

Abstract

Ethnic diversity of Indonesia produces Indonesian cultural diversity. They ancestor create culture that still exist until now. Ethnic have their own way to communicate like verbal communication and written communication. History found some script that used by ancestor like sansekerta script, Javanese script, Sudanese script and etc. Only people who have enough knowledge can translate that script. In this Final Task, knowledge about Sudanese script will be used for computer to recognize and translate Sudanese script. To recognize an image of Sudanese script, so it will be done preprocessing stage include cropping, resizing and thinning phase. Chain code is used as a method for feature extraction. Chain code describes an object to a sequence of directions which that direction is obtained by tracing an image. Then the result of chain code is being normalized. Normalized chain code will be assessed its similarity with the knowledge of the system using hamming distance method and then it's classified. Based on the experiment, average accuracy for font sundanese unicode is 78, 75% for font size 18, 20, 24 and 26 pt. For written sundanese script, accuracy from 7 data sample is 40%.

Key word: OCR, feature extraction, Sundanese script, normalized chain code, chain code, hamming distance.