Skewed Line Detection and Removal Preserving Handwritten Strokes A New Approach

Abstract

Text overlapping with lines poses serious problems for the optical character recognition systems. The dilemma becomes crucial for skewed and non-uniform thick line present in the word image. Although detection and removal of the straight underlines has been addressed but still skewed lines removal and restoration of the area after removal of lines persists to be a problem of interest. A new method is proposed to detect and remove skewed and straight line at any position inherited in the word image without characters distortion to avoid restoration stage by preserving strokes. The proposed technique is based on connected component analysis and is equally suitable to remove straight and skewed line from printed and handwritten words. Detailed experiments are conducted on manually filled forms of National Institute of Standard and Technology (NIST) special benchmark database19. Comparisons with other methods available in the literature exhibit potential of the new approach with accuracy up to 95.18%.