

Character Extraction and Recognition For Myanmar Script Signboard Images using Block based Pixel Count and Chain Codes

Kyi Pyar Zaw

kyipyarzaw08@gmail.com

University of Computer Studies, Mandalay, UCSM

Zin Mar Kyu

zinmarkyu.pp@gmail.com

University of Computer Studies, Mandalay, UCSM

This paper presents a very simple and efficient method for the text extraction and recognition of the Myanmar text from color natural signboard images taken by a mobile phone camera. Text extraction, line segmentation, character segmentation and recognition are the important steps in text understanding from natural signboard images. In this system, the color enhancement is firstly processed to overcome various illumination conditions. Background noises on the binary images are removed by four filtering features such as color threshold based filtering, aspect ratio based filtering, boundary based filtering and region area based filtering. After removing the noise, line segmentation and character segmentation are done. Horizontal projection profile is used for line segmentation and vertical projection profile and bounding box methods are used to segment the characters. These connected component characters are recognized by using 4×4 blocks based pixel density and total chain codes, 4-rows based pixel density, 4-columns based pixel density and count of eight directions chain code on the whole character image and on each block of character image. This system is investigated by feature based approach of template matching, and 83.15% character recognition accuracy is achieved on 2854 correctly extracted characters from 150 camera-captured Myanmar warning signboards.