

韩轶, 张弛朴, 王洪波. 基于人脸视频的非接触式心率测量研究[J]. 智能计算机与应用, 2026, 16(1): 1-7. DOI: 10.20169/j.issn.2095-2163.25081804

基于人脸视频的非接触式心率测量研究

韩轶, 张弛朴, 王洪波

(复旦大学工程与应用技术研究院, 上海 200433)

摘要: 传统接触式心率监测易造成皮肤不适与感染风险, 而非接触式技术存在成本高、环境适应性差的问题。现有方法在复杂光照或运动场景下存在精度限制, 难以满足连续无感监测需求。本文针对视频心率监测中的关键信号, 进行了视觉感知研究, 将时频分析与分类器检测 CHROM 算法相结合, 并通过多数据集实验验证了算法的有效性。目的是能够准确提取心率信号, 并后续基于该技术实现非接触式连续监测, 提高健康监护的质量稳定性和自动化程度, 为智能医疗背景下的无感生理参数监测提供新的技术路径。

关键词: 视频心率监测; 时频分析; CHROM 算法; 无感监测; 多数据库验证

中图分类号: R318.51

文献标志码: A

文章编号: 2095-2163(2026)01-0001-07

Research on non-contact heart rate measurement based on facial videos

HAN Zhan, ZHANG Chipu, WANG Hongbo

(Academy for Engineering and Technology, Fudan University, Shanghai 200433, China)

Abstract: Traditional contact-based heart rate monitoring carries risks of skin discomfort and infection, while non-contact technologies suffer from high costs and poor environmental adaptability. Existing methods face accuracy limitations in complex lighting or motion scenarios, failing to meet the demand for continuous, unobtrusive monitoring. This study conducts visual perception research on key signals in video-based heart rate monitoring, combining time-frequency analysis with the classifier-based CHROM algorithm. The effectiveness of the algorithm is validated through multi-dataset experiments. The research aims to accurately extract heart rate signals and subsequently enable continuous non-contact monitoring through this technology. This approach enhances the stability of healthcare monitoring quality and automation levels, providing a novel technical pathway for unobtrusive physiological parameter monitoring in intelligent healthcare contexts.

Key words: video-based HR monitoring; time-frequency analysis; CHROM algorithm; unobtrusive monitoring; multi-dataset validation

0 引言

传统接触式的监测方法, 例如心电图(ECG)^[1]与脉搏光量计(PPG)^[2], 要求直接将电极或感应器贴于皮肤上, 这一操作可能给老年人带来皮肤损伤和不适感, 且会提升交叉感染的可能性。与此相反, 非接触式的监测技术^[3], 比如雷达或热像仪, 尽管避免了接触的不便, 却因高昂的设备成本、复杂的安装过程和环境因素的影响而存在局限。相比而言通过视频监测心率, 利用现有的摄像头进行无侵入的实时心率测量, 对于失能老人而言更加方便友好, 降

低了使用难度, 实现了长期且连续的健康监测, 因此非常适合应用于护理场景。

基于视频的心率监测技术, 又称远程光电容积描记法(remote Photo Plethysmo Graphy, rPPG)^[4], 是一种非接触性的心率评估手段。该技术通过分析视频中皮肤颜色细微的波动, 间接推算出心率。每次心脏跳动时, 血液在面部血管内的循环导致皮肤色调出现微小变化。通过捕获并解析这些动态细节, rPPG 技术能从视频中识别与心跳节奏相关的关键信号, 进而估算出心率数值。利用普通摄像头实现心率信息的无感化提取, 具有非侵入性、低成本、

基金项目: 国家重点研发计划(2022YFB4703300)。

作者简介: 韩轶(1999—), 男, 硕士研究生, 主要研究方向: 护理机器人; 张弛朴(2002—), 男, 硕士研究生, 主要研究方向: 护理机器人。

通信作者: 王洪波(1956—), 男, 博士, 教授, 博士生导师, 主要研究方向: 医疗机器人。Email: Wanghongbo@fudan.edu.cn。

收稿日期: 2025-08-18

易部署等显著优势^[5]。

1 基本理论

对于人体表皮层和角质层而言,在将皮肤上面照射的光予以吸收后,就可以为人体提供必要的光学屏障。综观整个过程可知,在真皮层位置则表现出散射光特点。基于 rPPG 原理,构建了人体皮肤反射模型如图 1 所示^[6]。

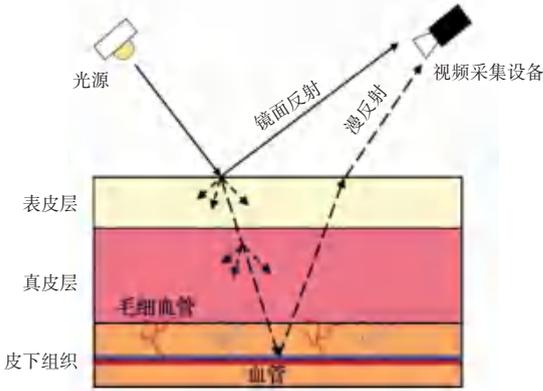


图1 人体皮肤反射模型

Fig. 1 Human skin reflection model

当人体表皮组织暴露于稳定光源照射时,视频采集装置不仅接收来自皮肤表面的反射光线,其余入射光会经历 2 种不同的物理过程:部分被表皮及皮下组织吸收转化,剩余部分通过次漫反射传播至体外空间。值得关注的是,正是这些经过组织反射出的光信号承载着重要的生命体征特征,通过视频采集设备捕捉这些漫反射出的光,即可从中解析出包含心率等生理参数的 rPPG 特征波形。基于视频的非接触式生理监测系统通过检测表皮层色度波动实现生命体征量化分析,皮肤反射模型整合了色度波动与运动伪影的数学表征。在假设照明源保持恒定光谱特性但存在强度波动的条件下,定义视频序列中第 K 个皮肤像素点的 RGB 值为:

$$C_k(t) = I(t) \cdot (v_s(t) + v_d(t)) + v_n(t) \quad (1)$$

其中, $C_k(t)$ 对应于 RGB 三通道的数值集合; $I(t)$ 表示光照强度系数,其数值受制于照明条件变化、表皮组织光学特性及距离改变等多重因素; $v_s(t)$ 和 $v_d(t)$ 分别反映镜面反射分量(表皮层直接反射)与漫反射分量(真皮层吸收散射效应)的叠加作用; $v_n(t)$ 表示图像传感器的随机噪声干扰。上述 3 个核心参数 $I(t)$ 、 $v_s(t)$ 和 $v_d(t)$ 均可通过线性分解方法区分为稳态分量与动态时变成分。对此可表示为:

$$v_d(t) = u_d \cdot d_0 + u_p \cdot p(t) \quad (2)$$

$$v_s(t) = u_s \cdot (s_0 + s(t)) \quad (3)$$

$$I(t) = I_0 \cdot (1 + i(t)) \quad (4)$$

其中, u_d 表示表皮组织光学特性的单位颜色向量; d_0 表示静态反射强度参数; u_p 表示由血红蛋白与表皮黑色素层光吸收效应引发的周期性强度波动; $p(t)$ 表示血容量脉冲(BVP)波形特征; u_s 表示入射光源频谱特性的单位颜色向量; s_0 和 $s(t)$ 分别表示镜面反射效应的稳态分量与动态波动分量; $s(t)$ 涵盖所有非生物性干扰因素(包括面部肌群运动、头部位姿变化、环境光频闪等); I_0 表示光照条件的静态强度参数; $I_0 \cdot i(t)$ 对应成像系统感知的光强动态变化。需特别指出,镜面反射动态分量 $s(t)$ 与环境光强 $I_0 \cdot i(t)$ 中生理与非生理信号的耦合作用呈现非线性关联特性。通过整合镜面反射与漫反射的稳态分量,可用来表征静态皮肤反射。由此得到:

$$u_c \cdot c_0 = u_s \cdot s_0 + u_d \cdot d_0 \quad (5)$$

其中, u_c 表示组织反射特性的单位颜色向量, c_0 表示综合反射强度参数。

将经过线性分解处理的 $I(t)$ 、 $v_s(t)$ 、 $v_d(t)$ 参数及表征静态反射的合成项 $u_c \cdot c_0$ 代入 $C_k(t)$ 原方程,推导得出:

$$C_k(t) = I_0 \cdot [1 + i(t)] \cdot [u_c \cdot c_0 + u_s \cdot (s(t) + u_p \cdot p(t)) + v_n(t)] \quad (6)$$

在基于视频的生理参数检测技术中,核心任务在于从复合信号 $C_k(t)$ 中分离出纯净的血容量脉冲(BVP)波形特征 $p(t)$ 。

2 非接触式心率测量算法

2.1 测量流程

本研究介绍了一种利用人脸视频估计人体心率的技术。该方法采取非接触式操作,基本流程如图 2 所示。对此做阐释分述如下。

(1) 视频帧序列的获取:整个过程由摄像头捕捉包含人脸部位的视频数据,并按帧逐一处理。

(2) ROI 定位:核心任务在于提取人脸视频中的关键数据。采用关键点检测算法来搜集这类核心信息,对人脸的关键部位进行逐帧分析,基于这些关键点的数据,进一步划定感兴趣区域(ROI),并获取在此区域内每帧的详细信息。

(3) rPPG 信号获取:依据 ROI 内收集的数据,计算了目标区域内各通道 RGB 像素强度的均值,从而获得了初始 rPPG 信号。

(4) rPPG 信号预处理:在收集到原始 rPPG 信

号后,对其进行了相应流程处理,旨在从信号中筛选并提取包含心率信息的核心部分,这些核心部分呈现周期性波动。本研究在预处理环节,对 rPPG 信号进行了归一化处理,并使用带通滤波去除噪声。

(5) 心率提取:鉴于 rPPG 信号具备明显节律特征,波峰定位算法与频谱解析技术作为经典心率检测手段被广泛采用。然而实际应用环境中,rPPG 信号易受运动伪影及环境干扰影响,本研究构建时频域联合表征模型,将频谱特性与周期规律进行耦合分析,进而生成具有生物特征辨识度的多维特征矩阵。该模型通过深度挖掘时域波形与频域分量的关联性,实现心动周期特征的鲁棒性提取,最终完成精准的心率参数量化输出。

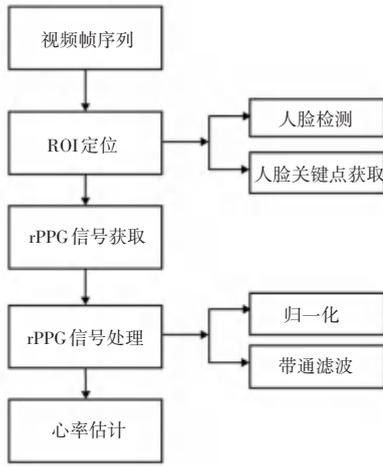


图2 非接触式心率测量基本流程

Fig. 2 Basic process of non-contact heart rate measurement

2.2 rPPG 信号获取

基于皮肤模型的 rPPG 信号提取算法摒弃了对光源色度参数的预设限制,转而运用多源信号中 RGB 三原色权重分布差异的物理先验知识实施信号解混。该策略有效解除了表皮像素信号对反射色度成分(即直流分量)的依赖性,涵盖外部光源色温属性与皮肤组织固有光谱特征。本研究通过时域归一化算法实现信号优化,将表皮色度信号除以其时域信号的平均值,通过数学运算将原始信号分解为动态变化的交流分量与相对恒定的直流分量。该处理选择性地调节直流分量,同时完整保留交流成分的波形特征。其中,时域均值作为基准参数,通过滑动窗口算法动态更新,有效表征监测时段内信号的稳态特征。用如下公式进行定义:

$$\overline{C_n(t)} \approx u_c \cdot I_0 \cdot c_0 \quad (7)$$

该表达式可唯一确定对角归一化变换矩阵 N :

$$N \cdot \overline{C(t)} = N \cdot u_c \cdot I_0 \cdot c_0 = 1 \quad (8)$$

矩阵 N 对原始信号 $C(t)$ 实施时域上的归一化处理:

$$C_n(t) = N \cdot C(t) = N \cdot u_c \cdot I_0 \cdot c_0 + N \cdot u_c \cdot I_0 \cdot c_0 \cdot i(t) + N \cdot u_s \cdot I_0 \cdot s(t) + N \cdot u_p \cdot I_0 \cdot c_0 \cdot p(t) = \mathbf{1} \cdot (1 + i(t)) + N \cdot u_s \cdot I_0 \cdot s(t) + N \cdot u_p \cdot I_0 \cdot c_0 \cdot p(t) \quad (9)$$

通过公式推导可得出以下关键结论:首先,光源强度波动 $\mathbf{1} \cdot (1 + i(t))$ 体现为沿 $\mathbf{1}$ 方向的光强变化,该方向 $\mathbf{1}$ 对应于时域归一化后表皮色度信号的矢量指向。值得注意的是,表皮色度分量通常构成信号 $C_n(t)$ 中的最大分量,即像素信号内噪声的主要来源,此类干扰通常在三通道中呈现均匀分布特性。镜面反射分量 $N \cdot u_s \cdot I_0 \cdot s(t)$ 表征了信号经缩放后沿镜面反射方向的时间波动特征。在标准白光照射条件下存在 $N \cdot u_s \cdot I_0 = N \cdot \mathbf{1} \cdot I_0$,表明其缩放尺度与肤色的倒数相关;而在非白光场景下, u_s 的取值受光源光谱特性和相机灵敏度共同影响,而 N 还额外涉及表皮光学特性(如黑色素吸收效应)的调制作用。目标心率分量 $N \cdot u_p \cdot I_0 \cdot p(t)$ 对应心脏搏动引发的时域色度波动,即算法关注的核心成分。在 RGB 空间中,该时域归一化心率信号的色度变化方向 $N \cdot u_p \cdot I_0$ 由光源光谱与相机传感器特性共同决定,但与个体肤色关联度较低。实验数据显示,在常规光照条件与相机传感器配置下, G 通道心率信号幅值最大, B, R 通道依次递减。

结合表皮反射的生理光学特性,可利用 $C_n(t)$ 去除直流分量后的色度波动信号进行心率特征提取,去除直流分量后的色度波动信号 $\widetilde{C}_n(t)$ 可以定义为:

$$\widetilde{C}_n(t) = \mathbf{1} \cdot (1 + i(t)) + N \cdot u_s \cdot I_0 \cdot s(t) + N \cdot u_p \cdot I_0 \cdot c_0 \cdot p(t) \quad (10)$$

本研究采用 CHROM 算法^[7]实现 rPPG 信号提取。该算法通过降维处理消除表皮像素信号中的镜面反射干扰,具体方法是将色度波动信号 $\widetilde{C}_n(t)$ 投影至与镜面反射变化方向正交的平面,用 M 来表示对 $\widetilde{C}_n(t)$ 进行映射的对角映射矩阵。为实现该目标,算法首先定义标准化肤色矢量 $\mathbf{1}$ 以达成自动白平衡,将标准化的肤色向量 $\mathbf{1}$ 定义为:

$$M^{-1} \cdot U_{skin} = \mathbf{1} \quad (11)$$

其中, U_{skin} 表示白光环境下的平均色度矢量, M 可以表示为:

$$M \cdot \widetilde{C}_n(t) = M \cdot \mathbf{1} \cdot i(t) + M \cdot N \cdot u_s \cdot I_0 \cdot$$

$$s(t) + \mathbf{M} \cdot \mathbf{N} \cdot u_p \cdot I_0 \cdot p(t) \quad (12)$$

该矩阵将时域归一化的表皮色度信号映射至白光条件下的标准化肤色方向,同时将镜面反射矢量 $\mathbf{N} \cdot u_s \cdot I_0$ 近似映射至白光方向:

$$\mathbf{M} \cdot \mathbf{N} \cdot u \cdot I_0 \approx \kappa \cdot \mathbf{1}$$

其中, κ 表示比例系数。随后算法将信号 $\mathbf{M} \cdot \widetilde{\mathbf{C}}_n(t)$ 投影至与 $\mathbf{1}$ 正交的平面,确保肤色校正后信号免受镜面反射波动影响,投影后得到的皮肤颜色信号 $S(t)$ 为:

$$S(t) = P_c \cdot \mathbf{M} \cdot \widetilde{\mathbf{C}}_n(t) \approx P_c \cdot \mathbf{M} \cdot \mathbf{1} \cdot i(t) + P_c \cdot \mathbf{M} \cdot \mathbf{N} \cdot u_p \cdot I_0 \cdot p(t) \quad (13)$$

上式相减可得:

$$P_c \cdot \mathbf{M} \cdot \mathbf{N} \cdot u_s \cdot I_0 \approx P_c \cdot \kappa \cdot \mathbf{1} = 0 \quad (14)$$

其中, P_c 表示初始投影矩阵,其行向量定义了时域归一化 RGB 空间的投影轴向,而 $P_c \cdot \mathbf{M}$ 则为最终应用的结果投影矩阵。

白光条件下的归一化肤色表征为:

$$\mathbf{U}_{\text{skin}} = [R_s, G_s, B_s] = [0.7682, 0.5121, 0.3841]$$

对于非白光场景需进行色度校正:首先对各通道信号进行均值归一化处理,再将表皮色度信号 R_n, G_n, B_n 与 $[0.7682, 0.5121, 0.3841]$ 相乘得到标准化 RGB 通道:

$$R_s = 0.7682R_n, G_s = 0.5121G_n, B_s = 0.3841B_n$$

基于此,可定义标准化映射矩阵 \mathbf{M} :

$$\mathbf{M} = \mathbf{U}_{\text{skin}} = [R_s, G_s, B_s] = [0.7682, 0.5121, 0.3841]$$

CHROM 算法的初始投影矩阵 P_c 定义为:

$$P_c = \begin{bmatrix} 1 & -1 & 0 \\ 0.5 & 0.5 & -1 \end{bmatrix}$$

最终结果投影矩阵 $P_c \cdot \mathbf{M}$ 定义为:

$$P_c \cdot \mathbf{M} = \begin{bmatrix} 3 & -2 & 0 \\ 1.5 & 1 & -1.5 \end{bmatrix}$$

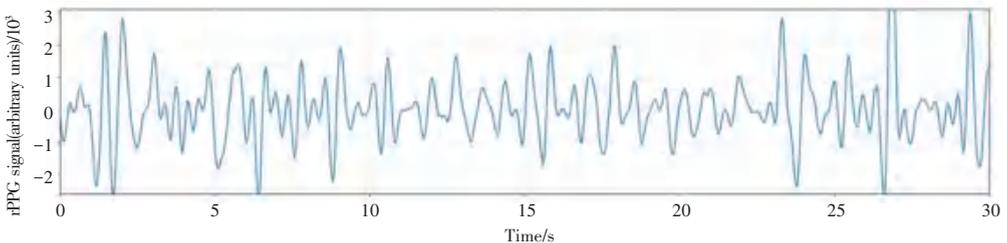


图3 提取出的 rPPG 信号示意图

Fig. 3 Schematic diagram of the extracted rPPG signal

CHROM 算法的核心优势在于投影方向选择的灵活性增强与先验知识依赖度降低,特别是在非白

光环境下的鲁棒性表现突出。但需注意其理论前提要求实际视频中的镜面反射分量 $\mathbf{N} \cdot u_s \cdot I_0$ 必须被

$$X_s = \frac{R_s - G_s}{0.7682 - 0.5121} = 3R_n - 2G_n \quad (15)$$

$$Y_s = \frac{R_s + G_s - 2B_s}{0.7682 + 0.5121 - 0.7682} = 1.5R_n + G_n - 1.5B_n \quad (16)$$

其中, $S = \frac{X_s}{Y_s} - 1$ 。由此可得 rPPG 信号表达式:

$$\log(1 + S) = \log(X_s) - \log(Y_s) \quad (17)$$

应用泰勒展开近似(各参数趋近于 $\mathbf{1}$)可简化为:

$$S \approx 1.5R_n - 3G_n + 1.5B_n \quad (18)$$

这表明 rPPG 信号可表征为独立通道的线性组合:

$$S = c_1R_n + c_2G_n + c_3B_n \quad (19)$$

原算法中脉冲信号 S 设计为 X_s 和 Y_s 的差值,但肤色标准化偏差将导致 X_s 和 Y_s 的幅值差异引发误差放大。所以修正上式为:

$$S = X_f - \alpha Y_f \quad (20)$$

其中, $\alpha = \frac{\sigma(X_f)}{\sigma(Y_f)}$; X_f 表示 X_n 的带通滤波版本; Y_f 表示 Y_n 的带通滤波版本; $\sigma(\cdot)$ 表示标准偏差算子。当心率先成分占主导时, X_f 和 Y_f 呈现相位反转特性,通过加权叠加即可提取有效 rPPG 信号。提取出的 rPPG 信号如图 3 所示。该修正确保即使存在标准化偏差,差值信号仍保持幅值平衡,将 rPPG 信号在 RGB 空间重写为:

$$S = 3\left(1 - \frac{\alpha}{2}\right)R_f - 2\left(1 + \frac{\alpha}{2}\right)G_f + \frac{3\alpha}{2}B_f \quad (21)$$

其中, R_f 表示 R_n 的带通滤波版本。

预设的单位颜色向量 u_s 充分补偿。若该条件不满足 ($M \cdot N \cdot u \cdot I_0 \neq \kappa \cdot 1$ 时), 算法处理后的信号仍可能残留微量镜面反射干扰。

2.3 心率值提取

传统频域分析法基于全局傅里叶变换, 假设信号具有平稳性, 难以适应动态心率场景下的非平稳特性。本研究提出一种基于汉宁窗的短时傅里叶变换 (Short-Time Fourier Transform, STFT) 时频联合分析方法, 通过时间-频率联合分辨率优化, 实现动态心率的精准跟踪与噪声抑制。

时频联合分析的核心思想在于将信号的能量分布映射至时间-频率二维平面, 通过局部频谱分析捕捉心率的瞬态变化特征。与全局频域分析不同, 该方法采用滑动时间窗口对信号进行分段加窗处理, 结合汉宁窗的旁瓣抑制特性, 有效平衡频谱泄漏抑制与时频分辨率。

(1) 汉宁窗的时频局部化。汉宁窗的数学表达式为:

$$\omega(n) = 0.5 \frac{\cos\left(\frac{2\pi n}{N-1}\right) + 1}{2}, n = 0, 1, \dots, N-1 \quad (22)$$

汉宁窗的能量集中特性使其 99.96% 的能量分布于主瓣, 适合准周期性的 rPPG 信号分析。

(2) 短时傅里叶变换。STFT 的表达式为:

$$X(t, \omega) = \int x(\tau) \omega(\tau - t) e^{-j\omega\tau} d\tau \quad (23)$$

其中, $\omega(\tau - t)$ 表示以时间 t 为中心的汉宁窗函数。

(3) 心率轨迹提取。时频能量密度定义为:

$$E(t, f) = |X(t, 2\pi f)|^2$$

逐帧搜索生理有效频段内的最大能量频率:

$$F_{HR}(t) = \arg \max_{0.75 \leq f \leq 4.0} E(t, f) \quad (24)$$

在时频联合分析中, 尽管汉宁窗和 STFT 能够有效提取心率频率轨迹, 但由于信号噪声和瞬时干扰的存在, 频率轨迹可能存在抖动或跳变。为进一步提升心率估计的稳定性, 引入卡尔曼滤波对频率轨迹进行动态平滑。卡尔曼滤波是一种递归状态估计算法, 能够通过融合预测值与观测值, 实现对动态系统状态的最优估计。最终心率值通过下式来计算:

$$HR(t) = 60 \times \hat{f}_{HR}(t) \quad (25)$$

其中, $\hat{f}_{HR}(t)$ 表示经卡尔曼滤波平滑后的频率轨迹。

设定窗长 $N = 3$ s (频率分辨率 $\Delta f \approx 0.33$ Hz), 理论上可区分频率差大于 0.33 Hz 的心率成分 (如相差大于 20 bpm)。3 s 窗长确保单帧包含至少 2 个心动周期 (心率大于等于 40 bpm), 避免频谱模糊。设定 75% 重叠率 (步长 22.5 帧, 帧率为 60), 使相邻帧共享 67.5 点数据, 在保证时间分辨率 (更新时间间隔 0.75 s) 的同时, 避免因帧间间隔过大导致心率突变检测延迟。基于汉宁窗的时频联合分析方法通过以下机制实现性能突破:

(1) 时频分辨率优化: 3 s 窗长与 75% 重叠率设定平衡了动态跟踪能力与频谱分辨率, 较传统频域法的阶梯响应时间缩短。

(2) 泄漏抑制: 汉宁窗控制频谱能量泄漏, 从而有效抑制运动伪影干扰。

3 非接触式心率测量验证实验

3.1 数据集与评估指标

UBFC^[8] 的数据集包含了 42 名被试者, 每位被试者对应 1 段独立的面部视频记录, 总计 42 段视频素材。UBFC 数据集实验场景示例如图 4 所示。视频以 30 帧/s 的恒定帧率进行录制。实验在受控的室内环境下完成, 引入了多样化的室内光照条件, 以模拟现实场景中可能遇到的光照变化对 rPPG 信号提取带来的挑战。被试者在测试阶段被引导进行数字游戏活动, 主动提升被试者的心率水平, 从而使采集到的 rPPG 信号包含更丰富的生理响应信息, 为评估心率检测算法的动态性能和鲁棒性提供了更具实际意义的数据基础。



图 4 UBFC 数据集实验场景示例

Fig. 4 Examples of experimental scenarios for the UBFC dataset

PURE^[9] 数据集收集了 10 位被试者的数据, 每位被试者展现出 6 种不同状态, 分别涉及静止、交谈、慢速头部平移、快速头部平移、轻微头部转动及中等程度的头部转动。参与者的座位位于窗户附近, 导致光线产生变化。交谈时面庞下部区域可能出现显著颜色变化, 而头部转动则引发光线比例的变动, 这些因素增加了测量的难度。PURE 数据集实验场景示例如图 5 所示。



图5 PURE数据集实验场景示例

Fig. 5 Examples of experimental scenarios for the PURE dataset

COHFACE^[10]数据集共有40位被试者,每位被试者在2种不同光照条件下被采集了4段视频,共计160段视频素材。2种光照条件中,一种是使用工作室的稳定光源,确保关闭所有窗户以避免自然光线的干扰,专注于通过室内光源均匀照射参与者的脸部进行录制。另一种条件则是自然光照,将窗户保持开放状态,与此同时关闭所有的室内照明设备。COFACE数据集实验场景示如图6所示。

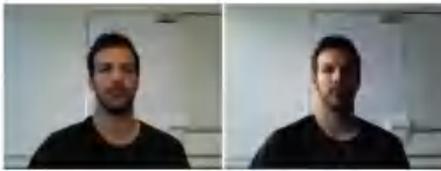


图6 COHFACE数据集实验场景示例

Fig. 6 Examples of experimental scenarios for the COHFACE dataset

采用3类量化指标进行算法性能评估:平均绝对误差(MAE)、均方根误差(RMSE)以及皮尔森相关系数(R)。各指标的定义论述如下。

(1)平均绝对误差(MAE):用于量化预测结果与参考设备测量值的绝对差异,其数学表达式如下:

$$MAE = \frac{1}{T} \sum_{i=1}^T |HR_i - HR'_i| \quad (26)$$

其中,HR表示接触式生理监测设备获取的心率、呼吸率基准值,HR'表示视频分析输出的预测结果。

(2)均方根偏差(RMSE):作为补充评估维度,着重考察预测值与真实值间的离散程度,其计算公式定义如下:

$$RMSE = \sqrt{\frac{1}{T} \sum_{i=1}^T (HR_i - HR'_i)^2} \quad (27)$$

(3)皮尔森线性相关系数:用于衡量预测值与基准值的线性关联强度,计算过程如下:

$$R = \frac{T \sum_{i=1}^T HR_i HR'_i - \sum_{i=1}^T HR_i \sum_{i=1}^T HR'_i}{\sqrt{(T \sum_{i=1}^T HR_i^2 - (\sum_{i=1}^T HR_i)^2) (T \sum_{i=1}^T (HR'_i)^2 - (\sum_{i=1}^T HR'_i)^2)}} \quad (28)$$

该系数通过协方差与标准差的比值关系,有效表征2组数据的同步变化趋势。

3.2 实验结果

表1展示了本研究提出的心率提取方法与现有方法在COHFACE数据集上的性能对比。COHFACE数据集包含复杂光照变化(200~1000 lux),对非接触式心率检测的鲁棒性提出了严峻挑战。实验结果表明,本研究方法在MAE(6.32 bpm)和RMSE(10.15 bpm)上显著优于传统方法(如2SR:MAE=20.89 bpm, RMSE=25.84 bpm; POS:MAE=13.43 bpm, RMSE=17.05 bpm),但皮尔逊相关系数仅为0.35。

表1 COHFACE数据集上的效果对比

Table 1 Comparison of effects on the COHFACE dataset

Method	MAE/bpm	RMSE/bpm	R
2SR ^[11]	20.89	25.84	-0.32
LiCVPR ^[12]	19.98	25.59	-0.44
POS ^[13]	13.43	17.05	0.07
HR-CNN ^[14]	8.10	10.78	0.29
DeepPPG ^[15]	3.07	7.06	0.86
本文方法	6.32	10.15	0.35

表2展示了本研究提出的心率提取方法与2SR、LiCVPR、POS、HR-CNN及PhysNet在PURE数据集上的性能对比。PURE数据集包含6种头部运动模式(如平移、旋转),对算法的运动鲁棒性提出极高要求。实验结果表明,本研究方法在MAE(1.95 bpm)与RMSE(2.43 bpm)上接近最优。且皮尔逊相关系数($R=0.98$)与2SR、HR-CNN、PhysNet持平。

表2 PURE数据集上的效果对比

Table 2 Comparison of effects on the PURE dataset

Method	MAE/bpm	RMSE/bpm	R
2SR ^[11]	2.44	3.06	0.98
LiCVPR ^[12]	28.20	30.96	-0.38
POS ^[13]	3.14	10.57	0.95
HR-CNN ^[14]	1.84	2.37	0.98
PhysNet ^[16]	1.90	3.44	0.98
本文方法	1.95	2.43	0.98

表3展示了本研究提出的心率提取方法与POS、MODEL、MAICA、BIC及CK在UBFC数据集上的性能对比。UBFC数据集因广域心率分布(45~180 bpm)对算法的动态适应能力提出挑战。实验结果表明,本研究方法以MAE=3.23 bpm、RMSE=4.30 bpm优于部分对比方法(如MODEL、MAICA),

且皮尔逊相关系数 ($R = 0.97$) 接近最优。

表 3 UBFC 数据集上的效果对比

Table 3 Comparison of effects on the UBFC dataset

Method	MAE/bpm	RMSE/bpm	R
POS	2.44	6.61	0.94
MODEL ^[17]	3.99	5.55	0.75
MAICA ^[18]	3.34	-	0.72
BIC ^[19]	1.21	2.41	0.95
CK ^[20]	2.30	3.80	0.98
本文方法	3.23	4.30	0.97

4 结束语

本文研究基于皮肤反射模型, 结合 CHROM 算法实现了抗运动干扰的 rPPG 信号提取, 针对传统时域峰值检测与频域分析法的局限性, 提出了基于汉宁窗短时傅里叶变换 (STFT) 的时频联合分析方法, 通过 3 s 窗长与 75% 重叠率设定 (时间分辨率为 0.75 s, 频率分辨率 0.33 Hz), 在动态场景下实现心率瞬态特征的精准捕捉, 提出的心率提取方法在 COHFACE、PURE 及 UBFC 数据集上均展现出不错性能, 验证了其在非接触式心率检测中的综合优势, 但在复杂光照环境下的表现不算理想。总体而言, 为非接触式健康监护提供了高精度、低成本的解决方案, 为临床场景下的连续生理监测奠定了技术基础。

参考文献

[1] 刘伟, 欧阳健飞, 闫勇刚. 基于 iPPG 的非接触式心率测量方法研究[J]. 计量技术, 2015(4): 3-5.

[2] 范强. 心血管生理参数非接触式检测关键技术研究[J]. 武汉: 武汉大学, 2017.

[3] 拜军, 张保华, 张珊, 等. 非接触式人体心率监测方法的研究进展[J]. 生物医学工程与临床, 2024, 28(2): 293-298.

[4] GIDEON J, STENT S. The way to my heart is through contrastive learning: Remote photoplethysmography from unlabelled video [C]//Proceedings of 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, NJ: IEEE, 2021: 3975-3984.

[5] 张震乾. 基于面部视频分析的非接触式心率实时检测研究[D]. 兰州: 兰州大学, 2023.

[6] SALIM A S, KHIDHIR A S M. A comprehensive review of rPPG methods for heart rate estimation [J]. Open Access Library Journal, 2024, 11(11): 1-18.

[7] HAAN D G, JEANNE V. Robust pulse rate from chrominance-based rPPG [J]. IEEE Transactions on Biomedical Engineering, 2013, 60(10): 2878-2886.

[8] BOBBIA S, MACWAN R, BENEZETH Y, et al. Unsupervised skin tissue segmentation for remotephotothysmography [J]. Pattern Recognition Letters, 2019, 124: 82-90.

[9] STRICKER R, MÜLLER S, GROSS H M. Non-contact video-based pulse rate measurement on a mobile service robot [C]//Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication. Piscataway, NJ: IEEE, 2014: 1056-1062.

[10] HEUSCH G, ANJOS A, MARCEL S. A reproducible study on remote heart rate measurement [J]. arXiv preprint arXiv, 1709.00962, 2017.

[11] WANG Wenjin, STUIJK S, HAAN D G. A novel algorithm for remote photoplethysmography: Spatial subspace rotation [J]. IEEE Transactions on Biomedical Engineering, 2015, 63(9): 1974-1984.

[12] LI Xiaobai, CHEN Jie, ZHAO Guoying, et al. Remote heart rate measurement from face videos under realistic situations [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2014: 4264-4271.

[13] WANG Wenjin, DEN BRINKER A C, STUIJK S, et al. Algorithmic principles of remote PPG [J]. IEEE Transactions on Biomedical Engineering, 2016, 64(7): 1479-1491.

[14] ŠPETLÍK R, FRANC V, MATAS J. Visual heart rate estimation with convolutional neural network [C]//Proceedings of the British Machine Vision Conference. Newcastle, UK: BMVA, 2018: 3-6.

[15] LIU S Q, YUEN P C. A general remotephotothysmography estimator with spatiotemporal convolutional network [C]//Proceedings of 2020 15th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2020). Piscataway, NJ: IEEE, 2020: 481-488.

[16] YU Zitong, LI Xiaobai, ZHAO Guoying. Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks [J]. arXiv preprint arXiv, 1905.02419, 2019.

[17] LI Peixi, BENEZETH Y, NAKAMURA K, et al. Model-based region of interest segmentation for remote photoplethysmography [C]//Proceedings of the 14th International Conference on Computer Vision Theory and Applications. Cham: Springer, 2019: 383-388.

[18] MACWAN R, BENEZETH Y, MANSOURI A. Heart rate estimation using remotephotothysmography with multi-objective optimization [J]. Biomedical Signal Processing and Control, 2019, 49: 24-33.

[19] BENEZETH Y, BOBBIA S, NAKAMURA K, et al. Probabilistic signal quality metric for reduced complexity unsupervised remotephotothysmography [C]//Proceedings of 2019 13th International Symposium on Medical Information and Communication Technology (ISMICT). Piscataway, NJ: IEEE, 2019: 1-5.

[20] SONG Rencheng, ZHANG Senle, CHENG Juan, et al. New insights on super-high resolution for video-based heart rate estimation with a semi-blind source separation method [J]. Computers in Biology and Medicine, 2020, 116: 103535.