

文章编号: 2095-2163(2021)02-0187-04

中图分类号: TP399

文献标志码: A

面向渐冻人设计的有意识地眨眼交互系统

杨俊, 高凡承, 王新龙, 宁磊

(南京工程学院 自动化学院, 南京 211167)

摘要: 渐冻症患者随着病情的发展会逐渐丧失语言功能, 无法通过语言表达内心需求, 造成较大的心理压力, 为了帮助渐冻人能够轻松地与外界交流, 本文为渐冻人设计了一种人机交互系统。由于患者会逐渐丧失行动能力, 本文设计了眨眼识别算法, 该识别算法能够识别出眨眼的面部动作, 并且区分接收到的眨眼信号是否为患者有意识地眨眼。交互系统配以图形界面, 图形界面内置了虚拟软键盘, 用户能够不通过敲击键盘实现打字输出。设计的交互系统以眨眼作为人机交互方式, 系统将接收到的眨眼信号传入图形界面, 图形界面通过虚拟软键盘实现输出功能。患者虽丧失了语言功能和行动能力, 但可以通过该交互系统实现眨眼打字, 舒缓心理压力。

关键词: 人脸对齐; 人脸特征点检测; 眨眼识别; 人机交互

Man-machine interface system for ALS patients

YANG Jun, GAO Fancheng, WANG Xinlong, NING Lei

(School of Automation, Nanjing Institute of Technology, Nanjing 211167, China)

[Abstract] In order to help ALS patients communicate with the outside world easily, this research designs a man-machine interaction system for ALS patients. Because the patient will lose the ability of action, this paper designs an eye blink recognition algorithm, which can recognize the facial movements of blinking and distinguish whether the received blink signal is conscious. The interactive system is equipped with a graphical interface, which has built-in virtual keyboard, users can type without hitting the keyboard. The interactive system takes blink as human-computer interaction mode, and the interactive system can receive the blink signal and send it to the graphic interface. Meanwhile, the graphic interface realizes expression function through virtual keyboard. Although the patient lost the language function and the action ability, the interactive system can help the patient relieve the psychological pressure.

[Key words] face alignment; face landmark detection; blink detection; man-machine interaction

0 引言

到2020年,全球已有超过一百万的“渐冻症”患者^[1]。研究可知,罹病患者由于神经损伤,丧失语言功能和行动能力,无法与外界交流以满足自身需求,因此承受巨大的心理压力和生理痛苦^[2]。

为了能让渐冻症患者可以自主地与外界交流,本文为渐冻症患者设计了一种人机交互系统,使用该系统患者将无需穿戴任何设备,只需面对摄像头,通过眨眼与计算机交互,并通过图形界面来完成打字输出,轻松地实现与外界交流。

1 交互方式的设计

由于渐冻症患者随着病情的加重会丧失语言和行动能力,所以无法通过控制手指敲击键盘与计算机交互。患者通常只能控制眼睛的运动,因此有面

向渐冻人设计的人机交互系统是通过识别眼球来展开研究,将人眼眼球的位置作为输入信号来与计算机进行交互^[3],但是这种交互方式存在一些不足。首先,计算机无法判断输入信号是否为患者有意识地交互,易造成误判。其次,人眼眼球运动幅度较小,不易被计算机辨别,增加了交互难度。并且,患者的视线长时间注视在显示器的同一位置,易造成视觉疲劳。

本次研究即以眨眼作为患者和计算机交互的方式,通过训练面部标志模型,设计眨眼检测算法和图形界面,来实现人机交互。同时,眨眼检测算法能够辨别患者是否为有意识地眨眼,避免了误判,提升了交互系统的稳定性,降低了操作难度。

基金项目: 江苏省大学生创新创业训练计划项目(202011276057Y)。

作者简介: 杨俊(2000-),男,本科生,主要研究方向:计算机视觉、人机交互。

通讯作者: 杨俊 Email:2399753110@qq.com

收稿日期: 2020-11-15

2 眨眼识别算法的研究

2.1 人脸特征点检测模型的选择

为了能够辨别患者是否眨眼,研究时需要有能够识别人眼关键点的深度学习模型,关键点包括左侧眼角关键点、右侧眼角关键点、左侧上眼睑关键点、右侧上眼睑关键点、左侧下眼睑关键点以及右侧下眼睑关键点^[4],如图1所示。过程中还要考虑到患者头部姿态的改变对识别精确度的影响,该模型也要能够克服头部姿态所带来的噪声,因此深度学习的神经网络模型采用深度对齐网络^[5](Deep Alignment Network)。深度对齐网络的输入为整张图片,充分利用了人脸的全局信息,使得检测人脸关键点时受头部姿态的影响较小。深度对齐网络每个阶段均有3个输入,分别是:被矫正过的图片、关键点热图和全连接层生成特征图。每个阶段有一个输出,是面部形状。每阶段的输出要经过连接层转换,得到下个阶段的3个输入。网络结构示意图如图2所示。

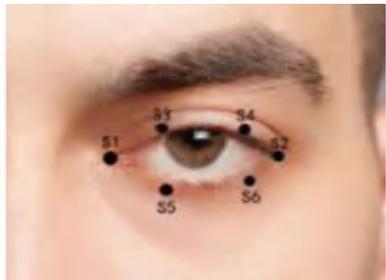


图1 眼部关键点示意图

Fig. 1 Schematic diagram of eyes key points

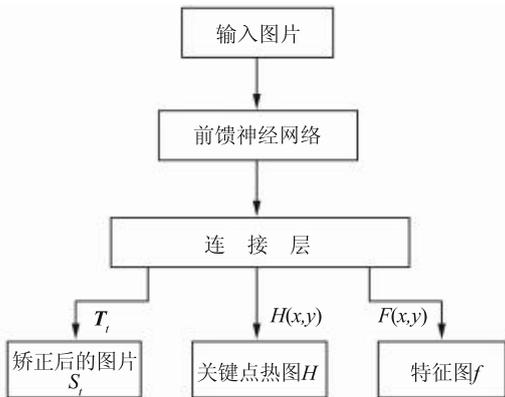


图2 神经网络结构示意图

Fig. 2 Schematic diagram of neural network structure

这里,输入图片 S_{i-1} 先通过相似变化矩阵 T_i 对图片进行矫正,将输入图片 S_{i-1} 变化得到 S_i ,变换公式如下:

$$S_i = T_i^{-1}(T_i S_{i-1} + \Delta S_i), \quad (1)$$

其中, ΔS_i 表示输入图片 S_{i-1} 与矫正完成图片 S_i 之间的偏移量。而在输入图3后,经过矫正处理即会得到图4。

再通过 $H(X,Y)$ 函数对矫正好的图片进行运算,得出关键点热图, $H(X,Y)$ 函数公式如下:

$$H(x,y) = \frac{1}{1 + \min_{S_i \in T_i(S_{i-1})} \|(x,y) - S_i\|}. \quad (2)$$

其中, S_i 为人脸关键点坐标。

分析 $H(x,y)$ 函数式可知,关键点处函数数值最大,距离关键点处越远则函数值越小,得到的关键点热图如图5所示。



图3 未矫正的输入图片

Fig. 3 Input image



图4 矫正后的图片

Fig. 4 Warped picture

同时,将图片输入全连接层,以线性整流函数作为激活函数,再进行上采样,得到用于输入的特征图。

经过多层网络迭代训练,最终得到能识别人脸面部关键点的深度学习模型。经由分析可知,深度对齐网络加入了关键点热图,神经网络能以整张图片作为输入,提取人脸的全局信息,所以深度学习模型能够较好地克服头部姿态导致的干扰,识别效果如图6所示。



图5 关键点热图

Fig. 5 Landmark heatmap

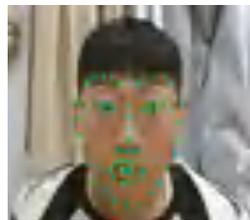


图6 识别结果图

Fig. 6 Image recognition result

2.2 眨眼检测算法的设计

在实现2.1节研发的基础上,本文就能得到人脸中眼睛的6个关键点坐标,如图7所示。

在患者闭眼时,关键点坐标会发生显著变化,如图8所示。

由于人脸与摄像头之间的距离会影响图片中2像素点之间坐标的差值^[6],因此通过计算出上下眼睑关键点距离之和与左右侧眼角关键点距离的比

值^[7], 来判断人眼的状态, 令函数 DR 如下式所示:

$$DR = \frac{\sqrt{(S_{3x} - S_{5x})^2 + (S_{3y} - S_{5y})^2} + \sqrt{(S_{4x} - S_{6x})^2 + (S_{4y} - S_{6y})^2}}{2\sqrt{(S_{1x} - S_{2x})^2 + (S_{1y} - S_{2y})^2}} \quad (3)$$

其中, x, y 分别表示关键点的横坐标和纵坐标。将眨一次眼过程中 DR 函数的值用图绘出, 如图 9 所示。

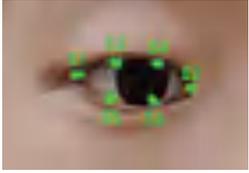


图 7 人眼关键点图

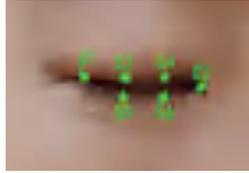


图 8 闭眼时关键点图

Fig. 7 Key points of eyes Fig. 8 Key points of closed eyes

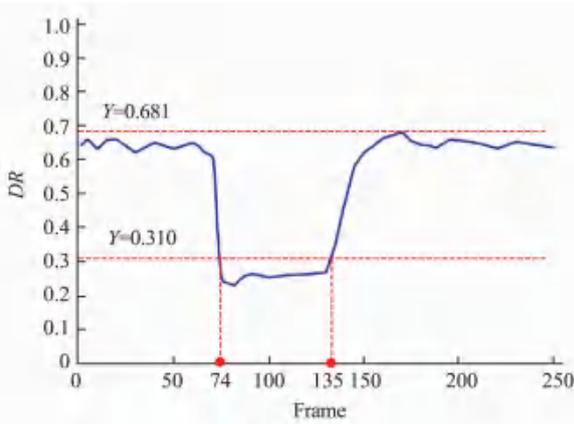


图 9 有意识地眨眼的 DR 曲线图
Fig. 9 Conscious blink results

分析图 9 中曲线可知, 在有意识地眨眼过程中, 睁眼时 DR 函数值稳定在区间 $[0.61, 0.69]$, 当闭眼时 DR 函数值稳定在区间 $[0.22, 0.27]$ 。因此可以通过人眼的 6 个关键点坐标计算 DR 函数值来判断是否眨眼, 当 DR 函数值从区间 $[0.61, 0.69]$ 下降到区间 $[0.22, 0.27]$ 、并再次回到区间 $[0.61, 0.69]$ 时, 认定为一次眨眼。

由于患者偶尔会无意识地眨眼, 易导致计算机误判, 为了避免这种情况的发生, 通过将图 9 与无意识地眨眼一次的 DR 曲线、即图 10 进行对比, 可知有无意识地眨眼最大的区别在于 DR 函数值处于区间 $[0.22, 0.27]$ 的时长, 也就是闭眼的时长。于是当 DR 值由区间 $[0.61, 0.69]$ 下降到区间 $[0.22, 0.27]$ 时, 计算机开始记录闭眼维持的帧数, 如果帧数低于 55 帧则认定为无意识地眨眼, 不会作为有效的交互信号, 否则如果 DR 值再次回到区间 $[0.61, 0.69]$, 就认定为一次有效的眨眼交互信号。

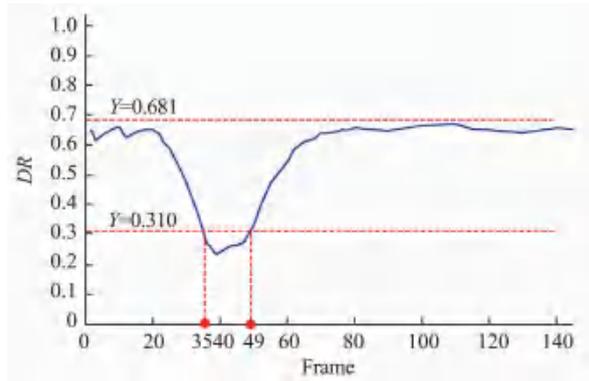


图 10 无意识地眨眼的 DR 曲线图
Fig. 10 Unconscious blink results

3 人机交互界面的设计

人机交互系统以眨眼作为人机交互方式, 图形界面均匀地排布着 26 个英文字母按键^[8], 作为虚拟软键盘的界面。并且加入了便捷按键, 能够直接在显示框中输出常用的交流短语, 例如: “I want to drink water”, “I am hungry” 等, 以提升交互界面的使用便捷度, 交互界面如图 11 所示。



图 11 图形界面
Fig. 11 Graphical interface

研究中为了降低患者使用该系统的复杂度, 减少眨眼次数, 图形交互界面会先循环选中每一行字母按键, 每行选中时停留 3 s 的时间, 若在停留时间内, 图形界面接受到眨眼输入信号, 则会退出行循环, 开始循环选中该行中的每一个字母按键, 每个按键选中时停留 3 s 的时间, 若在停留时间内, 图形界面接收到眨眼输入信号, 则输出框中会出现选中的字母, 并退出该行的字母按键循环。此时交互系统完成一次输出, 开始进行下一轮行循环, 等待眨眼信号的输入。

该图形界面自动循环选中按键, 不需要患者通过眨眼来操纵选中按键的位置, 极大程度上减少了眨眼操作的次数。并且图形界面先循环选中每行按
(下转第 194 页)