

毕业设计资料

Kangkang

2017 年 12 月 2 日

1 Dataset

2017.12.1

1.1 来源

<https://physionet.org/physiobank/database/noneeg/>

Reference: Birjandtalab, Javad, Diana Cogan, Mazyar Baran Pouyan, and Mehrdad Nourani, A Non-EEG Biosignals Dataset for Assessment and Visualization of Neurological Status, 2016 IEEE International Workshop on Signal Processing Systems (SiPS), Dallas, TX, 2016, pp. 110-114. doi: 10.1109/SiPS.2016.27

1.2 描述

用于推断 20 名健康人的神经状态（包括身体压力，认知压力，情绪压力和放松）。

使用非侵入式手腕佩戴的生物传感器收集数据,并且包括电活动(EDA), 温度, 加速度, 心率 (HR) 和动脉血氧饱和度 (SpO2)。

数据包括 20 个样本的 7 个阶段数据:

1. 放松 5min
2. 身体压力: 站立 1min, 以每小时一英里的速度步行 2min, 然后在跑步机上以每小时三英里的速度步行/慢跑 2min
3. 放松 5min

4. 小情绪压力：40s，告知被试在接下来的 3min 会计算从 2485 每次减 7 的结果（注意：这部分数据是在认知压力任务之前收集的，在本文中
没有解释）
5. 认知压力：计算 3min，进行 Stroop 测试 2min。Stroop 测试：读取用
不同颜色墨水的笔写的颜色名称，说出墨水的颜色
6. 放松 5min
7. 情绪压力：告知志愿者会在 1min 内从一部恐怖电影中看到一个 5min
的片段。经过一段时间后，播放《部落》的片段
8. 放松 5min（我们本来并不打算把指令的读数算作情绪压力。毕竟，每
个任务都有指示。然而，与其他指令集不同的是，这一项在许多志愿
者身上产生了压力反应，这对测试管理员来说是显而易见的。）

数据文件以 WFDB 格式提供，每个被试有两个 record：一个包含加速度，
温度和 EDA 信号，另一个包含 SpO2 和心率信号。.hea 文件包含有关该
被试的信息。每个记录有一个注释文件，用于指示转换状态的时间位置和标
签。subjectinfo.csv 文件还包含有关每个主题的信息。

1.3 数据读取

每个被试的数据包括以下五个文件：

- Subject * __AccTempEDA.atr
- Subject * __AccTempEDA.dat
- Subject * __AccTempEDA.he
- Subject * __SpO2HR.dat
- Subject * __SpO2HR.he

其中 * 表示 1 ~ 20.

1.3.1 读取 *AccTempEDA.atr* 文件

.atr 文件为 ATE 信号文件的注释文件，指示转换状态的时间位置和标
签。使用 WFDB 包中的 rdann 函数进行读取。

Function 1 *annotation = rdann(recordname, extension, sampfrom=0, sampto=None, shiftsamps=False, pbdir=None, return_label_elements=['symbol'], summarize_labels=False)*

```
import wfdb
ann = wfdb.rdann('C:/Users/KangKang/Desktop/Graduation-Project/Dataset/
                Subject1_AccTempEDA', 'atr')
```

ann 中包含以下参数，可以通过 *ann._* 调用。

extension	所处文件扩展名
sample	标记所处的位置
symbol	用于标记的符号
subtype	标记的类型
chan	标记所处的通道
num	每种标记的数目
aux_note	标记的辅助信息
fs	每条记录的采样频率
label_store	用于存储/编码每个注释标签的整数值
description	每个注释标签的描述性字符串
custom_labels	注释文件中定义的自定义注释标签
contained_labels	此注释中包含的唯一标签

查看 ann 的参数，发现在该数据库中有用的参数有以下几个：

ann_sample - NumPy array

	0
0	1
1	2401
2	5026
3	7427
4	7748
5	10661
6	13062
7	15943

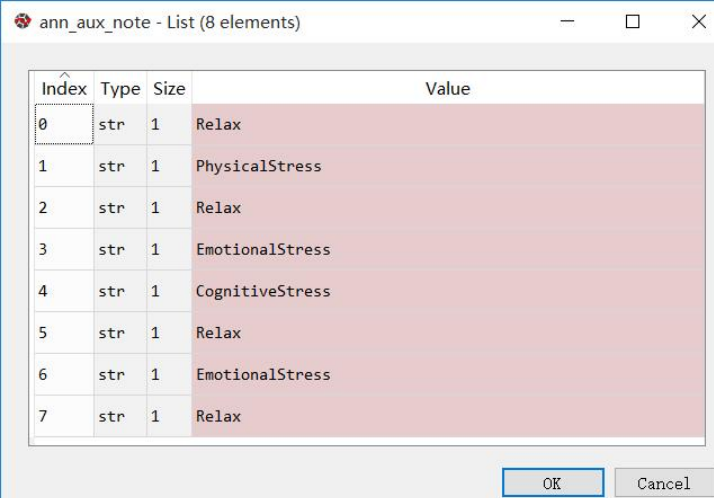
Format Resize ☒ Background color

OK Cancel

ann_symbol - List (8 elements)

Index	Type	Size	Value
0	str	1	"
1	str	1	"
2	str	1	"
3	str	1	"
4	str	1	"
5	str	1	"
6	str	1	"
7	str	1	"

OK Cancel



Index	Type	Size	Value
0	str	1	Relax
1	str	1	PhysicalStress
2	str	1	Relax
3	str	1	EmotionalStress
4	str	1	CognitiveStress
5	str	1	Relax
6	str	1	EmotionalStress
7	str	1	Relax

其中 *ann.sample* 记录标记所在的点的位置, *ann.symbol* 表明该数据库中的标记用"表示, *ann.aux_note* 表示标记的含义分别为: Relax-PhysicalStress-Relax-EmotionalStress-ConitiveStress-Relax-EmotionalStress-Relax.

1.3.2 读取 *AccTempEDA.dat* 文件和 *AccTempEDA.he*

.dat 文件为 ATE 信号文件的数据文件, .hea 为 ATE 信号的头文件。使用 WFDB 包中的 *rdsamp* 函数进行读取。

Function 2 *record = rdsamp(recordname, sampfrom=0, sampto=None, channels=None, physical=True, pmdir = None, m2s=True)*

```
import wfdb
ATE_dat = wfdb.rdsamp(root+'Subject2_AccTempEDA')
```

ATE_dat 中包含以下参数, 可以通过 *ATE_dat._* 调用.

p_signals	五个通道的数据
fs	采样频率
units	单位
signame	信号的名字
comments	其他信息

查看 ATE_dat 的参数，发现在该数据库中有用的参数有以下几个：

ATE_dat_sig - NumPy array

	0	1	2	3	4
0	0.969971	-0.670058	-0.0700455	28.7001	0.818044
1	0.969971	-0.680018	-0.0700455	28.9001	0.818044
2	0.979983	-0.670058	-0.0700455	28.9001	0.816017
3	0.969971	-0.670058	-0.0600322	28.9001	0.815003
4	0.979983	-0.670058	-0.0600322	28.9001	0.813024
5	0.969971	-0.670058	-0.0700455	28.9001	0.815003
6	0.969971	-0.670058	-0.0700455	28.9001	0.813024
7	0.969971	-0.670058	-0.0700455	28.9001	0.813024
8	0.969971	-0.670058	-0.0700455	28.9001	0.807039

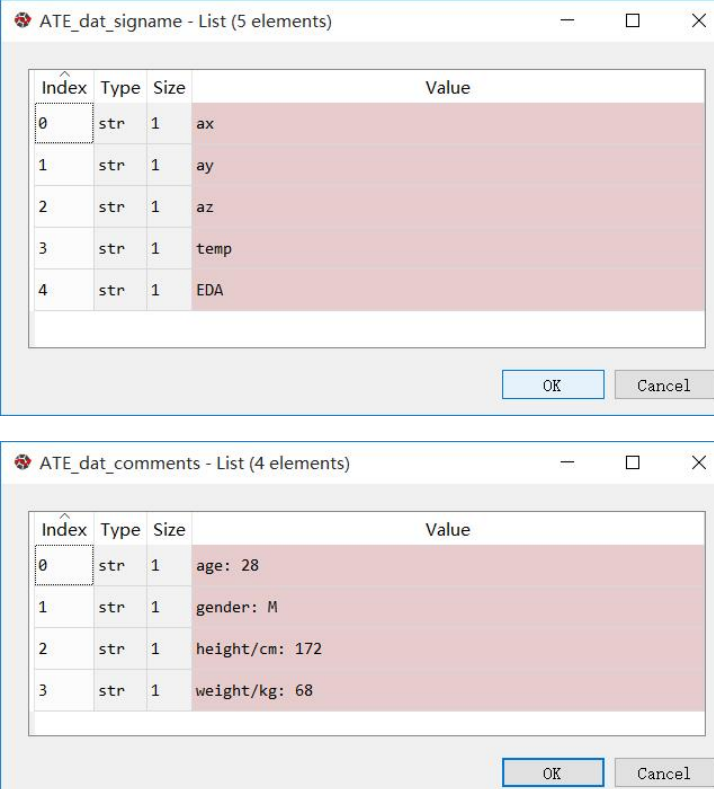
Format Resize ☒ Background color OK Cancel

ATE_dat_fs int 1 8

ATE_dat_units - List (5 elements)

Index	Type	Size	Value
0	str	1	NU
1	str	1	NU
2	str	1	NU
3	str	1	degC
4	str	1	NU

OK Cancel



ATE_dat_signame - List (5 elements)

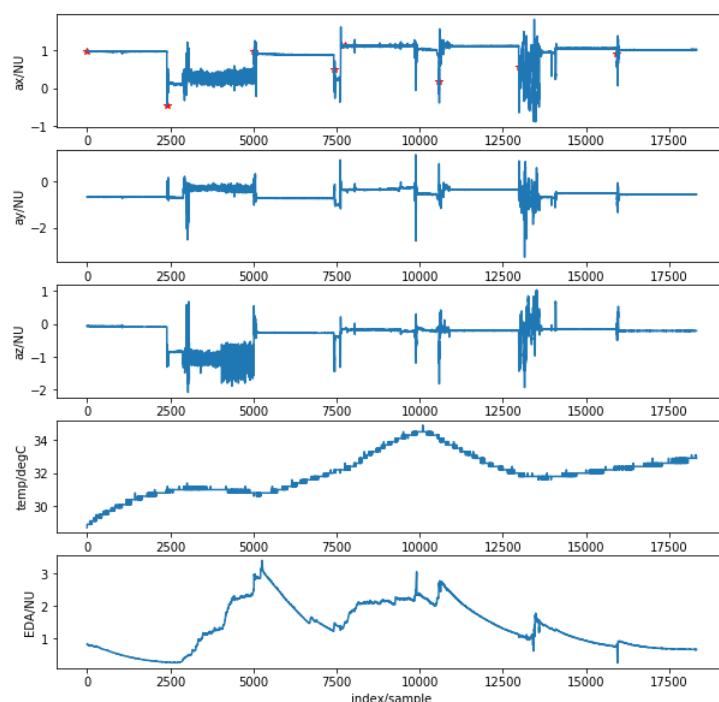
Index	Type	Size	Value
0	str	1	ax
1	str	1	ay
2	str	1	az
3	str	1	temp
4	str	1	EDA

ATE_dat_comments - List (4 elements)

Index	Type	Size	Value
0	str	1	age: 28
1	str	1	gender: M
2	str	1	height/cm: 172
3	str	1	weight/kg: 68

其中 *ATE_dat.p_signals* 记录五个通道的数据：三轴加速度、温度、EDA, *ATE_dat.fs* 记录采样率, 为 8Hz, *ATE_dat.units* 记录单位, 分别为: nv, nv, nv, degC, nv, *ATE_dat.signame* 记录信号的名称, 分别为: ax, ay, az, Temp, EDA, *ATE_dat.comments* 记录其他信息, 包括: 年龄、性别、身高、体重。

使用自带的作图包绘图, 结果如下:



1.3.3 读取 *SpO2HR.dat* 文件和 *SpO2HR.he*

.dat 文件为 SH 信号文件的数据文件，.hea 为 SH 信号的头文件。使用 WFDB 包中的 `rdsamp` 函数进行读取。

Function 3 `record = rdsamp(recordname, sampfrom=0, sampto=None, channels=None, physical=True, pmdir = None, m2s=True)`

```
import wfdb
SH_dat = wfdb.rdsamp(root+'Subject2_SpO2HR')
```

SH_dat 中包含以下参数，可以通过 `SH_dat._` 调用。

p_signals	两个通道的数据
fs	采样频率
units	单位
signame	信号的名字
comments	其他信息

查看 ATE_dat 的参数，发现在该数据库中有用的参数有以下几个：

SH_dat_sig - NumPy array

	0	1
0	97.9999	90.0005
1	97.9999	90.0005
2	99	90.0005
3	99	90.0005
4	99	90.0005
5	97.9999	90.0005
6	97.9999	91.0001
7	97.9999	95.0006
8	97.9999	95.0006
9	97.9999	97.0004

Format Resize ☒ Background color

OK Cancel

SH_dat_fs	int	1	1
-----------	-----	---	---

SH_dat_units - List (2 elements)

Index	Type	Size	Value
0	str	1	%
1	str	1	bpm

OK Cancel

SH_dat_signame - List (2 elements)

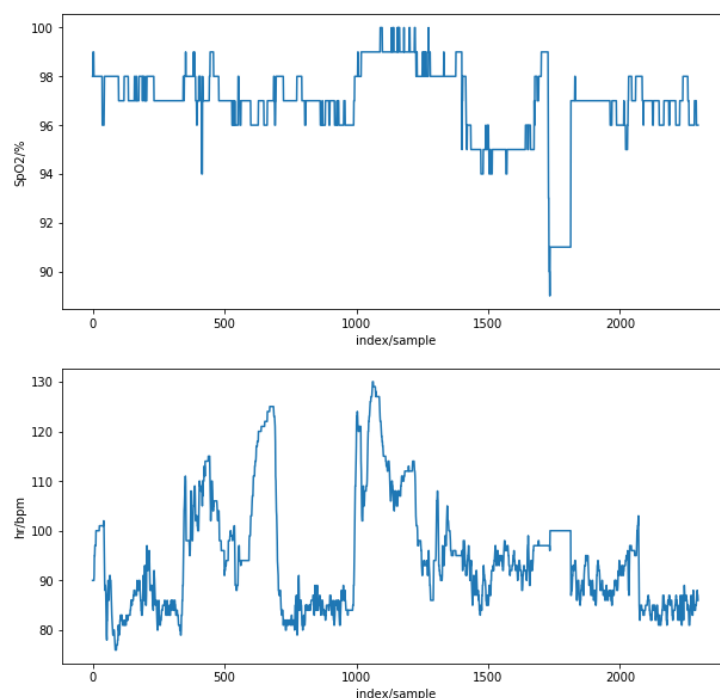
Index	Type	Size	Value
0	str	1	SpO2
1	str	1	hr

SH_dat_comments - List (4 elements)

Index	Type	Size	Value
0	str	1	age: 28
1	str	1	gender: M
2	str	1	height/cm: 172
3	str	1	weight/kg: 68

其中 $SH_dat.p_signals$ 记录两个通道的数据:血氧饱和度,心率, $SH_dat.fs$ 记录采样率,为 1Hz, $SH_dat.units$ 记录单位,分别为:%、bpm, $SH_dat.signame$ 记录信号的名称,分别为: SpO2, hr, $SH_dat.comments$ 记录其他信息,包括: 年龄、性别、身高、体重。

使用自带的作图包绘图, 结果如下:



2 数据整理

2017.12.2

将从数据库读取的每一条记录整理成可处理的格式。

2.1 变量处理

使用 root、root1、root2 进行记录的路径设置。

```
root='C:/Users/KangKang/Desktop/Graduation-Project/Dataset/'
root1='Subject1_AccTempEDA'
root2='Subject1_SpO2HR'
```

使用 `ReadData(root1,root2)` 函数，将阶段点位置、ATE 数据、SH 数据、个人信息整理成 tuple，用 `Origin_data` 接收返回值。

```
def ReadData(root1,root2):

    #Read Annotation
    ATE_ann = wfdb.rdann(root+root1, 'atr')
```

```

ATE_ann_sample = ATE_ann.sample
#ATE_ann_symbol = ATE_ann.symbol
#ATE_ann_aux_note = ATE_ann.aux_note

#Read ATE_Data
ATE_dat = wfdb.rdsamp(root+root1)
ATE_dat_sig = ATE_dat.p_signals
#ATE_dat_fs = ATE_dat.fs
#ATE_dat_units = ATE_dat.units
#ATE_dat_signame = ATE_dat.signame
#ATE_dat_comments = ATE_dat.comments

#Read SH_Data
SH_dat = wfdb.rdsamp(root+root2)
SH_dat_sig = SH_dat.p_signals
#SH_dat_fs = SH_dat.fs
#SH_dat_units = SH_dat.units
#SH_dat_signame = SH_dat.signame

comments = SH_dat.comments
age = int(comments[0].split(':')[1])
gender = int(comments[1].split(':')[1] == 'M')
height = int(comments[2].split(':')[1])
weight = int(comments[3].split(':')[1])
const_info = np.array([age,gender,height,weight])
return (ATE_ann_sample, ATE_dat_sig, SH_dat_sig, const_info)

```

使用字符串的 *split* 方法将个人信息整理成 numpy。

2.2 常量处理

整理系统所通用的数据信息，存储在 Constant_Value 中。

```

Constant_Value = {
    'ATE_ann_aux_note': ['Relax', 'PhysicalStress', 'Relax', 'EmotionalStress',
                        'CognitiveStress', 'Relax', 'EmotionalStress', 'Relax'],
    'ATE_dat_fs': 8,
    'SH_dat_fs': 1,
    'ATE_dat_units': ['NU', 'NU', 'NU', 'degC', 'NU'],
    'SH_dat_units': ['%', 'bpm'],
    'ATE_dat_signame': ['ax', 'ay', 'az', 'temp', 'EDA'],
    'SH_dat_signame': ['SpO2', 'hr']
}

```