

第7回

検証する条件: 録音機器(パソコン or スマホ)による照合制度への影響

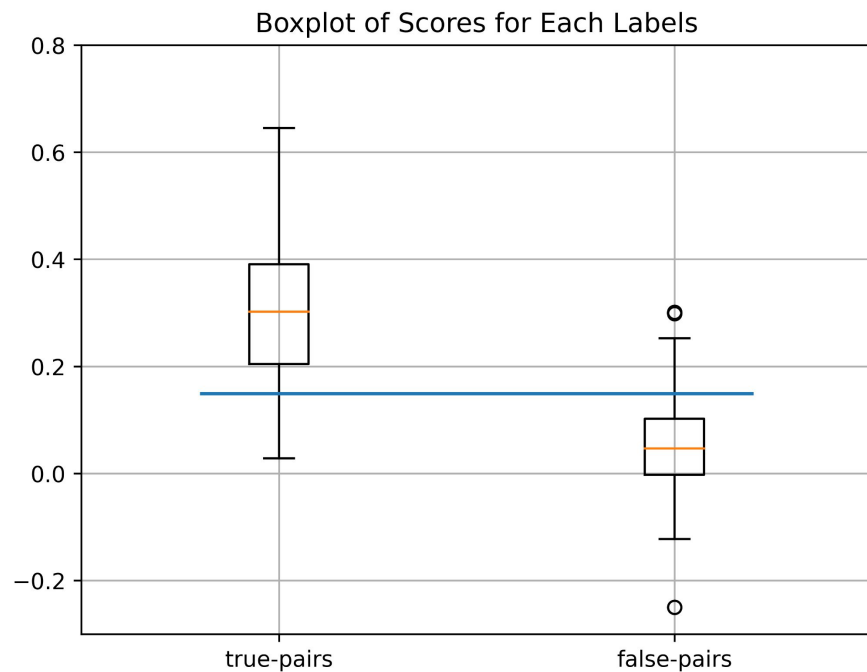
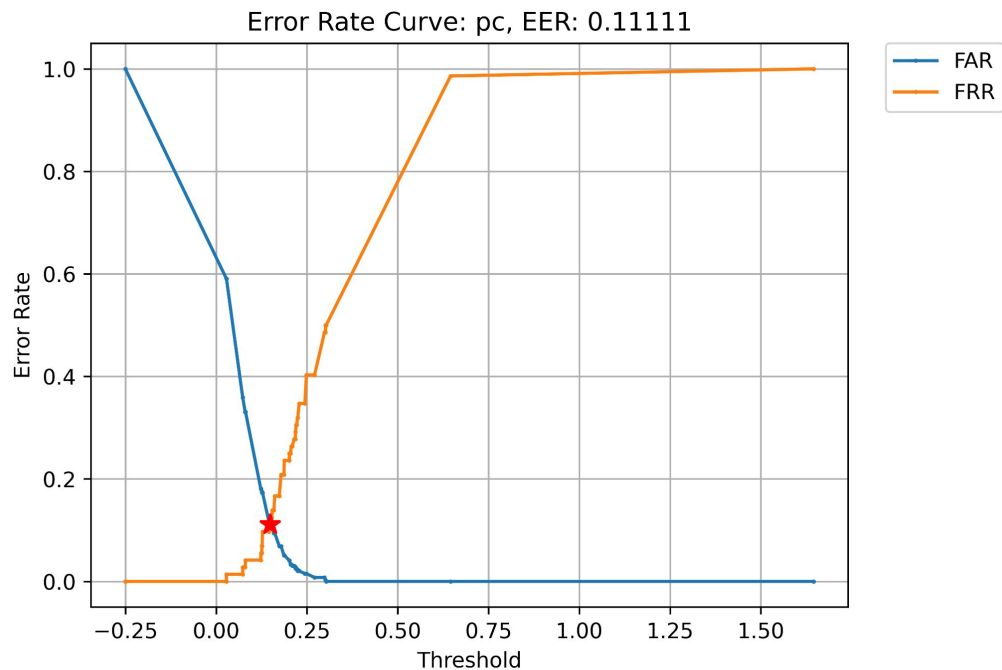
- パソコン録音orスマホ録音の各条件について, 発話対リストを作成し, 照合する.

音声:

- 条件間で発話内容発話数は同じ.
- 音声の切り出しタイミングは若干異なる.
- 男3人女3人
 - 計31発話
 - 男(6+7+3) + 女(6+6+3)
 - 465対
 - 同一話者(72)+異話者(393)

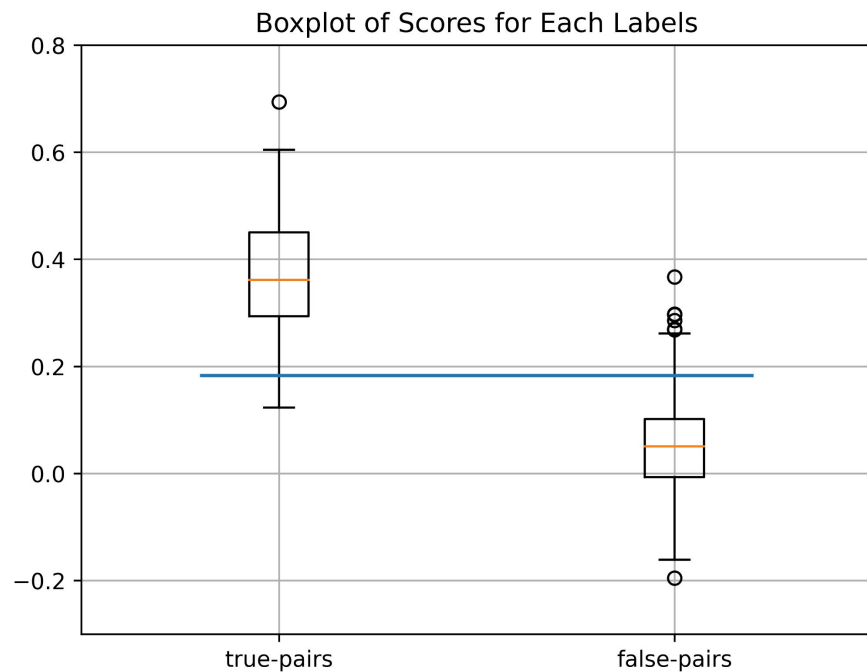
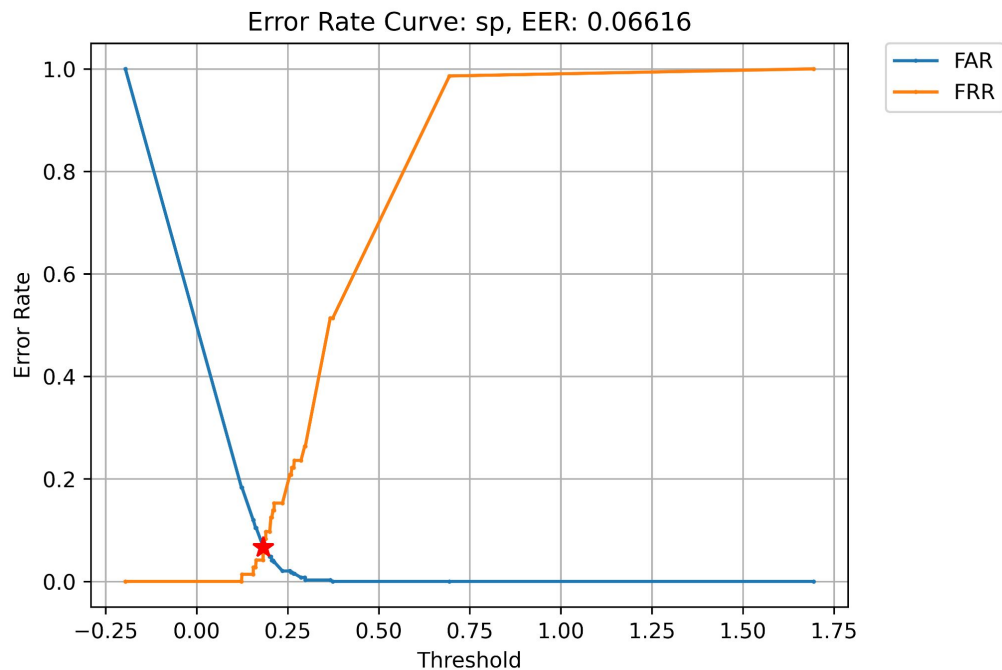
パソコン録音

EER = 0.111



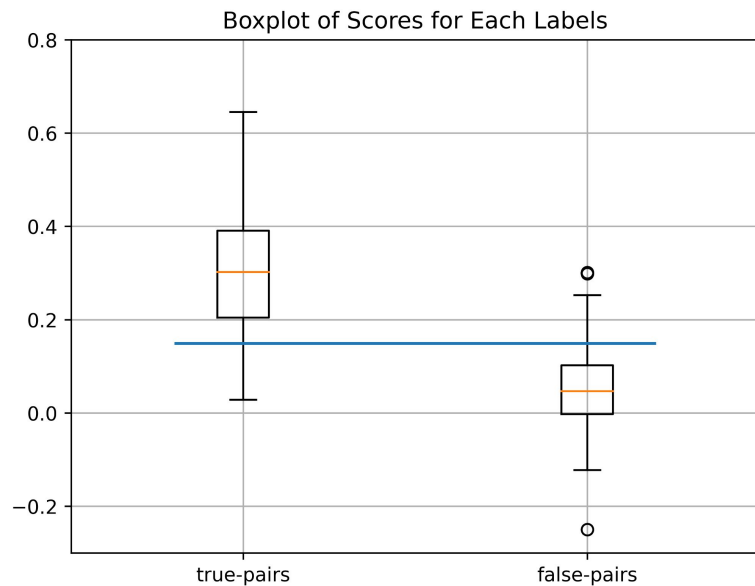
スマホ録音

EER = 0.066

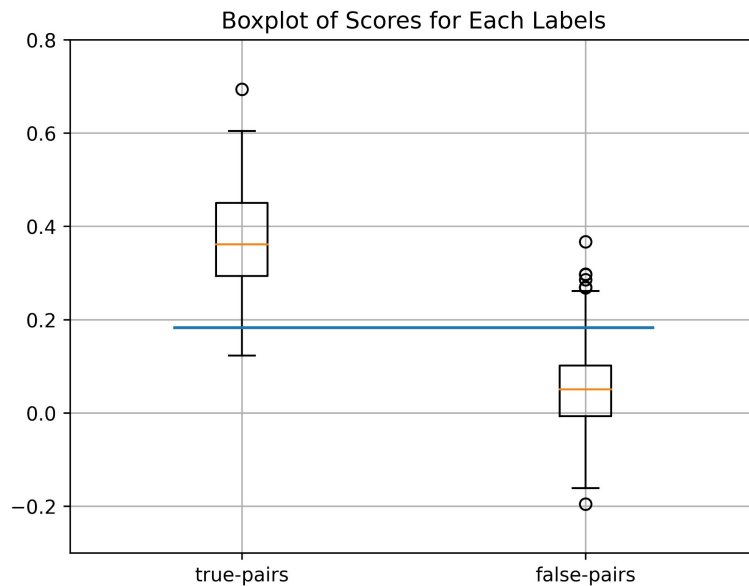


パソコン, スマホ, 比較

EER_{pc} = 0.111



EER_{sp} = 0.066



pcの方は, FNが多いように見える(true-pairの分散が大きい)

FNs

FN_pc

```
FalseNegatives(eer_th:0.14905281489126118)
score:0.12709996104240417 /content/drive/MyDrive/CS_seminar/data/wav/ymgt_A/pc/ymgt-A_pc_01.wav, /content/drive/MyDrive/CS_seminar/data/wav/ymgt_A/pc/ymgt-A_pc_02.wav
score:0.028317326679825783 /content/drive/MyDrive/CS_seminar/data/wav/ymgt_A/pc/ymgt-A_pc_02.wav, /content/drive/MyDrive/CS_seminar/data/wav/ymgt_A/pc/ymgt-A_pc_06.wav
score:0.12670762836933136 /content/drive/MyDrive/CS_seminar/data/wav/ymgt_C/pc/ymgt-C_pc_01.wav, /content/drive/MyDrive/CS_seminar/data/wav/ymgt_C/pc/ymgt-C_pc_02.wav
score:0.07980848848819733 /content/drive/MyDrive/CS_seminar/data/wav/ymgt_C/pc/ymgt-C_pc_01.wav, /content/drive/MyDrive/CS_seminar/data/wav/ymgt_C/pc/ymgt-C_pc_03.wav
score:0.0730365440249443 /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/pc/ymgt-D_pc_01.wav, /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/pc/ymgt-D_pc_05.wav
score:0.1228519007563591 /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/pc/ymgt-D_pc_01.wav, /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/pc/ymgt-D_pc_06.wav
score:0.12583622336387634 /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/pc/ymgt-D_pc_02.wav, /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/pc/ymgt-D_pc_05.wav
score:0.14505170285701752 /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/pc/ymgt-D_pc_04.wav, /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/pc/ymgt-D_pc_06.wav
```

FN_sp

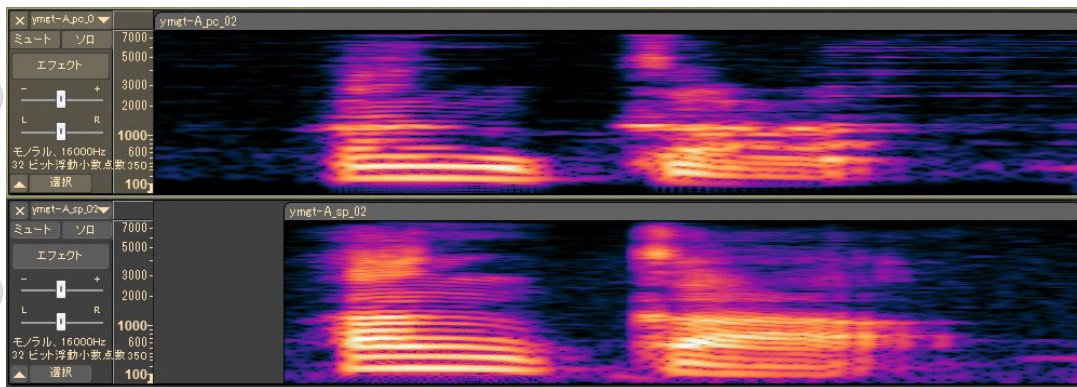
```
FalseNegatives(eer_th:0.1831274926658716)
score:0.15528130531311035 /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/sp/ymgt-D_sp_01.wav, /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/sp/ymgt-D_sp_05.wav
score:0.16189070045948029 /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/sp/ymgt-D_sp_02.wav, /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/sp/ymgt-D_sp_03.wav
score:0.18226751685142517 /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/sp/ymgt-D_sp_04.wav, /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/sp/ymgt-D_sp_06.wav
score:0.12314678728580475 /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/sp/ymgt-D_sp_05.wav, /content/drive/MyDrive/CS_seminar/data/wav/ymgt_D/sp/ymgt-D_sp_06.wav
```

- pcだとymgt_A, ymgt_Cが弾かれている
- 山口が用意した音声だけ弾かれており、僕のpc録音環境が劣悪なせいでpcのスコアが下がっているかも...
 - 設定は変えていないが、ノイズが乗ったり乗らなかったり、音量がまちまちであったりした

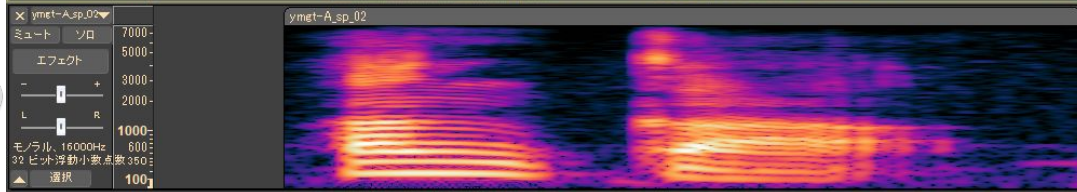
音声確認

pcは1500~5000Hzの成分が少なく、その違い？
wavの長さも違う

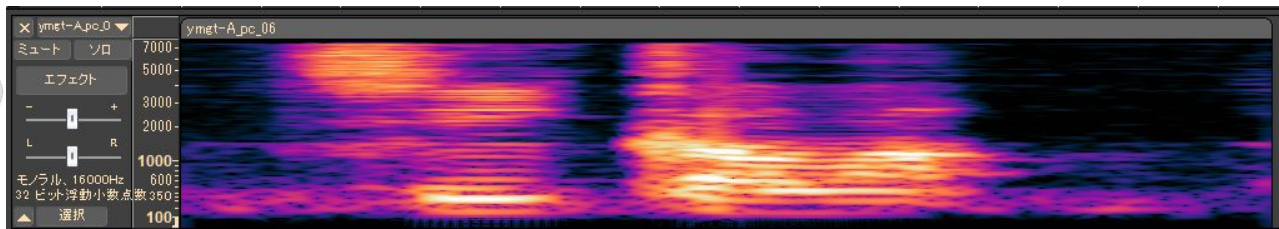
ymgt_A_pc_02: 



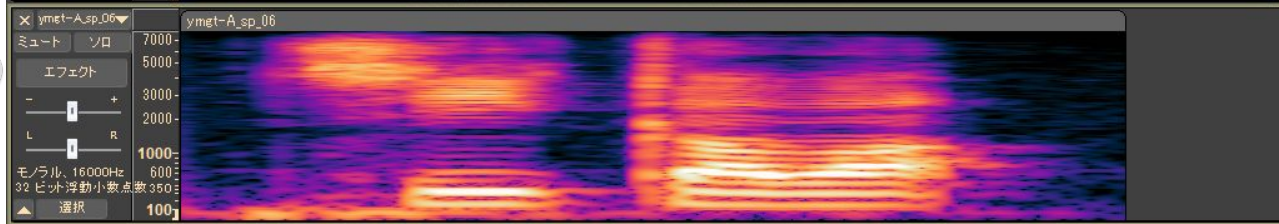
ymgt_A_sp_02: 



ymgt_A_pc_06: 



ymgt_A_sp_06: 



考察

- $pc_eer = 0.111$, $sp_eer = 0.066$
 - エラー率5pointは誤差レベル？
- スコア分布は, pc , sp で比較すると, $pc/true\text{-}pair$ の分散が大きい
 - pc の音声は1500~5000Hzあたりの成分が少なく, これが影響している？
 - 実際聞いてみるとかなり違う, sp の方が自然に聞こえる
 - pc , sp で発話内容は同じだが, wav の長さが違い, このため？
 - 完璧に同じ長さにするのは難しい
 - 時間が無く, できるだけ同じ長さにもできなかった (ごめんなさい)
 - FNのペアは山口が用意した音声だけ
 - えっ...私の録音環境, 劣悪すぎ ... ?
 - 設定は変えていないが, 音量やノイズがまちまちであった (なんで?)

ymgt_A_pc_06:



ymgt_A_sp_06:



hre_A_pc_1:



hre_A_sp_1:



next

フィルターを掛ける

いろいろな周波数を切ってみて, 精度を確認

全員の結果をmergeして1つのスライドにまとめる

soxで長さを揃える