

文章编号:1001-9081(2009)02-0456-03

基于 K 段主曲线算法的手绘形状识别

屠红蕾, 黄 静

(北京师范大学珠海分校 信息技术学院, 广东 珠海 519086)

(helentu_2006@sohu.com)

摘 要:提出了利用主曲线算法对人机交互中广泛使用的脱机手绘形状的识别。利用改进 K 段主曲线方法对学习样本点提取主干线,然后对被测试样本计算到主干线的投影点的距离的均方差,识别是基于对目标主曲线的距离均方差最小。实验证明该算法也是比较有效的一种形状识别算法。

关键词:主曲线;K 段主曲线;形状识别;降维

中图分类号: TP391.43 **文献标志码:** A

Hand-drawn shape recognition based on k-segment algorithm for principal curves

TU Hong-lei, HUANG Jing

(College of Information Technology, Beijing Normal University Zhuhai, Zhuhai Guangdong 519086, China)

Abstract: A hand-drawn shape recognition algorithm based on improved K-segment principal curve was presented. The algorithm was to generate a curve passing through the "middle" of the distribution. Recognition was based on the average distance from the sample points to the projected points on the generated principal curve. Experimental results show that the algorithm is feasible for hand-drawn shape recognition and exhibits good results.

Key words: principal curve; K-segment principal curve; shape recognition; dimension reduction

0 引言

形状识别是模式识别的一个主要课题,在机器视觉等领域有广泛使用。目前形状识别的主要算法^[1]有:边缘编码法^[2]、多边形拟合算法,基于变换域特征(例如不变矩^[5])、基于傅立叶描述子^[3]和基于多尺度下的小波描述子^[4]等。另外广泛使用的还有主成分方法,文献[5]使用主成分法对手写数字进行了识别。

主成分方法的一个衍生是主曲线和主曲面方法^[7],它们是利用样本点分布概率密度和最大似然估计的统计理论设计出的算法,对样本数据进行主曲线或者曲面的提取。近年来,主曲线方法在模式识别、数据挖掘等领域获得了广泛关注,同时出现了许多改进算法。本文就人机交互中的一个应用,手绘形状识别进行了研究。

手绘形状有任意性,个性化的特点,识别有一定难度,目前常用的算法识别准确率受到这些因素的影响。本文使用了 K 段主曲线算法来解决该类问题,给出了改进算法的详细描述和识别方法,同时给出了实验结果。实验表明该方法是一种比较有效和有自身特点的形状识别算法。

1 主曲线和 K 段主曲线算法

主成分法在模式识别中被广泛使用,它利用数据统计的特征,找到样本数据的主要分量。基于主成分概念,文献[7]提出了主曲线概念,它是对主成分的非线性衍生。主曲线是通过数据分布“中间”并满足“自相合”的光滑曲线。图 1 是一个简

单的二维空间的例子,从该图中可发现主曲线与第一主成分相比,主曲线对数据的信息保持好,贴近数据形态,它与数据间的距离均方差小。

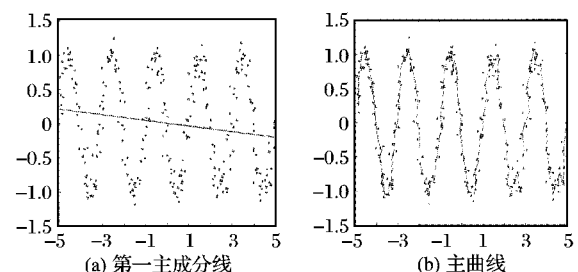


图 1 第一主成分和主曲线比较示例

1.1 主曲线定义

假设一个 d 维空间的点表达成向量 $\mathbf{x} = (x_1, x_2, x_3, \dots, x_d)$, n 个这样的点组成了一个向量集合,用 $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ 表示。一个该空间的光滑曲线表达为 $f(t) = (f_1(t), f_2(t), \dots, f_d(t))$, $t \in R$, 它是以 t 为参数的一条曲线。假设向量 \mathbf{x}_i 投影到该曲线的点是 $t_f(\mathbf{x}_i)$, 它可以形式化表达为: $t_f(\mathbf{x}_i) = \sup\{t: \|\mathbf{x}_i - f(t)\| = \inf_t \|\mathbf{x}_i - f(t)\|\}$ 。点 \mathbf{x}_i 到曲线的距离是点到曲线投影点之间的距离,表示为 $d(\mathbf{x}_i, f) = \|\mathbf{x}_i - f(t_f, \mathbf{x}_i)\|^2$ 。样本 X 到曲线的距离为可以表达为 $d(X, f) = \sum_{i=1}^n d(\mathbf{x}_i, f)$ 。

定义 1 R_d 空间的主曲线定义为一条光滑的曲线 $f(t)$, 该曲线最小化了样本 X 到曲线的距离 $f(t) = \arg_f(\min[d(X,$

收稿日期:2008-08-20;修回日期:2008-10-20。

基金项目:国家 863 计划项目(2007AA12Z156);澳门科技发展基金资助项目(018/2005A);北京师范大学珠海分校重点资助项目(Z06007)。

作者简介:屠红蕾(1970-),女,江苏南京人,讲师,硕士,主要研究方向:人工智能、模式识别;黄静(1967-),女,湖南衡阳人,副教授,博士,主要研究方向:图形学、图像处理。

$f))$ 。

主曲线上每个点是所有投影至该点的数据点的条件均值。它满足自相合性。

继 HS 主曲线提出后,具体求主曲线的算法很多,例如 BR 主曲线^[8]、D 主曲线^[9]、多边形算法主曲线^[10]等。这些算法基于比较平滑的,弯曲小而无自相交的样本数据。对任意性大、弯曲大、可能自相交的数据,前面算法效果不好,文献[11]用 K 段主曲线算法解决该类问题。鉴于手绘形状有自相交,弯曲度大的数据,我们选择了该算法,同时就算法进行了改进,以提高效率和准确性。

1.2 改进的 K 段主曲线算法

算法描述前,先就算法中使用到的术语给出定义。定义点集和线段的距离关系,有点集合 x_1, x_2, \dots, x_n 和线段集合 l_1, l_2, \dots, l_n , 线段 l_j 的两个端点表示为 v_{1j}, v_{2j} 。设 t_{ij} 是点 x_i 到线段 l_j 的投影点,则点 x_i 到线段 l_j 的距离表示为 $\text{dist}(x_i, l_j)$, 它是该点到该线段投影点的距离或者到线段两端点中最近的端点的距离:

$$\text{dist}(x_i, l_j) = \min(\|x_i - t_{ij}\|, \min(\|x_i - v_{1j}\|, \|x_i - v_{2j}\|))$$

线段 l_j 的 Voronoi 区域(VR)是由那些离该线段距离最近的点组成的集合:

$$VR(l_j) = \{x_i : \text{dist}(x_i, l_j) = \min\{\text{dist}(x_i, l_k)\}\}$$

下面就改进的 K 段主曲线算法进行详细描述:

步骤 1 初始化。假设数据样本为 $S_1(v_1, v_2, \dots, v_n)$, 它是初始的样本区,在其中找到它们的第一主成分,插入主成分线段时,以 3σ 为长度(σ 是同一样本区中样本点到区域质心的距离方差),令该线段为 l_1 , l_1 的 Voronoi 区域为 $VR_1 = \{v_1, v_2, \dots, v_n\}$ 。

步骤 2 插入新线段。1) 寻找偏远点: 在所有样本点中,选择离开自己主成分线段最远的点 v_j , 同时该点周围还有至少 2 个点到该点的距离小于它们到自己主成分线段的距离,即:

$$\begin{aligned} V_j &= \inf \{v_i : \text{dist}(v_i, l_i) = \max\{\text{dist}(v_j, l_j)\} \\ &\text{and } \text{Nearest}_v = \{v_k : \text{dist}(v_i, v_k) < \text{dist}(v_k, l_k) \\ &\text{and } |\text{Nearest}_V| \geq 3\} \end{aligned}$$

其中 $i, j, k = 1, \dots, n$, $\text{dist}(v_m, l_m)$ 表达点 v_m 到自己相应第一主成分线段 l_m 的距离。

假设本步开始样本区为 S_1, S_2, \dots, S_k , 偏远点和邻近它的点形成新的样本区域即为 S_{k+1} 。

2) Voronoi 区域的调整步骤:

2.1) 在样本区域 S_1, S_2, \dots, S_{k+1} 中找到调整了样本点的区域,计算第一主成分,以 3σ (σ 同上) 的长度插入主成分线段。

2.2) 由于产生了新主成分线段,对所有样本点调整它们的 Voronoi 区域,有的样本区被更新,返回 2.1)。

2.1) 和 2.2) 反复迭代,直到没有样本点被重新调整到新的 Voronoi 区域。

步骤 3 构造和优化:将步骤 2 得到的 l_1, l_2, \dots, l_k 个线段构成 Hamilton 路径。使用贪心算法。算法如下:

1) 令 $h = k$, 初始有 h 个哈密顿子图(sub-HP)。

2) 当 $h > 1$, 求每个子图的端点到其他子图端点的代价值。代价函数定义为 $C(e_i) = l(e_i) + \lambda\alpha(e_i)$, e_i 是连接两个 sub-HP 端点的边,其中 $l(e_i)$ 是边的长度, $\alpha(e_i)$ 是角度惩罚

(图 3 所示), λ 是角度惩罚系数,由用户设定。 h 个子图共计算 $C_{2h}^2 - h$ 个组合。连接使 $C(e_i)$ 最小的边的端点, $h = h - 1$, 返回 2), 直到 $h = 1$ 为止, 设该步中产生的所有连为 e_1, e_2, \dots, e_{k-1} 。

3) 2-opt 优化方法对 HP 路径优化。

4) 检查形成 HP 过程中产生的假边。对 e_1, e_2, \dots, e_{k-1} , 检查 e_i 连接的主干线 l_i, l_{i+1} , 将 l_i, l_{i+1} 被 e_i 连接的两端点的 VR 区域点形成一个点集合, 计算集合中每个点相对 e_i 边的 VR 区域, 如果 e_i 长度大于某阈值并且中间段 VR 区点数目小于某阈值, 则判断 e_i 为假边, 在 HP 路径中做假边记录。

步骤 4 终止条件判断: 计算目标函数, $F = n \lg l + \sum_{i=1}^k \sum_{j \in V_i} d(x_i, l_i)^2 / (2\sigma^2)$, l 是形成的 HP 的长度。终止条件是如果目标函数达到了最小, 或者线段数已经达到了预先设定的最大值。如果不满足终止条件, 返回步骤 2。

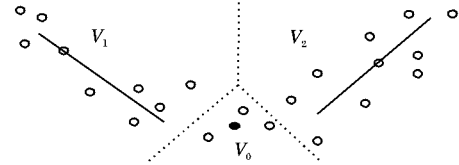


图2 选择离主成分线段最远点以产生新线段

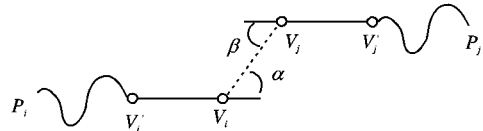


图3 子Hamilton路径形成中的角度惩罚示意图

本算法对 K 段主曲线算法做了改进, 主要在步骤 3 的第 4) 小步, HP 路径形成以后, 连接各个第一主成分的线段称为连接线段, 该算法是基于把所有主曲线用 Hamilton 路径算法全部相连, 但是对一些形状, 形成所有主成分线段完全连接是不合适的, 比如字母 E, F, X, i, j 等不连续形状, 会产生不必要的线段连接, 我们把这些边叫假边。为此算法对原始算法进行改进。对形成的 HP 连接进行测试, 检查线段是否有数据分布在周围, 即它并不是任何数据的主干线, 是条假边, 则把该边从 HP 路径中删除。改进的算法可以精确地提取任何原始形状的主干线, 无论形状是否连续, 如图 5 中的字母 E, F, i 就是几个例子。

该算法的目标函数 $F = n \lg l + \sum_{i=1}^k \sum_{j \in V_i} d(x_i, l_i)^2 / (2\sigma^2)$

的推导, 是假设数据云和主曲线上的每个点概率分布关系是正态分布, 然后根据数据点到自己主曲线上点的条件概率为最大而推导出的。对于有一定笔画宽度的手绘形状, 假设笔画宽度的一半为 w , 可以取目标函数中的方差 σ^2 为 $w^2/3$ 。针对算法的结束条件, 可以选择插入的线段条数超过了一定的上限, 或者目标函数达到最小。如果单纯选择让目标函数达到最小, 可能会达到局部最小。可以选择插入线条数目超过上限为结束条件, 然后在所有的主曲线中选择能够让目标函数最小的那个主曲线。

2 实验结果和分析

本文针对不同的手绘形状, 利用不同粗细的笔画进行了实验。其中包括常用的直线、圆型、方型、三角型和箭头, 还有一

些比较任意的图形,比如花瓶、小鸭子、气球、花边等,如图 4 所示。每个图中左边是原始形状,右边是算法形成的主曲线。实验数据表明,主曲线提取的是数据的骨干线,由若干线段组成。针对数字和字母,我们也进行了实验,图 5 列出一些数字和字母的实验结果。

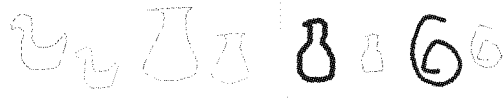


图 4 一些手绘任意形状的实验结果



图 5 一些手写字母和数字的实验结果

在具体的识别中,我们对不同的手绘形状,不同人绘制的个性化比较强的形状,如每个人手写数字的习惯不同,进行了先期训练。具体做法是把每个用户的训练样本的主曲线求出以后,保存在数据库中。在测试时,先把测试样本和模式曲线做比例缩放,调整到相同尺度,例如都在 16×16 的方格中进行数据对比。匹配是计算测试样本和数据库中每个模式形状的主曲线的距离均方差,均方差最小的是最匹配的模式。

假设测试样本为 $TS(x_1, x_2, \dots, x_n)$, 数据库中有模式 M_1, M_2, \dots, M_p , 模式 $M_j (j = 1, 2, \dots, p)$ 的主曲线为 $M_j_PS(l_1, l_2, \dots, l_q)$, 定义点 x_i 到模式 M_j 中主曲线的线段的距离为 $\text{dist}(x_i, M_j) = \min(\text{dist}(x_i, l_{q_j})) (i = 1, 2, \dots, n, q_j = 1, 2, \dots, q)$ 。测试样本和模式 M_j 的距离方差是:

$$\text{mean_dist}(TS, M_j) = \sqrt{(\sum_{i=1}^n \text{dist}(x_i, M_j)^2) / n}$$

表 1 给出一些形状的距离数据,数据表明对形状相同或接近的两个样本,距离会小,例如数字 3 和 8 的距离近,8 和字母 S 的距离近,所以本方法可以用来比较两个形状接近的程度。

表 1 部分被测试样本和训练样本均方差

样本	三角	箭头	圆形	字母 S	字母 W	数字 2	数字 3
三角	3.63	18.82	16.12	13.19	16.22	9.55	12.92
箭头	18.82	5.35	22.06	11.26	12.54	10.02	14.59
圆形	16.12	22.06	6.97	13.87	18.53	12.77	14.60
字母 S	13.19	11.26	13.87	5.08	16.84	8.86	6.49
字母 W	16.22	12.54	18.53	16.84	3.03	14.45	14.13
数字 2	9.55	10.02	12.77	8.86	14.45	3.44	11.65
数字 3	12.92	14.59	14.60	6.49	14.13	11.65	4.80

本算法经过自己建立的测试数据库的测试,该数据库包含仿 USPS 的数字的形状,字母和任意形状。针对每个不同的形状进行了主曲线提取,均能性能稳定地得到结果。该算法的算法复杂性在特征提取上代价比较高,假如 N 是数据维,在特征提取方面,需要的时间代价约 $O(N^4 + N^4/k^2 + k^3)$, k 为最后的主曲线段数,提取出的线段存储所需要的代价是 $O(k)$, 匹配的时间代价约 $O(kN)$, 相比 PCA 方法,特征提取部分时间代价约 $O(N^4)$, 匹配代价约 $O(dN^2)$, 特征存储代价约 $O(dN)$, d 为前主分量个数。可以看出前者特征提取方面消耗大,而匹配和特征存储消耗小。该算法基于提取每个样本主曲线,匹配是让测试样本和每个样本曲线比较,而 PCA 方法是基于多个样本的统计量做的主成分提取,所以本方法提取的数据特征比 PCA

更接近于个体数据。经过测试,该算法对个体数据的识别率比 PCA 高。

另外,基于均方差的算法要求测试形状和模式错位不能太大,在旋转情况下算法失效。但是由于提取的主曲线可以比较容易地根据匹配需要进行缩放旋转和移动,所以经过进一步的深入工作可以把该算法应用到有旋转缩放移动的情况中。除了计算距离均方差方法以外,识别技术还包括文献[12]中提到的分析曲线特征的方法,例如分析交叉点,回路数等。该方法对于已经知道模式特征的情况下是可行的,而且可以抓住模式主特征,在形状旋转等情况下部分有效,需要视特征而定。但是如果模式特征任意,分析特征的方法就不适合了,丢失了通用性。而计算距离方法可以比较任意形状而不用考虑形状的具体特征。

3 结语

本文用 K 段主曲线算法进行手绘形状识别,同时进行了改进,用此方法可以对弯曲度很大或有自相交的数据进行处理,改进点可以让算法对任意不连续的形状进行识别。实验对不同的手绘形状,包括数字,字母,任意形状,包括相对复杂形状进行了测试,同时对具有不同方差的数据分布进行了测试。利用提取的主干线,使用均方差方法对测试样本进行了识别,识别率比较高。由此我们认为把 K 段主曲线算法应用在脱机手绘形状识别上是可行的,比主成分方法,基于正交变换的算法更直观,虽然运算时间代价略高,但是数据特征提取准确。该方法可以比较容易推广到通用的形状识别中。

参考文献:

- [1] 丁险峰,吴洪. 形状匹配综述[J]. 自动化学报, 2001, 127(15): 678 - 694.
- [2] 郭克华,刘广海,刘传才,等. 微分几何编码识别物体的形状[J]. 计算机科学, 2007, 34(7): 219 - 221.
- [3] 王涛,刘文印,孙家广,等. 傅立叶描述子识别物体的形状[J]. 计算机研究与发展, 2002, 39(12): 1714 - 1719.
- [4] 夏敏,刘宏申. 基于小波描述子和神经网络的形状识别[J]. 计算机研究与发展, 2007, 17(3): 106 - 108.
- [5] HIT M K. Visual pattern recognition by moment invariant[J]. IEEE Transactions on Information Theory, 1962, 8(2): 179 - 187.
- [6] 芮挺,沈春林,丁健,等. 基于主分量分析的手写数字字符识别[J]. 小型微型计算机系统, 2005, 26(2): 289 - 292.
- [7] HASTIE T. Principal Curves and Surfaces[EB/OL]. [2008 - 06 - 17]. <http://handle.dtic.mil/100.2/ADA148833>.
- [8] BANFIELD J D, RAFTERY A E. Ice floe identification in satellite images using mathematical morphology and clustering about principal curves[J]. Journal of the American Statistical Association, 1992, 87(417): 7 - 16.
- [9] DELICADO P. Another look at principal curves and surfaces[J]. Journal of Multivariate Analysis, 2001, 77(1): 84 - 116.
- [10] KEGL B, KRZYSAK A. A polygonal line algorithm for constructing principal curves[C]// Proceedings of Neural Information Processing System, Cambridge: MIT Press, 1999: 501 - 507.
- [11] VERBEEK J J, VLASSIS N, KROSE B. A k-segments algorithm for finding principal curves[J]. Pattern Recognition Letters, 2002, 23(8): 1009 - 1017.
- [12] 焦娜,迟呈英,苗夺谦,等. 基于软 K 段主曲线算法的字符特征提取研究及实现[J]. 计算机科学, 2006, 33(1): 229 - 231.