Fifth International Conference on Construction in the 21st Century (CITC-V)
"Collaboration and Integration in Engineering, Management and Technology"
May 20-22, 2009, Istanbul, Turkey

An Automated Entities Extraction Model for Elements Recognition from the Structural 2D Drawing

Chitrakala M.
Universiti Teknologi Malaysia, JohorBahru, Malaysia
chitrakala_m@yahoo.com

Mohd Zaimi A. M., Abdul Hakim Mohammed
Universiti Teknologi Malaysia, JohorBahru, Malaysia
drzaimi@utm.my, dd1rmc@hotmail.com

Abstract
The players in construction industry are still relying on manual extraction from the drawing for the element's details for quantity estimating, planning and scheduling, and other related activities. Unlike industrialized construction, the conventional construction is under increased pressure to measure the quantities and under tight schedules exercise automation for a better project management. Since the design serves as a source for all construction processes, the building information modeling with the specifications and the details based on elements provides the possibility of accessing the required details from the drawing directly. Conversely, the 2D drawing is complex enough to automatically recognize and identify the elements, their shape, and the properties like length, height etc from the primitive drawing shapes such as lines, arcs and texts. However the development of fuzzy logic's rule based algorithm recognizes the elements through its properties. This paper presents the conceptual model which has been developed through extracting and recognizing the element's details from the structural 2D digital drawing effectively.

Keywords
2D structural drawing, Entities extraction, Element recognition algorithm

1. Introduction

The construction industry is information intensive. Since every project of this industry is unique because of its size, geological area properties, method of construction, technology used, etc., these constraints fragment the construction industry processes. However, the construction industry has been criticized for being slow to accept and apply modern management methods for planning, estimating and executing of projects and this is said to explain, in part, poor construction performance (Oglesby *et al.*, 1989). Marsh and Flangan (2000) also mentioned that the use of the Information Technology (IT) evidence suggests that the construction industry is yet to realize full benefits from IT utilization. Soetanto *et al*. (2005) describes that appropriate use of advanced IT will help to achieve performance efficiency and effectiveness in construction processes. Moreover numbers of research have been taken by large number of researchers to enhance the operational efficiency, accuracy and other issues such as time, man power, cost, and quality as an optimized solution to make construction industry under single roof. As a result, the Information and Communication Technology (ICT) must be viewed as potential resources to enhance the efficiency and accuracy in construction and approached in an integrated way among the fragmented sections of the construction industry.

2. Research Background

In Malaysian construction industry as observed, the design serves as the source for all construction phases from planning until facilities. Figure 1 clearly shows that the digital CAD drawing serves as source for all construction phases. But, all the construction activities are mostly done manually. Except architects, all the industry players from engineers to facility managers are forced to reproduce their activities manually from the CAD drawing that consumes cost and time. Chitrakala *et al.* (2008) clearly defines how the drawing is used in Malaysian construction industry and the details of manual extraction for the activities.

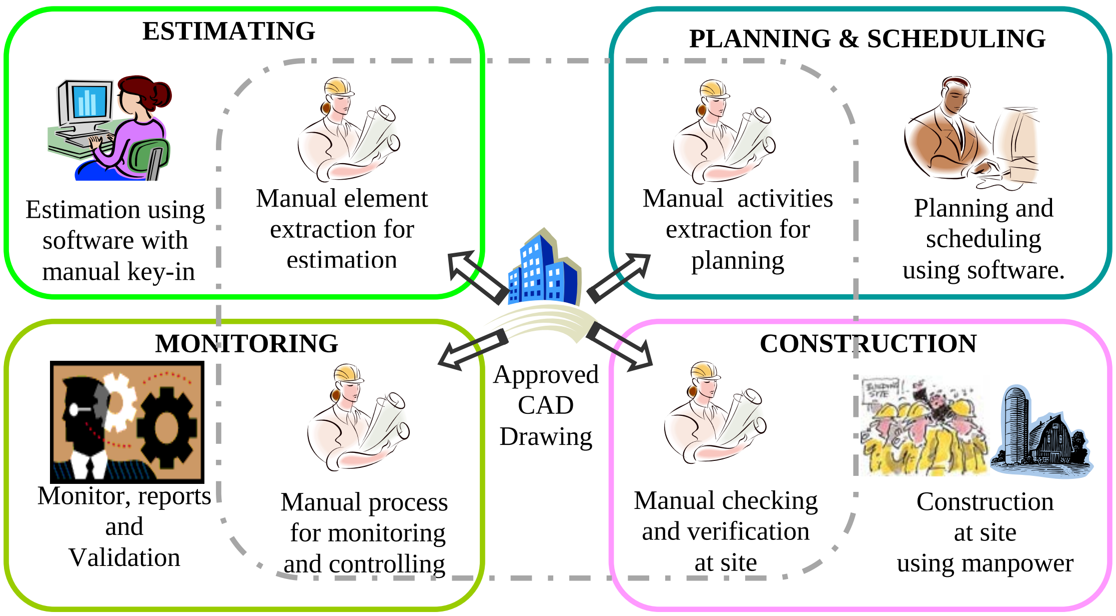

Figure 1: The Common Malaysian Construction Industry Practice

The industry still follows the practice of interpreting the drawing manually for their activities without an integrated automatic CAD elements reading from the drawing. Currently there is no system that can automatically identify the symbols of the drawing. This paper explains the conceptual model of the research framework that automatically read the drawing and identifies the element together with their details and eliminates the manual interference while processing the activities from the drawing details.

3. Requirement Analysis

Since the research is mainly objective with elements details extraction and recognition from the approved structural digital file, in order to proceed, the research is conducted with the construction drawings, drawing standards and formats, required software components, and the extraction methods to carry on are analyzed.

3.1 Construction Drawings

Drawings in construction are used in many ways at different stages in the life time of a building. A Computer Aided Drawing (CAD) is a system typically used to produce a lower level drawing following the vector coordinate system. Each has its own focus and level of detail. The detailed design requires coordination of the systems in the building; while accurate dimensioning and quantities are needed at the planning and construction level. The documentation of a construction project is not only in the form of drawings. Certain information is better expressed in the form of text, such as measurements, materials, and element details (Herzell, 1993). The vector drawings in construction are classified as 2D, 3D and 3D product models and represented as either architecture or structural plan of a building.

3.2 Drawing Standards

Beside the building details representation in the drawing, layering is a method to classify the information content of the drawing used by today's CAD systems. Layers are used to control consistency and allow different actors to edit certain layers from the core drawing and merge them back since nobody else can draw on those layers (Robert N., 2001). The naming conventions for every layer is typically defined by the drawing classification system, but not until with any standard codes for layers. The release of ISO 13567 provides a framework for structuring layer names but has only just been implemented in CAD systems and not practiced in the drawings generally in Malaysia (Bjork *et al*, 1997). In developed countries, international layer naming standards such as ISO STEP, IFC, and ISO 13567 are under practice and n order to standardize the automation processes, they specify their own standards. Malaysian Construction Industry Development Board (CIDB) has defined the layer naming standards called **Malaysian Standards CIDB Architecture Layer Standards** for layer and layer naming. This research follows the CIDB architecture layer standard in the 2D structural drawings.

3.3 Software Components

AutoCAD 2008 by Autodesk Inc with the COM and .OCX file gives the access to the .dwg drawing file format databases. This is mainly used by other programming software such as Microsoft Visual Basic, Visual Basic.NET to access the CAD database directly. The research programmed with this environment.

An alternative ObjectARX have been used to extract the AutoCAD drawing elements. They are provided with C++ libraries. Java 3D by Sun Microsystems also can be used to access the AutoCAD database.

3.4 Extraction Methods

Extraction from the drawing can be a sequence of processes trying to implement the knowledge of how those shapes are formed using primitive or solid or objects. Noack (2001) clearly identified that any drawing can be extractable in two ways namely:

Recognizing the elements by identification: This method is applicable more to paper based image identification. i.e. scanning the hard copy and recognize the elements.

Reading the elements by properties: Digital soft copy files can be readable through accessing their database and recognize by their property values.

Since our research objective to extract the elements from the digital drawing, the research follows this method. There is currently no system that can classify automatically all the primitives in a vector drawing of their appropriate shapes.

4. Process Model

Figure 2 depicts the research methodology of the research in IDEF0 format. Since the research is mainly objective with elements details extraction from the digital file, in order to proceed, the research conducted the analysis of extraction methods to carry on.

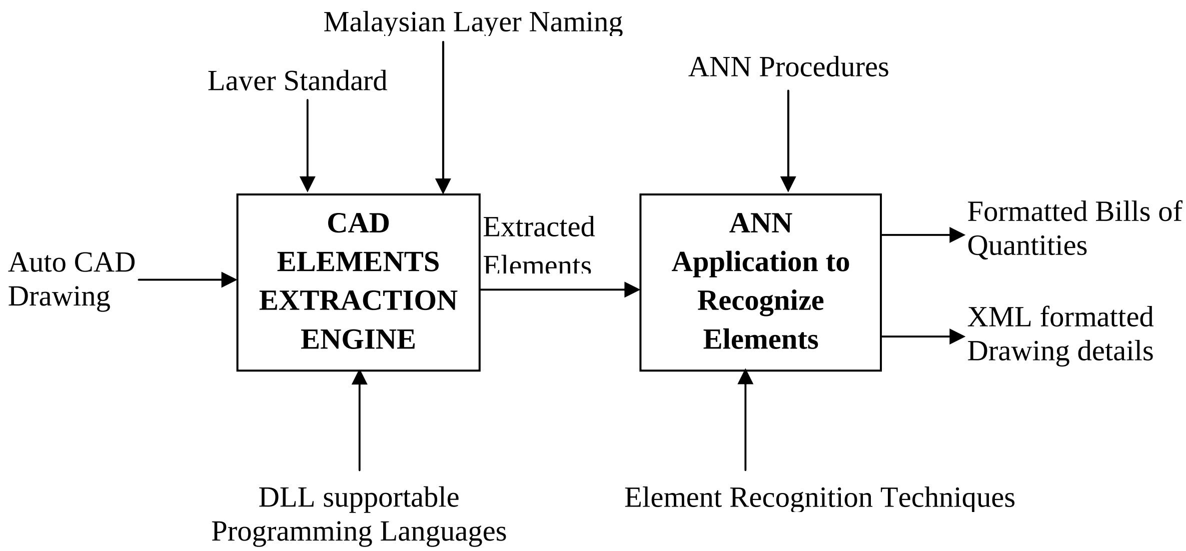

Figure 2: Extraction Algorithm Process in IDEF0 Format

4.1 The Prototype

The prototype has been developed with little user interaction. The main functionality of this prototype is to open the digital drawing file in CAD environment automatically and read the drawing elements automatically. The layer naming standard in the drawing eliminates the search time. Although layers can help to classify what a drawing represents, applying the Artificial Neural Network (ANN) on layer elements makes it is easy to recognize the element. But the problem still remains with identifying the rest of the geometry representing the elements. The ANN can be an element recognizer and together with the details of the relevant elements, the information can be stored in the temporary database. The overall prototype of the extraction process is given in Figure 2.

4.2 Extraction Process Flow

The extraction algorithm is a flow of processes used in the prototype to read, identify and interpret the elements from a conventional structural drawing. Some of the processes are done with different methods and techniques. In order to extract the elements, the AutoCAD, the front-end interface, and the C++ dynamic library as a pipeline between the interface and CAD are necessary at run time.

Since the element are drawn using a group of primitive symbols and through manual interpretation, it is identified as beam or column or other element. Using graph theory, the collective primitive symbols can be grouped into a single element and based the dimension; the element type can be automatically identifiable. This technique is named as shapes recognizer. Through shape recognizer, the element shape will be identified. Though texts surrounded that element, the research can get the dimensions and the reinforcement details. The process flow used in elements extraction can be outlined as follows:

1. Start.
2. Open CAD Application.
3. Read the CAD drawing.
4. Check for Layer name specifications and select a layer.
5. Select particular area of the drawing to get the details.
6. Read the selected area's 2D drawing details into temporary database.
7. Apply ANN technique to Identify the elements by their primitive shapes
8. Classify the identified shapes using properties.
9. Define the techniques to recognize the element.
10. Interpret the shapes & details of the recognize element into meaning format.
11. Apply the BQ Specification Format on the formatted data.
12. Convert the elements details into standard XML format.
13. Export the XML data.
14. Close the CAD application.
15. End.

5. Results

Currently, the extraction algorithm has developed to interpret the beam element and implemented on small sample drawing. They had continuously improved and updated until it can interpret the beam with the combination of drawing element properties and the text details. The sample source drawing of a beam is given in Figure 3.

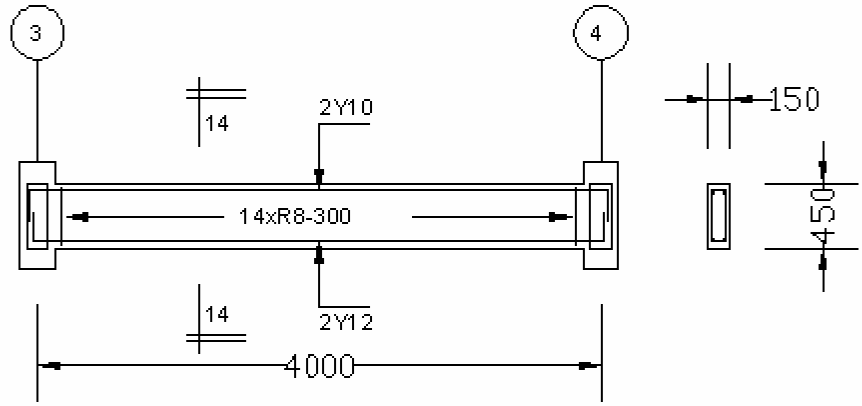

Figure 3: Sample Source 2D Drawing

The extraction algorithm reads all the primitive shapes such as lines, circles, etc with their properties. The layer naming standard implementation in the drawing reduces the time for shape recognizer to identify the element type. Through combining the drawing element recognized by the shape recognizer and the text details, the algorithm can define dimensions and reinforcement details. The output of the recognized beam in terms of its text details is given as Figure 4.

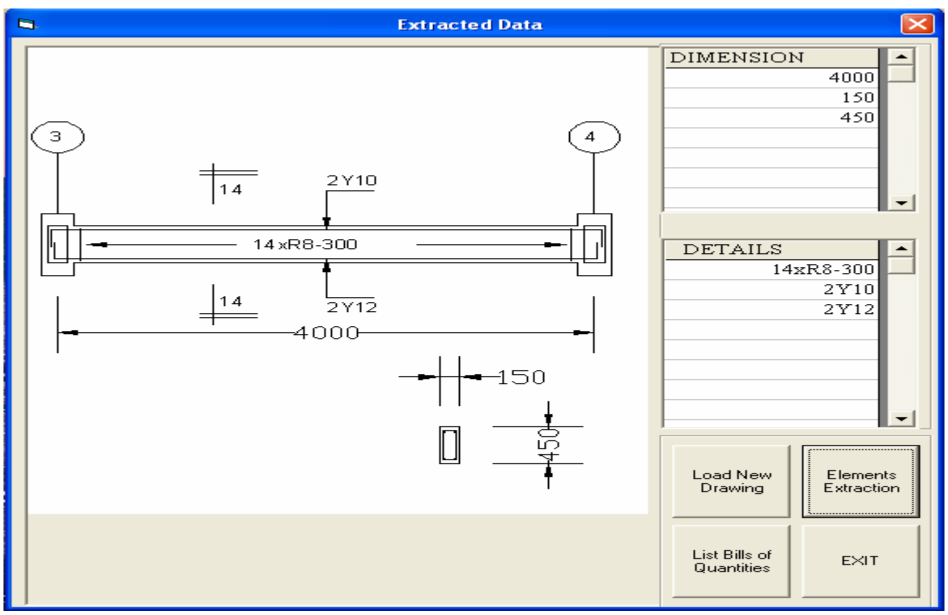

Figure 4: Output Screen Shot

6. Conclusion

Currently, the extraction algorithm has developed to read the elements and the text details. The shape recognizer is developed and under verification. Currently the algorithm is limited to read the drawing drawn based on layer specifications. At this time of this paper written, the algorithm is able to identify the beam elements of any structure drawing.

7. Acknowledgement

The “An automated project work activities extraction from the approved structural drawing” is funded by Construction Industry Development Board (CIDB), Malaysia and this support is gratefully acknowledged. The academic team consists of the Universiti Teknologi Malaysia (UTM). We are grateful to those who involved with this project.

8. References

Bjork, B-C., Lownert, K., and Kivinieni,A. (1997). “The proposed international standard for structuring in computer aided building design”, *Electronic Journal of Information Technology in Construction (http://itcon.org)*, Vol.2, No. 2.

Chitrakala M., Adbul Hakim Mohammed., and Muhd Zaimi Abd Majid (2008). “An automated quantities estimating framework from the structural 2D drawing”, *The 6th Regional Symposium on Infrastructure Development (RSID6) in Civil Engineering*, Kasetsart University, Thailand.

Herzell, T. (2000). *Documentation of Construction Projects*. Bygghandlingar 90, Stockholm, 1993.

Marsh, L. and Flangan, R., (2000). “Measuring the cost and benefits of information technology in construction”, *Journal of Engineering, Construction and Architecture Management*; 7(4): pp. 423-435.

Oglesby, C., Parker, H. and Howell, G. (1989). *Productivity Improvement in Construction*, First Edition, McGraw-Hill.

Noack R. (2001). *Converting CAD Drawings to Product Models*. Master Thesis, Royal Institute of Technology, ISBN 91-7283-067-0.

Soetanto,R., Dainty, A.R.J., Glass, J., Price, A.D.F., and Thorpe, A. (2005). “Improving the utility and value of CAD software for decision-making and design of structural frames”, *Proceeding of Computing in Civil Engineering (CICE) ASCE*, Cancun, Mexico, July 12-15.