The full standard initiative is located at [www.voicebiometry.org](http://www.voicebiometry.org)

Quick description

Standard manual with detailed description and a quick user guide to...

The reference demo package
  - Contains full speaker-recognition (demo) pipeline
• Information-rich
• Low-dimensional
• Fixed-length
• Vector of real numbers

• Based on statistical model
• Easy to compare
• Easy to store
• Not recoverable to speech

I-VECTOR EXTRACTOR (SITE 1)

SPEAKER IDENTITY AND SPEECH CONTENT

SPEAKER IDENTITY NO SPEECH CONTENT

I-VECTOR EXTRACTOR (SITE 2)

AUDIO 1

AUDIO 2

i-vector

COMPARISON

SCORE
Standardization objectives

- Acoustic feature extraction
- i-vector extraction algorithm
- i-vector extraction parameters (GMM parameters, i-vector extractor parameters)
- The data exchange formats
- (tuned for telephone speech)
Feature Extraction

- Pre-emphasis
- 25ms windowing with 10ms shift
- Hamming window
- 24 Mel filter-banks in the range of 125 – 3800 Hz
- 19-dimensional MFCC coefficients + $C_0$
- Delta + Double-delta
- Short-time Cepstral Mean and Variance Normalization
  - Over 3 second window
Universal Background Model

- 2048 Gaussian mixture components
- Diagonal covariances
- Trained on 1156 hours of the NIST SRE 2004-2008 data (gender independent)
- Trained using gradual Gaussian splitting with 10 EM steps in each split
- The UBM is used to extract the sufficient statistics for the i-vector extractor and to normalize (whiten) these statistics
i-vector extractor

• Trained on the same data as UBM + Switchboard 2 (phases 2 and 3) + Fisher English (phases 1 and 2)
• 600 dimensional
• 10 iterations of EM and MD steps
Reference implementation

- Python code
  - Readability and ease of understanding
  - Extensibility
- Standard Python packages
- Numpy + Scipy
System performance

- NIST SRE 2010, cond 5, female

<table>
<thead>
<tr>
<th></th>
<th>New DCF</th>
<th>Old DCF</th>
<th>EER</th>
</tr>
</thead>
<tbody>
<tr>
<td>2010/cond5/f</td>
<td>0.3877</td>
<td>0.1142</td>
<td>2.26</td>
</tr>
</tbody>
</table>
i-vector compatibility

• i-vectors produced by one system are incompatible with those generated by a different system
• We run experiments for training an i-vector transformation to migrate i-vectors of one systems to another
• Take it as an invitation for tomorrow’s talk:

“Migrating i-vectors Between Speaker Recognition Systems Using Regression Neural Networks”
THANK YOU