

# STMicroelectronics Singapore for NIST SRE2010



AA&VoIP  
HVD/IP-dev

*Evelyn Kurniawati*  
*pk459839@ntu.edu.sg*

# Objectives

- ▣ **Learning experience**
- ▣ **Jump start project on speaker recognition**
- ▣ **With limited time (registration confirmed on 26 February), use as much open source tool as possible.**
- ▣ **Database are priceless to see how system fare**

# Tools

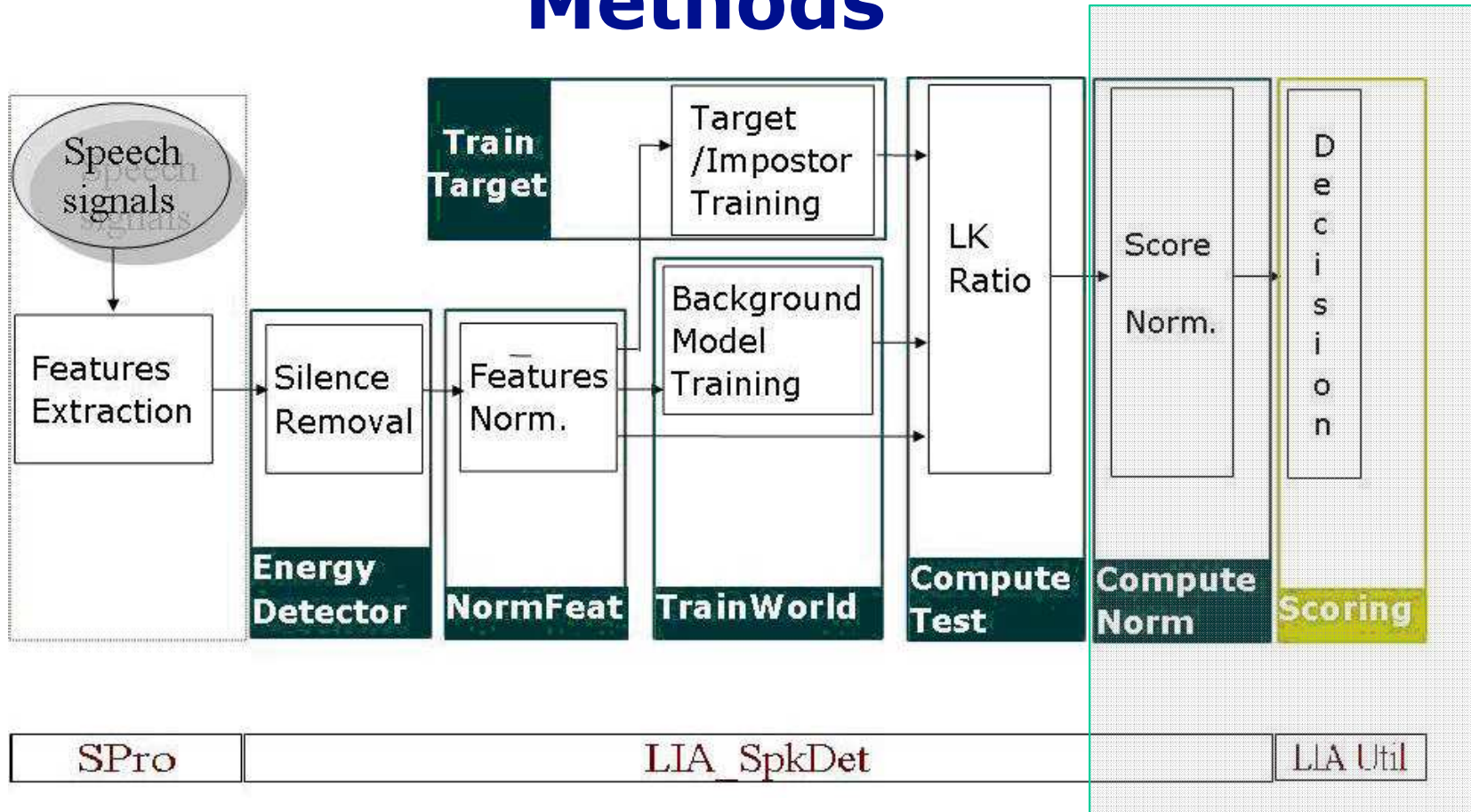
 **Front end : SPRO**

 **Statistical tool : ALIZE, LIA\_RAL**

 **Result analysis : Detware (matlab)**

- Apart from DETware, all tools compiled on windows (due to resource constraint).
- Statistical tool adjustment for windows : casting behavior and 48 bit random number generator
- Most useful guide : Biosecure “Reference system based on speech modality”

# Methods



Unfortunately the last two steps were not done

# System Description

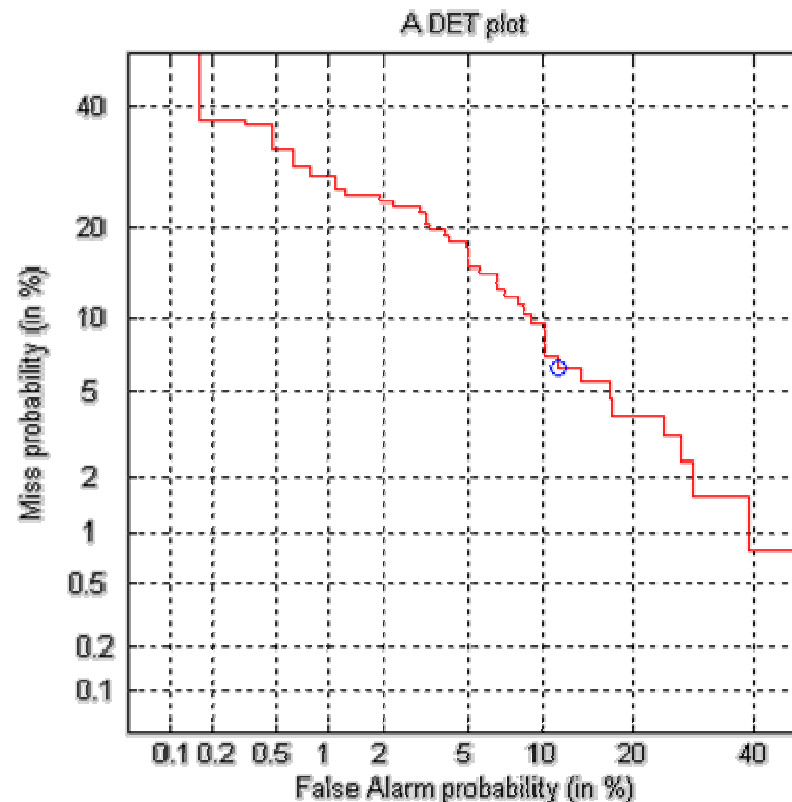
- ▣ GMM-UBM system
- ▣ Front end extraction : SPRO3.0  
(19 LFCC with  $\Delta$  and  $\Delta\Delta$  "- AD" option )
- ▣ Single world model from a subset of 2008 short2 data.
- ▣ 512 mixtures
- ▣ MAP adaptation strategy
- ▣ No channel compensation and score normalization
- ▣ Transcript is not used

# Development data

	Interview			Telephone_mic			Telephone_phn		
Train\Test	A	B	C	High	Low	Normal	High	Low	Normal
Interview A	1	2	2			9			8
Interview B	2	1	2			9			8
Interview C	2	2	1			9			8
Telephone_mic (Normal)				6	7				
Telephone_phn (Normal)							5	4	3

- Development data was used for verification (split audio files 85% for training 15% for testing)
- 9 testing scenario illustrated in above diagram

# Verification on Development Data



- Cdet : 0.4365, Min Cost : 0.2011  
(126 target 630 impostor representing the 9 evaluation condition)
- Verification on 2008 data was only partially done due to limited time and resource

# Conclusion

## Issues we face:

- how to judge the quality of world model
- deciding streams for impostor model in ZT norm
- the overwhelming amount of data to be processed (no resources to do a 2<sup>nd</sup> run before submission)

## Future work

- Channel compensation and SVM
- more in depth result interpretation

 Our first try here with modest results, but it's a start.

 Inputs are very very welcomed (author unable to attend workshop due to visa requirement ☹ )