

一种交互式识别几何图形的简易方法

杨 凯^{1,2}, 华庆一^{1,2}, 陈新胜^{1,2}

(1. 西北大学 信息学院 计算机科学系, 陕西 西安 710069;

2. 中科院 计算机科学国家重点实验室, 北京 100080)

摘 要:介绍了一种识别多笔画几何草图简单方法。它利用图形的暂时邻接关系和全局几何特征来识别一些简单的几何图形其中有实线和虚线类型, 选择和删除手势。随着图形的旋转和尺度的变化这些几何特征(凸包、规则多边形的最小面积、周长和面积的数量比)是不变的。尽管多笔画的方法在选择合适的超时时间上存在问题, 但这种方法可以得到令人满意的识别率。虽然在简单性和健壮性上更侧重于前者, 但该方法已被证实适合交互式应用。

关键词:符号识别; 工具和技术; 模糊逻辑; 多笔画图形

中图分类号: TP391.41

文献标识码: A

文章编号: 1673-629X(2008)03-0021-03

A Simple Approach to Recognise Geometric Shapes Interactively

YANG Kai^{1,2}, HUA Qing-yi^{1,2}, CHEN Xin-sheng^{1,2}

(1. Department of Computer Science, Institute of Information Science, Northwest University, Xi'an 710069, China;

2. State Key Laboratory of Computer Science, Institute of Software, Chinese Academy of Science, Beijing 100080, China)

Abstract: Presents a simple method to recognise multi-strokes sketches of geometric shapes. It uses temporal adjacency and global properties of figures to recognise a simple vocabulary of geometric shapes including solid and dashed line styles, selection and deleting gestures. The geometric features used (convex hull, smallest-area regular polygons, perimeter and area scalar ratios) are invariant with rotation and scale of figures. Found the method very usable with acceptable recognition rates although the multi-stroke approach poses problems in choosing appropriate values for time-outs.

Key words: symbol recognition; tools and techniques; fuzzy logic; multi stroke shapes

0 引言

随着个人数字助理(PDAs)的普遍, 采用手绘几何图形的识别有了新的关注方向。传统的离线识别几何图形方法一直集中在未被处理的分类性能上, 在线系统则产生了一系列不同的问题。首先, CAD产生的几何图形可以通过人们画草图或者外围的计算机终端来准确地输入。识别这些图形主要是处理噪音, 这些噪音是处理过程之外的图形本身引起的, 例如拷贝的退化和光学还原技术不成熟。但在线应用上, 噪音是信息收集过程中所固有的, 图形常常因媒体、操作员和处理过程的局限性被不充分地勾画, 产生未完成和不明晰的图形, 这些图形甚至人们自身都很难辨别。与离

线算法另一个不同是输入数据不是图像位图而是序列化的点。同样, 在线识别应该处理在手绘过程中产生的许多变化——所有的几何图形实现一笔绘制成功不大合理, 尤其是想要识别的虚线图形。

想要强调的是人们主要关注的是能够识别的预定义的符号, 也就是用户首先打算绘制的图形。同样地, 目的不是介绍这种极其简单的方法, 而是为了说明此为用户能交互式地构造图表和其他结构化图形提供一个工具。

1 识别算法

这种识别算法是主要基于三个思想: 首先, 它完全地依赖于从输入图形中提取的全局几何特征。因仅仅是为了识别几何实体, 识别主要是依靠几何信息; 第二, 为了加强识别性能, 通过一棵决策树用特有的规则来过滤掉不期望的图形; 第三, 为了克服草图的不确定和不精确性, 使用模糊逻辑^[1]来识别图形, 因此可以自然地处理模糊性。

收稿日期: 2007-06-09

基金项目: 中国科学院计算机科学国家重点实验室开放课题基金(SYSKF0704); 国家 863 高技术发展研究计划(2006AA01Z328)

作者简介: 杨 凯(1980-), 男, 陕西西安人, 硕士研究生, 研究方向为人机交互; 华庆一, 教授, 博士生导师, 研究方向为人机交互、软件方法学。

面积。这个比 $(Ach * Ach) / (Aer * Alt)$ 允许辨别椭圆, 对于椭圆该比的值要大于菱形该比的值(见图 3f)。因为菱形面积大概是三角形面积的两倍。若面积比 (Alt/Aer) 的比为 50% 那么这个对象则是菱形, 否则该图形则不能被识别。为了保持上面的关系, 识别菱形的时候必须对图形有约束的限定(例如不被旋转)。

该方法用 $((Pch * NStrks) / Len)$ 的比来辨别虚线和实线, 这里的 $NStrks$ 是草图的笔画数。因虚线图形由许多笔画组成, 每个笔画的长度都非常小, 虚线图形这个比的值比封闭图形的这个比的值要大。

为了选择最佳的模糊集来描述决策树上的图形, 三个小组分别用实线画每个图形十次, 虚线画每个图形五次, Joaquim A. Jorge and Manuel J. Fonseca 用这些图形来计算出一个训练集。Joaquim A. Jorge and Manuel J. Fonseca 用这些值来描述决策树(见图 3a~f)的模糊集。图 2 为每一次过滤显示不同的模糊集, 每次识别的图形和不同过滤器的应用顺序。模糊集得自于试验数据所产生的极端百分比, 降低错误的识别率和增加识别的种类, 举例来说增加圆比例因子(thinness)的最大值防止把某些椭圆误认为圆, 或者降低直线比例因子(thinness)的最小值防止薄的图形(例如矩形)误认为是直线。

1.2 试验计划

为了评估这个识别算法, Manuel J. Fonseca 等人组织了 7 个小组每个图形画 15 次, 其中实线 10 次, 虚线 5 次。每个小组都按照如图 4A 所示来画图, 产生 934 个图形。小组被告知试验是为了测试识别率, 因此他们不可以画不自然的图形。图 4B 显示了其中一个小组画的图形, 图 4C 显示的是对应的识别图形。

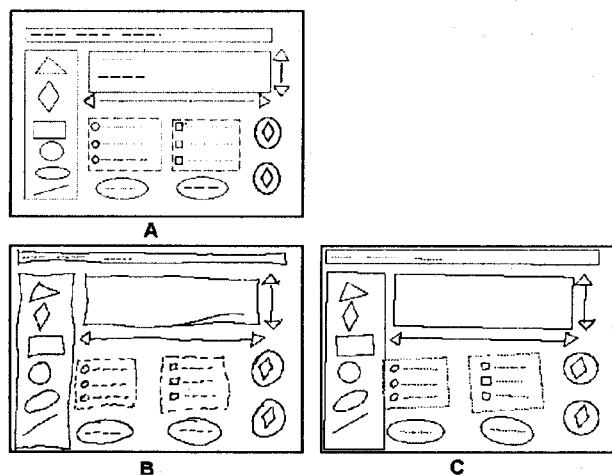


图 4 试验图

用 Wacom LCD digitizing tablet PL-300 和 cord-

less stylus 来画图。其中三个小组对笔和写字板很熟悉, 另外的小组没使用过数字化的写字板。Manuel J. Fonseca 等人把此种识别大致给用户描述了一下, 其中包括可以识别的图形。也告知他们可以识别多笔画的图形, 识别与图形的旋转和尺度是无关的, 另外还告知关于超时的概念。为了使新手小组能很好地结合笔和写字板, 他们有一个短时间的练习机会。在这期间识别被关闭以避免妨碍数据搜集。

实验结果表明草图的识别成功率达到 91%。图 5 交叉矩阵所示为每个图形的识别率和该图形错误的识别率。根据交叉矩阵粗略地分析圆很容易地被识别成椭圆, 但这种错误识别是可以接受的, 是人们的直觉行为。如果认为圆和趋近于圆的椭圆是同一种图形类, 识别率则增加为 93%。第二, 椭圆和菱形容易混淆, 所以椭圆识别率低。第三, 菱形的识别率也低, 因菱形与矩形和椭圆容易混淆。

		Recognised												
		Line S	Line D	Triangle S	Triangle D	Rectangle S	Rectangle D	Circle S	Circle D	Ellipse S	Ellipse D	Diamond S	Diamond D	Delete
Drawn	Line S	100%												
	Line D	3%	90%				5%							2%
	Triangle S			97%								2%		1%
	Triangle D				91%								9%	
	Rectangle S					96%				1%		3%		
	Rectangle D						96%				2%	2%		
	Circle S							73%		23%		3%		
	Circle D								83%		11%		6%	
	Ellipse S									85%		15%		
	Ellipse D										91%		9%	
	Diamond S					7%		1%		4%		88%		
	Diamond D						3%				9%		89%	
	Delete													100%

图 5 交叉矩阵

如果把识别出的 3 种图形中有一个是正确的认为是成功的分类, 识别率就会从 91% 增加到 94%。如果把圆和趋近于圆的椭圆看成同一种图形类, 那么识别率将会增加到 97%。

2 结论

文中描述了一种简单且快速识别基本几何图形的方法^[6]。这里的用意主要是提供用笔完成交互的一种手段, 而不是准确简单地得到想要得到的图形。经验表明交互式识别输入的识别效率是有效的, 是可行的。不管这个图形有多少笔画, 利用暂时的邻接来识别图形需要一个超时是迄今为止被共识的最不自然的约束和限制。

高识别率和识别的快速性特征使其在交互式应用中非常有用。通过更好的计算几何算法, 更自然分割输入手段应该被设计出来, 提高识别率和方法的健壮

(下转第 27 页)

由于查字典的时间可看作为常数时间 $\log(N)$, N 为每个字为首字成词的平均词数, 平均切分时间为 $4\log(N)$, 所以总的算法复杂度为 $O(n)$ 。

2 实验分析

用纯 C 语言实现该算法与反向最大匹配切分算法, 并下载了中科院免费版的 ICTCLAS, 将三种算法的函数封装成同样接口的 DLL 文件, 在 Java 中调用测试, 测试机器为 AMD2600 + /1024M PC, 首先进行速度测试, 对从《人民日报》上随机摘取的 55 195kB 语料进行切分速度测试, 测试结果如表 1。接着, 对国内最权威的著名美食网站 POCO 上随机挑选一百篇文章进行切分准确性测试, 平均文章长度为 1 000 汉字, 测试结果如表 2。

表 1 切分速度对比测试

算法	反向最大匹配切分	ICTCLAS	专有名词优先切分
总耗费时间	18.322s	1839.926s	29.932s
平均切分速度	3 012 kB/s	30kB/s	1844 kB/s
平均每秒切分字数	1 500 000 字/s	15 000 字/s	922 000 字/s

表 2 切分准确率对比测试

算法	反向最大匹配切分	ICTCLAS	专有名词优先切分
准确率	82%	88%	96%

最后, 对 152 930 条数据进行索引测试, 测试了使用本分词组件前后索引文件大小的变化, 并抽取了十个词进行检索测试, 得出前后不同的平均检索时间, 测试结果如表 3。

表 3 Lucene 索引检索测试

项目	不使用分词组件	使用本分词组件
索引文件大小	26 260kB	19 510kB
检索速度	0.10s	0.07s

从上面的实验数据证明, 本分词算法的速度极快, 准确率较高, 适应对时间响应快应用要求。与 Lucene 配合使用后, 能使 Lucene 的索引文件体积减少 30%,

检索速度提高 20%, 显著提高 Lucene 的索引检索效率。另外, 本组件运行时占用内存少, 支持多用户同时使用, 满足基于 Lucene 的搜索引擎的应用需要。

3 结束语

该分词方法的主要优势在于速度极快, 能达到 92 万字每秒的切分速度。对于商旅信息分词的准确率可达到 96%。不足之处在于它不能进行自动的新词识别。解决的办法是, 根据实际需要, 人工录入, 完善词库, 这样同时可有效控制词库的大小。本方法也不能进行真正的歧义识别, 真正的歧义识别必须根据中文的语法和词性分析, 这样一难以覆盖全部语法, 二耗费过多时间和系统资源, 得不偿失, 有待进一步研究突破。因此, 本算法有效满足主题信息处理的应用要求。

参考文献:

- [1] 朱德熙. 语法讲义[M]. 北京: 商务印书馆, 1982.
- [2] 孙茂松, 邹嘉彦. 汉语自动分词中的若干理论问题[J]. 语言文字应用, 1995, 4(4): 40-47.
- [3] Palmer D. A trainable rule-based algorithm for word segmentation[C]// The 35th Annual Meeting of the Association for Computational Linguistics (ACL'97). Madrid: [s. n.], 1997.
- [4] Choi A, Cheng C H, Ko Y L. Word extraction from Chinese documents by occurrence counts[C]//1988 International Conference on Computer Processing of Chinese and Oriental Languages. Toronto, Canada: [s. n.], 1988: 488-491.
- [5] Fan C K, Tsai W H. Automatic word identification in Chinese sentences by the relaxation technique[J]. Computer Processing of Chinese and Oriental Languages, 1988, 4(1): 33-56.
- [6] 孙茂松, 左正平, 黄昌宁. 汉语自动分词词典机制的实验研究[J]. 中文信息学报, 2000, 14(1): 1-6.
- [7] 刘群, 张华平, 俞鸿魁, 等. 基于层叠隐马模型的汉语词法分析[J]. 计算机研究与发展, 2004, 41(8): 1421-1429.

(上接第 23 页)

性, 以便识别更广泛, 可用性更好。

参考文献:

- [1] Bezdek J C, Pal S K. Fuzzy Models for Pattern Recognition [M]. [s. l.]: IEEE Press, 1992.
- [2] Kimura T D, Apté A, Van Vo. Recognizing Multistroke Geometric Shapes: An Experimental Evaluation[C]//In Proceedings of the ACM Conference on User Interface and Software Technology (UIST'93). Atlanta, GA: ACM Press, 1993: 121-128.

- [3] Ulgen A F F, Akamatsu N. Geometric shape recognition with fuzzy filtered input to a backpropagation neural network[J]. IEEE Trans. Inf. Syst., 1995, E788-D(2): 174-183.
- [4] O'Rourke J. Computational geometry in C[M]. 2nd edition. Cambridge: Cambridge University Press, 1998.
- [5] Freeman H, Shapira R. Determining the minimum-area enclosing rectangle for an arbitrary closed curve[J]. Communications of the ACM, 1975, 18(7): 409-413.
- [6] Fonseca M J, Jorge J A. Using Fuzzy Logic to Recognize Geometric Shapes Interactively[C]//Proceedings of the 9th Int. Conference on Fuzzy Systems (FUZZ-IEEE) 2000. San Antonio, USA: [s. n.], 2000.