Persay systems for NIST SRE 2005

Ran Gazit & Nir Krause

Persay Ltd., Israel

PerSay

General description

1. **GMM / SVM** (primary)
2. **GMM**
3. **1 & 2 fused**

All systems use acoustic information only

Baseline configuration

Features:

- **Echo suppression using 4wire data**
- **Energy-based VAD**
- **20 LPCC + 20 Delta** (25ms frames, 12.5ms overlap)
- **CMS, unit-variance**

Models:

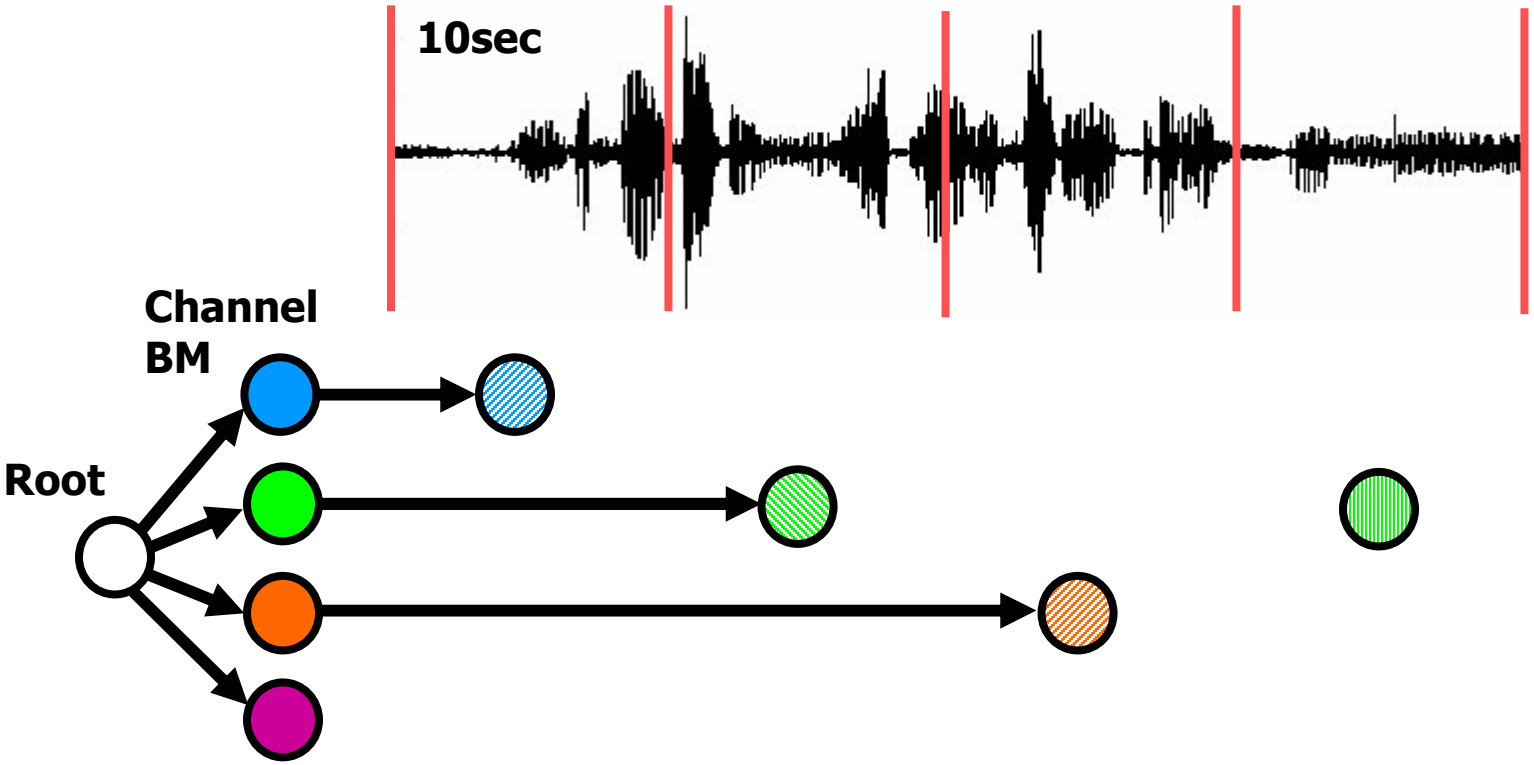
- **4 UBM per gender** (cdma, gsm, carb, elec)
- **Means-only adaptation from best UBM**

Development set

Background models	**99** **1side training**	**carb**	250
		elec	250
	03 **1side training**	**GSM**	250
		CDMA	250
Tnorm	**04** **1side training**		100
Parameter optimization	**04** **1side-1side**		
Threshold setting	**04**		

per gender …

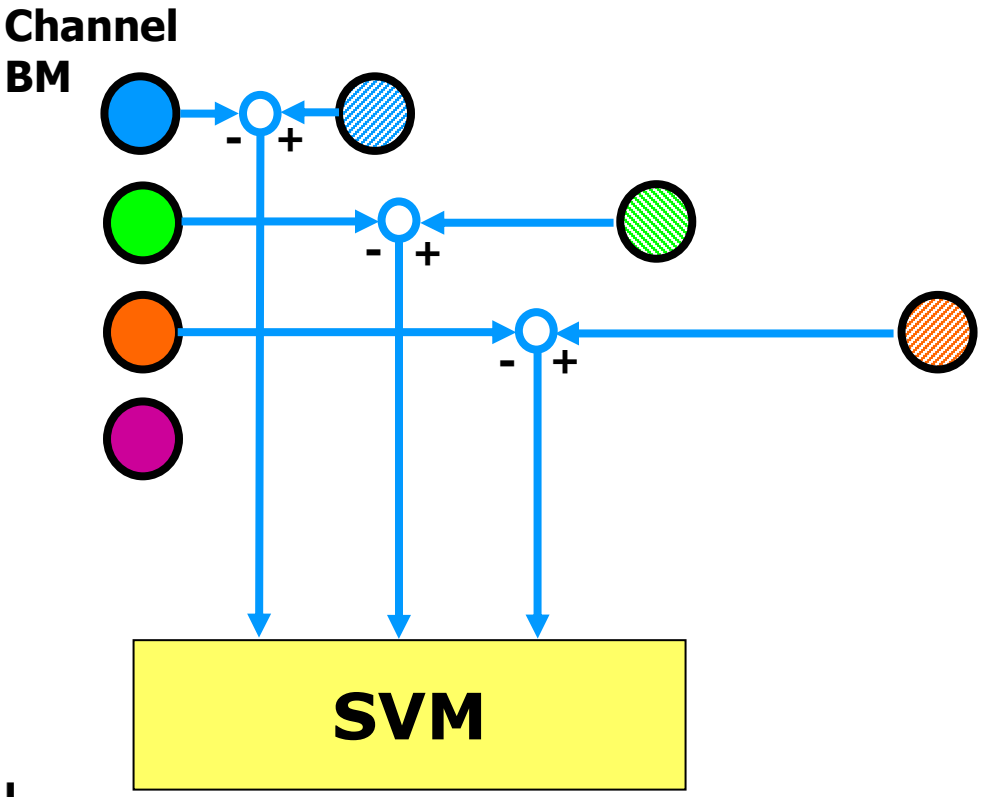
1. GMM/SVM system



1. GMM/SVM system

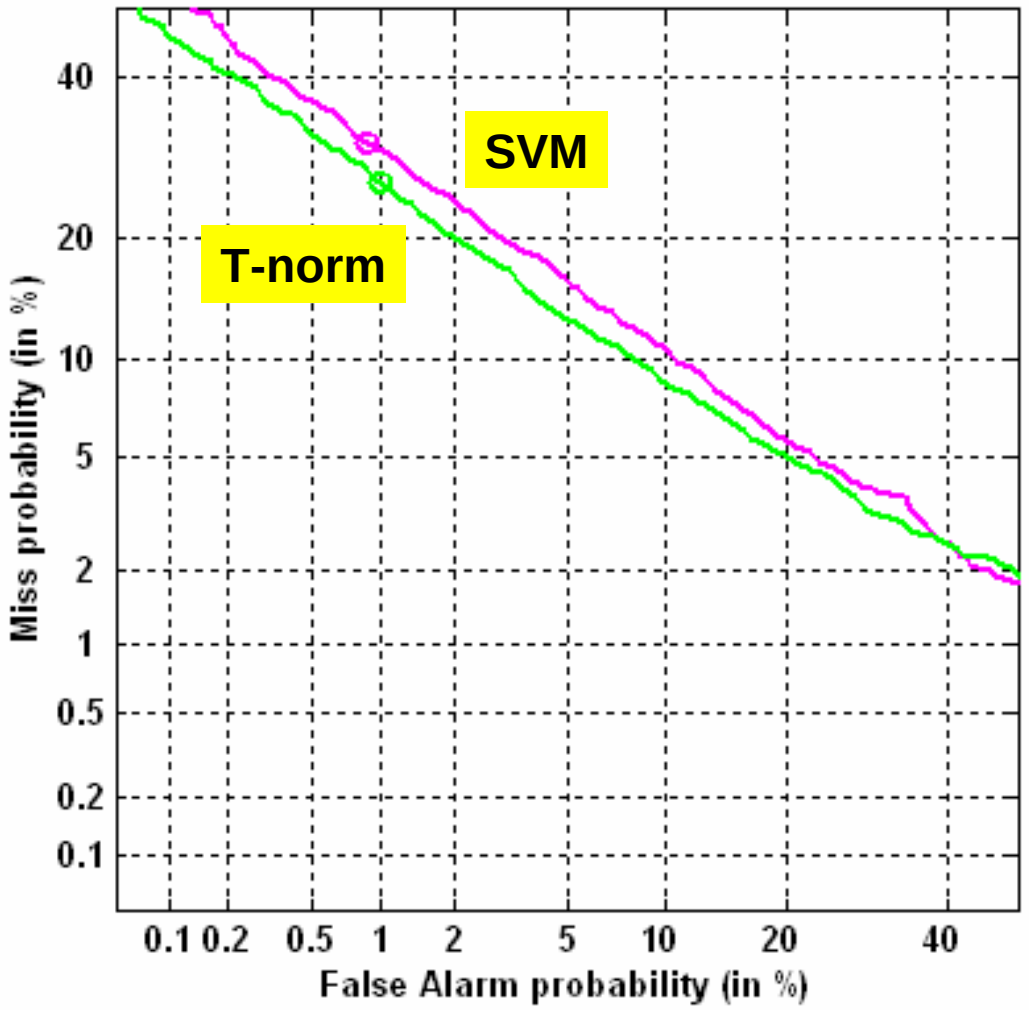
Classifier training:

- Positive examples from training segment
- Negative examples from first 10sec of all BM data



PerSay

T-norm effect

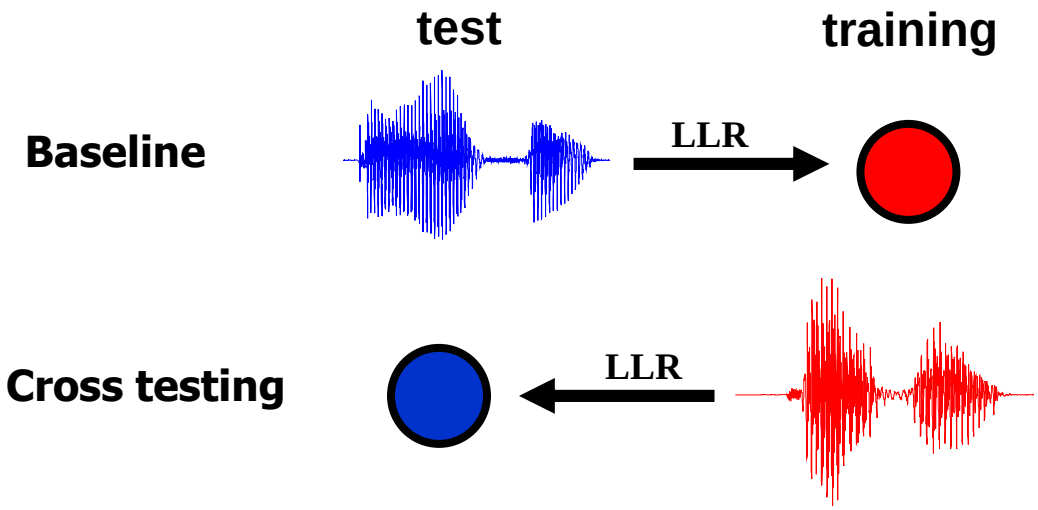


1conv4w-1conv4w, English

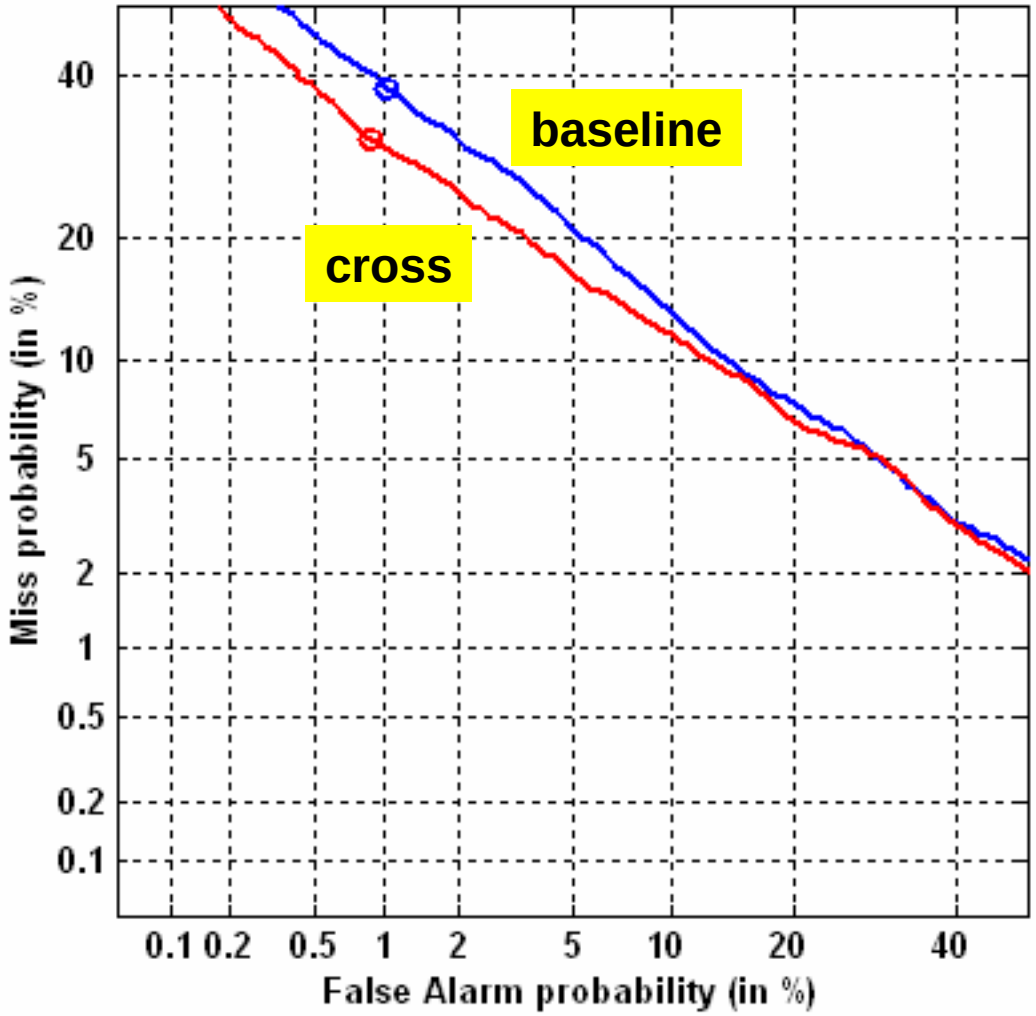
PerSay

2. GMM system

- **256-mixture GMMs**
- **Cross – testing** (on symmetric tasks):



Cross testing effect



1conv4w-1conv4w, English

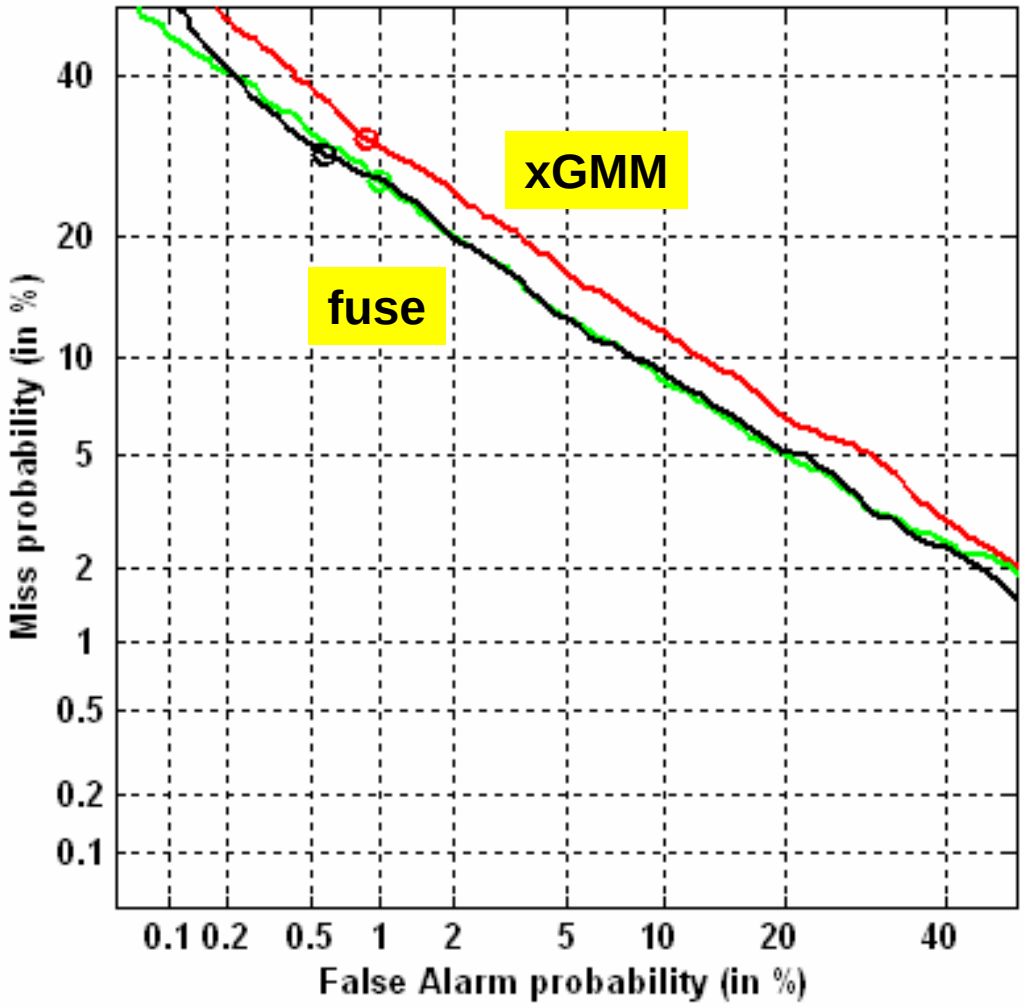
PerSay

3. Classifier fusion

- **SVM, Gaussian kernel**
- **Features: SVM-Tnorm and cross-GMM scores**
- **Training: NIST04, per task**

Fusion results

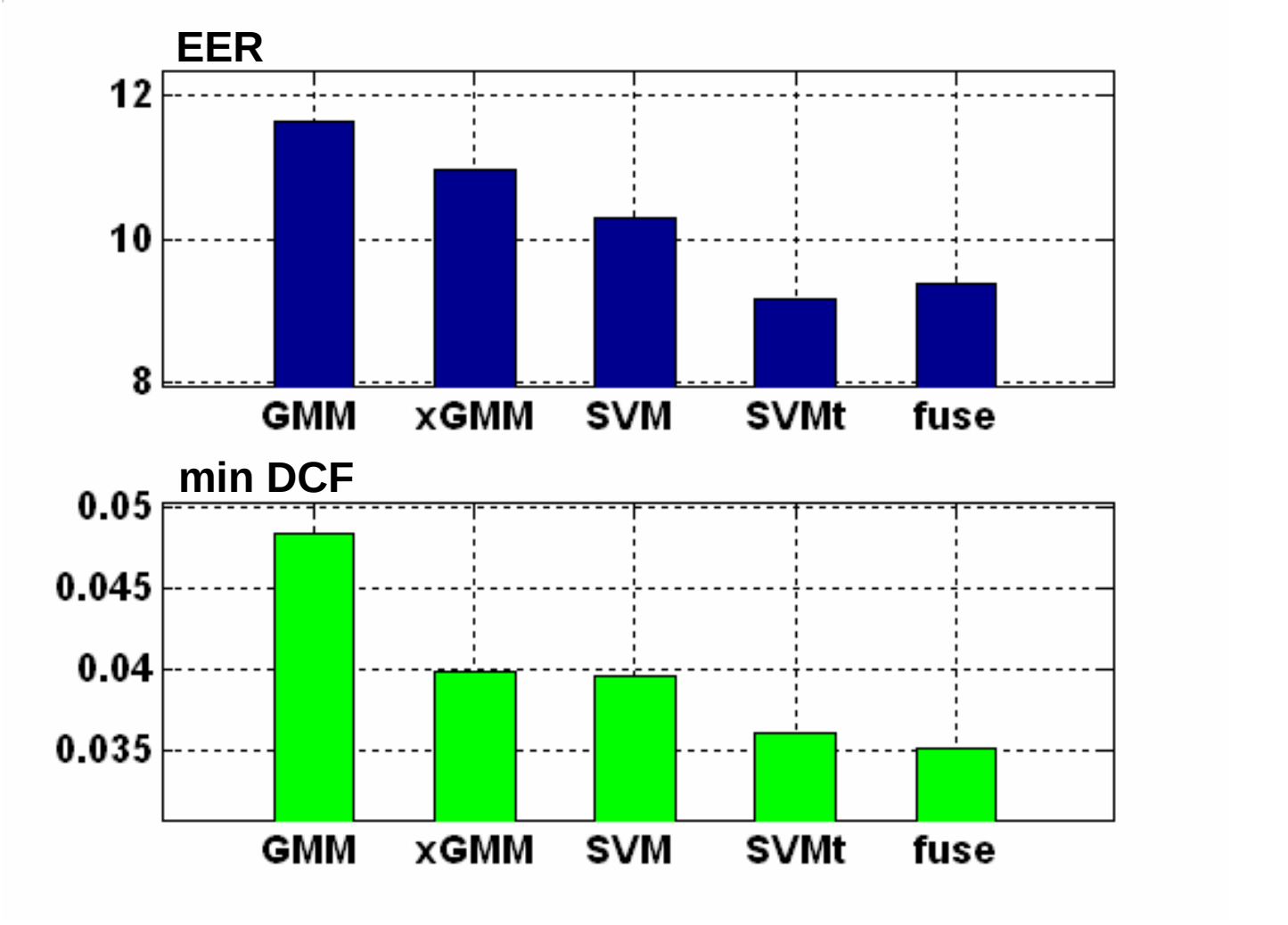
SVMt
xGMM
Fuse



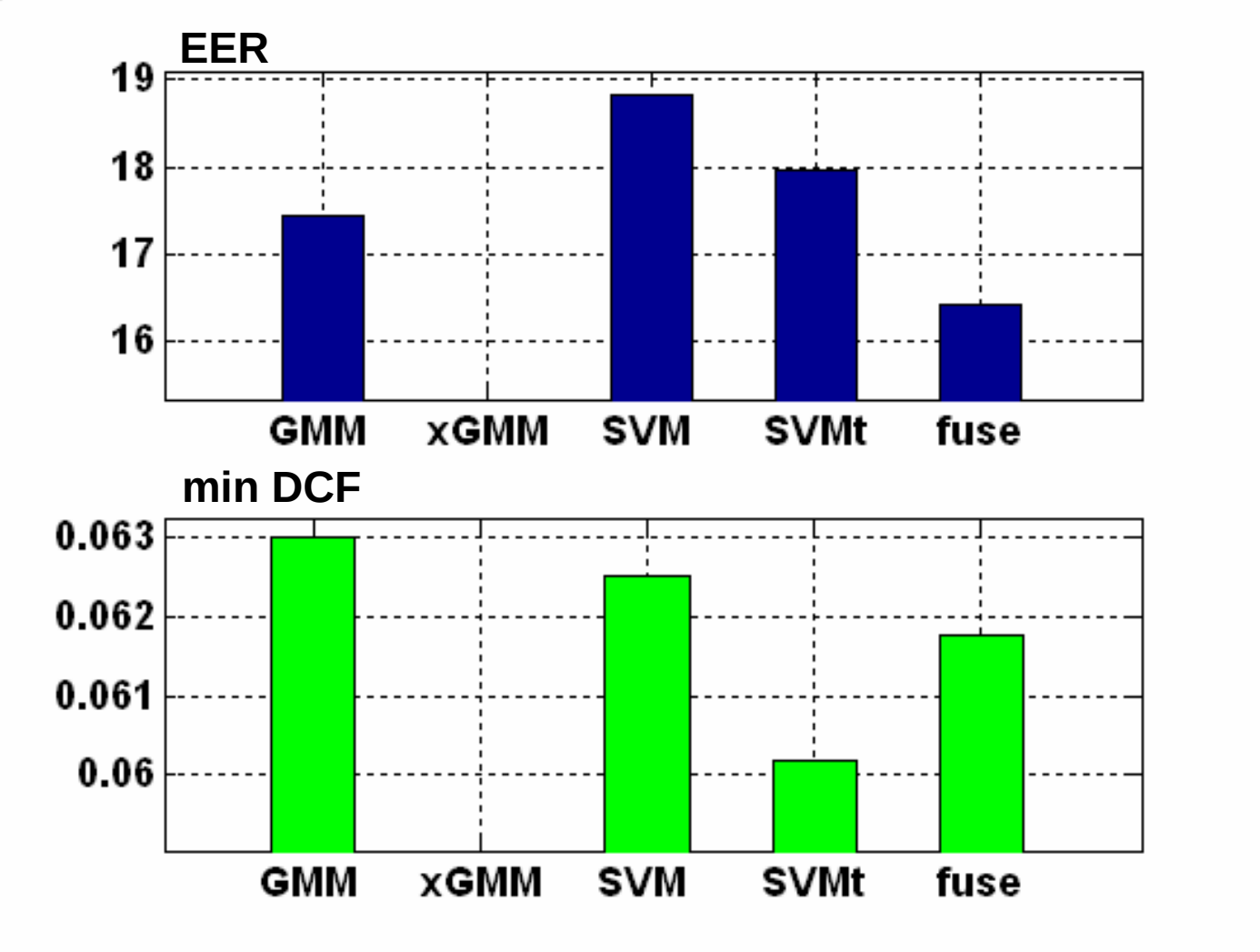
1conv4w-1conv4w,
English

PerSay

1conv – 1conv

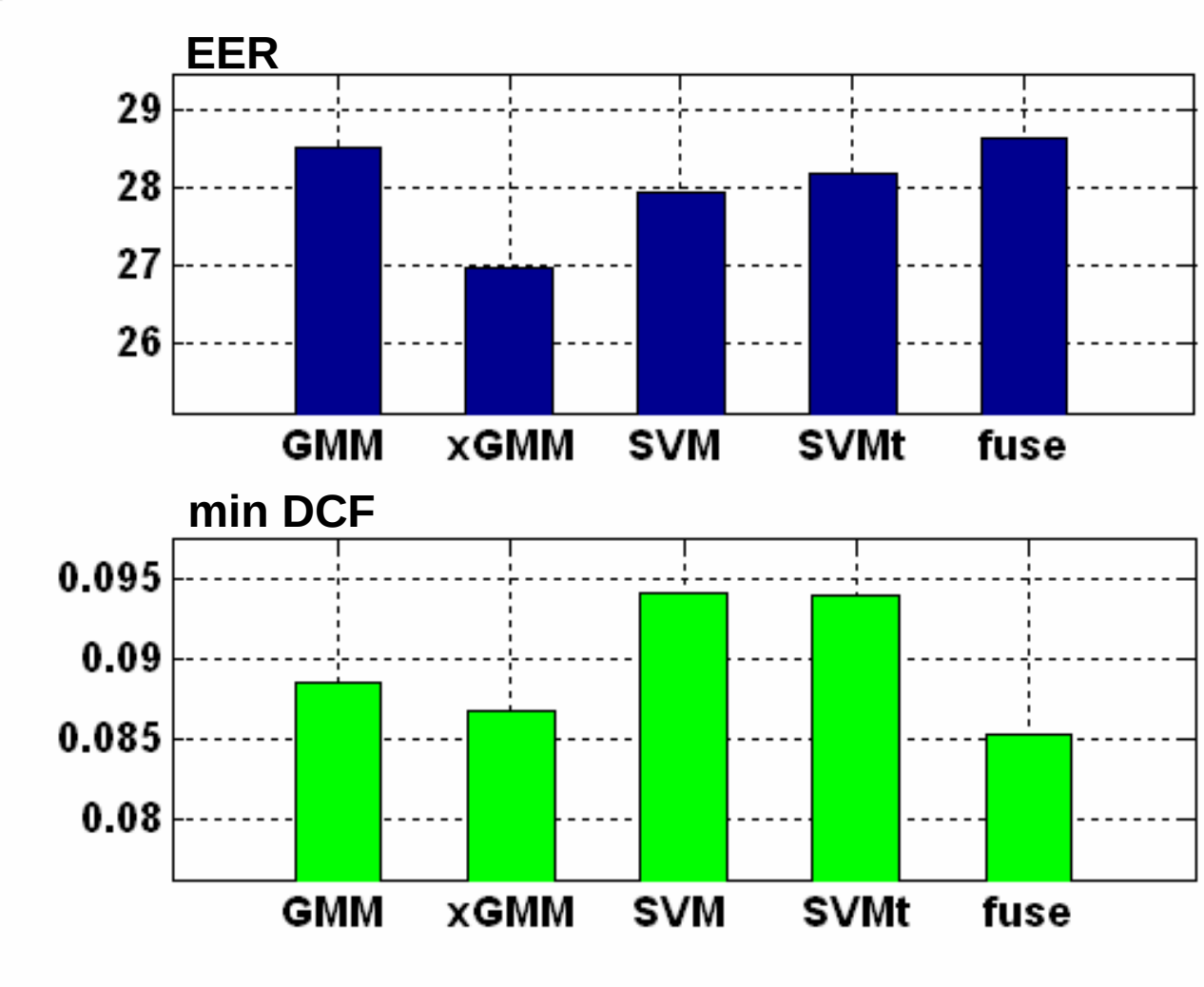

EER
12
10
8
GMM
xGMM
SVM
SVMt
fuse
min DCF
0.05
0.045
0.04
0.035
GMM
xGMM
SVM
SVMt
fuse

1conv – 10sec



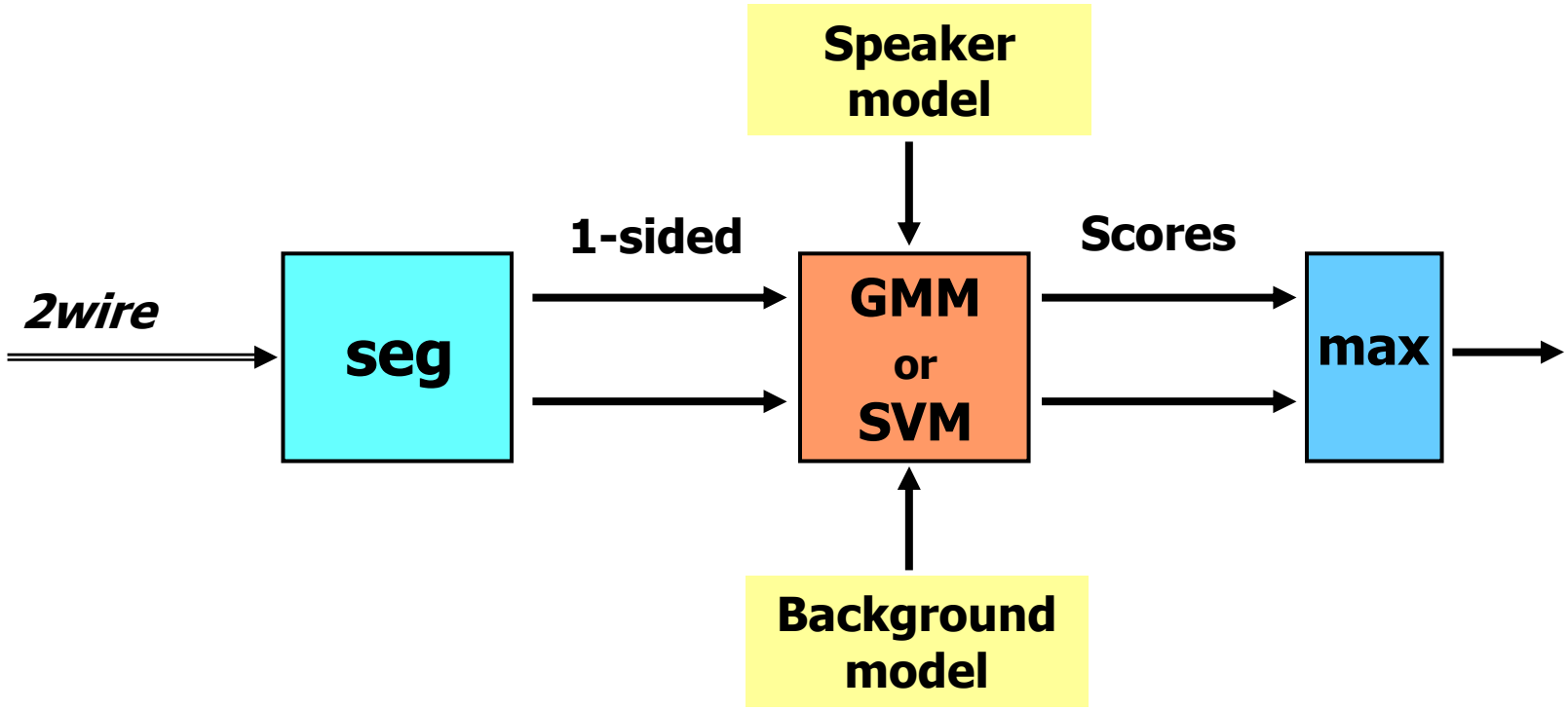
PerSay

10sec – 10sec

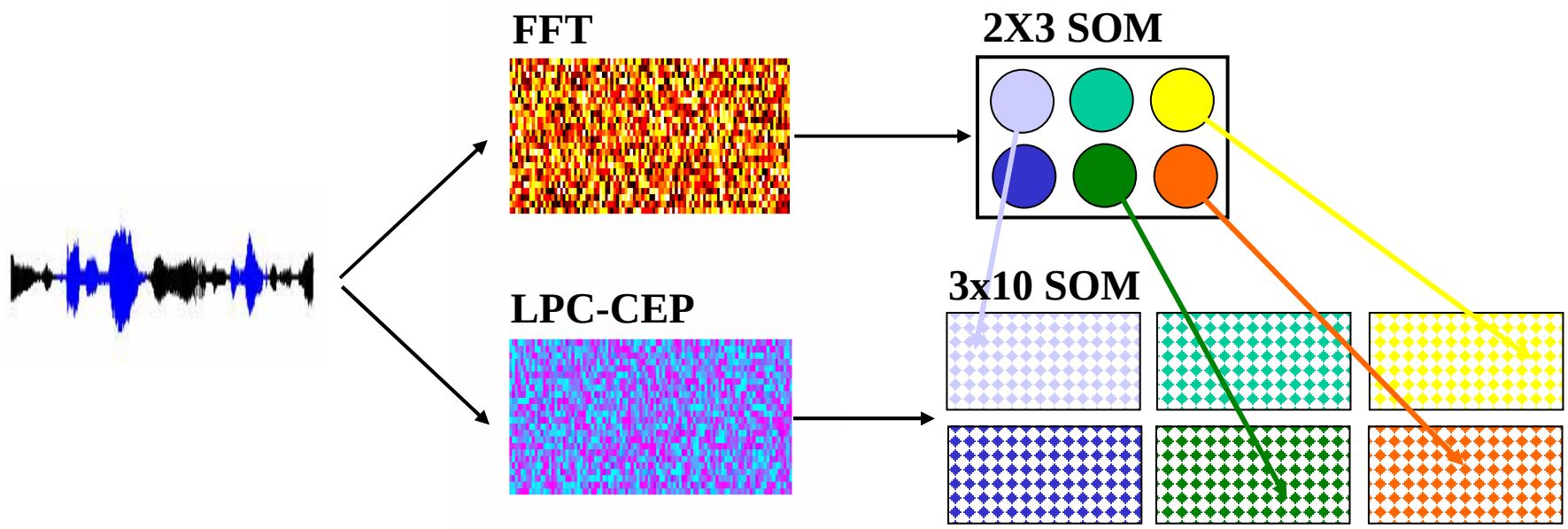


PerSay

2-speaker detection

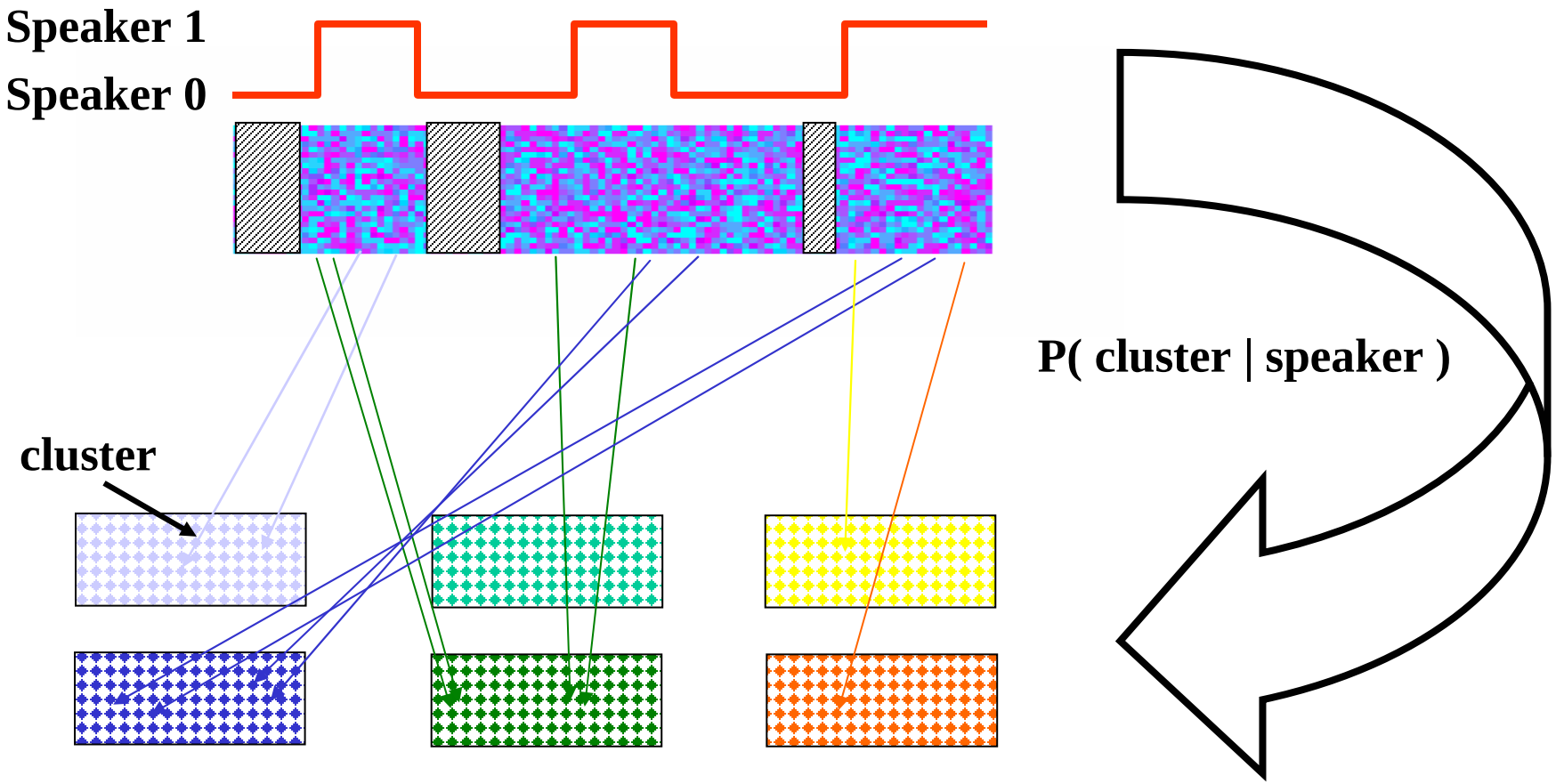


External segmentation

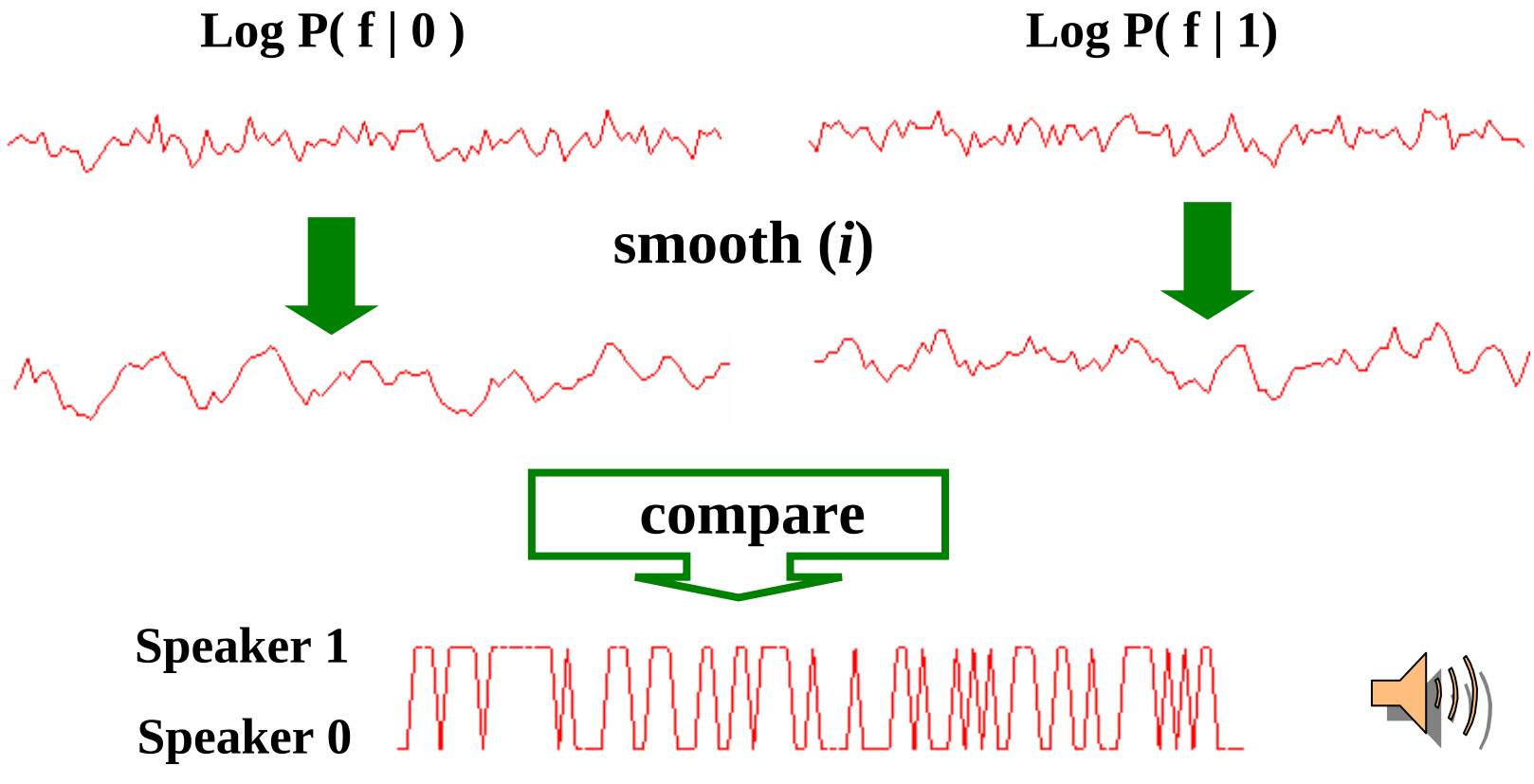


- **FFT:** 210-1510Hz 110Hz×12 +Δ×12
- **LPC-CEP:** 12CEP + 12ΔCEP

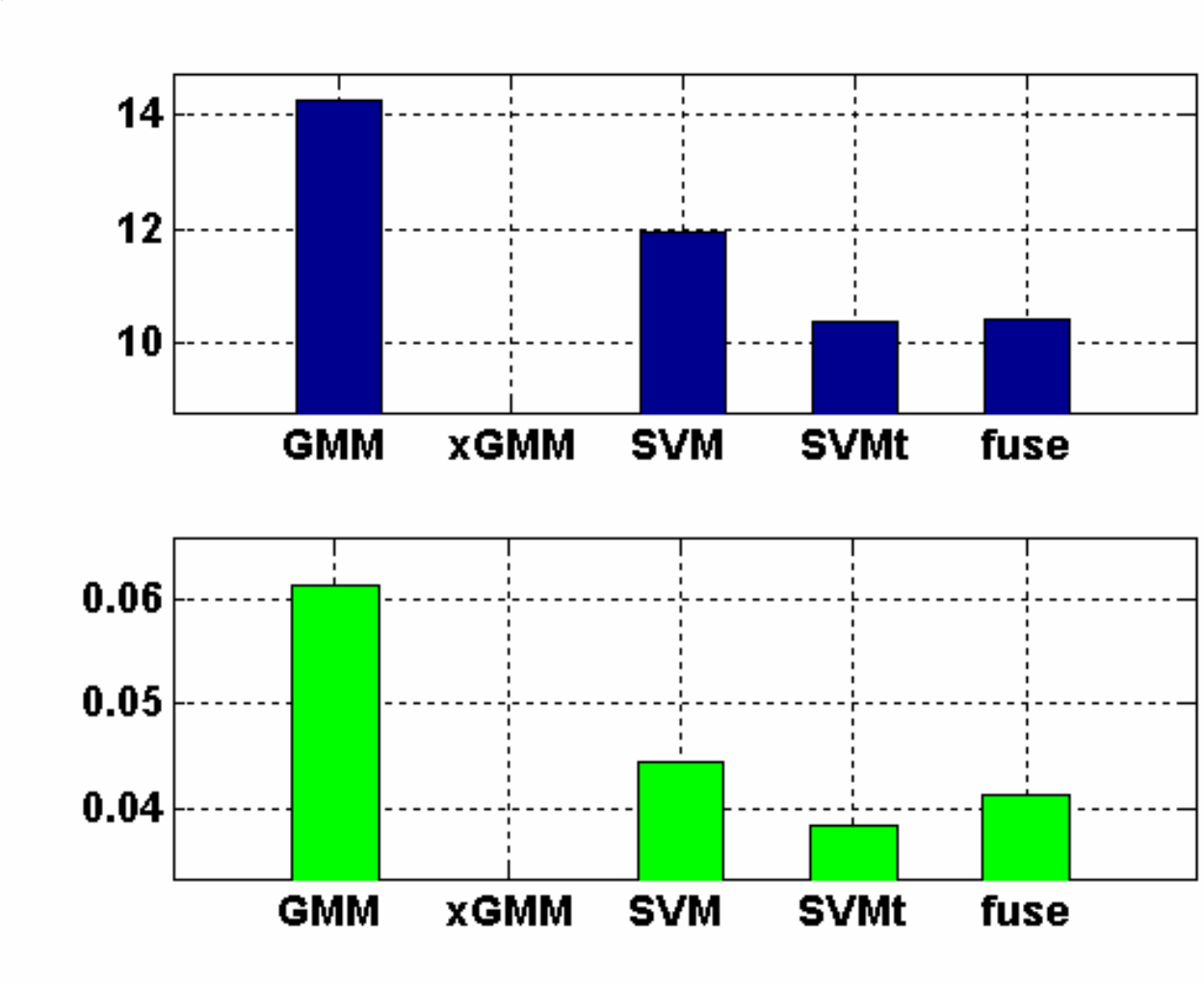
Find P(cluster | speaker)



Get new segmentation; iterate …



1conv4w – 1conv2w



Conclusion

Conclusions:

- SVM shows better results than GMM in 1conv tasks
- SVM more efficient for identification tasks

To do:

- Check SVM on cross-channel conditions
- Improve fusion
- Improve threshold setting procedure