



(12) 发明专利申请

(10) 申请公布号 CN 101930545 A

(43) 申请公布日 2010.12.29

(21) 申请号 200910146369.2

(22) 申请日 2009.06.24

(71) 申请人 夏普株式会社

地址 日本国大阪府

(72) 发明人 江淑红 吴波 吴亚栋 缪炜

李爱龙

(74) 专利代理机构 中科专利商标代理有限责任

公司 11021

代理人 王波波

(51) Int. Cl.

G06K 9/62 (2006.01)

G06K 9/72 (2006.01)

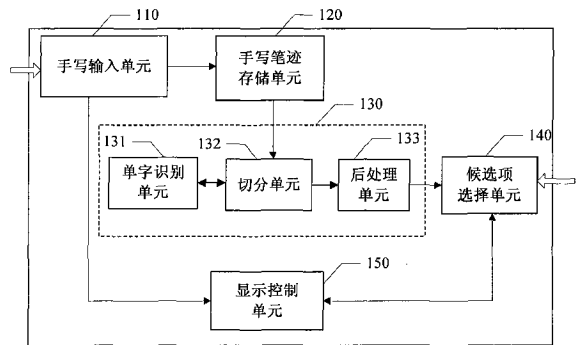
权利要求书 2 页 说明书 10 页 附图 9 页

(54) 发明名称

手写识别方法和设备

(57) 摘要

本发明公开了一种手写识别方法和设备,能够对用户连续手写输入的字符序列进行识别,为用户提供方便。该方法包括以下步骤:计算输入字符序列的各种特征,包括它的各种笔画组合的单字识别正确度特征和空间几何特征;采用概率模型,来计算各种切分方式中的各个笔画组合的切分可信度,概率模型中的各项系数,通过对已有样本的训练,采用参数估计方法来估计;在进行笔画组合的单字识别时,采用多模板匹配的方法来识别多种不同写法的字符;搜索出最佳的切分路径;进行后处理,对识别结果进行优化。本发明的方法和设备具有结构简单、硬件要求低、识别速度快,识别率高等优点,可在嵌入式系统等上实现。



1. 一种手写识别方法,用于对用户连续输入的字符序列进行识别,该方法包括步骤:  
基于不同笔画组合和对其所包含的笔画进行划分形成的“子笔画组合”的单字识别结果,计算与输入字符序列的不同笔画组合的单字识别正确度相关的特征;  
根据对不同笔画组合所包含的笔画进行划分形成的“子笔画组合”的空间几何关系来确定不同笔画组合的空间几何特征;  
基于与单字识别正确度相关的特征和空间几何特征,确定对输入字符序列的不同切分方式下各个笔画组合的切分可信度;  
基于所述切分可信度确定切分路径;以及  
向用户呈现:与确定的切分路径相关的字符序列识别结果。
2. 如权利要求1所述的手写识别方法,其中在得到所述单字识别结果时,采用多模板匹配方法来识别不同写法的字符。
3. 如权利要求1所述的手写识别方法,还包括步骤:利用字典数据库或者语言模型对字符序列识别结果进行处理。
4. 如权利要求1所述的手写识别方法,其中所述的与单字识别正确度相关的特征包括以下之一:“子笔画组合”合并后的单字识别正确度,“子笔画组合”合并后的单字识别正确度与“子笔画组合”的单字识别正确度之差,“子笔画组合”合并后单字识别的第一选择正确度与合并后单字识别的其他候选字正确度的比值;  
其中所述的笔画组合的空间几何特征包括以下之一:“子笔画组合”的外接矩形框的间隔,“子笔画组合”进行合并后的宽度,上一“子笔画组合”结束点与下一“子笔画组合”起始点之间的向量,上一“子笔画组合”结束点与下一“子笔画组合”起始点之间的距离,上一“子笔画组合”起始点与下一“子笔画组合”起始点之间的距离。
5. 如权利要求1所述的手写识别方法,其中确定切分可信度的步骤包括:通过逻辑回归模型来计算输入字符序列的各种切分方式中的各个笔画组合的切分可信度。
6. 如权利要求5所述的手写识别方法,其中逻辑回归模型中的危险因子是上述各种笔画组合特征。
7. 如权利要求5所述的手写识别方法,其中逻辑回归模型中的截断和各项回归系数,是通过对已有样本的训练来估计的。
8. 如权利要求1所述的手写识别方法,其中确定切分可信度的步骤包括:  
根据输入字符序列的特征,通过正态分布模型来计算输入字符序列的各种切分方式的切分可信度。
9. 如权利要求1所述的手写识别方法,其中基于所述切分可信度确定切分路径的步骤包括采用N-best方法或者动态规划法(DP)计算切分路径。
10. 如权利要求1所述的手写识别方法,其中所述呈现步骤包括向用户提供字符序列识别结果及针对该字符序列识别结果的至少一部分的候选项。
11. 如权利要求10所述的手写识别方法,其中响应于用户对候选切分方式的选择,向用户呈现与选择的切分方式相关的字符序列识别结果。
12. 如权利要求10所述的手写识别方法,其中响应于用户对单个字符的选择,向用户呈现与选择的字符相关的字符序列识别结果。
13. 一种手写识别设备,用于对用户连续输入的字符序列进行识别,该设备包括:

手写输入单元,采集用户连续输入的字符序列;

单字识别单元,对字符序列中的不同笔画组合进行识别,得到单字识别结果;

切分单元,基于不同笔画组合和对其所包含的笔画进行划分形成的“子笔画组合”的单字识别结果,计算与输入字符序列的不同笔画组合的单字识别正确度相关的特征,并根据对其“子笔画组合”的空间几何关系确定不同笔画组合的空间几何特征;根据与单字识别正确度相关的特征和空间几何特征,确定对输入的字符序列的不同切分方式下各个笔画组合的切分可信度;基于所述切分可信度确定切分路径;以及

显示控制单元,控制显示屏向用户呈现:与确定的切分路径相关的字符序列识别结果。

14. 如权利要求 13 所述的手写识别设备,其中所述单字识别单元采用多模板匹配方法来识别不同写法的字符。

15. 如权利要求 13 所述的手写识别设备,还包括:后处理单元,利用字典数据库或者语言模型对字符序列识别结果进行处理。

16. 如权利要求 13 所述的手写识别设备,其中所述的“与单字识别正确度相关的特征”包括以下之一:“子笔画组合”合并后的单字识别正确度,“子笔画组合”合并后的单字识别正确度与“子笔画组合”的单字识别正确度之差,“子笔画组合”合并后单字识别的第一选择正确度与合并后单字识别的其他候选字正确度的比值;

其中所述的笔画组合的空间几何特征包括以下之一:“子笔画组合”的外接矩形框的间隔,“子笔画组合”进行合并后的宽度,上一“子笔画组合”结束点与下一“子笔画组合”起始点之间的向量,上一“子笔画组合”结束点与下一“子笔画组合”起始点之间的距离,上一“子笔画组合”起始点与下一“子笔画组合”起始点之间的距离。

17. 如权利要求 13 所述的手写识别设备,其中切分单元通过逻辑回归模型来计算输入字符序列的各种切分方式中的各个笔画组合的切分可信度。

18. 如权利要求 13 所述的手写识别设备,其中切分单元根据输入字符序列的特征,通过正态分布模型来计算输入字符序列的各种切分方式的切分可信度。

19. 如权利要求 13 所述的手写识别设备,其中所述切分单元采用 N-best 方法或者动态规划法 (DP) 计算切分路径。

20. 如权利要求 13 所述的手写识别设备,其中所述显示控制单元还控制显示屏向用户提供字符序列识别结果及针对该字符序列识别结果的至少一部分的候选项。

21. 如权利要求 20 所述的手写识别设备,其中所述显示控制单元响应于用户对候选切分方式的选择,控制显示屏向用户呈现与选择的切分方式相关的字符序列识别结果。

22. 如权利要求 20 所述的手写识别设备,其中所述显示控制单元响应于用户对单个字符的选择,控制显示屏向用户呈现与选择的字符相关的字符序列识别结果。

## 手写识别方法和设备

### 技术领域

[0001] 本发明涉及文字输入,具体涉及一种手写识别方法和设备,能够识别用户连续手写输入的无框的多个字符,提高输入效率。

### 背景技术

[0002] 目前在诸如手机之类的电子设备中广泛应用了能够对用户的手写输入进行识别的模块,这使得用户不用再学习其他的通过按键进行字符输入的输入方法就能够与电子设备进行交互。

[0003] 非专利文献 1 (“Online Character Segmentation Method for Unconstrained Handwriting Strings Using Off-stroke Features” (Source:Hitachi Ltd.in the Tenth International Workshop on Frontiers in Handwriting Recognition,La Baule, France,2006)) 披露了一种手写识别方法,它设计了切分方式 (segmented pattern) 的物理特征 (‘无笔画 (off-stroke)’ 特征) 来识别无框手写输入的字符序列。在该方法中,‘无笔画’ 信息可以从前一笔画的最后一个采样点到后一笔画的第一采样点来获得,如图 1 中的虚线所示。该物理信息还包括诸如切分方式的高度 / 宽度和相应切分方式的手写时间等信息。该方法中,物理信息包括切分方式的形状特征、位置特征和间隙特征;笔画的长度;无笔画的平均距离;无笔画的平均时间;无笔画的距离;无笔画的角度的正弦和余弦值;无笔画的间隔。该方法主要针对书写前一笔画的结束点到书写当前笔画的起点之间的‘无笔画’ 过程来进行手写输入识别的。

[0004] 该手写识别方法假设,对于书写的字符而言,即使在不同的字符之间发生了连笔现象,字符之间的无笔画距离和时间间隔也要大于字符内的笔画之间的无笔画距离和时间,并且该方法假设每个笔画分布满足正态分布。基于上述的假设,该手写识别方法使用概率模型,根据特征的均值和方差计算不同切分方式之间的相似度。最后,该方法使用动态规划 (DP) 来确定最佳的切分路径。

[0005] 上述非专利文献 1 中存在的一个问题是对手写字符序列的切分依赖于每个笔画的书写时间。对该方法来说,无笔画的时间间隔是非常重要的特征。该方法假设切分方式之间的无笔画的时间间隔越大,则切分的正确度越高。当用户以较为恒定的速度进行书写时,这样的假设是合理的。但是在使用过程中,用户经常以不同的速度,例如一会儿快一会儿慢的速度进行书写。因此,如果用户在书写过程中改变书写速度,则非专利文献 1 所披露的方法将难以准确识别。

[0006] 上述非专利文献 1 中存在的另一问题是,仅仅使用了几何特征和时间特征来确定切分是否正确。该方法假设字符之间的无笔画距离大于字符内的笔画之间的无笔画距离。但是这样的假设并非总是正确的。非专利文献 1 列出了一些切分错误的典型示例,如图 2 所示。由图 2 可以看出,一些字符之间的无笔画距离小于字符内的笔画之间的无笔画距离。在图 2 所示的第一个例子中,‘5’ 被过切分了,这是由于字符内笔画之间的间隙过大造成的。在第二和第三个例子中,当一个输入字符序列的字符之间的距离变动较大以及字符的大小

不同时,出现了错误切分。

## 发明内容

[0007] 本发明的目的是提出一种手写识别方法和设备,能够对用户连续手写输入的多个字符进行识别,而与用户的书写速度的变化无关。

[0008] 在本发明的一个方面,提出了一种手写识别方法,用于对用户连续输入的非框(writing-box-free)的多个字符进行识别,该方法包括步骤:基于不同笔画组合和对其所包含的笔画进行划分形成的“子笔画组合”的单字识别结果,计算与输入字符序列的不同笔画组合的单字识别正确度相关的特征;根据对不同笔画组合所包含的笔画进行划分形成的“子笔画组合”的空间几何关系来确定不同笔画组合的空间几何特征;基于与单字识别正确度相关的特征和空间几何特征,确定对输入的字符序列的不同切分方式下各个笔画组合的切分可信度;基于所述切分可信度确定切分路径;以及向用户呈现:与确定的切分路径相关的字符序列识别结果。

[0009] 在本发明的另一方面,提出了一种手写识别设备,用于对用户连续输入的非框的字符序列进行识别,该设备包括:手写输入单元,采集用户连续输入的字符序列;单字识别单元,对字符序列中的不同笔画组合进行识别,得到单字识别结果;切分单元,基于不同笔画组合和对其所包含的笔画进行划分形成的“子笔画组合”的单字识别结果,计算与输入字符序列的各种笔画组合的单字识别正确度相关的特征,并根据其“子笔画组合”的空间几何关系确定不同笔画组合的空间几何特征;根据与单字识别正确度相关的特征和空间几何特征,确定对输入的字符序列的不同切分方式下各个笔画组合的切分可信度;基于所述切分可信度确定切分路径;以及显示控制单元,控制显示屏向用户呈现:与确定的切分路径相关的字符序列识别结果。

[0010] 由于采用非框输入,用户可以连续输入包含较多字符的一句话(或英文单词),提高用户的手写输入效率。对于传统的需要用户将字符写在手写框(writing-box)中的输入方法,手写字符之间的停顿常常会打断用户的思路从而影响输入速度,而要求每个字符都写在规定的手写框中(例如:目前手机上常用的两框输入法,要求用户在两个手写框之间来回切换)也改变了用户的书写习惯,降低了手写输入效率。本发明实施例的方法和设备允许用户实现连续输入,即时输出或者整体输出识别结果,无需改变书写习惯。

[0011] 由于本发明实施例的方法和设备在计算字符序列的切分可信度时,不仅仅考虑了现有技术中常用的空间几何特征,还充分考虑了笔画组合合并后的单字识别正确度以及子笔画组合的单字识别正确度,所以对于现有技术比较难以正确切分的情况,例如不同字符的笔画在空间上部分重叠,或同一个字符所包含的笔画分隔较大,本发明方法都能得到正确的切分和识别结果。

[0012] 而且,由于本发明实施例的方法和设备在进行字符序列切分时,并不依赖于用户写每一笔画的输入时间,所以可以适应用户的不同输入习惯,即使某用户输入字符的时间时快时慢,也不会影响本发明方法的切分正确性。

[0013] 另外,由于本发明实施例的方法和设备采用的笔画组合空间几何特征都是根据估计的字符平均宽(高)度进行规整化后的几何特征,所以该系统可以适应用户输入的任意大小的字符序列。同时,由于在单字识别时采用多模板训练和多模板匹配的方法,所以对

于不同用户输入的多种不同写法的字符（例如：汉字的简略字等），本发明方法都能准确识别。更进一步的，本发明实施例采用了语言模型和字典匹配，使得本识别设备还具有一定的拼写检查和纠错功能。

[0014] 最后，本发明实施例的方法和设备识别的字符序列可以为英文单词、日语假名组合、汉字组成的句子、韩文组合等等。进行手写识别判断的时机可以任意指定，既可以在用户输入字符序列的同时不断刷新识别结果，也可以在用户全部输入完字符序列后一次性进行手写识别。

### 附图说明

[0015] 从下面结合附图的详细描述中，本发明的上述特征和优点将更明显，其中：

[0016] 图 1 示出了根据现有技术的基于‘无笔画’特征进行字符识别的方法；

[0017] 图 2 示出了根据现有技术的基于‘无笔画’特征进行字符识别时出现的问题的例子；

[0018] 图 3 示出了根据本发明实施例的手写识别设备的结构示意图；

[0019] 图 4 示出了根据本发明实施例的手写识别设备的训练过程的流程图；

[0020] 图 5A、5B、5C 和 5D 示出了根据本发明实施例的手写识别设备中笔画组合及其“子笔画组合”的示意图；

[0021] 图 6A、6B、6C 和 6D 示出了根据本发明实施例的手写识别设备中笔画组合的空间几何特征的含义的示意图；

[0022] 图 7 是根据本发明实施例的同一字符的不同写法的一个示意图；

[0023] 图 8 是根据本发明实施例的同一字符的不同写法的另一示意图；

[0024] 图 9A、9B 和 9C 是根据本发明实施例的描述多模板训练和多模板匹配的示意图；

[0025] 图 10 示出了根据本发明实施例的逻辑回归模型的函数曲线；

[0026] 图 11 示出了根据本发明实施例的手写识别过程的流程图；

[0027] 图 12A、12B、12C 示出了根据本发明实施例的以不同切分路径进行切分的示意图；

[0028] 图 13A、13B、13C 和 13D 示出了根据本发明实施例的手写识别设备的手写输入识别结果的示意图；

[0029] 图 14 示出了根据本发明实施例的手写识别方法在电子词典上的应用；

[0030] 图 15 示出了向用户提供识别结果的至少一部分的候选项供用户选择和纠正的示意图；以及

[0031] 图 16A 和图 16B 示出了根据本发明实施例的手写识别方法在笔记本电脑和手机上的应用。

### 具体实施方式

[0032] 下面，参考附图详细说明本发明的优选实施方式。在附图中，虽然示于不同的附图中，但相同的附图标记用于表示相同的或相似的组件。为了清楚和简明，包含在这里的已知的功能和结构的详细描述将被省略，否则它们将使本发明的主题不清楚。

[0033] 图 3 示出了根据本发明实施例的手写识别设备的结构示意图。

[0034] 如图 3 所示，根据本发明实施例的手写识别设备用于对用户连续输入的非框

(writing-box-free) 的多个字符进行识别,它包括:手写输入单元 110,用于采集用户的笔迹,并且对其数字化,作为输入笔迹信号;手写笔迹存储单元 120,用于存储手写输入单元 110 产生的输入笔迹信号;字符序列识别单元 130,用于识别所输入的字符序列,该字符序列识别单元 130 包括三个子单元:切分单元 132、单字识别单元 131 和后处理单元 133。

[0035] 由于采用无框输入,用户可以连续输入包含较多字符的一句话(或英文单词),或者在用户输入过程中即时显示识别结果,或者在用户输入该句话后,再给出识别结果,提高用户的手写输入效率。对于传统的需要用户将字符写在手写框(writing-box)中的输入方法,手写字符之间的停顿常常会打断用户的思路从而影响输入速度,而要求每个字符都写在规定的手写框中(例如:目前手机上常用的两框输入法,要求用户在两个手写框之间来回切换)也改变了用户的书写习惯,降低了手写输入效率。本发明实施例的方法和设备允许用户实现连续输入,即时输出或者整体输出识别结果,无需改变书写习惯。

[0036] 切分单元 132 从输入笔迹信号中提取输入字符序列的各个笔画组合的各种空间几何特征,同时切分单元 132 调用单字识别单元 131,得到各个笔画组合的单字识别结果及其单字识别正确度,再通过逻辑回归模型来计算“切分可信度”,然后利用 N-best 算法得到最佳的 N 种切分方式,如后面详细说明。

[0037] 后处理单元 133 采用语言模型和字典数据库匹配,对切分单元 132 得到的字符序列识别结果进行校正。

[0038] 如图 3 所示,根据本发明实施例的手写识别设备还包括显示控制单元 150,在用户通过手写输入单元 110 输入笔画的同时,它一方面控制系统显示笔迹,通过显示屏呈现给用户,另一方面,在显示屏上显示识别单元 130 所产生的识别候选项,供用户选择;以及候选项选择单元 140,它在用户的操作下从候选项中选择要输入的字符序列或者单个字符,然后把识别结果显示给用户或者提供给其他应用,例如与字典中词条进行匹配,以便找出相应的释义等。

[0039] 根据本发明的实施例,字符序列识别单元 131 中采用的逻辑回归模型的截断(intercept)和各项回归系数(Regression Coefficients)是通过对已有样本的训练来估计得到的。

[0040] 图 4 示出了根据本发明实施例的手写识别设备的训练过程的流程图。

[0041] 根据本发明的实施例,样本训练中的样本既包括各个字符的单字样本,也包括各个字符包含的每个笔画样本,以及字符内若干笔画的组合,或是不同字符部分笔画的组合,这些统称为笔画组合类。

[0042] 如图 4 所示,在步骤 S10,采集用户的代表手写字符序列的手写轨迹数据。在步骤 S11,加入相应的笔画组合类。然后在步骤 S12 和 S13 进行预处理并计算笔画组合特征。

[0043] 样本训练中计算的特征即为逻辑回归模型中的  $m$  维特征  $(x_1, x_2, \dots, x_m)$ ,笔画组合的特征包括:“子笔画组合”的外接矩形框间隔;“子笔画组合”进行合并后的宽度;“子笔画组合”之间的向量和距离;合并后的单字识别正确度;合并后的识别正确度与“子笔画组合”的识别正确度之差;合并后单字识别的第一选择正确度与合并后单字识别的其他候选字正确度的比值,等等。

[0044] 在步骤 S13 进行特征计算之前,要在步骤 S12 进行“预处理”,根据字符序列的高度和宽度,估计字符平均高度  $H_{avg}$  和字符平均宽度  $W_{avg}$ ,为笔画组合的空间几何特征进行规整

化做准备,使本发明实施例的手写识别设备可以适应用户输入的任意大小的字符序列。

[0045] 下面以字符序列中的第  $k$  笔画至第  $k+3$  笔画的切分为例,解释本发明实施例中“子笔画组合”(以下简称“子笔画”)的概念。由第  $k$  笔画开始,可能的切分方式有如下四种,如图 5A、5B、5C 和 5D 所示:

[0046] 1) 对于一笔画组合,它只包括第  $k$  笔画,所以无子笔画。

[0047] 2) 对于二笔画组合,它包括第  $k$  和  $k+1$  两个子笔画。

[0048] 3) 对于三笔画组合,它有两种子笔画分类方式:

[0049] ◆方式一:上一子笔画为第  $k$  笔画,下一子笔画为  $k+1$  和  $k+2$  的笔画组合;

[0050] ◆方式二:上一子笔画为  $k$  和  $k+1$  的笔画组合,下一子笔画为第  $k+2$  笔画。

[0051] 4) 对于四笔画组合,它有三种子笔画分类方式:

[0052] ◆方式一:上一子笔画为第  $k$  笔画,下一子笔画为  $k+1$ 、 $k+2$  和  $k+3$  的三笔画组合;

[0053] ◆方式二:上一子笔画为  $k$  和  $k+1$  的笔画组合,下一子笔画为  $k+2$  和  $k+3$  的笔画组合;

[0054] ◆方式三:上一子笔画为  $k$ 、 $k+1$  和  $k+2$  的三笔画组合,下一子笔画为第  $k+3$  笔画。

[0055] 可见,根据本发明的实施例,“子笔画组合”可以是某个“笔画组合”中包含的笔画按照顺序划分成的不同组合。例如,书写顺序为“ $k, k+1, k+2$ ”的笔画组合,与其相关的“子笔画组合”可以从笔画“ $k$ ”和“ $k+1$ ”之间进行划分产生的第一类组合,也可以是从笔画“ $k+1$ ”和“ $k+2$ ”之间进行划分产生的第二类组合,如图 5C 所示。

[0056] 本发明实施例的设备中,对字符序列中的所有可能的笔画组合,计算笔画组合的各种特征,包括其单字识别正确度特征和子笔画组合的空间几何特征。各种具体特征如下:

[0057] (a) 子笔画合并后的单字识别正确度  $C_{\text{merge}}$ :该正确度越大,合并后为一个单字的可能性越大;

[0058] (b) 合并识别正确度  $C_{\text{merge}}$  与两个子笔画的单字识别正确度  $C_{\text{str1}}$ 、 $C_{\text{str2}}$  的差: $(2 * C_{\text{merge}} - C_{\text{str1}} - C_{\text{str2}})$ 。如果该值大于 0,表示两笔合并为单字的可能性比两个子笔画分别为一个单字的可能性更大,且这个差值越大,合并为单字的可能性越大;

[0059] (c) 合并后单字识别的第一选择正确度(即  $C_{\text{merge}}$ )与合并后单字识别的其他候选字正确度  $C_{\text{merge}1}$  的比值( $T$  表示第  $T$  候选字, $T$  值可设定):如果这个比值比较大,表示合并后的笔画组合与其单字识别的第一选择字的匹配距离很近,而与其他候选字的匹配距离较远,即表明合并后为单字的可能性较大;

[0060] (d) 两个子笔画的外接矩形框间隔  $\text{gap}/W_{\text{avg}}$ (或  $\text{gap}/H_{\text{avg}}$ ):子笔画之间的间隔越小,合并后为单字的可能性越大,如果间隔为负,合并后为单字的可能性就更大;

[0061] (e) 子笔画合并后的宽度  $W_{\text{merge}}/W_{\text{avg}}$ (或  $W_{\text{merge}}/H_{\text{avg}}$ ):合并后的宽度越小,合并为单字的可能性越大;

[0062] (f) 上一子笔画结束点与下一子笔画起始点之间的向量  $V_{s2-e1}/W_{\text{avg}}$ (或  $V_{s2-e1}/H_{\text{avg}}$ );

[0063] (g) 上一子笔画结束点与下一子笔画起始点之间的距离  $d_{s2-e1}/W_{\text{avg}}$ (或  $d_{s2-e1}/H_{\text{avg}}$ );

[0064] (h) 上一子笔画起始点与下一子笔画起始点之间的距离  $d_{s2-s1}/W_{\text{avg}}$ (或  $d_{s2-s1}/H_{\text{avg}}$ )。

[0065] 以上特征中,“/”为除法符号, $W_{\text{avg}}$  和  $H_{\text{avg}}$  为“预处理”中估计出的字符平均宽度和字符平均高度。第 (d) ~ (h) 这些空间几何特征参考图 6A ~ D 的图示,图中的圆点表示每



一笔画的起始点。

[0066] 对于上述特征 (a)、(b)、(c), 通过在步骤 S14 调用“单字识别单元”来得到: 子笔画合并后的单字识别正确度  $C_{\text{merge}}$  及其他候选字正确度  $C_{\text{mergeT}}$ , 两个子笔画的单字识别正确度  $C_{\text{str1}}$  和  $C_{\text{str2}}$ 。

[0067] 本发明实施例的“单字识别单元”采用模板匹配的方法来进行单字识别, 单字识别的正确度由模板匹配的距离来度量, 距离越小, 正确度越大。单字识别的样本训练中, 采用机器学习算法 (例如: GLVQ) 生成特征模板; 其单字特征向量包括: “笔画方向分布特征”、“网格笔画特征”和“周边方向特征”; 特征提取前, 要进行预处理, 包括“等距平滑”、“质心归一化”和“非线性归一化”等操作, 以便使得该样本的特征变得规整; 模板匹配时, 采用“分段式快速匹配”方法, 逐级滤除候选项, 提高匹配速度。单字识别的上述方法在中国专利申请公开 CN101354749A 披露, 该专利申请公开被整体引入本申请作为参考。

[0068] 在实际的书写过程中, 不同的用户对于同一个字符常常有不同的写法。例如: 英文字母“A”可能有如下多种写法, 如图 7 所示。

[0069] 再如, 日文汉字“機”可能有如下三种写法 (后两种是简略写法), 如图 8 所示。

[0070] 因此, 为了提高手写识别的鲁棒性, 本发明实施例的设备中采用“多模板训练”的方法对同一个字符的不同写法进行单独训练, 这样就可以采用“多模板匹配”的方法来识别多种不同写法的字符。为了进行“多模板训练”, 首先对采集到的样本根据它们的不同写法进行分类。例如: 对于上述提到的“機”字, 本发明实施例在样本训练时采用如图 9A、9B 和 9C 所示的三种形式的样本构成多模板训练。

[0071] 如图 4 所示, 在步骤 S15, 计算逻辑回归模型的系数。对字符系列进行正确的切分, 是实现多字符无框连续输入的手写识别的关键。本发明实施例的设备和方法根据输入字符序列的各种特征, 计算输入字符序列的各种切分方式中的各个笔画组合的切分可信度。本发明实施例的切分可信度公式采用逻辑回归模型 (Logistic Regression Mode), 逻辑回归模型为:

$$[0072] \quad f(Y) = \frac{1}{1 + e^{-Y}} \quad \dots\dots (1)$$

[0073] 上述逻辑回归模型的函数曲线如图 10 所示, 当  $Y$  在  $-\infty \sim +\infty$  变化时,  $f(Y)$  的值为  $0 \sim 1$ , 即切分可信度为  $0\% \sim 100\%$ , 且当  $Y = 0$  时,  $f(Y) = 0.5$ , 切分可信度为  $50\%$ 。

[0074] 在上述逻辑回归模型中:

$$[0075] \quad Y = g(X) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m \quad \dots\dots (2)$$

[0076] 其中,  $X = (x_1, x_2, \dots, x_m)$  是逻辑回归模型的危险因子 (risk factor), 在本发明实施例的设备和方法中计算切分可信度时,  $X = (x_1, x_2, \dots, x_m)$  表现为笔画组合的  $m$  维特征。 $(\beta_0, \beta_1, \beta_2, \dots, \beta_m)$  是逻辑回归模型的截断 (intercept) 和各项回归系数 (Regression Coefficients)。

[0077] 在计算出字符序列中的所有可能的笔画组合的  $m$  维特征后, 本发明实施例的设备和方法采用最大似然估计方法 (也可以用最小二乘估计等其他参数估计方法) 来估计切分可信度的逻辑回归模型中的截断  $\beta_0$  和各项回归系数  $(\beta_1, \beta_2, \dots, \beta_m)$ 。

[0078] 假设有  $n$  个笔画组合样本, 观测值分别为  $(Y_1, Y_2, \dots, Y_n)$ 。对于第  $i$  个笔画组合,  $m$  维特征  $X_i = (x_{i1}, x_{i2}, \dots, x_{im})$ , 观测值为  $Y_i$ 。 $n$  个回归关系可以写成:

$$[0079] \quad \begin{cases} Y_1 = \beta_0 + \beta_1 X_{11} + \beta_2 X_{12} + \cdots + \beta_m X_{1m} \\ Y_2 = \beta_0 + \beta_1 X_{21} + \beta_2 X_{22} + \cdots + \beta_m X_{2m} \\ \dots \\ Y_n = \beta_0 + \beta_1 X_{n1} + \beta_2 X_{n2} + \cdots + \beta_m X_{nm} \end{cases} \quad \dots\dots \quad (3)$$

[0080] 在样本训练时,对于给定的第  $i$  个笔画组合,如果该笔画组合可信:

$$[0081] \quad \text{令 } f_i = f(Y_i) = \frac{1}{1+e^{-Y_i}} \rightarrow 1, \text{ 至少 } f(Y_i) > 0.5 \text{ 即 } Y_i > 0 \quad \dots\dots (4)$$

[0082] 若该笔画组合不可信(即该种组合方式不正确):

$$[0083] \quad \text{令 } f_i = f(Y_i) = \frac{1}{1+e^{-Y_i}} \rightarrow 0, \text{ 至少 } f(Y_i) < 0.5 \text{ 即 } Y_i < 0 \quad \dots\dots (5)$$

[0084] 把  $Y = g(X) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_m x_m$  代入逻辑回归模型公式:

$$[0085] \quad f(Y) = \frac{1}{1+e^{-Y}} = \frac{1}{1+e^{-g(X)}} = \pi(X) \quad \dots\dots (6)$$

[0086] 设  $p_i = P(f_i = 1 | X_i)$  为  $f_i = 1$  的概率,则  $f_i = 0$  的条件概率为  $P(f_i = 0 | X_i) = 1 - p_i$ 。于是,得到一个观测值的概率为:  $P(f_i) = p_i^{f_i} (1 - p_i)^{(1-f_i)}$ 。

[0087] 因为各项观测独立,所以它们的联合分布可以表示为各边际分布的乘积:

$$[0088] \quad l(\beta) = \prod_{i=1}^n \pi(X_i)^{f_i} [1 - \pi(X_i)]^{1-f_i} \quad \dots\dots (7)$$

[0089] 上式称为  $n$  个观测的似然函数。我们的目标是能够求出使这一似然函数值最大的参数估计。于是,最大似然估计的关键就是求出参数  $(\beta_0, \beta_1, \beta_2, \dots, \beta_m)$ ,使上式取得最大值。对上述似然函数求对数,得到对数似然函数,再对此对数似然函数求导,得到  $m+1$  个似然方程。应用牛顿-拉斐森(Newton-Raphson)方法迭代求解  $m+1$  个似然方程,可以得到逻辑回归模型中的各项系数  $(\beta_0, \beta_1, \beta_2, \dots, \beta_m)$ ,这些系数存储在该设备中,供识别过程中使用。

[0090] 根据本发明的另一实施例,也可通过正态分布模型来计算输入字符序列的各种切分方式的切分可信度。

[0091] 图 11 示出了根据本发明实施例的手写识别方法的流程图。如图 11 所述,在步骤 S20,用户进行手写输入,通过手写输入单元 110 采集字符序列的笔画。然后,在步骤 S21,将采集的手写笔迹在存储单元 120 中存储,并且在步骤 S22 由显示控制单元 150 显示在用户界面上。

[0092] 然后,字符序列识别单元 130 对存储在手写笔迹存储单元中的笔画进行在步骤 S23、S24、S25、S26、S27 和 S28 所示的“预处理”、“计算笔画组合的特征”、“单字识别”、“计算切分可信度”、“选取切分最优路径”和“识别后处理”的操作。

[0093] 具体而言,步骤 S23、S24 和 S25 的执行过程与上述样本训练估计逻辑回归模型系数的方法中的相应各个步骤的操作类似。在步骤 S23,进行“预处理”,根据字符序列的高度和宽度,估计字符平均高度  $H_{avg}$  和字符平均宽度  $W_{avg}$ ,为笔画组合的空间几何特征进行规整化做准备,使本发明实施例的手写识别设备可以适应用户输入的任意大小的字符序列。

[0094] 在步骤 S24,对字符序列中的所有可能的笔画组合,计算笔画组合的各种特征,包括其单字识别正确度特征和子笔画组合的空间几何特征。

[0095] 在步骤 S25,调用“单字识别单元”来得到:子笔画合并后的单字识别正确度  $C_{merge}$

及其他候选字正确度  $C_{\text{merge}}$ , 两个子笔画的单字识别正确度  $C_{\text{str1}}$  和  $C_{\text{str2}}$

[0096] 在步骤 S26, 本发明实施例的方法根据输入字符序列的各种特征 ( $X = (x_1, x_2, \dots, x_m)$ ) 和样本训练得到的各项系数 ( $\beta_0, \beta_1, \beta_2, \dots, \beta_m$ ), 利用公式 (1) 和公式 (2), 采用逻辑回归模型, 来计算输入字符序列的各种切分方式中的各个笔画组合的切分可信度  $f(Y)$ 。

[0097] 在步骤 S27, 本发明实施例的方法采用 N-Best 方法计算最可能的 N 种切分路径。定义每个笔画的起始点为一个基元节点, 基元或基元组合构成的路径即为对应的笔画组合, 每个部分路径的代价函数为:  $C(Y) = 1 - f(Y)$ , 也就是说, 切分可信度越高, 部分路径的代价函数值越小。N-best 方法就是要选取最佳的 N 种路径, 使所经过的所有路径的代价函数的数值之和最小、第二小……第 N 小。

[0098] N-Best 方法可以用多种方式实现, 例如, 把动态规划 (DP) 方法与堆栈 (Stack) 算法相结合来产生多个候选项, 等等。本发明实施例中, N-Best 方法包括两个步骤: 前向搜索过程采用一种改进的维特比 (Viterbi) 算法 (维特比算法就是一种用于查找最可能的隐含状态序列的动态规划方法), 用来记录转移到每个基元节点的最优 N 个部分路径的状态 (即为所经过路径的代价函数值之和); 第 k 个基元节点的状态只和第 k-1 个基元节点的状态有关; 后向搜索过程采用一种基于 A\* 算法的堆栈算法, 对每一个节点 k, 它的启发函数 (heuristic function) 为下列两个函数的和: 一是“路径代价函数”, 表示从起始点到第 k 节点的最短路径的代价函数值之和, 二是“启发估计函数”, 表示从第 k 节点到目标节点的路径代价的估计。在后向搜索过程中, 堆栈中的路径得分是计算的全路径得分, 且最优的路径总是位于栈顶, 所以, 该算法是一种全局最优算法。

[0099] 假设用户输入的是图 6A 所示的手写字符序列“defne”, 图 12A 示出了本发明实施例对该手写字符序列进行切分的结果。采用 N-best 方法得到的最可能的三种切分方式依次如图 12A、图 12B 和图 12C 所示: 第一种切分方式的每个字符的第一单字识别结果为“define (即为正确答案)”, 第二种切分方式的一选结果为“ccefine”, 第三种切分方式的一选结果为“deftine”。

[0100] 在步骤 S28, 本发明实施例的方法最后通过和语言字典 (例如: 英文单词字典) 数据库的匹配, 或者使用语言模型 (例如: 二元模型 bigram) 对识别结果进行后处理, 纠正错误 (例如: 英文单词的拼写错误)。

[0101] 在步骤 S29, 显示控制单元 150 控制显示屏向用户呈现手写输入的识别结果及相关的候选项, 提供给用户在候选项选择单元 140 选择或确认 (默认的识别结果是第一切分方式的每个字符的第一单字识别结果); 用户可以从字符序列的候选切分方式中选择正确的切分方式; 也可以在各个字符的候选项中选择正确的字符, 手动纠正其中的部分识别字符, 例如选中单个字符或词组, 对作为字符序列的一部分的该字符或词组的候选识别结果进行选择。图 15 示出了根据本发明实施例的提供字符序列识别结果的一部分的候选项供用户选择和纠正的示意图。

[0102] 在步骤 S30, 对用户是否确认或选择某个候选项进行识别。如果用户没有确认或选择, 而是继续书写, 则流程转到步骤 S20, 继续进行上述的识别过程。如果识别到了对某个候选项的选择, 则在步骤 S31, 从候选项选择识别结果, 将识别结果显示出来或提供给其他的应用。同时, 在步骤 S32 对手写输入的识别结果进行更新。

[0103] 由于本发明实施例的方法和设备在计算字符序列的切分可信度时, 不仅仅考虑了

现有技术中常用的空间几何特征,还充分考虑了笔画组合合并后的单字识别正确度以及子笔画组合的单字识别正确度,所以对于现有技术比较难以正确切分的情况,例如不同字符的笔画在空间上部分重叠,或同一个字符包含的笔画分隔较大,本发明实施例的方法和设备都能得到正确的切分和识别结果。

[0104] 而且,由于本发明实施例的设备和方法在进行字符序列切分时,并不依赖于用户写每一笔画的输入时间,所以可以适应用户的不同输入习惯,即使某用户输入字符的时间快时慢,也不会影响本发明实施例的方法和设备的切分正确性。

[0105] 另外,由于本发明实施例的方法和设备采用的笔画组合空间几何特征都是根据估计的字符平均宽(高)度进行规整化后的几何特征,所以该设备可以适应用户输入的任意大小的字符序列。同时,由于在单字识别时采用多模板训练和多模板匹配的方法,所以对于不同用户输入的多种不同写法的字符(例如:汉字的简略字等),本发明实施例的方法和设备方法都能准确识别。更进一步的,本发明实施例的方法和设备采用了语言模型和字典匹配,使得本设备还具有拼写检查和纠错功能。

[0106] 最后,本发明实施例的方法和设备识别的字符序列可以为英文单词、日语假名组合、汉字组成的句子、韩文组合等等。进行手写识别判断的时机可以任意指定,既可以在用户输入字符序列的同时不断刷新识别结果,也可以在用户全部输入完字符序列后一次性进行手写识别。

[0107] 图 13A、13B、13C 和 13D 示出了根据本发明实施例的手写识别设备的手写输入识别结果的示意图。由于在识别过程中不仅考虑到了笔画组合的几何特征,而且考虑到了单字识别结果的正确度,因此对于现有技术比较难以正确切分的情况,包括:不同字符的笔画在空间上部分重叠,或者字符之间的距离小于字符内的笔画之间的距离,或者当用户在输入过程中出现字体大小不一的情况,本发明方法也能够做出正确的识别。例如:如图 13D 所示,“d”和“e”、“f”和“i”的笔画在空间上部分重叠;如图 13A 和图 13C 所示,“어”和“사”之间的间隔小于“사”内部笔画之间的距离,“日”和“本”之间的间隔也小于“語”内部笔画之间的距离;如图 13B 和图 13D 所示,“かいしゃいん”和“define”各个字符的字体大小是不等的。以上这些情况,本发明实施例的方法都能正确识别。

[0108] 图 14 示出了根据本发明实施例的电子词典。如图 14 所示,对用户输入的一连串英文字符进行识别,然后将识别结果显示出来。通过调用词典中的与该识别的英文字符串相关的条目,向用户展现手写输入的英文的日文释义。如图 15 所示,一旦用户选中了识别结果中的某单个字符,则向用户提供该字符的候选识别结果,供用户对其进行纠正。换言之,用户可以选择字符序列识别结果中的一个或者更多个字符,一旦系统确定用户进行了选择,就显示出与该选择的单个或者多个字符相关的候选项,供用户选择。

[0109] 可见,根据本发明的上述实施例允许用户对整个字符序列的识别结果进行整体纠正,也允许用户对识别结果中的任何一部分进行纠正。

[0110] 根据本发明的另一实施例,显示区域和手写输入区域可以被设置在不同的平面上,也可以设置在相同的平面上,如图 16A 和 16B 所示。例如,针对笔记本电脑,可以在键盘所在的平面上设置手写区域。

[0111] 如上所述,本发明的方法和设备可以应用于或者包含在各种能采用手写作为输入或控制方式的信息终端产品,包括个人电脑,手提电脑, PDA, 电子辞典, 复合机, 手机以及大

型触摸屏的手写设备等。

[0112] 说明书和附图仅示出了本发明的原理。因此应该意识到,本领域技术人员能够建议不同的结构,虽然这些不同的结构未在此处明确描述或示出,但体现了本发明的原理并包括在其精神和范围之内。此外,所有此处提到的示例明确地主要只用于教学目的以帮助读者理解本发明的原理以及发明人所贡献的促进本领域的构思,并应被解释为不是对这些特定提到的示例和条件的限制。此外,此处所有提到本发明的原则、方面和实施方式的陈述及其特定的示例包含其等同物在内。

[0113] 上面的描述仅用于实现本发明的实施方式,本领域的技术人员应该理解,在不脱离本发明的范围的任何修改或局部替换,均应该属于本发明的权利要求来限定的范围,因此,本发明的保护范围应该以权利要求书的保护范围为准。

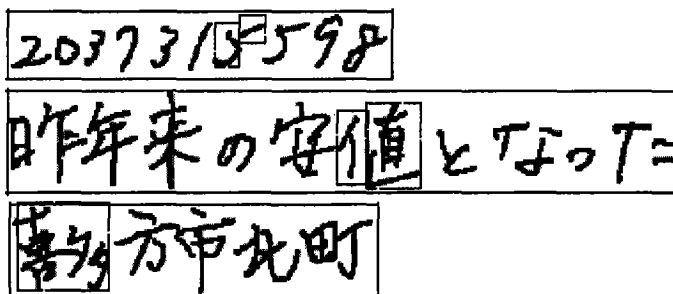
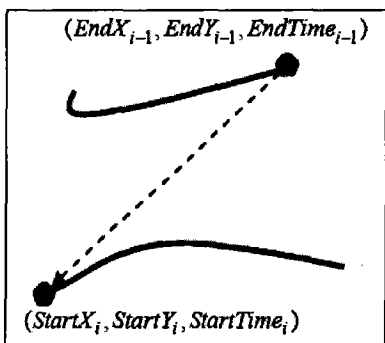


图 1

图 2

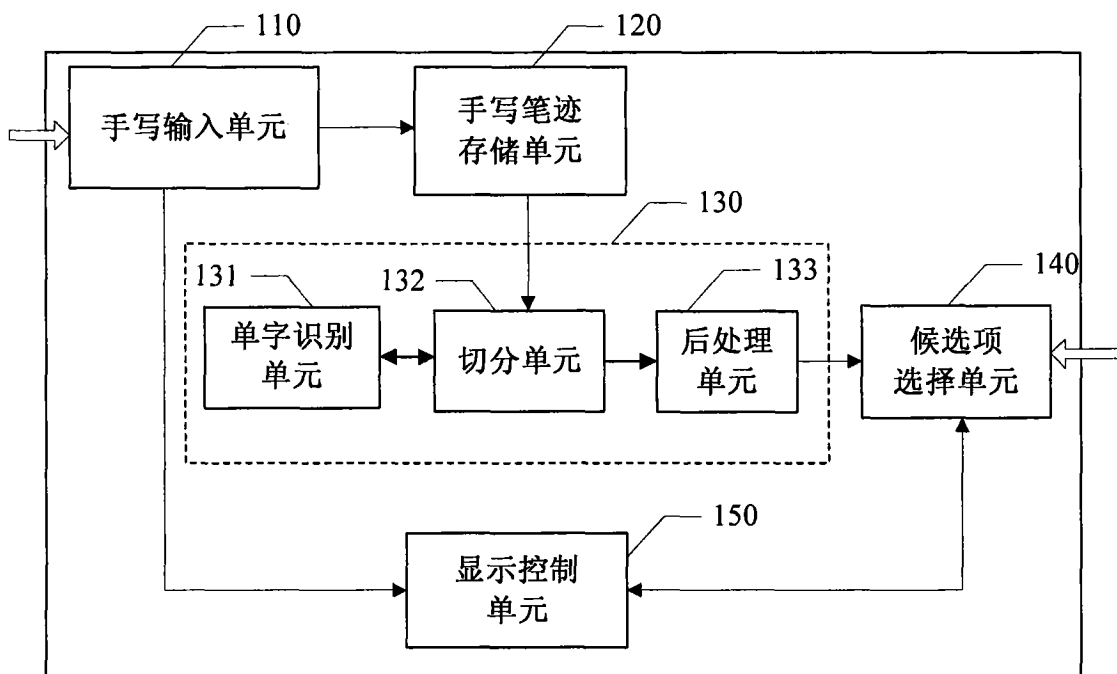


图 3

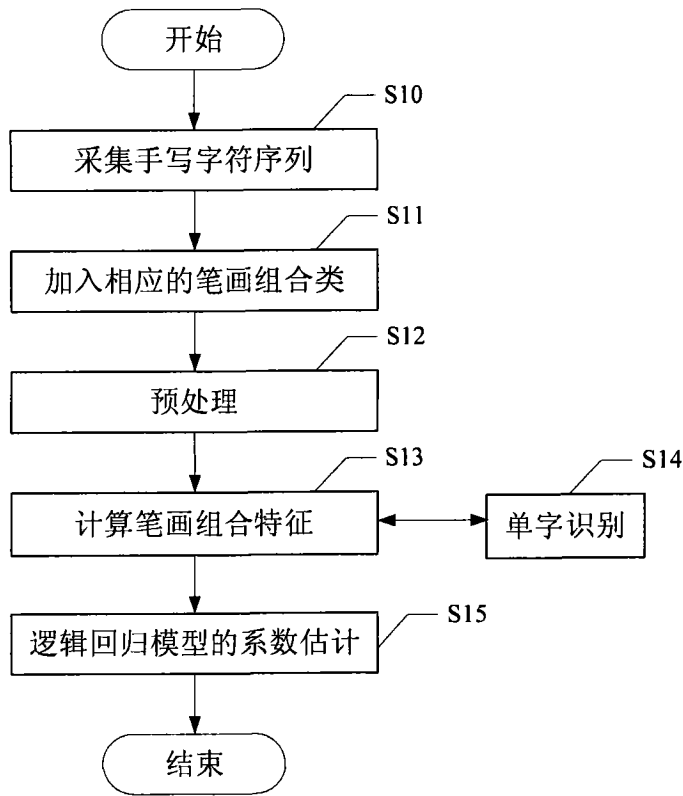


图 4

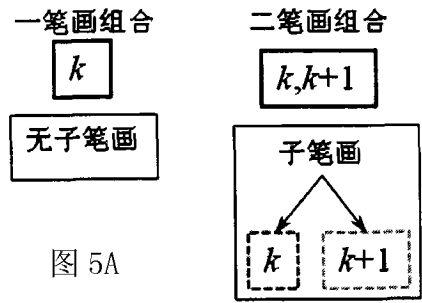


图 5A

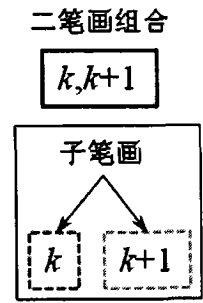


图 5B

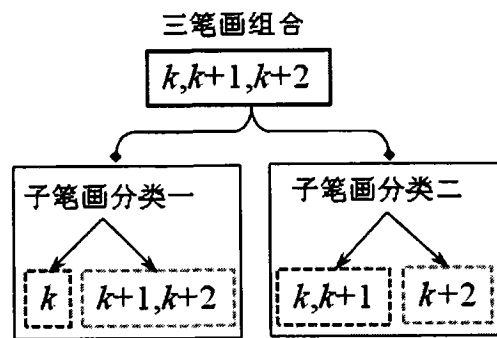


图 5C

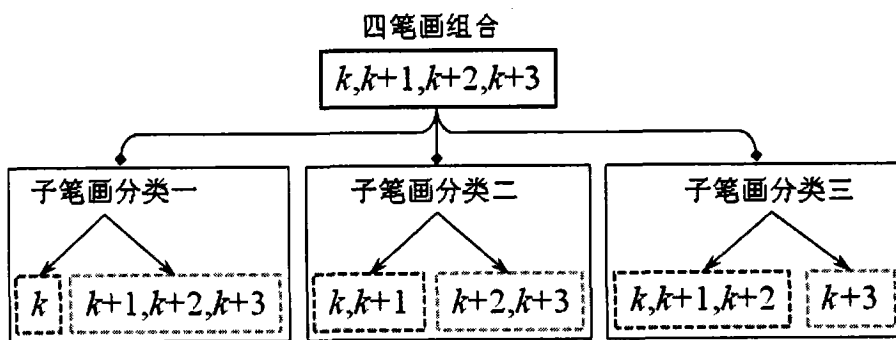


图 5D

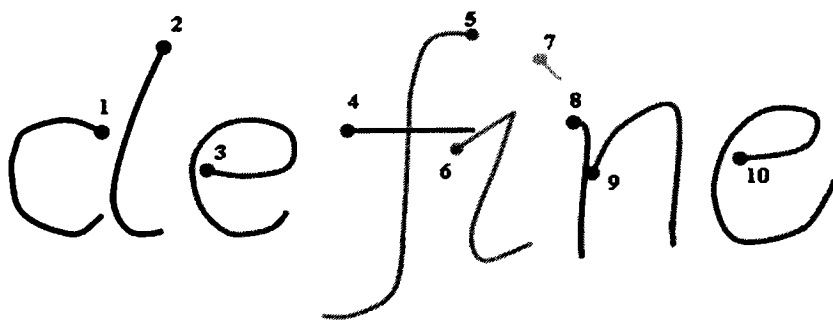


图 6A

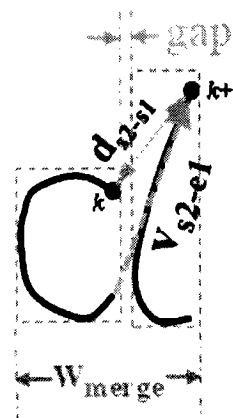


图 6B

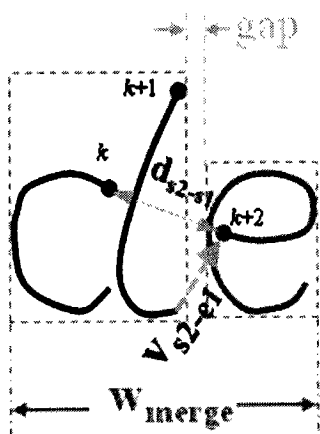


图 6C

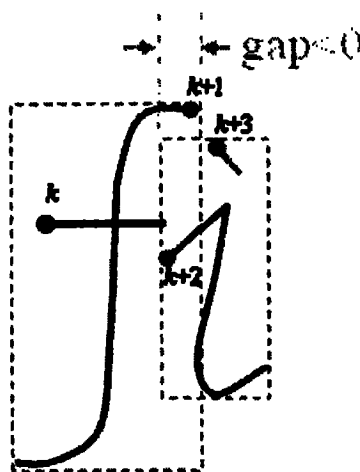


图 6D



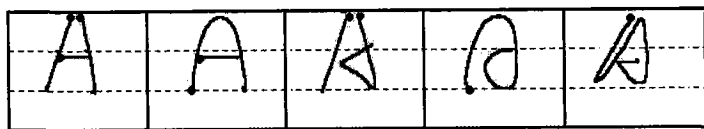


图 7



图 8



图 9A

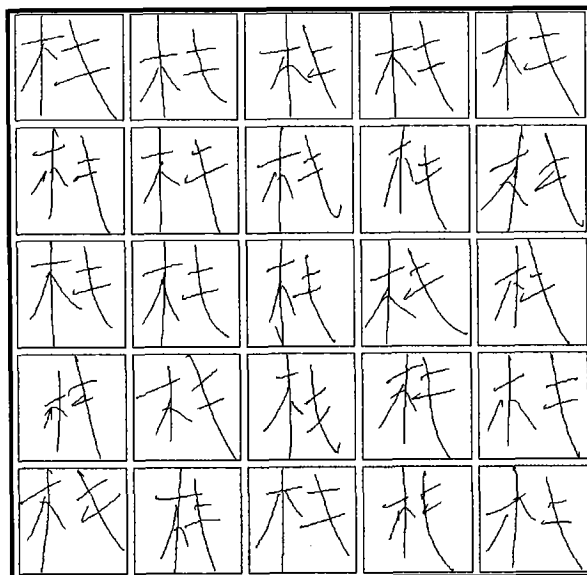


图 9B

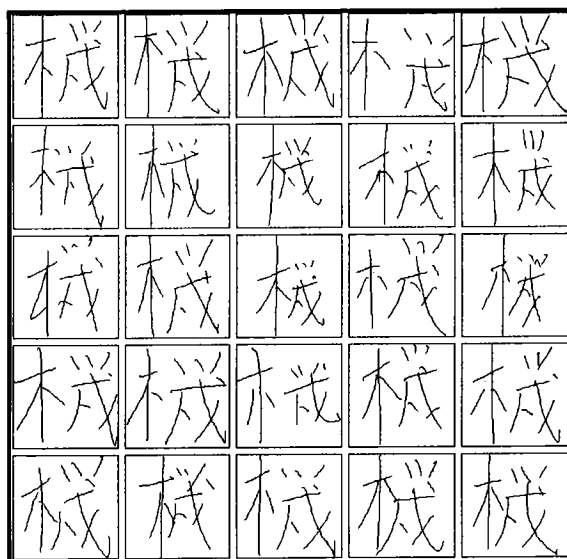


图 9C

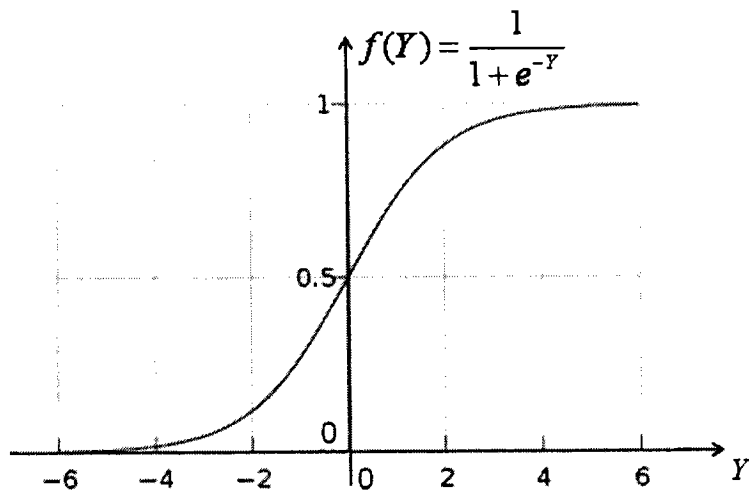


图 10

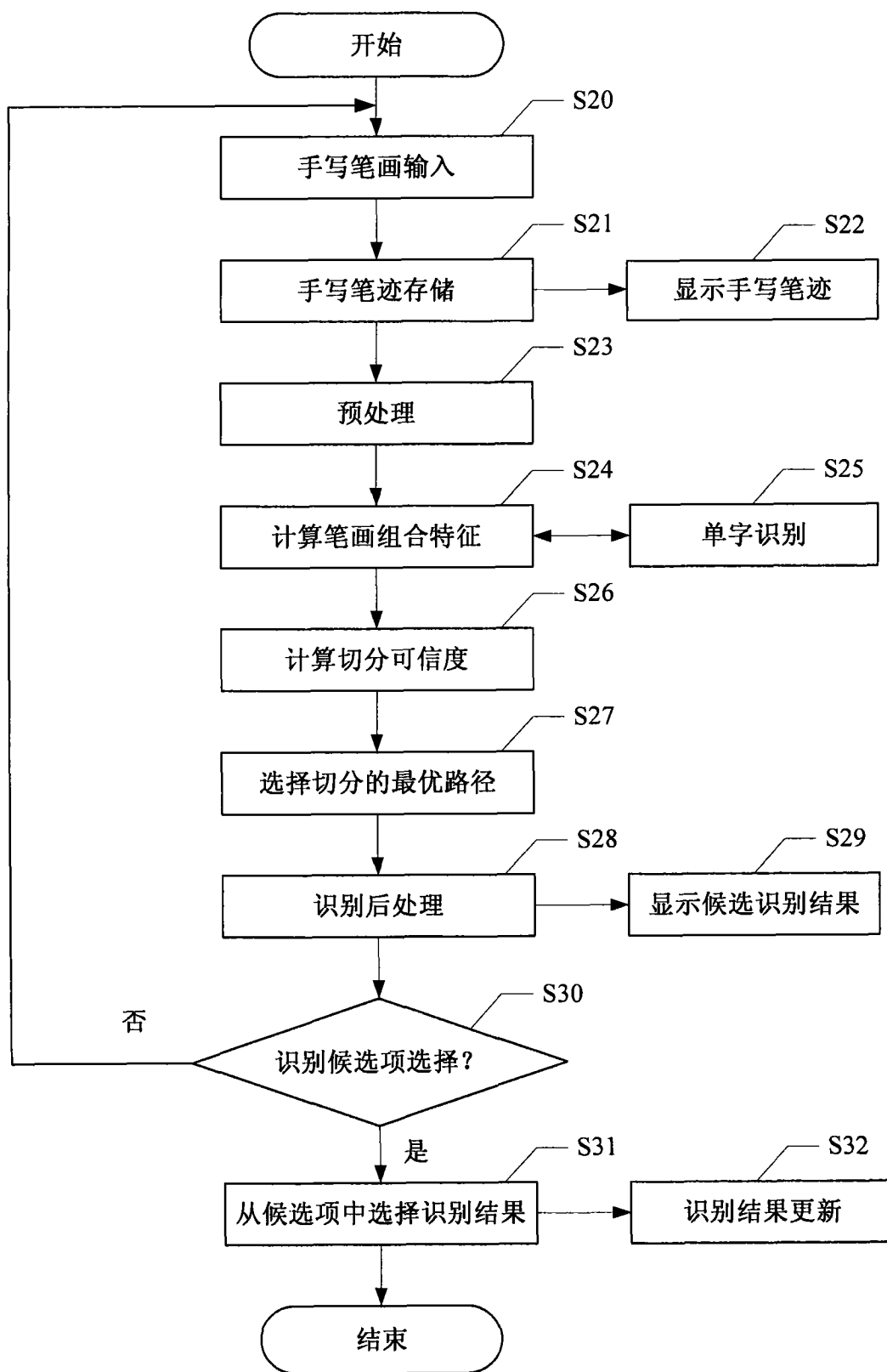


图 11

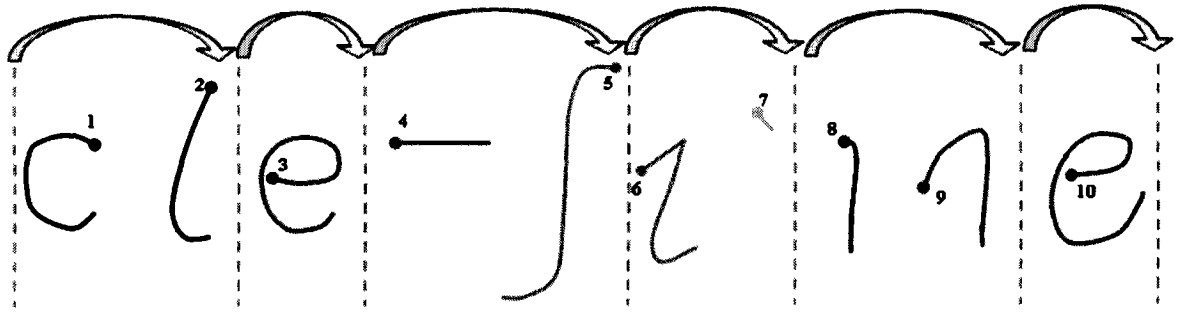


图 12A

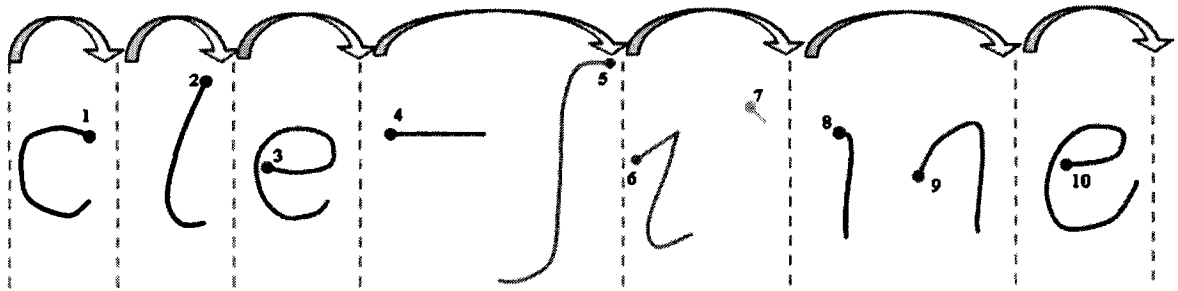


图 12B

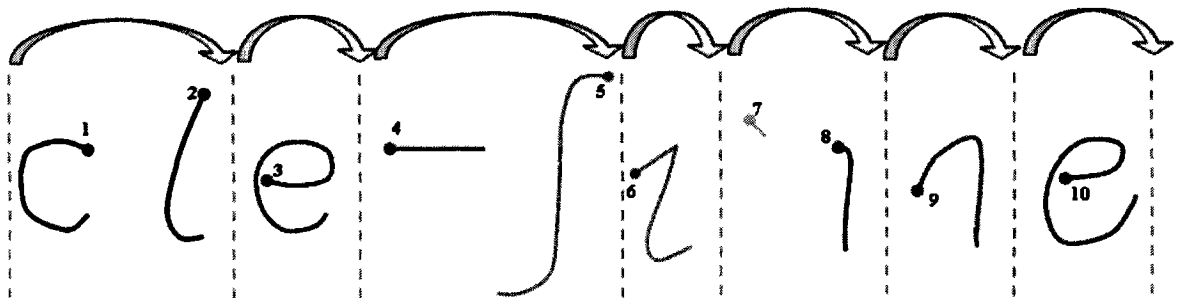


图 12C

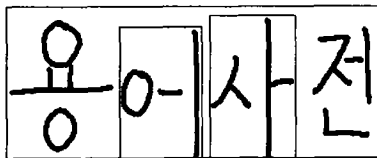


图 13A

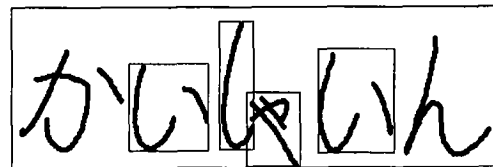


图 13B

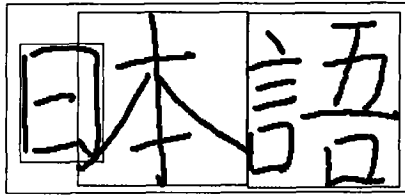


图 13C

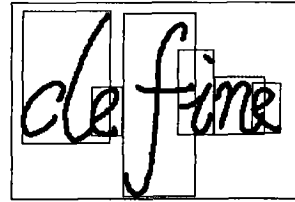


图 13D

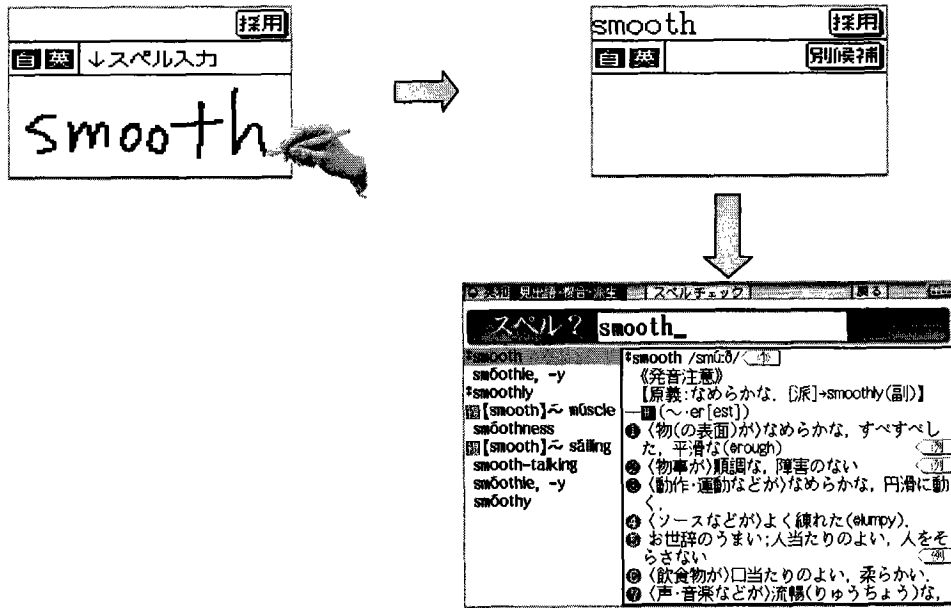


图 14

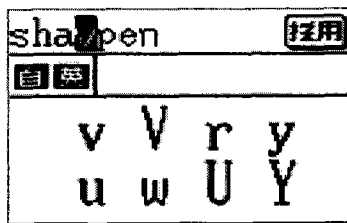


图 15

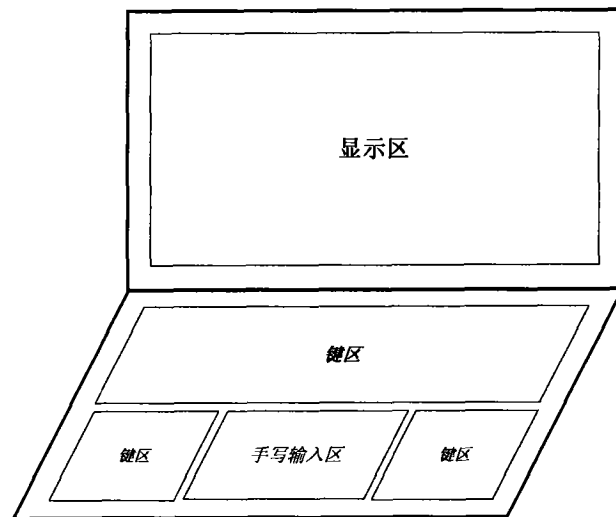


图 16A

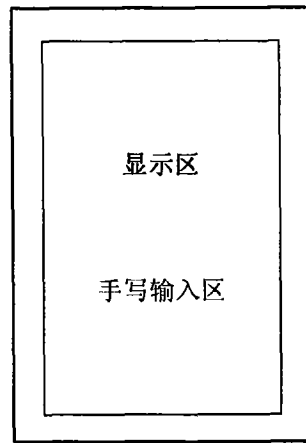


图 16B