

Digital Speech Processing HW2-1

楊正彥 B03901086 電機三

Part 1 : Run Baseline

```
===== HTK Results Analysis =====
Date: Mon May 15 23:52:07 2017
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]
=====
```

Part 2 : Improve Recognition Accuracy

```
===== HTK Results Analysis =====
Date: Tue May 16 16:01:55 2017
Ref : labels/answer.mlf
Rec : result/result.mlf
----- Overall Results -----
SENT: %Correct=94.79 [H=455, S=25, N=480]
WORD: %Corr=98.39, Acc=98.33 [H=1710, D=24, S=4, I=1, N=1738]
=====
```

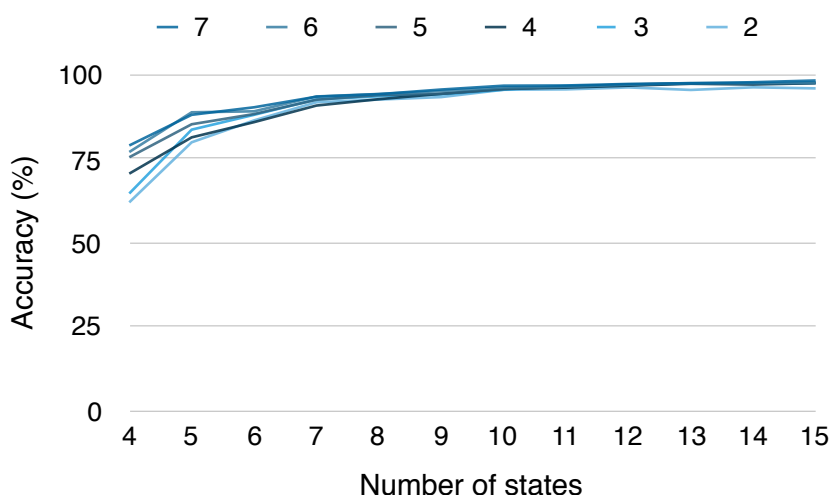
/lib/proto => State number = 15

/lib/mix2_10.hed => Number of Gaussian mixture = 15

/03_training.sh => HErest Iteration time = (8,8,14)

Part 3 : Discussion

- (1) The state number s have to be in the range of $4 \leq s \leq 16$, or else it can not execute properly.
- (2) In order to achieve higher accuracy, I wrote a script to help me tested all of the combination of different state number, number of Gaussian mixture and HErest iteration time. Generally, the accuracy increases simultaneously as the number



of state and Gaussian mixtures increase.

(3) The runtime of the program is approximately linear to the number of states and Gaussians.

(4) The trend of different number of states in any two lines share the same observation.