UWS Submission - NIST 2004

Speech and Image Group
University of Wales Swansea

Benoit Fauve

John Mason
Roland Auckenthaler
Neil Pearson
Nick Evans

Introduction

- UWS Participation: 03, 04 and 05
- UWS1: Roland Auckenthaler work
  - fast system, optimized for commercial use.
- UWS2: standard (in-house) GMM+TNorm
  - Plus LIA Alize/Spro4 features
- No class-depenedent optimization
  - ie «turn on and go systems»

UWS1

- UWS1- 05 is very similar to UWS 04 ***except:***
  - Faster processing (optimised FFT,…)
  - Score fusion
    - 1024 + 32
    - helps particularly for small trainging sets

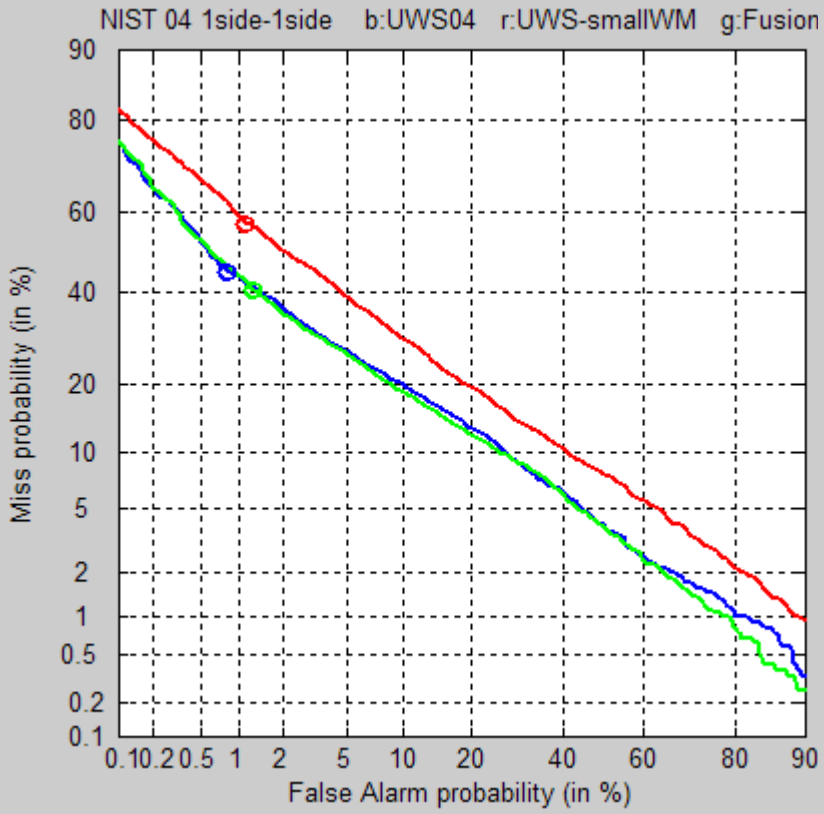
UWS1 : System Overview

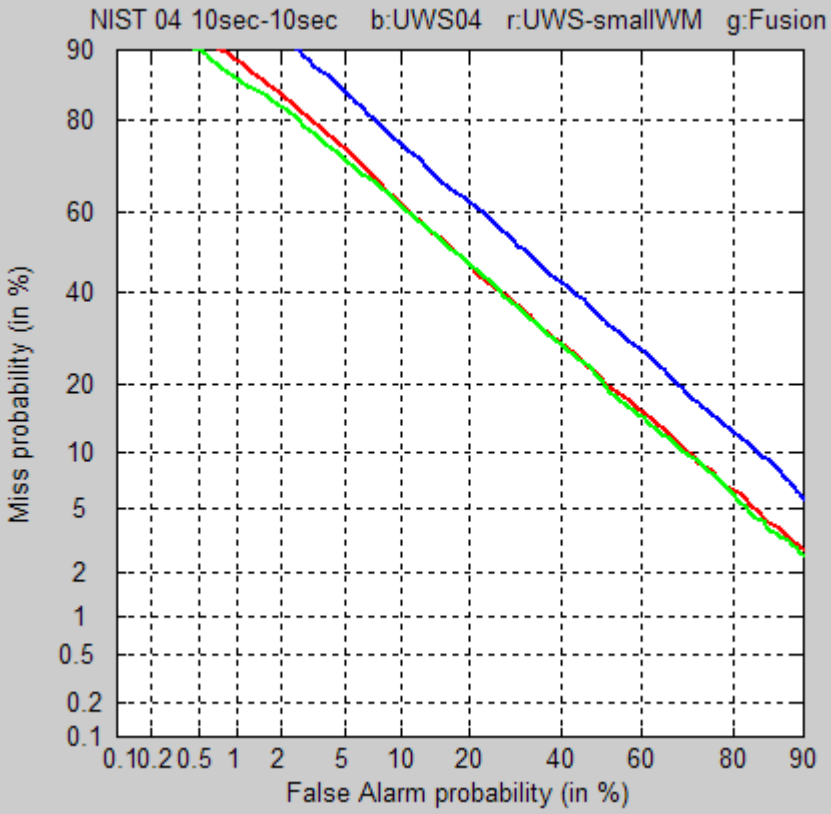
FFT: 62.5 fps
DCT/MEL: 19 Coefficients
RASTA/DELTA: 38 Features

UBM: 1024 Components, gender specific
UBM: 32 Components, gender specific
Data: Nist 2004
Adaptation: Mean-only
Scoring: T-Norm 75 speakers, common

Same System for all conditions except 10sec -10sec.

UWS1 fusion





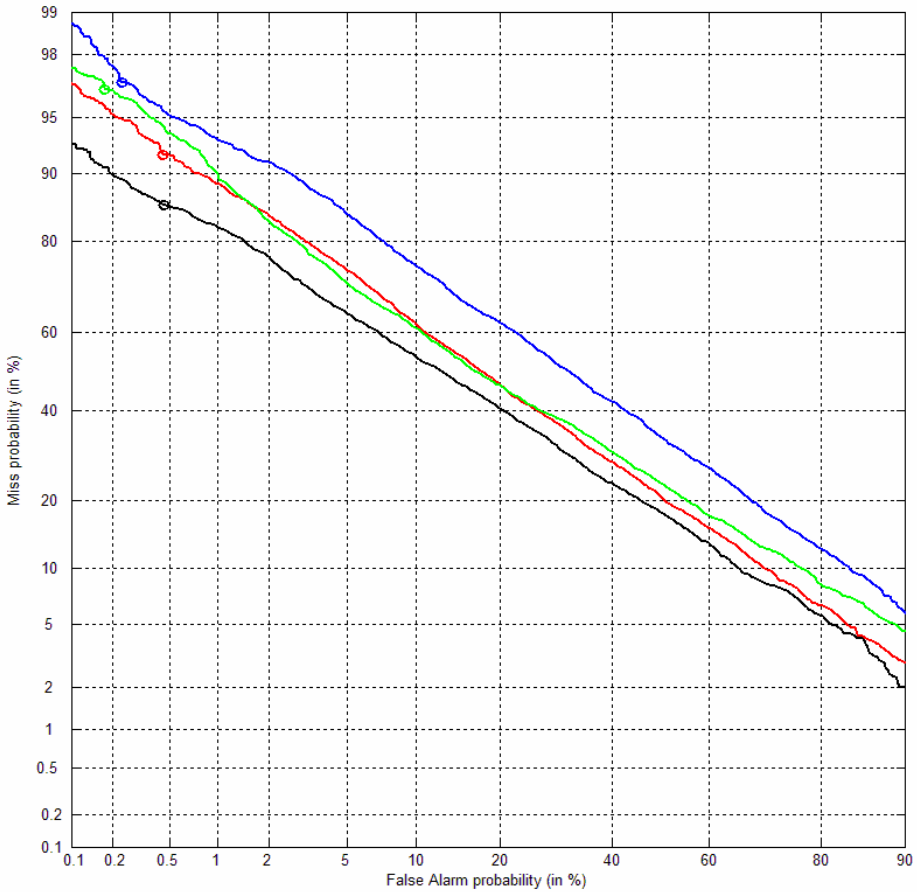
1: UWS04 2: UWS_UBM32 3: Fusion "1+2"

UWS1 10sec-10sec

1: UWS04 1024
2: UWS 32
3: UWS2 64

4: optimal fusion

1	of	1
0.7	of	2
0.8	of	3



UWS2 : System Overview

- Features SPRO 04 (INSEA), speech detector from LIA
- Vector Size: 34 16 Cepstra + energy, 17 Delta,
- GMM: 64 and 1024, from NIST 04
- Adaptation: mean only
- Testing: best 5 scoring components
- Normalisation:
  - TNorm targeted to training
  - 50 gender specific from NIST 04

Conclusion

- Only one development since last year
  - Fusion of small model with large model
- Need to look at conditions of larger data sets, particularly testing sets
- Next year.... We'll be there!

Further Information

http://www.speaker-verification.com/Projects

- Difference to NIST 2005 submission
  - Single UBM model with 256 Components