UWB system description - NIST10

Lukáš Machlica, Jan Vaněk
Department of Cybernetics, University of West Bohemia, Pilsen

Submitted Systems

- **Primary** = fusion of
  - GMM-UBM (+ TNorm)
  - SVM-GSV (+ NAP)
  - SVM-GLDS (+ NAP)
- 1st Contrastive = GMM based
- 2nd Contrastive = SVM based

Feature Extraction

- VAD - energy based
- **LFCC** - dim = 20, frame/10ms
- feature warping + dwnsmp 2:1
- *FeatSet1* - LFCC + Δ (dim = 40)
- *FeatSet2* - instead of Δ coeffs **Discrete Cosine Transformation** in the time domain (dim = 60)

GMM Systems

- feature extraction: *FeatSet2*
- GMM-UBM:
  - Fusion of **18 gender dependent** GMM-UBMs varying in **BG data & number of mixtures**.
  - data: SWB cell part 1, SRE04, 05, 06, 08-int, 08-tel.
  - #mixtures: 256, 512, 1024
- MLLR + MAP, $\tau = 14$
- Tnorm - gender & channel dependent, pre-cohort size = 600 (SRE 08), final cohort = 40

SVM Systems

- feature extraction: *FeatSet1*
- training set for each speaker divided into subsets with 1000 frames
- NAP trained on *SRE04, 05, 06* (co-rank 256)
- **3 impostor sets** ⇒ 3 SVM models
- linear kernels
- SVM Torch
  - **SVM-GSV**: involved 512mix UBM, MAP ($\tau = 5$), SV dim = 20480
  - **SVM-GLDS**: SV dim = 12341

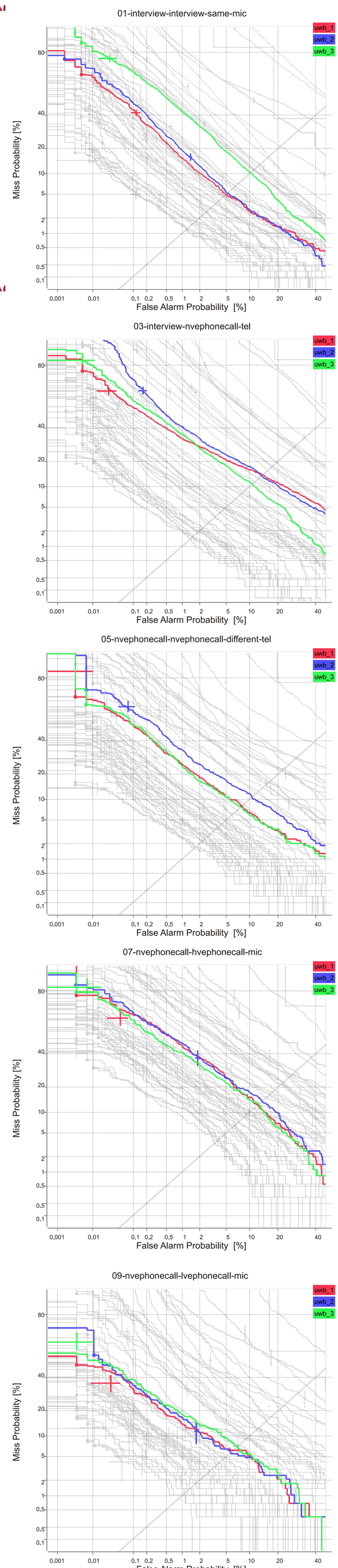
Fusion

- based on FoCal toolkit
- weights - Linear Logistic Regression - SRE08

CPU Execution Time

	enrollment [x RT]	memory demands	verify [x RT]
GMM-UBM	0.0072	2 MB	0.0019
SVM-GSV	0.1240	1.8 GB	6.6e-5
SVM-GLDS	0.0566	0.6 GB	3.7e-5
primary	0.5994	1.8 GB	0.0402
1st contrastive	0.0792	2 MB	0.0399
2nd contrastive	0.5238	1.8 GB	3.1e-4

Results NIST SRE 2010



Summary

- combination performs well
- SVM: lack in performance in interview cond.
- participating for the 1st time

This work was supported by the Grant Agency of the Czech Republic, project No. GAČR 102/08/0707, and by the Ministry of Education of the Czech Republic, project No. MŠMT LC536.