

使用几何特征与隐 Markov 模型的手绘笔画图元分解*

冯桂焕¹, 孙正兴¹⁺, Christian VIARD-GAUDIN²

¹(计算机软件新技术国家重点实验室(南京大学),江苏 南京 210093)

²(Laboratoire IRCCyN/UMR CNRS 6597, Ecole Polytechnique de l'Université de Nantes, France)

Stroke Fragmentation Using Geometry Features and Hidden Markov Model

FENG Gui-Huan¹, SUN Zheng-Xing¹⁺, Christian VIARD-GAUDIN²

¹(State Key Laboratory for Novel Software Technology (Nanjing University), Nanjing 210093, China)

²(Laboratoire IRCCyN/UMR CNRS 6597, Ecole Polytechnique de l'Université de Nantes, France)

+ Corresponding author: E-mail: szx@nju.edu.cn

Feng GH, Sun ZX, Viard-Gaudin C. Stroke fragmentation using geometry features and hidden Markov model. *Journal of Software*, 2009,20(1):1-10. <http://www.jos.org.cn/1000-9825/3383.htm>

Abstract: Stroke fragmentation is the core of the pen-based interaction. This paper presents a novel method of stroke fragmentation, which combines geometric features and HMM (hidden Markov model). Four geometric features are employed to describe the local geometry of strokes, and a HMM structure is designed to model the drawing context to describe the global. Furthermore, stroke data is compressed as much as possible with the least loss of information by means of global searching and the best matching algorithm. It can locate the segment point and judge the primitive type simultaneously with acceptable computation efficiency. Experimental results show the effectiveness of the proposed method.

Key words: pen-based interaction; freehand sketch; stroke fragmentation; hidden Markov model

摘要: 图元分解是笔交互的关键环节.提出一种将几何特征和隐马尔可夫模型 HMM(hidden Markov model)结合的笔画图元分解方法.该方法采用 4 种关键几何特征来描述笔画的局部几何信息,并通过 HMM 结构对绘制上下文的建模来描述笔画的全局几何特性,进而利用全局搜索与最佳匹配实现在尽可能不丢失笔画几何信息的前提下对数据进行最大限度的压缩,可在保证系统计算效率的前提下同时完成分割点的查找与图元类型的判定.实验结果表明了该方法的有效性.

关键词: 笔交互;手绘草图;图元分解;隐马尔可夫模型

中图法分类号: TP391 文献标识码: A

笔画图元分解是指将由点序列构成的笔画转化为物理含义更加明确的图元表示的过程^[1].图元获取是复

* Supported by the National Natural Science Foundation of China under Grant Nos.60721002, 60373065, 69903006 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2007AA01Z334 (国家高技术研究发展计划(863)); the Program for New Century Excellent Talents in University of the Ministry of Education of China under Grant No.NCET-04-0460 (国家教育部新世纪优秀人才资助计划)

Received 2007-12-19; Accepted 2008-05-19

杂手绘图形识别中必不可少的关键环节,一方面,相比于点序列,图元序列的数据量要小得多;另一方面,图元具有更加简明的物理含义,更利于用户把握图形的几何构成.此外,可靠的图元分解结果能够降低因抖动或采样误差而带来的模糊性,因而对提高草图识别系统的性能起到重要的作用^[2].但是,手绘图形的随意性以及用户对输入设备的不适应等等,都使得这一课题更具挑战性.

针对现有笔画分割工作所存在的能够处理的笔画类型有限及运算复杂度高等问题,本文提出了一种可对由任意方向直线和弧线所构成的笔画进行图元分解的方法,该方法综合利用了绘制的局部信息和全局信息,且选用的几何特征计算简单,因此,能够较为高效地对平滑过渡曲线进行准确分解.此外,文中所定义模型为领域无关,能够方便地拓展到其他相关领域.

本文第 1 节介绍相关工作.第 2 节给出隐马尔可夫模型的背景知识,并介绍图元类型与模型定义.第 3 节阐述特征选择及概率密度函数定义(probability density function,简称 PDF).笔画分解过程将在第 4 节详细论述.最后给出方法的性能评估结果,并对全文进行总结.

1 相关工作

笔画图元分解方法可分为 3 类.第 1 类是基于尺度空间(scaling space)的分解方法^[3,4],它将笔画以不同层次表示,并使用不同平滑方程进行平滑,最终选择最能表示用户意图的一层作为结果输出.尺度空间分解方法的优点是适合处理噪音比较大的笔画,缺点在于计算复杂度很高,所需时间也较长.第 2 类即为比较常用的基于曲率的分解方法^[2,5-7].该方法直接从输入笔画进行计算,并将曲率或速率的变化幅度作为分割点查找的主要依据,值小于设定阈值的点即被确认为分割点,不同方法间的区别在于各自使用的过滤方法不同.这类方法实现起来比较简单,且一般情况下都能够取得良好的效果.但是由于只关注绘制的局部上下文,因而容易产生过分割与弱分割的情况,同时,对于类似“J”、“S”等平滑过渡的曲线分解效果不佳.Paulson 等人^[8]在速率与曲率的基础上提出了两个新的特征,分别为方向极值点间的标准化距离(normalized distance between direction extremes,简称 NDDE)以及方向改变率(direction change ratio,简称 DCR),实验结果表明该方法具有较高的识别精度,但目前仅能对 8 种特定笔画类型进行处理.第 3 类方法可视对笔画层建模的方法,由于充分利用了绘制的上下文信息,因此这类方法精度较高.Cates 等人^[9]将每个笔画表示为一个马尔可夫随机场(Markov random field,简称 MRF),场中每个节点表示可能的拐点区域或两个拐点之间的片段.候选拐点特指那些方向改变大于某个阈值的点.由于候选拐点的筛选基于阈值定义,因而与第 2 类方法类似,存在对于平滑过渡曲线丢失拐点的情况.同时,对每一个笔画都需要重新建模,因此运算复杂度较高.Hse 等人^[1]使用动态规划算法将分割点查找与图元类型判定相结合,方法为每个目标图符建立模板,模板中预先定义图符的笔画分解结果,并通过查找输入与模板的最佳匹配实现笔画分解的目的.本文方法与之类似,但模型一经建立无须额外定义模板,且能够处理任意由直线和圆弧构成的笔画,因此,方法的可扩展性明显优于前者.

隐马尔可夫模型(hidden Markov model,简称 HMM)于 20 世纪 60 年代后期被提出之后,因其非常适合于时序相关模型的构造,因而在语音识别^[10]、Gesture 识别^[11]、手写文档理解^[12]等诸多领域产生了许多应用成果.但是,利用草图绘制的时序相关性而将 HMM 应用到草图识别领域的研究却很少,我们认为导致这一现象的主要原因是手绘草图并不像手写体那样具有相对稳定的时序信息.Sun 等人^[13]提出了用于描述笔画的代表性几何特征和动态特征,并据此构建多笔画图符模型用于手绘形状识别;Sezgin 等人^[14]在图元层时序信息之外还使用了图符层时序信息,因此能够对多个图符连续绘制的情况加以识别.但是,由于不同人对同一图符的绘制习惯不同,因此,这二者均需要为不同人定义不同的模型,其应用领域与应用范围存在限制.本文所提出的模型单纯从图元的几何构成着手,方法对用户的绘制习惯没有任何依赖.裴继红等人^[15]提出了一种自适应隐马尔可夫模型对手绘图形进行压缩,但其工作更像是一种对笔画的多边形拟合方法,因而无论从数据压缩比还是结果的物理含义方面,相比于笔画图元分解都存在劣势.

2 HMM 模型与状态定义

2.1 HMM 背景知识

离散 HMM 模型通常表示为一个五元组 $\lambda = \{A, B, \pi, N, M\}$. 矩阵 A 记录状态间的转移概率, B 为各状态所表征的观察向量, 模型初始状态分布以向量 π 表示, M 和 N 分别表示一阶 HMM 的状态数目和每个状态可能对应的观察向量. 又因 M 和 N 已经隐含在矩阵 A 和 B 之中, 因此可简记为 $\lambda = \{A, B, \pi\}$.

观察向量的量化过程可能导致模型退化, 因此很多情况下需要使用连续观察密度的 HMM 来提高系统性能. 同时, 为保证模型参数能够以一致的方式得以重估(reestimate), 模型概率密度函数的选择需满足若干限制. 较为常用的概率密度函数表示如下:

$$b_j(O) = \sum_{m=1}^M c_{jm} \Pi[O, \mu_{jm}, U_{jm}], 1 \leq j \leq N \quad (1)$$

这里, O 为观察向量, c_{jm} 代表状态 j 的第 m 个高斯分布的混合系数, Π 可以为任意具有对数凹性或椭圆对称性的密度函数, 如高斯分布等. μ_{jm} 和 U_{jm} 分别为状态 j 的第 m 个混合分布的均值和方差, 且满足 $\sum_{m=1}^M c_{jm} = 1$, $1 \leq j \leq N$, 且 $c_{jm} \geq 0$. 以上约束的目的是保证公式(1)成立, 即 $b_j(O)$ 为合法密度.

$$\int_{-\infty}^{\infty} b_j(x) dx = 1, 1 \leq j \leq N \quad (2)$$

在使用 HMM 解决实际问题的过程当中, 涉及到以下 3 个关键技术:

- (1) 使用模型进行识别, 即判断模型产生特定观察序列 O 的概率 $P(O|\lambda)$. 通常使用前向-后向算法(forward-backward algorithm);
- (2) 分析模型生成给定观察序列 O 的最佳状态序列 $Q = q_1, q_2, \dots, q_T$. 即著名的 Viterbi 算法;
- (3) 模型的构建或者学习. 使用 Baum-Welch 算法对模型参数进行调整, 使得给定观察序列的产生概率 $P(O|\lambda)$ 最大.

2.2 基本概念

定义 1. 笔画. 这是指采样设备由落笔到抬笔间电子笔所行径的轨迹, 可以点序列形式表示为 $Stroke = \langle p_0, p_1, p_2, \dots, p_n \rangle$, 其中 $p_i = (x_i, y_i)$ 为时刻 i 设备的采样坐标.

定义 2. 图元. 笔画点序列的子集, 且具有更为明确的物理意义. 文中图元类型集合为 $Prim = \{Line, Arc\}$.

定义 3. 笔画图元分解. 将笔画分解为一系列图元集合的过程. 设分解结果为 $Stroke = \{P_1, P_2, \dots, P_m\}$, $P_k \in Prim$, 且 $P_k = \langle p_{k0}, p_{k1}, \dots, p_{kn} \rangle$, 则存在 $\cup P_k = Stroke, P_k \cap P_{k+1} = \{p_{kn}\}$.

2.3 图元类型定义

模型定义与图元类型选择密切相关. 对于大多数应用领域, 如电路图绘制, 直线与圆弧的描述能力已经足够了. 这里, 圆弧既可以是封闭的圆, 也可以只是圆的一部分. 考虑到直线方向的不同, 我们将其划分为 8 种不同的图元类型; 对于圆弧, 根据旋转方向的不同分为顺时针与逆时针方向, 同时与直线相对应, 可以将圆划分为 8 个子片断(如图 1 所示). 在模型设计时, 对应于 $q^0 - q^{23}$ 的状态被称为基本状态, 用来描述构成笔画的基本图元. 状态定义有多种不同的方式, 这样划分的好处是尽管状态各不相同, 但是它们相互之间存在关联. 例如, q^0 和 q^{18} 除曲率信息不同以外, 方向信息是一致的, 这为计算与匹配降低了复杂度. 同时, 为了使用模型实现图元分解, 还需对局部信息进行描述, 且该信息应与方向及曲率信息相关联. 下文中我们还将引入一些新的模型状态($q^{24} - q^{32}$), 用来对基本状态之间的过渡进行建模.

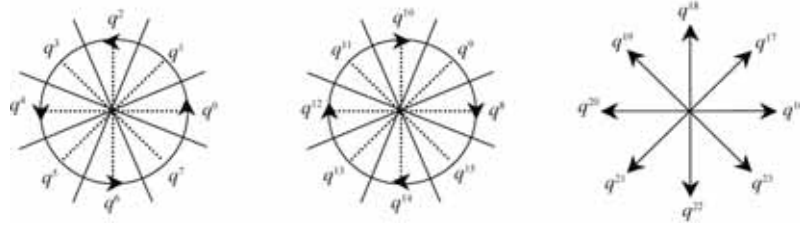


Fig.1 Definition of primitive type
图 1 图元类型定义

2.4 模型定义与参数设定

根据上文定义的基本状态即能构建用于描述高层图元的复杂模型.图 2 给出了一种比较直观的解决方案,使用该 ergodic 模型,能够实现状态之间的转换,其特点包括:

- 1) 状态间平等.不仅每个状态作为初始状态的概率相同,同时,从当前状态转换到其他任意状态的概率也是相等的;
- 2) q^s 和 q^f 为虚拟状态,只起连接作用,而不会发射任何观察向量;
- 3) 实现简单,且能够描述由文中定义的图元所构成的任意笔画.

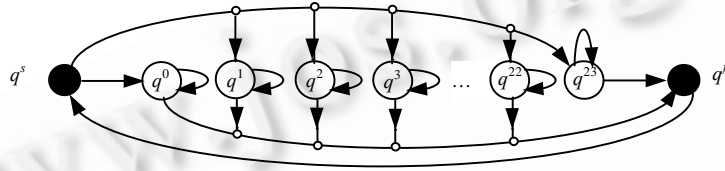


Fig.2 Generic ergodic model
图 2 各态历经的简单模型

问题在于,它否认了隶属于同一图元(这里特指弧)之间状态的一致性.如在逆时针弧线的绘制过程中, q^0 与 q^1 顺序出现的概率要明显高于其他状态,因此,这二者之间的转移概率也应该高于他者.此外,手绘图形具有不规则的特性,用户绘制的过程中难免会存在抖动的情況,状态间的绝对平等导致图 2 所示模型对噪音的抗干扰能力明显较弱.

文中使用的模型如图 3 所示.与图 2 不同,整个模型被分为 3 部分,分别是逆时针方向圆弧(最上层)、顺时针方向的圆弧(中间层)和 8 个不同方向的直线.与图 2 类似,这里的 q^s 仍然是起连接作用的虚拟状态,它不会向外发射任何观察向量;然而, q^{24}, q^{25} 不仅能够发射观察向量,同时与基本状态的区别在于它们能够以相同的概率发射任意观察向量.设置状态 24,25 的目的主要是为了处理两个图元之间的拐点,因为这类点的几何特征比较特殊,难以将其归属为基本状态;同时,为适应手绘图形的不规则性以及重采样过程可能引入的误差,模型中设置了两个拐点状态.

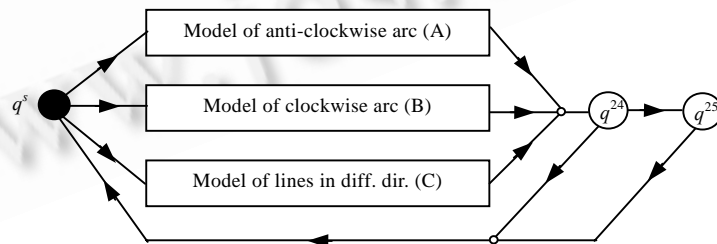


Fig.3 A sketch map of our model
图 3 本文所使用的笔画模型

以逆时针方向圆弧模型为例(如图 4 所示),与图 2 类似,模型能够以任意状态作为起始;不同的是,相邻状态

之间有一条转换边用于确保弧线绘制过程的连贯性.顺时针方向的圆弧可依此类推.

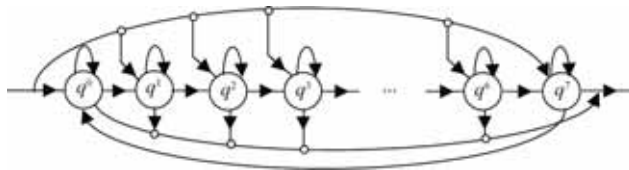


Fig.4 Model of anti-clockwise circle

图 4 逆时针方向圆弧模型

图 5 显示了直线模型状态之间的连接(模型 C).由于折线拐点的特殊性,模型中添加了用于表示特定折线拐点的节点状态($q^{26}-q^{32}$).如 q^{26} 用于描述水平向右的直线(q^{16})与右上方 45° 折线(q^{17})之间的拐点(为方便图形显示,隐去了其他直线状态之间的节点状态).与基本直线状态相比,这类状态的特点是曲率要大得多(有关曲率定义详见本文第 3 节),甚至与弧线上点的曲率近似.尽管图 3 中的状态 24 和状态 25 同样能够表征此类观察向量,但由于这里状态是特别定制的,因此,较 q^{24} 和 q^{25} 能够提供更大的概率,从而使系统具有更高的性能.

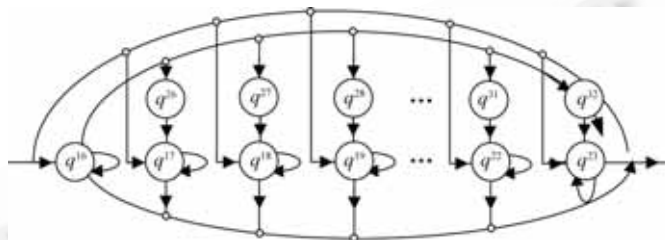


Fig.5 Model of lines in different orientations

图 5 不同方向的直线模型

与传统 HMM 针对单一模式的情况不同,本文所提出的模型包含了 3 种基本模式的任意组合,因此难以为其生成完备的训练样本.实验中用到的所有模型参数都是在先验知识与严格的数学运算的基础上得到的.为方便计算与推导,参数定义与计算过程基于如下假设:(1) 所有基本状态(q^0-q^{23})作为初始状态的概率是均等的,且其他状态作为初始状态的概率为 0;(2) 弧线状态自循环与到相邻后续状态的概率相等;(3) 所有基本状态(q^0-q^{23})到达拐点状态 q^{24} 的概率相同.

公式(3)给出了用于描述模型发射特定观察向量的观察概率.对于点 p_i ,本文使用四维向量(f_1, f_2, f_3, f_4)进行描述,有关特征计算的相关内容详见第 3 节.文字描述为:点 p_i 在状态 j 的观察概率为各特征在状态 j 的概率密度函数的连乘积.将公式(1)中的迭加修改为连乘运算的优点在于体现了相对 pdf 取值的重要性,避免了迭加过程中部分 pdf 分量很小却具有较高联合概率密度的情况.

$$pdf(o_i | q^j) = \prod_{k=1}^4 pdf(f_k | q^j) \quad (3)$$

3 特征计算

3.1 特征选择

本文使用局部方向信息与局部曲率信息区分模型中的不同状态.方向信息的计算相对简单,对于直线而言, p_i 的方向就是直线 $p_{i-1}p_{i+1}$ 的方向;而对于圆弧,可以证明在重采样后 p_i 的切线方向与直线 $p_{i-1}p_{i+1}$ 的方向一致.

直线 $p_{i-1}p_{i+1}$ 倾角的正余弦能够方便地通过两点坐标得到.相比较而言,曲率的定义与计算则显得比较复杂.通常在离散情况下曲率都是根据特定范围内采样点的分布获得的,这里我们将曲率定义为当前点到邻近点构成窗口的两个端点所在直线的距离.对于直线上的点而言,该距离值很小,近似为 0;而对于圆弧上的点,取值应相

对大一些.图 6 中给出了特征计算公式以及相应的物理含义.此外,曲率符号随绘制方向的变化而改变:对于顺时针方向其值为正;对于逆时针方向,其值为负.

$$f_1 \doteq \cos \varphi = \frac{x_{i+1} - x_{i-1}}{\sqrt{(x_{i+1} - x_{i-1})^2 + (y_{i+1} - y_{i-1})^2}}$$

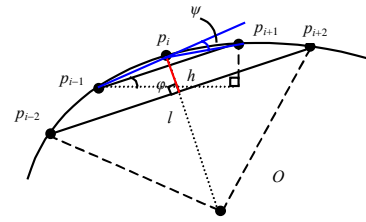
$$f_2 \doteq \sin \varphi = \frac{y_{i+1} - y_{i-1}}{\sqrt{(x_{i+1} - x_{i-1})^2 + (y_{i+1} - y_{i-1})^2}}$$

$$f_3 \doteq \text{curvature}(p_i) = h = \text{distance}(p_i, p_{i-2}p_{i+2})$$

$$f_4 \doteq \psi = \pi - \text{angle}(p_{i-1}p_i, p_i p_{i+1})$$

(a) Expression of feature computation

(a) 特征计算公式



(b) Physical meanings of the features

(b) 特征对应笔画的物理意义

Fig.6 Expression of feature computation and their physical meanings

图 6 特征计算公式及其物理意义

此外,我们引入了方向变化量特征(f_4),它同样能够被看作为曲率的一种度量机制.由于该特征仅与最近的两个点相关,所以能够平衡因曲率计算过程中引入过多上下文而有可能产生的问题.

曲率计算时,窗口大小的选择需要在直线与弧线之间进行权衡.通常情况下,窗口越大,模型的抗干扰能力就越强,但同时也会由于引入过多上下文而出现过平滑的现象;小窗口计算下精度较高,但其抗干扰能力相对较弱,同时存在区分力不强的缺点.在本文中,窗口大小定义为 5 个重采样点大小,而重采样距离则是在对笔画长度直方图进行分析的前提下确定的.

3.2 特征的 PDF 定义

本文的特征概率密度函数定义使用高斯分布.在定义时需满足如下要求:1) 函数在状态中心方向的取值最大;2) 函数取值与倾角与状态中心方向所成的角度成正比.

以状态 16 的 f_2 为例,其倾角覆盖范围为 $[-\pi/8, \pi/8]$.当角度为 0° 时位于状态 16 的概率最大;随着角度向两侧伸展,概率值逐渐降低.可以看到, pdf 取值是与倾角同状态中心方向(对于状态 16 而言,这里指的是 x 轴正向)所成角度成正比的,偏转角度越小,则概率越大.又由于函数 $|\sin(x) - \sin(0)|$ 在 $[-\pi/8, \pi/8]$ 区域内关于 y 轴对称,因此状态 16 的 sine 特征的 pdf 定义如图 7(a)所示:曲线在 $\sin(0)$ 的位置取得最大值,并且两侧对称.对于状态 17,其中心方向为 $\pi/4$,角度覆盖范围从 $\pi/8$ 到 $3\pi/8$.由于 $|\cos(x) - \cos(\pi/4)|$ 在 $\pi/4$ 两侧不是对称的,因此可以看到状态 17 的 cosine 特征的 pdf 函数由两条高斯曲线混合而成:函数在 $\cos(\pi/4)$ 处取得最大值,同时向两侧逐渐递减,且函数在 $\cos(\pi/8)$ 与 $\cos(3\pi/8)$ 处的取值相等.其他状态的方向 pdf 定义可依此类推.

曲率特征的 pdf 根据模型所属状态分为 3 类:Func1 用来表征正向曲率的圆弧状态($q^8 - q^{15}$),Func2 表征曲率为负时的圆弧状态($q^0 - q^7$),Func3 表征所有直线状态.由于在重采样距离一定的前提下,圆的半径越大则曲率越小,这意味着,当圆半径足够大时,曲率信息很难将直线与圆弧上的点进行正确划分.为此,我们将可处理的圆半径大小定义在 10~50 个重采样点的范围之内,且认为该范围能够满足大多数笔式交互应用的需要.实验结果表明,直线的曲率范围在 $[-d/8, d/8]$ 之间,并由 h 定义的几何性质可知, h 的有效取值范围在 $[-2d, 2d]$,这里, d 为重采样间距.图 7(c)给出了状态 17 的 pdf_{f3} 定义.图 7 只显示了 Func1 和 Func3,根据对称性可知,Func2 与 Func1 关于 y 轴对称.

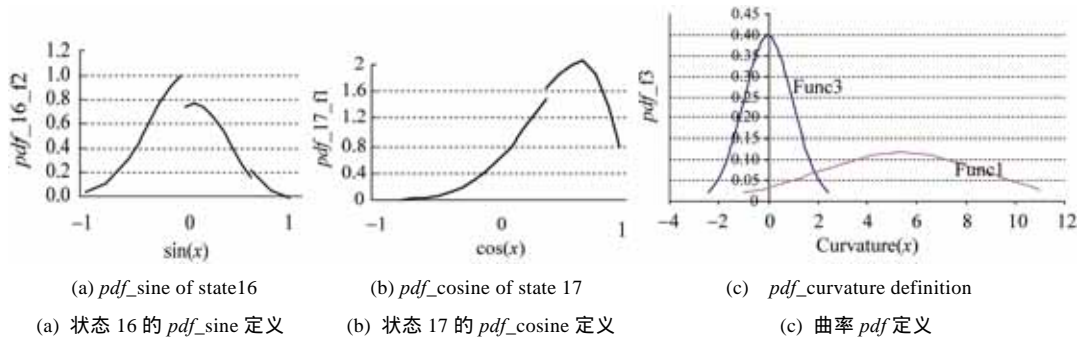
Fig.7 Examples of pdf definition

图 7 概率密度函数定义示例

4 图元分解步骤

笔式输入设备通常基于等时采样机制,在原始采样点上直接进行特征计算,一方面计算代价过高,另一方面也会由于采样频率以及绘制速度的差异而产生噪音,为此我们引入了重采样机制.由于对 HMM 而言,每个模式至少需要 4~5 个观察向量与之相对应,因此可以在对笔画图元长度分析的基础上确定重采样间距的大小.

图元分解过程如图 8 所示,算法步骤包括:

- Step 1. 笔画重采样后计算得到特征点序列;
- Step 2. 对特征点进行特征计算,作为 HMM 模型的输入;
- Step 3. HMM 模型使用 Viterbi 算法对特征序列进行识别,计算得到最优状态序列;
- Step 4. 对识别结果进行后处理,并转化为最终需要的图元序列.

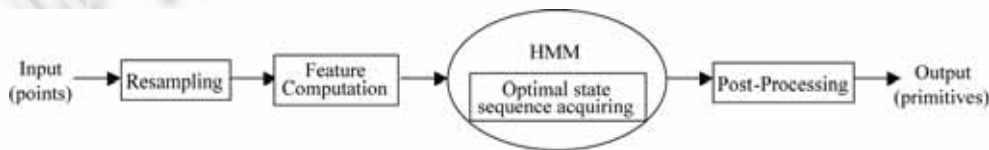


Fig.8 Stroke fragmentation process

图 8 笔画图元分解过程

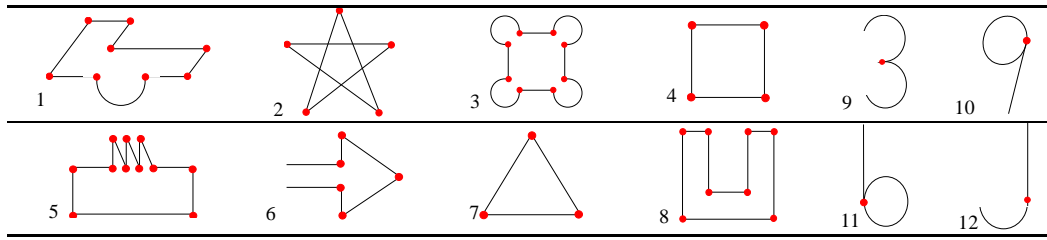
以形如“L”的笔画为例,根据上文的状态定义,模型输出为“22,...,22,68,16,...,16”.后处理过程对分割点进行准确定位,并以颜色信息区分不同的图元类型.对于上例,首先,状态 68 对应的特征点会被作为候选分割点,之后我们选择与该特征点最为接近的原始采样点作为最终分割点.这样做的目的是因为重采样之后不能保证重采样点能够覆盖所有的分割点,但却能保证真实分割点一定在候选分割点的附近.最后,根据图元类型以不同颜色将分解结果回显给用户.

5 性能评估

实验所需测试用例大部分来自已有笔画图元分解工作^[2,4],同时,在选择时避免了分解结果不唯一的样本.此外,我们添加了 9~12 这 4 个样本,用以检验方法对手写数字与文字的适用性.样本见表 1,图中圆点表示理想分割点,由于绘制时笔画的首、末两点自动地作为分割点来处理,因此不列入考察范畴.

Table 1 Samples used in our experiment

表 1 实验用测试样本



实验数据来自 12 个受试者,实验前我们首先向他们介绍采样设备的使用及注意事项,同时给每个人 5~10 分钟的时间学习使用;之后向他们说明实验目的及本文能够识别的图元类型,以避免出现绘制过于随意的情况.采样中受试者以非受限方式进行,对于图形绘制方向以及大小没有任何强制规定.每个图形以单笔画的形式绘制 5 遍~8 遍不等,且同一个图形的不同样本之间采样时间不连续.实验中部分用户的实际样本数目要多于规定样本个数,且由于部分图形与目标样本之间存在较大偏差,为保证评估的客观性未列入评估范畴,这样最终得到有效样本 930 个,全部作为测试样本.

图 2 所示模型同样被实现来与本文方法相对照,图 9 列出了二者性能的比较.为了尽可能多地发现分割点,方法不可避免地会出现过分割的情况,即将不是分割点的地方错误地标记为分割点;对于误否定也是如此.因此,仅以分割点定位的正确率来评估算法性能不是很全面,通常相关工作中都以误确认率(false-positive)和误否认率(false-negative)相结合的方式对算法进行评估.误确认率表示系统接受的伪分割点占有所有被接受为分割点的个数的比例;误否认率定义为被拒绝的真实分割点占有所有真实分割点的比例,二者的形式定义如下:

$$\begin{aligned} \text{false_positive} &= \frac{\text{no\#_accepted_false_segment_point}}{\text{no\#_all_accepted_segment_point}} \times 100\% \\ \text{false_negative} &= \frac{\text{no\#_rejected_true_segment_point}}{\text{no\#_all_true_segment_point}} \times 100\% \end{aligned} \quad (4)$$

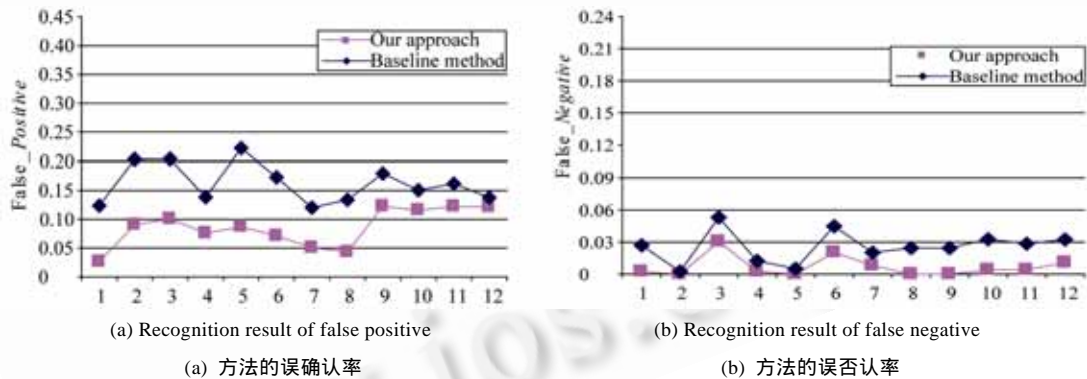


Fig.9 Stroke fragmentation results

图 9 笔画图元分解结果

由图 9 可知:

- (1) 本文所选模型较图 2 所示直观模型的性能要高许多,尤其体现在误确认率大约只为图 2 模型的一半.
- (2) 本文方法的误否认率很低,小于 1%,这表示该方法能够准确定义几乎所有的真实分割点;同时,误确认率在不同用户之间存在微小差异,证实方法具有一定的用户适应性.

实验结果表明,误确认主要发生在对圆弧的过分割方面,如对于图 10(a)中的笔画,圆被过分割为两条直线段和一条圆弧,这是由于 HMM 力图通过动态规划方式找到与模型的最佳匹配,因而若用户在绘制弧线的过程中

出现较长间隔的平滑片段,那么模型将该段识别为直线的概率要远大于将其识别为弧线的概率.从该例子也可以看出,若不考虑标准模板而单纯地从笔画的几何构成来看,系统的分割结果似乎更贴合用户的原始笔画,这实际上也说明用户的实际绘制与真实意图之间可能存在偏差.还有一些过分割发生在直线绘制过程中出现较大方向改变的情况,如图 10(b)中最下方笔画被过分割为两条不同方向的直线段.图 10(c)中列举了一些成功分解的例子(图中圆圈代表分割点).

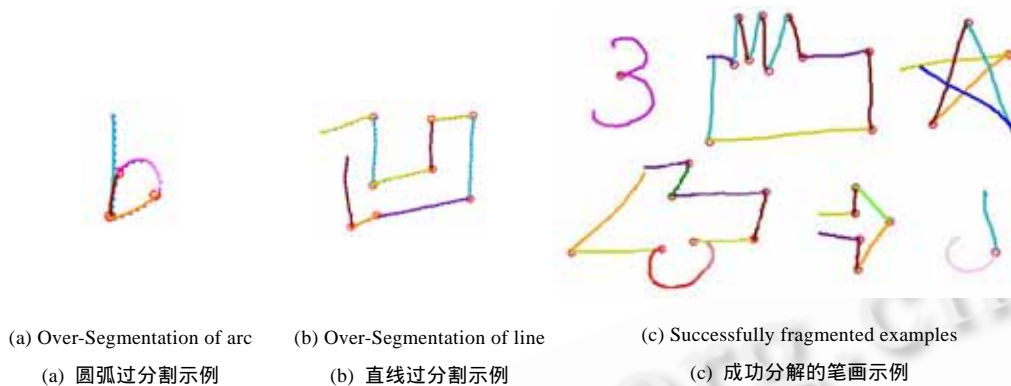


Fig.10 Fragmentation examples

图 10 部分分解结果示例

此外,方法对样本 1~样本 8 的处理精度较比样本 9~样本 12 要高.这是因为,对于图形绘制而言,大部分用户会按照模板给定的图元构成绘制,而数字和字母的绘制相对存在较多主观因素.由于就这些符号本身而言并不是由规整的圆弧所构成,因而往往出现过分割的情况.但是我们认为本文系统做到了在尽可能不丢失笔画原有信息的基础上,对数据进行最大限度压缩的目的(本文方法相比较文献[11]中提到的方法,数据压缩比要大得多).与此同时,实验也证实了在综合考虑了绘制的上下文的基础上,系统对平滑过渡的笔画(样本 12)能够实现较好的分割效果.

就处理时间而言,基于 HMM 方法的问题复杂度为 $O(N^2T)$,这里, N 为模型状态数, T 为观察序列的长度.实验中系统对每个笔画的处理时间为 89.6873ms,相比于用户绘制每个笔画的平均时间 2393.556ms 而言,基本上能够满足实时处理的需求.

6 结束语

本文提出了基于隐马尔可夫模型的笔画图元分解方法,通过寻找与模型的最佳匹配实现笔画图元分解的目的.文中模型能够描述任意由直线段和圆弧构成的笔画,且无须任何领域相关知识,从而具有良好的领域可扩展性.此外,方法能够将分割点查找与图元类型判定同时完成.由于它综合使用了局部信息和全局信息,因而能够处理传统方法不能处理的平滑过渡的曲线.简单几何特征保证了方法的运算效率.实验结果也表明,本文方法达到了在保证笔画几何信息的基础上对数据进行最大限度压缩的目的.

本文只是对综合利用绘制上下文信息的基础上进行笔画图元分解的初步尝试.实验中发现,上下文信息既可以对分解提供有效支持(先从整体把握笔画的几何构成然后进行分解),同时也可能为分解带来错误干扰(体现在曲率计算的连贯性与图元间曲率计算的非连续性之间的矛盾,这也是导致折线弱分割的主要原因).因此,如何恰当地利用上下文信息是笔画图元分解所需考虑的主要问题.

References:

- [1] Hse H, Shilman M, Newton AR. Robust sketched symbol fragmentation using templates. In: Vanderdonckt J, Nunes NJ, Rich C, eds. Proc. of the Intelligent User Interface 2004. Funchal: ACM Press, 2004. 156-160.

- [2] Kim DH, Kim M. A curvature estimation for pen input segmentation in sketch-based modeling. *Computer-Aided Design*, 2006, 38(3):238–248.
- [3] Dudek G, Tsotsos J. Shape representation and recognition from multiscale curvature. *Computer Vision and Image Understanding*, 1997,68(2):170–189.
- [4] Sezgin TM. Scale-Space based feature point detection for noisy digital ink. In: *Proc. of the SIGGRAPH: Computer Graphics and Interactive Techniques*. San Diego: ACM Press, 2004. 145–151. <http://portal.acm.org/citation.cfm?id=1281547>
- [5] Calhoun C, Stahovich TF, Kurtoglu T, Kara LB. Recognizing multi-stroke symbols. In: Alberta E, ed. *Proc. of the AAAI Spring Symp. 2002 on Sketch Understanding*. Palo Alto: AAAI Press, 2002. 15–23.
- [6] Sezgin TM. Feature point detection and curve approximation for early processing of free-hand sketches [MS. Thesis]. Cambridge: Massachusetts Institute of Technology, 2001.
- [7] Yu B. Recognition of freehand sketches using mean shift. In: *Proc. of the Intelligent User Interface 2003*. Miami: ACM Press, 2003. 204–210. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.14.4733>
- [8] Paulson B, Hammond T. A system for recognizing and beautifying low-level sketch shapes using NDDE and DCR. In: Shen C, Jacob R, Balakrishnan R, eds. *Proc. of the User Interface Software and Technology*. Newport: ACM Press, 2007. <http://srl.csdl.tamu.edu/publications.shtml>
- [9] Cates S, Davis R. A new approach to early sketch processing. In: *Proc. of the Making Pen-Based Interaction Intelligent and Natural*. Menlo Park: AAAI Press, 2004. 29–34. <http://rationale.csail.mit.edu/publications/Cates2004New.pdf>
- [10] Rabiner LR. A tutorial on hmm and selected applications in speech recognition. *Proc. of the IEEE*, 1989,77(2):257–286.
- [11] Marcel S, Bernier O, Viallet JE, Collobert D. Hand gesture recognition using input-output hidden Markov models. In: *Proc. of the 4th IEEE Int'l Conf. on Automatic Face and Gesture Recognition*. Grenoble: IEEE Communications Society Press, 2000. 456–461. <http://www.idiap.ch/resources/gestures/papers/marcel-fg-00.pdf>
- [12] Tay YH, Lallican PM, Khalid M, Viard-Gaudin C, Kneer S. An offline cursive handwritten word recognition system. In: *Proc. of the 10th IEEE Region Int'l Conf. on Electrical and Electronic Technology*. Singapore: IEEE Communications Society Press, 2001. 519–524. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.21.328>
- [13] Sun Z, Jiang W, Sun J. Adaptive online sketchy shape recognition based on HMM. *Lecture Notes in Artificial Intelligence*, 2006, 3930:948–957.
- [14] Sezgin TM, Davis R. Sketch interpretation using multiscale models of temporal patterns. *IEEE Computer Graphics and Applications*, 2007,27(1):28–37.
- [15] Pei JH, Li CY, Gong X. A novel hidden Markov model and its application to recognize hand-drawn graphic symbol. *Chinese Journal of Computers*, 2005,28(10):1745–1752 (in Chinese with English abstract).

附中文参考文献:

- [15] 裴继红,李翠芸,龚忻.一种新的隐马尔可夫模型及其在手绘图形识别中的应用. *计算机学报*,2005,28(10):1745–1752.



冯桂焕(1980 -),女,天津人,博士生,主要研究领域为智能人机交互,草图理解技术.



Christian VIARD-GAUDIN(1956 -),男,博士,教授,主要研究领域为动态规划,分割算法,手写体识别.



孙正兴(1964 -),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为多媒体计算,计算机视觉,智能人机交互.