

Truncated Speech based Speaker Recognition

Bie Fanhu
CSLT, RIIT, THU

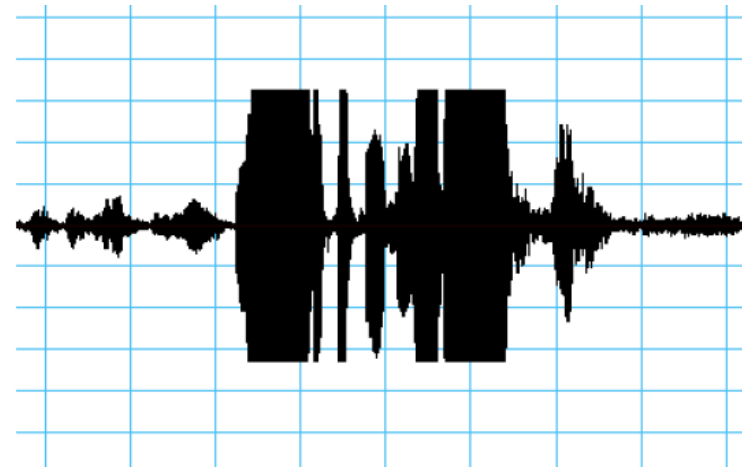
outline

- Truncated Speech Definition
- Analysis on Truncated Sine Signal
- Impact on Speaker Recognition
- Detection for Truncated Speech Segments
- Future Work

Truncated Speech

Figure for Truncated Speech

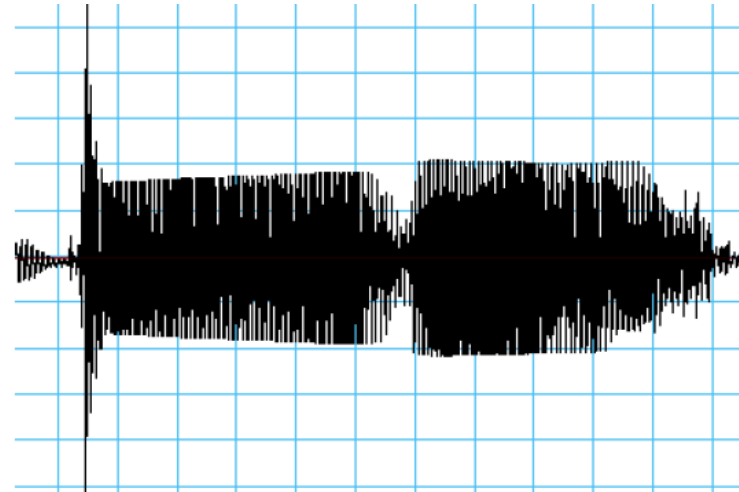
- The speech envelope is a straight line in some part
- The amplitude of the sample points exceed the range of AD
- The AD gives out the same the maximum value for the sample point.



Truncated Speech

Figure for Pseudo Truncated Speech

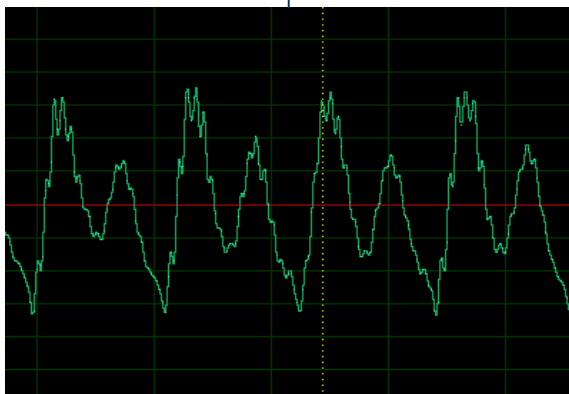
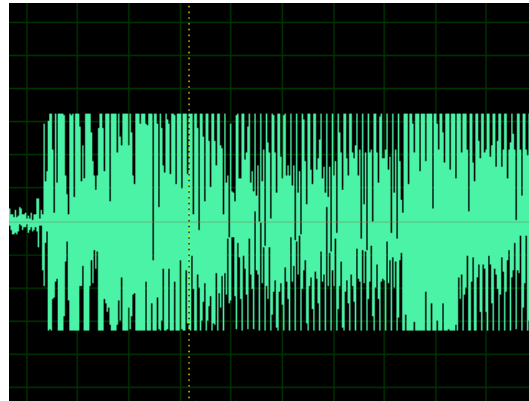
- The speech envelope is like a straight line in some part
- AGC (automatic gain control) has high gains on low volume and low gains on high volume



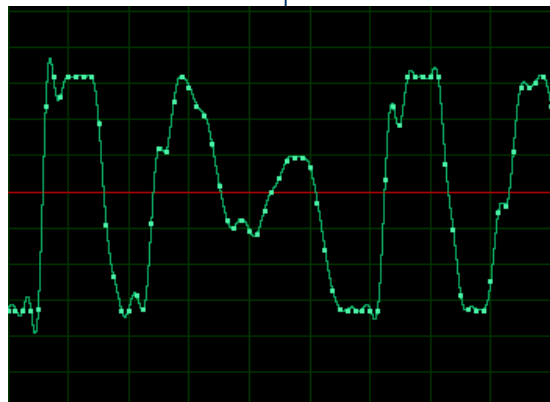
Real Truncated Speech

Different Types of Truncated Speech

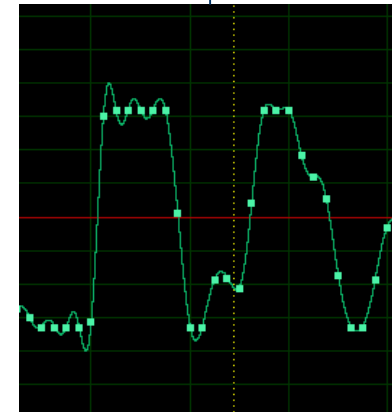
- Speech signal
Damage level



a) No damage



b) Little damage



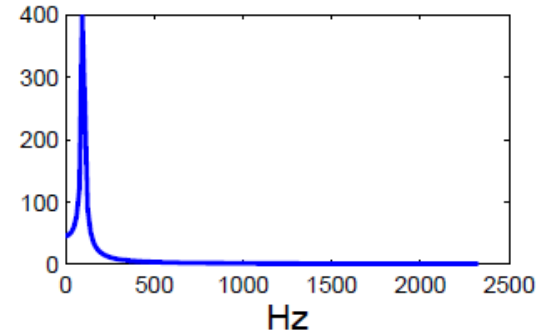
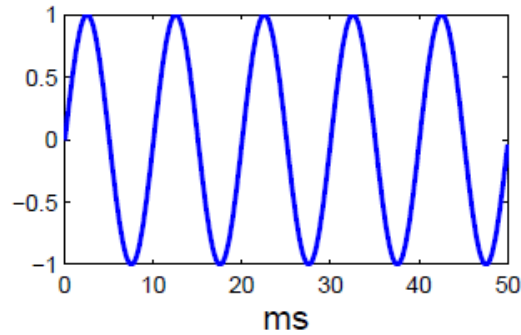
c) Serious damage

Analysis on Truncated Sine Wave

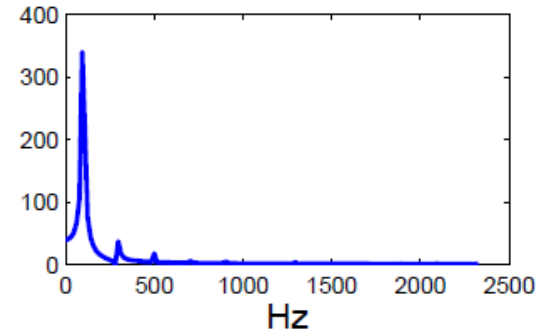
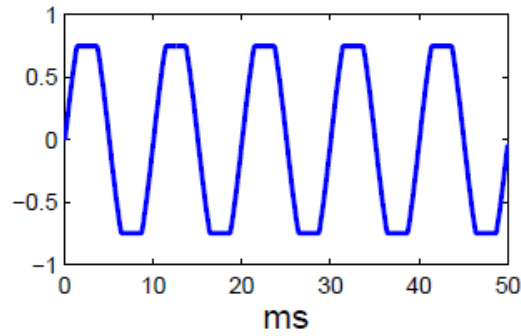
- Fourier analysis: The speech signal can be represented as a combination of simple sine waves
- Analyze on the sine wave, truncate the wave according to its maximum volume by 60% and 20%
- The energy extends to odd multiples of the basic frequency
- The higher the frequency of the harmonics is, the smaller energy has distributed to

Matlab Analysis

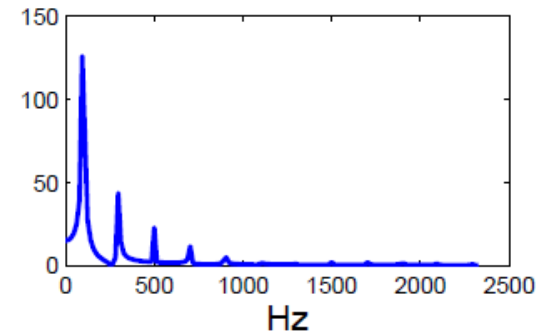
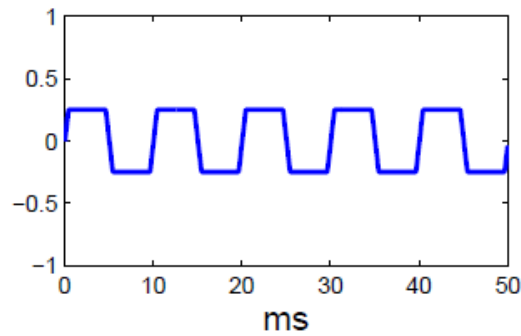
a) Original



b) 60%



c) 20%



Simulation on Speech

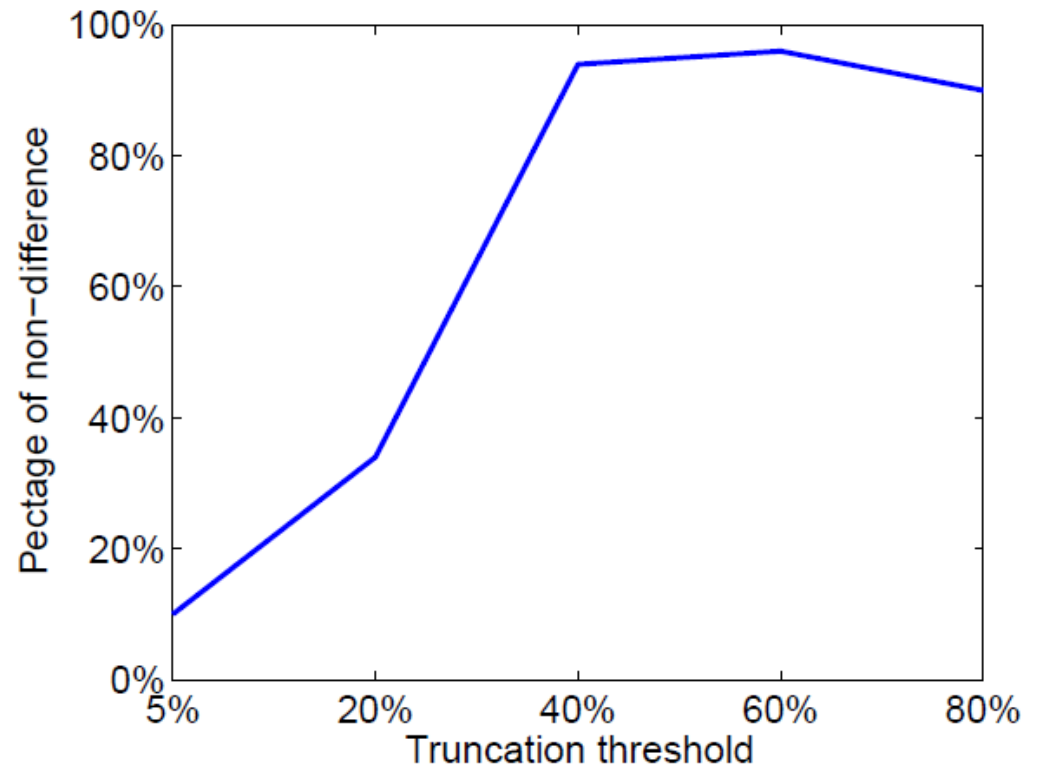
Simulation

- Database: Truncated the speech according to the maximum volume of each utterance by 10%, 20%, 40%, 60%. 100% stands for the original speech
- Hearing test: 5 persons listen to utterances from different truncated rate. Judge whether it is the original speech
- Performance: GMM-UBM, i-vector on different truncated rate database

Sine Wave Truncated

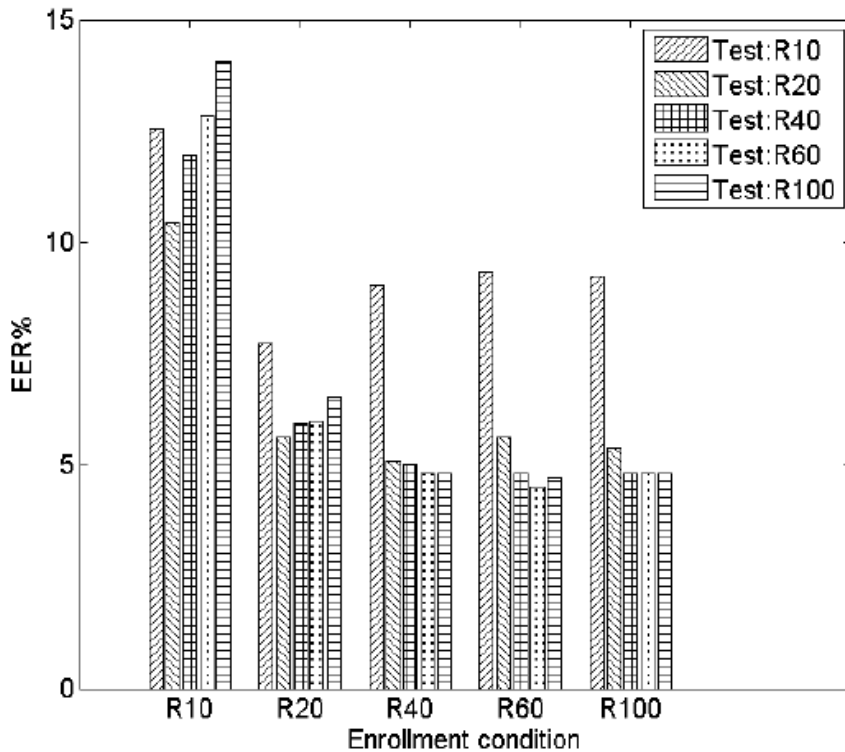
Hearing Test

- The lower truncated rate is, the greater the difference is
- Even truncated by 40%, there is still not great difference from the original speech

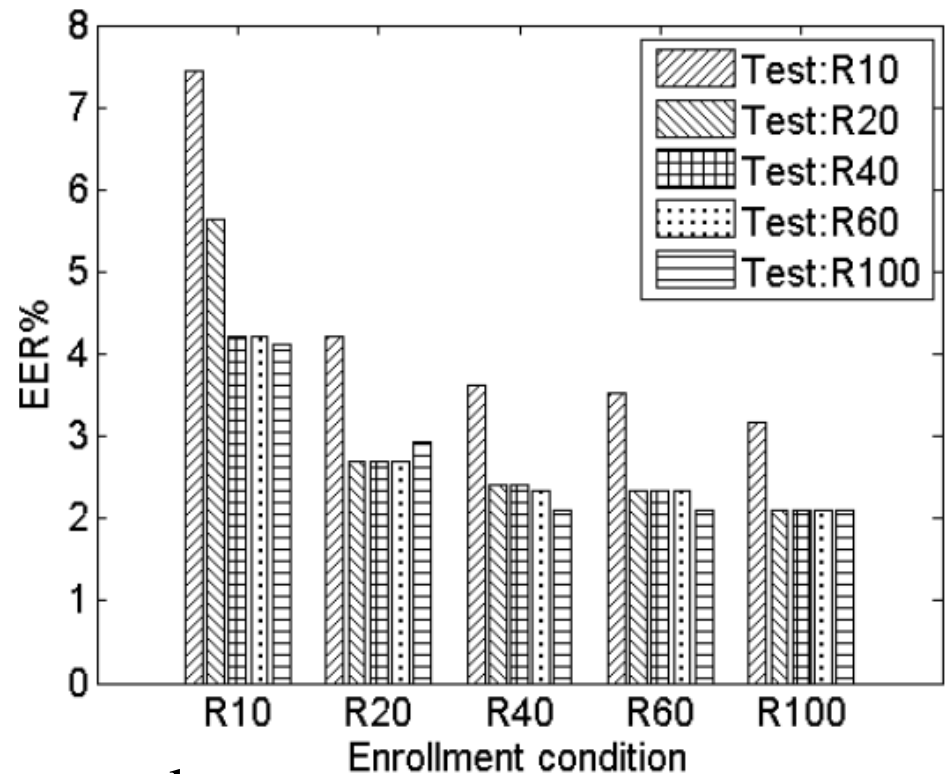


Speaker Recognition

GMM-UBM



i-vector



- Remain rate $> 40\%$: performance keeps
- Remain rate = 10%, performance decreases $> 50\%$
- Pseudo truncated condition: performance decreases sharply

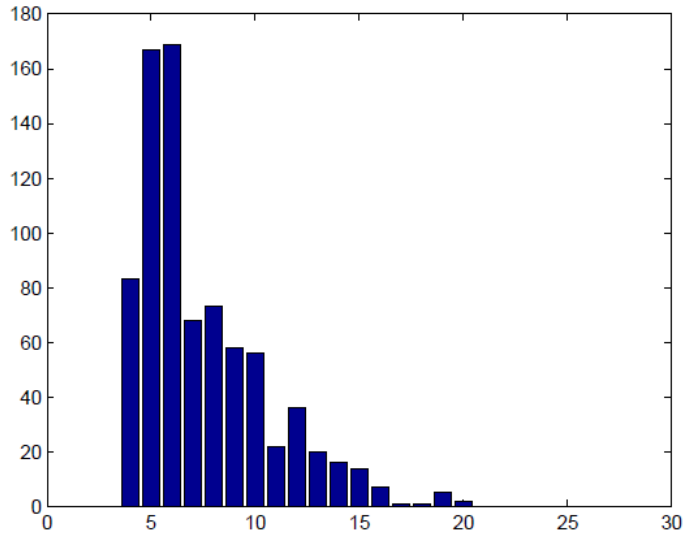
Sine Wave Truncated

Detection for Truncated Speech Segments

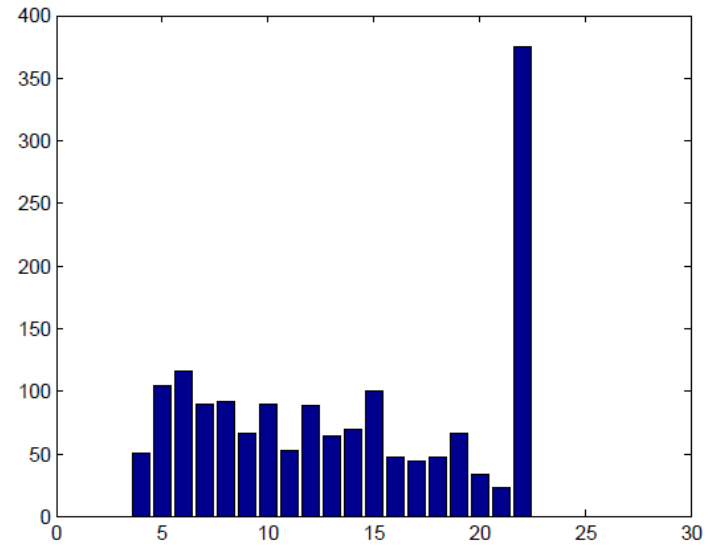
- I. Treat 0.5 s as a segment
- II. Separate into segments by the volume
- III. Count points to get the distribution for each segments
- IV. Judge if it is the truncated segment

Sine Wave Truncated

4) Histograms of distribution



Normal



Truncated

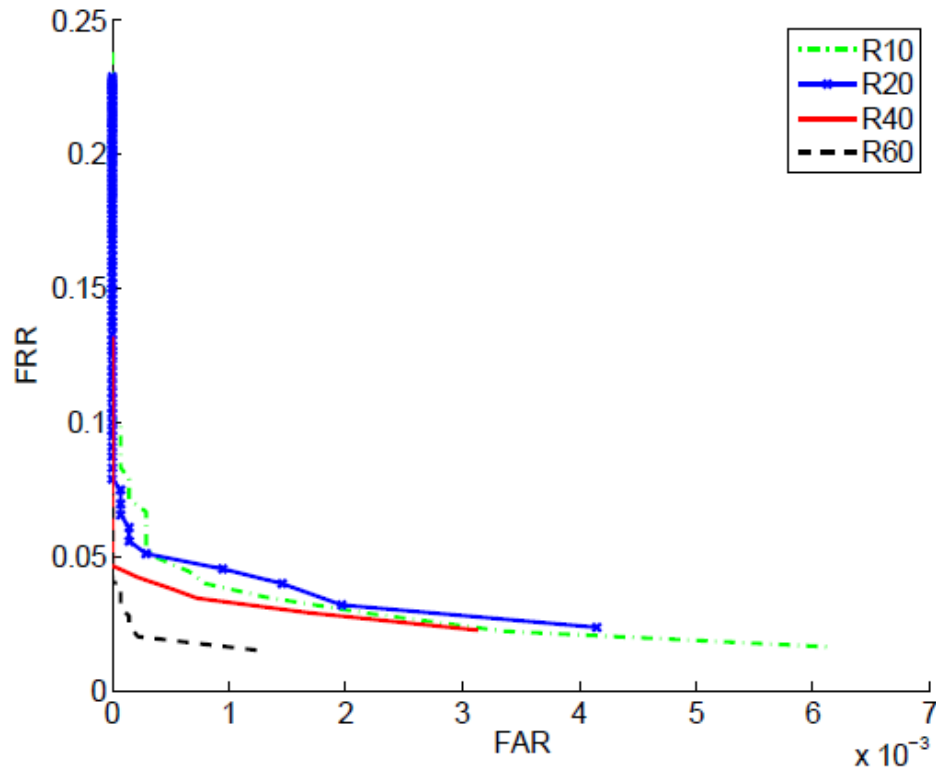
Sine Wave Truncated

Detection based on the histogram of distribution

- I. Find the maximum absolute value of the speech and separate into 20 intervals by the value
- II. Count the points located in the intervals of one 0.5s-segment
- III. If the point number of the last interval is larger than threshold, detect it as a truncated segment

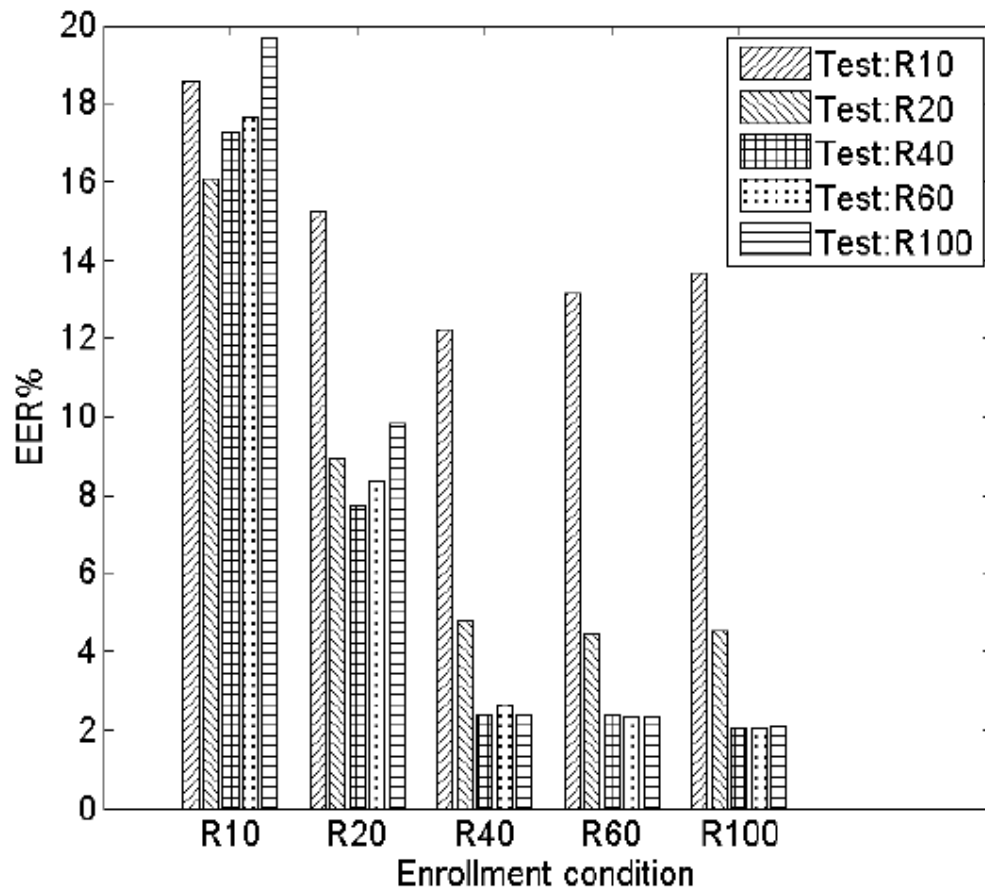
Sine Wave Truncated

6) DET of truncated segments detection



Sine Wave Truncated

Performance of discarding the truncated segments



Sine Wave Truncated

□ Conclusions

- When the truncated rate more than 80%, the performance will drop greatly
- We give out a proposal to detect the truncated speech by the distribution of sample points, which is able to detect the truncated speech
- Only by discarding the truncated segments, will the performance decrease

□ Future work

- Speech recovery
- DNN mapping

Experiments and results

Thank you!