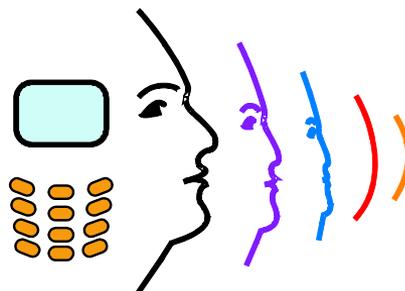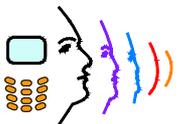# Speaker authentication

Speech@FIT, BRNO UNIVERSITY OF TECHNOLOGY

Honza Černocký and Pavel Matějka

MOBIO meeting, Martigny, January 8th-9th 2009

Publications

Publications - UNIS

ACCEPTED:

Paper Title: A Biometric Menagerie Index for Characterising Template/Model-specific Variation
Author List: Poh,Kittler
Conference: ICB2009

Paper Title: Challenges and Research Directions for Adaptive Biometric Recognition Systems
Author List: Poh,Wong,Kittler,Roli
Conference: ICB2009

Paper Title: Face Video Competition at ICB2009
Author List: Poh,Chan,Kittler
Conference: ICB2009

Under review:

Paper Title: Multimodal Biometric Fusion Exploiting the Doddington's Zoo Effect
Author List: Norman Poh, Samy Bengio and Arun Ross
Journal: IEEE Trans. Information Forensics and Security

Paper Title: Benchmarking Quality-dependent and Cost-sensitive Multimodal Biometric Fusion Algorithms
Author List: Norman Poh, Thirimachos Bourlai, Josef Kittler, Lorene Allano, Fernando Alonso, Onkar Ambekar, John Baker,Bernadette Dorizzi, Omolara Fatukasi, Julian Fierrez, Harald Ganster, Javier-Ortega Garcia, Donald Maurer, Albert Ali Salah, Tobias Scheidat and Claus Vielhauer
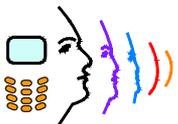Journal: IEEE Trans. Information Forensics and Security

Publications - IDIAP

ACCEPTED:

Title: Parts-Based Face Verification using Local Frequency Bands

Authors: C. McCool and S. Marcel

Conference: ICB'09

+ 1 given later.

Publications - BUT

Published:

Pavel Matějka, Lukáš Burget, Ondřej Glembek, Petr Schwarz, Valiantsina Hubeika, Michal Fapšo, Tomáš Mikolov, Oldřich Plchot, and Jan "Honza" Černocký: **BUT language recognition system for NIST 2007 evaluations**, in Proc. Interspeech 2008, Brisbane.

O. Glembek, P. Matejka, L. Burget, T. Mikolov: **Advances in Phonotactic Language Recognition**, in Proc. Interspeech 2008, Brisbane.
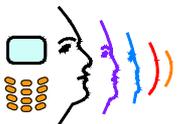
V. Hubeika, L. Burget, P. Matejka, P. Schwarz: **Acoustic Discriminative Training and Channel Compensation for Acoustic Language Recognition**, in Proc. Interspeech 2008, Brisbane.

O. PLchot, V. Hubeika, L. Burget, P. Schwarz, P. Matejka: **Acquisition of Telephone Data from Radio Broadcast with Applications to Language Recognition**, in Proc. TSD 2008, Springer Verlag LNCS.
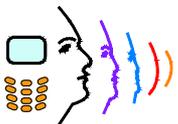
+1 at IEEE Goa: ….

Accepted:

Ondřej Glembek, Lukáš Burget, Najim Dehak, Niko Brümmer and Patrick Kenny: **COMPARISON OF SCORING METHODS USED IN SPEAKER RECOGNITION WITH JOINT FACTOR ANALYSIS**, accepted to ICASSP 2009.
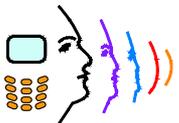
Publications

- We probably have more MOBIO-related publications …

- Need updates from all partners about published, accepted and submitted papers!

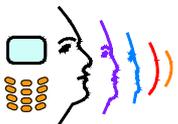- Also patents according to EC table sent by Valerie.
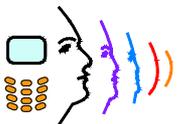
Progress report – last ½ year

WP2

- Contribution to fine tuning of specifications for MOBIO database (in Oulu)
- Big part of questions for "free" speech
- BUT's recording:
  - speakers found and recruited
  - came up with motivation strategy for BUT speakers
  - 2 batches of recordings done. 1 is in progress
    - 1st (with Chris)
    - 2nd "own",
    - this week 3rd batch is running (skiers will complete next week).

WP3

- participated at the definition of interface (Oulu)
- contributed VAD and speaker ID SW package for baseline system (D3.1) – sent to UMAN.
  - Generated results on BANCA, evaluated, submitted to SVN.
- wrote textual description to D3.2
- what does it include?
  - VAD based on phoneme recognizer with long temporal context and hierarchical structure of 3 nets.
  - GMM-based speaker system with eigen-channel compensation
  - both UBM and channel compensation trained on telephone data (from NIST), only enrollment and test done on BANCA DB.
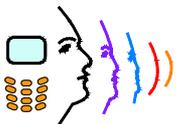 - EER 11% (set G1) and 8.2% (set G2), both protocol P.

WP3 - research

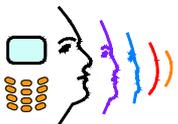mainly linked to the JHU workshop 2008

- Work-group **Robust Speaker Recognition Over Varying Channels** lead by Lukáš Burget (BUT)
- With some of the very good guys in Speaker ID (Patrick Kenny, Niko Brummer, Doug Reynolds…)
- and 2 other people from BUT: Ondrej Glembek (PhD), Valiantsina Hubeika (pre-grad, now PhD).
- Tasks
  - Diarization using JFA
  - Factor Analysis Conditioning
  - **SVM – JFA and fast scoring**
  - **Discriminative System Optimization**

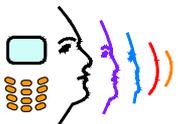www.clsp.jhu.edu/workshops/ws08/groups/rsrovc/

WP 4/5/6

- WP4
  - no planned activity, but communicated with Norman, sending scores for BANCA.
- WP5
  - Preparations so far, thinking what can be down-scaled.
- WP6
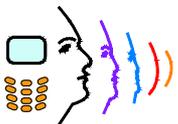 - no planned activity, but continuing development of BS-CORE with BS-API

WP7

- contributed to precise agreement on the evaluations on BANCA (building on D71).
- asked several companies about the CoI with moderate success.
- gathered info on publications
- in cooperation with Marek Janac prepared shots for Czech Radio (station Leonardo) and Czech TV, with broadcast on 12.11.2008 in the "PORT" series on science. Pavel was on chat after the broadcast with couple of interesting questions from the audience of CT.
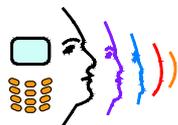 http://www.ceskatelevize.cz/port

Future work – research in speaker ID

- continuing with **full factor analysis** in speaker ID.
  - lots of research initiated at JHU, now on
    - modeling target/non-target distributions by SVM.
    - discriminative training
    - dot-product scoring (w/o GMM evaluation !)
  - one of outputs is "encoding" of the speaker into very limited amount of features.
  - an alternative way is to use "bottle-neck" NNs for speaker characterization - ongoing work.
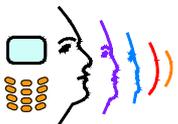
Future work ~ MOBIO

- D7.3 (M14) First report on dissemination activities – BUT lead, will need your contributions.
- D3.3 (M18) Advanced systems for uni-modal authentication: BUT contribution: implementation of full FA system.
- D4.1 (M14) Baseline bi-modal authentication systems: BUT contribution: theoretically nothing, but don't worry about our scores ;-) and other help.
- D4.5 (M16) Baseline model adaptation systems - BUT contribution: BUT has bunch of interesting techniques for adaptation from UBM, normalization (including investigation in norm. depending on length of training and test segments), needs more precise specification.
- D5.1 (M18) Scalable systems for uni-modal authentication; BUT contribution: VAD and speaker ID down-scaling foreseen.
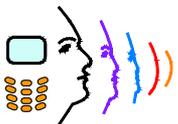
WP4 planning – possible BUT contribs.

- Score combination
  - Yes, already done, utterance-based
  - Experience with system calibration and fusion (multi-class logistic linear regression by Niko Brummer).
  - but do not have confidence measures nor quality measures to determine when to stop.
- Feature level
  - Yes, straight MFCC's probably not the best suited
  - but can provide features related to super-vectors of GMMs, directly (too big) or after projection to speaker space (couple of hundreds).
  - Are available for whole utterance or segment, not frame by frame.
  - Lip-features – in case face gyus can provide them, will be happy to test (although problem with UBM in this case !)
- Model combination
  - Not directly, but providing outputs for **phone conditioning**
  - EN phones or broad phonetic classes
  - Whatever output you want: phone strings with temporal labels, phone lattices, phone posteriors …
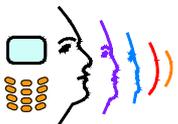
WP3 planning

- **VAD**: happy with our structure (also other sites in the world), will experiment with various training data to reach optimal results for mobile applications
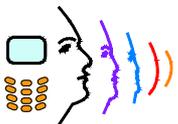- **Speaker Verification**: full Factor Analysis

WP5 planning - scalability

- scaling down existing systems without touching their architecture (e.g. less NN weights, less Gaussians in GMM...) – affect performance ☹

- but can also go for more dramatic change as dot-product scoring (very fast and minimum decrease of accuracy) - see [Glembek2009]
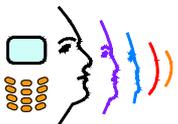
WP6 planning

- BUT/Phonexia production speaker ID system now being finished

- will be available for integration by EPM on PC-based architectures

- Mobiles ?

WP7- D7.2 (M14)

- D7.2 needs:
  - published publications supported by MOBIO
  - accepted papers supported by MOBIO.
    … note, that according to Consortium agreement, you should inform the consortium about such papers. Please send names of authors, title of the paper, conference/journal and abstract to MOBIO mail list. Thanks to all who have done so.
  - patents
  - media coverage
  - related projects
  - Invited talks
  - Competitions (NIST, ICB,…)
  - companies interested / using results / ...
  - needs input from Ideark (work to be done in 7.2).

Please send input till end January to Honza

WP7 – EC table (now!)

see Honza's mail from yesterday:

- published papers supported by MOBIO
  - please provide full BIB-TEX reference of such papers.

- applications for patents related to MOBIO - I need application number, title, and level of the patent (national, EU, US, EU and Japan, EU and US, US and Japan, EU and US and Japan level).

Please send this information by **next Tuesday, 13.1.,** as we have to deliver this info to the EC next week ☹