

December 9, 2019

はじめに
提案手法
音声認識
おわりに
コード

音声のテキスト化

沼田 賢一

富山県立大学 情報基盤工学講座

1. はじめに
2. セクション名
3. おわりに

December 9, 2019

はじめに

2/16

本研究の背景

現代、多くの人がスマートフォンやウェアラブルデバイスを持ち歩くことが一般的であり、急速な情報技術の発達から、個人の生活や行動をデータとして取得、記録することが可能となっている。

このようなスマートフォンやウェアラブルデバイスを使用して取得して得られる人間の活動の記録のデータをログデータという。ログデータは解析を行うことで、個人の健康管理に活かしたり、ビジネスとしてターゲティング広告にも使われ社会に活かしたりできると考えられている。

目的

従来の提案手法による行動パターン解析における分類精度の向上を図る。

2/16

提案手法

3/16

分類精度向上のための提案手法

- ① 入力信号によるラベル付け
- ② テキストデータのベクトル化

はじめに
提案手法
音声認識
おわりに
コード

3/16

入力信号によるラベル付け

4/16

ライログデータの収集時に画像のテキスト変換と同様に、同時に音声マイクを用いて、現在の行動を認識できるような簡易的なテキスト変換 API を用いて定期的に入力信号として収集することでライログデータの解析する際に入力信号を教師データとして含めることで分類精度を向上させる。

テキストデータのベクトル化

5/16

Word2vec は Mikolov らにより発表された単語群のベクトル化手法である。Word2vec を自然言語処理の分野で応用した例に、日本語動詞・形容詞に関する類似度データセットの構築やイベント情報の分類など数多くの利用例がある。本研究では、カメラから取得した画像を画像認識 API を通してテキスト化させたものに対して Word2vec を用いてベクトル化させる。そして家具、動作、食べ物といった 6 つの基準を設けそれらとの類似度の値を求める。

5/16

目的

6/16

前回と同様に目的は、行動識別の時に用いるラベルをマイクを通して音声をテキスト化したものを保存できることにする。

前回の研究会で紹介してもらった SpeechRecognition という API を使って試してみた。

6/16

はじめに

提案手法

音声認識

おわりに

コード

SpeechRecognition とは、Python Package Index にある音声認識を実行するためのライブラリ。

音声を扱うので python に追加で portaudio と pyaudio をインストールする。

```
sudo apt-get install portaudio19-dev
sudo apt-get install python-pyaudio python3-pyaudio
pip install pyaudio
```

SpeechRecognition API

8/16

ラベルのみをとりたいので常に音声認識を保存しないために、ホットワード検出後にラベル保存のための音声認識をするシステムを作った。

ホットワードは、なんでもよかったです「Ok Google」。
以下の画像が実行した結果。

ホットワード「Ok Google」が認識されたあとにラベル保存のための認識がされている。

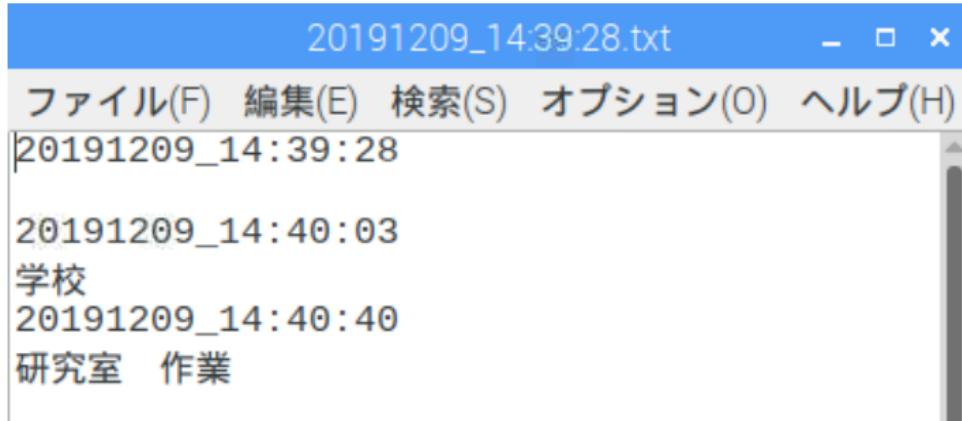
```
Start recognized by Ok Google ...
Now to recognize it...
Ok Google
Say something ...
Now to recognize it...
学校
```

8/16

SpeechRecognition API

9/16

ラベルの実行結果は日付・時間と音声認識の結果とセットでファイルに保存するようにした。



20191209_14:39:28.txt

ファイル(F) 編集(E) 検索(S) オプション(O) ヘルプ(H)

20191209_14:39:28

20191209_14:40:03

学校

20191209_14:40:40

研究室 作業

9/16

はじめに
提案手法
音声認識
おわりに
コード

ホットワードの認識が失敗した場合、ラベルが保存されない。
ラズベリーパイの画面を見ていれば実行の遷移がみれるが、実際にデータを集めるとときに画面を見るることはできない。
そこで LED(発光ダイオード) をブレッドボードに差し込んで光り方で遷移の状態をわかるようにした。

また、GPIO（デジタル信号の入出力インターフェース）を制御するために

```
sudo apt-get install python-rpi.gpio  
で GPIO を制御するライブラリを入れた。
```

SpeechRecognition API

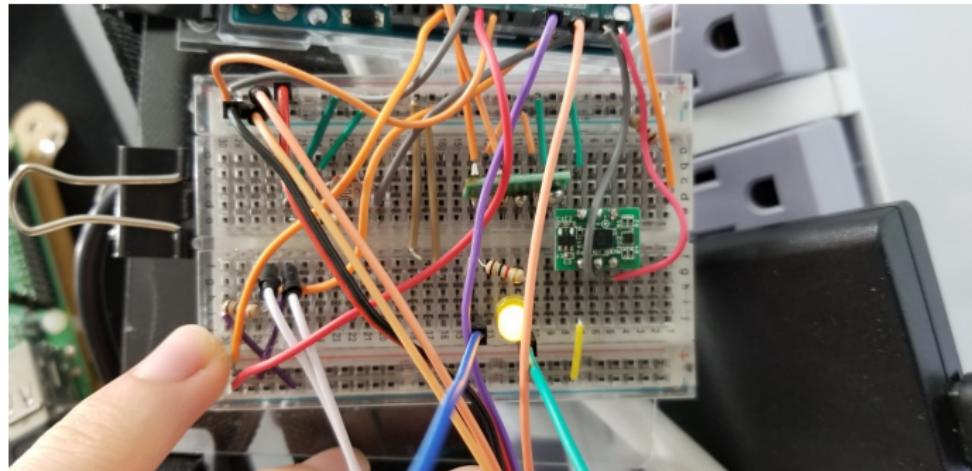
11/16

はじめに
提案手法
音声認識
おわりに
コード

使ったものは、LED(発光ダイオード)と $220\ \Omega$ の抵抗とラズベリーパイとブレッドボードの間が広めだったのでジャンプワイヤを4本使った。

光り方は、認識前が光っていて認識中は光らないようにした。

また、ホットワードが認識されたら5回点滅、保存がされたら2回点滅するようにした。



11/16

まとめ

- ① マイク（今回はカメラ）を使った音声認識ができるようになった。
- ② ラベルの保存を日付と一緒にテキストに自動で保存するようにした。
- ③ LEDにより、ラズベリーパイの画面を見ないで音声認識できるようにした。

コード

13/16

はじめに
提案手法
音声認識
おわりに
コード

```
1 # -*- coding: utf-8 -*-
2 import speech_recognition as sr
3 from datetime import datetime
4
5 import RPi.GPIO as GPIO
6 import time
7 GPIO.setmode(GPIO.BOARD)
8 GPIO.setup(11, GPIO.OUT)
9
10 nowtime = datetime.now().strftime('%Y%m%d_%H:%M:%S')
11 newtxt= nowtime+".txt"
12 with open(newtxt,'w') as f:
13     f.write(nowtime+"\n")
14
15 r = sr.Recognizer()
16 mic = sr.Microphone()
17
18 while True:
19
20     print("Start recognized by Ok Google ...")
21     GPIO.output(11, True)
22     with mic as source:
23         r.adjust_for_ambient_noise(source)
24         audio = r.listen(source)
```

コード

14/16

```
26     print ("Now to recognize it...")
27     GPIO0.output(11, False)
28
29     try:
30         print(r.recognize_google(audio, language='ja-JP'))
31         GPIO0.output(11, False)
32
33         if r.recognize_google(audio, language='ja-JP') == u"スト
34             GPIO0.output(11, False)
35             print("end")
36             break
37
38         if r.recognize_google(audio, language='ja-JP') == u"Ok 0
39             for i in range(5):
40                 GPIO0.output(11, True)
41                 time.sleep(0.05)
42                 GPIO0.output(11, False)
43                 time.sleep(0.05)
44
45
46         print("Say something ...")
47         GPIO0.output(11, True)
48         with mic as source:
49             r.adjust_for_ambient_noise(source)
50             audio = r.listen(source)
```

はじめに
提案手法
音声認識
おわりに
コード

コード

```
51
52     print ("Now to recognize it...")
53     GPIO.output(11, False)
54
55     print(r.recognize_google(audio, language='ja-JP'))
#seikou sitara 2kaitenmetsu
56
57     for i in range(2):
58         GPIO.output(11, True)
59         time.sleep(0.05)
60         GPIO.output(11, False)
61         time.sleep(0.05)
62
63     with open(newtxt, 'a') as f:
64         f.write("\n"+datetime.now().strftime('%Y%m%d_%H
65         f.write("\n"+r.recognize_google(audio, language=
66         GPIO.output(11, False)
67
68 except sr.UnknownValueError:
69     print("could not understand audio")
70 except sr.RequestError as e:
71     print("Could not request results from Google Speech Rec
```