

Oral presentation preferred

Natural Scene Text Extraction for Improved Recognition

Natural scenes are now acquired every second by handheld imaging devices whose market and use are increasing drastically. It leads to a volume of images to process and new applications appear attempting to reach the expected “life-log” one. Our context is mainly based on providing reading to blind or visually impaired users. Natural scene images imply strong degradations, absent in scanner-based acquisition, such as uneven lighting, highlights or shadows, non-formatted text, a wide range of fonts, very complex backgrounds, blur, sensor noise, low resolution and so on and our system proposes to handle these constraints with the larger versatility. Two parts will be detailed in this presentation: our selective metric clustering algorithm based on image formation models, which exploits hue and intensity in the same color space by using several similarities between pixels (Euclidean distance and Vector Angle similarity) and our log-Gabor based character segmentation into individual components to circumvent failures of traditional segmenters while drastically increasing recognition rates. In camera-based document analysis systems, errors of subsequent steps are propagating much more than in conventional document analysis and a particular point is given on relationships between these two parts to lead to an extraction-by-segmentation-by-recognition method.

Several results, comparisons with different algorithms, shall be finally described to enhance performance of our system.