

李若诗, 朱永同, 吴宝磊, 等. 基于深度相机的咽拭子机器人采样系统研究[J]. 智能计算机与应用, 2024, 14(4): 12-18. DOI: 10.20169/j.issn.2095-2163.240402

基于深度相机的咽拭子机器人采样系统研究

李若诗^{1,2}, 朱永同^{1,2}, 吴宝磊^{1,2}, 姚东昊^{1,2}, 李新潮^{1,2}, 刘娜^{1,2}

(1 上海理工大学 健康科学与工程学院, 上海 200093; 2 上海理工大学 机器智能研究院, 上海 200093)

摘要: 本文提出一个基于深度相机的咽拭子采样系统。首先, 基于 Mujoco 仿真平台搭建了一套机械臂协作平台。基于此仿真平台, 构建了一套视觉检测系统, 用来准确获取口腔的位姿信息, 并提出了两个视觉优化算法。一是基于关键点检测, 提出一种防抖算法, 其利用前后帧信息的差分滤波, 用来克服口腔位姿估计过程中因噪声和干扰而导致的抖动问题; 其二是基于非刚体的法向估计算法, 其融合面部关键点信息和先验面部三维信息进行 PnP 估计面部法向量。最后, 基于获得的鲁棒关键点和姿态信息完成了对采样信息的映射, 为后续咽拭子采样机器人高效、安全地使用机械臂完成采样任务提供条件。实验表明, 基于所提视觉检测系统可以准确、稳定的完成机械臂采样任务, 在保证实时检测准确率的基础上能够有效去抖, 在非限定环境下具有更高的鲁棒性和稳定性。

关键词: 咽拭子机器人; 采样系统; 位姿估计; 关键点检测; 防抖算法; 机械臂

中图分类号: TP241

文献标志码: A

文章编号: 2095-2163(2024)04-0012-07

Research on the throat swab robot sampling system based on depth camera

LI Ruoshi^{1,2}, ZHU Yongtong^{1,2}, WU Baolei^{1,2}, YAO Donghao^{1,2}, LI Xinchao^{1,2}, LIU Na^{1,2}

(1 School of Health Science and Engineering USST, University of Shanghai for Science and Technology, Shanghai 200093, China;

2 Institute of Machine Intelligence, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: The pandemic of respiratory infectious diseases has had a tremendous impact on all aspects of human society. The development of throat swab sampling robots is an important means to prevent large-scale outbreaks of respiratory infectious diseases and to liberate productivity. This paper proposes a throat swab sampling system based on a depth camera. First, we built a robotic arm cooperative platform based on the Mujoco simulation platform. Within this simulation platform, we constructed a vision detection system to accurately capture the positional information of the mouth, proposing two visual optimization algorithms. The first is an anti-shake algorithm based on keypoint detection, which uses differential filtering of frame-by-frame information to overcome jitter problems caused by noise and interference during the estimation of the mouth's pose. The second is a non-rigid normal estimation algorithm, which integrates facial keypoint information with prior 3D facial data to perform Perspective-n-Point (PnP) estimation of facial normals. Finally, based on the robust keypoint and pose information obtained, we completed the mapping of sampling information, providing conditions for the subsequent throat swab sampling robot to use the robotic arm efficiently and safely to complete the sampling task. Experiments show that the proposed vision detection system can accurately and stably complete the mechanical arm sampling task. It effectively eliminates jitter while ensuring real-time detection accuracy and exhibits higher robustness and stability in unconstrained environments.

Key words: throat swab robot; sampling system; pose estimation; keypoint detection; anti-shake algorithm; robotic arm

0 引言

人类与传染病之间的斗争从未停止, 例如 2002

年的非典和近期的新冠病毒等呼吸道传染病的大流行, 对人类社会的各个方面都造成了巨大的影响^[1]。呼吸道传染病主要通过飞沫传播、接触传

基金项目: 国家自然科学基金(92048205); 国家重点研发计划资助项目(2023YFC3605800)。

作者简介: 李若诗(1999-), 女, 硕士研究生, 主要研究方向: 计算机视觉, 深度学习; 朱永同(1998-), 男, 博士研究生, 主要研究方向: 点云配准, 深度学习; 吴宝磊(2000-), 男, 硕士研究生, 主要研究方向: 动作识别, 深度学习; 姚东昊(2000-), 男, 硕士研究生, 主要研究方向: SLAM, 深度学习; 李新潮(1999-), 男, 硕士研究生, 主要研究方向: 手势识别, 深度学习。

通讯作者: 刘娜(1985-), 女, 博士, 讲师, 主要研究方向: 目标跟踪, 深度学习。Email: liuna@usst.edu.cn

收稿日期: 2024-01-19

哈尔滨工业大学主办 ◆ 学术研究与应用

播、气溶胶传播^[2]。由于人类的疾病都与基因有直接或间接的关系,因此可以选择咽拭子检测,即基因检测的方式来判断是否感染,从而快速有效地控制传染源,切断传播的途径。面对检测人员紧缺、采样任务繁重且工作环境高危的大环境,研发一款自主、安全和智能操作的咽拭子采样机器人,对积极应对呼吸道传染病至关重要。此外,咽拭子在呼吸道病毒种类检测科室中,仍然有大量的检测需求,如扁桃体病毒检测等。因此,咽拭子检测机器人可以保障医护人员的安全,令采集更加高效规范,有效地缓解医务人员紧缺的问题,对积极应对新冠这类突发传染病具有重大意义。

咽拭子采样需准确定位口腔采样点,该过程极大地依赖于医护人员知识、经验及其智能决策与操作能力。对于咽拭子采样机器人来说,极大地依赖机器人的感知、定位和运动控制能力,因此咽拭子采样系统的设计和实现面临多项技术挑战。首先是确保视觉检测系统具备先进的感知能力和定位能力,能准确感知口腔结构并确定采样点,这需要高度精准的算法与传感技术支持;其次,机器人还需要高度精准的运动控制能力,以确保在采样过程中能够精准地使用机械臂操作棉签或采样工具,实现准确、无损伤的采样操作;最后,将视觉感知数据和机械臂控制相融合是一个复杂的过程。这一过程需要确保数据能够精确地传输和解读,以便机械臂能够在准确地感知其环境的基础上,能够实现精准的操作。

在 COVID-19 疫情期间,国内多个医工交叉团队相继启动了相关的应急性研究工作。如:钟南山院士团队与沈阳自动化研究所联合研发的咽拭子采样机器人^[3];2021年,深圳市罗湖医院集团联合杭州湖西云百生科技有限公司发布全球首台全自动鼻咽拭子核酸采样机器人;2022年,江苏科技大学的李顺君等^[4]提出了基于卷积神经网络的扁桃体咽拭子采样机器人,主要设计了一种基于卷积神经网络的两步检测模型。但是,目前所存在的咽拭子采样机器人检测模块计算量大,稳定性都需进一步增强。

深度学习在提取关键点方面应用广泛,常使用卷积神经网络(CNN)^[5]等结构进行特征提取,以实现关键点的准确定位。在关键点检测中,滤波器常用于图像预处理,降低噪声并改善检测结果。姿态识别中的最小二乘法则常用于模型拟合和参数估计,通过最小化残差实现对物体姿态的准确估计。由此可见,综合运用深度学习和数学模型,能够有效实现对复杂结构的关键点提取和姿态识别。

本文提出一个基于深度相机的咽拭子采样系统,搭建了一套协作机械臂采样系统,并基于深度相机提出了一种稳定的视觉检测系统,在该系统中克服非刚体位姿估计的难点以及关键点检测过程中因缺少时序信息而导致的抖动问题,显著提高了面部关键点识别的稳定性和准确性,使得模型输出的关键点位置更加平滑和稳定,达到了更好的实时检测效果,在一定程度上极大地保护了待采样者的安全。

1 相关工作

1.1 机械臂协作

随着现代社会的不断发展和对高度自动化的需求不断增长,单一控制手臂的应用逐渐显得有限,社会对更为灵活多变控制装置的需求日益凸显,这使得双臂协作技术自然而然地成为机器自动化领域中备受关注的焦点。双臂协作技术以其卓越的自由度和协同工作的能力,在各个领域中得到了广泛的应用。这种趋势的发展源于对更为高效、智能和多功能机器人系统的需求,推动了双臂协作技术的进一步发展。

特别是在高安全性需求的工作场景中(如:咽拭子采样等操作),对机械臂的要求更为严格。咽拭子采样作为一项重要的医学工作,对样本的准确采集和高度安全性的要求,使得传统的单臂控制技术显得力不从心。因此,双臂协作机械臂的引入成为解决这一问题的有效途径。双臂协作机械臂具备协同作业、高度灵活和精准的特点,能够满足咽拭子采样等工作中对多维度运动 and 高度精密控制的需求。通过将咽拭子采样与双臂协作技术相结合,不仅提高了工作效率,更确保了操作的安全性和准确性。

1.2 关键点检测

人脸关键点检测,是指通过一定方法将人脸面部各个部位的重要特征点进行定位。随着时间的推移,深度学习在人脸关键点检测领域取得了显著进展。与传统算法相比,在复杂环境下表现出更高的准确性与稳定性^[6]。为了提升关键点检测的精度,人脸关键点检测的数量也从5个点变为21点、68点、98点等^[7]。在实际咽拭子检测工程应用中,高安全性是第一位,因此口腔区域的关键点检测的高稳定性是本文考虑的重中之重。

2013年,Sun等^[8]首次在人脸关键点检测领域运用深度学习的方法,提出一种级联的卷积神经网络(Deep Convolutional Neural Network, DCNN)。该

方法具有强大的特征提取能力,使得关键点检测更为准确。Bulat 等^[9]提出的人脸对齐网络 (Face Alignment Network, FAN),实现了高精度的关键点定位,但对于复杂的网络结构需要较大的计算资源和较长的训练时间。Dong X 等^[10]提出的样式聚合网络 (Style Aggregated Network, SAN) 方法,通过堆叠 Hourglass 网络实现多尺度处理,提高了检测准确性,但在实时性方面存在着挑战。Xu Y 等^[11]提出的 CenterFace 以实时性和性能平衡著称,是一种用于实时人脸检测和关键点定位的方法,具有良好的速度和性能平衡。但在某些情况下,准确率低于专门针对关键点检测的算法。Guo X 等^[12]提出的实用面部关键点检测器 (Practical Facial Landmark Detector, PFLD) 算法,旨在实现实时性和准确性之间的平衡。主要思想是通过构建并联的子网络来加快检测速度,并在全局和局部两个层面上进行面部关键点定位。PFLD 算法的优点在于其通过并联的结构将全局和局部信息分别处理,从而提高了检测速度,在面部关键点定位上取得了较好的准确性。PFLD 是一个精度高,速度快,模型小的人脸关键点检测模型,在移动端达到了超实时的性能。但是在实验过程中发现,在口腔张开的情况下口腔处关键点的检测准确率较低,且出现了抖动的情況。

1.3 法向估计

1.3.1 通过 PnP 进行法向估计

通过透视 n 点法 (Perspective-n-Point, PnP)^[13]进行法向估计的实现原理,是基于物体表面上已知的三维点和其在相机坐标系中的对应投影点来计算相机的姿态,然后使用相机的视点和三维点来估计物体表面上的法向量。首先, PnP 算法使用已知的三维点和其在二维图像中的投影点,求解相机的姿态 (旋转和平移矩阵); 接下来, 根据相机的视点和三维点的位置, 计算出物体表面上每个点的法向量。PnP 进行法向估计之所以具有较高准确性, 是因为

其基于已知的三维点和相机的姿态来计算法向量。然而, PnP 方法也有一些缺点, 对大量标点和准确的相机姿态估计的依赖, 以及在特定情况下可能存在数值稳定性问题。此外, 对于复杂的场景和非均匀表面的物体, PnP 方法的性能可能会受到限制。

1.3.2 通过点云进行法向估计

通过点云^[14]进行法向估计, 其实现原理通常基于邻域分析和统计方法。首先, 对于每个点, 以其附近的邻域点构建一个局部区域, 然后通过拟合局部区域的平面或曲面来估计法向量。最常用的方法之一是最小二乘法, 即通过最小化法向量与邻域点的拟合平面之间的残差来估计法向量; 另一个常见的方法是基于主成分分析 (Principal Component Analysis, PCA)^[15], 其通过分析点云局部区域的协方差矩阵, 找到法向量的主方向。优点在于其适用于各种类型的点云数据, 提供高质量的法向估计, 对噪声和不规则性具有一定的鲁棒性, 可用于多种应用, 如三维重建、物体识别和虚拟现实。然而, 该方法也存在着一定的缺点, 对点云密度的依赖, 可能会受到噪声和异常值的干扰, 且计算复杂性较高, 需要大量计算资源。

2 关键实现技术

本文提出的协作机械臂采样系统整体流程如图 1 所示。首先, 通过深度相机获取二维信息和深度信息, 使用人脸关键点检测算法在二维图像上提取人脸关键点; 然后根据深度相机获取到关键点的深度信息, 得到口腔周围的点云, 再利用最小二乘法拟合口腔所在的平面, 从而估计出口腔平面的法向量; 最后根据口腔位姿的法向量, 驱动咽拭子机械臂轨迹规划, 完成口腔内部的咽拭子采样。研究针对提取关键点时关键点抖动、缺乏时序信息等问题, 设计了适应性的滤波器模块, 使得后续口腔法向量位姿估计准确性得到提升。

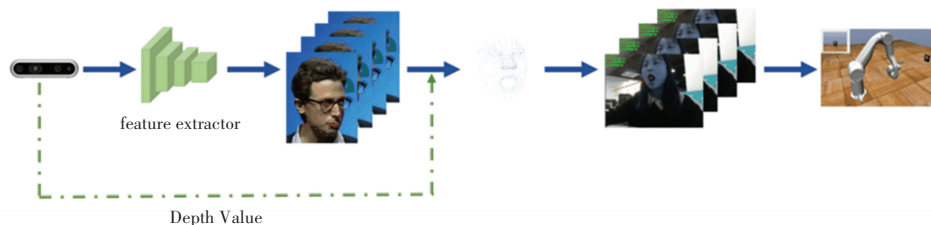


图 1 机械臂采样系统整体流程

Fig. 1 Overall process of the robotic arm sampling system

2.1 机械臂协作

本文采用 Mujoco 作为底层物理仿真引擎, 用于

构建仿真环境, 以便进行强化学习的训练。在仿真和实验环境中, 均选择了具有 7 个自由度的思灵机

器人 Diana 机械臂。

机械臂控制采用基本的导纳控制,导纳控制是基于位置的阻抗控制^[16],其将与外界环境的交互力作为输入,并转换为相应的运动偏移量作为输出。为了实现这一策略,通常需要在末端执行器上加入 F/T 传感器。本文中对笛卡尔空间下的导纳控制进行讨论,实现对机械臂末端执行器的位置控制。笛卡尔空间下的导纳控制可以用公式(1)表示:

$$\mathbf{M}(\ddot{x}_d - \ddot{x}) + \mathbf{B}(\dot{x}_d - \dot{x}) + \mathbf{K}(x_d - x) = F_d - F \quad (1)$$

其中, \ddot{x} 、 \dot{x} 、 x 分别对应机器人末端执行器当前在笛卡尔空间的加速度、速度和位置; \ddot{x}_d 、 \dot{x}_d 、 x_d 则分别表示机器人末端执行器在笛卡尔空间的期望加速度、期望速度和期望轨迹; F 表示末端执行器当前受到的力和力矩, F_d 表示末端执行器上期望保持的外力和外力矩; 矩阵 \mathbf{M} 、 \mathbf{B} 、 \mathbf{K} 为正定对角矩阵, 分别代表导纳系统中的虚拟惯性、虚拟阻尼和虚拟刚度参数。为了简化计算, 可将公式(1)改写为公式(2):

$$\mathbf{M}\ddot{x}_e + \mathbf{B}\dot{x}_e + \mathbf{K}x_e = F_e \quad (2)$$

其中, x_e 是机器人末端执行器的期望轨迹 x_d 与当前位置 x 的误差, F_e 是末端执行器上的期望所受外力和力矩 F_d 与当前传感器测量力和力矩 F 之间的误差。

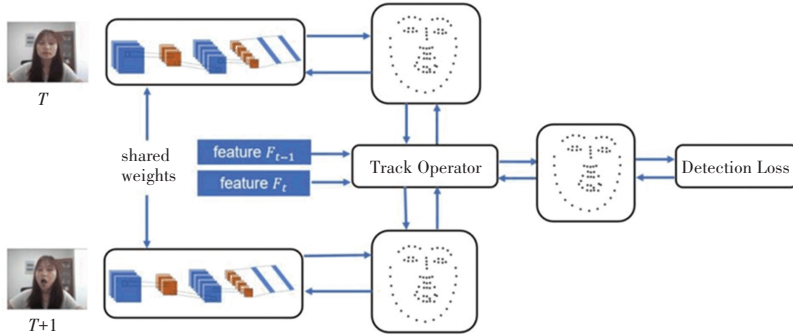


图2 网络结构

Fig. 2 Network structure

本文使用平均绝对误差 (Mean Absolute Error, MAE), 衡量预测的关键点位置与实际位置之间的距离。均方误差公式如式(1)所示:

$$MAE = \frac{1}{n} \sum_{i=1}^n \|y'_i - y_i\| \quad (4)$$

其中: MAE 表示均方误差; n 表示样本数; y_i 表示实际观测值(实际的数据点); y'_i 表示模型预测的值; $\|$ 表示绝对值运算符; Σ 表示对所有样本 i 进行求和操作。在本文中, 取每个关键点坐标与真值关

2.2 鲁棒关键点提取网络

本文提出用于提取人脸关键点的神经网络模型, 主要目的是在图像中能够准确、稳定地检测出人脸的各个关键点, 如眼睛、嘴巴、鼻子等部位的位置, 并通过滤波器模块解决检测过程中关键点随时间变化出现的抖动问题。模型整体结构如图2所示。

Backbone 网络负责提取特征, 用于学习图像的低级和高级特征, 使得模型能够理解图像中的各种信息。在此采用 MobileNet^[17] 作为基础, 输入为两帧图像, 根据特征提取网络得到 256×256 的特征图, 作为下一个模块的输入。

滤波器网络是专门用于关键点防抖的子网络, 结合线性加权函数对两个 256×256 的特征图进行滤波, 如公式(3)所示:

$$F'_i = \alpha F_{i-1} + \beta F_i \quad (3)$$

其中, F_i 为当前时刻的特征图; F_{i-1} 为上一时刻的特征图; α 、 β 分别为两个 256×256 的系数矩阵; 输出 F'_i 为当前时刻新的特征图。该网络在反向传播中不更新特征图, 只更新系数矩阵。最后根据输出的特征图, 通过全连接层处理, 最终输出关键点的坐标。该网络会在不同的分辨率上进行预测, 以保证模型能够在不同尺度下准确地识别关键点, 并防止关键点抖动。

键点坐标来计算损失, 可以最直接的反应网络的性能表现。

2.3 拟合口腔点云平面, 法向估计

由于数据点的分布具有较好的拟合性, 在定义一个合适的拟合范围和权重后, 使用最小二乘法估计, 便可以快速准确的获取拟合结果。考虑到口腔点云平面是一个柔性平面, 本文选取口腔中心点和嘴部周围点来近似代表口腔平面, 进而定义目标点为口腔中心点 (x_0, y_0, z_0) , 拟合平面的领域点集 (x_i, y_i, z_i) 。

使用最小二乘法,根据公式(5)构建误差函数,最小化所有数据点到平面的距离的平方和:

$$E(A, B, C, D) = \sum_i (Ax_i + By_i + Cz_i - D)^2 \quad (5)$$

式中参数估计通过微分求解,具体求解方式如公式(6):

$$\frac{\partial E}{\partial A} = 0, \frac{\partial E}{\partial B} = 0, \frac{\partial E}{\partial C} = 0, \frac{\partial E}{\partial D} = 0 \quad (6)$$

最后,关键的法向量 N 可以从上述平面方程系数中提取出来:

$$N = (A, B, C) \quad (7)$$

其原理是通过最小化实际观测值与模型预测值之间的残差平方和来确定模型参数。

3 验证

3.1 实验环境与参数设置

本实验框架搭建在 pytorch 框架上,运行在 GPU 上。实验运行环境的服务器配置 CPU: i7, 显卡: RTX2080 ti, 系统: 64 位 Ubuntu22.04, 深度学习框架: pytorch2.0。

对于超参数,设置批量大小为 256,采用自适应随机地图下降法(Adam)。总共训练 80 个 epoch,初始学习率为 0.01,后面每隔 10 个 epoch 就衰减为原来的 0.2。

3.2 数据集

本研究选择 300-VW^[18] (300 Videos in the Wild) 数据集作为评估人脸关键点定位准确性的测试数据集。该数据集包含 114 个人脸视频,视频长度从 649 帧到 3 650 帧不等,平均长度为 1 917.4 帧。这些视频的清晰度各不相同,人脸边框的尺寸变化范围也很大,从 46 像素到 753 像素不等。

将视频拆分成图像帧序列 ($V = \{I_f, f = 1, \dots, F_V\}$) 之后,为了定位面部的 ROI 使用 Yolov5 进行人脸检测,得到人脸边框序列 ($I_i = (x_i, y_i, l_i)$)。其中, x_i, y_i 表示人脸边框的左上角坐标, l 表示人脸边框的边长。将得到的人脸图片进行比例放大,可以得到更多的像素信息。

3.3 评价指标

为了准确分析实验结果,选取标准归一化平均误差(Normalized Mean Error, NME)^[19] 作为评估模型算法的指标。计算方式如公式(8)所示:

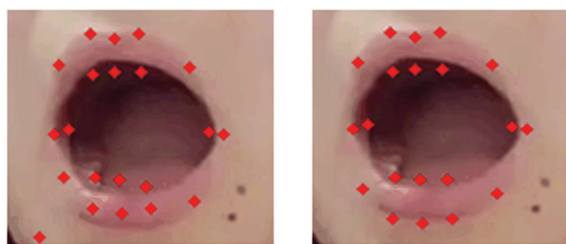
$$e = \frac{\sum_{i=1}^N \|x_i - x_i^*\|_2}{N \times d} \quad (8)$$

其中, N 代表预测关键点的总数; x_i 为预测第 i 个的坐标; x_i^* 为第 i 个坐标真实值; $\|\cdot\|_2$ 表示欧几里得距离(二范数); d 为两眼之间的距离,用作归一化的基准。

3.4 实验结果与分析

为了证明本文方法的有效性,通过实验与近几年的人脸关键点检测方法进行对比论证,分别从性能指标和检测结果上进一步证明所选模型的先进性。

如图 3 所示,在数据集 300-VW 测试时发现,当不添加防抖模块时,口腔处所识别的关键点会因人转头而产生抖动位移;而添加了防抖模块,口腔的抖动明显消除。

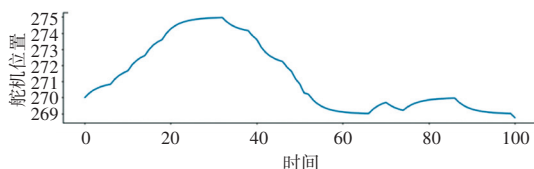


(a) 无防抖模块 (b) 加防抖模块

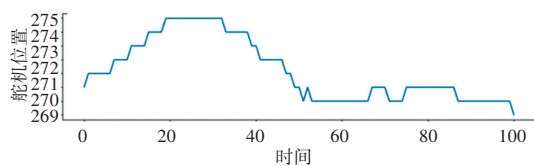
图 3 关键点检测可视化

Fig. 3 Keypoint detection visualization

添加防抖模块前后的舵机位置对比如图 4 所示。可以看出,未经滤波的网络输出舵机存在较大的突变现象,导致关键点发生抖动;加入滤波模块后,曲线变得更加平滑。



(a) 无防抖模块



(b) 加防抖模块

图 4 舵机位置对比图

Fig. 4 Comparison of servo position

在 300-VW 数据集上,通过 NME 指标将提出的方法与现有技术进行了比较;与经典的人脸关键点检测算法 TCDCN、DSRN、TSTN、TSCN、SDM、CFSS 等进行了对比。对比结果见表 1,显示了在 300-VW 数据集上人脸对齐性能的 NME 值。

表1 人脸对齐的 NME

Table 1 NME for face alignment

算法	类别 1	类别 2	类别 3
TCDCN	7.66	6.77	14.98
DSRN	5.33	4.92	8.85
TSTN	5.36	4.51	12.84
FHR+STA	4.40	4.16	5.96
SDM	7.41	6.18	13.04
CFSS	7.68	6.42	13.67
Ours	4.33	4.03	8.23

从表1中可以看出,本文提出的方法在3个类别中产生了2个最佳结果。相较于FHR+STA,在类别1和类别2分别提升了0.07和0.13。从而证明了本文所提出的方法在人脸关键点检测中的有效性。

3.5 咽拭子采样仿真

为了实现稳定准确并且高效采样的目的,本文基于Mujoco仿真平台^[20]进行实验验证。执行器采用思灵的Diana机械臂,口腔待采样点用采样块近似替代,如图5所示。

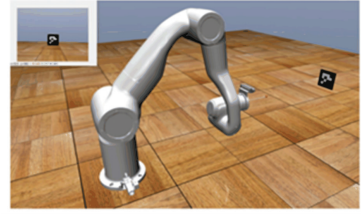


图5 Mujoco环境下机械臂咽拭子采样仿真

Fig. 5 Robotic arm throat swab sampling simulation in Mujoco environment

为精确评估采样效果,本文设计了动态咽拭子采样实验。随着待采样点的移动,同步记录机械臂末端咽拭子与待采样点的空间位置,全面衡量采样的准确性及响应速度。如图6所示,在图6(a)和图6(b)中可以清晰看出,当待采样点在空间平面内发生移动时,机械臂末端能够迅速而准确地作出相应调整。由于相机与待采样点之间存在一个固定的高度差 h ,从理论上讲,待采样点与机械臂末端传感器之间的距离应等于 h 。图6(c)则进一步验证了在空间纵轴上的采样精度。

综上所述,通过在空间3个轴上进行的这一系列实验,有力地证明了机械臂在响应待采样点位置变化时,展现出良好的准确性和时间响应性。

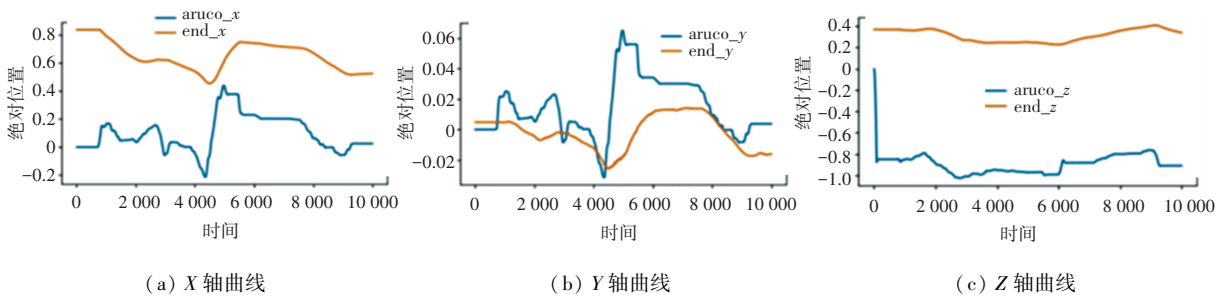


图6 机械臂咽拭子与待采样点空间位置动态响应曲线

Fig. 6 Dynamic response curve of the spatial position of the robotic arm throat swab and the sampling point

4 结束语

本文提出一个机械臂咽拭子采样系统,并针对采样系统的现有问题采用了一系列改进方法。首先,基于Mujoco仿真平台搭建了一套完整的机械臂协作框架。此外,利用提出的面部关键点防抖算法和法向估计算法准确高效的获得面部关键点,从而结合深度相机融合,生成人脸的三维点云。最后,创新性的通过嘴部中心点位置和通过最小二乘估计得到的人脸点云整体朝向,对棉签入点进行位姿估计,将得到的位姿信息传入机械臂的控制系统,完成整个采样流程。评价结果表明,本文所提模型在确保

高精度的前提下可以有效去抖,在非限定环境下具有更高的鲁棒性和稳定性,可以使机械臂稳定高效的完成整个采样过程。

参考文献

- [1] 王金龙,陈涛,任翔,等. 重要呼吸道传染病智能化监测预警与有效应对策略探讨[J]. 病毒学报,2021,37(5):1175-1178.
- [2] 王永华,李亚,刘志琛,等. 基于5M1E分析法的呼吸道传染病院内疫情防控现场管理研究[J]. 护理管理杂志,2020,20(11):823-827.
- [3] LI S Q, GUO W L, LIU H, et al. Clinical application of an intelligent oropharyngeal swab robot: Implication for the COVID-19 pandemic[J]. European Respiratory Journal, 2020, 56(2): 2001912.

- [4] 李顺君,钱强,史金龙,等. 基于卷积神经网络的扁桃体咽拭子采样机器人[J]. 计算机工程与应用,2022,58(15):324-329.
- [5] WANG A, CHEN H, LIN Z, et al. Repvit: Revisiting mobile cnn from vit perspective[J]. arXiv preprint arXiv:2307.09283, 2023.
- [6] 陈凌宇. 基于深度卷积神经网络的人脸特征点检测的研究[D]. 厦门:厦门大学,2017.
- [7] 李丹阳. 基于深度学习的人脸关键点检测算法研究[D]. 武汉:华中科技大学,2021.
- [8] SUN Y, WANG X, TANG X. Deep convolutional network cascade for facial point detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition . IEEE, 2013: 3476-3483.
- [9] BULAT A, TZIMIROPOLOS G. How far are we from solving the 2D and 3D face alignment problem? (and a dataset of 230.000 3D facial landmarks) [C]//Proceedings of the 2017 IEEE International Conference on Computer. Vision: ICCV2017.Venice, Italy: IEEE,2017:745-1479.
- [10] DONG X, YAN Y, OUYANG W, et al. Style aggregated network for facial landmark detection [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018:379-388.
- [11] XU Y, YAN W, YANG G, et al. CenterFace: Joint face detection and alignment using face as point [J]. Scientific Programming, 2020;1-8.DOI:10.1155/2020/7845384
- [12] GUO X, LI S, YU J, et al. PFLD: A practical facial landmark detector[J]. arXiv preprint arXiv:1902.10859, 2019.
- [13] LEPETIT V, MORENO - NOGUER F, FUA P. EPnP: An accurate solution to the PnP Problem[J]. Interantional Journal of Computer Vision, 2009,81:155-166.
- [14] WANG J, MA Z. Sparse tensor - based point cloud attribute compression [C]//Proceedings of 2022 IEEE 5th International Conference on Multimedia Information Processing and Retrieval (MIPR). IEEE, 2022: 59-64.
- [15] SHLENS J. A tutorial on principal component analysis[J]. arXiv preprint arXiv: 1404.1100, 2014.
- [16]王春明. 协作机器人阻抗/导纳混合系统框架与力控制研究[D]. 南京:南京航空航天大学,2019.
- [17]MARKS A, ZHU M L, et al. MobileNetV2: Inverted residuals and linear bottlenecks [C]//Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2018:4510-4520.
- [18]TANG Y, JIANG Z, XIE Z, et al. Breaking shortcut: Exploring fully convolutional cycle - consistency for video correspondence learning[J]. arXiv preprint arXiv:2105.05838, 2021.
- [19]TRIGEORGIS G, SNAPE P, NICOLAOU M A, et al. Mnemonic descent method: A recurrent process applied for end-to-end face alignment [C]//Proceedings of 29th IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016: 4177 - 4187.
- [20]TODOROV E. Convex and analytically-invertible dynamics with contacts and constraints; Theory and implementation in MuJoCo [C]//Proceedings of IEEE International Conference on Robotics and Automation. IEEE, 2014:6054-6061.