily parsed and recognized .The text and background should be monochrome and background-to-text contrast should be high .Thus this process generates an output image with white text against a black background.

### 4.1. CONNECTION COMPONENT

Usually, after become the image into binary image, we now deal with zero and one to represent the foreground and background that use Run-Length Coding Algorithm. The RUN mean take the pixels in same the row in blocks .That mean, every block of RUN represented in horizontal projection, and every horizontal projection calculated from run-length code. The labeling algorithm represent in equivalence table, Resolving the equivalence table has been the focus of most labeling algorithms, with little effort to implement and also minimizes use of memory. Process uses a run-length encoding representation. Conversion of the original binary image to run-length encoded format is easily parallelized by processing multiple rows in parallel. The run-length encoded format is much more compact than the binary image (individual runs have a single label), and so the sequential label propagation stage is much faster than the conventional algorithm. Details of the algorithm are given below. The stages involved in our implementation are as follows:

1. Pixels are converted to runs in parallel by rows,
2. Initial labeling and propagation of labels
3. Equivalence table resolution
4. Translating run labels to connected component

We have been used 8-neighbors where they share at least one corner, the position of it are  $[i+1,j]$  , $[i-1,j]$  , $[i,j+1]$  ,  $[i,j-1]$  ,  $[i+1,j+1]$  ,  $[i+1,j-1]$  ,  $[i-1,j+1]$  and  $[i-1,j-1]$  as shown in fig 9 below

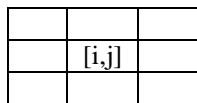


Fig.9, 8-neighborhoods for rectangular image tessellation pixel  $[i,j]$  is located in center of each figure

### 4.2. DETERMINE CHARACTER BY RUN-LENGTH

Here, we have to determine the characters and how many characters in all words and, how many words in each text or line. Finally, the text is extracted from the image .So when we have connected component and have identified pixel to it. As shown in Fig 10 below

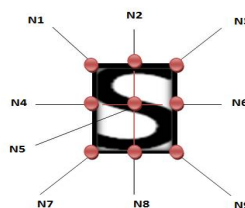


Fig.10 Identified pixel

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

let's suppose the following

N1=upper left N2=upper center on N3=upper right N4= center left N5=center center on N6=center right  
N7=lower left N8=lower center N9=lower right

Now, send identified pixel from N1 until N9 to fuzzy logic that has ten input and one output as shown in figure 11 and 12 below

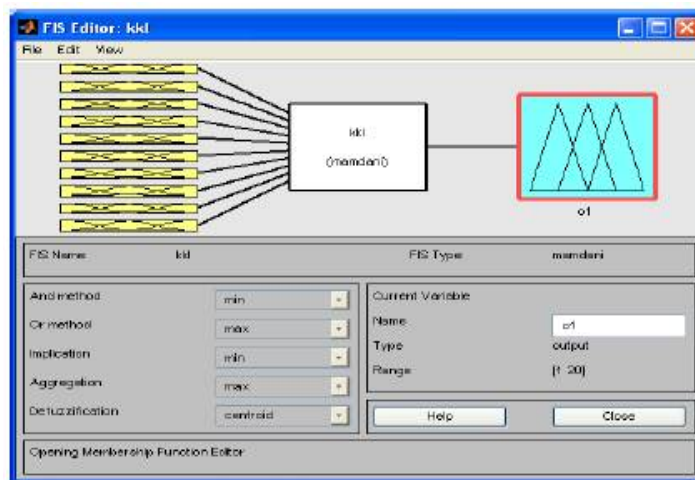


Fig.11 Ten Inputs and One Output

### 4.3. FUZZY LOGIC

We have been used fuzzy logic of identify the character or connection component in the image, after determine connection component which represented rectangle on it. Then we must know which pixel each corner and other pixels used in identify the connection component as shown in fig 12 below

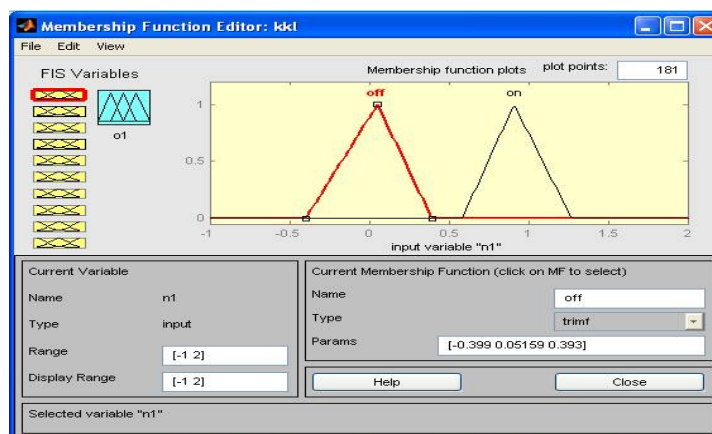


Fig.12 Input one

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

Afterward , we given of every pixel identify value between 0 and 1 of whose location that is nine pixel and send to fuzzy of make recognize the connection component then fuzzy receive this values of nine pixel and whose status “on” or “off”, after that fuzzy logic determine either character or noise. Below fig 13 design input fuzzy of location of pixel and status of pixel.

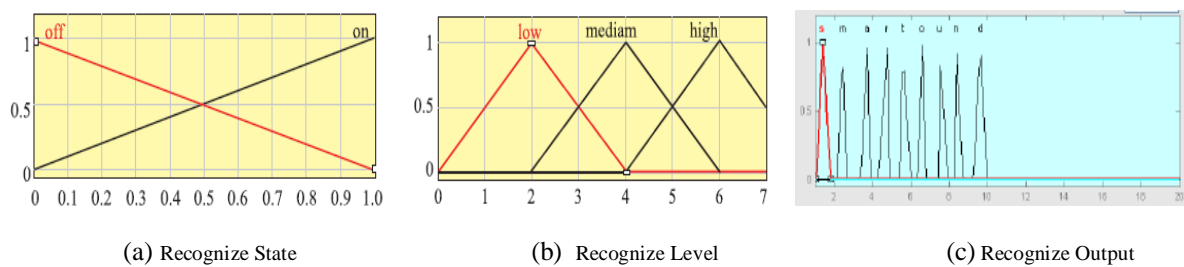


Fig.13 Membership function

As noticed above of Figure 14 shown output we can see output of character "s" is between 0 and 2 ,so we can know this is character "s" same scheme with all characters done with output of text is as shown in Fig 14 below

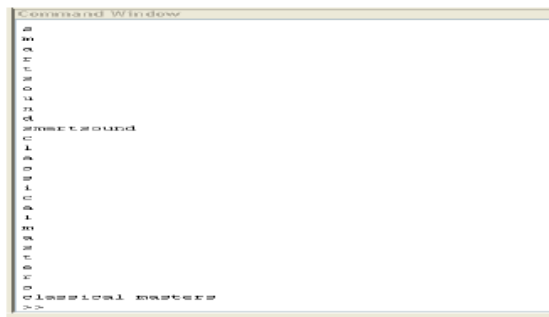


Fig.14 Output of extracted text

## IV. EXPERIMENT RESULTS AND DISCUSSION

In order to evaluate the performance of the proposed method to text-detection ,text-segmentation and extraction characters has been tested extensively on large, diverse, and difficult set of images. With different font sizes, perspective and alignment under different lighting conditions. Fig 15 show some of the results. We can see that our proposed method used in localization and extraction the text from with different font-size, orientation and perspective.



(a) Cover book

(b) Cover book

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018



Fig.15 Experiment results

Although there is no generally accepted method that can be used to evaluate the performance of the text-detection and extraction method, we use several metrics used by others to facilitate a meaningful comparison. In our proposed method, we compute the accuracy of the algorithm output by manually counting the number of correctly located characters, which are regarded as ground-truth. Precision rate and recall rate are quantified to evaluate the performance given by next equations (1),(2) and (3).

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

**Precision**=(correctedlocated/(correctedlocated+falsepositiv))\*100% (1)

**Recall**=((correctlylocated)/(correctedlocated+falsenegative))\*100%. (2)

**Falsepositive**=((falsepositive)/(correctlylocated+falsenegative))\*100% (3)

**Correctly located:** means the corrected localization of a text exists in an image .On other words, which is exactly located on place of the text.

**False positive:** relates to a localization process of the real intended text

**False negative:** relates to a localization process of non-text object.

Table 3 shows the performance evaluation of our proposed method with several existing methods, where our proposed method shows a clear improvement over existing methods. In this table, the performance statistics of other are cited from published work. Considering a 95% confidence intervals , it appears that [11][19].However , those methods are applied for different types of images.

Table 3 performance evaluation

No sample	Resolution	Precision rate	Recall rate	False positive
(a)	256x255	75%	66%	33%
(b)	150x190	68%	65%	53%
(c)	370x280	75%	58%	30%
(d)	169x170	80%	83%	80%
(e)	406x307	67%	67%	50%
(f)	250x190	80%	60%	90%
(g)	460x480	60%	70%	65%
(h)	290x220	62%	67%	60%
(i)	408x300	55%	62%	40%
(j)	373x498	60%	71%	65%

Now, we have been evaluated the performance of the text extraction .So we assessed the accuracy of our algorithm output by manually counting the precision of the number of falsely detected text and the number of detected text from the image. Given the marked ground-truth and detected result by our algorithm , we can the recall and false alarm rates by using the equations (4)and (5).

**Recall**=((num of correctly detect text)/(num of text))\*100% (4)

**False alarm rate**=((num of correctly detect text)/(num of detect text)\*100 % (5)

**Num of correctly detected text:** this is that number of how many characters detected correctly in the text.

**Number of text:** means the number of characters in the text

**Number of falsely detected text:** means the number of characters in the text algorithm failed to detect

Table 4 performance evaluation

No sample	Resolution	False alarm rate	Recall rate
(a)	256x255	25%	80%
(b)	150x190	5.26%	95%
(c)	370x280	33%	75%
(d)	169x170	2%	99%
(e)	406x307	44%	69%
(f)	250x190	1%	99%
(g)	460x480	53%	65%
(h)	290x220	20%	80%
(i)	408x300	40%	70%
(j)	373x498	30%	75%

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

Table 5 performance evaluation

No sample	Resolution	Feature	accuracy
(a)	256x255	Cover book	80%
(b)	150x190	Cover book	95%
(c)	370x280	Poster	75%
(d)	169x170	Poster	99%
(e)	406x307	Poster	69%
(f)	250x190	Note	99%
(g)	460x480	Poster	65%
(h)	290x220	Poster	80%
(i)	408x300	Cover product	70%
(j)	373x498	Cover book	75%

## V. CONCLUSION

In this paper, we propose an edge-based text region extraction and recognition algorithm which can localize, extract and recognize from real scenes. Experimental results show that our proposed method is very effective and efficient in text-detection and extracting text-based features. It is robust with respect to font sizes, styles, color/intensity, orientation, alignment/layout and perspective of text and can be used in images. Recently, optical character recognition (OCR) became widely used in extraction character from image. There are very large from approaches of text extraction from picture. Research on extract text via use edge-detection by use kernel with eight kernels also, identified pixel which use of recognition the character after determine the connected component. Then, fuzzy logic used as character classification after send identified pixel to fuzzy logic.

## VI. CONTRIBUTION

Although we did not conduct controlled experiments to study the behavior of the proposed method for each individual parameters, our test image set contains a wide variety of images and the performance was excellent overall. The proposed method is:

1. The kernels implemented are containing to kernels eight used of detect the edge of text in the image.
2. The proposed method was whose suitable performance in operation text detection in complex image.
3. Run-Length Coding was fast and suitable in determine connected component and reduce of time and small memory.
4. Distinguishes text regions from texture-like regions, such as windows frames, wall patterns, etc by using variance of edge orientations.
5. Robust with respect to aspect to font sizes, and styles, uneven illumination, perspective and reflection effects

## VII. FUTURE WORK

There are several suggestions and future works that can be done to improve this work in future the suggestion are:-

1. Improve in method that is focus on extract text from large complex image such as very similar color between text and background.
2. Improve the method that use in extract many text different it in style and formal.
3. Add more features to the software like: improve in function of separate characters and remove the noise.
4. Using a suitable existing OCR technique to recognize the extracted text from landmarks.
5. Applying an affine rectification to correct the perspective distortion caused by a camera view angle in order to increase the text recognition rate.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

## REFERENCES

- [1] J. Samarabandu, and X. Liu , "An Edge-Based Text Region Extraction Algorithm for in Door Mobile Robot Navigations". International Journal of Signal Processing, no. N6A 5B9, pp.273-280, 2007.
- [2] B.Gatos, I.Pratikakis, K.Kepene and J.perantonis. "Text Dectection in Indoor/Outdoor Scene Images". Proceedings to National center for scientific research "Demokritos",no. GR-153 10 Agia Paraskevi, Athens, Greece.pp.127-132,2003
- [3] T.Han, Y.Chien, C.Lun , "Comprehensive Motion Videotext Detection Localization and Extraction Method",National Central University ,Taoyuan County 320 , .no.0-7803-9584-0.,pp.113-116,2006
- [4] Q.Ye, W.Gao, W.Wang and W.Zeng, "Roust Text Tetection Tlgorithm in Tmages and Tideo Frames",ICICS-PCM, no 0-7803-8185-8,pp.1-5
- [5] B.Shanwar and N.Zarka , "Multiscale Edge-Based Text Extraction from Complex Image",ResearchGate,no 1-4244-0367-7.,pp.1-26 ,.2016.
- [6] R.Farhoodi and S.Kasaei, "Text Segmentation from Image with Textured and Colored Background. Sharif University of Technology, Tehran, Iran,PP.395-398,2005
- [7] J.Liu,J.Sun and S.Wang, "Pattern Recognition: An overview". IJCSNS International Journal of Computer Science and Network Security, VOL.6 No.6. 130012 Changchun, pp.57-61 June 2006
- [8] Holland and J.H, " Adaption in Natural and Artificial systems". University of MICHIGAN press, Vol.5 No.21, pp.256-347,December 1, 2014
- [9] V.Dutt,V.Chadhury and I.Khan, " Different Approach in Pattern Recognition", Computer Science and Engineering,no.10.5923,pp.32-35,2011.
- [10] S.Asht,R.Dass, "Pattern Recognition Techniques:Review":International Journal of Computer Sciences and Telecommunication,Vol.3,pp.25-29 2012.
- [11] C.Tu, J.Liang and T.D, "Adaptive Runlength Coding",IEEE ICIP,no. 0-7803-7622- 6,pp.665-668,2002.
- [12] M.Alata and M.Al-shabi, "Text Detection AND Character Recognition Using Fuzzy Image Processing",Journal of Electrical Engineering.Vol.57,No.5. Jordan.pp.258-267, 2006
- [13] S.N.Sivanandam, S.Sumathi and S.N.Deepa, "Introduction to Fuzzy Logic Using Matlab",Springer, Inc. 2006.
- [14] M.A.Rabbani and C.Chellappan, "Fast and New Approach to Gradient Edge Detection".International Journal of Soft Computing , Vol.2 :pp.325-330,2007
- [15] D.Kofi, R.Andrew, and P.Jonathan, "Run-Length Based Connected Component Algorithm for FPGA Implementation". University of Lincoln,England,pp245-250,2007
- [16] K.Wang and J.Kangas, "Character Location in Scene Images from Digital camera. Pattern Recognition".Journal of the pattern recognition ,pp 2287-2299,2003
- [17] K.Kim,K.Jung and J.Hyung, "Texture-Based Approach for Text Detection in Image Using Support Vector Machine and Continuously Adaptive Mean Shift Algorithm". IEEE Transtion On Pattern Analysis and Machine Intelligence, Vol. 25, No. 12,pp.1631-1639,2003
- [18] 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 Scene Images Using Hierarchical Feature Combining and Verification", Computer Sciences, 17th International Conference on Pattern Recognition,. no.1051-4651/04,pp.1-4,2009 .
- [19] R.Lienhart and A.Wernicke, "Localizing and Segmenting Text in Images and Videos", IEEE Transtion On Pattern Analysis and Machine Intelligence,Vol 12,No 4, PP.256-268,2002
- [20] C.Liu, C.Wang, R.Dai , "Text Detection in Image Based on Unsupervised Classification of Edge-Based Features",Eight International Conference on Document Analysis and Recognition.no. 0-7520-5263,pp.238-245,2005
- [21] D.Chen,H.Bourlard and J.Thiran, " Text identification n complex background using SVM" .Signal Processing,UTM,pp.621-626,2001.
- [22] Yu.zhong, k.kearu and A..K.Jain, "Locating Text in Complex Color Images", Pattern Recognition,Vol.28,pp1523-1535,1995
- [23] J.Gao and J.Yang, "An Adaptive Algorithm for Text Detection from Natural Scenes",Pittsburgh PA,no 15213,pp.1-6, 2000.
- [24] T.Yasuaki,Y.Nakano and M.Maruyama, "Digit Classification on Signboards for Telephone Number Recognition" . Seventh International Conference on Document Analysis and Recognition.pp.1-5,2003
- [25] X.Chen, J.Yang,J.Zhang and A.Waibel, "Automatic Detection of Signs with Affine Transformation". University Mobile Technologies.pp.1-5,2003.
- [26] H.Ezaki, M.Bulacu and E.Schomaker, "Text Detection from Natural scene Images",Snternational Son ference on Sattern Secognition (ICPR'04),pp.683-686.2004.
- [27] K.Matsuo,M.Ueda and C.Michio, "Extraction of Sharacter String from Scene Smage Binarizing Local Target Aarea.Transaction of the Anstitute of Electrical Engineers" .no 122,PP.232-241,2004.
- [28] V.Wu,R.Manmatha and E.M.Rissman, "Text finder:an Automatic System to Detect and Recognize Text in Image".IEEE Tansaction on Pattern Analysis and Machine Intelligence,vol.21, no.11, PP.165-171,1999.
- [29] Q.Ye,Q.Huang,W.Gao and D.Zhao, "Fast and Robust Text Detection in Images and Video Frames" .Image and Vision Computing,pp.566-576,2005.