

高级机器学习作业

室内行人移动方位推算技术

由于室内 GPS 信号受到阻隔，同时车库场景无线接入点 AP 较为稀疏，导致室内定位很难通过 GPS 或位置指纹等技术实现，但行人位推算技术（Pedestrian Dead Reckoning, PDR）可以有效解决这一问题，即在 GPS 阻隔条件下，利用手机内置传感器（包括不限于加速计、陀螺仪、磁强计），设计轻量级低功耗算法运行框架和技术，实现稳定可靠、高精度的方位推算计算。

主要工作

利用手机内置的传感器收集信息，推算设备的位置信息。

A. 自行收集数据

1. 数据收集部分可以多组合作完成。
2. 在实验过程中，手机大部分时间保持在同一个状态下（如一直拿在手上）。为了收集到的经纬度数据尽量准确，推荐在开阔的户外收集数据。
3. 自行决定收集训练数据的数量，但测试数据中每条数据不少于 2min，至少包含 10 条数据。
4. 收集的数据应该包括不同状态下采集到的数据，包括但不限于拿在手上，放在背包里，放在裤兜里等。
5. 在基础 PDR 算法使用的传感器可以根据自身设备情况选择，最终测试数据集中包含以 50HZ 频率收集到的如下传感器的数据：加速度计，线加速度，陀螺仪，磁力计，压力传感器。

B. 设计 PDR 算法

由于使用 GPS 定位高度误差较大，并且高度主要由压力传感器提供。所以本次作业只推测当前的经纬度'Latitude (°)', 'Longitude (°)'以及方位角'Direction (°)'。要求在测试数据上，通过训练数据训练出的模型可以获得前 10% 时间段的传感器信息和方位信息的情况下，在后 90% 时间段内依靠传感器信息预测出位置信息。

模型使用的算法没有限制，可以自行选择自己认为合理的算法。比如，可以直接把这个任务看成一个线性或者非线性的回归问题，也可以将此任务看成一个时间序列预测的问题，或者训练模型分别预测移动方向和移动速度，从而计算出当前位置，亦或利用强化学习解决这个问题等。但需要在实验报告中体现出该方法的合理性。

C. 考虑模型迁移

设计算法时，需要考虑模型可能运行在不同的设备上，所以实验将提供一个测试集上测试模型迁移性能。该测试集包含 10 组在多个设备上，不同状态下收集到的数据。数据的长度为 10min，前 10% 时间的数据有传感器和方位数据，模型需要做的是在后面 90% 时间段中仅根据传感器数据预测方位信息。

D. 评价指标

因为经纬度的尺度并不相同，所以在评测位置预测性能时，会计算每个 GPS 定位点下估计的位置和真实位置的直线距离的平均值作为模型的评价指标（具体参考 [test.py](#)）。

提交内容

1. 实验报告。实验报告中需要包含如下内容：
 - a. 数据的基本信息，即收集了多少数据，收集了什么状态下的数据，在实验中如何对数据进行划分。
 - b. 尝试使用了哪些方法，为什么使用这些方法，方法的效果怎么样，如果效果不好，可能的原因是什么，如何解决这些问题。
 - c. 如何运行代码，在自己收集数据上的性能如何。
 - d. 小组成员分工。

- 2. 数据及可执行的代码。
- 3. 在测试数据上的结果。

提交方式后续会公布在课程主页中。

评分标准

- 1. 50 pt 基础 PRD 算法的完成情况。
- 2. 30 pt 代码可运行性，包括算法在收集的数据上的运行情况，以及在迁移数据上的运行情况，包括模型推理准确性以及推理速度等。
- 3. 20 pt 迁移场景的完成情况。
- 4. 附加分 1：在测试数据集上，预测移动方位平均误差小于 15° 获得 5pt 附加分，超过 90% 误差均小于 15° 获得 15pt 附加分。
- 5. 附加分 2：性能最好的三组获得额外的加分，本附加分直接加在期末总成绩上。

注：1,3 两项主要分数主要取决与于实验报告；2 项主要由代码的运行情况决定基础的 PDR 算法与模型迁移分别占这部分成绩的 2/3 和 1/3。

备注

数据收集方法

数据收集采用 [phyphox](#) 软件，该软件可以提供了 Android 和 IOS 版本的安装包。在安装软件后，进入软件后点击“+”（Android 版本在右下角，IOS 版本在左上角）新建简易实验。最终测试数据集的实验设置如下图所示。

取消	新建简易实验	存储
标题		
50		
选择传感器		
加速度计		✓
线加速度		✓
GPS定位		✓
陀螺仪		✓
磁力计		✓
压力传感器		✓
距离传感器		

在点击右上角三个点之后，可以选择“定时运行”，推荐在收集数据时使用该功能。同时在该弹框中可以导出数据，在导出数据选项部分选择 CSV(Comma, decimal point)。

代码

参考 HW1 中的代码和数据。在提交的代码中，请勿使用绝对路径。

每一个测试数据中包含的如下文件

纯文本

```
Accelerometer.csv
Barometer.csv
Gyroscope.csv
Linear Accelerometer.csv
Magnetometer.csv
Location_input.csv
```

其中，Location_input.csv 为用于测试的方位信息，最后提交的结果参考 Location_output.csv。最后针对每一个 test case 提交一个 csv 文件，命名为 Location_test_case_i.csv, i 为第 i 个 test case。