

基于 LSTM 的眼动行为识别及人机交互应用^①



黄君浩, 贺 辉

(北京师范大学珠海分校 信息技术学院, 珠海 519087)

通讯作者: 贺 辉, E-mail: hehui@bnuz.edu.cn

摘 要: 眼动交互在人机交互领域中有广泛的应用前景, 针对传统的眼动交互传感设备具有普遍侵入性, 校准过程复杂且价格昂贵, 普通单目摄像头传感器分辨率低等问题. 提出一种基于前置摄像头视频源, 使用方向梯度直方图 (HOG) 特征+SVM+LSTM 神经网络的眼动行为识别方法, 进而实现了简单的人机交互应用. 该方法首先定位和跟踪人脸, 在人脸对齐操作后依据 4 个眼角关键点的坐标获取双眼区域, 使用 SVM 模型判断眼睛的睁闭眼及非眨眼状态, 再分析相邻帧之间眼球中心的位置粗略判断眼动情况, 将可疑的有意眼势帧间差分视频序列输入到 LSTM 网络中进行预测, 输出眼动行为识别结果, 进而触发计算机命令完成交互. 经过自制数据样本集 20 000 个样本 (其中约 10% 为负样本) 测试, 本文方法动态眨眼识别准确率优于 95%, 眼动行为预测准确率为 99.3%.

关键词: 眼动识别; LSTM; 瞳孔定位; 人机交互

引用格式: 黄君浩, 贺辉. 基于 LSTM 的眼动行为识别及人机交互应用. 计算机系统应用, 2020, 29(3): 206-212. <http://www.c-s-a.org.cn/1003-3254/7388.html>

Eye Movement Recognition and Its Human-Computer Interaction Application Based on LSTM

HUANG Jun-Hao, HE Hui

(College of Information Technology, Beijing Normal University, Zhuhai, Zhuhai 519087, China)

Abstract: Eye-movement interaction has a broad application prospect in the field of human-computer interaction. Aiming at the problems of traditional eye-movement interaction sensors, such as universal intrusiveness, complex calibration process and high price, low resolution of common monocular camera sensors, etc., an eye movement recognition method based on front-facing camera video using directional gradient histogram (HOG) features + SVM + LSTM neural network, and a simple human-computer interaction application are proposed in this study. Firstly, the region of eyes are localized and tracked after face alignment. Secondly, the open-close and non-blinking state of the eyes is judged by the SVM model. Then, the position of eye center between adjacent frames is analyzed to roughly judge the eye movements, and the suspicious interframe difference video sequence of intentional eye position is obtained, which is the input of the LSTM network for prediction, and then trigger computer commands to complete the interaction. Through the self-made data sample set (about 10% of which are negative samples), the accuracy of dynamic blink recognition is better than 95%, and the accuracy of eye movement behavior prediction is 99.3%.

Key words: eye movement recognition; LSTM; pupil localization; human-computer interaction

引言

眼球追踪技术^[1]可实时捕获人眼的运动并获取注

视点, 由于人的眼球运动行为与大脑的活动密切相关^[2],

因此这种技术可辅助用户在计算机等相关设备上完成

① 基金项目: 广东省教育厅特色创新类 (自然科学) 项目 (2017KTSCX207)

Foundation item: Special Innovation (Natural Science) Project of Education Bureau, Guangdong Province (2017KTSCX207)

收稿时间: 2019-09-03; 修改时间: 2019-11-04; 采用时间: 2019-11-14; csa 在线出版时间: 2020-02-28

简单的交互工作^[3-5],并广泛应用于网页测试、广告评估、人机交互、布局评价等诸多领域^[6,2].近年来有学者提出将眼动交互应用于多通道交互^[7],为眼动人机交互技术开发了更广阔的应用前景.利用眼动实现人机交互的方式^[1]主要有驻留,平滑跟踪及眼势^[8](凝视、眼睛上、下、左、右的运动、眨眼等).目前大多数眼动交互的工作都使用眼动仪辅助完成,其工作原理为使用红外摄像头拍摄高清的瞳孔图像,采用基于瞳孔和角膜反射的视频记录法^[9]以准确定位人眼注视点.多数研究眼动交互的学者一般也会使用红外摄像头作为采集设备,但在日常生活中,红外摄像头不易获得,眼动仪大多造价昂贵且需要预先进行复杂的校准工作,都不适合作为日常眼动交互的采集设备.有部分学者针对普通的网络摄像头,提出了双摄像头估计视线的办法^[10],但易受到头部姿态和外界环境的影响,无法替代传统使用红外光源辅助的视线跟踪系统.也有学者提出识别眼动行为的方法,但只是估计几种不同使用情景下的眼动行为,并未提出将识别出的眼动行为应用在人机交互中的解决方案^[11,12].在眼动交互中基于驻留和平滑跟踪的眼动交互方式存在米斯达(Midas)接触问题^[4]且需要较高的定位精度,有研究认为将眼动行为应用在人机交互中需要使用额外的高清摄像机^[8,13].

针对以上结论,本文利用前置网络摄像头作为采集设备,提出一种使用方向梯度直方图(HOG)特征+SVM和LSTM神经网络的方法,实现眼势识别并应用于人机交互中.相较于驻留和平滑跟踪的眼动交互方式,基于眼势的交互方式对精度和米斯达接触问题更具鲁棒性.并且随着计算机的普及,越来越多的家庭或公司职员都拥有笔记本电脑.大多数笔记本电脑自带前置网络摄像头,为低成本的基于视觉的人机交互技术提供了环境.

1 关键算法

1.1 人眼定位跟踪及预处理

1.1.1 人眼定位跟踪及预处理

在视频流的第一帧,使用方向梯度直方图(Histogram of Oriented Gradient, HOG)特征检测人脸位置,若检测人脸成功,采用文献^[14]所提出的基于回归树方法的人脸对齐进行人脸关键点定位.人脸部分关键点的标注如图1所示.

依据定位的脸部关键点,计算双眼区域,如图2所

示,具体计算方法如下:

4个眼角坐标点对应的编号为37、40、43、46.假设4个眼角坐标点的坐标从左到右分别是 (x_1, y_1) , (x_2, y_2) , (x_3, y_3) , (x_4, y_4) ,首先由式(1)~式(2)计算出双眼的平均宽度 w 和双眼之间的间隔 $eyelen$.

$$w = (x_2 - x_1 + x_4 - x_3) / 2 \quad (1)$$

$$eyelen = (x_3 - x_2) / 2 \quad (2)$$

假设含有双眼的矩形区域的左上角坐标和右下角坐标分别为 (X_1, Y_1) , (X_2, Y_2) 计算 X_1, Y_1, X_2, Y_2 的公式如式(3)~式(6):

$$X_1 = x_1 - eyelen \quad (3)$$

$$Y_1 = y_1 - w \times 2/3 \quad (4)$$

$$X_2 = x_4 - eyelen \quad (5)$$

$$Y_2 = y_2 + w \times 2/3 \quad (6)$$

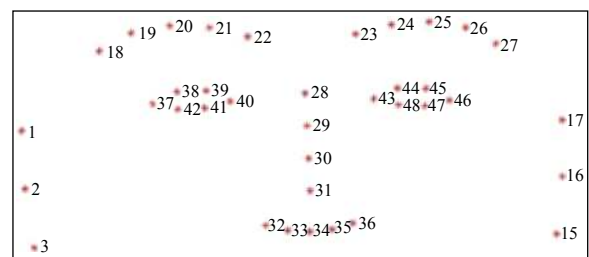


图1 人脸关键点编号

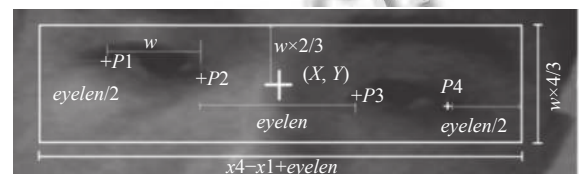


图2 双眼模板计算过程

1.1.2 人眼区域匹配跟踪

使用计算所得的双眼区域作为模板,在视频流的第二帧及随后的帧中使用模板匹配的办法来跟踪人眼.

模板匹配成功后,利用人脸三庭五眼的规律扩展出包含整个人脸的矩形区域,具体计算方法为:假设模板的宽 W 、高 H .以模板为中心向上扩宽 H 个像素,向下扩宽 $H \times 7/2$ 个像素,宽度不变,即可得到包含人脸区域的图像.重复进行人脸对齐关键点检测,以更新跟踪的双眼模板.

1.1.3 人眼区域处理

(1) 基于人脸关键点定位到的4个眼角坐标,按如

下计算方法得出单只眼睛的矩形区域。

以左眼为例, 首先计算眼睛的宽度 w , 两个眼角坐标的中心坐标 (x,y) , 以该坐标为中心, 左右扩展 $w \times 3/4$ 的像素区域, 上下扩展 $w \times 1/2$ 的像素区域, 即可计算出只包含单个眼睛的图像, 如图 3 所示。

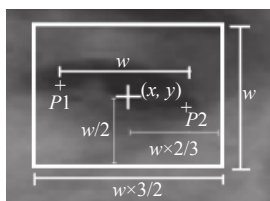


图 3 眼部区域裁剪过程

(2) 调整大小: 分别裁剪出双眼的图像后, 使用插值法调整成 36×36 像素大小。

(3) 对双眼图像进行图像处理: 使用边缘保留滤波去除噪声, 使用幂次变换提升区域整体对比度, 消除部分阴影, 结果如图 4 所示。

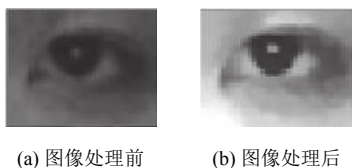


图 4 眼部图像处理结果

(4) 瞳孔光斑修复算法: 使用自适应阈值将眼睛图像分割成二值化图像, 瞳孔上的光斑与附近的对比度较高而形成孤立的区域, 其他大部分区域会联通。下一步标记联通区域, 将像素点总数在 14~30 的区域的像素点, 替换成该区域附近的像素值。瞳孔光斑修复结果如图 5 所示。



图 5 瞳孔光斑修复

1.2 眼动信息获取和行为识别

1.2.1 眨眼判断

(1) 粗略判断方法, 使用定位的人脸关键点计算人

眼长宽比, 区分睁闭眼状态, 计算方法: 在人脸关键点定位的基础上选取特征点编号为 37~48 的 12 个点. 按编号顺序设单眼周围的 6 个特征为 $P_1, P_2, P_3, P_4, P_5, P_6$, 如图 6 所示。

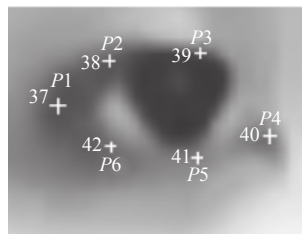


图 6 左眼特征点编号

使用式 (7) 可求出一只眼睛的长宽比, 其中分子表示眼睛的特征点在垂直方向上的距离, 分母表示眼睛的特征点在水平方向上的距离。

$$\frac{\|P_2 - P_6\| + \|P_3 - P_5\|}{2\|P_1 - P_4\|} \quad (7)$$

经过不同使用者测试实验, 当眼睛为完全睁开状态时, 长宽比平均在 0.30 左右, 眼睛完全闭合时, 长宽比平均在 0.19 左右. 因此以 0.25 为阈值, 当小于 0.25 时, 认为眼睛处于闭眼或半闭眼状态, 大于 0.25 时认为是睁眼状态。

(2) 精确判断方法, 使用 HOG+SVM 模型, 区分有意眨眼和无意眨眼: 判断为闭眼或半闭眼状态, 触发精确的动态眨眼判断。

具体判断方法为: 当连续 15 帧中, 有 10 帧的眼睛图像在提取完 HOG 特征, 并被 SVM 分类两只眼睛的状态分别为一睁一闭, 就判定其做出了一个单眨眼动作。

1.2.2 眼动轨迹获取

本文采用文献[15]的方法来定位眼球的中心, 以实时获取眼球的位置. 接下来计算眼动轨迹: 使用式 (8) 和式 (9) 计眼球中心坐标 (X_c, Y_c) 与眼睛图像长宽 36 像素的比值 t_x, t_y :

$$t_x = X_c/36 \quad (8)$$

$$t_y = Y_c/36 \quad (9)$$

计算出比值后, 比值作为纵坐标, 帧数作为横坐标, 可得到在发生眼动的 50 帧以内, 普通阅读扫视、向左眼动和向右眼动的完整轨迹, 如图 7 所示。

1.2.3 眼动行为识别

在获得左右眼动时瞳孔眼球中心的轨迹基础上,

根据在纵坐标上是否出现峰值来判断眼球是否可能正在进行一次眼势动作:

(1) 当眼球中心位置比值 t_x 小于 T_1 , 我们判断眼球做出了向左移动的眼势;

(2) 当眼球中心位置比值 t_x 大于 T_2 , 我们判断眼球做出了向右移动的眼势;

(3) 当眼球中心位置比值 t_y 大于 T_3 , 则判断眼球做出了向上移动的眼势.

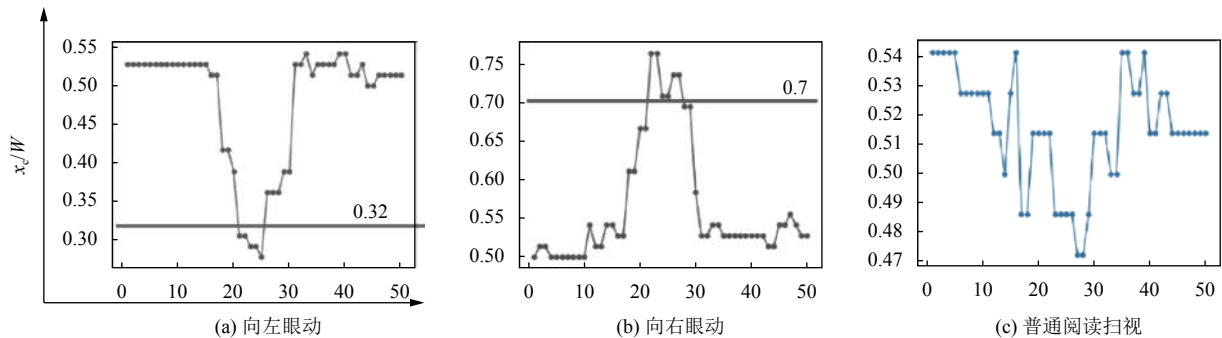


图7 眼球的横向运动轨迹

本文中 T_1, T_2, T_3 需根据实际不同的屏幕的大小, 推断出正常浏览时眼球的运动范围的范围, 调整出最佳的参数. 由图7可得, 本文在 27.2 英寸的屏幕正常浏览时的 t_x, t_y 的值集中在 0.45~0.55, 再经过实际体验, 本文选取的参数最佳为 $T_1=0.32, T_2=0.7, T_3=0.62$.

通过了以上 3 种判断的任一种, 说明在判断发生时眼球做出了一次眼动行为, 经过测试, 眼动行为的完成时间多在 1.5 秒以内, 因此将这 1.5 秒内的眼动视频截取并作为下一步输入. 本文在 30 帧/s 的摄像头采样速率下, 截取的眼动视频总帧数为 50 帧.

1.2.4 LSTM 行为识别

长短期记忆网络 (Long Short-Term Memory, LSTM) 的提出是为了解决循环神经网络 (Recurrent Neural Network, RNN) 的长依赖问题. LSTM 使用独特的输入门、输出门和遗忘门结构, 可以充分保留有效信息并丢弃不相关的信息, 将 CNN (Convolutional Neural Networks) 与 LSTM 结合起来, 可以应用于视频分类和视频描述^[16]. 应用在行为识别时, 为了去除背景光照等冗余信息, 可以提取原视频帧类似光流数据的运动信息作为输入^[17]. 本文将对原眼动视频帧进行帧间差分处理, 以提取眼动行为的运动特征, 消除由于人眼形状与光照不同而造成的误差. 考虑到识别的稳定性, 应当将两只眼睛的眼动视频序列一起作为输入. 识别的简要过程如图8所示.

本文将采取 LSTM 和 CNN 结合的网络模型来完成眼动行为的分类任务. 在 CNN 部分使用一个卷积层

和一个池化层, 卷积层使用两个卷积核, 尺寸为 4×4 , 步长为 1, 池化层的尺寸为 2×2 , 步长为 2, 还需在每个卷积层之后添加一个非线性激活函数 Relu. 在 LSTM 部分, 隐藏神经元个数与输入的帧长度相等, 为 50 个.

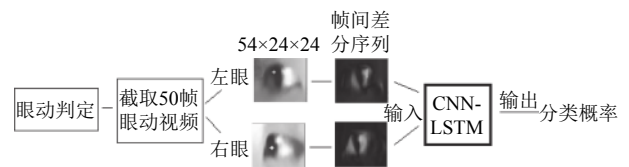


图8 基于 LSTM 的眼动行为识别流程

将两个眼动帧间差分视频同时作为输入, 在经过卷积层后, 将两张眼动视频提取的特征图通过 Concat 层进行特征融合操作, 特征维度是一个立方体, 使用 Flatten 层矢量化并输入进 LSTM, 最后经过 Softmax 分类器预测出每个眼动行为的概率.

本文采取的 LSTM 网络结构如图9所示.

2 实验及结果分析

2.1 数据采集和样本制作

数据采集环境如图10所示, 实验环境如下:

- (1) 摄像头: 罗技 C270i, 1280×720 像素 30 帧.
- (2) 显示器: 冠捷 AOC2701 Q27P1B (27.2 英寸).
- (3) 处理器: 英特尔 Core i7-8700 @ 3.20 GHz 六核.
- (4) 内存: 16 GB (金士顿 DDR4 2400 MHz).
- (5) 显卡: Nvidia GeForce GTX 1060 6 GB (6 GB/微星).
- (6) 操作系统: Windows 10 专业版 64 位 (DirectX 12).

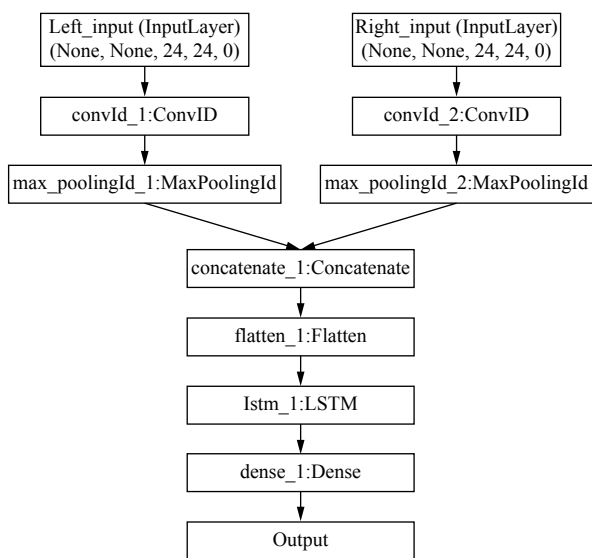


图9 LSTM网络结构图



图10 实验环境

使用 Python 语言并利用深度学习库 OpenCV, Keras, DLIB 进行实验. 被采集者距离屏幕的位置没有太多限制, 但应正对电脑屏幕, 并且摄像头能够拍摄到完整的人脸图像. 在正常均匀光照条件下, 使用上文提到的预处理方法采集训练数据.

2.2 眨眼数据集的训练

本文使用的训练数据集样本为长宽 36 像素的眼睛灰度图像. 采集总计 16 000 张单帧左右眼睁开图片. 其中睁开眼图片分别有 8000 张, 而 8000 张眼睛图像中左右眼各占 4000 张. 训练数据中的闭眼数据集包含半睁状态的人眼图片. 测试数据额外采集, 总计 4000 张单帧眼部图像.

2.3 眼势数据集的训练

单个样本为 50 帧长宽 36 像素的灰度视频, 采集 10 个不同的使用者总计 2000 张 6 分类的眼势视频, 加上双眼眨眼视频, 眼睛静止视频两种负样本各 100 张, 共有 2200 张 8 分类的训练数据. 额外录制 300 张 6 分类的眼动视频作为测试集.

由于采集到的样本数量有限, 本文使用如下方法扩充训练数据:

- (1) 调整帧长: 将一个完整的眼动视频帧序列, 增加或减少帧的长度.
- (2) 滑动窗口填充: 在数据帧的前或后加上长度为 10 以内的 0 填充帧, 填充后所有帧的长度一致.
- (3) 左右眼调换: 翻转左右眼, 运动方向也随之相反, 改变标签.
- (4) 加入随机噪声: 在图像中给不为零的像素点随机增减 5 以内的像素值.

经过上述步骤, 训练样本数量理论上可以扩充至 10 倍以上, 随后挑选正确率最高的组合训练模型来测试预测的准确率.

2.4 实验处理流程

本文定义了 6 种眼动行为, 分别是向左看、向右看、向上看、眨眼向左看、眨眼向右看和眨眼向上看. 并规定 2 种负样本, 分别是眼睛轻微眼跳或静止和普通无意眨眼.

实验总体技术流程如图 11 所示.

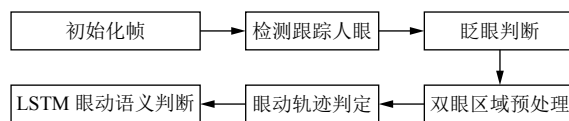


图11 眼动行为识别总体技术流程

其中眨眼判断时当识别到当前人眼状态为闭眼或半闭眼, 不管是有意眨眼还是无意眨眼, 都不会进入下一步眼动行为的判断, 结束本帧的处理, 并读取下一帧. 对于眼动语义预测, 由于存在米斯达接触问题, 因此眼球在向上、左和右运动的过程中, 视线都应尽量远离屏幕, 以区分有意眼动和无意眼动. 接下来将基于瞳孔中心轨迹的眼动行为识别到一个疑似的眼动行为的操作, 截取前后总计 50 帧的眼动视频序列, 输入到 LSTM 网络中进行最终的眼动行为语义判断. 需要说明的是在输入到 LSTM 网络前, 需要对截取的 50 帧眼动视频序列进行帧间差分处理, 以消除光照和人眼形状的影响.

2.5 实验结果及分析

在眨眼判断中, 训练时, 实际将图像调整成 24×24 像素, 一个 12 像素组成一个 cell, 一个 cell 构成一个 block, 其他参数设为默认, 使用线性核函数的

SVM 做分类. 将全体数据集样本训练一个轮次. 单张静态睁闭眼图片的分类在测试集中达到了 96.23% 的准确率. 在正常环境下当做出单眨眼行为, 单眨眼状态保持 0.3 s 以上, 动态识别准确率达到 95%.

在眼动语义判断中, 本文使用使用 Kreas 提供的接口, 除了上文提到的模型细节外其他参数都为默认. 训练时, 使用 adam 作为优化器, categorical_crossentropy 作为损失函数, 将 150 个样本作为一个批处理, 循环将所有样本训练 12 次. 经过各种扩充方法的组合训练, 最高在 50 帧的眼动视频在 LSTM 中的预测准确率达到 99.3%. 其中主要误差原因是基于瞳孔的移动轨迹来预测眼动行为会发生误报的情况, 因为眼中心定位有时会出错, 如图 12 所示.

经过判断, 识别错误的情况大多为无意眨眼时的

半闭眼状态瞳孔中心定位错误, 向右或向左眼动过程中中心定位到眼角从而做出了完全相反的预测.

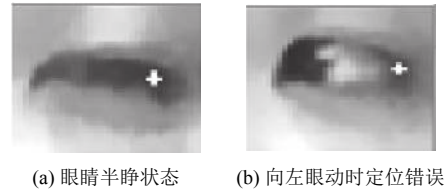


图 12 瞳孔中心定位错误情况

在为识别出的眼动行为的基础上我们分别添加滚动和翻页的命令, 应用于电脑阅读, 使得用户无需用手操作鼠标和键盘, 并且拥有良好的阅读体验.

以下选取了一些相似研究的文献与本文进行对比, 如表 1 所示.

表 1 各种眼动识别方法比较

来源	方法描述	识别种类	平均准确率 (%)	采集设备	平台	应用交互解决方案	区分眨眼
本文献	LSTM	6	99.3	网络摄像头	PC	是	是
文献[11]	时空上下文环境 STC	5	-	非红外	PC	否	否
文献[12]	虹膜跟踪算法	5	84.1	网络摄像头	PC	否	是
文献[8]	瞳孔跟踪	4	-	摄像机	PC	是	是
文献[13]	模糊推理	8	76.5	前置摄像头	移动端	是	否
文献[18]	模板匹配	3	96.8	非红外	-	否	否

从表 1 可以看出, 本文提出的基于 LSTM 的眼动行为识别方法兼顾识别数量、准确率与设备方便上的优势, 并拥有完整的眼动交互解决方案.

2.6 实际应用

本文一共尝试并实现了两种应用, 分别是网页阅读, 以及在 PC 上使用手机模拟器来模拟 APP 的使用.

2.6.1 网页阅读

如图 13 所示, 测试者正对电脑, 使用 4 种眼势 (向上、眨眼向上、向左、向右), 分别对应控制向下滚动、向上滚动、向左翻页和向右翻页. 测试表明: 在测试者双手占用的情况下, 可以使用眼势代替滚动和翻页等简单的操作, 并且在扫视浏览文字的过程中极少发生误触, 能够正常地阅读.

2.6.2 模拟的 APP 的使用

如图 14 所示, 使用者在电脑上运行手机模拟器, 模拟手机上的使用. 所有眼势均可对应一种操作. 分别为:

- 1) 左运动对应左翻页.
- 2) 右运动对应右翻页.
- 3) 上运动对应向下滚动.

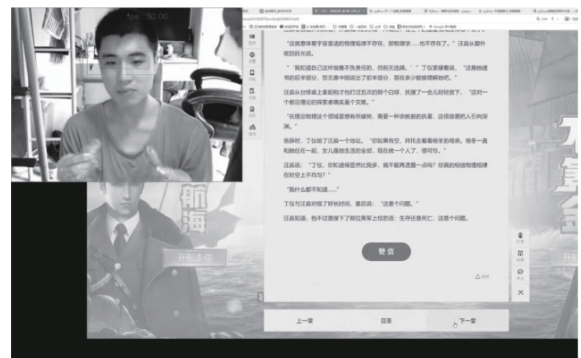


图 13 网页阅读应用



图 14 模拟手机 APP 应用

- 4) 左眨眼运动实现返回键.
- 5) 右眨眼运动则向下滚动.
- 6) 上眨眼运动触发大幅度向上滚动.
- 7) 左眨眼为模拟鼠标点击屏幕中央.
- 8) 右眨眼可轮流选择下方菜单模块.

手机上运行一款普通的新闻 APP, 经测试, 除在选择点击新闻时略不方便, 以及无法完成点击某个特定位置的需求, 大部分操作已可通过眼势完成.

2.7 使用体验

在上文提到的实验环境中测试, 程序运行时 CPU 占用率峰值不超过 10%, 内存占用 5%. 每秒帧数在 30, 目前的帧数处理上限由摄像头的采样频率的决定, 不会造成其他软件使用卡顿的情况. 此外, 使用 MacBook Air (2015 年初, 1.6 GHz 双核 i5) 笔记本搭建同样的环境运行, 每秒帧数也可以稳定在 25 左右, 可以达到实时性.

本文在录制不同测试者的数据集的同时, 收到了不同的用户反馈. 部分用户在初使用时, 存在眼势使用不熟练的情况. 经过训练, 如: 用某个物品模拟注视焦点, 让测试者的眼睛通过跟随焦点和移动焦点完成眼势, 最终都能达到比较好的效果.

3 结论与讨论

为实现简便实惠的眼动人机交互应用, 本文利用笔记本前置摄像头获取眼动信息, 提出一种使用人工提取特征分类+深度学习模型 LSTM 眼势预测的眼动行为识别方法, 并实现了基本人机交互操作.

其中人工提取特征分类的方法用来识别眼睛的静帧图像, 并将多个相邻帧之间的识别结果组合分析, 实现动态的眨眼识别过程; 采用 LSTM 网络的深度学习方法, 将 50 帧的眼动差分序列作为输入, 实现识别眼球向左右上运动等眼动行为. 将以上识别的结果分别触发相应的计算机命令, 实际应用在人机交互中. 实验结果验证在笔记本中基于前置摄像头视频源图像的眼动行为识别的可行性. 但本文方法虽然在测试集和训练集上都有良好的表现, 在实际应用时还要考虑存在部分遮挡、眼睛镜片反光、抖动模糊及眼球中心定位错误等情况下算法的完善和优化. 同时下一步将考虑加上视线追踪, 以期完成更为复杂的交互任务.

参考文献

- 1 陈庆荣, 周曦, 韩静, 等. 眼球追踪: 模式、技术和应用. 实验室研究与探索, 2012, 31(10): 10–15. [doi: 10.3969/j.issn.

- 1006-7167.2012.10.004]
- 2 王刚. 一种基于眼动轨迹的语义提取方法研究. 重庆师范大学学报(自然科学版), 2013, 30(1): 73–76.
- 3 朱麒文, 闫隆鑫, 张若兰, 等. 基于视线追踪的眼控鼠标设计. 电子器件, 2016, 39(2): 235–241. [doi: 10.3969/j.issn.1005-9490.2016.02.001]
- 4 李姗, 潘纲, 李石坚. 面向残障人士的基于眼动交互的绘画系统. 电子学报, 2011, 39(S1): 163–167.
- 5 董占勋, 许若楠, 顾振宇. 眼动交互在大屏手机操作中的应用趋势. 包装工程, 2015, 36(24): 57–60, 69.
- 6 刘青, 薛澄岐, Hoehn F. 基于眼动跟踪技术的界面可用性评估. 东南大学学报(自然科学版), 2010, 40(2): 331–334. [doi: 10.3969/j.issn.1001-0505.2010.02.022]
- 7 王增磊, 张树生, 白晓亮. 面向协同装配的混合现实手眼交互技术研究. 计算机应用与软件, 2019, 36(1): 251–258. [doi: 10.3969/j.issn.1000-386x.2019.01.045]
- 8 林敏. 基于眼动信息的人机交互关键技术研究[博士学位论文]. 上海: 上海大学, 2014.
- 9 王向军, 蔡方方, 刘峰, 等. 非接触动态实时视线跟踪技术. 计算机科学与探索, 2015, 9(3): 266–278.
- 10 温晴川, 达飞鹏, 方旭. 基于双目立体视觉的视线跟踪系统标定. 光学学报, 2012, 32(10): 1015002.
- 11 王际航. 基于视频图像的眼动系统算法研究与实现[硕士学位论文]. 长春: 吉林大学, 2016.
- 12 孟春宁, 白晋军, 张太宁, 等. 单摄像机下基于眼动分析的行为识别. 物理学报, 2013, 62(17): 174203. [doi: 10.7498/aps.62.174203]
- 13 Elleuch H, Wali A, Samet A, *et al.* A real-time eye gesture recognition system based on fuzzy inference system for mobile devices monitoring. Proceedings of the 17th International Conference on Advanced Concepts for Intelligent Vision Systems. Lecce, Italy. 2016. 172–180.
- 14 Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees. Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA. 2014. 1867–1874.
- 15 Rosch J L, Vogel-Walcutt J J. A review of eye-tracking applications as tools for training. Cognition, Technology & Work, 2013, 15(3): 313–327.
- 16 匡向阳, 苏学威. 基于 CNN 和 LSTM 混合模型的人体跌倒行为研究. 计算机应用研究, 2019, 36(12): 3857–3859, 3868. [doi: 10.19734/j.issn.1001-3695.2018.06.0424]
- 17 黄友文, 万超伦, 冯恒. 基于 CNN 与 LSTM 的多特征融合人体行为识别算法. 激光与光电子学进展. <http://kns.cnki.net/kcms/detail/31.1690.TN.20181113.1008.030.html>. [2019-07-25].
- 18 倪怡雯, 穆平安, 戴曙光. 快速眼动信息识别算法研究. 计算机应用研究, 2014, 31(3): 917–920. [doi: 10.3969/j.issn.1001-3695.2014.03.067]