

文章编号: 1003-0077(2009)05-0119-04

基于网络机顶盒的维吾尔语输入法研究与实现

程新方, 吾守尔·斯拉木, 张永才

(新疆大学 信息科学与工程学院, 新疆 乌鲁木齐 830046)

摘要: 该文介绍了一种在网络机顶盒遥控器上实现的维吾尔语输入法, 输入法采用遥控器上的数字键 2~9 和 5 个控制键来实现维吾尔语的字符输入。首先, 对网络机顶盒、维吾尔语字母特点进行了分析。然后, 对机顶盒遥控器维吾尔语输入法的框架、功能描述、键盘布局、处理流程和移植过程进行了研究。把网络机顶盒作为视频解码终端, 利用宽带网的基础设施, 以家用电视机为主要显示终端设备, 遥控器维吾尔语输入法在网络机顶盒上实现了多文种混合输入显示, 并成功地应用于双向有线网的交互业务中。

关键词: 计算机应用; 中文信息处理; 嵌入式系统; IP 机顶盒(IPSTB); 维吾尔语; 输入法

中图分类号: TP391

文献标识码: A

Research and Implementation of Uyghur IME for the IP Set-Top Box

CHENG Xinfang, WUShouer Silamu, ZHANG Yongcai

(Institute of Information Science and Engineering, Xinjiang University, Urumqi, Xinjiang 830046, China)

Abstract: A kind of Uyghur IME for the remote control of the IP Set-Top Box is presented, by which the input of Uyghur character can be accomplished by numeric keys 2~9 and five control keys on the remote control. Firstly, in this paper, the IP Set-Top Box and Uyghur letters characteristics are analyzed. Then, the Uyghur IME framework for the remote control, the function description, the keyboard layout and the handling and transplantation procedure are discussed. This mixed multi-language input and display technique is featured by the IP Set-Top Box as the video decoding terminal, the household TV as the main display terminal equipment and the direct employment of the broadband network infrastructure. This method has been successfully applied to interactive service of two-way cable networks.

Key words: computer application; Chinese information processing; embeded system; IP Set-Top Box; Uyghur; IME

1 引言

随着数字广播、因特网的迅速发展和宽带网络的普及, IP 多媒体技术日趋成熟, 多媒体内容提供商逐步以宽带网络平台作为其新的发布平台。近年来, 新疆维吾尔自治区政府加快信息化的发展步伐, 加速网络基础设施的建设, 随之而来的是网络机顶盒的各种交互业务在新疆地区的迅速发展, 同时也催生了广大少数民族客户在网络机顶盒的各种交互业务中对支持维吾尔语显示和遥控器输入法的迫切

需求。因此, 本文研究并开发了一种比较简单的网络机顶盒遥控器维吾尔语输入法, 在一定程度上满足了广大少数民族客户的需求并对新疆地区整个信息化进程起到了促进的作用。在对此输入法的研究过程中, 采用了任意一款家用电视机作为显示终端, 以网络机顶盒作为平台, 以网络机顶盒的红外遥控器作为输入设备, 实现了较快速地输入维吾尔语、汉字、英文字母、标点符号和一些比较常见的特殊符号, 满足了客户通过操作网络机顶盒的遥控器进行比较简单但又必须的文字交互需求。

收稿日期: 2008-09-27 定稿日期: 2008-12-25

基金项目: 新疆维吾尔自治区高技术研究与发展计划资助项目(200532112)

作者简介: 程新方(1983—), 男, 硕士生, 主要研究方向为多语种信息处理, 嵌入式系统; 吾守尔·斯拉木(1942—), 男, 教授, 博导, 主要研究方向为多语种信息处理; 张永才(1980—), 男, 硕士生, 主要研究方向为嵌入式系统。

2 IP 机顶盒概述

IPSTB(IP Set-Top Box)是宽带音视频解码的终端设备,中文名为宽带 IP 机顶盒,又称宽带网络 DVD;IPSTB 是一种基于宽带 IP 接入网,以电视机为显示终端、以流媒体视频等服务为主的网络终端设备^[1]。与传统的数字机顶盒相比,IPSTB 是计算机技术、通信技术、多媒体技术与家电产品之间的一种融合,实现了视频、语音、数据三者的融合,即所谓的三网合一业务(Triple Play Service)。IPSTB 的诞生大大降低了享受宽带多媒体的门槛,为宽带多媒体走进千家万户提供了一种经济、便捷的方式。

IPSTB 相当于一台自带嵌入式操作系统的计算机。从图 1 可以看出,IP 机顶盒一端通过 DSL 或以太网方式接入宽带网络,一端将通过宽带网络传输过来的媒体流等信息发送给 IPSTB 进行处理,并将其转换成模拟音视频输出至电视,IPSTB 具备计算机的数据交换等基本功能,此时用户在不需要计算机的情况下,只需简简单单的操作 IPSTB 遥控器即可完成许多的功能操作,例如:视频点播、电视直播、网页浏览、信息互动、信息检索等,用户通过电视机即可享受宽带网络带来诸多服务。

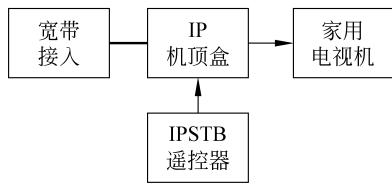


图 1 IPSTB 应用的一般架构

IPSTB 作为宽带音视频解码的终端设备,在应用架构中处于核心地位,那么从应用角度来看,它的主要功能有以下三方面:

(1) 支持目前的 LAN 或 DSL 网络传输,接收并处理 IP 数据和视频流;

(2) 支持 MPEG, Mp3 和 MPEG4 AAC 等音视频解码;

(3) 支持用户认证功能、通过与服务提供商的认证系统进行通信来实现用户的访问控制、计费等管理功能。

3 遥控器维吾尔语输入法的实现

3.1 维吾尔语的语言特征

维吾尔语系阿尔泰语系的拼音文字,它包含 32

个字母,维吾尔语与英文在书写形式上有些相似,都是在输入一个单词时必须逐个字母的进行输入。但是,维吾尔语在书写方向上与汉字和英文又有很大的不同,首先是书写方向相反,汉字和英文的书写方向是从左到右,称为右向输入文字;而维吾尔语的书写方向是从右向左,称为左向输入文字^[2]。其次是维吾尔语的同一个字母依在单词中的位置不同会有 4 种变形,在文字输入时要根据字母在词中的位置来确定使用何种形式。

在计算机存储、处理和显示维吾尔语的过程中,结合维吾尔语自身的特性把维吾尔语的显示形式分成了名义字符形式和变体显现形式。名义字符形式,是指字符在 Unicode 中有确定的编码。变体显现形式,是指在语义上与名义表现形式相同,但由于单字符在字词中所处位置的不同而输出显现为不同变体形式的字符。在 Unicode 标准编码中,诸如拉丁语、汉语等大部分文字的显现形式与 Unicode 中编码字符是一一对应的,即字符与它的字形是一一对应的,而且显示的顺序和在内存中存放的顺序是一样的。但是,由于维吾尔语自身的一些特性,使得维吾尔字符显示的顺序与在内存中存放的顺序截然相反。维吾尔语具有的特性有以下三方面:

(1) 与上下文内容相关的显现形式。维吾尔语字符在词的不同位置有着不同的显现形式,也就是说维吾尔语字符存在自动选形问题。一般来说,可以分为以下四种显现形式:独立形、词首形、词中形、词尾形。

(2) 双向性。维吾尔语书写方向既有从右向左,也有从左向右(例如:数字)。但在计算机的处理和存储过程中,无论是什么文字,Unicode 字符串都是以逻辑顺序存储的,即字符串的存储方向都是从左到右。因此计算机在处理多语言文字的 Unicode 字符串时,系统就必需识别出相应文字的方向。

(3) 连字形式。维吾尔语文字常用一个特殊的字形来显示(或代替)几个字形。

3.2 输入法总体框架

实现 IPSTB 遥控器维吾尔语输入法的总体思路是:采用普通的红外线遥控器作为输入设备,当用户按下网络机顶盒遥控器的数字按键后,遥控器的红外发光二极管发射出有编码的红外光^[3];然后通过某种特定的方式获取用户当前所输入的信息,当机顶盒的遥控接收端收到红外信号并对该信号进

行解码后,再把控制码送到机顶盒的应用程序中,经过搜索、选形引擎将用户当前所输入的信息转换成为其对应的维吾尔语字母,最后给用户提供维吾尔语字母的选择窗口,由用户选择所需要的字母,在应用程序与输入法搜索引擎之间进行交互以完成输入法的功能。

3.3 功能描述

IPSTB 遥控器既然要作为提供文字输入的工具,就必须使用户可以通过操作遥控器方便地完成汉字、英文以及维吾尔语、数字、符号的混合输入。

IPSTB 遥控器维吾尔语输入法的具体功能描述如下: 用户在文本编辑过程中,把输入法调整为维语输入法后,可以通过操作遥控器上 2~9 数字键和 5 个控制键来实现维吾尔语字母的输入,并且用户可以根据需要可在汉字、英文以及维吾尔语、数字、符号输入状态间进行切换,最终完成文本编辑工作。

IPSTB 遥控器维吾尔语输入法各个部分的具体任务描述如下:

(1) IPSTB 遥控器: 按键输入,发射一定编码的红外光。

(2) IPSTB 按键解码: 将红外光解码为计算机程序可以识别的数值,例如: 输入维吾尔语字母“ئ”时所对应的编码值为 0x0646。

(3) 处理按键消息: 处理不同数值的按键消息,与输入法搜索引擎交互,向搜索引擎发送字符,接收并在电视屏幕上显示从搜索引擎返回的候选字符。

(4) 输入法搜索、选形引擎: 根据 OSD(On-Screen Display)图层界面传进来的字符串,采用一定的算法从字库搜索与之匹配的维语字母并返回相应字符给界面显示,由于维语在输入中存在字母变形问题,所以选形引擎采用维吾尔语单字符选形算法进行名义字符到显现字符的转换。

(5) 字库: 专门用来存储汉字、维吾尔语字母、英文字母、数字、标点符号及特殊符号,按照常用性排列,与输入法搜索程序相互独立,可自由扩展。

3.4 输入法按键布局

IPSTB 遥控器维吾尔语输入法共使用了遥控器上其中的 13 个按键,即 2~9 数字键和 5 个控制键。根据调查与维吾尔语相关的电子产品的键盘布局,以方便用户按键输入为宗旨,同时考虑到很多民

族用户熟悉维吾尔语手机键盘的布局^[4],所以采用维吾尔语手机键盘布局作为 IPSTB 遥控器维吾尔语输入法的键盘布局,如表 1 所示的遥控器键盘布局方式,操作遥控器上的数字键 2~9 和 5 个控制键即可实现维吾尔语的字母输入。

表 1 维吾尔语字母在 IPSTB 遥控器键盘上的布局

数字 按键号	维吾尔语 字母布局	数字 按键号	维吾尔语 字母布局
键盘 2	غ چ ب ئا	键盘 6	ڭ ئۇن م
键盘 3	ف ئى ئى د	键盘 7	س رق ب
键盘 4	ئى خ ھ گ	键盘 8	ئ ئۇ ئۇ ت
键盘 5	ئۆل ك ج	键盘 9	ڙ ز ي ش ڦ

如您需要输入字母“ئ”,可以按照以下步骤进行:

- (1) 操作“输入法”按键并切换为维文输入法;
- (2) 按下“6”键;
- (3) 通过 5 个控制键中的方向键来进行选中,接着按一下确定键即可输入字母“ئ”。

3.5 输入法的处理流程

在需要使用 IPSTB 遥控器维吾尔语输入法的时候,首先,使用遥控器“输入法”按键开启输入法,配置相应的 GUI;其次,操作遥控器按键进行字符输入,机顶盒将收到的键值转换为字符串并将该字符串传给字符串查询模块,通过串比较,在字库中进行搜索,如果存在匹配的字符串,系统将返回相应的匹配维吾尔语字母,此时用户可根据列出的维吾尔语字母通过操作方向键来选择自己想要的字符;最后选形引擎根据字符所放置的位置进行自动选形并输出显示选形后的维吾尔语字母,然后进入下一个循环操作。在整个文本输入过程中,用户可以依据需求随时切换输入法状态来完成多文种的混合输入。

3.6 输入法的移植

在进行 IPSTB 遥控器维吾尔语输入法的移植时,首先将 IP 机顶盒通过串口与本地 PC 机进行连接^[5-6];其次,修改字库目录映射文件,并将维吾尔语字库、PC 机上对输入法源文件经过交叉编译后生成的应用程序文件添加到要烧写的目录下;然后通过远程登录进入 IP 机顶盒的根目录下,通过网络文件系统挂载要烧写的文件目录到 IP 机顶盒开发板上。

并进行测试,待测试成功后就可以进行最后的烧写,烧写成功后即可完成维吾尔语输入法的移植。

4 测试与结果

测试的基础平台采用 VCom 公司的 IP 机顶盒,该 IP 机顶盒采用的嵌入式操作系统为: Monta-Vista Linux (2.4.20 内核), 主处理器: IBM STB02500, DSP: ADI BlackFin BF533, 内存: 64MB 系统内存 + 8M 图像内存 + 8M DSP 内存, MPEG-2 解码器等单元。

开发完成的网络机顶盒遥控器维吾尔语输入法,已经成功的移植到了该型号网络机顶盒上,并通过了较为严格的测试,用操作遥控器的维吾尔语输入法在网页上输入维吾尔语的测试页面,如图 2 所示。本输入法已应用在该型号网络机顶盒的一些具体的交互业务中,例如: 网页浏览,信息发布,信息检索等。

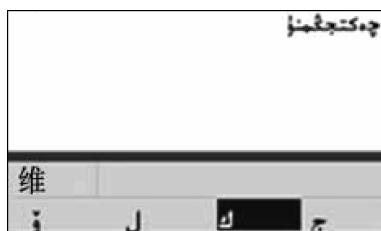


图 2 测试页面

5 结束语

互联网的飞速发展加快了宽带网络的普及,基于双向网络的很多交互业务也迎来了一个快速发展的时期,而在一些依靠网络机顶盒为平台的交互业务中都离不开遥控器输入法的支持,在新疆地区,尤其是维吾尔民族对通过遥控器进行维吾尔语输入有

迫切需求,因此,在 IP 机顶盒上能够设计开发出支持维吾尔语并能实现维吾尔语输入是很有必要的,同时对新疆地区的信息化、经济的快速发展也能起到一定的促进作用。本文研究了维吾尔语遥控器输入法,并给出了设计方案和实施步骤,实现了一种比较简单但能满足基本的文字交互需求的 IP 机顶盒遥控器输入法,在原有支持中英文、数字、标点符号及特殊符号输入的基础上增加了维吾尔语输入和混合显示功能,目前已经应用在了数种型号的 IP 网络机顶盒中,在新疆地区已经得到了广泛的应用并获得了较好的市场效益。

通过走访使用该输入法的用户,发现在用户使用过程中,该维吾尔语输入法也有些不足之处,譬如用户需要进行大量文本输入时,由于维吾尔语的单个词长度较长,逐个字母输入时速度较慢,用户耗费较长的时间才能完成文本输入,所以如果能够研究出具有联想功能的维吾尔语输入法并应用于 IP 机顶盒上是很有意义的,也是今后研究的重点所在。

参考文献:

- [1] 蒋艳. IP 机顶盒的功能和发展[J]. 世界电信, 2005, (11): 48-51.
- [2] 吾守尔·斯拉木, 努尔麦麦提·尤鲁瓦斯, 张玲. Windows CE. Net 嵌入式操作系统本地化研究及实现[J]. 新疆大学学报(自然科学版), 2007, 24(3): 269-273.
- [3] 张真波, 李俊, 倪宏, 等. 电视机遥控器中文输入法的实现[J]. 计算机工程, 2008, 34(2): 278-279, 282.
- [4] 热依曼·吐尔逊, 吾守尔·斯拉木. 维吾尔文手机输入关键技术研究与实现[J]. 中文信息学报, 2006, 20(2): 72-77.
- [5] 闵华松, 童学才, 陈启卷. 嵌入式中文输入法设计[J]. 微计算机信息, 2006, 22(23): 41-43.
- [6] 周佳淳, 陆以勤, 吕锦. 基于嵌入式 Linux 的汉字输入法[J]. 微计算机信息, 2006, 22(26): 60-62.