

DOI: 10.19650/j.cnki.cjsi.J2311085

模型和盲源分离融合的运动鲁棒视频心率检测*

成娟^{1,2}, 武梦影¹, 宋仁成¹, 刘羽¹, 陈勋³

(1. 合肥工业大学生物医学工程系 合肥 230009; 2. 合肥工业大学测量理论与精密仪器安徽省重点实验室 合肥 230009; 3. 中国科学技术大学电子工程与信息科学系 合肥 230026)

摘要: 远程光电容积描记法 (rPPG) 是一种基于视频的非接触心率检测技术, 由于其携带的血容量脉搏 (BVP) 信号幅值微弱易受运动噪声影响。本文提出一种融合色差模型和盲源分离的运动鲁棒视频心率检测技术。一方面, 对每个感兴趣区域 (ROIs) 采用色差模型并时延构建多通道数据集, 凸显准周期变量同时抑制无规律运动噪声; 另一方面, 采用联合盲源分离提取共同包含于两个数据集中的 BVP 源分量向量 (SCVs), 筛选合适的 SCV 确定为 BVP 信号, 从而计算心率值。所提方法在数据库 UBFC-RPPG 和 ECG-Fitness 上验证并与其他方法对比, 在剧烈运动场景下性能最佳, 心率 $HR_{\text{mac}} = 9.93$ bpm, $HR_{\text{rmse}} = 16.17$ bpm, 相关系数 $r = 0.75$, 为 rPPG 技术的实用化进程提供了解决思路。

关键词: 远程光电容积描记法; 非接触; 色差模型; 联合盲源分离; 运动噪声去除

中图分类号: TN911.7 TP751.1 TH79 **文献标识码:** A **国家标准学科分类代码:** 460.4099

Motion-robust heart rate measurement from videos via combining model with blind source separation

Cheng Juan^{1,2}, Wu Mengying¹, Song Rencheng¹, Liu Yu¹, Chen Xun³

(1. Department of Biomedical Engineering, Hefei University of Technology, Hefei 230009, China; 2. Anhui Province Key Laboratory of Measuring Theory and Precision Instrument, Hefei University of Technology, Hefei 230009, China; 3. Department of Electronic Engineering and Information Science, University of Science and Technology of China, Hefei 230026, China)

Abstract: Remote photoplethysmography (rPPG) is a video-based and the non-contact heart rate (HR) detection technique that is susceptible to motion noise due to weak amplitudes of blood volume pulse signal that carries, making it challenging to accurately detect video-based HR values in motion scenarios. This article studies HR estimation based on a combination of chrominance (CHROM) signals and joint blind source separation (JBSS). On the one hand, CHROM is applied to each facial region of interest (ROI) and time-delay operation is followed by constructing two multi-channel datasets, which can highlight quasi-periodic variables while suppressing irregular motion noises. On the other hand, JBSS technique is applied to the generated two datasets to extract the common underlying source component vectors (SCVs), where the one indicating BVP signal is selected and the HR is measured. The proposed method is evaluated on two public databases UBFC-RPPG and ECG-Fitness, and compared with several other typical methods. The results show that the method achieves the best performance of HR estimation during the dramatic sport situation, with $HR_{\text{mac}} = 9.93$ bpm, $HR_{\text{rmse}} = 16.17$ bpm and $r = 0.75$. It provides a solution for the practical applications of the rPPG technology.

Keywords: remote photoplethysmography; non-contact; chrominance; joint blind source separation; motion artifact removal

收稿日期: 2023-02-23 Received Date: 2023-02-23

* 基金项目: 国家自然科学基金面上项目 (62171176, 62271186)、中央高校基本科研业务费专项 (JZ2021HGPA0061)、安徽省重点研发计划项目 (202104d07020005) 资助

0 引言

生理参数可用于评估个体健康^[1-2],心率(heart rate, HR)作为最重要生理参数之一,对其进行检测具有重要意义。考虑到摄像头使用广泛、成本低,基于视频的非接触检测技术^[3],也称远程光电容积描记法(remote photoplethysmography, rPPG)引起了广泛关注^[4]。该项技术通过视频图像反映伴随心跳生理活动引起的面部毛细血管的血容量变化测量心率,可应用于移动健康监测^[5]。

由于rPPG携带的血容量脉搏(blood volume pulse, BVP)信号幅值微弱,容易受到运动噪声干扰,特别是运动幅度大且具有一定规律的运动场景,对视频心率检测造成挑战。现有很多研究工作致力于提高rPPG的运动鲁棒性,主要分为信号去噪法、基于模型的方法^[6]、基于盲源分离的方法^[7]、基于深度学习的方法^[8]、以及基于组合的方法^[9]。

信号去噪法对包含较强脉动信息的绿色通道进行多通道分解,筛选出表征心率信息的信号。例如采用集合经验模态分解方法对含噪绿色通道信号分解,获得包含心率信息的本征模态函数^[10]。基于模型的方法则是将引起人脸和相机之间距离或角度的变化的运动进行建模,例如色差模型(chrominance model, CHROM)^[11]和平面正交皮肤反射模型(plane orthogonal to skin, POS)^[6]。当运动幅值小,可以将脉动信号、运动噪声信号和其他信号看成线性组合的问题,此时采用盲源分离(blind source separation, BSS)方法进行心率源信号提取,最具代表性的有独立分量分析(independent component analysis, ICA)^[12]。然而,当运动噪声幅值过高时,BSS方法容易失效;同时经ICA分离后的独立成分顺序不固定。针对这个问题,考虑到面部不同区域共享HR信息,采用联合盲源分离(joint BSS, JBSS)技术进行HR检测^[13]。例如,独立向量分析(independent vector analysis, IVA)^[14]。伴随人工智能发展,深度学习技术也被用于rPPG中,例如,基于生成对抗网络(generative adversarial network, GAN)的BVP波形质量增强^[15]。此外,现有研究还提出将不同类型的方法结合以提高BVP信号质量。较为常见的是模型方法和盲源分离方法的融合,通常利用CHROM提供初步去噪后的BVP信号作为后续分析的输入。

基于深度学习的rPPG方法通常依赖较大规模的样本,同时其可解释性和泛化能力有待进一步加强;传统盲源分离方法当运动幅度剧烈时难以满足线性组合的先验假设,无法解决运动频率与感兴趣的HR频带范围重叠的心率准确提取问题。据此,本文提出一种基于色差模型和时延(time delay, TD)联合盲源分离技术相结合的

视频心率检测方法,用于提高健身环境下心率检测的运动鲁棒性,其主要贡献如下:

1)提出了一种运动鲁棒的视频心率检测框架CHROM-TDJBSS。该方法首先使用CHROM完成运动噪声的初步去除,并在TD的作用下,完成了从单通道信号到多通道信号的扩展,提供JBSS合适的输入的同时强化周期性变量。其次,利用JBSS提取共同包含于两个多通道数据集的源分量向量,并考虑周期性运动噪声和周期性心率信息的关系,提高了剧烈且周期性运动场景下视频心率检测的准确性。

2)在公开数据库UBFC-RPPG和ECG-Fitness上,对所提CHROM-TDJBSS进行性能评估,并与其他典型的rPPG方法对比,结果表明所提CHROM-TDJBSS在运动场景下取得了最好的结果。

1 相关工作

1.1 基于模型的方法

De等^[11]于2013年提出了色差模型CHROM,通过假设一个恒定的标准化肤色,并且归一化平均肤色信号与假设失真正交,以此来构建与肤色信号之间的线性组合,实现运动鲁棒的心率检测。不同于CHROM根据光学模型来去除运动伪迹,血容量脉冲特征模型^[16]则根据不同波长血容量变化的特征,区分颜色通道中脉冲引起的颜色变化和运动噪声。POS^[6]作为另一种基于光学模型的方法,选择与肤色正交的平面作为RGB的投影平面以完成脉冲信号的分离。

1.2 基于盲源分离的方法

基于BSS的方法的处理对象一般为单个数据集和多个数据集,后者通常称之为联合盲源分离。

当处理单个数据集时,ICA是典型的BSS方法。Poh等^[17]提出3个颜色通道为3个独立信号的线性组合,使用ICA分离出BVP信号。也有学者提出采用主成分分析(principle component analysis, PCA)从3个颜色通道中提取包含心率信息的主要成分^[18]。除了信号时域分析,BSS方法也进行频域分析。例如非负矩阵分解^[19]利用BVP信号的时间稀疏性,将包含心率的源信号从混合信号中分解出来。此外,还可以使用非线性模式分解从面部视频中估计出心率^[20]。但是,当运动伪迹幅值剧烈不能满足线性假设条件时,采用基于ICA的BSS方法提取的BVP信号质量较差,不能满足心率估计要求。

当处理多数数据集时,JBSS方法能够提取不同数据集共同包含的源信息。Guo等^[21]首次将JBSS引入rPPG领域,通过将面部皮肤区域划分成多个子区域,构建多个多变量时间序列数据集,并应用IVA提取不同数据集间

共享的源向量。Song 等^[7]利用多集典型相关分析 (multiset canonical correlation analysis, MCCA) 实现了非均匀光照条件下运动伪迹的去除。JBSS 通过同时处理多个数据集更好地协同了不同面部区域的信息,获得的 SCVs 按照相关性由高到低排序,解决了基于 ICA 的 BSS 方法排列不确定性问题^[14,21-22],但是当运动伪迹也是共同源时,JBSS 方法需要先进行运动伪迹去除才能确保心率检测的准确性。

1.3 基于模型和 BSS 组合的方法

基于模型的方法和基于 BSS 的方法相结合可以提高视频心率检测的准确性。例如 Song 等^[23]对 CHROM 信号应用快速核密度独立向量分析 (kernel density independent component analysis, KDICA),利用 KDICA 不需要隐源分布的先验假设,替代 α 调谐提取更高质量的 BVP 信号,解决当脉冲和镜面反射信号强度相似时 CHROM 的 α 调谐失败问题。但是该方法并不适用于剧烈健身环境下的心率检测。据此,本文提出融合 CHROM 和 JBSS 的运动鲁棒视频心率检测方法,实现剧烈且具有周期性运动伪迹情况下视频心率的有效检测。

2 方 法

所提基于 CHROM-TDJBSS 视频心率检测流程主要包含三大步骤。首先确定两个面部皮肤 ROIs,分别获取每个 ROI 的 CHROM 信号并时延,构建两个多通道的数据集;其次对这两个数据集采用 JBSS 提取共同包含的 SCVs,筛选合适的 SCV 为目标 BVP 信号;最后对该 BVP 信号进行傅里叶变换,其主频即为所测心率主频,从而计算 HR 值。

2.1 面部 ROI 确定和追踪

使用 OpenFace 工具箱提供的卷积专家约束局部模型进行面部特征点的检测和追踪^[24],选取其中相对稳定的 4 个特征点来确定前额 ROI 和面颊 ROI,如图 1 中矩形框所示。其中,前额 ROI 的宽 w_1 为左右眉毛最边缘两个特征点的 x 坐标值之差;高度 h_1 则设置为 30 pixel。脸颊 ROI 的宽 w_2 为眉毛中间和鼻尖上方这两个特征点的 x 坐标值之差;高度 h_2 则取这两个特征点的 y 坐标值之差的一半。

2.2 融合色差和联合盲源分离的 rPPG 算法

1) 色差模型

通过对第 q 帧视频图像中的 ROI 进行像素平均,逐帧连接获得三通道 RGB 时间序列。基于皮肤反射模型,利用标准化的 RGB 通道时间序列的投影来定义色度信号 $C_{q,1}^{[m]}$ 和 $C_{q,2}^{[m]}$,并通过 α 调谐技术进一步分离光强变化和脉冲信号得到色差信号 $C_q^{[m]}$, $1 \leq m \leq M$,本文 $M=2$ 对

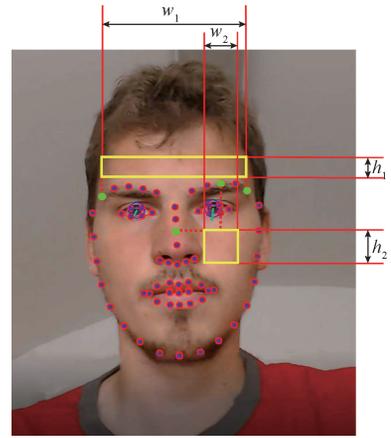


图 1 面部两个 ROI 示意图

Fig. 1 Diagram of two facial ROIs

应于两个面部 ROIs,具体计算如式(1)所示。

$$\begin{cases} C_{q,1}^{[m]} = 3R_q^{[m]} - 2G_q^{[m]} \\ C_{q,2}^{[m]} = 1.5R_q^{[m]} + G_q^{[m]} - 1.5B_q^{[m]} \\ C_q^{[m]} = C_{q,1}^{[m]} - \alpha \cdot C_{q,2}^{[m]}, \text{其中 } \alpha = \frac{\sigma(C_{q,1}^{[m]})}{\sigma(C_{q,2}^{[m]})} \end{cases} \quad (1)$$

其中, $R_q^{[m]}$, $G_q^{[m]}$ 和 $B_q^{[m]}$ 分别是归一化带通滤波后的第 m 个数据集的第 q 帧 R、G、B 像素均值, $\sigma(\cdot)$ 表示给定信号的标准偏差。每一通道的色差信号去趋势处理,根据运动场景的变化,选择不同的去趋势参数 λ 。

2) 联合盲源分离

将上述去趋势后的每个 CHROM 信号进行时延^[14],分别构建一个 3 通道数据集作为 JBSS 的输入。有研究表明,时延操作可以凸显 rPPG 信号中准周期性的心率变量或具有周期性的运动伪迹变量。

假设需要处理的视频图像帧总数 Q ,对应于第 m 个 ROI 去趋势后的单通道色差信号如式(2)所示。

$$\tilde{\mathbf{C}}^{[m]} = [\tilde{C}_{(1)}^{[m]}, \tilde{C}_{(2)}^{[m]}, \dots, \tilde{C}_{(q)}^{[m]}, \dots, \tilde{C}_{(Q)}^{[m]}] \quad (2)$$

时延向量构建的三通道数据集,其矩阵形式:

$$\mathbf{X}^{[m]} = \begin{bmatrix} \tilde{C}_{(1)}^{[m]} & \tilde{C}_{(2)}^{[m]} & \dots & \tilde{C}_{(Q)}^{[m]} \\ \tilde{C}_{(Q)}^{[m]} & \tilde{C}_{(1)}^{[m]} & \dots & \tilde{C}_{(Q-1)}^{[m]} \\ \tilde{C}_{(Q-1)}^{[m]} & \tilde{C}_{(Q)}^{[m]} & \dots & \tilde{C}_{(Q-2)}^{[m]} \end{bmatrix} \quad (3)$$

利用 JBSS 提取时延后的两个 3 通道数据集中共同存在的源分量向量,给定 M 个数据集 ($M \geq 2$),每个数据集包含 P 个通道, Q 个图像帧,第 m 个数据集 $\mathbf{X}^{[m]}$ 可以由其对应的列向量表示为:

$$\mathbf{X}^{[m]} = [x_{(1)}^{[m]}, x_{(2)}^{[m]}, \dots, x_{(q)}^{[m]}, \dots, x_{(Q)}^{[m]}] \quad (4)$$

其中, $1 \leq m \leq M$, $x_{(q)}^{[m]}$ 表示大小为 $P \times 1$ 的列向量 $\mathbf{x}^{[m]}$ 的第 q 个实现,假设每个数据集是 L ($L \leq P$) 个独立源的线性混合,即:

$$\mathbf{X}^{[m]} = \mathbf{W}^{[m]} \mathbf{S}^{[m]}, 1 \leq m \leq M \quad (5)$$

其中, $\mathbf{W}^{[m]}$ 是混合矩阵, $\mathbf{S}^{[m]} = [S_{(1)}^{[m]}, S_{(2)}^{[m]}, \dots, S_{(L)}^{[m]}]^T$ 是源矩阵, T 表示转置。JBSS 能够提取多个数据集间共同包含的 SCVs。第 l 个 SCV 是独立于其他 SCV 的随机向量, 每个 SCV 内的分量相互依赖。JBSS 旨在找到混合矩阵 $\mathbf{W}^{[m]}$ 或其逆矩阵 $(\mathbf{W}^{[m]})^{-1}$, 以及相应的源向量估计 $\mathbf{y}^{[m]} = (\mathbf{W}^{[m]})^{-1} \cdot \mathbf{X}^{[m]}$ 。第 l 个 SCV 的估计为 $\mathbf{y}_l = [y_l^{[1]}, y_l^{[2]}, \dots, y_l^{[m]}, \dots, y_l^{[M]}]^T$ 。 $y_l^{[m]}$ 表示 $y_l^{[m]} = ((\mathbf{W}_l^{[m]})^{-1})^T \cdot \mathbf{X}^{[m]}$ 。IVA 作为 JBSS 的典型方法, 其目标是从 M 个数据集中识别 L 个独立的 SCV, 并通过最小化估计 SCV 之间的互信息 I_{IVA} 来实现, 互信息的定义为:

$$\begin{aligned} I_{IVA} &\triangleq I[\mathbf{y}_1; \mathbf{y}_2; \dots; \mathbf{y}_L] = \sum_{L=1}^L H[\mathbf{y}_L] - H[\mathbf{y}_1; \mathbf{y}_2; \dots; \mathbf{y}_L] \\ &= \sum_{L=1}^L H[\mathbf{y}_L] - H[\mathbf{W}^{[1]}]^{-1} \mathbf{X}^{[1]}, \dots, (\mathbf{W}^{[M]})^{-1} \mathbf{X}^{[M]}] = \\ &= \sum_{L=1}^L \left(\sum_{m=1}^M H[y_l^{[m]} - I[\mathbf{y}_l]] \right) - \\ &= \sum_{m=1}^M \log |\det((\mathbf{W}^{[m]})^{-1})| - C_1 \end{aligned} \quad (6)$$

其中, $H[\cdot]$ 表示熵, C_1 是常数项 $H[\mathbf{X}^{[1]}, \mathbf{X}^{[2]}, \dots, \mathbf{X}^{[M]}]$ 。每一组 SCV 按照相关性从高到低排列, 最终可以解决排列歧义问题。

IVA 算法包括拉普拉斯分布和高斯分布, 分别记为 IVA-L 和 IVA-G。大多神经生理学应用中数据集间的二阶相关性表现明显, 故本文采用 IVA-G。

2.3 心率计算

JBSS 之后, 从每个数据集中恢复出与原通道数相同的 SCVs, 按照相关性由高到低排列。由于前额 ROI 面积大, 非刚性运动少, 且追踪效果好等优点, 我们选取前额 ROI 中恢复出的 SCVs 作为目标 BVP 信号的候选。经过实验观测, 当有明显运动伪迹存在时, 例如 ECG-Fitness 数据库, 第 1 相关的 SCV 仍是运动噪声占主导, 而第 2 相关的 SCV 通常是目标 BVP 信号。因此, 将前额 ROI 恢复出的排序第 2 的 SCV 确定为目标 BVP 信号。之后, 采用巴特沃斯滤波器滤波 (滤波器阶数设置为 3、截止频率为 0.75~2.5 Hz)。然后, 使用快速傅里叶变换 (fast Fourier transform, FFT) 计算该 SCV 的主频 f_{\max} , 最终估计的心率值为 $60 \times f_{\max}$ 次/分钟 (beat per minute, bpm)。本文采用的第 2 个数据库 UBFC-RPPG 涉及的运动伪迹通常为非周期性, 且运动幅值相对微弱, 为了在 UBFC-RPPG 数据库上获得精确的心率值, 对 JBSS 提取的所有 SCVs 进行 FFT, 并将峰值信噪比最大的 SCV 的主频确定为所测心率的主频, 从而得到最终的心率值。

3 实验与讨论

3.1 数据库介绍

1) UBFC-RPPG 数据库: 采集 42 名受试者的面部视频, 每个视频约 1 min。视频使用 Logitech C920 HD pro

摄像头采集, 信号采样率为 30 FPS, 分辨率为 640×480, 视频采用无压缩 8 位 RGB 格式保存。心率参考信号是由 ContecMedical CMS50E 采集的 PPG 信号, 采样率为 30 Hz。

2) ECG-Fitness 数据库: 由 204 个健身主题视频组成, 每个视频 1 min。17 名受试者 (年龄从 20~53 岁) 进行 4 种活动 (说话、划船、固定自行车、椭圆机), 摄像机以 30 FPS, 1 920×1 080 pixel 进行拍摄, 并以未压缩的 YUV 平面像素格式存储。采用 CC5 导联的双导联 Vitom CheckMeTMPPro 心电图仪同步采集 ECG 信号, 采样率为 125 Hz。

数据库中每个视频的处理窗口设置为 30 s, 窗口的滑动步长设置为 5 s, UBFC-RPPG 数据库共产生 320 段数据, ECG-Fitness 数据库共产生 332 段数据。

3.2 评价指标

采用常规评价指标评估心率检测性能^[25], 包括平均绝对误差 HR_{mae} (mean absolute error, MAE), 均方根误差 HR_{mse} (root mean square error, RMSE) 和皮尔逊相关系数 r (Pearson's correlation coefficient, r)。

3.3 实验结果

为评估所提 CHROM-TDJSS 方法的性能, 将其与其他几种典型的 rPPG 算法进行比较, 具体有 POS^[6]、CHROM^[11]、ICA^[17] 和 JBSS^[14] 方法; 同时, 也与几种最新的深度学习方法进行性能对比。此外, 我们也做了消融实验。

1) 静止场景下实验结果

表 1 给出了 UBFC-RPPG 数据库对比实验结果。从表 1 中不难看出, 相较于传统方法, 所提 CHROM-TDJSS 方法除了 HR_{mae} 比取得最佳性能的 POS 方法低 0.13 bpm 之外, 其他两个指标都取得了最好的性能。具体地, $HR_{\text{mae}} = 1.47$ bpm, $HR_{\text{mse}} = 3.77$ bpm, $r = 0.98$ 。相较于深度学习方法, 所提方法可以达到与之相当的性能。PulseGAN 方法获得了最好的性能指标, 而所提方法的 HR_{mae} 仅比 PulseGAN 高 0.28 bpm, 却比 MetaPhys 方法低 0.43 bpm, 说明了所提方法在 UBFC-RPPG 数据库上性能优异。

表 1 UBFC-RPPG 数据库对比实验结果

方法	$HR_{\text{mae}}/\text{bpm}$	$HR_{\text{mse}}/\text{bpm}$	r
POS ^[6]	1.34	4.10	0.98
CHROM ^[11]	1.65	4.82	0.97
ICA ^[17]	3.26	9.99	0.86
JBSS ^[14]	2.77	9.34	0.88
CHROM-TDICA	1.83	5.56	0.96
PulseGAN ^[9]	1.19	2.49	0.98
MetaPhys ^[26]	1.90	2.62	0.96
CHROM-TDJSS	1.47	3.77	0.98

由于 UBFC-RPPG 数据库中受试者相对静止,其运动干扰幅值较小且视频保存质量高(无压缩),一些传统方法,例如基于模型的 CHROM 和 POS 就能取得非常好的性能。此时,基于 BSS 方法的性能通常低于基于模型的方法,这与当前很多研究结果是吻合的^[6]。而单纯基于 JBSS 的方法要优于基于 BSS 的方法,具体地,JBSS 方法获得的 HR_{mae} 相比于 ICA 方法下降 0.49 bpm, HR_{mse} 下降 0.65 bpm,而 r 提升 0.02。而模型和 BSS 组合方面,JBSS 的组合同样优于 BSS 的组合。具体地,CHROM-TDJBSS 的 HR_{mae} 相比于 CHROM-TDICA 下降 0.36 bpm, HR_{mse} 下降 1.79 bpm,而 r 提升 0.02。CHROM-TDICA 相

比于 ICA,其性能有了显著提升, HR_{mae} 下降 43.9% (1.43 bpm)、 HR_{mse} 下降 44.3% (4.43 bpm)、 r 提高 0.1,充分说明了 CHROM 方法与 BSS 方法结合的有效性,尤其是 CHROM 外加时延可以有效地抑制幅值相对较小的运动伪迹干扰。与此同时,CHROM-TDJBSS 与 JBSS 的性能对比同样说明了 CHROM 外加时延对抑制运动的有效性。

图 2 给出了所提方法和 5 种传统方法的 Bland-Altman 图。图 2 表明,这 5 种方法获得的 1.96 倍的标准差最低为 8.53 bpm,最高为 21.35 bpm。然而所提 CHROM-TDJBSS 方法可以将 1.96 倍的标准差降低到 7.42 bpm,平均误差为 0.02 bpm。

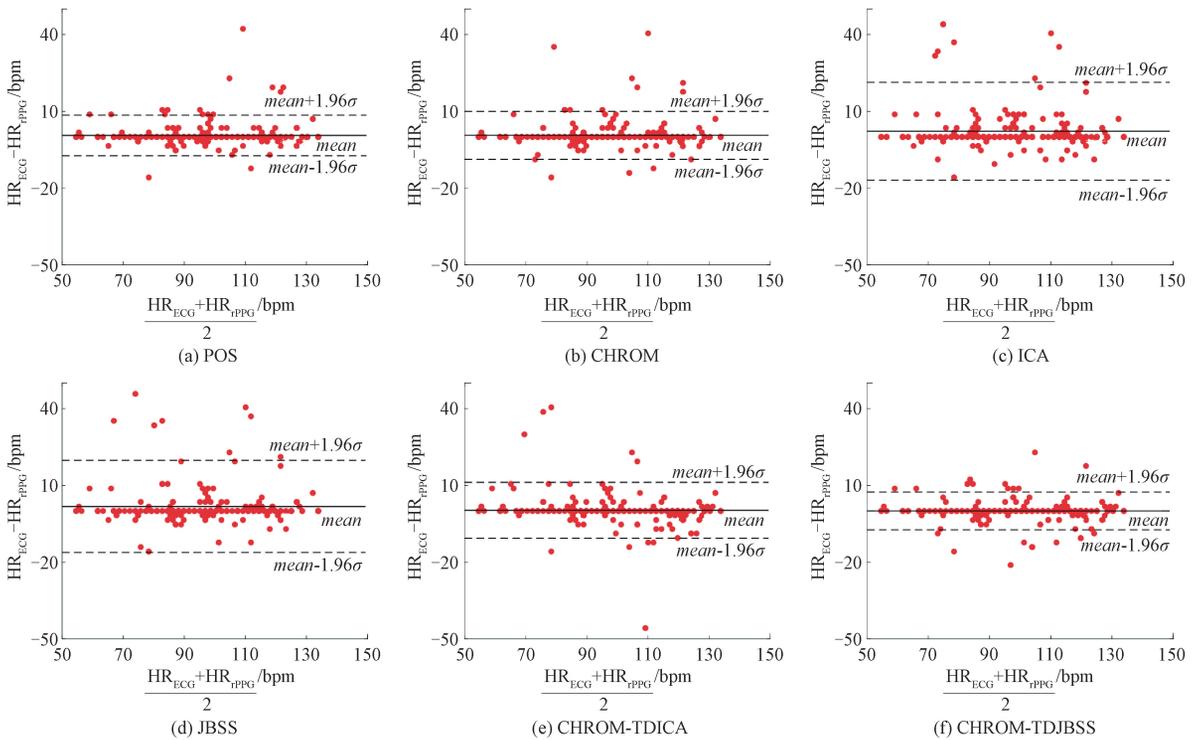


图 2 UBFC-RPPG 数据库上不同方法测量 HR 的 Bland-Altman 图

Fig. 2 Bland-Altman plots of HR measured by different methods on UBFC-RPPG database

2) 运动场景下实验结果

表 2 所示为 ECG-Fitness 数据库上所提方法与其他几种方法的实验结果对比。从表 2 中看出,所提 CHROM-TDJBSS 的所有指标均取得了最佳结果, HR_{mae} 降低到 9.25 bpm, HR_{mse} 降低到 16.12 bpm, r 提升到 0.75。基于模型方法(CHROM 和 POS)的性能相比基于盲源分离方法(ICA 和 JBSS)而言较优, HR_{mae} 至少降低 33.70% (6.66 bpm), HR_{mse} 至少降低 25.03% (6.91 bpm), r 至少提高 0.17。ECG-Fitness 数据库包含周期性且运动幅度较大的不同运动,这也印证了当运动幅度较大时,运动噪声与脉冲信号的组合难以符合线性模型^[6],同时也说明了基于模型的方法对运动伪迹具有更好的鲁棒性。

表 2 ECG-Fitness 数据库对比实验结果

Table 2 Results comparison on ECG-Fitness database

方法	$HR_{\text{mae}}/\text{bpm}$	$HR_{\text{mse}}/\text{bpm}$	r
POS ^[6]	12.87	20.70	0.58
CHROM ^[11]	13.10	19.17	0.59
ICA ^[17]	19.76	27.61	0.41
JBSS ^[14]	18.98	25.79	0.44
CHROM-TDICA	12.99	20.93	0.62
CHROM-TDJBSS	9.93	16.12	0.75

JBSS 方法获得的 HR_{mae} 相比 ICA 方法下降 0.78 bpm, HR_{mse} 下降 1.82 bpm, r 提升 0.03,这说明

JBSS方法性能优于BSS方法,主要原因在于多个面部ROI共同包含的BVP信号质量好于来自单个面部ROI的BVP信号。同时,CHORM-TDJBSS和CHROM-TDICA的结果对比更加说明了JBSS相比于ICA的优势。CHROM-TDJBSS相比于JBSS,其 HR_{mae} 下降47.68%(9.95 bpm), HR_{rmse} 下降37.50%(9.67 bpm), r 提升

0.31,说明CHROM加时延操作的有效性。CHROM-TDICA与ICA的性能对比也验证了该结论。

图3所示为在ECG-Fitness数据库上所提方法与其他rPPG方法的Bland-Altman图。对比方法取得的1.96倍的标准差最低为38.23 bpm,最高为60.38 bpm,而所提CHROM-TDJBSS将1.96倍的标准差降低到28.83 bpm。

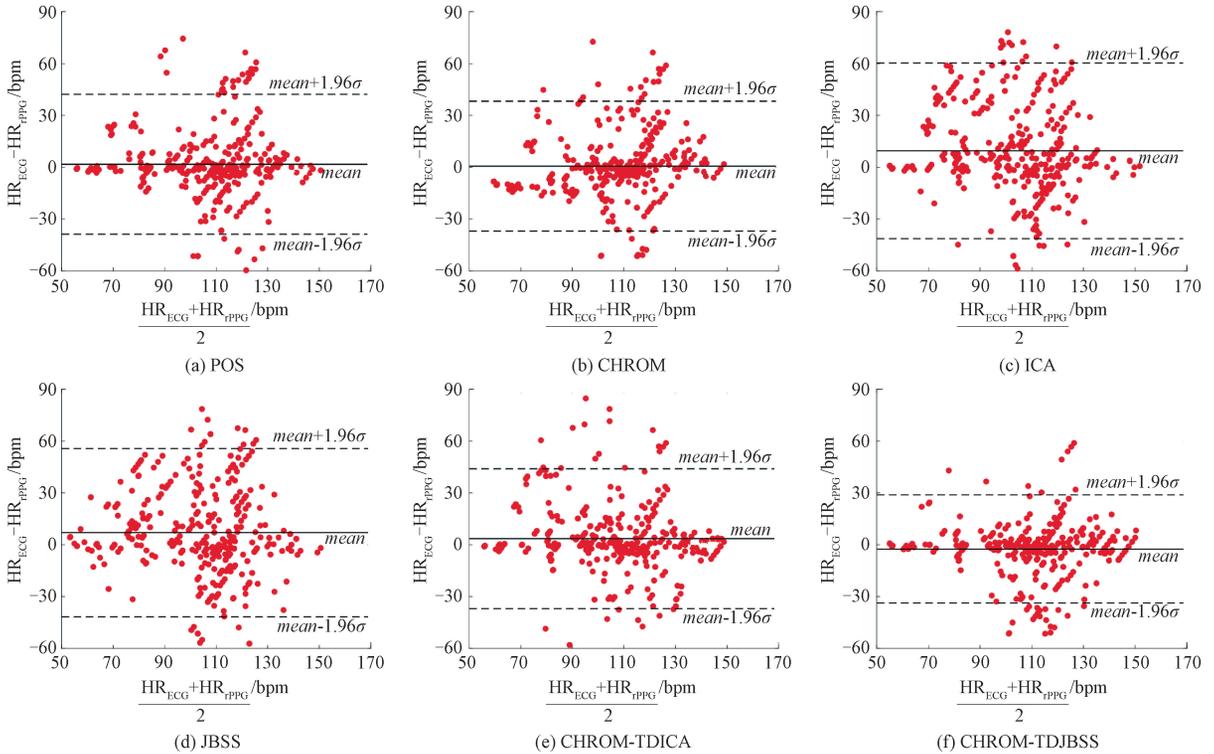


图3 ECG-Fitness数据库上测量HR的Bland-Altman图

Fig. 3 Bland-Altman plots of HR measured by different methods on the ECG-Fitness database

进一步地,图4所示为所提方法与其他rPPG方法测量HR的散点图。从图4中看出,所提CHROM-TDJBSS相较于其他五种方法,其HR测量的线性回归系数更集中在基线 $y=x$ 附近,其与HR参考测量值的相关性更强。

表3所示为在ECG-Fitness数据库上所提方法现有基于深度学习方法的结果对比,不难看出本文所提方法可以达到媲美基于深度学习方法的性能。

表3 ECG-Fitness数据库与深度学习方法的实验结果对比
Table 3 Results Comparison with deep learning methods on ECG-Fitness database

方法	$HR_{\text{mae}}/\text{bpm}$	$HR_{\text{rmse}}/\text{bpm}$	r
HR-CNN ^[27]	14.48	19.15	0.50
Song et al. ^[25]	10.34	12.99	0.57
CHROM-TDJBSS	9.93	16.12	0.75

图5所示为ECG-Fitness数据库上HR误差分布。根据中华人民共和国医药行业标准中心率检测误差要小

于5 bpm,不难发现,CHROM绝对HR误差低于5 bpm的比例为37.35%,而CHROM-TDJBSS将该比例提高到52.41%。CHROM的绝对HR误差在10 bpm以内的百分比为54.52%,而CHROM-TDJBSS将该比例提高到71.99%。

表4还细节展示了ECG-Fitness数据库不同运动场景下,所提方法和其他多种方法的实验结果对比。从表4中可以看出所提CHROM-TDJBSS在4种场景下均取得了最优结果。POS仅在椭圆机场景中性能表现较差,说明POS更适用于运动频率较低或非刚性运动去噪;而划船机和说话场景下的心率测量性能略低于所提方法。CHROM在四种运动场景中的性能差异不大,且从表4中看出CHROM相较于POS更能从总体上提升BVP信号质量。以ECG-Fitness 6号受试者划船机场景的30 s视频片段为例,图6所示为所提方法提取的BVP信号示意图,第2行虚线是提取的绿色通道信号,实线为参考ECG信号。经CHROM-TDJBSS处理后,原来由运动占主导的绿色通道信号优化为BVP信号(第3行虚线)

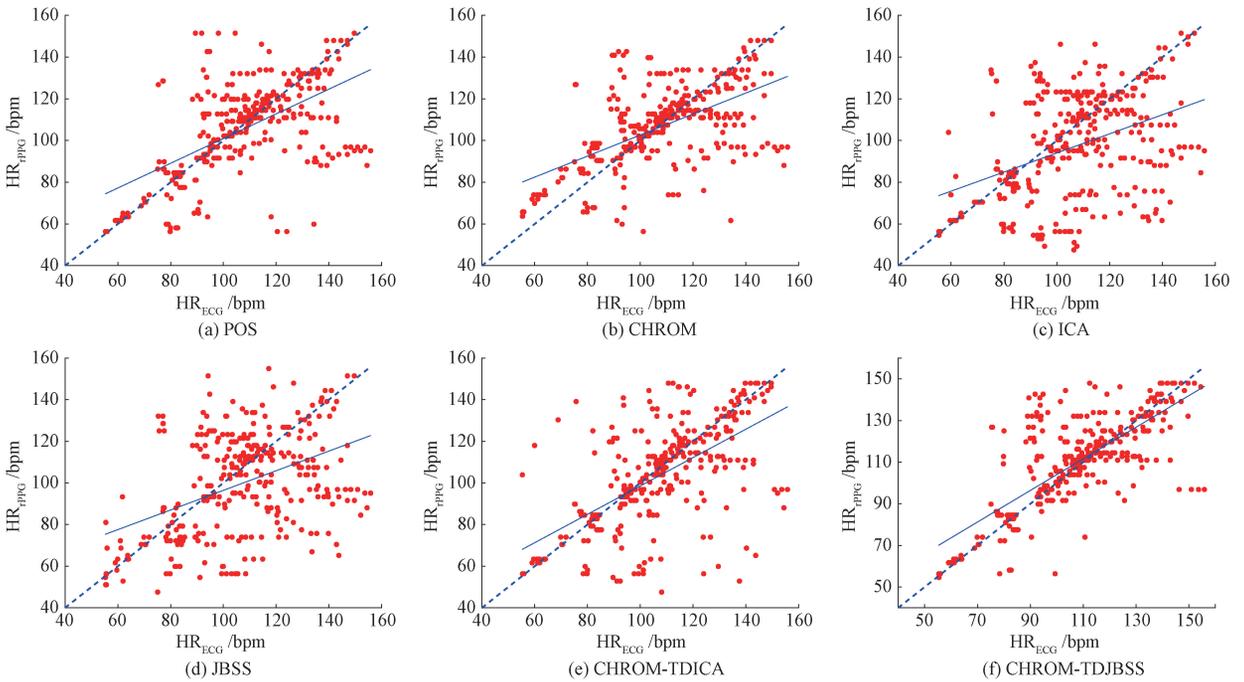


图 4 ECG-Fitness 数据库上不同方法测量 HR 的散点图

Fig. 4 Scatter plot of HR measured by different methods on ECG-Fitness database

表 4 ECG-Fitness 数据库上不同运动场景下 HR 测量结果对比

Table 4 Results comparison of different methods under different exercise scenarios on ECG-Fitness database

方法	划船机			椭圆机			固定自行车			说话		
	HR_{mae} /bpm	HR_{rmse} /bpm	r									
POS ^[6]	10.67	15.90	0.39	19.53	27.39	-0.06	12.12	20.61	0.29	9.78	17.59	0.59
CHROM ^[11]	10.36	15.08	0.37	13.45	21.58	0.37	14.19	21.73	0.41	14.17	17.61	0.58
ICA ^[17]	21.80	29.13	-0.14	18.16	26.16	0.22	21.97	29.09	0.31	12.90	20.83	0.47
JBSS ^[14]	18.23	25.58	0.07	25.19	31.49	-0.22	21.19	28.47	0.26	12.13	16.34	0.67
CHROM-TDICA	11.99	17.89	0.30	12.36	19.26	0.61	15.01	23.44	0.37	12.56	22.25	0.38
CHROM-TDJBSS	9.25	14.01	0.54	11.61	18.20	0.64	10.44	17.64	0.53	8.61	14.36	0.78

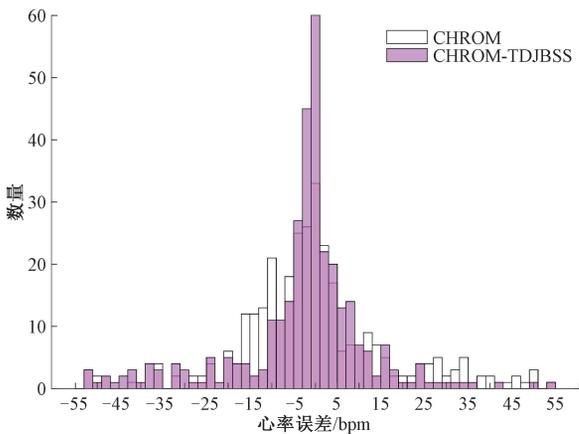


图 5 ECG-Fitness 数据库上 HR 误差分布

Fig. 5 HR error distribution on ECG-Fitness database

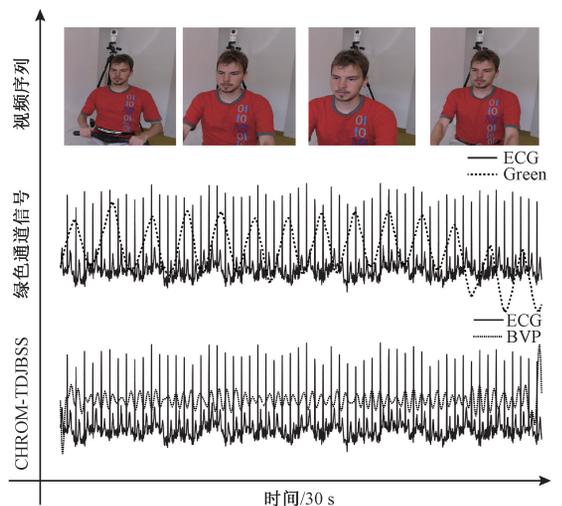


图 6 CHROM-TDJBSS 提取的 BVP 信号示意图

Fig. 6 Illustration of BVP signal extracted by CHROM-TDJBSS

凸显,再次印证了所提方法在去除运动幅值较大甚至规律性运动伪迹的优越性。

4 结 论

本文提出了一种基于模型和盲源分离相结合的运动鲁棒视频心率检测方法,使用色度信号代替原始 RGB 信号,初步实现运动伪迹去噪,后通过时延操作构建两个多通道数据集,一方面适合 JBSS 输入,另一方面凸显准周期性的 HR 变量和运动伪迹变量;利用 JBSS 提取两个数据集中共同存在的 BVP 分量向量,最终实现了运动伪迹的有效去除。

实验结果表明,运动伪迹是影响视频心率检测的重要因素。由于本文采用公开数据集可以预先知道运动的强度和周期性与否,可以根据信号特点有针对性地处理。然而,现实应用中通常无法获得全部的运动伪迹信息,因此,提出不依赖预先运动类型先验知识的运动鲁棒视频心率方法具有重要意义。其次,所提方法总体上取得了最好的结果,但是在源筛选方面尚不够智能,包括实验验证了前额 ROI 排序第二的 SCV 为目标 BVP 信号。未来,可以通过充分利用脉冲信号的分布规律,提出自适应筛选最佳源分量向量技术完善心率检测性能。此外,当运动剧烈时可能导致特征点追踪存在偏差,可以采用更先进的人脸特征点检测和追踪算法,促进该项技术的实用化进程。

参考文献

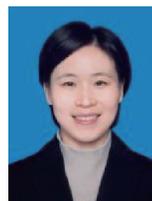
- [1] 刘今越,刘浩,贾晓辉,等. 基于视觉的非接触呼吸频率自动检测方法[J]. 仪器仪表学报, 2019(2): 51-58.
LIU J Y, LIU H, JIA X H, et al. Vision-based method for automatic non-contact respiratory rate detection[J]. Chinese Journal of Scientific Instrument, 2019(2): 51-58.
- [2] 淳新益,郑秀娟,张畅,等. 遗传算法优化 BP 神经网络的非接触式血压估计方法[J]. 电子测量与仪器学报, 2023, 35(7): 53-59.
CHUN X Y, ZHEN X J, ZAHNG CH, et al. Genetic algorithm optimization of BP neural network for non-contact blood pressure estimation [J]. Journal of Electronic Measurement and Instrumentation, 2023, 35(7): 53-59.
- [3] PREMKUMAR S, HEMANTH D J. Intelligent remote photoplethysmography-based methods for heart rate

estimation from face videos: A survey[J]. Informatics, 2022: 57.

- [4] 牛雪松,韩琥,山世光. 基于 rPPG 的生理指标测量方法综述[J]. 中国图象图形学报, 2022, 25(11): 2321-2336.
NIU X S, HAN H, SHAN SH G. Remote photoplethysmography-based physiological measurement: A survey [J]. Journal of Image and Graphics, 2022, 25(11): 2321-2336.
- [5] CHEN Q, JIANG X, LIU X, et al. Non-contact heart rate monitoring in neonatal intensive care unit using RGB camera[C]. 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), 2020: 5822-5825.
- [6] WANG W, DEN BRINKER A C, STUIJK S, et al. Algorithmic principles of remote PPG [J]. IEEE Transactions on Biomedical Engineering, 2016, 64(7): 1479-1491.
- [7] SONG R, LI J, WANG M, et al. Remote photoplethysmography with an eemd-mcca method robust against spatially uneven illuminations[J]. IEEE Sensors Journal, 2021, 21(12): 13484-13494.
- [8] LOKENDRA B, PUNEET G. AND-rPPG: A novel denoising-rPPG network for improving remote heart rate estimation [J]. Computers in Biology and Medicine, 2022, 141: 105146.
- [9] SONG R, CHEN H, CHENG J, et al. PulseGAN: Learning to generate realistic pulse waveforms in remote photoplethysmography[J]. IEEE Journal of Biomedical and Health Informatics, 2021, 25(5): 1373-1384.
- [10] LIN K Y, CHEN D Y, TSAI W J. Face-based heart rate signal decomposition and evaluation using multiple linear regression[J]. IEEE Sensors Journal, 2016, 16(5): 1351-1360.
- [11] DE H G, JEANNE V. Robust pulse rate from chrominance-based rPPG [J]. IEEE Transactions on Biomedical Engineering, 2013, 60(10): 2878-2886.
- [12] THARWAT A. Independent component analysis: An introduction[J]. Applied Computing Informatics, 2021, 17(2): 222-249.
- [13] RYU J, HONG S, LIANG S, et al. A new framework for

- robust heart rate measurement based on the head motion state estimation[J]. *IEEE Journal of Biomedical Health Informatics*, 2021, 25(9): 3428-3437.
- [14] CHENG J, WANG P, SONG R, et al. Remote heart rate measurement from near-infrared videos based on joint blind source separation with delay-coordinate transformation[J]. *IEEE Transactions on Instrumentation and Measurement*, 2021, 70: 5005313.
- [15] YIN R N, JIA R S, CUI Z, et al. PulseNet: A multitask learning network for remote heart rate estimation[J]. *Knowledge-Based Systems*, 2022, 239: 108048.
- [16] DE H G, VAN L A. Improved motion robustness of remote-PPG by using the blood volume pulse signature[J]. *Physiological Measurement*, 2014, 35(9): 1913-1926.
- [17] POH M Z, MCDUFF D, PICARD R W. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation[J]. *Optics express*, 2010, 18(10): 10762-10774.
- [18] LEWANDOWSKA M, RUMI ŃSKI J, KOCEJKO T, et al. Measuring pulse rate with a webcam— A non-contact method for evaluating cardiac activity[C]. 2011 federated Conference on Computer Science and Information Systems (FedCSIS), 2011: 405-410.
- [19] SONAWANI S, PATIL K, NATARAJAN P. Biomedical signal processing for health monitoring applications: A review[J]. *International Journal of Applied Systemic Studies*, 2023, 10(1): 44-69.
- [20] DEMIREZEN H, EROGLU ERDEM C. Heart rate estimation from facial videos using nonlinear mode decomposition and improved consistency check [J]. *Signal, Image Video Processing*, 2021, 15(7): 1415-1423.
- [21] GUO Z, WANG Z, SHEN Z. Physiological parameter monitoring of drivers based on video data and independent vector analysis[C]. 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2014: 4374-4378.
- [22] CHENG J, CHEN X, XU L, et al. Illumination variation-resistant video-based heart rate measurement using joint blind source separation and ensemble empirical mode decomposition [J]. *IEEE Journal of Biomedical and Health Informatics*, 2017, 21(5): 1422-1433.
- [23] SONG R, ZHANG S, CHENG J, 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.
- [24] BALTRUSAITIS T, ZADEH A, LIM Y C, et al. Openface 2.0: Facial behavior analysis toolkit[C]. 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), 2018: 59-66.
- [25] SONG R, ZHANG S, LI C, et al. Heart rate estimation from facial videos using a spatiotemporal representation with convolutional neural networks [J]. *IEEE Transactions on Instrumentation and Measurement*, 2020, 69(10): 7411-7421.
- [26] LIU X, JIANG Z, FROMM J, et al. MetaPhys: Few-shot adaptation for non-contact physiological measurement[C]. *Proceedings of the Conference on Health, Inference, and Learning*, 2021: 154-163.
- [27] ŠPETLIČ R, FRANČ V, MATAS J. Visual heart rate Estimation with convolutional neural network [C]. *Proceedings of the British Machine Vision Conference*, 2018: 3-6.

作者简介



成娟,2008年于中国科学技术大学获得学士学位,2013年于中国科学技术大学获得博士学位,现为合肥工业大学仪器科学与光电工程学院生物医学工程系教授,主要研究方向为智能健康监护。

E-mail: chengjuan@hfut.edu.cn

Cheng Juan received her B.Sc. degree from University of Science and Technology of China in 2008, and received her Ph.D. degree from University of Science and Technology of China in 2013. She is currently a professor at Hefei University of Technology. Her main research interest is intelligent health monitoring.



武梦影, 2020 年于中北大学获得学士学位, 现为合肥工业大学仪器科学与光电工程学院生物医学工程系硕士研究生, 现为主要研究方向为智能健康监护。

E-mail: wumengying@hfut.edu.cn

Wu Mengying received her B. Sc. degree from North University of China in 2020. She is currently a postgraduate student at Hefei University of Technology. Her main research interest is intelligent health monitoring.



宋仁成 (通信作者), 2005 年于吉林大学获得学士学位, 2010 年于浙江大学获得博士学位, 现为合肥工业大学仪器科学与光电工程学院生物医学工程系副研究员, 主要研究方向为非接触智能感知和情感人机交互。

E-mail: rcsong@hfut.edu.cn

Song Rencheng (Corresponding author) received his B. Sc. degree from Jilin University in 2005, and received his Ph. D. degree from Zhejiang University in 2010. He is currently an associate professor at Hefei University of Technology. His main research interests include non-contact intelligent perception and human-computer emotional interaction.



刘羽, 2011 年于中国科学技术大学获得学士学位, 2016 年于中国科学技术大学获得博士学位, 现为合肥工业大学仪器科学与光电工程学院生物医学工程系副教授, 主要研究方向为图像处理、计算机视觉。

E-mail: yuliu@hfut.edu.cn

Liu Yu received his B. Sc. degree from University of Science and Technology of China in 2011, and received his Ph. D. degree from University of Science and Technology of China in 2016. He is currently an associate professor at Hefei University of Technology. His main research interests include image processing and computer vision.



陈勋, 2009 年于中国科学技术大学获得学士学位, 2014 年于不列颠哥伦比亚大学获得博士学位, 现为中国科学技术大学教授, 主要研究方向为医学人工智能和移动健康监护。

E-mail: xunchen@ustc.edu.cn

Chen Xun received his B. Sc. degree from University of Science and Technology of China in 2009, and received his Ph. D. degree from the University of British Columbia in 2014. He is currently a professor at University of Science and Technology of China. His main research interests include medical artificial intelligence and mobile health monitoring.