基于摄像头的心率估计数据集创建流程

Mohamed Moustafa[1,2], Amr Elrasad[2], Joseph Lemley[2],

Peter Corcoran [1,2]

[1]戈尔韦大学，爱尔兰戈尔韦；[2]
Xperi Corporation，爱尔兰戈尔韦；

摘要

心率是最重要的健康指标之一，可用于研究并深入了解各种人体生理和心理信息。因此，不受接触式传感器限制的心率估计成为一个极具吸引力的研究领域，因为它能够在更广泛的场景中实现健康状况监测。由此，人们开发了多种基于摄像头的心率估计技术，其范围从传统图像处理到复杂的深度学习模型和架构。这类研究工作的核心在于健康与视觉数据的采集、清理、转换和标注。本文探讨如何为开发或测试基于面部区域图像进行心率估计的算法或机器学习模型准备数据。所准备的数据包括摄像头帧以及来自心电图传感器的传感器读数。所提出的处理流程分为四个主要步骤：移除错误数据、消除帧和心电图时间戳的抖动、信号去噪与滤波，以及创建帧标注。我们的主要贡献是一种消除健康传感器和摄像机时间戳抖动的新技术，以及一种对视觉帧和心电图传感器数据进行精确时间对齐的方法，该方法也适用于其他类型的传感器。

关键词：数据准备、心率、面部图像、机器学习、深度学习、时间序列分析

1. 引言

心率被认为是健康监测中最重要的指标之一。心率会受到多种因素的影响，例如药物[1], 、体温变化[2], 、认知状态[3], 和体能水平。监测心率有助于了解寿命、代谢综合征[4], 、心脏病发作风险[5], 、猝死风险[6],、情绪[7], 和社会行为[8]。

因此，即使在医院和医疗机构之外，心率估计也具有重要意义。在这些场景中，依赖接触式传感器要么过于不便，甚至可能带来危险（例如，传感器可能分散驾驶员的注意力，延缓其反应）。在驾车[9] 或于危险环境中工作[10] 等特定且频繁的活动期间监测个人心率尤为重要。

基于摄像头的心率估计是一种便捷可靠的传统接触式心率估计替代方案，适用于紧急医疗护理场景之外（例如驾车、工作或逛商店时） [11]。这种方法通过将视觉传感器（通常是摄像机）对准受试者来估计其心率。由此，视觉心率估计的研究与应用不断取得进展[12,13] ，所用数据也不再局限于可见光摄像头，还扩展到了近红外[14] 和热成像[15] 摄像机。

要实现这些改进，关键在于更容易获取大量详尽的心率数据集 [16–18] ，以便在新技术投入使用之前对其进行训练、调优和测试。对于机器学习和深度学习方法而言，这一点尤为重要，因为它们能够充分利用大量数据，在超越传统技术性能的同时，更好地应对环境和受试者姿势的变化，而传统图像处理方法在这方面表现较差[12]。然而，无论研究人员选择使用公开数据，还是自行收集数据，了解数据集准备和数据处理都至关重要。遗憾的是，已发表的研究通常要么省略某些数据准备步骤，要么只提及特定算法实现所必需的数据准备步骤，而对非创新性步骤一带而过。

在本研究中，我们提出了一种新方法，可利用采集到的视觉数据和健康传感器数据，创建时间对齐准确的心率数据集。我们的方法包括对齐时间，并对所采集样本的时间戳进行去抖动处理，而大多数研究人员都忽略了这一步。我们还提出了一种新颖的方法，用于消除健康传感器和摄像机时间戳中的抖动，并提出了一种方法，可准确地对齐视觉帧与心电图（ECG）传感器数据的时间。此外，我们的方法不仅适用于创建心脏信号数据集，也可用于其他采用已知采样率传感器的应用。据我们所知，我们提出的算法和处理流程此前尚未发表。

本文其余部分的组织结构如下：第2节探讨相关文献。第3节和第4节介绍我们提出的处理流程及实验结果。第5节给出结论性说明。

2. 相关工作

与完全聚焦于数据准备处理流程的出版物相比，在主要贡献为新数据集或估计方法的论文中，更常见的是找到数据准备细节。即便如此，时间对齐和数据同步的细节也常常被略过或讨论得不够充分。

[19] 将数据准备步骤作为其主要贡献之一进行了详细说明。使用Pan-Tompkins算法将心电图读数转换为RRI信号[20]，随后通过处理检测并移除异位搏动，即由异常脉冲引起的异常搏动。之后，使用小波法对RRI信号进行去趋势处理，再通过三次样条插值进行重采样。

在 [21], 中，视频序列被分解为时间上保持一致的超像素，并从中提取初步的远程光电容积描记（rPPG）信号。随后，进一步处理该信号，以增强从皮肤区域提取的生理信息：利用傅里叶变换和双阶跃函数，针对第一和第二谐波计算信噪比（SNR）。然后，根据SNR计算加权平均值，以重建rPPG信号。

在[22], 中，作者提出了一种基于卡尔曼滤波器的方法，用于解决多个传感器异步测量的时间同步问题。[23, 24] 进一步对此进行了改进，以考虑软件延迟和丢失的测量值。

在数据挖掘中的数据准备背景下，一项系统映射研究（SMS）[25] 对临床数据集中预处理技术的应用进行了分析与分类，涵盖了100多篇论文。该系统映射研究得出结论：医学领域最常见的预处理任务是数据简化和数据清洗。

3. 数据准备

本节将介绍为基于摄像头的心率估计任务准备数据集所涉及的各项任务。

3.1 错误数据的剔除

数据采集后，首先应确保数据已正确保存且未发生任何数据损坏。随后应检查数据（必要时可进行人工检查），以确保其质量令人满意。利用人脸视频估计心率时，每一帧中都应至少部分可见人脸。尤其是在长时间、多摄像机的设置中，即使受试者被要求不要移动，他们也并不少见会移出视野。某些健康传感器（例如脉搏血氧仪（SpO2）传感器）在从原位置移动或受到晃动时，可能无法正确读取数据。在某些情况下，生成的信号不仅噪声更大，还可能缺乏提取心率所需的足够信息。检测并剔除此类数据片段，是创建可靠数据集的重要步骤；否则，这些数据会给数据引入不必要的噪声，妨碍算法利用输入数据中的趋势来提取或预测信息。

3.2 时间戳去抖

由于最终需要将帧与传感器读数同步，以便为帧添加标注，因此时间戳在数据集准备中发挥着重要作用。如果数据不同步，帧可能会被赋予不正确的标签，导致模型训练和评估出现错误。

在数据采集过程中，设备发热、传感器与计算机之间的通信通道，以及使用实验室／公司编写的软件进行采集（通常用于同时从不同制造商生产的传感器和摄像机采集数据）等因素，会导致采样和捕获速率波动、数据丢失，还会使数据采集与时间戳生成之间产生轻微延迟（当时间戳并非由传感器本身生成，而是由软件生成时）。减少此类波动有助于提高使用传感器标签标注帧的准确性，我们将这一过程称为时间戳去抖动。

我们在图1中展示的时间戳去抖动方法，基于将捕获的时间戳映射到合成生成的等间隔时间戳。该方法对采样率非常敏感，因此计算时应格外谨慎，确保达到较高的准确度，且不依赖任何假设（例如设备标称采样率）。

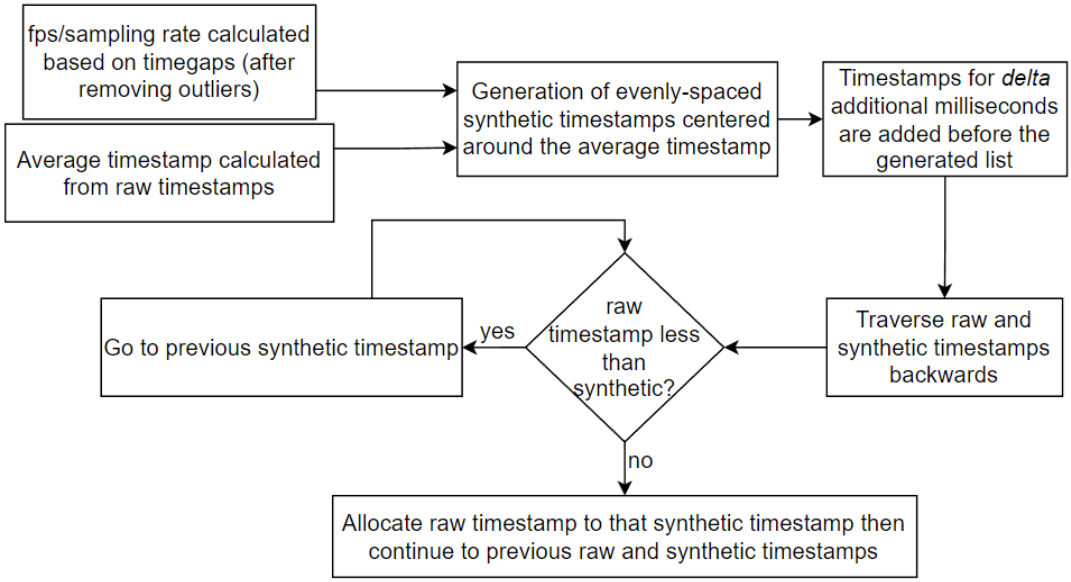

图1. 时间戳去抖动整体流程示意图。分配完所有捕获的时间戳后，流程即告结束。

首先，计算数据采样率：排除超出 m 个标准差的时间步长异常值后，求出相邻时间戳之间的平均差值（即每对连续时间戳之间的时间间隔），下文将其称为时间步长。采样率等于平均时间步长的倒数，通常略低于摄像机或传感器的标称帧率/采样率（例如，对于一台 30fps 的摄像机，计算得到的采样率为 29.9fps）。

接下来，先计算平均时间戳，再据此生成其前后两侧的合成时间戳。具体来说，通过不断从平均时间戳中减去或加上此前计算出的平均时间步长，生成等间隔的合成时间戳，使其覆盖捕获的原始时间戳原本覆盖的时间段。合成时间戳还会向捕获时间戳的起始位置之前额外延伸 delta 毫秒，以便去除采集开始时的时间戳延迟（也就是说，如果第一个时间戳有所延迟，校正后应将其映射到更早的时间点）。

将捕获的时间戳映射到合成时间戳，需要从后向前遍历捕获的时间戳。对于每个原始时间戳，都将其映射到第一个尚未分配且数值相等或更小的合成时间戳。这是因为时间戳的正确值不可能大于记录值，而记录值中包含了我们希望去除的延迟和抖动。扩展合成时间戳覆盖的时间段，是为了补偿第一个时间戳中可能存在的延迟（这意味着合成时间戳不能从第一个记录的时间戳开始）。

3.3 信号处理

为了更好地从捕获的健康传感器记录中提取整体趋势和信息，捕获的信号会经过若干处理算法，如图5所示。

首先，进行平滑处理以降低噪声并突出整体峰值。为此，我们使用萨维茨基-戈莱滤波器，对每个窗口中的点拟合w个多项式函数。我们发现，使用大小为 101 的窗口并拟合二次函数，对于我们的使用场景而言效果令人满意，实验结果部分将进一步讨论这一点。

接下来，我们对数据应用巴特沃斯带通滤波器，得到一个周期波，用于将帧标注为真实标注。由于我们希望滤除心率信号中的噪声，因此根据文献中人体心率范围的平均值，使用截止频率 [0.7、2.5] Hz [12]。

3.4 标注创建

由于健康传感器的采样率通常高于摄像机帧率，因此需要进行降采样，为每一帧创建标签。具体通过时间戳完成，如图 2 所示。

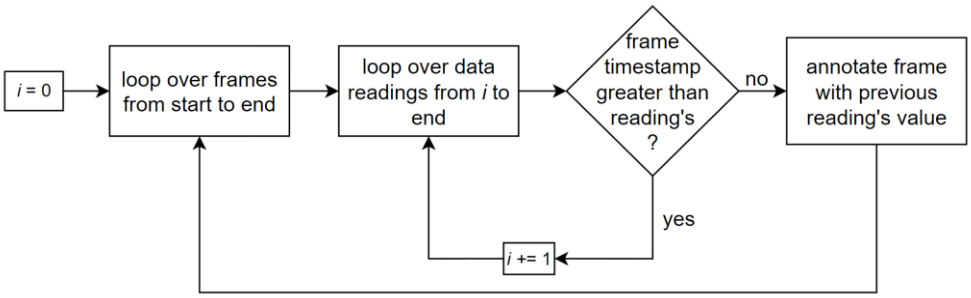

图 2. 帧标注示意图

对于每一帧，使用时间戳值小于该帧时间戳值且最接近的心电图信号来标注该帧。此方法基于两个假设：读数不会在时间戳之后产生，且时间上最接近的传感器读数能够最准确地代表该帧的生理信息。我们使用全局变量存储上一次遍历的心电图读数的索引，从而将算法的时间复杂度从 $O(n^2)$ 降低到 $O(n)$。

4. 实验结果

本节旨在说明将上述数据准备步骤应用于一个内部数据集后所获得的结果。该数据集使用帧率为每秒 30 帧的摄像机和 1kHz 心电传感器采集。在采集的数据中，帧时间间隔的均值和标准差分别为 33.33 ms 和 11.52 ms，心电读数时间间隔的均值和标准差分别为 1.00 ms 和 2.63 ms。这些结果使用 m=3 得出，并排除了与均值相差 3 个标准差的时间步。移除所有错误数据后，例如噪声信号区段（通过目视观察发现）或受试者移出视野时对应的帧，再对摄像机和传感器获取的时间戳进行去抖动处理，以便更可靠地在时间上对齐摄像机数据与传感器数据。

此过程通过将采集的时间戳映射到合成生成的时间戳来完成。然而，并非所有合成时间戳都会分配到捕获的时间戳（即并非所有合成时间戳都会被使用）。这意味着与这些时间戳对应的读数或帧被丢弃（即未被采样或捕获），如图 3 所示，这表现为时间间隔峰值。

丢失帧或读数可能由多种因素造成，例如采集系统无法处理传入的数据流、硬件无法跟上，或实验装置中的其他瓶颈；实验装置包括传感器/摄像机、连接（无线连接或有线连接）、采集计算机和采集软件。

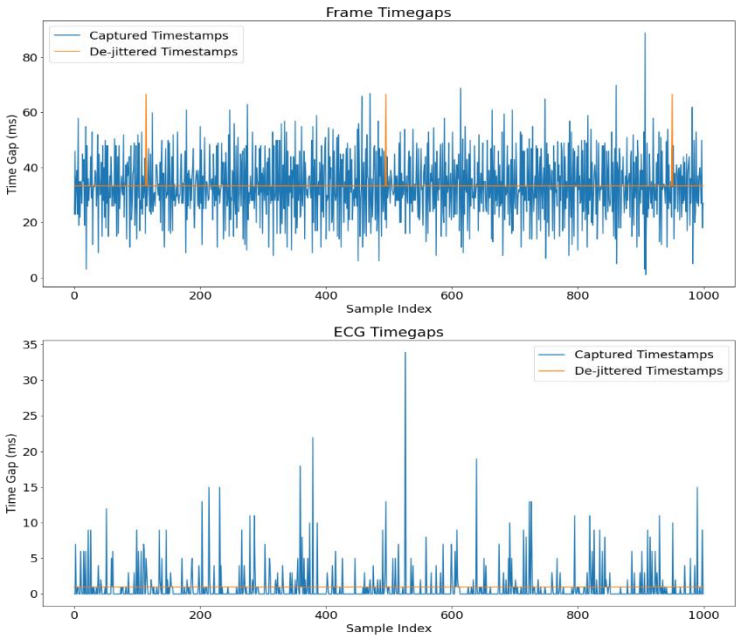

图3. 捕获和去抖动后的帧时间步（上）以及捕获和去抖动后的心电图时间步（下）样本

将我们的方法和 [23,24] 中介绍的基于卡尔曼滤波器的方法应用于我们的数据后，我们的方法将帧时间间隔的标准差降至 1.64 ms。与此同时，卡尔曼滤波器方法将均值提高至 54 ms，同时将标准差降至 5.47 ms。应用于心电图时间间隔时，我们的方法将标准差降至 0.005 ms，而卡尔曼滤波器方法将均值提高至 1.62 ms，同时将标准差降至 2.05 ms。由于数据中存在丢失的帧/读数，我们的方法会产生非零标准差。此特性使我们的方法能够轻松识别读数或帧丢失的时间点。下方的表1将两种算法的结果与原始时间戳的均值和标准差进行比较，而图4则比较了数据子样本上的输出。

	帧时间戳		心电图时间戳	
	均值（毫秒）	标准差（毫秒）	均值（毫秒）	标准差（毫秒）
Raw	33.33	11.517	1.00	2.629
卡尔曼滤波方法	53.92	5.470	1.62	2.053
我们的方法	**33.33**	**1.642**	**1.00**	**0.005**
理论最优值	33.33	0.000	1.00	0.000

表1. 基于卡尔曼滤波器的算法与所提出算法的实验结果比较

如图所示，卡尔曼滤波器方法在时间间隔抖动变化较大的情况下效果不佳。如图所示，他们的方法会受到较高时间间隔抖动的严重影响，不适合用于滤除或校正时间间隔抖动，以使时间戳更接近预期值。在他们的论文中，用于展示该方法性能的数据总体抖动似乎约为 0.2 毫秒，这显著低于我们的实验数据中的抖动。

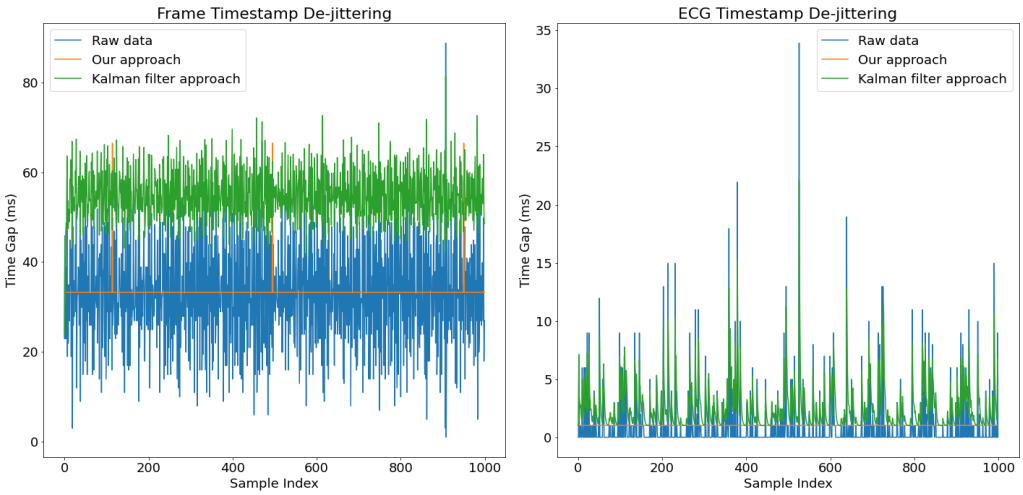

图 4. 我们的方法与卡尔曼滤波器方法在帧数据和心电图数据样本上的比较 le

随后，对心电图信号进行平滑处理，以去除噪声，同时保留主要趋势。我们发现，对每 101 个数据点拟合一个二次函数可以获得最佳结果。窗口大小为 101 是在尝试了其他各种宽度后确定的。窗口较大则会导致过度平滑，从而削弱数据中存在的主要峰值；窗口较小则不足以去除噪声。但是，如果传感器采样率发生变化，就必须选择不同的窗口大小。

平滑处理后，使用截止频率为[0.7Hz - 2.5Hz] 的带通滤波器，以获得最终的标签值，如图5所示，这些值将映射到摄像头帧上。该频率范围在心脏信号相关文献中十分常见。最后，由于摄像机帧率通常低于传感器采样率，因此需要进行降采样；为确保降采样后的传感器数据与高采样率处理后的读数在时间上正确对齐，必须格外谨慎。一个快速验证方法是绘制两个信号的时域图，如图6所示。

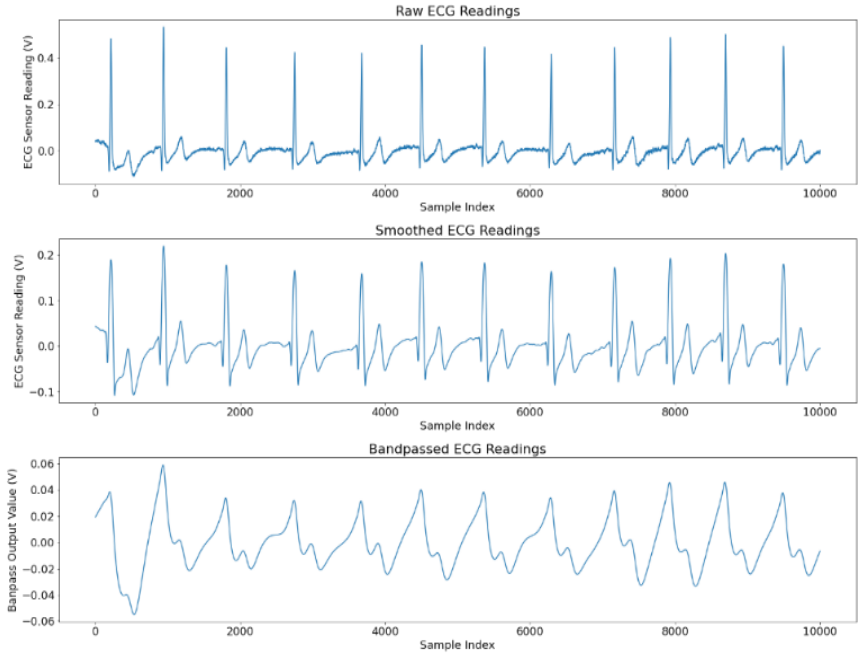

图 5. 采集到的原始 心电图信号（上）、使用萨维茨基–戈莱滤波器平滑后的心电图信号（中）以及带通滤波器输出 ut（下）

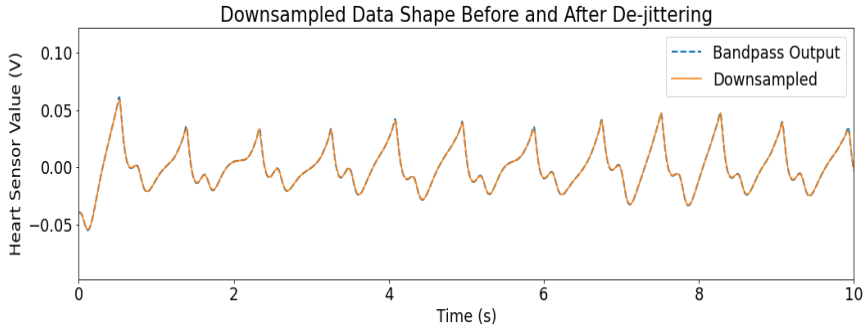

图 6。降采样至摄像机帧率前后心电图数据形状的比较。

5. 结论

在这项工作中，我们提出了一种用于去除传感器和摄像机时间戳抖动的新技术，并介绍了为机器学习和图像分析技术准备视觉心率数据所涉及的工作和基本任务。所提出的处理流程首先移除所有不可用数据，随后进行时间戳校正、信号分析与处理，最后进行帧标注。虽然我们是在心率信号估计这一背景下介绍去抖动方法，但只要已知传感器采样率，该方法就可用于校正任何采集时间戳中的抖动。

我们提出的处理流程可确保视觉传感器和脉搏传感器的时间戳均经过准确的去抖动处理并完成对齐，从而为每一帧准确标注脉搏信号。根据所开发的方法，还可以采取其他步骤，或使用不同的算法。例如，某些深度学习模型会将连续帧之间的差异作为输入 [12]。不过，即使发表的结果中没有明确提及，我们在本文概述的步骤通常也是大多数数据集准备处理流程的一部分。

致谢

本出版物所述研究由爱尔兰研究委员会通过项目编号 EBPPG/2021/92 资助，属于与 Xperi Corporation 合作开展的就业计划研究生奖学金项目。

参考文献

[1] R. P. Nolan、P. Jong、S. M. Barry-Bianchi、T. H. Tanaka 和 J. S. Floras，“药物、生物行为和运动疗法对冠状动脉疾病患者心率变异性的影响：系统综述”，《欧洲预防心脏病学杂志》15(4)，386–396（2008年）。[2] A. D. Jose、F. Stitt 和 D. Collison，“运动和体温变化对男性固有心率的影响”，《美国心脏杂志》79(4)，488–498（1970年）。[3] N. Hjortskov、D. Rissen、A. K. Blangsted、N. Fallentin、U. Lundberg 和 K. Søgaard，“精神压力′ 对计算机工作期间心率变异性和血压的影响”，《欧洲应用生理学杂志》92(1)，84–89（2004年）。[4] X. Jiang、X. Liu、S. Wu、G. Q. Zhang、M. Peng、Y. Wu、X. Zheng、C. Ruan 和 W. Zhang，“代谢综合征与静息心率相关，且可由静息心率预测：一项横断面和纵向研究”，《心脏》101(1)，44–49（2015年）。[5] K. Fox、I. Ford、P. G. Steg、M. Tendera、M. Robertson、R. Ferrari 等，“心率作为冠状动脉疾病和左心室收缩功能障碍患者的预后风险因素（BEAUTIFUL研究）：一项随机对照试验的亚组分析”，《柳叶刀》372(9641)，817–821（2008年）。

[6] M. Woodward、R. Webster、Y. Murakami、F. Barzi、T.-H. Lam、X. Fang、I. Suh、G. D. Batty、R. Huxley 和 A. Rodgers，“静息心率、心血管疾病与死亡率之间的关联：来自12个队列中112,680名男性和女性的证据”，《欧洲预防心脏病学杂志》21(6)，719–726（2014年）。[7] W. Yu、S. Ding、Z. Yue 和 S. Yang，“利用知识图谱，从面部表情和非接触式心率识别情绪”，载于《2020年IEEE国际知识图谱会议（ICKG）》，64–69页，IEEE（2020年）。[8] T. Colasante 和 T. Malti，“静息心率、内疚与同情：一项关于身体攻击行为的发展心理生理学研究”，《心理生理学》54(11)，1770–1781（2017年）。[9] H. Rahman、M. U. Ahmed、S. Barua 和 S. Begum，“利用生理和车辆参数进行基于非接触方式的驾驶员认知负荷分类”，《生物医学信号处理与控制》55，101634（2020年）。[10] R. Sharma、D. Goel、M. Srivastav 和 R. Dhasmana，“在重症和非重症监护病房工作的护士，其心率和皮肤电反应的差异”，《临床与诊断研究杂志》12(11)（2018年）。[11] B. Huang、C.-M. Chang、C.-L. Lin、W. Chen、C.-F. Juang 和 X. Wu，“基于卷积神经网络的面部视频视觉心率估计”，载于《2020年第15届IEEE工业电子与应用会议（ICIEA）》，1658–1662页，IEEE（2020年）。[12] X. Liu、J. Fromm、S. Patel 和 D. McDuff，“用于设备端非接触式生命体征测量的多任务时间移位注意力网络”，《神经信息处理系统进展》33，19400–19411（2020年）。[13] X. Niu、X. Zhao、H. Han、A. Das、A. Dantcheva、S. Shan 和 X. Chen，“利用时空注意力从面部进行鲁棒远程心率估计”，载于《2019年第14届IEEE国际自动人脸与手势识别会议（FG 2019）》，1–8页，IEEE（2019年）。[14] W. Chen 和 D. McDuff，“Deepphys：利用卷积注意力网络进行基于视频的生理测量”，载于《欧洲计算机视觉会议论文集（ECCV）》，349–365页（2018年）。[15] K. Hamedani、Z. Bahmani 和 A. Mohammadian，“用于心率估计的热成像视频序列时空滤波”，《专家系统与应用》54，88–94（2016年）。[16] Z. Zhang、J. M. Girard、Y. Wu、X. Zhang、P. Liu、U. Ciftci、S. Canavan、M. Reale、A. Horowitz、H. Yang 等，“用于人类行为分析的多模态自发情绪语料库”，载于《IEEE计算机视觉与模式识别会议论文集》，3438–3446页（2016年）。[17] S. Koelstra、C. Muhl、M. Soleymani、J.-S. Lee、A. Yazdani、T. Ebrahimi、T. Pun、A. Nijholt 和 I. Patras，“DEAP：用于情绪分析的数据库；使用生理信号”，《IEEE情感计算汇刊》3(1)，18–31（2011年）。[18] M. Soleymani、J. Lichtenauer、T. Pun 和 M. Pantic，“用于情感识别和隐式标注的多模态数据库”，《IEEE情感计算汇刊》3(1)，42–55（2011年）。[19] V. Timothy、A. S. Prihatmanto 和 K.-H. Rhee，“基于心率变异性分析和机器学习的自动诊断数据准备步骤”，载于《2016年第6届国际系统工程与技术会议（ICSET）》，142–148页，IEEE（2016年）。[20] J. Pan 和 W. J. Tompkins，“实时QRS检测算法”，《IEEE生物医学工程汇刊》(3)，230–236（1985年）。[21] S. Bobbia、R. Macwan、Y. Benezeth、A. Mansouri 和 J. Dubois，“用于远程光电容积描记的无监督皮肤组织分割”，《模式识别快报》124，82–90（2019年）。[22] J.-O. Nilsson 和 P. Handel，“多传感器导航系统中异步传感器测量的时间同步与时间排序”，载于《IEEE/ION PLANS 2010论文集》，897–902页（2010年）。[23] T. Huck、A. Westenberger、M. Fritzsche、T. Schwarz 和 K. Dietmayer，“多传感器融合中的精确时间戳标记与时间同步”，载于《2011年IEEE智能车辆研讨会（IV）》，242–247页，IEEE（2011年）。[24] A. Westenberger、T. Huck、M. Fritzsche、T. Schwarz 和 K. Dietmayer，“面向未来驾驶辅助系统的多传感器融合时间同步”，载于《2011年IEEE测量、控制与通信精密时钟同步国际研讨会》，93–98页，IEEE（2011年）。[25] A. Idri、H. Benhar、J. Fernandez-Alem´ an 和 I. Kadi，“知识´ 发现中的医疗数据预处理系统综述图谱”，《生物医学计算方法与程序》162，69–85（2018年）。