Research and Implementation of a new postprocessing algorithm to the engineering drawings vectorization and recognition, based on strip cutting method

Yalin Wang
P O Box 84-187, Tsinghua University Postoffice
Beijing 100084, P.R.China. Email: wyl@sun5.thit.tsinghua.edu.cn

Abstract

The research in automatic input and understanding systems for engineering drawings has lasted for many years, and there are many progresses having been made. The THRV system which was approved days ago is one of the progressing systems. However, the performance of vectorization is still beyond practical acquirements. Postprocessing procedures are required definitely.

In this thesis, we try to set up a platform for postprocessing, which contains positional indexing functions which will facilitate the postprocessing of vectorization. The concepts of Object-Oriented programming and B-tree are employed to implement the platform

For our strip cutting method, the entire drawing area is divided into adjacent horizontal strips of equal width. Each strip has a strip number and each strip contains a node list holding zero or more nodes. Each index node is a rectangle area in the drawing with a left boundary and right boundary. The nodes in each strip are sorted by their boundaries. The sorted index node sequence is realized as a balanced tree(B-tree).

The applications presented in this thesis have shown the effectiveness and efficiency of the platform. Based this platform, We propose a generic graphic object recognition method to recognize various graphic objects from the vectorized results. The algorithm works bottom-up and consists of two main steps: finding the first component and extending it. In the first step, finds the first key component of the object to be recognized by following the clue types. Then in the second step, uses the key component to extend the detection of the object in the proper directions to the possible maximum extent in order to find all the other components.

Real-time pick-up operation and recognition of dashed lines is described as the cases in point. In the end of this thesis, we have also discussed further uses of it.

Key Words: Engineering Drawings Recognition and Understanding, Vectorization, Postprocess, Data Structure, B-tree