

文章编号:1004-9037(2012)06-0737-05

基于眼视线跟踪的打字系统在 TMS320DM6446 EVM 上的实现

崔 耀 何 迪 于明轩 王军宁

(西安电子科技大学通信工程学院,西安,710071)

摘要:基于TMS320DM6446 EVM 硬件平台,使用Codec Engine 模块协调双核工作,ARM 核完成眼图的采集、显示及与计算机间的通信,DSP 核处理眼图、计算眼视线注视点。计算机将视线注视点转换为其所处的字符输出到文档中。DSP 端的核心算法采用C 语言与汇编混合编程,经过C 语言与汇编的优化,在C6416 下从4~7 s,优化到10~20 ms,移植到 TMS320DM6446 下达到15 帧/s,实现了视线打字功能,字符输入速率为40 个/min,准确率90%。

关键词:视线打字;视线注视点;程序优化

中图分类号:TP391.41

文献标识码:A

Realization of Eye-Tying System Based on Eye-Tracking on TMS320DM6446 EVM

Cui Yao, He Di, Yu Mingxuan, Wang Junning

(School of Telecommunications Engineering, Xidian University, Xi'an, 710071, China)

Abstract: Eye-typing system uses gaze point for text input, providing a kind of intelligent human-machine interaction method. Using Codec Engine module to coordinate the dual-core work on TMS320DM6446 EVM, the ARM side captures eye video, displays it on the monitor and communicates with host PC, while the DSP side processes eye images and calculates gaze points. The host PC translates gaze points to characters, and outputs them to document. The core algorithm on the DSP side is the integrated C language and Assembly language. Optimized by these two languages, the time cost is reduced from 4~7 s to 10~20 ms on C6416, and when porting to TMS320DM6446, it reaches 15 frames per second, with the input rate of 40 characters per minute and the accuracy of 90%.

Key words: eye-typing; gaze point; program optimization

引 言

眼睛是外界信息进入人脑的主要通道,对其活动情况的研究有助于理解人类对信息加工的过程。

将视线跟踪技术用于人机交互,可以帮助渐冻人、高位截瘫、脑瘫等残疾病友使用计算机。根据第二次全国残疾人抽样调查,中国目前各类残疾人总数为8 296 万人。目前,计算机打字输入系统有基于手写输入的,有基于语音识别以及基于视线跟踪技术输入的。对于缺乏语言能力的重度残疾人,不能

通过手写或者语音识别来完成打字输入,因此可利用视线跟踪技术来完成打字输入。视线跟踪的打字输入也适用于输入密码等保密信息,以防止手写或者语音输入的信息被窃取。

基于视线跟踪的打字输入方法中,最常采用的方法是在计算机屏幕上设置键盘,通过视线跟踪系统分析用户注视的字符,从而实现文本输入^[1]。在现有视线跟踪输入技术中主要存在以下缺陷:

(1)注视点识别精度低。由于注视点识别精度低,屏幕键盘仅由有限个键组成,不仅输入字符内容受限,也大大降低了字符输入速度。如文献[2]研

制一种基于注视点文本输入系统,屏幕仅被分为12个方格,由6个常用的英文字母和最可能的6个单词键组成,这样仅仅对最常用的英文字母和英文单词字有较高的输入速度,遇到其他字母和单词的输入较为困难。

(2) 输入字符过程复杂,对预期字符需要进行二次选择,从而导致字符输入速度较低。如文献[3],以视线点为中心,先确定一个小区域,该区域的字符键组成候选字符集,并计算候选字符集中各字符中用户期望概率确定空间域候选区,然后经过诱发脑电信号中的P300电位来确定时间域候选区,最后根据两个候选区联合计算出可能性最高的字符,字符输入过程操作复杂。

(3) 用户头部运动受限,使用舒适度低。再如文献[4],在不影响精度的情况下,用户头部运动补偿有限,头部运动范围很小,使用舒适度不高。

为克服上述现有技术存在的不足,本文研究设计了一种基于视线跟踪的字符输入方法,同步获取眼图和场景图像,处理并计算出瞳孔-光斑矢量和标志块中心坐标值,通过标定得到瞳孔-光斑矢量和场景图像之间的映射关系,再通过坐标转换得到场景图像和界面单元坐标之间的关系,从而获取界面单元上人眼注视的字符。由实验测得视线跟踪系统的注视精度为 $0.5\sim1^\circ$,能够达到主流的眼动跟踪技术水平,实现了视控打字,对其进行性能测试,在操作熟练情况下,打字速度为每40个字符/min,准确率可达90%。实现了基本的视线输入人机交互功能。

1 算法设计

图1为视线打字的原理框图,测试者注视计算机显示器,显示器上绘制有键盘(图2)及相应的控制模块,完成标定过程之后,就可以通过注视界面来进行眼动字符输入了。光学子系统安装在头盔上,实时采集人眼图像,并将其送入TMS320DM6446 EVM平台处理,处理结果如图3所示。

实现视线输入关键是跟踪人眼的视线方向。基于瞳孔-角膜光斑法,设计一种眼图处理算法,实时提取眼睛的瞳孔中心以及角膜反射光斑中心,使用基于矩阵分解的二次方程拟合的标定算法,将瞳孔角膜矢量转换为场景摄像机中的坐标,然后通过射影变换坐标转换,计算出屏幕键盘坐标系中的眼注视点,确定要输入的字符,将其输出到打印文档中。

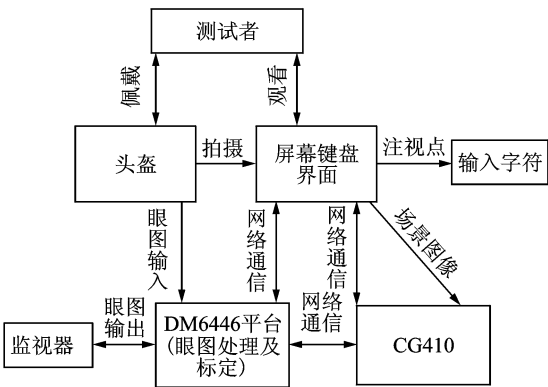


图1 视线打字原理框图

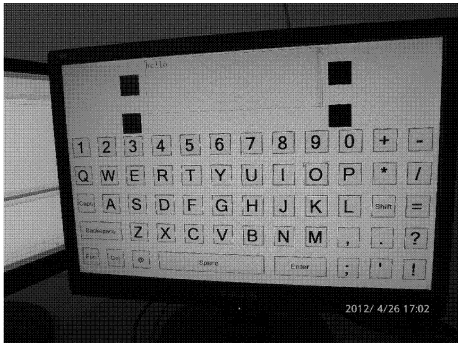


图2 系统界面

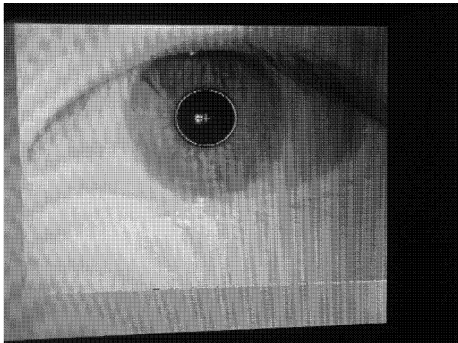


图3 眼图处理结果

1.1 眼图处理算法

1.1.1 瞳孔中心定位的算法设计

瞳孔中心定位的方法多种多样,不同的方法有其自身的优缺点,因此在工程实践中可根据实际需求及所具备的条件进行选择。瞳孔中心定位首先要进行瞳孔边缘的检测。根据边缘检测的原理和方法,结合瞳孔的形状等特点,在现有的瞳孔检测中,常用的方法如下:基于Hough变换圆检测方法、形心法和边界拟合法相结合的方法、瞳孔亚像素边缘检测与中心定位、局部阈值分割的方法^[5]等。

上述方法优缺点比较:眼部图像中瞳孔准确的

说应该是一个椭圆,这样 Hough 变换圆检测方法在准确性方面就有着先天性的缺陷。形心法和边界拟合法相结合的方法要求对亮斑的位置及瞳孔的遮蔽状况进行准确的判断,然后根据判断的结果确定采用的方法,因此该方法的鲁棒性不高。瞳孔亚像素边缘检测与中心定位在实时视线跟踪系统中,由于人眼的内在机制与眼动的复杂性,眼睛图像中含有较大噪声,若不很好地剔除干扰,精确到亚像素级非常困难。局部阈值分割的方法中采用圆形边缘探测算法,与 Hough 变换圆检测方法一样,在准确性方面有着先天性的缺陷。

通过对比分析可知,常用算法在实时性和准确性方面不易兼顾。因此,本文借鉴和参考了相关算法,采用一种瞳孔-角膜矢量算法,该算法是从初始的最佳瞳孔中心两次向外延展星射线,寻找边缘特征点,然后用椭圆拟合的方法定位瞳孔中心。该方法只对瞳孔周围的一定区域进行处理,避免了对整幅图像进行处理,大大减小了所需处理的数据量;该算法中采用的两步特征点探测过程改善了在较差起始点猜测的情况下算法的鲁棒性;且采用基于随机样本一致性的椭圆拟合方法允许杂质点的存在,确保了拟合定位算法的准确性。

1.1.2 瞳孔中心定位的实现

瞳孔中心定位的实现步骤如下:

(1)图像预处理,对输入的图像进行中值滤波,减少输入图像的噪声。

(2)瞳孔中心粗定位,计算图像的灰度直方图,利用直方图求得最佳的灰度阈值对图像进行分割,求其几何中心为瞳孔中心的粗定位。再以第一次定位的瞳孔中心为中心加矩形框进行瞳孔中心的二次粗定位直至收敛。

(3)角膜反射光斑的中心定位,在每一帧中都使用自适应的阈值来进行角膜反射光斑的定位。式(1)比率函数是一个先增后减的函数,此函数最大极值点所对应的阈值就是自适应后的最佳值。图像中最大四连通的部分被认为是真正的角膜反射光斑,其余的为镜像反射光斑。找到四连通面积最大的反射光斑,然后求取其几何中心坐标 (X_c, Y_c) ,即为角膜反射光斑中心

$$\gamma = \frac{S_{\max}}{S_t - S_{\max}} \tag{1}$$

式中: S_{\max} 为最大光斑; S_t 为总光斑。

(4)瞳孔边界特征点的提取,本文采用一种改进的两次延展星射线的方法来获取瞳孔边界特征点,即从预定义的初始点延展出的有限条星射线来

探测瞳孔边缘,使用梯度特征来判断边缘特征点。以第 1 次星射线求得特征点为顶点,以特征点到瞳孔粗定位中心连线为中线,左右各间隔 α° 的两条边所围成的扇形内,等角度地发出 n_1 条射线,每条射线上再次逐像素求梯度以提取二次延展星射线特征点。最后再对眼皮、眼睫毛及光斑干扰而产生的虚假特征点进行剔除,然后依据特征点的位置进行特征点分区处理,使下一步的椭圆拟合更加精确。

(5)椭圆拟合

这里用椭圆一般方程来表示瞳孔轮廓曲线

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0 \tag{2}$$
$$B^2 - 4AC < 0 \tag{3}$$

从上一步求出的特征点中分区域随机选取 6 个点进行最小二乘法拟合,求出椭圆的系数,进一步计算出瞳孔中心坐标。然后依次计算所有特征点到拟合出来的椭圆的欧式距离,记录距离小于 n 个像素的特征点(认为落在椭圆上)的个数,当匹配点个数占总特征点个数一定比率时,认为椭圆拟合成功,否则更改梯度阈值,继续进行边缘特征点提取及椭圆拟合。

1.2 场景图像处理算法

首先对场景图像进行高斯滤波和形态学开运算,再对图像进行二值化,在二值化图像上使用 Hu 矩匹配的方法搜索目标标志区块的轮廓,并计算轮廓的周长和面积参数检测其合理性,对满足合理性的轮廓求其最小外接矩形来定位轮廓中心坐标,并对满足合理性的轮廓进行计数,如果轮廓数目满足既定条件,则视为检测成功,否则更改阈值重新对图像进行二值化和目标标志区块的轮廓检测。如果在确定的阈值范围内检测成功,则输出排序后的各待检测区域的中心点坐标,否则视为检测失败,使用上一帧数据代替。

1.3 标定算法

根据标定的工作内容,将其分为两个部分,即标定部分和跟踪部分。对二维映射模型的标定方法进行了一些改进,加入场景摄像机实现了标定及跟踪过程中,头部均可移动。

标定的过程中首先注视 9 个标定点,眼图提取算法依次计算 9 点的瞳孔角膜,同时图像采集卡程序计算标定点在场景图像中的位置。场景图像上提取出的标定点坐标是 (X_s, Y_s) ,而同时瞳孔-角膜光斑矢量为 (X_e, Y_e) ,用如下的方程

$$\begin{cases} X_s = a_0 + a_1X_e + a_2Y_e + a_3 + X_e^2 + a_4X_eY_e + a_5Y_e^2 \\ Y_s = b_0 + b_1X_e + b_2Y_e + b_3X_e^2 + b_4X_eY_e + a_5Y_e^2 \end{cases} \tag{4}$$

式中,多项式系数矢量 $\mathbf{a}^T=[a_0 \ a_1 \ a_2 \ a_3 \ a_4 \ a_5]$ 和 $\mathbf{b}^T=[b_0 \ b_1 \ b_2 \ b_3 \ b_4 \ b_5]$ 为待求参数。

标定完成后,使用矩阵分解法求解方程组[4]的系数。待求出系数后,就可以实时地由瞳孔-角膜光斑矢量,计算出人眼视线落点在场景图像中的位置,场景摄像机和眼摄像机均固定在头盔上,二者与测试者之间均没有相对运动,因此这个映射关系式是不变的,在头动的过程中也可以跟踪人眼视线。

1.4 坐标转换及字符输入

如图2所示,在头部运动的过程中,实时求解4个黑色标志块中心在场景图像中的坐标值,进而求出射影变换的关系系数矩阵 \mathbf{H} ,再使用 \mathbf{H} 矩阵将场景图像中的坐标值转换为键盘界面上的值,并将结果发送给屏幕键盘。计算机屏幕上的点和对应的场景图像中的点满足

$$\mathbf{X}_i' = \mathbf{H}_{3 \times 3} \mathbf{x}_i \tag{5}$$

在电脑显示器上绘制有字母键盘,以屏幕的左上角为坐标原点,可以知道每个键盘字符的坐标值范围,接收图像采集卡程序发送过来的坐标,依据每个键盘按钮的位置范围判断人眼注视的字符,用红色波门圈出来作为注视反馈给用户,用户进行输入确认。该界面还可以完成简单的计算功能。

2 算法在DSP上的实现

将算法从matlab 仿真移植到TMS320DM6446上,需要3个步骤:C6416下DSP算法的实现与优化;TMS320DM6446下生成Codec包、构建DSP Server,实现DSP端算法;编写GPP端程序,配置Codec Engine。

2.1 DSP 算法的优化

将 matlab 程序及其相应的子函数改写成C语言的形式,经过测试,时间花销为4~7 s,这需要对眼图处理程序进行改写和优化。在改写和优化之前,经过对程序的分析,总结出眼图处理程序的几个特点如下:处理的数据量大;计算复杂;循环情况复杂。

根据这些特点,从如下的几个方面着手对算法进行优化。

(1)使用TI提供的dsplib和imglib。

(2)通过对程序的调整实现优化,使用查表法来代替正余弦函数的调用;改写循环结构;循环的合并;将二维图像矩阵当一维数组处理;使用距离的平方代替距离进行比较;将循环内部反复使用的简单函数定义为内联函数。

(3)浮点运算的处理:①采用TI提供的FastRTS Library和C FastRTS Library处理浮点运算。②先将汇编里面用到的浮点数放大,比如说1 000倍,或者是小数点右移多少位,等汇编处理完成之后,再缩小、左移,具体的放大倍数,需要结合精度考虑,另外还需考虑溢出。

(4)改写线性汇编,分析程序,一共选取出9段代码改写为线性汇编。处理时常常使用些特殊指令,一次读写、比较或处理4个甚至8个像素的数据(LDDW, CMPGTU4, XPND4, PACK2, BITC4等),注意数据对齐以及处理浮点时的溢出问题。

2.2 DSP 端算法实现

DSP端的算法实现先要将算法按照xDM标准封装成Codec包,在运行make命令编译前将需编译的算法文件路径和文件名添加至文件夹内的名为package.bld的文本文件中。修改其他相关文件,保证名称一致,编辑user.bld文件,此文件位于CE的安装目录下,选择所需的平台、编译工具、目标等。修改CE目录下的xdcpaths.mak文件,编辑所用平台、编译器、工具安装目录等,编译生成.a64p的库文件。

当算法编译成.a64p的文件之后就可以进行DSP Server的编译了。编译前需要的编译文件有main.c,link.cmd,* *.tcf,* *.cfg文件。

2.3 编写GPP端程序,配置Codec Engine

在ARM端,除了主线程之外,使用了3个线程,分别负责图像的采集、处理、显示,用DMAI模块提供的API,编写图像的采集及显示模块,使用VISA的函数调用远端的Codec算法,实现图像的处理,使用套接字来协调TMS320DM6446板和计算机间的通信。编写完C代码后,还需要对*.cfg的配置文件进行配置。

3 实验结果与分析

经测试,系统核心算法经过优化,在C6416DSK 1 GHz平台上从之前的4~7 s,优化到

10~20 ms,移植到TMS320DM6446 平台上之后,其DSP 端主频只有594 MHz,片内内存也小,因此在TMS320DM6446 上实时性较差一些。经过测试,平均帧率为15 帧/s。眼图处理的结果如图3 所示。

通过以上的⒉作,研究实现了基于TMS320DM6446 EVM 的眼动跟踪技术,要将其用于视线打字,还必须要保证一定的注视精度。为此进行了专门的大量精度测试实验,典型实验结果如下:其中的1,3,5,7,9 是标定点的序号。

从表1 可以看出,本系统的平均误差精度在0.5~1°,达到了主流眼动跟踪技术的性能指标^[6],可以用于视线打字,若操作熟练,打字速度为40 个/min,准确率可达90%,实现了利用眼睛来打字,为残疾人提供了一种便利的输入方式,让残疾人也能够编辑文档和操作电脑。

表1 系统性能指标 (°)					
	1	3	5	7	9
平均					
误差角	0.504	0.266	0.057	0.104	0.214
绝对误差					
角平均值	0.504	0.266	0.109	0.154	0.214
误差角					
标准差	0.098	0.118	0.119	0.160	0.099
误差角					
最大值	0.791	0.567	0.356	0.576	0.382
误差角					
最小值	0.302	0.148	0.002	0.008	0.029

4 结束语

本文深入研究了眼动跟踪关键技术及其在视线输入中的应用设计方案,设计一套视线跟踪打字系统,实现了眼动视线输入控制的文本编辑及打印,试验验证了整套系统的可行性。本系统相比国内外其他视线输入系统具有以下优点:嵌入式平台实现,具备较好的便携性;高精度,具有较好的实用性;标定和跟踪过程中头部均可动,简化了操作和

使用难度。

参考文献:

[1] Arlos H M, Marcio R M M. Eye gaze tracking techniques for interactive applications[J]. Computer Vision and Image Understanding, 2005, 98(1): 4-24.

[2] 龚秀锋. 头戴式视线跟踪系统关键技术研究 [D]. 合肥: 中国科学技术大学, 2010.

Gong Xiufeng. The key techniques research of head-mounted eye tracking system [D]. Hefei: Circuits and System, University of Science and Technology of China, 2010.

[3] 北京理工大学. 一种基于视线跟踪和P300 脑电电位的字符输入装置[P]. 中国: 200910080852. 5, 2011-01-05.

Beijing Institute of Technology. Character input device based on eye-tracking and P300 EEG potential [P]. China: 2009100808525, 2011-01-05.

[4] 蒋春燕. 视线跟踪技术研究及其在人机交互字符输入中的应用[D]. 上海: 上海交通大学, 2008.

Jiang Chunyan. Research on light-tracking technique and its use in the man-machine interact character input system [D]. Shanghai: Shanghai Jiaotong University, 2008.

[5] You J K, Eui C L, Kang R P. A robust gaze detection method by compensating for facial movements based on corneal specularities[J]. Pattern Recognition Letters, 2008, 29(10): 1474-1485.

[6] 张昀. 视线跟踪技术及可计算眼动建模研究[D]. 西安: 西北工业大学, 2010.

Zhang Yun. Eye tracking technology and computational eye morement model [D]. Xi'an: Northwestern Polytechnical University, 2010.

作者简介:崔耀(1987-),男,硕士研究生,研究方向:视线跟踪技术、DSP 技术, E-mail: wscyllw@139.com; 何迪(1980-),男,讲师,研究方向:图像信号处理、视线跟踪技术、DSP 技术; 于明轩(1988-),男,硕士研究生,研究方向:视线跟踪技术、DSP 技术; 王军宁(1958-),男,教授,研究方向:图像信号处理、视线跟踪技术、DSP 技术。