

Research Article

NGFICA Based Digitization of Historic Inscription Images

Indu Sreedevi,¹ Rishi Pandey,¹ N. Jayanthi,¹ Geetanjali Bhola,¹ and Santanu Chaudhury²

¹ *Electronics and Communication Engineering Department, Delhi Technological University, Formerly Delhi College of Engineering, Delhi 110042, India*

² *Electrical Engineering Department, Indian Institute of Technology Delhi, Delhi 110006, India*

Correspondence should be addressed to Indu Sreedevi; s.indu@rediffmail.com

Received 28 February 2013; Accepted 8 April 2013

Academic Editors: L.-M. Cheng, W.-L. Hwang, and P.-Y. Yin

Copyright © 2013 Indu Sreedevi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper addresses the problems encountered during digitization and preservation of inscriptions such as perspective distortion and minimal distinction between foreground and background. In general inscriptions possess neither standard size and shape nor colour difference between the foreground and background. Hence the existing methods like variance based extraction and Fast ICA based analysis fail to extract text from these inscription images. Natural gradient flexible ICA (NGFICA) is a suitable method for separating signals from a mixture of highly correlated signals, as it minimizes the dependency among the signals by considering the slope of the signal at each point. We propose an NGFICA based enhancement of inscription images. The proposed method improves word and character recognition accuracies of the OCR system by 65.3% (from 10.1% to 75.4%) and 54.3% (from 32.4% to 86.7%), respectively.

1. Introduction

A significant amount of research has been carried out in the direction of reading inscriptions from monuments around the world. Several methods have been proposed for detection of text, localization and extraction of text from images of inscriptions [1, 2]. But, the problem of text extraction intensifies when the difference in the text (foreground) and the background is very marginal, the background is textured, or the background and foreground are similar. Such is the case of camera-held images of inscriptions at the sites of historical monuments. Figure 1 shows an image of inscription found in world heritage site “Hampi.” These inscriptions are generally found engraved into/projected out from, stone, or other durable materials. However, due to effects of uncontrolled illuminations, wrapping, multilingual text, minimal difference between foreground and background images, and the distortion due to perspective projection as well as the complexity of image background, extracting text from these images is a challenging problem.

The commercially available Optical Character Recognition (OCR) has very poor recognition accuracy of images of the inscriptions on monuments. The images of English

inscriptions from the monuments were passed through the commercial OCR for text extraction, but the OCR failed to recognize these images. These images can be recognized by OCR only after proper enhancement. Fast ICA [3] based enhancement method has given good results for inscription images with a reasonable colour difference between text and the background. Most of the ancient inscriptions do not have such reasonable colour distinction between the two regions. Therefore to digitize such inscriptions we have to enhance the difference between the two regions. This paper proposes a method to enhance the minimal difference between text and non-text regions of such inscription images.

Natural gradient-based flexible ICA (NGFICA) has been extensively used in separating highly correlated signals [4] as it minimizes dependency among the different signals present in the source signal using gradient descent optimization approach. For minimizing the dependency among the foreground and background of historical inscription images, we used NGFICA for obtaining independent components of the images. This paper presents a novel enhancement technique to separate the text part of the inscription image by processing NGFICA output of inscription image.

2. Related Work

Text extraction from document images has been of interest for the research community over a decade, but there has been very few work done in digitizing inscription images of historical monuments. High contrast edges between text and background are obtained using the red color component in the approach by Agnihotri and Dimitrova [5]. In [6], the “uniform color” blocks within the high contrast video frames are selected to correctly extract text regions. Kim et al. [7] used 64 clustered color channels for text detection where cluster colors are based on Euclidean distance in the RGB space. The method based on variance by Babu et al. [8] makes use of the variance in the text and nontext regions. The variance is high at the text edges and vice versa. Variance method to extract the text as in [8] did not prove successful due to blurred edges of text and minimum distinction between text and nontext regions. The text in inscription images does not consist of a uniform color, and there is low contrast between text and background thus making the use of [6] unsuitable. Simple edge-based approaches [9] are also considered useful to identify regions with high edge density and strength. This method performs well if there is no complex background, but the inscription images have complex background, thus these methods cannot be used directly.

The authors of [10] estimate the intensity of nontext region (background) and do binarization in comparison with a threshold intensity. Laplacian of Gaussian filters in Sobolev space using different α factors for different images is used in [11] for enhancement of text images. Curvelet transform is proposed in [12] for denoising the degraded historical documents. Adaptive binarization technique for Palm Leaf Manuscripts proposed in [13], where authors used Wiener filter for noise removal and contrast adaptive binarization for segmentation of text from the background [14], proposes a wavelet based enhancement/smearing algorithm for the removal interfering strokes in archiving handwritten document images. In [15] authors proposed a hybrid approach which includes both local and global thresholding techniques for cleaning background noise from the ancient documents. The results in [15] show that enhancement has been achieved but cannot be read by OCR. The above-said methods were based on binarization, text extraction using variance, or edge detection based methods. These methods depend upon pixel's threshold value based on difference between foreground and background parts.

Garain et al. [3] describe how to enhance image using Fast ICA algorithm which results in three independent components or layers which correspond to the contribution of text in them. The method is an enhancement method, which; however, is unable to enhance inscription images which are weak or highly spatially correlated sources. More recently, its convergence has been shown to slow down or even fail in the presence of saddle points, particularly for short block sizes [16]. Moreover it is proved in [17] that Fast ICA fails in separating the sources for weak or closed sources.

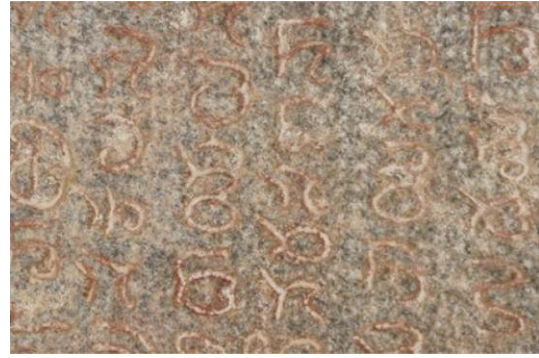


FIGURE 1: Inscription found at Hampi.

In case of unclear and complex archaeological inscription images, there is no sharp distinction between foreground and background. Natural gradient-based independent component analysis learning algorithm with flexible nonlinearity as described in [18, 19] gives better results than other algorithms as it is more efficient in minimizing dependence among correlated signals. In the proposed method we used NGFICA for minimizing the dependency between foreground, middle layer, and background of such inscription images and further the characters are retrieved from the foreground.

3. Motivation

Many of the inscriptions are couched in extravagant language, but when the information gained from inscriptions can be corroborated with information from other sources such as still existing monuments or ruins, inscriptions provide insight into world's dynastic history that otherwise lacks contemporary historical records. Digital archiving of these images is necessary for conservation and accessibility. The major challenges of digitization of such images are blurred edges of the text and minimum distinction between text and nontext parts. NGFICA algorithm deals with Gaussian, sub-Gaussian, and super-Gaussian source signals as is the case with the said inscription images [18]. We are proposing a novel method to enhance degraded inscription images using NGFICA in this paper.

3.1. Methodology. We separated the text (foreground), nontext (background), and noise of an image as three different components using NGFICA as it minimizes dependency among the components. We further refined all ICA outputs using morphological operation. The ICA output image with average threshold farther from average threshold of original image gave good results.

3.1.1. Finding the Independent Components. The images of inscriptions were complex because of the high correlation of foreground pixels with the background pixels. This merging of pixels deteriorates the clarity of the inscription images. The noise from the background due to illumination, shadows, and so forth added to the problem of clarity of regions



FIGURE 2: (a) Source image. (b) Final enhanced image.

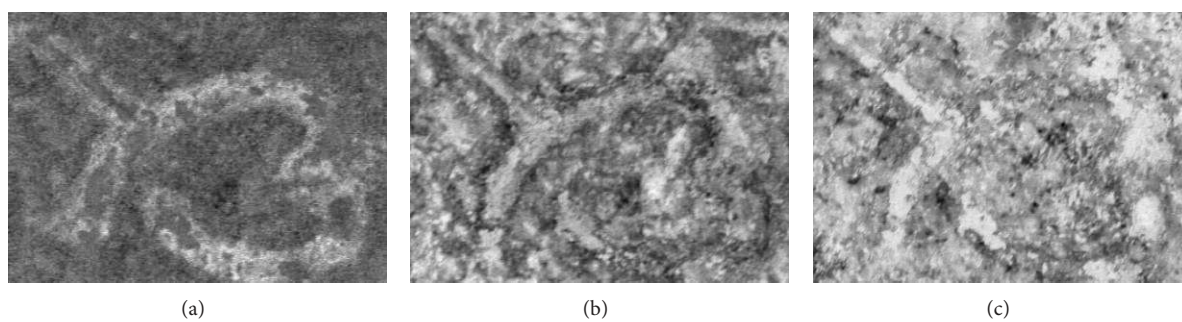


FIGURE 3: (a), (b), and (c) NGFICA output images.

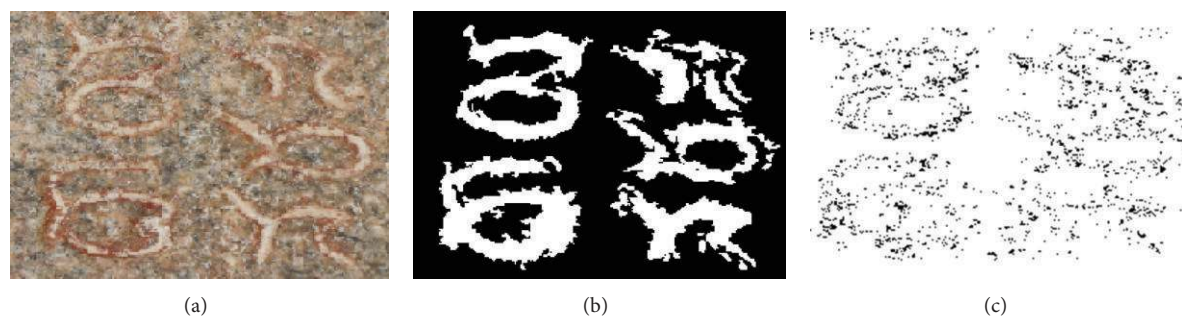


FIGURE 4: (a) Original image, (b) output image of proposed method, and (c) output after Fast ICA based enhancement.

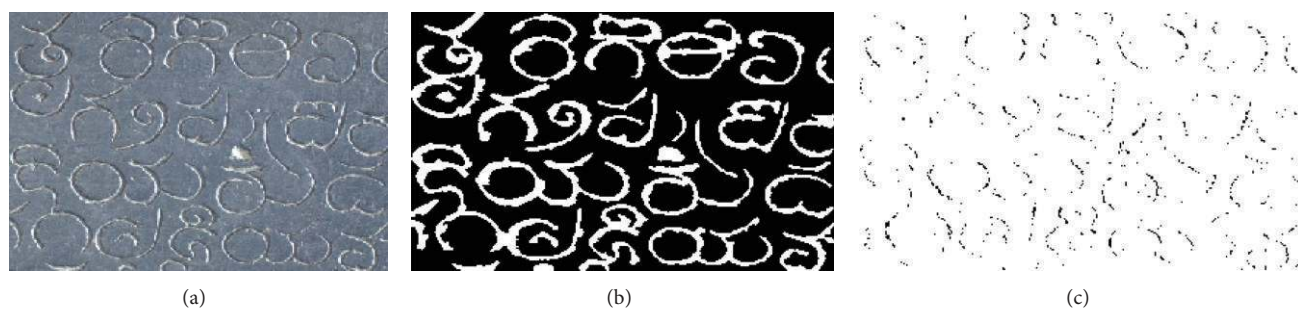


FIGURE 5: (a) Original image, (b) output image of proposed method, and (c) output image of Fast ICA based enhancement.



FIGURE 6: Other language results.

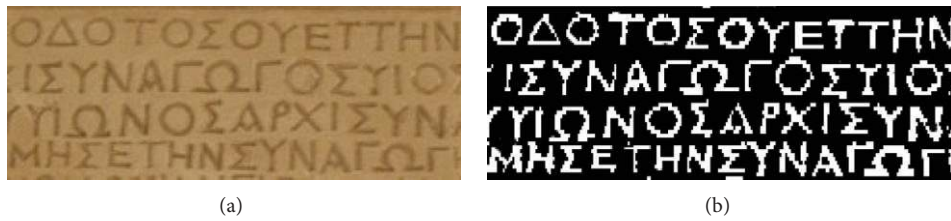


FIGURE 7: Other language 1. (a) Source and (b) result.



FIGURE 8: Other language 2. (a) Source and (b) result.



FIGURE 9: (a) Image of inscriptions. (b) Output of proposed method. (c) Corresponding OCR output.



FIGURE 10: (a) Image of inscriptions. (b) Output of proposed method. (c) Corresponding OCR output.



FIGURE 11: (a) Source (b) result.

(text, nontext). So, we performed Gaussian smoothing of the colored image using a 5×5 kernel. This helped to remove small scale noise and irrelevant image details. Then R, G, and B components of the smoothed image were extracted. Three independent components of the colored image were obtained by performing NGFICA on the extracted R, G, and B components. For reference purpose these independent components were named as text layer, nontext layer, and mixed layer on the basis of their contribution to the text.

The NGFICA [20] can be explained in the simplest possible way as follows. An image can be considered as a mixture of foreground and background and a common part for both. Let the mixing model be

$$X = AY, \quad (1)$$

where $X[n \times T]$ is the original image, $Y[n \times T]$ is the unknown mutually independent portion of the image, and $A[n \times n]$ is the mixing matrix. For a 3-channel image, (1) can be written as follows.

The de-mixing model is defined as

$$\begin{bmatrix} x_r \\ x_g \\ x_b \end{bmatrix} = \begin{bmatrix} r_1 & r_2 & r_3 \\ g_1 & g_2 & g_3 \\ b_1 & b_2 & b_3 \end{bmatrix} \times \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}, \quad (2)$$

$$Z = WX,$$

where $Z[n \times T]$ is the separated sources and $W[n \times n]$ be the de-mixing matrix

$$\nabla L(W) = W^{-T} - E[g(Y)X^T], \quad (3)$$

where $g(Y)$ is given by

$$g(Y) = -\frac{d}{dY} (\log P(Y)). \quad (4)$$

The randomized gradient $\nabla L(W)$ of formula given by (3) expresses the steepest drop direction of the Euclidean space cost function $L(W)$. The natural gradient $\bar{\nabla} L(W)$ is the steepest drop direction in Riemann space of parameters W . The natural gradient $\bar{\nabla} L(W)$ can be calculated by modifying the random gradient, which is obtained by multiplying $W^T W$

TABLE 1: Comparison with Fast ICA method.

Input image	Number	Fast ICA accuracy in %	Proposed method accuracy in %
Words	550	0.9%	75.4%
Character	2578	1%	86.7%

in the random gradient as given by (5). Thus NGFICA algorithm gives faster convergence and better performance. Faster convergence is due to the fact that decorrelation is performed together with separation, and the better performance is due to the nonlinear function $g(Y)$ controlled by the Gaussian exponent

$$\bar{\nabla} L(W) = (I - E[g(Y)Y^T])W. \quad (5)$$

To minimize dependency among output components we have to minimize cost function $L(W)$.

As explained in [21] gradient adaptation is a useful technique for adjusting a set of parameters to minimize a cost function. The natural gradient is based on differential geometry and employs knowledge of the Riemannian structure of the parameter space to adjust the gradient search direction. Unlike Newton's method, natural gradient adaptation does not assume a locally quadratic cost function. Moreover, for maximum likelihood estimation tasks, natural gradient adaptation is asymptotically Fisher efficient. The three independent components of NGFICA are shown in Figure 3.

3.1.2. Character Extraction from Foreground. NGFICA output image can be considered as foreground, back ground, and noise images. We compared the average threshold of each of these output images with that of original image. The one which is farther from original image is identified as the foreground as it had only text. The image shown in Figure 3(b) is identified as the foreground, and the image is further enhanced using Sobel edge detection and then dilation using disc-shaped structuring element to retrieve the characters (Figure 2).

TABLE 2: OCR Accuracy before and after enhancement.

Input image	Number	Rec. by OCR before enhancement	Accuracy in % before enhancement	Rec. by OCR after enhancement	Accuracy in % after enhancement
Words	550	56	10.1%	415	75.4%
Characters	2578	835	32.4%	2235	86.7%

4. Results and Discussion

The dataset for validating the proposed method was prepared by gathering images of inscriptions belonging to historical monuments (India Gate, new Delhi, India), heritage sites (Hampi, Karnataka), ancient temples (Vishnu temple, Tamil Nadu), and so forth. Such inscriptions are common at almost every monument and normally found engraved into/projected out from stone or other durable materials. Some images were clicked manually using a 10-megapixel camera, and some others were taken from Internet. The images demonstrated several processing difficulties like uneven illumination, wrapping, perspective distortion, multilingual text, text with foreground and background images, and so forth. The images of India Gate (English) without enhancement were tested on web-based OCR [22], and the results are shown in Table 2. The enhanced outputs of India Gate images using the proposed method are shown in Figures 9 and 10. We have also compared the proposed method with Fast ICA based enhancement [3], and results are shown in Table 1 in Figure 4. Other results of Hampi inscription images are shown in Figures 5 and 11.

The proposed method also worked equally well for other languages too. The results are shown in Figures 6, 7, and 8. We have tested the method on 650 images of which 550 were English word images which were passed through OCR and word accuracy of 75.4%, and character accuracy of 86.7% was achieved. The remaining 100 images of different languages gave very good results. Natural gradient algorithm not only deals with symmetric distribution of the signal but also can deal with asymmetrical distribution of the signal.

5. Conclusion

A novel method for enhancement of complex and unclear archaeological inscription images has been enhanced and validated using 650 word images. The proposed method establishes the important role of NGFICA in digitizing inscription images which has minimal distinction between text and nontext region, and blurred edges for the text. The method improved the word and character recognition accuracies from 10.1% to 75.4% and from 32.4% to 86.7%, respectively. The method proved successful in efficiently enhancing multilingual inscription images too. This method can be further extended for digitization of ancient coins, manuscripts, and archaeological sculptures.

References

- [1] S. S. Kuo and M. V. Ranganath, "Real time image enhancement for both text and color photo images," in *Proceedings of the IEEE International Conference on Image Processing (ICIP '95)*, vol. 1, pp. 159–162, October 1995.
- [2] C. Wolf, J.-M. Jolion, and F. Chassaing, "Text localization enhancement and binarization in multimedia documents," in *Proceedings of the 16th International Conference on Pattern Recognition (ICPR '02)*, vol. 4, pp. 1037–1040, September 2002.
- [3] U. Garain, A. Jain, A. Maity, and B. Chanda, "Machine reading of camera-held low quality text images: an ICA-based image enhancement approach for improving OCR accuracy," in *Proceedings of the 19th International Conference on Pattern Recognition (ICPR '08)*, pp. 1–4, December 2008.
- [4] X. Wen and D. Luo, "Performance comparison research of the fecg signal separation based on the bss algorithm," *Research Journal of Applied Sciences Engineering and Technology*, vol. 4, no. 16, pp. 2800–2804, 2012.
- [5] L. Agnihotri and N. Dimitrova, "Text detection for video analysis," in *Proceedings of IEEE International Workshop on Content-Based Access of Image and Video Libraries*, pp. 109–113, June 1999.
- [6] X. S. Hua, P. Yin, and H. J. Zhang, "Efficient video text recognition using multiple frame integration," in *Proceedings of the International Conference on Image Processing*, vol. 2, pp. 22–25, September 2004.
- [7] K. C. Kim, H. R. Byun, Y. J. Song et al., "Scene text extraction in natural scene images using hierarchical feature combining and verification," in *Proceedings of the 17th International Conference on Pattern Recognition (ICPR '04)*, vol. 2, pp. 679–682, gbr, August 2004.
- [8] G. R. M. Babu, P. Srimayee, and A. Srikrishna, "Heterogenous images using mathematical morphology," *Journal of Theoretical and Applied Information Technology*, vol. 15, no. 5, pp. 795–825, 2008.
- [9] E. Nadernejad, S. Sharifzadeh, and H. Hassanpour, "Edge detection techniques:evaluations and comparisons," *Applied Mathematical Sciences*, vol. 2, no. 31, pp. 1507–1520, 2008.
- [10] M. Seeger and C. Dance, *Binarising Camera Images for OCR*, Xerox Research Centre, Meylan, France, 2000.
- [11] S. Buzykanov, "Enhancement of poor resolution text images in the weighted sobolev space," in *Proceedings of the 19th International Conference on Systems, Signals and Image Processing (IWSSIP '12)*, pp. 536–539, Vienna, Austria, April 2012.
- [12] B. Gangamma, K. Srikanta Murthy, and A. V. Singh, "Restoration of degraded historical document image," *Journal of Emerging Trends in Computing and Information Sciences*, vol. 3, no. 5, pp. 36–39, 2012.
- [13] S. Cherala and P. Rege, "Palm leaf manuscript/color document image enhancement by using improved adaptive binarization method," in *Proceedings of the 6th Indian Conference on Computer Vision, Graphics and Image Processing (ICVGIP '08)*, pp. 687–692, December 2008.
- [14] C. L. Tan, R. Cao, and P. Shen, "Restoration of archival documents using a wavelet technique," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 10, pp. 1399–1404, 2002.

- [15] A. V. S. Rao, N. V. Rao, L. P. Reddy, G. Sunil, T. S. K. Prabhu, and A. S. C. S. Sastry, "Adaptive binarization of ancient documents," in *Proceedings of the 2nd International Conference on Machine Vision (ICMV '09)*, pp. 22–26, December 2009.
- [16] P. Tichavský, Z. Koldovský, and E. Oja, "Performance analysis of the FastICA algorithm and Cramér-Rao bounds for linear independent component analysis," *IEEE Transactions on Signal Processing*, vol. 54, no. 4, pp. 1189–1203, 2006.
- [17] P. Chevalier, L. Albera, P. Comon, and Ferreol, "Comparative performance analysis of eight blind source separation methods on radio communications signals," in *Proceedings of the International Joint Conference on Neural Networks*, vol. 8, pp. 251–276, July 2004.
- [18] S. Choi, A. Cichocki, and S. I. Amari, "Flexible independent component analysis," *Journal of VLSI Signal Processing Systems for Signal, Image, and Video Technology*, vol. 26, no. 1, pp. 25–38, 2000.
- [19] S. Choi, "Independent component analysis," in *Proceedings of the 12th WSEAS International Conference on Communications*, pp. 159–162, July 2008.
- [20] S. Amari, A. Cichocki, and H. H. Yang, "A new learning algorithm for blind signal separation," in *Advances in Neural Information Processing Systems*, vol. 8, pp. 752–763, MIT Press, Cambridge, Mass, USA, 1996.
- [21] S. Amari and S. Douglas, *Why Natural Gradient? Brain Style Information Systems*, Saitama, Japan, 2001.
- [22] "Optical character recognition," <http://www.onlineocr.net/>.

