Robustness of Quality-based Score Calibration of Speaker Recognition Systems

with respect to low-SNR and short-duration conditions

Andreas Nautsch[*], Rahim Saeidi[†],
Christian Rathgeb[*], Christoph Busch[*]

[*]Hochschule Darmstadt, CRISP, CASED, da/sec Security Research Group
[†] Department of Signal Processing and Acoustics, Aalto University, Finland

Odyssey'16, Bilbao, 24.06.2016

Quality-based Calibration and low-Quality Robustness

Calibration schemes in unconstrained conditions

- ▶ Conventional: train on full/clean $\Rightarrow$ use on all   df: 2
- ▶ Matched: train and use for each condition   df: $2\,C$
- ▶ Quality Measure Function: duration/SNR/...   df: 3
  in linear fusion: $S' = w_0 + w_1\,S + w_2\,Q\,(\lambda_{ref}, \lambda_{prb})$

Unified Audio Characteristics

- ▶ Gaussian models for conditions on raw i-vector domain
  as vector representation of quality estimates $\vec{q}$
- ▶ Bi-linear combination matrix: $w_2\,\vec{q}_{ref}\,\boldsymbol{\Sigma}\,\vec{q}_{prb}$ df: $C^2 + 3$

- ▶ **Proposed: Function of Quality Estimates (FQEs)**   df: 3
  e.g., by cosine distance: $S' = w_0 + w_1\,S + w_2\,\cos\,(\vec{q}_{ref}, \vec{q}_{prb})$

Robustness Analyses in 55 combined conditions

Excerpt: