

# 基于上下文的在线草图识别方法<sup>\*</sup>

谢强 冯桂焕 孙正兴

(南京大学计算机软件新技术国家重点实验室 南京 210093)

(南京大学计算机科学与技术系 南京 210093)

**摘要** 支持用户自由连贯地绘制草图是在线草图识别研究的目标之一。本文提出一种使用空间关系等上下文和贝叶斯分类器进行草图识别的方法,其主要特点包括两个方面:第一,使用笔画空间关系进行自动笔画成组,可以在不打扰用户绘图的情况下完成候选符号的选取;第二,利用贝叶斯分类器实现在线草图识别,可以解决识别方法的领域相关性及画法敏感性问题。实验验证了方法的有效性。

**关键词** 在线草图识别,空间关系,上下文,贝叶斯分类器,自动笔画成组

## Online Sketch Recognition Using Contexts

XIE Qiang FENG Gui-Huan SUN Zheng-Xing

(State Key Lab for Novel Software Technology, Nanjing University, Nanjing 210093)

(Department of Computer Science and Technology, Nanjing University, Nanjing 210093)

**Abstract** The online sketch recognition for continuous drawing is one of the most important researches. This paper presents an online sketch recognition method which uses spatial relationships and Bayesian classifier. It consists of two aspects: first, spatial relationships between strokes are used to do stroke grouping, and the candidate symbols can be automatically selected without disturbing the users; second, a Bayesian classifier is designed to implement the domain-independent online sketch recognition, it can recognize the inputting symbols insensitive to drawing styles of different users with highly precision. The experiment results prove the effectiveness of our method.

**Keywords** Online sketch recognition, Spatial relationships, Context, Bayesian classifier, Automatic stroke grouping

## 1 引言

手绘草图是人类一种自然而直接的表达思想和进行交流的方式,可以有效地记录和传递各种信息,可以帮助用户将注意力集中在问题解决上<sup>[1~3]</sup>。在线草图识别方法的目标就是在不干扰用户绘图过程前提下对用户输入进行分析处理,并正确捕捉用户输入意图。已有在线草图识别方法<sup>[4~6]</sup>大多面向单个图形符号识别(Symbol Recognition),只能支持用户逐个输入图形。然而,大多数情况下,用户需要自由连贯地输入图形,因此,支持用户自由连贯地绘制草图是在线草图识别研究需要解决的问题。

上述问题的解决就是要在不中断用户绘图过程的情况下,自动完成笔画成组和符号识别。笔画成组需要解决的问题是:在用户没有明确指明绘制草图对象的起始条件下,如何将用户绘制的笔画聚合成一个草图对象。已有工作证明,将笔画聚集到一个领域有意义的集合并不是一个容易的问题<sup>[3,8]</sup>。目前已经有许多的方法完成笔画的成组工作,例如 Rubin<sup>[9]</sup>设计了一种草图手势的方法来解决成组问题并依次帮助识别,这种方法的特点在于它要求每个草图对象都要一笔完成,即每个成组和识别单元都是一笔勾成的手势。这种方法的优点在于它的准确性,缺点是过于限制了用户的绘图方式。此外 Long<sup>[10,11]</sup>对手势的方式做了改进。QuickSet<sup>[12]</sup>

突破了单笔画的限制,通过笔画间的停顿和多通道信息(如语音)的帮助完成对笔画的成组;Wilcox<sup>[13]</sup>使用了一种时间阈值的方法完成笔画的分组:当停顿时间超过阈值,则之前笔画归为一组。使用时间阈值的缺点在于其难于确定,对用户限制仍然很大。除了使用时间阈值方法,还可以利用笔画间的空间关系进行成组,如 Viola<sup>[14]</sup>使用了一种纯粹的空间方法对手绘符号进行成组和识别,不需要限制用户绘图的笔画顺序,甚至没有限制符号间的关系。他们的方法用于文本图形取得了很好的效果,但对于其他的领域图形(如电路图)则有很大限制。Wan 和 Watt<sup>[3]</sup>则将空间和时间相结合的方法应用到 Pocket PC 上。此外,还有启发式<sup>[15,16]</sup>等其他的成组方法。

符号识别的一般解决方法是将候选符号笔画分割成基本的集合图元(如线、弧或者圆),然后根据某种特定结构方式进行计算,得到其内在属性和相互关系,并依此将其识别成领域符号。根据不同的识别特征和分类器,现在已经出现了很多的识别策略。其中有些受制于草图的大小、方向或者符号的笔画数<sup>[9]</sup>;有些则强调其对缩放和旋转的不变性<sup>[17]</sup>,还有些对训练样本有要求<sup>[15,16]</sup>。虽然已有工作在某些方面取得了不少的成就,但大多对符号上下文信息的利用不足,即大部分方法不能很好地适用于自由连贯的草图绘制。

本文提出一种利用空间关系和贝叶斯分类器完成基于上

<sup>\*</sup> 本文得到国家自然科学基金项目(项目编号:69903006 和 60373065)和新世纪优秀人才支持计划(项目编号:NCET-04-0460)资助。谢强 硕士生,研究方向为智能人机交互自动笔画成组;冯桂焕 博士生,主要研究方向为智能人机交互和草图理解;孙正兴 教授,博士生导师,主要研究领域:多媒体计算、计算机视觉和感知人机交互。

下文的在线草图识别方法,它允许用户自然、连贯地进行绘图。其中主要讨论了两个问题:第一,使用笔画的空间关系完成笔画的自动成组,相对于使用时间阈值和笔画约束的方法,使用空间关系更符合人们理解草图的过程,且对用户的绘图限制比较小,可以更好地保证用户自由、连贯地绘图;第二,本文设计了一个贝叶斯分类器,完成对样本的训练以及对成组笔画的识别。在贝叶斯理论的支持下,快速、高效地向用户提供识别结果。

## 2 基于上下文的草图识别方法

在诸如电路图及 UML 图设计等应用中,各图形符号相互独立,没有重叠;且用户绘图过程中,在完成某一符号的绘制前,不能绘制其他的符号。大量调查表明,这种限制符合人们日常的绘图习惯,对于某一符号的笔画顺序没有特别要求。为了便于说明,本文选取了电路图作为实验背景。其组成符号如图 1。

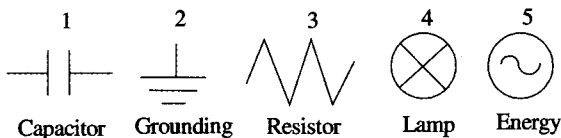


图 1 电路符号示意

我们设计了如图 2 所示的在线草图识别流程。当用户输入新笔画后,首先进行笔画成组过程,即将用户输入的笔画聚集成用于识别的候选笔画集合。方法是用笔画间的空间关系来判断是否新笔画属于符号。如果新笔画属于符号,那么继续等待新笔画;如果判断新笔画不属于符号,则认为已经找到完整的符号,此时将其传送给识别过程,进行符号识别。

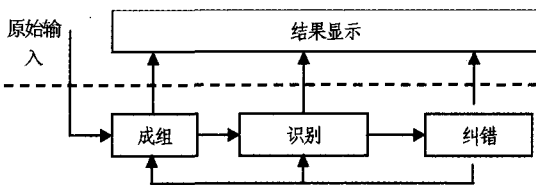


图 2 在线草图识别过程框架

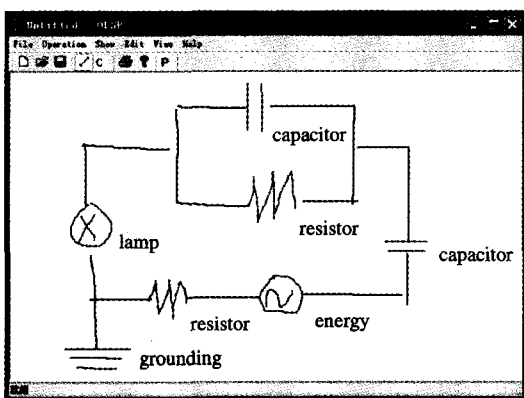


图 3 草图识别效果图

符号识别过程以成组过程得到的笔画集合作为输入,并找到领域中的最佳匹配对象。方法是用一个贝叶斯分类器完成笔画集的识别工作。当识别过程从成组过程得到一个笔画集合,首先计算该笔画集合的特征向量,然后与训练样本进行比较,从而确定最佳的识别结果。

此外,我们设计一个简单的纠错过程。当识别过程没有为笔画集找到可以匹配的分类标签时,需要将其返回到成组阶段,并重新进行分析。在这些过程中,用户可以选择查看中间过程。图 3 展示了一个完整的识别效果图。

## 3 使用空间关系进行自动笔画成组

自动笔画成组的目的是在不打断用户绘图的情况下,将草图笔画聚集成可用于识别的笔画集合。这是一个在线的过程。在本文方法中,我们使用原始笔画进行成组,因为在第 2 节中我们已经约定,在用户完成一个符号的绘制前,不会绘制其他的符号;并且,相比较于分割后的基本图元,原始笔画可以更好地反映笔画间的空间关系,提高效率并减少错误率。

使用笔画空间关系进行笔画成组更符合用户理解草图的过程。同一符号内的各笔画,其邻接程度相对于符号内外笔画邻接程度要紧密。在具体方法实现中,我们使用关联属性(RSP: Relative Sub-Parameter)和界限属性(BSP: Bounded Sub-Parameter)完成对笔画集搜索。不同的符号对不同的 RSP 和 BSP 函数敏感。针对电路图,分别设计 RSP 和 BSP。

### 3.1 关联属性(Relative Sub-Parameter)

RSP 用来确定何时开始搜索候选笔画集。主要的方法是通过判断当前笔画的几何特征以及它和前一笔画的组合几何特征,决定当前笔画或者当前笔画与前一笔画的组合是否属于某一领域符号。

针对电路图符号,我们定义 RSP 如公式(1)所示:

$$BSP = \langle ratio, \frac{Cdist}{L2}, \frac{S2}{L2}, \frac{L1}{L2}, ldiff \rangle \quad (1)$$

这是一个五维向量,其中每个向量表示意义如下:ratio 表示当前笔画长度与闭包矩形周长的比率,该值可用于判断笔画是否为圆或密集折线等;Cdist 表示当前笔画与前一笔画的闭包矩形的中心点距离;L1 表示当前笔画和前一笔画中较长的长度;L2 表示当前笔画与前一笔画闭包矩形长宽中的较长者;S2 表示当前笔画与前一笔画闭包矩形长宽中的较短者;前三维向量需要考虑当前笔画与前一笔画的各种几何关系,如笔画长度、闭包矩形等等,通常用于判断向电阻、接地等符号。

ldiff 向量和前三维向量不同,它表示的是当前笔画闭包矩形、前一笔画闭包矩形以及两笔画闭包矩形的差异较小者。该向量在判断向电灯、电源等含有封闭图形的符号时比较有用。例如图 4,添加笔画 2 后,1、2 共同的闭包矩形和 1 的闭包矩形相同,依此判定 1、2 属于同一符号。

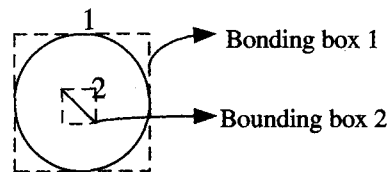


图 4 封闭图形示例

### 3.2 界限属性(Bounded Sub-Parameter)

BSP 在通过 RSP 确定符号 S 存在后有两个用途:第一,如果之前还有笔画,通过 BSP 确定 S 的起始绘制笔画;第二,当用户新添加一笔笔画后,使用 BSP 确定新笔画是否属于 S,即确定 S 是否终结。

我们使用笔画聚合度的变化率来计算 BSP。首先,定义当前笔画聚合度 density,如公式(2):

$$density = \frac{length\_of\_strokes}{area\_of\_bounding\_box} \quad (2)$$

笔画的长度包括隐藏笔画,即如果不抬笔时,连接相邻笔画起始的线段长度。

然后,定义添加了新笔画之后的笔画聚合度  $density'$ ,如公式(3):

$$density' = \frac{length}{area \times chang\_of\_Cdist} \quad (3)$$

在式(3)中,  $chang\_of\_Cdist$ (定义见公式(4))表示闭包矩形中心点距离的变化率,对于符号内笔画来说,其变化不明显;对于符号外笔画,其加权作用明显。

$$chang\_of\_Cdist = \frac{new\_Cdist}{max\_Cdist} \quad (4)$$

$new\_Cdist$  表示新添加笔画与之前笔画集最后一笔笔画的闭包矩形中心点距离,  $max\_Cdist$  表示未添加新笔画前已有笔画闭包矩形中心点距离的最大值。

最后计算 BSP 如式(5):

$$BSP = \frac{density - density'}{density} \quad (5)$$

相同闭包矩形范围内,笔画的长度反映了笔画间的紧密程度;而相同数目笔画的闭包矩形大小则反映笔画离散程度的不同。如果添加本属于符号内的笔画,则其对长度的加权要大于对闭包矩形面积的加权,因此引起聚合度上升;反之,符号外笔画会导致笔画聚合度的显著下降。因此,当 BSP 计算值超出阈值时,则认为找到了符号的起始或结束。

### 3.3 笔画自动成组过程

如图 5 所示,我们使用 RSP 和 BSP 完成对接地符号的成组过程如下:

添加笔画 1 时,不构成符号;添加笔画 2 后,计算 1、2 的 RSP,根据其值判断二者不构成符号;添加笔画 3,计算 3、2 的 RSP,得到符号 S,因为之前有笔画 1,所以计算添加 1 后的 BSP,用于确定 1 是否属于符号 S,结果将 1 排出,即 2 为符号 S 的第一笔画;添加笔画 4 后,因为已经确定了 S 的起始,计算添加 4 后的三个笔画的 BSP,确定笔画 4 属于符号 S;最后添加笔画 5,在计算新的 BSP 后确定 5 不属于符号 S,即笔画 4 是 S 的结束笔画,所以找到符号 S(2,3,4)。

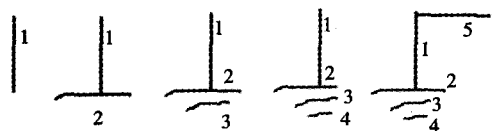


图 5 笔画成组举例

## 4 基于贝叶斯分类器的符号识别

笔画自动成组过程得到候选符号的笔画集合,并将其发送到识别过程进行识别。识别器的工作就是将笔画集合映射到领域内的相关符号。本文中,我们建立一个基于贝叶斯理论的分类器来完成识别工作。首先,将候选笔画集合中的笔画分割成基本的几何图元(直线、弧和圆)。在训练过程中,我们计算样本的特征值,并以此来得到符号的先验概率;在用户绘图过程中,根据和训练过程相同的方法将成组得到的笔画集合进行分割和计算特征,在计算相应后验概率后确定符号集对应的领域符号。符号笔画的先后顺序对识别没有影响,并且本方法对大小、方向不敏感,对于不均匀缩放也有很好的鲁棒性。

### 4.1 朴素贝叶斯分类器

朴素贝叶斯分类器基于贝叶斯理论。在统计学中,贝叶斯理论是讨论在已发生事件的条件下,特定事件发生概率的理论。它提供了一个有效利用先验概率计算后验概率的方法,其定义见公式(6):

$$P(H|X) = \frac{P(X|H)P(H)}{P(X)} \quad (6)$$

我们使用朴素贝叶斯分类器作为识别工具,主要包括两个部分的内容:训练和分类。

### 4.2 训练

在正式使用分类器之前需要对其进行训练。首先,针对每个符号,向用户收集若干样本;将这些样本进行分割,得到基本图元,然后计算一个六维几何特征向量。各特征向量分别为:笔画数目、直线段数目、弧线段数目、平行线数目、交叉线数目、封闭线数目。对于封闭线数目,我们可以从原始笔画中计算得到,在判断电灯等符号时,其作用明显。为了减少计算时的计算量,我们假设特征相互独立。虽然这种假设并不是一个完美的假设,但是 Domingos and Pazzani<sup>[18]</sup> 已经证明朴素贝叶斯分类器并不需要具有完备的特征独立性。

在确定符号训练样本的特征后,我们按照公式(7)构造一个高斯分布:

$$P(x_k | C_i) = \frac{1}{\sqrt{2\pi}\sigma_{C_i}} \exp\left(-\frac{(x_k - \mu_{C_i})^2}{2\sigma_{C_i}^2}\right) \quad (7)$$

$x_k$  表示特征值,  $C_i$  表示符号类别,  $\mu_{C_i}$ ,  $\sigma_{C_i}$  表示均值和方差。

对于公式(6),  $P(X)$  是恒定不变的,因此我们只需要计算  $P(x_k | C_i)$  和先验概率  $P(C_i)$  即可。

训练过程需要离线完成。我们在收集样本后即开始训练过程,得到用于分类比较的特征向量,然后将其用于在线的分类过程。

### 4.3 分类

当识别器从成组过程得到待识别的笔画集合后,第一步工作是将笔画集合中的笔画分割成基本图元,然后从中计算六维特征向量,该过程和训练过程一致。然后,将计算得到的特征向量和训练过程得到的特征向量分布进行比较。最后将最符合的符号标签赋给笔画集合,即完成对笔画集的识别。用公式表示即我们的目标是找到类别  $C_i$ , 对候选符号 S 具有最高的匹配概率,见公式(8):

$$P(X|C_i)P(C_i) > P(X|C_j)P(C_j) \quad (8)$$

$1 \leq j \leq m, j \neq i$ . 其中  $X$  是 S 的特征向量,  $m$  是类别的个数。

即如公式(9)所示,计算  $C^*$ :

$$C^* = \arg \max_j \prod_j P(x_j | C_i) P(C_i) \quad (9)$$

### 4.4 自动纠错

我们考虑的纠错过程发生在识别出现不确定值的时候。在该情况下,系统对从前一识别符号到当前的笔画集合重新启动成组和识别过程。

如果识别器不能为候选符号提供一个标签类别,那么我们认为在成组过程中出现了某些问题。此时,纠错过程会重新启动成组过程。再次的成组过程可以利用更多的信息来决定是否存在一个新的符号或者不存在符号,例如大致确定的符号集合,利用最可能的起始笔画开始新的符号搜索。

自动纠错是系统自动完成的信息反馈。通过纠错反馈:第一,系统重新完成笔画成组和符号识别,以尽可能准确地找到用户绘制的符号,或者确定并不存在符号;第二,通过反馈信息,可以对阈值进行选择性优化,用于训练用的样本对阈值

的确定并不是绝对完备的,我们可以通过纠错返回的信息对其进行调整。

## 5 实验和分析

我们选取电路图作为实验背景。图符的选取如图 1 所示。在实验过程中,我们请 13 位用户每人至少绘制了 20 幅草图,且每个符号至少出现 20 次。最后我们收集了 260 幅草图,符号总数为 1626 个。

在用户绘图的时候,我们约定用户在完成一个符号的绘制前不可以绘制其他的符号,即需要用户连续绘图;对符号的笔划顺序,是否连笔没有特别要求。

实验中从绘制后的草图挑选绘制比较规整的符号作为训练样本。本文认为用户对相同符号的绘制在一般情况下不会有太大的差别,因此选取的每个符号训练样本属于小样本,数目不定,但基本不超过 20 个。

### 5.1 实验一:不同用户的草图识别率实验

图 6 显示了针对不同用户的平均识别率。从图 6 中我们可以看到,13 位用户的平均识别率达到 84.3%,最高到 95% 左右,最低也在 78% 以上。这表明本文方法对不同用户具有很好的适应性。

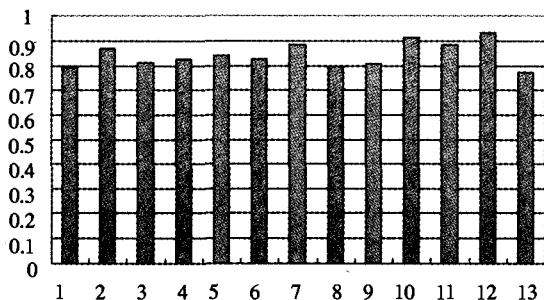


图 6 针对不同用户的识别率  
(横坐标:用户;纵坐标:识别精度)

### 5.2 实验二:每个符号的识别率实验

图 7 显示了 5 个电路符号的平均识别率。从中可以看出,不同的符号,其识别效果相差有可能比较大。如电容和电源相比,识别率分别为 60% 多和 90% 以上。其他符号识别率都比较高。表明,对于不同符号来讲,本文的方法有效,同时也有待改进的地方。

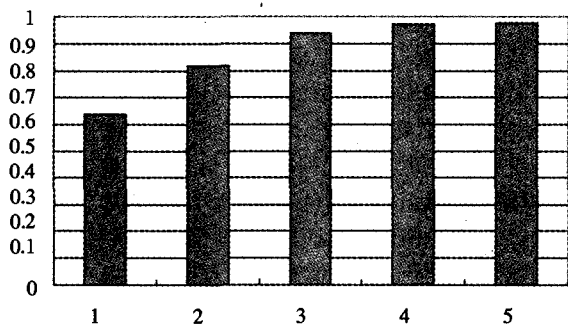


图 7 每个符号的识别结果  
(横坐标:符号标示;纵坐标:识别结果)

**结论** 本文提出了一种使用了空间关系和贝叶斯分类器进行基于上下文的草图识别方法。首先,方法使用笔画的空间关系进行自动笔画成组,在不干扰用户绘图的情况下,快速、高效、自动地从连续的笔画中抽取候选符号;其次,设计

并实现了朴素贝叶斯分类器,并用于候选符号的识别。该识别方法对符号笔画顺序不敏感,对用户不同的绘图方式也具有很好的鲁棒性。但是,由于我们所提出的 BSP 和 RSP 函数根据图符结构紧密程度来确定,对于结构松散符号,存在阈值确定不准确的情况,从而影响识别效果。我们将使用领域知识并将成组过程和识别过程结合,来进一步改进识别效果。

## 参考文献

- Alvarado C, Oltmans M, Davis R. A framework for multi-domain sketch recognition. In: Proc. of AAAI Spring Symposium on Sketch Understanding, 2002. 1~8
- Moran T P, Chiu P, van Melle W. Pen-based interaction techniques for organizing material on an electronic whiteboard [A]. In: Proceedings of the Annual ACM Symposium on User Interface Software and Technology, New York, 1997. 45~54
- Fish J, Scrivener S. Amplifying the mind's eye: Sketching and visual cognition [J]. Leonardo, 1990, 23(1): 117~126
- 孙建勇, 金翔宇, 彭彬彬, 等. 一种快速在线图形识别与规整化方法. 计算机科学, 2003, 30(2):172~176
- 蒋维, 张斌, 孙正兴. 基于自适应 HMM 的在线草图识别方法. 计算机科学, 2005, 32(5):185~189
- 张莉莎, 孙正兴, 周若鸿, 等. 基于图匹配的复杂草图识别方法. 计算机科学, 2004, 31(4):147~150
- Wan B, Watt S. An Interactive Mathematical Handwriting Recognizer for the Pocket PC. In: MathML Conference, 2002
- Yaeger L S, Webb B J, Lyon R F. Combining neural networks and context-driven search for online, printed handwriting recognition in the Newton. AI Magazine, 1996, 19(1): 73~89
- Rubin D. Specifying Gesture by Example. SIGGRAPH '91, 1991, 25(4):329~337
- Long A C Jr, Landay J A, Rowe L A, et al. Visual similarities of pen gesture. In: Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems, 2000
- Long A C. Quill: a Gesture Design Tool for Pen-based User Interface. Eecs Department, Computer Science Division, U C Berkeley, Berkeley, California, December 2001
- Cohen P, Johnston M, McGee D, et al. Quickset: Multimodal interaction for distributed applications. In: Proceedings of the Fifth ACM International Multimedia Conference (New York, NY), ACM Press, NY, 1997. 31~40
- Wilcox L D, Bill N S, Nitin S. Dynamite: A Dynamically Organized Ink and Audio Notebook. In: Proceedings of CHI'97, 1997. 186~193
- Shilman M, Viola P. Spatial Recognition and Grouping Text and Graphics. In: EUROGRAPHICS Workshop on Sketch-based Interface and Modeling, 2004
- Kara L, Stahovich T. Sim-U-Sketch: A Sketch-Based Interface for Simulink. In: AVI, 2004. 354~357
- Veselova O, Davis R. Perceptually Based Learning of Shape Descriptions for Sketch Recognition. In: The Nineteenth National Conference on Artificial Intelligence. (AAAI-04), July 2004
- Hse H, Newton A R. Sketched Symbol Recognition Using Zernike Moments. In: International Conference on Pattern Recognition, 2004
- Domingos P, Pazzani M. Beyond Independence: Conditions for the Optimality of the Simple Bayesian Classifier. In: Proc. Thirtieth International Conference on Machine Learning, 105~112