

{tag}

{/tag}

Digital Image and Signal Processing  
© 2016 by IJCA Journal

IJCA Proceedings on National Conference on

NCDISP 2016 - Number 1

Year of Publication: 2016

Authors:

G. G. Rajput

Suryakant B. Ummapure

Panditkumar Patil

{bibtex}ncdisp201627.bib{/bibtex}

## Abstract

A generic approach for the separation of overlapping and touching lines within handwritten text document images is proposed in this paper. Presence of touching or skewed that arises due to ascenders or descenders and style of writer makes text line extraction a difficult task. The approach is based on histogram and connected component analysis. The proposed method is a three stage approach wherein non overlapping lines are extracted during the first stage and separation of oriented and touching lines occurs during second and third stages respectively.

Average height of a text line computed using histogram profile forms the basis for text line segmentation. The proposed method has been evaluated on 120 handwritten documents written in English, Devanagari, Kannada, Telugu, and Malayalam scripts containing non-overlapping and overlapping or touching occurrences.

## Refer

## ences

- Lemaitre, Aurélie, and Jean Camillerapp. "Text-line extraction in handwritten document with Kalman filter applied on low resolution image". Document Image Analysis for Libraries, 2006. DIAL'06. Second International Conference on. IEEE, 2006.
- Anusree. M and Dhanya. M. Dhanalakshmy. "Text-line Segmentation of Curved Document Images". Anusree. M et al. Int. Journal of Engineering Research and Applications ISSN : 2248-9622, Vol. 4, Issue 5( Version 5), May 2014, pp. 32-36
- Sunanda Dixit, Sneha, Nilotpal Utkalit and Suresh . H. N. "Text-line Segmentation of Handwritten Documents in Hindi and English". International Journal on Recent and Innovation Trends in Computing and Communication ISSN: 2321-8169 Volume: 2 Issue: 4 733 – 739.
- Vikas J Dongre and Vijay H Mankar. "DEVNAGARI DOCUMENT SEGMENTATION USING HISTOGRAM APPROACH". International Journal of Computer Science, Engineering and Information Technology (IJCSIT), Vol. 1, No. 3, August 2011. 4
- Neha Sahu. "DEVANAGIRI DOCUMENT SEGMENTATION USING HISTOGRAM BASED APPROACH". International Journal of Electronics, Electrical and Computational System IJEECS ISSN 2348-117X Volume 3, Issue 3 May 2014.
- Saiprakash Palakollu, RenuDhir and Rajneesh Rani. "A New Technique for Line Segmentation of Handwritten Hindi Text". Special Issue of International Journal of Computer Applications (0975 – 8887) on Electronics, Information and Communication Engineering - ICEICE No. 5, Dec 2011.
- Saiprakash Palakollu, RenuDhir and Rajneesh Rani. "Segmentation of Handwritten Devanagari Script". SaiprakashPalakollu et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (3), 2011, 1244-1247. ISSN: 0975-9646.
- Rahul Garg and Naresh Kumar Garg. "An algorithm for Text-line Segmentation in Handwritten Skewed and Overlapped Devanagari Script". International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 5, May 2014).
- Varsha Hole, LeenaRagha and Pravin Hole. "Text-line and Word Segmentation of Indian Script Handwritten Document". International Conference & Workshop on Recent Trends in Technology,(TCET) 2012 Proceedings published in International Journal of Computer Applications@IJCA).
- M. Ravi Kumar, B. P. Pragathi and Nayana N Shetty. "Text-line Segmentation of Handwritten Documents using Clustering Method based on thresholding Approach". International Journal of Computer Applications (0975 – 8878),on National Conference on Advanced Computing and Communications - NCACC, April 2012
- Nazih Ouwayed, Abdel Belaid and Francois Auger. "General Text-line Extraction

Approach based on Locally Orientation Estimation". Author manuscript, published in "Document Recognition and Retrieval XVII - DRR 2010, 17th Document Recognition and Retrieval Conference, San Jose, CA : United States (2010)".

- Saiprakash Palakollu, RenuDhir and Rajneesh Rani. "Handwritten Hindi Text Segmentation Techniques for Lines and Characters". Proceedings of the World Congress on Engineering and Computer Science 2012 Vol IWCECS 2012, October 24-26, 2012, San Francisco, USA. 12

- Jayant Kumar, Le Kang David, Doermann Wael ,Abd-Almageed. "Segmentation of handwritten text lines in presence of touching components. " Document Analysis and Recognition (ICDAR), 2011 International Conference on. IEEE, 2011.

- NazihOuwayed, Abdel Belaid. "Separation of Overlapping and Touching Lines within Handwritten Arabic Documents". Xiaoyi Jiang and Nicolai Petkov. The 13th International Conference on Computer Analysis of Images and Patterns - CAIP 2009, Sep 2009, Munster, Germany. Springer Berlin / Heidelberg, 5702, pp. 237-244.

- Ram Sarkar, Nibaran Das, Subhadip Basu, Mahantapas Kundu, Mita Nasipuri and Dipak Kumar Basu. "CMATERdb1: a database of unconstrained handwritten Bangla and Bangla-English mixed script document image". IJDAR DOI 10. 1007/s 10032-011-0148-6 Published online:24 February 2011.

- Rafael C. Gonzalez and Richard E. Woods " Digital Image Processing"; Third Edition, Published by Pearson Education, Inc. and Dorling Kindersley Publishing, Inc. ISBN 978-81-317-1934-3.

- Shafali Goyal, Ashok Kumar Bathla " Method for Line Segmentation in Handwritten Documents with Touching and Broken Parts in Devanagari Script". International Journal of Computer Applications (0975 – 8887) Volume 102– No. 12, September 2014.

- Rahul Garg, Naresh Kumar Garg " An algorithm for Text Line Segmentation in Handwritten Skewed and Overlapped Devanagari Script". International Journal of Emerging Technology and Advanced Engineering Website: www. ijetae. com (ISSN 2250-2459, ISO 9001:2008 Certified Journal, Volume 4, Issue 5, May 2014).

- PRAMOD S. MALGI & SHAILJA GAYAKWAD "LINE SEGMENTATION OF DEVNAGRI HANDWRITTEN DOCUMENTS". International Journal of Electronics, Communication & Instrumentation Engineering Research and Development (IJEIERD) ISSN (P): 2249-684X; ISSN (E): 2249-7951 Vol. 4, Issue 2, Apr 2014, 25-32© TJPRC Pvt. Ltd.

- Abdollah Amirkhani-Shahraki, Amir Ebrahimi Ghahnavieh and Seyyed Abdollah Mirmahdavi. "A Morphological Approach to Persian Handwritten Text Line Segmentation". 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation.

- Raid Saabni , Abedelkadir Asi , Jihad El-Sana . "Text line extraction for historical document images" Pattern Recognition Letters 35 (2014) 23–33 0167-8655/\$ - see front matter \_ 2013 Elsevier.

- Naresh Kumar Garg, Lakhwinder Kaur and M. K. Jindal. "Segmentation of Handwritten Hindi Text". ©2010 International Journal of Computer Applications (0975 – 8887) Volume 1 – No. 4

- Abderrazak Zahour, Brunco Taconet, Laurence Likforman-Sulem and Wafa Bousellaa. "Overlapping and multi-touching text-line segmentation by Block Covering analysis". Pattern Anal Applic (2009) 12:335–351 DOI 10. 1007/s10044-008-0127-9 Springer.

- Satadal Saha, Subhadip Basu, Mita Nasipuri and Dipak Kr. Basu . "A Hough

Transform based Technique for Text Segmentation". JOURNAL OF COMPUTING, VOLUME 2, ISSUE 2, FEBRUARY 2010, ISSN 2151-9617 .

- KALYAN TAKRU and GRAHAM LEEDHAM "SEPARATION OF TOUCHING AND OVERLAPPING WORDS IN ADJACENT LINES OF HANDWRITTEN TEXT".

Proceedings of the Eighth International Workshop on Frontiers in Handwriting Recognition (IWFHR'02) 0-7695-1692-0/02 \$17. 00 © 2002 IEEE.

- G. G. Rajput, Suryakant B. Ummature and Preeti N Patil. "Text-Line Extraction from Handwritten Document images using Histogram and Connected Component Analysis". International Journal of Computer Applications (0975 – 8887) National conference on Digital Image and Signal Processing, DISP 2015.

Computer Science

### **Index Terms**

Image Processing

### **Keywords**

Handwritten Document; Text-line; Segmentation; Histogram; Connected Component.