

# 基于 Web 页面的藏文在线输入技术

于洪志, 何向真

(西北民族大学中国民族语言文字教育部重点实验室, 兰州 730030)

**摘要:** Web 页面藏文在线输入技术, 能够在浏览器中脱离本机输入法而进行藏文输入, 实现藏文网络在线文字交互, 为网络系统提供了跨平台的藏文输入解决方案。阐述了基于 Web 页面藏文在线输入技术的工作原理及基本设计思想, 介绍藏文在线输入法的组成、内码外码设计原则和输入法流程, 对藏文输入法进行系统分析, 给出实现模型, 论述浏览器内嵌藏文字体信息技术, 达到在线、即时的藏文输入。采用内嵌法和外挂法, 实现藏文在线输入技术与主流网页编辑器的整合。

**关键词:** Web 页面; 藏文; 在线输入; 字体嵌入

## Tibetan Online Input Technology Based on Web Page

YU Hong-zhi, HE Xiang-zhen

(Ministry of Education Key Laboratory for Chinese Ethnic Minority Language, Northwest University for Nationalities, Lanzhou 730030)

**【Abstract】** By using the technologies of Tibetan online input on the Web, Tibetan characters can be input to a Web browser directly without using local input methods. It accomplishes the online character interaction of Tibetan language, and provides a cross-platform solution of Tibetan input method for a network system. This paper explains the basic design philosophies and the principles of online input technologies of Tibetan language which is based on the Web pages. It establishes the components of the online input method of Tibetan language, the principles to design internal and external codes and the flows of the input method. It applies a system analysis and provides a real model of Tibetan input method, and addresses the information technologies of Tibetan typeface which are embedded inside browsers to achieve online and real-time inputs of Tibetan language. It also implements the combination between online input technologies of Tibetan and main webpage editors by using embedded and external methods.

**【Key words】** Web page; Tibetan; online input; font embedded

Web 页面藏文在线输入技术是基于网页浏览器的藏文输入平台。普通藏文用户在全球联网计算机上, 不安装任何藏文输入系统, 能即时在线输入藏文。该输入技术解决了未安装藏文输入系统的计算机输入藏文的问题, 实现藏文网络在线文字交互。本文以现代藏文拼写输入法为例, 论述在线输入现代藏文所有字符和常用梵音藏文技术。

### 1 Web 页面藏文在线输入法的组成

Web 页面藏文在线输入设计利用 Web 页面编程技术, 藏文的输入输出都是在 Web 浏览器上来实现, 它不归 Windows 的输入法管理器管理。Web 页面藏文在线输入技术和 Windows 本机藏文输入法原理的相同之处在于: 都具有键盘事件, 编码字典等; 差别在于: 系统的键盘事件不通过输入法管理器, 而直接对应用户编码字典, 输入法再翻译键盘事件为对应藏文, 然后将键盘事件传给 Web 浏览器的文字编辑框中, 这样就实现了 Web 页面的藏文输入。

键盘事件监控用户按键, 来预测用户可能需要的藏文字符, 并提供一张备选藏文字符列表以供用户从中选出所需藏文字符, 备选藏文字符列表就是用户编码字典。在默认情况下, 程序为用户提供一个可以键入并选择备选藏文字符的 IME 窗口。通过使用 IME 变换功能提供的定制接口, Web 输入法提供类似 Windows 输入法中的输入法管理器 (Input Method Manager, IMM)、函数和消息创建自己的输入法编辑 (IME) 窗口<sup>[1]</sup>。

#### 1.1 状态、编码和备选窗口

用户接口为程序提供状态、编码和备选窗口。状态窗口

指示程序被打开, 并向用户提供设置转换模式的方法, 要让 Web 藏文输入法打开, 必须检测 Windows 本机输入法是否为英文状态输入法, 如果为非英文输入法, 在程序输入框内启用 Windows 本机输入法输入文本, 而非 Web 在线输入法。在用户输入文本时, 编码窗口启动, 依赖转换模式显示输入文本或转换后的文本。备选窗口同编码窗口也一同出现, 包含待选的备选字符或编码窗口中的字符, 用户可以在备选列表中选择所需字符, 然后返回至编码窗口, 此时发送藏文字符至文本输入窗口。用户可以在线录入所需藏文文本, 直至录入藏文字符串结束且窗口关闭。图 1 所示为系统与输入法的关系。

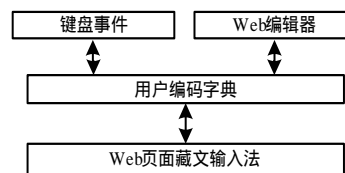


图 1 操作系统与 Web 页面藏文输入法的关系

#### 1.2 IME 窗口类

“IME”窗口类是预定义的定义标准 IME 窗口行为和

**基金项目:** 国家“863”计划基金资助项目(2003AA115530); 国家自然科学基金资助项目(02BYY038); 教育部科研基金资助重点项目(105172)

**作者简介:** 于洪志(1947-), 女, 教授、博士生导师, 主研方向: 藏文信息处理; 何向真, 硕士研究生

**收稿日期:** 2007-10-30 E-mail: yhz@xbmu.cn

观系统全局类。程序用 IME 类创建自己的 IME 窗口。这就允许程序利用默认的 IME 窗口进程以及窗口位置布局。

### 1.3 IME 消息

当某个键盘事件发生时,系统送出 IME 窗口消息到程序窗口程序来影响 IME 窗口。例如,当激活一个窗口,系统给文本输入窗口送出 IMEDOWN(EVENT)消息。

### 1.4 录入字符串

录入字符串是录入窗口的当前文本。即 IME 转换到最终字符的文本。每个录入字符串由一个或多个 Clauses 组成,这里 Clauses 是 IME 能转换到最终字符的字符最小组合。用户在录入窗口输入文本时,IME 跟踪录入窗口的状态,包括属性信息、clause 信息、键入信息和光标位置。

### 1.5 热键

热键使用户可以快速地改变 IME 输入模式或切换到英文 IME。程序为用户增加了几个很必要的热键,这些键可以帮助用户快速的来使用此输入法,但在不同的 BROWSER 中,键盘事件有相冲突的地方,此时针对不同的 BROWSER 而设置不同的热键。

## 2 藏文输入流程

### 2.1 Web 页面藏文输入法键盘布局

Web 页面下藏文输入法键盘布局,以藏文书写规则为主,考虑键盘的常用习惯,在藏文键盘字母布局表中,合理设计了上加字、下加字、反字、基字及常用字符的位置<sup>[2]</sup>,如图 3 所示。

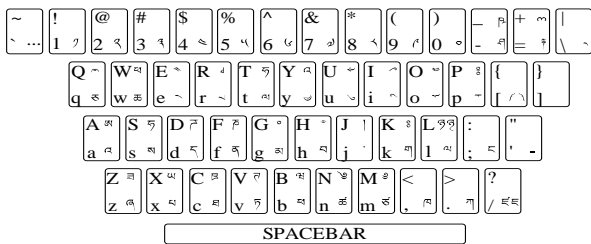


图 2 Web 页面藏文在线输入法键盘布局

### 2.2 藏文输入的基本结构

程序系统结构及藏文输入流程如图 3 所示。

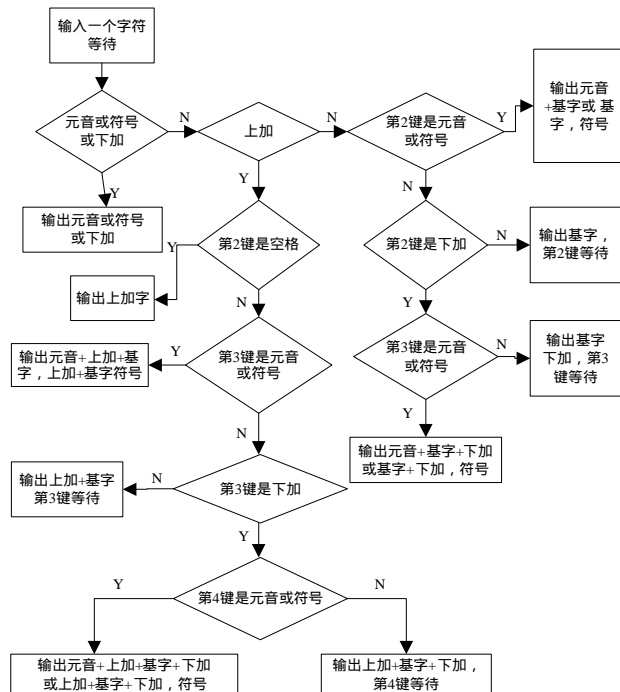


图 3 程序系统结构及藏文输入流程

藏文输入流程的程序有 2 个窗口:输入法转换接口和输入法用户接口。输入法转换接口由转换函数提供,这些函数被类似 Windows 系统输入法中的 IMM 调用。输入法用户接口由一组窗口提供,这些窗口接收消息并输出相应藏文<sup>[3]</sup>。

程序的多平台支持是由一组判断函数提供,程序运行首先判断用户浏览器类型,然后跳转至 Internet Explorer Browser 或 Netscaped Browser 相应的代码,实现多浏览器多平台的支持,如图 3 所示。

键盘事件监控用户的按键,来预料用户可能需要的字符,并提供一张备选字符列表以供用户从中选出所需字符,备选字符列表就是用户编码字典。

当用户敲击键盘时,程序作出判断,如果用户敲击的是元音或符号或下加字时,则程序直接输出字符至文本输入窗口;如果用户敲击的不是元音或符号或下加字时,程序再次进行判断,如果是上加字,即用户按下 q, w, e 按键时,并且第 2 键是空格的话,则输出上加字;如果第 2 键是元音或符号时,程序再次进行判断,依图中所示的那样,依次输出藏文或者符号。

### 3 藏文输入流程

藏文输入流程的程序有 2 个窗口:输入法转换接口和输入法用户接口。输入法转换接口由转换函数提供,这些函数被类似 Windows 系统输入法中的 IMM 调用。输入法用户接口由一组窗口提供,这些窗口接收消息并输出相应藏文。

键盘事件监控用户的按键,来预料用户可能需要的字符,并提供一张备选藏文字符列表以供用户从中选出所需藏文字符,备选藏文字符列表就是用户编码字典。

#### 3.1 藏文输入字符的选取原则

Web 页面藏文输入法字符选取依据国家标准藏文编码字符集扩充集 A。

#### 3.2 外码设计原则

根据藏文基字的使用频率,并参考了 GB/T 17543-1998 《信息技术 藏文编码字符集(基本集)键盘字母数字区的布局》标准<sup>[4]</sup>,结合藏文的书写特点和习惯,设计了藏文键盘布局,将现代藏文和梵音藏文分布在一个键盘表中,有效地离散了重码。

#### 3.3 用户编码字典

考虑到程序的执行速度、后续维护、重码率等方面,程序采用独立的用户编码字典文件,并以 JS 文件存放。为程序的维护带来方便,并且可以自由扩充。

#### 3.4 藏文输入法系统分析

藏文属于拼音文字,每个词都是由字符上下左右拼写叠加组成。

在设计 Web 藏文输入法时采用键盘映射方式,即将每个按键按下后传送给浏览器的输入控件,程序扫描码映射为需要输出的藏语文字符编码。这主要是针对世界上基于字母的语言使用的。在键盘抽象中,需要给出一个转换表(符号表),并有状态相关的转换规则,而且系统需要确定状态的表示方法;此外还需要定义特殊按键的控制作用。输入法中将英文与藏文分别设计成一组键,用一个修饰键在组之间切换,同时其他如汉字输入法仍然使用 Windows 本机输入法的方式。从而使 Web 编辑器可以同时输入藏文、汉字与英文。设定切换藏汉英文输入的快捷方式,可以达到顺利输入所需文字的目的。当系统输入法默认是汉字输入法时,此时,Web 编辑器中则为汉字输入,当系统输入法默认为英文输入状态时,

