

基于摄像头的手写输入法^①

邓俊

(江苏科技大学 计算机科学与工程学院, 镇江 212003)

摘要: 手写输入法有显著优点, 但目前手写输入法普遍依赖于“手写板”这种特殊设备。改进了目前手写输入法设计思路, 利用摄像头拍摄书写的过程, 生成手写字体图像, 为识别成文字做准备。该输入法特点就是不用“手写板”, 而充分利用了日益普及的摄像头。实验表明, 该方法有较强的可行性和可用性。

关键词: 摄像头; 手写; 输入法

Handwriting Input Method Using Camera

DENG Jun

(School of Computer Science and Technology, Jiangsu University of Science and Technology, Zhenjiang 212003, China)

Abstract: Handwriting input method has obvious advantage. However, current handwriting input depends on tablet. In this paper, we improve the design idea. We get an image of handwriting font shooting writing process by camera, which could get ready for recognizing character. The advantage of the handwriting input does not depend on tablet and fully use of camera which become universal. The experiments demonstrate the feasibility and usability.

Keywords: camera; handwriting; input method

1 引言

目前个人电脑普遍使用键盘来输入文字, 对于不喜欢使用键盘或者不熟悉使用各种中文输入法的人而言, 手写输入法是很好的选择; 然而目前手写输入法依赖于“手写板”这种特殊的设备, 这会给用户增加了经济负担。本文提到的手写输入法, 唯一依赖的硬件设备是摄像头, 而摄像头目前已经广泛应用在个人电脑上, 因此这套手写输入法不会给用户增加额外的经济负担, 充分利用了现有资源。

这套输入法是用摄像头记录下“笔头”的运动轨迹, 然后在对轨迹进行识别的方法。该方法示意图如图1。该输入法的主要构成是: 摄像头和“笔”。这里的“笔”并不是真正意义上的笔, 只要是顶端有某种与周围不一样颜色的杆就可以。杆的一端也就是“笔头”, 要求有一厘米的某种鲜明的颜色, 这样便于摄像头跟踪。例如这样的“笔”可以是顶端有红色的筷子。

该输入法工作过程如下: 当用户用这样的笔在桌面上书写时, 摄像头从适当的角度去拍摄书写过程,

记录笔头的位置。书写结束后, 把这个手写字体交给识别模块, 我们就可以得到一个文字。

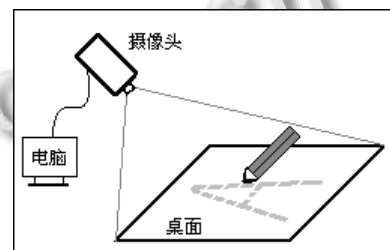


图1 系统示意图

2 手写字体的轨迹记录

在该研究中使用了 DirectShow 驱动摄像头及截取图片^[1]。截获的每一帧图像只能拍下“笔头”在某一点的位置, 把若干帧原始图像中“笔头”的位置叠加在同一个缓存上, 直到一个字书写的完成, 这样缓存上就保存了一个手写字体。

事实上直接使用原始图像的效果并不好, 因为在

^① 收稿时间:2010-08-22;收到修改稿时间:2010-09-22

于原始图像上噪声是不可避免地存在的, 另外还有“笔头”反光等因素; 因此并不能很好地形成一个字体。图 2(a)是未经处理的英文字母“a”的手写字体。从图中可以看出, 图片的下方存在若干斑点, 这样的噪声是我们不希望看见的。还有字体本身的轨迹也不清晰的问题, 这也会影响最终的识别。鉴于上面的两点不足, 对记录过程进行了改进。改进后效果如图 2(b)所示。



图 2 (a) 未经处理的字体 (b) 处理后的字体

改进方法的主要思想是: 由于一帧图像中“笔头”的颜色在某处最集中; 用聚类的方法^[2]找出其位置, 然后在缓存对应的位置记录并增强该点, 然后与前一点相连。

2.1 “笔头”的识别方法

对“笔头”的跟踪是依靠对其特殊颜色的识别。在截取的一帧图像中, 如果某个像素灰度值在一定的范围内, 则认为这个像素是“笔头”的, 就准备对其进行聚类处理; 否则不对该点做任何处理。可描述如下:

$$I_{pen}(x, y) = \begin{cases} 1 & T_{Min} < I(x, y) < T_{Max} \\ 0 & \text{Otherwise} \end{cases}$$

这里 $I_{pen}(x, y)$ 表示图像中, 任意坐标处像素是否为“笔头”对应的灰度, 1 表示“笔头”对应的灰度, 该点将做聚类的样本, 0 表示其它的灰度, 对其不做处理; $I(x, y)$ 表示任意坐标处像素的灰度; T_{Max} 和 T_{Min} 是“笔头”灰度的上下阈值, 阈值的确定要依据笔头的颜色而定。

2.2 字体轨迹合成

合成字体轨迹的具体步骤如下:

(1) 图像中像素点的 $I_{pen}(x, y)$ 若为 1, 则当作一个聚类样本 X_i , 计算它与聚类中心 Z_j 的欧式距离 D_{ij} , $D(X_i, Z_j) = \|X_i - Z_j\|$, 若 D_{ij} 小于阈值 T , X_i 加入到 Z_j 中。

对图像中所有的像素都做这样的处理。具体算法步骤如下:

第一步: 任取一个样本 X_i 作为第一个聚类中心的初始值, 例如令 $Z_1 = X_1$ 。

第二步: 计算样本 X_2 到 Z_1 的欧式距离 $D_{21} = \|X_2 - Z_1\|$ 。若 $D_{21} > T$ 定义一个新的聚类中心 $Z_2 = X_2$; 否则把 X_2 归入 Z_1 为中心的聚类。

第三步: 假如已有聚类中心 Z_1, Z_2 , 计算 $D_{31} = \|X_3 - Z_1\|$ 和 $D_{32} = \|X_3 - Z_2\|$ 。若 $D_{31} > T$ 且 $D_{32} > T$, 则建立第三个聚类中心 $Z_3 = X_3$; 否则 X_3 归入 Z_1 和 Z_2 最近者。

以此类推, 直到所有的样本都被分析为止。

(2) 聚类处理后, 找出 $\text{Max}\|Z_j\|$, 记录该聚类的中心位置。

(3) 在后台缓存上, 在该点位置与前一点的位置之间“画”一条线。

经过上述步骤处理后, 过滤了噪声并增强了“笔头”的信息, 可以得到非常清晰的轨迹。但是到此不能算得到了手写字体, 因为其中夹杂了无用的轨迹。

3 手写字体生成

该输入法工作过程中有三个难点:

(1) 如何识别出用户开始书写文字。识别出书写的起始点, 下面才能开始记录。

(2) 如何识别出有效笔画和无效笔画。并非所有笔画都要记录下来, 我们只记录有效笔画, 这里举例解释这两个概念。例如书写汉字“二”, 第一笔从左到右, 认为是有效笔画; 而书写完第一笔准备写第二笔前, 要将笔头移至左下角, 这样的移动认为是无效笔画。如果把有效笔画和无效笔画都记录下来, 必然对后面的手写字体识别带来麻烦, 如图 5(a)所示, 这样的字体显得杂乱。

(3) 如何识别出用户书写文字结束。在结束点识别手写字体, 然后清空缓存。

3.1 书写开始与结束的识别

这里对书写起始点、结束点的判别通过“笔头”状态来判别^[3]。这里把“笔头”分为: 等待状态、准备书写和正在书写这三个状态。状态转换关系如图 3 所示。这种方法无形中给用户增加了条规则, 即: 书写前后都要有停顿, 以示书写的开始与结束。但这是一个简单易行的方法。

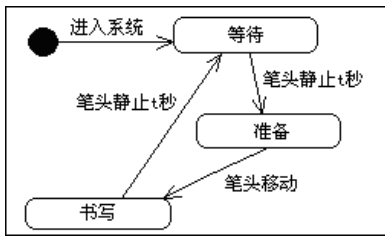


图3 “笔头”状态图

3.2 有效笔画和无效笔画的识别

对于有效笔画和无效笔画的识别，通过文字特征加以判别。汉字书写顺序有一个特点：从上到下、从左到右。因此当笔头往某些方向移动时我们就可以视为无效笔画，进而将其舍弃。

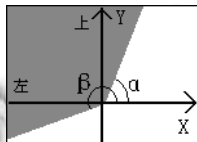


图4 灰色区域是汉字无效笔画方向

具体而言向左、向上、向左上方都是汉字书写中的无效笔画。数字和字母也有类似规律。图4中从中心原点移动到灰色区域是汉字的无效笔画。在本研究中 $\alpha = 70^\circ$ ， $\beta = 200^\circ$ 。

由图5(b)对比可知，该方法忽略了大部分无效笔画，字体更加明了，有助于后面的文字识别；而没有忽略无效笔画的字体显得杂乱。

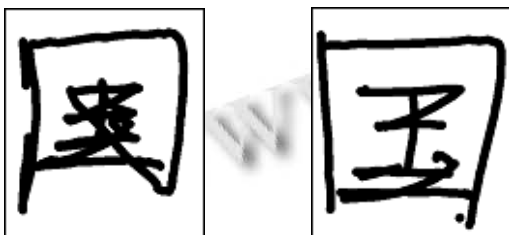


图5(a)含有无效笔画的汉字 (b)忽略无效笔画的汉字

4 实验结果

图6是在实验中记录下来的汉字手写字体：

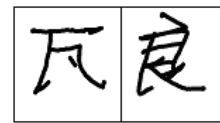


图6 实验中记录的手写字体

从上面几张图片可以看出，对于汉字能得到清晰的字体。图7是通过现有的软件^[4]，对记录下来手写字体识别的结果。由此可以证明，该方法得到的手写字体，是可以被识别成文字的。



图7 对手写字体的识别验证

5 结语

对于手写输入法而言，得到手写字体后，还有两项工作就是：进行手写字体识别和针对平台的输入法软件设计。这是相对成熟的技术，这里不做详解。

本文对利用摄像头实现手写输入法进行了研究，该研究为实现基于摄像头的手写输入法提供了前提条件。本文对于有效笔画和无效笔画的识别还有改进空间。一旦有了更好的识别方法，就有可能摆脱书写前后都要停顿的限制。

参考文献

- 1 陆其明.DirectShow 开发指南.北京:清华大学出版社, 2003.166-189.
- 2 齐敏,李大健,郝重阳.模式识别导论.北京:清华大学出版社,2009.13-36.
- 3 刘芳,林学闯,史元春.基于激光笔的远程人机交互技术.中国图像图形学报,2003,8(11):1356-1360.
- 4 静子.脱机手写体汉字识别演示程序.[2010-05-28].
http://www.pudn.com/downloads6/sourcecode/others/detail21066.html