

Application for English text finding from an image file

Natpreya Tangudamkarn and Rajalida Lipikorn

Department of Mathematics and Computer Science, Faculty of Science, Chulalongkorn University,
Tel: 086-3784877, e-mail: Natpreya.t@gmail.com

Abstract: The topic of the classroom action research is “Application for English text finding from an image file”. The objectives are to reduce the time which user use to search text within image, to improve the accuracy of Tesseract OCR and to improve the accuracy of searching method. The aim of this project is to create program which can search text from image files in Windows. This project applied the knowledge from “Domain Transform for Edge-Aware Image and Video Processing”, proposed by E. S. L. Gastal and M. M. Oliveira, to smoothing image while preserving edges, “Color Reduction for Complex Document Images”, proposed by N. Nikolaou and N. Papamarkos, used for color segmentation and “Font and background color independent text binarization”, proposed by T. Kasar, J. Kumar and A. G. Ramakrishnan, to binarize color image to binary image. These researches used for improving efficiency in image processing and accuracy in text conversion (OCR). The overall results from text conversion is good, its accuracy being improved. However, its speed should be improved to be faster. The ways to do is use other algorithms which have more efficiency, or do more research about old module and find another module that can solve the problem.

Keywords: OCR, image processing, computer vision

TA 012
