

中图分类号: TN911.7; TP751.1 文献标识码: A 文章编号: 1006-8961(2024)07-2024-11

论文引用格式: Cheng J, Yin C C, Song R C, Fu J and Liu Y. 2024. Facial video-based heart rate measurement against irregular motion artifacts. Journal of Image and Graphics, 29(07):2024-2034(成娟, 殷辰楚, 宋仁成, 付静, 刘羽. 2024. 非规律运动伪迹干扰鲁棒的人脸视频心率检测. 中国图象图形学报, 29(07):2024-2034)[DOI: 10.11834/jig.230428]

非规律运动伪迹干扰鲁棒的人脸视频心率检测

成娟^{1,2}, 殷辰楚¹, 宋仁成¹, 付静¹, 刘羽^{1*}

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

摘要: 目的 基于远程光电容积脉搏波描记法(remote photoplethysmograph, rPPG)的非接触人脸视频心率检测广泛应用于移动健康监护领域, 由于其携带的生理参数信息幅值微弱, 容易受到运动伪迹干扰。据此, 提出了一种结合非负矩阵分解(nonnegative matrix factorization, NMF)和独立向量分析(independent vector analysis, IVA)的非规律运动伪迹去除的视频心率检测方法, 记为NMF-IVA。方法 首先, 将面部感兴趣区域(region of interest, RoI)分为多个子区域(sub RoIs, SRoIs), 利用平均光照强度、光照强度变化、信噪比这3个指标筛选出3个最优的SRoIs, 并获取每个SRoI的绿色通道时间序列。其次, 将3个绿色通道时间序列去趋势、带通滤波后送入NMF-IVA进行盲源分离。然后, 对分离后的源信号进行功率谱密度分析, 并且将峰值信噪比最高且主频落在心率感兴趣范围内的源信号确定为血容量脉冲(blood volume pulse, BVP)信号。最后, 将BVP信号的主频确定为所测量心率的主频, 从而计算出心率值。结果 实验在2个公开数据集UBFC-RPPG和UBFC-PHYS, 及1个真实场景自采数据集上与最相关的7种典型的rPPG方法进行比较, 在UBFC-RPPG数据集上, 相比于性能第2的单通道滤波(single channel filtering, SCF)方法, 均方根误差提升了1.39 bpm(beat per minute)、平均绝对误差提升了1.25 bpm、皮尔逊相关系数提升了0.02; 在UBFC-PHYS数据集上的T2情况下, 其性能提升最为显著, 相比于性能第2的独立向量分析(IVA)方法, 均方根误差提升了16.42 bpm、平均绝对误差提升了9.91 bpm、皮尔逊相关系数提升了0.64; 在自采数据集上, 除了低于深度学习性能之外, 所提NMF-IVA方法在传统方法中取得了最好的结果。结论 所提NMF-IVA方法对规律信号提取具有敏感性, 即便是在头部存在剧烈非规律运动情况下, 相比于传统方法亦能取得最优结果, 该结果能够媲美基于深度学习的方法。

关键词: 远程光电容积脉搏波描记法(rPPG); 非接触式心率检测; 盲源分离(BSS); 非负矩阵分解(NMF); 独立向量分析(IVA)

Facial video-based heart rate measurement against irregular motion artifacts

Cheng Juan^{1,2}, Yin Chenchu¹, Song Rencheng¹, Fu Jing¹, Liu Yu^{1*}

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

Abstract: Objective Heart rate (HR) is one of the most important physiological parameters that can reflect the physical and mental status of individuals. Various methods have been developed to estimate HR values using contact and noncontact

收稿日期: 2023-07-11; 修回日期: 2023-10-17; 预印本日期: 2023-10-23

* 通信作者: 刘羽 yuliu@hfut.edu.cn

基金项目: 国家自然科学基金项目(62171176, 62176081, 62271186); 中央高校基本科研业务费专项资金资助(JZ2021HGPA0061)

Supported by: National Natural Science Foundation of China (62171176, 62176081, 62271186); Fundamental Research Funds for the Central Universities (JZ2021HGPA0061)

sensors. The advantage of noncontact methods is that they provide a more comfortable and unobtrusive way to estimate HR and avoid discomfort or skin allergy caused by conventional contact methods. The pulse-induced subtle color variations of facial skins can be measured from consumer-level cameras. Thus, camera-based non-contact HR detection technology, also called remote photoplethysmograph (rPPG), has been widely used in the fields of mobile health monitoring, driving safety, and emotion awareness. The principle of camera-based rPPG measurement is similar to that of traditional PPG measurement, that is, the pulsatile blood propagating in cardiovascular systems changes blood volumes in microvascular tissue beds beneath skins with each heartbeat, thus producing a fluctuation. However, such technology is susceptible to motion artifacts due to weak amplitudes of the physiological parameter information it carries. For instance, subjects' heads may move involuntarily during interviews, presentations, and other socially stressful situations, thus degrading rPPG-based HR detection performance. Accordingly, this paper proposes a novel motion-robust rPPG method that combines nonnegative matrix factorization (NMF) and independent vector analysis (IVA), termed as NMF-IVA, to remove irregular motion artifacts. **Method** First, the whole facial region of interest (RoI) is divided into several sub RoIs (SRoIs), among which three optimal SRoIs are selected based on three indicators: average light intensity, light intensity variation of a certain SRoI, and signal-to-noise ratio (SNR) of the green-channel signal derived from the SRoI. Afterwards, three green-channel time series are derived from the corresponding three optimal SRoIs. Second, the three channels of time series are detrended, bandpass filtered, and then sent to the proposed NMF-IVA as input. After the NMF-IVA operation, three source signals are extracted and then processed by power spectral density analysis. The one with the highest peak SNR and the corresponding dominant frequency falling within the interested HR range will be identified as the blood volume pulse (BVP) signal, whose dominant frequency is identified as that of the estimated HR. **Result** We compare the proposed NMF-IVA method with seven typical rPPG methods on two publicly available datasets (UBFC-RPPG and UBFC-PHYS) as well as one in-house dataset. On the UBFC-RPPG dataset, compared with the second-best performance of the single channel filtering (SCF) method, the proposed NMF-IVA achieves better performance, with an improved root mean square error (RMSE) of HR measurement by 1.39 beat per minute (bpm), an improved mean absolute error (MAE) by 1.25 bpm, and a higher Pearson's correlation coefficient (PCC) by 0.02. Although both the MAE and the RMSE achieved by the proposed NMF-IVA method are lower than those of deep learning-based methods, the PCC of the NMF-IVA is comparable to that of deep learning-based ones, which demonstrates the effectiveness of the proposed NMF-IVA method. As for the UBFC-PHYS dataset when compared with traditional rPPG methods, during the T1 condition, the performance of the proposed NMF-IVA method is better than that of the second-best SCF method, with an improved RMSE by 6.45 bpm, an improved MAE by 2.53 bpm, and a higher PCC by 0.18. When compared with deep learning-based ones, the proposed NMF-IVA method achieves the second-best performance. The performance improvement of the proposed NMF-IVA is most noticeable during the T2 condition on the UBFC-PHYS dataset. Specifically, when compared with the second-best performance of IVA, the above three metrics are improved by 16.42 bpm, 9.91 bpm, and 0.64, respectively. As for the UBFC-PHYS dataset, when during the T3 condition, the best performance is still achieved by the proposed NMF-IVA method. When compared with the second-best performance of the independent component analysis method, the corresponding three metrics are improved by 8.54 bpm, 6.14 bpm, and 0.37, respectively. The performance of the proposed NMF-IVA method can be comparable to that of deep learning-based ones both in T2 and T3 conditions. As for the in-house dataset, the proposed NMF-IVA method achieves better performance compared with the traditional methods, except for deep learning-based methods. **Conclusion** The proposed NMF-IVA method achieves the best results on all the three datasets when compared with traditional rPPG methods, and the performance improvement is most noticeable during irregular motion artifact conditions involving head motions with large amplitudes. However, the performance of the proposed NMF-IVA method is slightly poorer than that of deep learning-based methods possibly because deep learning technology has excellent abilities in learning and extracting effective features. However, sufficient training samples and generalization should be considered when adopting deep learning-based methods. In addition, before the high-quality BVP source is derived, upsampling is employed, which leads to a relatively large time consumption. In the future, the HR estimation performance and the upsampling rate should be traded off. The proposed NMF-IVA method has advantages in extracting regular signals. Thus, our study can provide a new solution for promoting the practical application ability of rPPG technology.

Key words: remote photoplethysmograph (rPPG); non-contact heart rate measurement; blind source separation (BSS); nonnegative matrix factorization (NMF); independent vector analysis (IVA)

0 引言

心率(heart rate, HR)是心血管疾病诊断中的重要生理指标,可通过光电容积脉搏波描记法(photo plethysmograph, PPG)进行无创检测(Biswas等, 2019),但是PPG传感器需要与人体皮肤接触,往往造成不便。研究表明,血流变化会导致面部肤色相应变化,这种变化可以采用消费级摄像头捕捉到(牛雪松等, 2020),为心率检测提供了一种低成本、广泛应用的非接触检测方法,这种基于视频的心率检测技术可称为远程光电容积脉搏波描记法(remote PPG, rPPG)。

众所周知,通过摄像头捕捉到的生理信号幅值十分微弱,容易受到光照、运动伪迹等噪声的干扰,造成血容量脉冲(blood volume pulse, BVP)信号失真(杨昭等, 2018)。因此,如何提高rPPG技术的噪声鲁棒性,特别是运动伪迹鲁棒性是保障rPPG技术性能的关键。针对rPPG运动伪迹去除问题,现有研究已经提出了许多方法,可以分为基于模型的方法和基于盲源分离(blind source separation, BSS)的方法(Chen等, 2019)。伴随人工智能的发展,基于深度学习的方法应运而生,具体参见Yu等人(2021)的综述。

基于模型的方法利用颜色向量信息进行分量的推导和分离,共同点是消除皮肤反射模型对平均皮肤反射颜色通道的依赖性。典型的方法有基于色度模型(chrominance, CHROM)方法(den Brinker和Wang, 2022)、基于血容量脉冲方法(Wang等, 2020)以及基于与皮肤正交的平面(plane orthogonal to skin, POS)方法(Zhao等, 2022)。杨昭等人(2018)通过引入色度特征和空间梯度来消除脸部运动干扰,使用粗估计和精估计相结合,成功提取出准确的BVP波形。

基于BSS的典型方法是独立成分分析(independent component analysis, ICA)。将从单个面部感兴趣区域(region of interest, RoI)分别获取的RGB三通道信号,或从多个RoIs分别提取的单通道绿色通道信号作为ICA的输入,从而获得BVP信号(Poh等, 2011)。基于BSS的方法在相对静止或运动幅度不

剧烈的情况下能有效抑制运动伪迹,但若多个源之间不满足独立性等假设,如运动噪声幅值过大,其性能会急剧下降。

之后,Cheng等人(2021)和Song等人(2021c)利用多个面部RoI的血流信息,使用了联合盲源分离(joint BSS, JBSS)算法,包括独立向量分析(independent vector analysis, IVA)和多集典型相关分析(multiset canonical correlation analysis, MCCA),均取得了比基于BSS方法更优的性能。但是,当不同的RoIs存在相同的刚性运动时,该运动也会被作为共同的源成分矢量(source component vector, SCV)提取,影响心率检测结果。同时,如何确定目标BVP信号是基于JBSS方法需要考虑的问题,通常需要谱聚类或设定某种准则来实现。

Singh和Raj(2020)采用非负矩阵分解(nonnegative matrix factorization, NMF)来获得BVP信号。NMF能够将原始信号分解为特征矩阵和激活矩阵的组合,但是,单纯使用NMF只包含分解过程,无法将特征聚类,不适合直接处理盲源分离问题。而上述IVA算法虽然能够利用不同数据集间的共同信息进行源分离,但是它的源先验模型通常为超高斯球面分布(Ruan等, 2022),无法实现具有特定频谱信号的有效分离。据此,该文提出了一种结合NMF和IVA的非规律运动伪迹去除的人脸视频心率检测方法,记为NMF-IVA。该方法利用NMF分解混合信号的特征矩阵,将其组合构建IVA的源先验模型,提高IVA对于规律性BVP信号的源分离性能,同时达到去除非规律性运动伪迹干扰的目的。

本文的主要贡献如下:1)提出了一种NMF和IVA相结合的抑制非规律运动伪迹干扰的人脸视频心率检测方法,实现了具有准周期性特性BVP信号的有效提取,从而提高非规律运动伪迹情况下rPPG的运动鲁棒性。2)所提方法在公开数据集UBFC-RPPG和UBFC-PHYS,及自采数据集上进行了测试,场景涉及相对静止、轻微运动和大幅度运动,并与几种典型的传统rPPG算法以及几种最新的基于深度学习的rPPG算法进行性能对比。结果表明,所提方法相比传统rPPG方法性能最佳,能够媲美一些深度学习方法。

1 方法

1.1 独立向量分析

IVA是ICA的多元扩展(Algumaei等,2023),为解决ICA的置换问题。IVA通常假定通过 N 个通道捕捉到 M 个信号,而这 M 个信号是目标信号和噪声信号的混合信号。假定混合信号为

$$\mathbf{X} = \mathbf{A}_i \mathbf{S} \quad (1)$$

式中, $i = 1, 2, \dots, I$ 是单个通道多维向量索引, $j = 1, 2, \dots, J$ 是帧索引, $\mathbf{X} = [x_{ij,1}, x_{ij,2}, \dots, x_{ij,M}]^T$ 是每个通道的多维向量组成的三维矩阵,是 N 个通道的观测信号。 $\mathbf{S} = [s_{ij,1}, s_{ij,2}, \dots, s_{ij,M}]^T$ 为真实源信号, $\mathbf{A}_i = [\mathbf{a}_{i,1}, \mathbf{a}_{i,2}, \dots, \mathbf{a}_{i,M}]^H$ 是混合矩阵(其中 $\mathbf{a}_{i,m}$ 是第 m 个源的方向向量,上标H表示赫米特转置)。

基于IVA的源分离过程表示为

$$\mathbf{S} \approx \mathbf{Y} = \mathbf{W}_i \mathbf{X} \quad (2)$$

式中, $\mathbf{Y} = [y_{ij,1}, y_{ij,2}, \dots, y_{ij,M}]^T$ 代表源信号的估计矩阵, $\mathbf{W}_i = [w_{i,1}, w_{i,2}, \dots, w_{i,M}]^H$ 为解混合矩阵。

IVA的源向量变量的先验假设 $p(s_{j,m})$ 为超高斯球面多元分布(通常为球形拉普拉斯分布),并不能包含源向量变量的频谱结构信息,具体为

$$p(s_{j,m}) = \rho \exp\left(-\sqrt{\sum_i \left|\frac{s_{ij,m}}{r_{j,m}}\right|^2}\right) \quad (3)$$

式中, ρ 是归一化项, $r_{j,m}$ 是频率区间上的源模型方差。

1.2 非负矩阵分解

NMF是一种稀疏表示算法,将一个非负矩阵分解为两个非负矩阵,具体为

$$\mathbf{D} = \mathbf{T}\mathbf{V} \quad (4)$$

式中, $\mathbf{D} (\in \mathbf{R}_{\geq 0}^{I \times J})$ 是非负矩阵, $\mathbf{T} (\in \mathbf{R}_{\geq 0}^{I \times L})$ 是基矩阵,即特征矩阵。而 $\mathbf{V} (\in \mathbf{R}_{\geq 0}^{L \times J})$ 是激活矩阵,它以矩阵 \mathbf{T} 中每个基的时变增益作为行向量。 L 表示分解出的基的数量,即特征的数量。这些基必须经由聚类手段才能进行盲信号分离。

1.3 NMF-IVA

NMF的损失函数可以表示为

$$Q_{\text{NMF}} = \sum_{i,j} \left(\frac{d_{ij}}{\sum_l t_{il} v_{lj}} + \log \sum_l t_{il} v_{lj} \right) \quad (5)$$

式中, t_{il}, v_{lj} 和 d_{ij} 分别代表 \mathbf{D}, \mathbf{T} 和 \mathbf{V} 中的非负值。

IVA的损失函数可以表示为

$$Q_{\text{IVA}}(\mathbf{W}_i) = \sum_m \frac{1}{J} \sum_j G(\mathbf{Y}) - \sum_i \log |\det \mathbf{W}_i| \quad (6)$$

式中, $G(\mathbf{Y})$ 为对比函数,当源先验模型假设为拉普拉斯球分布时,可得到 $G(\mathbf{Y}) = -\log p(s_{ij,m})$ 。

所提NMF-IVA方法充分发挥两者的优势,通过结合两者的损失函数,实现利用NMF分解的特征矩阵作为IVA的源先验来捕获源信号的频谱结构信息,NMF-IVA新的损失函数为

$$Q = \left[\sum_m \log \sum_l t_{il,m} v_{lj,m} + \sum_m \left(\frac{|y_{ij,m}|^2}{\sum_l t_{il,m} v_{lj,m}} - 2 \log |\det \mathbf{W}_i| \right) \right] \quad (7)$$

式中, $t_{il,m}, v_{lj,m}$ 和 $y_{ij,m}$ 分别为特征矩阵 $\mathbf{T}_m \in \mathbf{R}^{I \times J}$ 、激活矩阵 $\mathbf{V}_m \in \mathbf{R}^{L \times J}$ 和估计矩阵 \mathbf{Y} 中的非负值。

改进后的IVA对比函数为

$$G(\mathbf{Y}) = \sum_m \left(\frac{|y_{ij,m}|^2}{\sum_l t_{il,m} v_{lj,m}} + \sum_m \log \sum_l t_{il,m} v_{lj,m} \right) \quad (8)$$

由式(8)可以看出,NMF-IVA通过分解混合矩阵 \mathbf{D} ,得到各信源信号的特征 $t_{il,m}$ 和激活 $v_{lj,m}$,并将得到的特征和激活重构为源模型中的方差项 $r_{j,m} = \sum_l t_{il,m} v_{lj,m}$,完成信源信号特征融入IVA算法的计算过程,使得IVA能够有效提取具有明显特征的源信号。

2 实验与结果

2.1 实验流程

图1所示为所提基于NMF-IVA的人脸视频心率检测流程图,主要包含4个步骤。1)确定人脸完整的RoI,筛选出3个最优面部子RoIs,分别提取单通道绿色通道信号;2)对每个绿色通道信号进行预处理;3)将预处理后的绿色通道送入NMF-IVA进行源信号提取;4)对源信号进行功率谱分析,确定目标BVP信号,计算心率值。

2.1.1 RoI区域筛选和信号预处理

采用通用的RoI区域筛选技术(Song等,2021b),首先,对第1帧视频图像使用Viola-Jones人脸检测器获得面部区域矩形框,通过比例缩放去除背景和毛发等非皮肤区域。其次,将该完整的RoI划分为 $N \times N$ 个子区域。采用Kanade-Lucas-Tomasi

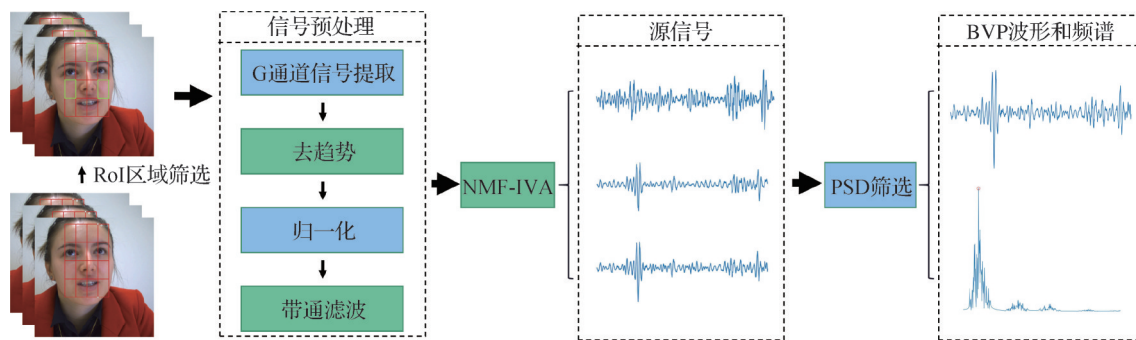


图1 基于NMF-IVA人脸视频心率检测流程图

Fig. 1 Flow chart of NMF-IVA based facial video heart rate detection

算法跟踪后续视频帧中的每个子RoI (sub RoI, SRoI)。之后,采用所提方法确定3个最优SRoIs。如图2所示,绿色矩形框即为确定的SRoIs。最后,对每个SRoI逐帧像素平均获得绿色通道信号,并做预处理,包括去趋势和带通滤波,去趋势的对应参数 $\lambda=20$,带通滤波的通带范围设置为 $0.75\sim 3$ Hz,对应的感兴趣心率值是 $42\sim 180$ bpm (beat per minute)。

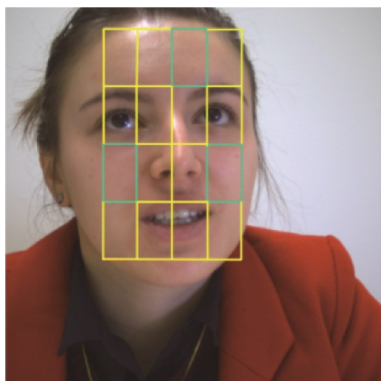


图2 筛选后的高质量RoI子区域示意图

Fig. 2 Illustration of the selected high-quality sub RoIs

2.1.2 基于NMF-IVA的BVP信号获取

将预处理后的三通道信号作为NMF-IVA的输入,经过NMF-IVA处理,生成三通道源信号,并进行功率谱密度(power spectral density, PSD)分析,将频率落在感兴趣心率范围内最多的源信号确定为BVP信号,之后计算BVP信号的主频,即为测量心率的主频,从而计算平均心率值。

2.2 实验数据库

为了验证所提方法的性能,采用两个公共数据库UBFC-RPPG和UBFC-PHYS及一个自采数据集进行测试。

1) UBFC-RPPG数据库(Bobbia等,2019)。包含

42个受试者的42个视频。这些视频时长约1 min,由罗技C920 HD pro RGB相机使用未压缩的8位avi格式保存,分辨率为 640×480 像素,帧率(frame per second, fps)为30帧/s。参考PPG信号由Contec Medical CMS50E记录。由于要求在电脑前玩数学游戏,许多受试者有明显的头部运动。

2) UBFC-PHYS数据库(Sabour等,2023)。包含来自56位受试者的3个部分168个视频,每个视频持续时间为3 min,使用Edmund Optics的EO23121C RGB数码相机进行拍摄,帧率是35帧/s,分辨率为 1024×1024 像素。采用Motion JPEG压缩保存。参考BVP采用Empatica E4腕带同步采集。数据库分为3部分:1)静止情况(T1):受试者被要求保持静止并禁言;2)模拟演讲情况(T2):受试者被要求模拟求职面试或是进行假期的美好回忆分享;3)数字游戏情况(T3):受试者被要求进行等差数列形式的报数游戏。

两个公共数据库中的每个视频均截取成30 s的片段进行HR测量,其中UBFC-RPPG数据库截取的时间段为 $10\sim 40$ s,共形成42个视频片段;UBFC-PHYS数据库截取的时间段为 $120\sim 150$ s,同时剔除参考BVP信号质量较差的片段,最终形成101个视频片段(T1:41个,T2:26个,T3:33个)。

3) 自采数据集。包含来自12个受试者的12个视频,每个视频持续约4 min。由罗技C920 HD pro RGB相机使用未压缩的8位avi格式保存,分辨率为 640×480 像素,帧速率为30帧/s。参考PPG信号由Contec Medical CMS50E记录。受试者被要求参加面试对话,每个受试者均有大量明显的面部表情和头部运动。

该自采数据集中每个视频均截取成30 s的片段

进行HR测量,并采用5 s的步长进行重叠分段,最终共形成618个视频片段。

该文采用3种评价指标评价视频心率检测性能(Song等,2021b),具体为均方根误差 HR_{rmse} 、平均绝对误差 HR_{mae} 和皮尔逊相关系数 r 。

2.3 实验结果

2.3.1 UBFC-RPPG实验结果

对比的传统rPPG方法有单通道滤波(single channel filtering, SCF)(Chen等,2016)、CHROM(De Haan和Jeanvea,2013)、POS(Wang等,2017)、ICA(Poh等,2011)和IVA(Cheng等,2021)。其中与IVA的对比可作为NMF-IVA的消融实验。

表1是不同方法在UBFC-RPPG数据库上的对比实验结果。从表1可以看出,得益于NMF算法对于规律信号的特征分解能力和IVA的盲源分离能力,所提NMF-IVA取得了最好的结果。相较于IVA, HR_{rmse} 提升了1.39 bpm、 HR_{mae} 提升了1.25 bpm、相关系数 r 提升了0.02。SCF方法在该数据集上也取得了较好的检测性能,这主要归功于UBFC-RPPG的视频质量非常好(未压缩);其次,该数据集包含的运动伪迹幅值并不剧烈,也验证了绿色通道包含BVP信号质量较好。IVA相较于ICA方法, HR_{rmse} 提升了6.61 bpm、 HR_{mae} 提升了2.41 bpm、相关系数 r 提升了0.12,可能的原因是IVA能够提取共同存在于多个SRoI中的心率信息,一定程度上可以消除

非刚性运动伪迹。基于光学反射模型的方法CHROM和POS通常在UBFC-RPPG数据集上会取得很好的性能(Song等,2021a),但在该实验中,基于CHROM和POS并没有取得那么好的性能。可能的原因在于该文对RoI区域进行了划分,选择的SRoI面积相对较小,难以充分发挥光学反射模型的优势。

表1 UBFC-RPPG数据库上与传统rPPG方法的对比实验结果

Table 1 Results comparison with traditional rPPG methods on UBFC-RPPG database

方法	HR_{rmse} /bpm	HR_{mae} /bpm	r
SCF(Chen等,2016)	5.52	4.06	0.95
CHROM(De Haan和Jeanne,2013)	10.79	8.63	0.86
POS(Wang等,2017)	8.20	6.60	0.88
ICA(Poh等,2011)	11.74	6.47	0.83
IVA(Cheng等,2021)	5.58	4.06	0.95
NMF-IVA(本文)	4.19	2.81	0.97

注:加粗字体表示各列最优结果。

图3展示了每种方法的Bland-Altman图,体现不同方法估计心率值的一致性分布。从图3中看出,所提NMF-IVA获得的心率值与真实值具有最好的一致性,其1.96倍标准差为7.6 bpm,而性能次好的IVA方法的1.96倍标准差为10 bpm左右。

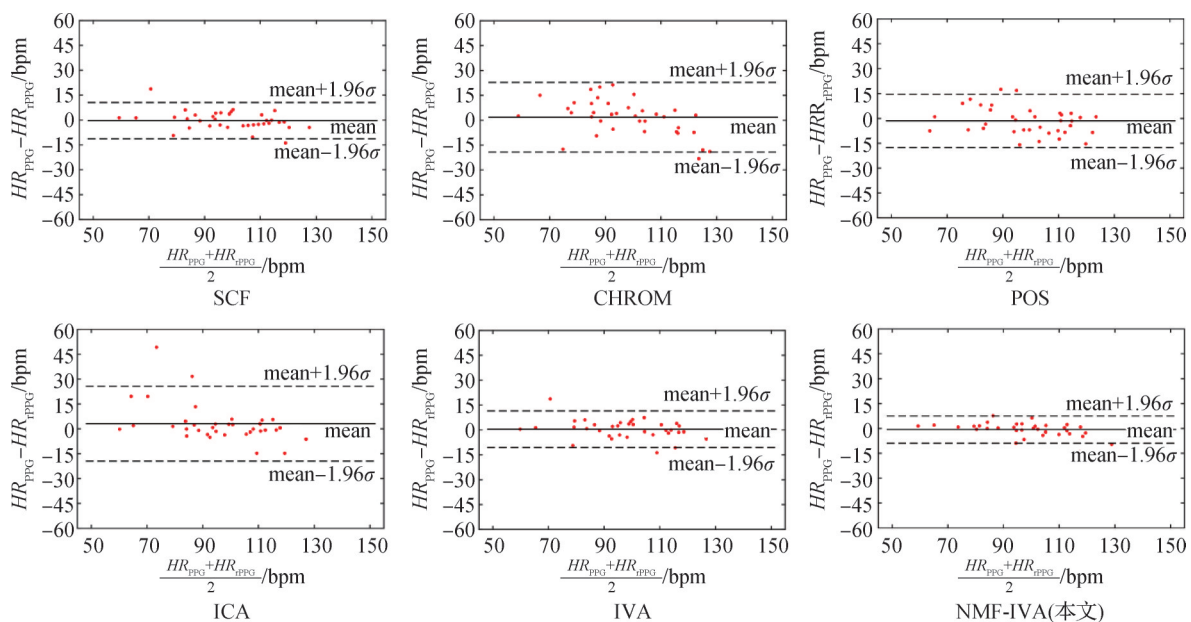


图3 NMF-IVA及对比算法在UBFC-RPPG数据库的Bland-Altman图

Fig. 3 Bland-Altman plots of NMF-IVA and comparison algorithms on UBFC-RPPG database

表2是UBFC-RPPG数据库上所提方法与几种最新的深度学习方法的结果对比。对比方法分别是Song等人(2021a)提出的PluseGAN、Hu等人(2022)提出的时空注意力网络、Li等人(2023)提出的多层神经网络、Tsou等人(2020)提出的Siamese-rPPG,以及Liu等人(2021b)提出的MetaPhys,这些结果来自于每种对比方法的论文中。

表2 UBFC-RPPG数据库上与深度学习方法的结果对比
Table 2 Results comparison with deep learning methods on UBFC-RPPG database

方法	$HR_{\text{rmse}}/\text{bpm}$	$HR_{\text{mae}}/\text{bpm}$	r
Song等人(2021a)	2.49	1.19	0.98
Hu等人(2022)	3.31	1.43	0.97
Li等人(2023)	3.82	2.15	0.97
Tsou等人(2020)	0.97	0.48	-
Liu等人(2021b)	2.62	1.90	0.96
NMF-IVA(本文)	4.19	2.81	0.97

注:加粗字体表示各列最优结果。“-”表示该处无数值。

从表2中不难发现,Tsou等人(2020)提出的方法取得了最小均方根误差 HR_{rmse} 和平均绝对误差 HR_{mae} ,分别为0.97 bpm和0.48 bpm。PluseGAN取得了最优的相关系数($r = 0.98$)。PluseGAN、Hu等人(2022)以及MetaPhys均采用了传统方法和深度学习方法结合的方式,共同点在于都使用了基于光学模型方法对RoI中捕获的颜色通道信号进行预处理,之后对网络进行训练,得到的心率检测性能优于Li等人(2023)提出的单纯深度学习方法。所提NMF-IVA的 HR_{rmse} 和 HR_{mae} 两项指标性能不如深度学习方法,再次说明基于深度学习的心率检测准确性普遍优于传统方法。但所提NMF-IVA在相关系数 r 上得到了与深度学习方法媲美的结果,说明了NMF-IVA的有效性。然而,面向实际应用时,基于深度学习的方法需要着重考虑模型训练时间长、可解释性差以及跨数据集的泛用性较差等问题。

2.3.2 UBFC-PHYS实验结果

表3为不同方法在UBFC-PHYS数据库不同子集T1、T2和T3上的对比实验结果。可以看出,CHROM方法没有取得良好的效果,主要原因是压缩视频会破坏皮肤光学反射模型假设(Song等,2021b)。SCF方法由于运动伪迹,其性能随着伪迹

剧烈程度增大急剧下降,以 HR_{mae} 为例,静止时SCF的 HR_{mae} 为8.53 bpm,存在少量运动时的 HR_{mae} 下降到21.46 bpm,而存在大量运动时的 HR_{mae} 下降到24.67 bpm。ICA方法在相对静止情况下可以取得较好的结果。ICA和IVA在T2、T3数据集上取得了较为相近的性能,但在T2存在较大幅值运动时,IVA算法性能比ICA方法略显稳定,体现为相关系数 r 分别为0.06和0.24。而本文所提NMF-IVA方法在3种情况下都取得了最好的效果,尤其是在包含大量运动的T2情况下, HR_{rmse} 为7.14 bpm、 HR_{mae} 为5.06 bpm,相关系数 r 为0.88。而结果次好的CHROM方法得到的 HR_{rmse} 为20.26 bpm,比NMF-IVA高出13.12 bpm;相关系数 r 为0.26,比NMF-IVA低0.62,充分说明了所提NMF-IVA方法抑制运动伪迹的有效性。

以T2情况为例,图4给出几种方法的Bland-Altman图。从图4可以看出,NMF-IVA方法具有最好的一致性,同时每个样本点的估计值与真实值的误差也最小。相较于IVA,NMF-IVA的1.96倍标准差为15 bpm,而IVA的1.96倍标准差为32 bpm。

表4为UBFC-PHYS数据库上采用深度学习方法和所提NMF-IVA结果对比。深度学习方法来自于工具箱,且调用在PURE数据集上预训练的模型(<https://github.com/ubicomplab/rPPG-Toolbox>),测试集与本文方法一致。不难发现,在T1场景下,EfficientPhys取得了最佳结果,而TSCAN(temporal shift convolutional attention network)在T2和T3场景下性能相对更好,说明TSCAN的运动鲁棒性相对更好。而所提方法也取得了不错的性能,在T1场景下取得了次好的结果,在T2和T3场景下其性能也能媲美深度学习方法。

为了进一步说明所提方法的运动鲁棒性,以UBFC-PHYS-T2-S55受试者为例,将经过NMF-IVA分离后获得的BVP源信号与参考PPG信号进行时域波形和频域频谱对比,同时与IVA获得的BVP源矢量分量对比,如图5所示。从图5中不难看出,当头部存在明显运动时,NMF-IVA分离出来的BVP信号与参考PPG信号的拟合度更高,而此时,IVA获得的BVP波形与参考PPG信号的拟合度较差。对比功率谱图可以发现,相较于IVA,NMF-IVA获得的BVP信号拥有更清晰的主频,其心率值也更接近真实心率值。

表 3 UBFC-PHYS 数据库上与传统 rPPG 方法的对比实验结果

Table 3 Results comparison with traditional rPPG methods on UBFC-PHYS database

方法	HR_{mse}/bpm			HR_{mac}/bpm			r		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
SCF(Chen 等, 2016)	8.53	24.67	21.46	3.69	17.65	12.70	0.80	0.10	0.27
CHROM(De Haan 和 Jeanne, 2013)	10.35	20.26	23.25	7.62	14.47	18.63	0.78	0.26	-0.02
POS(Wang 等, 2017)	10.50	28.92	25.15	7.98	22.35	18.59	0.74	-0.07	-0.02
ICA(Poh 等, 2011)	5.41	23.80	15.29	2.83	15.07	11.30	0.91	0.06	0.59
IVA(Cheng 等, 2021)	10.30	23.56	16.87	5.23	14.97	11.66	0.70	-0.24	0.57
NMF-IVA(本文)	2.08	7.14	6.75	1.16	5.06	5.32	0.98	0.88	0.91

注:加粗字体表示各列最优结果。

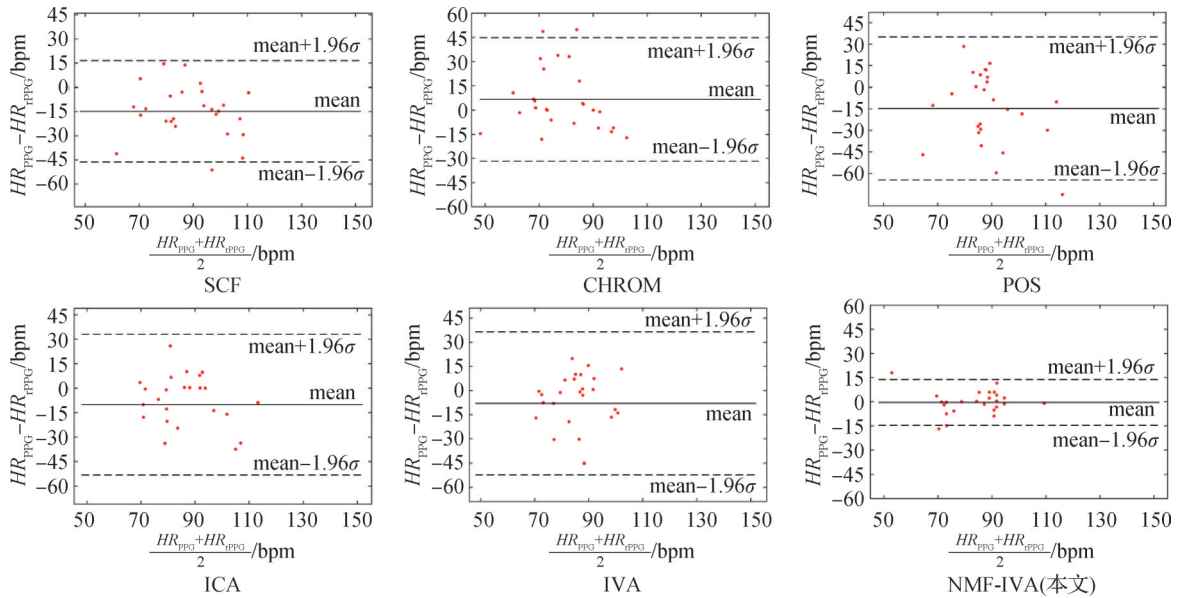


图 4 不同方法在 UBFC-PHYS-T2 上的 Bland-Altman 图

Fig. 4 Bland-Altman plots of different methods on UBFC-PHYS-T2

表 4 UBFC-PHYS 数据库上与深度学习方法的结果对比

Table 4 Results comparison with deep learning methods on UBFC-PHYS database

方法	HR_{mse}/bpm			HR_{mac}/bpm			r		
	T1	T2	T3	T1	T2	T3	T1	T2	T3
EfficientPhys(Liu 等, 2023)	1.23	<u>6.92</u>	<u>5.04</u>	0.73	<u>4.67</u>	<u>3.30</u>	0.99	0.83	<u>0.92</u>
TSCAN(Liu 等, 2021a)	2.15	5.43	4.44	1.41	3.13	3.08	0.99	0.91	0.96
NMF-IVA(本文)	<u>2.08</u>	7.14	6.75	<u>1.16</u>	5.06	5.32	<u>0.98</u>	<u>0.88</u>	0.91

注:加粗、下划线字体分别表示各列最优、次优结果。

2.3.3 自采数据集实验结果

表 5 给出了本文方法在真实场景自采数据集上与传统方法和深度学习方法的结果对比。不难发现,除了性能低于深度学习方法之外,本文所提 NMF-IVA 方法取得了最优的结果。值得注意的是,

由于自采数据集中,摄像头与受试者的距离进一步变大,约 1.5 m,导致视频图像中人脸 RoI 区域的进一步缩小,由 RoI 区域再分割产生的 SRoI 区域上进行 CHROM 或 POS 的性能并不好,但是本文所提 NMF-IVA 方法的性能仍然有保障。

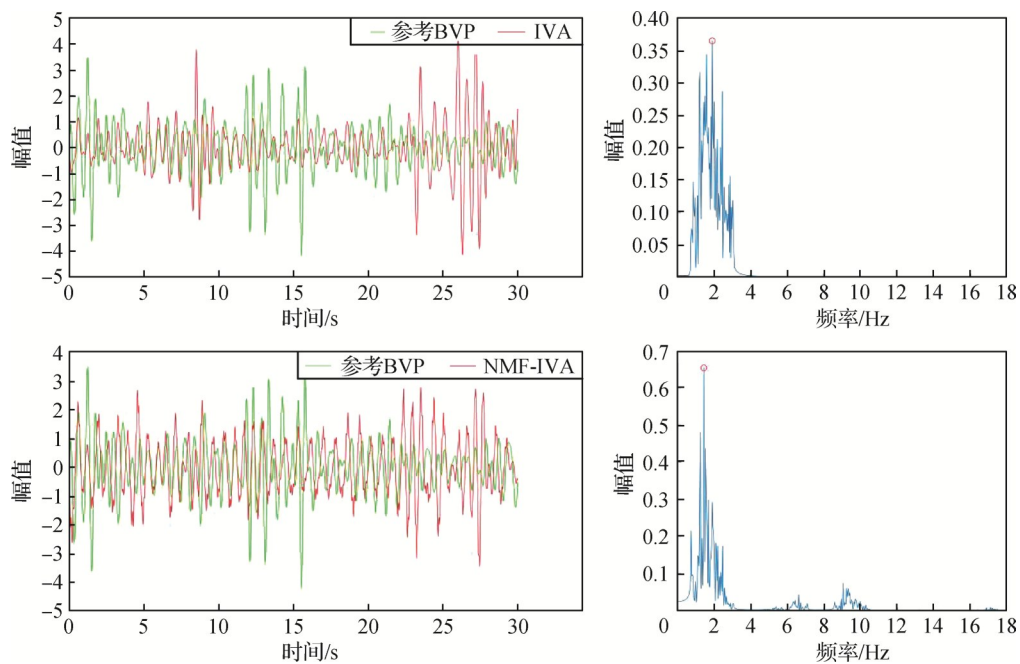


图5 NMF-IVA 和 IVA 算法与参考 BVP 信号的波形对比及各自频谱图

Fig. 5 Comparison of the BVP signals derived by NMF-IVA and IVA algorithms and the reference, as well as the spectrograms

表5 自采数据库上实验结果对比

Table 5 Result comparison on the house-in database

方法	HR_{rmse} /bpm	HR_{mae} /bpm	r
SCF(Chen 等, 2016)	17.99	12.83	0.57
CHROM(De Haan 和 Jeanne, 2013)	17.82	14.16	0.57
POS(Wang 等, 2017)	23.13	17.48	0.45
ICA(Poh 等, 2011)	12.19	5.35	0.76
IVA(Cheng 等, 2021)	7.10	3.09	0.91
EfficientPhys(Liu 等, 2023)	1.99	0.91	0.99
TSCAN(Liu 等, 2021a)	2.26	1.03	0.99
NMF-IVA	4.72	2.77	0.94

2.3.4 RoI 选择对视频心率检测结果的影响

表6给出了采用和未采用最优 SRoI 筛选规则对人脸视频心率检测结果的影响,此时3个 SRoI 是随机确定的。从表6不难发现,无论是在哪一个数据集上,采用最优 SRoI 筛选得到的指标均优于未采用的。现有研究也证明了不同 RoI 包含生理参数信息的强弱程度不同、受运动伪迹干扰的程度不同,导致最终的视频心率检测结果有差异(Kwon 等, 2015)。该消融实验证明了本文所提 RoI 筛选机制是有效的。

表6 最优子 RoI 筛选与否结果对比

Table 6 Result comparison of optimal SRoI

数据集	screening or not					
	HR_{rmse} /bpm		HR_{mae} /bpm		r	
	采用	未采用	采用	未采用	采用	未采用
UBFC-RPPG	4.19	4.77	2.81	3.62	0.97	0.96
UBFC-PHYS-T1	2.08	5.53	1.16	3.26	0.98	0.90
UBFC-PHYS-T2	7.14	12.87	5.06	8.94	0.88	0.45
UBFC-PHYS-T3	6.75	10.09	5.32	8.64	0.91	0.81
自采数据集	4.72	9.01	2.77	5.69	0.94	0.84

2.3.5 运算时间实验结果

表7是各种方法进行视频心率检测的运算时间结果对比。其中前6种传统方法采用 MATLAB 2019b 编程,运行环境为 Intel i5-7500 @ 3.40 GHz; 后2种深度学习方法采用 PyCharm 编程,运行环境为 Intel i9-10900X,搭载 NVIDIA GeForce RTX 3090 显卡。本文 NMF-IVA 相比于传统方法运行时间最长。主要原因是为了获得质量好的 BVP 信号源,需要将 RoI 中获得的时间序列进行升采样(通常高于 256 Hz),导致时间序列长度显著增加,从而增加时间开销。在真实场景下使用时,可以结合生理参数检测性能和时间开销设置合适的升采样率。

表7 运行时间结果对比

Table 7 Result comparison about elapsed time

方法	运行时间/s
SCF(Chen等,2016)	0.019 5
CHROM(De Haan和Jeanne,2013)	0.068 3
POS(Wang等,2017)	0.072 1
ICA(Poh等,2011)	0.117 5
IVA(Cheng等,2021)	0.823 7
NMF-IVA(本文)	6.007 4
EfficientPhys(Liu等,2023)	0.470 7
TSCAN(Liu等,2021a)	0.486 3

3 结 论

本文创新性地提出了一种基于NMF-IVA的非规律运动伪迹干扰去除的人脸视频心率检测方法,该方法充分利用NMF挖掘混合信号的特征,将其用于构建IVA的源先验模型,从而解决单纯IVA算法对规律性BVP信号分离性能不佳的问题,同时解决单纯NMF在没有聚类等手段前提下不能直接进行盲源分离的问题,为非规律运动伪迹情况下的运动鲁棒视频心率检测提供解决思路。

在公开数据集UBFC-RPPG和UBFC-PHYS,及自采数据集上对本文方法和其他几种典型的rPPG方法进行了性能对比,结果表明所提方法与传统方法对比时性能最佳,当数据集中存在大量不规则运动伪迹的场景时,本文所提NMF-IVA方法的性能提升最为明显。本文所提方法在一定程度上可以媲美深度学习方法。得益于深度学习方法强大的特征学习和挖掘能力,通常取得了比传统方法更好的性能,但是训练样本量和泛化性是该类方法重点考虑的问题之一。

值得注意的是,本文取得较好性能的假设是BVP信号具有准周期性而运动伪迹具有非规律性,当真实环境下运动伪迹也具有周期性,例如运动健身场景,如跑步、划船、骑单车等规律运动时(Špetlík等,2018),本文所提方法往往会将幅值更大的周期性运动噪声当做有效源信号进行提取,此时,如何提高视频心率检测方法的运动鲁棒性是未来的研究方向。此外,本文方法为了获得质量好的BVP信号

源,需要对从SRoI中获得的绿色通道时间序列进行升采样(通常高于256 Hz),导致时间序列长度的显著增加,从而带来较大的时间开销。在真实场景下使用时,可以结合生理参数检测性能和时间开销设置合适的升采样率。

参考文献(References)

- Algumaei A, Azam M, Najar F and Bouguila N. 2023. Bounded multivariate generalized Gaussian mixture model using ICA and IVA. *Pattern Analysis and Applications*, 26(3): 1223-1252 [DOI: 10.1007/s10044-023-01148-w]
- Biswas D, Simões-Capela N, Van Hoof C and van Helleputte N. 2019. Heart rate estimation from wrist-worn photoplethysmography: a review. *IEEE Sensors Journal*, 19(16): 6560-6570 [DOI: 10.1109/JSEN.2019.2914166]
- Bobbia S, Macwan R, Benezeth Y, Mansouri A and Dubois J. 2019. Unsupervised skin tissue segmentation for remote photoplethysmography. *Pattern Recognition Letters*, 124: 82-90 [DOI: 10.1016/j.patrec.2017.10.017]
- Chen J, Chang Z Q, Qiu Q, Li X B, Sapiro G, Bronstein A and Pietikäinen M. 2016. RealSense = real heart rate: illumination invariant heart rate estimation from videos//*Proceedings of the 6th International Conference on Image Processing Theory, Tools and Applications (IPTA)*. Oulu, Finland: IEEE: 1-6 [DOI: 10.1109/IPTA.2016.7820970]
- Chen X, Cheng J, Song R C, Liu Y, Ward R and Wang Z J. 2019. Video-based heart rate measurement: recent advances and future prospects. *IEEE Transactions on Instrumentation and Measurement*, 68(10): 3600-3615 [DOI: 10.1109/TIM.2018.2879706]
- Cheng J, Wang P, Song R C, Liu Y, Li C, Liu Y and Chen X. 2021. Remote heart rate measurement from near-infrared videos based on joint blind source separation with delay-coordinate transformation. *IEEE Transactions on Instrumentation and Measurement*, 70: #5005313 [DOI: 10.1109/TIM.2020.3041083]
- De Haan G and Jeanne V. 2013. Robust pulse rate from chrominance-based rPPG. *IEEE Transactions on Biomedical Engineering*, 60(10): 2878-2886 [DOI: 10.1109/TBME.2013.2266196]
- den Brinker A C and Wang W J. 2022. Model-based camera-PPG: Pulse-rate monitoring in fitness//Wang W J and Wang X Y, eds. *Contactless Vital Signs Monitoring*. [s.l.]: Academic Press: 51-78 [DOI: 10.1016/B978-0-12-822281-2.00011-1]
- Hu M, Guo D, Jiang M X, Qian F, Wang X H and Ren F J. 2022. rPPG-based heart rate estimation using spatial-temporal attention network. *IEEE Transactions on Cognitive and Developmental Systems*, 14(4): 1630-1641 [DOI: 10.1109/TCDS.2021.3131197]
- Kwon S, Kim J, Lee D and Park K. 2015. ROI analysis for remote photoplethysmography on facial video//*Proceedings of the 37th Annual*

- International Conference of the IEEE Engineering in Medicine and Biology Society. Milan, Italy: IEEE: 4938-4941 [DOI: 10.1109/EMBC.2015.7319499.PMID:26737399]
- Li B, Zhang P P, Peng J Y and Fu H. 2023. Non-contact PPG signal and heart rate estimation with multi-hierarchical convolutional network. *Pattern Recognition*, 139: #109421 [DOI: 10.1016/j.patcog.2023.109421]
- Liu X, Fromm J, Patel S and McDuff D. 2021a. Multi-task temporal shift attention networks for on-device contactless vitals measurement [EB/OL]. [2023-06-25]. <https://arxiv.org/pdf/2006.03790v2.pdf>
- Liu X, Jiang Z H, Fromm J, Xu X H, Patel S and McDuff D. 2021b. MetaPhys: few-shot adaptation for non-contact physiological measurement//*Proceedings of 2021 Conference on Health, Inference, and Learning*. Virtual Event, USA: ACM: 154-163 [DOI: 10.1145/3450439.3451870]
- Liu X, Hill B, Jiang Z H, Patel S and McDuff D. 2023. EfficientPhys: enabling simple, fast and accurate camera-based cardiac measurement//*Proceedings of 2023 IEEE/CVF Winter Conference on Applications of Computer Vision*. Waikoloa, USA: IEEE: 4997-5006 [DOI: 10.1109/WACV56688.2023.00498]
- Niu X S, Han H and Shan S G. 2020. Remote photoplethysmography-based physiological measurement: a survey. *Journal of Image and Graphics*, 25(11): 2321-2336 (牛雪松, 韩琥, 山世光. 2020. 基于 rPPG 的生理指标测量方法综述. *中国图象图形学报*, 25(11): 2321-2336) [DOI: 10.11834/jig.200341]
- Poh M Z, McDuff D J and Picard R W. 2011. Advancements in noncontact, multiparameter physiological measurements using a webcam. *IEEE Transactions on Biomedical Engineering*, 58 (1): 7-11 [DOI: 10.1109/TBME.2010.2086456]
- Ruan H X, Lei T, Chen K and Lu J. 2022. An explicit connection between independent vector analysis and tensor decomposition in blind source separation. *IEEE Signal Processing Letters*, 29: 1277-1281 [DOI: 10.1109/LSP.2022.3176534]
- Sabour R M, Benezeth Y, De Oliveira P, Chappé J and Yang F. 2023. UBFC-Phys: a multimodal database for psychophysiological studies of social stress. *IEEE Transactions on Affective Computing*, 14(1): 622-636 [DOI: 10.1109/TAFFC.2021.3056960]
- Singh N and Raj A A B. 2020. Estimation of heart rate and respiratory rate using imaging photoplethysmography technique//*Proceedings of 2020 International Conference on System, Computation, Automation and Networking (ICSCAN)*. Pondicherry, India: IEEE: #9262320 [DOI: 10.1109/ICSCAN49426.2020.9262320]
- Song R C, Chen H, Cheng J, Li C, Liu Y and Chen X. 2021a. PulseGAN: learning to generate realistic pulse waveforms in remote photoplethysmography. *IEEE Journal of Biomedical and Health Informatics*, 25(5): 1373-1384 [DOI: 10.1109/JBHI.2021.3051176]
- Song R C, Li J J, Cheng J, Li C, Liu Y and Chen X. 2021b. Motion robust imaging ballistocardiography through a two-step canonical correlation analysis. *IEEE Transactions on Instrumentation and Measurement*, 70: #4001710 [DOI: 10.1109/TIM.2020.3034974]
- Song R C, Li J J, Wang M D, Cheng J, Li C and Chen X. 2021c. Remote photoplethysmography with an EEMD-MCCA method robust against spatially uneven illuminations. *IEEE Sensors Journal*, 21(12): 13484-13494 [DOI: 10.1109/JSEN.2021.3067770]
- Špetlík R, Franc V, Čech J and Matas J. 2018. Visual heart rate estimation with convolutional neural network//*Proceedings of the 29th British Machine Vision Conference*. Newcastle, UK: [s.n.]: 3-6
- Tsou Y Y, Lee Y A, Hsu C T and Chang S H. 2020. Siamese-rPPG network: remote photoplethysmography signal estimation from face videos//*The 35th Annual ACM Symposium on Applied Computing*. Brno, Czech Republic: ACM: 2066-2073 [DOI: 10.1145/3341105.3373905]
- Wang W J, den Brinker A C and De Haan G. 2020. Discriminative signatures for remote-PPG. *IEEE Transactions on Biomedical Engineering*, 67(5): 1462-1473 [DOI: 10.1109/TBME.2019.2938564]
- Wang W J, Den Brinker A C, Stuijk S and Haan G D. 2017. Algorithmic principles of remote PPG. *IEEE Transactions on Biomedical Engineering*, 64 (7): 1479-1491 [DOI: 10.1109/TBME.2016.2609282]
- Yang Z, Yang X Z, Huo L, Liu X N and Li J S. 2018. Heart rate estimation from face videos against motion interference. *Journal of Electronics and Information Technology*, 40(6): 1345-1352 (杨昭, 杨学志, 霍亮, 刘雪南, 李江山. 2018. 抗运动干扰的人脸视频心率估计. *电子与信息学报*, 40(6): 1345-1352) [DOI: 10.11999/JEIT170824]
- Yu Z T, Li X B and Zhao G Y. 2021. Facial-video-based physiological signal measurement: recent advances and affective applications. *IEEE Signal Processing Magazine*, 38(6): 50-58 [DOI: 10.1109/MSP.2021.3106285]
- Zhao C C, Zhou M H, Han W R and Feng Y J. 2022. Anti-motion remote measurement of heart rate based on region proposal generation and multi-scale ROI fusion. *IEEE Transactions on Instrumentation and Measurement*, 71: #5012413 [DOI: 10.1109/TIM.2022.3169567]

作者简介

成娟,女,教授,主要研究方向为医学信号与图像处理 and 模式识别。E-mail: chengjuan@hfut.edu.cn

刘羽,通信作者,男,副教授,主要研究方向为图像处理、医学图像分析和信息融合。E-mail: yuliu@hfut.edu.cn

殷辰楚,男,硕士研究生,主要研究方向为图像处理。

E-mail: 2020110095@mail.hfut.edu.cn

宋仁成,男,副研究员,主要研究方向为医学信号与图像处理。E-mail: rcsong@hfut.edu.cn

付静,女,实验师,主要研究方向为图像处理。

E-mail: fujing@ustc.edu.cn