CHAPTER 1

INTRODUCTION

Chapter 1

Introduction

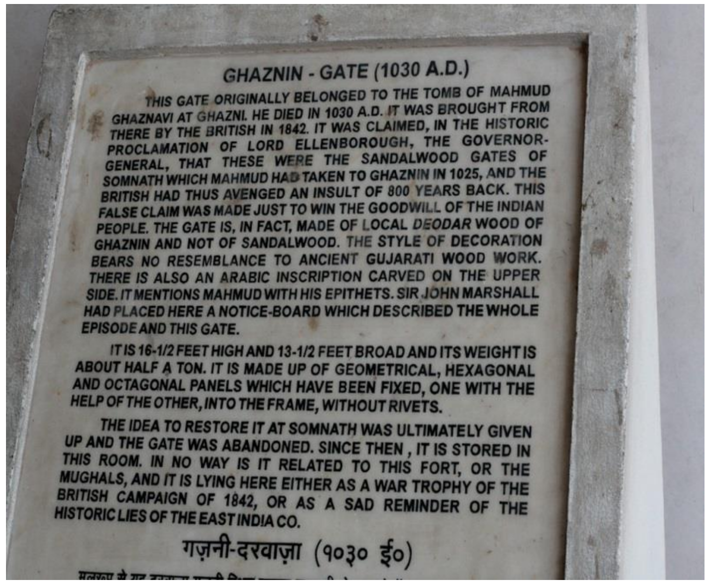
Ancient monuments and ancient documents like inscription of monuments have a lot of information in them about the history all we have to do is to extract that information from them.We will be able to know much about the historical dynasty if we become able to extract the text which written no these historical monuments.So from past decade a lot of work has been carried out for this purpose but the main work which has been carried ot is basically related to the degraded document and a little work has been done for the inscription images.Such kind of inscription images have many problems related to it and such problem may be due to ageing.These problems or we can say such degradation includes missing of the edges of the text or may be the complete text or it may be the fading of the color of the text or it may be the poor condition of the monument.Due to such conditions the text written on such monuments or we can say the text engraved on such monuments may have the poor contrast condition in the region which contain text.For the regions having high edge density and strength simple edge-based approaches are also considered useful. This edge-based method give good result if background is not complex,but for the inscription images background is complex,thus this method cannot be used directly.

Badly degraded images which is having high inter/intravariation between the background and the text region,the segmentation of the text becomes a big challenge.

This thesis is basically about the method which is used for the binarization of documents as well as for the monuments inscription imagesThe motivation for the development of this method come from the need which is basically the digitization of such documents and inscription and allow the public for their easy access such as in meuseum and libraries.This method which we have proposed here uses two two methods which are local image gradient and local image contrast in adaptive manner.Both the mothed are used according to the need of the image.First with the help of these method we develop an contrast image of the input image then to incorporate the edge information of the text to make the result more better we calculate the edge map.The binary image of the contrast image is developed by using the well known technique of the thresholding that is is otsu's global thresholding.Now with the help of canny edge detection we develop the edge map and in next step we combine both of these maps.Now in the next step the local thresholding is also used to separate the text from background,and finally we will use some post processing technique for the betterment of the result.These post processing process involve the use of the morphological operation on the final result.in this way the method is having the less no of parameter and involve the less calculation and it is quite easy and robust as well.

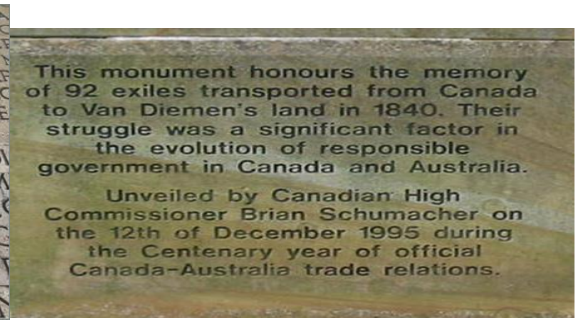
Binarization of image is performed in the preprocessing stage for the purpose of document analysis and it is performed to separate the foreground text from the document background.As we know the fast and very accurate image binarization technique is very important for the image processing purpose which as the optical character recognition.The problem become more complex when there is very small difference in the foreground(text) and the background or we can say is very marginal or can be said that background and foreground are very much similar.This type of case be seen in the images of the inscription of the historical monuments which are taken by the camera such image is shown in the fig.2(d).We can see from these images that such inscription found engraved into stone and in many cases projected out from stone or like material of this etc..Some images have reasonable color difference between text and the background so they can easily give good result but some of the images do not have the color difference between text and background in such cases previously available method does not give good result.

Fig. 1. Five degraded document image examples (a)–(d) are taken from DIBCO series datasets and (e) is taken from Bickley diary dataset.



(a) (b)

(c) (d)



(e) (f)

Fig1.1.2 six monument inscription images taken at different conditions(a)-(f)

As shown in fig.1,for the text in the handwritten documents which is degraded generally shows a variations in terms of the width of the stroke,as well as brightness,connection and background also.As shown,in historical documents which are genrally ruined by the bleedthrough as in fig.1(a) and (c) here the ink of the other side of the image seeps through to the front which makes document unclear.In addition,historical documents are having differnet types of documents degradations which induce the document thresholding error and hence making degradation of document image binarization a big challenge to all the techniques which are already present.

This thesis presents a method of binarization which is basically the extention of the previous local maxima-minimum method and the technique which is used in the latest DIBCO 2011.The proposed method first has been applied on the different degrade documents images and then it has been applied on the monuments inscription images. The method proposed here is very simple,as well robust and it is having the capability of handling different types of ruined document images with least parameter tuning.This method makes use of the adaptively obtained contrast image that is an adaptive combination of the local image contrast and the local image gradient and therefore it is immune to text as well as background variation which is caused by differnet kind of degradations of document,but when we apply the same proposed method on the inscriptions it ives good result in many cases but fails in some cases which is basically the limitation of the method.The proposed method addresses the problem which is over-normalization of the local maximum minimum algorithm.On the other hand,the parametes which has been used in the algorithm can be obtained adaptively .

CHAPTER 2

LITERATURE REVIEW

Chapter 2

Literature review

2.1 Introduction

For historical document and ancient monuments incription image collections,it is quit common for libraries to provide public access . These document and inription images require specialized processing treatment for removing the background noise and become more clear to understand or to read. For most of the document image analysis,image binarization is the very initial step which is basically the conversion of grey image into binary image and it motto to differentiate the text from background regions.binarization play a very important role in document processing task because the extent of success which is character recognition as well as segmentation is very crucial.

Generally the image binarization methods for documents are categorized in main two class:

(i) global thresholding

(ii)local thresholding.

In the case of global approach,one value of threshold is selected for the whole image.Thresholding which is global give the good results in case when there is good separation between the text and the background.While in the case of historical monument inscription documents, there exist such degradation that dimish robustness of this class of binarizationExample of such kind of degradation include shadows and non-uniform illumination, ink seeping, smear and strains. To handle such degradations,now-a-days we use local information or we can say local threshold value pixelwise in an adaptive manner.in case of local binarization it ignores the edge property and give result that is erroneous due to the creation of fake shadows. For such kind of case,there exist other approaches that incorporate edge information whereas there they find seeds near the image edges and present an edge connection technique.In order to obtaine an effective solution for various degradation problems,use the closed edge map,obtaine binary map and fillit with seeds.The adaptive local binarization five satisfactory results but there is not a particular standard method exist for this.

2.2 Adaptive binarization

As its name shows this method uses the global thresholding as well as local thresholding but in adaptive manner.Here when we are having the input document image which is historical then we apply the global thresholding to the entire image and this global thresholding is applied in an iterative manner to the entire image now we detect the region which still containing the noise after this a local threshold value is calculated and apply to those area which is noisy in nature.In this manner we keep on checking if there is noise present or not and apply the iterative global threshold .This technique is having the following steps.

(a) First apply the IGT to entire image.
(b) Detect the area which still contain noise
(c) finally re-apply IGT to the area which is noisy.

The noisy area contain more black pixels on average in comparision with other areas,this fact is utilized for detecting the area which still contain noise.The entire image is divided into fixed nxn size.In every segment frequency of black pixels is calculated.The area which satisfy the following condition is selected as noisy area.

$$F(s)>m+ks$$

Where f(s) is the number of frequency of black pixels in segment 'S','m' and s' are known as the mean and standard deviation of black pixel in the whole image.By utilizing row by row labelling algorithm the selected segments form areas.The obtained noisy areas are re-applied separately by using IGT to each area.If the iteration exceeds the number of iterations required for the IGT on entire image,the iteration stops or when the following condition is satisfied.

$$\left| T_i - T_i - 1 \right| < 0.001$$

T_i is threshold in i_{th} iteration

Here 'k' is the sensitivity parameter.If k is large it means less area will be selected,hence the area which still needs improvement will not be selected.Whereas ,if k is small then more are will be selected.In this case the area still need improvement will be selected along with the area in which noise has been removed already.If we further apply the IGT to these noiseless area,it will increase computational time as well as the cost of application of IGT.The size of the segment nxn is an important factor for successful re-application.Small segment result more area but small area.hence it become possible to adapt the area which is having noise in more detailed manner.But these resulting small areas will not provide sufficient information for successfully re-applying the IGT algorithm.Bigger ara are selected for large sizes.the part of image which is still having noise,these bigger area cannot be easily adapted.As a result,the final resultant image may contain neighboring areas that have dissimilar amount of background noise.

2.2.1 Iterative global thresholding

In this iterative global thresholding,it selects a global threshold for a document image based on iterative procedure.Following steps are followed in each iteration.

From the past decade,the transformation of greyscale images to black and white is one of the major problem.Suppose greyscale image is applied as input,here tones of the foreground is out of range over background.The input image can be expressed as follow

$$I(x, y) = r, \quad r \epsilon s[0,1] \quad \text{-----------} \quad (1)$$

Here x and y represent the horizontal and vertical coordinates of the above described image,where r can have the value between 0 to 1,where r=1 reprent white colour and r=0 represent black colour.The intermediate tones are shifted to foreground or background.Adocument image includes very few pixels of useful or we can say foreground as compared to the size of the entire image i.e.(foreground+background),the presented algorithm is based on this fact.Here we can say that,rarely the amount of black pixels exceeds the 10% of the total pixels in the entire image.By utilizing this fact,it is supposed that the average value of pixel values of document image is determined basically by the foreground even if the document image is quit clear.By utilizing this fact the prposed method consist of the two procedures and these are applied alternately.In the initial stage the average pixel value of the entire image is calculated and then subtracted from the image,now in the next part of the algorithm the histogram is stretched so that remaining pixels are distributed in all the grey scale tones.Some of the pixels are moved from the foreground to background after each iteration.Iteration stops when the following criterion satisfies

$$\left| T_i - T_i - 1 \right| < 0.001 \quad \text{-------------}(1)$$

This explained algorithm works good on the historical document images in which foreground is darker than background.This takes very less time as well as does not require any complex calculation.

2.2.2 Mathematical analysis

From equation (1) the value of T_p threshold for the image MxN is calculated and it is given as follow:

$$T_i = \frac{\sum_x \sum_y I_i(x,y)}{MxN} \quad \text{-------------} \ (2)$$

The above equation is used to find the mean value of the image I(x,y),in this I_p is the grey scale image.As we know the fact that 1s stand for background and 0s represent foreground,the equation (3) is used for the subtraction and gives the image after subtraction which represented by $I_s(x,y)$.

$$I_s\,(x,\,y) = 255 - T_p + I_p(x,\,y) \qquad \text{------------ (3)}$$

Here equation (4) is used to perform image equalization after subtraction

$$I_i\,(x,y) = 1 - \frac{1 - I_S(x,y)}{1 - Ei} \qquad \text{------------- (4)}$$

In the above equation I_s is just as given by the equation (3) and E_i gives the minimum pixel value in the I_s during i-th repetition just before the histogram equalization.

2.3 Morphological approach

Binarization is one of the important step in the image processing,it basically refer to the extraction or separation of the text from its background region.Many binarization algorithm have been proposed since decades but till now it remains an unsolved issue as there are different types of problem of degradations and such degradations are nonuniform illumination condition of backgrund,shadows,smearing,these are some of the geradations which creates problem while converting gray image into binary.In case of the handwritten document the ink of the written text of one paper side generally seeps from the other side,variation in brightness of the stroke and its connection are some of the features of handwritten documents.For such kind of the documents here we have presented one of the approaches and the methodology follow the following steps:

1. first is the preprocessing,
2. second is the determining the foreground or text areas and
3. finally the localization of the text area.

2.3.1 Pre-processing for approach

This step is impotant step in case when we are performing the binarization of historical document images.By doing such pre processing we generally make the background region as well as the text area smooth or even,this is a kind of pre filtering by which the uniformity or the smoothness can be increased.To do so, we first apply the well known filter which is median filter such as Y=med(X),here Y represents the resultant image,W is the structural element which we will use and which is having dimensions 3x3 and X represents the original image.In the original image possible noise in the intensities are suppressed as the result.Howere this non linear spatiall filter is having one advantage that it does not blur the edges of the regions which is having the text,as a default averaging filter do.To compensates for non-uniform background illumination and for each text region to provide similar intensities, we apply the top-hat- reconstruction procedure.By erodind Y^C,we specifically make initial estimation of the background,i.e. the complement its which is Y and this can be done by using a sctructuring element which is of quite large size such thar the disk which is having the radius which is equal to 25 pixels.As we all know that in the case of hand written documents it is having quite different nature so there is very less chances of any structuring element to be fitted in such documents thus the application of the erosion process will discard the desired text area and will result in the image which is represented as Z and this image Z can be used as the as the marker so there must not be any value which is left and which belongs to Y^C.

Now our aim is to obtain the background and this is done by a particular procedure.rst we have to calculate the image Y^C then the resultant image is subtracted from this Y^C.The main task is to construct the Y^c and it s obtained from the Z.So the final result which we have is basically the filtered image.

2.3.2 Foreground area estimation

Here we have to detect the text area and for this we will use the best known thresholding technique which is otsu's technique.So first we have to determine the value of the threshold which we will use as the global threshold.9*t is the value which we have used as the global threshold value in this case and we will use this value where we are confident that this region belongs to the text region.So most of the text is covered by this threshold but few are not covered but the large number of the pixel are covered by this.

To do so,we have to apply the reconstruction transform beginning from the seed image and we go on developing the regions upto where the neighboring pixels which is having intensity lower than a predefined value which is equal to 1.1*T.The initial assumption of the text is included in the resultant image.So from all this we can say that ground truth and the results which are intermediate are quite different.

2.3.3 localization of text area

Now we have to localize the text with the accuracy such that it must contain its edges also so to deal with such problem we have to first obtain the foreground that is containing the text area after that we have to apply some more processing technique to make the result more clear.Initially what we have as foreground is covering most of the the pixels that belong to the text area.We are know this fact that all text pixels are surrounded by the negative text pixels so we have to use this fact.As it is very much clear that BW_1UI is having some of the text pixels which are missing that's why we use BW_1UI as the mask or we can say the marker after doing all this what we get is the binary image wich is represented by the R and it is having starting text pixels,noisy area and also the parts which are missing.Therefore we keep only those area which are only connected components of $BW_{1,}$here the benefit of using the second derivative is to keep the connected component which are 8-connected component.

By calculating the smoothness of R,these area can be easily highlighted just as follow

$$M = 1 - 1/(1+\sigma^2),$$

Here σ stand for the local standard deviation.Now want to calculate the global threshold and also probability density function which is required here,so for this we have to plot the histogram and this is basically the normalized histogram and it represents the probability density function.We will get a new binary image and this is obtain by assigning the value of the pixel of BW_1 to the area which is of R equal to some zero value,we will get another binary image which is represented by BW_2

Now in the final stage we do some post processing things to improve the result and this is done for those pixels which are detected as the text pixels.For this purpose what we do is basically convert pixel which is OFF of the 1 into a ON pixel if and only if it is very close to ON.On the otherside convert the ON pixel to OFF if it is of the BW_2

2.4 Efficient method of binarization

This method basically consist following four steps just as shown in the flow chart.in the first step,Wiener filter is applied for the pre-processing. In the next step certain numbers of binariztion results are combined,which result into a binarized image and this is basically the resultant image of combination of several binarization technique.In next step, grey level image's edge information is combined with the binary which is obtained in the previous step.According to a criterion only those pixels are selected which are most probably represents to the text region from all the all edge pixels.After that particular type of smoothing procedure is applied and its basic purpose is to fill the skeletone of the text or we can say the edge map.Finally as we want to improve our result performance or the quality of the result for this we uses a post processing method which is basically the morphological operation,along with retaining the strokeinformation.The whole procedure is described as follow.

2.4.1.Pre-processing for the approach

For the historical and degraded documents grey scale source image pre-processing stage is important,to eliminate noisy areas,smooth the background texture as well as contrast enhancement within background and text areas.The Wiener filter is the most suitable for the above mentioned conditions.it is basically used for the image restoration.Statistics estimated from the local neighborhood around each pixel is used for the pre-processing which is adaptive wiener method.

Let us consider the source image which is grey:

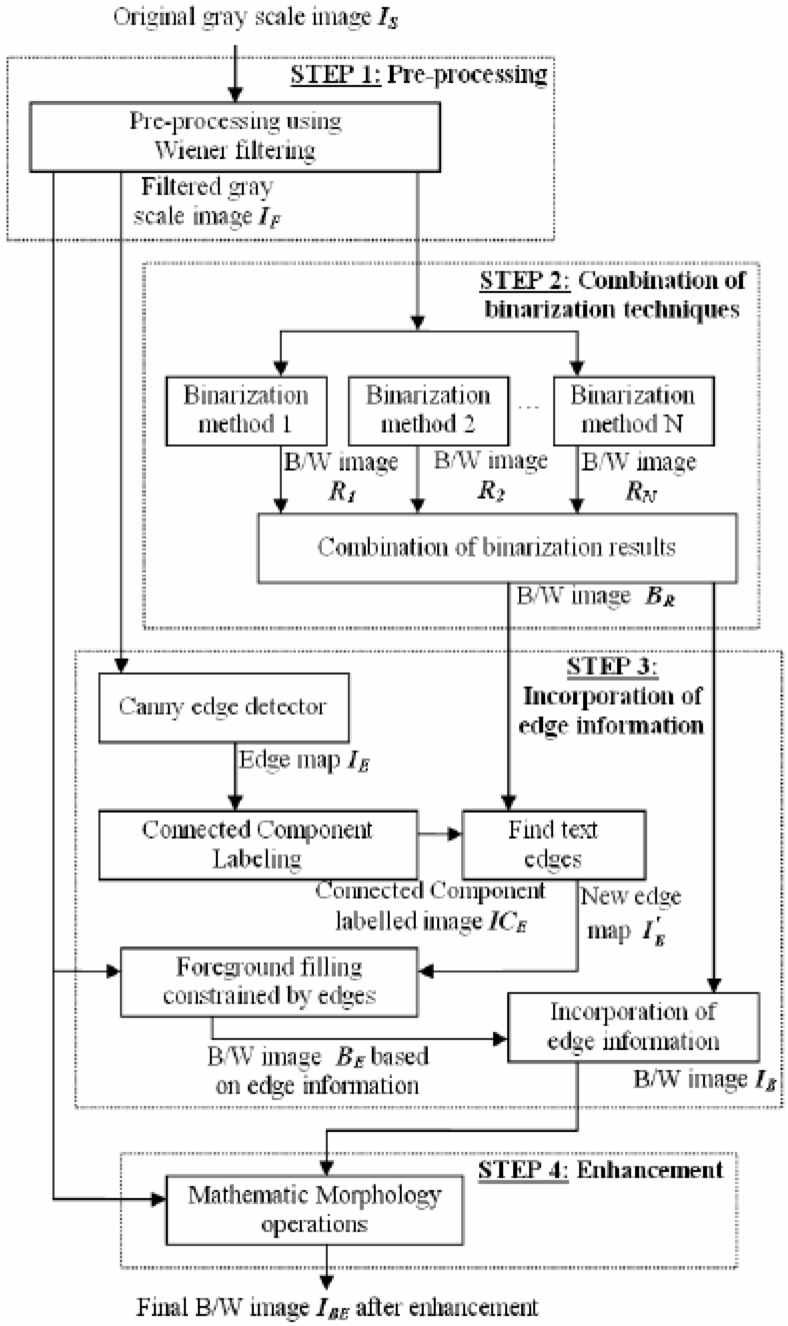
$$I_S(x,y) = \{0,1, \ldots,255\}, \ 1 \le x \le I_x, \ 1 \le y \le I_y$$

Here 0 represent black and white corresponds to 255

To obtain the grey image which is basically filtered and is represented as I_F from the grey source image which is represented as I_S, the following formula is used which is as:

$$I_F(x,y) = \mu + \frac{(\sigma^2 - \nu^2)(I_S(x,y) - \mu)}{\sigma^2}$$

Here μ represent the mean,σ^2 represent the its variance in a size of 5x5 mask surrounding every pixel and γ^2 represents the value of variance for all those pixels which are present in the particular mask.



Block diagram representation of method

2.4.2 Mixture of several binarization technique

Suppose $R_1(x,y)$,$R_2(x,y)$,………..$R_N(x,y)$ corresponds to the result of N different binarization methods,which we have applied on the image which is $I_F(x,y)$.Here we have selected N as an odd number (N=2m+1).Image R_i are expressed as follow:

$$R_i(x,y) = \begin{cases} 1, \text{ foreground} \\ 0, \text{ background} \end{cases}, \text{ where } 1 \le i \le 2m+1$$

Now our purpose to to separate the foreground or to mark the foreground text and this is done by taking only the common portion of the foreground text from several already existing binarization methods.The area is represented as the text if and only if most of the method which are taken into account represent it as text and the result is represented by B_R:

$$B_R(x,y) = \begin{cases} 1, \text{ if } \sum_{i=1}^{2m+1} R_i(x,y) > m \\ 0, \text{ otherwise} \end{cases}$$

2.4.3.Inclusion of edge information

Now at this stage, we will generate the edge map of the filtered grey scale image I_F.there are various method for calculating the edge map of an image.But in our case we use the canny edge detector.Canny uses the sobel operator for calculating the edge magnitude of the grey scale image and one of the reason which makes it better than other is,it uses non-maxima suppression and also hysteresis thresholding.The resultant edge map obtained is expressed as follow:

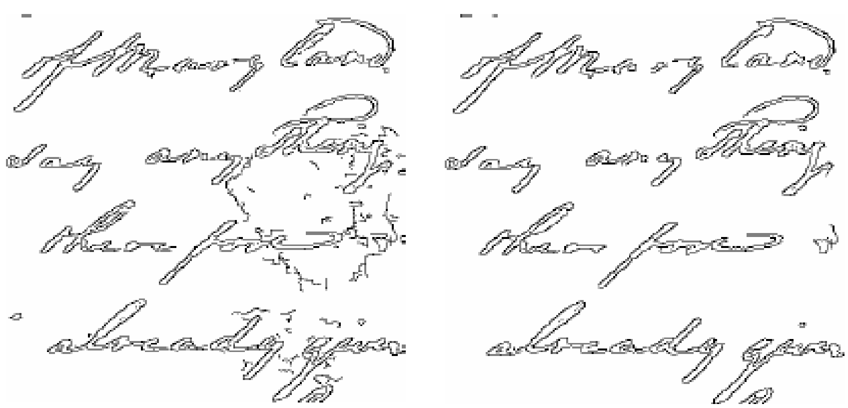
$$I_E(x,y) = \begin{cases} 1, \text{ if } (x,y) \in \partial I_F \\ 0, \text{ otherwise} \end{cases}$$

So we start from the edge map I_E and than we obtain an adapted edge map I'_E and it is only containing the connected components of the I_E which is having the significant overlapping with the binary image B_R.In this manner we remove the edges that do not belong to text regions.Following are the assumptions:

a) main focus is on the image BR so its foreground pixels must be overlap with the edge pixels of he edge pixel of the image IE,for this we use a mask of 3x3 of edge pixels.

b) Here binary image BR and the connected component of I_E has the maximum overlap if and only if the connected component have the overlap which is atleat more than 10 percent

This whole concept can be explained by an example as follow



(a) (b)

Figure.2.4.1Various edge images(a)edge map obtained from original grey scale imageI_F after application of the canny edge detector(b)adapted edge map I'_E which is containing only connected components of I_E that is having significant overlap with the binary image B_R.

Finally to incorporate the edge information in the foreground text to improve the performance or to make results more accurate we have to apply an algorithm.This algorithm is generally called as the run length algorithm.Now how does this algorithm perform,first we are having the edge map of the foreground,now what we can do is to fill this edge map or convert the inside pixels of this edge map which are white in color into black.All this is done by the run length algorithm or we can say that convert the 0 which are inside of the edge map into 1,to perform this there are some conditions which must be satisfy for this suppose starting pixel be (X_1,Y) and it ends at pixel (X_2,Y),so region inbetween these two is converted to black if following condition are satisfied:

i. If the distance in these pixels is short.
ii. If the value of the average intensity of the pixels in between these two ends is smaller than the mean value of the grey level which are outside these end pixels

The above explained algorithm can best be understood by the diagram as shown in figure 2.4.2,we can also explain the above mentioned conditions mathematically he sas follow:

$$x_2 - x_1 + 1 < th \quad \text{AND}$$

$$\frac{\sum\limits_{\substack{i=x_1-2\\ j=y-1}}^{\substack{x_1\\ y+1}} I_F(i,j) + \sum\limits_{\substack{i=x_1-2\\ j=y-1}}^{\substack{x_1\\ y+1}} I_F(i,j)}{18} - \frac{\sum\limits_{\substack{i=x_1\\ j=y}}^{\substack{x_2\\ y}} I_F(i,j)}{x_2 - x_1 + 1} < 0$$

Here th represent half the average height of the character and which is calculated as shown in below figure:

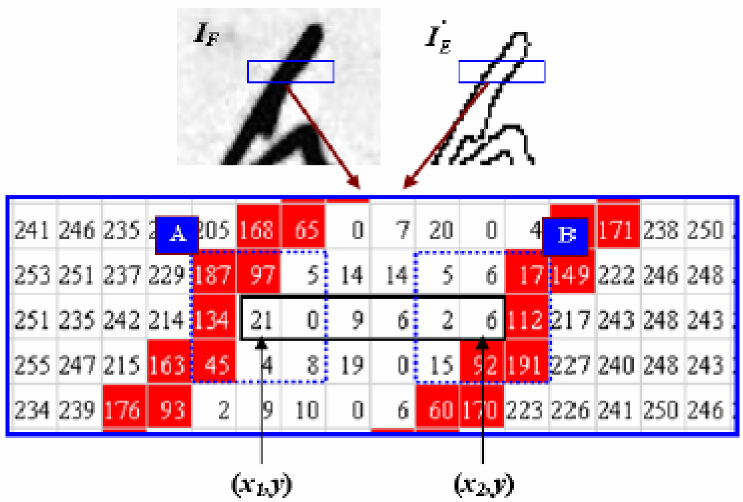


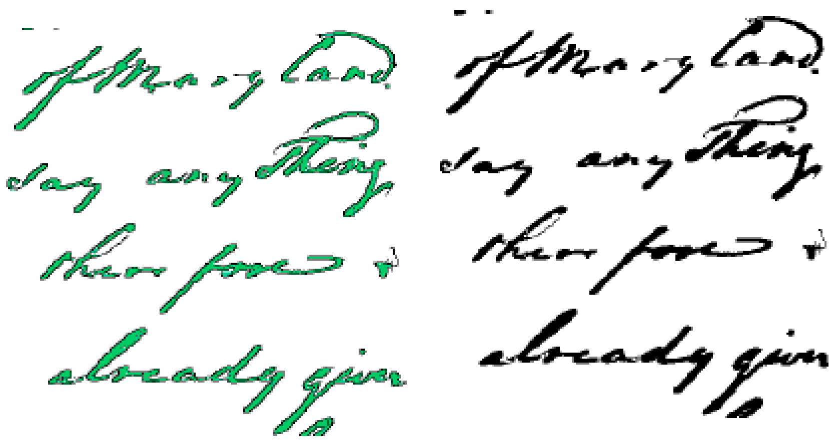
Figure2.4.2 The extension of the run length smoothing algorithm and this is guided by grey scale image I_F.It shows the example how a white run is converted to a black run assuming the value of th=10

As shown in above figure,we convert the white pixels starting from (X_l,Y) and ends to (X_2,Y) to a black pixels because it fulfill the following conditions.

1. The distance between these two pixels is of small length when we compare the distance with the threshold value which is taken as the mean value here th=10
2. X_2-X_1+1=6
3. The total average of the white pixel intensities which lie inside these two end pixels is small as compared to the mean value of the intensities of pixels which are outside these pixels as explained below:

[21+0+9+6+2+6]/6=7.3

As explained above we have applied the run length algorithm in the horizontal direction so in the same manner we can proceed for the vertical direction to completely fill the color in the edge map and resultant image is B_E and this operation of color filling is guide by the grey image which is represented as I_E.but finally for the better results we incorporate the edge information by simply applying the logical OR operation between two images which are represented by B_E and B_R.The final result of the procedure is as shown below.



(a) (b)

Figure2.4.3 Example of incorporating the edge information in the binary image B_R which is based on the adapted edge image I'_E(a)Represents the binary image B_E which is having the filled text area(b)final image containing the edge information

2.4.4 Enhancement

The image I_B as obtained in the previous step,to enhance this image we first apply the conditional dilation with a 3x3 4-connected structuring element.A pixel is considered as the foreground pixel if it satisfy the condition that is if the corresponding grey value does not have difference with the grey value of the already existing foreground pixel in the 3x3 mask.The explained condition ensure that the thickness of the text character remains as it is or there will be no difference in the thickness of the new obtained character but the advantage is that the pixels which is creating the gap in character are removed by applying this condition.The pixel which is represented by I(X,Y) is assigned the pixel value 1 if and only if it satisfy the following given conditions :

$$
\begin{aligned}
& I_B(x,y) = 0 \text{ AND} \\
& (I_B(x-1,y) = 1 \text{ AND} \\
& |I_F(x-1,y) - I_F(x,y)| < 0.05 I_F(x,y)) \text{ OR} \\
& (I_B(x+1,y) = 1 \text{ AND} \\
& |I_F(x+1,y) - I_F(x,y)| < 0.05 I_F(x,y)) \text{ OR} \\
& (I_B(x,y-1) = 1 \text{ AND} \\
& |I_F(x,y-1) - I_F(x,y)| < 0.05 I_F(x,y)) \text{ OR} \\
& (I_B(x,y+1) = 1 \text{ AND} \\
& |I_F(x,y+1) - I_F(x,y)| < 0.05 I_F(x,y))
\end{aligned}
$$

Finally to improve the obtained result quality we apply the post processing which is generally at the end so we use the morphological operator for this purpose and these are the shrink and swell filter. These operator have the advantage of preserving the connectivity and it removes or filter the isolated pixels.$I_B(x,y)$ corresponds to 0 value according to the shrink filter if the following condition satisfy:

$$I_B(x,y) = 1 \text{ AND} \sum_{\substack{ix=x-d \\ iy=y-d}}^{\substack{x+d \\ y+d}} I_B(ix,iy) < t_1$$

Here d corresponds to the size of thewindow and here t_1 represents the value of threshold for the pixel density in the size of the assumed window.In a similar manner swell filter works,pixel $I_B(x,y)$ is converted 1 if the condition satisfy

$$I_B(x,y) = 0 \text{ AND} \sum_{\substack{ix=x-d \\ iy=y-d}}^{\substack{x+d \\ y+d}} I_B(ix,iy) > t_2$$

We have used d=2 and $t_1=t_2=16$ which have been considered after the experimentation.

CHAPTER 3

TECHNIQUES USED

Chapter 3

Techniques Used

3.1 Introduction

There are many techniques[6]-[13] for the thresholding for the binarization of the document images.Global thresholding is suitable where there is clear bimodal pattern in the image,but as we all know we generally do not have such images so it is not worthy to use global thresholding for binarization.While on the other hand local thresholding can be better approach in comparision to the global because it can deal more btterely with images which are having the variantion of contrast in the document images.For examplelocal threshold can be obtainesd or estimated by using the standard deviation as well as mean of the images pixels.These images pixels are considered only in particular windowThese window-dependet thresholding techniques have the main drawback is that this kind of threshold result depends heavily on size of the window used and the stroke width of the character .

3.2 Local image gradient

The local image gradient and local image contrast are very crucial parameter for the extracting the foreground from the document background and this is because generally there is contrast between the foreground pixels and the background pixels.Many of the already present binarization techniques use this fact as we know it is very effective.Ihe local contrast is defined as follows:

$$C(i, j) = I_{\max}(i, j) - I_{\min}(i, j) \qquad (1)$$

here C(i, j) denotes the contrast of the image pixel (i, j),Imax(i, j) and Imin(i, j) denote the maximum and minimum intensities within a local neighborhood windows of (i, j),respectively. If the local contrast C(i, j) is less than a threshold, the pixel is set as background directly. Otherwise it will be considered into text or background by comparing with the mean of Imax(i, j) and Imin(i, j).Although Bernsen’s method is simple,but it cannot work properly on the degraede images with a complex document background.

1.3 Local image contrast

A novel method for document binarization is described as follow and it is known as the by local image contrast and is given as:

$$C(i, j) = \frac{I_{\max}(i, j) - I_{\min}(i, j)}{I_{\max}(i, j) + I_{\min}(i, j) + \epsilon} \quad (2)$$

In the above expresson ε is a very small positive value and its value comes in existence only if the value of the local maxima is equal to 0.In the equation-1,we can see there is no denominator while in above equation there is present and this is known as the normalization factor.This factor compensates the result in case when there is variation within the background.As shown in fig take only the shaded area of the document as in sample image.As expressed in equation-2,small normalization factor is used for the image contrast which are present at the text edge of the strokes.

3.4 Motivation

Monuments,manuscripts and inscriptions which are still present represents the history or we can say it give us the related information of the dynastic history.For conservation and accessibility digitization of these images is very necessary.There are many challenges in digitizing such images,some of the such challenges are the text which are engraved on it are having blurred edge of the textand these are having very less difference between the region which represents the text and the non text part.Different monument inscritptions are engraved in different languages,but we can combined the information gained by inscription with the information from other sources and we can get the information about the history.By digitizing such text,this information which is obtained after digitization, can be easily available in the future if the monuments is not present there.

CHAPTER 4

PROPOSED METHOD

Chapter 4

Proposed method

4.1 Introduction

In this section we describes the proposed binarization technique.For any input image,if it is color image then convert it into grey image then we develop an contrast map which is adaptive in nature and then the canny edge detection is applied and finally edges of the strokes are determined by combining the contrast map which is adaptive in nature also binarized with the edge map obtained from the canny edge detection.After that local thresholding is applied to segment the text,edge pixels of the text is used for the estimation of the local threshold.Further to improve the binarization performance the post processing is used.

4.2 Formation of Contrast Image

We generally use the gradient technique for the edge detection and here we use it for the detection of the stroke edge pixels of the manuscript and also images of the documents which are having the background which is uniform in nature.By using this we generally get the many non-stroke edge pixels from background of the input images which is having certain amount of variations which may be in light contrast or it may be non-uniformity in light or it may be bleed through,in case of degraded documentd image while in case of the inscription lighting variation or contrat variantion,due to aging effect certain noise also present.To get only the edges of the strokes it needs to normalize the image gradient to compensate the image variation in the image background region.

The expression given in equation-2,here its numerator is known as the image gradient which is basically the difference of the maxima and minima of locals.Here the purpose of the use of the local maxima and minima is to suppress the variation of the background.

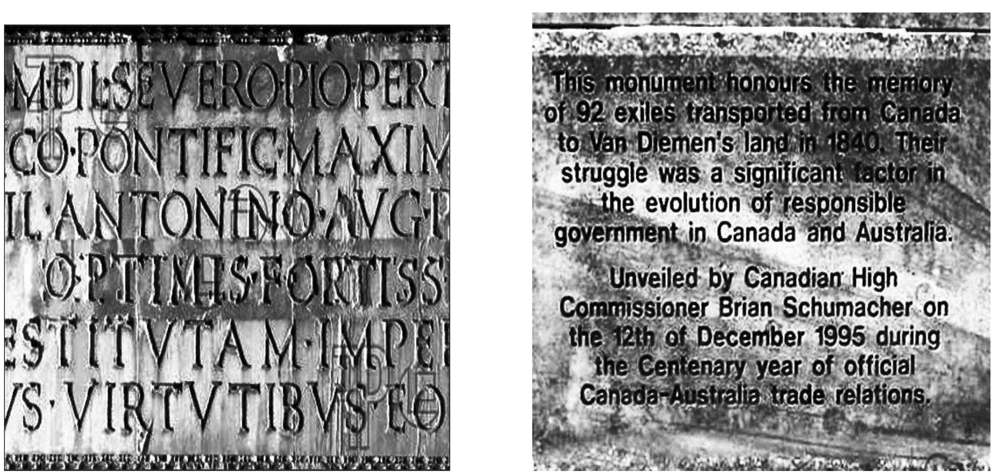
In this expression denominator represents the normalization factor and it is used to netralise the variation in the background.Its values vary according to the situations,if the region or area is dark it will give a small value of denominator and if the region is quit bright then it will give a value which is large for large contrast.So this normalization factor adapt according to the situation of the background region.

(a)

(b)

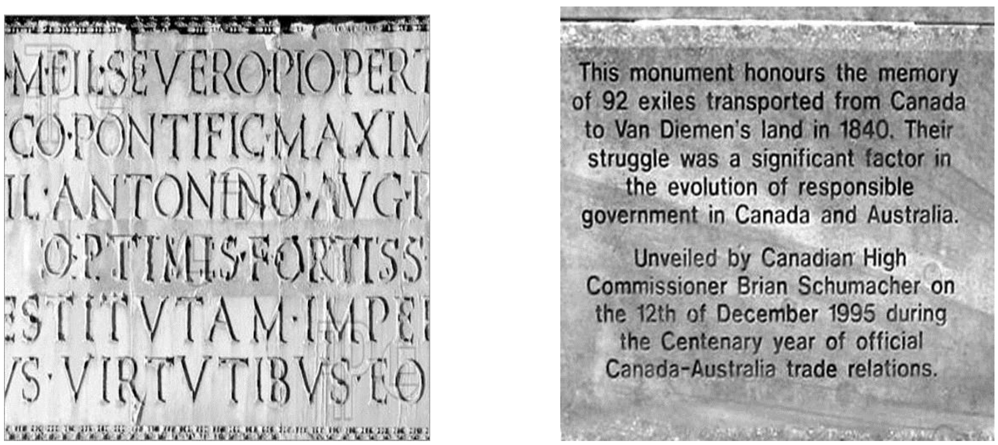
(c)

Fig.4.2.1. Formation of contrast image (a)local image gradient.b)local image contrast.c)our method contrast image for fig. 1(b) and (d)



(a)

(b)



(c)

Fig.4.2.2contrast images(a)result of local image gradient(b)result of local image contrast(c)result of our method

Now we can see from the above images which are of degraded document are inscription images that here we have shown the result of the all the three methods.In figure fig4.2.1(a),this represents the output of the local image gradient method and this is called the contrast image,so we can say that from above images that local gradient method work good for the images where the background is complex.Now from the figure 4.2.1(b),this represents the result of the local image contrast so from this we can see that this technique give the better result when there is a large variation in the stroke of the text.The same thing can also be seen for the inscription images,here the first method is giving quite poor result and from figure 4.2.2(b),we can say that it give a little improved result while the images which are resulted from our proposed method is much better than the other two and it can be seen from the figure 4.2.2(c)Our proposed method basically uses the both method but in adaptive manner because previous methods are giving individually good result but in different conditions,so we use them in adaptive manner and it will adapt according to the condition of the image present.

The above explained method can not deal with the images which are having the text in bright color and this is its major problem.This problem is known as overnormalization and to outcome this problem we can use the combination of the local image contrast along with the image gradient and will get an contrast map which is adaptive in nature just shown follow,while in the case of inscription it does not work where there is a shadowy region in the image.

$$C_a(i, j) = \alpha C(i, j) + (1 - \alpha)(I_{\max}(i, j) - I_{\min}(i, j)) \quad (3)$$

Here $C_a(i,j)$ corresponds to the local contrast in eqn-2 and [Imax(i,j)-Imin(i,j)]denotes the gradient of local image which is normalized in between [0,1].Generally local window measure is taken as 3.Here α represent the weight which give situation according weightage to local image contrast and local image gradient and its control is based on the document image statistical information.High weight that means large α will be assigned to image contrast,when the image has significant intensity variation.In our previous method we were having the over-normalization problem,but in the method proposed here we use two technique in an adaptive manner.So if in the given input the variation in the intensity is large then it will adapt to the local image contrast or we can say the weight assign to the local image contrastis more and it will produce a better result on the other hand if the input image condition is quit different means if the variation is in the text strokes then the more weight is assigned to the local image gradient,and hence in this manner it will avoide the over-normalization limitation.

By using the power function expressed as below the mapping of the variation in intensity with the α is done in this:

$$\alpha = \left(\frac{Std}{128}\right)^{\gamma}.$$

Here the value of γ is already defined and std is known as the standard deviation of image intensity.We know that a power function which is as defined in the above expression has the property that its value increases very smoothly and in a monotonic manner and its total control depends on the value of γ,and this fact has been used here. Γ may have its value inbetween[0,∞],and the power function become the linear function when γ=1.Hence now we can conclude that the method proposed here will depend on the value of ,because if the value of γ is having the large magnitude then the contribution of the local image gradient is more and on the other hand if it is having small value then the local image contrast will contribute more.

Fig.1(b),shows the sample document with complex background,as explained earliar the normalization factor will help to remove noise when it is present in upper left side region as shown in fig 4.2.1(a).The weightage to the local image contrast will be given more for the image shown in fig4.2.1(b),because in this there is complex background.The document images where the variation of the intensity is very small in the background region and the variation is large in the strokes of the text,in such cases the use of local image contrast will suppress a lot of the lightly visible strokes of the text,just as shown in the fig4.2.1(b),but if we use local image gradient it shows the good results just shown in the fig4.2.1(a) because it is havig the capability to reserve the strokes of the text which is light.

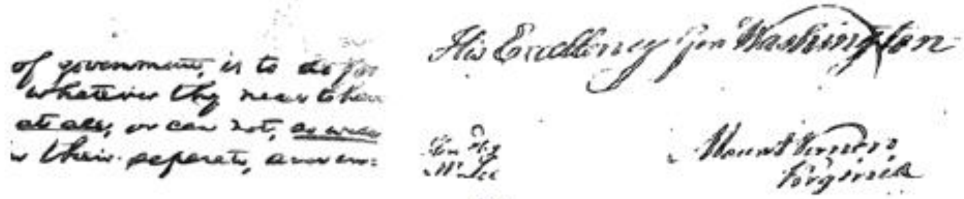
When we compare,contrast maps of the different methods as shown in fig4.2.1,we find that the combination of the previous method in adaptive manner obtained the best result,it can obtained the contrast map for input image which in degraded in several manner just asshown in figure fig4.2.1(c).Here it work on the fact that when there is large variation in intensity of the document image background it provide more weightage to the local image gradient while in other case if variation in stroke is large it provige more weightage to local image contrast.

While when we compare the combined image which is obtained adptively applying local image contrast and the other local image gradient it produces good contrast map for the image which is having good lightining condition and also it produce good result for the images which is having the same color of the background as well as for foreground.This can be seen from the fig.4

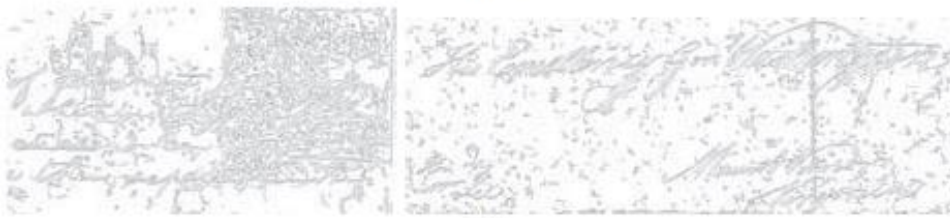
4.3Text edge pixel Estimation

Our main aim is to obtain the bi modal map of the image and this can be obtained from the contrast map.All this thing is done to obtain the edges pixels of the strokes.Here we can say that computed contrast image is having the large value at the stroke edges but its value is small when compared to value in background.The well known method which we generally use for detection of the edges are the otsu's method.It is basically called the otsu's global thresholding technique.So we can see from the figure fig2.4.1(c)that pixels of the edge strokes are obtained properly.

In the case of inscription images the contrast image which is shown in fig.4 (a)and(b) will have the binary image as shown in the fig 6(a) and (b) which also obtained by applying the otsu's algorithm.



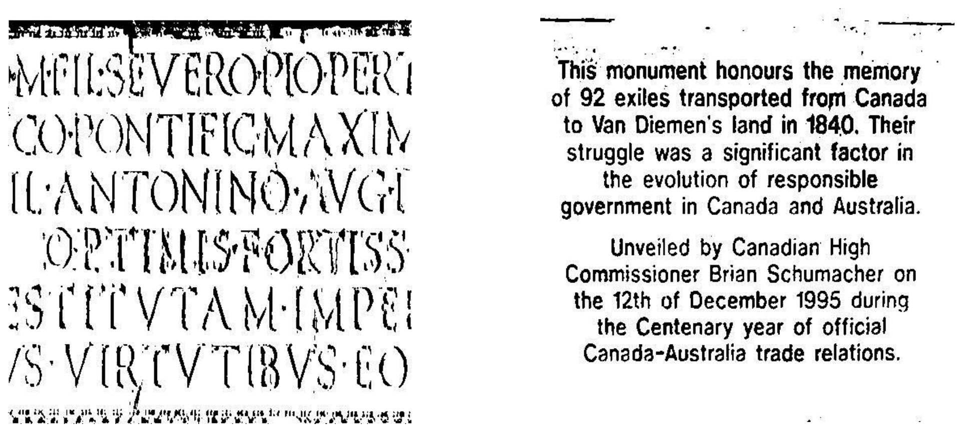
(a)



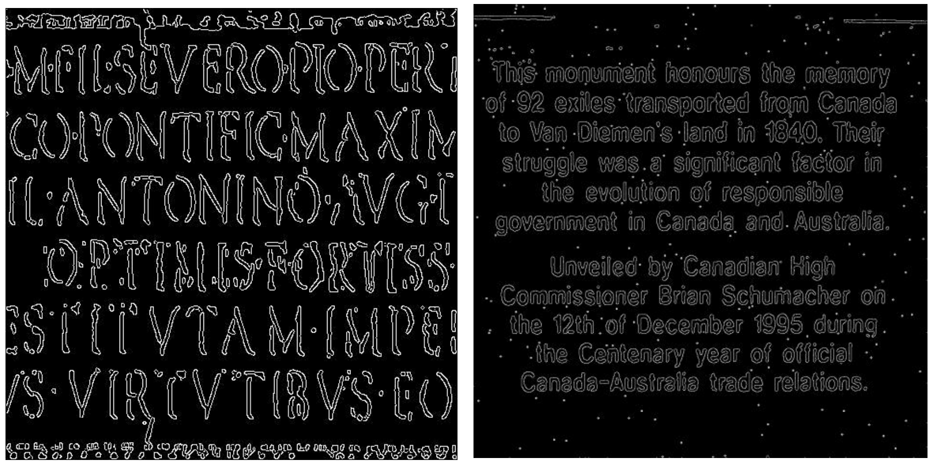
(b)

(c)

Fig.4.3.1(a)Binary contrast maps,(b)edge map from canny (c)combined resultant image of both the images of fig.1(b)and (d)



(a)



(b)

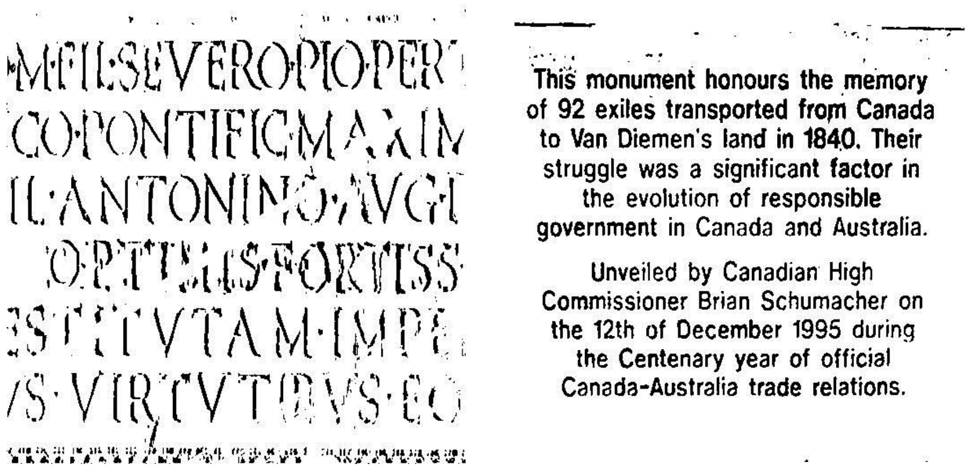


Fig4.3.2(a)Binary contrast maps ,(b) edge maps from canny (c)combined resultant image of both images fig.(d) and (f)

As we know in both cases local image contrast as well as in local image gradient their final value depends on the difference of the intensities which of basically the local maxima and local minima and these are particularly taken in the window and also the pixels on either side is taken as the pixels which are of high contrast.To further improve the performance or result we can introduce here the edge information and this can be done by combing it with the edge map.The best technique known for it is the caany edge detection, this technique is quite tolerant to different types of degradations present in in the images as it uses two adaptive thresholds.We can see from fig4.3.1(b) and fig4.3.2(b),it gives many unnecessary edges or nonedge pixels in the resultant image.In the next step of proposed technique we combine both maps one which is canny edge map and other is otsu's binary map,the portion which is common in both the images such as in pixel map of contrast image and edge map which is obtained from canny.The resultant image obtained from it will be helpful in determining the edge pixels of the text properly which is shown in figure fig4.4

4.4 Estimation of Local Thresholding

First we have to obtain pixels of the stroke edges accurately then the region which containing the text can be separated from the back-ground pixels of the document.From observing different kind of images following characteristics can be obtained:

1. The original text pixels are very much similar to the obtained pixels of the stroke edges.
2. In edge pixels of the stroke and the pixels of the surrounding background,there is a lot of difference within there intensities.

Now we can obtain the text region from the edge pixels of the strokes just as shown below:

$$R(x, y) = \begin{cases} 1 & I(x, y) \leq E_{\text{mean}} + \frac{E_{\text{std}}}{2} \\ 0 & \text{otherwise} \end{cases} \qquad (5)$$

Here E_{mean} represent the mean and E_{std} is the value of the standard deviation which is determined in particular window of particular size W where the text edge pixels of stroke are present.

To contain stroke edge pixels in the neighborhood window, its size should be at least larger than the stroke width.The stroke width which is represented as EW play a important role in determining the size of the mask W,and this stoke width can obtained from edges of the strokes just as shown in the fig4.3.1(b) and it has been explained in algorithm-1.

Now we have to calculate the width of the stroke and this is done by finding the distance which is having the highest frequency which denote the both side of the stroke edge .Initially the edges of the image is scanned horizontally row wise and we can see from the step-3 how the pixels are selected.Now how to determine if the edge map is correctly detected or not for this we verify that,the pixel of the edge which we leveled as 0 represents the background and aslo if the coming next pixel is having value 1 which is the edge of the pixel then we say correctly detected.One thing which they must follow is that the value of the pixels which belongs to the strokes pixels the intensity value must be high.Now in the next step which is step 4,the remaining pixels which are improper edge pixels are suppressed or we can say removed.So we now go on to check the entire row and if it is found that two consecutive edge pixels are very much similar then we mark them as the sides of the edge stroke,like this the remaining adjacent pixels are also matched to the pairs and the distance among these two are calculated which is explained in the next step.Now we have to plot the histogram and this is drawn between the frequency of the calculated distance which is inbetween two alike adjacent pixels.So now we can easily estimate the width of the stroke edge which is represented by EW,and this is equal to the value wich is having the highest frequency and it can be calculated from the following figure.

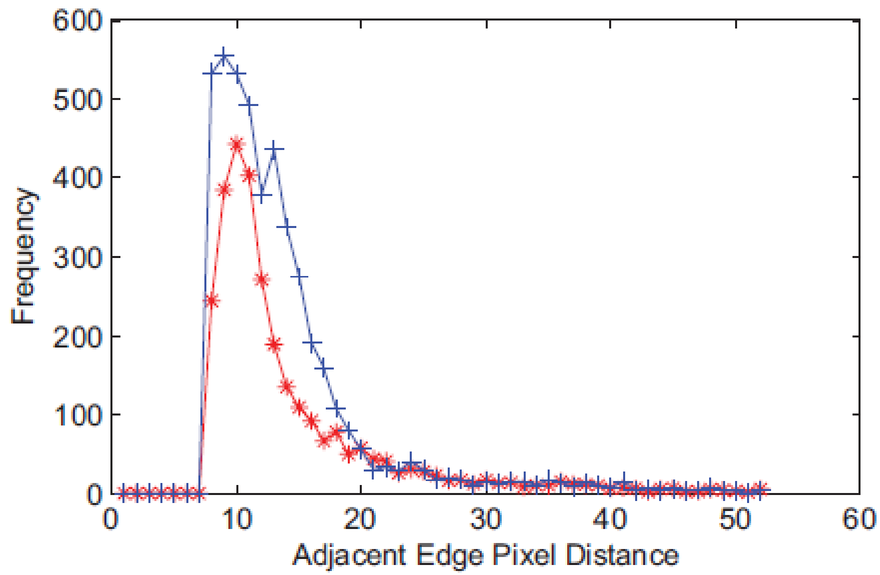


Fig4.histogram representation of the most frequent adjacent edge pixels where ++ gives the histogram for fig.1(d).

Algorithm 1 Edge Width Estimation

Require: The Input Document Image I and Corresponding Binary Text Stroke Edge Image Edg

Ensure: The Estimated Text Stroke Edge Width EW

1: Get the $width$ and $height$ of I

2: **for** Each Row $i = 1$ to $height$ in Edg **do**

3: Scan from left to right to find edge pixels that meet the following criteria:

a) its label is 0 (background);

b) the next pixel is labeled as 1(edge).

4: Examine the intensities in I of those pixels selected in Step 3, and remove those pixels that have a lower intensity than the following pixel next to it in the same row of I.

5: Match the remaining adjacent pixels in the same row into pairs, and calculate the distance between the two pixels in pair.

6: **end for**

7: Construct a histogram of those calculated distances.

8: Use the most frequently occurring distance as the estimated stroke edge width EW.

4.5 Post-processing

As mentioned in the previous section,if one time we find the initial binarization result from equation5,the binarization result may further be better by introducing some related knowledge just as decribed in the algorithm-2.

Algorithm 2 Post-Processing Procedure

Require: The Input Document Image I, Initial Binary Result B and Corresponding Binary Text Stroke Edge Image Edg

Ensure: The Final Binary Result B_f

1: Find out all the connect components of the stroke edge pixels in Edg.
2: Remove those pixels that do not connect with other pixels.
3: **for** Each remaining edge pixels (i, j): **do**
4:   Get its neighborhood pairs: $(i-1, j)$ and $(i+1, j)$; $(i, j-1)$ and $(i, j+1)$
5:   **if** The pixels in the same pairs belong to the same class (both text or background) **then**
6:     Assign the pixel with lower intensity to foreground class (text), and the other to background class.
7:   **end if**
8: **end for**
9: Remove single-pixel artifacts [4] along the text stroke boundaries after the document thresholding.
10: Store the new binary result to B_f.

1. Firstly,to make a precise set of the edge pixels what we do is the processes in which the pixels which are other then the foreground text pixels or we can say that these are non text pixels and these are removed or filtered.
2. There are basically two classes one is foreground and other is background,so the pixel pair which are on edge pixel of text stroke or lie on symmetry are classify in these two classes

   Here we know that pixel pairs are present if both pixels of the pair pixel represents the same pair then one pixel out of the pair pixel belong to the different class which are mentioned above,and the remaining isolated or pixels which are present only one are removed by using some filtering.

<u>CHPATER 5</u>
EXPERIMETS AND DISCUSSION

Chapter 5
Experiments and Discussion
5.1 General
The dataset for the inscription images for the proposed method was made by collecting the images of inscriptios which belongs to different historical monuments,some images was taken from the internet.These kind of inscription are very common at almost every monument and usually found engraved into/projected out from stone or other durable materials.These images have certain processing difficulties like uneven illumination wrapping,perspective distortion,multilingual text text with foreground and background images.While the dataset for degraded document images was public dataset that was used in the document image binarization contest(DIBCO)2009&2011 and handwritten-DIBCO 2010.The separation of the text from such images is a very challenging task and it is due to high inter/intra variation between the document background and the foreground.

5.2 Parameter Selection
First we will try to find the best suitable value for the γ,for the good result.Here the power function is an important parameter and is used as,by using different values of γ for this power function.After applying the different values to power function we then evaluate for its performance on dataset.the variation of the values of γ is as shown in the below figure.

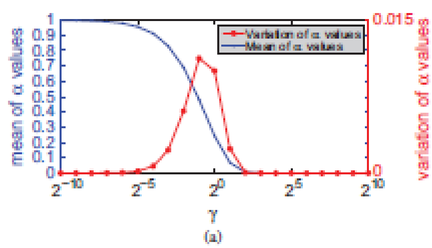


Fig5.2.1(a) values of the α of the 20 images on dataset for different values of γ and its mean

As we have mentioned earliar that here two methods are being utilized adaptively so for the same reason the value of α which is very near to 1 when there is value of γ is very small and in such cases the method which is having the more weight is local image contrast for the contrast of the image which is represented by C_a.the domain of the variation for the values of α become large when the value of γ_s is very close to 1.under these conditions for global intensity of document image become more sensitive for the power function and the document images which are having the different characteristics the value of the suitable weights is allotted.

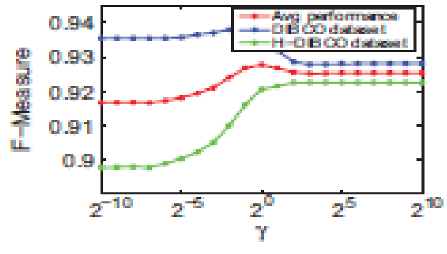


Fig5.2.2(b)F-measure performance on dataset images using different γ power functions.

Now we have to decide what should be the value of γ and for this purpose we use the above shown figure.From the above shown figure we can conclude the best suitable value for γ is very close to 1,for the case when its values is very near to 1 here we can say that local image contrast method is given more weightage.Its performance becomes more better in terms of f-measure aslo.So for the better results and more stability for the binarization the value of the γ is set very close to 1.

CHPTER 6

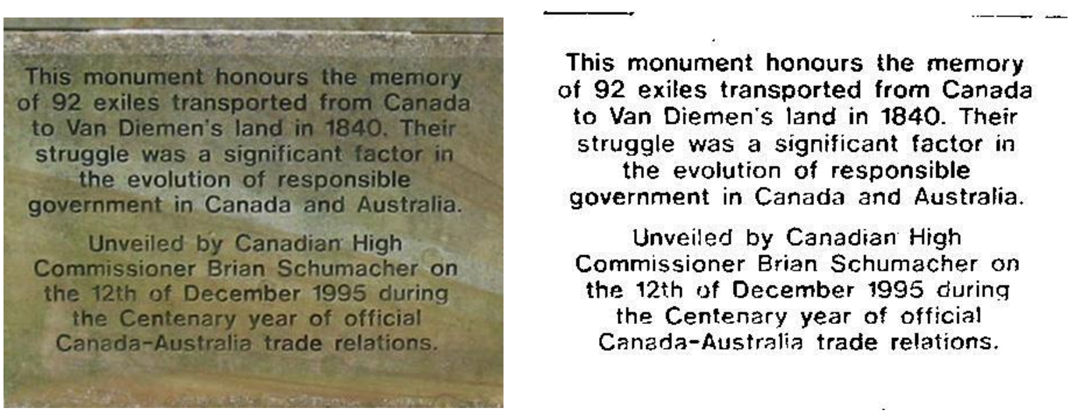
RESULTS

Chapter 6

Results

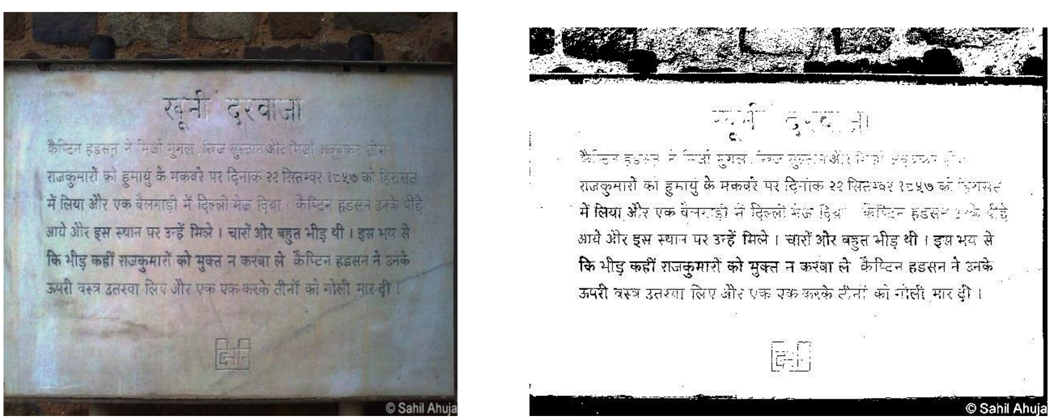
6.1 Results on inscription images

In this thesis we have proposed the robust binarization technique for the degraded document and inscription images,it basically performs adaptively by using the local image gradient and local image contrast.The proposed method uses the improved form of the ealiar available method and due to this it do not have the over normalization problem.Below shown is the resultant images after applying the proposed method,we can see the results are quite good and having very good visibility and it perform good irrespective of the language used in inscription and it does not depend if the foreground and background are of the same color or of different color.



(a) (b)

Fig6.1.1(a)Input inscription image,and(b)Output binarized image for given input



(a) (b)

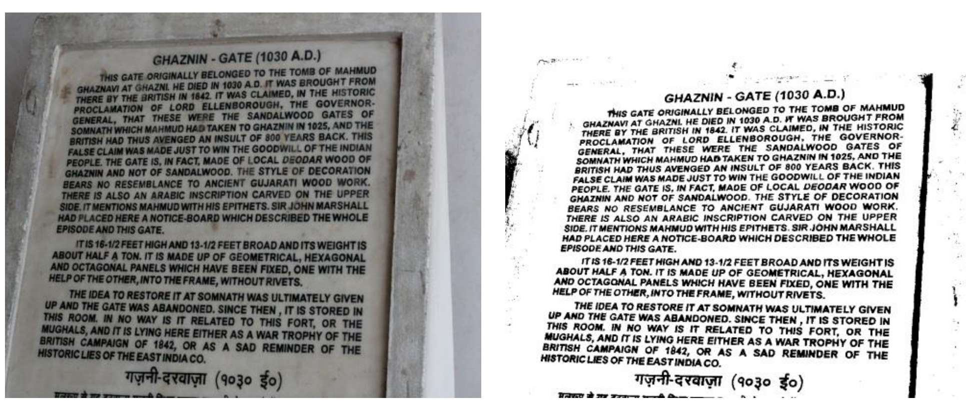
Fig6.1.2(a)Input inscription image,(b)Output of the given inscription image

(a) (b)

Fig6.1.3(a)Input image,(b)output binarized image

(a) (b)

Fig6.1.4(a)Input image,(b)output binarized image



(a) (b)

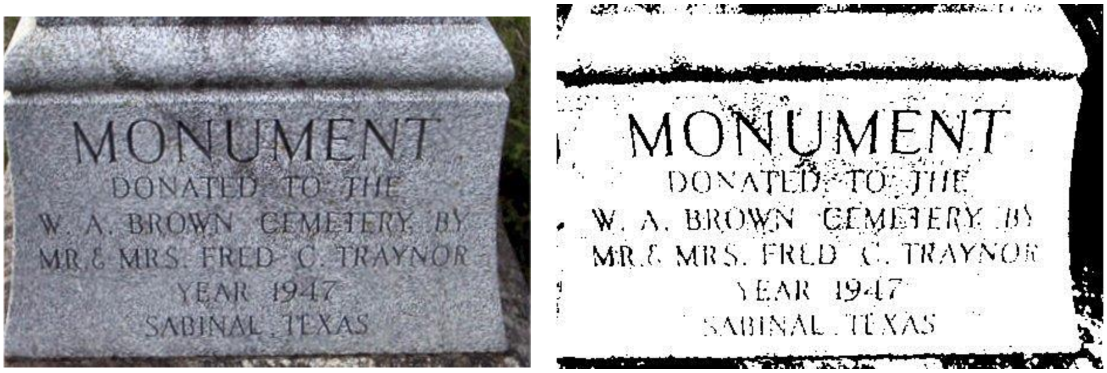
Fig6.1.5(a)input image,(b)output binarized image

(a) (b)

Fig6.1.6(a)Input image(b)Output binarized image

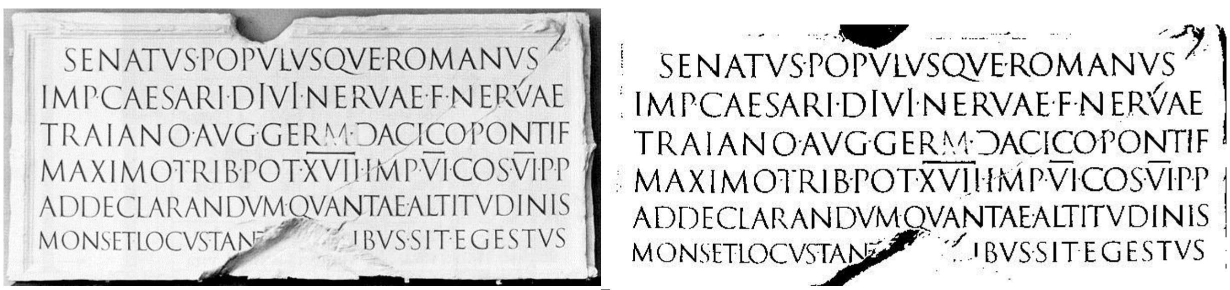
(a) (b)

Fig6.1.7(a)Input image,(b)Output binarised image



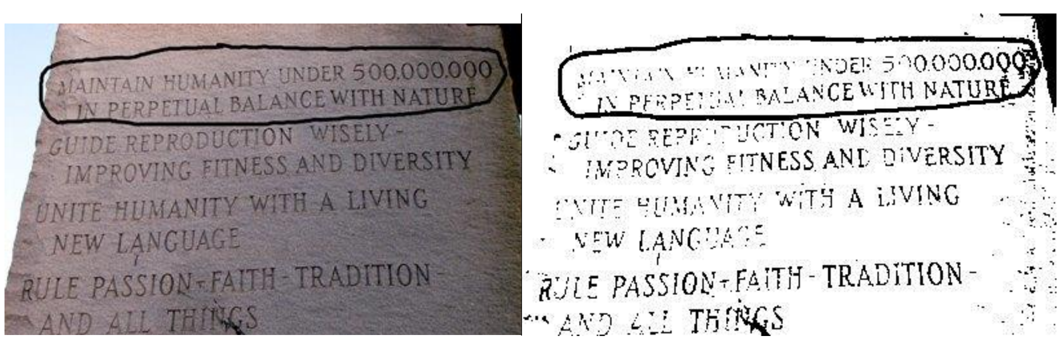
(a) (b)

Fig6.1.8(a)input image(b)output image



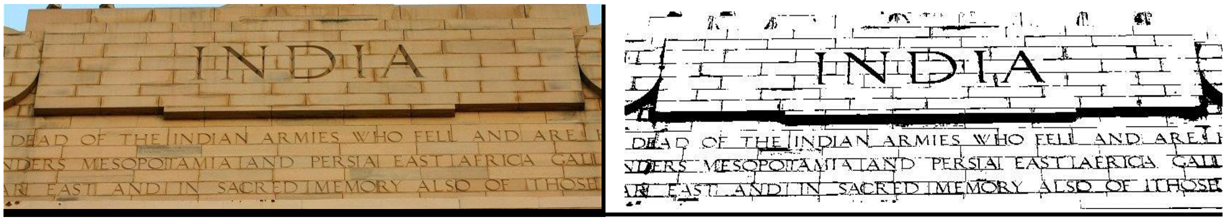
(a) (b)

Fig6.1.9(a)input image(b)output image



(a) (b)

Fig6.1.10(a)input image(b)output image



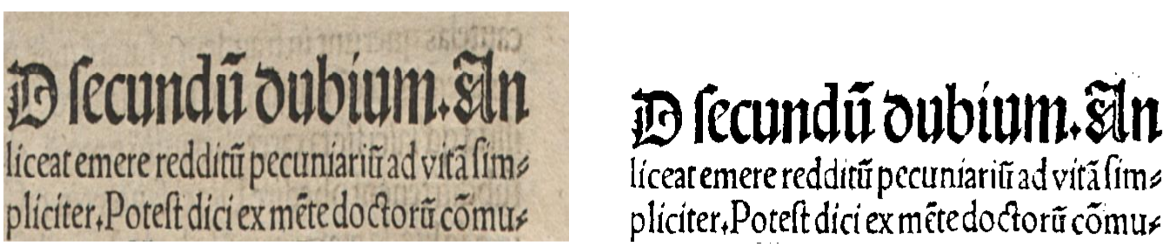
(a) (b)

Fig6.1.11(a)input image(b)output image

6.2Result on the degraded document images

(a) (b)

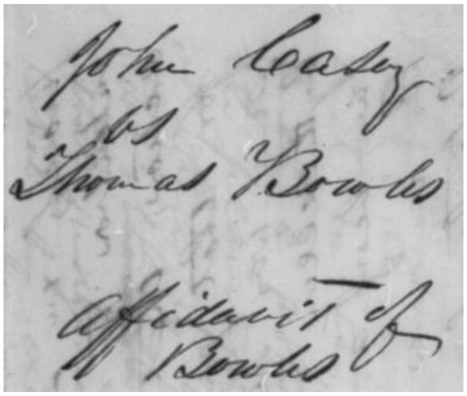
Fig6.2.1(a)Input image,(b)Output image



(a) (b)

Fig6.2.2(a)Input image,(b)Output image

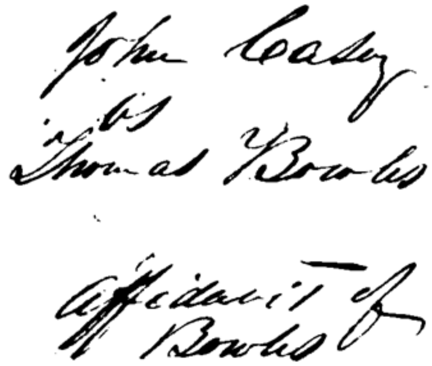
In the above image which is printed degraded document image,we can see the bleedthrough effect of the back side ink.After applying the proposed method we can see the resultant image is quite good i.e no noise is present in the output image,hence it shows the effectiveness of the proposed method.



John Casey
vs
Thomas Bowles

Affidavit of
Bowles

(a)



John Casey
vs
Thomas Bowles

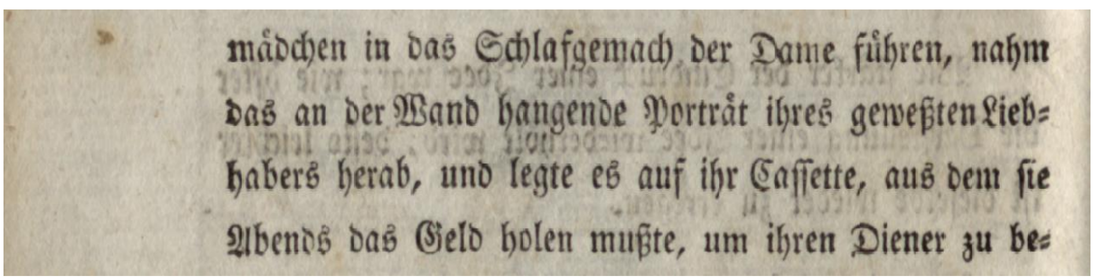
Affidavit of
Bowles

(b)

Fig6.2.3(a)input image(b)output image

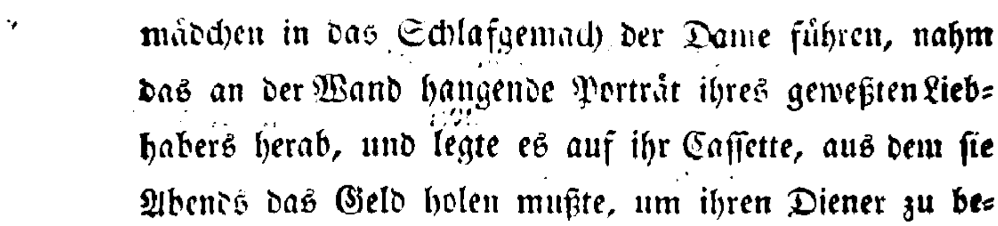
The above image is the handwritten degraded document image,we can see the noise present in the input image,other is the output image and here we can see that

The resultant image is free from the all noise present in the input image.



mädchen in das Schlafgemach der Dame führen, nahm
das an der Wand hangende Porträt ihres geweßten Lieb=
habers herab, und legte es auf ihr Cassette, aus dem sie
Abends das Geld holen mußte, um ihren Diener zu be=

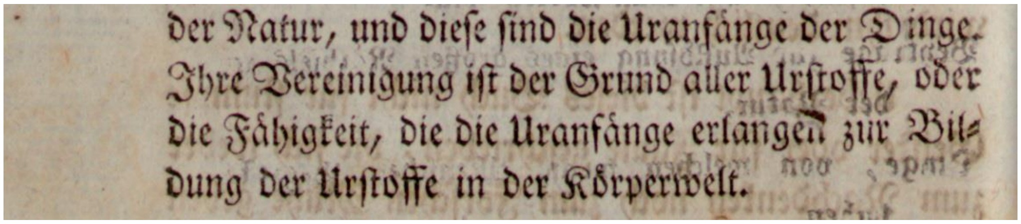
(a)



mädchen in das Schlafgemach der Dame führen, nahm
das an der Wand hangende Porträt ihres geweßten Lieb=
habers herab, und legte es auf ihr Cassette, aus dem sie
Abends das Geld holen mußte, um ihren Diener zu be=

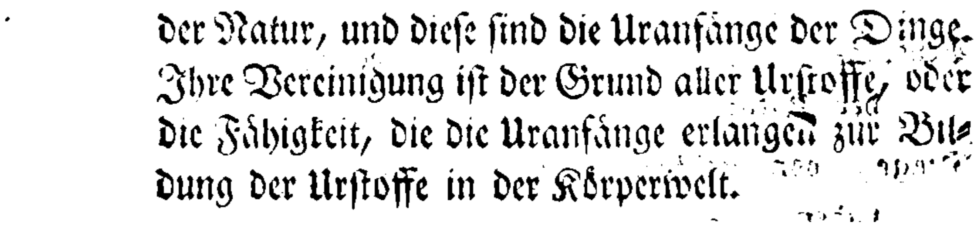
(b)

Fig6.2.4Printed document(a)Input image,(b)output image



der Natur, und diese sind die Uranfänge der Dinge. Ihre Vereinigung ist der Grund aller Urstoffe, oder die Fähigkeit, die die Uranfänge erlangen zur Bildung der Urstoffe in der Körperwelt.

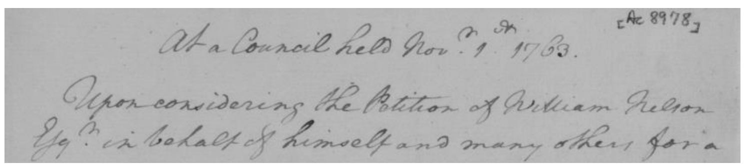
(a)



der Natur, und diese sind die Uranfänge der Dinge. Ihre Vereinigung ist der Grund aller Urstoffe, oder die Fähigkeit, die die Uranfänge erlangen zur Bildung der Urstoffe in der Körperwelt.

(b)

Fig6.2.5Printed document(a)input image,(b)output image

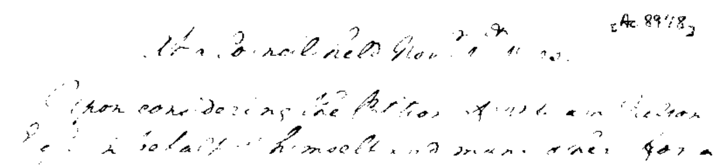


[Ac 8978]

At a Council held Nov.r 9th 1763.

Upon considering the Petition of William Nelson Esq.r in behalf of himself and many others for a

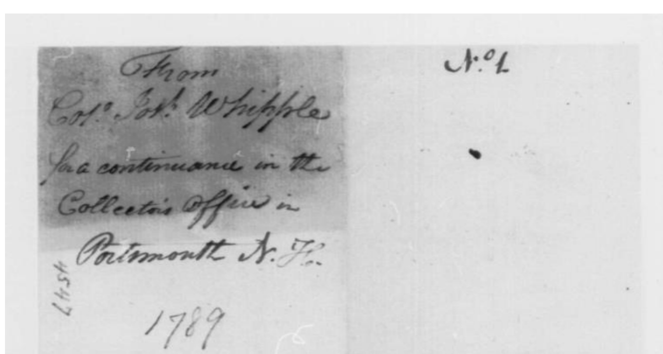
(a)



[Ac 8978]

(b)

Fig6.2.6Handwritten document(a)Input image,(b)output image

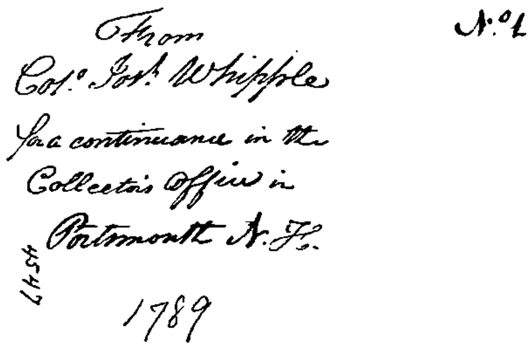


From Col.o Jos.h Whipple for a continuance in the Collector's Office in Portsmouth N. H.

1789

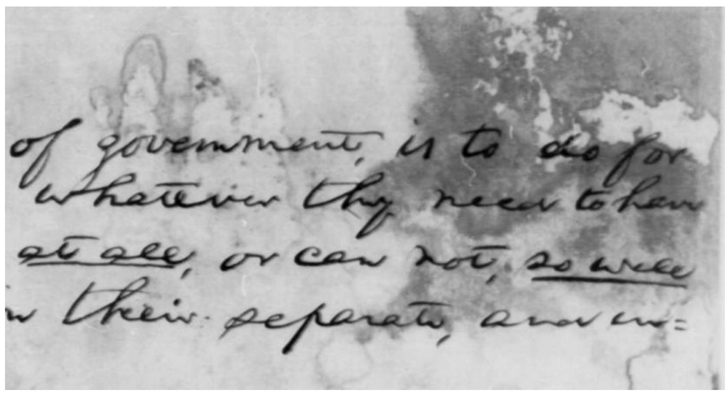
N.o 1

(a)

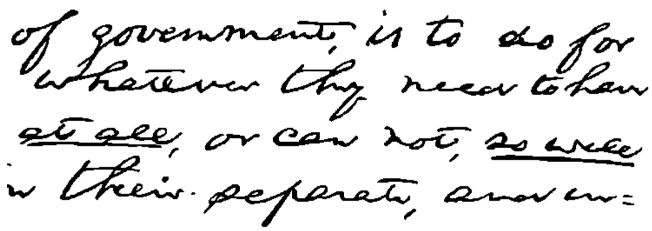

From
Col.o Jos.h Whipple
for a continuance in the
Collector's Office in
Portsmouth N.H.
1789
N.o 1
4547

(b)

Fig6.2.7Handwritten document(a)input image(b)output image

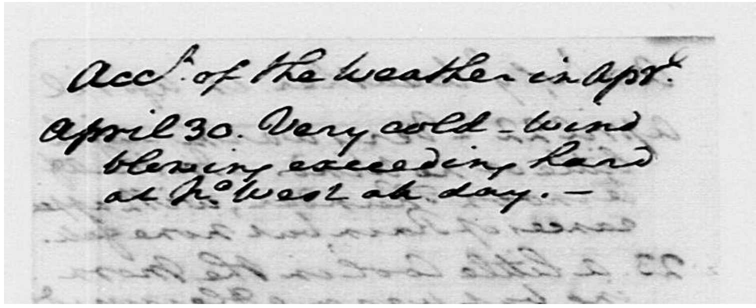

of government, is to do for
whatever they need to have
at all, or can not, so well
in their separate, and in=

(a)


of government, is to do for
whatever they need to have
at all, or can not, so well
in their separate, and in=

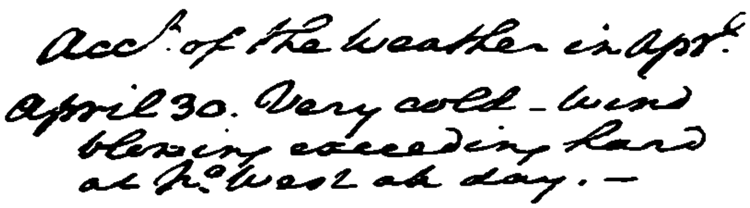
(b)

Fig6.2.8Handwritten document(a)input image(b)output image



Acct. of the weather in Apr.
April 30. Very cold – wind
blowing exceeding hard
at No West all day.—

(a)



Acct. of the weather in Apr.
April 30. Very cold – wind
blowing exceeding hard
at No West all day.—

(b)

Fig6.2.9Handwritten document(a)input image(b)output image

Hence from above images we can see that the proposed method work well for the degraded document images wether it is printed or handwritten and for the inscription as well,it give very very good results for both the cases.

6.3 Discussion

As we have mentioned earlier that the method proposed involve less parameters,some of them can be automatically determined which is based on the statistics of the images which are taken as input.Due to this the method proposed is more stable and easy to use for images having different variaty of degradation.This method is having the superior performance and it can be said because of several factors.

1. The method proposed here have many advantage over the previously present method,as we know the local image contrast and gradient both give good result but in different conditions and our method uses these two in adaptive manner avoiding the over normalization problem.
2. It is having the advantage that the result obtained is having the good egde information as it utilizes the edge map for obtaining the result.

3. In this method we have used the edge information which makes this method quite robust and the it accurately separates the foreground text from its background.

It gives good result for the inscription images as well as for the degraded document images,also it gives good result in inscription images irrespective of its language.But the performance on bickey diary and on the inscription images which are having the shadow regions still needs improvement.

CHAPTER 7

CONCLUSION AND FUTURE SCOPE

Chapter 7

Conclusion and future scope

7.1General

Historical and ancient document image collection requires specialized processing, because such kind of images have quite complicated kind of degradations which makes bibarization difficult task.We know that for libraries it is very common task to provide access for such images.Special treatment of images involves removal of background noise and to make it more clear.

7.2 Main Conclusion

This thesis present a method which is basically the binarization technique and this can be used for the binarization of the degraded document images and for the binariztion of monument inscriptions images. This method uses two previously available method in an adaptive manner.These previously available methods are

1) local image contrast
2) local image gradient.

In case of document images it is tolerant to degradations such as uneven illumination and document smear while in the case of inscription images it gives good results regardless the foreground and background are of the same the same color or not.In the proposed method it deals with the various issue by

- a)Developing an adptive image contrast.This adptive image contrast is basically the combination of the local image contrast and the local image gradient.
- b)Now this contrast map is converted to the binary map and this binary map is combined with the edge map that is obtained after the application of the canny edge detection technique.
- c)Finally document text is futher segmented by local threshold,this threshold is estimated from the ontensities of the detected text stroke edge pixels,which is within local window.

The proposed method is simple and robust and involve less parameters.The proposed method has been tested on the various degraded document dataset and inscription images.For the document images it outperforms most binarization methods while in the ase of the inscription images it gives quite good result in most conditions.

7.3 Future Work

The proposed method work good for the degraded document images and the inscription images,but this method does not give that much good result for the Bickley diary dataset and the inscription images where there is large variation in the contrast due to light conditions or we can say it does not give good result if the image is shadowy in nature.So in future a lot work can be done in this field to make result good for these conditions.Further the result obtained may also be improved by applying several pre or post processing technique.

REFERENCES

References

[1] B. Gatos, K. Ntirogiannis, and I. Pratikakis, "ICDAR 2009 document image binarization contest (DIBCO 2009)," in Proc. Int. Conf. Document Anal. Recognit., Jul. 2009, pp. 1375–1382.

[2] I. Pratikakis, B. Gatos, and K. Ntirogiannis, "ICDAR 2011 document image binarization contest (DIBCO 2011)," in Proc. Int. Conf. Document Anal. Recognit., Sep. 2011, pp. 1506–1510.

[3] I. Pratikakis, B. Gatos, and K. Ntirogiannis, "H-DIBCO 2010 handwritten document image binarization competition," in Proc. Int. Conf.Frontiers Handwrit. Recognit., Nov. 2010, pp. 727–732.

[4] S. Lu, B. Su, and C. L. Tan, "Document image binarization using background estimation and stroke edges," Int. J. Document Anal. Recognit., vol. 13, no. 4, pp. 303–314, Dec. 2010.

[5] B. Su, S. Lu, and C. L. Tan, "Binarization of historical handwritten document images using local maximum and minimum filter," in Proc.Int. Workshop Document Anal. Syst., Jun. 2010, pp. 159–166.

[6] G. Leedham, C. Yan, K. Takru, J. Hadi, N. Tan, and L. Mian, "Comparison of some thresholding algorithms for text/background segmentation in difficult document images," in Proc. Int. Conf. Document Anal.Recognit., vol. 13. 2003, pp. 859–864.

[7] M. Sezgin and B. Sankur, "Survey over image thresholding techniques and quantitative performance evaluation," J. Electron. Imag., vol. 13,no. 1, pp. 146–165, Jan. 2004.

[8] O. D. Trier and A. K. Jain, "Goal-directed evaluation of binarization methods," IEEE Trans. Pattern Anal. Mach. Intell., vol. 17, no. 12,pp. 1191–1201, Dec. 1995.

[9] O. D. Trier and T. Taxt, "Evaluation of binarization methods for document images," IEEE Trans. Pattern Anal. Mach. Intell., vol. 17, no. 3, pp. 312–315, Mar. 1995.

[10] A. Brink, "Thresholding of digital images using two-dimensional entropies," Pattern Recognit., vol. 25, no. 8, pp. 803–808, 1992.

[11] J. Kittler and J. Illingworth, "On threshold selection using clustering criteria," IEEE Trans. Syst., Man, Cybern., vol. 15, no. 5,pp. 652–655, Sep.–Oct. 1985.

[12] N. Otsu, "A threshold selection method from gray level histogram," IEEE Trans. Syst., Man, Cybern., vol. 19, no. 1, pp. 62–66, Jan. 1979.

[13] N. Papamarkos and B. Gatos, “A new approach for multithreshold selection,” Comput. Vis. Graph. Image Process., vol. 56, no. 5,pp. 357–370, 1994.

[14] J. Bernsen, “Dynamic thresholding of gray-level images,” in Proc. Int.Conf. Pattern Recognit., Oct. 1986, pp. 1251–1255.

[15] L. Eikvil, T. Taxt, and K. Moen, “A fast adaptive method for binarization of document images,” in Proc. Int. Conf. Document Anal. Recognit.,Sep. 1991, pp. 435–443.

[16] I.-K. Kim, D.-W. Jung, and R.-H. Park, “Document image binarization based on topographic analysis using a water flow model,” Pattern Recognit., vol. 35, no. 1, pp. 265–277, 2002.

[17] J. Parker, C. Jennings, and A. Salkauskas, “Thresholding using an illumination model,” in Proc. Int. Conf. Doc. Anal. Recognit., Oct. 1993,pp. 270–273.

[18] J. Sauvola and M. Pietikainen, “Adaptive document image binarization,”Pattern Recognit., vol. 33, no. 2, pp. 225–236, 2000.

[19] W. Niblack, An Introduction to Digital Image Processing. Englewood Cliffs, NJ: Prentice-Hall, 1986.

[20] J.-D. Yang, Y.-S. Chen, and W.-H. Hsu, “Adaptive thresholding algorithm and its hardware implementation,” Pattern Recognit. Lett., vol. 15,no. 2, pp. 141–150, 1994.

[21] B. Gatos, I. Pratikakis, and S. Perantonis, “Adaptive degraded document image binarization,” Pattern Recognit., vol. 39, no. 3, pp. 317–327, 2006.

[22] Y. Liu and S. Srihari, “Document image binarization based on texture features,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 19, no. 5,pp. 540–544, May 1997.

[23] M. Cheriet, J. N. Said, and C. Y. Suen, “A recursive thresholding technique for image segmentation,” in Proc. IEEE Trans. Image Process.,Jun. 1998, pp. 918–921.

[24] S. Kumar, R. Gupta, N. Khanna, S. Chaudhury, and S. D. Joshi,“Iterative multimodel subimage binarization for handwritten character segmentation,” IEEE Trans. Image Process., vol. 13, no. 9,pp. 1223–1230, Sep. 2004.

[25] Y. Chen and G. Leedham, “Decompose algorithm for thresholding degraded historical document images,” IEE Proc. Vis., Image Signal Process., vol. 152, no. 6, pp. 702–714, Dec. 2005.

[26] Q. Chen, Q. Sun, H. Pheng Ann, and D. Xia, “A double-threshold image binarization method based on edge detector,” Pattern Recognit., vol. 41,no. 4, pp. 1254–1267, 2008.

[27] R. Cao, C. L. Tan, Q. Wang, and P. Shen, “Double-sided handwritten archival documents,” in Proc. Int. Workshop Doc. Anal. Syst., 2000,pp. 147–158.

[28] I. Blayvas, A. Bruckstein, and R. Kimmel, “Efficient computation of adaptive threshold surface for image binarization,” Pattern Recognit.,vol. 39, no. 1, pp. 89–101, 2006.

[29] S. Nicolas, J. Dardenne, T. Paquet, and L. Heutte, “Document image segmentation using a 2D conditional random field model,” in Proc. Int.Conf. Doc. Anal. Recognit., Sep. 2007, pp. 407–411.

[30] C. Wolf and D. Doermann, “Binarization of low quality text using a markov random field model,” in Proc. Int. Conf. Pattern Recognit., 2002,pp. 160–163.

[31] T. Lelore and F. Bouchara, “Document image binarisation using Markov field model,” in Proc. Int. Conf. Doc. Anal. Recognit., Jul. 2009,pp. 551–555.

[32] J. G. Kuk, N. I. Cho, and K. M. Lee, “Map-MRF approach for binarization of degraded document image,” in Proc. Int. Conf. Image Process., 2008, pp. 2612–2615.

[33] S. Kumar, R. Gupta, N. Khanna, S. Chaudhury, and S. D. Joshi,“Text extraction and document image segmentation using matched wavelets and MRF model,” IEEE Trans. Image Process., vol. 16, no. 8,pp. 2117–2128, Aug. 2007.

[34] A. Dawoud, “Iterative cross section sequence graph for handwritten character segmentation,” IEEE Trans. Image Process., vol. 16, no. 8,pp. 2150–2154, Aug. 2007.

[35] B. Su, S. Lu, and C. L. Tan, “A self-training learning document binarization framework,” in Proc. Int. Conf. Pattern Recognit., Aug. 2010,pp. 3187–3190.

[36] N. Howe, “A Laplacian energy for document binarization,” in Proc. Int.Conf. Doc. Anal. Recognit., Sep. 2011, pp. 6–10.

[37] F. Deng, Z. Wu, Z. Lu, and M. S. Brown, “Binarizatioinshop: A userassisted software suite for converting old documents to black-and-white,”in Proc. Annu. Joint Conf. Digit. Libraries, 2010, pp. 255–258.

[38] H. Yi, M. S. Brown, and X. Dong, “User-assisted ink-bleed reduction,”IEEE Trans. Image Process., vol. 19, no. 10, pp. 2646–2658, Oct. 2010.

[39] E. Badekas and N. Papamarkos, “Optimal combination of document binarization techniques using a selforganizing map neural network.”Eng. Appl. Artif. Intell., vol. 20, no. 1, pp. 11–24, Feb. 2007.

[40] B. Gatos, I. Pratikakis, and S. Perantonis, “Improved document image binarization by using a combination of multiple binarization techniques and adapted edge information,” in Proc. Int. Conf. Pattern Recognit.,Dec. 2008, pp. 1–4.

[41] M. van Herk, “A fast algorithm for local minimum and maximum filters on rectangular and octagonal kernels,” Pattern Recognit. Lett., vol. 13,no. 7, pp. 517–521, Jul. 1992

[42] D. Ziou and S. Tabbone, “Edge detection techniques—An overview,”Int. J. Pattern Recognit. Image Anal., vol. 8, no. 4, pp. 537–559, 1998.

[43] J. Canny, “A computational approach to edge detection,” IEEE Trans.Pattern Anal. Mach. Intell., vol. 8, no. 6, pp. 679–698, Jan. 1986.

[44] T. Lindeberg, “Edge detection and ridge detection with automatic scale selection,” Int. J. Comput. Vis., vol. 30, no. 2, pp. 117–156, 1998.

[45] T. Lelore and F. Bouchara, “Super-resolved binarization of text based on the fair algorithm,” in Proc. Int. Conf. Document Anal. Recognit.,Sep. 2011, pp. 839–843.

[46] E. Saund, J. Lin, and P. Sarkar, “Pixlabeler: User interface for pixel-level labeling of elements in document images,” in Proc. Int. Conf. Document Anal. Recognit., Jul. 2009, pp. 646–650

Appendix-1

File name:-

```
function [histo,pixval,IDX,sep] = minor_p_algo_1(I,n)

%function [IDX,sep] = (minor_p_algo_1 (1) (I,n))

% I should be an rgb image
%   I = ind2rgb (image_path)
% n should be between 4 to 255


if isRGB(I)
    sizI = size(I);
    I = reshape(I,[],3);
    [V,D] = eig(cov(double(I)));
    [~,c] = max(diag(D));
    I = reshape(double(I)*V(:,c),sizI(1:2));
end


I = I-min(I(:));
I = round(I/max(I(:))*255); %Normalize

unI = sort(unique(I));
nbins = min(length(unI),256);
disp('nbins ');disp(nbins);
disp('n ');disp(n);
if nbins==n
    IDX = ones(size(I));
    for i = 1:n, IDX(I==unI(i)) = i; end
    sep = 1;
    return
elseif nbins<n              % errror return NAN
    IDX = NaN(size(I));
    sep = 0;
    return
elseif nbins<256
    [histo,pixval] = hist(I(:),unI);
else
    [histo,pixval] = hist(I(:),256);
end

P = histo/sum(histo);
clear unI

k0 = linspace(0,1,n+1); k0 = k0(2:n);
[k,y] = fminsearch(@sig_func,k0,optimset('TolX',1));
k = round(k*(nbins-1)+1);
IDX = ones(size(I))*n;
IDX(I<=pixval(k(1))) = 1;
```

```
for i = 1:n-2
    IDX(I>pixval(k(i)) & I<=pixval(k(i+1))) = i+1;
end
sep = 1-y;

IDX(~isfinite(I)) = 0;

    function y = sig_func(k)

        muT = sum((1:nbins).*P);
        sigma2T = sum(((1:nbins)-muT).^2.*P);

        k = round(k*(nbins-1)+1);
        k = sort(k);
        if any(k<1 | k>nbins), y = 1; return, end

        k = [0 k nbins];
        sigma2B = 0;
        for j = 1:n
            wj = sum(P(k(j)+1:k(j+1)));
            if wj==0, y = 1; return, end
            muj = sum((k(j)+1:k(j+1)).*P(k(j)+1:k(j+1)))/wj;
            sigma2B = sigma2B + wj*(muj-muT)^2;
        end
        y = 1-sigma2B/sigma2T; % within the range [0 1]

    end

end




% TODO - create the histogram using the above values and find the most
frequently occuring distance as the stroke size
% TODO - for now use the output image to visually estimate the stroke size
and try different values
```

Appendix-2

File Name:-

```
% should create a variable with the image handle (eg H01)
%uiopen('file_location' , 1);
img=imread('./p02.bmp');
im = rgb2gray(img);
figure,Imshow(img);
im = im2double(im);

% should try different values for n (1.7 for H04.. etc)
n = 1.7;
f_makebw = @(I) im2bw(I.data, double(median(I.data(:)))/n);
bw = blockproc(im, [128 128], f_makebw);

figure, Imshow(bw);
% TODO - this is basic binarization, need to support the proposal from the
paper..

%[IDX,sep] = minor_p_algo_1(bw,150);
%disp(IDX);
```

Appendix-3

File name:-

```
clear all;
clc;
% disp(pwd);
images=imread('C:\Users\hemu dobhal\Documents\MATLAB\P02.bmp');
[histo,pixval,IDX,sep] = minor_p_algo_1(images,6);       % Function call
minor_p_algo_1

figure,Imshow(IDX,[0 3]);                                % Represent image
figure,plot(pixval,histo);
disp(' minor_p_algo_1 Completed ');
```

Appendix-4

File name:-

```
%-----------------------------------
%--Step 1: Read the image------
%-----------------------------------
clc;                   %------Clear the command window  data-------
clear all;               %------Clear the storage  data-------
close all;             %-----Close all the existing opened figures
%------Read the image-----
for iiii=1:12
    str=sprintf('C:\\hemu\\images\\%d.jpg',iiii);
images=imread(str);
%------Read Image---------
imagen=images;
imagen=imresize(imagen,[1024,1024]);
%----Show image---------
% figure(1);
% imshow(imagen);
% title('Input Image');
%-----Check if it is colored image and then conver it to gray scale-----
if size(imagen,3)==3           %--check the third element of the input image
and if is 3 then convert it into gray scale
    imagen=rgb2gray(imagen);   %--rgb2gray is to convert it into gray scale
end

%Contrast Image Construction%
imr=imagen;
imr1=imr./255;
J = adapthisteq(imr);

figure(1);
imshow(J,[]);
title('contrast Adjustment');
%-----------------------------------
%--Step 2: Add noise to image------
%-----------------------------------
noiselevel=0.02;   % Change to add noise level
imagen = imnoise(imagen,'salt & pepper', noiselevel);
%--------------------------------------------------------------------
%--Step 3: Convert the image from grayscale to binary image------
%--------------------------------------------------------------------
threshold = graythresh(imagen);  %---Measure the threshold----
imagen =im2bw(imagen,threshold); %---Apply the threshold and convert to
binary----
% figure(2);
% imshow(imagen);
% title('Binary Version of Input Image');
%--------------------------------------------------------------------
%--Step 4: Region Extraction from the binary image------
%--------------------------------------------------------------------
% Remove all object containing fewer than 30 pixels
imagen = bwareaopen(imagen,15);
%--Step 4: Apply morphological operation on the binary image------
imagen1 =bwmorph(imagen,'fill');
```

```
imagen1=~imagen1;
stat=regionprops(imagen1,'ALL');
st=size(stat,1);
edge1=edge(imagen1,'canny');

% checking no. of connecting objects
if(st<205)
imagen1=~imagen1;
stat=regionprops(imagen1,'ALL');

st=size(stat,1);
end
count=0;
% saving the sizeof the objects
for i=1:st
    object_size(i)=stat(i).Area;

end

if(st>30)
for i=1:st
     if(object_size(i)>7800)
        count=count+1;
        pixel_no{count}=stat(i).PixelIdxList;
    end
end
    if(count>0)
        for i=1:size(pixel_no,2)
        ind=pixel_no{i};
        imagen1(ind)=0;
        end
    end
end
figure(2);
imshow(imagen)
title('Binary Version of Input Image After Region Extraction');

% figure(4);
% imshow(imagen)
% title('Binary Version of Input Image After Region Extraction');
figure(3);
imshow(edge1);
title('canny edge operation ');
figure(4);
imshow(imagen1)
title('morphological operation ');
pause(0.01);
des=sprintf('C:\\hemu\\output\\%d.jpg',iiii);
des1=sprintf('C:\\hemu\\edge\\%d.jpg',iiii);
des2=sprintf('C:\\hemu\\contrast\\%d.jpg',iiii);
imwrite(imagen1,des);
imwrite(edge1,des1);
imwrite(J,des2);
end
```