

# Windows Mobile 平台下智能手机输入法研究

殷芳刚<sup>1</sup>, 吴建国<sup>1,2</sup>, 吴海辉<sup>1</sup>, 李 炜<sup>1,2</sup>

(1. 安徽大学 计算机科学与技术学院, 安徽 合肥 230039;

2. 安徽大学 计算智能与信号处理教育部重点实验室, 安徽 合肥 230039)

**摘 要:**笔画输入相比拼音输入在重码率方面具有强大的优势。主要对传统笔画输入法在编码方案、字词库构造等方面进行优化,使得用户能够运用笔画方式在手机上实现快速的文字输入。对于笔画输入法的编码方案,主要结合传统方案给出了一种新的编码方式,在字词库构造方面,采用空格分隔符方式实现笔画的词组、语句输入。最后结合输入法设计流程,给出 Windows Mobile 平台下输入法的实现流程,使得用户能够连续地输入笔画字符串,实现单字、词组以及语句的输入,提高汉字的输入效率。

**关键词:**Windows Mobile; 笔画; 输入法; IME

**中图分类号:**TP39

**文献标识码:**A

**文章编号:**1673-629X(2011)05-0075-04

## Research of Input Method for Intelligent Mobile Phone on Windows Mobile Platform

YIN Fang-gang<sup>1</sup>, WU Jian-guo<sup>1,2</sup>, WU Hai-hui<sup>1</sup>, LI Wei<sup>1,2</sup>

(1. School of Computer Science and Technology, Anhui University, Hefei 230039, China;

2. Ministry of Education Key Laboratory of Intelligent Computing and Signal Processing, Anhui University, Hefei 230039, China)

**Abstract:** Stroke input has a strong advantage relative to the Pinyin input in weight rate. Traditional stroke input method will be optimized in the coding scheme, word thesaurus construction, etc. Then, the user can use stroke to achieve fast text input on the phone. For stroke input coding scheme, a new coding method is designed based on the traditional program. In the structural aspects of the word thesaurus, using a space delimited to realize the strokes phrase, sentence input. Based on the input method design process, the input method of approach on the windows mobile platform was given, users can input a continuous string of strokes, words, phrases and statements can be entered, Chinese character input efficiency is improved.

**Key words:** Windows Mobile; stroke; input method; IME

## 0 引言

随着科学技术的发展,以手机为代表的移动通讯产品得到了快速的发展,手机终端的功能也从单一语音通话,发展到手机短信、手机上网、订阅资讯、软件安装等各种多样化的功能,综合了手机与 PDA 功能的智能手机无疑代表了未来手机的发展趋势<sup>[1]</sup>。

文字输入作为手机的重要组成部分,随着短信息、电子邮件以及文档处理等在智能手机上的普及,用户对快速文字输入的需求日趋强烈,在这样巨大的市场背景下,研究手机输入具有重要的意义。

Windows Mobile 作为微软为手持设备推出的“移动版 Windows”,它将熟悉的 Windows 桌面系统扩展到个人设备中,其基础为 Win 32 API,界面类似于 PC 机的 Windows,易操作。然而伴随 Windows Mobile 系统的输入法只有拼音和笔画等几种,虽然它们都各具一些优点,但还普遍存在效率不高——重码率太高或击键次数太多等缺点。鉴于 Windows Mobile 是一个开放平台,且支持第三方软件,因此研究 Windows Mobile 平台下的输入法变的十分必要。

## 1 输入法系统的设计

Windows Mobile 平台下手机输入法系统通过 IME 接口获取用户的按键事件,然后调用相应的转换模块,转换为相应的内码,最后把输入的按键和获取的内码发送到应用程序,通过用户界面显示出来<sup>[2]</sup>,见

收稿日期:2010-09-15;修回日期:2010-12-21

基金项目:安徽省教育科研重点计划项目(kj2010A023)

作者简介:殷芳刚(1987-),男,安徽六安人,硕士研究生,研究方向为中文信息处理;吴建国,教授,博士生导师,研究方向为中文信息处理、智能 CAD/EDA。



表 1 汉字外码表

汉字	笔画	外码
大	一ノ、	134
小	ノ丨、	324
...	...	...

字库文件的数据结构为:

```
struetRecord
{
String outcode;//汉字外码
String incode;//汉字内码
};
```

字库建立后,输入的字符串与内码进行匹配转换才能输入文字。对于单字符串的模式匹配问题可描述为:

已知:有限符号集合  $\Sigma$

文本串  $T, T = t_1t_2 \cdots t_n \quad t_i \in \Sigma \quad (1 \leq i \leq n)$

模式串  $P, P = p_1p_2 \cdots p_m \quad p_i \in \Sigma \quad (1 \leq i \leq m)$

求出位置集合  $O = \{i\}, (1 \leq i \leq n)$ , 使得  $\forall i \in O, t_{i+k} = p_k (1 \leq k \leq m)$ 。即在一个有限符号集合上的文本串和模式串,找出文本串与模式串完全相等的所有子串出现位置集合。

2.2.3 词库的建立与匹配

词库的建立同字库相类似,词组的外码由该词的每个汉字的外码组成,并在字与字的外码之间加入分隔符“#”(用户输入时为空格符)。表 2 为词组外码表。

表 2 词组外码表

词组	笔画	外码
汉字	丶、一フ、 丶、フフ丨一	44154#445521
安大	丶、フフ丨一 一ノ、	435531#134
...	...	...

当用户输入带有分隔符“#”时,程序自动在词库中搜索,并输出所有匹配成功的外码串所对应的词组。本输入法采用模糊匹配算法进行匹配。首先将输入码串按“\*”分段,若某子段中有分隔符“#”,标出其出现的次序。然后将输入码各段与词库中的外码逐个匹配,其中“?”与“#”不进行匹配。若某子段中含有的分隔符“#”是输入码的第*i*个分隔符,则必须使其与词库外码中的第*i*个分隔符匹配。匹配结束后,在候选窗口中显示结果<sup>[8]</sup>。

3 IME 输入法接口的研究

Windows Mobile 系统中输入法程序的核心构件是 IME,其主要接收键盘事件传递过来的各种消息,并对

这些消息进行分类处理,最终把键盘输入的外码转换为能够被系统识别的内码<sup>[9]</sup>。其中键盘 HOOK 函数用于传递键盘消息给 IME,通过此接口函数,可以将应用程序与 IME 之间的关系描述如下:

当按下键盘后,系统会首先调用接口函数 ImeProcessKey,通过它判别是否需要处理此次键盘事件。

若需要 IME 处理此次键盘事件,则此时将 WM\_KEYDOWN 键盘消息传递给应用程序,然后应用程序调用函数 Translate-Message,将处理后的键盘消息传送到接口函数 ImeToAsciiEx 中;若不需要处理此次键盘事件,则直接反馈键盘消息给应用程序。

IME 处理的各种事件中,会将各种事件产生的消息存放在一个缓冲区中,函数 ImeToAsciiEx 中的参数会以参数形式调用该缓冲区。应用程序也从 IME 产生的消息区域中获得所需要处理的各种消息,若不对消息进行处理,则会直接反馈给默认的 IME 窗口。

因此,在 IME 的实现中,为了完成文字的输入工作,主要是通过编码实现 ImeProcessKey 函数以及 ImeToAsciiEx 函数的功能,通过它们实现外码到内码的转换过程,以达到文字输入的目标<sup>[10]</sup>。

4 输入法的实现

笔画输入作为汉字输入的一种输入方式,为了打破目前手机笔画输入尴尬局面,实现词组、语句级输入方式,文中设计一种笔画输入方法,使用户可以连续地输入笔画字符串,实现汉字的语句输入功能;同时,该输入法具有学习模式,通过自然学习或者用户自定义学习方式,增加总体词库容量,实验表明,通过多次学习,能够明显提高文字输入效率<sup>[11]</sup>。

通过对输入法系统框架的研究,可以将其实现方式划分为 3 个阶段:词组生成模块、语句搜索模块和学习模块。流程图见图 3。

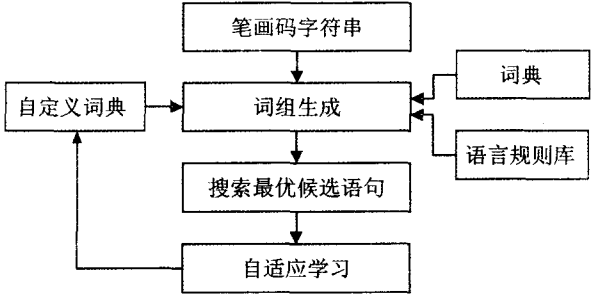


图 3 汉字输入法系统结构流程图

词组生成模块:当用户输入一连串的笔画字符串时,首先将字符串以空格符分离成各段字符串,然后对每段字符串查询字典,查找到每段笔画串对应的汉字,再查询系统内的词典,根据语言规则库中的语言规则形成所有能组合的词组。

语句搜索模块:当形成词组时,需要对形成的词组进行输出以供用户选择,为了提高文字输入效率,需将常用词组排列在前面,此时可运用动态规划中的 Viterbi 算法,统计系统内的各个词组频率值,按照该值进行排列输出。

学习模块:当用户输入修正后的词组或语句时,该模块会自动查询该语句是否已在字典中,若不在,可进行记忆学习,将该语句添加到自定义词典中,提高输入法输入的效率。

## 5 结束语

文中首先分析了 Windows Mobile 平台下智能手机输入法的构成和输入法用户接口,并设计了输入法的编码方案和字词库的创建与匹配,然后对实现输入法的接口原理进行了详细分析,最后给出了笔画输入法实现原理,从此可看出,本输入法可使用户连续进行笔画输入汉字,减少了停留时间和重码率等,提高了输入效率<sup>[12]</sup>。

随着 3G 时代的来临和智能手机用户数量的增多,它提供的强大娱乐功能逐渐丰富了人们的生活,而目前在 Windows Mobile 平台下笔画输入效率低下,且不具有词组以及语句输入。因此文中提出的笔画输入法设计模式对于智能手机输入技术的研究具有重要的意义。

(上接第 74 页)

## 4 结束语

文中给出一种基于多核环境下蚁群的并行优化算法,利用多核的优势并行处理蚁群算法中耗时较长的循环迭代和循环赋值部分。实验结果表明经过改进后的蚁群算法在多核环境下解决大规模 TSP 问题时所需运行时间明显少于单核机上的运行时间,程序执行效率明显提高。

### 参考文献:

- [1] OpenMp C and C++ Application Program Interface (Version 2.0) [EB/OL]. 2002. <http://www.openmp.org>.
- [2] 段海滨. 蚁群算法原理及其应用[M]. 北京:科学出版社, 2005:34-37.
- [3] 付延友. PC 机群环境下蚁群算法的并行化研究[D]. 天津:河北工业大学,2007:29-43.
- [4] Stutzle T, Hhoos H. The MAX-MIN ant system and local search for the traveling salesman problem[C]// Proceedings of the IEEE International Conference on Evolutionary Computation (ICEC'97). USA: Indianapolis, 1997:309-314.
- [5] 王娟,王建. 一种求解 TSP 问题的改进蚁群算法[J].

### 参考文献:

- [1] 欣闻. 手机文字输入技术及其发展趋势[J]. 现代通信, 2005(2):34-35.
- [2] 傅曦,齐宇,徐俊. Windows Mobile 手机应用开发[M]. 北京:人民邮电出版社,2005.
- [3] 李培峰,朱巧明,钱培德. 一个应用于手持设备的汉字通用输入模型[J]. 计算机工程, 2006,32(18):258-260.
- [4] 乔建良,赵增建. 谈谈手机的中文输入法[J]. 现代通讯, 2004,12(3):41-42.
- [5] 张晋. 汉字信息处理研究[M]. 北京:北京语言学院出版社, 1992:4-21.
- [6] Po Lai-Man, Wong Chi-Kwan. Six-Digit Stroke-based Chinese Input Method[C]//Proceedings of the 2009 IEEE International Conference on Systems, Man, and Cybernetics. San Antonio, TX, USA:[s. n.], 2009:818-823.
- [7] 吴海辉. 笔画码汉字输入法[D]. 合肥:安徽大学,2004.
- [8] 高升,王晓龙. 语句级汉字输入系统中语义规则研究[J]. 计算机工程与应用,2003(4):80-82.
- [9] Microsoft. Win32 Multilingual IME Application Programming Interface[M]. [s. l.]:[s. n.], 2003.
- [10] Microsoft Corp. Win32 Multilingual IME Application Programming Interface[M]. [s. l.]:[s. n.], 1998.
- [11] 徐志明,王晓龙,姜守旭. 一种语句级汉字输入技术的研究[J]. 高技术通讯,2000(1):51-55.
- [12] 李政. 计算机汉字输入技术的现状和发展趋势[J]. 松辽学刊,1999,12(2):33-36.

计算机技术与发展,2008,18(12):50-52.

- [6] 崔明义,张新祥,苏白云,等. 用蚁群算法实现地理信息系统空间曲线的描述[J]. 计算机工程与应用, 2008, 44(30):160-162.
- [7] 赖金富. GIS 和蚁群算法及其在城市交通分配中的应用研究[D]. 昆明:昆明理工大学,2008:22-23.
- [8] Stützle T. Parallelization strategies for ant colony optimization[J]. Lecture Notes in Computer Science, 1998, 1498:722-741.
- [9] Marcus, Randall, Andrew, et al. A parallel implementation of ant colony optimization[J]. Journal of Parallel and Distributed Computing, 2002, 62:1421-1432.
- [10] Ellabib I, Calamai P, Basir O. Exchange Strategies for Multiple Ant Colony System[J]. Information Sciences: An International Journal, 2007, 177(5):1248-1264.
- [11] Middendorf M, Reischle F, Schmech H. Multi Colony Ant Algorithms[J]. Journal of Heuristics: Special Issue on Parallel Metaheuristics, 2002, 8(3):305-320.
- [12] 何丽莉,王克森,白洪涛,等. 基于 CMP 的多种并行蚁群算法及比较[J]. 吉林大学学报(理学版), 2010, 48(5):787-792.