



CCNT SYSTEM DESCRIPTION: NIST SRE 2010

Yingchun Yang1, Zhenchun Lei2, Ting Huang1, Li Chen 1, Wenxiang Chen1, Qi Yu1, Zhenyu Shan1, Rui Huang1 yyc@zju.edu.cn

1CCNT Lab, College of Computer Science and Technology, Zhejiang University Hangzhou, Zhejiang, P.R.China 2 College of Computer Science and Technology, Jiangxi Normal University Nanchang, Jiangxi, P.R. China

1. Abstract

CCNT submitted three systems to NIST SRE 2010 evaluations, only to the core test. The primary system is a fusion of two sub-systems (also as the alternate systems) using two different cepstral features and 2 different classifiers. The Alternate System I is based on the generative GMM-UBM [1] approach, the Alternate System II is based on discriminative SVM techniques [2]. Only the Alternate System I was submitted to the later extended core test. We mainly outline in this document the CCNT system including feature extraction, classifiers, fusion methods. Some implementation details such as the processing speed and the utilized software are also introduced.

•4. Classifiers •GMM LFA

• 2.Submitted Systems

•NIST SRE 2010 core test

GMM LFA (Alternate System I) Gender dependent GMM LFA system with MFCC13/LPCC18 features and gender dependent z-norm.

CD UBM SVM GSV NAP (Alternate System II)

Universal background models (UBM):

(1) The two gender-dependent UBMs.

(2) Tel-UBM : only the data from telephone channel are used for training. (3)Mix-UBM : the data from all channels are used for training, including channels including telephone, microphone and interview.

• Latent Factor analysis

(1) channel matrix rank 50, the regulation factor 14 (2) interview and microphone Eigen-channel matrices : trained on all the data from NIST SRE 2008 and follow up data , the speakers with more than 5 utterances selected.

(3)telephone Eigen-channel,:

NIST 2004, 2006, 2008 SRE as the training data and those speakers with more than 6 utterances selected.

(4) the Eigen-channel used only on the enrollment data, not on the test trails.

Normalization

Z-norm. Imposters are drawn from NIST SRE08 data

•SVM GSV NAP

 CD UBM(Channel Dependent Universal background) models)

Channel Dependent UBM (CD UBM) SVM GSV NAP with MFCC13/LPCC18 features and gender dependent t-norm.

Primary System

fusion of the two Alternate Systems

•NIST SRE 2010 extended core test

GMM LFA (Alternate System I)

Gender dependent GMM LFA system with MFCC13/LPCC18 features and gender dependent z-norm.



(1) The whole evaluation set is split into 10 different type according to gender and train-test type. UBM for each type is trained seperately.

(2) All utterances in 2008 short2-short3 evaluation corpus are used to train CD UBM. For mic-UBM and tel-UBM, 2010 tarball data with different vocal effort are also included. For each utterance in training corpus, GMM is adapted from different UBM according to the channel type of the utterance. In the evaluation step, GMM is chosen according to the train-test type.

• SVM GSV

(1) training data for positive samples :2010 core evaluation data are split into five subsets according to train-test channel type.

(2) training data for negative samples : NIST 2004 1-side training corpus

NAP (Nuisance Attribute Projection)

The training data for NAP matrix selected form 2008 SRE core evaluation corpus, selecting the speakers with more than 5 utterances. Principal components number 50

Normalization

T-norm. Cohorts are drawn from NIST SRE08 data

• Fig. 1. CCNT system framework

•3. Feature extraction

•MFCC13/LPCC18:

energy-based voice activity detector (VAD), CMS and variance normalization

•5. Fusion

•**Stage1** – Linear fusion with equal weights

The score of the two Alternate Systems is generalized by fusing the scores from their individual subsystems.

•**Stage2**– Linear fusion with different weights

The score of the primary system is linear fusion of the two Alternate Systems, and the weights are obtained by FOCAL Toolkit with the minimum equal-error-rate (EER) criterion on the **NIST 2008 SRE.**

•6. Processing Speed

• PC Intel Xeon E5420 processor with 4 Gb RAM

- 1.02×RT training , 0.019×RT testing for Alternate System I
- 0.43×RT training, 0.11×RT testing for Alternate SystemI I

ZJU CCNT LAB -