MVA'94 IAPR Workshop on Machine Vision Applications Dec. 13-15, 1994, Kawasaki

UNDERSTANDING ELECTRONIC CIRCUIT DIAGRAMS AND LINKING WITH ORCAD

Shi Zesheng, Jiang Cunhong, Li Sifang, Yang Jing

Department of Electronic Engineering University of Science and Technology of China Hefei, Anhui, China(230026)

① ABSTRACT

This paper presents a rapid detecting method to tell the external contours from internal ones, and adopt an aid measure of adding " black points". These two methods result in obtaining the high reliability of pattern separation. For the recognizing hand-written symbols, we propose an extracting feature method projecting four sidehands and bring it in the decision tree. It can lead to increasing the recognition rate and speed obviously. After the system understands the topologic relationship, it can still link with ORCAD which is an international current CAD software.

Keywords: Automatic input system, Pattern separation, Sybol recognition, Decision tree, Vector encoding.

I. INTRODUCTION

Automatic input technique of graphics is an intersecting technique of several courses. It involves image processing, patter recognition, artificial intelligence and compputer graphics. Because this technique will have its wide application prospect in the fields of CAD, OA and GIS etc. So people are Paying more and more attention to this research.

Input method in the current CAD systems are more backward than output method. Data entry by human operators is time-consuming and error-prone. An idea solution is to let the computer see and understand the diagrams directly. From late 80's to early 90's, there were some reports on automatic input systems for electronic circuit diagrams from Japan, U.S.A, Taiwan, hungary, etc.

Recently we have developed an automatic input system of electronic circuit diagrams, which can be used not only to recognize hand-written symbols, characters and connecting line, but also to understand the topologic relationship of the whole diagram and to link with ORCAD (an international current CAD software). Now this system has been implemented on an AST 486 / 25 microcomputer.

This paper describes the structure and the features of the system briefly. We shall discuss following three contents:

1) The pattern separation of electonic circuit diagrams;

2) The recognition methods of symbols based on the raster data in a circumscribed rectangle.

3) The technique of linking with ORCAD.

Finally the experiment results and conclusions are given.

II. THE TECHNIQUES SEPARATING THREE PRIMITIVE ELEMENT

In general, engineering drawings consist of three parts: characters, symbols and connecting lines. To understand engineering drawings, we have to separate it into three independent parts. Fig.1 is the flow diagram for this automatic input system, It is obvious that the separations of three primitive elements are implemented by two steps: firstly, separating characters; Seconedly, separating symbols.

1) Character separation

The characters have two feature: (1) independence and connectivity; (2) small size. We use the contour following technique of Freeman

①This work was supported in part by the 863 plan and in part by the National Nature Science Fundation of China.

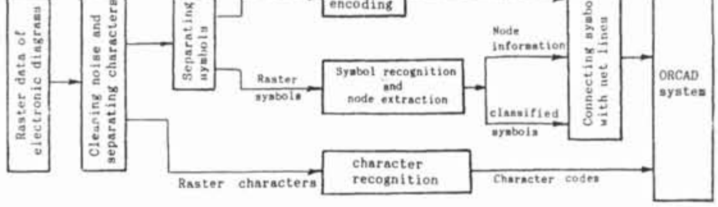

Fig.1 Flow diagram for the automatic input system

Chain−Codes to follow the whole diagram and calculate the width and height of each closure contour according to Eq.(1).

$$\begin{aligned} \text{width} &= \Delta x = \max\ x_j - \min\ x_j \\ \text{height} &= \Delta y = \max\ y_j - \min\ y_j \end{aligned} \qquad (1)$$

In the light of first feature for the character, an independent and connecting character must possess an external closure contour. sometimes, besides the external closure contour, it still possesses one or two internal closure contours(see Fig.2). But for character separation, we only need the external one. So we propose a rapid detecting method to tell the external closure contours from the internal ones.

For the external closure contours,the following direction of Freeman Chain−codes is counter clockwise, and for the internal ones the following direction is just the opposite—clockwise.(see Fig.2)

Fig.2 The external and internal contours of a hand−written character.

In this way, the decision of internal and external contours depends directly on the following direction of Freeman Chain−Codes. After analyzing carefully, we found out that the Chain−Codes for the external closure contours at the x_{max} or y_{max} possess their inherent rule. Fig.3 shows the variational rule of Chain−Codes at the X_{max} for the external contours. Here s(i) expresses the Chain−codes arriving at x_{max}; s(i+1) expresses the Chain−Codes leaving x_{max}.

The closure contours satisfying variationed rule of Fig.3 are external closure contours, otherwise,are internal ones. For the all external closure contours, if Δx and Δy calculated according to Eq.(1) satisfy Eq.(2) then the circumscribed rectangles composed of Δx and Δy are characters separated.

$$\begin{cases} \Delta x_{nmax} < \Delta x < \Delta x_{max} \\ \Delta y_{nmax} < \Delta y < \Delta y_{max} \end{cases} \qquad (2)$$

here ΔX_{max} and Δy_{max} are the maximum width and height of a character respectively; they are the selected thresholds according to the second feature of characters, ΔX_{max} and ΔY_{max} are the maximum width and height of the noise respectively.

2) Symbol separation

After characters are separated, the remain is the image consisted of symbols and connecting lines. The goal of symbol separation is to take them apart into two independent parts. Whole symbols can be divided into two regiments (circular symbols and non−circular symbols). They are called circular symbols, for they have internal closure contour in a symbol, we can use approach in Fig.3 to extract them.

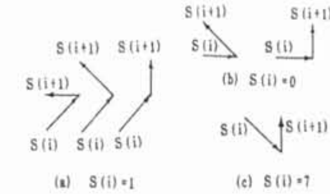

Fig.3 The variational rule of Chain−Codes for the external contours

For the non−circular symbols, it is more difficult to separate them from connecting lines, because the demarcation between non−circular symbols and connecting lines is not clear. So we adopt an aid measure, adding " black points" on the terminals of each non−circular symbol and use a vector encoding method based on the edge detection along two−orthogonal LAG (Line adjacency−graph)[2] to extract "black points". Once the "black points" are found, the demarcation of non−circular are fond. In order to separate non−circular symbols, firstly, " black points" found should be deleted from raster image, at this moment, non−circular symbols are isolated just like the characters isolated. Secondly, the approach separatinag character can be used to separate non−circular symbols.

As to the circular symbols, their common feature is to have internal circulars. But some internal circulars are produced by several lines intersected each other. In order to remove it, we may test four corners of a circumscribed rectangle, if they have "corner lines", they must not belong to the symbols. Moreover, circular symbols not only have one internal circular, but also have two or three ones sometimes. Therefore the circular symbols may be divided three classes: Mon−circle, Bi−circle and Tri−circle(see Fig.4). For the

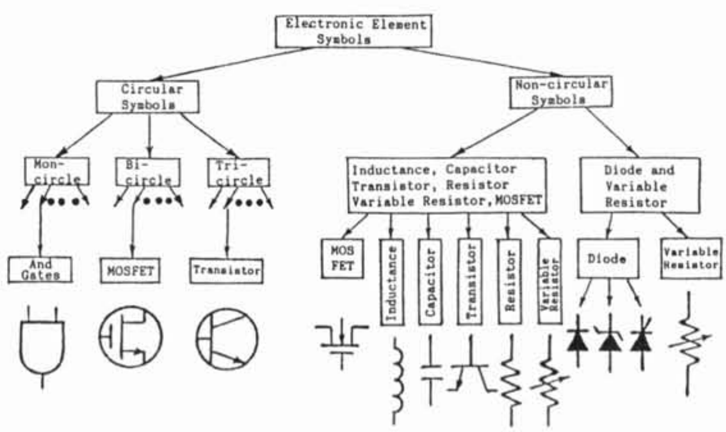

Fig.4 The simple graph of decision tree

symbols of Bi–circle and Tricircle, as circumscribed rectangleof each internal circle must overlop partially. We can utilize this feature to merge them into a complete symbol.

III. SYMBOL RECOGNITION USING DECISION TREE

After characters and symbols are separated, the electronic circuit diagram has been divided into three independent parts. Follow–on is to recognize them separately. This paper only discusses symbol recognition.

(1) The extracting feature method projecting four sidebands.

The symbols separated appear in form of the raster image of its circumscribed rectangle,from which geometrical features are extracted.

Except the four edges of circumscribed rectangle themselves, each raster image of symbol circumscribed rectangle can be represented as matrix Eq.(3)

$$F=\begin{bmatrix} f_{11} & f_{12} & \cdot & \cdot & \cdot & f_{1N} \\ f_{21} & f_{22} & \cdot & \cdot & \cdot & f_{2N} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ f_{M1} & f_{M2} & \cdot & \cdot & \cdot & f_{MN} \end{bmatrix} = [f_{ij}] \qquad (3)$$

here i = 1,2,3,...,M; j = 1,2,3,...,N;

and f_{ij} is the discrete value of image pels for the ith line and the jth column.

Here the raster data are binary images:

$$f_{ij}=\begin{cases} 1 : \text{(black)} \\ 0 : \text{(white)} \end{cases}$$

We found that contour features of various symbols mostly centralize all around their circumscribed rectangle. In order to extract their features, we propose an extracting feature method projecting on four sidebands which are formed by contracing four sides of circumscribed rectangle 6 rels. While four sidebands are projecting on the X and Y coodinated axes, four projected maximums are got and expressed as xm1, xm2, ym1 and ym2:

$$\begin{cases} xm1 = \max_j \{\sum_{i=1}^{M} fij \quad j=1,2,\cdots,6\} \\ xm2 = \max_j \{\sum_{i=1}^{M} fij \quad j=N-5,N-4,\cdots,N\} \\ ym1 = \max_i \{\sum_{j=1}^{N} fij \quad i=1,2,\cdots,6\} \\ ym1 = \max_i \{\sum_{j=1}^{N} fij \quad i=M-5,M-4,\cdots,M\} \end{cases} \qquad (4)$$

They are four important feature parameters which can be used not only to recognize most symbols, but also to datemine symbols directions.

2) Using decision tree to recognize electronic element symbols

After being separated electronic element symbols will be divided into two regiments (circular symbols and non–circular symbols). Each regiment has its sub–tree. Fig.4 is the simple graph of decision tree made of two subtrees. The whole decision procedure is sequentied from top to bottom. Every time a node arrives, a decision is made according to the condition, which is the result calculating parameters. This decision procedure will go on until arriving at leaves.

Fig.5 show the decision process of inductance and transistor. Suppose decision tree has already arrived at node A of Fig.5, which consists of inductance and transister. We give a horizontal line through the middle of a symbol and test the numbers of intersecting points, if they are more than 2, it will arrive at leave of Fig.5 (inductance); or it will belong to transistor (node B). In order to recognize transistor types further, the arrow position should be determined. Here rthe rectangle region given by dotted lines in Fig.5 are quartered, then the black pels of each quarter are calculated and the region of most black pels is the arrow region, thus PNP or NPN are indetified.

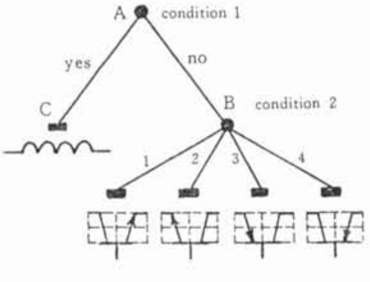

Fig.5 Decision process of inductance and transistor. (condition 1: Are the numbers of intersecting points more than 2? condition 2: Where is the arrow region 1,2,3 or 4?)

IV. LINKING WITH ORCAD

To understand electronic circuit diagrams, besides to recognize their symbols, the connecting lines have to be processed and the connecting relationship of symbols and lines have to be made clear.

The processing contents on the connecting line include vector encoding, processing of intersecting points, joining two lines together and connecting lines with symbols, etc.

Through the processing mentioned above, the topologic relationship of whole electronic circuit diagram has already be understood. In order to link with ORCAD from the bottom, we have to solve following two key questions:

1) The file format of binary principle graph must be dissected. Its contents consist of file head, file tail and represented formats of every element symbol, etc.

2) We know the symbol size of original input

graph is variable, but the symbol size of ORCAD is fixed, both are no identieal. So this contradiction have to be solved, otherwise, the principle file established can not be displayed or printed well.

Now, two problems mentioned above have been solved throughly. System can form the format needed for the ORCAD and link with the ORCAD from bottom. Finally, the principle file of ORCAD is establishs. Fig.6 is the input original hand–written diagram.

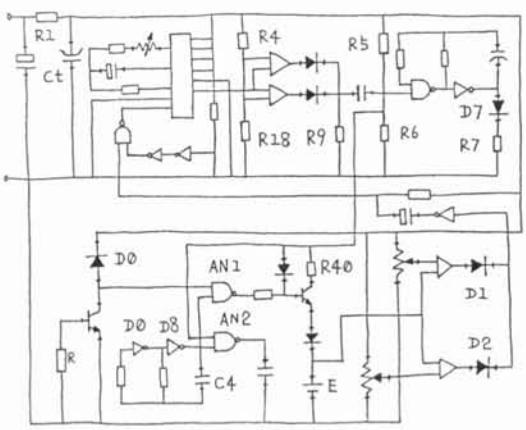

Fig.6 The input original hand–written diagram

V. PERFORMANCE AND CONCLUSION

We have implemented an automatic input system of electronic circuit diagram on an AST 486 / 25 microcomputer. It has following functions:

(1) This system can recognize 62 hand–written characters and 47 hand–written element symbols, each symbol is allowed to varify at four directions (0,90,180,270). Thus it can recognize about 170 patterns of the element symbols;

(2) The recognition rate of characters and symbols are over 95%, 97% recpectively.

(3) The recognition speed of characters and symbols is over 10 symbols (characters) / sec.;

(4) This system can understand the topotogic relationship of the whole graph and link with ORCAD from bottom. Finally the principle graph file of ORCAD is established.

The featunes of this paper are:

(1) In the pattern separation, this paper proposes a rapid detecting method to tell the external contours from internal ones, and adopts an aid measure, adding "block points" on the terminals of each non–circular symbols. These methods lead to a high separating rate near 100%.

(2) The recognition process of symbols and characters is based on the raster data. Thus we can take advantage of primitive information as far as possible.

(3) For the symbol recognition, this paper proposes an extracting feature method projecting four sidebands, and uses the decision tree to recognize circular and non–circular symbols. Therefore it has obvious advantages such as high speed, less memory and easily being realized in microcomputers.

(4) Beacuse this system can link with ORCAD and establish the principle graph file of ORCAD, It means that besides the layout and the line routing of printed circuit board, this automatic input system can also support the circuit simulation.

It is important that the system can utilize the editing functions of ORCAD to correct the errors in the recognition processing.

VI. REFENCES

[1] S.K.Chang, "Principles of Pictorial Information System Design" , Prentice–Hall, inc. 1989, pp.18–20.

[2] Shi Zesheng, "A New Method of Vector Encoding Based on the Two–Orthogonal LAG", Acta Electronics Sinica, (Published by the Chinese Institure of Electronics, Beijing), Vol.22, No.1, Jan. 1994, pp.101–104

[3] Y.Fnkada, "A Primary Algorithm for the Understanding of Logic Circuit Diagrams" , Pattern Recognition, Vol 17 No.1, 1984, pp.125–134.

[4] horst Bunke, "Experience with Several Methods for the Analysis of Schematic Diagrams", PROC. of the6–th conf. on Pattern Recognition, Oct. 19–22 / 1982, pp.710–712.

[5] Zesheng Shi, Jianhua Ding, " A New Input Method for CAD —– Automatical Read–in Technique of Paper Drawings" , CHINA 1991 INTERNATIONAL CONF. ON CIRCUIT AND SYSTEMS, CONF.PROC. Vol, 1 of 2, 16–17 June, 1991 pp.431–433.