

基于主元分析的空间手写平面化预处理技术

谭啸峰, 沈海斌*

(浙江大学 超大规模集成电路设计研究所, 浙江 杭州 310027)

摘要:为了将空间手写识别回归为平面手写识别问题,提出了一种新型的空间手写识别平面化预处理技术,采用基于主元分析(PCA)的投影算法对空间手写字符轨迹进行了平面化处理。该算法中,投影平面的确定仅依赖于手写字符轨迹采样点集本身的统计特征,故当书写角度发生变化时,投影平面也会随之产生适应性变化,以产生最佳的投影效果;最后在实验中,对比了指定初始平面投影法和主元分析投影法在不同书写角度下的投影效果,实验结果直观地证明了该投影算法的有效性。

关键词:空间(3D)手写识别;主元分析;平面化;预处理

中图分类号:TP301

文献标志码:A

文章编号:1001-4551(2011)08-0965-05

Flatten preprocessing technology of 3D space handwriting based on PCA

TAN Xiao-feng, SHEN Hai-bin

(Institute of VLSI Design and Research, Zhejiang University, Hangzhou 310027, China)

Abstract: In order to change 3D handwriting problem to 2D handwriting problem, a novel flatten preprocessing technology of 3D space handwriting was proposed, the locus of 3D handwriting was flatten by the algorithm based on principal component analysis (PCA). In the algorithm, the projection plane was only determined by statistic character of the locus. The projection plane would adapt to the degree change of handwriting to generate the best projection. At last, the projecting results of both fixed plane projecting and PCA analyzed plane projecting are compared to prove the effectiveness of PCA based projecting algorithm.

Key words: 3D space handwriting recognition; principal component analysis (PCA); flatten; preprocessing

0 引 言

手写识别技术已经发展了十几年,大体可分为脱机识别和联机识别^[1]。脱机手写识别的研究对象是字符图像,是个静态的二维点阵,而联机手写识别的研究对象是按时间顺序依次输入的笔迹点坐标,不但有空间信息,还有时间信息。因为联机识别具有更多的信息,在即时输入的场合下,应首选联机识别,本研究主要讨论联机识别。

联机手写识别技术按照书写空间的不同可以分为平面手写识别和空间手写识别,其中平面手写识别已经发展了十几年,一般是用手写笔在书写板上书写,因为其接近人们传统的书写习惯并且技术发展相对成熟,已经成为现阶段手写识别的主流输入方式。虽然平面手

写识别符合人们的传统手写习惯,但也具有与生俱来的缺陷:①手写笔和手写板之间不可避免的有直接接触,从而造成设备磨损;②输入设备既需要手写笔又需要手写板,增加了设备成本,同时也造成手写设备体积大难于携带;③书写者需在一个固定书写平面上书写,对书写姿势有约束,容易产生疲劳。因此,空间手写识别的出现和发展成为了手写领域的必然趋势^[2]。

目前对空间手写识别的相关研究比较少,相应的专用于空间手写识别的预处理方法也不多见,较之平面手写识别成熟的预处理技术^[3]研究基本处于相对空白的状态。一般的预处理包括很多环节^[4],但由于空间手写有其自身的特点,其预处理过程中比平面手写预处理多了一个重要环节,即将手写轨迹由空间字符投影为正确的平面字符,从而回归为平面手写识别

收稿日期:2010-12-20

作者简介:谭啸峰(1986-),男,湖南长沙人,主要从事电路与系统方面的研究。E-mail:tanxf@vlsi.zju.edu.cn

通信联系人:沈海斌,男,副教授。E-mail:shb@vlsi.zju.edu.cn

问题进一步处理。在手写、手势识别的投影技术方面,文献[5]在研究空间手写输入设备的基础上,利用空间手写笔初始位置来确定投影平面,简单易行,但是要求书写者在书写过程中保持对初始平面的稳定;文献[6]在研究人机交互的手势识别过程中,使用了仿射曲线矩来尽量消除面对摄像头的不同视角对手势识别的影响,一定程度上放宽了对输入平面不变的要求,但是其前提在于手势的深度变化相对其到摄像头的距离比较小,仍然要求使用者在正对摄像头的一个不大的视角范围内动作,不能充分发挥空间手势识别的自由性。

本研究提出了一种基于主元分析的空间手写平面化预处理技术,运用该技术使得投影平面不由手写笔初始位置确定,而是由字符采样点集本身的特征确定,对字符书写角度变化有较好的适应性,在不同的书写角度下都能获得比较理想的投影效果。

1 空间手写字符特点

空间手写、手势识别按照输入设备可以分为两种:一种基于视频采集;另一种基于加速度传感器的手写笔输入。前者成本低,在摄像头普及的今天甚至不需要另购专门的输入设备,但是该方法精度差、算法复杂,仅适合相对简单的手势识别,无法满足手写字符识别更精确细微的要求。因此本研究的预处理方法是基于加速度传感器的输入设备^[7-8]所采样的字符信息。

基于加速度传感器的手写输入设备采集到的字符信息是一个坐标序列,即将字符笔画采样点的三维空间坐标按时间顺序排列所得。空间手写识别与平面手写识别的最大区别在于,平面手写轨迹可以直接作为识别器的输入,而空间手写轨迹是一条空间曲线,需要先投影成平面轨迹再作识别。如果投影平面选取不当,则得到的轨迹可能跟书写者的意图相去甚远。空间手写符及各方向投影如图1所示。图1(a)的三维轨迹,是用加速度传感器采样书写者垂直于地面进行空间手写时得到的数字“8”的轨迹。其余3图是分别从正面、侧面和上面进行投影得到的效果,可见在该例中只有从正面投影才能得到最为理想的“8”字,从上面和侧面投影都无法识别。因此选取一个合适的投影面是空间手写平面化预处理必须解决的关键问题,为此需要考察空间手写字符轨迹的特点。

为了了解空间手写字符轨迹的特点,本研究对0~9这10个数字字符进行了手写轨迹采集,每个字符采集10个样本,定义最能体现手写轨迹正确字型的方

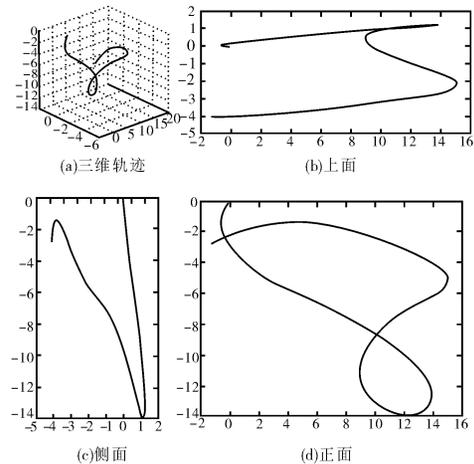


图1 空间手写字符及各方向投影

向为正面,测量空间字符的宽度和高度,再从垂直方向测量空间字符厚度,对每种字符各方向尺寸的10组样本数据取平均值,如表1所示。

表1 字符空间尺寸实验统计平均值

数字字符	各方向尺寸平均值		
	宽度/cm	高度/cm	厚度/cm
0	8.18	9.94	3.77
1	2.15	9.69	2.20
2	20.56	17.66	4.42
3	9.24	22.33	4.33
4	14.04	16.39	7.31
5	8.94	14.94	4.36
6	6.86	12.39	2.43
7	9.37	13.74	2.03
8	13.19	16.30	3.85
9	8.58	22.56	4.54

通过观察和分析上述实验数据容易看出,字符轨迹在空间不同方向上的尺寸有明显差异,除数字“1”以外,所有空间字符轨迹的厚度平均值明显小于宽度、高度的平均值。这是因为书写者在进行空间手写输入时会延续日常习惯,在一个假想平面上书写,这样字符轨迹就会体现出两个特点:一方面由于没有实际书写平面的约束,提笔、按压这类控制笔画轻重的习惯性书写动作以及人手在书写过程中本身存在不可避免的轻微抖动,这些都会造成字符轨迹包含垂直于假想书写平面的偏移量,从而造成手写曲线空间化;但是另一方面,习惯会使书写者尽量保持字符轨迹贴近其假想书写平面,只要字符本身尺寸超过一定大小,则偏移量总是小于其自身尺度(在空间手写过程中,字符大小的要求一般是容易满足的),即空间字符一般总是“扁平”的。

基于以上实验结果,本研究可以确立这样一个投影原则,即:只要找到使得字符轨迹最为“扁平”的观

察方向并向垂直于它的方向投影,就可以得到理想的投影效果;而对于字符“1”,其厚度、宽度相差不大,向这两个方向投影出来的字型区别不大,都可以识别为“1”,运用上述投影原则同样没有问题。

2 基于主元分析的投影方法

确立了上述投影原则后,需要找出一种数学方法来实现它。经考察,主元分析法能够很好地达到这一目的,以下是主元分析法的基本原理和算法实现。

2.1 主元分析法

主元分析法(PCA)是一种数学变换的方法,它把给定的一组相关变量通过线性变换转换成另一组不相关的变量,这些新的变量按照方差依次递减的顺序排列。主元分析法的几何意义是,对以 X_1, X_2, \dots, X_p 为坐标轴的原坐标系进行旋转,得到一个以 X_1', X_2', \dots, X_p' 为轴的新坐标系,新坐标轴指向的方向代表数据变性大小依次递减的方向^[9]。而在本研究中,数据变性最大的方向对应于字符轨迹具有最大尺寸(最宽、最高)的方向;数据变性最小的方向对应于字符轨迹具有最小尺寸(最“扁平”),即理想投影面的法向量方向。因为采样点集实质上是三维坐标点集,只要对该点集进行主元分析,将以 X_1, X_2, X_3 为轴的原始三维坐标系变换成 X_1', X_2', X_3' 为轴的新坐标系(其中 X_1' 指向数据变性最大的方向, X_2', X_3' 所指方向数据变性依次递减),则 X_1', X_2' 就是体现字符尺寸最大的两个方向, X_3' 就是投影面法向量方向。

基于上述方法得到的手写字符曲线的投影面完全由点集本身的特征确定,而不是把点集向一个初始指定的面上投影。因此,在书写者姿势变动的情况下,投影面也会随之向最合适的方向变动,从而具有较好的适应性。

2.2 算法实现

在每一个采样时间点上,手写输入设备都记录了一组三维坐标,每一组可以写成一个三维列向量,这是测量所得的空间坐标,如果有 n 个采样点,则该点集可以表示为 $P = [p_1, p_2, \dots, p_n]^T$

$$P = [p_1, p_2, \dots, p_n]^T = \begin{bmatrix} x_1 & x_2 & \dots & x_n \\ y_1 & y_2 & \dots & y_n \\ z_1 & z_2 & \dots & z_n \end{bmatrix}$$

现需将每个点坐标 $p_i = [x_i, y_i, z_i]^T$ 转换为按数据变性依次递减的方向为三轴方向的空间坐标系所表示的坐标 $q = [x_i', y_i', z_i']^T$,坐标变换形式为 $Q = CP$,其中 C 为变换矩阵,则问题转化为如何求解变换矩阵

C 。

考察协方差矩阵 $Conv = [p_1 - \bar{p}, p_2 - \bar{p}, \dots, p_n - \bar{p}] \cdot [p_1 - \bar{p}, p_2 - \bar{p}, \dots, p_n - \bar{p}]^T$,其中 \bar{p} 是采样点集的均值,求它的特征值、特征向量 $Conv \cdot v_k = \lambda_k \cdot v_k, \{k \in 1, 2, 3\}$,则对应的3个特征向量 v_k 构成了新的三维直角坐标系 x', y', z' 轴方向,如果有 $\lambda_1 \geq \lambda_2 \geq \lambda_3$,则经过 \bar{p} 点的平面 $(p_i - \bar{p}) \cdot v_3 = 0$ (p_i 为空间中任意点)所表示的几何意义是:点集 P 中所有点到该平面的距离的平方和最小, v_3 就是数据变性最小的方向;换句话说,从垂直于 v_3 的角度观察点集 P 最为“扁平”^[10]。因此可以确定: v_1, v_2 确定的平面就是投影面, v_3 即投影方向也即 z' 轴方向。

本研究用主元分析法确定方向为 v_3 后,已经可以投影出字形准确的字符,但是由于手写字形长宽尺寸不确定性大,而 v_1 的方向实际是指向投影后的二维字符的最宽方向,如果直接把 x', y' 轴方向定为 v_1, v_2 的方向就无法保证字符的方向按阅读习惯正确摆放,这样投影出来的字符往往是经过一定旋转的字符。而在原 x, y, z 坐标系中,字符的方向是正确的,所以应将 x' 方向取为与原 z 和新 z' 同时垂直的方向, y' 自然取为与 x', z' 都垂直的方向,这样投影在 x', y' 平面上的字符就是正确摆放的。

本研究在确定 x', y', z' 方向后,由 $[x, y, z]^T = [x', y', z']^T \cdot C$ 可求出变换矩阵 C ,从而可以由 $Q = CP$ 得出采样点集变换后的坐标。在新坐标系中,投影操作已经简化为去掉 z' 轴方向分量, x', y' 平面上就已经是正确的字符图形。

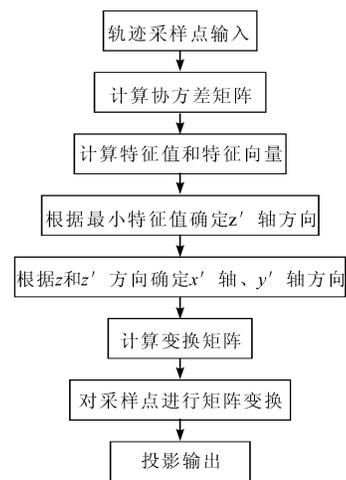


图2 平面化预处理算法流程图

综上所述,整个平面化预处理过程可以归结为如图2所示流程。

3 性能分析

本研究现在对初始平面投影法和主元分析投影法的效果进行比较,指定初始平面为地面,考察相对于地面采用不同书写角度时,两种方法的投影结果有什么不同。本研究在实验中采用 ADI 公司的三轴加速度传感器评估板 EVAL - ADXL345Z - M(如图 3 所示)进行轨迹捕捉。ADXL345 是一款超低功耗 3 轴加速度计,分辨率高(13 位),测量范围达 $\pm 16\text{ g}$ 。数字输出数据为 16 位二进制补码格式,可通过 SPI(3 线或 4 线)或 I²C 数字接口访问。

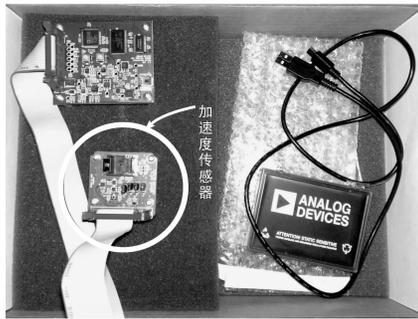


图 3 加速度传感器评估板

本研究使用该评估板时,将其连接至计算机,打开图形界面,手持加速度传感器在空中进行书写,如图 4 所示,可以通过图形界面观察三轴加速度的波形(图中粗实线、细实线、虚线三组波形分别代表 x 、 y 、 z 三轴方向的加速度变化)并且保存采样到的加速度数据。用 Matlab 进行数据处理时,先将各方向加速度两次积分得到字符的空间位移信息,再运用上文所述主元分析算法进行投影,对比效果。

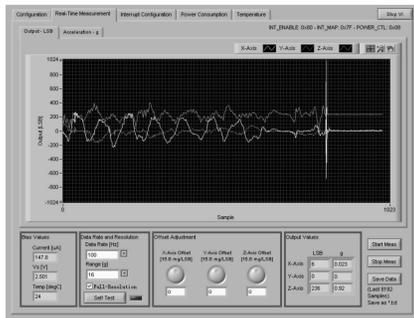


图 4 加速度传感器采样图形界面

在平行于地面进行书写时,两种投影法选取的投影平面相同,不用比较。本研究主要针对以下两种情况对 0~9 这 10 个空间手写数字字符进行比较:①假想书写平面与地面呈小于 90° 的倾角;②假想书写平面与地面

垂直。通过这两种情况比较,可以看出投影效果随倾角变化而发生的变化。实验结果如图 5、图 6 所示。从上述两组实验结果可以看出,当假想书写平面与地面呈小于 90° 的倾角时,初始平面投影法的投影形变已经比较

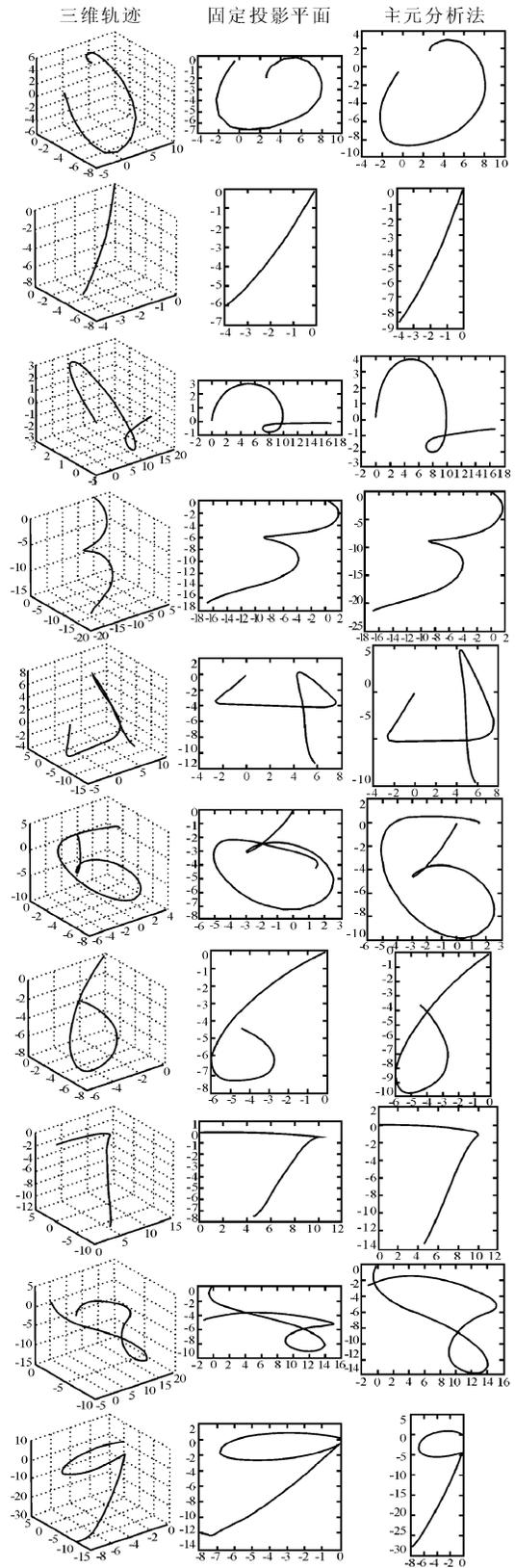


图 5 相对地面倾斜时书写的投影效果对比

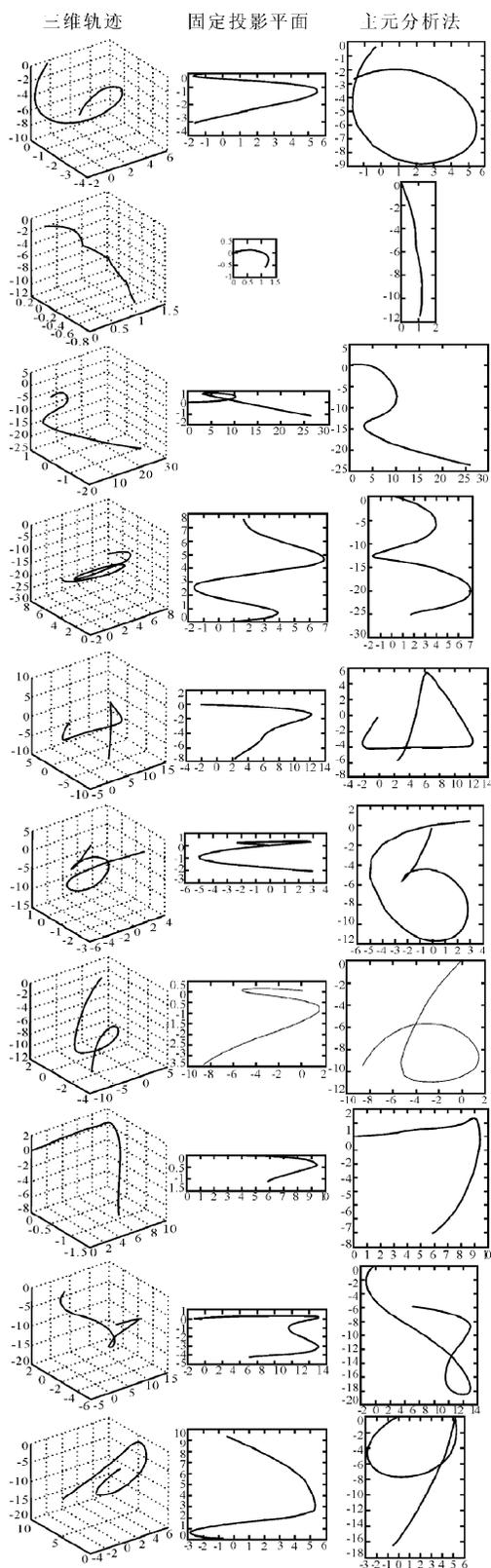


图6 相对地面垂直时书写的投影效果对比

严重,主要体现在字型被纵向压缩了,部分字符已经比较难识别,如数字“5”、“8”等;而当书写平面与地面垂直时,采用该投影方法则完全得不到正确的字型。而本研究采用主元分析法的投影方法,在两种情况下都表现出良好的投影性能。实验结果证明了本研究基于主元

分析法的投影算法完全不依赖于初始投影平面的选取,只根据空间手写轨迹采样点集本身的特征来确定投影平面,对书写者动作的变化具有良好的适应性。

4 结束语

本研究提出了一种新的空间手写平面化预处理技术,采用基于主元分析(PCA)的算法确定投影平面,改善了书写角度变化时的空间手写字符轨迹的投影效果。

在今后的工作中,仍有以下几个方面可以改进:①本研究中的投影仅用到了位移信息,加速度和速度信息在处理连笔、拐点和手写开始和结束判定时,应该还可以发挥应有的潜力;②实验时可以采用更多种类和数量的样本,如手写中文字符和手写英文字符等等,以把握各种手写字符的更多细节特征。

参考文献(References):

- [1] 赵继印,郑蕊蕊,吴宝春,等. 脱机手写体汉字识别综述[J]. 电子学报, 2010, 2(2):405-145.
- [2] KIM D H, CHOI H, KIM J H. 3D Space Handwriting Recognition with Ligature Model[C]. Proceedings of UCS 2006. Seoul: [s. n.], 2006:41-56.
- [3] 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.
- [4] 马立权,李 维,蔡韩辉,等. 手写数字识别中的预处理技术研究[J]. 仪器仪表学报, 2001, 22(3):263-265.
- [5] 王玮,李德胜,高宗余,等. 空中手写笔迹检测系统[J]. 光学精密工程, 2009, 17(6): 1286-1292.
- [6] 顾立忠. 基于表现的手势识别及人机交互研究[D]. 上海:上海交通大学自动化系,2008.
- [7] ZHANG Shi-qi, YUAN Chun, YAN Zhan-gu. Wave: Accelerometer-based Personalized Gesture Recognition and Its Applications[C]//Proceedings of the 2009 IEEE International Conference on Pervasive Computing and Communications. PERCOM: [s. n.], 2009:1-9.
- [8] BOUWHUISEN C F, MEULENBROEK R G J, THOMASSEN A J W M. A 3D motion-tracking method in graphonomic research: possible applications in future handwriting recognition studies [J]. **Pattern Recognition**, 2002, 35(5): 1039-1047.
- [9] [美] JOHNSON R A, WICHERN A W. 实用多元统计分析[M]. 陆璇,赵衡管,叶 俊,等译. 4版. 北京:清华大学出版社,2005.
- [10] PAULY M, GROSS M, KOBELT L P. Efficient Simplification of Point-Sampled Surfaces[C]//Proceedings of VIS 2002, IEEE. Boston: [s. n.], 2002:163-170.

[编辑:李 辉]