

Myanmar Character Recognition Using Eight Direction Chain Code Frequency Features

Authors : Kyi Pyar Zaw, Zin Mar Kyu

Abstract : Character recognition is the process of converting a text image file into editable and searchable text file. Feature Extraction is the heart of any character recognition system. The character recognition rate may be low or high depending on the extracted features. In the proposed paper, 25 features for one character are used in character recognition. Basically, there are three steps of character recognition such as character segmentation, feature extraction and classification. In segmentation step, horizontal cropping method is used for line segmentation and vertical cropping method is used for character segmentation. In the Feature extraction step, features are extracted in two ways. The first way is that the 8 features are extracted from the entire input character using eight direction chain code frequency extraction. The second way is that the input character is divided into 16 blocks. For each block, although 8 feature values are obtained through eight-direction chain code frequency extraction method, we define the sum of these 8 feature values as a feature for one block. Therefore, 16 features are extracted from that 16 blocks in the second way. We use the number of holes feature to cluster the similar characters. We can recognize the almost Myanmar common characters with various font sizes by using these features. All these 25 features are used in both training part and testing part. In the classification step, the characters are classified by matching the all features of input character with already trained features of characters.

Keywords : chain code frequency, character recognition, feature extraction, features matching, segmentation

Conference Title : ICIAP 2017 : International Conference on Image Analysis and Processing

Conference Location : Dubai, United Arab Emirates

Conference Dates : November 24-25, 2017