

分类号 TP393

学号 13063149

U D C

密级 公 开

工程硕士学位论文

基于智能手机传感器的用户  
情绪识别技术研究

硕士生姓名 陈茜

工 程 领 域 软件工程

研 究 方 向 移动感知

指 导 教 师 史殿习 研究员

国防科学技术大学研究生院

二〇一五年十一月

基于智能手机传感器的用户情绪识别技术研究

国防科学技术大学研究生院

# **User Emotion Recognition Based on Smartphone Sensors**

**Candidate: Chen Xi**

**Advisor: Prof. Shi Dianxi**

**A thesis**

**Submitted in partial fulfillment of the requirements  
for the professional degree of Master of Engineering  
in Software Engineering  
Graduate School of National University of Defense Technology  
Changsha, Hunan, P.R.China  
(November, 2015)**



## 独 创 性 声 明

本人声明所呈交的学位论文是我本人在导师指导下进行的研究工作及取得的成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表和撰写过的研究成果，也不包含为获得国防科学技术大学或其它教育机构的学位或证书而使用过的材料。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示谢意。

学位论文题目： 基于智能手机传感器的用户情绪识别技术研究

学位论文作者签名： 陈浩 日期： 2015 年 11 月 2 日

## 学位论文版权使用授权书

本人完全了解国防科学技术大学有关保留、使用学位论文的规定。本人授权国防科学技术大学可以保留并向国家有关部门或机构送交论文的复印件和电子文档，允许论文被查阅和借阅；可以将学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文。

(保密学位论文在解密后适用本授权书。)

学位论文题目： 基于智能手机传感器的用户情绪识别技术研究

学位论文作者签名： 陈浩 日期： 2015 年 11 月 2 日

作者指导教师签名： 史殿习 日期： 2015 年 11 月 2 日

## 目 录

摘 要 .....	i
ABSTRACT .....	ii
第一章 绪论 .....	1
1.1 研究背景和意义 .....	1
1.2 研究现状 .....	3
1.2.1 基于面部表情的情绪识别研究现状 .....	3
1.2.2 基于语音语调的情绪识别研究现状 .....	4
1.2.3 基于生理信号的情绪识别研究现状 .....	5
1.2.4 基于人体姿态的情绪识别研究现状 .....	5
1.2.5 基于智能手机的情绪识别研究现状 .....	6
1.3 研究内容 .....	6
1.4 论文结构 .....	8
第二章 相关技术研究 .....	9
2.1 情绪分类模型 .....	9
2.1.1 基本情绪模型 .....	9
2.1.2 维度空间模型 .....	10
2.1.3 分析比较 .....	11
2.2 数据处理过程 .....	11
2.2.1 数据预处理 .....	12
2.2.2 特征提取 .....	13
2.3 分类算法 .....	14
2.3.1 支持向量机(SVM) .....	15
2.3.2 k-近邻(kNN)算法 .....	16
2.3.3 决策树 .....	16
2.3.4 随机森林 .....	17
2.3.5 AdaBoost .....	18
2.3.6 梯度树提升(GTB) .....	19
2.4 本章小结 .....	19
第三章 数据收集和存储机制研究 .....	20
3.1 预调查 — 网络问卷调查 .....	20
3.1.1 问卷调查内容的设计 .....	20

3.1.2 问卷调查的发放和回收 .....	21
3.1.3 问卷调查结果分析 .....	21
3.2 数据收集 .....	22
3.2.1 手机传感器数据收集 .....	23
3.2.2 情绪数据收集 .....	25
3.3 数据存储机制 .....	26
3.3.1 数据存储格式 .....	26
3.3.2 基于 JSON 格式的数据实现 .....	27
3.4 本章小结 .....	28
<b>第四章 多维移动数据的处理机制 .....</b>	<b>29</b>
4.1 数据的预处理 .....	29
4.1.1 JSON 解析 .....	29
4.1.2 离散数据的完整性和一致性处理 .....	31
4.1.3 连续数据的离散化处理 .....	31
4.2 运动类感知数据的特征提取 .....	32
4.2.1 加速度计、陀螺仪和磁力计的特征提取 .....	32
4.2.2 活动状态的特征提取 .....	32
4.3 环境类感知数据的特征提取 .....	32
4.3.1 光传感器的特征提取 .....	32
4.3.2 GPS 的特征提取 .....	33
4.4 手机使用数据的特征提取 .....	35
4.4.1 社交信息的特征提取 .....	35
4.4.2 手机使用信息的特征提取 .....	36
4.4.3 手机状态信息的特征提取 .....	37
4.5 数据特征归纳 .....	37
4.6 数据归一化 .....	37
4.7 本章小结 .....	38
<b>第五章 多维移动数据的用户情绪识别 .....</b>	<b>39</b>
5.1 数据集、情绪识别方法和评估基准 .....	39
5.1.1 数据集 .....	39
5.1.2 情绪识别方法和评估基准 .....	41
5.2 基于特征融合的情绪识别 .....	42
5.2.1 特征级融合 .....	42
5.2.2 实验结果与分析 .....	44

---

---

5.3 基于结果融合的情绪识别.....	47
5.3.1 结果级融合 .....	47
5.3.2 实验结果与分析.....	49
5.4 相关工作比较 .....	53
5.5 本章小结.....	54
第六章 总结与展望.....	55
致 谢 .....	57
参考文献.....	59
作者在学期间取得的学术成果 .....	63

## 表 目 录

表 2.1 两种情绪分类模型的区别 .....	11
表 2.2 常用时域特征计算公式 .....	14
表 3.1 问卷调查：情绪变化对行为的影响 .....	20
表 3.2 问卷调查选择题统计结果 .....	22
表 3.3 传感器数据收集内容 .....	24
表 3.4 数据存储格式比较 .....	26
表 3.5 JSON 格式举例 .....	27
表 3.6 传感器数据 JSON 存储格式 .....	27
表 4.1 WiFi 数据 JSON 解析核心代码 .....	30
表 4.2 传感器数据提取特征表 .....	37
表 5.1 不同志愿者同一天数据容量统计 .....	40
表 5.2 愉悦度和活跃度相关性分析 .....	42
表 5.3 多传感器数据的特征融合算法描述 .....	43
表 5.4 多传感器数据的结果融合算法描述 .....	48
表 5.5 运动类感知数据分类概率（基本情绪模型） .....	49
表 5.6 环境类感知数据分类概率（基本情绪模型） .....	49
表 5.7 手机使用数据分类概率（基本情绪模型） .....	50
表 5.8 结果融合后分类概率（基本情绪模型） .....	50
表 5.9 运动类感知数据分类概率（环状情绪模型） .....	50
表 5.10 环境类感知数据分类概率（环状情绪模型） .....	50
表 5.11 手机使用数据分类概率（环状情绪模型） .....	50
表 5.12 结果融合后分类概率（环状情绪模型） .....	50



## 图 目 录

图 1.1 三星 GALAXY 系列传感器演化图 .....	2
图 1.2 情绪识别系统框架流程图 .....	8
图 2.1 不同学者对基本情绪的分类定义 <sup>[49]</sup> .....	9
图 2.2 基本情绪聚类图 .....	10
图 2.3 环状情感模型 .....	10
图 2.4 数据预处理方法 .....	12
图 2.5 滑动窗口数据分割 .....	13
图 2.6 分类算法流程图 .....	15
图 2.7 SVM 最大间隔超平面 .....	15
图 2.8 kNN 算法决策过程 .....	16
图 2.9 决策树模型 .....	17
图 2.10 AdaBoost 算法示意图 .....	18
图 3.1 数据收集框架主界面 .....	23
图 3.2 StarLog 通知栏消息 .....	23
图 3.3 StarLog 文件存储形式 .....	24
图 3.4 情绪记录软件 .....	25
图 3.5 对象的格式设置 .....	27
图 4.1 WiFi 数据 JSON 解析 .....	30
图 4.2 光传感器数据图 .....	33
图 4.3 GPS 原始数据图 .....	34
图 4.4 GPS 数据分类图 .....	34
图 5.1 手机手机数据与情绪记录数据映射图 .....	39
图 5.2 某志愿者数据容量分布 .....	40
图 5.3 特征融合过程 .....	43
图 5.4 特征融合部分结果 .....	43
图 5.5 混合数据特征融合情绪识别准确率 .....	44
图 5.6 志愿者三号个人数据特征融合情绪识别准确率 .....	45
图 5.7 志愿者十二号个人数据特征融合情绪识别准确率 .....	45
图 5.8 混合数据和个人数据特征融合的基本情绪模型对比 .....	46
图 5.9 混合数据和个人数据特征融合的环状情绪模型对比 .....	46
图 5.10 结果融合过程 .....	48
图 5.11 混合数据情绪结果融合识别准确率 .....	51

---

---

图 5.12 志愿者三号个人数据结果融合情绪识别准确率 .....	51
图 5.13 志愿者十二号个人数据结果融合情绪识别准确率 .....	52
图 5.14 混合数据和个人数据结果融合的基本情绪模型对比 .....	52
图 5.15 混合数据和个人数据结果融合的环状情绪模型对比 .....	53

## 摘 要

目前,智能手机已经成为一个集计算、存储与感知于一体的移动计算平台,成为人们日常生活当中不可或缺的交流 and 感知设备。通过智能手机可以感知应用使用模式以及人们的各种日常行为,可以为人们提供各种各样的个性化服务。最近的研究表明,人们的情绪与日常的行为和智能手机的使用模式密切相关。使用智能手机进行情绪识别的优点在于,在不影响用户日常使用情况下,可以实时推测人们的情绪,从而更好地了解人们的精神状况,能够促进心理健康和人们之间更好地交流。因此,基于智能手机进行情绪识别具有非常重要的理论意义和实用价值。

本文针对目前基于智能手机的情绪识别研究中所用数据较为单一,并不能全面的反应用户行为模式,进而并不能真实反应用户情绪这一问题展开研究,从多个维度全面收集反应用户日常行为的细粒度感知数据,采用多维数据特征融合方法和结果融合方法,利用多种分类方法和情绪分类模型进行情绪识别。本文的工作主要体现在以下三个方面:

(1) 针对如何采集和采集哪些与情绪密切相关的感知数据这一问题,首先设计了一个网络调查问卷,对情绪如何影响人的行为模式进行预调查,并对调查结果加以分析,作为选择手机数据采集种类的依据;进而设计实现了一个实时手机数据采集工具,收集与用户情绪密切相关的各种感知数据,与此同时设计开发了一个交互式日常情绪记录工具,以便及时记录和采集用户情绪情况。

(2) 针对通过多种传感器收集的感知数据具有复杂度高和种类多样性的问题,本文采用不同的预处理方法和特征提取方法,并对数据进行归一化处理,以求最大程度提取传感器数据中能够体现出用户行为模式的多维特征,从而为基于多维的数据特征融合和结果融合方法进行情绪识别奠定基础。

(3) 采用多维数据特征融合和结果融合两种方法,利用支持向量机(Support Vector Machine, SVM)、随机森林(Random Forest)等六种分类方法,基于离散情绪模型和环状情绪模型两种情绪分类模型,对 14 名志愿者的混合数据和个人数据分别进行情绪识别,并进行了对比实验。实验结果表明,本文提出的全面反应用户行为的多维数据特征融合方法,能够更好地对用户的情绪进行识别,其中使用个人数据进行情绪识别的准确率最高可达到 79.78%,而且环状情感模型分类结果明显优于离散分类模型。

**关键词:** 智能手机; 情绪识别; 分类算法; 特征融合; 结果融合

## ABSTRACT

Currently, the smartphone has become a mobile computing platform of computing, storage and sensing. It becomes an essential equipment of communication and sensing in people's daily life. We can provide people with a wide range of personalized services through the smartphone's application usage patterns and people's daily behaviors. And recent studies show that people's emotions are closely related to daily behavior and smartphone usage patterns. The advantages of using smartphones emotion recognition is that people's emotions can be speculated in real time without affecting their daily use to better understand people's mental state and to promote better communication between each other. Therefore, emotion recognition based on smartphones has very important theoretical and practical value.

In this paper, we study about the problem of how to recognize the user emotion based on smartphone data more really. With single data used in the previous studies, it cannot make a comprehensive response of user behavior patterns. So we collected fine-grained sensing data which could reflect user daily behavior fully from multiple dimensions based on smartphone, and then used multidimensional data feature fusion method and result fusion method, using a variety of classification methods and emotion classification model to recognize emotion. Work in this paper is mainly reflected in the following three aspects:

(1) The first problem is how to collect data and choosing which kind of sensing data closely related to emotions to collect. Firstly, we designed a network questionnaire on how emotions affect people's behavior patterns as the pre-survey, and analyze the survey results as the basis for selecting the type of phone data. Then, we design and implement a real-time data collection tools for collect all kinds of sensing data which are closely related with the user's emotion, at the same time developed an interactive tool to record daily emotion for users timely.

(2) With the issues of mobile data scale expansion and type complicated for multi-sensor data collection, we use different preprocess methods and feature extraction methods to different types of data, and make data normalization at last in order to maximize the extraction of sensor data's features which can reflect the user behavior patterns. In this part, we lay the foundation for emotion recognition with the methods of multi-dimensional data fusion and result fusion .

(3) Finally, we launched a study of emotion recognition based on smartphone with the use of feature fusion method and result fusion method. In this experiment, we used six classification methods such as Support Vector Machine and Random Forest and two emotion classification models. Then we carried out contrast

experiment with 14 volunteers' hybrid data and personal data respectively to recognized user emotion based on discrete emotion model and circumplex emotion model. The results showed that the multidimensional data feature fusion method we mentioned which could better recognize user emotion, among them personal data training the accuracy rate can reach 79.78%, and circumplex model of affect is better than discrete emotion model.

Key Words: Smartphone, Emotion Recognition, Classification Algorithm, Feature Fusion, Result Fusion

## 第一章 绪论

近些年来, 智能手机发展迅速。诺基亚的塞班系统(Symbian OS)从独占鳌头到退出市场成为历史, 谷歌的安卓系统(Android OS)以及苹果的 iOS 从崭露头角到占有市场最大份额, 智能手机操作平台的更新换代似乎从未停止。与其操作系统发展相同步, 智能手机的功能也日益丰富。随着手机内嵌传感器的增多, 智能手机已经不仅只是作为简单的通信工具存在, 而是逐渐成为集通信、办公、社交、娱乐、健康等多种功能为一体的综合性个人手持终端设备。如今, 智能手机与人们的日常生活密不可分, 它在给人们提供便利生活的同时也忠实的记录着人们的生活轨迹, 成为了一种新式可穿戴计算<sup>[1]</sup>(Wearable Computing)的载体。随着我国经济水平的持续增长与医疗卫生条件的不断提高, 人们对健康的认知从只知生理健康逐步提高到生理健康与心理健康并重。而人的情绪变化作为衡量心理健康的一个重要判断因素也逐渐被人们关注, 因此利用手机传感器数据对用户情绪进行识别已成为学术界一个新的研究热点。

### 1.1 研究背景和意义

人类在对外界事物进行探索和认知的过程中, 会产生诸如喜悦、悲伤、愤怒、恐惧等主观情感。我们把人对客观事物的态度体验以及相对应的行为反应, 称之为情绪(Emotion)<sup>[2]</sup>。它作为一种不同于认识和意识的心理形式, 不同程度上影响着人的学习、工作效率以及行为模式, 在日常生活中扮演着重要的角色。情绪主要表现为一种瞬间的、自发的生理及心理过程, 常常不受主观意识的控制, 但和生理变化紧密联系。通常情况下, 情绪的变化受外界影响, 常常伴随着表情、声音、行为的短暂变化。也正是因为这些变化, 才使得情绪能够通过图像和语音等进行识别。

情绪识别并不是指计算机能够直接测量或识别出用户的情绪状态, 而应该解释为“通过观察表情、行为和情感产生的前提环境来推断情感状态”<sup>[3]</sup>。因为情绪通常是经由一些外在因素刺激而产生的主观体验(如喜、怒、哀、惧等情感), 并伴有外部表现的变化(如面部表情、身体行为和声音语调等)以及生理反应的变化(如心率的节奏、皮下的特定活动等), 因此我们可以通过摄像机、麦克风等电子设备以及加速度等相关传感器获得有关情绪状态的一些观测值。假设这些数据的观测值有效可靠, 那么就可以根据这些数据把潜在的情绪状态推测出来。随着智能手机的普及和发展, 诸如照相机、麦克风、加速度等传感器早已内嵌至其中, 这样不仅方便了用户的使用, 也为我们进行情绪识别相关研究提供了数据。

智能手机为了实现丰富的功能和增强用户体验而增加了多种传感器,如图 1.1 所示,该图展示了三星旗下 GALAXY 系列手机传感器的演化过程。可以看到该系列手机除了拥有现今智能手机必备的加速度计(Accelerometer)和磁力计(Magnetometer)等惯性传感器,还持续增加了包括光传感器(Ambient light)、温度传感器(Temperature)、湿度传感器(Humidity)、压力传感器(Pressure)等在内的用于感知周围环境情况的传感器以及测量心率(Heart Rate)和识别指纹(Fingerprint)的生物识别传感器,这使得手机能够探测和感知多种外界信号。随着智能手机内嵌传感器的不断增加,通过手机传感器收集人们的日常行为已经变得触手可及。智能手机现在已经集通信、存储、计算等多种能力为一体,发展成为了新一代感知平台。

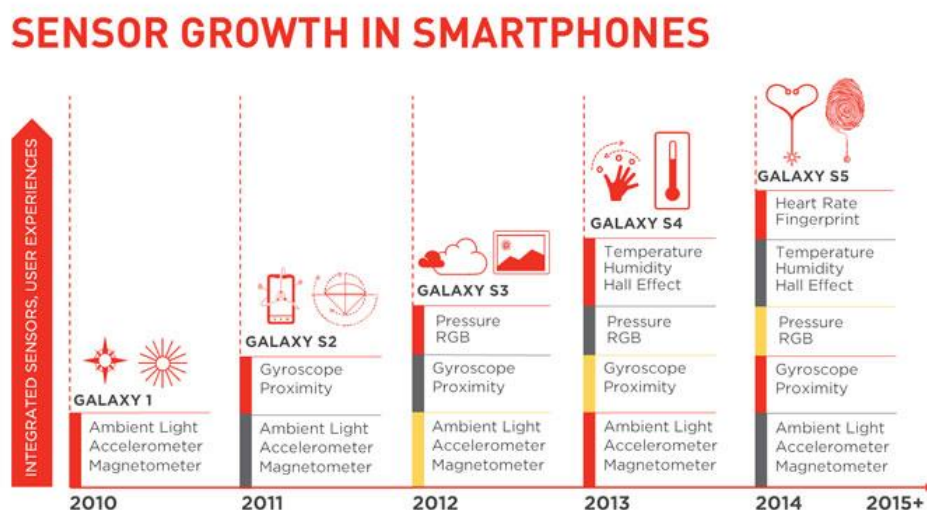


图 1.1 三星 GALAXY 系列传感器演化图

采用智能手机作为传感器数据收集的平台,优势如下:首先,智能手机的用户众多,如此大规模的潜在实验数据收集对象是以前的相关研究未能达到的,这为大规模感知提供了可能。其次,随身携带手机已经逐渐成为一种生活习惯,并且基于智能手机的感知是非干扰式的,用户的参与度低,相较传统的捆绑式传感器基本不会对用户造成困扰,这也为持续感知提供了前提。最后,随着手机内嵌传感器的增多和自身运算、存储能力的提高,捕捉使用者的身体行为变化也变得更加容易,这比以前的其他平台更具优势。

正是因为使用智能手机作为感知平台具有上述优点,因此基于智能手机的研究和应用近些年来大量涌现,涉及包括人体健康监测、交通路况监测、电子商务、环境监控(天气情况、环境污染等)、人类行为监测以及社交活动等多种领域<sup>[4]</sup>。例如 EyePhone<sup>[5]</sup>系统通过手机的前置摄像头对用户的眼球位置进行捕捉,以眨眼模拟鼠标点击过程开启应用程序; LiveCampare<sup>[6]</sup>系统通过手机

的后置摄像头将商品的二维码进行扫描然后返回给用户该商品在周围商场的价格；TrafficSense<sup>[7]</sup>系统通过手机的加速度计、麦克风、GMS 信号和 GPS 传感器监测路面情况（如地面凹陷等）和交通情况（如急刹车、追尾、堵车等）；EmotionSense<sup>[8]</sup>系统通过麦克风和距离传感器等感知用户的情绪和社交活动。

移动群体感知技术领域<sup>[9][10]</sup>中的研究涉及方方面面，而本文的研究重点在于情绪识别技术。情绪识别作为人工智能和普适计算的重要组成部分，受到了业内人士的高度重视，已经成为当下的研究热点，在人机交互、远程教育、医疗保健、心理治疗等多个领域均有广阔的应用前景<sup>[11]</sup>。随着智能手机感知平台的兴起，利用智能手机对用户情绪进行识别成为热点。在这样的研究背景之下，本课题将移动群体感知技术与情绪识别相结合，通过对用户的行为模式相关移动数据的分析，分别建立了离散情绪模型和环状情绪模型，力求充分利用采集到的各种传感器文本数据对用户的情绪进行较为准确的推测与识别，为情绪识别技术提供新的研究方向。

## 1.2 研究现状

自从 1997 年麻省理工学院(MIT)媒体实验室的 Picard 教授提出“情感计算” (Affective Computing)<sup>[12]</sup>这个概念之后，越来越多的学者对情绪识别展开研究。因为一般情况下情绪的产生都会伴随着身体的某些变化，诸如面部表情、身体动作、肌肉紧张、心率加快等，所以情绪的基本识别方法即是利用这些信号的变化直接进行情绪识别。目前情绪识别的主要方法包括面部表情识别、语音语调识别、生理信号识别和人体姿态识别<sup>[13]</sup>，分别通过情绪产生过程中的表情变化、声音变化、生理信号变化和姿势动作变化来推断所处的情绪状态。下面将对这些基本识别方法的研究现状分别进行分析和描述。

### 1.2.1 基于面部表情的情绪识别研究现状

面部表情是指通过眼部肌肉、脸颊肌肉和口唇肌肉的变化所表现出的各种情绪<sup>[14]</sup>，由于面部表情能精细的表达出各种不同类型的情绪，因此被作为识别情绪的主要标志之一。面部的各个器官都能传递情绪，尤其是眼睛、眉毛和嘴巴。不同的眼神可以表达出不同的情感，瞳孔的扩大或缩小反映出人的心理变化；眉间肌肉皱纹的舒展或紧缩蕴含着不同的意味；嘴部表情则体现在口型的变化上，高兴时嘴角的上扬、伤心时的下撇、紧张或痛苦时的咬紧下唇都代表着不同的情绪。虽然某些情绪的表现方式由单个面部感官完成，但是更多的情绪需要多个感官相互配合，并且由于各个部分对识别的贡献率并不相同，因此更加大了面部表情情绪识别的难度。



关于面部表情的研究自 19 世纪达尔文在《人类和动物的表情》一书中的阐述而拉开序幕，之后 1971 年美国的心理学家 Ekman 等人<sup>[15]</sup>开发了面部运动编码系统(Facial Action Coding System, FACS)，系统性的建立了上千幅不同的人脸表情图像库，分别对高兴、悲伤、愤怒、厌恶、惊讶和恐惧这六种面部表情进行研究。随后，研究人员又在 FACS 的基础上相继采用 MPEG-4 脸部参数运动法<sup>[16][17][18]</sup>、主分量分析法<sup>[19][20][21]</sup>等多种算法进行识别。这些识别算法可以分为基于静态图像和基于动态图像两类，其中基于静态图像的方法使用单一面部表情图像，只考虑图像的图像的空间信息而没有考虑时间信息，计算相对简单但识别准确率普遍不高；基于动态图像的方法则采用图像序列，虽然充分考虑了图像的时间与空间信息，准确率较高，但是计算量大成为其发展的需要克服的一个因素。

目前基于面部表情的情绪识别使用的表情数据库与实际应用场合有很大差别，受约束较高，例如夸张化的面部表情、图像背景单一、头发全部扎起或没有饰物干扰等，这样的算法运用到实际生活场合中识别率必然受到影响。并且在两种主流的识别算法中，或识别率低，或计算复杂度高，这些问题都亟待解决。

### 1.2.2 基于语音语调的情绪识别研究现状

语言是人类相互交流的重要工具之一，我们可以通过语音的变化来识别不同的情绪。人们在打电话的时候虽然并不能看到对方的表情与动作，但却可以通过言语的交流明显感知到对方的情绪状态。这正是因为语言的表达不仅包括表明情绪的词语和语句（如：我很开心），还包括说话时能反映出说话人情绪状态的语音语调（如：用高兴的语气说“今天天气真好！”）。因此，在研究语音语调时，我们要特别注重交谈时的语音、语调、语速等声调特征。

语音语调作为人类表达自己情感的一种重要方式，早在上个世纪八十年代就开始了系统性的研究<sup>[22]</sup>。随着 1985 年人工智能创始人之一的 Minsky 教授关于“让计算机具有情感能力”这一观点的提出，越来越多的研究者们认识到了情感智能的重要性并开始了语音情感识别的探索。目前，基于语音进行情绪识别最重要的声学特征包括音频 F0(Fundamental Frequency, F0)、能量、说话速率、流利程度和音质特征等<sup>[23][24][25]</sup>，不断地提取新的声学特征是语音情绪识别的发展方向。因为语音信号个体差异性较明显，因此样本越大，受个体和背景影响的因素越小，而庞大语音数据库的建立又给识别增加了新的难度。现有数据库中的语音片段主要有三种方式获取，专业演员朗读材料获取语音样本<sup>[26]</sup>，收集自然状态下的语音样本<sup>[27]</sup>和诱发情绪取得语音样本<sup>[28]</sup>。这几种方法获取的

语音样本各有优劣，第一种方法操作简单，虽然可以保证语音材料内容的一致性，但和现实情绪表达有一定差距；第二种方法收集的语音材料最接近现实情绪并且样本中包含情绪类型较少，但背景音复杂，处理难度相对较大；最后一种方法得到的诱发情绪强度往往比较低，并且不同被测试者对同一诱发产生的情绪可能由于认知评估等因素的影响并不相同。

目前基于语音语调的情绪识别使用的语音数据库大多采用演员表演录制的语音材料，这种方法相对来说对其他刺激因素产生的影响更容易控制，但建立全面的语音数据库与增强对声学特征的分析仍是语音情绪识别的重点研究方向。

### 1.2.3 基于生理信号的情绪识别研究现状

当人的情绪发生变化时，身体会不由自主的产生一系列生理变化。相较于面部表情、语音语调、人体姿态等这类外显且易被掩饰的特征，生理信号的变化由神经和内分泌系统控制与支配，不易受主观意识控制，因而使用生理信号进行情绪识别相对来说更为客观<sup>[29]</sup>。

学术界对于情绪状态是否能在生理信号上有所体现一直存在争议，但自从 Picard 教授带领 MIT 媒体实验小组证明了“应用生理信号可以对情绪进行识别”<sup>[30]</sup>后，国际上陆陆续续又展开了关于生理信号情绪识别的新一轮研究。德国 Augsburg 大学的 Johannes Eagner 等人<sup>[31]</sup>、韩国 Kim Bang 等人<sup>[32]</sup>都在基于生理信号的情绪识别研究中取得了很好的效果。目前，国内外用于情绪识别的生理信号主要有心电信号 (Electrocardiography, ECG)、肌电信号 (Electromyography, EMG)、呼吸信号 (Respiration, RSP)、皮电信号 (Skin Conductivity, SC) 等<sup>[33][34][35][36]</sup>。

虽然生理信号携带的情绪状态更为客观和准确，但存在着信号强度微弱以及干扰源多、噪声大等各种问题，因此在信号提取和去除噪声过程中要求较高。并且由于实验条件较为复杂，因此关于生理信号的研究大多局限于实验室内部，还没有足够的实际应用条件。

### 1.2.4 基于人体姿态的情绪识别研究现状

实际上，人体的动作与姿势不仅包含着丰富的情绪信息，更能传达出动作意图。譬如说，当一个人感到恐惧时，不仅身体会摆出恐惧的姿势，更能透露出他是否会选择逃跑这种行为。因此，人体姿态可以看作是不同情绪状态下的身体表情，并呈现出一种预见性。

对面部表情和语音语调进行的情绪识别由来已久，但基于人体姿态的研究

则远远滞后于前者,原因在于后者大多是以身体动作与姿态作为研究对象,而学术界对于不同的情绪状态是否表现为特定的动作尚不确定。但研究发现身体动作的大小反映了情绪的强烈程度并且特定的情绪能够引起肌肉的紧张,从而改变动作和姿态的呈现方式<sup>[37][38]</sup>。

目前,对于人体姿态的情绪识别主要以对各种情绪状态下的典型样例进行提取并建立特征库为主<sup>[39][40]</sup>,并以动作持续时间的长度、频率等数据作为评判依据<sup>[41]</sup>。但由于很多动作或姿势并不具备明显的情绪特征,因而在识别过程中并不能全面的分辨出结果,实验难度相对更大,准确率也更低。

### 1.2.5 基于智能手机的情绪识别研究现状

随着手机内嵌传感器的不断增加,智能手机如今已可以为研究人员提供丰富的用户数据和相关信号,利用手机中的照相机、麦克风、加速度等传感器可以轻易捕捉到用户的图像、声音以及心率等信号,因此基于智能手机的情绪识别开始逐渐兴起。

LiKamWa 等人建立了心情传感器 MoodScope<sup>[42][43]</sup>,他们将用户的通话记录、短信记录、邮件、上网历史、应用程序使用情况以及位置的历史数据进行收集,采用多元线性回归的方法用来进行心情的推测。Bogomolov 等人<sup>[44]</sup>也利用手机作为数据采集终端来对用户的幸福感进行识别,以通话记录、短信记录和蓝牙信息作为采集数据辅佐以天气因素和人格测试,采用随机森林的方法进行训练。EmotionSense<sup>[8]</sup>系统通过麦克风、距离传感器、加速度计等通过声音和对用户活动状态的判断感知用户的情绪以及社交活动。

而以上的相关研究也正是本课题研究的出发点,通过智能手机的传感器可以真实的记录用户的生活日志,从而识别出每个人的行为模式并建立不同行为模式与情绪之间的映射关系,完成情绪识别的目标。基于智能手机的用户情绪识别相较于其他四种情绪识别方法可操作性高、难度低,更易于普通受众参与到实验中来,可以更加方便快捷的对用户进行实时的情绪识别。

## 1.3 研究内容

在对人工智能领域中基于面部表情、语音语调、生理信号、人体姿态这四种传统的情绪识别方法进行深入分析和了解后上节又对新兴的基于智能手机的方法进行了介绍,同时阐明了本文使用智能手机进行情绪识别的优势以及意义。本课题力求设计一个以智能手机作为感知平台收集用户行为模式数据并最终可以完成情绪识别的系统框架,在课题研究过程中,为了达到预期目标,从以下三个方面展开了研究:

---

### (1) 传感器数据的收集和存储机制研究

为了使用传感器数据进行用户的情绪识别研究，本课题提出了一个针对多维移动数据进行处理与分析的系统框架，该系统由手机客户端和服务端组成。由于移动设备的资源配置较低，所以手机只负责数据的采集与传输功能，由服务端负责处理和分析数据等功能的实现。本文在数据收集前使用了网络问卷调查的形式以“情绪如何影响人的行为模式”为主题进行预调查，并对调查结果进行相关统计作为选择哪些类型的手机数据进行收集的依据。然后，采用一个基于 Android 平台的数据收集工具实时采集可以表示用户行为模式的数据进行记录，同时开发了一个情绪记录软件供用户每天记录情绪。在移动端所收集的全部数据将采用 JSON 格式进行存储并上传至服务端进行后续处理。

### (2) 多维移动数据的处理机制研究

多传感器数据的收集带来了移动数据规模的扩大化与数据类型复杂化等问题，增加了数据的处理难度，因此在这种情况下，本文针对不同类型的数据采用了不同的预处理方法和特征提取方法。一般来说预处理的方法相对固定，但特征提取方法具有多样性，并且不同的特征方法所提取的结果对最终的识别准确率影响很大，因此选择相对来说更合适的特征就显得格外重要。在各种传感器数据完成特征提取后，为了消除因为量纲和量纲单位不同引起的问题，不影响到数据分析的结果和程序运行时的收敛速度，还需要对数据进行归一化处理。

### (3) 基于多维移动数据的用户情绪识别研究

随着智能手机作为感知平台的优势日益明显，基于智能手机的情绪识别逐渐成为普适计算领域研究的热点问题，并取得了一系列研究成果。但目前基于智能手机的情绪识别研究中所使用的传感器较为单一，大多数数据为通讯信息的记录以及位置信息，并不能全面的反应用户行为模式，因此本文使用智能手机从多个维度全面收集反应用户日常行为的细粒度感知数据，采用多维数据特征融合方法，利用支持向量机、k-近邻、决策树、随机森林、AdaBoost、梯度树提升六种分类器以及离散情绪模型和环状情绪模型两种分类方式，对 14 名志愿者连续一个月的混合数据和个人数据分别进行基于特征融合与基于结果融合的情绪识别和对比实验。

本文基于上述三个方面展开研究，首先通过网络问卷调查确定移动数据收集的种类并相继展开数据收集工作和数据传输工作，然后在系统服务端进行数据分析与处理功能的实现，最后将基于特征融合和结果融合两种方式使用基本情绪模型和环状情绪模型完成情绪识别的预期目标，课题的系统框架流程图如图 1.2 所示。

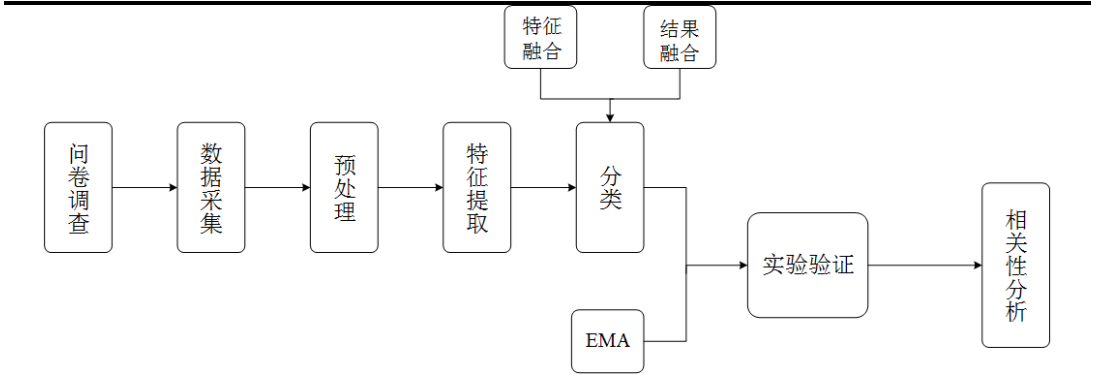


图 1.2 情绪识别系统框架流程图

### 1.4 论文结构

本文总共分为六章，各章的主要内容概述如下：

第一章描述了课题的研究背景和意义，并对情绪识别的四种主要方式即基于面部表情、基于语音语调、基于生理信号和基于人体姿态的研究现状分别进行了简单的描述，并指出现有研究中存在的一些问题。在归纳总结当前研究工作的基础上，提出了课题的三个主要研究内容以及本文的组织结构。

第二章主要描述了与课题工作密切相关的三个技术。首先介绍了情绪分类模型的种类和特点；然后描述了传感器数据处理过程，预处理和特征提取技术，着重介绍了噪声消除和数据分割相关方法以及常用的特征提取方法；最后对情绪识别中常用的分类算法进行了分析和描述，包括支持向量机、k-近邻、决策树、随机森林、AdaBoost 和梯度树提升六种。

第三章首先详细描述了网络问卷调查“情绪变化对行为的影响”，并对调查结果进行了统计和分析；之后对传感器数据收集平台和情绪记录软件分别作了详细的描述；最后对数据存储方式进行描述。

第四章对基于智能手机传感器的用户情绪识别技术中的数据处理部分进行研究，主要包括数据预处理、特征提取和数据归一化。文中针对不同类别传感器的数据处理过程进行了详细的阐述。

第五章对基于智能手机传感器的用户情绪识别技术中的数据分析部分进行研究，主要是实现了对用户的情绪识别。在实验验证过程中，对不同数据源的数据和不同情绪分类模型所取得的实验效果分别进行了描绘和准确性分析，并对算法性能进行评估。

第六章总结了本文的主要工作，并指出了未来的研究方向。

第二章 相关技术研究

上一章详细描述了课题的研究背景和意义，随后对基于面部表情、语音语调、生理信号和人体姿态的情绪识别研究现状分别进行分析，并对课题的主要研究内容进行描述。在本章中将会详细分析与课题研究相关的技术方法，包括情绪分类模型、数据处理过程以及分类算法。

2.1 情绪分类模型

进行情绪识别，面临的首要问题是建立情绪分类模型，即如何对情绪进行划分。通常上讲，采用一个国际通用的标准便于对不同的研究结果进行比较，虽然学术界关于如何有效的划分情绪状态一直存在争议，但其共通点是都认为情绪与生理反应之间存在一定的映射关系。目前使用较多的分类模型有两种，一种是将情绪以离散的模式划分为几种基本情绪<sup>[45][46]</sup>，如喜、怒、哀、惧；另一种则是认为情绪是连续的，使用维度空间模型进行划分<sup>[47]</sup>。

2.1.1 基本情绪模型

基本情绪模型认为情绪是离散可分的，由数种基本情绪类型组成，具有形容词标签的形式，如愉快、悲伤、愤怒等。这种情绪分类方式在人们的日常生活中普遍使用，并被广泛应用于早期的情感研究中。我国古代有所谓喜、怒、忧、思、悲、恐、惊之说的“七情论”，而美国的心理学家 Ekman 将情绪划分为喜悦、悲伤、恐惧、惊奇、愤怒、厌恶六种<sup>[48]</sup>。基于文化差异和不同的研究背景，因此在基本情绪的分类问题上，研究人员还没有达成普遍共识，图 2.1 总结并列举了关于基本情绪分类的研究成果<sup>[49]</sup>。

学者	基本情感
Arnold	Anger, aversion, courage, dejection, desire, despair, dear, hate, hope, love, sadness
Ekman, Friesen, Ellsworth	Anger, disgust, fear, joy, sadness, surprise
Fridja	Desire, happiness, interest, surprise, wonder, sorrow
Gray	Desire, happiness, interest, surprise, wonder, sorrow
Izard	Anger, contempt, disgust, distress, fear, guilt, interest, joy, shame, surprise
James	Fear, grief, love, rage
McDougall	Fear, disgust, elation, fear, subjection, tender-emotion, wonder
Mower	Pain, pleasure
Oatley, Johnson-Laird	Anger, disgust, anxiety, happiness, sadness
Panksepp	Anger, disgust, anxiety, happiness, sadness
Plutchik	Acceptance, anger, anticipation, disgust, joy, fear, sadness, surprise
Tomkins	Anger, interest, contempt, disgust, distress, fear, joy, shame, surprise
Watson	Fear, love rage
Weiner, Graham	Happiness, sadness

图 2.1 不同学者对基本情绪的分类定义<sup>[49]</sup>

情绪图书馆(emotion library)<sup>[50]</sup>作为情绪分类的一个标准，其中定义了 14 种基本情绪，如图 2.2 所示。但这些情绪中很多具有一定的相似性，区分起来难度很高。因此，本文采取了被社会心理学家广泛使用的一种划分方式：即将情绪图书馆中所有种类的情绪通过聚类的方式划分到五种标准情绪库中，分别为高兴(happy)、悲伤(sad)、害怕(fear)、生气(anger)、中性(neutral)。这五种广义情绪区分度明显，辨识度较高，因此将其作为基本情绪模型中的具体分类进行研究。

广义情绪	狭义情绪
Happy	Elation, Interest, Happy
Sad	Sadness
Fear	Panic
Anger	Disgust, Dominant, Hot anger
Neutral	Neutral normal, Neutral conversation, Neutral distant, Neutral tete, Boredom, Passive

图 2.2 基本情绪聚类图

2.1.2 维度空间模型

维度空间论则认为，情绪之间不是独立存在的，而是存在着一种连续的、渐变的关联，不同情绪之间可以平滑、逐渐地进行过渡。这里将情绪空间表述为一个笛卡尔空间，空间中的每一个维度都对应情绪的某一属性，因此每一种情绪状态都可以被描述为该笛卡尔空间中的一个映射点，并且各维度的坐标相对应的数值大小反映情绪在该属性下的强弱属性。由于情绪在维度空间模型中使用实数值进行描述，因此又被称为连续情感描述模型<sup>[51]</sup>。

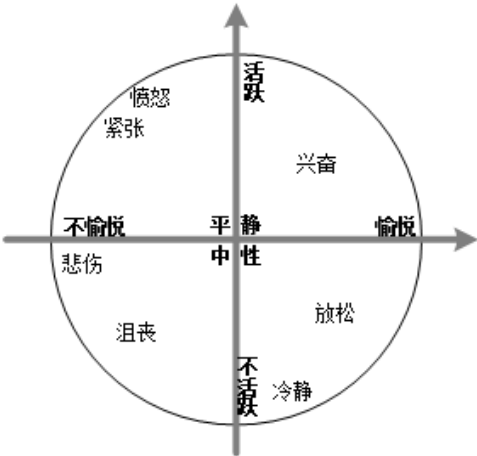


图 2.3 环状情感模型

其中，较为经典的维度空间模型有环状情感模型(circumplex model of affect)<sup>[52]</sup>和 Plutchik 的情绪轮模型(Emotion Wheel)<sup>[53]</sup>。在本文中我们采取了使用较为广泛的环状情感模型，如图 2.3 所示，这个模型由愉悦(pleasure)维度和

活跃(activeness)维度两部分组成。其中愉悦维度是水平轴，描述的是情绪的正面或负面程度，具体用来衡量一个人的情绪是积极还是消极；活跃维度是垂直轴，描述的是情绪的激烈程度，具体表示为一个人在某种情绪的下的行为是主动或是被动。通过对情绪状态的理解和估计，就可以将情绪在该二维空间中一一映射出来，很方便的进行情绪标签和空间坐标的相互转化。

2.1.3 分析比较

从表达情绪状态能力的角度看，基本情绪模型在大多数情况下只能刻画出较为单一的有限情绪类型，然而人们在日常生活中的展示的情绪却并不单单只是某一种，有可能是两种或者是多种的混合体，即使是同一种情绪也会有不同的激烈程度，因此基本情绪模型描述方式和真实情感之间还存在着较大的障碍，显示出了一定的局限性；维度空间模型则可以从多角度、连续地进行情绪的描述，并且能够以较为精确的数值对情绪状态进行描述，回避了基本情绪模型中的模糊性问题。从模型复杂度来看，基本情绪模型较为简洁，更有利于相关工作的研究；维度空间模型则需要面对如何从定性情绪状态转换为定量空间坐标的问题，研究工作相对不易展开。表 2.1 总结了两种情绪模型之间的区别，可以看出基本情绪模型和维度空间模型各有千秋。因此，在课题研究中分别使用了以为高兴、悲伤、害怕、生气、中性为衡量标准的基本情绪模型和以愉悦度和活跃度为衡量标准的二维空间模型进行对比试验，以寻求更合适的分类方法，得到更高的准确率。

表 2.1 两种情绪分类模型的区别

比较项	基本情绪模型	维度空间模型
情绪描述方式	形容词标签	笛卡尔空间中的点
情绪描述能力	离散描述，描述类型有限	连续描述，可描述任意类型
优点	简洁、易于研究工作展开	多角度、连续进行情绪描述，情绪描述能力高
缺点	单一、有限进行情绪描述，情绪描述能力低	将情绪从主观状态转化为客观具体数值是一个复杂的过程，不易于研究工作展开

2.2 数据处理过程

情绪识别需要收集能够反映人们日常情绪的各种数据，本文在基于智能手机收集移动数据进行情绪识别的过程中需要采集包括加速度、GPS、通话记录等各种类型的原始数据，在采集数据之后，因为收集到的数据存在不完整或不



一致的地方，无法直接进行数据处理，或是并不能得出有效的结果，因此需要先对数据进行预处理。只有经过预处理的数据才能在完整度、整齐度、冗余度等方面达到情绪识别的最低要求，在经过计算机的计算后提取出具有该类数据代表性的特征，即是特征提取过程。下面将分别对数据处理过程中的预处理和特征提取予以介绍。

2.2.1 数据预处理

原始数据库中的数据在从现实提取的过程中会出现很多空缺值或是错误和冗余的记录，不可避免的有一些缺陷，而很多成熟的算法都会对数据集的完整性、冗余性作出要求，因此原始数据很难直接满足数据挖掘的最低要求。并且随着近些年大数据研究的兴起，GB、TB 乃至 PB 级的海量数据计算越来越多，如此海量的数据中无意义的成分很多，如果不加以处理会严重影响算法的执行效率。在多种因素促使下，预处理的重要性不言而喻。

数据预处理<sup>[54][55]</sup>的典型方式有数据清理、数据集成、数据变换、数据规约四类，如图 2.4，下面将对这四类方法进行简介。

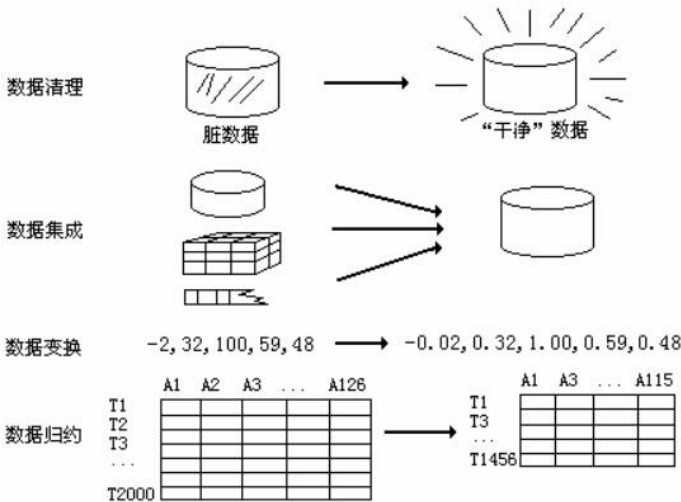


图 2.4 数据预处理方法

数据清理通过填写数据缺失值、光滑噪声数据、识别或删除离群点来达到数据格式标准化、清除异常和重复数据、纠正错误的目的，包括遗漏值处理、噪声数据处理、不一致数据处理和冗余数据处理四种方法。

数据集成是指将多个数据源的数据结合到一起并统一存储，最终建立数据仓库的过程，包括实体识别、数据冗余和数据值冲突的检测与处理。

数据变换是指通过数据平滑、数据聚集、数据概化、数据规范化、属性构造等形式将数据转化为适用于数据挖掘的形式，主要表现为从高维数据向低维数据的转化。

数据规约的目标是把握主要数据，一般在减少存储空间的同时尽量保证数据的完整性，具体方法有数据立方体聚集、维规约、数值规约、数据压缩、离散化和概念分层。

在采集数据过程中收集到的数据最常遇到数据缺失和存在噪音的问题。针对数据缺失的问题，我们通常使用某一全局常量、同类样本均值或是临近数据均值等数据对缺失位置进行填充。对于数据存在噪音这个问题，通常采用均值平滑或滤波器的方法对数据进行处理。均值平滑方法<sup>[56]</sup>是使用和原始数据相邻的两个数据点的均值对原始数据进行替换，目的是为了减少数据中的尖峰噪声。人体活动产生的数据多为中频，因此需要使用带通滤波器<sup>[57]</sup>对高频噪声和低频重力进行过滤，去除无关影响。

由于很多分类算法只能对离散数据集进行分类，因此连续数据离散化也是预处理中很重要的一个步骤。数据离散化最重要的是要合理的分割数据，通常采用滑动窗口分割技术，图 2.5 所示即为具有 50% 重叠的滑动窗口。通过滑动窗口方法，连续数据被分割为固定长度的片段，长度和窗口大小一致，长度多集中在 0.25s 到 10s 不等。

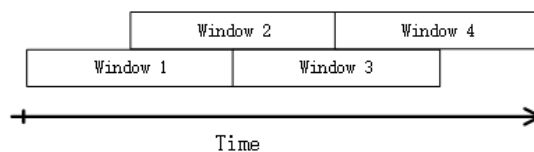


图 2.5 滑动窗口数据分割

### 2.2.2 特征提取

特征提取<sup>[58]</sup>是指通过对大量原始数据的处理，提取出一组最能代表该数据集模式并且具有明显物理意义或统计意义的少量特征。在特征提取过程中，对数据进行时域特征和频域特征的提取方式最为常见。

时域特征也称为信号的统计特征，主要是使用概率统计的方式对数据的均值、标准差（或方差）、最大值、最小值、皮尔逊相关系数、偏态值、四分位数值等衡量标准进行计算，部分特征计算方式如表 2.2 所示。

- 均值，又称为平均数，由一组数据中所有数据之和除以这组数据的个数所得，反映了数据的集中趋势。
- 标准差，又称均方差，方差的算术平方根，由各数据与其平均数之差的平方和进行算术平均数的平方根计算所得，反映了数据的离散程度。
- 最大值，数据中的最大数值，其余值都小于或等于该值，表示为  $\max(X)$ 。
- 最小值，数据中的最小数值，其余值都大于或等于该值，表示为  $\min(X)$ 。
- 皮尔逊相关系数，由两个变量与各自均值之间差值的乘积所得，反映了两个

变量之间的线性相关关系，即密切程度。

- 偏态值，等于均值与中位数的差值除以标准差所得数值的 3 倍，反映数据分布服从正态分布的程度，一般来说当偏态值计算结果(-1,1)之间就可以认为数据基本呈正态分布。
- 四分位数值，将所有数据由小到大排列并分成四等份，处于三个分割点位置（即 25%，50% 和 75%）的数据值即是。

表 2.2 常用时域特征计算公式

特征	计算公式
均值	$\mu = \frac{1}{n} \sum_{i=1}^n x_i$
标准差	$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}$
皮尔逊相关系数	$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$
偏态值	$3(\mu - m_{0.5}) / \sigma$

实际运用中，时域特征计算方法简单，便于应用。但由于计算方法较为简单，因此存在着丢失数据中有效信息的风险，进而影响到分类结果，所以频域特征是我们提取到尽量多的有效信息的另一个保证。

频域特征描述了信号在频率方面的特征，通常使用快速傅里叶变换(Fast Fourier Transformation, FFT)对分割后的数据进行频域转换，转换后的频域数据代表着信号在对应频率的振幅以及能量分布，主要对主频率和频域能量进行计算。原先的数据经过快速傅里叶变化后变为一组系数，取其中最大的 n 个系数值作为特征并记录它们的频率和振幅即为主频率。将所有系数的平方和与系数个数相处得到的即为频域能量。

很多情况下，在时域上很难看出信号的特征，但经过 FFT 变换到频域中时，特征则显而易见，这也是特征提取技术中总是将时域特征与频域特征相结合使用的原因。

## 2.3 分类算法

分类问题就是确定对象属于哪个预定义的目标类的过程，它作为生活中一个普遍存在的问题，早已受到学术界的关注并有了广泛的应用。例如医生可以根据核磁共振的结果区分病人的肿瘤是良性的还是恶性的，电子邮件系统可以根据邮件的标题、发件人等信息判断邮件是否为垃圾邮件以及附件是否安全可靠。分类算法属于有监督的学习算法，是一个通过利用分类器对已知类别的数

据样本进行训练和分析以发现分类规则，并根据发现的规则预测新数据类别的过程，处理问题的大致流程图如 2.6 所示。本节将对在课题实验过程中使用的几种分类算法进行介绍，包括支持向量机、k-近邻、决策树、随机森林、AdaBoost 和梯度树提升六种。

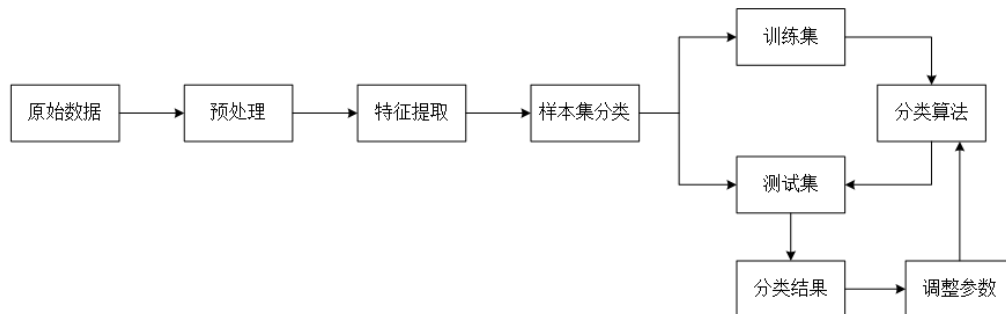


图 2.6 分类算法流程图

### 2.3.1 支持向量机(SVM)

支持向量机(Support Vector Machine, SVM)<sup>[59]</sup>是 Cortes 和 Vapnik 于 1995 年提出具有坚实的统计学理论基础的一种有监督的学习模型。SVM 可以很好地解决小样本数据、非线性数据等问题，并且能够很好的应用于高维数据，避免维数灾难问题。

当数据在  $n$  维空间中并不能很好的被区分开时，就需要映射到更高维的空间（如  $n+1$  维）进行分类。通过构造一个分隔超平面  $H$ ，如图 2.7 所示，使得与该平面平行的超平面  $H_1$  和  $H_2$  可以将两个不同类别的数据分开。当  $H$  在某一方向的平行超平面  $H_1$  和  $H_2$  之间的距离最大时即为最大间隔超平面，即为使用 SVM 分类的最优情况，此时分类误差最小。为了构造出具有良好性能的 SVM，关键在于核函数的选择，常用的核函数包括线性核函数、径向基核函数(RBF)、多项式核函数和 Sigmoid 函数。

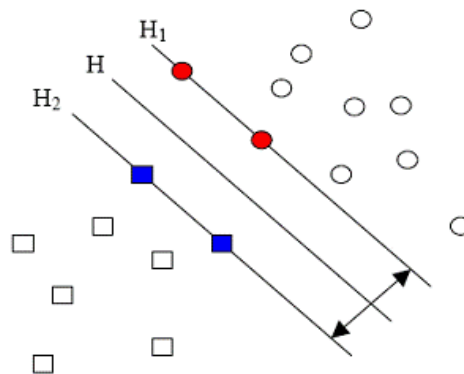


图 2.7 SVM 最大间隔超平面

**SVM** 作为一种小样本学习算法,将计算的复杂性从样本空间的维数转移到核函数中支持向量的数目,这样很好地避免了“维数灾难”,并且具有较好的鲁棒性。但是 **SVM** 存在对于大规模训练样本难以实施的缺点;并且由于只能进行二类分类,因此在对多类别样本进行分类时,需要使用多个二类支持向量机的组合,具有一定的困难性。

### 2.3.2 k-近邻(kNN)算法

k-近邻 (k-Nearest Neighbor, kNN)算法<sup>[60]</sup>是机器学习分类算法中最简单的算法之一,其核心思想是每个样本可以用离它最近的  $k$  个邻居中大多数所属的类别来表示,前提是这  $k$  个邻居都必须是已经正确分类的样本。如图 2.8 所示,现要判断红色圆所属类别是蓝色正方形还是绿色三角形。根据  $k$  取值的不同,判断结果也可能产生变化。当  $k=3$  时,由于蓝色正方形所占比例为  $2/3$ ,红色圆属于蓝色正方形类别;当  $k=7$  时,由于绿色三角形所占比例为  $4/7$ ,红色圆属于绿色三角形类别。

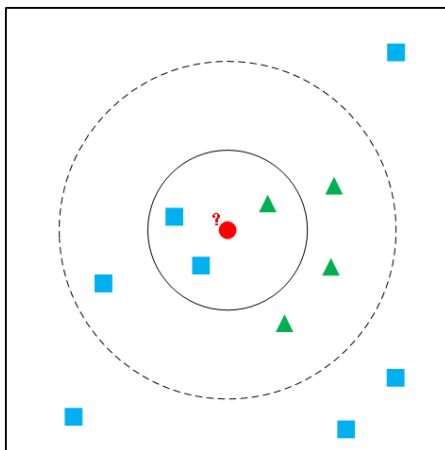


图 2.8 kNN 算法决策过程

作为理论上较为成熟的分类算法, **kNN** 简单、易于操作和实现,尤其在多分类问题上比 **SVM** 算法的表现要好。但 **kNN** 的计算量较大,因为待分类样本  $k$  个最邻近点的选择需要计算它到全体已知样本的距离,因此当样本数量较大并且待分类样本个数较多时,对分类效率是种考验。

### 2.3.3 决策树

决策树(Decision Tree)算法<sup>[61]</sup>是一种经典的分类算法,产生于上个世纪六七十年代,是现在应用最为广泛的分类算法之一。决策树是一种从无规则的样本中推理出具有决策树表示形式的分类规则的算法,即从根节点开始,根据不同的属性选择不同的非叶节点向下分支,直到某一叶节点为止,该叶节点即为

样本所属的分类，如图 2.9 所示。决策树算法采用了自顶向下的递归方式对样本分类，从根节点到每一个叶节点的对应路径都代表着一合取规则，整个决策树就是一个析取表达式。

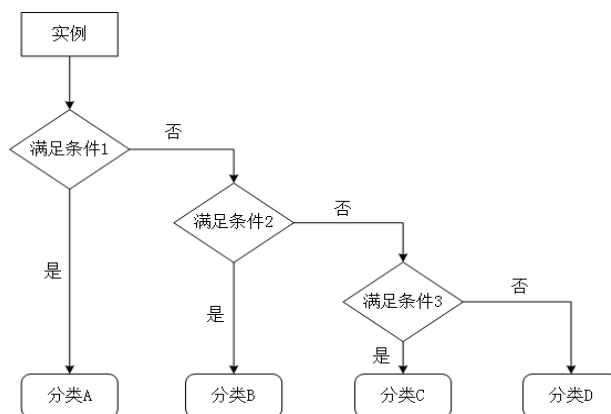


图 2.9 决策树模型

决策树的构造主要由两部分构成：决策树的生成和决策树的剪枝。训练样本生成决策树的过程被称为决策树的生成。决策树的剪枝则是指使用测试数据集对上一阶段生成的决策树进行验证和修正的过程，通过将影响准确性的分支剪除而提高准确率。

决策树的典型算法有三种，ID3，C4.5 和 CART，其中 C4.5 算法由 ID3 改进而得。ID3 的优势在于减少了树的深度，但是却忽略了对叶节点数目的研究。C4.5 对此进行改进，产生的分类规则易于理解，准确率高，曾在国际权威学术组织评选的数据挖掘领域十大经典算法中位列第一名；但是在生成树的过程中需要对数据集反复顺序扫描和排序，从而对算法效率有所影响。虽然如此，决策树算法的优势依然比较明显，它的生成模式简单、对噪声数据具有很好的健壮性并且分类准确率高。

#### 2.3.4 随机森林

随机森林(Random Forest)<sup>[62]</sup>是由 Breiman 在 2001 年提出的一种新型分类算法，可以理解为由许多随机生成没有关联的决策树构成的森林，实质是对决策树算法的一种改进。

单棵决策树的分类能力较为有限，但当随机产生大量的决策树后，分类能力将成倍增加。随机森林就是通过自助法重采样技术，从初始数据样例集  $M$  中有放回的随机抽取  $k$  个样例生成训练样例集  $N$ ，然后根据训练样例集  $N$  中的  $k$  个样例随机生成  $k$  个决策树，并由这  $k$  棵树组成随机森林。这些决策树之间彼此并没有关联，当一个新的样例输入时，由这  $k$  个决策树做出分类判断，哪一种分类结果被选择的最多，那么样例就被预测为哪个类别。

随机森林算法在很多数据集上相对于其他的算法都有着很大的优势，具体表现如下：数据上表现良好，不易陷入过拟合；具有较好的抗噪能力；对数据集适用能力强，可灵活处理连续型数据，离散型数据和高维数据，并且不用做特征选择；训练速度快；在决定类别时能够评估变量的重要性；实现相对简单等。

### 2.3.5 AdaBoost

AdaBoost 算法<sup>[63]</sup>是一种迭代算法，通过将若干弱分类器集合形成强分类器从而提高对数据的分类能力。初始时每个数据都被赋予了同等的权重进行训练得到一个弱分类器，并通过这个弱分类器对准确率进行判断。如果某个样本在此次分类过程中已被准确分类，那么在构造下一个数据集中它的权重相应就会降低，而没有被正确分类的样本点权重会增大。这么做的目的在于突出被分错的样本，从而在新的训练中得到新的样本分布。周而复始最终得到  $N$  个弱分类器，再通过一定的算法将其全部融合起来最终得到一个强分类器。如图 2.10 所示，需要对(1)中的数据进行分类得到(2)，分析(2)中结果，降低分类结果正确的样本权重并对分类错误的样本权重进行增加（图中以对样本的放大或缩小来代表加权或减权）。如此迭代，并最终对所有弱分类器进行整合，得到(5)中的分类结果。

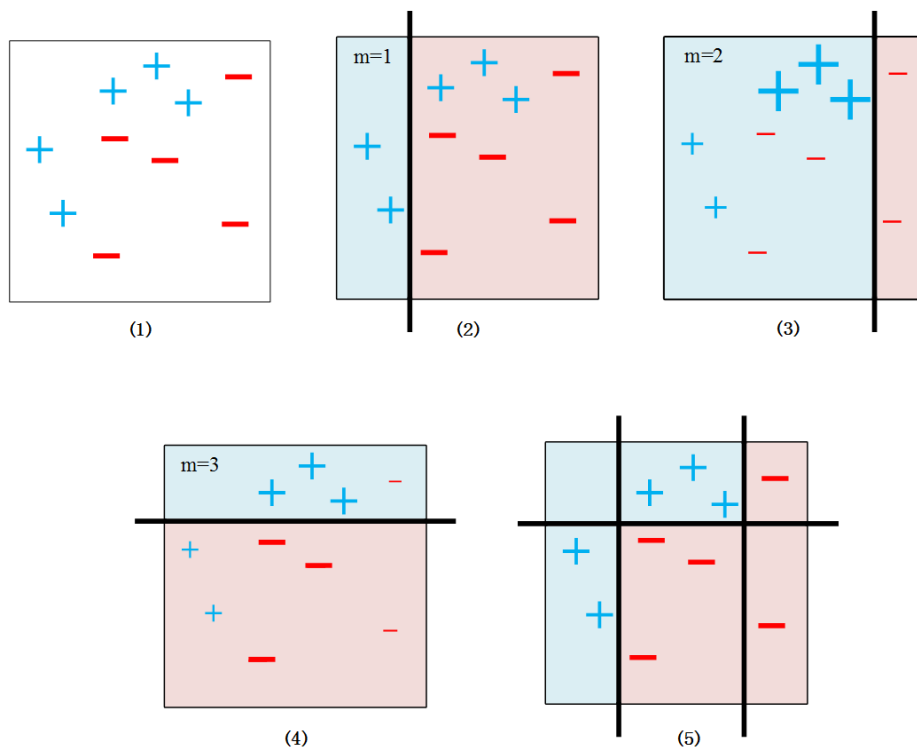


图 2.10 AdaBoost 算法示意图

AdaBoost 作为一种高精度的分类器使用起来较为简单，而且也不用考虑

过拟合的问题，优点较为明显。但是在最终的强分类器整合过程中需要考虑如何将各个弱分类器进行组合形成强分类器，如果只是将各个弱分类器进行简单的叠加，分类效果可能不尽如人意。这时不如考虑根据之前的分类结果准确率对各个弱分类器赋予权重，准确率高的权重高，反之，则权重低，这样组合形成的强分类器将具有较高的准确率。

### 2.3.6 梯度树提升(GTB)

梯度树提升(Gradient Tree Boosting, GTB)<sup>[64][65]</sup>是一种组合算法，又被称为梯度回归提升树(Gradient Boosted Regression Trees, GBRT)，它与随机森林、AdaBoost 并称为最常用的三个基于决策树的组合算法。GTB 是对 Boost 算法的一种改进，其核心思想是每棵树从先前所有树的残差（指实际观察值和拟合值之间的差）中进行学习，以减少残差为目的，在残差减少的梯度方向建立新模型。也就是说，每个新模型的建立都是使上一个模型残差在梯度方向可以减少。

GTB 在分类性能上可以和随机森林相媲美，在某些数据集上可能表现会更好，它适合处理混合类型的数据，具有很强的预测能力，对输出空间的异常值具有很好的鲁棒性。然而，GTB 在可扩展性方面表现并不理想，因为提升时序性的原因并不能进行并行处理。

## 2.4 本章小结

本章对课题研究过程中涉及的三方面技术进行了详细的分析和研究。2.1 节介绍了常用的两种情绪分类模型，并对它们的优势与劣势进行了分析比较。2.2 节对数据处理过程进行描述，分别详细的介绍了数据预处理和特征提取的过程。2.3 节描述对数据挖掘领域常用的六种算法进行分析和研究，所研究算法包括支持向量机、k-近邻、决策树、随机森林、AdaBoost 和梯度树提升。



### 第三章 数据收集和存储机制研究

本文首先需要确定情绪如何影响用户的行为模式并根据它们之间的关系选择需要收集的数据，因而作为整个研究前提的移动数据，如何对其进行获取和存储就显得尤为重要。本章将基于网络问卷调查的结果展开讨论并确立需要收集的传感器数据，然后对数据收集过程中手机传感器数据收集和情绪记录数据收集的设计和实现分别予以阐述，最后对手机中的数据存储机制进行描述。

#### 3.1 预调查 — 网络问卷调查

问卷调查作为社会调查中的一种数据收集方式，通常以书面提出问题的形式进行数据收集。研究者首先将要调查的问题制作成问题表格，然后交由调查对象填写，最后回收问卷并分析整理得出结果。问卷调查通常有纸质问卷调查和网络问卷调查两种形式，由于网络问卷调查具有成本相对低廉、无地域限制、便于统计结果等多种优势，因此在此次课题研究中选择网络问卷调查作为预调查的形式。

##### 3.1.1 问卷调查内容的设计

在数据采集前，作为前期准备，首先需要确定收集的数据类别。因为情绪变化会导致人的行为模式发生变化，因此只需选定能反映出用户行为变化的手机数据进行采集，然后通过对这些数据的分析就能识别用户情绪。为了全面选出反映个人行为变化的传感器数据，本文依据大部分智能手机中存在的功能和传感器对问卷调查的题目进行设计，并在网上进行发布，问卷调查具体内容如表 3.1 所示。

表 3.1 问卷调查：情绪变化对行为的影响

问卷调查说明：本次实验旨在通过手机传感器数据采集的结果对用户的情绪进行推测，因此设计问卷调查筛选出情绪变化会导致发生变化的行为，和这些行为相关的传感器数据将被记录。	
问题 1	您的性别： A. 男 B.女
问题 2	您的出生年份： A. 1980 年以前 B. 1980 年至 1989 年 C. 1990 年以后
问题 3	请问您是已经工作还是仍在上学： A. 工作 B. 上学
问题 4	您每天携带手机的时间： A. 手机基本随身携带 B. 从不随身携带手机 C. 介于二者之间

问题 5	您每天使用手机的时间： A. 小于 2 小时 B. 2 ~ 4 个小时 C. 大于 4 小时
问题 6	您是否认为情绪的变化会影响到您的行为活跃度？ 譬如说情绪低落时比较喜欢独自一个人呆着，情绪高涨时喜欢和朋友一起出去玩。 A. 是 B. 否
问题 7	您是否认为情绪变化会影响到您使用手机应用程序的种类以及时间长短？ 譬如说心情不好时喜欢打游戏或心情不错时喜欢和朋友聊天。 A. 是 B. 否
问题 8	您是否认为情绪变化会影响到您平时和朋友或同事之间的交流？ 譬如说会影响到打电话的频率和时长、发短信的数量、以及微信、QQ 等社交工具的使用。 A. 是 B. 否
问题 9	您是否认为情绪变化会影响到您的活动范围？ 譬如说情绪高涨时喜欢外出旅游或到人员密集的地方，而情绪低落时更喜欢在家独自一人呆着。 A. 是 B. 否
问题 10	您是否认为情绪变化会影响到您的运动情况？ A. 是 B. 否
问题 11	您是否认为情绪变化会影响您使用开/锁屏功能？ 譬如说心情烦躁时会频繁的开屏或锁屏。 A. 是 B. 否
问题 12	您觉得如果需要记录您的情绪，多长时间记录一次比较合适？ 即您能承受的最高频率。 A. 2 小时的间隔 B. 4 小时的间隔 C. 6 小时的间隔
问题 13	您认为还有哪些可以通过手机捕捉的行为会受到情绪变化的影响？ _____

该份问卷调查从用户的性别、年龄、手机使用习惯和可能被传感器捕捉到的行为的角度对智能手机使用者进行调查。问题设置以选择题为主，问题数量保持在一页之内，方便用户填写；此外，还设置了一个问答题，被调查者开放性的回答有助于补充问卷调查题目设置的不足。

### 3.1.2 问卷调查的发放和回收

此次主题为“情绪变化对行为影响”的问卷调查自 2015 年 2 月 1 日发布于问卷星网站<sup>[66]</sup>，至 2015 年 2 月 15 日回收，总共历时十五天，总共有 125 人匿名参加，最终回收有效问卷 117 份。

### 3.1.3 问卷调查结果分析

117 份有效问卷的前 12 道选择题统计结果如表 3.2 所示, 经过数据统计和分析, 参与本次网上问卷调查活动的男女人数比例维持在 1.2:1, 即人数大致相等, 因此其余 11 道题的选择结果基本可以排除因某一性别比例较高导致选择结果有针对性这一影响因素。在被调查者的年龄和身份问题上, 14.5% 出生于 1980 年以前, 33.3% 是 80 后, 有超过一半的人都是 90 后, 其中 72.6% 的被访者仍是学生, 这间接表明了此次问卷调查的主体以及他们潜在的生活方式和作息习惯等信息。超过 85% 的人处于 35 岁以下, 他们对于新事物和新科技的接受能力较强, 并且这部分参与者随身携带手机的比例超过 90%, 因此本文选择智能手机作为传感器数据收集的平台完全合理。

表 3.2 问卷调查选择题统计结果

单位: 人	选项 A	选项 B	选项 C
问题 1	64	53	
问题 2	17	39	61
问题 3	32	85	
问题 4	104	0	13
问题 5	23	55	39
问题 6	112	5	
问题 7	109	8	
问题 8	114	3	
问题 9	110	7	
问题 10	69	48	
问题 11	89	28	
问题 12	5	77	35

结果显示, 47% 的参与者每天使用手机的时间在 2 小时至 4 小时之间, 33.3% 的参与者每天使用手机的时间超过 4 小时, 只有不足 20% 的参与者使用手机时间少于 2 小时。大部分参与者都认为包括正在使用的应用程序、通话记录、短信、微信、运动情况等在内的多项数据都可以反映出使用者的行为。另外, 根据第 13 道非选择题的统计答案, 又对需要收集数据类别做了补充。最终, 我们根据问卷调查的结果决定了需要收集的传感器数据项以及用户手动记录情绪的最适频率, 以实现准确率高且干扰性低的预期目标。

### 3.2 数据收集

根据 3.1 节调查问卷的统计结果, 本课题最终选择了包括通话记录、短信记录、微信消息、QQ 消息、蓝牙、WiFi、GPS、加速度计、陀螺仪、磁力计、光传感器、手机应用程序使用情况、屏幕解锁和锁屏在内的十余种传感器数据

进行采集,并确定每五小时记录一次情绪。下面将对传感器数据收集软件和情绪记录软件分别予以介绍。

### 3.2.1 手机传感器数据收集

现今智能手机操作系统中最为广泛使用的有谷歌的安卓系统、苹果的 iOS 系统和微软的 Windows Phone 系统,其中安卓系统以开放性著称,成为很多第三方开发者的乐园。鉴于安卓系统在全球超过 80% 市场份额的现状<sup>[68]</sup>以及易于开发应用的特点,课题中所有数据收集工具的开发都是基于安卓平台。本文使用 StarLog 作为传感器数据收集工具,StarLog 是根据需要对开源应用程序 UbiqLog<sup>[67]</sup>进行扩展和改进的一款数据收集软件,其主界面如图 3.1 所示。打开 StarLog 应用时,需要点击下方按钮使其切换至蓝色才会开始数据收集工作。StarLog 只需从后台运行,不会影响用户正常使用手机;当切换到后台运行时,StarLog 将在通知栏显示一条通知提醒用户,通过点击此条通知可直接进入主界面查看或停止数据采集工作,如图 3.2 所示。



(a)

(b)

图 3.1 数据收集框架主界面

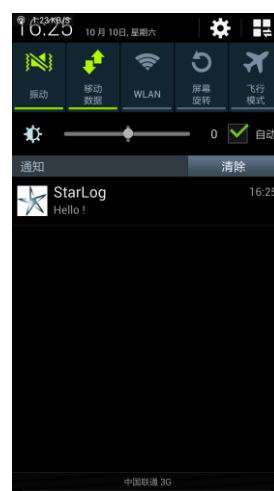


图 3.2 StarLog 通知栏消息

根据 3.1 节统计分析后得到的结果,我们确定收集以下种类的数据:①运动类感知数据,包括加速计、陀螺仪、磁力计以及活动状态(可判定用户处在平稳、缓慢或快速状态);②环境类感知数据,包括光传感器和 GPS 信息;③手机使用数据,包括通话、短信、微信、QQ 以及蓝牙记录等社交记录, WiFi、App 使用、屏幕解锁和锁屏等使用记录以及手机状态记录(手机处于铃声/静音/飞行、是否联网、是否充电等)。其中,需要打开手机中 WLAN、蓝牙和 GPS 三项功能,并对部分数据的收集频率进行设定,收集频率的不同取值会影响到数据的精度,最终对情绪识别准确率产生影响。通过对收集频率的反复试验,在对手机耗电量影响尽可能小与实验数据精度尽可能高的双重前提下,我们设

定应用采集频率每 5 秒一次，蓝牙扫描频率每 10 分钟一次，WiFi 扫描频率每 15 分钟一次，位置扫描频率每 30 秒一次，其他数据如通话、短信、微信、QQ、划屏等当手机接收消息或捕捉到变化时进行记录，加速度计、陀螺仪、磁力计、活动状态等则是软件调用的安卓系统接口，根据不同手机的内置信息采样频率可能有所不同。表 3.3 中详细描述了每种传感器具体采集的数据。

表 3.3 传感器数据收集内容

传感器	描述	数据内容
应用	记录当前主屏幕正在运行的程序（前台应用程序）	应用名称、开启时间、关闭时间
通话	记录所有来电/去电电话记录	通话时长、通话时刻、接拨类型
短信	记录来往短信记录	接发类型、时间
蓝牙	扫描到的蓝牙设备	设备名称、Mac 地址、状态、时间
WiFi 热点	扫描到的 WiFi 热点	名称、Mac 地址、信号强度、连接状态、时间
位置（高频）	利用 GPS 或者网络获取当前位置	经纬度、海拔高度、GPS 速度、时间、精确度、GPS 定位/网络定位
活动	利用加速度计进行的简单活动识别	开始时间、结束时间、活动状态（平稳/缓慢/快速）
加速度	加速度传感器	x,y,z 轴的数值、时间
光度	光敏传感器	光度值、时间
陀螺仪	陀螺仪传感器	x,y,z 轴的数值、时间
磁力计	磁力计传感器	x,y,z 轴的数值、时间
划屏	屏幕的开关及解锁记录	操作（亮屏/暗屏/解锁）、时间
微信/QQ	接收到的微信/QQ 用户发给自己的消息	名称、类型（文字/图片/语音/小视频）、时间
手机状态	手机的状态	飞行模式是否开启、铃声/静音、网络状态（无网络/WiFi/移动网络）、是否在充电

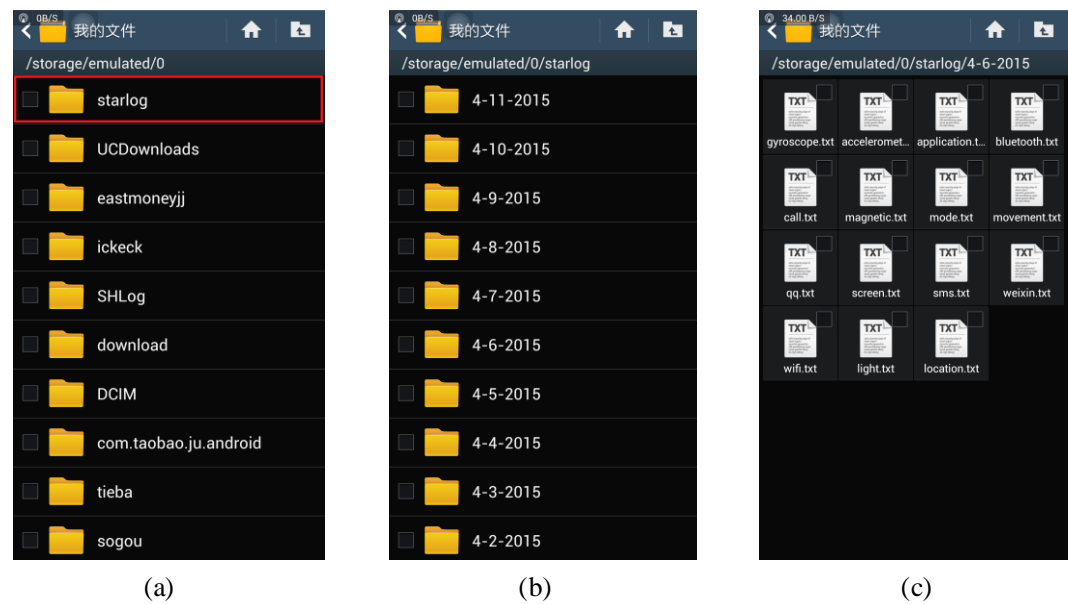


图 3.3 StarLog 文件存储形式

程序 StarLog 会在手机的存储空间中自动生成名为“starlog”的文件夹，并在该文件夹中以“mm-dd-yyyy”的命名规则为每天采集的数据集合命名，并在子文件夹下以 txt 文档形式进行存储，具体表现形式如图 3.3 所示。

因为当用户睡觉后手机会处于闲置状态，并且根据实验开始前对志愿者的询问，大多数人会在早上 7 点至 7 点半起床，晚上 10 点半至 11 点睡觉，且睡觉有关机或者使用飞行模式的习惯，因此本文中数据收集时间是每天从早上 7 点半到晚上 10 点半，总计 15 个小时。

### 3.2.2 情绪数据收集

根据 3.1 节问卷调查统计结果，被调查者选择愿意每 4 小时或每 6 小时记录一次的结果总和超过 90%，因此，以此作为参考，同时本着对参与者干扰最少的原则，本文最终选择每隔 5 个小时记录一次用户情绪状态，即每天记录三次，分别代表当日上午、下午和晚上的情绪。

为了方便参与者对其情绪状况进行记录，设计并开发基于安卓平台的情绪记录收集工具，如图 3.4 所示。其中，位于左侧的 a 图上部分有两行滑动条分别代表愉悦度和活跃度，取值范围为 1~5，分别代表非常不愉悦(非常不活跃)、不愉悦(不活跃)、中性(平静)、愉悦(活跃)、非常愉悦(非常活跃)；a 图下部分是一个离散的基本情绪输入对话框。因此用户需要在 a 图中采用两种情绪模型分别进行记录，如前文所述，完成记录后需要点击下方保存情绪按钮。b 图是一个展示历史情绪记录的界面，用户可以对已记录的情绪进行查看、删除等操作。c 图则为文件导出界面，可以将情绪记录数据导出为 txt 文档或 sql 数据库格式，并具有清除数据的功能。



图 3.4 情绪记录软件

3.3 数据存储机制

不论是从计算能力还是从运算速度方面考虑，在手机上处理数据都不具备优势，因此本课题中手机只作为一个数据收集和暂时存储的终端，等数据收集工作结束后需要将数据上传到服务器端进行处理和分析。因此在移动端数据存储格式的选择就显得格外重要，合适的数据存储格式有利于服务器端数据预处理工作的展开。

3.3.1 数据存储格式

根据 3.2 节中描述可见，课题中所需要收集的数据不仅类型具有多样性，每种数据的具体内容也不尽相同，因此需要选择一种合适的数据存储格式对各类传感器数据统一描述，它应该具有占用存储空间小，封闭性好，便于传输等优点。目前较为流行的两种数据传输格式为 XML 和 JSON，下面将分别对这两种格式进行介绍，并从空间占用、数据传输、数据解析等方面对二者进行比较。

可扩展标记语言 XML(Extensible Markup Language)<sup>[69]</sup>是 1998 年由 W3C(World Wide Web Consortium)批准的一种用于标记文档和数据使其具有结构性的标记语言，用于数据传输和数据存储，它非常适合在万维网进行跨平台的远程数据交互，数据共享方便。不过作为一种传统的传输格式，XML 格式虽统一却复杂，且文件过于庞大，在解析过程中会花费较多时间和资源，不易维护。

JSON(JavaScript Object Notation)<sup>[70]</sup>作为一种新兴起的轻量级数据交换语言，由于它采用了类 C 语言家族（C、C++、C#、JAVA 等）的习惯，因此不仅易于人进行阅读和编写，而且也易于机器传输和解析。

表 3.4 对 XML 和 JSON 各自的优势与不足进行了总结和对比，不难发现，JSON 格式简单，易于读写，并且从空间占用和数据传输方面都更占优势。当用于存储和传输的数据为轻量级时，明显使用 JSON 更为合适，因此我们采用 JSON 作为移动端数据存储格式。

表 3.4 数据存储格式比较

	XML	JSON
可读性	规范的标签形式	语法简易
可扩展性	具有很好的扩展能力	除了具备 XML 具备的扩展能力，还可存储 Javascript 对象
数据解析	解析复杂，需要考虑父节点、子结点，易使服务器和客户端代码异常复杂、不易维护	易于解析，只需考虑键值对或值

空间占用	冗长	格式都是压缩的，冗余小
通用性	通用性高	比 XML 低，处于推广阶段

3.3.2 基于 JSON 格式的数据实现

JSON 有对象和数组两种结构，课题中使用的 JSON 结构为对象形式，该结构允许相互嵌套。对象结构是由一组无序的“键/值”对组成的集合，其中键(Key)指所要描述的数据的名字，值(Value)指该数据的取值。每个对象都以“{”（左括号）标记开始，并以“}”（右括号）标记结束，键和值之间用“:”（冒号）标记隔开，两个“键/值”对之间则用“,”（逗号）标记分隔，数据结构具体表现形式为{key: value, key: value,...}, 如图 3.5 所示。

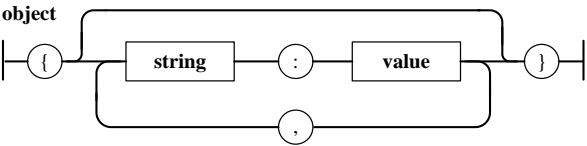


图 3.5 对象的格式设置

其中“值”的取值可以是数值（整数或浮点数）、布尔值（true 或 false）、字符串（由双引号所引）、对象（由花括号所引）、数组（由方括号所引）和 null 值，如表 3.5 所示。示例中，“Name”是 Key，“Jack”是“Name”的 Value；“Student”是 Key，冒号后面的花括号中所有内容都是“Student”的 Value。

表 3.5 JSON 格式举例

```
{ "Student": { "Name": "Jack ",
               "Age": 14,
               "IsAdult": false,
               "Course": { "a.m.": " math", "p.m.": "history" },
               "Friend": [ "Lucy", "Jim" ],
               "Address": null }
}
```

根据上文中 JSON 的定义对传感器数据的存储结构进行格式化，比如位置信息由纬度(Latitude)、经度(Longitude)、海拔高度(Altitude)、时间(time)、精确度(Accuracy)、GPS 速度(Speed)、定位方式(Provider)等内容构成。表 3.6 详细展示了每种传感器数据的 JSON 存储格式，Value 具体数值以“#”进行代替，其中光传感器、加速度计、陀螺仪、磁力计四种传感器采集数据为连续值，并不以 JSON 作为存储格式，而是直接记录具体数字。

表 3.6 传感器数据 JSON 存储格式

传感器	JSON 格式
应用	{ "Application": { "ProcessName": "#", "Start": "#", "End": "#" } }
通话	{ "Call": { "Number": "#", "Duration": "#", "Time": "#", "Type": "#" } }



短信	{ "SMS":{ "Address":"#", "type":"#", "date":"#" } }
蓝牙	{ "Bluetooth":{ "name":"#", "address":"#", "bond status":"#", "time":"#" } }
WiFi 热点	{ "Wifi":{ "SSID":"#", "BSSID":"#", "level":"#", "state":"#", "time":"#" } }
位置（高频）	{ "Location":{ "Latitude":"#", "Longitude":"#", "Altitude":"#", "time":"#", "Accuracy":"#", "Speed":"#", "Provider":"#" } }
活动	{ "Movement":{ "start":"#", "end":"#", "state":"#" } }
加速度	accelerometer-x, accelerometer-y, accelerometer-z, time
光度	LightValue, time
陀螺仪	gyroscope-x, gyroscope-y, gyroscope-z, time
磁力计	magnetic-x, magnetic-y, magnetic-z, time
划屏	{ "SCREEN_ON":{ "time":"#" } } { "KEYGUARD_DEBLOCKED":{ "time":"#" } } { "SCREEN_OFF":{ "time":"#" } }
微信	{ "weixin":{ "id":"#", "type":"#", "time":"#" } }
QQ	{ "qq":{ "id":"#", "type":"#", "time":"#" } }
手机状态	{ "Mode":{ "time":"#", "ringer":"#", "airplane":"#", "network":"#", "power":"#" } }

3.4 本章小结

本章就传感器的数据收集和存储问题进行了探讨，首先设计了主题为“情绪变化对行为影响”的问卷调查，通过对问卷调查结果的统计与分析确定了需要收集的传感器数据种类。接下来对传感器数据收集软件和情绪记录软件的功能进行了介绍。最后简述了移动端数据的存储机制，主要采用 JSON 格式对数据进行存储和交换。

## 第四章 多维移动数据的处理机制

第三章描述了数据的收集和存储机制，本章将着重描述传感器数据的处理过程。根据传感器数据的不同属性，我们将收上来的数据分为三个类别，分别是运动类感知数据、环境类感知数据和手机使用数据，下面将分别对这三类数据的预处理和特征提取过程进行详细的描述。

### 4.1 数据的预处理

我们用 StarLog 收集的原始数据经上传至服务器才能进行预处理工作，原始数据会存在不完整、不一致以及含噪声的情况，因此数据预处理工作的必要性不言而喻。数据的预处理不仅可以显著提高数据质量，更可以为后续处理过程节省大量的时间和空间。因为我们选择 JSON 格式对数据进行存储和传输，因此预处理的第一步是要先对数据进行 JSON 解析，然后再按照通用的流程根据需要选择数据清理、数据集成、数据变换、数据规约等方法中的一个或几个进行后续处理。

#### 4.1.1 JSON 解析

在 3.2 节中已经对收集的数据种类进行了明确，在这 15 种数据中，根据数据的类型可以分为连续数据和离散数据两种，其中光传感器、加速度计、陀螺仪和磁力计这四种数据属于连续数据，在采集过程中只记录了相对应的数值和时间，并未进行 JSON 格式化，因此只有离散类数据需要进行 JSON 解析，经过 JSON 解析后的数据格式标准，便于进行预处理的其他操作。

下面以 WiFi 数据为例进行 JSON 解析介绍。图 4.1(a)中为 WiFi 原始数据，可以看出其存储格式是以 SSID、BSSID、level、state、time 为 Key 的 JSON 格式，经过 JSON 解析后的标准数据格式如图 4.1(b)所示，JSON 解析的核心代码如表 4.1 所示。在解析 WiFi 数据过程中，由于部分文件中的"SSID"值为中文或是特殊符号，因此出现不能解析的问题；这种情况在蓝牙文件中的"name"Key 中也出现。针对此种问题，我们选择对 WiFi 中的"SSID"和蓝牙中的"name"这种不能被 JSON 解析的 Key 值进行忽略，最终解析结果为图 4.1 (c)。采用这种方法主要是由于"SSID"表示无线网络名称，存在重复可能性；而"BSSID"表示无网线路接入点的 MAC 地址，具有唯一性。因此 BSSID 作为可以唯一标示每个无线网络的地址，更具有标志性。在蓝牙的文件中我们选择"address"而忽略"name"的意义也在于此，因为"address"表示蓝牙的 MAC 地址，

由手机制造商为每台手机单独分配，同"BSSID"一样具有唯一性。

```
1 {"Wifi":{"SSID":"ChinaCCTV","BSSID":"c8:3a:35:45:61:60","level":-49,"state":"connected","time":"4-14-2015 07:23:10"}}
2 {"Wifi":{"SSID":"","BSSID":"c8:3a:35:45:61:60","level":-46,"state":"connected","time":"4-14-2015 07:23:10"}}
3 {"Wifi":{"SSID":"cmcc-edu-5","BSSID":"1c:fa:68:46:f7:32","level":-43,"state":"not_connected","time":"4-14-2015 07:23:10"}}
4 {"Wifi":{"SSID":"CMCC-Edu","BSSID":"d4:ee:07:0d:17:4e","level":-55,"state":"not_connected","time":"4-14-2015 07:23:10"}}
5 {"Wifi":{"SSID":"FAST_xxtty","BSSID":"28:2c:b2:d4:ab:f2","level":-67,"state":"not_connected","time":"4-14-2015 07:23:10"}}
6 {"Wifi":{"SSID":"cmcc-7","BSSID":"14:75:90:7b:74:38","level":-85,"state":"not_connected","time":"4-14-2015 07:23:10"}}
7 {"Wifi":{"SSID":"Tenda_532BD8","BSSID":"c8:3a:35:53:2b:d8","level":-76,"state":"not_connected","time":"4-14-2015 07:23:10"}}
8 {"Wifi":{"SSID":"CMCC-EDU","BSSID":"28:2c:b2:21:7a:9e","level":-91,"state":"not_connected","time":"4-14-2015 07:23:10"}}
9 {"Wifi":{"SSID":"CMCC-EDU","BSSID":"14:75:90:75:67:92","level":-91,"state":"not_connected","time":"4-14-2015 07:23:10"}}
10 {"Wifi":{"SSID":"TP-LINK_IPTV_7B7438","BSSID":"1a:75:90:7b:74:38","level":-76,"state":"not_connected","time":"4-14-2015 07:23:10"}}
11 {"Wifi":{"SSID":"TP-LINK_IPTV_756792","BSSID":"1a:75:90:75:67:92","level":-91,"state":"not_connected","time":"4-14-2015 07:23:10"}}
12 {"Wifi":{"SSID":"TP-LINK_7AA04A","BSSID":"d0:c7:c0:7a:a0:4a","level":-91,"state":"not_connected","time":"4-14-2015 07:23:10"}}
13 {"Wifi":{"SSID":"ChinaCCTV","BSSID":"c8:3a:35:45:61:60","level":-49,"state":"connected","time":"4-14-2015 07:23:10"}}
14 {"Wifi":{"SSID":"","BSSID":"c8:3a:35:45:61:60","level":-46,"state":"connected","time":"4-14-2015 07:23:10"}}
15 {"Wifi":{"SSID":"cmcc-edu-5","BSSID":"1c:fa:68:46:f7:32","level":-43,"state":"not_connected","time":"4-14-2015 07:23:10"}}
16 {"Wifi":{"SSID":"CMCC-Edu","BSSID":"d4:ee:07:0d:17:4e","level":-55,"state":"not_connected","time":"4-14-2015 07:23:10"}}
17 {"Wifi":{"SSID":"FAST_xxtty","BSSID":"28:2c:b2:d4:ab:f2","level":-67,"state":"not_connected","time":"4-14-2015 07:23:10"}}
18 {"Wifi":{"SSID":"cmcc-7","BSSID":"14:75:90:7b:74:38","level":-85,"state":"not_connected","time":"4-14-2015 07:23:10"}}
19 {"Wifi":{"SSID":"Tenda_532BD8","BSSID":"c8:3a:35:53:2b:d8","level":-76,"state":"not_connected","time":"4-14-2015 07:23:10"}}
20 {"Wifi":{"SSID":"CMCC-EDU","BSSID":"28:2c:b2:21:7a:9e","level":-91,"state":"not_connected","time":"4-14-2015 07:23:10"}}
21 {"Wifi":{"SSID":"CMCC-EDU","BSSID":"14:75:90:75:67:92","level":-91,"state":"not_connected","time":"4-14-2015 07:23:10"}}
22 {"Wifi":{"SSID":"TP-LINK_IPTV_7B7438","BSSID":"1a:75:90:7b:74:38","level":-76,"state":"not_connected","time":"4-14-2015 07:23:10"}}
23 {"Wifi":{"SSID":"TP-LINK_IPTV_756792","BSSID":"1a:75:90:75:67:92","level":-91,"state":"not_connected","time":"4-14-2015 07:23:10"}}
24 {"Wifi":{"SSID":"TP-LINK_7AA04A","BSSID":"d0:c7:c0:7a:a0:4a","level":-91,"state":"not_connected","time":"4-14-2015 07:23:10"}}
25 {"Wifi":{"SSID":"ChinaCCTV","BSSID":"c8:3a:35:45:61:60","level":-55,"state":"connected","time":"4-14-2015 07:38:10"}}
26 {"Wifi":{"SSID":"","BSSID":"c8:3a:35:45:61:60","level":-55,"state":"connected","time":"4-14-2015 07:38:10"}}
27 {"Wifi":{"SSID":"cmcc-edu-5","BSSID":"1c:fa:68:46:f7:32","level":-46,"state":"not_connected","time":"4-14-2015 07:38:10"}}
28 {"Wifi":{"SSID":"CMCC-Edu","BSSID":"d4:ee:07:0d:17:4e","level":-64,"state":"not_connected","time":"4-14-2015 07:38:10"}}
29 {"Wifi":{"SSID":"FAST_xxtty","BSSID":"28:2c:b2:d4:ab:f2","level":-91,"state":"not_connected","time":"4-14-2015 07:38:10"}}
30 {"Wifi":{"SSID":"cmcc-7","BSSID":"14:75:90:7b:74:38","level":-70,"state":"not_connected","time":"4-14-2015 07:38:10"}}
```

(a) WiFi 数据 JSON 格式

1 ChinaCCTV,c8:3a:35:45:61:60,-49,connected,4-14-2015 07:23:10	1 c8:3a:35:45:61:60,-49,connected,4-14-2015 07:23:10
2 ,c8:3a:35:45:61:60,-46,connected,4-14-2015 07:23:10	2 c8:3a:35:45:61:60,-46,connected,4-14-2015 07:23:10
3 cmcc-edu-5,1c:fa:68:46:f7:32,-43,not_connected,4-14-2015 07:23:10	3 1c:fa:68:46:f7:32,-43,not_connected,4-14-2015 07:23:10
4 CMCC-Edu,d4:ee:07:0d:17:4e,-55,not_connected,4-14-2015 07:23:10	4 d4:ee:07:0d:17:4e,-55,not_connected,4-14-2015 07:23:10
5 FAST_xxtty,28:2c:b2:d4:ab:f2,-67,not_connected,4-14-2015 07:23:10	5 28:2c:b2:d4:ab:f2,-67,not_connected,4-14-2015 07:23:10
6 cmcc-7,14:75:90:7b:74:38,-85,not_connected,4-14-2015 07:23:10	6 14:75:90:7b:74:38,-85,not_connected,4-14-2015 07:23:10
7 Tenda_532BD8,c8:3a:35:53:2b:d8,-76,not_connected,4-14-2015 07:23:10	7 c8:3a:35:53:2b:d8,-76,not_connected,4-14-2015 07:23:10
8 CMCC-EDU,28:2c:b2:21:7a:9e,-91,not_connected,4-14-2015 07:23:10	8 28:2c:b2:21:7a:9e,-91,not_connected,4-14-2015 07:23:10
9 CMCC-EDU,14:75:90:75:67:92,-91,not_connected,4-14-2015 07:23:10	9 14:75:90:75:67:92,-91,not_connected,4-14-2015 07:23:10
10 TP-LINK_IPTV_7B7438,1a:75:90:7b:74:38,-76,not_connected,4-14-2015 07:23:10	10 1a:75:90:7b:74:38,-76,not_connected,4-14-2015 07:23:10
11 TP-LINK_IPTV_756792,1a:75:90:75:67:92,-91,not_connected,4-14-2015 07:23:10	11 1a:75:90:75:67:92,-91,not_connected,4-14-2015 07:23:10
12 TP-LINK_7AA04A,d0:c7:c0:7a:a0:4a,-91,not_connected,4-14-2015 07:23:10	12 d0:c7:c0:7a:a0:4a,-91,not_connected,4-14-2015 07:23:10
13 ChinaCCTV,c8:3a:35:45:61:60,-55,connected,4-14-2015 07:38:10	13 c8:3a:35:45:61:60,-55,connected,4-14-2015 07:38:10
14 ,c8:3a:35:45:61:60,-55,connected,4-14-2015 07:38:10	14 1c:fa:68:46:f7:32,-46,not_connected,4-14-2015 07:38:10
15 cmcc-edu-5,1c:fa:68:46:f7:32,-46,not_connected,4-14-2015 07:38:10	15 d4:ee:07:0d:17:4e,-64,not_connected,4-14-2015 07:38:10
16 CMCC-Edu,d4:ee:07:0d:17:4e,-64,not_connected,4-14-2015 07:38:10	16 28:2c:b2:d4:ab:f2,-91,not_connected,4-14-2015 07:38:10
17 FAST_xxtty,28:2c:b2:d4:ab:f2,-91,not_connected,4-14-2015 07:38:10	17 14:75:90:7b:74:38,-79,not_connected,4-14-2015 07:38:10
18 cmcc-7,14:75:90:7b:74:38,-70,not_connected,4-14-2015 07:38:10	18 c8:3a:35:53:2b:d8,-88,not_connected,4-14-2015 07:38:10
19 Tenda_532BD8,c8:3a:35:53:2b:d8,-79,not_connected,4-14-2015 07:38:10	19 28:2c:b2:21:7a:9e,-88,not_connected,4-14-2015 07:38:10
20 CMCC-EDU,28:2c:b2:21:7a:9e,-88,not_connected,4-14-2015 07:38:10	20 74:d0:2b:4e:33:a4,-91,not_connected,4-14-2015 07:38:10
21 ChinaNet.edu,74:d0:2b:4e:33:a4,-91,not_connected,4-14-2015 07:38:10	21 1a:75:90:7b:74:38,-70,not_connected,4-14-2015 07:38:10
22 TP-LINK_IPTV_7B7438,1a:75:90:7b:74:38,-70,not_connected,4-14-2015 07:38:10	22 c8:3a:35:45:61:60,-49,connected,4-14-2015 07:53:10
23 ChinaCCTV,c8:3a:35:45:61:60,-49,connected,4-14-2015 07:53:10	23 1c:fa:68:46:f7:32,-49,not_connected,4-14-2015 07:53:10
24 ,c8:3a:35:45:61:60,-49,connected,4-14-2015 07:53:10	24 d4:ee:07:0d:17:4e,-52,not_connected,4-14-2015 07:53:10
25 cmcc-edu-5,1c:fa:68:46:f7:32,-49,not_connected,4-14-2015 07:53:10	25 28:2c:b2:d4:ab:f2,-76,not_connected,4-14-2015 07:53:10
26 CMCC-Edu,d4:ee:07:0d:17:4e,-52,not_connected,4-14-2015 07:53:10	26 14:75:90:7b:74:38,-79,not_connected,4-14-2015 07:53:10
27 FAST_xxtty,28:2c:b2:d4:ab:f2,-76,not_connected,4-14-2015 07:53:10	27 c8:3a:35:53:2b:d8,-88,not_connected,4-14-2015 07:53:10
28 cmcc-7,14:75:90:7b:74:38,-79,not_connected,4-14-2015 07:53:10	28 1a:75:90:7b:74:38,-79,not_connected,4-14-2015 07:53:10
29 Tenda_532BD8,c8:3a:35:53:2b:d8,-88,not_connected,4-14-2015 07:53:10	29 28:2c:b2:21:7a:9e,-91,not_connected,4-14-2015 07:53:10
30 TP-LINK_IPTV_7B7438,1a:75:90:7b:74:38,-79,not_connected,4-14-2015 07:53:10	30 1a:75:90:75:67:92,-91,not_connected,4-14-2015 07:53:10

(b)WiFi 数据标准格式

(c) WiFi 数据标准最终格式

图 4.1 WiFi 数据 JSON 解析

表 4.1 WiFi 数据 JSON 解析核心代码

```
import json
def preprocess_json (fileName,fname):
    f_wifi = open(fileName,'r')
    f_wifinew = open(fname,'w')
    f_wifi_json = []
    for i,f_wifi_line in enumerate(f_wifi.readlines()):
        #将文件中的数据按行取出并去除多余字符
        f_wifi_line = f_wifi_line.strip()
        #JSON 格式解析
        f_wifi_line = json.loads(f_wifi_line)
```

```
#按 Key 取出 Value 具体数据
each_line = str(f_wifi_line['Wifi']['BSSID']) + ',' + str(f_wifi_line['Wifi']['level'])\
+ ',' + str(f_wifi_line['Wifi']['state']) + ',' + str(f_wifi_line['Wifi']['time']) + '\n'
if each_line not in f_wifi_json:
    f_wifi_json.append(each_line)
for each_json_line in f_wifi_json:
    f_wifinew.write(each_json_line)
f_wifi.close()
f_wifinew.close()
```

#### 4.1.2 离散数据的完整性和一致性处理

经过 JSON 解析后的离散数据以二维数组的形式存储为新文件，在新文件中可以明显看到某些数据有空缺值，造成这种情况的原因有很多种，比如说某些收集数据的志愿者没有打开手机软件或者是其手机自动杀死了后台进程等，这样造成数据收集的不完整性。

针对此种情况，需要区分离散数据值的空缺是属于没有数据还是没有采集到数据。例如通话记录，该类传感器只有当有新的通话时才写入新的数据，如果某用户一天都没有接听或打出电话，那么该天没有通话记录文件，这属于没有数据情况，可为之赋值为 0。但其他数据比如位置信息，在 StarLog 中 GPS 的采样频率设置为每 30 秒一次，所以每小时会有 120 个数据，如果记录文件中出现大量数据空缺，则一定是属于没有采集到数据，程序出现异常。这种情况下就要为空缺值赋值并且赋值方法的选择要合理可靠，以保证数据的完整性和一致性。在数据处理过程中，我们采用了临近值赋值方法，即用缺失值前后数据的平均数为其赋值。因为用户行为的变化连续的、渐变的，所以缺失值前后采集的数据原则上与空缺值差异不会太大并呈现出一定规律性，以此赋值能较好地保证数据的完整性和一致性。

#### 4.1.3 连续数据的离散化处理

机器学习的分类算法基本都是针对离散数据进行分类，因此以光传感器、加速度计、陀螺仪、磁力计为代表的连续数据在预处理的过程中必须进行离散化处理。

由于不同厂家、不同型号的安卓手机内置的这四类传感器的采样频率并不相同，多则每秒采样几十次，少则每秒采样几次，为了对各种手机的硬件传感器数据进行统一的滑动窗口分割，需要先对这四类数据进行一个简单的均值处理，具体表现为以秒为单位对同秒内所有经过滤噪的数据取均值得到的数值作为该秒的最终数据。然后根据 2.2.1 小节中对滑动窗口数据分割的介绍，使用

窗口长度 10s 的具有 50%重叠的滑动窗口进行分割。

## 4.2 运动类感知数据的特征提取

运动类感知数据包括加速度计、陀螺仪、磁力计和活动状态这四类数据，总共提取了 84 个特征值。其中，加速度计、陀螺仪和磁力计的数据收集信息包括时间以及相应的三轴数据；活动状态的数据收集信息包括开始时间、结束时间以及这段时间内的状态(平稳、缓慢或快速)。

### 4.2.1 加速度计、陀螺仪和磁力计的特征提取

因为加速度计、陀螺仪和磁力计收集的数据格式相同，都是时间和相应时刻的  $x$ ,  $y$ ,  $z$  三轴数值，因此可以采用相同的特征值表示并采取同一种特征提取方式。

经过 4.1 节的预处理，我们得到了长度 10s 的具有 50%重叠的数据样例集，然后对每一个长度 10s 的样例集都提取最大值、最小值、均值、标准差、波形个数、波峰均值、波谷均值、波峰与波谷差值的最大值、波峰与波谷差值的最小值这 9 个特征值；又因为这三种传感器都是三轴传感器，因此每种传感器可以提取 27 个特征，总共提取 81 个。

### 4.2.2 活动状态的特征提取

活动状态的数据是根据加速度传感器进行简单计算得出，可以判断出用户的具体状态。该传感器以每分钟一次的频率对数据进行计算得出用户在这一分钟内的状态是处于平稳、缓慢或是快速。因此，我们以用户分别处于静止、慢速和快速的时间占总时间的比例作为活动状态的特征值，总共提取 3 个特征值。

## 4.3 环境类感知数据的特征提取

环境感知数据包括光传感器数据和 GPS 数据两类，总共提取了 6 个特征值。其中，光传感器数据收集信息包括时间以及相应的光强度；GPS 的数据收集信息包括经度、纬度、海拔高度、GPS 速度、时间、精确度、定位方式（GPS 定位或网络定位）。

### 4.3.1 光传感器的特征提取

从数据集中随意选取一天光传感器数据，以时间为  $x$  轴，光的强度为  $y$  轴

绘制光传感器数据，如图 4.2 所示。绘制成图后，可以发现 y 轴数据数值大小差异明显，基本分为三档，弱光、灯光和强光。通过和其他光传感器文件数据进行对比，基本可以确定弱光、灯光和强光的取值大致范围。当手机放置在衣物口袋、包中或在黑暗情况下使用，数值较小，大多集中在 0~100 之间；当在室内时，大多数是受到灯光照射的情况，数值多集中在 100~1000 之间；当处于户外时，受到太阳直射，采集到的数值很大，多集中在 1000~20000 之间。根据这些特点，我们将光传感器数据进行分类并以光度值 100 和 1000 作为分类标准，光强度低于 100 视为弱光，光强度大于 1000 视为强光，光强度处于两者之间视为灯光。因为手机只有在使用时光传感器采集的数据数值大于 0，因此光传感器数据数值不为 0 的时间等同于使用手机的时间，我们可以提取出 4 个特征，用数据值为 0 和处于弱光、灯光、强光下的时间比例分别表示手机未被使用、在弱光下（黑暗中）使用、在室内使用以及在户外使用的比例。

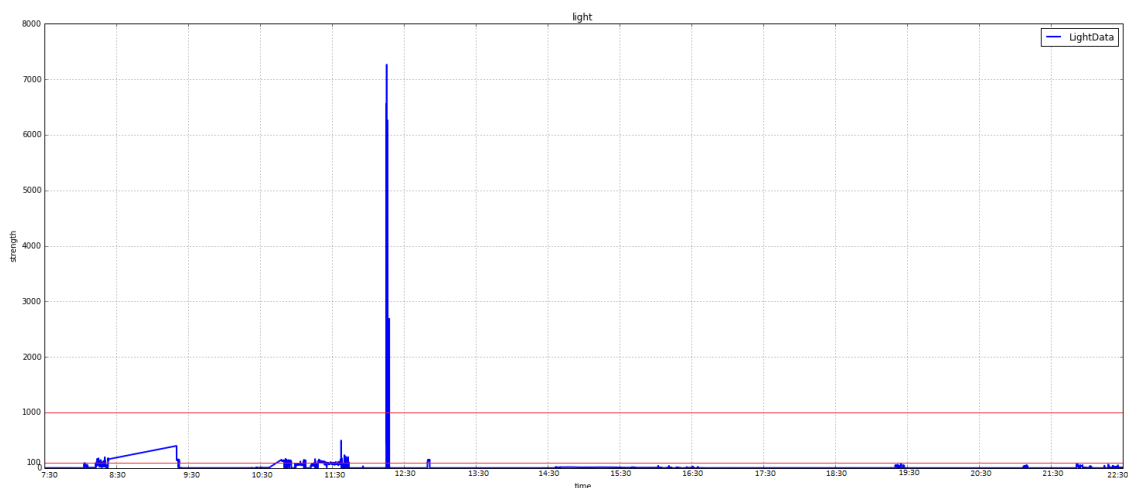


图 4.2 光传感器数据图

### 4.3.2 GPS 的特征提取

我们希望通过 GPS 信息获知用户每天去过的位置的数量以及在每个位置停留时间的长短，并以此作为用户活跃度的一个判断依据。因此，为了更准确地记录用户的行动轨迹，GPS 采用高频进行数据采集(每 30s 记录一次)。收集的数据经过 JSON 解析后得到每个采样点的纬度和经度，并以它们分别作为 x 轴和 y 轴对数据进行绘制，结果如图 4.3 所示。可以明显看出，图 4.3 中有三个成型的类簇，说明用户在这三个点停留时间较长，因此数据采样足够多；但实际上用户是去过四个位置的，最后一个位置由于停留时间较短，采样数量不够聚集成类，但是可以清晰的看出从图中位置②到位置④的路径。

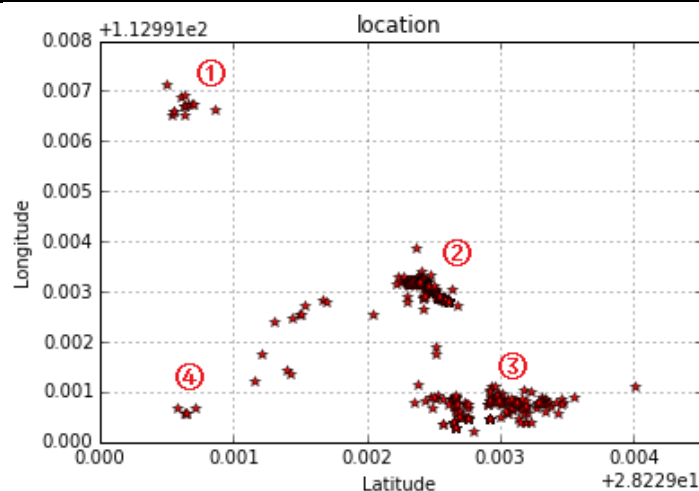


图 4.3 GPS 原始数据图

在本文中采用 K 均值聚类算法(K-means)<sup>[71]</sup>对数据进行聚类分析, K 均值算法以其简洁和高效的特点成为所有聚类算法中使用最为广泛的一种, 这也是本文选择使用它的原因。K-means 作为一种基于距离的典型聚类算法, 以距离作为相似性的评价标准, 该算法认为距离越近的两个对象其相似度越大, 被分到同一个簇的概率也越高。在对数据进行分类前, 需要给定聚类的数目 K 和需要聚类的数据集合, 其中 K 由用户指定。算法开始时随机选取 K 个数据对象作为初始聚类中心; 然后对剩余的数据计算其到每个中心的距离, 并把它归到距离最短的聚类中心中; 接下来不断重复计算各个类的聚类中心并对数据进行分类直到每个聚类稳定、不再发生变化为止, 此时算法已经收敛。图 4.4 就是当  $K = 4$  时的聚类结果, 通过对不同类簇的不同颜色标记, 可以明显看出每个点属于的类别。

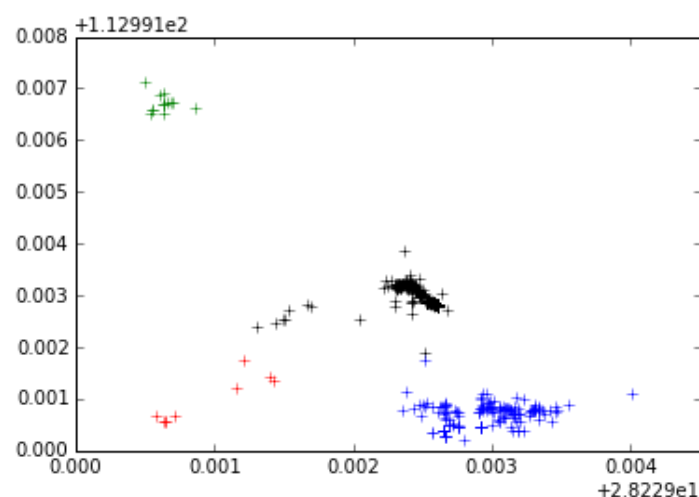


图 4.4 GPS 数据分类图

为了衡量不同用户每天在不同位置停留时间的长短, 我们采用信息熵  $H(u)$

进行计算。信息熵  $H(u)$  表示信息出现的概率或不确定性，可以看做是对信息的有序化度量，计算公式如(4-1)所示：

$$H(u) = -\sum_{i=1}^n p_i \cdot \log_2 p_i \quad (4-1)$$

其中  $p_i$  为用户数据属于第  $i$  个聚类中心的数量占总记录数据数量的比例。根据公式可知，当数据采集的位置越多、在每个位置停留的时间越接近则熵  $H(u)$  越大，表示数据越混乱，即用户出现在某一个位置的不确定性越大；数据采集的位置越少、在每个位置停留的时间差别越大则熵  $H(u)$  越小，表示数据越有序，用户出现在某一个位置的不确定性小，即能准确预测出用户位置。假如某个用户一天都在家中，经过聚类后只有一个位置，那么熵就是 0，表示数据处于一种有序状态；若用户坐火车前往某地，因为位置一直在变化，熵值会很大，表示数据处于一种无序的混乱状态。

因此，根据表示位置信息的 GPS 数据提取出的特征值有 2 个，分别为经过聚类的位置个数以及表示数据出现概率的信息熵。

## 4.4 手机使用数据的特征提取

手机使用记录的数据可以分为社交信息、手机使用信息和手机状态信息三类，总共提取了 27 个特征值。

### 4.4.1 社交信息的特征提取

社交信息包括通话记录、短信记录、微信、QQ 消息和蓝牙。其中，通话的数据收集信息包括时间、经过加密处理的电话号码、通话时长和通话类别（接听、拨出和未接）；短信数据收集信息包括时间、通信类别（接收或发送）、经过加密处理的通信号码和通信号码状态（数字或汉字）；微信、QQ 的数据收集信息包括时间、通信类别（接收或发送）、经过加密处理的通信 ID；蓝牙的数据收集信息包括手机名称、蓝牙 MAC 地址、连接状态和时间。

通过对通话记录信息中电话号码的统计可以计算出不同时间段最频繁通话的号码，因此以通话个数，通话时长，最频繁通话个数占总个数的比例，最频繁通话时长占总时长的比例 4 个值为特征。

短信、微信和 QQ 消息三类数据性质相似，并且日志文件的格式也大致相同，因此特征提取方法类似。短信记录的第四个属性主要是用来区分信息是否为广告短信，如果同时满足以下四个条件就认为此条短信属于广告短信，相关数据将不会作为影响情绪的因素被统计：①通信号码并不是 11 位，国内手机号码统一为 11 位，多于或少于 11 位都不是手机发送；②通信号码状态以数字



形式存在，这表示用户并未存储该号码；③通信类型为接收；④用户并未回复该信息，即用户和此通信号码间并未有一个完整地会话过程。

通过对短信记录第四个属性筛选的结果将作为识别情绪的数据与微信和 QQ 消息一起提取特征。与通话记录相类似，这三类数据每种提取的特征为 2 个，总共有 6 个，包括信息条数和最频繁联系人发送消息总数占总条数的比例。

蓝牙记录文件中的第一个属性手机名称因为包含中文不能被 JSON 解析，但是蓝牙 MAC 地址可以唯一标识手机，因此采用蓝牙 MAC 地址区分不同的手机。蓝牙信息可以提取 2 个特征，分别为手机检测到的蓝牙个数以及其中连接过的个数，其中手机检测到的蓝牙个数能够表示出用户所处的人群密度并且和用户连接过的手机持有者明显亲密度高于其他人。

在社交信息部分，五种传感器总共提取 12 种特征。

#### 4.4.2 手机使用信息的特征提取

手机使用信息包括 WiFi、应用和划屏。其中，WiFi 数据收集信息包括时间、SSID、BSSID、信号强度以及连接状态；应用的数据收集信息包括开始时间、结束时间以及进程名；划屏的数据收集信息包括屏幕开启时间、屏幕解锁时间、屏幕关闭时间。

与蓝牙记录文件相似，WiFi 文件中的属性 SSID 也出现了不能被 JSON 解析的情况，不过同一个 SSID 由于覆盖面广泛存在具有几个 BSSID 的可能性，而 BSSID 作为无线网络接入点的 MAC 地址具有唯一性，可以清楚地分辨出用户所在的网络，因此采用 BSSID 属性值作为区分无线网络的标准。针对 WiFi 文件提取了 2 个特征，分别为手机周边 WiFi 个数以及连接过的个数。

应用记录文件中记录了手机使用过的所有进程和时间，因为这些进程中除了包含用户使用的 APP 程序还包含了一些系统应用和手机运营商提供的应用，这些程序手机自动运行，并不属于用户使用记录，因此需要过滤掉。对过滤掉系统应用和运营商应用的数据，我们以程序的使用时间占总时间的比例作为特征值提取，用来衡量用户使用手机程序的频率，这个数值可以间接反映出用户对手机的依赖程度。

通常情况下用户都会设置屏幕锁，屏幕亮后需要对其解锁，因此文件会记录屏幕开启、屏幕解锁、屏幕关闭三种数据；对于不设置屏幕锁的用户，只会记录屏幕开启和屏幕关闭时间。但这两者情况中手机使用时间计算方式一致，都为屏幕开启时间与屏幕关闭时间的差值，因此我们对划屏信息提取了 3 个特征，分别为屏幕亮、屏幕暗的时间比例以及解锁次数。屏幕亮的时间比例反映出用户使用手机的时间长短，与应用记录文件相呼应；而解锁次数的频繁性也

能侧面反映出情绪的稳定性。

在手机使用信息部分，三种传感器总共提取 6 种特征。

4.4.3 手机状态信息的特征提取

手机状态信息是对手机的一些基本信息进行采集，包括时间，铃声模式(铃声开、铃声关)，飞行模式(飞行模式开、飞行模式关)，网络连接模式(数据流量连接、WiFi 连接或者未连接)，充电状态(充电、未充电)五种数据。根据各类数据计算手机处于不同模式的时间比例，共提取了 9 个特征。

4.5 数据特征归纳

本节主要是对 4.2 节至 4.4 节所描述的特征提取结果作以归纳总结。如表 4.2 所示，运动类感知数据从加速度计、陀螺仪、磁力计、活动状态四类传感器中共提取了 84 个特征，环境感知类数据从光传感器和 GPS 中提取 6 个特征，手机使用数据分别从社交信息、手机使用信息、手机状态信息三个方面九种传感器中提取 27 个特征，总计提取 117 个特征。

表 4.2 传感器数据提取特征表

数据类型		传感器	特征数量	特征总数	
运动类感知数据		加速度计	27	84	
		陀螺仪	27		
		磁力计	27		
		活动状态	3		
环境感知类数据		光传感器	4	6	
		GPS	2		
手机使用数据	社交信息	通话记录	4	12	27
		短信记录	2		
		微信	2		
		QQ 消息	2		
		蓝牙	2		
	手机使用信息	WiFi	2	6	
		应用	1		
		划屏	3		
	手机状态信息	手机基本信息	9	9	

4.6 数据归一化

根据 4.5 节所描述，经过对 StarLog 收集的数据的处理，我们一共提取了

117 个特征，这些特征具有不同的量纲和量纲单位，这往往会影响到数据分析的结果和程序运行时的收敛速度。为了消除因为量纲引起的问题，需要对数据进行标准化处理。数据归一化就是一种数据标准化处理，经过归一化处理，数据各指标都将处于同一数量级，即统一到同一个参考系下，更适合进行综合比较或计算。数据归一化有两种常用的方式，一种是离差标准化(Min-Max Normalization)方法，另一种是 Z-score 标准化方法。

Min-Max 方法是指通过对原始数据的线性变化，使变化后的结果值映射到区间[0, 1]之间，转换函数如(4-2)所示：

$$x^* = \frac{x - \min}{\max - \min} \quad (4-2)$$

其中 min 为原始数据样本中的最小值，max 为原始数据样本中的最大值。这种方法存在一个问题，就是当样本中加入新数据时可能导致最大值 max 和最小值 min 发生变化，需要重新定义具体数值。

Z-score 标准化方法需要对原始数据样本的均值(mean)和标准差(standard deviation)进行计算，然后通过公式(4-3)对数据进行转换，转换后的数据符合均值为 0，标准差为 1 的标准正态分布。

$$x^* = \frac{x - \mu}{\sigma} \quad (4-3)$$

其中  $\mu$  为原始数据样本的均值， $\sigma$  为原始数据样本的标准差。

## 4.7 本章小结

本章就多维移动数据的处理机制展开了研究，4.1 节对数据的预处理过程进行了详细的描述，其中包括对原始数据的 JSON 解析、离散数据的完整性和一致性处理以及连续数据的离散化处理。4.2 节、4.3 节和 4.4 节分别从运动类感知数据、环境类感知数据、手机使用数据三个角度对数据进行分类并进行特征提取，总共提取了 117 个有效特征。4.5 节就数据归一化进行讲解，通过对已提取特征的归一化处理，将数据都统一到同一个参考系下，便于数据分析，同时可以加快程序的收敛速度。

## 第五章 多维移动数据的用户情绪识别

第四章描述了多维移动数据的预处理、特征提取和归一化过程，本章将着重描述使用经过处理的数据进行用户情绪识别的实验过程以及结果分析。本章首先对数据集进行分析，并对评估基准进行了描述；其次，采用支持向量机(SVM)、随机森林(Random Forest)、梯度树提升(GTB)等六种分类算法，从特征融合以及结果融合两个角度进行情绪识别，并对实验结果进行分析和说明。

### 5.1 数据集、情绪识别方法和评估基准

#### 5.1.1 数据集

为了验证本文所提出的采用多维移动数据进行情绪识别的有效性，我们招募了 14 名研究生作为志愿者（包括 8 名男性和 6 名女性），进行连续一个月的数据收集工作（全部使用安卓手机）。在数据开始收集前为这 14 名志愿者进行了数据收集软件和情绪记录软件的安装与使用培训，以确保每个人收集到的数据尽可能有效、可利用。作为对这 12 名志愿者的回报，在实验开始时给每人发放了一个容量为 32GB 的手机存储卡。

实验使用 StarLog 数据收集平台采集每天上午 7:30 至夜间 22:30 总计 15 个小时的数据，同时将每天记录的三次情绪（每间隔 5 小时记录一次）分别与上午、下午和晚上的数据相对应，作为已提取特征的标签，便于后期进行学习训练，如图 5.1 所示。其中使用数字 1~5 分别代表高兴、悲伤、恐惧、生气、中性；同样使用数字 1~5 分别代表愉悦度和活跃度的五种级别，数字越高表示愉悦度和活跃度越高。

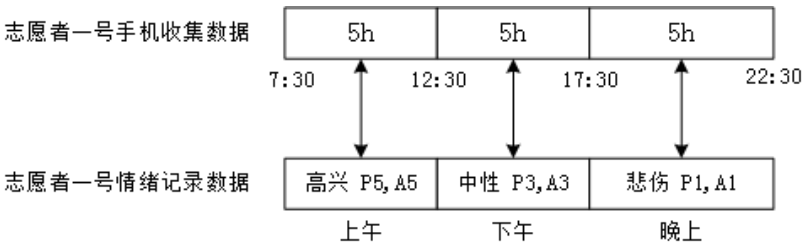


图 5.1 手机数据与情绪记录数据映射图

我们采用 14 名志愿者连续一个月的数据，总计 1260 条(以每 5 个小时采集的数据整体作为一条记录)作为实验数据集。图 5.2 显示的是某志愿者从 4 月 1 日至 4 月 10 日之间的数据容量分布饼图，这里的数据容量指的是志愿者每天记录的所有日志文件大小总和，单位为 MB。从图中可以看出该志愿者在这段时间内每天记录的日志文件大小相对稳定，维持在 70MB 左右。

但通过统计不同志愿者每天的日志文件，发现不同志愿者同一天的数据总量差别明显，统计结果如表 5.1 所示。经过分析得知，除却因个人使用手机较少而导致的数据采集量少这个因素，单天数据量相对较大的用户和单天数据量相对较少的用户的数据量差值主要在于光传感器、加速度计、陀螺仪和磁力计这四类硬件传感器收集的文件日志大小，数据量大的用户这四类连续数据的采样频率偏高，因此整体日志文件的容量偏大，数据量小的用户这四类连续数据的采样频率偏低，因此整体日志文件的容量偏小。又因为这四类硬件传感器的采样频率均为手机制造商设定，因此不同手机厂商生产出的手机存在采样频率不一致且差距较大的可能性，日志记录文件大小也就差异较大，由此可见 4.1 节数据的预处理中对硬件传感器采集的连续数据进行均值处理的重要性。

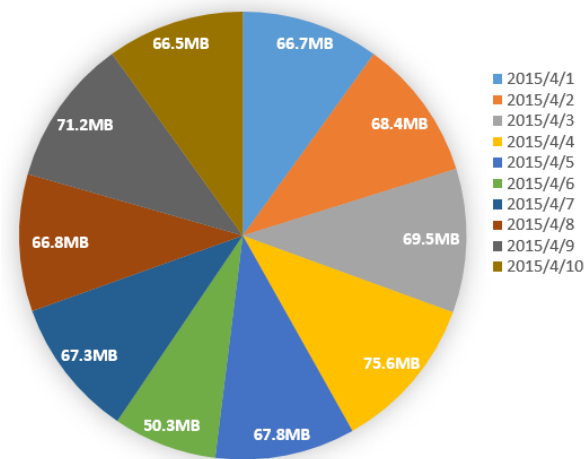


图 5.2 某志愿者数据容量分布

表 5.1 不同志愿者同一天数据容量统计

志愿者序号	数据容量
志愿者 1	74.6MB
志愿者 2	86.1MB
志愿者 3	246MB
志愿者 4	191MB
志愿者 5	169MB
志愿者 6	72MB
志愿者 7	141MB
志愿者 8	47.1MB
志愿者 9	2.41MB
志愿者 10	66.8MB
志愿者 11	85.6MB
志愿者 12	53.4MB
志愿者 13	174.6MB

志愿者 14

77.2MB

经过统计,我们发现部分用户的数据存在缺少和丢失的现象严重,尤其是日志文件中没有加速度等数据的用户,每天的日志记录文件容量只有几 MB 甚至几百 KB。调查发现,数据收集过程中存在以下问题:首先,WiFi、蓝牙和 GPS 需要每天开启 15 个小时,以及数据收集软件的持续收集造成了耗电量高的现象,因此当不能及时充电时,部分用户会选择暂时停止数据的收集,这就造成了传感器数据的缺失;其次,用户并没有对记录情绪这一行为形成习惯,可能有时忘记记录,造成情绪标签的缺失;最后,存在部分手机中安装的安全卫士等软件将 StarLog 后台进程杀死,造成强制终止的情况出现。这三个原因造成了部分数据的不完整,经过筛选,我们最终选出 1027 条有效数据作为数据样例。

### 5.1.2 情绪识别方法和评估基准

在有关情绪识别的相关研究中,不同的研究人员选用了不同的分类器进行情绪识别,比如在文献[72]中使用了支持向量机(SVM)对情绪(Emotion)进行识别,文献[73]和[74]中使用随机森林(Random Forest)对压力(Stress)程度和幸福感(Happiness)进行判断。学术界如今并未对情绪识别中究竟哪种分类算法更好得出统一的结论,因此本文采用了在情绪识别和机器学习领域中较为广泛使用的六种分类算法,分别是支持向量机、k-近邻、决策树、AdaBoost、随机森林和梯度树提升。

在分类实验中,以用户的情绪记录结果作为相应时间段手机数据的情绪标签,即本实验的评估基准(ground truth)。随机选择有效数据样例的 80%作为训练数据,其余 20%作为测试数据。后文实验结果中的准确率指的是情绪识别的平均准确率,即测试数据中分类结果与实际标签相符的数据个数占全部测试数据的百分比。

为了验证环状分类模型中愉悦度和活跃度的相关性,我们对通过情绪工具收集的、用户报告的情绪状况进行统计分析,如表 5.2 所示,计算方法为某一愉悦度和活跃度的记录条数占总记录条数的比例。表中 P1 ~ P5 分别代表非常不愉悦、不愉悦、中性、愉悦、非常愉悦五种愉悦度级别,A1 ~ A5 分别代表非常不活跃、不活跃、平静、活跃、非常活跃五种活跃度级别。根据统计结果,发现用户情绪处于(P3, A3)的概率最大,其次是(P4, A3)和(P4, A4),这表明用户大多数情况下处于平静状态,并且高兴的概率要远远大于悲伤的概率,与实际生活经验相吻合。其中(P1, A4)、(P1, A5)、(P5, A1)、(P5, A4)的概率为 0,充分表明愉悦度和活跃度的统一,二者呈正相关关系。因为本文主要探讨用户

的情绪，因此后文中如无特殊提及，环状情绪模型的准确率都是指愉悦度的准确率。

表 5.2 愉悦度和活跃度相关性分析

<div>活跃度</div> <div>愉悦度</div>	A1	A2	A3	A4	A5
P1	0.23%	0.22%	0.23%	0.00%	0.00%
P2	0.00%	6.67%	2.07%	0.23%	0.46%
P3	0.46%	4.83%	46.7%	1.61%	0.00%
P4	0.69%	1.38%	18.85%	11.03%	0.69%
P5	0.00%	0.00%	0.46%	1.61%	1.61%

5.2 基于特征融合的情绪识别

5.2.1 特征级融合

随着信息技术的快速发展，传感器的研发和使用也逐渐开始受到重视，越来越多的研究已经开始利用多个传感器对数据进行全方位、多角度的采集，以期能获取到更加全面和完整的信息。对于多传感器系统来说，信息的多样性和复杂度较高，数据量大，信息融合技术通过对多方面信息的规律性组合从而推导出更全面和更有价值的结果，其重要性不言而喻。本节将采用信息融合技术中的特征级融合对信息进行处理。特征级融合是由每个传感器先提取出自己的特征向量，然后再对所有的特征向量进行融合的一种方法，具有处理速度快、可实时处理等优势，下面将对特征级融合进行详细描述。

如第四章中所述，本文共提取了 15 种数据的 117 个特征，为了更好地进行情绪识别工作，我们将所有数据的特征以等权重的方式进行特征拼接并融合成为一个 117 维的新特征，具体融合过程如图 5.3 所示。其中，应用数据提取的特征向量为 1 维向量  $A_1$ ，例如 4 月 30 日志愿者二号的应用数据特征向量为 0.1202；通话数据提取的特征向量为 4 维向量  $(A_2, B_2, C_2, D_2)$ ，例如 4 月 30 日志愿者二号的通话数据特征向量为(0.057154, 0.010875, 0.039125, 0.033103)；以此类推直到最后一个传感器数据手机状态数据提取的特征向量为 9 维向量  $(A_{15}, B_{15}, C_{15}, D_{15}, E_{15}, F_{15}, G_{15}, H_{15}, I_{15})$ 。通过将它们等权重拼接成为新的 117 维特征向量 $(A_1, A_2, B_2, C_2, D_2, \dots, A_{15}, B_{15}, C_{15}, D_{15}, E_{15}, F_{15}, G_{15}, H_{15}, I_{15})$ 获得训练样例的输入样本，经过支持向量机、k 临近、决策树、随机森林、AdaBoost、梯度树提升六种分类器的训练，就能得到基本情绪模型和环状情绪模型的识别准确率，具体算法描述如表 5.3 所示，特征融合部分结算结果如图 5.4 所示。

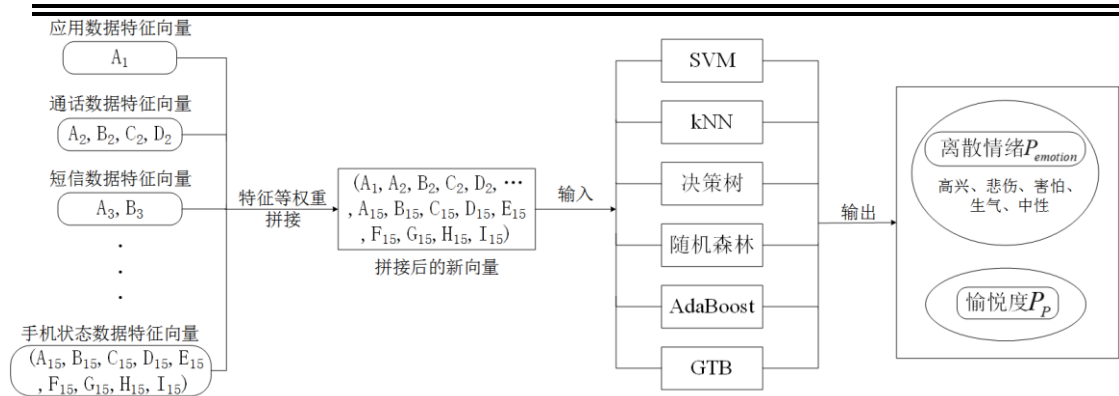


图 5.3 特征融合过程

表 5.3 多传感器数据的特征融合算法描述

<b>输入：</b> 1、文件存储路径； 2、各个传感器特征提取文件；
<b>输出：</b> 经过融合的 117 维新特征文件
<b>算法描述：</b> Step1: 设置标准日期匹配格式； Step2: 对输入的文件存储路径内符合标准日期格式的文件根据前文所述的传感器顺序计算特征值； Step3: if 某个传感器特征值 = null 成立： if 传感器 = 蓝牙 or 电话记录 or 短信记录 or QQ 记录 or 微信记录 or WiFi 记录 成立： 将特征值赋值为 0； else： 采取临近值赋值方法，即用缺失值前后数据的均值为其赋值； Step4: 将计算过的特征值依次写入特征文件进行等权拼接并进行存储； Step5: 重复 Step2 至 Step4 直至遍历并完成文件夹内所有文件特征值计算； Step6: 将融合后的新特征值作为输入值分别输入支持向量机、k 临近、决策树、随机森林、AdaBoost、梯度树提升六种分类器中进行训练，得出识别准确率。

121	5-22-2015	am,-0.163416803154,-1.48755206265,6.59030300389,1.74026598206,24.3477894821,21.9962014956,1.31919141221,
122	5-22-2015	pm,0.0725214229223,1.92249674047,-2.90787887119,7.49159739633,15.8664066277,66.0935770959,2.73707825908,
123	5-22-2015	nit,2.7046281002,1.30952428876,3.55378949913,20.5111424421,18.0340301656,37.1386572416,4.5289228788,4.24
124	5-23-2015	am,1.30380466843,0.596298654023,0.17233081628,12.6674172022,3.96070682839,78.2017036759,3.55913152359,1.
125	5-23-2015	pm,-0.0910873209607,-0.246110681267,8.44378679598,6.29853236365,5.14914596019,7.42787219296,2.5096877024
126	5-23-2015	nit,0.197318423488,2.02287457204,7.54094781969,2.5624423358,14.5094422542,12.4565175847,1.60076304799,3.
127	5-24-2015	am,0.486002723209,1.48564260693,7.90739913663,6.15725317385,6.30124738416,11.2504251597,2.4813813036,2.5
128	5-24-2015	pm,5.18160439926,0.18672517507,7.4624675043,3.28229177523,0.383813308554,0.314510427548,1.81170962774,0.
129	5-24-2015	nit,1.44471542925,0.48854316142,7.17916361494,19.5481099882,3.01567159246,14.3046059881,4.42132446086,1.
130	5-25-2015	am,0.383225272649,0.847147393558,6.19555198606,6.21139203506,27.1608883248,22.688396739,2.49226644544,5.
131	5-25-2015	pm,1.09810845161,2.90631537701,5.08603771951,24.1356840923,11.5028080118,22.625802585,4.91280816767,3.39
132	5-25-2015	nit,0.402493599184,-0.21365956671,8.75461578636,4.39023151802,4.20691631192,4.83777440305,2.09528793201
133	5-26-2015	am,-0.0317598319236,-2.01526867663,6.79745161517,1.05110548749,22.2338617629,20.562311123,1.02523435735,
134	5-26-2015	pm,0.22112523073,-0.664443602293,-3.64375080683,1.50647179504,5.22225774464,79.4051567641,1.22738412693,
135	5-26-2015	nit,-0.424760514238,0.250963275617,8.04443677746,2.7756899689,6.67975522382,16.3022550077,1.66604020627,
136	5-27-2015	am,0.0859846683018,-1.14582178557,6.98040665301,1.6066631592,26.4598350195,15.5685185572,1.26754217255,5
137	5-27-2015	pm,0.263720236976,0.25142984517,7.96330067414,2.01367656255,10.7026810185,13.3873148239,1.41904071913,3.
138	5-27-2015	nit,-0.366778888998,-0.0647162746686,8.67308609027,1.87325068882,6.34289024853,5.33011938746,1.368667486
139	5-28-2015	am,-0.477795771409,2.9940046281,4.40252232588,3.53359103411,20.0530178679,42.8334198816,1.87978483718,4.
140	5-28-2015	pm,0.873320606879,0.0942829742539,6.15794483126,9.77185013199,21.0958705831,23.4257424258,3.12599586244,
141	5-28-2015	nit,-0.294138168033,0.930455536189,8.46799446073,2.00118205712,6.75435957732,6.08166489504,1.41463142094
142	5-29-2015	am,-0.0225441975374,-0.164681270351,5.56169851087,2.4667580396,16.159707347,44.5880988345,1.57059162089,
143	5-29-2015	pm,-0.575054734026,1.09313656619,4.47379639563,21.9907288866,19.231803711,33.4328205889,4.68942735167,4.
144	5-29-2015	nit,0.186970701598,0.304979713328,7.61822998694,1.11973588217,13.8525527956,17.6927398603,1.05817573312,
145	5-30-2015	am,0.850140211996,2.80823270557,-1.71673545054,19.6262797462,12.6712435378,56.1302365056,4.43015572483,3

图 5.4 特征融合部分结果



### 5.2.2 实验结果与分析

在实验验证部分，我们将数据分为两类进行对比实验，一类是由全部志愿者数据构成的混合数据集，另一类是由个人数据较完整的志愿者数据构成的个人数据集。我们将采用两类不同的数据分别对用户情绪进行基于特征融合的识别，并对识别结果进行比较和分析。

首先，使用混合数据进行分类识别，实验结果如图 5.5 所示。从图中结果可以看出，当采用基本情绪模型作为情绪分类标准时，随机森林分类器识别的准确率最高，可以达到 65.91%；采用环状情绪模型作为情绪分类标准时，也是随机森林的识别率排在首位，可以达到 72.73%。

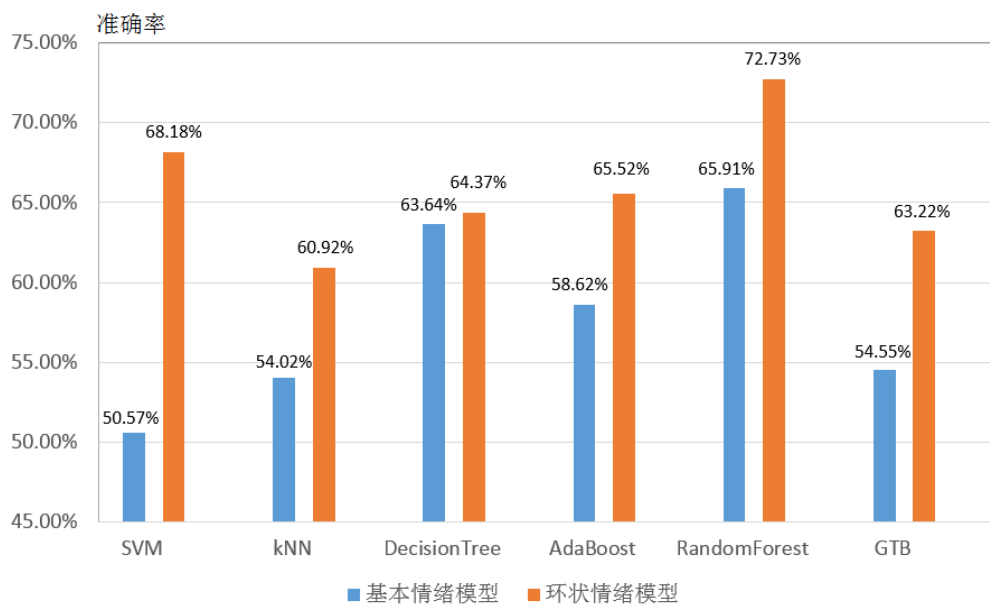


图 5.5 混合数据特征融合情绪识别准确率

通过对比两种分类模型准确率，我们发现在混合数据试验中六种分类器使用环状情绪模型识别的结果都要比基本情绪模型准确率更高。对此实验结果，我们认为环状情绪模型比基本情绪模型识别率更高的原因在于连续的情绪是一种模糊的衡量方式，它比具体地将情绪划分到某一类更人性化，便于用户度量 and 选择。在后期回访中，一部分志愿者也曾经表示有时需要记录的情绪并不在备选的五种基本情绪中，不知道该如何选择，因此存在影响准确率的可能性。关于如何划分类型才能更全面衡量情绪的问题一直是学术界争论的焦点，这同时也是离散情绪识别过程中需要克服的一个难题。

然后使用个人数据进行分类，将 14 名志愿者的全部个人数据分别进行训练后发现，由于每名志愿者最多有 30 天的数据量，因此部分数据缺失对实验结果造成的影响较大，据此原因我们选取了其中数据量最为完整的两名志愿者数据

进行分析，这两名志愿者每人连续 30 天共计 90 条 450 小时的数据记录完整，是参与实验的 12 名志愿者中有效数据最多的两人，实验结果如图 5.6 和图 5.7 所示。其中一人具有 2.65GB 数据，另一人具有 5.53GB 数据，二人因手机硬件传感器采样频率不同导致数据收集量相差较大，但数据记录完整，不会影响分析结果。

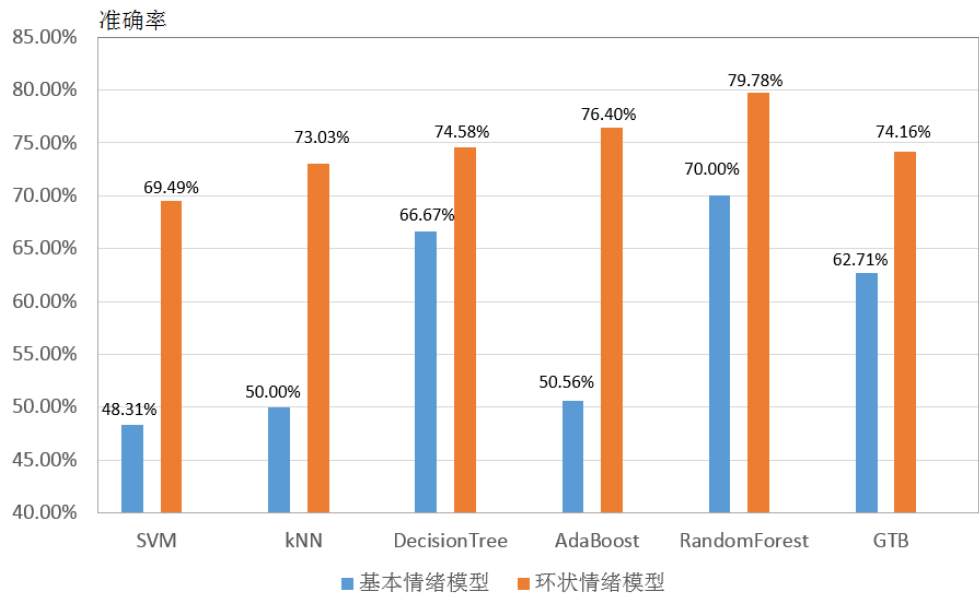


图 5.6 志愿者三号个人数据特征融合情绪识别准确率

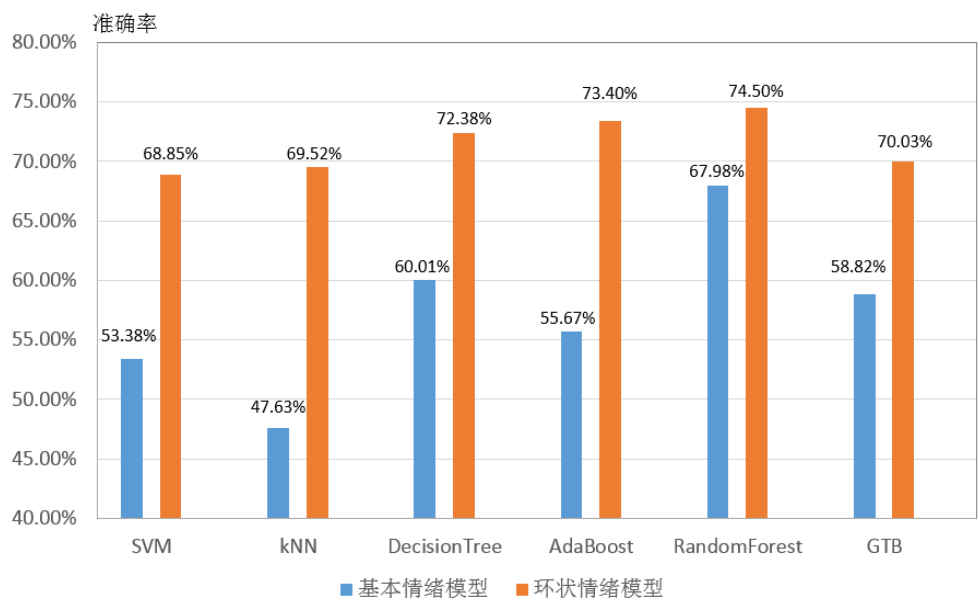


图 5.7 志愿者十二号个人数据特征融合情绪识别准确率

通过对比发现在这两幅图中个人数据识别结果与混合数据识别结果均呈现出相似性，随机森林依然是识别率最高的分类器，对基本情绪模型和环状情绪模型的识别最高可以达到 70.00%和 79.78%；并且使用在个人数据的对比实

验中环状情绪模型识别的结果也比基本情绪模型准确率更高,甚至在某些情况下基本情绪模型的识别准确率并不理想。

通过对比混合数据、个人数据在基本情绪模型下的识别率以及混合数据、个人数据在环状情绪模型下的识别率(如图 5.8 和图 5.9 所示),我们发现,不论是在基本情绪模型中还是在环状情绪模型中个人数据的最高识别率都比混合数据高,并且分类效果最好的分类器为随机森林;在基本情绪模型中,少数分类器存在混合数据识别准确率比个人数据更高的情况,而环状情绪模型中的所有分类器分类结果都是个人数据比混合数据准确率更高。

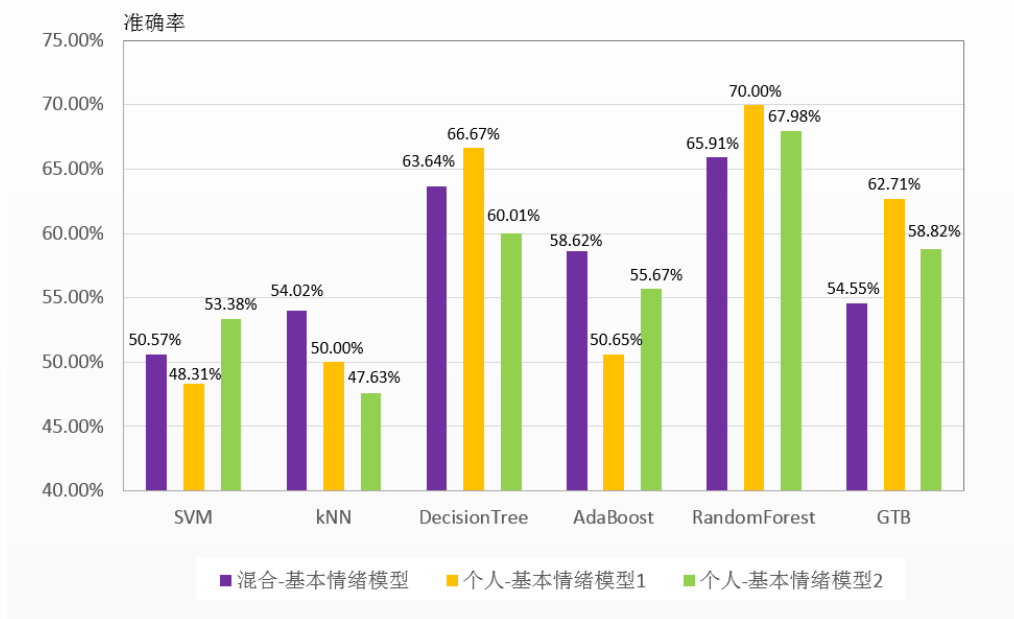


图 5.8 混合数据和个人数据特征融合的基本情绪模型对比

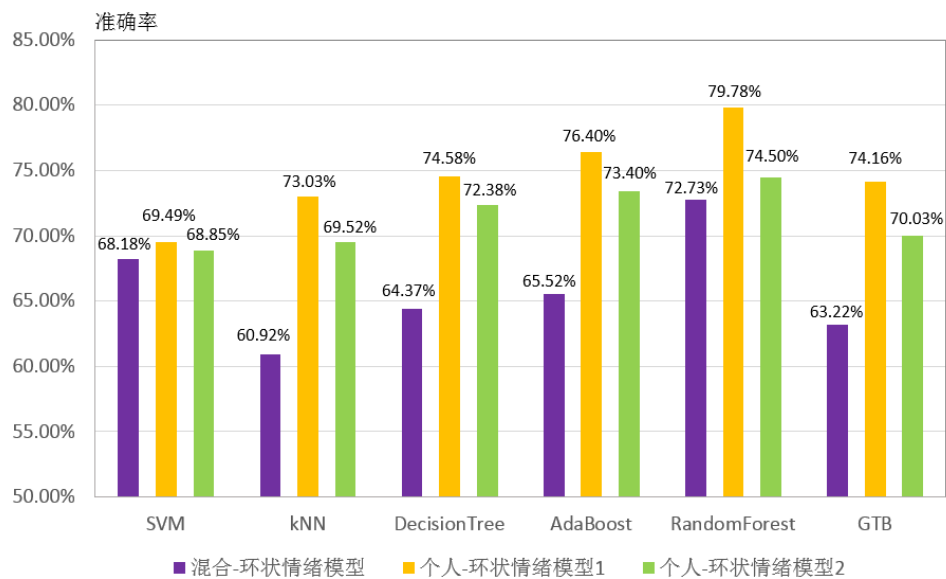


图 5.9 混合数据和个人数据特征融合的环状情绪模型对比

对此,我们认为每个用户的行为模式都有其独特的特点,使用手机的模式同样存在差异性,用单独的个人数据进行训练理论上更容易得到适合用户本身的模型,并且这样比混合数据更具有针对性,因此得到准确率更高的结果。然而在本实验中,由于存在个人数据量相对较少、基本情绪模型分类精度高等原因,某些分类器在基本情绪模型下对较少数据的分类结果可能并不理想,所以在基本情绪模型中有时会出现混合模型准确率比个人数据准确率更高的情况。但无论使用哪种情绪模型,我们选取的六种分类器中随机森林分类器的分类结果都远远高于其他五种分类器,这也可由随机森林算法中随机生成的决策树森林做出解释。

### 5.3 基于结果融合的情绪识别

#### 5.3.1 结果级融合

上一节中对信息融合技术中的特征级融合进行了介绍,本节将对信息融合中的另一类主要方法结果级融合进行介绍。结果级融合又称为决策级融合,是由每个传感器基于自己收集的数据做出结果判断,然后再对所有判断出的结果进行投票完成最终决策的过程,具有抗干扰能力强、对传感器依赖程度小等优势,下面将对结果级融合的过程进行详细描述。

本文共提取了 15 种传感器的数据,如果以每个传感器为单位进行决策级融合,不仅计算复杂度高,准确率也将相对降低。因为单类传感器与用户情绪之间的关联度并不会很高,例如我们并不能以用户某天的位置信息决断出用户的具体情绪,因此本节将以第四章中对传感器数据的分类为单元进行决策级融合,具体融合过程如图 5.10 所示。其中特征向量分为三类,分别是运动类感知数据特征向量,环境类感知数据特征向量和手机使用数据特征向量,我们先将这三类数据内的特征向量进行等权重拼接即特征融合以获得训练样例,然后分别对其使用 5.2 节中分类效果最好的随机森林分类器进行分类训练得到三组分类结果 $(P_{a_1}, P_{b_1}, P_{c_1}, P_{d_1}, P_{e_1})$ 、 $(P_{P_{11}}, P_{P_{21}}, P_{P_{31}}, P_{P_{41}}, P_{P_{51}})$ 和 $(P_{a_2}, P_{b_2}, P_{c_2}, P_{d_2}, P_{e_2})$ 、 $(P_{P_{12}}, P_{P_{22}}, P_{P_{32}}, P_{P_{42}}, P_{P_{52}})$ 以及 $(P_{a_3}, P_{b_3}, P_{c_3}, P_{d_3}, P_{e_3})$ 、 $(P_{P_{13}}, P_{P_{23}}, P_{P_{33}}, P_{P_{43}}, P_{P_{53}})$ ,最后将这三类特征向量的每种分类结果一一对应相加得到最终的分类结果,概率最高的情绪种类即为识别的结果,具体算法描述如表 5.4 所示。

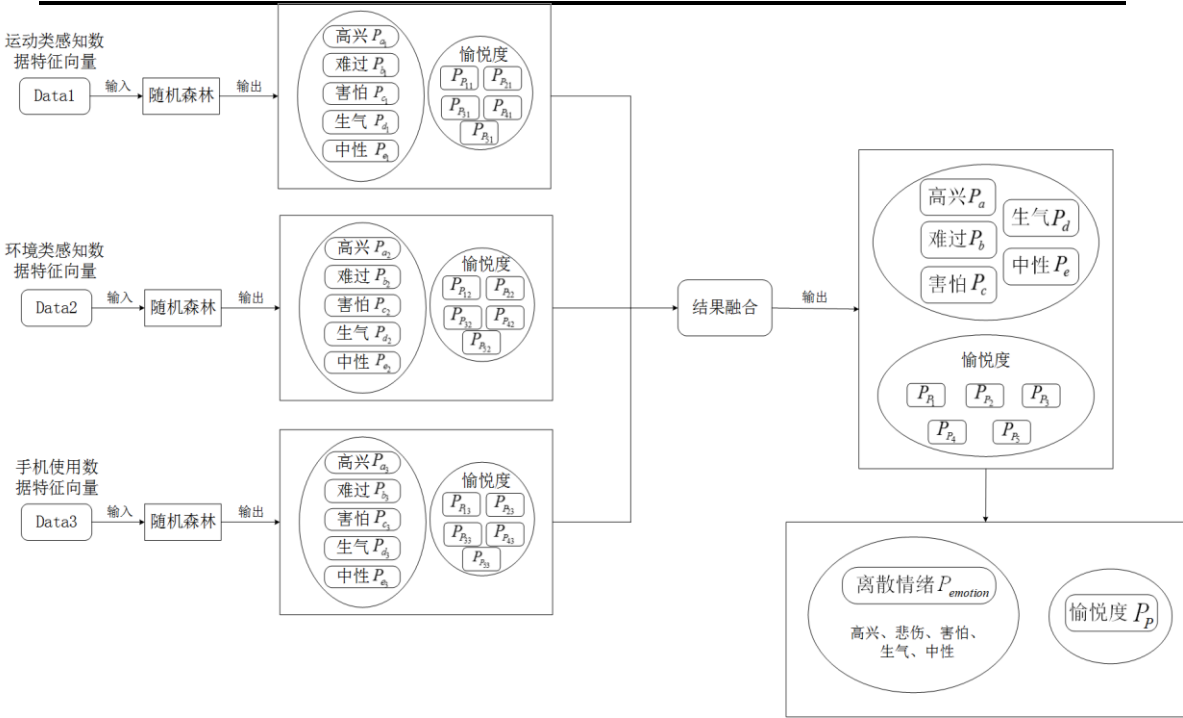


图 5.10 结果融合过程

表 5.4 多传感器数据的结果融合算法描述

<b>输入：</b> 1、文件存储路径； 2、各个传感器特征提取文件；
<b>输出：</b> 经过结果融合的情绪识别准确率模型
<b>算法描述：</b> Step1: 设置标准日期匹配格式； Step2: 对输入的文件存储路径内符合标准日期格式的文件根据前文所述的传感器顺序计算特征值； Step3: if 某个传感器特征值 = null 成立： if 传感器 = 蓝牙 or 电话记录 or 短信记录 or QQ 记录 or 微信记录 or WiFi 记录 成立： 将特征值赋值为 0； else： 采取临近值赋值方法，即用缺失值前后数据的均值为其赋值； Step4: 将计算过的特征值按照运动类感知数据、环境类感知数据和手机使用数据的类别依次写入三个不同的特征文件中并进行等权重拼接和存储； Step5: 重复 Step2 至 Step4 直至遍历并完成文件夹内所有文件特征值计算； Step6: 将融合后的三个新特征值作为输入值分别输入随机森林中进行训练，得出每类数据的情绪识别准确率； Step7: 将得到的情绪识别准确率（基本情绪模型和环状情绪模型中的愉悦度）分别一一对应相加得到最终的分类概率，即 $P_a = P_{a_1} + P_{a_2} + P_{a_3}$ ，

$$P_b = P_{b_1} + P_{b_2} + P_{b_3} \quad , \quad P_c = P_{c_1} + P_{c_2} + P_{c_3} \quad , \quad P_d = P_{d_1} + P_{d_2} + P_{d_3} \quad ,$$
$$P_e = P_{e_1} + P_{e_2} + P_{e_3} \quad , \quad P_{f_1} = P_{f_{11}} + P_{f_{12}} + P_{f_{13}} \quad , \quad P_{f_2} = P_{f_{21}} + P_{f_{22}} + P_{f_{23}} \quad ,$$
$$P_{f_3} = P_{f_{31}} + P_{f_{32}} + P_{f_{33}} \quad , \quad P_{f_4} = P_{f_{41}} + P_{f_{42}} + P_{f_{43}} \quad , \quad P_{f_5} = P_{f_{51}} + P_{f_{52}} + P_{f_{53}} \quad ,$$

其中  $P_a$ 、 $P_b$ 、 $P_c$ 、 $P_d$ 、 $P_e$  和  $P_{f_1}$ 、 $P_{f_2}$ 、 $P_{f_3}$ 、 $P_{f_4}$ 、 $P_{f_5}$  中准确率最高的即作为本种情绪分类模型的最终识别情绪以及准确率。

5.3.2 实验结果与分析

在实验验证部分，与特征融合相类似，我们依然将数据分为两类进行对比实验，一类是由全部志愿者数据构成的混合数据集，另一类是由个人数据较完整的志愿者数据构成的个人数据集。下面将采用两类不同的数据分别对用户情绪进行基于结果融合的识别，并对识别结果进行比较和分析。

首先，使用混合数据进行分类识别，在实验开始前，对所有的数据随机选择 80%作为训练数据，20%作为测试数据。对每一个数据样本进行特征向量的提取，并将特征向量的分量按照来源划分为运动类感知数据特征特征向量、环境类感知数据特征向量和手机使用数据特征向量三个类别。其中表 5.5 至表 5.8 分别展示了某五个样本数据在单独使用运动类特征向量、环境类特征向量、手机使用特征向量过程中以及经过结果融合后基于基本情绪模型的各种情绪分类概率，表 5.9 至表 5.12 展示了同样五个样本在单独使用运动类特征向量、环境类特征向量、手机使用特征向量过程中以及经过结果融合后基于环状情绪模型的各种情绪分类概率。可以从这八个表中看出，虽然各类感知数据的具体情绪分类概率并不完全相同，但是总体趋势一致，即同一样本数据经不同特征向量分类后概率最高的情绪类型基本一致，如样本 A 无论是在运动类感知数据、环境类感知数据、手机使用数据中还是在结果融合后都是“高兴”这一情绪概率最大。这个结果也从侧面验证了结果融合实验的可靠性，即三类传感器数据不论是融合前的单独判断还是融合后的整体判断分类结果都是相同的，只是具体数值不同。

表 5.5 运动类感知数据分类概率  
(基本情绪模型)

	高兴	悲伤	恐惧	生气	中性
数据 A	0.5	0.1	0.0	0.0	0.4
数据 B	0.2	0.1	0.0	0.0	0.7
数据 C	0.3	0.0	0.0	0.0	0.7
数据 D	0.5	0.0	0.0	0.0	0.5
数据 E	0.1	0.8	0.0	0.0	0.1

表 5.6 环境类感知数据分类概率  
(基本情绪模型)

	高兴	悲伤	恐惧	生气	中性
数据 A	0.6	0.0	0.0	0.0	0.4
数据 B	0.0	0.0	0.0	0.0	1.0
数据 C	0.2	0.0	0.0	0.0	0.8
数据 D	0.2	0.1	0.0	0.0	0.7
数据 E	0.3	0.3	0.0	0.0	0.4

表 5.7 手机使用数据分类概率  
(基本情绪模型)

	高兴	悲伤	恐惧	生气	中性
数据 A	0.6	0.0	0.0	0.0	0.4
数据 B	0.1	0.0	0.0	0.0	0.9
数据 C	0.2	0.3	0.0	0.0	0.5
数据 D	0.6	0.0	0.0	0.0	0.4
数据 E	0.2	0.3	0.0	0.0	0.5

表 5.9 运动类感知数据分类概率  
(环状情绪模型)

	P1	P2	P3	P4	P5
数据 A	0.0	0.0	0.8	0.2	0.0
数据 B	0.0	0.0	1.0	0.0	0.0
数据 C	0.0	0.21	0.68	0.11	0.0
数据 D	0.0	0.3	0.6	0.1	0.0
数据 E	0.0	0.8	0.2	0.0	0.0

表 5.11 手机使用数据分类概率  
(环状情绪模型)

	P1	P2	P3	P4	P5
数据 A	0.0	0.1	0.9	0.0	0.0
数据 B	0.0	0.2	0.8	0.0	0.0
数据 C	0.0	0.1	0.6	0.2	0.1
数据 D	0.0	0.3	0.5	0.1	0.1
数据 E	0.0	0.2	0.3	0.5	0.0

表 5.8 结果融合后分类概率  
(基本情绪模型)

	高兴	悲伤	恐惧	生气	中性
数据 A	0.567	0.033	0.0	0.0	0.4
数据 B	0.1	0.033	0.0	0.0	0.867
数据 C	0.23	0.1	0.0	0.0	0.67
数据 D	0.43	0.03	0.0	0.0	0.54
数据 E	0.2	0.467	0.0	0.0	0.333

表 5.10 环境类感知数据分类概率  
(环状情绪模型)

	P1	P2	P3	P4	P5
数据 A	0.0	0.0	0.9625	0.0375	0.0
数据 B	0.0	0.1	0.8	0.1	0.0
数据 C	0.0	0.05	0.75	0.2	0.0
数据 D	0.2	0.13	0.6	0.067	0.0
数据 E	0.0	0.8	0.2	0.0	0.0

表 5.12 结果融合后分类概率  
(环状情绪模型)

	P1	P2	P3	P4	P5
数据 A	0.0	0.033	0.8875	0.0795	0.0
数据 B	0.0	0.1	0.867	0.033	0.0
数据 C	0.0	0.12	0.677	0.17	0.033
数据 D	0.067	0.244	0.567	0.089	0.033
数据 E	0.0	0.6	0.233	0.167	0.0

前面八个表是对每个样本的具体情绪类别进行了识别，而全部测试样本的实验准确率结果如图 5.11 所示，这里的准确率指的是平均准确率，即测试数据中分类结果与实际标签相符的数据个数占全部测试数据的百分比。我们得到三类不同传感器数据测试样本基于基本情绪模型的准确率分别为 47.13%、49.08%、56.32%以及基于环状情绪模型的准确率分别为 59.77%、60.92%、64.37%。接下来根据这三类传感器特征向量的分类结果进行结果融合，可以看出经过结果融合的情绪识别准确率比单一类别数据的准确率有所提高，分别为基于基本情绪模型的 58.62%和基于环状情绪模型的 67.82%。通过对数据的分析，我们得出混合数据在经过结果融合后往往会比单一类别的传感器数据特征向量的识别准确率更高的结论，并且与前文一致，环状情绪模型识别准确率普遍高于基本情绪模型。

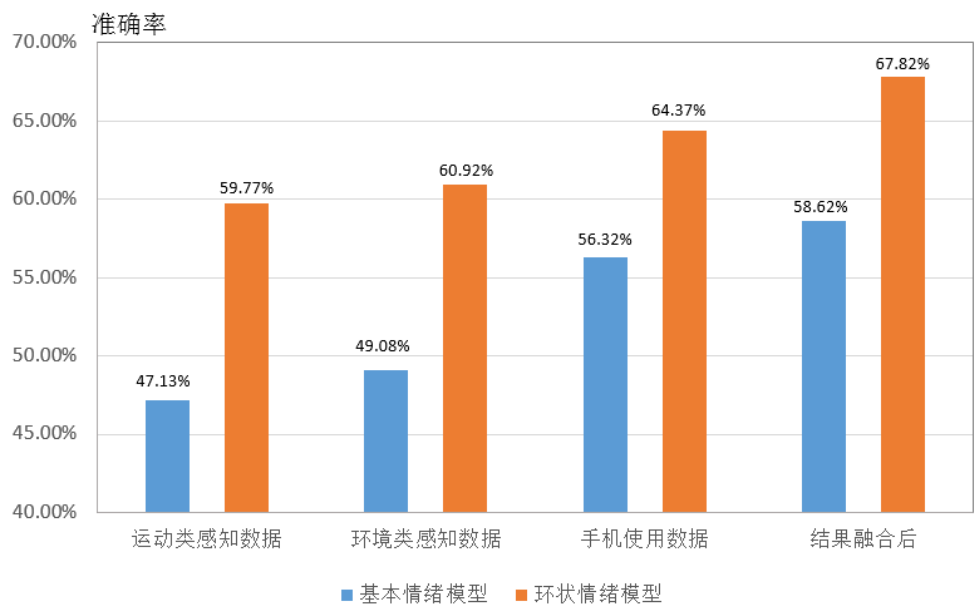


图 5.11 混合数据结果融合情绪识别准确率

接下来使用个人数据进行分类，与 5.2 节中一致，在实验中依然选取了有效数据最多的三号志愿者和十二号志愿者作为个人数据分析对象，实验结果如图 5.12 和图 5.13 所示。图中个人数据识别结果与混合数据识别结果呈现出相似性，经过结果融合的认可准确率要高于任何单独一类的传感器；并且使用环状情绪模型识别的结果也比基本情绪模型准确率更高。

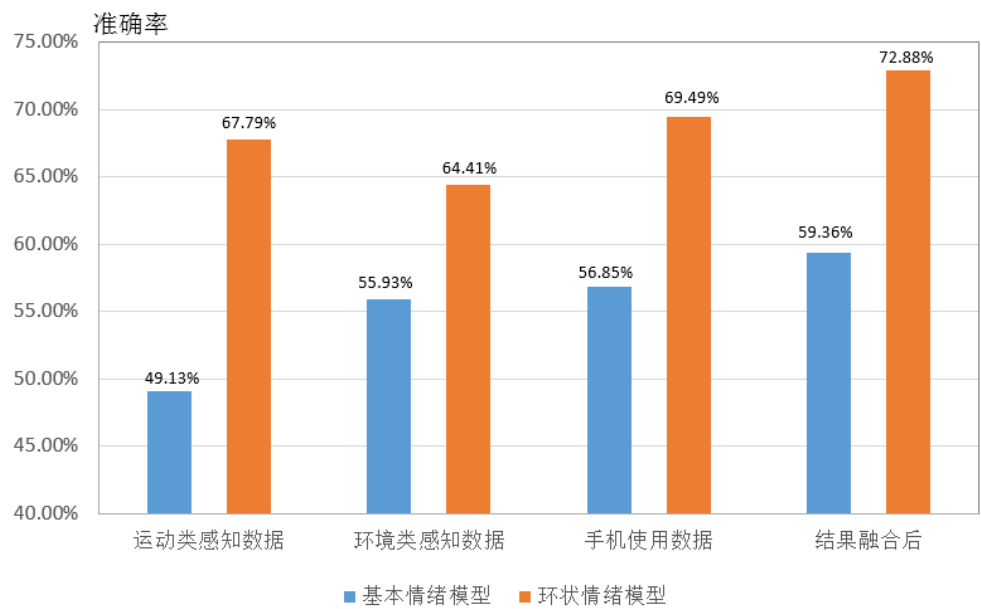


图 5.12 志愿者三号个人数据结果融合情绪识别准确率



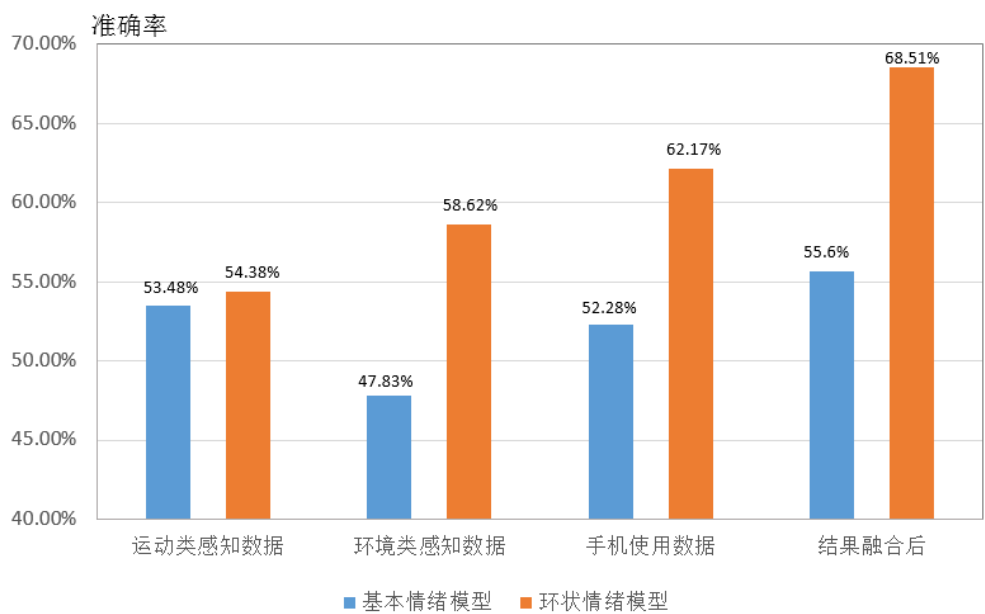


图 5.13 志愿者十二号个人数据结果融合情绪识别准确率

通过对比混合数据、个人数据在基本情绪模型下的识别率以及在环状情绪模型下的识别率(如图 5.14 和图 5.15 所示)，我们发现，不论是在基本情绪模型中还是在环状情绪模型中，不论是混合数据还是个人数据，经过结果融合的情绪识别准确率都会比单一传感器类别的识别准确率高，这是因为情绪的产生和变化受到多种因素影响并且会表现为多种形式，运动类数据、环境类数据和手机使用数据都只是其中的一种表现方式而已，只有将它们有机的结合起来才能更加全面和准确的进行情绪识别。

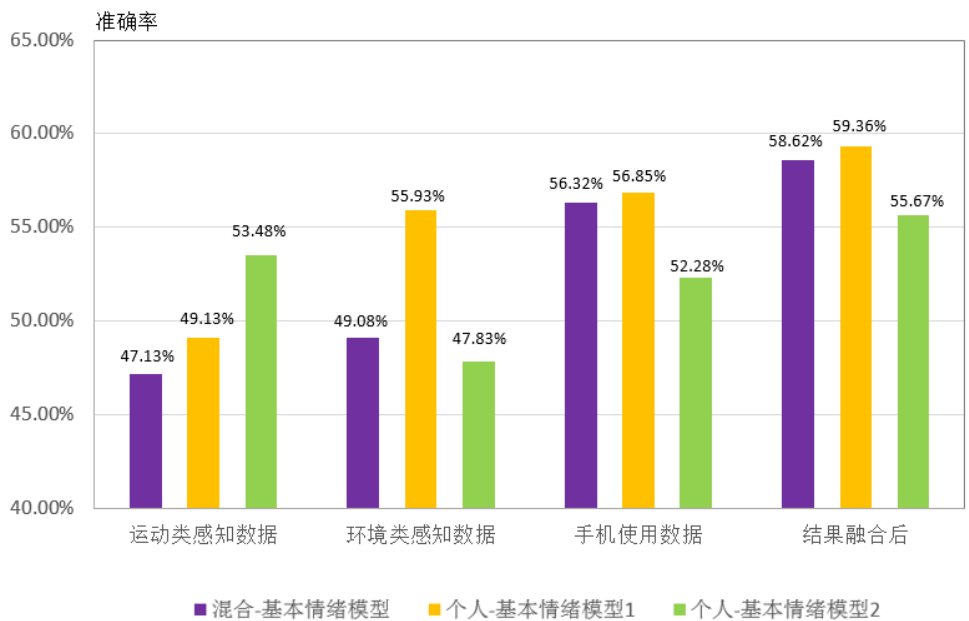


图 5.14 混合数据和个人数据结果融合的基本情绪模型对比

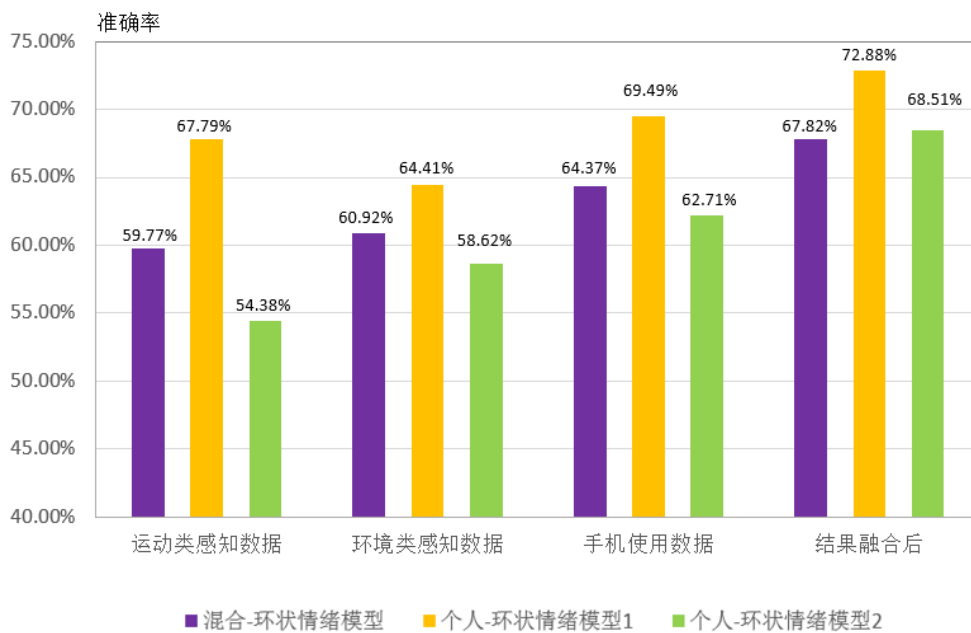


图 5.15 混合数据和个人数据结果融合的环状情绪模型对比

但是由于结果融合过程中数据损失量相较特征融合更大,存在精度相对较低的原因,因此结果融合的最高识别准确率并不如特征融合高。因为每个人的情绪变化所体现的行为模式的变化都是不一样的,因此个人数据相较混合数据更具有代表性。但由于数据收集过程中存在的一些问题导致出现个人有效数据量较小以及不均衡,因此并不能很好的挖掘出每个用户的个人模式,会有个人数据情绪识别准确率高于混合数据以及低于混合数据的情况同时出现,关于如何提高个人数据的有效率也是本文需要继续研究的方向。

## 5.4 相关工作比较

随着移动数据感知技术的不断发展,利用手机作为数据收集平台进行情绪识别的工作越来越多,其中最具代表性工作的包括 EmotionSense<sup>[8]</sup>、StressSense<sup>[75]</sup>、MoodScope<sup>[42]</sup>等。

EmotionSense 选取专业演员进行多种情绪的音频录制并使用高斯混合模型进行训练,将训练后的声音情绪识别模型装载到手机上并对高兴、悲伤、恐惧、生气和中性五类情绪进行识别。EmotionSense 中的数据样本是由专业演员对不同情绪类别进行的等数据量录制,因此这些样本不同于实际生活产生的情绪数据。但由于这些录制的数据为均衡样本,因此可以对每种情绪进行分类的识别,准确率依次达到 58.67%, 60%, 60%, 64%, 84.66%。与本文相比较,EmotionSense 中的实验数据存在两个问题,一是由于不是使用真实数据,因此情绪表达可能会较为夸张,存在影响准确率的可能性;二是日常生活中大多数

人的情绪是以中性和高兴居多<sup>[42]</sup>，而 EmotionSense 使用的是均衡样本，又因为本文中实验采集到的数据也大多集中在中性和高兴两类，将较来说本文的识别结果更具有真实意义。

StressSense 与 EmotionSense 类似，都是以声音的语音语调分析作为识别基础，它通过手机内嵌的麦克风记录声音，并对用户的压力程度进行分析，达到室内识别率 81% 和室外识别率 76% 的成效。虽然它只研究了压力一种情绪，但也为本文提供了值得研究的方向。

MoodScope<sup>[42]</sup>是和本文较为类似的研究，它通过对社交记录(电话、短信和邮件)和日常活动记录(浏览器历史、应用使用记录和历史位置)这两类数据的数据集训练得到用户的心情模型，并最终可以达到 93% 的准确率。但在具体实现过程中，本文和 MoodScope 的区别主要体现在以下三个方面：

首先，使用的情绪分类模型不同。MoodScope 中只使用了维度空间模型；而本文中采用了心理学分类中最为普遍使用的情绪分类模型，不仅包括了维度空间模型，还使用了基本情绪模型，并对它们的分类结果进行了分析。

其次，使用的数据集不同。环状情绪模型作为一种连续的情绪分类方式，重点描述了愉悦度和活跃度之间正相关关系，虽然 MoodScope 中也使用了该模型，但并未收集能够表示活跃度的相关数据；本文从多个角度对数据进行收集，其中运动类感知数据能够较好的对用户的活跃度进行表示。因此，在数据收集方面本文更为全面的收集了与用户情绪相关的感知数据，提取的数据特征相较来说类型丰富，更有价值性

最后，使用的分类器不同。MoodScope 中使用了多元线性回归的分类方法，而本文则是选取了 SVM、kNN、随机森林等六种分类器。

## 5.5 本章小结

本章就多维移动数据的用户情绪识别技术具体展开了研究，首先先对实验中的数据集进行了统计和分析，筛选出可供实验使用的有效数据并制定了实验的评估标准。接下来分别从特征融合以及结果融合两种角度对实验数据进行分类，由于数据收集和处理过程中会存在某些无法避免的噪声和损失，因此实验结果的准确率会受到一定影响。具体来讲，结果融合由于数据损失量相对较大，最终的情绪识别准确率比特征融合偏低；在情绪模型方面，由于环状情绪模型中的情绪属性具有连续性、模糊性，因此具有较高的识别率；每个人的情绪变化所体现的行为模式的变化都是不一样的，因此个人数据相较混合数据更具有代表性，会出现个人数据最高情绪识别准确率高于混合数据的情况。

## 第六章 总结与展望

随着智能手机领域的蓬勃发展,越来越多的传感器嵌入手机内部用以实现复杂的功能。作为移动感知领域中的感知平台,智能手机以其用户的广泛性、使用的简易性和感知方式的非干扰性著称,因此基于智能手机的相关研究近些年来大量涌现。而情绪识别工作作为人工智能和普适计算中的重要组成部分,一直受到业内人士的高度重视。在这样的背景之下,基于智能手机的情绪识别技术作为一种崭新的领域发展迅速,引起了人们的广泛关注。作为移动感知领域中的新型研究,基于智能手机的情绪识别收集了多种传感器数据,通过对这些移动数据的分析,分别建立不同的情绪模型,力求充分利用采集到的各种传感器文本数据对用户的情绪进行较为准确的推测与识别,为情绪识别技术提供新的研究方向。课题将多维移动数据的情绪识别作为主要的研究点,而进行情绪识别的前提是获取实验所需的移动数据以及数据的处理工作,因此文中也针对数据的收集和处理工作展开了研究。综上所述,本文的组要研究内容和贡献如下:

(1) 针对多样化的移动数据,究竟其中哪些能够反应出用户情绪变化后的行为模式值得研究,对此设计了网络问卷调查对情绪如何影响人的行为模式进行预调查,并对调查结果加以分析作为选择手机数据采集种类的依据。然后,采用数据收集工具实时采集手机数据记录用户的行为模式,同时开发了一个情绪记录软件供用户每天记录情绪,解决了移动数据收集的问题。

(2) 针对多传感器数据的收集带来移动数据规模的扩大化与类型复杂化的问题,本文针对不同类型的数据采用了不同的预处理方法和特征提取方法,并对数据进行归一化处理,解决了多维移动数据处理难度高的问题。

(3) 最后课题对基于智能手机的情绪识别展开了研究,采用特征级融合与结果级融合两种方法进行了实验验证。结果表明通过算法中参数的合理选择,能够较准确的对用户情绪进行识别,具有一定的可用性。

由于实验条件的限制,本课题的工作在未来还有以下几个需要改进的地方,值得进一步展开研究:

(1) 在实验验证部分使用的数据集规模较小,如何动员志愿者收集到更大规模的数据集亟待解决。下一步准备在大规模数据集上进行实验验证,并分析算法效果。

(2) 实验数据收集过程中会存在很多噪声,如何在数据预处理过程中最大程度的去除噪声和降低噪声影响,以及提取出更有效、更具有代表情绪作用的特征向量是下一步的研究目标。

(3) 对于情绪识别的分类阶段的研究还不够深入，下一步将对特征级融合和结果级融合算法进行更深入的研究。

## 致 谢

知止而后有定，定而后能静，静而后能安，安而后能虑，虑而后能得。

——《大学》

天下难事，必作于易；天下大事，必作于细。

——《老子》

时光荏苒，日月如梭，犹如白驹过隙。回忆起刚刚到长沙时对学校的憧憬，离家半年后对家乡和父母亲友的各种思念，犹如发生在昨天，历历在目。“勿谓寸阴短，既过难再获”，转眼间两年半的硕士求学之路即将结束，虽有些如释重负，但却不复之前期待毕业时的满满喜悦感，反倒有些怅然若失。两年多的求学之路让我更加深刻的体会了这两句话的含义：若想终有收获，必先确立目标；制定出每一步的计划，化整为零，化难为易，这样才能夯实基础、稳扎稳打、取得成功。

离别在即，尽管在毕业之后会有新的工作、生活、学习环境，会有新的朋友，但现在我的心中却不是充满了对未来的憧憬之情，而是满怀对导师的不舍，对同窗的不舍，对科大的不舍。很感谢导师、同学们、亲友们的教导、关心和帮助，是你们的鼓励和支持陪伴着我走过了每一个春夏秋冬，度过了每一个或愉快或气馁的日子。因此，在论文即将完成之际，我衷心的感谢你们，感谢一路有你们的相伴，感谢在你们的帮助下我能逐渐成长。

衷心感谢我的导师史殿习教授！我做这个课题根本是毫无基础可言，感谢史老师在课题之初为我指点迷津，以及开题之后对我每一步进展的肯定和提出的诸多宝贵建议。您严谨的治学态度和始终如一的工作热情一直让我深深的敬佩，受益良多，尤其是您在研究问题时的那种精益求精的探索精神，将一直影响我之后的学习和工作。俗话说，“一日为师，终身为父”，在课题研究过程中，您就像一位“严父”般一直督促着我不断前进；而在生活中，您又像朋友一样平易近人，关心着我们。正因如此，我收获了很多，在此谨向您表示最衷心的感谢和最诚挚的敬意。

衷心感谢丁博老师、刘惠老师与尹刚老师，感谢您们在我课题之初对我提出的宝贵的意见和建议，使我更加明确了我的研究方向和研究思路，为后期课题的展开提供了很多指导。

感谢李勇谋、何炫辰、丁涛杰、陈富霞、金星、姜龙泉等师兄师姐，你们在我课题之初提供的宝贵建议使我受益良多；感谢同级的杨若松、吴渊、李寒、樊泽栋、谭杰夫和周荣在学习和生活上对我的帮助，感谢你们在每次开会作报告时提出我课题中的不足，每次团体活动时对我的照顾，感谢你们这一路的陪

伴，毕业在即，愿友谊长存；感谢莫晓赟、李中秋、王峰、刘帆、陈晓鹏、赵邦辉、魏箐等学弟学妹，从你们身上我学到了很多，是你们拓宽了我的知识面，也丰富了我的硕士生活；感谢舍友黄杨琛、刘大雪和宋倩，是你们一直陪伴着我度过了在科大的每一天。

感谢朱涛政委、郑勇辉政委、孙友佳队长，是你们队干部亲切的关怀帮助我迅速的适应了科大的生活；同样感谢学员八队的每一位同学，和你们在一起的时间虽然不长，只有两年半，但却是我一生中最宝贵的财富。

最后我要深深的感谢我的父母，感谢你们这么多年来对我无微不至的照顾，感谢你们在我人生成长过程中每一个阶段对我的教育，感谢你们一直迁就我的任性。无论现在还是将来，你们对我的支持和认同将是我最大的精神支持和最强的前进动力。爸爸妈妈，你们辛苦了！

最后的最后，我想说乘风破浪终有时，直挂云帆济沧海，希望每个同学今后都能在工作中大展宏图，为实现自己的人生目标而努力！

## 参考文献

- [1] Hong J, Baker M. Wearable Computing[J]. IEEE Pervasive Computing, 2014, 13(2): 7-9.
- [2] 孟昭兰. 人类情绪[M]. 上海人民出版社, 1989.
- [3] 张颖, & 罗森林. (2004). 情感建模与情感识别. 计算机工程与应用, 39(33), 98-102.
- [4] Khan, Wazir Zada, et al. "Mobile phone sensing systems: A survey." Communications Surveys & Tutorials, IEEE 15.1 (2013): 402-427.
- [5] Miluzzo, Emiliano, Tianyu Wang, and Andrew T. Campbell. "EyePhone: activating mobile phones with your eyes." Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds. ACM, 2010.
- [6] L. Deng and L. P. Cox, "Livecompare: grocery bargain hunting through participatory sensing," in Proceedings of the 10th workshop on Mobile Computing Systems and Applications, ser. HotMobile '09. New York, NY, USA: ACM, 2009, pp. 4:1-4:6. [Online]. Available: <http://doi.acm.org/10.1145/1514411.1514415>
- [7] P. Mohan, V. N. Padmanabhan, and R. Ramjee, "TrafficSense: Rich monitoring of road and traffic conditions using mobile smartphones," Tech. Rep. no. MSR-TR-2008-59, April 2008.
- [8] Rachuri, Kiran K., Mirco Musolesi, et al. , EmotionSense: a mobile phones based adaptive platform for experimental social psychology research[C]//In Proceedings of the 12th ACM international conference on Ubiquitous computing, pp. 281-290. ACM, 2010.
- [9] Khan, Wazir Zada, et al. , Mobile phone sensing systems: A survey[J]. Communications Surveys & Tutorials, IEEE 15.1 (2013): 402-427.
- [10] 熊英, 史殿习, 丁博, 邓璐. 移动群体感知技术研究 [J]. 计算机科学, 2014, 04: 1-8.
- [11] Calvo, Rafael, and Sidney D'Mello. Affect detection: An interdisciplinary review of models, methods, and their applications[J]. Affective Computing, IEEE Transactions on 1.1 (2010): 18-37.
- [12] Picard, Rosalind W., and Rosalind Picard. Affective computing[M]. Vol. 252. Cambridge: MIT press, 1997.
- [13] 党宏社, 郭楚佳, 张娜. 信息融合技术在情绪识别领域的研究展望[J]. 计算机应用研究, 2013, 12: 3536-3539.
- [14] 张琰. 三种情绪表情的识别及其相互影响研究. MS thesis. 苏州大学, 2010.
- [15] Ekman, Paul, and Wallace V. Friesen. "Facial action coding system." (1977).
- [16] Lavagetto Fabio, Pockaj Roberto. An Efficient Use Of MPEG-4 FAP Interpolation for Facial Animation at 70 bits/Frame [J]. IEEE TRANSACTIONS. ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, 2001-10, 11(10).
- [17] Sarris Nikos, Michael Nikos Grammalidis. Gerassimos Strintzis FAP Extraction Using Three-Dimensional Motion Estimation [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002-10, 12(10): 865-876.
- [18] Kuo Chung J, Huang Ruey Song, Lin Tsang Gang. 3-D Facial Modal



- 
- Estimation From Single Front -View Facial Image [J]. IEEE TRANSACTIONS. ON CHIRCUITS AND SYSTEMS FOR VIDEO TECHINOLOGY, 2002-03, 12(3).
- [19]Calder Andrew J, Burton A Mike, Miller Paul, Young Andrew W. A Principal Component Analysis of Facial Expressions [J]. Vision research, 2001, 41: 1179-1208.
- [20]Georghiades Athinodoros S, et al. From Few to Many: Illumination Cone Models for Faces Recognition under Variable Lighting and Pose [J]. IEEE TRANSACTIONS ON PAMI., 2001-06, 23(6).
- [21]Martin Lades, et al. Distortion Invariant Object Recognition in the Dynamic Link Architecture [J]. IEEE TRANSACTIONS ON COMPUTERS, 1993-03, 42(3).
- [22]C.M. Lee and S.S. Narayanan. Toward detecting emotions in spoken dialogs[J]. IEEE Trans. Speech and Audio Pro-cessing, 13(2), pages 293-303, 2005.
- [23]Morrison, Donn, Ruili Wang, and Liyanage C. De Silva. "Ensemble methods for spoken emotion recognition in call-centres." Speech communication 49.2 (2007): 98-112.
- [24]Spector, Paul E. "Organizational frustration: A model and review of the literature." Personnel Psychology 31.4 (1978): 815-829.
- [25]Cahn, Janet E. "The generation of affect in synthesized speech." Journal of the American Voice I/O Society 8 (1990): 1-19.
- [26]Yoon, Won-Joong, Youn-Ho Cho, and Kyu-Sik Park. "A study of speech emotion recognition and its application to mobile services." Ubiquitous Intelligence and Computing. Springer Berlin Heidelberg, 2007. 758-766.
- [27]Greasley, Peter, Carol Sherrard, and Mitch Waterman. "Emotion in language and speech: Methodological issues in naturalistic approaches." Language and Speech 43.4 (2000): 355-375.
- [28]Bachorowski, Jo-Anne, and Michael J. Owren. "Vocal expression of emotion: Acoustic properties of speech are associated with emotional intensity and context." Psychological science 6.4 (1995): 219-224.
- [29]Wu, Naiyu, Huiping Jiang, and Guosheng Yang. "Emotion recognition based on physiological signals." Advances in Brain Inspired Cognitive Systems. Springer Berlin Heidelberg, 2012. 311-320.
- [30]Picard R.W., Vyzas E. and Healey J. Toward Machine Emotional Intelligence: Analysis of Affective Physiological State. IEEE Transactions on Patern Analysis and Machine Intelligence. 2001, 23(10): 1175-1191.
- [31]Johannes Wagner and Jonghwa Kim and Elisabeth Andre. From Physiological Signals to Emotions: Implementing and Comparing Selected Methods for Feature Extraction and Classification[C]. In IEEE International Conference on Multimedia & Expo (ICME 2005), 2005, 940-943.
- [32]Kim K H, Bang S W, Kim S R. Emotion recognition system using short-term monitoring of physiological signals. Med Biol Eng compute, 2004, 42: 419-427.
- [33]谢晶, 方平, 姜媛. 情绪测量方法的研究进展[J]. 心理科学, 2011, 34(2), 488-493.
- [34]Kim, Jonghwa, and Elisabeth André "Emotion recognition based on physiological changes in music listening." Pattern Analysis and Machine
-

- Intelligence, IEEE Transactions on 30.12 (2008): 2067-2083.
- [35]Soleymani, Mohammad, et al. "A multimodal database for affect recognition and implicit tagging." Affective Computing, IEEE Transactions on 3.1 (2012): 42-55.
- [36]Koelstra, Sander, et al. "Deap: A database for emotion analysis; using physiological signals." Affective Computing, IEEE Transactions on 3.1 (2012): 18-31.
- [37]Ekman, Paul, and Wallace V. Friesen. "Detecting deception from the body or face." Journal of personality and Social Psychology 29.3 (1974): 288.
- [38]Scherer, Klaus R., and Harald G. Wallbott. "Ausdruck von emotionen." Psychologie der Emotion (1990): 345-422.
- [39]Shaarani, Ahmad S., and Daniela M. Romano. "Perception of emotions from static postures." Affective Computing and Intelligent Interaction. Springer Berlin Heidelberg, 2007. 761-762.
- [40]De Silva, P. Ravindra, and Nadia Bianchi - Berthouze. "Modeling human affective postures: an information theoretic characterization of posture features." Computer Animation and Virtual Worlds 15.3 - 4 (2004): 269-276.
- [41]Kapur, Asha, et al. "Gesture-based affective computing on motion capture data." Affective computing and intelligent interaction. Springer Berlin Heidelberg, 2005. 1-7.
- [42]LiKamWa R, Liu Y, Lane N D, et al., Moodscope: build-ing a mood sensor from smartphone usage patterns[C] //Proceeding of the 11th annual international conference on Mobile systems, applications, and services. ACM, 2013: 389-402.
- [43]LiKamWa R, Liu Y, Lane N D, et al. Can your smartphone infer your mood?[C]//PhoneSense workshop. 2011.
- [44]Bogomolov, Andrey, Bruno Lepri, and Fabio Pianesi. Hap-piness recognition from mobile phone data[C]//Social Computing (SocialCom), 2013 International Conference on. IEEE, 2013.
- [45]P. Ekman. Universal and cultural differences in facial ex-pressions of emotion[M]. In Proc. Nebraska Symp. Motiva-tion, 1971.
- [46]S. S. Tomkins. Affect, Imagery, Consciousness[M]. Springer, 1962.
- [47]Russell, James A. A circumplex model of affect[J]. Journal of personality and social psychology 39.6 (1980): 116
- [48]Ekman P, Power MJ. Handbook of Cognition and Emotion. Sussex: John Wiley & Sons, 1999
- [49]Ortony, Andrew, and Terence J. Turner. "What's basic about basic emotions?." Psychological review 97.3 (1990): 315.
- [50]L. Feldman Barrett and J. Russell. Independence and Bipo-larity in the Structure of Current Affect[J]. Journal of Per-sonality and Social Psychology, 74:967-984, 1998.
- [51]Xie, B. "Research on key issues of Mandarin speech emotion recognition [Ph. D. Thesis]." Hangzhou: Zhejiang University (2006).
- [52]Russell, James A. A circumplex model of affect[J]. Journal of personality and social psychology 39.6 (1980): 1161.
- [53]Plutchik, Robert. "A general psychoevolutionary theory of emotion." Theories of emotion 1 (1980).
- [54]方洪鹰. 数据挖掘中数据预处理的方法研究. Diss. 硕士, 西南大学, 2009.

- 
- [55]关大伟, and 李雄飞. "数据挖掘中的数据预处理." (2006): 21-27.
- [56]Khan A M, Lee Y-K, Lee S Y, et al. A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recog-nizer [J]. Information Technology in Biomedicine, IEEE Transactions on. 2010,14 (5): 1166–1172.
- [57]Kozina S, Gjoreski H, Gams M, et al. Efficient Activity Recognition and Fall Detection Using Accelerometers [M] // Kozina S, Gjoreski H, Gams M, et al. Evaluating AAL Systems Through Competitive Benchmarking. Springer, 2013: 2013: 13–23.
- [58]Huynh T, Schiele B. Analyzing features for activity recognition [C]. In Proceedings of the 2005 joint conference on Smart objects and ambient intelligence: innovative context-aware services: usages and technologies. 2005: 159–163.
- [59]Cortes C, Vapnik V. Support-vector networks [J]. Machine learning. 1995, 20 (3): 273–297.
- [60]Peterson, Leif E. "K-nearest neighbor." Scholarpedia 4.2 (2009): 1883.
- [61]Safavian, S. Rasoul, and David Landgrebe. "A survey of decision tree classifier methodology." IEEE transactions on systems, man, and cybernetics 21.3 (1991): 660-674.
- [62]Breiman, Leo. "Random forests." Machine learning 45.1 (2001): 5-32.
- [63]R äsch, Gunnar, Takashi Onoda, and K-R. Müller. "Soft margins for AdaBoost." Machine learning 42.3 (2001): 287-320.
- [64]Friedman, Jerome H. "Greedy function approximation: a gradient boosting machine." Annals of statistics (2001): 1189-1232.
- [65]<http://scikit-learn.org/stable/modules/ensemble.html#gradient-tree-boosting>
- [66]<http://www.sojump.com/jq/4414533.aspx>
- [67]Rawassizadeh R, Tomitsch M, Wac K, et al. UbiqLog: a generic mobile phone-based life-log framework[J]. Personal and ubiquitous computing, 2013, 17(4): 621-637.
- [68]<http://tech.163.com/15/0225/07/AJ9HL4H7000915BD.html>
- [69]Bray T, Paoli J, Sperberg-McQueen C M, et al. Extensible markup language (XML)[J]. World Wide Web Consortium Recommendation REC-xml-19980210. <http://www.w3.org/TR/1998/REC-xml-19980210>, 1998.
- [70]Crockford D. The application/json media type for javascript object notation (json)[J]. 2006.
- [71]Hartigan, John A., and Manchek A. Wong. Algorithm AS 136: A k-means clustering algorithm[J]. Applied statistics (1979): 100-108.
- [72]Kwon O W, Chan K, Hao J, et al. Emotion Recognition by speech signal[c]/INTERSPEECH. 2003.
- [73]Shi T, Horvath S. Unsupervised learning with random forest predictors[J]. Journal of Computational and Graphical Statistics, 2006, 15(1).
- [74]Bogomolov A, Lepri B, Pianesi F. Happiness recognition from mobile phone data[C]//Social Computing (SocialCom), 2013 International Conference on. IEEE, 2013: 790-795.
- [75]Lu H, Frauendorfer D, Rabbi M, et al. Stresssense: Detecting stress in unconstrained acoustic environments using smartphones[C]//Proceedings of the 2012 ACM Conference on Ubiquitous Computing. ACM, 2012: 351-360.
-

## 作者在学习期间取得的学术成果

- [1] 陈茜,史殿习,杨若松. 基于多维数据特征融合的用户情绪识别.计算机科学与探索(已录用)