

基于单义域邻接图的圆弧与圆识别

张习文 欧宗瑛

(大连理工大学机械系 CAD & CG 研究所, 大连 116024)

摘要 工程图纸扫描输入与识别理解是 CAD 推广和普及的关键步骤之一, 主要解决已有大量图纸再利用问题. 在工程图纸扫描图象识别研究中, 圆弧识别是识别算法中的重点和难点. 传统的圆弧识别多是基于线段逼近. 该文提出一种基于单义域邻接图的圆弧及圆识别算法, 可以直接提取圆弧. 对二值图象作水平黑游程编码, 相关游程基于线宽与拓扑的一致性构成条形域, 对其中多义域进行分裂得单义域(线段域和圆弧域). 单义域邻接图可较好描述图象的几何属性与拓扑关系. 单义域具有明显的形状意义(线段、圆弧、箭头等), 提高了识别的整体性. 圆弧及圆的识别先从邻接图顶点中抽取圆弧域, 作为种子圆弧, 然后从此出发遍历图, 按照同圆来建立路径, 进行整弧和整圆增长, 最终获得圆弧和圆的几何表达. 实例表明, 本算法可以较好地处理圆弧与线段及圆弧的相交与相切, 适应性较强、识别率较高.

关键词 工程图纸 矢量化 圆弧识别 条形域 单义域邻接图

中图分类号: TP391.41 **文献标识码:** A **文章编号:** 1006-8961(2000)01-0070-05

An Algorithm for Recognizing Circular Arcs and Circles Using Primitive Region Adjacency Graph

ZHANG Xi-wen, OU Zong-ying

(CAD & CG Institute, Dalian University of Technology, Dalian 116024)

Abstract The scanning input and recognition of engineering drawings is a key step in CAD, and is to reuse lots of engineering drawings. In study on recognition of scanned image of engineering drawings, the recognition for circular arcs is an important and difficult problem. Recent algorithms of recognizing arcs are mainly about approximation with lines. This paper presents an algorithm for recognizing arcs and circles using Primitive Regions Adjacent Graph, which can directly extract arcs. The binary image is encoded with black horizontal runlength. A stripe region consists of correlative runlengths with the same width and topology. The stripe regions then can be segmented as some primitive regions (line and arc). The graph is used to describe geometrical property and topological constraint. The primitive region supplies shape information (line, arc, arrow etc.) improving integrality of recognition. After extraction of the arc region from regions, the seed for an arc is obtained. By traversals for the graph, the seed arc grows by constrains for the same circle. Some applications to recognize arcs and circles are finally provided, which show that the algorithm is effective and robust, can solve well intersection and tangency of between an arc and a line or an arc.

Keywords Engineering drawings, Vectorization, Recognition for circular arc, Stripe region, Primitive Region Adjacency Graph

0 引言

圆弧和圆是工程图形中的重要图元, 已有多种识别算法, 可分为两种: 逼近法和直接法. 细化方法先跟踪中心骨架象素得到短小线段, 再用来逼近圆弧和

圆^[1]. 正交扫描法(orthogonal zig-zag)是先获得条, 再用中垂线跟踪(perpendicular bisector tracing)分割圆弧^[2]. 文献[3]以梯形域来逼近圆弧和圆. 这3种方法都是以线段来逼近圆弧和圆, 如果线段过短, 会造成数据冗余; 如果线段过长, 将难以识别短小圆弧, 需要后续处理. 轮廓匹配法可直接获得圆弧和圆, 但

轮廓获取及其匹配都很复杂^[4]. 文献[5]采用图段与圆进行模式匹配, 确定圆的种子图段, 然后跟踪其它图段, 最终获得圆弧和圆的图形表示.

本文提出一种新的识别方法, 以相关游程线宽和拓扑为约束生成条形域, 对其中多义域作分裂获得单义域: 线段域和圆弧域, 并建立其邻接图, 选取弧形域, 以此为起始点, 基于同圆几何要求, 通过深度优先搜索来遍历图, 完成圆弧和圆识别.

1 条形域构建和多义域分裂

1.1 游程分类

在同一线宽的二值图象中, 一个游程唯一属于单一图元, 所以, 对工程图纸扫描图象先作图文分离再作粗细分离^[3], 然后对同一线宽图象进行处理. 本文采用自上而下和从左到右扫描, 定义同一行中连通黑像素为一水平黑游程, 并以此对图象进行编码. 如果当前游程是参考游程的上(或下)一行, 且左右至少一端搭接(包括邻接), 则称两个游程相关, 具有父子(或子父)关系. 游程可以根据其父子游程进行分类^[6], 本文针对工程图纸扫描图象特点分为 7 类(如图 1 所示):

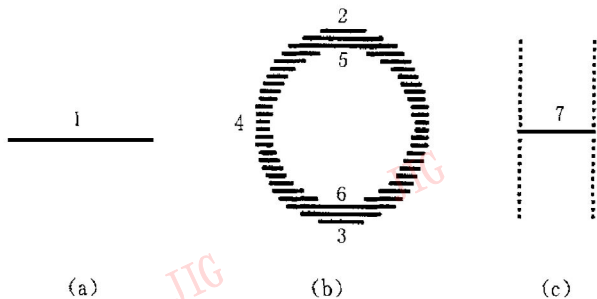


图 1 游程分类

- (1) 孤立游程: 没有父游程和子游程;
- (2) 起始游程: 没有父游程, 只有一个子游程;
- (3) 终止游程: 没有子游程, 只有一个父游程;
- (4) 单一游程: 只有一个父游程和一个子游程;
- (5) 分叉游程: 有多个子游程, 至多一个父游程;
- (6) 汇合游程: 有多个父游程, 至多一个子游程;
- (7) 交叉游程: 有多个父游程和多个子游程.

1.2 构建条形域

相关游程基于宽度和拓扑的一致性组合为条形域, 具体条件如下:

- (1) 游程相关;
- (2) 首游程不是分叉游程和交叉游程, 末游程不是汇合游程和交叉游程;

(3) 游程宽度没有较大变化;

(4) 孤立游程和交叉游程单独构成条形域.

条件(1)保证条形域的连通性; (2)确保交叉域、汇合域和分叉域的准确建立; (3)是条形域宽度一致性的要求; (4)处理特殊情况.

下面给出条形域构建算法:

(1) 移游程到新建条形域中, 设为当前游程, 如果是孤立游程、分叉游程或交叉游程, 则转到(4);

(2) 取当前游程的子游程, 如果与当前条形域游程平均宽度相差较大, 则转到(4), 否则, 设为当前游程;

(3) 如果当前游程是单一游程, 则加入条形域, 并返回(2); 如果是分叉游程或终止游程, 则加入条形域;

(4) 完成条形域构建.

1.3 分裂多义域

从几何意义上讲, 根据上述方法构建的条形域包括两种情况: (1) 表达单一几何实体, 如线段、圆弧和箭头等; (2) 线段与圆弧组合, 如折线等. 所以, 将条形域分为单义域和多义域. 一般来讲, 多义域可看作由线段和圆弧连接而成. 因此, 对多义域进行分裂, 提取线段域和圆弧域.

在这里, 多义域是开环的, 且单调, 相对一般曲线拟合要简单些, 本文采用首尾相连最大距离法来分裂^[7]. 对多义域, 先用一条直线连接首尾游程中点, 计算多义域游程中点与直线的最大距离差, 并以此点为一分裂点, 然后判定所分裂的两段是否为单义域, 如仍为多义域则继续分裂, 一直进行到都是单义域为止, 这是一个递归的过程. 下面给出分裂步骤:

(1) 输入条形域, 如为单义域, 则转到(5);

(2) 对多义域以最大距离法分为两段: 域 1 和域 2, 输出分裂点;

(3) 如果域 1(域 2)为多义域, 则返回(2);

(4) 对分裂点以其 y 坐标作从小到大排序;

(5) 根据分裂点, 输出单义域.

2 单义域邻接图的建立及其拓扑分类

2.1 建立单义域邻接图

多义域分裂后, 图象描述单元变为单义域. 但, 单义域描述只是表达几何信息, 而单义域之间还有连接关系, 表达拓扑信息. 如果当前单义域末游程与参考单义域首游程(或当前单义域首游程与参考单义域末游程)相关, 则称两个单义域相关, 具有父子

(或子父)关系.采用域邻接图(如梯形域和图段等)可以较好地表达图象的几何与拓扑信息,本文采用单义域邻接图描述图象,下面给出建立步骤:

- (1) 加单义域到邻接图的顶点表中;
- (2) 取顶点 i, i 从 1 到 n, n 为图的顶点总数,
 - {
 - 取顶点 j, j 从 1 到 $n,$
 - {
 - 如果顶点 j 是顶点 i 的父域,则加顶点 j 到顶点 i 的父邻接表,且加顶点 j 到顶点 i 的邻接表;
 - }
 - }

如果顶点 j 是顶点 i 的子域,则加顶点 j 到顶点 i 的子邻接表,且加顶点 j 到顶点 i 的邻接表;

图 2 给出圆的图表示,顶点 1 的父邻接表为空,子邻接表包含顶点 2 和 3; 顶点 2 的父邻接表包含顶点 1,子邻接表包含顶点 4; 顶点 3 的父邻接表包含顶点 1,子邻接表包含顶点 4; 顶点 4 的父邻接表包含顶点 2 和 3,子邻接表为空.

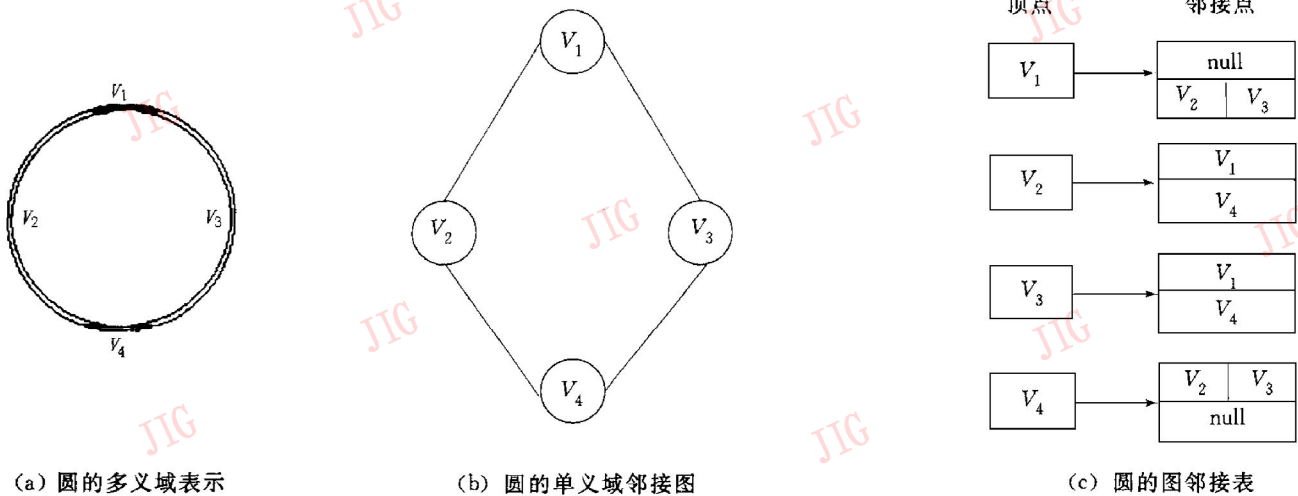


图 2 圆的图表示

2.2 单义域拓扑分类

从单义域邻接图中很容易就可获得某一顶点的邻接点情况,单义域(顶点)根据其父子域(邻接点)可以分为 7 种拓扑类型(如图 3 所示):

- (1) 孤立域: 没有父域和子域;
- (2) 起始域: 没有父域,只有一个子域;
- (3) 终止域: 没有子域,只有一个父域;
- (4) 单一域: 只有一个父域和一个子域;
- (5) 分叉域: 有多个子域,至多一个父域;
- (6) 汇合域: 有多个父域,至多一个子域;
- (7) 交叉域,有多个父域和多个子域.

此.可以看出,单义域不仅为单一完整图元的获得提供几何数据,而且也为图元之间的拓扑关系提供线索,单义域邻接图比较完整地表达了图象的几何数据与拓扑关系.

3 圆弧识别

3.1 确定种子圆弧

本文采用假设验证法从单义域中提取圆弧域,主要思路是,先假设候选域为圆弧域,采用最小二乘法来计算其几何定义数据(圆心和半径)^[8],并计算平均径向误差和最大径向误差,如果小于阈值,则判定为圆弧域,作为种子圆弧,以指导圆弧和圆的识别.

在确保计算精度的前提下,为提高速度,平均抽取 5 个游程的中点,如果单义域游程数小于 5,则全部选取.基于径向误差最小,对样点采用最小二乘法拟合,下面给出计算公式:

设给定样点为 $pt_i(x_i, y_i)$, 所求圆心为 $pt(x_o, y_o)$, 半径为 R , 平均径向误差为 δ , 最大径向误差为 σ_{max} , 则

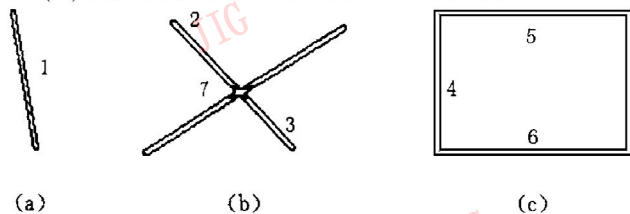


图 3 单义域分类

在图 3(a)中,孤立域 1 表示一个图元,即一条线段.在图 3(b)中,交叉域 7 连接起始域 2 和终止域 3,组成一条线段,还表示该线段与另一线段相交于

$$\begin{aligned}
x_0 &= \frac{\left| \frac{2\sum_{i=1}^m x_i \sum_{i=1}^m y_i}{m} - 2\sum_{i=1}^m x_i y_i \right| \left| \sum_{i=1}^m y_i^3 + \sum_{i=1}^m x_i^2 y_i - \frac{\sum_{i=1}^m x_i^2 + \sum_{i=1}^m y_i^2}{m} \sum_{i=1}^m y_i \right|}{\left| \frac{2\left| \sum_{i=1}^m x_i \right|^2}{m} - 2\sum_{i=1}^m x_i^2 \right| \left| \frac{2\left| \sum_{i=1}^m y_i \right|^2}{m} - 2\sum_{i=1}^m y_i^2 \right| - \left| \frac{2\sum_{i=1}^m x_i \sum_{i=1}^m y_i}{m} - 2\sum_{i=1}^m x_i y_i \right|^2} \\
&\quad \frac{\left| \frac{2\left| \sum_{i=1}^m y_i \right|^2}{m} - 2\sum_{i=1}^m y_i^2 \right| \left| \sum_{i=1}^m x_i^3 + \sum_{i=1}^m x_i y_i^2 - \frac{\sum_{i=1}^m x_i \left| \sum_{i=1}^m x_i^2 + \sum_{i=1}^m y_i^2 \right|}{m} \right|}{\left| \frac{2\left| \sum_{i=1}^m x_i \right|^2}{m} - 2\sum_{i=1}^m x_i^2 \right| \left| \frac{2\left| \sum_{i=1}^m y_i \right|^2}{m} - 2\sum_{i=1}^m y_i^2 \right| - \left| \frac{2\sum_{i=1}^m x_i \sum_{i=1}^m y_i}{m} - 2\sum_{i=1}^m x_i y_i \right|^2} \\
y_0 &= \frac{\left| \frac{2\sum_{i=1}^m x_i \sum_{i=1}^m y_i}{m} - 2\sum_{i=1}^m x_i y_i \right| \left| \sum_{i=1}^m x_i^3 + \sum_{i=1}^m x_i y_i^2 - \frac{\sum_{i=1}^m x_i \left| \sum_{i=1}^m x_i^2 + \sum_{i=1}^m y_i^2 \right|}{m} \right|}{\left| \frac{2\left| \sum_{i=1}^m x_i \right|^2}{m} - 2\sum_{i=1}^m x_i^2 \right| \left| \frac{2\left| \sum_{i=1}^m y_i \right|^2}{m} - 2\sum_{i=1}^m y_i^2 \right| - \left| \frac{2\sum_{i=1}^m x_i \sum_{i=1}^m y_i}{m} - 2\sum_{i=1}^m x_i y_i \right|^2} \\
&\quad \frac{\left| \frac{2\left| \sum_{i=1}^m x_i \right|^2}{m} - 2\sum_{i=1}^m x_i^2 \right| \left| \sum_{i=1}^m y_i^3 + \sum_{i=1}^m x_i^2 y_i - \frac{\sum_{i=1}^m y_i \left| \sum_{i=1}^m x_i^2 + \sum_{i=1}^m y_i^2 \right|}{m} \right|}{\left| \frac{2\left| \sum_{i=1}^m x_i \right|^2}{m} - 2\sum_{i=1}^m x_i^2 \right| \left| \frac{2\left| \sum_{i=1}^m y_i \right|^2}{m} - 2\sum_{i=1}^m y_i^2 \right| - \left| \frac{2\sum_{i=1}^m x_i \sum_{i=1}^m y_i}{m} - 2\sum_{i=1}^m x_i y_i \right|^2} \\
R &= \sum_{i=1}^m \frac{(x_i - x_0)^2 + (y_i - y_0)^2}{m}, \\
\delta &= \frac{\sum_{i=1}^m \left| \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} - R \right|}{m}, \\
\sigma_{\max} &= \max \left\{ \left| \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2} - R \right| \right\}
\end{aligned}$$

3.2 识别圆弧与圆

在单义域邻接图中, 从顶点中选取圆弧域, 作为种子圆弧, 以此为起始点, 遍历得其邻接域(或邻接域的邻接域), 如果属于同一圆, 则种子圆弧生长, 并继续遍历, 否则终止, 即可得一弧, 如果首末闭合, 则得一圆. 圆弧及圆的识别算法如下所述:

- (1) 从单义域邻接图的顶点集中提取弧形域, 作为种子圆弧, 设为当前域;
- (2) 如果当前域无邻接域, 则转到(4);
- (3) 取当前域的邻接域 i, i 从 1 到 n, n 为邻接域总数,

{ 如果邻接域 i 与种子圆弧同圆, 则种子圆弧生长, 并设邻接域 i 为当前域, 同时返回(2);
取邻接域 i 的邻接域 j, j 从 1 到 m, m 为邻接域 i 的邻接域总数

- { 如果邻接域 j 与种子圆弧同圆, 则种子圆弧生长, 并设邻接域 j 为当前域, 同时返回(2);
}
(4) 如果该路径为回路, 则得一圆, 否则得一圆弧, 计算其几何数据, 获得矢量表达.

本算法不仅能够识别单独的圆弧和圆, 而且还兼顾圆弧(圆)与线段、圆弧(圆)的相交和相切等情况. 在图 4 中, 原始图象(a)经过图文分离和粗细分离, 包含圆和多种圆弧; (b) 对图象多义域进行分裂; (c) 为图象的单义域表达; (d) 为矢量化图形. 由结果可以看出, 该法识别能力很强, 能处理多种圆弧和圆.

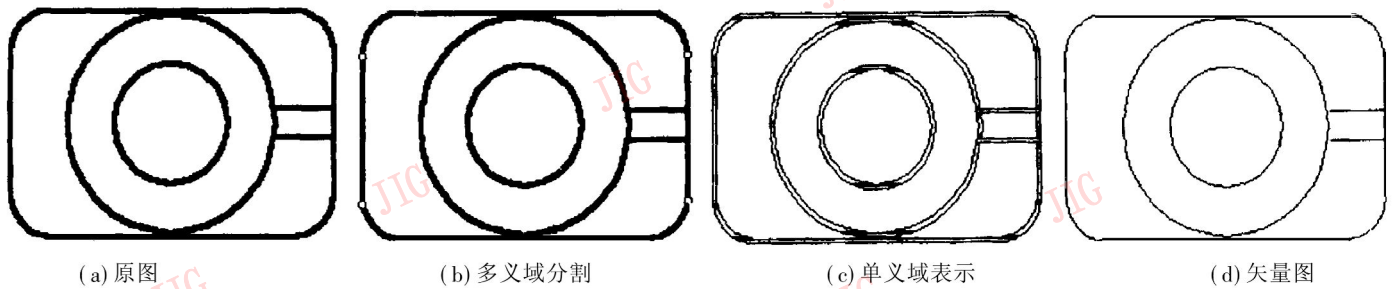


图4 圆弧和圆的识别示例

4 结束语

本文介绍了一种基于单义域邻接图的圆弧与圆识别算法,与以前方法相比,无需线段逼近和模式匹配,可以直接提取圆弧,对短小圆弧也有较高识别率,对圆弧(圆)与线段、圆弧(圆)的相交和相切等也同样适用,该方法已被应用于我们开发的工程图纸扫描图象识别与理解系统之中,效果较好.但是仍需进一步完善,研究各种复杂情况,以提高识别范围.单义域邻接图的描述模型也可用于线段完整识别及字符识别等方面.

参考文献

1 Vijay Nagasamy, Noshir A Langrana. Engineering drawing processing and vectorization system. *Computer Vision, Graphics and Image Processing*, 1990, 49: 379~ 397.



张习文 1971年生,博士生.主要研究方向为模式识别和图纸理解.

- 2 Dov Dori. Vector-based arc segmentation in the machine drawing understanding system environment. *IEEE Transactions on Pattern and Machine Intelligence*, 1995, 17(11): 1057~ 1068.
- 3 周辉. 扫描工程图纸识别输入处理与联机手绘图形输入技术的研究[学位论文]. 大连:大连理工大学,1998.
- 4 Han C C, Fan K C. Skeleton generation of engineering drawings via contour matching. *Pattern Recognition*, 1994, 27(2): 261~ 275.
- 5 李伟青,彭群生.一种基于模式的圆的识别算法. *软件学报*, 1999, 10(2): 129~ 132.
- 6 Zeno S Di, Cinque L, Levialdi S. Run-based algorithms for binary image analysis and processing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1996, 18(1): 83~ 89.
- 7 郑南宁著. *计算机视觉与模式识别*. 北京:国防工业出版社, 1998, 3: 160~ 168.
- 8 吴仲科,焦海星等.一种线段和圆弧的逼近方法及其在工程图纸矢量化中的应用. *计算机辅助设计与图形学学报*, 1998, 10(4): 328~ 332.

欧宗瑛 1936年生,教授,博士生导师.主要从事智能CAD、图象理解、人工智能、分形和小波等研究.