

基于笔触特征三角形的手写汉字书法效果美化方法

于东 吕晓晨 荀恩东

(北京语言大学汉语国际教育技术研发中心 北京 100083)

摘要 提出笔触特征三角形方法,它为手写字字形骨架添加毛笔书法效果,以达到美化手写字的目的。该方法首先提取手写字数据中的书法特征和字形结构特征点集,构成字形骨架;然后通过添加笔触特征三角形序列生成笔画宽度信息;最后利用B样条曲线对生成的手写字轮廓线进行平滑,得到最终美化结果。该方法仅利用计算机采集到的手写字的坐标信息,以三角形方法为手写字添加书法特征,适用于各种手写字采集设备。

关键词 字形骨架,书法效果,手写字美化,笔触特征三角形

中图分类号 TP391 **文献标识码** A

Soft Pen Calligraphy Beautification for Handwriting Characters Based on Stroke Feature Triangle

YU Dong LV Xiao-chen XUN En-dong

(International R&D Center for Chinese Education, Beijing Language and Cultural University, Beijing 100083, China)

Abstract This paper presented an algorithm based on Stroke Feature Triangle Model, adding calligraphy effect to points and lines sampled by computer, in order to beautify hand writing characters. Strokes are formed by point set with calligraphy and character features. Algorithm based on Stroke Feature Triangle Model offers width to each part of strokes. B-Spline Curves are used in smoothing of handwriting stroke outline. This method can offer precise calligraphy simulation in the fly of character handwriting, and has good portability between various equipments.

Keywords Character stroke, Calligraphy effects, Handwriting character beautification, Stroke feature triangle

1 引言

现代社会中,“文房四宝”逐渐远离日常生活,人们广泛采用硬笔书写汉字,但毛笔书法仍然为大众所欣赏。随着计算机技术的发展,利用手写板等设备采集硬笔手写字得到广泛应用。在此基础上研究硬笔手写字的毛笔书法效果美化方法,生成具有个人风格的电子书法作品,成为非真实感艺术仿真领域的重要课题。在这方面,文献[1]在事前确定汉字笔画轮廓的前提下,通过控制样条曲线来实现笔速、墨量的模拟;文献[2]提出填充毛笔笔画算法来模拟毛笔临摹效果;文献[3]对笔尖和笔杆的运动建模,根据笔尖运动变化表现笔画粗细;文献[4]提出笔画相似度匹配算法,建立基本笔画库,根据用户输入笔画相似度挑选合适笔画进行书法效果模拟;文献[5]分析真实的书法图像,提取出基本笔画集用于构造书法字;文献[6]利用草书汉字轮廓和骨架信息,合成产生草书风格作品;文献[7]提出一种基于统计模型的书法创作模拟方法,其通过调节参数可以得到带有书法效果的笔画;文献[8]提出基于模糊支持向量机的曲线类比学习方法,其能够根据用户设定的参数生成不同风格的书法作品。此外,文献[9, 10]以笔段和连接件为单位,实现了联机的楷书、隶书的书法效果美化。以上研究均涉及书法效果的计算机模拟方法,但算法复杂度普遍较高,算法的实现需要大量学习样本,或者仅

应用于特定字帖字,影响了算法的实用性。

本文针对手写字提出一种简单有效的毛笔书法效果美化方法,即首先对采集得到的手写字数据按照时序信息计算字形骨架;然后以字形骨架中每两点为单位计算具有书法特征的笔触特征三角形,产生带有宽度的笔画轮廓;最后通过平滑算法,对所有笔画轮廓进行平滑,生成最终的书法效果美化字。该方法对各种手写字均可有效生成具有书法效果的美化字体,并且能够保留原字的书法特征。实验表明,该方法对行书、隶书字体均有较好的适应性,简单可行,能基本满足目前电子设备中采集、美化手写字的需求。

2 手写字的字形骨架计算

本文采用手写板作为采集手写字的主要设备,用户在手写板上书写汉字,计算机将该书写信息转为一系列带有时间信息的坐标点保存。采集得到的原始数据往往非常密集,包含大量冗余和书写过程中产生的抖动数据,因此需要对原始数据进行预处理,去除抖动和冗余点,在简化数据的同时保留写字中的书法特征,得到带有书法特征的字形拓扑结构,即字形骨架。

通常,一个手写字在计算笔画拐点后,其拓扑结构可简单地将笔画起点、拐点、终点相连来表示。然而仅用以上3类信息无法准确描述书写过程中的书法特征。为此,本文在计

到稿日期:2012-03-20 返修日期:2012-06-20 本文受国家自然科学基金项目(60973062, 61170162),北京语言大学青年自主科研支持计划项目(中央高校基本科研业务专项资金)(11JBB037, 12YBG02)资助。

于东(1982—),男,博士,讲师,主要研究方向为中文信息处理、人工智能, E-mail: yudong@blcu.edu.cn;吕晓晨(1988—),女,硕士生,主要研究方向为字形计算;荀恩东(1967—),男,博士,教授,主要研究方向为中文信息处理。

算字形骨架时不仅考虑到笔画拐点,还提取了一系列能够代表运笔笔势特征的点共同组成字形骨架,并对起点、拐点、终点进行处理,在简化数据同时保留书法特征。

2.1 笔画拐点和笔画中间点计算

拐点是汉字笔画中笔段与笔段间的转折点。在拐点计算方面,应用较广泛的有 Douglas-Peucker 算法^[11];在此基础上,文献[12]利用笔画动态射线方法计算拐点。本文采用类似动态射线的算法计算拐点,所不同的是在计算拐点的同时能够根据运笔笔势提取运笔中间点,并能够消除写字过程中抖动对计算的影响。

如前所述,采集得到的手写汉字中的点可记为 $p_i = (x_i, y_i, t_i)$, 其中 (x_i, y_i) 为该点坐标, t_i 为采集时间, 且对任意 $i < j$, 有 $t_i < t_j$ 。手写字笔画记为:

$$S_k^{(u,v)} = (p_u, p_{u+1}, \dots, p_v) \quad (1)$$

式中, k 为该笔画的书写顺序, 整个手写汉字可由笔画描述, 记为:

$$C = \{S_1^{(1,u)}, S_2^{(u+1,v)}, \dots, S_n^{(h,m)}\} \quad (2)$$

式中, m 为采集到的点数量, n 为总笔画数。为方便描述, 记笔画点集 $S_k^{(u,v)}$ 中所有有效拐点构成集合 G_i , 有效中间点构成集合 M_i , 无效点集合为 D_i , 则笔画拐点、中间点计算过程描述如下。

(1) 计算中间点。对笔画 $S_k^{(u,v)}$, 必有 $p_u, p_v \in M_i$ 。对其任意一点 $P_l, u+N \leq l \leq v-N$, 计算向量夹角:

$$\theta_l = \angle(\overrightarrow{p_{l-N}p_l}, \overrightarrow{p_l p_{l+N}}), 0 \leq \theta < \pi \quad (3)$$

式中, N 为间隔点数。令中间点角度阈值为 θ_{lim1} , 若 $\theta_l \leq \theta_{lim1}$, 则有 $p_l \in M_i$, 否则 $p_l \in D_i$, 如图 1 所示。

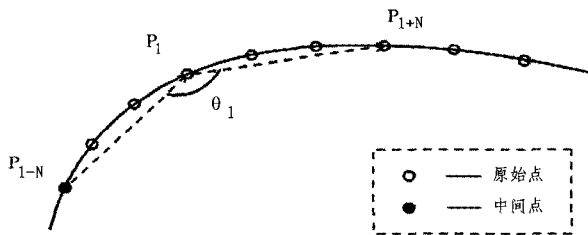


图 1 笔画中间点的计算

(2) 去除抖动。如果存在两点 $p_u, p_v \in M_i$, 且有 $u < v, v-u \leq N$, 则 $p_u \in D_i$ 。

(3) 计算拐点。记 M_i 中第 j 点为 m_j , 计算相邻 3 个中间点构成的两个向量之间的夹角:

$$\theta_{ij} = \angle(\overrightarrow{m_{j-1}m_j}, \overrightarrow{m_j m_{j+1}}), 0 \leq \theta < \pi \quad (4)$$

并通过设定阈值来判断 m_j 是否是一个合理拐点, 若有 $\theta_{ij} \leq \theta_{lim2}$, 则有 $m_j \in G_i$, 如图 2 所示。

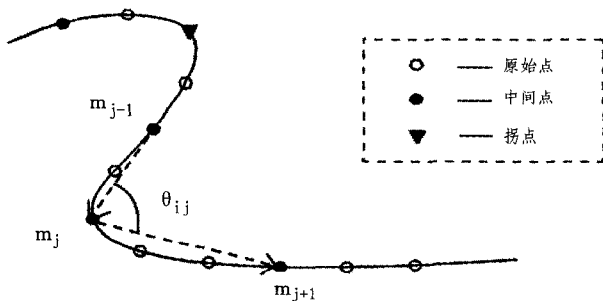


图 2 笔画拐点的计算

2.2 笔画起点、拐点、终点的书法特征处理

从书法的角度来讲, 汉字书写过程可大致分为起笔、运笔、转笔、收笔 4 个过程。其中汉字拓扑结构特征主要通过运笔过程体现, 而汉字的书法特征主要蕴含在起笔、转笔、收笔过程中。上文所描述的字形骨架计算方法主要针对运笔过程, 得到的字形骨架只保留了运笔过程中具有一定变化趋势的点, 而对起点、拐点、终点处的细节信息仍有欠缺, 因此不足以涵盖所有书法特征。为实现书法效果美化, 必须对起、转、收笔 3 个过程进行处理, 其分别对应于手写汉字数据中的起点、拐点、终点。

为解决该问题, 可在以上 3 类数据点附近进行细化, 将原本舍弃的无效点恢复, 使信息更加完善。处理过程可描述如下。

(1) 起点处理。对笔画 $S_k^{(u,v)}$ 的起点, 若其满足任意的 $p_j \in D_i, u < j \leq u+N$, 则令 $p_{u+N} \in M_i$ 。

(2) 终点处理。对笔画终点, 若其满足任意的 $p_k \in D_i, v-N \leq k < v$, 则令 $p_{v-N} \in M_i$ 。

(3) 拐点处理。对笔画 $S_k^{(u,v)}$, 若有点 $p_j \in G_i$, 且对任意点 $p_k, j-N \leq k \leq j+N, k \neq j$ 满足 $p_k \in D_i$, 则令 $p_{j-N} \in M_i, p_{j+N} \in M_i$ 。

通过以上方式, 可将原本舍弃掉的点恢复至中间点, 从而保留了更多的细节信息。经过以上处理, 字形骨架最终由所有笔画的拐点和中间点组成。

3 笔画骨架的书法效果建模

毛笔书法中, 正楷字体以其优雅敦厚、严谨的特点成为书法样本。而正楷字体由于其起、转、承、合的约束过强, 不易把握, 因而行楷成为书写中的常用字体。因此, 本文以毛笔行楷字体作为研究对象, 研究在字形骨架基础上进行毛笔书法美化的方法。相对于正楷, 毛笔行楷字的笔画间连接和笔画简化较多, 使得运笔更加圆滑灵活; 行楷的起笔以露锋居多, 起笔尖锐; 收笔则以藏锋居多, 有变捺为点的倾向。针对这 3 条特征, 本文提出利用笔触特征三角形算法对毛笔运笔时产生的笔触轮廓建模, 将手写字按照字形骨架点中的所有点划分为若干段, 依次在每一段上添加首尾相接的笔触特征三角形, 然后利用 B-样条曲线对产生的轮廓进行平滑, 最终得到具有毛笔书法效果的美化字。

3.1 笔触特征三角形建模

采集得到的字形骨架能够反映写字过程中的结构特征和书法特征, 模拟毛笔书法效果需要进一步为字形骨架添加笔画宽度信息, 产生毛笔运笔时留下的轮廓效果, 也即笔触的书法效果建模。

针对上文所述的行楷字体起笔露锋、运笔圆润、收笔藏锋的特点, 本文提出利用笔触特征三角形方法对书法效果建模, 即通过将字形骨架中每两点之间笔段的宽度变化表示为三角形, 来实现宽度信息的模拟。为方便描述, 将一个笔画的字形骨架点集记为 $K = (k_1, k_2, \dots, k_m)$, 则笔触特征三角形生成方法可描述如下。

(1) 笔触特征三角形底边。记 $k_i k_{i+1} (i \neq 1, i \neq m-1)$ 为 K 上任意两个连续的非首、尾点构成的线段, 则笔触特征三角形以该两点为底边, 顶点在线段 $k_{i-1} k_i$ 上, 如图 3 所示。

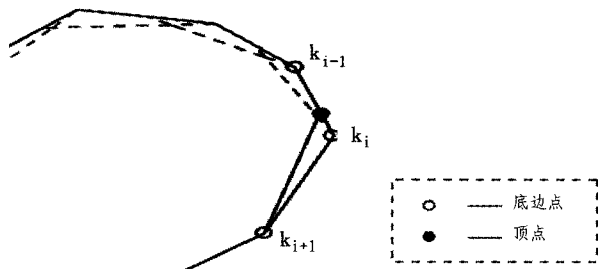


图3 笔触特征三角形:底边建模

(2)起笔顶点方向参数 δ 。对于起笔 k_1k_2 ,笔触特征三角形仍以两点为底,顶点方向按照 k_1k_2 斜率划分为横、竖、撇、提、捺5类,其中横、捺、提情况下令 $\delta=1$,顶点在底边所确定的轴向左侧;竖、撇情况下令 $\delta=-1$,顶点在底边轴向右侧,如图4所示。

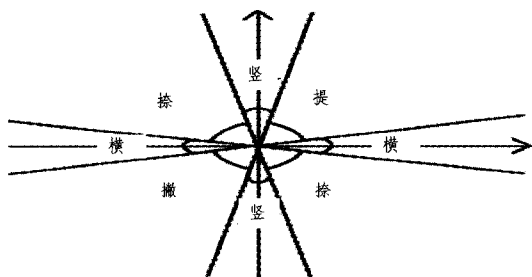


图4 笔触特征三角形:笔划起始斜率划分

(3)起笔顶点位置。在给定 δ 的情况下,笔触特征三角形的顶点位置可由起笔宽度参数 γ 和笔锋参数 φ 表示,其中 γ 为顶点到底边的距离。记点 k' 为顶点在底边的投影点,笔锋参数 $\varphi=k_1k'/k'k_2$,如图5所示。

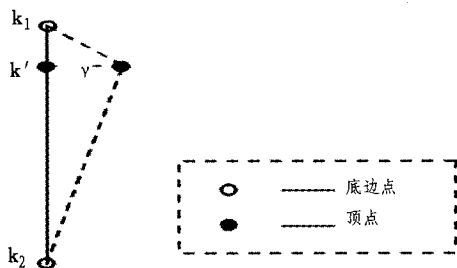


图5 笔触特征三角形:起笔顶点位置

(4)中间笔宽度参数 τ 的建模。对于笔画的中间笔段 $k_i k_{i+1}, i \neq 1, i \neq m-1$,其笔触特征三角形顶点在线段 $k_{i-1}k_i$ 中,记为 k' ,则 $\tau=k'k_i$,如图6所示。

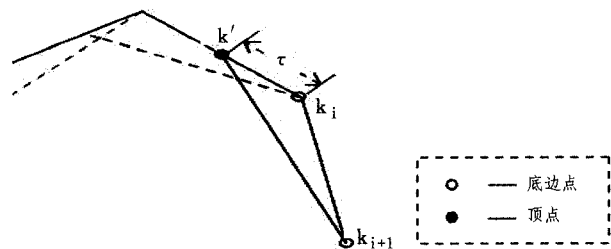


图6 笔触特征三角形:中间笔顶点位置

对于一个手写汉字的字形骨架,给定其五元组参数 $(K, \delta, \gamma, \varphi, \tau)$,即可确定整个骨架子的笔触特征三角形,从而实现

对毛笔书法效果的建模。该方法的优点在于,给定任意字形骨架,美化后的笔画宽度、起笔形态有规律、运笔收笔大小形状基本相同。而且,由于笔段间采用拼接方式形成宽度,使得该笔段的内外边缘都会呈现一种均匀变化的姿态。相对地,在其他有着突兀转笔的笔段则会因为笔触特征三角形形状的变化而有着明显的转笔特征。同时,由于每个笔段都是一个三角形,因此在收笔处都将会有一个变细的笔锋变化,而以横、竖、点为收笔的笔画因为有顿笔的存在而消除了这种笔锋的变化。

3.2 基于B-样条曲线模型的笔画轮廓平滑算法

在为字形骨架添加笔触特征三角形后,生成笔触宽度信息。每个笔画的宽度信息都可表示为一条封闭轮廓线。此时,笔画轮廓线仍然由若干三角形首尾相接构成,需要进一步对轮廓进行平滑,以获得起笔笔锋特征和圆润运笔特征。在这方面,文献[13]提出采用区间样条方法绘制毛笔笔触,可实现有粗细变化的毛笔笔触绘制。与之不同,本文选择B-样条曲线作为平滑方法,直接对笔画轮廓进行建模。B-样条曲线由Bezier曲线改进得到,与Bezier曲线类似,B-样条曲线可通过控制点位置来控制曲线形状,此外还解决了Bezier曲线生成复杂曲线时需要拼接的问题,这就使得平滑地绘制一些连笔的复杂笔画成为可能。当B-样条曲线的一个控制点发生改变时,只影响曲线的局部形状而非整体,这对于描述运笔过程中细节的书法效果非常重要。在本文中,用底边点和定点作为控制点集合。

通常,B-样条曲线按照其逼近的控制点分段定义。对含有 m 个控制点的曲线 $Q=(q_1, q_2, \dots, q_m)$,定义 $m+1$ 段 n 次B-样条曲线,将其表示为:

$$P_{i,n}(t) = \sum_{k=0}^n P_{i+k} F_{k,n}(t) \quad (5)$$

式中, $0 \leq t \leq 1, i=0, 1, 2, \dots, m, P_{i,n}$ 为控制点, $F_{k,n}(t)$ 为B-样条基函数,也称B-样条分段混合函数。基函数的阶次决定了B-样条曲线的阶次。随着阶次的增加,曲线对于控制点构成的封闭曲线逼近效果越差,而对相连曲线段的连续性越好。模拟笔锋变化时,要求其尽可能与拟合前形状逼近,又需要产生运笔的平滑效果,因此选用2次B-样条曲线,可表示为:

$$P_i(t) = F_{0,2}(t)P_i + F_{1,2}(t)P_{i+1} + F_{2,2}(t)P_{i+2} \quad (6)$$

选择基函数为:

$$F_{0,2}(t) = \frac{(t-1)^2}{2} \quad (7)$$

$$F_{1,2}(t) = \frac{(-2t^2 + 2t + 1)}{2} \quad (8)$$

$$F_{2,2}(t) = \frac{t^2}{2} \quad (9)$$

4 实验结果和分析

本文设计实验,测试字形骨架计算和毛笔书法效果美化方法的性能。实验采用脱机方式进行,利用电磁屏手写板采集行书、隶书手写字作为测试样本,得到原始书写数据。分别设计实验,测试在不同参数条件下的字形骨架计算和美化效果。

4.1 行书书法效果美化实验

图7给出了采集得到的行书字体以及美化各个步骤的运

行结果。手写汉字自上而下依次是:思、望、永、安、秋;从左至右依次给出原始采集字、字形骨架、笔触特征三角形、美化结果。



图7 行书法效果美化效果

原始数据经过数据简化后,冗余数据大大减少,字形骨架与原书特征基本一致,在添加笔触特征三角形后,字形骨架已经初步具有书法特征,经过平滑后,得到书法效果的轮廓信息。与原字相比,书法美化后的结果中,书法效果较为明显。

4.2 隶书书法美化效果实验

本文描述方法不仅适用于一般手写体行书,对隶书字体也有较好的适应性,图8中给出了对一句手写隶书的古诗字体所进行的美化效果。可以看出,美化结果对于整句诗句具有良好的—致性,能够较好地体现出书写者的书写特点和意图,在细节处理方面能够较好的把握书写中的书法特点。



图8 隶书法效果美化效果

结束语 本文实现了一种手写汉字书法效果美化方法。该方法能够为计算机采集得到的手写字添加具有书法特征的笔触特征三角形,产生带有宽度的笔画轮廓,进而通过平滑算

法,生成书法效果美化字,方法简单可靠,对行书、隶书字体均有较好地适应性;整个算法只需用到骨架汉字的坐标信息,对设备依赖性要求较低,能基本满足目前电子设备中采集手写字的需求;提出的笔触特征三角形算法能够较好地模拟毛笔运笔宽度信息,产生书法美化效果。在未来的工作中,将尝试改进数据采集模式,利用更多信息如压力和速度信息,对起笔、行笔和收笔的笔触特征三角形进行参数自适应调整,使该算法具有更好的拓展性。

参考文献

- [1] 王钰旋,庞云阶. 一个计算机书法系统 CCCS[J]. 计算机辅助设计与图形学学报,1991,3(1):35-40
- [2] 王钰旋,庞云阶. 一个笔划填充算法及其在计算机书法中的应用[J]. 计算机辅助设计与图形学学报,1994,6(3):168-173
- [3] 方遼,吴涛. 书法艺术的计算机模拟及其实现[J]. 模糊系统与数学,1996,10(1):71-74
- [4] 王征,孙济洲,孙美君,等. 采用自回归模型的计算机书法的研究[J]. 工程图学学报,2006,5:38-43
- [5] Wong S T S, Leung H, Ip H H S. Model-based analysis of Chinese calligraphy images[C]// Proceedings of the 9th International Conference on Information Visualization. London, UK, 2005:221-226
- [6] Yu Jin-hui, Peng Qun-sheng. Realistic synthesis of Cao Shu of Chinese calligraphy[J]. Computer & Graphics, 2005, 29(1): 145-153
- [7] 董军,徐森,潘云鹤. 基于统计模型的书法创作模拟[J]. 计算机学报,2008,31(7):1276-1282
- [8] 陈颖,朱福喜. 结合模糊支持向量机的曲线类比在书法风格仿真中的研究[J]. 中国图像图形学报,2010,15(2):340-345
- [9] 夏伟平. 联机手写体汉字楷书美化方法研究[D]. 广州:华南理工大学,2010
- [10] 刘丽娟. 个性化联机手写汉字隶书美化方法研究[D]. 广州:华南理工大学,2011
- [11] Douglas D, Peucker T. Algorithms for the reduction of the number of points required to represent a digitized line or its caricature[J]. The Canadian Cartographer, 1973, 10(2): 112-122
- [12] 黄襄念,程萍,杨波,等. 自然手写汉字预处理子系统[J]. 重庆大学学报:自然科学版,2000,23(4):33-37
- [13] Su Sara L, Xu Ying-qing, Shum H-Y, et al. Simulating artistic brush strokes using interval splines[C]// Proceedings of the 5th IASTED International Conference on Computer Graphics and Imaging. Kauai, Hawaii, 2002:85-90

(上接第269页)

- [14] Ganter B, Wille R. Formal concept analysis, mathematical foundations[M]. Springer Verlag, 1999
- [15] Nilsson J F. A logico-algebraic framework for ontologies[C]// Proceedings of the 1st International OntoQuery Workshop. 2001:32
- [16] 时百胜,余泓. 概念知识表示和推理[J]. 小型微型计算机系统,

2006,27(9):1618-1622

- [17] 敦仁,周立柱. 概念代数—新一代数据库系统的理论[J]. 计算机研究与发展,1996,33(1):32-38
- [18] Wang Y. Cognitive informatics: towards future generation computers that think and feel[C]// Proceedings of the 5th IEEE International Conference on Cognitive Informatics, ICCI 2006. 2006:3-7