

基于单目视觉的空中手写数据采集系统

屈喜文^{1,2}, 韩瑶妹¹, 胡冕军¹

1. 安徽工业大学计算机科学与技术学院, 安徽 马鞍山 243000;

2. 合肥综合性国家科学中心人工智能研究院, 安徽 合肥 230071

基金项目: 国家自然科学基金项目(61906003, 61971004); 安徽省高校协同创新项目(GXXT-2021-004)

通信作者: 屈喜文, qxw_ahut@ahut.edu.cn 收稿/录用/修回: 2023-03-09/2023-06-21/2023-08-08

摘要

针对基于视觉的空中手写系统使用的 3 维传感器体积庞大、价格昂贵的缺点, 提出了一种新型的基于单目视觉的空中手写数据采集系统, 该系统用体积小、价格便宜的单目摄像头代替 Kinect、Leap Motion 等 3 维传感器, 更适用于手机、空调、电视等家用设备集成和大规模推广使用。为了解决目前基于单目视觉的空中手写系统每次只能写入单个字符或短文本, 用户书写文本长度受到限制的问题, 提出了一种坐标系虚拟滑动技术。首先, 该系统使用背景差分法和 HSV (hue, saturation, value) 颜色空间阴影消除法从 2 维视频帧中获取完整的手部轮廓; 其次, 利用手的结构特点将手掌和手指进行分割, 利用构建的指尖模板从分割出的手指中进行匹配来确定指尖的位置; 最后, 连接连续视频帧中的指尖位置形成空中手写字符轨迹, 通过提出的坐标系虚拟滑动技术使用户可以进行连续书写从而形成文本。实验结果表明, 所提系统可以让用户自由地在空中连续书写, 对指尖的检测准确率达到了 96.8%。

关键词

空中手写

单目摄像头

数据采集

背景差分法

中图法分类号: TP391

文献标志码: A

In-air Handwriting Data Acquisition System Based on Monocular Vision

QU Xiwen^{1,2}, HAN Yaomei¹, HU Mianjun¹

1. School of Computer Science and Technology, Anhui University of Technology, Maanshan 243000, China;

2. Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230071, China

Abstract

Aiming at the shortcomings of the large and expensive three-dimensional (3D) sensor used in a vision-based in-air handwriting system, we propose a new type of in-air handwriting data acquisition system based on monocular vision. The system uses a small and cheap monocular camera to replace the 3D sensor such as Kinect and Leap Motion, which is more suitable for integration and large-scale promotion with household devices such as mobile phones, air conditioners, and televisions. To solve the problem that existing in-air handwriting systems based on monocular vision can only write a single character or short text each time, and the length of the user's written text is limited, we propose a coordinate system virtual sliding technology. First, the system uses the background difference method and the HSV color space shadow elimination method to obtain the complete hand contour from the two-dimensional video frame. Second, we separate the palm and finger

Keywords

in-air handwriting;

monocular camera;

data acquisition;

background difference

using the structural characteristics of the hand. Using the constructed fingertip template, we determine the position of the fingertip by matching the segmented fingers. Finally, we connect the fingertip positions in continuous video frames to form in-air handwritten character trajectories. Through the proposed coordinate system virtual sliding technology, users can write continuously to form texts. The experimental results show that the system can allow users to write continuously in the air freely, and the detection accuracy of fingertips reaches 96.8%.

0 引言

文字是人机交互的主要信息载体之一,传统的基于文字的人机交互系统是利用键盘、触摸屏、书写板等接触式设备进行信息交互。此类设备受到按键数量、屏幕大小等限制,文字输入方式复杂,不利于长文本输入,携带也不方便。人机交互的发展趋势是更加智能且对人更友好,即对人的约束越少越好。随着计算机软硬件技术的发展,空中手写交互技术因其书写方式更加自然,不受按键数量、空间大小等限制,没有摩擦力影响,更加人性化,吸引了越来越多研究者的兴趣^[1-4]。空中手写也有着广泛的应用前景,比如:指挥机器人工作;代替遥控器通过隔空书写电视频道、空调温度等切换电视台、调节温度等;用于军事隐蔽渗透过程中代替手势进行隔空书写交流;还可以用于课堂的趣味教学等。近年来,对空中手写系统研究已经取得了不少研究成果,大体分为基于穿戴/手持设备、视觉传感器(如 Leap Motion、Kinect、摄像头等)、WiFi 信号这 3 类系统。

基于穿戴/手持设备的系统^[5-6]在进行人机交互时,通过将传感器戴在手指上或者拿在手里进行隔空书写,相较于键盘、书写板等,这些系统允许用户使用熟悉的语言进行人机交互,而不必熟悉键盘输入法,也不受书写板大小的限制,但是仍然要依赖于设备的灵敏度,携带也不方便,另外书写时也受到连接线和手持设备大小的影响,无法灵活地进行空中书写,所以完全脱离穿戴/手持设备,使用裸手的书写系统更加方便实用。

基于 WiFi 信号的系系统^[7-10]和基于视觉的空中手写系统^[3-4, 11-12]允许用户裸手书写。基于 WiFi 信号的空中手写系统最大的优势是可以让用户在黑暗嘈杂环境中进行空中手写人机交互,不受光照、行人等的影响,但是此类系统过分依赖于 WiFi 环境,在 WiFi 信号不好或者没有信号时系统就失去了作用。

基于视觉的空中手写系统更符合人们日常交流

环境,目前基于视觉的空中手写系统主要使用的传感器是 Leap Motion、Kinect 等 3D 传感器^[3, 12-13],此类传感器具有深度信息,可以准确地定位指尖的空间位置,从而允许用户流畅地进行人机交互,由于空中手写字符/文本都是一笔完成,没有起笔和落笔信息,相对于传统手写字符/文本,空中手写识别难度较大,为此发展出了更适用于空中手写识别的算法^[14-15]。整个系统从数据获取到识别日趋成熟。但是此类传感器体积大(如图 1 所示),不适合与手机、电视、空调等家用电器集成,价格昂贵也不利于大规模推广使用。单个摄像头(如手机/笔记本等自带的摄像头)体积小,价格便宜,通用性好,更有利于集成推广,因此开发基于单目视觉的空中手写系统更具有应用价值。



图 1 Kinect 3 维传感器示例

Fig.1 Example for 3D sensor Kinect

不同于 3 维视觉传感器可以获得 3 维坐标,单个摄像头只能获得 2 维视频序列,从 2 维视频序列中稳定地跟踪、定位指尖是实现空中手写系统的难点和关键。在基于单目视觉的空中手写方面,MALIMA 等^[16]通过肤色统计分割手部区域,并利用质心距离法定位指尖,但是此方法不能检测出与背景颜色相近的肤色目标,并且对光照较为敏感。GURAV 等^[17]使用 AdaBoost (Adaptive Boosting) 结合类 Haar 特征对手势进行检测、识别,利用缺陷分析法定位检测指尖,但是该方法对光照干扰也较为敏感。ROY 等^[18]提出了一个单笔手写数字识别框架,通过将统一固定颜色的标记物放在摄像头之前进行书写,在书写过程中通过颜色分割技术将标记物从背景中分割出来,并将标记的尖端作为书写的笔尖进行书写,并通过训练的卷积神经网络(CNN)模型对手势进行识别,但是该系统在书写过程中用户需要手持固定颜色的标记物,书写起来非常不方

便, 在书写过程中会将与标记物颜色相近的目标检测为前景目标, 影响分割结果。MUKHERJEE 等^[19]提出了一种书写手检测算法, 使用 R-CNN (Region with CNN feature) 框架对书写手进行检测, 利用皮肤分割和背景减法进行手部分割, 使用距离加权曲率熵定位指尖, 使用深度卷积神经网络 (DCNN) 识别空中手写数据集中的 36 个字符 (26 个小写英文字母、10 个数字), 但是该系统主要适用于手指和手掌位于同一平面、手掌垂直或接近垂直握持的手势, 当书写手不完全处于视频帧中时, 用于指尖检测的手的几何特征也不再有效。YADAV 等^[20]通过语义手分割 (HandSNet) 模型对裸手进行分割, 应用点跟踪器和卡尔曼滤波器对裸手进行跟踪, 通过移动的裸手的质心位置形成书写字符的粗略轨迹, 最终使用 4 个独立的 DCNN 模型识别手势字符, 在语义分割手过程中需要对多张手部图片标记并进行训练, 此方法比较耗费时间。

在日常实际应用中, 用户往往是连续书写文本, 而不是单个字符输入, 目前基于单目视觉的空中手写系统每次只能写入大小写字母、数字 0~9 等单个字符或两个字符, 用户书写时受到了文本行长度的限制。针对上述基于视觉的空中手写系统受到设备体积大小、价格昂贵和输入字符长度限制的情况, 提出了一种基于单目视觉的空中手写数据采集系统, 通过使用体积小、价格便宜、通用性好的单个摄像头进行数据采集。采集过程中利用手部运动信息对书写手进行检测, 结合手的几何结构特点对手指和手掌进行分割, 然后通过对手指部分进行指尖模板匹配从而对指尖进行定位, 记录连续视频帧中的指尖位置形成字符轨迹。由于单个摄像头的感受范围有限, 书写者书写时总是在摄像头前一个相对固定的范围内书写, 当书写文本时会产生大量的字符坐标重叠, 造成文本剧烈形变, 不利于长文本的采集, 因此提出一种坐标系虚拟滑动技术, 即实际书写范围保持不变, 按照一定的速度使坐标系逐步累加, 从而避免了文本中字符重叠, 解决了用户书写时受到文本行长度限制的问题。最终将形成的文本送入高精度分类器进行识别。本文的主要贡献有:

1) 提出了基于单目视觉的指尖检测与定位算法, 该算法对环境干扰有较好的鲁棒性, 能够稳定、准确地检测和定位指尖;

2) 基于指尖检测与定位算法开发了一个基于单目视觉的空中手写字符采集系统;

3) 提出了坐标系虚拟滑动技术, 借助该技术, 空中手写系统允许用户输入长文本, 从而更贴近实用;

4) 通过手势控制书写开始与结束。

1 空中手写数据采集系统整体设计

基于单目视觉的空中手写数据采集系统的总体设计如图 2 所示。系统首先通过提出的指尖检测与定位算法检测书写指尖, 通过判断指尖的数量 N 确定书写是否开始, 以伸出一个手指的手势表示开始书写, 即 $N=1$ 时开始书写。开始书写后以一定的采样频率从视频图像中检测和定位指尖, 连接检测到的指尖位置构成空中手写字符。在实际应用中, 用户往往不是单个字符输入, 而是连续书写文本, 为了使系统更贴近实际应用, 提出了一种坐标系虚拟滑动技术, 通过该技术构造空中手写文本, 书写结束与否仍然是通过手势确定, 以手掌完全张开表示书写结束, 即指尖数量 $N=5$ 时书写结束, 最后将文本送入高精度卷积神经网络进行识别。指尖检测与定位算法流程如图 2 所示, 具体地, 系统从单目视觉的 2 维视频图像序列中检测和定位指尖, 首先通过平均背景差分法检测书写手, 由于光照射到

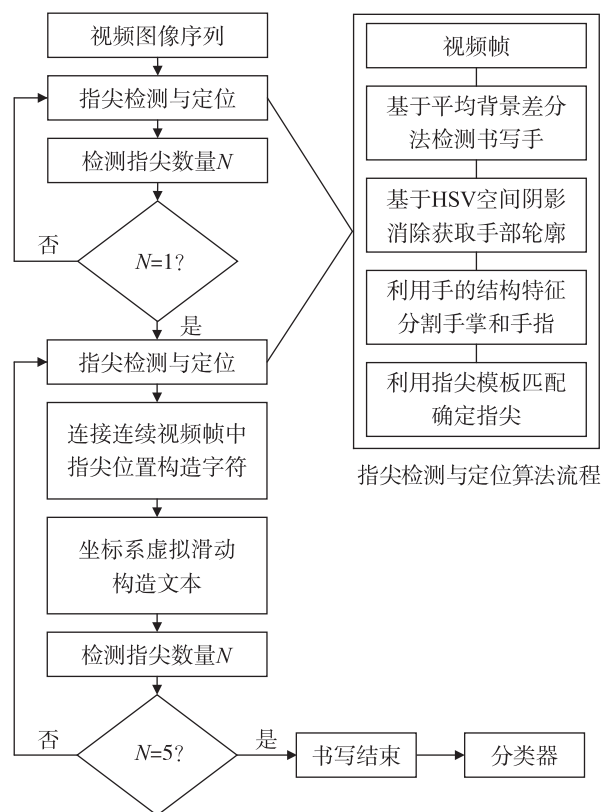


图2 基于单目视觉的空中手写数据采集系统框图

Fig.2 System diagram of in-air handwriting data acquisition system based on monocular vision

手部会产生阴影,影响检测精度,通过使用 HSV 空间阴影去除法从检测到的手部区域中去除光照阴影以获得清晰完整的手部轮廓,然后利用手的几何结构特点将手掌和手指进行分割,利用制作好的指尖模板去匹配图像中的手指部分,找到匹配度最高的区域确定为指尖。

1.1 基于背景差分法和 HSV 空间阴影消除的手部检测算法

颜色特征^[21]和运动信息在运动目标检测和跟踪中常常用到。利用基于颜色的运动目标检测方法虽然可以提取较为完整的手部轮廓,但是这种方法对类肤色背景和光照等干扰比较敏感,对于空中手写而言,要形成稳定的文字轨迹,需要系统有较强的抗干扰能力。基于运动信息的检测方法有:帧差法^[22]、背景差分法^[23]、光流法^[24]等。帧差法简单易行,但很难获得完整的运动目标轮廓,容易产生空洞和双影现象。光流法的检测效果好但是算法复杂,计算量较大所以很难实现实时处理。背景差分法作为一种有效的运动目标检测方法,其算法简单、易于实现并且可以较好提取出运动目标的轮廓,但对场景的变化较为敏感。针对上述检测方法的不足之处,将平均背景差分法和 HSV 空间阴影检测和消除^[25]进行结合,可以很好地消除光照和场景变化对运动目标检测的影响,提取出较为完整的手部轮廓。

1.1.1 基于平均背景差分的人手检测

由于提出的空中手写数据采集系统是在场景相对简单的室内进行,用户利用空中手写系统书写时,一般位于摄像头前,身体与摄像头位置相对固定,最常见的干扰是光照变化、光照角度造成的阴影、偶尔走过的行人等。因此本文采用简单且快速的平均背景模型^[26]分割前景运动目标,为了适应场景中光照和背景环境的变化,在运动目标检测过程中对背景进行实时更新并采用 HSV 颜色空间阴影消除算法消除场景中光照角度的变化带来的阴影从而可以更好对前景运动目标进行检测和分割。首先在没有进行书写时,也就是手部没有进入摄像头书写区域之前,对摄像头采集到的前 n 帧图像求均值作为初始背景 B_0 ,即:

$$B_0(x, y) = \frac{1}{n} \sum_{i=1}^n f_i(x, y) \quad (1)$$

其中, $f_i(x, y)$ 表示第 i 帧图像在 (x, y) 处的像素值, $B_0(x, y)$ 表示 B_0 在 (x, y) 处的像素值。利用背景差分获取前景运动目标过程中,由于背景会因

为光照等外部因素的干扰而产生变化,当外界的干扰因素发生比较频繁,累积到一定程度时,会影响运动前景目标的分割效果,所以需要对背景图像进行实时更新,背景更新如下:

$$B_k(x, y) = (1 - \alpha)B_{k-1}(x, y) + \alpha f_k(x, y) \quad (2)$$

其中, $B_k(x, y)$ 为第 k 次采样时的背景图像; $B_{k-1}(x, y)$ 为 $k-1$ 次采样时的背景图像; α 为更新权重。随着 α 的增大,背景图像中当前帧的比重相应增大,也意味着背景图像的变化更加明显。当 α 的取值太小时,背景更新速率稳定,但无法及时更新,满足不了在背景差分过程中光照等因素的变化;当 α 取值较大时,场景可以迅速变化,噪声也会混入,前景目标会被更新到背景中从而影响检测效果。如图 3 所示,通过设置不同 α 值下前景目标检测正确率以获得合适的 α 值,前景目标检测正确率计算方法为

$$T = N_c / N_m \quad (3)$$

其中, N_c 是正确检测出来的前景目标数量, N_m 是人工标注出来的前景目标数量, T 代表前景目标检测的正确率。计算 80 帧图像的前景目标检测的正确率,由图 3 中可以看出当 α 取 0.05 时效果较好。

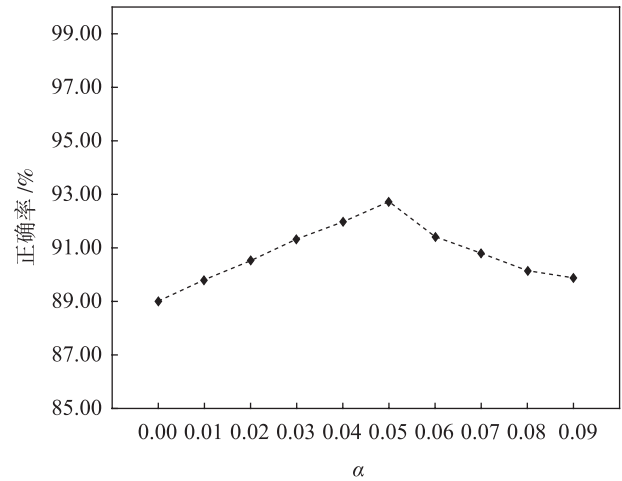


图 3 不同调节因子 α 下前景目标检测的正确率

Fig.3 Accuracy of foreground target detection under different adjustment factors α

经过背景更新后,对第 k 次采样时刻的视频帧 $F_k(x, y)$ 进行背景差分,差分图像 $D_k(x, y)$ 为

$$D_k(x, y) = |F_k(x, y) - B_k(x, y)| \quad (4)$$

通过设定合适的阈值 t 来判断像素点是背景区域还是前景运动目标,当差分图像中 (x, y) 处的像素值 $D_k(x, y)$ 大于等于设定的阈值 t ,则该像素点属于运动区域并将该像素点的值置 1,当 $D_k(x, y)$

小于设定的阈值 t 时该像素点属于背景区域并将该像素点的值置 0, 这样就得到了第 k 次采样时刻运动手势区域的二值化图像 R_k 的像素值:

$$R_k(x, y) = \begin{cases} 1, & D_k(x, y) \geq t \\ 0, & D_k(x, y) < t \end{cases} \quad (5)$$

如图 4 所示, 通过计算不同阈值下 80 帧图像的前景目标检测的正确率来选取合适的阈值 t , 从图 4 中可以看出当 $t = 15$ 时检测正确率最高。

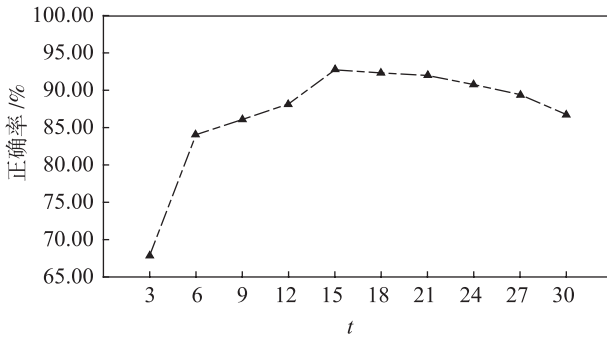


图 4 不同调节因子 t 下前景目标检测的正确率

Fig.4 Accuracy of foreground target detection under different adjustment factor t

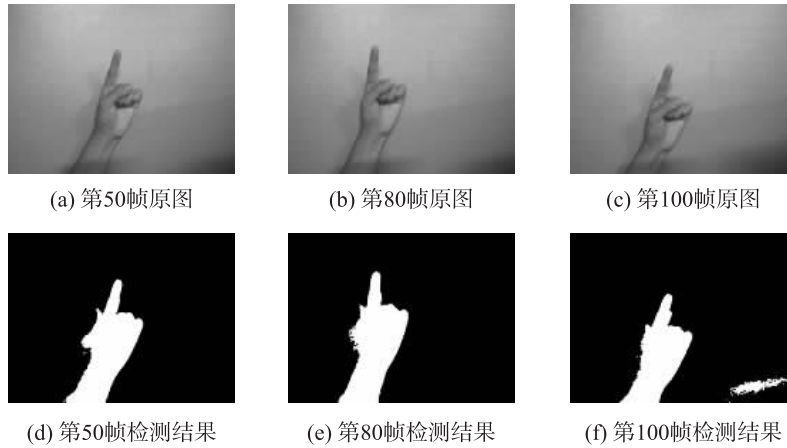


图 5 手部检测示例

Fig.5 Hand detection example

接近光谱色的程度, 取值范围是 $0 \sim 1$, 亮度 V 由物体的反射系数决定, 表示颜色明暗的程度, 通常取值范围为 0 (黑) 到 1 (白)。与 RGB 颜色空间相比, HSV 颜色空间中色彩的表示更接近于人的色觉感受。因此采用 HSV 阴影检测和消除算法消除运动目标的阴影, 具体地通过将阴影区域和背景区域的像素在 H (色度)、 S (饱和度)、 V (亮度) 三个分量上进行对比, 选择合适的阈值来判断是阴影还是运动区域。RGB 到 HSV 颜色空间的映射关系为^[27]

使用上述平均背景差分模型对视频序列中的书写手的检测效果如图 5 所示。从图 5 可以看出, 当光照射到手上形成阴影后, 随着手的移动, 阴影也在移动, 因此第 100 帧的检测结果中有阴影部分被当做运动前景目标检测了出来。

1.1.2 基于 HSV 颜色空间的阴影消除

在背景差分的过程中由于光照角度的影响, 容易将运动目标遮挡产生的阴影检测为目标, 从而会影响运动目标的检测效果, 早在 1824 年 HELMHOLTZ 等便提出了三原色学说, 由于人类的视网膜中存在 3 种锥体细胞可以分别感知 3 基色 R (红)、 G (绿)、 B (蓝), RGB 色彩空间中 3 个通道 R (红)、 G (绿)、 B (蓝) 的取值范围都在 $0 \sim 255$ 之间, 虽然在现实生活中人们见到的颜色大多数是由三基色混合而成的色彩, 但是人眼直观感受到的是一张图片什么颜色, 深浅如何, 明暗如何? 1978 年, SMITH 等提出了 HSV 颜色模型, 模型参数分别为 H (色度)、 S (饱和度)、 V (亮度)。色度 H 表示色彩信息, 所处的光谱颜色的位置, 取值范围是 $0^\circ \sim 360^\circ$, 饱和度 S 指颜色的鲜明程度, 表示颜色

$$V = \frac{\max(R, G, B)}{255} \quad (6)$$

$$S = \frac{\max(R, G, B) - \min(R, G, B)}{\max(R, G, B)} \quad (7)$$

$$H_1 = \arccos \left(\frac{((R - G) + (R + G))}{2 \sqrt{(R - G)^2 + (R - B)(R - G)}} \right) \quad (8)$$

$$H = \begin{cases} H_1, & B \leq G \\ 360 - H_1, & \text{其他} \end{cases} \quad (9)$$

式中, R 、 G 、 B 分别表示转换前图像在 RGB 颜色空

间中 3 个通道的值, H 、 S 、 V 分别表示图像经过转换后在 HSV 颜色空间中 3 个通道的值。

由于光线影响产生的阴影区域的色度和饱和度变化很小, 但是亮度相对较暗, 因此根据这一特性对运动目标区域进行阴影检测, 检测算法如式 (10) 所示。式 (10) 中, $B_H(x, y)$ 、 $B_S(x, y)$ 、 $B_V(x, y)$ 分别表示背景图像像素在 HSV 颜色空间的 3 个分量; $F_H(x, y)$ 、 $F_S(x, y)$ 、 $F_V(x, y)$ 分别表示当前帧图像像素在 HSV 颜色空间的 3 个分量; $S(x, y)$ 表示阴影检测结果:

$$S(x, y) = \begin{cases} 1, & \text{if } T_{sv} \leq \frac{F_V(x, y)}{B_V(x, y)} \leq T_{bv} \text{ and} \\ & |F_S(x, y) - B_S(x, y)| \leq T_s \text{ and} \\ & |F_H(x, y) - B_H(x, y)| \leq T_h \\ 0, & \text{otherwise} \end{cases} \quad (10)$$

当前帧中像素点是阴影时, $S(x, y)$ 为 1, 否则为 0; T_{sv} 代表光照因子, 光照越强 T_{sv} 越小, 因为阴影区域的亮度一般小于非阴影区, 亮度参数 T_{sv} 、 T_{bv} 的取值范围为 $0 \leq T_{sv} \leq T_{bv} \leq 1$; T_s 、 T_h 分别表示饱和度和色度的阈值。通过计算 10 组不同参数下阴影消除后的前景目标检测正确率, 来获得合适的 T_{sv} 、 T_{bv} 、 T_s 、 T_h , 选择检测正确率最高的参数组中的参数值作为 HSV 阴影检测算法中的阈值。不同参数组在 80 帧图像的前景目标检测的平均正确率如图 6 所示, 由图 6 可以看出第 7 组参数值下, 即 $T_{sv} = 0.7$, $T_{bv} = 0.85$, $T_h = 0.1$, $T_s = 0.25$ 时前景目标检测的正确率最高。 T_{sv} 、 T_{bv} 、 T_s 、 T_h 参数组选择:

$$T_{sv} = 0.9, T_{bv} = 0.95, T_h = 0.1, T_s = 0.25$$

$$T_{sv} = 0.8, T_{bv} = 0.95, T_h = 0.1, T_s = 0.25$$

$$T_{sv} = 0.7, T_{bv} = 0.95, T_h = 0.1, T_s = 0.25$$

$$T_{sv} = 0.6, T_{bv} = 0.95, T_h = 0.1, T_s = 0.25$$

$$T_{sv} = 0.5, T_{bv} = 0.95, T_h = 0.1, T_s = 0.25$$

$$T_{sv} = 0.7, T_{bv} = 0.90, T_h = 0.1, T_s = 0.25$$

$$T_{sv} = 0.7, T_{bv} = 0.85, T_h = 0.1, T_s = 0.25$$

$$T_{sv} = 0.7, T_{bv} = 0.95, T_h = 0.3, T_s = 0.25$$

$$T_{sv} = 0.7, T_{bv} = 0.95, T_h = 0.5, T_s = 0.25$$

$$T_{sv} = 0.7, T_{bv} = 0.95, T_h = 0.3, T_s = 0.4$$

通过对背景差分得到的运动区域进行 HSV 阴影检测, 将运动目标的阴影和干扰阴影从运动区域中消除从而得到更为完整、清晰的前景运动目标。通过 HSV 阴影消除方法对视频帧进行实验, 实验结果如图 7 所示。从图 7 可以看出, HSV 阴影消除

方法可以很好地去除阴影获得较为完整清晰的书写手和手臂。

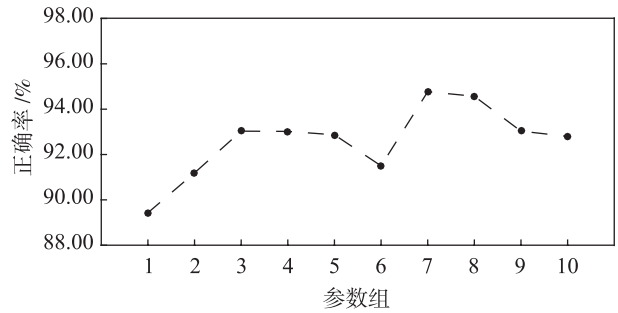


图 6 不同参数组下前景目标检测的正确率

Fig.6 Accuracy of foreground target detection under different parameter groups

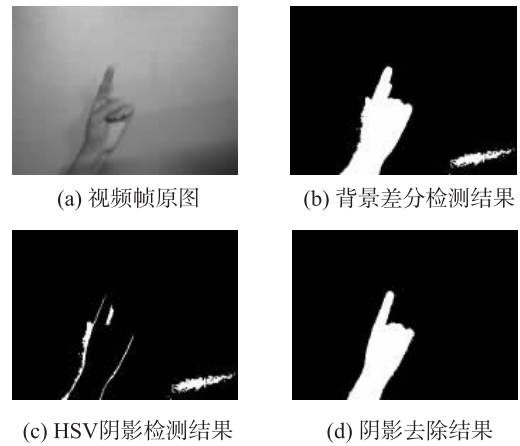


图 7 基于 HSV 颜色空间的阴影去除效果

Fig.7 Shadow removal result based on the HSV color space

1.2 书写手分割

1.2.1 手部与手臂分割

由于背景差分得到的是所有运动目标区域, 书写者在空中手写时胳膊也在运动, 多种运动目标的几何结构各不相同, 结合手的几何结构特征对书写者的手和手臂进行分离^[28], 从 1.1 节所述方法中可以获得较为完整二值化运动目标图像 (如图 7(d) 所示), 根据其轮廓拓扑结构可以得到图 8 所示的书写手轮廓。选择书写手轮廓在 x 轴和 y 轴中坐标值最大和最小的 4 个候选点 $A(x_A, y_A)$ 、 $B(x_B, y_B)$ 、 $F(x_F, y_F)$ 、 $E(x_E, y_E)$, 连接 4 个候选点从而构成轮廓的最小外接矩形。为了得到轮廓区域的最大内切圆, 首先画出如图 8 所示的书写手轮廓, 然后计算书写手轮廓区域内每一个像素点到最近的轮廓线上像素点的欧氏距离, 进行距离统计, 寻找所有欧氏距离中的最大值作为书写手轮廓的最大内切圆的

半径, 书写手轮廓区域内相对应的像素点的位置作为最大内切圆圆心 $m(x_m, y_m)$ 的坐标。

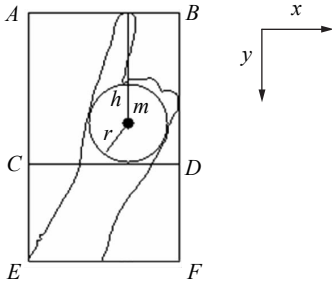


图8 手部和手臂结构图

Fig.8 Structure diagram of hands and arms

由上文已知 A 和 B 两点的坐标, 则 A 和 B 两点之间构成的直线方程如式(11)所示, 化简之后得到的直线方程如式(12)所示, 点到直线的距离公式如式(13)所示, 其中 $Ax + By + C = 0$ 为直线方程, L 为坐标点 (x_0, y_0) 到直线的距离。

$$\frac{y - y_A}{y_B - y_A} = \frac{x - x_A}{x_B - x_A} \quad (11)$$

$$(y_A - y_B)x + (x_B - x_A)y + x_A y_B - x_B y_A = 0 \quad (12)$$

$$L = \left| \frac{Ax_0 + By_0 + C}{\sqrt{A^2 + B^2}} \right| \quad (13)$$

如式(14)所示, 通过式(12)中 AB 构成的直线方程和式(13)中点到直线的距离公式可以计算得到最大内切圆圆心 m 的坐标点 (x_m, y_m) 到 AB 构成的直线的距离 h , 手掌的长度 d 可以近似由式(15)得到, 如图8所示由于坐标原点置于2维视频图像的左上角, 从而 C 和 D 两点的坐标可以近似为 $(x_A, y_A + d)$ 和 $(x_B, y_B + d)$ 。

$$h = \left| \frac{(y_A - y_B)x_m + (x_B - x_A)y_m + x_A y_B - x_B y_A}{\sqrt{(y_A - y_B)^2 + (x_B - x_A)^2}} \right| \quad (14)$$

$$d = h + r \quad (15)$$

将矩形 $ABDC$ 以外的像素点的值全部置0, 便可得到只有手部的二值化图像, 手部分割结果如图9所示。



(a) 二值化图像

(b) 分割后结果

图9 手掌和手臂分割结果

Fig.9 Palm and arm segmentation results

1.2.2 手掌和手指分割

上述算法得到的图9(b)所示手部的运动轨迹覆盖范围太大, 不适合构建空中手写字符, 由于最终目的是要检测和跟踪指尖, 因此需要进一步从图9(b)所示结果中分割出手指。将图9(b)分割后得到手部二值化图像以 m 为圆心, 如图10所示, 将最大内切圆的半径扩大至 $1.5r$ 在二值化图像上画圆, 使手部轮廓和圆相交并将内切圆内部像素点置0从而获得只剩手指的二值化图像。手掌和手指的分割结果如图11所示。如图2所示, 通过统计只有手指轮廓的二值化图像中像素值为255且位置相邻的像素组成的区域即连通区域的(图11(b))的个数 N 来控制书写开始与结束。

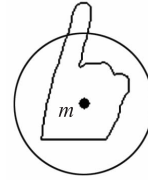


图10 手掌和手指的分割结构图

Fig.10 Segmentation structure diagram of the palm and finger



(a) 手部二值化图像

(b) 手部和手掌分割结果

图11 手掌和手指的分割结果

Fig.11 Results for palm and finger segmentation

1.3 指尖定位

图11(b)中分割得到的手指目标仍然太大, 为精准地定位指尖, 预先制作了指尖模板, 通过指尖模板对图11(b)中二值图像进行指尖模板匹配, 从而找到指尖的位置。在制作指尖模板时首先将图12所示指尖模板顺时针和逆时针每隔 2° 旋转一次, 各旋转8次从而获得17张覆盖多个角度的指尖模板, 可以较好地适应运动中手指的旋转带来的变化。

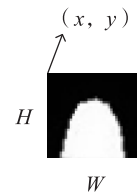


图12 指尖模板

Fig.12 Fingertip template

通过将指尖模板图像在图 11(b) 所示分割后的手指二值化图像的左上角坐标原点开始按照从左到右、从上到下的顺序逐像素进行滑动, 计算每次滑动时指尖模板图像中每个像素点的像素值与该时刻手指二值化图像中相对应位置像素点的像素值的差值平方和作为匹配值, 模板匹配中的差值平方和可由式(16)计算得到。

$$M(x, y) = \sum_{x', y'} (T(x', y') - I(x + x', y + y'))^2 \quad (16)$$

其中, $T(x', y')$ 为指尖模板图像中点 (x', y') 处的像素值, $I(x + x', y + y')$ 为指尖模板滑动到二值化图像中 (x, y) 处时二值化图像与指尖模板图像的重叠区域相对应像素点的像素值, $M(x, y)$ 为指尖模板滑动到 (x, y) 处的匹配值。通过将图像和模板的各个像素进行匹配, $M(x, y)$ 越小匹配效果越好。

为了适应在书写过程中手指旋转带来的变化, 通过将上文中的若干个不同角度的指尖模板依次和图 11(b) 所示的分割后的手指二值化图像进行模板匹配, 记录所有指尖模板进行模板匹配时的匹配值, 查找所有匹配值中的最小值对应的指尖模板, 并记录该模板在匹配值最小时在手指二值化图像中的位置 (x, y) , 将该位置的指尖模板的中心作为指尖的位置, 指尖位置的坐标为

$$(X, Y) = \left(x + \frac{W}{2}, y + \frac{H}{2} \right) \quad (17)$$

其中, (X, Y) 为指尖位置的坐标, (x, y) 为模板匹配值最小时指尖模板在二值化图像上的位置, W 和 H 为指尖模板图像的宽度和高度。指尖检测和定位效果如图 13 所示, 图 13 中指尖位置用实心圆标出。

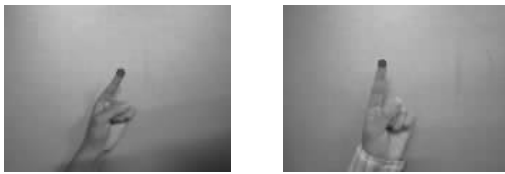


图 13 指尖定位结果

Fig.13 Fingertip positioning results

1.4 生成空中手写轨迹

使用上文中指尖检测和定位算法从视频序列中检测和定位指尖, 然后连接从视频序列中检测到的指尖位置, 从而形成空中手写字符。

书写者正对着摄像头进行隔空虚拟手写, 受单个摄像头的检测范围限制, 用户所能书写的文字长度也受到了限制, 用户书写文本时, 由于手始终位

于摄像头和身体之间的空间中, 导致写成的文本相互叠加非常严重, 空中手写字都是一笔完成, 没有起笔和落笔信息, 只有一个笔画, 所以坐标叠加在一起的文本中字符结构扭曲严重, 导致识别非常困难。为解决这个问题, 使得系统可以让用户自由书写文本, 提出了坐标系虚拟滑动算法。由于用户的书写轨迹是由记录的指尖坐标连线所得, 所以当用户书写时, 每检测一帧图像就将当前帧之前记录的所有指尖的横坐标进行平移, 纵坐标不变, 可以根据平移的像素大小控制滑动的速度, 通过连接平移后的指尖坐标从而形成滑动的字符轨迹, 由于这种横坐标平移是由程序中对坐标加或者减完成, 而不是真实的窗口滑动, 所以是一种虚拟的坐标系移动。当不使用坐标系虚拟滑动时, 空中手写文本 P 可以写成:

$$P = \begin{bmatrix} x_1 & x_2 & \cdots & x_t & \cdots & x_T \\ y_1 & y_2 & \cdots & y_t & \cdots & y_T \end{bmatrix} \quad (18)$$

其中, (x_t, y_t) 是空中手写文本 P 的第 t 个坐标, $t = 1, \dots, T$, T 是 P 的长度, 随着书写样本的不同而变化。当使用坐标系虚拟滑动时, 也就是说, 假设坐标系以速度 V 在与书写方向相反的方向上滑动, 那么 P 的第 t 个点的坐标 (x_t, y_t) 可以记录为

$$\begin{aligned} x_t &= \bar{x}_t + v \Delta t, \quad t \geq 2 \\ y_t &= \bar{y}_t, \quad t \geq 2 \end{aligned} \quad (19)$$

其中, Δt 为两个相邻点之间的时间间隔, 当书写完成后保存最终得到的指尖坐标序列, 形成空中手写文本; $\bar{x}_t$ 和 $\bar{y}_t$ 是第 t 个点在当前帧中的真实坐标。空中手写文本“广东卫视”的手写轨迹如图 14 所示。从图 14 左边可以看出, “广”字并没有显示在屏幕上, 就像显示窗口在向右“滑动”一样。



图 14 “广东卫视”的轨迹

Fig.14 The trajectory of “Guangdong TV”

2 空中手写数据采集系统

2.1 系统设计

通过对空中手写数据采集过程进行可视化可以很好地帮助书写者观察书写效果, 便于实时调整书写姿态从而更好地进行数据采集。使用 PyQt5 对系

统进行界面设计, 实时显示视频帧, 系统界面设计如图 15 所示, 用户使用该系统进行书写时不需要触碰键盘、触摸屏等硬件设备, 只需要通过鼠标触动“打开相机”按钮打开摄像头进行空中手写, 书写过程中系统通过判断指尖的数量来控制书写的开始和结束, 当指尖数量为 1 时即书写开始。系统通过前述指尖定位算法对指尖进行定位, 书写过程中通过连接指尖坐标从而可视化字符轨迹。当用户伸出五指时摄像头关闭即书写结束保存最终得到的指尖坐标, 当书写完成时用户也可以通过鼠标触动“数据处理”按钮关闭摄像头保存指尖坐标。



图 15 基于单目视觉的空中手写数据采集系统界面
Fig.15 Interface of in-air handwriting data acquisition system based on monocular vision

空中手写文字都是一笔完成, 没有起笔和落笔信息, 只有一个笔画。书写过程中的短暂停顿时常发生, 由于指尖的轻微抖动会产生大量的重复点, 这些重复点由于集中于一个小区域, 在文本识别算法预处理中通过区域均值平滑即可消除。空中手写轨迹明显的抖动问题本身也为识别算法带来挑战, 也正是空中手写这些特点, 使得研究专门针对空中手写识别的算法显得极为必要。使用该系统在空中进行书写时如果出现大量的错误则需要重新书写, 无法进行回溯或修改。如图 15 所示, 书写时由于系统使用单目摄像进行数据采集, 其检测范围受到限制, 且不包含深度信息, 所以书写者进行空中书写时正对摄像头。系统可以正向或逆向书写, 逆向书写图 14 中的“广东卫视”仍然可以正常形成书写

轨迹, 从视觉效果看仍然可以形成“广东卫视”。但是与正向完成的“广东卫视”从识别算法的角度已经不再是同一文本类。因此在数据采集过程中书写者需按照一般书写习惯从左往右进行书写。书写过程中书写手应该尽量完全出现在摄像头检测范围内。

2.2 创建数据集

基于本文开发的单目视觉空中手写数据采集系统, 初步构建了一个空中手写文本数据集 IAHT-AHUT2022。为了避免书写字体过于统一而导致过拟合, 组织多个参与者来采集手写样本。该数据集包括中央和各省的电视台标, 如北京卫视、山东卫视、CCTV1、CCTV2 等共 47 个类, 每个类包含 40 个样本。数据集 IAHT-AHUT2022 中样本示例如图 16 所示, 空中手写样本“山东卫视”和“CCTV3”在书写完成之后保存为文本格式, 第一行为该样本包含的轨迹坐标总数, 之后为轨迹坐标序列。

191 -1 0.09589 -0.983193 0.09589 -0.966387 0.09589 -0.94958 0.123288 -0.941176 0.123288 -0.92437 0.09589 -0.92437 0.123288 -0.94958 -0.260274	80 -0.569444 0.886792 -0.583333 0.660377 -0.597222 0.660377 -0.625 0.622642 -0.666667 0.660377 -0.777778 0.584906 -0.888889 0.509434 -0.944444 0.320755

图 16 数据集 IAHT-AHUT2022 中样本示例
Fig.16 Sample example in dataset IAHT-AHUT2022

3 实验

在本节中, 为了验证基于单目视觉的空中手写系统的指尖检测和定位算法的性能, 所有实验在 Intel® Core™ i5-8300H, 内存为 8 GB 的 PC 机上进行。使用笔记本电脑自带的摄像头在不同复杂性和光照水平环境下采集空中手写视频, 例如在白天或者晚上拉开窗帘和拉上窗帘, 在一个灯光和两个灯光等不同环境下进行视频采集。视频图像的分辨率为 320 × 240, 总共采集 30 个不同的空中书写轨迹视频, 其中包含单个字母、数字及电视频道等, 部分实验数据集如图 17 所示。



图 17 部分实验数据集
Fig.17 Part of the experimental dataset

3.1 指尖定位实验和分析

空中手写系统的关键在于对书写者指尖的准确检测和定位。设数据集中指尖总数为 N_0 ，通过提出的检测算法正确检测到的指尖数量为 N_1 ，则系统采集准确率为

$$P = N_1 / N_0 \times 100\% \tag{20}$$

为了验证提出的算法对指尖定位的准确性，通过收集上文采集的不同空中手写视频中的 700 个视频帧图像作为指尖定位实验数据；使用工具手动标注每一帧图像中的指尖位置；通过计算每一帧手动标注的指尖位置和文中提出的指尖检测算法，所得到的指尖位置之间的间隔的像素点数作为误差距离；在指定像素点内(如两者之间的像素点数 ≤ 5)，被视为指尖定位准确。通过对上文收集到的 700 个视频帧图像进行指尖定位实验，可以得出当提出的指尖检测与定位算法在误差为 10 个像素点时指尖定位准确率可以达到 96.85%，在误差为 5 个像素点时指尖定位准确率也能达到 95% 以上，证明了算法的可靠性。

将本文方法与现有的运动目标检测算法进行比较，包括质心距离法^[16]、凸包和缺陷分析法^[17]、K 曲率法^[29]。实验结果如表 1 所示，提出的指尖检测算法的指尖定位准确率明显优于其他算法。

表 1 不同算法的指尖定位结果

Tab.1 Fingertip localization results of different algorithms

方法	准确率 /%
质心距离法	91.0
K 曲率法	93.3
凸包和缺陷分析法	92.0
本文算法	96.8

由于质心距离法、K 曲率法、凸包和缺陷分析法等容易受光照、行人等的干扰，导致指尖定位的准确性会受到影响，而本文算法一方面通过采用平均背景差分法可以有效滤除扰动干扰，另一方面通过 HSV 颜色空间的阴影消除法有效去除了光照角度造成的运动阴影干扰。采用的模板匹配方法可以忽略手指轮廓上一些缺失，对光照明暗变化有了一定的鲁棒性，可以在一定程度上提升指尖定位的准确性。

3.2 采集效果及分析

为了验证系统使用的算法在实际应用情况下的有效性和实时性，在 Intel® Core™ i5-8300H，内存为 8 GB 的 PC 机上采集的视频序列中进行验证。本实验中，增加了光照强度，整个用于测试的视频序列共 1 005 帧，图 18 是在测试视频的第 173 帧、第

367 帧、第 565 帧、第 876 帧的检测效果。

从图 18(a)和图 18(b)中可以看出,在前景目标靠近光照干扰区域时的检测结果会存在一定缺失但是仍然可以检测出较为完整的手部轮廓,并且可以很好地将手指和手掌进行分割,通过提出的算法可以定位不同角度下的指尖位置(见图 18(d))。从算法的整体流程执行速度来看,可以达到 12.5 帧/s 的平均执行速度,通过统计该视频序列共 1 005 帧中手指与手掌分割和指尖检测的结果,本文提出的算法在手指与手掌分割的漏检率为 4.57%,指尖检测的漏检率为 4.73%,误检率为 3.24%,并

且通过手动标注 1 005 帧视频图像中的指尖位置,计算手动标注的实际指尖坐标位置与算法检测的指尖坐标位置之间间隔的像素点,当误差范围在 10 个像素点时,指尖坐标定位准确率达到 96.4%,在实际应用情况下算法仍然可以很好地检测和定位指尖。

图 19 展示了基于单目视觉的空中手写数据采集系统连续书写 11 个字符的图片,并按照书写的先后顺序连接最后得到的指尖坐标序列形成完整的书写轨迹。从图 19 可以看出,系统在实际应用条件下也可以使书写者稳定、实时地进行连续书写。

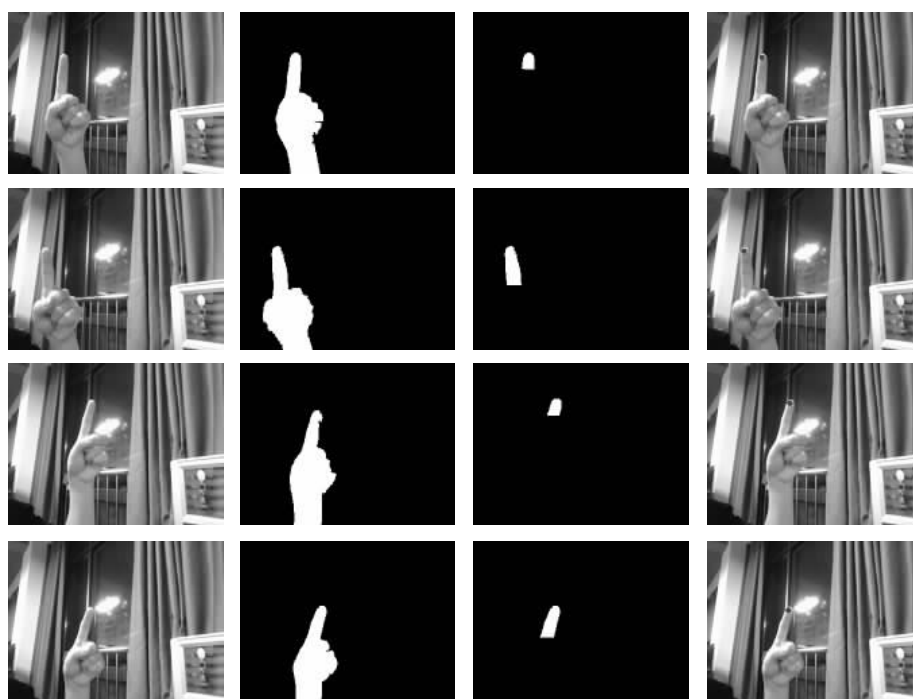


图 18 检测效果示例

Fig.18 Example of detection effect



图 19 连续书写示例

Fig.19 Example of continuous writing

然而，提出的空中手写数据采集系统在实际应用时存在一定的局限性，本文使用的指尖定位算法是在对书写手检测的基础上，利用书写手的几何结构特征对书写者的手部和手指进行分割，然后对分割后的手指使用模板匹配算法进行指尖定位。所以在书写过程中当书写手没有完全出现在摄像头的检测范围内时，书写手无法完全的被检测出来从而会对手指分割造成影响导致指尖定位效果不佳。书写手被遮挡时的检测效果如图 20 所示，从图中可以看到由于书写手被遮挡没有完全出现在视频帧图像

中，虽然检测出书写者的部分手部轮廓，但是利用手的几何结构特征对手指进行分割失去了作用，导致指尖定位时会出现定位错误或者漏检的现象。

系统的另一个局限性是书写者在书写时不能将指尖垂直指向摄像头，不利于指尖检测。如图 21 所示，由于单目摄像头不具有深度信息，当将指尖垂直指向摄像头时，手掌与手指处于同一水平方向时书写者的手指轮廓被手掌轮廓遮盖住，导致无法正确分割手指轮廓，从而在使用模板匹配进行指尖定位时会出现误差甚至无法定位到指尖。

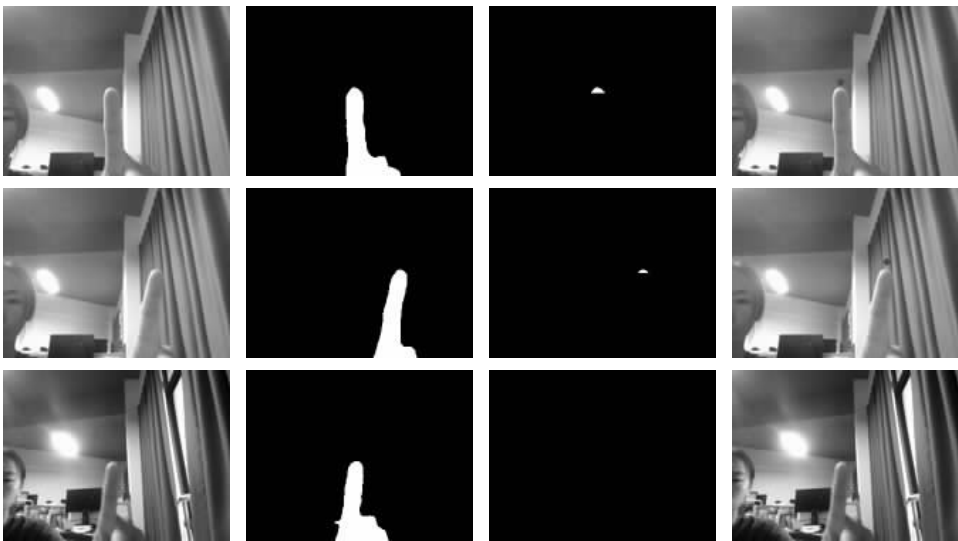


图 20 遮挡书写手的检测效果

Fig.20 Detection effect of obstructing the writing hand

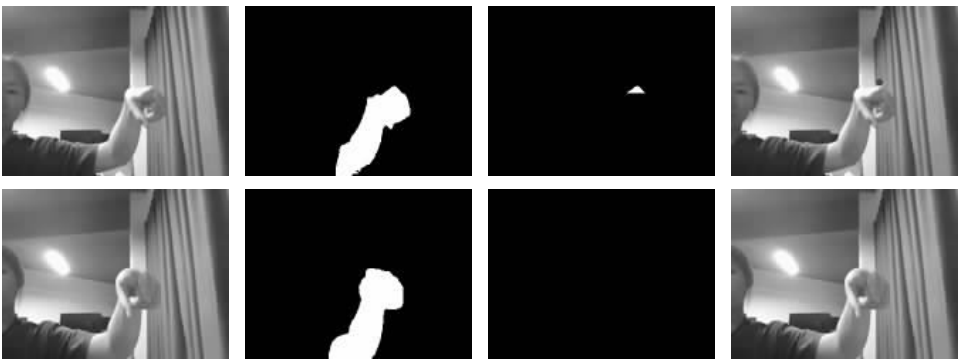


图 21 指尖垂直指向摄像头的检测效果

Fig.21 Detection effect of fingertip pointing vertically towards the camera

4 结论

本文提出一种基于单目视觉的空中手写系统。该系统使用了体积小的摄像头，因此适用于手机、

笔记本电脑、冰箱、电视机等家用电器。由于摄像头比 3 维视觉传感器廉价很多，因此该系统适合大规模应用推广。空中手写系统采用平均背景差分 and HSV 颜色空间阴影消除，在指尖定位中使用了模板

匹配的策略, 因此对于光照明暗变化、光照角度造成的运动阴影、行人等扰动具有较好的鲁棒性, 由于使用了坐标系虚拟滑动技术, 可以让书写者在空中自由地连续书写长文本, 例如诗歌。相对于此前系统每次只能写入一个或两个字符来说该系统的设计贴近实际应用, 比如可以通过空中手写电视频道进行电视台的切换、远距离书写空调温度进行温度调节等。实验结果表明该系统可以稳定可靠地定位

书写者指尖的位置, 从而让用户在空中自由的连续书写任意长度的文本。

提出的系统目前只适合于裸手书写, 系统不适合戴手套书写或者背景环境较为复杂、色彩很丰富的场合。在接下来的工作中, 将针对上述问题开发多种类模板匹配的深度学习指尖定位与检测模型。

参考文献

- [1] JIN L W, YANG D D, ZHEN L X, et al. A novel vision based finger-writing character recognition system[J]. Journal of Circuits, Systems and Computers, 2007, 16(3): 421–436.
- [2] GAN J, WANG W Q, LU K. A new perspective: Recognizing online handwritten Chinese characters via 1-dimensional CNN [J]. Information Sciences, 2019, 478: 375–390.
- [3] XU N, WANG W Q, QU X W. On-line sample generation for in-air written Chinese character recognition based on leap motion controller[C]//16th Pacific-Rim Conference on Multimedia. Berlin, Germany: Springer, 2015: 171–180.
- [4] YE Z C, ZHANG X, JIN L W, et al. Finger-writing-in-the-air system using kinect sensor[C/OL]//IEEE International Conference on Multimedia and Expo Workshops. Piscataway, USA: IEEE, 2013[2023-02-24]. <https://ieeexplore.ieee.org/document/6618236>. DOI: 10.1109/ICMEW.2013.6618236.
- [5] XIA Q X, HONG F, FENG Y, et al. MotionHacker: Motion sensor based eavesdropping on handwriting via smartwatch[C]//IEEE Conference on Computer Communications Workshops. Piscataway, USA: IEEE, 2018: 468–473.
- [6] YAO Y X, LAI J J, LUO C W. Automatic handwriting inference via motion sensor embedded wrist-worn device[C]//5th International Conference on Advanced Robotics and Mechatronics. Piscataway, USA: IEEE, 2020: 364–367.
- [7] FU Z J, XU J S, ZHU Z D, et al. Writing in the air with wifi signals for virtual reality devices[J]. IEEE Transactions on Mobile Computing, 2019, 18(2): 473–484.
- [8] FANG Y Y, XU Y, LI H J, et al. Writing in the air: Recognize letters using deep learning through wifi signals[C]//6th International Conference on Big Data Computing and Communications. Piscataway, USA: IEEE, 2020: 8–14.
- [9] LIN C, XU T T, XIONG J, et al. WiWrite: An accurate device free handwriting recognition system with COTS wifi[C]//40th International Conference on Distributed Computing Systems. Piscataway, USA: IEEE, 2020: 700–709.
- [10] CHEN X, ZHANG H B, CHEN S. Write-Fi: An accurate handwriting recognition system using multimodal net[C]//Asia-Pacific Conference on Image Processing, Electronics and Computers. Piscataway, USA: IEEE, 2022: 115–119.
- [11] KANE L, KHANNA P. Vision-based mid-air unistroke character input using polar signatures[J]. IEEE Transactions on Human-Machine Systems, 2017, 47(6): 1077–1088.
- [12] YIN F, LIU P P, HUANG L L, et al. Lexicon-driven recognition of one-stroke character strings in visual gesture[C]//13th International Conference on Document Analysis and Recognition. Piscataway, USA: IEEE, 2015: 421–425.
- [13] PRADEEP K, SAINI R, BEHERA S K, et al. Real-time recognition of sign language gestures and air-writing using leap motion [C]//Fifteenth IAPR International Conference on Machine Vision Applications. Piscataway, USA: IEEE, 2017: 157–160.
- [14] QU X W, WANG W Q, LU K, et al. In-air handwritten Chinese character recognition with locality-sensitive sparse representation toward optimized prototype classifier[J]. Pattern Recognition, 2018, 78: 267–276.
- [15] QU X W, WANG W Q, LU K, et al. Data augmentation and directional feature maps extraction for in-air handwritten Chinese character recognition based on convolutional neural network[J]. Pattern Recognition Letters, 2018, 11: 9–15.
- [16] MALIMA, OZGUR, CETIN. A fast algorithm for vision-based hand gesture recognition for robot control[C/OL]//14th Signal Processing and Communications Applications. Piscataway, USA: IEEE, 2006[2023-03-08]. <https://ieeexplore.ieee.org/document/1659822>. DOI: 10.1109/SIU.2006.1659822.
- [17] GURAV R M, KADBE P K. Real time finger tracking and contour detection for gesture recognition using opencv[C]//2015 In-

- ternational Conference on Industrial Instrumentation and Control. Piscataway, USA: IEEE, 2015: 974–977.
- [18] ROY P, GHOSH S, PAL U. A cnn based framework for unistroke numeral recognition in air-writing[C]//16th International Conference on Frontiers in Handwriting Recognition. Piscataway, USA: IEEE, 2018: 404–409.
- [19] MUKHERJEE S, AHMED S A, DOGRA D P, et al. Fingertip detection and tracking for recognition of air-writing in videos[J]. Expert Systems with Applications, 2019, 136: 217–229.
- [20] YADAV K S, KIRUPAKARAN A M, LASKAR R H. GCR-Net: a deep learning-based bare hand detection and gesticulated character recognition system for human-computer interaction[J/OL]. Concurrency and Computation Practice and Experience, 2022[2023-01-19]. <https://onlinelibrary.wiley.com/doi/10.1002/cpe.7578>.
- [21] SUN J H, JI T T, ZHANG S B, et al. Research on the hand gesture recognition based on deep Learning[C/OL]//12th International Symposium on Antennas, Propagation and EM Theory. Piscataway, USA: IEEE, 2018[2023-01-09]. <https://ieeexplore.ieee.org/document/8634348>. DOI: 10.1109/ISAPE.2018.8634348.
- [22] CUI X X, ZHANG W B, LIU D. Improved frame difference algorithm based on CNN for moving target detection[C]//39th Chinese Control Conference. Piscataway, USA: IEEE, 2020: 7595–7598.
- [23] ZUO J H, JIA Z H, YANG J, et al. Moving target detection based on improved Gaussian mixture background subtraction in video images[J]. IEEE Access, 2019, 7: 152612–152623.
- [24] HUANG J J, ZOU W, ZHU Z, et al. An efficient optical flow based motion detection method for non-stationary scenes[C]//Chinese Control and Decision Conference. Piscataway, USA: IEEE, 2019: 5272–5277.
- [25] WU M H, CHEN R, TONG Y. Shadow elimination algorithm using color and texture features[J/OL]. Computational Intelligence and Neuroscience, 2020, 2020[2023-02-10]. <https://www.hindawi.com/journals/cin/2020/2075781/>. DOI: 10.1155/2020/2075781.
- [26] 杨鸽, 郑嘉龙, 王莹. 基于 HSV 与 RGB 颜色空间的人体检测与跟踪算法[J]. 自动化技术与应用, 2022, 41(9): 17–20, 28.
- YANG G, ZHENG J L, WANG Y. Human detection and tracking algorithm based on HSV and RGB color space[J]. Techniques of Automation and Applications, 2022, 41(9): 17–20, 28.
- [27] 冷建伟, 沈芳婷. 基于 HSV 色彩模型与区域生长法的水文图像分割[J]. 计算机工程, 2017, 43(7): 223–228.
- LENG J W, SHEN F T. Hydrological image segmentation based on HSV color model and region growing algorithm[J]. Computer Engineering, 2017, 43(7): 223–228.
- [28] 付潇聪, 王浩. 一种基于视觉的手势识别系统[J]. 电子设计工程, 2017, 25(17): 26–30.
- FU X C, WANG H. A vision based gesture recognition system[J]. Electronic Design Engineering, 2017, 25(17): 26–30.
- [29] LI C Q, ZHANG R, LIU Z, et al. Algorithm of fingertip detection and its improvement based on Kinect[C]//International Conference on Industrial Informatics-Computing Technology, Intelligent Technology, Industrial Information Integration. Piscataway, USA: IEEE, 2017: 63–66.

作者简介

屈喜文(1984–), 男, 博士, 讲师, 硕士生导师。研究领域为模式识别, 图像处理, 机器学习。

韩瑶妹(1997–), 女, 硕士生。研究领域为图像处理, 人机交互。

胡冕军(1996–), 男, 硕士生。研究领域为深度学习, 模式识别。