

一种在矢量基础上进行图形识别的通用方法

刘文印 唐龙 唐泽圣

Dov Dori

(清华大学计算机科学与技术系 北京 100080)

(以色列工程技术学院工业工程与管理系)

摘要 本文描述了一种在矢量基础上进行图形识别的通用方法. 该方法包括 2 个步骤: 第 1 步, 寻找组成该图形的第 1 个关键图素; 第 2 步, 根据该图形的构造模式, 从第 1 个关键图素出发, 不断在其周围区域寻找组成该图形的其它图素. 本文还提出了一种新的平面空间数据结构——位置索引. 这种数据结构把平面图上的图形及图素用它们的平面位置索引组织起来, 从而提高了按区域位置查找的效率. 本文还介绍了该方法的 2 个应用实例: 虚线识别和剖面线识别.

关键词 图形识别, 空间数据结构, 索引, 工程图理解, 虚线识别, 剖面线识别.

中图分类号 TP391

图形识别是图形理解的基础. 在线型图理解, 特别是工程图理解中, 图形识别尤为重要, 因此受到普遍重视. 近年来, 人们已提出了针对某些图形的识别方法, 如字符分离与识别^[1,2]、箭头识别^[3]、圆弧识别^[4]、虚线识别^[5~7]、标注识别^[8~10]及方框识别.^[11]但是, 这些图形识别方法都是专用的, 只能用于各自种类的图形识别, 还没有一个统一的通用的图形识别方法. 对于某些干净而简单的图纸, 用这些方法可取得较好的识别效果, 但对于中等复杂程度以上的图纸, 其识别效果尚不令人满意.

为了统一对各种图形的识别方法并提高识别率, 我们提出了一种通用的图形识别方法. 因为基于矢量的图形识别方法更方便易行, 所以我们的通用图形识别方法也是建立在矢量基础上的, 即在矢量化后进行. 该方法根据已经找到的图素的几何位置信息渐进地、逐个地寻找组成图形的所有其它图素, 而不是直接使用模型匹配的方法. 在明确表示某种图形的构成语法之后, 通用识别方法即可用于识别这种图形. 因此, 此方法可用于多种类型的图形识别.

此方法中, 最主要的、大量的操作是查找几何图素. 我们提出了一种新的空间数据结构——位置索引. 这种数据结构把平面图上的图形及图素用它们的平面位置索引组织起来, 从而提高了按区域位置查找图素的效率. 用这种数据结构所实现的通用图形识别方法具有很

• 本研究得到国家自然科学基金、国家 863 高科技项目基金和国家八五攻关项目基金资助. 作者刘文印, 1966 年生, 讲师, 主要研究领域为人工智能, 工程图输入及处理. 唐龙, 1938 年生, 副教授, 主要研究领域为计算机应用, 工程图输入及处理, 科学计算可视化. 唐泽圣, 1932 年生, 教授, 博士生导师, 主要研究领域为计算机图形学, 科学计算可视化, 计算几何, 工程图输入及处理. Dov Dori, 1953 年生, 教授, 博士生导师, 主要研究领域为计算机图形学, 工程图输入及处理, 信息系统等.

本文通讯联系人: 刘文印, 北京 100084, 清华大学计算机科学与技术系

本文 1996-05-27 收到修改稿

高的识别速度。

第 1、2 节讨论该通用图形识别方法及所用数据结构——位置索引。第 3 节介绍该通用图形识别方法的 2 个应用实例,即虚线识别和剖面线识别。用该方法实现的虚线识别软件在 1995 年 8 月 10 日美国宾西法尼亚州立大学举办的“图形识别”国际讨论会期间所举办的“虚线识别竞赛”上取得了第 1 名的佳绩。最后是一些评价。

1 通用图形识别方法

这是基于矢量的方法,当然只能在矢量化后应用。但是,作为一种图形识别的通用方法,该方法对矢量化的要求很低,一般只要求保留有线宽信息的矢量线段和折线集。虽然,象圆弧、箭头等稍高层次的图素将会提供更多线索,但它们并不是必不可少的。

1.1 基于语法的图形线索

每种图形都是由一些较低层次的图素或图形按其自身的语法和语义规则组成的。例如,一条虚线是由一组共线的、不连续的、等粗细的、等长的和等间隔的短线段组成的;又如,箭头通常被矢量化成一段粗线段,但其一头要连着一条与之共线的细线段作为它的尾部,另一头顶着一条与之垂直的线段作为引证线。本文所提出的通用识别方法的思想,就是根据每类图形的语法规则来寻找其组成图素。如果能找到足够的组成图素,就能确认为某种图形。

我们首先要知道该图形是否存在。一般每类图形都有其存在的线索。其线索可分为 2 类。一类是直接的,即其组成图素的类型可作为其存在的线索,称之为组成线索。例如,共线的短粗线段可作为虚线存在的线索。然而,有时这类线索帮助不大,主要原因是该类组成图素对该类图形的表征性不够好,还需要大量地测试候选图素。例如,字符可能由一些短线段组成,很多其它类图形也含有短线段。若以短线段作为线索来测试字符的存在,势必会增加寻找测试的工作量和难度。因此,该类型的线索要有突出的特殊性才好。另一类线索是间接的,即用其常相邻的图形或图素作为其存在的线索,称之为相邻线索。例如,工程图中的箭头经常出现在字符附近(因为它们都是标注的组成部分),因此,箭头可以作为字符存在的线索。搜索箭头附近的区域很可能会找到字符,反之亦是如此。

在图形识别中,这 2 种线索(组成线索和相邻线索)可能要同时用到。这是因为有时其中一类不存在或不好用,而且多一种线索有助于降低误识率。我们的通用图形识别方法就是从这 2 类线索出发,用渐进搜索方式来完成的。

1.2 图形识别方法的渐进搜索方式

我们把能作为线索的图素称为关键图素。由于关键图素的存在,能表明某些图形存在的概率很大,因此我们可以从寻找关键图素开始,逐步找到组成某种图形的其它图素,从而识别该图形。

我们的通用图形识别方法就是这样一种自底向上的渐进搜索方式,它主要包含 2 步:第 1 步,找到第 1 个关键图素;第 2 步,在关键图素周围区域搜索组成该图形所应有的其它图素。

第 1 步,用函数 FindFirstComponent(),根据上述 2 种线索,在所要识别的图纸的图素数据库中寻找第 1 个关键图素。当然,首先要知道识别什么图形(假设该图形存在),然后根据上述 2 种线索来选定其第 1 个关键图素的类型。第 1 个关键图素类型的选定要依照以下

原则:①要保证所有这类图形中必定都含有该类图素,否则,有些不含该类图素的这类图形就不可能被识别出来;②要保证该类图素对这类图形的一定的表征性,即该类图素在一定程度上能代表这类图形的存在,或只有这类图形中有该类图素,一般选用更有代表性的关键图素类型,会使候选的第1关键图素减少,从而可减少测试量.因此,第1个关键图素的类型的选定很重要.若条件过紧,则只能识别少数图素;若条件过松,则可能导致误识过多.当然,若找不到该类图素,则应认为这类图形不存在.

第2步,用函数 Extend(),从找到的第1个关键图素开始,在它周围所有可能的区域中,由近至远地进行搜索,逐个地找出组成该图形所需的其它图素.实际上,这一步是完成对第1步中所作假设的测试.如果能在预期的区域找到预期的其它图素,则说明假设成立,从而证实了该图形的存在.经过了第1步的严格筛选,该图形存在的可能性已经增加,因此,第2步中的条件可以适当放宽.找到的组成图素越多,该图形存在的概率就越大.当这个概率增加到一定程度(例如达到90%)或找到了足够的图素时,就足以确定该图形的存在.确定了该图形的存在以后,还可以继续在其周边区域查找其它图素,直到找不到其它图素为止.

由上述2个主要步骤组成的通用图形识别方法,可用C++语言概括描述算法如下:

```
while ((FirstComponent = Graphic.FindFirstComponent()) != NULL)
{
    Graphic.InitializeWith(FirstComponent, (Likelihood)0);
    while ((Component = Graphic.Extend()) != NULL) Graphic.Add(Component);
    if (Graphic.Likelihood() > 0.9) Graphic.AddToGraphicDataBase();
}
```

这是由上述2个主要步骤组成的一般的框架模式,可用于多种图形的通用识别方法.将它用于某类具体图形的识别时,当然,还需要加入其具体的构造语法规则,即函数 FindFirstComponent()和 Extend()要用该图形的构造语法规则加以实现.

2 一种有效的数据结构——位置索引

2.1 空间查找

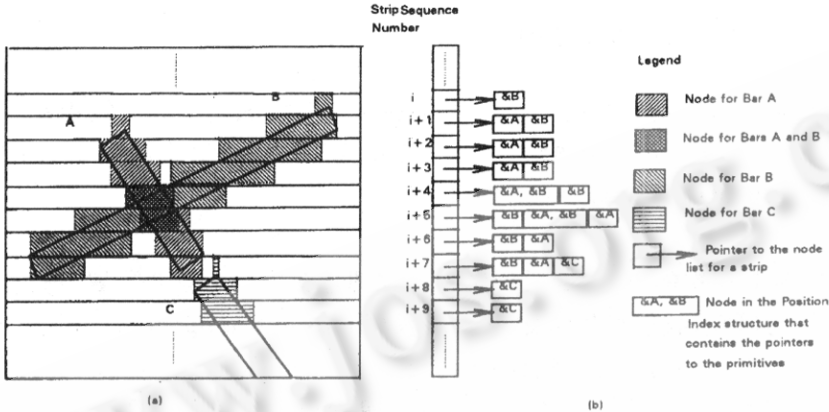
在图形识别过程中,查找是非常频繁进行的操作,占用了大部分时间.一般的工程图纸图形识别系统中,图素和图形都是用线性表来存储的,它们之间的几何位置关系无法表示,因此,按位置查找图素或图形时,需要在该线性表中进行顺序查找.这种全区域顺序查找很费时.有时,我们只要在很小的一个区域查找,若仍做全区域顺序查,则既不必要也不方便.所以,以这种线性表形式组织的数据库对图形识别来说已远远不够,需要设计更合适、更有效的数据结构.

在图形系统中,图形的空间表示(在图纸上是平面表示)是一个重要的问题.不仅要表示图形本身,还要表示图形之间位置和拓扑等关系.每个图形都存在于空间某个位置上.每个图形都可以从其位置坐标来实现其自身到其空间位置的映射.从其空间位置到其自身的映射并不容易简单地表示清楚,而按位置查找正需要这种从其空间位置到其自身的映射能快速实现.线性表结构不能实现这种映射,因而,在线性表中进行按位置查找很慢.为了实现这种从其空间位置到其自身的映射,我们提出给图形数据库中的图素和图形加一个按位置

的索引. 采用位置索引可以更有效地组织和管理图形数据.

2.2 位置索引

为了实现从平面位置到图形图素自身的映射, 图形数据库中的图形图素都以下述方式, 用它们的平面几何位置索引组织起来, 给图形数据库建立一个位置索引.



(a) 3 条线段 A,B,C 及其经过的索引结点区域

(b) 位置索引结构

图 1 位置索引

把整个图形区域(一张图纸)划分为若干等宽的横格, 如图 1 所示. 横格的宽度取决于平行线间的最小距离. 每一横格都有一个序号和一个索引结点表. 该索引结点表可存放若干索引结点. 每个结点是一个长方形区域, 有一个左边界和一个右边界, 其高度就是该结点所在横格的宽度. 每个结点都有一个图素指针表, 存放与该结点区域相交的所有图素的指针. 每一横格的索引结点表中的所有结点是以结点的左右边界为升序排列的. 每一横格的索引结点表中的所有相邻结点的区域不能相交, 但可以连续.

在图 1 中, 共有 3 条线段. 索引结点的左右边界由通过该区域的线段决定. 索引结点的左边界是线段的左边框与该横格上下两边交点的最小值. 索引结点的右边界是线段的右边框与该横格上下两边交点的最大值. 在线段相交的区域, 原来的索引结点分裂成几个结点, 使得每个线段都占满所有与之相交结点区域, 即线段与结点相交的左右边界要与该结点的左右边界一致. 这一分裂是在向该位置索引结构中不断加入新的图素过程中逐渐发生的.

2.3 复杂性分析

每一横格的索引结点表中的所有结点以结点的左右边界按升序排列. 因此, 可用对有序表的查找方法进行按位置查找, 其时间复杂度为 $O(\log(n))$, 其中 n 为该横格内索引结点表的所有结点数, 约为通过该横格内的图素数量. 若搜索某一区域, n 则为通过该搜索区域的图素数量. 如果不用位置索引, 只在线性图素表中进行按位置查找, 需要查找整个表, 那么, 查找的时间复杂度应为 $O(N)$, 其中 N 为整个图纸上全图形区域中的图素数量. 一般来说, N 远远大于 n . 显然, 利用位置索引可使查找速度大大提高.

当然, 位置索引的空间复杂度稍高一些. 原因是整个位置索引都需要另增空间予以存放. 位置索引所需的空间大小取决于横格的宽度和图素与横格的相交角度. 横格越宽, 所需的空间越小, 但是图素位置的表示越欠精确. 与水平图素相交的横格比与等长的垂直图素相交的横格要少得多. 因此, 水平的图素比垂直的图素占用的空间要少. 对于非水平的图素, 与

横格的交角越小,所需的空间就越少.这一点可以从图 1 中看得很清楚.最好的情况是图中所有图素都是水平的,此时,只需 $O(N)$ 空间来存储索引结点.而最坏的情况是图中所有图素都是垂直的且其长度与图高相等,此时,需要 $O(mN)$ 空间来存储索引结点,平均空间复杂度仍是 $O(mN)$,其中 m 是图中的横格数量.

3 通用图形识别方法的实际应用

本节给出通用图形识别方法的 2 个应用实例——虚线识别和剖面线识别.要想实现某一类图形的识别,只要根据其构造语法规则,实现此通用图形识别方法中 2 个主要步骤的函数 FindFirstComponent() 和 Extend() 即可.

3.1 虚线识别

虚线是一类较高层次的图形,虚线识别在工程图理解中占有重要位置.通过对虚线的具体构造语法规则的研究,我们可以确定,执行函数 FindFirstComponent() 所需的第 1 关键图素的类型应当是符合下列条件的线段:① 该线段的长度小于一定值;② 该线段的宽度大于一定值;③ 该线段的 2 个端点悬空,即该线段的两个端点处无其它图素经过.

在第 2 步中,执行函数 Extend(), 寻找其它图素的搜索区域应被限定为从当前已找到的图素向该虚线的两个延长方向延伸.在该搜索区域中,所找到的图素(线段)除了要满足第 1 步中的 3 个条件外,还要经过筛选,还应当满足以下 2 个补充条件:④ 该线段要与该虚线的已找到的其它线段共线;⑤ 该线段要与离它最近的虚线段有一定间隙,该间隙应大致等于该虚线上其它各线段间隙的平均值.

在第 2 步 Extend() 中,条件③(即要求线段的两个端点悬空)可以放宽.其原因是虚线有可能与其它图形(如剖面线)相交.

把上述规则加入到通用图形识别方法中的 2 个主要步骤 FindFirstComponent() 和 Extend() 后,就实现了虚线识别.这个虚线识别程序取得了很好的识别效果.1995 年 8 月 10 日在美国宾夕法尼亚州立大学召开“图形识别”国际讨论会期间,举办的“虚线识别竞赛”上,我们这个虚线识别程序取得了第 1 名.^[12]图 2 是 1 个随机生成的带噪音的虚线图象的一部分.黑象素是原始图象,其上的灰色线段是已识别的虚线重画的结果.该图象是由美国华盛顿大学 R. M. Haralick 教授领导的小组编写的图象自动生成程序生成的.经过由美国宾夕法尼亚州立大学 R. Kasturi 教授领导的小组编写的虚线识别自动评价程序的评价,我们这个虚线识别程序对此图象的识别率是 100%,其中只有 1 个误识.图 3 是评价的结果.

3.2 剖面线识别

在工程图纸中,剖面线是更高层次的图形.剖面线识别有助于理解物体的剖面结构.通过对剖面线性质的研究,我们认为剖面线通常是由一组等距平行的斜细线组成的,并由轮廓线包围着.因此,我们可以确定执行函数 FindFirstComponent() 所需的第 1 关键图素的类型应当是符合下列条件的线段:

① 该线段是水平倾角为 $30\sim 60^\circ$ 的细线;② 该线段的长度应在一定范围之内;③ 该线段的两个端点不能悬空.

在第 2 步中,执行函数 Extend(), 寻找其它图素的搜索区域被限定为从当前已找到的剖面线组向与该剖面线垂直的两侧方向延伸.在该搜索区域中,所找到的图素(线段)除了要

满足第 1 步中的 3 个条件外,还要经过筛选,还应满足以下 3 个补充条件:④ 该线段平行于已找到的其它剖面线;⑤ 该线段与最近的剖面线的距离应大致等于该组剖面线之间的平均距离;⑥ 该线段与已找到的其它剖面线共享同一轮廓线。

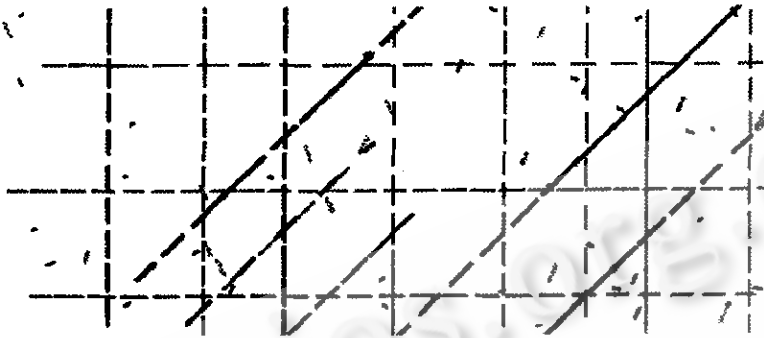


图 2 虚线识别的结果

```

.....
:: Detection Rate ::
.....

```

	Detected Lines		
	same as line-type i	Not line-type i	Not detected
Ground truth line-type i	100.00% correct	0.00% mis-label	0.00% mis-detect
Not ground truth	4.35% false-alarm		

图 3 对图 2 中虚线识别的自动评价结果

在第 2 步 Extend() 中,条件③(即要求在线段的 2 个端点不悬空)可以放宽,其原因是有些剖面区域轮廓线并不封闭。

将上述规则加入通用图形识别方法后,实现的剖面线识别如图 4 所示,图 4 是一张已经经过矢量化了的机械图的一小部分(原图由德国西门子公司提供),其中的浅色线段区域是识别出的剖面线,共有 2 组,上面的一组中,共识别出 5 条剖面线,另外,还有一条小线段由于太短或其角度不合适而未被识出,下面的一组中,共有 6 条剖面线全部识别出来。

4 结束语

本文描述了一种在矢量基础上进行图形识别的通用方法及所用的一种新的空间数据结构——位置索引,这个图形识别的通用方法把许多种图形的识别归纳成为一种统一的、通用的渐进搜索模式,不需要对各种图形、各种情况(如水平、垂直或斜的图形)进行分门别类地特殊处理,这种统一的通用方法不仅使基于知识的集成的图形识别系统成为可能,还可使一些复杂的推理方法,例如,非单调推理得以方便实现,我们的图形识别的通用方法虽是基于

矢量的,但对矢量化要求并不高,只要求保留有线宽信息的线段和折线集即可.诸如 OZZ^[13]和 SPT^[14]之类的矢量化方法均可作为其前处理.某些不保留线宽信息的基于细化的矢量化方法则不适用.当然,对矢量化的图形文件格式有所要求,但只要转换成比较流行的 DXF 格式便可进行处理.位置索引实现了从物体(图形)空间位置到物体(图形)本身的映射,因此大大提高了按位置查找物体(图形)的效率.位置索引的使用有效地提高了图形识别速度.本文列举了 2 个应用实例:虚线识别和剖面线识别,表明该方法是成功的.实际上,该方法还可用于其它图形识别,例如,图文分离^[15]、箭头识别和标注识别以及轮廓线识别等.

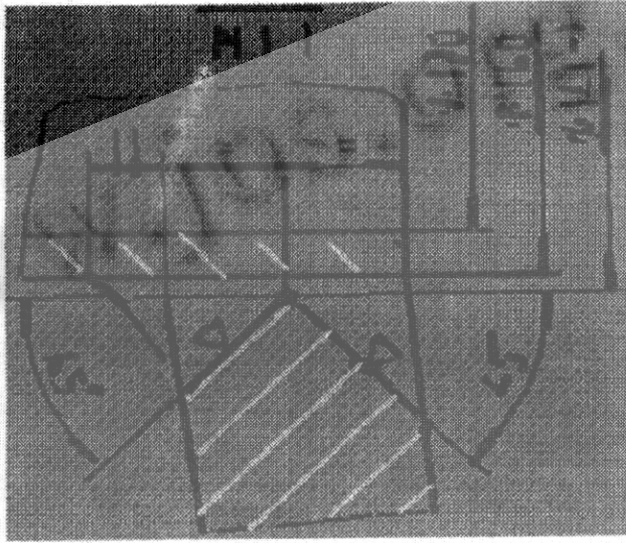


图 4 剖面线识别结果

该方法的研究和开发还有待进一步深入和完善.它已提供了一个统一的、通用的图形识别框架,还需要具体的图形构造语法规则予以填充,才能完成对特定图形的识别.此外,本文所提出的位置索引,其空间复杂度也有待进一步降低.

参考文献

- 1 Dori D, Chai I. Extraction of text boxes from engineering drawings. In: Proc. SPIE/IS&T Symposium on Electronic Imaging Science and Technology, Conference on Character Recognition and Digitizer Technologies, San Jose(CA, USA), SPIE, 1992:1661,38~49.
- 2 Gao J, Tang L, Liu W *et al.* Segmentation and recognition of dimension texts in engineering drawings. In: Proceedings of 3rd International Conference on Document Analysis and Recognition, Montreal, Canada, August 1995. 528~531.
- 3 Dori D, Liang Y, Dowell J *et al.* Spare pixel recognition of primitives in engineering drawings. Machine Vision and Applications, 1993, 6:69~82.
- 4 Dori D. Vector-based arc segmentation in the machine drawing understanding system environment. IEEE Transactions on PAMI, 1995, 17(11):1159~1071.
- 5 Kasturi R, Bow S T, El-Masri W *et al.* A system for interpretation of line drawings. IEEE Transactions on PAMI, 1990, 12(10):978~992.
- 6 Jodeph S H, Pridmore T P. Knowledge-directed interpretation of mechanical engineering drawings. IEEE Trans.

- on PAMI, 1992, **14**(9).
- 7 Boatto L *et al.* An interpretation system for land register maps. *IEEE Computer Magazine*, 1992, **25**(7):25~32.
 - 8 Min W, Tang Z, Tang L. Using web grammar to recognize dimensions in engineering drawings. *Pattern Recognition*, 1993, **26**(9):1407~1416.
 - 9 Lai C P, Kasturi R. Detection of dimension sets in engineering drawings. *IEEE Trans. on PAMI*, 1994, **16**(8):848~855.
 - 10 Das A K, Langrana N A. Recognition of dimension sets and integration with vectorized engineering drawings. In: *Proceedings of 3rd International Conference on Document Analysis and Recognition*, Montreal, Canada, 1995. 347~350.
 - 11 Vaxiviere P, Tombre K, Celesstin; CAD conversion of mechanical drawings. *IEEE Computer Magazine*, 1992, **25**(7):46~54.
 - 12 Dori D, Liu W, Peleg M. How to win a dashed line detection contest. To appear in *Lecture Notes in Computer Science*, Springer-Verlag, 1996.
 - 13 Chai I, Dori D. Orthogonal zig-zag: an efficient method for extracting lines from engineering drawings. Arcelli C, Cordella L P, Sanniti G di Baja eds. *Visual Form*, Plenum Press, New York and London, 1992. 127~136.
 - 14 Liu W, Dori D. Sparse pixel tracking: a new algorithm for basic vectorization of engineering drawings. To appear in *Proc. of the 12th Israeli Symposium on Artificial Intelligence, Neural Networks and Computer Vision*, Tel Aviv, Feb. 4~5, 1996.
 - 15 Liu W, Dori D. Vector-based text segmentation in engineering drawings. To appear in *Proc. of the 12th Israeli Symposium on Artificial Intelligence, Neural Networks and Computer Vision*, Tel Aviv, Feb. 4~5, 1996.

A GENERIC METHOD FOR VECTOR BASED GRAPHIC RECOGNITION

LIU Wenyin TANG Long TANG Zesheng

(*Department of Computer Science and Technology Tsinghua University Beijing 100084*)

Dov Dori

(*Faculty of Industrial Engineering and Management Technion-Israel Institute of Technology Haifa 32000 Israel*)

Abstract A unified generic method for the vector based graphic recognition is presented. The scheme consists of two main steps: finding the first key component and extending it according to the specific graphic syntax. A new spatial data structure, the position index, which facilitates the search for the constituent primitives in given areas, is devised to realize the mapping from planar positions to graphic primitives. Two applications, dashed line detection and hatching lines detection, are also presented as two cases in point.

Key words Graphic recognition, spatial data structure, indexing, engineering drawings interpretation, dashed line detection, hatching lines detection.

Class number TP391