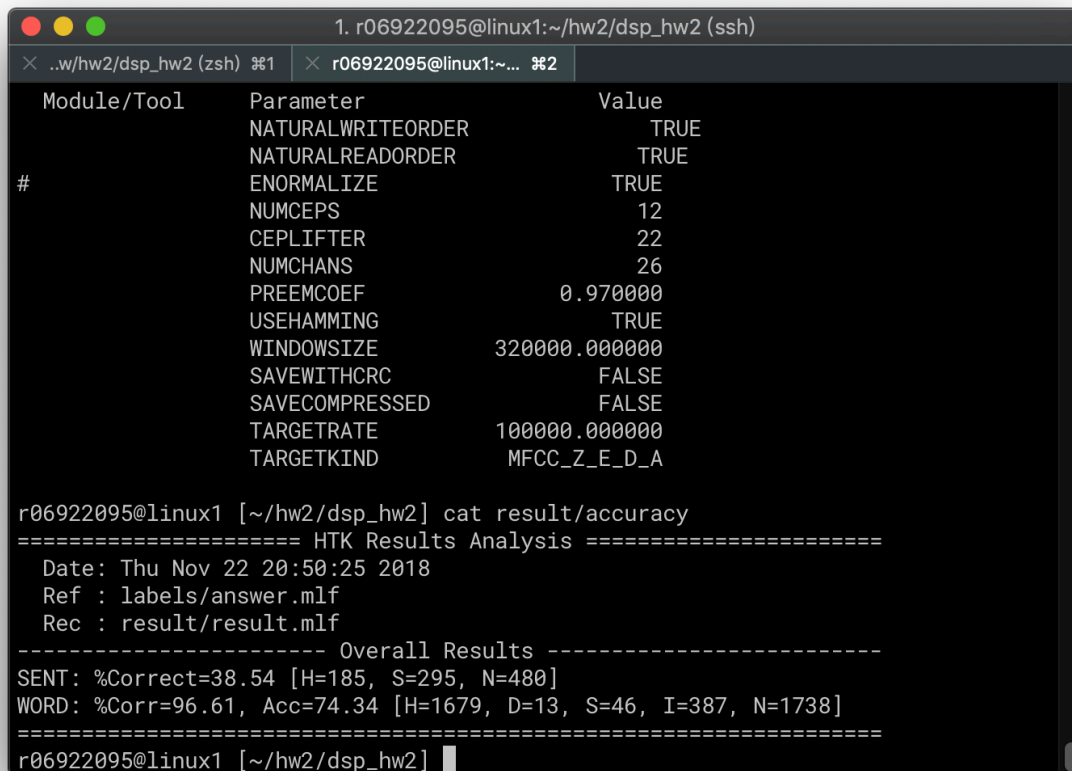


# Digital Speech Processing hw2

## Part 1: Run Baseline



```
1. r06922095@linux1:~/hw2/dsp_hw2 (ssh)
x ../hw2/dsp_hw2 (zsh) %1 x r06922095@linux1:~... %2
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

r06922095@linux1 [~/hw2/dsp_hw2] cat result/accuracy
===== HTK Results Analysis =====
Date: Thu Nov 22 20:50:25 2018
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]
=====
r06922095@linux1 [~/hw2/dsp_hw2] █
```

## Part 2: Improve Accuracy

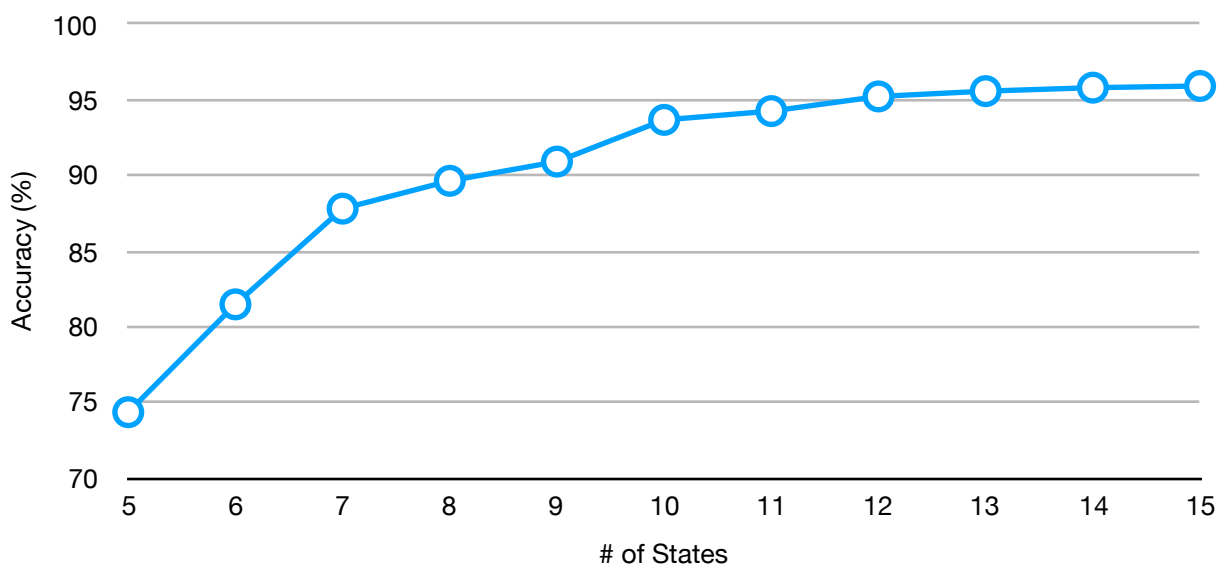
### lib/proto

可以改的參數為States的數量、初始Gaussian mean及variance、初始Transition

Probability。試過改變初始值，但最終影響不大，所以主要調整States的數量。下圖為

States數量對Accuracy做圖，其他參數為Baseline參數。可以看到States的數量影響很

大，在數量為15時，大約收斂在95%。



### lib/mix2\_10.hed

可以改變Gaussian的數量，Baseline的設定是每個維度用2個Gaussian Mixture Model表

示，Tuning參數時一次改變0~9及sil共11個model的Gaussian數量，sp的數量則保留

Baseline的參數不改變。實驗兩種改變的策略，第一種：1-step，一次增加Gaussian的數

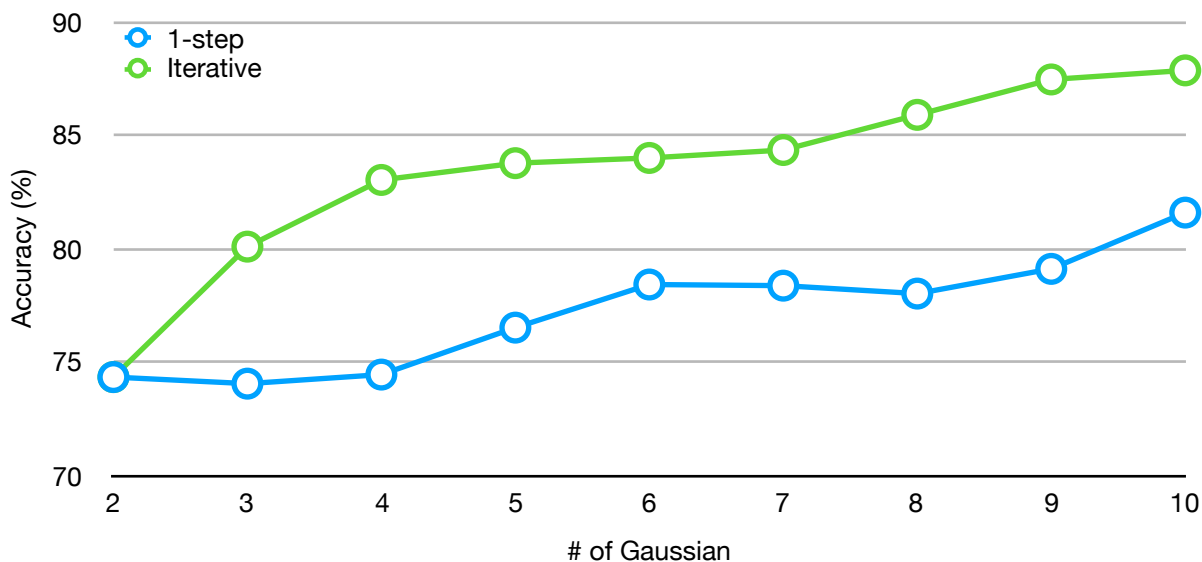
量，從一個直接增加到指定數量。第二種：Iterative，利用Fine-tuning的技巧，一次只增

加一個Gaussian，Re-estimate後再增加一個Gaussian，重複上述步驟直到指定數量。下

圖為Gaussian數量對Accuracy做圖，其他參數為Baseline參數。從中可以發現，Iterative

效果明顯比1-step好，在5個Gaussian時卡在84%，但在8個Gaussian後又持續上升。反

之1-step上升緩慢，效果也較差。Iterative每次都會重新Re-estimate，效果較好也是合理的。



## Final Model

最終參數設置為：

# of States	# of Gaussian	# of Re-estimate Iteration	Gaussian Adding Method
15	10	6	Iterative

最終準確率如下：

Accuracy為98.04%

```
===== HTK Results Analysis =====
Date: Fri Nov 23 01:34:12 2018
Ref : labels/answer.mlf
Rec : result/result.mlf
----- Overall Results -----
SENT: %Correct=93.75 [H=450, S=30, N=480]
WORD: %Corr=98.10, Acc=98.04 [H=1705, D=28, S=5, I=1, N=1738]
=====
```