

红外图像下的手部姿态估计

SZ2316132 郑晓豪
计算机科学与技术学院

1 具体要解决的问题

针对红外图像下的手部姿态估计任务进行解决方案的设计。手部姿态估计任务类似人体姿态估计任务，手部姿态估计的主要特点在于手部目标检测的是大目标。目前，人体姿态估计作为研究的热点已经被广泛研究，但是针对手部姿态估计任务的研究较少。同时，针对红外图像的手部姿态估计任务的更少，主流研究集中于可见光图像的研究。而红外图像天然具有温度信息，现有的针对于可见光图像的解决方案不适用于红外图像。因此，针对红外图像下的手部进行姿态估计是具有现实意义的，作为红外图像下的手势识别任务的上游任务，为其提供研究的基础。

2 拟使用的方法

针对手部姿态估计任务的特点，采用深度学习的方法对问题尝试解决。首先，通过目标检测的方法对红外图像上的手部目标进行检测。确保手部目标的完整性的同时，为后续姿态估计任务提供基础。然后，建立手部关键点检测的网络模型，并尝试将温度相关的特征结合进图卷积网络中。在搭建网络模型时，采用 mmdet 和 mmpose 等开源框架帮助快速的构建模型的主干网络。通过两个阶段，手部姿态估计任务最终会形成红外图像上手势对应的骨架序列。

3 具体数据来源

数据集采用的是自建数据集，采集设备为智能手机热像仪 Infray。数据集共包含 6756 张带标注红外手势图像，由九名实验人员共同采集数据，包含十六种以上常见的军用手势。

不同成员采集的手势类别并不固定，且不同成员采集的样本数量也不固定。实验人员的手部特征均不相同，基本能够涵盖绝大多数特征的手部。数据集中的手势，基本能够涵盖绝大多数特种兵手语交流场景。所有手势的带标注图像都在 300 张以上，能够保证每种手势都有足够的带标注数据。