ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

Improved Degraded Images using Hybrid Inpainting Binarization Techniques

Rajwinder Kaur[1], Upasna Rani[2]

M. Tech Scholar, Dept. of Computer Science and Engineering, Guru Kashi University, Talwandi Sabo, Punjab, India[1]

Assistant Professor, Dept. of Computer Science and Engineering, Guru Kashi University, Talwandi Sabo, Punjab, India[2]

ABSTRACT: Wide collections of ancient historical documents printed or handwritten in native languages are kept in libraries and museums. In recent years, libraries have begun to digitize historical document that are of interest to a wide range of people, with the goal of preserving the content and making the documents available via electronic media. But for historical documents suffering from degradation due to damaged background, stained paper, holes and other factors, the recognition results drop appreciably. These recognition results can be improved using binarization techniques. Binarization techniques can distinguish text from background. The simplest way to get an image binarized is to choose a threshold value, and classify all pixels with values above this threshold as white, and all other pixels as black. Our main objective is to effectively binarize the document images suffering from strain & smear, uneven backround, holes & spot and various illumination effect by applying Inpainting Binarization Techniques. The ink spots, mud spots and other spots hide various image details and make it almost unreadable, which is considered as a major loss of the information in the historical documents. Various types of impulsive noises also exist in the historical documents. Now days, the historical documents are stored in the computer memories in the form of digital historical documents. The digitized historical documents are easy to restore than the manual restoration process on the historical documents. The historical documents take years in the manual restoration, which causes higher latency in the useful information retrieval. During the same period, the historical documents also receive more degradation elements. The algorithms are capable of restoring most of the digitized historical documents accurately and removing the most of the noises and spots from the digitized historical documents. This algorithm is the hybridization of morphological operations, image de-noising filters, image dilation, OTSU thresholding, and Sauvola thresholding, Bernsen's all Methods, Proposed method (OSTU and BOLAN SU) techniques. The Mean Square Error (MSE) and peak signal to noise ratio (PSNR), Coefficient of Correlation (Coc), Elapsed Time, Accuracy, Sensitivity, Precision are taken as the performance parameters. The performance of the OSTU & Bolan Su means proposed algorithm is based on these parameters is better and the results are higher than the Sauvola algorithms, when compared.

KEYWORDS: Historical Documents, OSTU, Sauvola, PSNR, MSE, Coc.

I. INTRODUCTION

Binarization is the starting step of most document image analysis systems and refers to the conversion of the gray-scale image to a binary image. Binarization is a key step in document image processing modules since a good binarization sets the base for further document image analysis. Binarization usually distinguishes text areas from background areas, so it is used as a text locating technique. Binarization is usually reported to be performed either using Global, Local techniques.

Inpainting and Degraded Documents

Inpainting is the process of reconstructing lost or deteriorated parts of images and videos. For instance, in the museum world, in the case of a valuable painting, this task would be carried out by a skilled art conservator or art restorer. In the digital world, inpainting (also known as image interpolation or video interpolation) refers to the application of sophisticated algorithms to replace lost or corrupted parts of the image data (mainly small regions or to remove small defects).

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

Degradations appear frequently and may occur due to several reasons. Examples of degradation influence may include the appearance of variable background intensity caused by non-uniform intensity, and low contrast. Adaptive degraded document image binarization gives a reliable result. It is mainly performed by clearing up any unneeded objects appeared in the document, hiding background, remove noise and fill possible breaks, gaps or holes in the foreground and finally improve the quality of the character strokes before start converting it to editable text.

II. LITERATURE SURVEY

P. S. Jonesherine (2015), here in this a Mean shift algorithm is proposed for ancient document images, as well as a post processing method that can improve any Binarization method. it introduce an approach, a local-global Mean Shift based colour image segmentation. In this steps works such as, the first step consists in shifting each pixel in the image according to its R-Nearest Neighbor Colors (R-CC) in the spatial domain. The second step process shifts only the previously extracted local modes according to the entire pixels of the image. Binarized model is made efficient by including mean shifting technique in the image. While in the post processing step, specialized adaptive Gaussian and median filters are considered. The result of proposed method shows the output binarized image which is more efficient and provide better computed PSNR values comparing to the prior art [4]. Sushilkumar N. Holambe (2015), Image binarization is the separation of each pixel values into two collections, black as a foreground and white as a background. Thresholding technique is used for document image binarization. Here proposed a technique to address the issues of degraded images using adaptive image contrast with combination of the local image contrast and the local image gradient as to tolerant the variation of text and background. The proposed technique, constructs adaptive contrast map for degraded image that is combined with Canny's edge map, for the identification of text stroke edge pixels with global and local thresholding technique on recent document image binarization contest (DIBCO) 2009 & 2011 and handwritten-DIBCO 2010 [6]. Jagroop Kaur (2014), author describes that Image binarization is the method of separation of pixel values into dual collections, black as foreground and white as background. Thresholding technique used for binarization of document images is has found to be best technique. In degraded documents, where extensive background noise or difference in contrast and brightness exists i.e. there exists many pixels that cannot be effortlessly categorized as foreground or background. They said in future new algorithm which will use as nonlinear enhancement pre-processing technique to improve the results further [3]. Prof. S. P. Godse (2014), author said here that Recovering of text from badly degraded document images is a very difficult task due to the very high inter/intra variation between the document background and the foreground text of different document images. In this, a robust document image binarization technique that addresses the issues by using inversion gray scale image contrast by first converting the input image to invert image and then finding the contrast of the inverted image to differentiate text and background variation caused by different types of document degradations. This paper presents an adaptive image contrast based document image binarization technique that is tolerant to different types of document degradation such as uneven illumination and document smear with few parameters by use of the local image contrast as maximum and minimum. The proposed method has been tested on the various datasets [5]. VIKUL J. PAWAR (2014), here authors presents that an improved binarization technique is used for improving the quality of text from poorly degraded document images with good efficiency and accuracy. The degraded document images are consisting of various noises, illumination because of old age and many more reason. The proposed method here addresses issues by using accommodative image contrast which is a combination of the local image contrast and the local image gradient for constructing binarized contrast map then pooled with Canny's edge map detection to identify the text stroke edge pixels by using Segmentation to perform with in minimum parameters and efficiency [7]. BOLAN SU (2013), degraded document of Segmentation text images is a very challenging difference between the document background and the foreground text of different document images. In this paper, a novel document image binarization technique is addresses to adaptive image contrast i.e is a combination of the local image contrast and the local image gradient that is tolerant to text and background variation caused by different types of document degradations. In this, document text is estimated based on the intensities of detected text stroke edge pixels within a local window. The proposed method is simple, robust, and involves minimum parameter tuning. Here they, achieves accuracies of 93.5%, 87.8%, and 92.03%, for Dibco-2009, 2010, 2011. This paper presents an adaptive image contrast based document image binarization technique that is tolerant to different types of document degradation such as uneven illumination and document smear [1]. Md. Iqbal Quraishi et. al (2013) said that the old degraded historical documents carry various important information regarding our culture, economics etc. Proper restoration of these documents is very necessary. They proposed a novel approach to

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

enhance ancient historical documents. To enhance these digital format documents a two-way approach is considered. At first Particle Swarm Optimization (PSO) and bilateral filter is applied. At second level, Non-Linear Enhancement with bilateral filter is applied. Both the approaches are then tested visually and quantitively to show the effectiveness of the approach [8]. Weisheng Dong et al. (2013) said that sparse representation models code an image patch as a linear combination of a few atoms chosen out from an over-complete dictionary, and they have shown promising results in various image restoration applications. However, due to the degradation of the observed image (e.g., noisy, blurred, and/or down-sampled), the sparse representations by conventional models may not be accurate enough for a faithful reconstruction of the original image. They introduced concept of sparse coding noise to improve the performance of sparse representation-based image restoration, and the goal of image restoration turns to how to suppress the sparse coding noise. They exploited the image nonlocal self-similarity to obtain good estimates of the sparse coding coefficients of the original image, and then centralize the sparse coding coefficients of the observed image to those estimates. The so-called non-locally centralized sparse representation (NCSR) model is as simple as the standard sparse representation model, while they extensive experiments on various types of image restoration problems, including denoising, deblurring and super-resolution, validate the generality and state-of-the-art performance of the proposed NCSR algorithm [9]. Akihito Kitadai et al. (2012) said that shape features of character patterns on the documents are unstable or missing because most of the documents have been stained and degraded deeply. Digital archives of the documents with accurate character pattern retrieval methods are helpful for archaeologists and historians. They proposed a similarity evaluation method for character patterns with missing shape parts. It collaboratively works with non-linear normalization for such patterns, and modifies the templates for each trial of the retrieval efficiently. In the experiences using 4,911 Kanji (Chinese origin) character patterns from the Japanese historical documents called mokkans, the method shows improvements of the retrieval accuracy. They also presented a simple implementation of gradient feature extraction to compare the chain code feature with the gradient feature in the retrieval [10]. N. Sandhya et al. (2012) analyzed the different types of noise that occur in printed and handwritten historical documents mainly based on Kannada (Kannada is a language used in Karnataka, a southern state in India) documents and created a taxonomy for the same. They characterized each noise type based on factors such as their source, their effect on characters and the associated challenges in character recognition. They also catalogued the different noise detection, removal and restoration techniques that are reported in the literature for each of the prominent noise types, and identified areas relating to noise detection, removal for further research focus [11]. B Gangamma et al. (2011) proposed a method that mainly deals with enhancing the historical document image of palm scripts using gray scale morphological operations with the combination of spatial filters. Morphological opening operation is applied to compensate for non-uniform background intensity and suppress bright details smaller than the structural element, while closing suppresses the dark details. They proposed a method that works well for the images with dark background and low contrast and enhances the image with clear white background [12]. Chuche Fung (2010), Survey is performed on different binarization techniques such as Otsu, Sauvola, Ni-Black algorithm and different evaluation measurements have been taken. These different measures have been under taken by different researchers such as: (i) Zhang studied three different method of measurement: analytical, empirical goodness, empirical discrepancy methods. He found that discrepancy method are more effective than the other methods as these methods have to compare with ground truth image so these are more complex. (ii) Sezgin and Sankur describe different performance criteria for binarization algorithms .They used five criteria's: Misclassification Error (ME), Edge Mismatch (EMM), Region Non-Uniformity (NU), Relative foreground area error (RAE), Shape distortion penalty via Hausdorff distance. This technique was implemented to measure the quality of 40 thresholding algorithms over two different context of method and result that Kittler and Illingworth is the best quality of thresholding technique in both type of image. (iii) Pavlos and et al. surveyed the evaluation of binarization on historical documents proposed by statistical measures of image quality. They found that the global method applied with histogram give good results as local methods. All the above research concluded that there is no single binarization technique that is suitable for all images. They proposed a framework that has to be applied for the processing of ancient manuscript written on media such as palm leaves. There algorithm has following 4 step: (i) select appropriate binarization algorithm (ii) cluster training and testing image set for edge detection evaluation for estimating best binary technique of each image (iii) extract the features from the noise reduction image by using value from binarization method (iv) classify the optimal algorithm for image by machine learning technique. This technique result in noise reduction and in selection of automatic optimal binarization algorithm [16]. Rachel Mabanag Chong et al. (2010) considered that degradation in images is usually undesirable but unavoidable. The most common cause of this is motion. Given only the image degraded by motion blur, blind

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

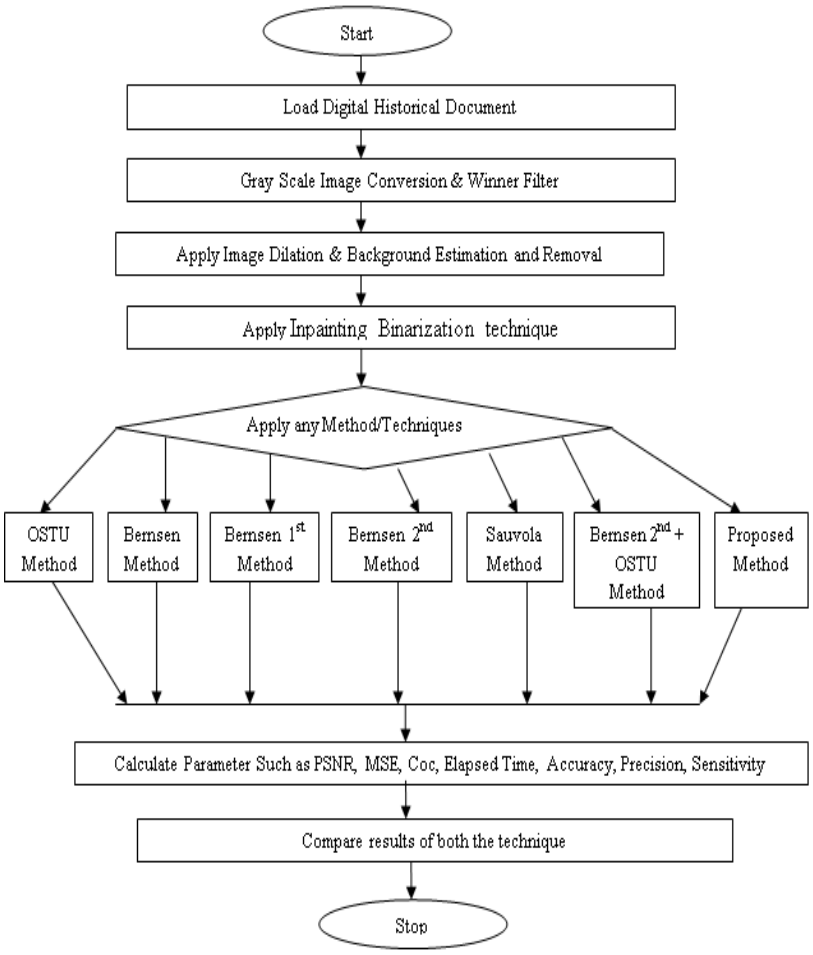
International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

restoration aims to determine the original un-blurred image and the blurring function. They proposed an approach to simultaneously determine these two quantities by utilizing a reference point spread function (RPSF) as a learning basis in the alternating minimization cycle. Based on the behavior of maxima in the presence of blurs and assuming a blurring function support size, RPSF can be extracted from the given degraded image. This implies that learning is not dependent on estimated values and as a result, it will not vary as restoration iteration progresses. The cost function is formulated such that spurious colour artifacts can be avoided, and the contour of the piecewise blurring function is also accounted for. They proposed a method that is capable of determining an RPSF that closely resembles the actual blur despite the presence of noise and a small image size. Reconstruction with the proposed cost function resulted in better image quality in contrast to other methods [13].

III. METHDOLOGY



IV. RESULTS

A procedure is followed and experimental results are designed to display the effectiveness and robustness of our proposed method. We first examine the performance of the proposed technique on public datasets for parameter selection. The proposed technique is then experienced and compared with state-of-the-art methods over on standard datasets of Dibco 2009, 2010, 2011, 2012 and so on. The binarization performance is evaluated by using Parameters such as PSNR, MSE, Coc, Time Elapsed, Accuracy, Precision, Sensitivity by using OSTU, Sauvola, Bernsen, Bernsen 1st, Bernsen 2nd, Combined Bernsen 2nd and Ostu, and our proposed(Bolan Su and Ostu) Method.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

Figure 1.1: Degraded Original Documents or Images

The model has been designed to restore the degraded historical document of importance. The historical documents are considered very important because they contain a lot of general and secret information about our predecessors, their learning's, discoveries, literature, sociology and other important information. This information is helpful in the case of understanding and making our projections about the time they were living into. The historical documents are stored on the online/offline repositories in the form of digital documents, which makes it adaptable to use the computer algorithms to replace the human work of document restoration. The proposed algorithm has been developed using MATLAB environment. The MATLAB environment is considered a flexible and powerful research and development tool to write computer algorithms.

GUI Design

Main window as GUI part showing below Fig 1.2, which describes about the techniques on which we are working to recover degraded images using Inpainting binarization. Select Push button of OSTU Method the working of this algorithm starts to degrade the images nad recover them. Same way on selecting Push button of Sauvola Method then its algorithm is worked. Third Push button is used for Existing from GUI window.

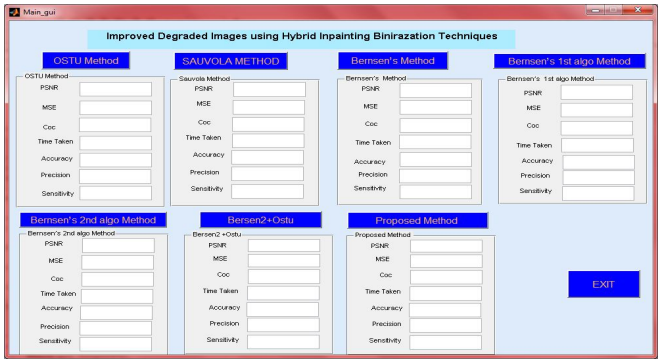


Figure 5.2 GUI for selecting Method

In above figure 1.2 Click on either Method/ Techniques Push button for selecting common part first as Browsing window for selection of any types of historical degraded images such as shown in figure 1.3 below:

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

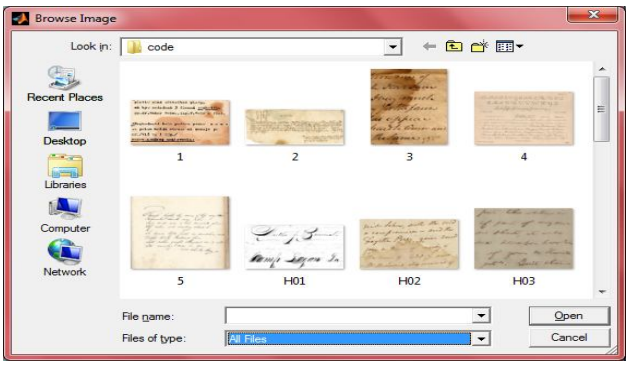


Figure 1.3 Browsing Window for selecting Historical degraded Image

Below show at first, the Otsu method reads an image of degraded historical documents from database. Then converts the original image into gray scale image. In third step, Wiener filtered image , fourth as image dilation is performed that subtracts the image background from foreground and then converted into Inpainting binarized form as fifth image. At last, restored image by Otsu method is obtained as sixth image shown in figure 1.4 (a) & (b).

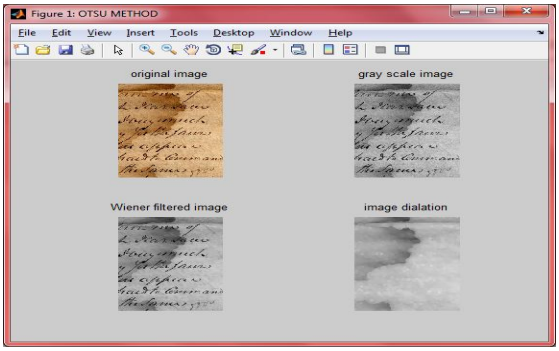


Figure 1.4 (a) Results of OSTU method

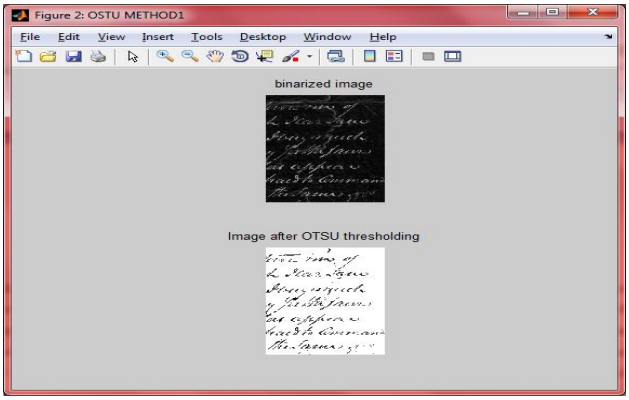


Figure 1.4: (b) Results of OSTU method

Below show at second, the Sauvola method which reads an image of degraded historical documents from database. Then converts the original image into gray scale image. In third step, Wiener filtered image, fourth as image dilation is performed that subtracts the image background from foreground and then converted into Inpainting binarized form as fifth image. At last, restored image by Sauvola method is obtained as sixth image shown in figure 1.5 (a) & (b).

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

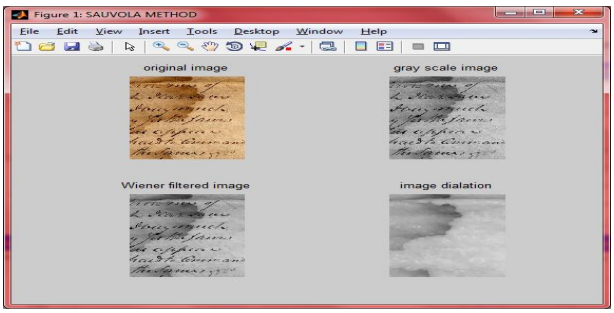


Figure 1.5: (a) Results of Sauvola method

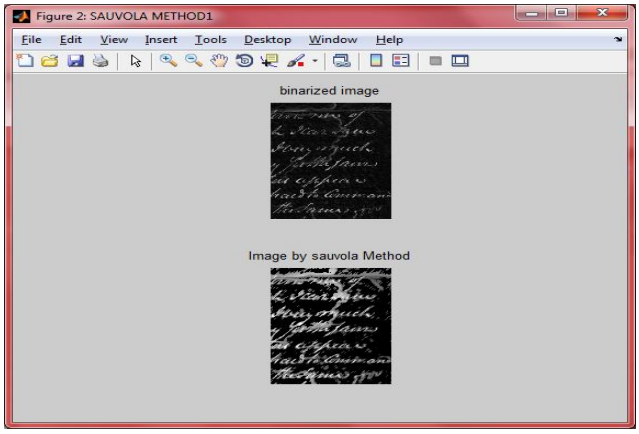


Figure 1.5: (b) Results of Sauvola method

Below show at third, the Bernsen method which reads an image of degraded historical documents from database. Then converts the original image into gray scale image. In third step, Wiener filtered image, fourth as Inpainting binarized restored image by Bernsen method is obtained shown in figure 1.6

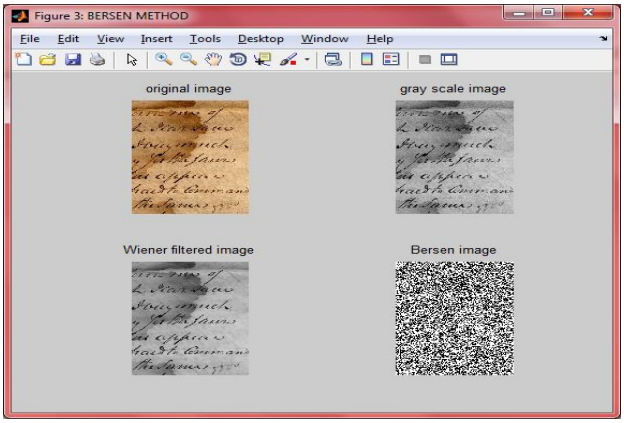


Figure 1.6: Results of Bernsen method

Below show at Fourth, the Bernsen 1st method which reads an image of degraded historical documents from database. Then converts the original image into gray scale image. In third step, Wiener filtered image, fourth as Inpainting binarized restored image by Bernsen 1st method is obtained shown in figure 1.7

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

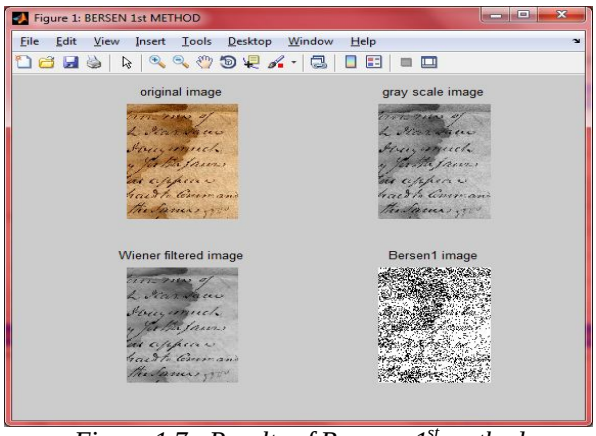


Figure 1.7: Results of Bernsen 1st method

Below show at Fifth, the Bernsen 2nd method which reads an image of degraded historical documents from database. Then converts the original image into gray scale image. In third step, Wiener filtered image, fourth as Inpainting binarized restored image by Bernsen 1st method is obtained shown in figure 1.8

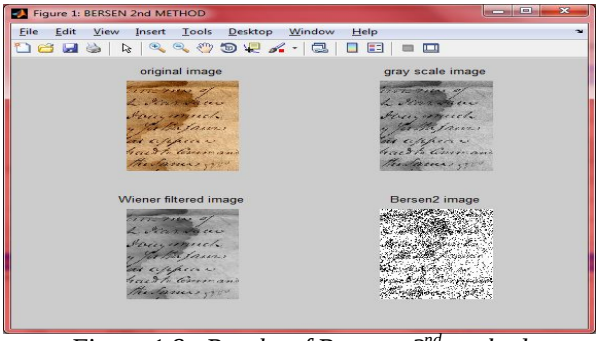


Figure 1.8: Results of Bernsen 2nd method

Below show at Sixth, the combined Bernsen 2nd and OSTU method which reads an image of degraded historical documents from database. Then converts the original image into gray scale image. In third step, Wiener filtered image, fourth as Inpainting binarized restored image by Bernsen 1st method is obtained shown in figure 5.9 (a) and in Figure 5.9 (b) shows the images degraded by OSTU method than by canny edge detecting technique image and at last Combined Bernsen 2nd and OSTU method image.

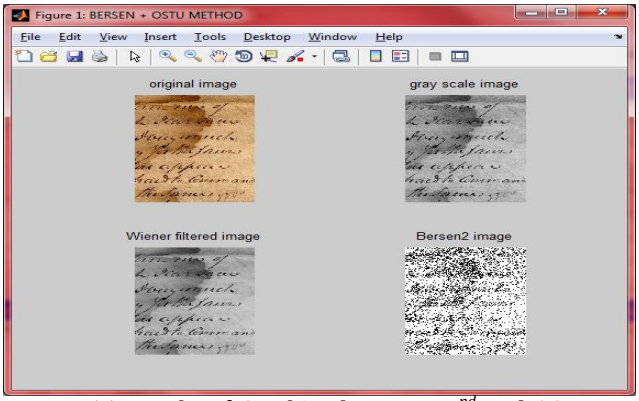


Figure 1.9: (a) Results of Combined Bernsen 2nd and OSTU method

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

Figure 1.9: (b) Results of Combined Bernsen 2nd and OSTU method

Below show at Seventh, the combined Bolan Su and OSTU method which is our proposed work reads an image of degraded historical documents from database. Then converts the original image into gray scale image. In third step, Wiener filtered image, fourth as Inpainting binarized restored image by using method combined as Bolan Su and OSTU is shown in figure 1.10.

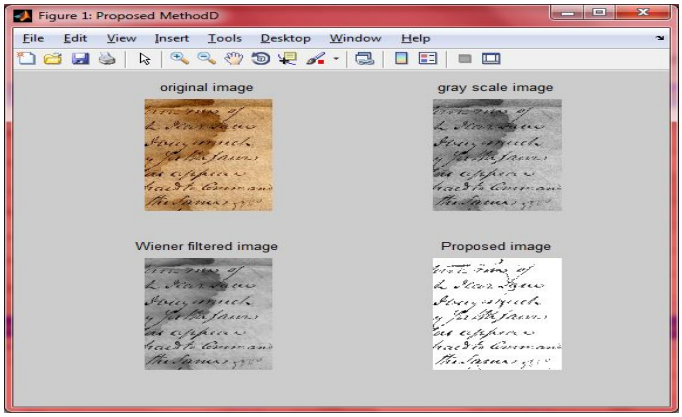


Figure 1.10: Results of Combined Bolan Su and OSTU method (Proposed Method)

In above image 1.4 to 1.10 results are obtained from various method of local and global thresholding method. When results are compared with each other through parameters, proposed method combined with OSTU method performs better than all other methods as shown in above image and discussed below.

V. DISCUSSION

The Below table 1.1 is displaying PSNR, MSE, Coc, Elapsed Time, Accuracy, Precision, Sensitivity as parameters to perform and obtained value which gives out results of best technique that is applied on degraded Inpainting documents from methods such as OTSU, Sauvola, Bernsen, Bernsen 1st, Bernsen 2nd, Bernsen 2nd + OSTU, Proposed (Bolan Su + OSTU) Method. In this, the values of each method are compared with each other. The PSNR values testify the quality of the images produced before and after the algorithm processing. Higher PSNR shows the better quality of the results. A lower value for MSE means lesser error that means minimum MSE value generally indicates that the reconstruction is of higher quality. Coc should be the maximum in order to minimize the difference between original and filtered image. Maximum Coc value shows the better quality of the results. Elapsed time values obtained shows the time taken by each method to process an image. Lesser time will make execution faster but results in better quality through PSNR. Accuracy means the correctness that should be near or less. It is measured in form of true and false or with Boolean value 1 or 0. Again we can say that precision and sensitivity is also measured in true (positive) or false (Negative) condition means 1 or 0 respectively. All these three parameters such as Precision, Accuracy and Sensitivity should be positive means true for best results is shown in table 1.2.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

Table 1.1: Comparing values of PSNR, MSE, Coc, and Elapsed time to performs on degraded Documents

Parameters		PSNR	MSE	Coc	Elapsed Time
Images Name	Techniques				
1.jpg	OSTU Method	34.5902	2627.55	0.886525	0.00676
	Sauvola Method	28.8959	36174.2	-0.844436	1.46312
	Bernsen Method	28.7334	38986.4	0.64225	0.12491
	Bernsen 1st Method	28.7334	38967	0.799805	0.19773
	Bernsen 2nd Method	28.7335	38966.5	0.802894	0.13224
	Bernsen 2nd + OSTU Method	28.7335	38966.5	0.802894	0.13313
	Proposed (Bolan Su + OSTU) Method	86.2546	0.970593	0.697613	1.14008
H02.png	OSTU Method	43.2552	2620.68	0.851875	0.0479883
	Sauvola Method	38.2041	26831.1	-0.827105	13.3065
	Bernsen Method	37.3505	39753	0.415499	0.744543
	Bernsen 1st Method	37.3533	39700.3	0.673247	1.41519
	Bernsen 2nd Method	37.3534	39698.8	0.674152	3.13495
	Bernsen 2nd + OSTU Method	37.3534	39698.8	0.674152	2.88943
	Proposed (Bolan Su + OSTU) Method	83.9477	0.957167	0.780438	8.08383
P03.bmp	OSTU Method	40.8239	2210.92	0.91363	0.0253398
	Sauvola Method	35.2554	28726.5	-0.963741	6.48519

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

	Bernsen Method	34.6013	3882.6	0.327098	0.437545
	Bernsen 1st Method	34.6065	3873.1	0.593192	0.851409
	Bernsen 2nd Method	34.6067	3872.1	0.592309	0.86985
	Bernsen 2nd + OSTU Method	34.6067	3872.1	0.592309	0.864924
	Proposed (Bolan Su + OSTU) Method	96.7176	0.80028	0.82213	3.96526

Table 1.2: Comparing values of Accuracy, Precision, and Sensitivity to performs on Degraded Documents

Parameters		**Accuracy**	**Precision**	**Sensitivity**
Images Name	**Techniques**			
1.jpg	**OSTU Method**	0.999993	0	NaN
	Sauvola Method	0.390169	0	0
	Bernsen Method	0.853532	1	4.62856e-05
	Bernsen 1st Method	0.901342	1	6.87144e-05
	Bernsen 2nd Method	0.902786	1	6.9735e-05
	Bernsen 2nd + OSTU Method	0.897647	0	6.62339e-05
	Proposed (Bolan Su + OSTU) Method	0.999993	0	NaN
H02.png	**OSTU Method**	1	NaN	NaN
	Sauvola Method	0.999504	NaN	0
	Bernsen Method	0.799819	NaN	0
	Bernsen 1st Method	0.932832	NaN	0
	Bernsen 2nd Method	0.937262	NaN	0
	Bernsen 2nd + OSTU Method	0.954672	NaN	0

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

	Proposed (Bolan Su + OSTU) Method	1	NaN	NaN
P03.bmp	**OSTU Method**	0.999569	0	NaN
	Sauvola Method	0.998169	0	0
	Bernsen Method	0.590834	1	0.00105228
	Bernsen 1st Method	0.812267	1	0.00229062
	Bernsen 2nd Method	0.824358	1	0.00244792
	Bernsen 2nd + OSTU Method	0.912336	1	0.00489256
	Proposed (Bolan Su + OSTU) Method	0.999569	0	NaN

From above table 1.1, overall result in all parameters showed Proposed method combing OSTU with Bolan Su method is best as compared to all other OSTU, Sauvola, Bersen's or its Modified Methods. PSNR has high resolution, MSE has less error, Coc has higher quality and Elapsed time is less in all images of degraded historical documents for Proposed Method.

From above table 1.2, again results in best as through parameters showed that Proposed method combing OSTU with Bolan Su method is best as compared to all other OSTU, Sauvola, Bersen's or its Modified Methods. Accuracy is near by 1, Precision is not applicable to judge as true or false. Sensitivity is again related to Precision means 1 or 0. Some time it may be not applicable to find out details for postiveness and negativeness.

VI. CONCLUSION

The algorithm has been worked in the MATLAB using a combination of morphological operations, de-noising filters, image dilation technique and Inpainting binarization, etc. These algorithms have been aimed to solve the problem of digital historical document restoration and to produce better results. The OSTU, Sauvola, Bernsen 2nd Modified + OSTU and Lastly Proposed method algorithm has proved to be efficient when tested with a set of historical images. The historical image restoration algorithm has been tested with the images of various sizes, and the results have been collected and compared with the each other based on Coc (Coefficient of Correlation), Mean Square Error (MSE) and Peak Signal to Noise Ratio (PSNR), Elapsed Time, Accuracy, Precision, and Sensitivity. The Proposed algorithm performed well in the terms of all parameters. Next to it is OSTU algorithm alone proven to be best? The image restoration using Proposed Method than OSTU method has been performed better than Sauvola and Bernsen's all methods in the terms of PSNR, MSE, Coc, Elapsed Time, Accuracy, Precision, and Sensitivity. The PSNR value is used to measure the quality of the image after image processing images. The working technique has been proved to be better on both Local and Global techniques such as, Sauvola and OSTU methods. But combining both gives us best result which in proposed method has be done with Bolan Su as Local and Ostu method as global.

VII. FUTURE SCOPE

Our algorithm is capable of producing the quality results in terms of document clarity and character shape evaluation, but Sauvola algorithm has a critical set back of elapsed time. In the future, the authors/researchers can improve or develop a new model to enhance the performance better than the existing model.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 8, August 2016

REFERENCES

[1] Bolan Su, el.at, "Robust Document Image Binarization Technique for Degraded Document Images", IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 22, NO. 4, APRIL 2013.
[2] Faisal Shafait, Daniel Keysers, Thomas M. Breuel: "Efficient Implementation Of Local Adaptive Thresholding Techniques Using Integral Images", Proceedings of the 15th Document Recognition and Retrieval Conference, Proc. SPIE, Vol. 6815, 2008.
[3] Jagroop Kaur, Dr.Rajiv Mahajan, "A Review of Degraded Document Image Binarization Techniques", International Journal of Advanced Research in Computer and Communication Engineering, Vol. 3, Issue 5, May 2014.
[4] P.S. Jonesherine, J.Paul Richardson Gnanaraj, "Ancient Degraded Document Binarization
Using Mean Shift Technique", International Journal of Advanced Research in Electrical,
Electronics and Instrumentation Engineering (An ISO 3297: 2007 Certified Organization)
Vol. 4, Issue 5, May 2015.
[5] Prof. S. P. Godse, Samadhan Nimbhore, Sujit Shitole, Dinesh Katke, Pradeep Kasar , "Recovery of badly degraded Document images using Binarization Technique", International Journal of Scientific and Research Publications, Volume 4, Issue 5, May 2014, ISSN 2250-3153.
[6] Sushilkumar N. Holambe, Ulhas B. Shinde, Bhagyashree S. Choudhari, "Image Binarization for Degraded Document Images", International Journal of Computer Applications (0975 – 8887)
Volume 128 – No.15, October 2015.
[7] Vikul J.Pawar, "An Improved Binarization Technique for Degraded Document Images using Local Thresholding Method", International Journal of Pure and Applied Research in Engineering and Technology, Volume 2 (8): 71-82, 2014, ISSN: 2319-507x.
[8] Md. Iqbal Quraishi, Mallika De, Krishna Gopal Dhal, Saheb Mondal, and Goutam Das, "A Novel Hybrid Approach To Restore Historical Degraded Documents", ISSP, Vol. 1, pp. 185-189, 2013.
[9] Weisheng Dong, Lei Zhang, Guangming Shi, and Xin Li, "Nonlocally Centralized Sparse Representation for Image Restoration", IEEE Transactions On Image Processing, Vol. 22, No. 4, pp. 1620, April 2013.
[10] Akihito Kitadai, M. Nakagawa, H. Baba, and A. Watanabe, "Similarity evaluation and shape feature extraction for character pattern retrieval to support reading historical documents", in Proc. IAPR Intl. WDAS, pp. 359–363, 2012.
[11] Sandhya N, R. Krishnan and D. R. Ramesh Babu, "A language independent Characterization of Document Image Noise in Historical Scripts", International Journal of Computer Applications, Vol. 50, No. 9, pp. 11-18, 2012.
[12] B. Gangamma and Srikanta Murthy K, "Enhancement of Degraded Historical Kannada Documents", International Journal of Computer Applications, Vol. 29, No.11, pp. 1-6, 2011.
[13] Rachel Mabanag Chong and Toshihisa Tanaka, "Motion Blur Identification Using Maxima Locations for Blind Colour Image Restoration", Future Technology Research Association International, Vol. 1, No. 1, pp. 49-56, December 2010.
[14] Pradeepa D. Samarasinghe, and Rodney A. Kennedy, "Minimum Kurtosis CMA Deconvolution for Blind Image Restoration", IEEE, pp. 271-276, 2008.
[15] Ntogas, Nikolaos and Ventzas, Dimitrios, "A Binarization Algorithm for Historical Manuscripts", International Conference on Communications, pp. 41-51, 2008.
[16] Fung, C.C. and Chamchong, R., "A Review of Evaluation of Optimal Binarization Technique for Character Segmentation in Historical Manuscripts", In: 3rd International Conference on Knowledge Discovery and Data Mining, Phuket, IEEE, PP: 236 – 240, 2010.
[17] Tony F. Chan, Ke Chen, And Jamylle L. Cart," Iterative Methods For Solving The Dual Formulation Arising From Image Restoration", Electronic Transactions on Numerical Analysis, Vol. 26, pp. 299-311, 2007.
[18] R. F. Moghaddam, D. R.-H´enault, and M. Cheriet, "Restoration and segmentation of highly degraded characters using a shape-independent level set approach and multi-level classifiers," in Proc. IAPR ICDAR, pp. 828–832, 2009.
[19] Chen Yan, and Graham Leedham," The Multistage Approach to Information Extraction in Degraded Document Images", IEEE, pp. 2128-2132, 2004.
[20] D. Tschumperl´e and R. Deriche, "Vector-valued image regularization with PDE's: A common framework for different applications," IEEE Trans. PAMI, Vol. 27, No. 4, pp. 506–517, 2005.
[21] Extrapolation, interpolation, and smoothing stationary time series. New York: Wiley, 1949.
[22] Fadoua Drira, Frank LeBourgeois, and Hubert Emptoz," A New PDE-based Approach for Singularity-Preserving Regularization: Application to Degraded Characters Restoration", International Journal of Document Analysis and Recognition, pp. 6-42, 2010.
[23] Hiroko Furuya and Shintaro Eda, Testuya Shimamura," Image Restoration via Wiener Filteringin the Frequency Domain", Wseas Transactions On Signal Processing, Issue 2, Vol. 5, pp. 63-73, February 2009.
[24] K. Shirani Y. Endo, A. Kitadai, S. Inoue, and N. Kurushima, "Character Shape Restoration of Binarized Historical Documents by Smoothing via Geodesic Morphology", ICDAR, Vol. 12, pp. 1285-1289, 2013.
[25] L. Xu, Q. Yan, Y. Xia, and J. Jia, "Structure extraction from texture via relative total variation," ACM TOG (Proc. SIGGRAPH Asia), Vol. 31, No. 6, pp. 139:1–139:10, 2012.
[26] Ms. Ketki R. Ingole and Prof. V K. Shandilya, " Image Restoration of Historical Manuscripts", International Journal of Computer Science & Engineering Technolo, Vol. 2, No. 4, pp. 102-107.