



RESEARCH ARTICLE

Multioriented and Curved Text Lines Extraction from Documents

Vaibhav Gavali¹, B. R. Bombade²

¹M. Tech Student, Department of CSE, SGGS IE&T, Nanded, India

²Assistant professor, Department of CSE, SGGS IE&T, Nanded, India

¹ vaibhavgavali@gmail.com; ² b.r.bombade@gmail.com

Abstract— There is need of the robust algorithm to extract text lines from script independent documents, color independent, font and size independent segmentation algorithms. This paper presents simple method to extract curved and multioriented text lines from the documents. The input is may be colored or grayscale image. Discrete wavelet transform is applied on input image to get four sub-bands. Thresholding is applied on the three sub-bands (horizontal, vertical, diagonal). Edge detection is applied on the each sub-band after thresholding. Morphological dilation using different dilation operators for different sub-bands are used. Finally logical and operator is on the 3 dilated components to get the curved and Multioriented text lines as output.

Key Terms: - Curved; edge; text lines; dilation; wavelet transform

Full Text: <http://www.ijcsmc.com/docs/papers/June2013/V2I6201383.pdf>