International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 06 | June 2019 www.irjet.net p-ISSN: 2395-0072

Enhancement of Old Images and Documents by Hybrid Binarization Techniques

P Kalyanchakravarthi[1], P Krishna Rao[2]

[1]Assistant professor, Dept. of ECE, GMRIT, Rajam, Andhrapradesh, India
[2]Assistant Professor, Dept. of ECE, AITAM, Tekkali, Andhrapradesh, India

---

Abstract - *Documents can be valuable source of information but often they suffer degradation problems, especially in the case of historical documents, such as strains, background of big variations and un-even illumination, ink seepage, etc.... Binarization techniques should be applied to remove the noise and improve the quality of the documents.*

Binarization is a process of converting the document image into binary image containing text as foreground and plain white as background or vice versa. Characters from the document image should be extracted from the binarized image, in order to recognize them. So performance of the character recognition system is completely depends on the binarization quality. A hybrid binarization approach is proposed in this project for improving the quality for the old documents. Combination of global and local thresholding techniques are used for the same.

Initially, a technique named global thresholding is applied to the whole image. The image area that still has background noise are detected and the technique is again re-applied to each area separately. Therefore, a better adaptability is achieved for the algorithm where various kinds of noise re exist in different areas of same image.

The proposed technique is efficient to tolerate the high inter and intra intensity variation in the degraded document image. The proposed method is based on spatial domain techniques: Laplacian operator, Adaptive Bilateral filter and Gaussian filter and works well for degraded documents and palm leaf manuscript images.

Advantage of applying global thresholding, is that it avoids the computational and time cost of applying a local thresholding in the entire image. Hence it is indicated that this technique is pretty effective in removing background noise and improving the quality of degraded images.

Key Words: component; formatting; style; styling; insert

1. INTRODUCTION

The existence of ancient palm leaf manuscripts in Southeast Asia is very important both in term of quantity and variety of historical content. In the last five years, ancient palm leaf manuscripts have received great attention from researchers in the field of document image analysis.

The ancient palm leaf manuscript is written on a dried palm leaf by using some sort of sharp pen or small knife, which is then scrubbed with natural dyes. Due to these specific characteristics, ancient palm leaf manuscripts are providing new challenges in document analysis.

Recently, large scale projects from various organizations and universities are mainly focusing on making the written material available in a digital form on-line. Many reputed universities around the globe, libraries and research organizations are working towards this goal and a lot of challenges are still to be addressed in the processing of these documents. The main purpose of these projects is to digitize all the word written documents that are spread around the globe and facilitate access to their digital contents.

The Optical Character Recognition (OCR) technology has helped in changing document images into machine editable format. Even though the OCR system adequately recognizes the documents, the recognition of handwritten documents is not completely reliable and is still an open challenge to researchers.

Inaccurate recognition is due to many factors like scanning errors, lighting conditions, quality of the documents etc. Further inaccuracies stem from the age of these documents and the condition of the materials these documents are inscribed upon. In addition to this, historical documents are also preserved in the digital form as these documents play a vital role in connecting present to past. Hence upholding of such resources is crucial. Preserving and maintaining historical document in its original form is highly impossible as these documents are prone to get affected by environmental conditions. Due to degraded quality of the document, the captured images also carries all the degraded contents in digital form and making it highly undesirable for further stages of character recognition system such as binarization, segmentation, character extraction, feature extraction and recognition. Hence efficient binarization is required to binarize degraded document image properly.

Binarization process as one of the early and important stage document analysis pipeline, is also a real challenge for ancient palm leaf manuscripts. To evaluate and to select an optimal binarization method, the ground truth binarized image is necessary.

Image binarization is usually performed in the pre-processing stage of image processing; because binarization is one of the main phases of the Optical Character Recognition

(OCR) and consists of transforming a grayscale image (i.e. Image of up to 256 grey levels) into a binary image (i.e., black and white image), our proposed approach focuses on the foreground extraction (i.e., useful information part of an image) in order to be used by an OCR application.

2. PROPOSED METHODS

There are many methods proposed in THRESHOLDING. Otsu's method calculates best possible threshold by distinguishing the two classes so that their collective intra-class variance is minimal. Local Binarization Algorithms such as Niblack, Sauvola's, Bemsen, Wolf, Fegn's, use local window to estimate the threshold value of each window and binarize local window individually. Leedham compares some of the traditional methods. While a method using a combination of existing techniques is proposed by Gatos, Shi and Yan.

In computer vision and image processing, Otsu's method, named after Nobuyuki Otsu is used to automatically perform clustering-based image thresholding or, the reduction of a gray level image to a binary image. The algorithm assumes that the image contains two classes of pixels following bi-modal histogram (foreground pixels and background pixels), it then calculates the optimum threshold separating the two classes so that their combined spread (intra-class variance) is minimal, or equivalently (because the sum of pairwise squared distances is constant), so that their inter-class variance is maximal. Consequently, Otsu's method is roughly a one-dimensional, discrete analog of Fisher's Discriminant Analysis.

Otsu's method exhibits the relatively good performance if the histogram can be assumed to have bimodal distribution and assumed to possess a deep and sharp valley between two peaks. But if the object area is small compared with the background area, the histogram no longer exhibits bimodality.

There are many improvements focusing on different limitations for Otsu's method. One famous and effective way is known as two-dimensional Otsu's method. In this approach, the gray-level value of each pixel as well as the average value of its immediate neighborhood is studied so that the binarization results are greatly improved, especially for those image corrupted by noise.

Fig -1: Otsu's Binarization

Fig -2: Original Image

Another similar method called Bradley Local Thresholding. It also examines the neighborhood of each pixel, setting the brightness to black if the pixels brightness is t percent lower than the average brightness of surrounding pixels. The corresponding paper can be found here. This stack overflow mentions a local (adaptive) thresholding method called Niblack.

The output of Niblack's method is significant and has most acceptable result out of all thresholding techniques in segmenting text documents. In this work the same method is applied on images keeping one of the variables i.e. weight k of Niblack's method constant while changing the other (window size) from images to images. The output image is better segmented but the background is noisy. Improvements in the resultant images are demonstrated by applying the morphological operations of opening and closing. Opening and closing are combination of two fundamental morphological operations dilation and erosion. Dilation thickens objects in a binary image by adding pixels to the boundaries of the objects, while erosion shrinks objects in a binary image.

The method can be described like:

```
if pixel > (mean + k * standard deviation)
pixel = object;
else
pixel = background;
end
```

It performs relatively well on classical documents. However, three main defects remain: the window parameter of Sauvola's formula do not fit automatically to the content; it is not robust to low contrasts; it is non-invariant with respect to contrast inversion. Thus, on documents such as magazines, the content may not be retrieved correctly which is crucial for indexing purpose.

Pixel-based accuracy and OCR evaluations are performed on more than 120 documents. This implementation remains notably fast compared to the original algorithm. For fixed parameters, text recognition rates and binarization quality are equal or better than other methods on small and medium text and is significantly improved on large text. Thanks to the way it is implemented, it is also more robust on textured text and on image binarization. This implementation extends the robustness by making the results almost insensible to the window size whatever the object sizes. Its properties make it usable in full document analysis toolchains.

Sauvola's method takes a grayscale image as input.

Since most of document images are color images, converting color to grayscale images is required. For this purpose, we choose to use the classical luminance formula, based on the eye perception:

Luma = 0.299 × R + 0.587 × G + 0.114 × B.

3. PROPOSED FILTERS

There are various filters like Mean filter, Median filter, Gaussian filter, Bilateral filter, Wiener filter, Shrink and Swell filter. These filters will eliminate the noise and smooth the image giving blurring effect. In degraded document, the text information is very important for further stages of character recognition, losing text information while smoothing, is unacceptable.

Mean Filter :

The idea of mean filtering is simply to replace each pixel value in an image with the mean (`average') value of its neighbors, including itself. This has the effect of eliminating pixel values which are unrepresentative of their surroundings. Mean filtering is usually thought of as a convolution filter. Like other convolutions it is based around a kernel, which represents the shape and size of the neighborhood to be sampled when calculating the mean. Often a 3×3 square kernel is used, as shown in Figure 4.1, although larger kernels (e.g. 5×5 squares) can be used for more severe smoothing. (Note that a small kernel can be applied more than once in order to produce a similar but not identical effect as a single pass with a large kernel.)

Table -1: 3×3 averaging kernel often used in mean filtering

1/9	1/9	1/9
1/9	1/9	1/9
1/9	1/9	1/9

There are two main problems with mean filtering, which are:

A single pixel with a very unrepresentative value can significantly affect the mean value of all the pixels in its neighborhood.

When the filter neighborhood straddles an edge, the filter will interpolate new values for pixels on the edge and so will blur that edge. This may be a problem if sharp edges are required in the output.

Both of these problems are tackled by the median filter, which is often a better filter for reducing noise than the mean filter, but it takes longer to compute..

Median Filter

The median filter is normally used to reduce noise in an image, somewhat like the mean filter. However, it often does a better job than the mean filter of preserving useful detail in the image. Like the mean filter, the median filter considers each pixel in the image in turn and looks at its nearby neighbors to decide whether or not it is representative of its surroundings. Instead of simply replacing the pixel value with the mean of neighboring pixel values, it replaces it with the median of those values. The median is calculated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value. (If the neighborhood under consideration contains an even number of pixels, the average of the two middle pixel values is used.)

Gaussian Filter

The Gaussian smoothing operator is a 2-D convolution operator that is used to `blur' images and remove detail and noise. In this sense it is similar to the mean filter, but it uses a different kernel that represents the shape of a Gaussian (`bell-shaped') hump. This kernel has some special properties which are detailed below.

The Gaussian distribution in 1-D has the form:

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}}$$

where σ is the standard deviation of the distribution. We have also assumed that the distribution has a mean of zero

(i.e. it is centered on the line x=0). The distribution is illustrated in Figure

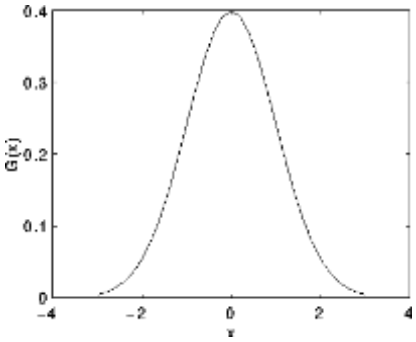

Fig- 1: 1-D Gaussian distribution with mean 0 and σ=1

4. Methodology & Results

The input image which is filled with noise and is to be binarized is taken. A series of steps starting from reading the input image to applying morphological techniques are applied over the image. Every step has its own advantages. Resizing is particularly done based on the original size of the image. This step plays a very important role because all the operations cannot be performed either on very small images or on very large images. Desired operations cannot be performed if the image is small because the pixel size may not support particular operation. The filters like Gaussian and bilateral help remove the noise from the images.

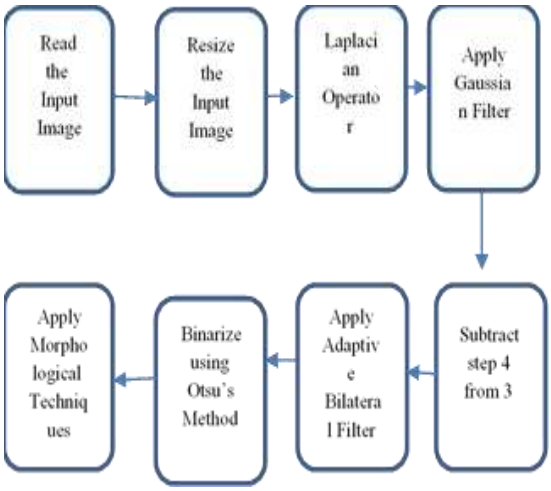

Fig- 4: Block diagram

Laplacian operator can be used to enhance the image. The already discussed Otsu's method is further used along the erosion and dilation techniques which make the last step but the main step in giving the characters their correct shape and format

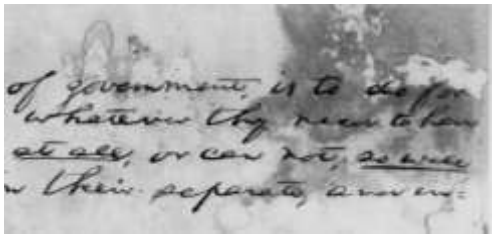

Fig- 5: input Image

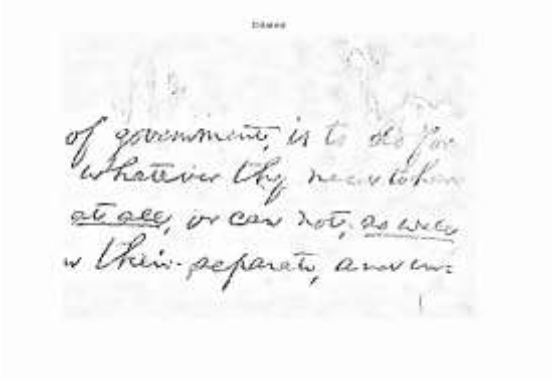

Fig- 6: output Image

REFERENCES

[1] E.Kavallieratou. (2015). A Binarization Algorithm Specialized on Document Images and Photos.Eighth International Conference on Document Analysisand Recognition, 463-467.

[2] Otsu, N. (2015). A threshold selection method from gray-level histograms. IEEE Trans. Systems Man Cybernet. , 62-66.

[3] Ganga Holi, Ranganath D. (2005). Hybrid Binarization Technique for Degraded Document Images IEEE International Advance Computing Conference (IACC).

[4] Mrs.Preeti.Kale,Prof.G.M.Phade,Prof.Pravin.A.Dhulekar. (2005). Enhancement of old images and documents by Digital Image Processing Techniques. International Conference on Communication,Information &Computing Technology (ICCICT).

[5] G C Tomasi and R Manduchi, "Bilateral Filtering for Gray and Color Images", 6th IEEE International Conference on Computer Vision, 1998,p ages 839-846.

[6] Y Zhang, Xuemin Tiann, and Peng Ren, "An Adaptive Bilateral Filter Based Framework for Image Denoising", Elsevier Neurocomputing, 2014, pages 299-316.

[7] Buyue Zhang, and Jan P Allebach, "Adaptive Bilateral Filter for Sharpness Enhancement and Noise Removal", IEEE 2007 pages 417 -420.

[8] B Gangamma, Srikanta Murthy K, Arun Vikas Singh, "Restoration of Degraded Historical Document Image", Journal of Emerging Trends in Computing and Information Sciences, Vol. 3, Issue. 5, May 2012, ISSN 2079-8407, pages 792-798.

[9] T. Q. Pham and L. J. V. Vliet, "Separable bilateral filtering for fast video preprocessing," in Proc. IEEE Int. Conf. Multimedia Expo, 2005, pp. 454–457.

[10] M. Elad, "On the origin of the bilateral filter and the ways to improve it," IEEE Trans. Image Process., vol. 11, no. 10, pp. 1141–1151, Oct. 2002..