NIST'05 Speaker Recognition Evaluation

Israel National Police / Bar-Ilan University

Yosef Solewicz
Moshe Koppel
Dan Shachoach
Uri Biran

Automatic forensic speaker verification

Objective:

To supply consistency in examination reports for the typical irregular forensic environment even at the expense of decreasing the weight of the result.

Basic system used in current forensic casework

- Standard GMM/UBM [Reynolds et al. 2000]
- Scores expressed in terms of p-value:
- Choose closest reference population available according to suspect / evidence recordings.
- Estimate impostor hypothesis PDF through suspect model normalization (‘Z-norm’).

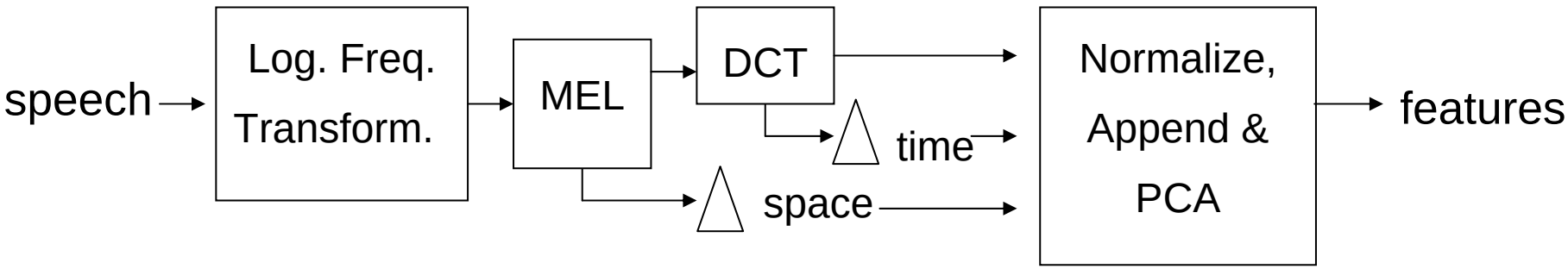
[Solewicz & Koppel 2004]

System used for ’05 eval.

Same as forensic except for:

- Scores are T-normed
- Same reference population (100 highest ranking of 150 models from ’03 eval.)
- Combined cepstrum/spectrum-slope features

Improved parameterization by combining cepstrum, deltas and spectral slope



19 cepstrum

19 delta-cep

19 spectral slope

38 combined features

PCA and UBM training

- Gender dependent PCAs trained on regular/cellular/cordless-phone recordings from Nist’04 eval.
- Gender dependent UBMs formed by agglomerating phone-type dependent UBMs (180x3=540 components each)

Performance of combined features (no T-norm) - DCF(x10^{-5}) on male portion of:

	CEP*	Combined	Reduction (%)
'00 eval	483	436	9.7
'01 eval	513	472	8.0
'05 1conv x 1conv	516	536	**? -3.9 ?**
'05 1conv x mic	859	836	2.7
→ (EER:	15.4	13.8	10.6)

---

* 19 cep. + 19 delta cep.
  540 (3 phone types x 180) component UBM

Current goal

Automatic assessment of bias due to inconsistency between the specific forensic case and previously collected statistics.