# 计步器算法测试报告

**版本：1.1**

By 张琛

2013-11-04

目录

[计步器算法测试报告 1](#_Toc371702704)

[1. Changelog 3](#_Toc371702705)

[2. 编写目的 3](#_Toc371702706)

[3. 测试环境 3](#_Toc371702707)

[4. 数据预处理 3](#_Toc371702708)

[4.1 捕获 3](#_Toc371702709)

[4.2 插值 3](#_Toc371702710)

[4.3 分割 4](#_Toc371702711)

[5. 参数设定 4](#_Toc371702712)

[5.1 基础阈值 4](#_Toc371702713)

[5.2 滤波参数 4](#_Toc371702714)

[5.3 数据轴选择 4](#_Toc371702715)

[5.4 阈值-步频关系参数 5](#_Toc371702716)

[6. 测试结果 5](#_Toc371702717)

[6.1 单一文件训练&&测试 5](#_Toc371702718)

[6.2 单一文件训练&&对多文件测试 7](#_Toc371702719)

[6.3 所有数据文件集合训练并测试 8](#_Toc371702720)

[6.4 加入不同用户、不同步行速度数据 9](#_Toc371702721)

[7. 误差分析 10](#_Toc371702722)

[8. 参考资料 10](#_Toc371702723)

## Changelog

|  |  |  |  |
| --- | --- | --- | --- |
| 版本号 | 变更人 | 变更说明 | 变更时间 |
| V1.0 | 张琛 | 初稿 | 2013/10/23 |
| V1.1 | 张琛 | 重点标注文中的误差对比、结论与分析 | 2013/11/8 |

## 编写目的

本文旨在测试根据论文[1]实现的计步算法准确度，并对算法实现细节和改进[[1]](#footnote-1)[1]进行测试。

## 测试环境

平台：PC，

操作系统：windows7 x64

编程语言：python

## 数据预处理

### 捕获

手机放置在右裤口袋，启动程序采集 -> 静止等待3s，行走指定的步数，停下静止3s -> 停止采样。数据保存为指定格式的xml [[2]](#footnote-2)[2] 。

数据命名格式为：<Name><step-cnt><speed>\_a9\_0[1,2,…].xml

举例：

ZC30fast\_a9\_3.xml 表示张琛以正常（包括较快）步速，行走30步，第四次采集；

HY60slow\_a9\_0.xml 表示胡钰以较慢（间隔 >1s）步速，行走60步，第一次采集。

### 插值

因为算法中实际计算频率的方法为 1/(index2-index1)，如果采样不等间隔，(index2-index1)不能准确对应时间间隔。为此，将传感器数据插值为 30Hz。

### 分割

因为手机在放入衣兜、取出的时候存在较大幅度的震动，通过简单的分割，使得数据只包括静止与行走两种状态。

## 参数设定

### 基础阈值

数据训练时，假定振动幅度大于此基础阈值才可能为一步；若小于此阈值，则不计入样本点。数值过小，则误将“余震”计入样本点；数值过大，则有的步被遗漏。

### 滤波参数

行走过程中产生的“余震”，使得一步并非只有一个峰值，为了训练时自动寻找峰值对应区间，此处使用python的scipy.signal.firwin方法生成hamming窗，用scipy.signal.lfilter方法进行低通滤波。Firwin方法参数设定为经验值：

窗口长度numtaps=10, （不能太大，否则振幅太低）

截断频率cutoff=40,

奈奎斯特频率nyq=1600/2

滤波效果如下图1所示，其中洋红色为原始数据，靛青色为滤波后的数据，可以看到，原先的多个峰值在滤波之后大多减为单峰值。

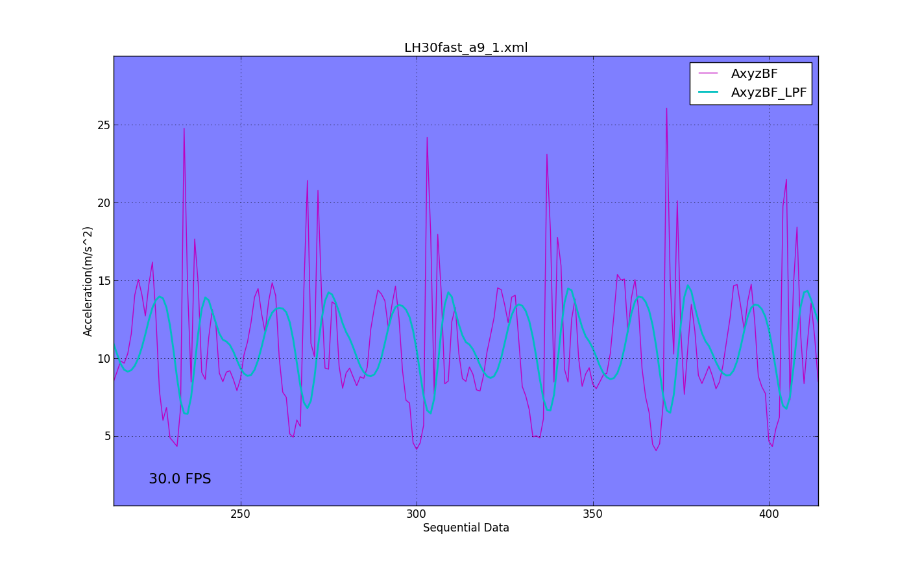


图 1 低通滤波效果

### 数据轴选择

分别对机身坐标合加速度（含重力），对应图1中**AxyzBF\_LPF**曲线，以及世界坐标重力方向Z轴加速度**AzWF\_LPF**，进行训练和测试。

理论上我们应该依赖世界坐标重力方向Z轴加速度AzWF\_LPF，但是基于以下三个原因，我们选择AxyzBF\_LPF作为训练、测试使用数据：

1. 坐标转换需要依赖多种传感器，耗电量增加，并且有些设备只具有加速度传感器；
2. 有些设备因为传感器之间采样频率不一致，转换为世界坐标后，反而有误差；
3. 对于我们测试中的应用场景，手机位于裤兜时，步行产生的加速度主要表现在竖直分量上。包含重力的机身坐标合加速度数据能较好的模拟实际竖直分量数据。

实际测试结果表明，AxyzBF\_LPF不比 AzWF\_LPF测试准确度低。

### 阈值-步频关系参数

依据论文[1]中算法，训练目的在于求解中的。如果（步幅，步频）点集分布不理想，即无法得到二者的正相关下确界，解决办法见设计报告V1.3[[3]](#footnote-3)[1]。

## 测试结果

### 单一文件训练&&测试

**主要目的**：对比AxyzBF\_LPF与AzWF\_LPF准确度,

示例文件：ZC60fast\_a9\_4.xml

基础阈值设定 baseTh=2

表 1测试结果对比：

|  |  |  |  |
| --- | --- | --- | --- |
|  | Step |  |  |
| AzWF\_LPF | 66 | 0.339072663486 | 1.99681171853 |
| AxyzBF\_LPF | 60 | 13.0370776228 | 2.38212155379 |

**AzWF\_LPF训练结果：**

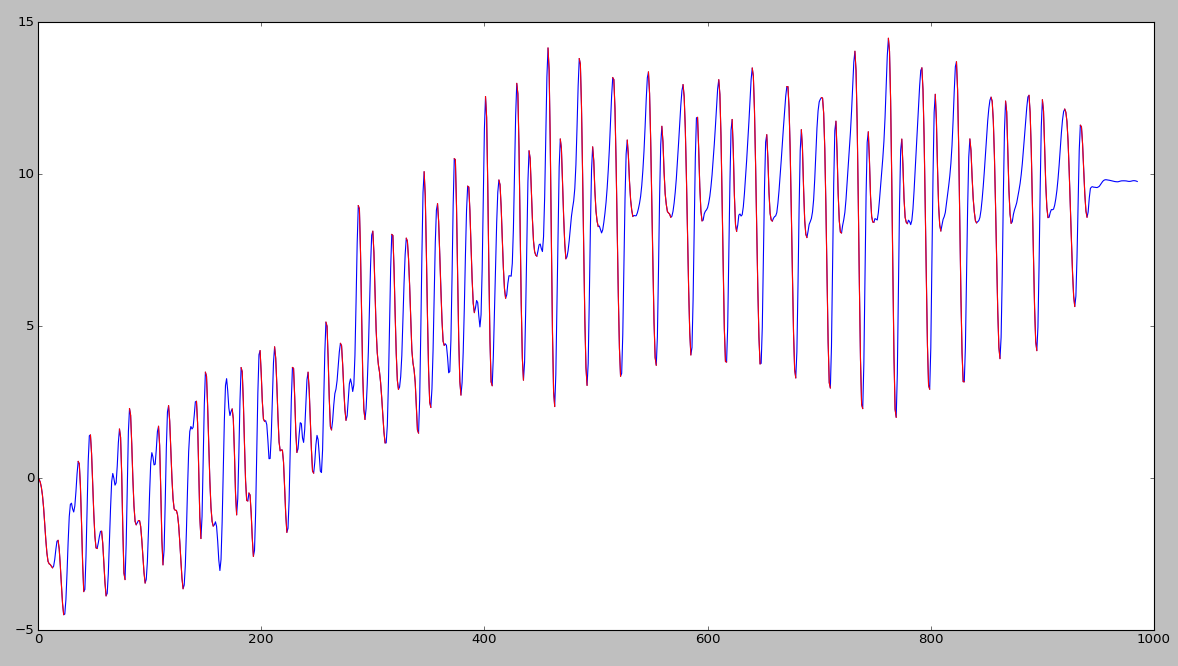
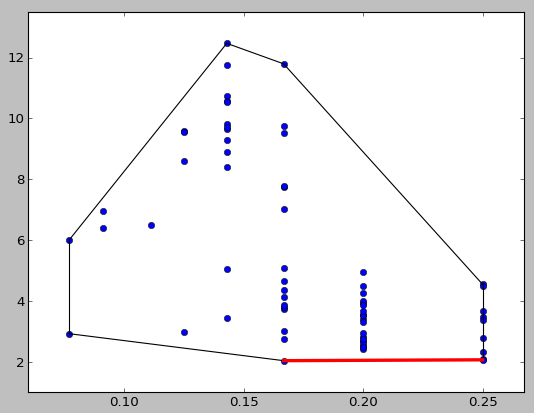
 

图 2 训练过程，自动标记波峰-波谷区间 图 3 计算下确界（图中红线），即

**AzWF\_LPF测试结果：**

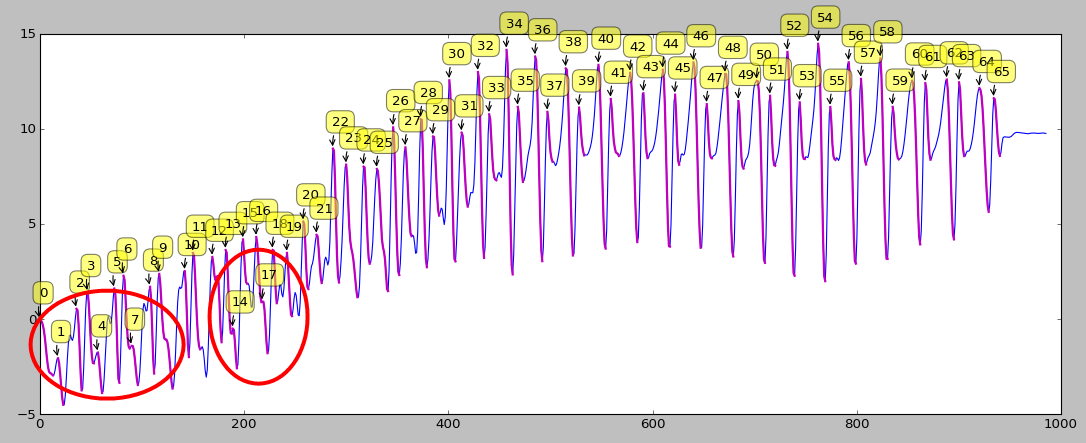
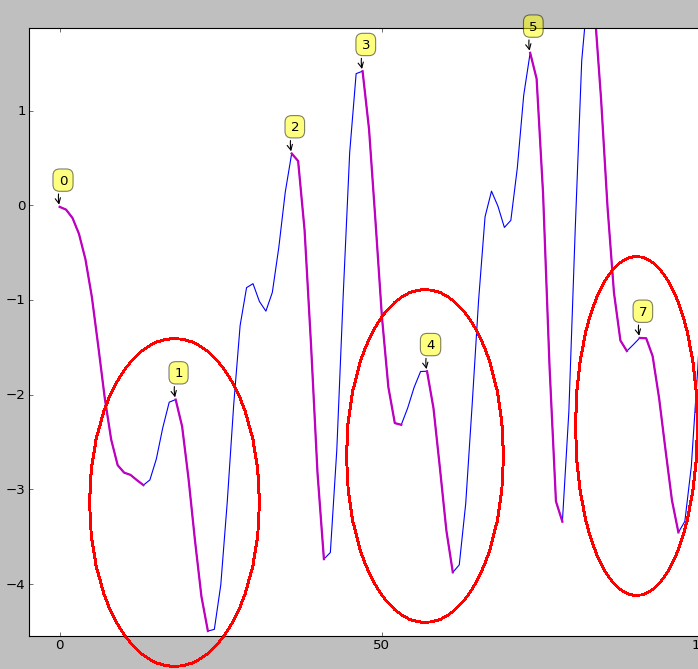
 

图 4 测试结果，存在错误计步 图 5 错误计步的局部放大图

**AxyzBF \_LPF训练结果：**

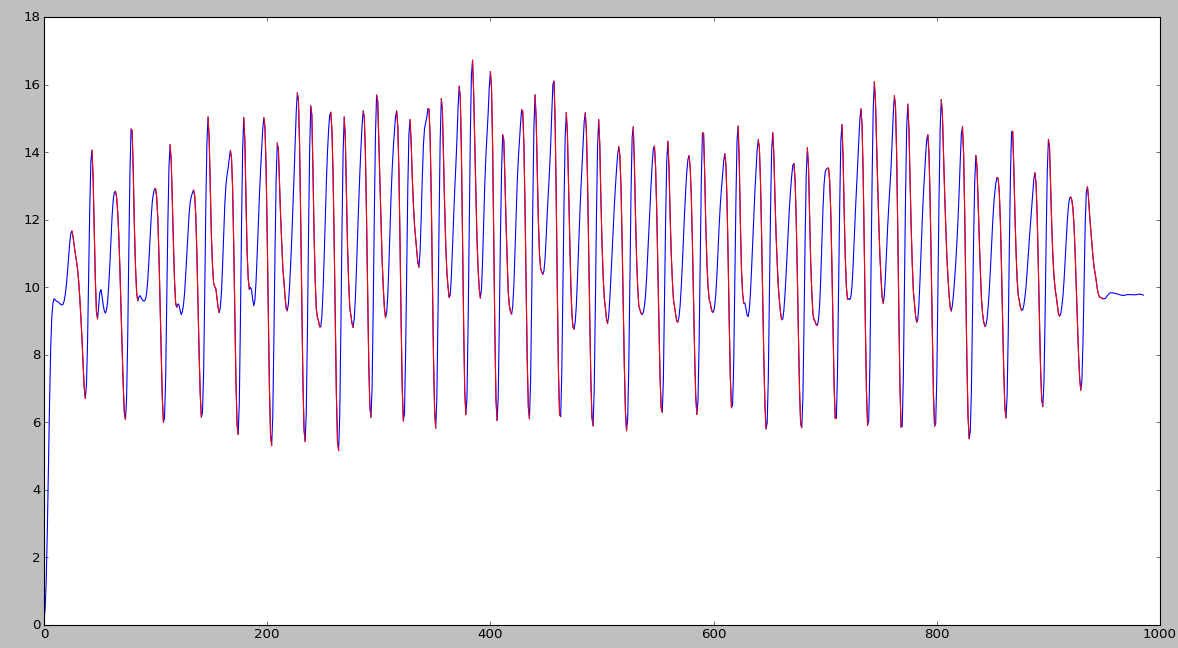
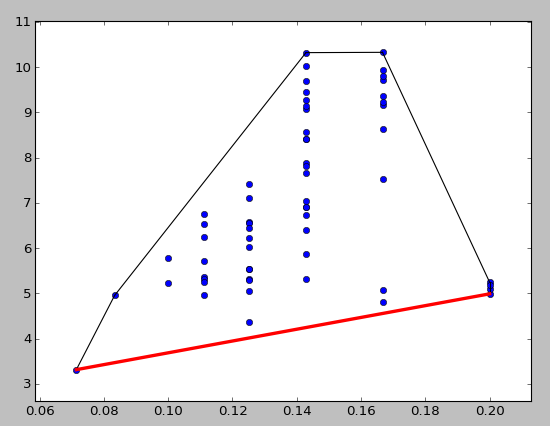
 

图 6 训练过程，自动标记波峰-波谷区间 图 7 计算下确界（图中红线），即

**AxyzBF \_LPF测试结果：**

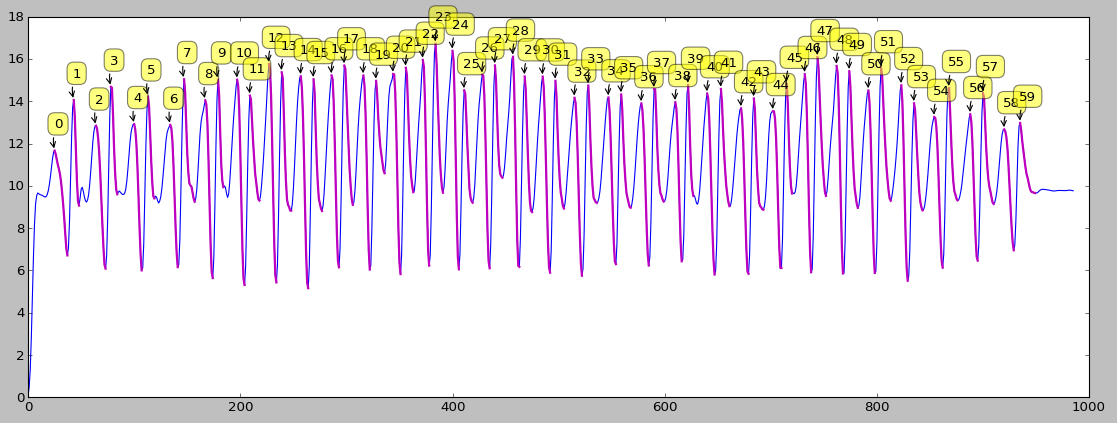


图 8 测试结果，不存在误计步

**结论**：AxyzBF \_LPF 训练以及测试效果优于 AxyzBF \_LPF。

**分析：**对于AzWF\_LPF数据，由于坐标转换误差，导致即使滤波之后，一步之中仍然存在多峰值，结果训练时误将一步的多峰值纳入样本点，测试时一步记成多步。由图4可以看到，坐标转换的误差较大，加速度波动范围存在较大变动。

### 单一文件训练&&对多文件测试

使用单文件求解α,β，将其用于其余所有文件（包括自己）的计步测试。下表1、表2分别为AzWF\_LPF、AxyzBF \_LPF 计步结果， 对角线红色字体为文件对自身的测试结果，即6.1小节进行的测试。

基础阈值设定 baseTh=2

**数据集**：只包含同一个采集人员，以正常&&较快步速行走的数据。实际步数为采样人员人工计数结果。

表 2 **AzWF\_LPF**计步结果

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 文件 | 实际步数 | 计算步数 | | | | | | | | | | | | | | 误差均值 |
| ZC30fast\_a9\_0.xml | 30 | **30** | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 32 | 60 | 60 | 60 | 60 | 66 | 1% |
| ZC30fast\_a9\_1.xml | 30 | 29 | **30** | 30 | 30 | 30 | 29 | 30 | 27 | 29 | 59 | 59 | 60 | 59 | 63 | 2% |
| ZC30fast\_a9\_2.xml | 30 | 28 | 29 | **30** | 30 | 30 | 29 | 28 | 26 | 25 | 58 | 59 | 58 | 58 | 60 | 4% |
| ZC30fast\_a9\_3.xml | 30 | 24 | 22 | 24 | **30** | 23 | 22 | 21 | 19 | 19 | 52 | 59 | 48 | 41 | 48 | 22% |
| ZC30fast\_a9\_4.xml | 30 | 27 | 27 | 27 | 30 | **30** | 29 | 28 | 25 | 23 | 57 | 59 | 58 | 55 | 56 | 8% |
| ZC30fast\_a9\_5.xml | 30 | 29 | 30 | 30 | 30 | 30 | **30** | 30 | 28 | 31 | 60 | 59 | 60 | 59 | 63 | 2% |
| ZC30fast\_a9\_6.xml | 30 | 31 | 30 | 31 | 30 | 31 | 31 | **31** | 32 | 33 | 61 | 62 | 61 | 61 | 65 | 4% |
| ZC30fast\_a9\_7.xml | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | **30** | 31 | 60 | 59 | 60 | 60 | 65 | 1% |
| ZC30fast\_a9\_8.xml | 30 | 29 | 30 | 30 | 30 | 30 | 30 | 30 | 28 | **31** | 60 | 59 | 60 | 60 | 63 | 1% |
| ZC60fast\_a9\_0.xml | 60 | 33 | 31 | 32 | 31 | 32 | 35 | 33 | 34 | 36 | **63** | 64 | 63 | 63 | 68 | 9% |
| ZC60fast\_a9\_1.xml | 60 | 20 | 19 | 19 | 25 | 20 | 19 | 15 | 18 | 16 | 43 | **59** | 39 | 33 | 40 | 34% |
| ZC60fast\_a9\_2.xml | 60 | 29 | 29 | 30 | 30 | 30 | 29 | 29 | 27 | 29 | 59 | 59 | **60** | 59 | 63 | 3% |
| ZC60fast\_a9\_3.xml | 60 | 29 | 30 | 30 | 30 | 30 | 30 | 30 | 28 | 30 | 59 | 59 | 60 | **60** | 66 | 2% |
| ZC60fast\_a9\_4.xml | 60 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 31 | 60 | 59 | 60 | 60 | **67** | 1% |

表 3 **AxyzBF \_LPF** 计步结果

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 文件 | 实际步数 | 计算步数 | | | | | | | | | | | | | | 误差均值 |
| ZC30fast\_a9\_0.xml | 30 | **30** | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 60 | 60 | 60 | 60 | 60 | 0% |
| ZC30fast\_a9\_1.xml | 30 | 29 | **30** | 29 | 30 | 30 | 29 | 28 | 30 | 28 | 59 | 59 | 59 | 58 | 59 | 2% |
| ZC30fast\_a9\_2.xml | 30 | 29 | 30 | **30** | 30 | 30 | 30 | 30 | 30 | 28 | 59 | 59 | 59 | 60 | 59 | 1% |
| ZC30fast\_a9\_3.xml | 30 | 29 | 30 | 30 | **30** | 30 | 30 | 29 | 30 | 29 | 59 | 59 | 59 | 60 | 60 | 1% |
| ZC30fast\_a9\_4.xml | 30 | 29 | 30 | 29 | 29 | **30** | 29 | 28 | 30 | 28 | 59 | 59 | 59 | 58 | 59 | 3% |
| ZC30fast\_a9\_5.xml | 30 | 31 | 30 | 31 | 30 | 30 | **31** | 31 | 31 | 32 | 60 | 60 | 61 | 61 | 61 | 2% |
| ZC30fast\_a9\_6.xml | 30 | 29 | 30 | 30 | 30 | 30 | 30 | **30** | 30 | 28 | 59 | 59 | 59 | 60 | 59 | 1% |
| ZC30fast\_a9\_7.xml | 30 | 28 | 28 | 22 | 30 | 30 | 29 | 25 | **30** | 28 | 58 | 60 | 50 | 57 | 57 | 7% |
| ZC30fast\_a9\_8.xml | 30 | 30 | 30 | 31 | 30 | 30 | 30 | 30 | 30 | **32** | 60 | 60 | 60 | 60 | 60 | 1% |
| ZC60fast\_a9\_0.xml | 60 | 29 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | **60** | 60 | 60 | 60 | 60 | 0% |
| ZC60fast\_a9\_1.xml | 60 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 29 | 60 | **60** | 60 | 60 | 60 | 0% |
| ZC60fast\_a9\_2.xml | 60 | 29 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 30 | 59 | 60 | **60** | 60 | 60 | 0% |
| ZC60fast\_a9\_3.xml | 60 | 30 | 30 | 31 | 30 | 30 | 30 | 30 | 30 | 31 | 60 | 60 | 60 | **60** | 60 | 0% |
| ZC60fast\_a9\_4.xml | 60 | 29 | 30 | 30 | 30 | 30 | 30 | 29 | 30 | 29 | 59 | 59 | 59 | 60 | **60** | 1% |

**结论：**由误差均值可以看出AxyzBF \_LPF 的计步准确率确实优于 AzWF\_LPF。

### 所有数据文件集合训练并测试

6.1，6.2小节均为单一文件进行训练求解α,β，将所有文件获得的点集合并，统一计算参数，结果如下图：

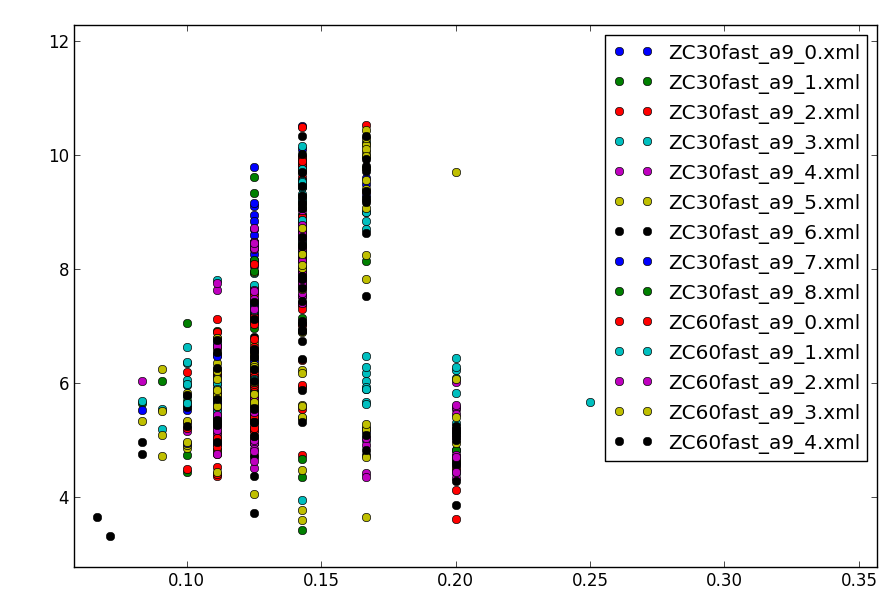
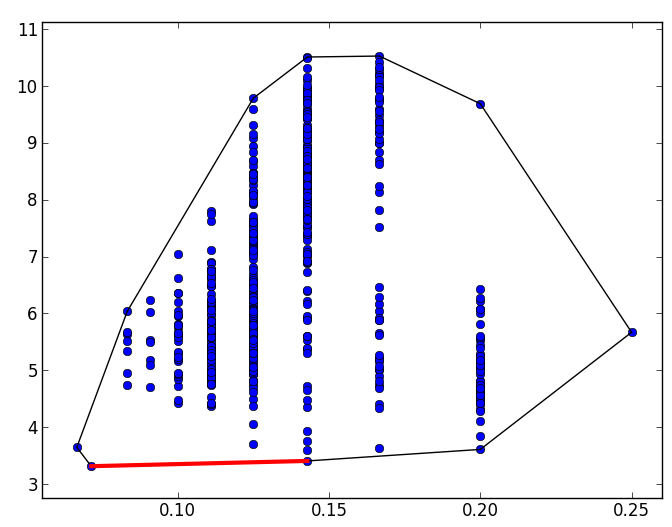
 

图 9 点集合并 图 10 计算合并后点集下确界

得到

进而计算个文件包含的步数，如表2所示。

表 4 误差统计

|  |  |  |  |
| --- | --- | --- | --- |
| 文件 | 实际步数 | 计算步数 | 误差 |
| ZC30fast\_a9\_0.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_1.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_2.xml | 30 | 31 | 3.23% |
| ZC30fast\_a9\_3.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_4.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_5.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_6.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_7.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_8.xml | 30 | 31 | 3.23% |
| ZC60fast\_a9\_0.xml | 60 | 60 | 0.00% |
| ZC60fast\_a9\_1.xml | 60 | 60 | 0.00% |
| ZC60fast\_a9\_2.xml | 60 | 60 | 0.00% |
| ZC60fast\_a9\_3.xml | 60 | 60 | 0.00% |
| ZC60fast\_a9\_4.xml | 60 | 60 | 0.00% |

**结论：**对单个采集人员，多组数据统一训练、测试准确度非常好，上表中3.23%误差，仅仅是因为计算步数与真实值相差一步所致。

### 加入不同用户、不同步行速度数据

如下表3所示，加入多人，不同速率数据之后，误差有一定增加。很明显，标有slow的数据文件计步误差非常大。原因在于，即便做低通滤波，缓慢踱步在同一步产生的多峰值也无法消除，如图11所示。

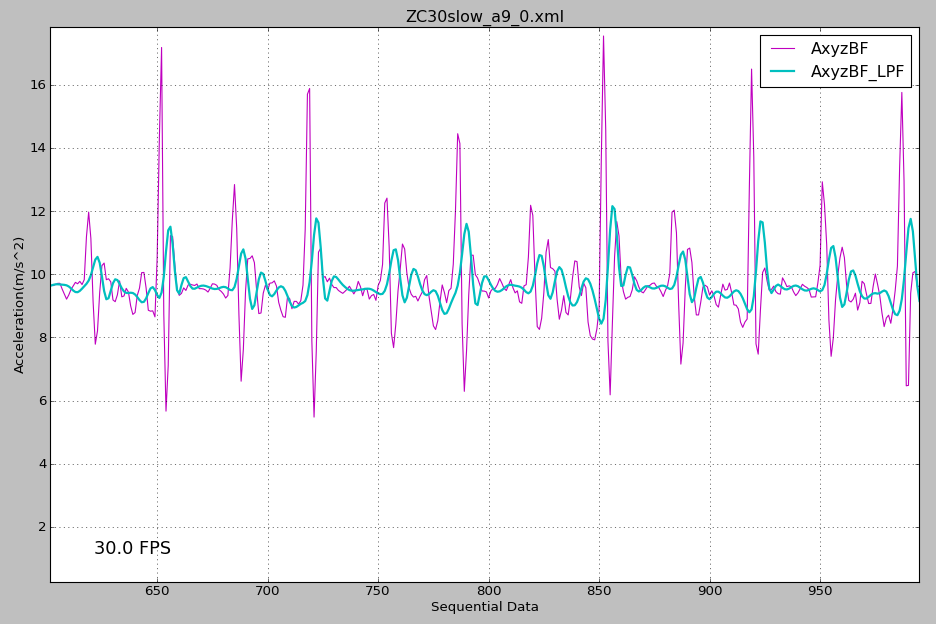


图 11 缓慢踱步滤波之后仍存在多峰值

表 5 加入不同用户、不同步行速度数据的测试结果

|  |  |  |  |
| --- | --- | --- | --- |
| 文件 | 实际步数 | 计算步数 | 误差 |
| HY60fast\_a9\_0.xml | 60 | 60 | 0.00% |
| HY60fast\_a9\_1.xml | 60 | 60 | 0.00% |
| HY60slow\_a9\_0.xml | 60 | 60 | 0.00% |
| HY60slow\_a9\_1.xml | 60 | 60 | 0.00% |
| LH30fast\_a9\_0.xml | 30 | 31 | 3.33% |
| LH30fast\_a9\_1.xml | 30 | 31 | 3.33% |
| LH30fast\_a9\_2.xml | 30 | 31 | 3.33% |
| LH30fast\_a9\_3.xml | 30 | 35 | 16.67% |
| ZC30fast\_a9\_0.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_1.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_2.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_3.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_4.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_5.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_6.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_7.xml | 30 | 30 | 0.00% |
| ZC30fast\_a9\_8.xml | 30 | 31 | 3.33% |
| ZC30slow\_a9\_0.xml | 30 | 34 | 13.33% |
| ZC40slow\_a9\_1.xml | 40 | 46 | 15.00% |
| ZC60fast\_a9\_0.xml | 60 | 60 | 0.00% |
| ZC60fast\_a9\_1.xml | 60 | 60 | 0.00% |
| ZC60fast\_a9\_2.xml | 60 | 60 | 0.00% |
| ZC60fast\_a9\_3.xml | 60 | 60 | 0.00% |
| ZC60fast\_a9\_4.xml | 60 | 60 | 0.00% |

**结论：**对不同用户，或者同一用户缓慢行走的数据，计步算法不是十分精确。原因在于不同采样人员的行走习惯可能各不相同，统一训练得到的参数不能通用。

## 误差分析

不同用户行走，或同一用户以差别较大的速度行走，产生的（步幅，步频）分布是不同的，计算得到的阈值-步频关系参数也是不同的，类似闭集识别。训练好的参数，拿到新的数据集进行测试，可能得到的结果并不理想。

## 参考资料

[1] Naqvi, N. Z., Kumar, A., Chauhan, A. and Sahni, K.: Step Counting Using Smartphone-Based Accelerometer. International Journal on Computer Science and Engineering (IJCSE), (May 2012) 1-7

[2] 张琛 《基于智能手机加速计的计步器设计》 V1.3

[3] 杜宇《时序数据文件格式》详述

1. [1] 张琛 《基于智能手机加速计的计步器设计》 V1.3 [↑](#footnote-ref-1)
2. [2] 杜宇《时序数据文件格式》详述 [↑](#footnote-ref-2)
3. [1] 张琛 《基于智能手机加速计的计步器设计》 V1.3， part 4.4.3 - d) [↑](#footnote-ref-3)