

ABSTRAK

Dokumen teks sastra Jawa merupakan akar budaya bangsa Indonesia yang perlu dilestarikan. Pengenalan huruf pada sebuah citra dokumen teks sastra Jawa adalah salah satu cara untuk melestarikan kebudayaan tersebut. Segmentasi adalah salah satu proses dalam pengolahan citra pada citra dokumen teks yang bertujuan untuk memisahkan obyek-obyek yang ada pada citra dokumen teks. Hasil pemisahan obyek tersebut akan digunakan sebagai masukan untuk proses pengenalan huruf citra dokumen teks sastra Jawa.

Pada tugas akhir ini dibuat perangkat lunak untuk melakukan segmentasi citra dokumen teks sastra Jawa menggunakan algoritma *watershed*. Algoritma *watershed* menganggap sebuah citra merupakan bentuk tiga dimensi yaitu posisi x, posisi y dan tinggi dengan tingkat warna piksel paling gelap menjadi dasarnya. Data masukan untuk perangkat lunak ini adalah sebuah file gambar dengan format *.jpg. Perangkat lunak ini dikembangkan dengan menggunakan bahasa pemrograman Java dan bahasa pemrograman *MATLAB 6.5* yang dihubungkan dengan *JMATLink*.

Berdasarkan pengujian terhadap 10 citra dokumen teks sastra Jawa, diperoleh segmentasi obyek *watershed* dengan rata-rata prosentase kebenaran dan rata-rata prosentase akurasi masing-masing sebesar 97,123% dan 84,932% sehingga dapat disimpulkan bahwa pemilihan metode segmentasi obyek ini relatif baik. Untuk segmentasi karakter *watershed* rata-rata prosentase kebenaran sebesar 57,443% sehingga dapat disimpulkan bahwa pemilihan metode segmentasi karakter ini kurang baik dan rata-rata waktu yang dibutuhkan adalah 39,7267 detik yang menjelaskan bahwa metode segmentasi yang digunakan relatif singkat.

ABSTRACT

The Javanese literature texts are the roots of Indonesian people's culture which need to be preserved. Knowing the image document texts of the Javanese literatures is one of the methods to process image in the image document texts which is intended to split the objects on the image document text. The result of the splitting the objects would be used as the entry for the process to know the letters of image document texts in the Javanese literature texts.

In this final project, a software was made to do the segmenting of the image document texts from the Javanese literature documents using Watershed algorithm. Watershed algorithm considers an image as a tridimensional form of X position, Y position and its height with the deepest cell colour as its background. The entry of this software was an image file with .jpg format. This software was developed using Java language programming and MATLAB 6.5 language programming which was linked to JMATLink.

Based on the test result on the 10 document image texts of the Javanese literatures, the watershed object segmentation achieved the average of 97,123% validity and average of 84,932% accuracy which then concluded that this selection of the object segmentation method was quite good. Because the average accuracy of Watershed segmentation character achieved 57,443% it was concluded that the choice of character segmentation method was not really good and the average time needed was 39,7267 seconds which showed that this used segmentation method was quite short.