

交互式儿童视力保护视频智能终端设计

李清霞

(广东理工学院, 广东 肇庆 526100)

摘 要: 文章采用语音命令控制和蒙层屏幕显示相结合的方式, 提供一种有效监视儿童屏幕距离、观看屏幕时间的语音控制智能视频终端的交互方法。该设计提供语音控制智能视频终端的交互方案; 提供屏幕交互方案, 包括如何获取屏幕图像、如何获取背景色、如何识别可点击区域、蒙层页面的生成以及距离监控的实现等。整个系统设计智能化, 提供自动识别屏幕背景色和寻找各按钮点击区域块的方法, 将适用于各类智能视频终端, 观看者只需要安装语音控制软件便可以对现有的视频终端进行语音控制, 不需要各视频终端开发商进行二次开发。

关键词: 语音控制; 屏幕交互; 蒙层页面; 距离监控

中图分类号: TP39 **文献标识码:** A

Design of the Interactive Intelligent Video Terminal for Children Eyesight Protection

LI Qingxia

(Guangdong Polytechnic College, Zhaoqing 526100, China)

Abstract: By means of combining of voice command and layer screen display, this paper provides an interactive way of voice-controlled intelligent video terminal, which effectively monitors the distance between children and screen as well as the screen watching time. This design provides an interactive program of voice-controlled intelligent video terminal, including voiceprint recognition, voice-recognition and analysis. This design also provides screen interactive program which includes how to get screen images and background colors, how to recognize clickable area, generation of covered pages and realization of distance monitoring. The entire system design provides ways of automatically recognizing the screen background color and searching area blocks, which will be applicable to various intelligent video terminals, and by installing voice-controlled software, viewers can control the existing video terminals by speech without further development by software suppliers.

Keywords: voice control; screen covered page; distance monitoring

1 引言(Introduction)

据国家教育部、卫生部最新调查表明: 目前我国共有4亿多近视眼患者, 近视发病率达33.3%, 小学生近视率超过25%, 初中生近视率达到70%, 高中生近视率达到85%, 而且每年还以8%的速度增长, 居世界第一位。90%的近视都是后天近距离用眼过度造成的, 智能终端设备产品的普及, 是加剧近视低龄化趋势的重要原因之一。

使用平板电脑、智能手机等智能终端导致近视的一个重要原因是使用习惯不当, 特别是长时间距离屏幕太近, 孩子的眼睛在屏幕发出的强光长时间刺激下, 不仅容易使孩子视觉的敏锐度和适应性降低, 而且轻易造成眼睫状肌调节功能的降低,

晶状体逐渐变凸, 导致儿童视力下降。眼科专家建议看屏幕时要注意屏幕树立高度应与视线相平, 眼与屏幕的距离大于屏幕尺寸的五倍。

长期注视荧光屏会导致包括视疲劳、干眼、视力模糊、视力下降等眼部症状出现。眼科专家建议尽量不要让孩子玩电脑和手机超过45分钟。

2 设计方案(Designing scheme)

本课题所要解决的技术问题在于, 提供一种有效监视儿童屏幕距离、观看视频时间的方法, 让儿童在观看智能终端中的视频时必须在预设的距离之外、有效的时间内才能流畅观看屏幕, 一旦违反这些原则将出现音乐并黑屏提醒屏幕, 只要修

正,则可以继续返回观看。为了便于描述,本课题拟用“智能视频终端”代替智能终端设备中的视频客户端软件系统。

为了解决上述技术问题,本课题实施例提供了一种语音控制智能视频终端的交互方法,所述方法允许大人设定开始观看和结束观看的语音命令,以及设定可观看的时长和距离,在系统接收到大人输入的开始观看语音命令后便开始计时,一旦到达设定的观看时长,便黑屏并播放音乐提醒休息,直到大人进行语音干涉。在开启语音控制系统并输入开始观看语音命令后,将通过智能终端上的摄像头监视儿童的距离与预设距离相比较,一旦小于预设距离则黑屏并音乐提醒,修正后则恢复屏幕画面。在语音控制系统运行过程中,只允许用户进行语音控制屏幕,屏蔽手指触摸屏幕的消息。设计图如图1所示。

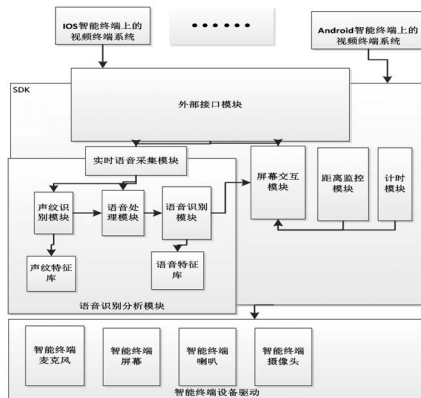


图1 视频智能终端设计图

Fig.1 Video intelligent terminal design drawing

2.1 语音控制智能视频终端的交互方法

语音控制智能视频终端的交互方法采用以下技术方案:

声纹识别模块:在本地保存1—2个大人、1—4个小孩的声纹特征,在开启语音控制系统时需要选择本次默认的一个大人和一个小孩的声纹特征作为本次观看过程中将要识别的声纹特征,不选择则默认上一次的设置。

语音识别分析模块:包括实时语音采集模块、语音处理模块、语音识别模块。

实时语音采集模块^[1]在开启语音控制后便开始在后台实时监听语音输入信号转换为数字信号,并且提取监听到的语音的声纹特征发送给声纹识别模块进行识别,声纹识别模块将所述声纹特征和本次选择的大人声纹特征及小孩声纹特征进行匹配,并将匹配结果返回给语音采集模块。

如果语音采集模块接收到的返回结果表明是预设小孩的声纹特征,则将该语音信号转换后的数字信号发送给语音处理模块,语音处理模块进行降噪处理消除环境音得到净音数字信号,并将其发送给语音识别模块。

语音识别模块^[2]将在本地语音特征库中匹配所述净音数字

信号,一旦匹配成功,便将所匹配的语音特征所对应的文字发送给屏幕交互模块,如果匹配失败则什么也不做。所述本地语音特征库为已经训练好的语音特征库,可以包含多个常用的语音特征(例如从01到09,1到100,A01到A99,B1到B9这些语音的标识符,以及ok、no、back、hello等特定命令语音的标识符),可以满足一般视频终端界面的导航按钮数量和每一屏显示的节目画面数量。

2.2 屏幕交互模块

屏幕交互模块包含点击区域识别模块。如果屏幕交互模块接收到语音识别模块发送过来的文字消息时,将判断是否为特定的命令文本,例如判断为“hello”文本,则在屏幕上显示预设透明度的蒙层页面,如图2所示,该蒙层页面按照点击区域识别模块识别出的每个矩形区域和文本序号显示每个矩形框及对应文本,同时让播放中的视频终端暂停,并且触发实时语音采集模块开始等待观看者发出语音指令。这时观看者可以根据想观看的界面画面上标识的文本(例如图2中“葫芦娃”节目对应的文本为“2”)来发出语音指令,实时语音采集模块把此语音指令转换后的数字信号发送给语音处理模块,语音处理模块进行降噪处理消除环境音得到净音数字信号,并将其发送给语音识别模块。

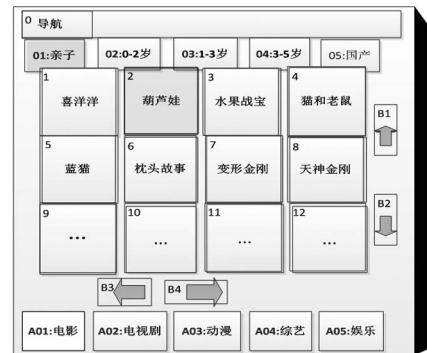


图2 蒙层页面

Fig.2 The covered pag

语音识别模块将在本地语音特征库中匹配所述净音数字信号,一旦匹配成功,便将所匹配的语音特征所对应的文字发送给屏幕交互模块,如果匹配失败则什么也不做。例如图2中,匹配出的命令文本为“2”,则屏幕交互模块在文本“2”对应的矩形区域中心点模拟点击消息发送给视频终端系统,进入到对应节目的播放画面。特别地,如果匹配出的命令文本为“B1”“B2”,则屏幕交互模块模拟滚动屏幕的消息发送给视频终端系统,其中点击区域识别模块识别出每个点击区域(矩形)的方法如下:

获取屏幕图像:获取当前屏幕图像,转换成像素矩阵数据到内存中^[3],假设该像素矩阵有N行M列数据。

获取背景色:根据智能视频终端通常在两边留空白的特点,在所述像素矩阵的中间行($N/2$)处获取第0列至第2列和第 $M-3$ 列至 $M-1$ 列的像素颜色值,如果这6列的颜色值相同,记录该颜色值为 $V1$,则继续在这6列上的第 $N/2+1$ 至第 $N/2+20$ 行上获取颜色值,如果每一行上这6列的颜色值都为 $V1$,则将颜色值 $V1$ 记录为背景色颜色值。如果以上过程中有颜色值不同,则根据智能视频终端软件通常以列表模式排版显示并且一行排3列或4列节目画面的特点,在所述像素矩阵的第 $M/3$ 、第 $M/2$ 列、第 $M*2/3$ 列处,从中间行($N/2$)开始向第 $N-1$ 行方向前进分别获取对应行上的颜色值,直到3列上的颜色值连续相等 $Y0$ 次(即这3列上连续有 $Y0$ 行的颜色值为同一个值),其中 $Y0$ 为一个预设值(例如5,通常列表的上下边距会大于的一个值),则结束背景色搜寻,并且记录连续相等 $Y0$ 次的颜色值 $V1$ 为背景色颜色值。

预划分搜索区域^[4,5]:然后将整个屏幕划分为5行3列的矩形区域,其中第3行的矩形高度占屏幕高度的 $7/10$,第1行的矩形高度占屏幕高度的 $(3/10)*(1/8)$,第2行、第4行的矩形高度均占屏幕高度的 $(3/10)*(2/8)$,第5行的矩形高度均占屏幕高度的 $(3/10)*(3/8)$,中间列的矩形宽度占屏幕宽度的 $9/10$,左右两列的矩形宽度分别占屏幕宽度的 $1/20$ 。给每个矩形区域(坐标位置)设定序号对应关系。例如:第2行从左到右的矩形区域设定的序号为01、02、03;第3行从左到右的矩形区域设定的序号为1、2、3;第4行从左到右的矩形区域设定的序号为B1、B2、B3;第5行从左到右的矩形区域设定的序号为A1、A2、A3。

识别可点击区域:从S303得到的4*3个矩形区域中(去掉第1行)开始按以下顺序搜索操作按钮区域(可点击区域)。

进行第4行矩形的(标记为B1、B2、B3的矩形)搜索方法为:当 $V1$ 为空时,直接将当前矩形区域划分成4行5列的小矩形区域,给每个小区域标记序号文本(从左到右、从上到下依次标记为B01至B20);当 $V1$ 不为空时,将第4行整个组合成一个大矩形R,从左到右、从中间行向上下两端渐进式进行搜索,搜索步骤为:

在当前被搜索的矩形区域R中,从该矩形区域的中间行 $midY$ 第1列开始向右搜索像素颜色值。

假设搜索完整行,所有颜色值都和 $V1$ 相同,则从第 $midY+midY/2$ 行第1列开始向右搜索像素颜色值。

假设搜索到第 $X1$ 列像素上获取的颜色值不等于背景色颜色值 $V1$,则以该点为出发点,分别向下、向上和向右查找像素的颜色值,直到和背景色颜色值连续 $Y0$ 次相同的两行($Y1$, $Y2$)和列($X2$),则记录下矩形的左顶点坐标为($X1$, $Y1$),右下角坐标为($X2$, $Y2$),为该矩形标记序号文本(序号以1为基础累加。

设置 $X3=X1$ 、 $X4=X2+Y0$ 、 $Y3=Y1-Y0$ 、 $Y4=Y2+Y0$,从 $X3$ 列上 $Y3$ 行处向上搜索,每向上一行则 $Y3=Y3-1$,假设搜索到第 $X3$ 列第 $Y3$ 行(重新赋值)像素上获取的颜色值不等于背景色颜色值 $V1$,此时计算两行之间的行间距为 $spaceY=Y1-Y3$;在 $Y2$ 行上 $X4$ 列处向右搜索,每向右一行则 $X4=X4+1$,假设搜索到第 $Y2$ 行第 $X4$ 列(重新赋值)像素上获取的颜色值不等于背景色颜色值 $V1$,此时计算两列之间的行间距为 $spaceX=X4-X2$ 。

在大矩形R中,根据每个将要标识的小矩形的宽度 $W=Y2-Y1$ 及高度 $H=X2-X1$,以及行间距 $spaceY$ 和列间距 $spaceX$,以及已经找到的一个矩形(序号为1)坐标位置($X1$, $Y1$, $X2$, $Y2$),可以以这个矩形1为基础分别向上、向下、向右辐射搜索,很快地标识其它相同大小的矩形框位置。

最后在大矩形R的最右边设置两个用于上下翻页的小矩形,假设R的最右边列为 $maxX$,中间行是 $midY$,则这两个矩形分别为 $B1(maxX-spaceX, midY-H, maxX, midY)$ 和 $B2(maxX-spaceX, midY+Y0, maxX, midY+Y0+H)$ 。

第2、3、5行矩形的搜索方法为:对每一行的矩形依次从左到右进行搜索,在当前被搜索的矩形区域中,获取该矩形区域的中间行第0列处的像素颜色值设置为背景色 $V1$,然后从中间行第1列开始向右搜索像素颜色值,假设搜索到第 $X1$ 列像素上获取的颜色值不等于背景色颜色值 $V1$,则以该点为出发点,分别向下、向上和向右查找像素的颜色值,直到和背景色颜色值连续 $Y0$ 次相同的两行($Y1$, $Y2$)和列($X2$),则记录下矩形的左顶点坐标为($X1$, $Y1$),右下角坐标为($X2$, $Y2$),为该矩形标记序号文本;再以中间行第 $X1+Y0$ 列为起始搜索列,向右边进行搜索,重复上述过程,直到找到最右边的一个矩形。

生成蒙层页面:在当前屏幕最顶端显示设定透明度的蒙层页,根据“识别可点击区域”识别出的各可点击区域的矩形区域坐标及其对应的序号文本,在蒙层页上显示这些矩形框及其序号文本。

2.3 距离监控模块

在开启语音控制时,将检测智能终端设备上是否具备摄像头装置,如果有则询问用户是否需要进行距离监控。如果用户选择了需要进行距离监控,则提示用户让观看者的脸在设定的距离处进行图像拍摄,拍摄完成后距离监控模块将对该参考图像中的人脸进行识别,取最小的人脸(有时可能识别到多个人脸)生成矩形区域O,将该矩形区域O的大小和范围记录下来,之后开启智能终端设备上的摄像头开始在后台实时录像。距离监控模块将在后台不停(假设1分钟一次)获取视频流中的图像,并对获取到的图像进行人脸识别,取最小的人脸生成矩形

(下转第22页)