Available online at www.sciencedirect.com
ScienceDirect
Engineering Applications of Artificial Intelligence 21 (2008) 1073–1079
Engineering Applications of ARTIFICIAL INTELLIGENCE
www.elsevier.com/locate/engappai

# Graph theory-based approach for automatic recognition of CAD data

H.C. Huang[a], S.M. Lo[a,*], G.S. Zhi[b], R.K.K. Yuen

[a]*Department of Building and Construction, City University of Hong Kong, Tat Chee Avenue, Kowloon Tong, Hong Kong*
[b]*Xi'an Jiaotong University, Xi'an, PR China*

Received 24 February 2006; accepted 2 December 2007
Available online 6 March 2008

**Abstract**

CAD architectural plans basically contain original geometrical information of graphical primitives. However, in many applications, such as building's 3D reconstruction, auto-detecting errors of design, original CAD data are very hard to be directly utilized. It is always a time-consuming and exhaustive task to extract useful information like the coordinates of a line from CAD files. To make this task performed in more efficient way, a way to develop an automatic method to extract spatial information from architectural plans produced in the form of computer-drawn CAD drawings is proposed in this article. The aim of the proposed method is to provide automatic transformation of architectural drawings into the spatial and topological information of the enclosure in a building. To auto-understand the 'meaning' of the graphic elements in the drawings such as walls, doors and rooms, the approach employs algorithms in graph theory, which can identify every functional component in the enclosure and establish their connectivity relationships. The method has been implemented by using object-oriented C++ language and is found to be able to produce satisfying results.


*Keywords:* Graph theory; Information extraction; Computer-aided design

## 1. Introduction

Nowadays, computer technology has been widely applied in various engineering application fields. In mechanical/manufacturing and architectural fields, design of products or buildings depends more and more on computer-aided systems. In the past, a CAD system was relatively simply established and mainly acted as an interactive graphics platform. Recently, more intelligent tools have been added into the system or separately developed, providing the capability of automatically finishing some works such as building's 3D reconstruction, auto-detecting errors of design, plan evaluation, plan checking, etc. However, prior to using these tools, we usually need to carry out pre-process works, which require extensive preparation of input data. It is always a time consuming and exhaustive task, especially for complex setting. For example, the detection of errors from a complex building plan requires the extraction of extensive information about the configuration of every room and door.

Generally, a CAD drawing contains original geometrical information of graphical primitives. However, the original CAD data is very hard to be directly utilized. In traditional pre-process modules, input data relating to the CAD drawings are required to be manually extracted and stored in a file. In fact, the information represented by the data is naturally embraced in the CAD drawings. To make the pre-process task performed more efficiently and accurately, it is necessary to develop a method to automatically extract meaningful information from the original drawings.

The study on automatic capturing CAD drawing is closely related to the specific engineering application. The development of generic recognition of CAD plans originated from the mechanical/manufacturing field (Prabhu and Pande, 1999; Meeran and Taib, 1999). Related information about the attributes of the products is captured to provide the input for manufacturing/production process. In the recent decades, some works have been

*Corresponding author. Tel.: +852 27887683; fax: +852 27887612.
*E-mail address:* bcsmli@cityu.edu.hk (S.M. Lo).

0952-1976/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.engappai.2007.12.001

done on the intelligent understanding (IU) of architectural plans with the aim of assisting building designers in architectural, structural or cost analysis.

Ah-Soon and Tombre (1997) and Ah-Soon (1997) and have performed architectural plan analysis on the basis of symbol detection, geometrical and spatial analysis. In their approach, graphic symbols are represented by a set of geometrical features constraints. These constraints are propagated through the network hierarchy. Similar to Ah-soon's constraint network, Messmer's (1996) network for exact and inexact graph matching also has shown a good performance in recognition of the architectural symbols. Lewis and Sequin (1998) and Bukowski and Sequin (1997) have presented a method called building model generator (BMG) to build 3D building models from 2D architectural plans form a smoke-spread prediction model, CFAST. The BMG uses room label as seed point to find interior space contours, and can provide topological correction and semantic enrichment architectural model for CFAST. The BMG have largely reduced the modeling time; however, it is a semi-automatic processor that requires users to point out the location of each vertex. Other models like building EXODUS (Owen et al., 1996), SIMULEX (Thompson and Marchant, 1995) and SGEM (Zhi et al., 2003) for evaluating the evacuation system of a building also have been tried to incorporate automatic function to capture CAD information from architectural plans. These models are not fully automatic and require pre-treatment of the CAD files.

This article further describes the study on the use of graph theory algorithms to analyze and capture the spatial information of the building layout from architectural plans. In a CAD file, the vertices and lines connecting them are first identified; and then a graph is established, in which the nodes denote vertices and arcs denote lines. On the basis of the graph theory, the topological relationship of the nodes is analyzed according to some rules. As the result, the spatial information of the enclosures (rooms, passages, stairs, obstacles, etc.) inside the building can be obtained for specific applications. The proposed method has been implemented in C++ and integrated into a building evacuation simulation model as the pre-process module.

## 2. Problem statement and difficulties

The goal of the study is to extract spatial information inside a building from architectural plans. In a CAD drawing, a floor plan can be decomposed into many enclosure components with openings or without opening, such as compartments, corridors, halls, stairwells, walls, columns, etc. The capturing process needs to identify all the enclosures, their types and connectivity relationships. Each enclosure has its boundary represented by a series of vertices. So the process should start with recognizing all the key vertices including endpoints and intersection points. After determining the boundaries of every enclosure, their relationships should be further recognized. On such basis, we attempt to conclude the type and function of each enclosure. The task of information reorganization can be defined as: "Given a floor plan, identify the boundary of each enclosure and decide its type and relations with other enclosures." For example, Fig. 1 presents a possible floor plan of a building consisting of four rooms, one passage and one column. After identifying all the vertices, enclosures A–F can be recognized. The doors, common vertices can be used to decide connectivity relationships of these enclosures; and then the type of each enclosure may be presumable. Table 1 lists the output of capturing. Details of this process are further described in Section 3 and 4.

To capture information from architectural plans, there are several difficulties that need to be considered (Zhi et al., 2003):

1. drafting errors in CAD files may cause misunderstanding;
2. useless information such as dimensions, text notes, and other technical information/specifications must be removed;
3. it is hard to locate the openings of a enclosure usually represented as the doors(single-leaf, double-leaf);

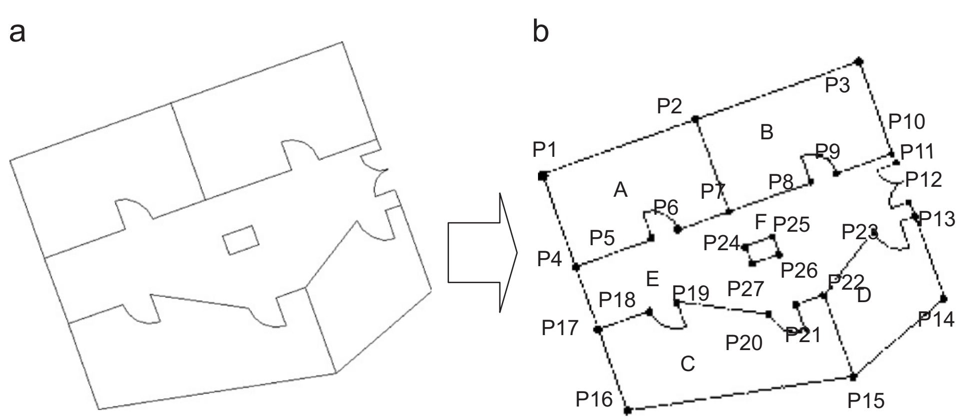


Fig. 1. Task of recognizing a floor plan (a) given floor plan and (b) result after processing.