

自适应人脸多区域分析的视频心率检测^①



陈宇^{1,2}, 杨学志^{2,3}, 方帅^{1,2}, 李龙伟⁴, 王定良^{1,2}

¹(合肥工业大学 计算机与信息学院, 合肥 230009)

²(工业安全与应急技术安徽省重点实验室, 合肥 230009)

³(合肥工业大学 软件学院, 合肥 230009)

⁴(中国科学技术大学附属第一医院, 合肥 230036)

通信作者: 杨学志, E-mail: xzyang@hfut.edu.cn

摘要: 心率是衡量人体心血管健康状况和情绪压力的重要生理参数。然而, 基于视频的非接触式心率检测技术在真实场景中, 会由于人脸运动和光照变化等导致检测准确性的降低。为了解决上述问题, 考虑到心率检测算法中感兴趣区域 (region of interest, ROI) 的选取与检测准确度高度相关。故提出一种自适应超像素分割多区域综合分析的心率检测新方法。首先利用人脸检测和追踪算法, 裁切获得人脸图像; 之后采用自适应超像素分割算法将 ROI 划分成互不重叠的子块; 再通过色度特征提取构建各子块原始血液容积脉搏矩阵; 最后对脉搏矩阵使用多指标综合分析并挑选出最佳区域进行心率估计。实验结果表明, 通过自适应超像素分割和多区域分析优选可以有效提升心率检测准确性。在静止状态下和运动干扰条件下准确性分别达到 99.1% 和 95.6%, 光照干扰条件下准确性相对传统方法最高提升 8.2%。增强了真实场景下心率检测的鲁棒性。

关键词: 视频成像技术; 自适应分割; 多区域分析; 心率检测

引用格式: 陈宇, 杨学志, 方帅, 李龙伟, 王定良. 自适应人脸多区域分析的视频心率检测. 计算机系统应用, 2022, 31(10): 175-183. <http://www.c-s-a.org.cn/1003-3254/8702.html>

Video Heart Rate Detection with Adaptive Face Multi-region Analysis

CHEN Yu^{1,2}, YANG Xue-Zhi^{2,3}, FANG Shuai^{1,2}, LI Long-Wei⁴, WANG Ding-Liang^{1,2}

¹(School of Computer and Information, Hefei University of Technology, Hefei 230009, China)

²(Key Laboratory of Industrial Safety and Emergency Technology of Anhui Province, Hefei 230009, China)

³(School of Software, Hefei University of Technology, Hefei 230009, China)

⁴(The First Affiliated Hospital of USTC, Hefei 230036, China)

Abstract: Heart rate is an important physiological parameter for measuring human cardiovascular health and emotional stress. However, video-based non-contact heart rate detection techniques can degrade the detection accuracy in real scenarios due to facial movements and lighting changes. To solve the problem, this study proposes a new method of heart rate detection based on adaptive superpixel segmentation and multi-region integrated analysis depending on the high correlation between the selection of the region of interest (ROI) in a heart rate detection algorithm and its detection accuracy. Firstly, a face detection and tracking algorithm is used to crop the face image. Then the ROI is divided into non-overlapping sub-blocks by an adaptive superpixel segmentation algorithm. The original blood volume pulse matrix of each sub-block is constructed by chromaticity feature extraction. Finally, the pulse matrix is analyzed using multiple indicators, and the best region is selected for heart rate estimation. The experimental results show that the heart rate detection accuracy can be effectively improved by adaptive superpixel segmentation and optimal selection through multi-region analysis. The accuracy reaches 99.1% and 95.6% under stationary and motion disturbance conditions, respectively, and the accuracy is improved by up to 8.2% under illumination disturbance conditions compared with that of the

① 基金项目: 安徽高校协同创新项目 (GXXT-2019-003); 合肥工业大学“智能互联系统安徽省实验室”开放基金 (PA2021AKSK0111)

收稿时间: 2021-12-31; 修改时间: 2022-01-29; 采用时间: 2022-02-15; csa 在线出版时间: 2022-06-28

traditional method. The proposed method enhances the robustness of heart rate detection in real scenarios.

Key words: video photoplethysmography; adaptive segmentation; multi-region analysis; heart rate detection

心率作为人的基本生命体征,不仅可以反应人体的健康状态,也是评估人体心脑血管疾病最重要和有效的参数之一.其心脑血管疾病、高血压、心理和精神疾病等领域都被证实是一种有效的诊断手段^[1].

临床上传统的心率监测技术主要是利用心电图机(electrocardiogram, ECG),其通过设备上的十二导联绑定在体表不同的裸露部位获取精确的心电数据进行分析处理^[2].这种方式可以得到较为精准的心率,但是其存在成本高昂、操作繁杂、专业要求高,普通家庭无法独立使用.除此以外,家用的脉搏血氧仪也被广泛用来测量受试者心率,其是利用光电传感器夹在受试者指尖、耳垂等体表末梢通过接收反射或透射的光信号进行心率测量.该方法需要测试部位和设备紧密贴合且不能够有相对运动,使用不便,舒适度较差不利于长时间实时检测.针对传统方法的在实际应用中的局限性,研究一种使用方便的非接触式的心率检测技术就显得十分重要.

基于视频的光容积描记术(video photoplethysmography, VPPG),是一种非接触式心率检测技术.其可以利用普通彩色相机采集面部或指尖等区域表皮视频,通过检测皮下毛细血管由于心脏周期泵血使得血液血红蛋白浓度变化产生的微弱颜色变化,提取血液容积脉冲信号(blood volume pulse, BVP).通过分析该信号的频率即可进行心率估计.Verkruysse等人^[3]最早使用普通彩色摄像头,在可见光作为光源的条件下证明了使用VPPG技术进行非接触式心率检测的可行性.随后,Poh等人^[4]提出对视频帧进行人脸感兴趣区域检测并计算RGB三通道像素均值,然后使用独立成分分析(ICA)进行BVP信号提取和心率估计.尽管这些算法都可以比较好的进行心率估计,但需要测试者全程保持静止或人脸和环境背景的稳定才能保证良好的实验结果,这不仅会给测试者带来不便,而且会限制算法的使用场景.由于人脸运动(如说话、表情变化)、头部晃动和光照不均匀等干扰因素的存在,会影响VPPG技术BVP信号的提取以及心率估计的准确性.因此,如何保证在真实的生活场景中提升心率检测算法的稳定性和准确性,已经成为国内外学者研究的重点.

研究者们主要做出的改进,有如下几个方面:其一,是根据皮肤光照反射特点采用不同的颜色空间进行BVP信号提取,通过不同颜色通道信号的差异性和互补性来增强原始BVP信号.如de Haan等人^[5]利用不同颜色通道信号之间的互补性,提出了基于色度的颜色空间投影分解算法(chrominance-based rPPG, CHROM)以减轻头部晃动产生的影响.其二,是根据BVP信号特性以及先验知识等进行噪声滤除和信号增强等操作.如Wang等人^[6]通过分析BVP信号频域上心动信号和噪声信号的区别,采取频域修正的方式来去除非心率范围内的干扰.Tasli等人^[7]通过估计头部运动频率并进行剔除,取得了更鲁棒的心率估计.

除此以外,近年来研究表明通过同步分析多个不同的面部感兴趣区域的信号可以用来提升原始BVP信号的提取效果.如Kumar等人^[8]使用信噪比来对不同的感兴趣区域信号进行加权.Feng等人^[9]将ROI区域分块后设计并提取了两种特征,最后使用K均值算法自适应选择有效的感兴趣区域.Li等人^[10]使用人脸特征点裁剪获得固定的人脸皮肤区域.Favilla等人^[11]使用ICA算法分析面部3个区域的多通道信息进行心率检测.周双等人^[12]采用矩阵低秩分解来对不同感兴趣区域的信号进行降噪处理.然而,由于以上的多区域分析的划分均是在人脸检测框的基础上进行简单的矩形分割,此方式虽然操作简单,但实际上并没有考虑到分块内人脸轮廓、阴影和皮肤纹理等信息,以及视频帧之间分块区域位于人脸位置的一致性,ROI分块划分存在一定的不合理之处.

为了解决上述ROI分块划分和挑选中存在的问题,本文受传统的超像素分割算法的启发,提出一种基于人脸视频的自适应超像素分割和ROI优选方法.使得分块区域拥有相同或近似的特征集,一方面提升子块划分质量,从源头提升了BVP信号提取的质量,另一方面通过优选算法寻找面部信号最佳区域.增强了面部运动干扰以及环境光分布不均时算法的抗干扰能力.

1 心率检测方法

本文心率检测算法是以ROI超像素分割和脉搏矩

阵信号优选为基础. 通过裁切视频帧内面部 ROI, 随后根据轮廓、皮肤纹理和阴影等特征将面部 ROI 通过超像素聚类分割成互不重叠的子块, 再对分块结果进行色度 BVP 矩阵提取, 最后进行多指标综合分析优选出最佳面部子区域进行信号处理与心率估计.

本方法的核心步骤包括以下几个部分: 1) 人脸与特征点检测; 2) 人脸自适应超像素分割; 3) 色度 BVP 矩阵提取; 4) 脉搏信号评估优选与心率估计. 具体算法框图如图 1 所示.

1.1 人脸与特征点检测

在真实生活场景中, 由于检测过程中头部晃动以及表情变化等会影响摄像头对面部颜色周期性变化的

提取效果, 进而导致 BVP 信号提取质量的大幅降低. 因此对原始视频进行人脸与特征点检测算法, 来获取稳定的人脸 ROI 区域 (图像帧). 首先根据视频第 1 帧图像, 使用 Viola 等人^[13]的人脸检测器输出视频中人脸检测矩形区域的顶点坐标; 然后采用鲁棒判别响应图拟合 (DRMF)^[14]方法对人脸坐标区域内主要器官进行特征点定位, 记录人脸 66 个特征点位置 $C1$, 包含眼睛、鼻子、眉毛、脸颊和嘴巴等区域. 之后对后续的视频帧进行 KLT 跟踪特征点的位置, 获取连续帧面部特征点矩阵 C , 记作:

$$C_i = [C1(i), C2(i), \dots, C_m(i)] \quad (1)$$

其中, i 为视频帧数, m 为特征点个数, 范围为 1-66.

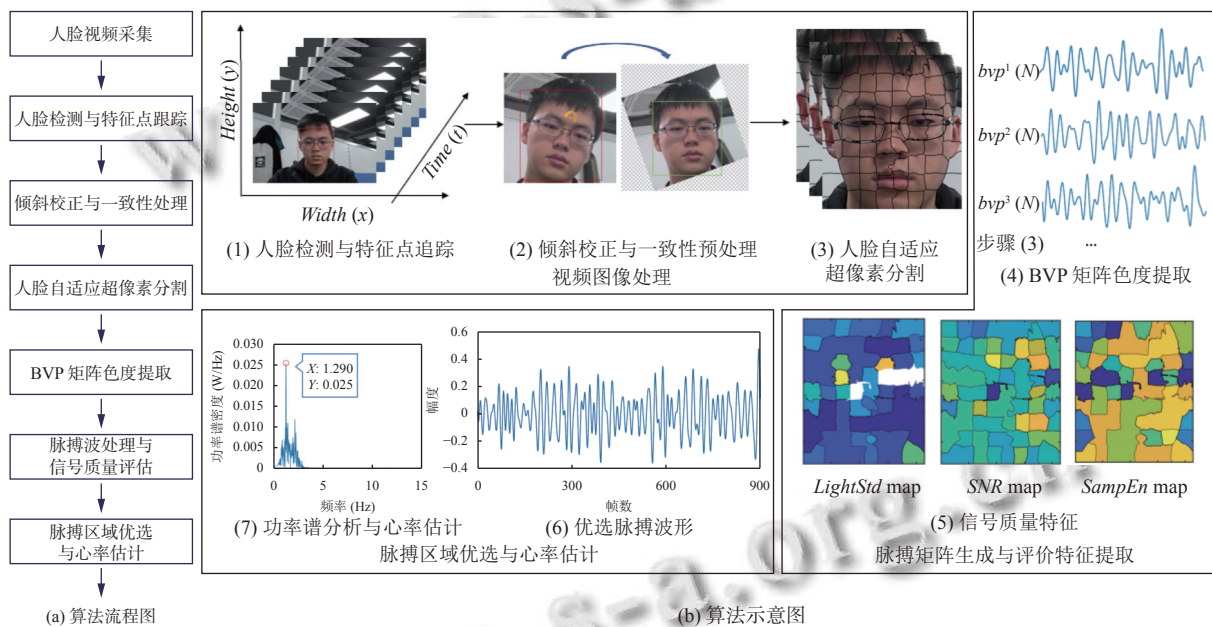


图 1 算法流程及示意图

1.2 人脸自适应超像素分割

传统的 ROI 分割算法均是人脸检测框内区域简单的分割成 m 行 n 列的矩形. 如图 2 所示. 这样分割, 首先并没有考虑到头部运动导致分块对应人脸位置的变化, 其次也没有考虑到分块内像素间的特征联系, ROI 分割存在不合理之处. 例如: 平面内头部偏转会造成检测框改变, 分块对应面部位置变化, 导致连续帧之间分块不连续; 光照分布不均时阴影区域和正常亮度区域若被随机划分为一个子块, 会导致子块内像素均值增大, 颜色变化幅度减小进一步降低 BVP 信号提取质量.

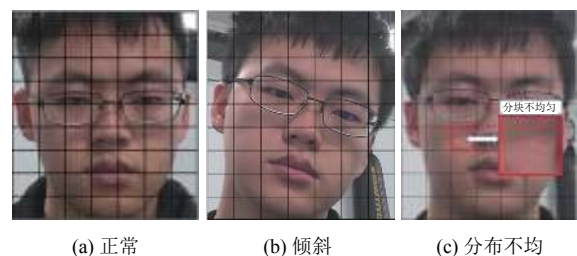


图 2 传统 ROI 分割效果演示

超像素分割算法是通过像素间共享特征 (如亮度、纹理、阴影、轮廓等) 将图像分割为包含特定图像结构的子区域^[15]. 该算法可用于对面部区域进行分割, 但

是传统的超像素分割仅对图像进行处理,未考虑连续帧之间的分块个数及位置的连续性.故针对心率检测场景对传统方法进行改进提出人脸自适应超像素分割算法.

方法首先针对视频图像中人脸位置运动做出位置自适应处理,再结合简单的线性迭代聚类 (simple linear iterative clustering, SLIC)^[16] 实现连续帧间人脸区域分块的一致性.能够很好地应对视频中人脸位置移动、远近距离变化和平面内旋转等情况下人脸分块的稳定性,保证子块 BVP 信号提取区域对应人脸位置的稳定性和连续性.具体步骤为:

(1) 视频人脸获取.根据第 1.1 节获取的视频人脸位置对原始图形进行裁切获取原始人脸区域图像帧序列 F .

(2) 自适应处理.倾斜校正:如图 3 所示根据获取的图像帧序列 F 以及第 1.1 节获取的人脸特征点位置 C 使用戚刚^[17] 的倾斜校正方法校正视频中旋转的人脸获得校正人脸序列 F_1 .



图 3 人脸倾斜校正

人脸尺寸一致处理:由于头部运动会影响图像中面部像素区域大小^[18],若采用固定的分割个数可能会使得分割子块像素区域过小,导致 BVP 信号提取质量的急速下降^[19].因此通过设置初始化超像素分块尺寸,根据视频初始帧人脸像素尺寸自适应决定分块个数.由于人脸五官位置在不同尺寸图片中比例保持相对稳定,因此根据面部特征点 C 对倾斜校正序列 F_1 进行对比缩放,获得尺寸一致的人脸图像序列 F_2 ,以确保分块个数一致.

(3) 初始化种子点.设序列 F_2 中图像像素个数为 K ,为了确保分块 BVP 提取质量,设置初始超像素大小为 $S = 50 \times 50$ 的矩形区域,则超像素个数 $M = K/S$,超像素种子点间的距离 $L = \text{sqrt}(K/S)$.

(4) 相似度衡量.针对序列 F_2 图像帧中的每个像素,分别计算与距离最近的种子点间的相似程度,为该像素点赋予最相似种子点标签.相似度衡量公式如下所示:

$$\begin{cases} d_{lab} = \sqrt{(l_k - l_i)^2 + (a_k - a_i)^2 + (b_k - b_i)^2} \\ d_{xy} = \sqrt{(x_k - x_i)^2 + (y_k - y_i)^2} \\ D_i = d_{lab} + \frac{e}{L} d_{xy} \end{cases} \quad (2)$$

其中, d_{lab} 为像素点之间的颜色差异, d_{xy} 是像素点间的空间距离. D_i 用来衡量两像素间的相似程度; L 为种子点的间距, e 为平衡参数,用来反应颜色差异与空间距离在相似度衡量中的比重. D_i 越大,表明两像素之间相似度越高.

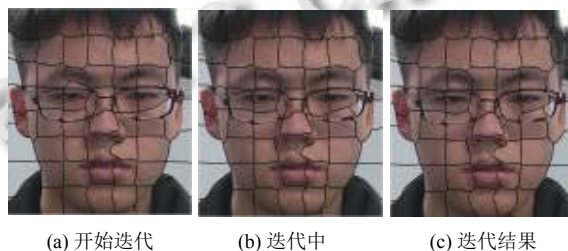


图 4 自适应超像素分割迭代过程

自适应超像素分割迭代过程如图 4,通过不断迭代上述过程直至达到收敛,分块之间充分考虑到面部像素之间距离和颜色特征等联系,并且经过了面部倾斜校正和一致性处理,可以确保连续帧之间分块位置的一致性,获得最终自适应分割结果标签 $Label$,范围 $1-M$.

1.3 分块色度脉搏矩阵提取

对获取的连续帧人脸自适应分块结果,根据 $Label$ 标记区域进行 BVP 信号提取.研究表明^[20],HSV 颜色空间 H 通道色度特征对面部运动具有良好的鲁棒性,可以抵抗面部运动光线反射角度变化造成的皮肤亮度和颜色变化.因此采用 H 通道进行色度脉搏提取.首先根据分块 $Label$ 结果,分别计算 M 个 ROI 分割子区域像素均值,记为 $b^m(N)$ 其中 m 表示第 m 个分块区域, N 表示的视频总帧数.随后将计算得到的 M 个区域的像素均值矩阵 B 通过颜色空间转换到 H 通道获得 H 通道色度 BVP 信号矩阵,记为:

$$H = [h^1(N), h^2(N), h^3(N), \dots, h^m(N)]^T \quad (3)$$

其中, $h^m(N)$ 表示第 m 个分块区域色度脉搏提取信号.

H 通道颜色转换方法在以下公式中进行描述.使用以下公式对 RGB 值进行标准化与颜色空间转换:

$$R' = R/255, G' = G/255, B' = B/255 \quad (4)$$

$$C_{\max} = \max(R', G', B'), C_{\min} = \min(R', G', B') \quad (5)$$

$$\Delta = C_{\max} - C_{\min} \quad (6)$$

$$H = \begin{cases} 0, & \text{if } \Delta = 0 \\ 60^\circ \times ((G' - B')/\Delta), & \text{if } C_{\max} = R' \\ 60^\circ \times ((B' - R')/\Delta + 2), & \text{if } C_{\max} = G' \\ 60^\circ \times ((R' - G')/\Delta + 4), & \text{if } C_{\max} = B' \end{cases} \quad (7)$$

其中, R', G', B' 为RGB三通道的归一化值, 范围 0-1; C_{\max} 和 C_{\min} 分别为归一化RGB通道的最大值和最小值, 以及归一化RGB通道的差值 Δ ; H 即为 HSV 颜色空间的色度分量.

1.4 脉搏信号筛选与心率估计

获取到面部 M 个子块色度 BVP 信号后, 对各子块的原始脉搏信号质量进行评估与优选. 在进行脉搏波质量评估的特征选取时, 必须要符合以下准则. 首先, 因为真实场景下无心率参考真值, 质量评估指标不依赖真值对比; 其次, 特征数值大小需体现信号质量优劣; 最后, 特征值应方便从简短的波形获得, 确保特征及时更新实现 ROI 优选的时效性. 考虑到上述条件, 因此选取幅度、频域和非线性 3 个方面指标综合评判脉搏波质量.

(1) 幅度特征: 分块区域亮度变化标准差 ($light\ std$), 表示分块内像素亮度与均值之间的偏离程度, 可以体现分块内亮度变化程度. 公式如式 (8):

$$light(m) = \sqrt{\frac{\sum_{i=1}^N (lvalue_i^m - \mu)^2}{N}} \quad (8)$$

其中, $light(m)$ 表示的是第 m 个分块的亮度变化程度指标, $lvalue_i^m$ 代表第 m 分块中第 i 帧的亮度值, μ 代表分块中所有帧的亮度均值, N 为视频帧总数.

(2) 频域特征: 分块脉搏波信噪比 (signal noise ratio, SNR), 表示脉搏信号平均功率与噪声平均功率的比值. 可以反应心率信号能量占正常心率区间内的总能量大小, 公式如式 (9):

$$SNR(m) = \frac{S(f_{HR} \pm 0.01)}{\sum_{f=0}^{f_s/2} S(f) - S(f_{HR})} \quad (9)$$

其中, $SNR(m)$ 表示的是第 m 个分块的信噪比指标, f_{HR} 表示信号频谱峰值频率, f_s 表示信号采样率, 函数 S 表示频谱密度函数.

(3) 非线性特征: 分块内样本熵. 熵值作为一种非线性特征, 可以体现随机变量不确定性的程度, 反应信号的规则与不规则性. 近来已经被越来越多人应用于对生理信号分析上^[21], 样本熵 (sample entropy, $SampEn$) 对心率信号的高斯噪声尤其敏感, 故选取作为波形质

量评估特征.

以上 3 种指标由于评价尺度不一无法直接对比数值大小使用, 现分别对指标进行有效性归一化处理, 公式如下:

$$Light_{norm}(m) = \frac{\max(light) - light(m)}{\max(light) - \min(light)} \quad (10)$$

$$SNR_{norm}(m) = \frac{SNR(m) - \min(SNR)}{\max(SNR) - \min(SNR)} \quad (11)$$

$$SampEn_{norm}(m) = \frac{\max(SampEn) - SampEn(m)}{\max(SampEn) - \min(SampEn)} \quad (12)$$

其中, \min 和 \max 分别为各分块区域中对应指标的最小值和最大值. $Light_{norm}(m)$ 、 $SNR_{norm}(m)$ 、 $SampEn_{norm}(m)$ 分别为第 m 块有效性归一化后的指标, 数值越大表明分块内 BVP 信号质量越好.

综合有效性归一化后的指标, 得到分数 $Score$ 值, 公式如下:

$$Score(m) = SNR_{norm}(m) + Light_{norm}(m) + SampEn_{norm}(m) \quad (13)$$

其中, $Score(m)$ 为第 m 块分区 BVP 质量最终分数, 分数越高说明综合质量越好.

如图 5 所示, 对 $Score(m)$ 数值排序后选取最佳分块区域对应 BVP 信号, 对 BVP 信号进行带通滤波, 选取在 0.8-3.5 Hz 区间最大功率对应的信号频率 f_{HR} 得到受试者心率. 其中对应关系如下:

$$HR = f_{HR} \times 60 \quad (14)$$

2 实验与结果

2.1 实验设备及数据

本文实验和数据采集是在白色 LED 照明环境中进行. 使用普通彩色摄像头 (罗技 C920) 连接微软 Surface 笔记本 (Windows 10, 16 GB 内存, 512 GB 硬盘) 进行面部视频采集, 同步使用指夹式脉搏传感器 (紫金电子 PWS-20D) 采集左手食指指尖 ppg 信号. 摄像头设置采集分辨率为 1080 P、帧率 30 fps, 每次采集受试者 30 s 的人脸视频, 并以 RGB 颜色空间保存成 mp4 格式. 脉搏传感器设置采样率为 200 Hz, 采样精度为 10 位作为脉搏真值保存. 共有 20 名受试者参与本次实验, 其中男性 12 名、女性 8 名, 年龄分别为 20-43 岁、19-35 岁. 数据采集方式和采集界面如图 6 所示, 受试者需正对摄像头并且保持距离 40-60 cm 左右而坐, 同时采集面部视频与指尖 ppg 信号.

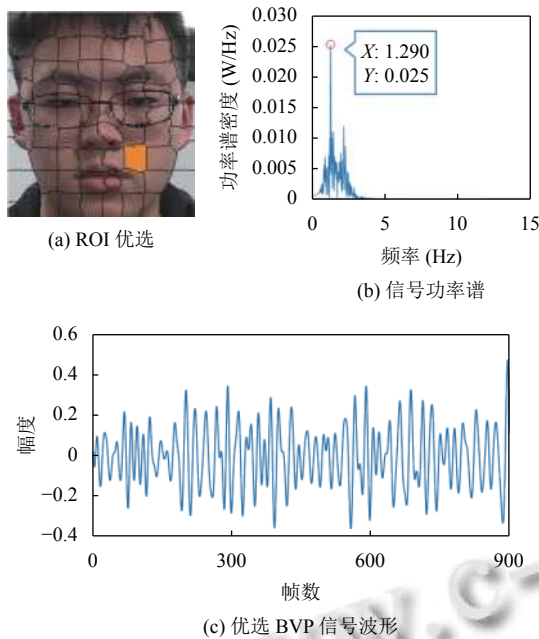


图5 优选 BVP 信号及功率谱密度

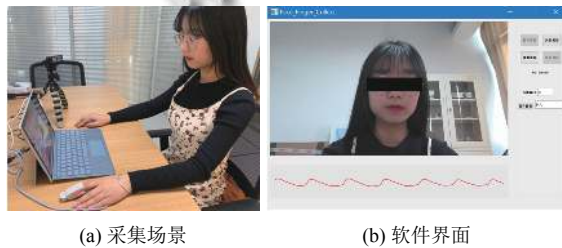


图6 采集场景演示

2.2 实验结果

本文针对算法的效果复现了文献 [10–12] 中的 3 种心率检测算法作为对比, 主要从实际场景下对算法进行评估. 其中采用了 5 种误差指标对准确性进行比较. 第 1 个指标为平均绝对误差, 代表心率检测值与真实值之间误差的绝对平均值, 记为 Me ; 第 2 个指标为误差标准差 SDe , 用来反映误差之间相互离散的程度; 第 3 个指标为误差均方根记作 $RMSE$, 用来评估面部测量值和指尖 ppg 参考值之间的偏离程度; 第 4 个指标为心率平均准确率 HR_{ac} , 主要是反映心率检测结果的平均准确性, 具体公式如式 (15):

$$HR_{ac} = 1 - \frac{1}{T} \sum_{t=1}^T \frac{|HR_{test}^t - HR_{ref}^t|}{HR_{ref}^t} \quad (15)$$

其中, T 为实验样本总数, HR_{test}^t 为第 t 组面部心率检测的心率测量值, HR_{ref}^t 为第 t 组脉搏传感器心率测量的参考值; 第 5 个参数是皮尔森相关系数记作 r , 是用来反

应心率测量值和真实值之间线性相关性的统计量. r 值通常范围在 -1 至 1 之间, 且绝对值越大说明两变量之间的相关程度越高.

为了全面验证本文方法在实际场景下的性能, 设计了 3 个典型场景验证算法的准确性. 首先分析受试者保持面部静止场景, 其次是受试者在镜头内存在头部运动和表情变化等动态干扰场景. 随后使用不同光照场景下视频验证算法的稳定性, 最后对受试者进行 10 min 的连续心率检测.

2.2.1 静止状态下心率检测

在本次实验中, 为检验静止状态下本文算法心率检测的准确性, 受试者需端坐在采集摄像头前 50 cm 左右位置, 保持环境光照稳定及受试者头部静止采集 2 组人脸视频. 每组视频时长为 30 s, 实验样本总数为 40 组. 实验结果如表 1 所示, 其中单位为 bpm (beat per minute) 即心脏每分钟跳动次数.

表1 静态场景下算法准确性分析

方法	Me (bpm)	SDe (bpm)	$RMSE$ (bpm)	HR_{ac} (bpm)	r
文献[10]	1.52	1.86	1.92	97.56	0.962
文献[11]	1.32	1.42	1.48	98.03	0.977
文献[12]	0.92	1.32	1.12	98.26	0.988
本文	0.74	1.17	1.13	98.56	0.991

注: 加粗字体显示为最优结果

如表 1 所示, 展示了 4 种方法在受试者保持静止时的准确性情况. 可以发现 4 种方法都可以很好的应对静态场景下的心率检测, 其中, Me 均低于 1.6 bpm, SDe 和 $RMSE$ 低于 2 bpm, HR_{ac} 可以达到 98%, r 接近 1. 通过观察可以发现, 本文方法中的 Me 最低、 r 最高, 分析原因可能是: 在视频采集中, 受试者虽然身体保持静止但是由于呼吸动作的存在会引入轻微的起伏干扰, 其次眨眼等动作的存在也会导致面部某些区域存在轻微的干扰. 而本文采用的人脸自适应超像素分割和综合指标 ROI 优选方法可以从面部区域中自动寻找最佳区域进行心率检测, 表现出来更好的准确性.

2.2.2 动态干扰场景下心率检测

由于检测过程中, 头部的晃动和表情变化等, 会使得镜头内人脸位置和背景发生变化造成皮肤与光源反射角度的变化, 最终导致 BVP 信号的失真进一步影响心率检测的准确性. 因此本文特地设计一组头部移动和表情变化干扰场景. 在这次的实验中, 运动干扰包括在镜头内以一定的速率水平或竖直移动头部以及进行

平面内旋转等操作;表情变化为视频采集时受试者随机朗读 1-9 号数字. 在本组实验中, 每名受试者采集 3 组数据, 样本总数为 60 组. 具体的实验结果如表 2 所示. 为了直观展示实验结果, 实验数据通过散点图和 Bland-Altman 图进行展示, 如图 7 所示.

由表 2 结果可以得知, 由于运动干扰和表情变化的引入导致 4 种方法的准确性均表现出不同程度的降低. 经过对比发现, 文献 [10-11] 方法的性能明显偏低, 其中, Me 、 SDe 、 $RMSE$ 均大于 7 bpm, HR_{ac} 下降至 90% 以下, 相比较静止状态下最差下降幅度高达 11.2%, 与心率的相关性明显降低; 文献 [12] 的效果虽然有稍许提升, 但是 Me 依然大于 5 bpm 超过实际场景下心率

检测的误差上限. 而本文的方法, 在 4 种方法中受到运动干扰和表情变化的干扰幅度最小, 尤其是心率准确性指标 HR_{ac} 达到了 95.34% 相关系数 r 高达 0.956. 上述数据表明本文的人脸自适应超像素分割方法可以有效地从运动的人脸中提取信号最佳的区域进行心率估计, 具有良好的性能表现.

表 2 动态干扰场景下算法准确性分析

方法	Me (bpm)	SDe (bpm)	$RMSE$ (bpm)	HR_{ac} (bpm)	r
文献[10]	8.73	8.81	8.92	86.67	0.723
文献[11]	7.67	7.24	7.18	89.63	0.887
文献[12]	5.23	5.42	5.49	91.41	0.902
本文	3.74	3.52	3.69	95.34	0.956

注: 加粗字体显示为最优结果

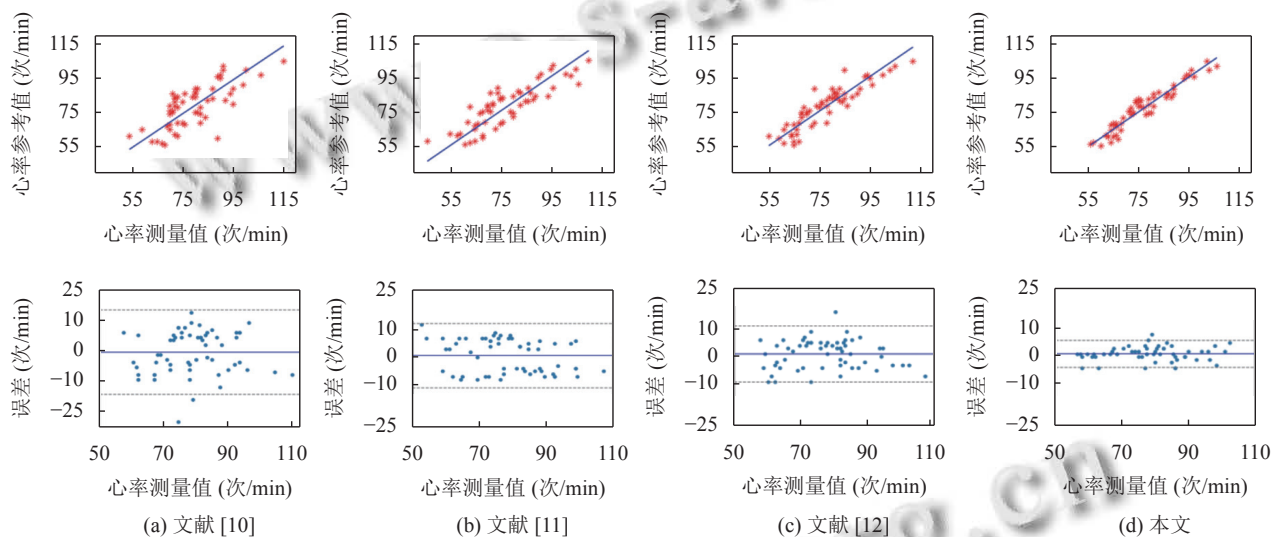


图 7 运动结果散点和 Bland-Altman

通过对比图 7 第 1 行中 4 种方法的散点图可以看出文献 [10-12] 的方法存在较多的数据误差较大明显偏离线性回归线, 而本文方法的数据集中在线性回归线附近, 且线性回归线的斜率接近于 1, 表明本文方法的测量值与心率参考值的相关最强.

图 7 第 2 行是运动结果的 Bland-Altman 图, 其在医学领域经常被用来分析新方法与标准方法测量结果的一致程度. 虚线内的区域是 95% 置信区间, 中心实线代表心率测量值与参考值之间的相对平均误差. 从图中可知, 本文方法的虚区域远远窄于其他 3 种方法且中心实线最接近于 0, 表明本文方法较其他 3 种方法的测量值与标准参考值有着更好的一致性. 进一步分析原因可能如下:

1) 文献 [10] 采用的自适应滤波算法, 能够较好地

处理高斯白噪声干扰. 但由于使用的是整张人脸作为 ROI, 而运动时人脸引起的尖锐干扰并不能被自适应滤波器很好地滤除. 因此该方法无法提取到纯净的 BVP 信号, 进而无法在运动干扰条件下进行准确的心率估计.

2) 文献 [11] 采用的多区域分析方法, 主要选取额头和两颊 3 个独立区域作为心率检测的 ROI. 该方法从一定程度上可以从运动干扰中提取优质 BVP 信号进行心率估计. 但是由于选取的是固定区域, 固定区域无法确保不受到偶然干扰影响, 并且额头部位在女性中容易受到刘海覆盖干扰, 影响实际应用场景下的表现.

3) 文献 [12] 也是通过将面部进行分割提取不同区块的色度特征信号进行心率估计. 但是, 首先其对人脸区域仅进行简单的矩形分割没有考虑区域内像素间关系; 其次由于固定了分割矩形块大小为 20×20 的像素

区域,并没有对不同帧内人脸图像尺寸进行一致性处理,无法确保运动中分块个数的稳定;最后由于没有对人脸旋转进行校正,难以保证倾斜状态下分块位置的连续性。

4) 本文方法与文献 [12] 采取的方法类似,都是通过将脸部区域分割成不同的子块通过对子块信号的分析进行心率估计。不同的是,本文在简单矩形分割的基础上引入了自适应超像素聚类算法,确保区域内像素间的特征相关性。同时本文算法针对人脸视频分割后分块位置的一致性和连续性进行了人脸校正操作。一方面提升原始 BVP 信号的提取质量,另一方面通过综合指标优选 ROI 方法进一步提升了算法在实际场景中的抗干扰能力。

2.2.3 不同光照场景下心率检测

考虑到真实场景中视频采集环境的复杂性,位于不同光照条件下心率检测的准确性会受到影响。因此本次实验分别采集了受试者静止状态 4 种光照场景下的视频进行对比实验,具体为面向光源(正)、侧面对光源(侧)、逆向光源(逆)、以及暗光条件(暗)。在本次实验中,每名受试者均在特定场景下采集 2 组视频,视频时长 30 s,样本总数为 160 组。实验的结果如图 8 所示。

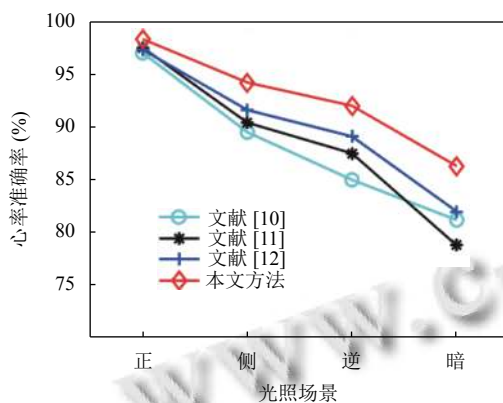


图 8 不同光照场景下的心率准确性变化

由图 8 可知,在正面向光源的场景中 4 种方法的准确率均可以达到 97% 以上,表明在光照条件良好的状态下 4 种方法均可以准确地进行心率检测。但是针对侧光和逆光这类光照分布不均匀的场景,文献 [10-12] 方法的准确性均有较大幅度的下降,其中文献 [10] 更是分别下降了 7.8%、12.4%。本文方法展示出较高的准确性优势,其中侧光准确率达到 94.3%,逆光准确率达到 92.1%,对比最差方法提升了 5.2% 与 8.2%,但是

针对暗光环境 4 种方法均表现不佳。分析原因为侧光、逆光时光照分布不均匀,本文方法可以根据光照分布的情况从面部选取像素特征相近区块进行分析并挑选质量最佳子块进行心率估计。面对暗光这样的光照场景,由于射入皮肤的光照强度变低导致面部反射的光强变低,无法从 ROI 挑选上对光照进行增强,故以上方法在暗光场景中均表现欠佳。

2.2.4 连续时间心率检测

长时间连续的心率检测,可以用来评估用户的心率变化趋势以及对一些心血管疾病的产生进行早期预警。为了最终评价本算法的准确性,设计一组连续 10 min 的心率检测实验。受试者正坐在摄像机前 50 cm 的位置,保持 10 min 内面部一直处于视频采集画面内,期间不限制受试者头部移动和表情等变化。并同时使用脉搏血传感器每 10 s 记录一次实时心率值做对比分析,实验结果如图 9 所示。

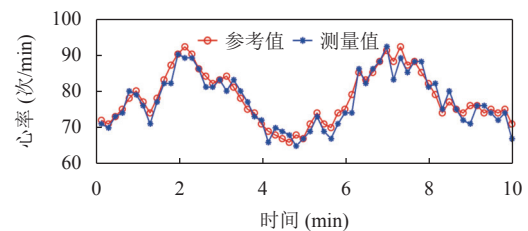


图 9 心率变化折线图

在图 9 中的红线和蓝线分别代表指尖 PPG 传感器采集的心率参考值和面部心率检测的测量值。从整体上看,蓝线的变化趋势随着红线有着相似的起伏波动,说明测量值和参考值之间有着良好的一致性 ($r=0.9558$)。综上所述,本文的算法可以有效地进行心率长期检测,提升真实场景下算法的鲁棒性。

3 结论与展望

近年来由于视频图像处理技术的快速发展,基于视频的非接触心率检测技术发展迅速,但在实际场景应用仍有诸多挑战。本文针对面部运动和光照变化干扰问题,提出一种人脸自适应分割和多区域多指标综合评价算法,提升脉搏波质量,提高心率检测可靠性,增强实际场景鲁棒性。主要由以下几部分组成:面部区域自适应超像素分割为互不重叠的子块;提取色度特征构建脉搏信号矩阵;根据综合指标评估各分块质量,最后对最优子块进行心率估计。实验结果表明,相比传

统 ROI 分割方法, 本文算法分割后脉搏矩阵信号质量有显著提升. 在面部运动和光照不均匀场景下心率检测准确性和稳定性均有明显提升. 然而, 在暗光和环境光照变化剧烈等场景下, 本文心率检测算法性能出现比较明显的下降. 在以后的工作中, 将对环境光照变化造成的干扰问题进行研究, 进一步提升复杂光照场景下的心率检测鲁棒性.

参考文献

- 1 吴学思. 心率在心血管疾病中的意义. 中华内科杂志, 2006, 45(7): 601-602. [doi: 10.3760/j.issn:0578-1426.2006.07.033]
- 2 黄宛. 黄宛临床心电图学. 中国医刊, 2009, 44(11): 80.
- 3 Verkruyse W, Svaasand LO, Nelson JS. Remote plethysmographic imaging using ambient light. Optics Express, 2008, 16(26): 21434-21445. [doi: 10.1364/OE.16.021434]
- 4 Poh MZ, McDuff DJ, Picard RW. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. Optics Express, 2010, 18(10): 10762-10774. [doi: 10.1364/OE.18.010762]
- 5 de Haan G, Jeanne V. Robust pulse rate from chrominance-based rPPG. IEEE Transactions on Biomedical Engineering, 2013, 60(10): 2878-2886. [doi: 10.1109/TBME.2013.2266196]
- 6 Wang WJ, den Brinker AC, Stuijk S, et al. Color-distortion filtering for remote photoplethysmography. Proceedings of the 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017). Washington DC: IEEE, 2017. 71-78.
- 7 Tasli HE, Gudi A, den Uyl M. Remote PPG based vital sign measurement using adaptive facial regions. Proceedings of 2014 IEEE International Conference on Image Processing. Paris: IEEE, 2015. 1410-1414.
- 8 Kumar M, Veeraraghavan A, Sabharwal A. DistancePPG: Robust non-contact vital signs monitoring using a camera. Biomedical Optics Express, 2015, 6(5): 1565-1588. [doi: 10.1364/BOE.6.001565]
- 9 Feng LT, Po LM, Xu XY, et al. Dynamic ROI based on K-means for remote photoplethysmography. Proceedings of 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). South Brisbane: IEEE, 2015. 1310-1314.
- 10 Li XB, Chen J, Zhao GY, et al. Remote heart rate measurement from face videos under realistic situations. Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE, 2014. 4264-4271.
- 11 Favilla R, Zuccalà VC, Coppini G. Heart rate and heart rate variability from single-channel video and ICA integration of multiple signals. IEEE Journal of Biomedical and Health Informatics, 2019, 23(6): 2398-2408. [doi: 10.1109/JBHI.2018.2880097]
- 12 周双, 杨学志, 金兢, 等. 采用自适应信号恢复算法的非接触式心率检测. 中国图象图形学报, 2019, 24(10): 1670-1682. [doi: 10.11834/jig.190013]
- 13 Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001. Kauai: IEEE, 2001. I.
- 14 Asthana A, Zafeiriou S, Cheng SY, et al. Robust discriminative response map fitting with constrained local models. Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland: IEEE, 2013. 3444-3451.
- 15 Stutz D, Hermans A, Leibe B. Superpixels: An evaluation of the state-of-the-art. Computer Vision and Image Understanding, 2018, 166: 1-27.
- 16 Achanta R, Shaji A, Smith K, et al. SLIC superpixels compared to state-of-the-art superpixel methods. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(11): 2274-2282. [doi: 10.1109/TPAMI.2012.120]
- 17 戚刚. 基于人脸视频处理的心率检测方法研究 [博士学位论文]. 合肥: 合肥工业大学, 2017.
- 18 牛雪松, 韩璇, 山世光. 基于 rPPG 的生理指标测量方法综述. 中国图象图形学报, 2020, 25(11): 2321-2336. [doi: 10.11834/jig.200341]
- 19 Rapczynski M, Werner P, Saxen F, et al. How the region of interest impacts contact free heart rate estimation algorithms. 2018 25th IEEE International Conference on Image Processing (ICIP). IEEE, 2018. 2027-2031.
- 20 Ernst H, Scherpf M, Malberg H, et al. Color Spaces and Regions of Interest in Camera Based Heart Rate Estimation. 2020 11th Conference of the European Study Group on Cardiovascular Oscillations (ESGCO). IEEE, 2020. 1-2.
- 21 王帅, 赵钟瑶, 张翔宇, 等. 穿戴式心电信号质量的三分类评估方法. 中国生物医学工程学报, 2020, 39(5): 550-556. [doi: 10.3969/j.issn.0258-8021.2020.05.005]

(校对责编: 牛欣悦)