International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-7, Issue-6, March 2019

Restoration of Characters in Degraded Inscriptions using Phase Based Binarization and Geodesic Morphology

Sachin Bhat, Seshikala G

Abstract: ***It is the requirement of the time to store and conserve the ancient manuscripts for the use of next generation. Epigraphists find it hard to decrypt the information present in inscriptions due to variety of reasons including erosion of letters, noise and many more. Here, we present a new binarization and postprocessing technique to efficiently extract and reconstruct the foreground text from heavily degraded documents. The proposed method uses combination of phase based feature maps and geodesic morphology with anisotropic filtering. Phase feature maps will binarize the text by removing background noise and geodesic operators will reconstruct the deteriorated characters. Statistical performance evaluation is done on different datasets and efficiency of the proposed method is demonstrated by comparing with many state of art algorithms.***

Index Terms**:** ***Document Analysis, Image binarization, Morphological operators, Phase congruency***

I. INTRODUCTION

Currently, there is a huge growing pursuit in the domain of document image analysis. Many researchers are trying to develop the systems for the extraction of pertinent information from these documents. It is a dynamic research area being studied from years for tasks like optical character recognition (OCR).

Documentation existed even in the ancient ages without papers or gadgets. Information was created on palm leaves, clothes, metal plates or on stone surfaces. These writings are generally termed as inscriptions or epigraphs. It is the key tool used in history to be able to study the life of ancient time. Manuscripts like epigraphs attained across the globe expose the details of lifestyle, sociocultural environment, political position, fine art and even about the society regarding that time frame and location. These manuscripts became a crucial part of passing the information and procedures from one generation to another. To save this information for the future use to save our culture and heritage is necessity of the day. Preservation of any manuscript generally requires digitization, preprocessing, data extraction and recognition.

Revised Manuscript Received on 30 March 2019.
* Correspondence Author
Sachin Bhat, Department of Electronics and Communication Engineering, Reva University, Bengaluru/ SMVITM, Udupi (Karnataka), India.
Seshikala G, Department of Electronics and Communication Engineering, Reva University, Bengaluru (Karnataka), India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license http://creativecommons.org/licenses/by-nc-nd/4.0/

Primary technique used in any document analysis is preprocessing and binarization which helps in the extraction of useful information from these manuscripts. Document image binarization(BZ) aims at segmenting the foreground characters of a manuscript from the noisy background during the preprocessing stage. manuscripts generally suffer from various degradations over time making BZ an intimidating task. Typically, an epigraph can be heavily degraded due to erosion of characters over time as they are exposed to different environmental conditions, low or varying contrast, stains, intra/inter variation between text and background, missing data noise due to variation of light during image acquisition. Thou' document analysis is being studied across-the-board in the recent past, BZ and postprocessing of highly degraded images is still a largely undiscovered problem which is mainly because of the difficulties in moulding different types of degradations and varying noise. As far as the field of inscriptions and olden manuscripts are concerned, no standard existing methods deal with the issues of foreground extraction, uneven noise removal and character reconstruction. The Document Image Binarization Competition (DIBCO) aims to cover this problem by bringing in datasets of handwritten degraded documents to assess the recent advancement in image binarization. However, enhancement and binarization of ancient manuscripts[1] in still considered to be a tricky task.


(a) (b) (c) (d) (e)

Fig 1(a)-(e). Sample inscription images

Few epigraphs used in this work are shown in Fig 1(a)-(e). The degree of degradation in these materials is too severe(fig 1.b) that it becomes highly challenging to draw out the character shape by binarization. Noisy background adds more complexity while processing and result of binarized characters are fractionated. Sometimes, the images of palm leaf and copper plate inscriptions which were scanned several years ago are available because of further deterioration of the original materials like in fig 1(e).

II. LITERATURE REVIEW

In this part, we briefly depict some of the enhancement and BZ methods used by earlier researchers. Generally, the techniques used for BZ can be either local or global. The global BZ techniques allot a single threshold for the entire image whereas threshold for individual or group of pixels in the document image will be identified in local BZ. Histogram shape based global BZ methods[1][2] tries to estimate a global threshold to minimize intra-class variance. It requires a bimodal histogram pattern and therefore, cannot handle the document images with high variation in background. Though local BZ methods comparatively yield a better result, it is still an unsolved problem in case of ancient manuscripts. This is mainly due to different variety of noise, degradation and unclear foreground. Adaptive thresholding methods like Sauvola[3] which is an improvement over Niblack's[4] and Bernsen[5] will either generate a certain quantity of noise or fail to identify the text with a low contrast[6]. All these algorithms use mean, variance or standard deviation and contrast information of local region to calculate different thresholds. They have also failed to handle the images with light texture background. Feng[7], sets a standard in contrast adaptive thresholding by dynamically calculating the threshold depending on the gray-scale average and variance of current pixel in the neighborhood. Paper [8] demonstrates a real time adaptive thresholding using integral image of the input. This is robust to illumination changes in the image which makes it worthy for the processing of video applications. Lu in [9] used an adaptive image contrast combining the local image contrast and gradient which is tolerant to background and foreground variation caused by different types of degradations. This has been tested on 3 public databases achieving an overall accuracy of 90%. Valizadeh et al in [10] has proposed another method for document image binarization. Here, the image is mapped into a feature space to separate the text and background. This is then partitioned into smaller regions and each region is labelled using result of a basic BZ algorithm. There are few methods based on Independent Component Analysis(ICA) and its variations like Fast ICA and NGFICA that have been used for enhancement of historical documents by maximizing text layer information[11] [12] [13]. This improves the recognition accuracy of word and characters by 65.3% and 54.3% respectively. It uses gradient descent optimization approach to minimize the dependency among the various components present in the source.

Image restoration and retrieval is an active area of research. Different methods involving comparison using local features such as color, shape[14] and other low-level features [15] have been proposed but are unsuccessful in providing good results under addition of correlated noise or change in scale.

There are other approaches being developed in recent times. Some domain knowledge such as texture feature and cross-section sequence graph-analysis [16] can be used to produce better results. However, they need some prior knowledge of the testing document images. [17] proposes an energy-based method using graph-cut algorithm to segment text information by minimizing Laplacian energy. Few works have shown that preprocessing the document images in frequency domain is efficient compared to spatial domain. Frequency domain based phase features are widely used in computer vision, biomedical applications like face recognition, palm-print verification and object identification as feature detectors. Moghaddum et all in [18][19] have shown that foreground of document images can be better handled using phase features. [20] and [21] have also shown that phase is one of the dynamic ways to process the handwritten and printed text. Most of these are based on the work of Kovesi [22]. To the best of our knowledge, compared to the efforts devoted to spatial domain, very little attention is given to use phase as one of the features in document analysis. However, the BZ techniques are not fully explored for all the possible type of degraded documents and still needs further research.

III. PROPOSED METHODOLOGY

Here, we have used 2 types of phase congruency(PC) feature maps[23] of the document image for preprocessing, Expectation Maximization algorithm(EMA) for denoising and geodesic morphological operators for restoration of broken text. Brief flowchart of the method is shown in fig 2. Individual steps are discussed separately below.

A. Preprocessing and Binarization

Phase information of an image in frequency domain always overweighs its magnitude information in spatial domain. There are two approaches to the phase based processing of images. One is Kovesi's PC model[22][24] and another is monogenic scale-space method[25]. We have understood in our experiment that [24] works better for our proposed binarization. PC manifests the nature of image in Fourier domain. Here, the points where the Fourier-components(FC) are maximal are considered as pixels of interest. PC is a ratio of weighted alignment of Fourier components to the sum of Fourier components. Weighted-alignment is the combination of weighted mean function and phase deviation function. Extending this to a 2 dimensional function with scale(p) and orientation(o) is calculated by

$$PC_2(x) = \frac{\sum_p \sum_o W_r(x) \left\lfloor \lfloor A_{po}(x) \cos(\varphi_{po}(x) - \varphi'_{po}(x)) \rfloor \right\rfloor}{\sum_p \sum_o A_{po}(x) + \varepsilon} \quad (1)$$

PC manifests the nature of image in Fourier domain.
Here, the points where the Fourier-Components(FC) are maximal are considered as pixels of interest.

PC is a ratio of weighted alignment of Fourier components to the sum of Fourier components. Weighted-alignment is the combination of weighted mean function and phase deviation function. Extending this to a 2 dimensional function with scale(p) and orientation(o) is calculated by

$$PC_2(x) = \frac{\sum_p \sum_o W_r(x)\left[\lfloor A_{po}(x)\cos(\varphi_{po}(x)-\varphi'_{po}(x))\rfloor\right]}{\sum_p \sum_o A_{po}(x)+\varepsilon} \quad (2)$$

At a given scale and orientation, $\Phi(x)$ is a local phase and A_{po} (x) is local amplitude of nth component in Fourier series expansion. ε is a constant for small fourier amplitudes. Phase deviation function is represented by

$\cos(\Phi n(x) - \Phi'(x))$.

Weighting mean function (W(x)) of PC with is defined as

$$W(x) = \frac{1}{1+e^{g(c-s(x))}} \quad (3)$$

c is cut off value of the filter response below which PC values are penalized. 'g' is gain factor and s(x) is a spread function represented by

$$s(x) = \frac{1}{N}\left(\frac{\sum_p A_p(x)}{A_{max}(x)}\right) \quad (4)$$

PC is noise sensitive. To overcome which, normal distribution is used.

$$G(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \quad (5)$$

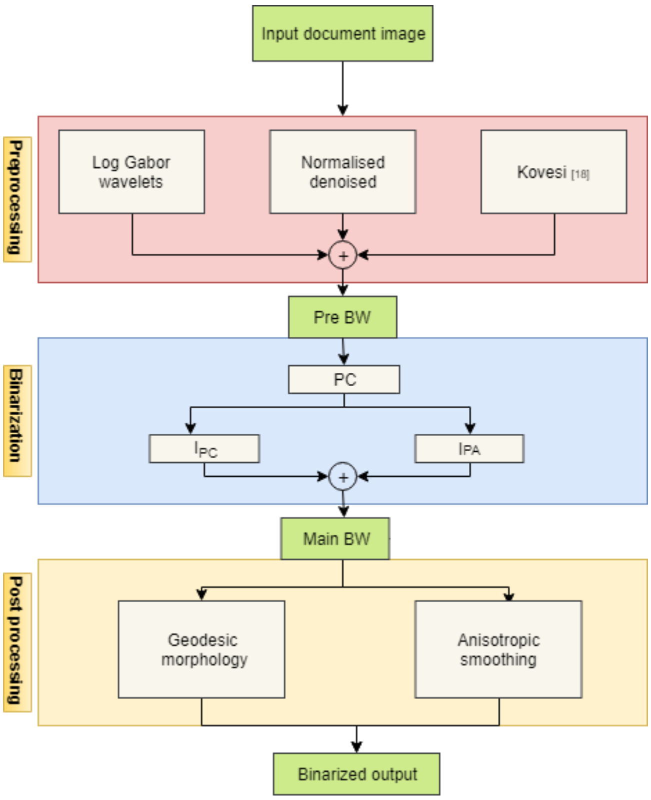


Fig 2. Proposed methodology

Measureof edge strength (I_{PC}) is denoted by

$$I_{PC} = \max PC_2(x), 0 \le I_{PC} \le 1 \quad (6)$$

Sum of all filter responses (I_{PA}) is given as

$$I_{PA} = arctan2\left[\sum_{p,o} u(x), \sum_{p,o} v(x)\right] \quad (7)$$

where $-\pi/2 \le I_{PA} \le \pi/2$

Fig 3 shows an example of I_{PC} andI_{PA} maps for image 1a.

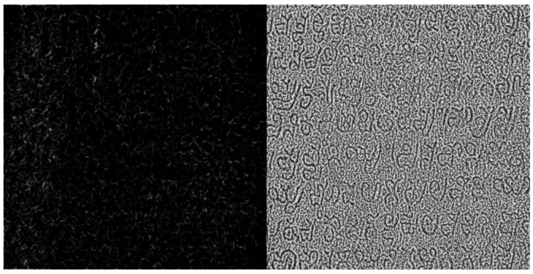


Fig. 3. I_{PC} and I_{PA} maps

Non-orthogonal Gob-gabor wavelets are used to extract the local phase information of each point in the image. Noise threshold of each level is estimated so that magnitudes of filter response vector can be shrunk by soft thresholding while phase information will be untouched. Noise shrinking threshold (T) is calculated by using mean and variance of Gaussian distribution in (5)

$$T = \mu + k\sigma^2 \quad (8)$$

By eliminating the estimated noise, (1) is modeled as

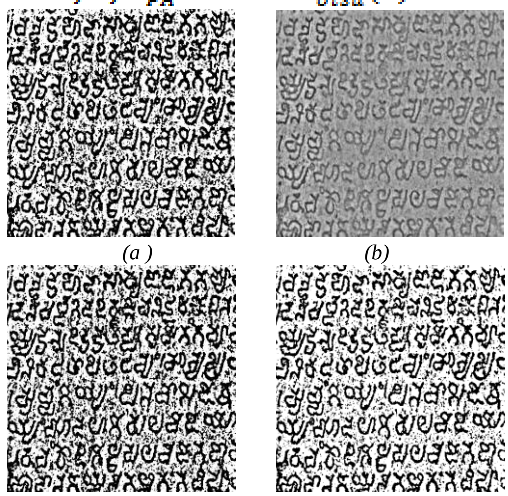
$$PC_2(x) = \frac{\sum_p \sum_o W_r(x)\left[\lfloor A_{po}(x)\cos(\varphi_{po}(x)-\varphi'_{po}(x)) - T\rfloor\right]}{\sum_p \sum_o A_{po}(x)+\varepsilon} \quad (9)$$

Otsu's method is used on normalized denoised image. Otsu's approach will also eliminate the degraded parts of images, because the denoising method tries to reduce the amplitude of the noise component. As it misses false negatives, it is considered as a weak denoising method. With this, I_{PC} used to separate foreground from background pixels can reject most of the badly degraded background pixels by means of a noise modeling method. Log Gabor filter scale 'p' is set to 2 with 10 orientations. Number of standard deviations(k) to reject noise is estimated by the binarization result of Otsu's method. An example of this procedure is shown in fig 4. Output of the preprocessing step given by

$$K = 2 + \left[\frac{\alpha(\sum_{m,n} I_{otsu}(m,n))}{\sum_{m,n} I_{pre}(m,n)}\right] \quad (10)$$

We have considered the following assumption while classifying the pixels of foreground and background using I_{PA}.

$$P(x) = \begin{cases} 1 \ if \ I_{PA} \le 0 \\ 0 \ if \ if \ I_{PA} > 0 \ and \ I_{otsu}(x) = 0 \end{cases} \quad (11)$$



(a) (b)

(c) (d)

(e) *(f)*

Fig 4. Example of the procedure used in pre-processing a) denoised image using Kovesi method. b) denoised normalization. c) Binarization using Otsu's. d) Binariztion using proposed. e) Edge enhancement using Log Gabor wavelets. f) Final binarized image

B. Postprocessing

In post-processing, we focus on the geodesic morphological operators and the inter-convertibility of BZ-image and the distance-transformed(DT) image. Because the DT image has continuous intensity. Well developed edge-aware smoothing filters can be employed on this. Geodesic morphological operators mainly consist of distance transform and transformed image which is then embedded with an anisotropic smoothing filter.

i. Geodesic morphology

Lets consider the input binarized image for $I(p) \in (0,1)$ at pixel p. We apply the Quasi-Euclidean distance transformation[26] with a fast raster scanning algorithm performed once in lexicographical order, and then in inverse lexicographical order to obtain the distance transformed image. Boundary pixel is defined as a black pixel which is adjacent to white pixel(p_b). Let DT, the distance transformed image initializes the distances by using the boundary pixels as

$$DT(p) = \begin{cases} 0 \; if \; p = \; pb \\ \infty \; otherwise \end{cases} \quad (12)$$

non-adaptive thresholding filter with a single threshold for all pixels is employed. Let the binarized image converted back be I'(p) whose pixel values are given by

$$I'(p) = \begin{cases} 0 \; if \; DT(p) \leq t \\ 1 \; otherwise \end{cases} \quad (13)$$

Pixelwise adaptive dilation is performed if threshold t>0 and erosion if t<0.

i. Anisotropic smoothing

Anisotropic smoothing filtering method is applied on distance transformed image. Gradients are smooth for distance transformed images in the range direction. In spatial direction because of the curvature of the edges, variations can be observed[27]. So, local gradient based anisotropic diffusion is suited for the smoothing of distance transformed image.

The edge direction vectors (say θ_+ and θ_-) and gradients (say λ_+ and λ_-) are obtained from single value decomposition(SVD) of the structure tensor G $\in R^{2X2}$ consisting of gradient images of distance transformed image {Dx,Dy} = ∇D as

$$UAU^T = SVD \begin{bmatrix} KD_x^2 & KD_xD_y \\ KD_xD_y & KD_y^2 \end{bmatrix} \quad (14)$$

Where K is a Gaussian filtering to smooth the gradient image. A=diag(λ_+, λ_-) is highest and lowest eigenvalues of G and U=(θ_+, θ_-) is the accompanying eigenvector.

Using eigenvalue and eigenvectors, gradient tensor T $\in R^{2X2}$ is calculated as

$$T = f_+ \left(\sqrt{\lambda_+ + \lambda_-}\right)\theta_+\theta_+^T + f_+ \left(\sqrt{\lambda_+ + \lambda_-}\right)\theta_-\theta_-^T \quad (15)$$

In the above equation, f_+ and f_- control the smoothness of image. Using this gradient tensor, smoothing equation for t+1 iterations is derived as

$$D_{t+1} = D_t + \Upsilon \, T_r(TH) \quad (16)$$

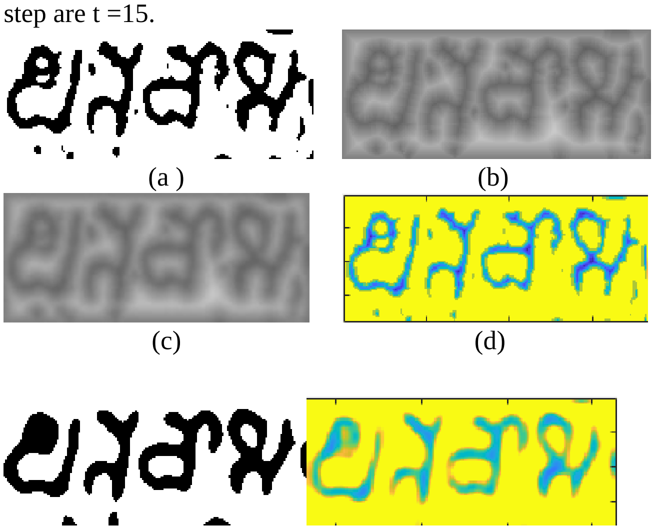
Here, H is a Hessian

$$\nabla(\nabla D)^T = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \quad (17)$$

and ϒ controls the convergence set to 0.5

IV. RESULTS AND DISCUSSION

Fig. 5(a) shows a sample of input binarized image to which distance transformed image and smoothed distance transformed image are displayed in fig. 5(b) and (c). Color enhancement of this image is shown in (d) and (e). Final reconstructed characters are shown in (f). These images are obtained by setting the parameters as follows: resized with a scale of 1/4. Variance of Gaussian filter for a window size of 5 is 1.25. Number of iterations are 40 for ϒ = 1/2 and its time step are t =15.



(a) (b) (c) (d) (e) (f)

Fig.5. Results and evaluation

The proposed method of binarization and postprocessing is first tested on number of available standard datasets like DIBCO11, DIBCO12, DIBCO13 and BICKLEY DIARY. Images in these datasets have suffered from different types of degradations. They are enough challenging in terms of evaluating and formulating purposeful examination of various algorithms. We then fine tune the parameters to work well on our customized dataset of inscription images. Around 200 images have been collected from different places of south India. It is observed that proposed method suits well for low intensity, low contrast type of images with unclear background and undefined type of noise. This code is executed in MATLAB 2018b running on windows 10.

For the inscription image in 1(e), I_{PC} $and I_{PA}$ are shown in fig.6. Comparison of binarized image using [18] and proposed method is shown in fig 7(a) and (b). Image is Normalized by linear transformations with Otsu's algorithm is used to remove some amount of noise from the binarized output. Though this is a weak denoise technique, misclassified objects below threshold T are removed as in fig 8(a). Final reconstructed image using anisotropic smoothing is shown in fig 8(b).

Fig 6. I_{PC} and I_{PA} maps

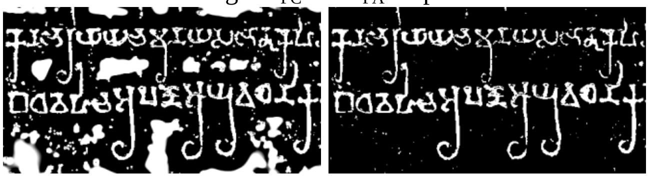


Fig 7. Binarization using (a) Kovesi (b) proposed method

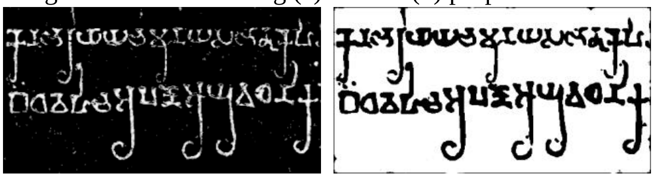


Fig 8. (a) Final binarized image (b) Final reconstructed image

We have used several traditional state-of-art-algorithms and top ranking methods of DIBCO contest for comparison. Many matrices are available for the evaluation of binarization model in the algorithms reported. Out of which, 4 of them have been used in our study. They are F-measure, Peak signal-to-noise ratio(PSNR), Negative-Rate-Metric(NRM), Misclassification Penalty Metric (MPM). These parameters are widely used in the objective analysis of image-quality-measurement. Higher values of PSNR, F-measure and lower values of MRM, NRM provide better binarization results. Proposed method of phase congruency model is compared with different classical binarization methods in the literature namely Otsu, Bernsen, Sauvola, Niblack, Gatos. DIBCO11, DIBCO12, DIBCO13, BICKLEY and our customized inscription database of 50 images are considered for evaluation. Comparison results are shown in Figures 9-14. Proposed method produces better results compared to most o the algorithms.

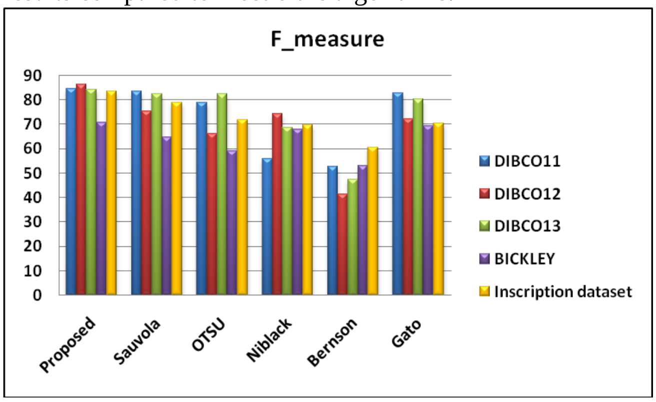


Fig 9. Comparison of different datasets against F-measure

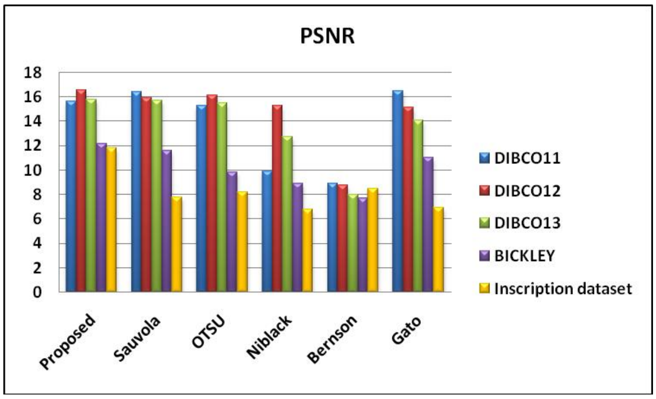


Fig 9. Comparison of different datasets against PSNR

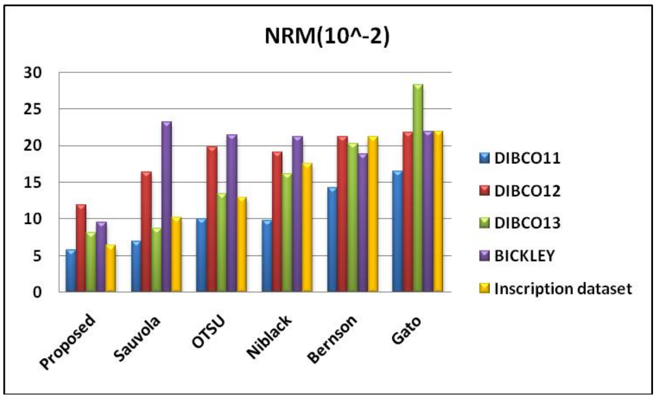


Fig 9. Comparison of different datasets against NRM

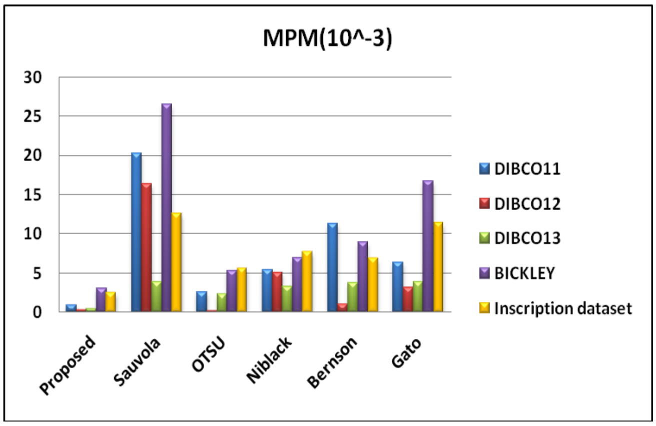


Fig 9. Comparison of different datasets against MPM

V. CONCLUSION

A new binarization and postprocessing technique is presented here to efficiently extract and reconstruct the foreground text from heavily degraded documents. Proposed method uses combination of phase based feature maps and geodesic morphology. Statistical performance evaluation is done by calculating PSNR,NRM,FM and MPM. These parameters are compared with the results obtained from inscriptions and other classical methods to demonstrate efficiency of the work. This plays an important role segmentation, recognition and classification of characters from inscriptions and epigraphs.

REFERENCES

1. N. Chaki, S. H. Shaikh, and K. Saeed, "Exploring image binarization techniques," *Studies in Computational Intelligence*, Springer 2014
2. N. Otsu, " A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics,* vol. 9, no. 1, pp. 62–66, 1979.
3. J. Sauvola and M. Pietikainen, "Adaptive document image binarization," Pattern recognition, vol. 33, no. 2, pp. 225–236, 2000.
4. W. Niblack, "An introduction to digital image processing", vol. 34, Englewood Cliffs: Prentice-Hall, 1986.
5. J. Bernsen, "Dynamic thresholding of grey-level images," *In Proc. 8th Int. Conf. on Pattern Recognition*, 1986, pp. 1251–1255
6. B. Gatos, K. Ntirogiannis, and I. Pratikakis, "ICDAR 2009 document image binarization contest (DIBCO 2009)," *International Conference on Document Analysis and Recognition*, pp. 1375–1382, July 2009.
7. Feng, Meng-Ling, and Yap-Peng Tan, "Contrast adaptive binarization of low quality document images," *IEICE Electronics Express* 1, no. 16 (2004): 501-506.
8. Bradley, Derek, and Gerhard Roth. "Adaptive thresholding using the integral image." *Journal of graphics tools* 12, no. 2 (2007): 13-21.
9. 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).
10. M. Valizadeh and E. Kabir, "Binarization of degraded doc- ument image based on feature space partitioning and classification." International Journal on Document Analysis and Recognition, pp. 1–13, December 2010.
11. Sreedevi, Indu, et al. "Enhancement of inscription images." *2013 National Conference on Communications (NCC)*. IEEE, 2013.
12. Sreedevi, Indu, Rishi Pandey, N. Jayanthi, Geetanjali Bhola, and Santanu Chaudhury. "NGFICA based digitization of historic inscription images." *ISRN Signal Processing* 2013 (2013).
13. Sachin Bhat, Avinash N J, "ICA algorithm for image enhancement and improvement of word and character recognition in epigraphs", International Journal of Current Engineering and Scientific Resrarch", vol.4, no.5, pp. 72-76, 2017
14. Rui, Y., Huang, T.S., Chang, S.F.: Image retrieval: Current techniques, promising directions, and open issues. Journal of visual communication and image representation 10 (1999) 39–62
15. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Interna- tional journal of computer vision 60 (2004) 91–110
16. A. Dawoud, "Iterative cross section sequence graph for handwritten character segmentation," IEEE Trans. Image Process., vol. 16, no. 8, pp. 2150–2154, Aug.2007
17. N. Howe, "A Laplacian energy for document binarization," in Proc. Int. Conf. Doc. Anal. Recognit., Sep. (2011), pp. 6–10
18. R. F. Moghaddam and M. Cheriet, "A multi-scale framework for adaptive binarization of degraded document images," Pattern Recognit., vol. 43, no. 6, pp. 2186–2198, 2010
19. H. Z. Nafchi, R. F. Moghaddam, and M. Cheriet, "Historical document binarization based on phase information of images," *In Proc. Asian Conference on Computer Vision*, pp. 1-12. Springer, Berlin, Heidelberg, 2012.
20. H. Z. Nafchi and H. R. Kanan, "A phase congruency based document binarization," *In Proc. International Conference on Image and Signal Processing*, pp. 113-121. Springer, Berlin, Heidelberg, 2012.
21. Nafchi, Hossein Ziaei, Reza Farrahi Moghaddam, and Mohamed Cheriet, "Application of phase-based features and denoising in postprocessing and binarization of historical document images," *In Proc. Document Analysis and Recognition (ICDAR), 2013 12th International Conference on*, pp. 220-224. IEEE, 2013.
22. P. Kovesi, "Phase preserving denoising of images," signal, 1999
23. Sachin Bhat, and G. Seshikala. "Preprocessing and Binarization of Inscription Images using Phase Based Features." *2018 Second International Conference on Advances in Electronics, Computers and Communications (ICAECC)*. IEEE, 2018.
24. H. Z. Nafchi, R. F. Moghaddam, and M. Cheriet, "Phase-based binarization of ancient document images: Model and applications," *IEEE transactions on image processing*, vol. 23, no. 7, pp. 2916–2930, 2014.
25. M. Felsberg and G. Sommer, "The monogenic scale-space: A unifying approach to phase-based image processing in scale-space," J. Math. Imag. Vis., vol. 21, nos. 1–2, pp. 5–26, 2004.
26. K. Shirai, Y. Endo, "Character Shape Restoration of Binarized Historical Documents by Smoothing via Geodesic Morphology", 12th International Conference on Document Analysis and Recognition, 2013, p:1285-1289
27. 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.

AUTHORS PROFILE

Sachin Bhat received his M.Tech degree in Digital Electronics and Communication from NMAMIT, Nitte. He is a research scholar in the School of Electronics and Communication Engineering, Reva University, Bengaluru and working as a senior Assistant Professor in SMVITM, Udupi. His areas of interest are Document Image Processing, Deep Learning, Remote Sensing and Object Oriented Programming. He has published over 30 research papers in national and international journals.

Dr. G. Seshikala has 25 years of teaching and research experience. She holds PhD degree in Biomedical Signal Processing from JNTU-A. BE in ECE and and ME in Digital Electronics. She has published over 20 research papers in international journals and conferences. Her areas of specialization are Communication Engineering, Image Processing, Pattern Recognition and Signal Processing.