

Digital Speech Processing HW2

網媒所 吳旻昇 r07944005

Part 1: Run Baseline

```
vincentwu@vincentwu-Aspire-T6000: ~/NTU2019DSP/hw2
vincentwu@vincentwu-Aspire-T6000:~/NTU2019DSP/hw2$ bash 04_testing.sh

HTK Configuration Parameters[13]
Module/Tool      Parameter      Value
#               NATURALWRITEORDER      TRUE
#               NATURALREADORDER      TRUE
#               ENORMALIZE      TRUE
#               NUMCEPS      12
#               CEPLIFTER      22
#               NUMCHANS      26
#               PREEMCOEF      0.970000
#               USEHAMMING      TRUE
#               WINDOWSIZE      320000.000000
#               SAVEWITHCRC      FALSE
#               SAVECOMPRESSED      FALSE
#               TARGETRATE      100000.000000
#               TARGETKIND      MFCC_Z_E_D_A

WARNING [-8232] ExpandWordNet: Pronunciation 1 of sp is 'tee' word in HVite

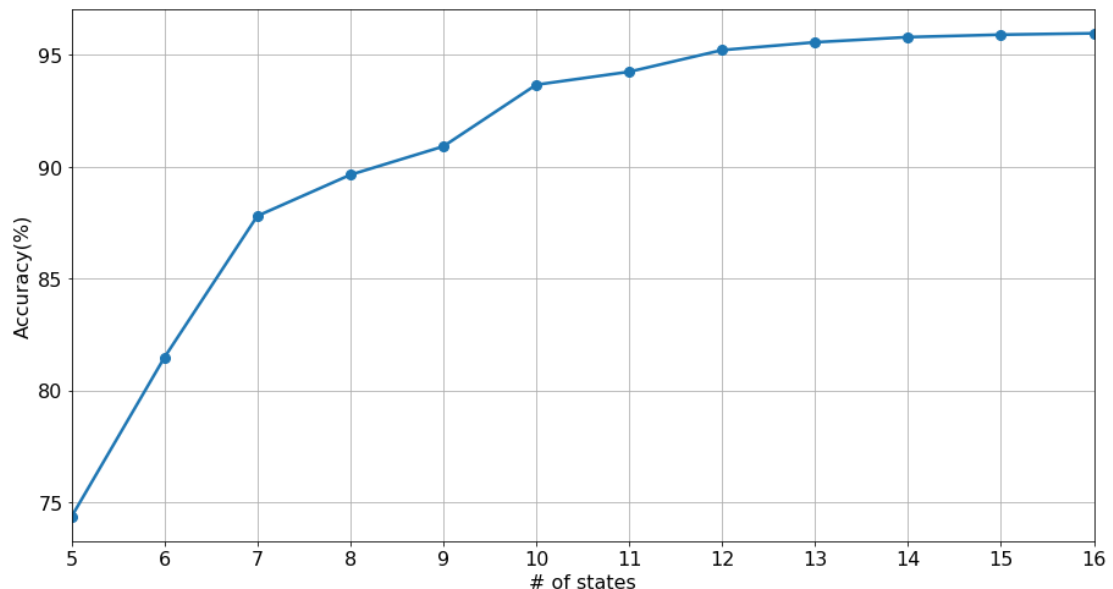
HTK Configuration Parameters[13]
Module/Tool      Parameter      Value
#               NATURALWRITEORDER      TRUE
#               NATURALREADORDER      TRUE
#               ENORMALIZE      TRUE
#               NUMCEPS      12
#               CEPLIFTER      22
#               NUMCHANS      26
#               PREEMCOEF      0.970000
#               USEHAMMING      TRUE
#               WINDOWSIZE      320000.000000
#               SAVEWITHCRC      FALSE
#               SAVECOMPRESSED      FALSE
#               TARGETRATE      100000.000000
#               TARGETKIND      MFCC_Z_E_D_A

vincentwu@vincentwu-Aspire-T6000:~/NTU2019DSP/hw2$ cat result/accuracy
===== HTK Results Analysis =====
Date: Sat Apr 13 17:22:59 2019
Ref : labels/answer.mlf
Rec : result/result.mlf
----- Overall Results -----
SENT: %Correct=38.54 [H=185, S=295, N=480]
WORD: %Corr=96.61, Acc=74.34 [H=1679, D=13, S=46, I=387, N=1738]
=====
vincentwu@vincentwu-Aspire-T6000:~/NTU2019DSP/hw2$
```

Part 2: Improve Accuracy

(1) Modify the model format in "lib/proto" to change the # of states

此部份的實驗為更改 states 數量，觀察 accuracy 的變化情形，在固定 baseline 的初始值 (Gaussian mean, variance, and transition probability values) 的情況下，逐步調整 states 數量，可繪出如下的曲線圖， accuracy 從 baseline 的 74.34% 上升至 95.97% 。



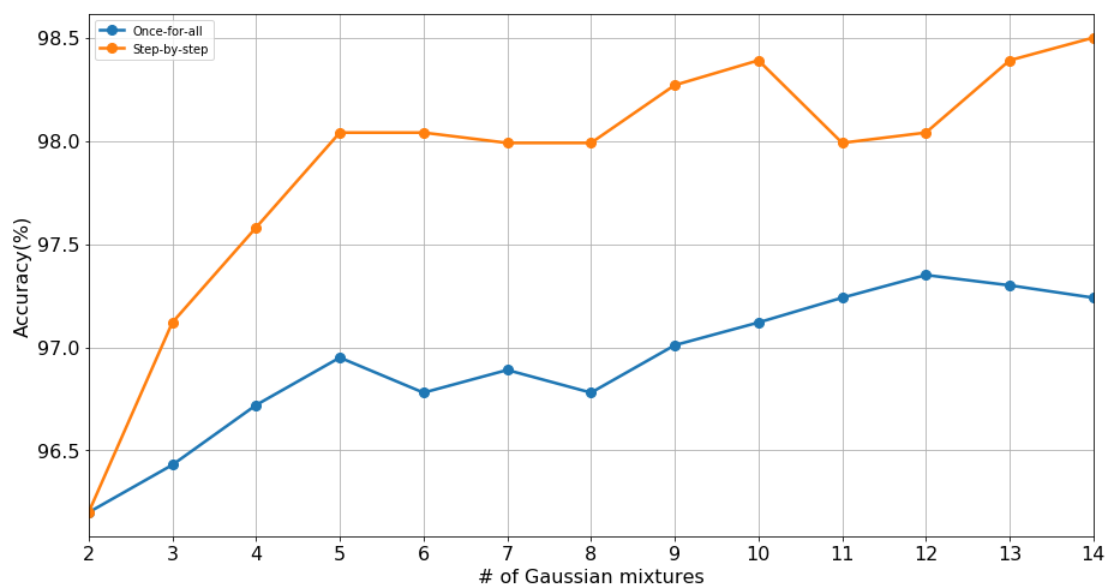
(2) Modify the # of Gaussian mixtures in "lib/mix2_10.hed"

此部份的實驗為更改 Gaussian mixture models 的數量，觀察 accuracy 的隨之變化的情形，而初始值的參數以及 states 數量我採用上個實驗的最佳值 (# of states = 16) ，另外，我在這裡將欲增加 Gaussian mixtures 的 states 均設為第 2~16 個 state 。

一開始的設定是在 "零~九" 每個 MFCC 維度以 2 個 Gaussian mixtures 來表示，而 "sil" 以 3 個 Gaussian mixtures 來表示，此時的 accuracy 為 96.20% ，而增加 Gaussian mixtures 的方式在本實驗分為兩種策略：

- 1) Once-for-all: 一次加到指定的 Gaussian mixtures 。
- 2) Step-by-step: 再每次 Re-estimate 之前增加 1 個 Gaussian mixtures ，逐步遞增至指定數量。

下圖為兩種策略的 accuracy 結果 (x 軸為 "零~九" 的 Gaussian mixtures 數量， "sil" 則加 1)，可以發現有 Re-estimate 做 tune-tuning 的緣故，Step-by-step 比 Once-for-all 的方法來的好，增加到 14 個 Gaussian mixtures 時，可達到 98.50% 。



Part 3: Training details

# of States	16
# of Gaussian Mixture Models	14
Gaussian Adding Strategy	Step-by-step
# of Re-estimate Iteration	6

最終達到的結果為：

```

===== HTK Results Analysis =====
Date: Sat Apr 20 14:32:37 2019
Ref : labels/answer.mlf
Rec : result/result.mlf
----- Overall Results -----
SENT: %Correct=95.21 [H=457, S=23, N=480]
WORD: %Corr=98.50, Acc=98.50 [H=1712, D=21, S=5, I=0, N=1738]
=====

```