第1章 绪论

1.1 研究背景及意义

手势识别作为一项关键的人机交互技术，在过去几十年中取得了显著的发展。最早期的手势识别系统主要基于传感器和相机技术，但受限于计算能力和算法的限制，其性能相对较弱。随着计算机视觉、深度学习和传感技术的飞速发展，手势识别进入了一个新的阶段。计算机视觉的高度进步使得系统能够更准确地捕捉和理解手势，而深度学习的应用为手势的复杂特征提取提供了更强大的工具。在过去的几年中，随着智能手机、AR/VR设备和智能家居的普及，手势识别技术得到了广泛应用。例如，智能手机的解锁手势、游戏中的姿势控制、虚拟现实中的手势操作等，都展示了手势识别在日常生活和工作中的潜力。手势识别技术也在医疗1–3、康复4–6和工业领域7–9发挥着越来越重要的作用。在医疗方面，手势识别可用于手术室中的非接触式操作，以减少感染风险。在康复中，通过追踪患者的手势，系统可以提供实时的反馈和指导。在工业领域，手势识别被应用于生产线上的操作和控制。未来，随着技术的不断创新，手势识别有望在更多领域发挥作用，为人机交互提供更加自然和智能的方式。

而近年来，随着可穿戴技术的迅速发展，智能手环和智能手表等生态链产品逐渐跻身成为用户生活中不可或缺的一部分。这些小巧、轻便且可穿戴的设备通过集成各类传感器如加速度计、陀螺仪以及光电容积描记术（Photoplethysmography，PPG）传感器等，拓展了它们的功能，使其不仅仅局限于传统的健康监测和通知提醒，而且成为了实现手势交互的理想平台。

其中，基于PPG传感器的手势识别技术备受研究者关注。PPG传感器的主要功能是测量皮肤微血管区域的血流情况，而这项技术最初被广泛应用于心率监测。然而，由于PPG能够捕捉到细微的血流变化，使其不仅在心率监测领域表现出色，同时在手势识别方面也展现出了巨大的潜力。通过在可穿戴设备上引入PPG传感器，用户的手势动作可以通过对皮肤微血管区域的光电信号进行分析而得以识别。这种基于PPG的手势识别技术不仅为用户提供了一种自然而直观的交互方式，而且避免了对传统输入设备的依赖，如鼠标和触摸屏。因此，基于PPG的手势识别系统成为当前研究的热点之一，为可穿戴设备带来了更广泛的应用前景。

1.2 国内外研究现状

根据手势信息获取方式的不同，目前主要存在的手势识别技术可以分为以下四类10：计算机视觉技术、射频技术、声学技术和可穿戴式传感器技术，其中声学和射频技术的基本原理类似，可以归为一类。下面将分别对这三类手势识别技术的研究现状展开介绍。

1.2.1 手势交互系统

1. 基于计算机视觉的手势交互系统

计算机视觉的手势识别技术通过对摄像头捕获的图像进行分析，实现对用户手部动作的实时识别。这项技术的发展为人机交互带来了革命性的变革，使得用户可以通过自然而直观的手势与计算机进行沟通。其基本原理是通过深度学习和图像处理技术，提取和分析手部在图像中的特征，从而理解用户的意图。

而计算机视觉的手势识别技术经过长久的发展，已有大量研究取得了不错的成绩，且已经有许多基于计算机视觉的手势识别类型的电子产品投入商用。例如，Microsoft于2010年推出的Kinect设备广泛运用了计算机视觉手势识别技术，Kinect可以捕捉用户的身体动作和手势，使其成为Xbox等游戏平台上的互动控制方式11。2018年，Amazon Go便利店采用了基于计算机视觉的手势识别技术，实现了“无人收银”概念，用户只需在入口扫描二维码，然后系统通过对用户手势和动作的追踪，自动识别用户拿取的商品，并完成支付流程，无需人工干预12。此外，在医疗保健领域，2024年，SYK公司使用Blueprint混合现实系统完成了首例远程肩关节置换术手术，该系统由微软的HoloLens 2头戴显示设备、Blueprint混合现实软件和外围显示器组成，利用HoloLens 2配备的前置100万像素深度摄像头实现了全关节捕捉的手势识别，从而辅助医生完成了手术13。

然而，这一技术仍然面临一些挑战。首先，计算机视觉系统对光照和背景比较敏感，可能在复杂环境中导致识别准确性下降。其次，在多人交互等复杂场景下，系统可能难以准确识别手势，尤其是涉及到遮挡的情况。使用深度学习模型对图像进行处理需要较大的计算资 源，可能导致系统运行效率下降。

2. 基于可穿戴式设备的手势交互系统

基于可穿戴式设备的手势识别系统主要依赖外部的穿戴式设备，如肌电臂环、数据手套等，收集手势信息实现手势的检测和识别。其原理在于利用设备中集成的肌电、惯性等传感器单元采集手势相关的生理信号或动态数据，然后通过数据处理和特征提取的方式实现对手势的识别。例如，肌电臂环可以测量肌肉的电活动，而数据手套则通过柔性伸展传感器和惯性传感器捕捉手部运动的动态信息。这些信号经过一系列处理和分析，提取其关键的特征，最后输入到机器学习模型中进行手势的识别。

此类手势识别系统在医疗康复、虚拟现实、体感游戏等领域都已经存在商业应用。例如，

2006年，任天堂公司推出了Wii家用游戏机，其通过加速度计和陀螺仪首次实现了体感游戏，开启了游戏交互方式的新篇章14。加拿大Thalmic Labs公司于2013年推出的MYO手环集成了肌电传感器和加速度传感器，能够捕捉用户手臂的肌肉运动和手势，用户可以通过手势控制计算机、游戏或其他智能设备15。Facebook于2021年提出的EMG腕带交互，通过集成多个EMG传感器阵列，并结合外部摄像头可以捕捉整个手部在三维空间的动作，是Facebook未来AR、VR重要的交互方式之一16。

但目前此类手势识别系统仍然存在很多没有解决的问题。例如，可穿戴式设备的传感器受到尺寸和功耗等因素的制约，精度可能不如实验室级的设备；由于设备尺寸和传感器数量的限制，可穿戴式设备对于手势的表示空间相对有限；对于某些手势，如手指之间的微小动作，设备可能受到局部干扰，导致误识别。这些问题也导致了此类系统在实际应用过程中的不良效果。

3. 基于声学和射频信号的手势交互系统

基于声学和基于射频信号的手势识别系统都是利用波的传播和反射特性。一般步骤包括发送波信号和接收其反射，借助手势对反射信号的干扰，通过分析反射信号的时间延迟、强度和频谱等信息，系统可以识别手势动作。

因为该类手势识别系统相比于基于计算机视觉的手势识别系统而言，对环境要求并不高，且其可以实现较远距离的手势识别，所以长期以来，有不少研究者致力于此类手势识别系统的研究。Luo17等人利用声学传感器，在Android系统上实现了特征提取和手势识别的实时处理，并在不同的实验设置下充分测试了其抗噪性能，使用7种典型手势，识别准确率可达91%。2019年，Vincent Becker18等人利用智能手表中的麦克风和加速度计采集到的信号作为输入，提出了一种用于手势识别的轻量级卷积神经网络架构，专门设计用于在资源受限的设备上本地运行，该架构对九个不同手势实现了 97.2% 的用户独立识别准确率。2020年，Dian19等人提出了基于射频技术的细粒度手势识别系统，他们在深度学习架构中引入对抗模型，达到了16个常用美国手语平均90%的识别准确率。Yang20等人于2023年提出了一个中文手势识别系统，该系统基于射频技术，在相应的识别场景下准确率达到98%左右。

基于声学的手势交互系统对环境噪音和杂音比较敏感，可能导致性能下降。且声波传播存在一定范围和方向性，导致系统在一定范围内才能有效识别手势。而且其需要设备配备有足够数量的麦克风、扬声器或其他音频组件，这也致使该类系统难以商用。而基于射频信号的手势识别系统则需要提前布置专用设备，且极容易出现视距干扰问题。

1.2.2 PPG手势识别

1. PPG技术原理

根据朗伯－比尔定律21（Beer-Lambert Law），当一束平行单色光垂直通过均匀非散射的吸光介质时，光强和介质之间存在如式（1）所示的定量关系。

式中，为出射光强度，为入射光强度，为吸光系数（与吸光介质的性质及入射光的波长有关），为吸光介质的浓度，是吸光介质的厚度。

光学体积描记术（Photoplethysmography, PPG）技术的理论基础就是Beer-Lambert Law。典型的PPG信号采集包括两个主要步骤：光源发射和光电检测。首先利用发光二极管向人体皮肤表面发射一定频率的光源，一般是红外或者绿光，光在皮肤表面处会发生反射及透射。根据Beer-Lambert Law，光经过的人体的各个部分都会对光有所吸收，其中肌肉、骨骼和静脉等组织对光的吸收基本保持不变，而动脉里的血液浓度会随心跳周期性变化。因此，透射入人体的光线会产生一定的损失，并且会随着血液浓度的变化而变化。再通过光电二极管将反射或透射后的光信号转换为电信号，提取其中的交流成分，就可以得到人体血液流动的相关信息。

2. PPG手势识别现状

PPG信号不仅用于心率监测，而且越来越多地被应用于手势识别领域。手势动作导致了血流流动，这也会在PPG信号中留下特定的模式，从而可以实现对手势动作的识别。与传统的手势捕获技术相比，PPG手势识别具有独特的优势：其不容易受环境因素影响，且PPG传感器轻巧、低成本，在大部分商业穿戴式设备（如Apple Watch、HuaWei Watch GT系列）上都已集成。这为开发便携、鲁棒的手势交互系统提供了更为可行的解决方案。

过去几年中，已经有研究者对基于PPG信号的手势识别进行了初步探索。2018年，Zhao等人最早尝试使用PPG信号进行手势识别，他们采用梯度提升树实现了9种不同手势的识别，在10名受试者上可以达到88%的平均识别率，这也证明了PPG信号用于手势识别的可行性22。后续研究中，Zhao23等人又引入ACC信号，并设计了基于ResNet的神经网络，将识别率提升到了98%。Subramanian24等人通过开展一系列实验，比较了PPG信号和sEMG信号在手势识别中的表现，使用自制的三通道 PPG 传感器腕带对4位受试者的4种简单手势进行分类的准确率为93.36%，与使用MYO臂带的识别准确率（92.80%）基本相当，得出了 PPG技术有望成为手势识别应用中一种新的可替代方法的结论。Zhang25等人对 20 位受试者的 10 种手势开展识别研究，获得了90.55%的识别准确率和 90.73%的召回率，是领域内首次在商业腕带式可穿戴设备上开展的基于PPG 技术的手势识别研究。Ling26 等人进行了一项多角度的比较研究，对比了基于加速度计和PPG传感器的手势识别技术,通过在四种不同的运动场景中涉及手腕和手指运动的14种手势的识别实验，他们得出结论，PPG 技术更适合在可穿戴式设备上实现手势交互。Li27等人探索了使用PPG信号同时测量力水平及手势类别的可行性，他们设计了３种力水平及４类手势，组合等到12种识别类别，并进行了相应的识别实验，得到了90%以上的识别率。Zhou28等人则验证了PPG手势识别用于用户身份验证的可行性，通过基于PPG手势识别设计的认证机制，合法用户的通过率达96.67%，而非法用户则仅有0.62%。Li和Zhou的研究验证了PPG手势识别可以适用于各类交互场景。

3. PPG手势识别的主要问题

PPG手势识别技术在取得显著成果的同时，仍存在一些尚待克服的不足之处。在成果方面，已经成功地将PPG技术应用于手势识别领域，实现了在理想条件下的高效识别。但其在模型训练和实际应用等方面仍然存在一系列问题亟待解决。

在模型训练方面，主要挑战是如何培养一个具有鲁棒性的模型，以有效缓解甚至解决传感器位置偏移问题。PPG传感器位置偏移问题源于手势执行过程中，肌肉和肌腱对动脉的直接压迫，这种直接的血流变化在靠近动脉的PPG传感器位置上更为显著23。实际应用中，像手表、手环这样的可穿戴设备经常需要被反复佩戴，比如在洗澡后再次佩戴或者在擦拭汗液后再次佩戴。即使手表没有经历重复佩戴，长时间佩戴和不同强度的运动也可能导致表盘背面与手腕皮肤之间的微小偏移。这种传感器位置偏移的情况会导致训练数据和测试数据之间的特征分布产生显著差异，从而降低手势识别分类器的性能。为了在每次使用时都能保持较高的识别性能，用户必须在每次使用时重新采集足够的数据用于训练分类器，以确保系统的稳定性能。然而，这种方式会带来沉重的训练负担和糟糕的用户体验。因此，从实际应用的角度来看，模式识别分类算法的鲁棒性和用户训练负担等问题是决定基于PPG技术的手势交互产品是否能够成功应用的关键因素。在解决这一问题上，需要深入研究和开发有效的方法，以确保模型能够在不同传感器位置偏移的场景中保持稳定的性能。

而在实际应用方面，PPG手势识别也面临着实时性、算法模型移植等方面的挑战。实时性是指在识别手势时需要快速响应，确保用户在交互中获得即时的反馈。PPG信号的采集和处理可能涉及大量的数据，而实时性要求系统在处理这些数据时能够迅速而准确地做出响应，这对于设计高效的实时算法提出了挑战。另一个问题是算法模型的移植。在实际应用中，PPG手势识别算法可能需要在不同的设备或平台上部署，例如从研究阶段的计算机环境到嵌入式系统或移动设备。在这个过程中，需要解决不同平台之间的兼容性和性能适配问题，确保算法在各种硬件条件下都能够有效运行。这牵涉到算法的优化、轻量化以及对不同硬件架构的适应性研究。因此，为了实现PPG手势识别技术在实际场景中的广泛应用，需要在提高实时性的同时，关注算法模型的通用性和可移植性，以便在各种设备和应用场景中都能够稳定可靠地运行。这也是未来研究和开发中需要集中精力解决的关键问题。

1.3 本文研究目标与研究内容

如上所述，基于PPG的手势识别技术有望克服传统手势识别技术所面临的一系列问题，且由于大部分商业手环、手表等设备都已集成了PPG传感器，所以无需额外硬件，即可实现基于PPG的手势交互系统。因此，本研究旨在设计并实现一种基于多通道PPG技术的便携式手势交互系统，实现在不同运动状态下对手势的高鲁棒性识别，提供一种具有实用性和新颖性的交互应用示范。具体而言，研究内容将分为三个主要方面：

（1）采集前端设计

成功构建了一个多通道PPG数据采集系统。通过合理布局和排列多通道PPG传感器，我们确保了对手势变化的高度敏感性。针对市面上现有设备可能遇到的问题，如传感器位置偏移和佩戴时的变化，我们设计了相应的解决方案，提高了采集系统的稳定性和准确性。

（2）克服传感器偏移的手势分类算法研究

针对PPG传感器位置偏移带来的手势分类算法鲁棒性差的问题，提出了一套手势识别算法流程。本套算法流程充分结合了采集前端的特殊设计，首先通过将位置作为标签进行分类，以剔除与测试位置数据分布差异过大的训练位置，再通过时间窗采样、梯度提取等一系列预处理流程，减小测试数据与训练数据的特征分布差异。本套算法流程有效地解决了传感器偏移问题对手势分类器性能的影响。通过多轮实验和优化，成功提升了系统的鲁棒性，使其在传感器偏移场景下都能保持卓越的手势识别性能。

（3）交互终端程序设计及性能验证

最后，进行了交互终端程序的设计和性能验证。借助先前两个阶段的研究成果，设计了一款用户友好、实用性强的交互应用程序。该程序能够实时响应用户的手势，并提供直观的界面以增进用户体验。通过一系列性能验证实验，充分展示了系统在实时性和对手势的高鲁棒性识别等方面的能力。用户可以在日常生活中进行各类手势，系统都能稳健地进行准确的识别，展示了该系统在实际应用中的可行性和实用性。