

# LSI 芯片图象的双子图同构匹配法

王吉仁 李介谷  
(华东师范大学) (上海交通大学)

## 摘 要

本文提出一种用于 LSI 芯片图象的双子图同构匹配方法。该方法是在芯片分割图象的区域邻接图上搜索同构匹配结点间的对应关系。

同构匹配算法利用启发信息提高搜索的效率。算法根据专门建立的区域相似性度量和环境串匹配值这些判据选取同构匹配起始结点。

文中将详细说明双子图同构匹配法的原理,并提供匹配结果。本算法已在实际中采用。

## 一、引 言

用图象处理和模式识别技术对 LSI 芯片解剖分析时,芯片图象经数字化后输入计算机。

在数字化采样中,为了不丢失芯片图象的信息,需要保证一定的采样精度,整枚芯片被表示成很大的象素矩阵。受现有采样设备分辨力的限制,整枚芯片图象必须分解成若干幅子图后再输入。

整枚芯片分解成若干幅子图后,为获得整幅芯片版图就出现了重新拼成整图的问题。

如果各幅子图的上下左右之间恰好完全衔接,则拼接只是简单的接续。此时要求安放芯片的平台具有很高的进给精度。高精度的机械进给装置价格通常比较昂贵。

采样时使各幅子图的上下左右之间都包含一定的搭接部分,可以降低对机械进给装置的精度要求。

用图象处理和模式识别技术代替价格昂贵的高精度机械进给装置,这在经济上是合宜的。为去除各幅子图之间的搭接重叠部分,就出现确定重叠部位即芯片图象的匹配问题。

由于两幅交叠图象的待匹配部分是通过两次采样得到的,它们在显微镜的视场中分别处在不同的位置上,因光照和光学畸变等原因,使应该匹配的部分灰度的分布不尽相同,分割处理后也难以保证分割图象在细节上完全相同。这给灰度图象的相关匹配法<sup>[1]</sup>和敏感于图象中细小变化的符号匹配法<sup>[2]</sup>带来了困难。

为此应寻求另外一种匹配方法,它应建立在两幅图的结构相同基础上。这就是本文

提出的图同构匹配方法。

## 二、区域邻接图与图同构

图同构匹配法是在分割图象的区域邻接图上搜索同构匹配结点间的对应关系。

由灰度图象向分割图象的转化,其间经过一个图象分割处理的过程,这不属于本文的讨论范围,因而这里从略。本文的讨论都是以分割图象为起点。

从分割图象转换成其区域邻接图的过程如下:把分割图象中的区域映射成区域邻接图中的结点,此结点可以赋以原区域的各种属性;如果分割图象中两区域相邻,则把这种邻接关系映射成连接区域邻接图中区域对应结点的边,此边作为区域邻接图边集中的元素。这里相邻是指区域之间有公共的外侧边界。外侧边界指并非由两包含区域所形成的边界。

图 1 中示出分割图象的一个局部幅面及与之相应的区域邻接图。

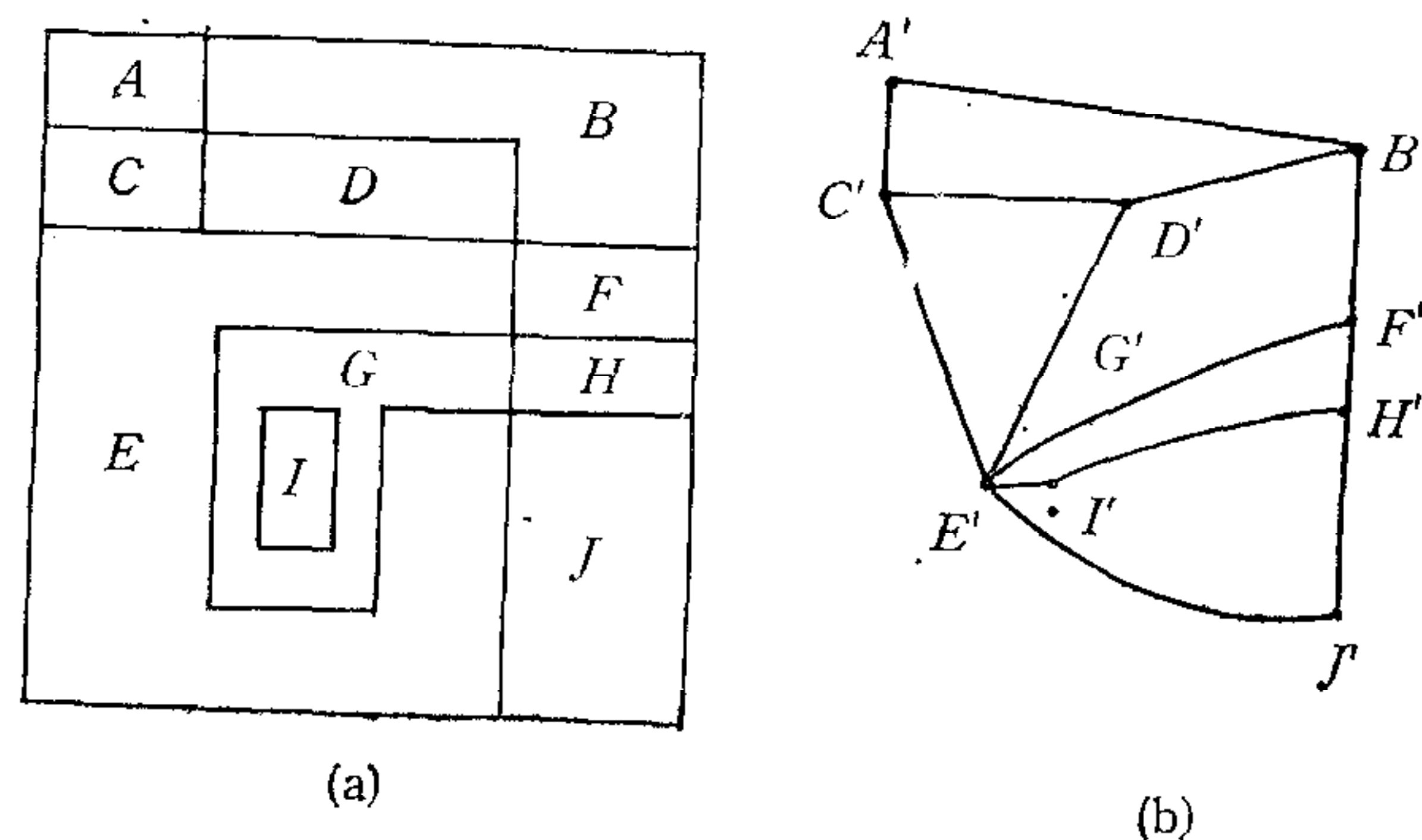


图 1 分割图象(a)及与之相应区域邻接图(b)

图(a)是一幅分割图象,它由区域  $A, B, C, D, E, F, G, H, I, J$  所组成。图 1(b) 是与之相应的区域邻接图,它由结点  $A', B', C', D', E', F', G', H', I', J$  所组成。如分割图象中区域  $A, B$  间含有公共的外侧边界,则在区域邻接图中有连接结点  $A'$  和  $B'$  的边;区域  $G$  与  $I$  无公共的外侧边界,则区域邻接图中结点  $G'$  和  $I'$  不存在边。区域邻接图中  $I'$  是孤立结点,则说明  $I'$  的对应区域  $I$  是一个被其它区域所包含的区域。

区域邻接图以邻接表的形式存放。

两个区域邻接图  $RAG_1 = (V_1, E_1)$  与  $RAG_2 = (V_2, E_2)$  同构时应满足下述条件:

1)  $F(v_i) = F(v_j)$ , 式中  $F(v)$  是指结点  $v$  的特征集,  $v_i \in V_1$  和  $v_j \in V_2$  为同构时对应结点。

2)  $|v_i| = |v_j|$  和  $V_1^\# = V_2^\#$ , 式中  $|v|$  指结点  $v$  的度数,  $V^\#$  指图结点总数。

3)  $E_1^\# = E_2^\#$ , 式中  $E^\#$  指图中边的总数。

同构匹配就是按照这些要求来建立其判据的。

整个匹配过程由三个主要步骤所组成,它们是:多边形区域数据的裁剪和交换,结点特征的提取及区域的相似性度量和双子图的同构匹配。下面分别对它们加以说明。



### 三、多边形区域数据的裁剪和变换

LSI 芯片图象经分割形成的区域一般都是以水平直线、垂直直线、与水平线成  $45^\circ$  和  $135^\circ$  直线为边界的多边形区域。其中每个多边形区域被表征为如图 2 所示的数据形式。

多边形 边数 $n$	顶点1 状态	顶点1 坐标	...	顶点 $n$ 状态	顶点 $n$ 坐标	多边形 质地
---------------	-----------	-----------	-----	--------------	--------------	-----------

图 2 表征一个多边形区域的数据形式

图 2 中顶点的状态只有两种,其一是处在图象的边缘上,表示这一顶点可能是由图象割裂而形成的;其二是处在图象的内部。这里多边形区域质地是指芯片图象中的层叠情况,例如铝和多晶叠层,多晶和扩散的叠层等。质地信息的获取是在本文以外的芯片图象识别过程中完成的。该过程根据区域灰度、形状、邻接区域的情况等信息识别出各多边形区域的层叠情况。

通常一幅芯片的分割图象由数百个多边形区域所组成,由此将构成一个庞大的区域邻接图。但是从衔接匹配的任务来说,应该进行匹配的只是前一幅图的末尾部分与后一幅图的开始部分,称之为重叠范围,它的大小在采样中是可以加以控制的。数据的裁剪是把可能的重叠范围从整幅图象中裁取出来,这样既可以缩小区域邻接图的尺寸,又可以提高同构匹配搜索的效率。据[3]可实现多边形区域数据的裁剪。

为构成正确的区域邻接图,如图 2 所示的多边形区域数据还须加以变换,这一步骤的必要性可由图 3 来加以说明。

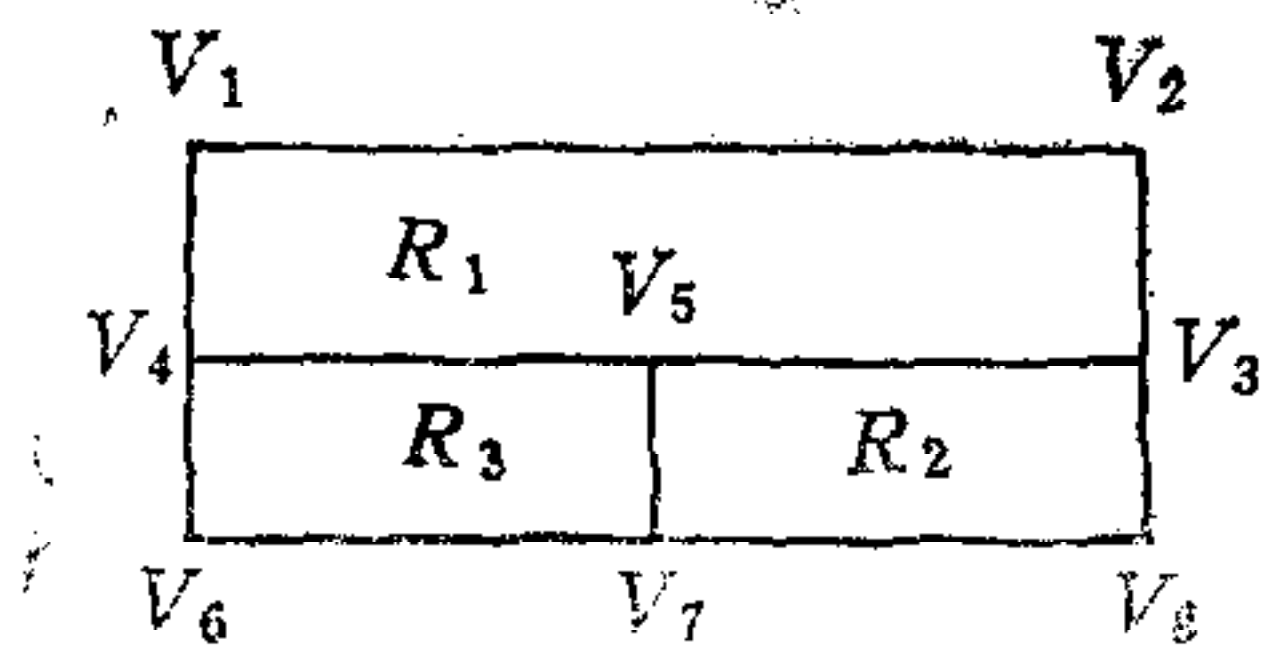


图 3 多边形区域数据的变换

在图 3 中含有三个区域,按照用各自顶点表达时,它们是:  $R_1(V_1V_2V_3V_4)$ 、 $R_2(V_5V_3V_8V_7)$  和  $R_3(V_4V_5V_7V_6)$ 。区域邻接图是反映区域邻接关系的图。构造区域邻接图时,判别两区域是否邻接是根据两区域是否存在公共的外侧边界;对于多边形区域而言,则判别是否存在公共边。对于上述三个区域来说,因为  $R_1$  中并没有包含顶点  $V_5$ ,因而没有反映出  $R_1$  和  $R_2$ 、 $R_1$  和  $R_3$  之间具有公共边。因而多边形区域数据变换  $T$  对于上述三个区域的例子而言,它所起的作用是:

$$T: R_1(V_1V_2V_3V_4) \longrightarrow R_1(V_1V_2V_3V_5V_4).$$

多边形区域数据变换算法分标记和跟踪这两个步骤。第一步是构造标记平面,在其上标出所有多边形的顶点;第二步是沿着各多边形的每条边  $\overline{v_i v_j}$  在标记平面上跟踪。在  $\overline{v_i v_j}$  的路径内除  $v_i$ 、 $v_j$  本身以外,如出现其它标记点,则该标记点作为夹角等于  $180^\circ$  的新顶点引入。

多边形区域数据的裁剪和变换为构造正确的区域邻接图做好了准备。

## 四、结点特征提取与区域相似性度量

区域邻接图中结点表示分割图象中某区域, 可以把区域有关属性作为特征赋给相应结点。

这里把层叠质地、区域多边形边数、多边形面积、多边形状态和多边形最大最小中心作为结点的特征。

层叠质地前面已经介绍了, 多边形边数在此不必多加叙述, 现在给出其它几个特征的提取方法。

多边形  $P_i$  的最大最小中心是这样提取的: 设多边形  $P_i$  的各顶点为  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , 令  $\min X(P_i) = \min\{x_1, x_2, \dots, x_n\}$ ,  $\max X(P_i) = \max\{x_1, x_2, \dots, x_n\}$ ,  $\min Y(P_i) = \min\{y_1, y_2, \dots, y_n\}$ ,  $\max Y(P_i) = \max\{y_1, y_2, \dots, y_n\}$ , 则多边形  $P_i$  的最大最小中心定义为:

$$(X_{\min \max}(P_i), Y_{\min \max}(P_i)),$$

式中  $X_{\min \max}(P_i) = (\min X(P_i) + \max X(P_i))/2$ ,

$$Y_{\min \max}(P_i) = (\min Y(P_i) + \max Y(P_i))/2.$$

该多边形  $P_i$  面积(证明从略)为:

$$A_{P_i} = \left| \frac{1}{2} \begin{vmatrix} X_1 & X_2 \\ Y_1 & Y_2 \end{vmatrix} + \frac{1}{2} \begin{vmatrix} X_2 & X_3 \\ Y_2 & Y_3 \end{vmatrix} + \dots + \frac{1}{2} \begin{vmatrix} X_n & X_1 \\ Y_n & Y_1 \end{vmatrix} \right|.$$

多边形区域状态有两种, 其一是处在图象内部, 其二是处在图象边缘。如果多边形顶点都处在图象的内部, 则此多边形区域状态便处在图象内部; 否则此多边形状态便处在图象边缘。

定义多边形区域的相似性度量是为了用数值来表示两区域的相似程度。反映区域的相似性度量的相似函数定义为:

$$R_{ij} = - \sum_{k=1}^n |F_{ik} - F_{jk}| W_k S_k,$$

式中  $i, j$  是被比较的两个区域的序号,  $n$  为有效特征数,  $F_{ik}$  和  $F_{jk}$  是区域  $i$  和  $j$  第  $k$  个特征值, 负号表示在良好匹配时取较大值。  $W_k$  是对取值范围不同的各特征进行规范化的加权因子,  $S_k$  是对在不同情况下各特征所起作用赋予的强度因子。例如当区域状态为边缘时, 它可能是从某一区域割裂出来的一部分, 此时面积特征是不可信的。

对于 CMOS 集成电路芯片来说  $F$  是质地、多边形区域边数、多边形区域面积、多边形区域状态等集合。

## 五、双子图的同构匹配

一幅图的一部分与另一幅图的一部分同构匹配问题是一个双子图的同构匹配问题。

CMOS 芯片分割图象的区域邻接图的双子图同构匹配法是建立在以下三个假设的基



基础上:

- 1) 各多边形区域的质地信息基本可信;
- 2) 待匹配两图的重叠部分至少含有一对不处于图象边界(即未被割裂)的匹配对应区域;
- 3) 该对应区域满足同构匹配起始点两个选择条件.

对于两幅区域邻接图的双子图的同构匹配,为了节省同构匹配的搜索时间,这里不采用穷举法而采用启发式的方法,即利用启发信息求得同构匹配的起始结点,然后沿此起始结点作广度搜索.

同构匹配的起始结点根据区域的局部特性和全局特性来选取.区域的局部特性指多边形区域质地、边数、面积和状态等,通过区域相似性函数比较两区域的局部特性.区域的全局特性指区域的邻接区域情况,通过区域环境串的匹配值来比较两区域的全局特性.

为了说明区域环境串的匹配值,现给出区域间夹角的定义.

区域  $P_i$  和  $P_j$  之间的夹角即为它们各自最大最小中心的连线与水平线的夹角.

设区域  $P_i$  有  $n$  个邻接区域,按它们与区域  $P_i$  的夹角大小排序为:  $P_{j_1}, P_{j_2}, \dots, P_{j_n}$ , 按此次序对它们的特征集  $F$  进行排列,就可以组成区域  $P_i$  的区域环境串  $L_{P_i}$ , 它是:

$$L_{P_i} = \{F(P_{j_1}), F(P_{j_2}), \dots, F(P_{j_n})\}.$$

区域环境的匹配与否取决于区域环境串的匹配情况. 设另有一待匹配的区域  $P'_i$ , 其区域环境串为:

$$L_{P'_i} = \{F(P'_{j_1}), F(P'_{j_2}), \dots, F(P'_{j_m})\},$$

且  $1 \leq m \leq n$ , 则区域环境串的匹配值为:

$$M = \sum_{i=1}^n D_i.$$

式中  $D_i$  为决策函数,且有:

$$D_i = \begin{cases} \frac{1}{n}, & F(P_{j_i}) = F(P'_{j_i}); \\ 0, & F(P_{j_i}) \neq F(P'_{j_i}). \end{cases}$$

因而只有当  $M = n \times \frac{1}{n} = 1$  时,才意味着两个环境串的匹配.

因此,同构匹配搜索起始结点的两个选择条件为:

- 1)  $R_{ij} \geq T$ ,  $T$  为某一阈值;
- 2)  $M = 1$ .

由于区域处在边缘时,强度因子  $S_k$  变弱,则前一条件说明对应同构起始结点所代表区域必须相似且不处在图象边缘,而后一条件则说明对应同构起始结点所代表区域所处环境即邻接区域情况必须相同.

求得同构匹配搜索的起始结点以后,区域邻接图的双子图同构结点对应关系的建立过程实际上是沿此起始结点在区域邻接图上作广度搜索的过程.在搜索时可以把结点特

征作为启发信息。

在广度搜索过程中对作为同构起始结点的结点进行标识。如果建立的对应结点满足选择同构起始点的两个条件且未被标识,则该结点作为新的同构起始点,重复上述广度搜索过程。上述扩展子图结构这一搜索过程一直进行,直到对应结点中再也没有未标识结点满足选择同构起始点的两条件为止。

有时待匹配图某一结点和另一待匹配图中,多个结点之间满足同构起始点的选择条件,这时说明待匹配图中存在较多的相似结构。此时选择较大的同构子图,即按同构关系建立较多对应结点的结果作为所求的结果。

在得到两区域邻接图结点对应关系即分割图象中区域的对应关系时,最后应求得匹配时两图间的相对位移。该相对位移值可用最小二乘误差方法进行估计。设应该求得的相对位移值为  $(\Delta x, \Delta y)$ , 如以  $(x_i, y_i)$  和  $(x'_i, y'_i)$  分别代表已匹配且不处在图象边缘上的第  $i$  个对应区域  $(P_i, P'_i)$  的最大最小中心, 这样的区域有  $n$  个, 则因为

$$\delta(\Delta x) = \sum_{i=1}^n (x_i + \Delta x - \Delta x'_i)^2,$$

以 
$$\frac{\partial \delta(\Delta x)}{\partial \Delta x} = 0,$$

可得

$$\Delta x = \frac{1}{n} \sum_{i=1}^n (x'_i - x_i).$$

同理有

$$\Delta y = \frac{1}{n} \sum_{i=1}^n (y'_i - y_i).$$

## 六、实验结果

图 4 和图 5 是两幅待匹配的 CMOS 芯片的分割图象, 图上各区域都给以相应的符号标记。表 1 为图 4 和图 5 匹配时区域间的对应关系, 这是用前面所介绍的方法得到的。匹配时的相对位移值为:  $\Delta x = -73, \Delta y = 3$ 。

表 1 图 4 和图 5 匹配区域对应关系

(5,2)	(+,')	(4,1)	(6,3)	(A,>)	(",")
(*,&)	(3,0)	(@,=)	(7,4)	(B,?)	(J,I)
(,;)	(,%)	(2,1)	(?,<)	(I,H)	(8,5)
(C,@)	(K,J)	(R,Q)	(-,)	((,\$)	(>;)
(H,G)	(Q,P)	(D,A)	(L,K)	(S,R)	(T,S)
(X,U)	(M,L)	(Y,V)	(Z,W)	(W,T)	([,X)

本文所介绍的方法实际上不仅适用于集成电路芯片图象的匹配, 因为任何一种形状的对象都可以用一些多边形来加以近似, 因而本方法可以推广到其他应用领域中去。

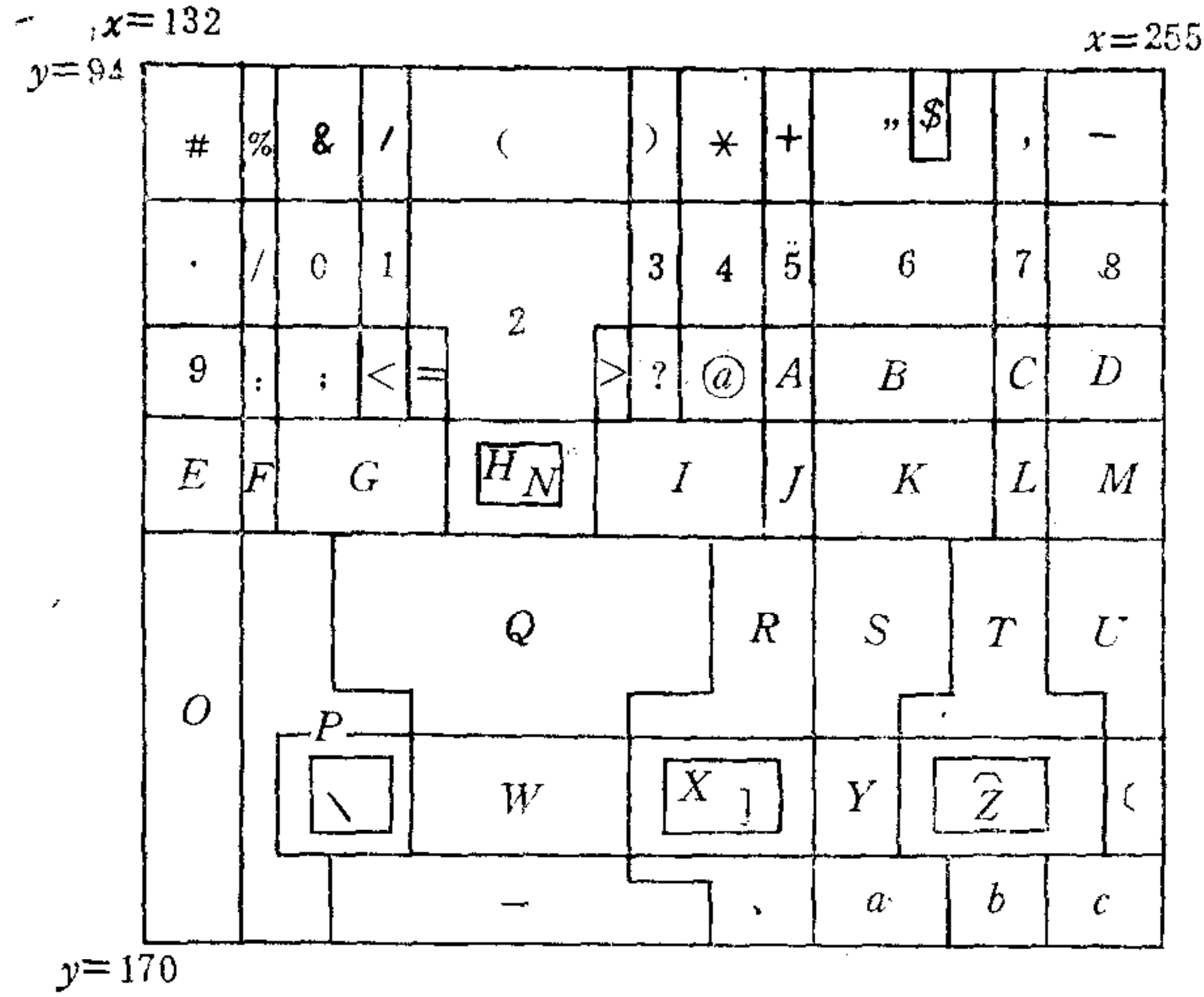


图4 待匹配图左幅

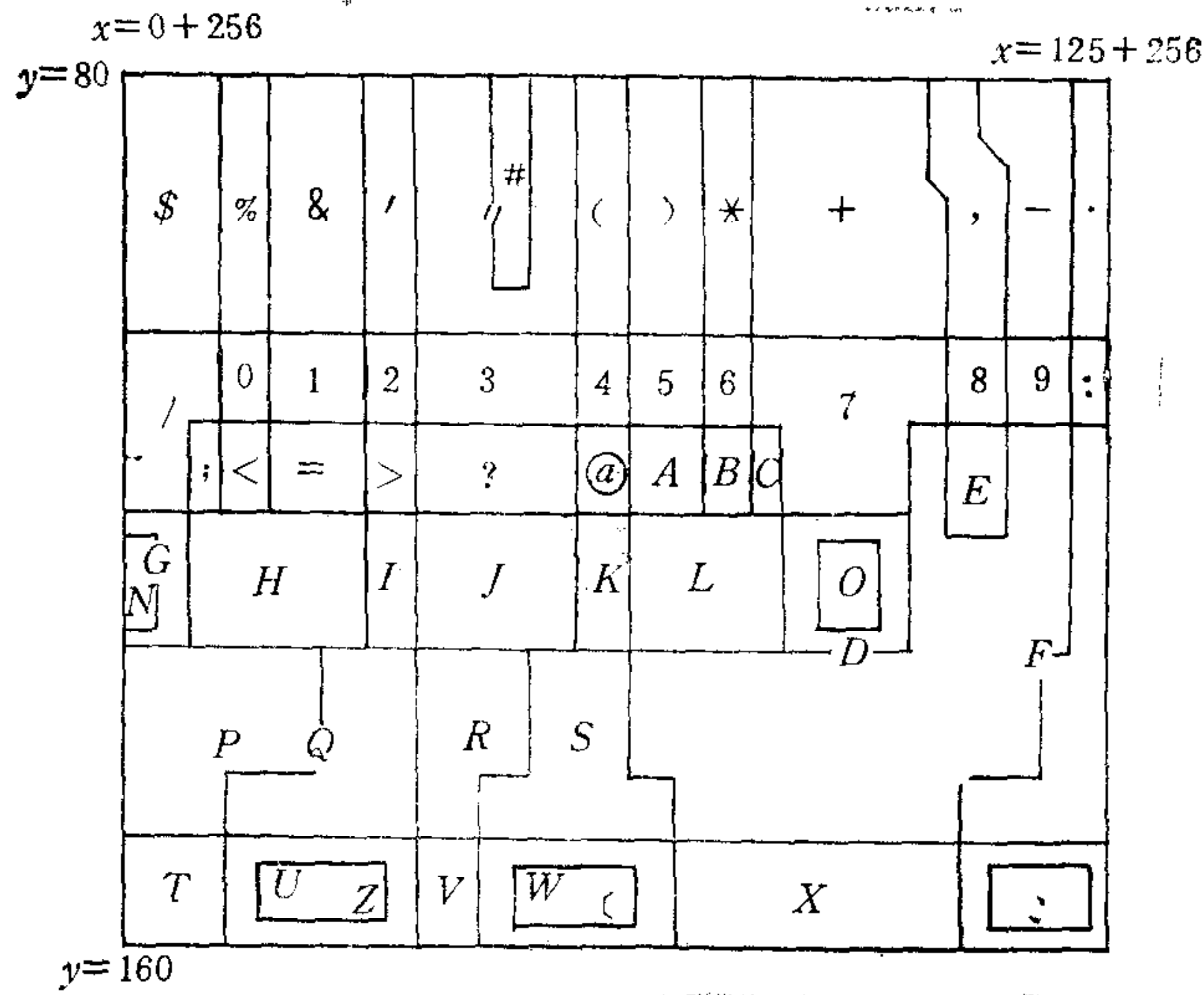


图5 待匹配图右幅

### 参 考 文 献

- [1] William K. Pratt, Correlation Techniques of Registration, *IEEE Trans. on AES.*, Vol. AES-10, No. 3, May 1974, 353—358.
- [2] C. Granger, Symbolic Scene Matching, 7th International Conference on Pattern Recognition, 1984, 883—885.
- [3] Liang, Y. D. & Barsky, B. A., An analysis and algorithm for polygon clipping, *Comm.ACM*, Vol. 26, No. 11, Nov., 1983, 868—877.



## “DOUBLE” SUBGRAPH ISOMORPHISM METHOD FOR MATCHING LSI CHIP IMAGES

WANG JIREN

*(East China Normal University)*

LI JIEGU

*(Shanghai Jiaotong University)*

### ABSTRACT

This paper presents a “double” subgraph isomorphism method for matching LSI chip images, by which the corresponding relation between isomorphic matched nodes is searched from the region adjacency graph of the segmented chip images.

In order to perform the algorithm of isomorphism matching, heuristic information is extracted to improve the searching efficiency. The algorithm uses special criteria of region similarity measure and matching figure of surrounding string to select the starting node of isomorphism matching.

In the paper, we present not only the principle of “double” subgraph isomorphism matching but also the matching result. This algorithm has been employed in practice.