

中图法分类号:TP391 文献标识码:A 文章编号:1006-8961(2020)11-2321-16

论文引用格式: 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]

基于 rPPG 的生理指标测量方法综述

牛雪松^{1,2}, 韩琥^{1,3}, 山世光^{1,2,4}

1. 中国科学院计算技术研究所智能信息处理重点实验室, 北京 100190; 2. 中国科学院大学, 北京 100049;
3. 鹏城实验室, 深圳 518066; 4. 中国科学院脑科学与智能技术卓越创新中心, 上海 200031

摘要: 远程光电容积脉搏描记法(remote photoplethysmography, rPPG)是指通过摄像头等传感器来捕捉由心动周期造成的皮肤颜色周期性变化的技术。利用rPPG技术可以提取血液体积脉冲信号,进而测量心率、呼吸率和心跳变异性等心动周期相关的生理指标。近年基于rPPG的生理指标测量方法取得了飞速发展,准确性和鲁棒性已得到了大幅提高。该类技术的算法流程主要包括图像/视频的获取和感兴趣区域提取、血液体积脉冲信号提取和生理指标测量等步骤。基于这一算法流程,本文从算法所依据的假设或先验知识出发,对相关文献进行了总结和讨论。此外,还从评测任务、评测数据、评测指标及评测协议等4个方面系统整理针对基于rPPG的生理指标测量方法的评价体系。最后,本文讨论了该领域当前所面临的挑战并展望了可能的技术路线。

关键词:远程光电容积脉搏描记法(rPPG);心动周期;生理指标测量;文献综述;算法评测

Remote photoplethysmography-based physiological measurement: a survey

Niu Xuesong^{1,2}, Han Hu^{1,3}, Shan Shiguang^{1,2,4}

1. Key Laboratory of Intelligent Information Processing of Chinese Academy of Sciences (CAS), Institute of Computing Technology, CAS, Beijing 100190, China; 2. University of Chinese Academy of Sciences, Beijing 100049, China;
3. Peng Cheng Laboratory, Shenzhen 518066, China; 4. Chinese Academy of Sciences (CAS) Center for Excellence in Brain Science and Intelligence Technology, CAS, Shanghai 200031, China

Abstract: Physiological signals, such as heart rate (HR), respiration frequency (RF), and heart rate variability (HRV), are important clues to analyze a person's health and affective status. Traditional measurements of physiological signals are based on the electrocardiography (ECG) or contact photoplethysmography (cPPG) technology. However, both technologies require professional equipment, which may cause inconvenience and discomfort for subjects. Remote photoplethysmography (rPPG) technology for remote measurement of physiological signals has progressed considerably and recently attracted considerable research attention. The rPPG technology, which is based on skin color variations due to the periodical optical absorption of skin tissue caused by cardiac activity, demonstrates high potential in many applications, such as healthcare, sleep monitoring, and deflection detection. The process for rPPG-based physiological measurement can be divided into three steps. First, regions of interest (ROIs) are extracted from the face video. Second, blood volume pulse (BVP) signal is reconstructed from signals generated from the ROIs. Finally, the reconstructed BVP signal is used for physiological

收稿日期:2020-07-01;修回日期:2020-08-14;预印本日期:2020-08-21

基金项目:国家重点研发计划项目(2017YFA0700804);国家自然科学基金项目(61672496)

Supported by: National Key Research and Development Program of China (2017YFA0700804); National Natural Science Foundation of China (61672496)

measurements. The reconstruction of the BVP signal is the key step for rPPG-based remote physiological measurements. A detailed review of methods for rPPG-based remote physiological measurement is presented in this study from the aspect of assumptions they use, which can be categorized into three kinds, i. e., methods based on the skin reflection model, methods based on the BVP signal's physical characteristics, and data-driven methods. Studies on the skin reflection model-based methods can be further categorized into spatial skin and skin reflection models of different color channels. Studies on methods that using the BVP signal's physical characteristics can be further categorized into blind signal separation, manifold projection, low rank factorization, and frequency domain constraint. Studies on data-driven methods can be further categorized into methods based on hand-crafted features and deep learning. A detailed review of evaluations of different rPPG-based physiological measurement methods is also presented from the aspects of tasks, databases, metrics, and protocols. Evaluation tasks used for remote physiological measurement include average heart rate measurement, respiration frequency measurement, and heart rate variability analysis. Databases of rPPG-based physiological measurements are summarized according to database scale and variations. Evaluation metrics for remote physiological measurement can be categorized into statistics of error, correlation, and signal quality. Evaluation protocols for data-driven methods are summarized into fixed partition, subject-independent division, subject-exclusive division, and cross-database protocols. Finally, we discuss the challenges of the rPPG-based remote physiological measurement and put forward the potential research directions for future investigations. Challenges include video quality (i. e., video compression and pre-processing of frames), influence of subject's head movements, variations of lighting conditions, and lacking data. Future research trends include designing hand-crafted methods for different challenge scenarios and exploring technologies, such as self-supervised, semi-supervised, and weakly-supervised learning, for data-driven methods.

Key words: remote photoplethysmography (rPPG); cardiac cycle; physiological measurement; literature survey; algorithm evaluation

0 引言

心动周期(cardiac cycle)是指从一次心跳的起始到下一次心跳的起始心血管系统所经历的过程。通过对心动周期的分析,可以提取心率(heart rate, HR)、呼吸频率(respiration rate, RF)和心率变异性(heart rate variability, HRV)等重要的生理指标。这些信息可以帮助有效地判断一个人的健康状态和情绪状态,在医学诊断、健康监控和刑侦检测等领域有着重要的应用。

现有的心动周期信号的测量方法主要有两种,一种是心电图(electrocardiogram, ECG)(Noble, 1984)。该方法通过记录心脏活动产生的生物电信号来获得,是当前最为精确的测量方式,但其需要专业人士利用专业设备,通过在人身上连接复杂的电极来进行测量,使用起来十分不便。另一种是光电容积脉搏波描记法(photoplethysmography, PPG),该方法通过实时测量特定身体部位因心脏搏动而产生的对不同波长光照吸收/反射量的周期性变化,即血液体积脉冲(blood volume pulse, BVP)信号(Manzzone等,2007),并以此来推断心动周期情况以及心

率、呼吸频率和心跳变异性等生理指标。该方法仍然需要通过专业的设备进行接触式测量,在实际使用过程中仍会带来诸多不便。

基于远程PPG(remote PPG, rPPG)的心动周期信号测量方法应运而生。这类方法采用普通摄像头在较远距离上进行信号采集,无需专业设备,也无需与被检测者发生接触,因而更加方便,具有更加广泛的应用前景。该类方法的基本原理是:利用摄像头捕捉皮肤区域(通常选取人脸或胳膊上的皮肤)的视频,分析皮肤区域因心跳带来的血流脉动而产生的周期性颜色变化,进而恢复出相应的BVP信号并进行生理指标的测量。但是由于远程摄像头所能记录下来的肤色的周期性颜色变化十分微弱,自然场景下的环境光照、被试头部运动等变化因素会对最终的生理指标测量结果产生很大的影响。研究者们针对这些难题,提出了许多稳定的BVP信号恢复方法和鲁棒的生理指标测量方法。

如图1所示,现有基于rPPG的生理指标测量算法流程主要包含3个步骤:视频获取及感兴趣区域划分、BVP信号提取和增强以及生理指标测量。其中涉及如下关键性问题:

1)如何选取合适的感兴趣区域(region of inter-

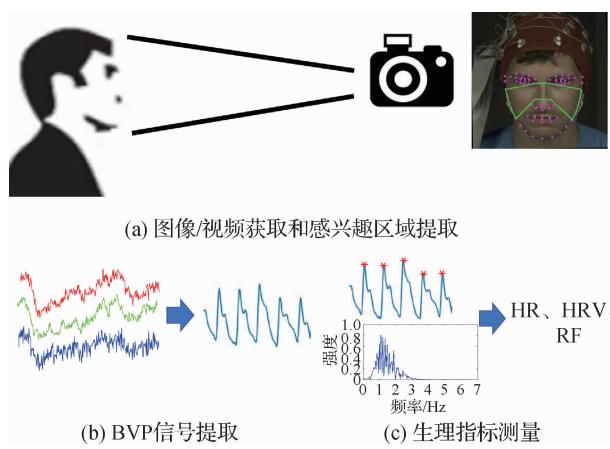


图1 基于rPPG的生理指标测量算法流程

Fig. 1 The pipeline for rPPG based physiological measurement
 ((a) capturing video frames and extracting region of interests; (b) extracting BVP signals;
 (c) physiological measurement)

est, ROI)来减小背景噪声的引入;

- 2) 对目标BVP信号做出什么样的假设能更有效地帮助恢复出稳定的BVP信号;
- 3) 如何通过提取到的BVP信号来计算各种生理指标。

本文围绕以上几个问题对过去十余年的rPPG技术加以总结和讨论。同时,对文献中使用的不同评测体系从评测任务、评测数据、评测标准和评测协议4个维度加以总结和归纳。最后,讨论了该领域所面临的挑战以及未来可能的发展方向。

尽管已有综述对非接触式生理信息测量相关工作做过总结,但该领域缺少全面的算法层面的综述论文。例如,有些研究者针对包括普通彩色摄像头在内的各种传感器在非接触式生理测量上的应用进行了文献归纳和总结(Teichmann等,2012;Kranjec等,2014;Al-Naji等,2017)。Khanam等人(2019)则对rPPG生理指标检测的应用场景做了总结。这些文献并不是从算法角度归纳基于rPPG的生理信号测量技术。部分研究者从算法设计角度进行了广泛的文献调查(Sun和Thakor,2016;McDuff等,2015;Hassan等,2017;Chen等,2019),或是从面对不同的挑战所设计的不同策略,或是从相关算法所用技术类别角度进行文献总结,但对最近出现的新思路、新算法、新技术关注不够。此外,现有的文献综述都没有对评测体系进行总结和讨论。

相较于已有的综述文章,本文贡献在于:1)基

于已有算法流程,从算法所依据的假设或先验知识这一角度详细总结了近年来出现的基于rPPG的心动周期相关生理指标检测算法,包括了在现有综述中未被总结的最新算法和应用;2)对已有文献所使用的评测体系做了充分的调研和总结;3)详细讨论了该领域当前面临的挑战,展望了未来发展趋势。

需要说明的是,本文仅聚焦于在自然环境下利用普通彩色摄像头来进行心动周期相关生理信息测量的算法。利用其他模态的专业传感器,例如近红外摄像头(van Gastel等,2015)、热成像摄像头(Garbey等,2007)、语音传感器(Mesleh等,2012)及雷达传感器(Li等,2009),或者基于其他算法原理,例如心冲描记法(ballistocardiograph, BCG)(Balakrishnan等,2013)等相关方法在本文中将不做详细总结和讨论。此外,本文也将只涉及与心动周期信号相关的生理指标检测,对于其他可被测量的生理指标,例如血氧饱和度(Tarassenko等,2014)等,不进行深入总结和讨论。

1 算法原理

如图2所示,人类皮肤可划分为3层:表皮层、真皮层和皮下组织层,在真皮层和皮下组织层中分布着丰富的毛细血管。人类皮肤对光照的吸收主要来自于黑色素和毛细血管中的血红蛋白。心脏跳动所造成的血流的涌动会导致毛细血管的血流量及其中的血红蛋白含量发生周期性变化,进而导致皮

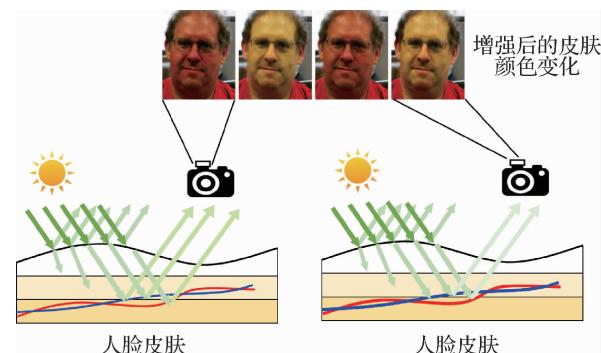


图2 自然光照在皮肤上的反射和吸收模型及通过摄像头捕捉到的皮肤颜色变化(皮肤的颜色变化通过Wu等人(2012)方法放大)

Fig. 2 The optical reflection and absorption model of skin and the skin color variations captured by the camera (skin color variations have been magnified using Wu et al. (2012) method)

肤对光照的吸收量的周期性变化。这一变化虽然很微弱,但仍可以被摄像头所捕捉。

rPPG 技术就是基于这一原理,通过摄像头来捕捉固定皮肤区域由于光照吸收量不同而产生的周期性颜色变化,并进一步用捕捉到的皮肤颜色变化来提取 BVP 信号并测量心动周期相关的生理指标。由于由摄像头所记录下来的这一皮肤周期性颜色变化十分微弱,环境光照、被试者的头部运动等因素会对最终的生理指标测量造成较大的影响。

2 现有方法

自 Verkruyse 等人(2008)首次提出利用 rPPG 技术来进行生理指标测量以来,研究者提出了许多方法来提取稳定的 BVP 信号并进行鲁棒的生理指标测量。这些方法的算法流程可以被总结为图 1 所示的 3 个步骤。其中,BVP 信号的提取是基于 rPPG 的生理指标测量算法设计的核心步骤,依据其所使用的假设或先验知识可以分为基于皮肤光照模型的方法、基于 BVP 信号特性假设的方法和基于数据驱动的方法。依据这一算法流程,从算法所依赖假设及先验知识这一角度详细总结现有方法。

2.1 图像/视频的获取以及感兴趣区域的提取

根据第 1 节算法原理可知,基于 rPPG 的生理指标测量算法所需要提取的是由血流涌动所造成的皮肤周期性颜色变化。在日常生活中,最容易被摄像头所获取的皮肤区域是人脸和手部。相较于手部区域,人脸区域没有大幅度的形变,更加稳定。因此,大部分算法都是采用人脸视频作为输入。

在获取到人脸的视频/图像后,接下来要考虑的问题便是如何选取人脸上合适的感兴趣区域来减小背景噪声的引入。在早期研究中,常用的人脸感兴趣区域包括图像中的固定区域(Verkruyse 等,2008)或者人脸检测框内区域(Poh 等,2010,2011)。这两种选择方式容易引入非皮肤区域(例如背景或眼部)的噪声。为了解决这一问题,研究者提出采用皮肤分割的方法来获取感兴趣区域。常用的皮肤分割方法主要有两种:一种是利用人脸的关键点裁剪出固定的人脸皮肤区域(Li 等,2014);另一种是利用肤色分割算法来划分出人脸的皮肤区域(de Haan 和 Jeanne,2013;Wang 等,2015)。

除了选择固定的感兴趣区域外,研究者还对利

用算法自动选取高质量的皮肤区域进行了探索。Tasli 等人(2014)在利用人脸关键点定义的兴趣区域的基础上自动将颜色离群点加以剔除。Feng 等人(2015a)依据生理指标估计这一任务的特点设计了两种特征,并使用 K 均值算法自适应选择有效的感兴趣区域。Po 等人(2018)在 Feng 等人(2015a)的基础上进一步使用均值漂移的聚类方法挑选更鲁棒的兴趣区域。Bobbia 等人(2019)依据皮肤颜色变化的周期性,使用信噪比来挑选需要的兴趣区域。而 Kumar 等人(2015)则使用信噪比来对不同的兴趣区域信号进行加权。

总之,图像/视频的获取以及感兴趣区域的提取作为基于 rPPG 生理指标测量的第一步,主要作用是为后续算法剔除信噪比低的像素或区域,提供鲁棒的输入信号。不论是选择定义好的感兴趣区域,还是通过算法自动挑选感兴趣区域,都能够在一定程度上过滤掉背景区域,从而有效提高算法的鲁棒性。

2.2 BVP 信号的提取

在获得了尽可能过滤掉背景噪声的兴趣区域之后,需要提取 BVP 信号来进行生理指标的测量。虽然摄像头直接捕捉到的人脸皮肤颜色变化可以作为一种初步的 BVP 信号来测量各种生理指标,但是这种初步的 BVP 信号非常容易受到外部环境的干扰。为了提取鲁棒的 BVP 信号用于生理指标测量,需要通过进一步的建模来增强 BVP 信号,去除其他因素的干扰。现有 BVP 信号提取算法可以根据其所依赖的假设或先验知识分为基于皮肤光照模型的方法、基于 BVP 信号特性假设的方法和基于数据驱动的方法。下面分别对这 3 种类型的 BVP 信号提取方法进行总结。

2.2.1 基于皮肤光照模型的方法

由第 1 节算法原理可知,基于 rPPG 的生理指标测量算法依据人体皮肤对外部光照的吸收和反射所造成的变化。因此,通过对皮肤的光照反射模型进行建模,依据摄像头捕捉到的皮肤颜色变化,可以进一步获得皮肤毛细血管内的血红蛋白含量的变化,从而推断出血流量的变化,并以此来测量出相应的生理指标。常用的对皮肤模型的建模方法分为两种:一种是建模皮肤对不同波长光照(不同颜色通道)的吸收和反射情况,利用皮肤对不同波长光照吸收的差异性和互补性来提取和增强 BVP 信号;另

一种是考虑不同头部的姿态和运动状况对光照在皮肤上的吸收和反射情况的影响,进而对BVP信号进行校正和增强。

de Haan 和 Jeanne(2013)首先建立不同颜色通道的皮肤光照反射模型,利用不同颜色空间的信号的互补性,提出了在白光假设下的基于色度的颜色空间投影分解算法 Chrominance-based rPPG (CHROM) 来消除头部运动所带来的影响。Wang 等人(2015)将 CHROM 方法推广到了逐像素操作并取得了更加鲁棒的结果。de Haan 和 van Leest (2014)则在给定相机特性的前提下将 CHROM 方法进一步推广到其他非白光光照条件下。此外 Wang 等人(2017a)还将 CHROM 所使用的皮肤模型推广到更加普通的场景,并提出了动态寻找最优的颜色空间投影的方法。

除了建模不同颜色通道捕捉到的皮肤周期性颜色变化,研究者们还探究了在不同的头部运动情况下的皮肤光照模型。Feng 等人(2015b)将人脸建模为标准朗博体模型,并以此为基础建模不同头部运动状态下皮肤对光照的吸收和反射情况,并根据模型来消除头部运动对于 BVP 信号的影响。

总体而言,基于皮肤光照模型的方法能够依据皮肤光照模型的先验有效缓解由头部运动等因素所带来的信号不稳定。但是由于皮肤光照模型是基于对皮肤特性的假设,在复杂的场景下可能并不适用。在复杂场景下如何对皮肤的光照模型进行建模将是未来重要的研究方向。

2.2.2 基于 BVP 信号特性假设的方法

除了利用皮肤的光照模型来提取 BVP 信号外,研究者们还探索了利用 BVP 信号的特性去除噪声、增强信号鲁棒性的方法。经常使用的特性假设包括:线性混合假设、流形假设、信号同步假设以及周期性假设。下面将从这几个方面具体总结现有方法。

线性混合假设认为不同颜色通道信号或者不同感兴趣区域内信号是由 BVP 信号及噪声信号线性组合而成。因此,可以通过分离 BVP 信号和噪声信号来获得鲁棒的 BVP 信号的估计。为了最大程度地表征噪声信息,经常假设构成了所有颜色通道信号的噪声信号以及 BVP 信号的数目与颜色通道数相同。以 RGB 摄像头为例,摄像头所捕捉到的不同颜色通道的颜色变化信号为 $\mathbf{x}(t) = [\mathbf{R}(t), \mathbf{G}(t),$

$\mathbf{B}(t)]^T$, t 为时间,线性混合假设认为 $\mathbf{x}(t) = \mathbf{As}(t)$, 其中, $\mathbf{s}(t) = [\mathbf{BVP}(t), \mathbf{S}_1(t), \mathbf{S}_2(t)]^T$, 代表了 BVP 信号和其他两种噪声信号, $\mathbf{S}(t)$ 为噪声信号,而 \mathbf{A} 为 3×3 的线性混合矩阵。基于这一假设,可以通过盲源信号分离的技术来求得 $\mathbf{BVP}(t)$ 。Poh 等人(2010,2011)首先提出这一假设并成功使用独立成分分析 (independent component analysis, ICA) 分离出 BVP 信号,取得了鲁棒的生理指标测量结果。此后,McDuff 等人(2014)将 ICA 应用到可以捕捉红绿蓝青橙 5 种颜色的摄像机上。Lam 和 Kuno(2015)则假设 BVP 信号是来自不同皮肤区域的信号的线性组合,并基于这种假设,对脸颊内随机采样的多对区域使用 ICA 来分离 BVP 信号,然后使用投票机制来求取心率值。除 ICA 之外,其他盲源信号分离的方法,例如主成分分析(principle component analysis, PCA) (Lewandowska 等,2011), 约束独立成分分析(constrained independent component analysis, cICA) (Tsouri 等, 2012), 也被成功地用于 BVP 信号的提取。

除了假设不同颜色通道信号是由 BVP 信号和噪声信号线性组合而成之外,研究者们还提出了不同颜色通道信号是由 BVP 信号和噪声信号通过非线性组合而成的可能性。进而提出 BVP 信号是不同颜色通道信号在原始颜色空间的流形子空间上的投影的假设。Wei 等人(2012)选取了拉普拉斯特征映射来求解 BVP 信号所在颜色子空间。而 Pilz(2019)则假设 BVP 信号在 RGB 颜色空间的黎曼流形上,并以此来提取 BVP 信号。相较于盲源信号分离的方法,这类方法所使用的假设更加复杂,也能更好地应对复杂的现实场景。

除了假设不同颜色通道的信号是 BVP 信号和噪声信号的混合之外,研究者们还考虑不同感兴趣区域之间的关系来增强 BVP 信号。由于不同感兴趣区域内的皮肤颜色变化都是由被试者的心动周期引起的,因此不同的感兴趣区域内的信号应该是同步的。基于这一假设,可以认为所有感兴趣区域的信号中不同步的部分即为噪声信号。因此,这一假设可以用来消除不同感兴趣区域信号的噪声。Tulyakov 等人(2016)首先提出这一假设,并利用自适应矩阵补全的方式来同步不同的感兴趣区域,提取 BVP 信号。周双等人(2019)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。除了对

信号做低秩分解,黄继风等人(2020)提出对不同感兴趣区域的频域特征做低秩分解来进行鲁棒的生理指标测量。

除了上述对于 BVP 信号的假设外,研究者们还利用 BVP 信号在频域内的特点来帮助增强 BVP 信号。只包含心动周期信息的 BVP 信号有着较强的周期性,并且这一周期性具有一定的先验特点,例如其周期在人正常心率范围内且在相邻的几次估计内不容易发生周期的变化。而噪声信号往往不具有这些特性。因此,这些特性可以用来去除噪声,增强 BVP 信号的鲁棒性。频域处理工具是用来处理信号周期性的有效工具,因此常常用于 BVP 信号的去噪和增强。Wang 等人(2017b)分析了在频域内心动周期信号与非心动周期信号的特点,并使用频域修正的方式来去除非正常心率范围内的噪声。此后 Wang 等人(2017c)还结合人类正常心率范围这一先验知识,在频域内去除了非心率信号的噪声。除了利用 BVP 信号的频域特点外,研究者们还通过去除其他因素的频域干扰来增强 BVP 信号。Gudi 等人(2019)在频域上剔除了头部运动的频率,并取得了更鲁棒的估计结果。Niu 等人(2017)考虑了连续估计下相邻的估计结果的频域关系,并利用分布学习的方式来提高心率估计的准确性。

总结而言,基于对 BVP 信号特性假设的方法主要是对 BVP 信号可能的状态做出合理的假设,并依据所采用的假设来提取 BVP 信号。这些方法能够有效去除无关的噪声,提高算法鲁棒性。但同时,在复杂场景下,由于所依赖的假设未必成立,算法可能失效。如何寻找复杂自然场景下良好的模型假设和建模方式也是未来重要的发展方向。

2.2.3 基于数据驱动的方法

除了基于皮肤模型或特定假设来增强 BVP 信号外,研究者们还尝试通过数据驱动的方法来进行 BVP 信号的提取和生理指标的测量,希望在有专业仪器记录的纯净 BVP 信号或者生理指标参考值的情况下,能够通过机器学习建模的方式来过滤掉噪声,训练出一个良好的生理指标测量模型。研究者们早期尝试了利用手工设计特征来训练心率估计模型。Hsu 等人(2014)将原始的 RGB 三通道信号的频域特征以及 ICA 分解后的信号频域特征作为输入,利用支持向量回归心率值。Osman 等人(2015)利用手工设计的特征来分类 BVP 信号的峰值位置,

然后利用峰值测量心率。

除了手工设计特征的数据驱动方法外,研究者们也探索了将具有强大建模能力的深度学习方法应用到基于 rPPG 的生理指标测量上。Hsu 等人(2017)将不同的时序特征和时频表示作为输入,利用卷积神经网络来分类心率值。Qiu 等人(2019)使用多尺度的时间空间表示作为输入,并使用卷积神经网络来回归心率值。Chen 和 McDuff(2018)提出了人脸图像的动态表征,并将这一表征作为输入,利用结合了注意力机制的卷积神经网络来回归 BVP 信号。Niu 等人(2018a)提出了人脸视频的时间空间表示,并针对训练数据缺乏的问题,提出了利用模拟生成的 BVP 信号预训练的方式来更好地初始化模型。Niu 等人(2020)将时间空间表示加以改进,并将时域相邻心率估计结果之间的关系引入到模型的训练过程中。此外,Niu 等人(2019)还针对训练数据不足和不均衡的问题,提出了上下采样的数据增广方式和基于注意力机制的心率估计模型。除了使用不同的表示输入,还有一些研究者尝试直接使用人脸视频作为输入来进行生理指标的测量。Spetlik 等人(2018)通过级联全卷积网络来实现对心率值的估计。Yu 等人(2019a)则使用 3 维卷积网络来训练端到端的心率估计模型。

总结而言,相较于基于手工设计特征的方法,基于深度学习的方法能够从数据中学习各种生理指标测量相关的特征而不受限于各种假设,因此有着更大的应用潜力。但同时,数据规模和多样性对算法的鲁棒性和泛化能力有着重要的影响,并且预测结果也缺乏可解释性。如何充分利用现有数据以及如何提高模型的泛化性能和可解释性将是未来重要的发展方向。

2.3 生理指标测量

提取出 BVP 信号后,可以进一步测量各种生理指标。生理指标测量的方法主要可以分为两种:一种是频域分析,另一种是时域分析。

在频域分析中常用的方法是傅里叶变换。当求得 BVP 信号所对应的频谱后,就可以选取频谱中对应最大频谱强度的周期作为平均心率的估计结果。Poh 等人(2011)还提出 BVP 信号在呼吸频率范围内的最大频谱幅值对应的周期就是被试者的呼吸频率。利用频域分析的方法可以快捷有效地求得心率以及呼吸率,但由于频域分析只适用于分析一段时

间内的平均心率,直接使用频域分析将无法得到如心跳变异性等瞬时心率的特性。

要想测量心跳变异性等生理指标,需要对BVP信号做时域分析。主要方法为检测BVP信号的峰值点,并计算相邻峰值点间的时间间隔,进而计算瞬时心率以及瞬时心率时序变化,并以此来测量心跳变异性各项指标。常用的心跳变异性的测量指标包括低频能量(*low frequency*, LF)、高频能量(*high frequency*, HF)以及低高频能量比(LF/HF)。其中LF反映了交感和副交感神经的活性,HF则反映了副交感神经的活性,而LF/HF反映了自律神经活性平衡情况(Akselrod等,1981)。同时,HF和呼吸频率RF表现出高度的相关性(Yasuma和Hayano,2014),因此也可以通过计算HF的峰值点对应频率来计算呼吸率。此外,除了峰值点检测之外,研究者们也探索了利用时频分析的方法来进行心率估计和心跳变异分析(Bousefsaf等,2013)。

3 算法评测

基于rPPG的生理指标测量在近年获得了快速的发展。但不同文献中对方方法的评价体系各不相同。下面从评测任务、评测数据、评测指标以及评测协议4个方面总结现有方法的评测体系。

3.1 评测任务

根据2.3节的分析可知,利用BVP信号可以测量的生理指标包括:心率、呼吸率和心跳变异性(低频能量、高频能量和低高频能量比)。如附表所示,已有文献中大部分方法只考虑使用平均心率值作为算法的评测任务,只有较少的方法同时对平均心率、呼吸率及心跳变异分析的测量准确性做出评测。

平均心率能够在一定程度上反映算法的有效性,但由于现有的算法评测中使用的视频时长为30 s左右,只计算较长时间内的平均心率缺乏实际应用的价值。研究者们提出了使用基于短时长人脸视频(4 s,6 s,8 s)的心率估计作为测评任务(Tulyakov等,2016;Qiu等,2019)。

虽然平均心率能够有效评价所提取的BVP信号的周期性,但是其并不能测评算法是否提取出每一个心动周期的波形变化。基于这一考虑,研究者们提出同时使用平均心率测量、心跳变异性分析和

呼吸率测量作为评价任务(Poh等,2011;Yu等,2019a)。其中心跳变异性分析和呼吸率测量能够较好地反映算法是否提取了BVP信号中每个心动周期的波形变化。此外,Wang等人(2017b,c)还通过直接比较BVP信号及ECG信号的频域差别来评价提取到的BVP信号的好坏。

总结而言,基于rPPG信号的生理指标测量主要使用3种任务作为评测:平均心率测量、心跳变异性分析及呼吸率测量。其中心跳变异性分析及呼吸率测量能更细致地反映算法的有效性。但由于现有数据集中较少提供用于计算心跳变异性及呼吸率参考值的ECG信号(见表1),这两个评测任务在文献中较少被考虑。在未来,构建包含在各种自然场景下丰富的生理指标参考值的数据集将有助于评测算法的有效性。

3.2 评测数据

早期基于rPPG的生理指标测量算法主要是在自采数据集上进行验证,这使得不同算法之间无法进行有效比较。研究者们发布了一些用于生理指标测量的公开视频数据集,为不同算法的公平对比提供了数据基础(见文后附表)。

如表1所示,从数据规模的角度来看,受限于采集设备、采集成本等因素,现有的公开数据集规模都较小,被试者人数在十人到百人左右,视频长度为30 s~1.5 min左右。较小规模的数据一方面不能全面地评测算法在不同自然场景下的鲁棒性,另一方面无法保证数据驱动方法所需要的训练数据量。因此,构建更大规模数据集对基于rPPG的生理指标测量有重要意义。

从数据多样性的角度看,现有的公开数据集多是针对具体的算法应用场景或现阶段挑战所构建的。常被考虑的数据多样性包括被试者的头部运动、光照条件以及被试者运动前后的心率变化。由于现有数据集主要是通过指挥被试者完成指定动作或控制现场的环境光照来获取不同场景的数据,其人脸视频和实际应用场景还有所不同。此外,现有数据集中提供的生理指标参考值多为BVP信号及平均心率值。只有较少的数据集提供了更为精确的ECG信号。构建包含更多生理信号参考值、更加接近实际应用场景的数据集将有助于评估现有算法的准确性和适用范围,推动该领域的发展。

表1 基于rPPG的心动周期相关生理指标检测公开数据集总结

Table 1 Summary of the public databases for remote physiological measurement

数据集名称	文献	被试人数	视频数	单视频时长	摄像头	生理信号参考值	录制场景多样性	被试者多样性	心率范围/dpm	是否压缩
MAHNOB-HCI	Soleymani 等人(2012)	27	527	30 s	网络摄像头	ECG 信号	只包含被试者表情变化场景	15 女性, 12 男性, 年龄范围 19~40	N/A	是
PURE	Stricker 等人(2014)	10	60	1 min	机器人用摄像头	心率值, 血氧值, BVP 信号	6 种头部运动场景	2 女性, 8 男性	N/A	否
VicarPPG	Tasli 等人(2014)	10	20	1.5 min	网络摄像头	BVP 信号	被试者基本不运动, 运动前后录制	年龄范围 20~35	运动前 50~90, 运动后大于 120	N/A
DistancePPG	Kumar 等人(2015)	不同子集: 12、5、2、8	不同子集: 40 s、80 s、40 s、N/A	不同子集: 40 s、80 s、40 s、N/A	单色摄像头	BVP 信号	4 种录制场景	5 女性, 7 男性	N/A	N/A
MMSE-HR	Tulyakov 等人(2016)	40	102	30 s	网络摄像头	心率值	只包含被试者表情变化场景	不同种族	N/A	是
PFF	Hsu 等人(2017)	10	104	2~3 min	网络摄像头	心率值	2 种光照, 2 种头部运动	N/A	N/A	否
COHFACE	Heusch 等人(2017)	40	106	1 min	网络摄像头	BVP 信号, 呼吸信号	2 种光照	12 女性, 28 男性, 平均年龄 35.6, 方差 11.5	N/A	否
LGI Multi-Session	Pilz 等人(2018)	25	100	平均 2 min	网络摄像头	心率值, BVP 信号	1 种运动场景, 2 种头部运动场景, 2 种光照场景	5 女性, 20 男性, 年龄范围 25~42	N/A	否
VIPL-HR	Niu 等人(2018b, 2020)	107	2 378	约 30 s	网络摄像头、RGBD 摄像头, 手机前置摄像头	心率值, 血氧值, BVP 信号	3 种不同头部运动, 3 种光照条件以及手持手机的不同场景	28 女性, 79 男性, 年龄范围 22~41	47~146	是
ECG-Fitness	Spetlik 等人(2018)	17	204	1 min	彩色摄像头、热度摄像头	ECG 信号	运动场景、3 种不同光照条件	3 女性, 14 男性, 56~159	~	否
OBF	Li 等人(2018)	100	200	5 min	彩色摄像头、头、近红外摄像头	ECG 信号, 呼吸率, BVP 信号	被试者基本不动, 运动前后录制	39 女性, 61 男性, 年龄范围 18~68	N/A	是
UBFC-RPPG	Bobbia 等人(2019)	42	42	约 1.5 min	网络摄像头	BVP 信号	两种光照场景	N/A	N/A	否

注:N/A 表示论文中未提供相关数据,RGBD 表示 red-green-blue-depth。

3.3 评测指标

在评价基于 rPPG 的生理指标测量算法时, 常用的评价指标主要是从生理指标测量的准确性和提取的 BVP 信号的可靠性两个角度来对算法结果进

行度量。

在度量生理指标测量的准确性时, 最常考虑的评价指标就是生理指标测量值与参考值之间的误差。在求得测试数据集所有样本的测量误差之后,

一般用误差的统计量来作为算法的最终评测指标。常用的误差统计量包括平均绝对误差 (mean absolute error, MAE)、均方根误差 (root mean square error, RMSE)、绝对误差的方差 (standard deviation of absolute error, SD)、误差的平均值 Mean、误差的方差 (standard deviation of error, Std) 以及平均百分比误差 (mean of error-rate percentage, MER)。这些指标主要反映了算法在测试数据上的平均估计准确性。

但是,由于人类的心率和心跳变异性个体差异较小,当测试集内每个样本的测量结果均被预测为测试集所有样本的平均结果时,也能够取得较低的评测误差。这种现象表明单独使用整个评测数据集的误差统计量作为评测指标无法完整地反映算法的优劣。为了解决这一问题,除了使用误差统计量指标之外,研究者们还提出将测量误差小于特定阈值(例如每分钟心跳数误差在5次以内)样本占总样本的比例作为测量准确性的指标;或者计算生理指标测量值和参考值之间的皮尔逊相关性系数 r 或Bland-Altman图来度量测量值和参考值之间的一致性。目前研究者多是综合利用不同的评测指标全面地比较不同方法的优劣。

除了度量生理指标测量的准确性,还通过评估所提取的BVP信号的质量来评价算法的优劣。常用到的评测指标是信噪比 (signal-to-noise ratio, SNR),即在提取到的BVP信号的频域内,预测心率频率所对应的频域能量占整个频域能量的比值。需要指出的是,这里的信噪比是指预测心率值而非参考心率在频域内的能量占比。所有测试样本信噪比的均值常被用来度量算法的有效性。SNR虽然能够有效反映提取到的BVP信号的周期性,但并不能细致地评测BVP信号中每个周期是否准确。此外,除SNR外,研究者们还使用成功率 (success-rate) 来评价提取的BVP信号的好坏。成功率是指提取到的BVP信号和参考BVP信号在频域内的绝对误差。相较于SNR,这一指标能更细致地反映算法提取到的BVP信号的质量。但由于传感器和测量皮肤区域不同,人脸的BVP信号与参考BVP信号间有一定的差距,会在评测时带来一定的偏差。

总结而言,基于rPPG的生理指标测量算法所使用的评价指标较多,且不同评价指标的侧重点不同。综合考虑不同评测指标,将能够有助于更客观、

全面地评价不同的算法。

3.4 评测协议

非数据驱动的算法一般将全部数据集作为评测集,不同算法的对比较为公平客观。而对于数据驱动的方法,数据集划分对评测的结果有着较大的影响。如文后附表所示,现有的数据驱动方法所采用的评测协议各不相同。下面总结基于数据驱动方法所使用的评测协议的优劣(见附表)。

常用到的评测协议包括训练/测试集固定划分、与人无关的多折测试、按人划分的多折测试以及跨数据集测试。在这些测试协议中,固定划分的测试协议容易受到训练的偏差影响,未被广泛采用。与人无关的多折测试是指同一个被试者可以同时出现在训练和测试数据中。由于被试者的生理指标在数据录制时基本不会发生较大变化,这一测试协议并不能判断算法是否能基于心动周期相关特征做出正确的预测,因而应该是后续研究中尽量避免采用的评测协议。按人划分的多折测试是指训练集与测试集中的被试者不同。虽然这一评测协议能够去除个体特征的影响,但由于同一数据集录制的环境、被试者的动作较为相似,这样的评测协议不能有效帮助判断模型是否过拟合到训练集的实验环境上,建议未来研究者不应继续将其作为技术论文中唯一的评测协议。跨数据集测试是指训练集和测试集来自不同的数据集。这种评测协议相较于其他测试协议,能够避免被试者个体特征和数据集录制等因素的影响,进而能够更好地帮助评测算法的泛化性能。因此,强烈建议该方向的研究者未来重点采用跨数据集的评测协议。

4 挑战及展望

虽然基于rPPG的生理指标测量在近年取得了较大的进步,但是当前技术仍然无法满足实际应用的需求。急需解决的主要挑战包括:输入视频的质量较低、被试者的头部运动复杂、实际环境的光照条件多变及现有数据规模和多样性不足等。

4.1 输入视频质量较低

输入视频的质量会受到视频压缩和图像预处理两个方面的影响。摄像机在成像时,会出于存储和成像视觉效果等方面的考虑,在底层进行视频压缩及图像预处理,而这一过程很容易破坏皮肤上

微弱的周期性颜色变化信息,对生理指标测量造成影响。

视频压缩对基于 rPPG 的生理指标测量的影响已经被研究者所关注,并出现了一些方法来解决这一因素所造成的信号损失。McDuff 等人(2017)首先研究了不同的编码方式所造成的算法性能损失。之后研究在训练过程引入未压缩的视频数据,使用监督学习的方法来减弱视频压缩所带来的影响(Nowara 和 McDuff,2019;Yu 等,2019b)。如何利用未压缩视频或视频压缩算法的原理来指导生理指标测量是值得重视的研究方向。

除了视频压缩外,成像设备尤其是手机摄像头会在输出视频前对视频图像进行预处理。例如有些手机摄像头为了获得更好的视觉效果,会在输出视频前对视频图像的局部区域进行亮度或对比度的调整。这些处理会改变皮肤区域原有的像素值大小,对生理指标测量造成灾难。现阶段研究者们还未针对该问题进行系统研究,因而非常值得关注。理论上,某些重度的预处理甚至会导致 BVP 信号的完全消失。因此,从系统设计层面尽量基于未经预处理的原始视频数据进行测量更为现实。

4.2 被试者头部运动复杂

被试者在头部运动时,会造成人脸皮肤反射光的变化,进而影响摄像头对人脸皮肤周期性颜色变化的捕捉。常见的被试者头部运动包括头部的平移、旋转以及由眨眼和说话所引起的皮肤的非刚性形变。其中,头部的平移以及平面内旋转可以通过人脸检测和跟踪来消除,对最终的生理指标测量影响较小。而由于眨眼和说话引起的皮肤非刚性形变所涉及到的人脸区域较小,当选取较大的感兴趣区域或剔除形变区域后,其对生理指标测量的影响也可以被抑制。但头部的俯仰和偏航旋转由于会造成皮肤区域的遮挡和皮肤光照反射的变化,对生理指标测量会造成较大的影响。如何消除这一影响,是当前基于 rPPG 的生理指标测量的重要挑战,也是值得关注的未来方向。

此外,现有的评测数据仍然是要求被试者做出某种特定的头部运动,与实际的应用场景有差距。未来,构建在各种实际应用场景下的人脸视频数据集将能帮助研究者们对算法做出更有效的评测和分析。从方法层面讲,通过高精度的人脸对齐或 3D 人脸跟踪技术消除各类运动的影响是值得关注的研

究思路。

4.3 环境光照条件多变

除了输入视频的质量和被试者的头部运动外,环境光照也对算法的鲁棒性有着重要的影响。环境光照可以分为稳定的环境光照和变化的环境光照两种情况。在稳定且明亮的环境光照下,一般可以获得较好的测量结果。但当环境光照较暗时,皮肤对光照的反射和吸收量较小,并且摄像机在成像时会引入较多的噪声,这些因素会给生理指标的测量带来一定的挑战。此外,在一些环境光照会发生变化的场景,例如驾驶场景下,环境光照的变化会叠加到皮肤的周期性颜色变化上,对最终的生理指标测量造成较大的影响。如何在较暗环境以及变化环境下提取 BVP 信号将是未来重要的研究方向。

此外,现阶段所用的评测数据多为实验室采集场景,其环境光照多为可控且单一,且很少考虑环境光照发生变化的情况。因此,构建包含各种挑战性环境光照场景的人脸视频数据集有助于研究者们更有效地分析并提升算法鲁棒性。从方法层面讲,除了通过大规模数据驱动的机器学习方法获得光照鲁棒性之外,对环境光照进行更准确的估计和建模,从而消除其对 BVP 信号提取的影响是非常值得关注的方向。

4.4 现有数据规模和多样性不足

基于 rPPG 的生理指标测量数据规模和多样性严重不足是制约该领域发展的重大障碍。一方面,由于生理指标参考值的获取需要专业的医用设备,现有数据集的数据规模都较小;另一方面,由于现有数据集多为实验室在可控条件下采集的数据,数据多样性与现实应用场景有较大的差距。数据量和数据多样性的不足不仅使得对算法的评测无法做到客观完整,而且就数据驱动的方法而言,更因缺乏足够的训练数据而导致性能难以大幅提高。如何使用相对有限的数据获得泛化性能良好的生理指标测量模型是一个重要的挑战。

针对这一挑战,对于数据驱动的方法而言,未来可能的研究方向有:1) 使用无监督、半监督和弱监督等机器学习方法并充分利用包括无标注样本在内的各种数据,从而提高算法的鲁棒性和泛化性能;2) 充分利用心动周期的生物物理模型和相关知识,从大规模模拟数据生成、生理指标的先验知识约束等方面引导机器学习模型的设计和模型训练过程,

从而提高这类方法的鲁棒性和泛化能力。

5 结语

基于rPPG的生理指标测量研究有着重要的理论意义和应用价值。现有方法通过基于不同假设或先验知识的算法建模,实现了在受限场景下较为准确的生理指标测量。但在更加复杂的现实场景中,基于rPPG的生理指标测量仍然面临较大的挑战,需要通过更加精细的算法建模或利用更大规模的数据来进一步提高算法的鲁棒性。与此同时,现有方法的评测体系并不统一,不同方法间的比较较为困难。本文对现有方法的评测体系从评测任务、评测数据、评测指标和评测协议4个方面进行了系统的归纳,并给出了今后应重点采用跨数据集评测协议的建议。最后,本文详细讨论了基于rPPG的生理指标测量当前所面临的主要挑战,并展望了相应的技术路线或研究思路。

参考文献(References)

- Akselrod S, Gordon D, Ubel F A, Shannon D C, Berger A C and Cohen R J. 1981. Power spectrum analysis of heart rate fluctuation: a quantitative probe of beat-to-beat cardiovascular control. *Science*, 213(4504): 220-222 [DOI: 10.1126/science.6166045]
- Al-Naji A, Gibson K, Lee S H and Chahl J. 2017. Monitoring of cardio-respiratory signal: principles of remote measurements and review of methods. *IEEE Access*, 5: 15776-15790 [DOI: 10.1109/ACCESS.2017.2735419]
- Balakrishnan G, Durand F and Guttag J. 2013. Detecting pulse from head motions in video//Proceedings of 2013 IEEE Conference on Computer Vision and Pattern Recognition. Portland: IEEE: 3430-3437 [DOI: 10.1109/CVPR.2013.440]
- Bobbia S, Macwan R, Benerezeth Y, Mansouri A and Dubois J. 2019. Unsupervised skin tissue segmentation for remote photoplethysmography. *Pattern Recognition Letters*, 124(6): 82-90 [DOI: 10.1016/j.patrec.2017.10.017]
- Bousefsaf F, Maaoui C and Pruski A. 2013. Continuous wavelet filtering on webcam photoplethysmographic signals to remotely assess the instantaneous heart rate. *Biomedical Signal Processing and Control*, 8(6): 568-574 [DOI: 10.1016/j.bspc.2013.05.010]
- Chen W X and McDuff D. 2018. DeepPhys: video-based physiological measurement using convolutional attention networks//Proceedings of the 15th European Conference on Computer Vision. Munich: Springer: 349-365 [DOI: 10.1007/978-3-030-01216-8_22]
- 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]
- 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]
- de Haan G and Van Leest A. 2014. Improved motion robustness of remote-PPG by using the blood volume pulse signature. *Physiological Measurement*, 35(9): 1913-1926 [DOI: 10.1088/0967-3334/35/9/1913]
- Feng L T, Po L M, Xu X Y, Li Y M, Cheung C H, Cheung K W and Yuan F. 2015a. Dynamic ROI based on K-means for remote photoplethysmography//Proceedings 2015 IEEE International Conference on Acoustics, Speech, and Signal Processing. Brisbane: IEEE: 1310-1314 [DOI: 10.1109/ICASSP.2015.7178182]
- Feng T, Po L M, Xu X Y, Li Y M and Ma R Y. 2015b. Motion-resistant remote imaging photoplethysmography based on the optical properties of skin. *IEEE Transactions on Circuits and Systems for Video Technology*, 25(5): 879-891 [DOI: 10.1109/TCSVT.2014.2364415]
- Garbey M, Sun N F, Merla A and Pavlidis I. 2007. Contact-free measurement of cardiac pulse based on the analysis of thermal imagery. *IEEE Transactions on Biomedical Engineering*, 54(8): 1418-1426 [DOI: 10.1109/TBME.2007.891930]
- Gudi A, Bittner M, Lochmans R and Van Gemert J. 2019. Efficient real-time camera based estimation of heart rate and its variability//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshop. Seoul: IEEE: #00196 [DOI: 10.1109/ICCVW.2019.00196]
- Hassan M A, Malik A S, Fofi D, Saad N, Karasfi B, Ali Y S and Meriaudeau F. 2017. Heart rate estimation using facial video: a review. *Biomedical Signal Processing and Control*, 38: 346-360 [DOI: 10.1016/j.bspc.2017.07.004]
- Heusch G, Anjos A and Marcel S. 2017. A reproducible study on remote heart rate measurement [EB/OL]. [2020-06-01]. <https://arxiv.org/pdf/1709.00962.pdf>
- Hsu G S, Ambikapathi A and Chen M S. 2017. Deep learning with time-frequency representation for pulse estimation from facial videos//Proceedings of 2017 IEEE International Joint Conference on Biometrics. Denver: IEEE: 383-389 [DOI: 10.1109/BTAS.2017.8272721]
- Hsu Y C, Lin Y L and Hsu W. 2014. Learning-based heart rate detection from remote photoplethysmography features//Proceedings of 2014 IEEE International Conference on Acoustics, Speech, and Signal Processing. Florence: IEEE: 4433-4437 [DOI: 10.1109/ICASSP.2014.6854440]
- Huang J F, Bai G C, Xiong N X and Wei J G. 2020. Non-touch heart

- rate estimation based on the low-rank and sparse matrix decomposition. *Journal of Graphics*, 41(1) : 66-72 (黄继风, 白国臣, 熊乃学, 魏建国. 2020. 基于低秩稀疏矩阵分解的非接触心率估计. 图学学报, 41(1) : 66-72) [DOI: 10.11996/JG.j.2095-302X.2020010066]
- Khanam F T Z, Al-Naji A and Chahl J. 2019. Remote monitoring of vital signs in diverse non-clinical and clinical scenarios using computer vision systems: a review. *Applied Sciences*, 9(20) : #4474 [DOI: 10.3390/app9204474]
- Kranjec J, Beguš S, Geršak G and Drnovšek J. 2014. Non-contact heart rate and heart rate variability measurements: a review. *Biomedical Signal Processing and Control*, 13 : 102-112 [DOI: 10.1016/j.bspc.2014.03.004]
- Kumar M, Veeraghavan A and Sabharwal A. 2015. DistancePPG: robust non-contact vital signs monitoring using a camera. *Biomedical Optics Express*, 6 (5) : 1565-1588 [DOI: 10.1364/BOE.6.001565]
- Lam A and Kuno Y. 2015. Robust heart rate measurement from video using select random patches//Proceedings of 2015 IEEE International Conference on Computer Vision. Santiago: IEEE: 3640-3648 [DOI: 10.1109/ICCV.2015.415]
- Lewandowska M, Rumiński J, Kocejko T and Nowak J. 2011. Measuring pulse rate with a webcam—a non-contact method for evaluating cardiac activity//Proceedings of 2011 Federated Conference on Computer Science and Information Systems. Szczecin: IEEE: 405-410
- Li C Z, Cummings J, Lam J, Graves E and Wu W. 2009. Radar remote monitoring of vital signs. *IEEE Microwave Magazine*, 10(1) : 47-56 [DOI: 10.1109/MMM.2008.930675]
- Li X B, Alikhani I, Shi J G, Seppanen T, Juntila J, Majamaa-Voltti K, Tulppo M and Zhao G Y. 2018. The OBF database: a large face video database for remote physiological signal measurement and atrial fibrillation detection//Proceedings of the 13th IEEE International Conference on Automatic Face and Gesture Recognition. Xi'an: IEEE: 242-249 [DOI: 10.1109/FG.2018.00043]
- Li X B, Chen J, Zhao G Y and Pietikäinen M. 2014. Remote heart rate measurement from face videos under realistic situations//Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE: 4264-4271 [DOI: 10.1109/CVPR.2014.543]
- Manzone T A, Dam H Q, Soltis D and Sagar V V. 2007. Blood volume analysis: a new technique and new clinical interest reinvigorate a classic study. *Journal of Nuclear Medicine Technology*, 35(2) : 55-63 [DOI: 10.2967/jnmt.106.035972]
- McDuff D, Gontarek S and Picard R W. 2014. Improvements in remote cardiopulmonary measurement using a five band digital camera. *IEEE Transactions on Biomedical Engineering*, 61(10) : 2593-2601 [DOI: 10.1109/TBME.2014.2323695]
- McDuff D J, Blackford E B and Estepp J R. 2017. The impact of video compression on remote cardiac pulse measurement using imaging photoplethysmography//Proceedings of the 12th IEEE International Conference on Automatic Face and Gesture Recognition. Washington: IEEE: 63-70 [DOI: 10.1109/FG.2017.17]
- McDuff D J, Estepp J R, Piasecki A M and Blackford E B. 2015. A survey of remote optical photoplethysmographic imaging methods//Proceedings of the 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Milan: IEEE: 6398-6404 [DOI: 10.1109/EMBC.2015.7319857]
- Mesleh A, Skopin D, Baglikov S and Quteishat A. 2012. Heart rate extraction from vowel speech signals. *Journal of Computer Science and Technology*, 27(6) : 1243-1251 [DOI: 10.1007/s11390-012-1300-6]
- Niu X S, Han H, Shan S G and Chen X L. 2017. Continuous heart rate measurement from face: a robust rPPG approach with distribution learning//Proceedings of 2017 IEEE International Joint Conference on Biometrics. Denver: IEEE: 642-650 [DOI: 10.1109/BTAS.2017.8227252]
- Niu X S, Han H, Shan S G and Chen X L. 2018a. SynRhythm: learning a deep heart rate estimator from general to specific//Proceedings of the 24th International Conference on Pattern Recognition. Beijing: IEEE: 3580-3585 [DOI: 10.1109/ICPR.2018.8546321]
- Niu X S, Han H, Shan S G and Chen X L. 2018b. VIPL-HR: a multimodal database for pulse estimation from less-constrained face video//Proceedings of the 14th Asian Conference on Computer Vision. Australia: Springer: 1-16 [DOI: 10.1007/978-3-030-20873-8_36]
- Niu X S, Zhao X Y, Han H, Das A, Dantcheva A, Shan S G and Chen X L. 2019. Robust remote heart rate estimation from face utilizing spatial-temporal attention//Proceedings of the 14th IEEE International Conference on Automatic Face and Gesture Recognition. Lille: IEEE: #8756554 [DOI: 10.1109/FG.2019.8756554]
- Niu X S, Shan S G, Han H and Chen X L. 2020. RhythmNet: end-to-end heart rate estimation from face via spatial-temporal representation. *IEEE Transactions on Image Processing*, 29 : 2409-2423 [DOI: 10.1109/TIP.2019.2947204]
- Noble D. 1984. The surprising heart: a review of recent progress in cardiac electrophysiology. *The Journal of Physiology*, 353 (1) : 1-50 [DOI: 10.1113/jphysiol.1984.sp015320]
- Nowara E and McDuff D. 2019. Combating the impact of video compression on non-contact vital sign measurement using supervised learning//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshop. Seoul: IEEE: #211 [DOI: 10.1109/ICCVW.2019.00211]
- Osman A, Turcot J and El Kaliouby R. 2015. Supervised learning approach to remote heart rate estimation from facial videos//Proceedings of the 11th IEEE International Conference and Workshops on Automatic Face and Gesture Recognition. Ljubljana: IEEE:

- #7163150 [DOI: 10.1109/FG.2015.7163150]
- Pilz C S, Zaunseder S, Krajewski J and Blazek V. 2018. Local group invariance for heart rate estimation from face videos in the wild// Proceedings of 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. Salt Lake City: IEEE; 1335-13358 [DOI: 10.1109/CVPRW.2018.00172]
- Pilz C S. 2019. On the vector space in photoplethysmography imaging// Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshop. Seoul: IEEE; #197 [DOI: 10.1109/ICCVW.2019.00197]
- Po L M, Feng L T, Li Y M, Xu X Y, Cheung T C H and Cheung K. W. 2018. Block-based adaptive ROI for remote photoplethysmography. *Multimedia Tools and Applications*, 77 (6): 6503- 6529 [DOI: 10.1007/s11042-017-4563-7]
- Poh M Z, McDuff J D and Picard W R. 2010. Non-contact, automated cardiac pulse measurements using video imaging and blind source separation. *Optics Express*, 18 (10): 10762-10774 [DOI: 10.1364/OE.18.010762]
- Poh M Z, McDuff J D and Picard W R. 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]
- Qiu Y, Liu Y, Arteaga-Falconi J, Dong H W and El Saddik A. 2019. EVM-CNN: real-time contactless heart rate estimation from facial video. *IEEE Transactions on Multimedia*, 21 (7): 1778-1787 [DOI: 10.1109/TMM.2018.2883866]
- Soleymani M, Lichtenauer J, Pun T and Pantic M. 2012. A multimodal database for affect recognition and implicit tagging. *IEEE Transactions on Affective Computing*, 3(1): 42-55 [DOI: 10.1109/T-AFC.2011.25]
- Spetlik R, Franc V, Cech J and Matas J. 2018. Visual heart rate estimation with convolutional neural network//Proceedings of British Machine Vision Conference. Newcastle: BMVA Press; #329044207
- Stricker R, Müller S and Gross H M. 2014. Non-contact video-based pulse rate measurement on a mobile service robot//Proceedings of the 23rd IEEE International Symposium on Robot and Human Interactive Communication. Edinburgh: IEEE; 1056-1062 [DOI: 10.1109/ROMAN.2014.6926392]
- Sun Y and Thakor N. 2016. Photoplethysmography revisited: from contact to noncontact, from point to imaging. *IEEE Transactions on Biomedical Engineering*, 63 (3): 463-477 [DOI: 10.1109/TBME.2015.2476337]
- Tarassenko L, Villarroel M, Guazzi A, Jorge J, Clifton D A and Pugh C. 2014. Non-contact video-based vital sign monitoring using ambient light and auto-regressive models. *Physiological Measurement*, 35 (5): 807-831 [DOI: 10.1088/0967-3334/35/5/807]
- Tasli H E, Gudi A and Den Uyl M. 2014. Remote PPG based vital sign measurement using adaptive facial regions//Proceedings of 2014 IEEE International Conference on Image Processing. Paris: IEEE; 1410-1414 [DOI: 10.1109/ICIP.2014.7025282]
- Teichmann D, Brüser C, Eilebrecht B, Abbas A, Blanik N and Leonhardt S. 2012. Non-contact monitoring techniques-principles and applications//Proceedings of 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. San Diego: IEEE; 1302-1305 [DOI: 10.1109/EMBC.2012.6346176]
- Tsouri G R, Kyal S, Dianat S A and Mestha L K. 2012. Constrained independent component analysis approach to nonobtrusive pulse rate measurements. *Journal of Biomedical Optics*, 17 (7): 077011 [DOI: 10.1117/1.JBO.17.7.077011]
- Tulyakov S, Alameda-Pineda X, Ricci E, Yin L J, Cohn J F and Sebe N. 2016. Self-adaptive matrix completion for heart rate estimation from face videos under realistic conditions//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE; 2396-2404 [DOI: 10.1109/CVPR.2016.263]
- Van Gastel M, Stuijk S and De Haan G. 2015. Motion robust remote-PPG in infrared. *IEEE Transactions on Biomedical Engineering*, 62 (5): 1425-1433 [DOI: 10.1109/TBME.2015.2390261]
- Verkruyse W, Svaasand L O and Nelson J S. 2008. Remote plethysmographic imaging using ambient light. *Optics Express*, 16 (26): 21434-21445 [DOI: 10.1364/oe.16.021434]
- Wang W J, Stuijk S and De Haan G. 2015. Exploiting spatial redundancy of image sensor for motion robust rPPG. *IEEE Transactions on Biomedical Engineering*, 62 (2): 415-425 [DOI: 10.1109/TBME.2014.2356291]
- Wang W J, Den Brinker A C, Stuijk S and De Haan G. 2017a. Algorithmic principles of remote PPG. *IEEE Transactions on Biomedical Engineering*, 64 (7): 1479-1491 [DOI: 10.1109/TBME.2016.2609282]
- Wang W J, Den Brinker A C, Stuijk S and De Haan G. 2017b. Color-distortion filtering for remote photoplethysmography//Proceedings of the 12th IEEE International Conference on Automatic Face and Gesture Recognition. Washington: IEEE; 71-78 [DOI: 10.1109/FG.2017.18]
- Wang W J, Den Brinker A C, Stuijk S and De Haan G. 2017c. Amplitude-selective filtering for remote-PPG. *Biomedical Optics Express*, 8 (3): 1965-1980 [DOI: 10.1364/BOE.8.001965]
- Wei L, Tian Y H, Wang Y W, Ebrahimi T and Huang T J. 2012. Automatic webcam-based human heart rate measurements using Laplacian eigenmap//Proceedings of the 11th Asian Conference on Computer Vision. Daejeon: Springer; 281-292 [DOI: 10.1007/978-3-642-37444-9_22]
- Wu H Y, Rubinstein M, Shih E, Guttag J, Durand F and Freeman W. 2012. Eulerian video magnification for revealing subtle changes in the world. *ACM Transactions on Graphics*, 31 (4): #65 [DOI: 10.1145/2185520.2185561]
- Yasuma F and Hayano J I. 2004. Respiratory sinus arrhythmia: why does

- the heartbeat synchronize with respiratory rhythm? Chest, 125(2): 683-690 [DOI: 10.1378/chest.125.2.683]
- Yu Z T, Li X B and Zhao G Y. 2019a. Remote photoplethysmograph signal measurement from facial videos using spatio-temporal networks// Proceedings of the 30th British Machine Vision Conference. Cardiff: BMVA Press
- Yu Z T, Peng W, Li X B, Hong X P and Zhao G Y. 2019b. Remote heart rate measurement from highly compressed facial videos: an end-to-end deep learning solution with video enhancement//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision. Seoul: IEEE: 151-160 [DOI: 10.1109/ICCV.2019.00024]
- Zhou S, Yang X Z, Jin J, Fang S and Liu X N. 2019. Non-contact heart rate detection using self-adaptive signal recovery algorithm. Journal of Image and Graphics, 24(10): 1670-1682 (周双, 杨学志, 金兢, 方帅, 刘雪南. 2019. 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, 24(10): 1670-1682) [DOI: 10.11834/jig.190013]

作者简介



牛雪松,1994年生,男,博士研究生,主要研究方向为计算机视觉、机器学习及情感计算。

E-mail: xuesong.niu@vipl.ict.ac.cn



山世光,通信作者,男,研究员,博士生导师,主要研究方向为计算机视觉与模式识别。

E-mail: sgshan@ict.ac.cn

韩琥,男,副研究员,主要研究方向为计算机视觉与模式识别、医疗影像分析。E-mail: hanhu@ict.ac.cn

附表(Aappendix)

表 基于 rPPG 的心动周期相关生理指标测量代表性算法总结
Table Summary of the rPPG based remote physiological measurement methods

文献	感兴趣区域选择	BVP 信号恢复方法	生理指标检测方法	评测任务	评测指标	测试数据集名称(测试协议)
Verkruyse 等人(2008)	固定区域	直接使用绿色通道信号	频域分析	HR/ RF	定量分析具体 案例,无统计结果	自采数据
Poh 等人(2010)	人脸检测框内区域	ICA	频域分析	HR	RMSE/Bland- Altman 图	自采数据
Poh 等人(2011)	人脸检测框内区域	ICA	时域分析	HR/ RF/ HRV	MAE/SD/RMSE/r HRV	自采数据
Li 等人(2014)	脸颊区域	绿色通道信号 + 剔除背景及运动影响	频域分析	HR	Mean/ Std/ RMSE/ MER/r/小误差 结果占比	自采数据/ MAHNOB-HCI
de Haan 和 Jeanne (2013)	人脸皮肤区域	CHROM	频域分析	HR	SNR/Bland-Altman 图	自采数据
Wang 等人(2015)	人脸皮肤区域	逐像素 CHROM + PCA	频域分析/ 时域分析	HR	SNR/Bland-Altman 图	自采数据
Tasli 等人(2014)	自适应选择人脸区域	直接使用绿色通道信号	频域分析/ 时域分析	HR	MAE	自采数据
Feng 等人(2015a)	自适应选择人脸区域	直接使用感兴趣区 域内颜色均值	频域分析	HR	RMSE/Bland- Altman 图	自采数据
Po 等人(2018)	自适应选择人脸区域	ICA/小波变换 修正	频域分析	HR	r/Bland-Altman 图	自采数据/ PURE

续表

文献	感兴趣区域选择	BVP信号恢复方法	生理指标检测方法	评测任务	评测指标	测试数据集名称(测试协议)
Bobbia 等人(2019)	自适应选择人脸区域	CHROM	频域分析	HR	r/RMSE/小误差结果占比/SNR	UBFC-RPPG
Kumar 等人(2015)	人脸检测框内区域,针对不同局部感兴趣区域加权求和	直接使用感兴趣区域内颜色均值	频域分析/时域分析	HR/HRV	SNR/RMSE/Bland-Altman 图	Distance-PPG
De Haan 和 Van Leest(2014)	人脸皮肤区域	给定相机的特性下的颜色空间投影	频域分析	HR	SNR/小误差结果占比	自采数据
Wang 等人(2017a)	人脸皮肤区域	颜色空间投影	—	—	SNR	自采数据
Feng 等人(2015b)	脸颊区域	人脸运动状态下的反射模型	时域分析	HR	r/Bland-Altman 图	自采数据
McDuff 等人(2015)	脸颊、额头区域	五通道视频上使用ICA	时域分析	HR/HRV	r	自采数据
Lam 和 Kuno(2015)	人脸检测框内区域	不同像素间使用ICA	频域分析 + 投票机制	HR	MAE/RMSE/小误差结果占比	MAHNOB-HCI
Lewandowska 等人(2011)	脸颊、额头区域	PCA	频域分析	HR	定量分析具体案例,无统计结果	自采数据集
Tsouri 等人(2012)	人脸检测框内区域	cICA	频域分析	HR	定量分析具体案例,无统计结果	自采数据集
Wei 等人(2012)	人脸检测框内区域	拉普拉斯特征映射	频域分析	HR	Bland-Altman 图	自采数据集
Pilz(2019)	人脸皮肤区域	黎曼流形空间降维	频域分析	HR	RMSE/r	UBFC-RPPG/LGI Multi-Session
Tulyakov 等人(2016)	脸颊区域	CHROM + 自适应矩阵补全 + PCA	频域分析	HR	Mean/Std/RMSE/MER/r	MAHNOB-HCI/MMSE-HR
黄继风等人(2020)	脸颊区域	CHROM	频域低秩分解	HR	Mean/Std/RMSE/MER/r/Bland-Altman 图	MAHNOB-HCI
周双等人(2019)	人脸检测框内区域	CHROM + 矩阵低秩分解	频域分析	HR	MAE/SD/RMSE/MER/r/	自采数据集
Wang 等人(2017b)	人脸皮肤区域	原始信号频域修正	频域分析	HR	SNR/Success-rate	自采数据集
Wang 等人(2017c)	人脸皮肤区域	基于正常心率范围的频域修正	频域分析	HR	SNR/Success-rate	自采数据集
Gudi 等人(2019)	脸颊、额头区域	颜色空间投影 + 剔除头部运动在频域的影响	时域分析	HR/HRV	MAE	自采数据/VicarPPG/PURE/MAHNOB-HCI
Niu 等人(2017)	脸颊区域	CHROM + PCA	频域分析 + 分布学习	HR	Mean/Std/RMSE/MER/r/Bland-Altman 图	MAHNOB-HCI
Hsu 等人(2014)	人脸检测框内区域	原始颜色信息 + ICA	频域特征回归	HR	RMSE/r/Bland-Altman 图	自采数据(两折)

续表

文献	感兴趣区域选择	BVP 信号恢复方法	生理指标检测方法	评测任务	评测指标	测试数据集名称(测试协议)
Osman 等人(2015)	人脸检测框内区域	绿色通道 + 手工设计特征分类波峰所在位置	时域分析	HR	Mean/Std/RMSE/r	自采数据集(固定训练测试划分)
Chen 等人(2019)	人脸检测框内区域	运动表征 + 含注意力分支的卷积神经网络	频域分析	HR/ RF	MAE/SNR	自采数据集(与人无关划分, 按人划分, 按场景划分)/ MAHNOB-HCI (跨数据测试)
Niu 等人(2018a)	人脸皮肤区域	时空表示 + 模拟数据预训练的卷积神经网络	回归	HR	Mean/Std/RMSE/ MER/r/Bland-Altmann 图	MAHNOB-HCI/ MMSE-HR (均为与人无关三折测试)
Niu 等人(2020)	人脸皮肤区域	时空表示 + 考虑时序信息的网络结构	网络回归	HR	Mean/Std/RMSE/ MER/r/Bland-Altmann 图	MAHNOB-HCI/ MMSE-HR (均为与人无关三折测试, 跨数据集测试)/ VIPL-HR(按人划分五折测试)
Qiu 等人(2019)	脸颊区域	多尺度时空表示 + 卷积神经网络	网络回归	HR	Mean/Std/RMSE/ MER/r	MMSE-HR (视频段与人无关的固定划分)
Niu 等人(2019)	人脸皮肤区域	时空表示 + 数据增广 + 注意力机制卷积网络	网络回归	HR	Mean/Std/RMSE/ MER/r/Bland-Altmann 图	VIPL-HR(按人划分五折测试)/ MMSE-HR (跨数据集测试)
Hsu 等人(2017)	人脸皮肤区域	时频表示 + 卷积神经网络	网络分类	HR	Mean/Std/RMSE/ MER/r	PURE/PFF(训练)/ MAHNOB-HCI (跨数据集测试)
Spetlik 等人(2018)	人脸检测框内区域	级联全卷机网络	网络回归	HR	MAE/RMSE/r	COHFACE/ MAHNOB-HCI/ PURE/ECG-Fitness (均为固定划分)
Yu 等人(2019a)	人脸检测框内区域	3 维卷积网络	时域分析	HR/ RF/ HRV	Std/MAE/RMSE/r	OBF(按人划分十折测试)/ MAHNOB-HCI (跨数据集测试)

注:对于非监督学习的方法,测试数据集将全部用于测试,在表中将不做专门标注;与人无关的测试协议是指训练和测试集中可以出现同一个人,“—”代表原文献中未进行最终的生理指标测量。