International Journal of Applied Engineering Research ISSN 0973-4562 Volume 13, Number 10 (2018) pp. 8561-8567

Support Vector Machine Based Approaches For Real Time Automatic Speaker Recognition System

Satyanand Singh

Assistant Professor, School of Electrical and Electronics Engineering,
Fiji National University, Fiji Island.

Abstract

It is known that the Percentage of Identification Accuracy (PIA) of Automatic Speaker Recognition (ASR) systems is increasingly vulnerable, such as noise and channel degradation in real-time. This study presents a novel class SVM and i-vector based speaker specific feature extractor **(**SVM-IVSSFE) algorithms based ASR system. I introduce SVM based approach at two stages within state-of-the-art ASR system: i-vector extraction estimation and at the Probabilistic Linear Discriminant (PLDA). SVM based approaches have classified more challenging subsets on the basis of training using a proper difficulty criterion. After this, the proposed training algorithm is initialized with a subset that reaches a clean set and then moves into subsets which more are challenging for training because the algorithms move forward. Based on the SVM-IVSSFE approach on noisy background I evaluated the PIA of ASR on critically degraded datasets with NIST 2010 database with standard i-vector extractor. The multi session shows a consistent and significant improvement over several testing sets on a standard baseline with SVM-PLDA. SVM-IVSSFE ASR system shows 1.99% of improvements compare to SVM-PLDA ASR system as well as SVM-IVSSFE ASR system shows 1.5% of improvements compare to ISVM-PLDA ASR system.

Keywords: Automatic Speaker Recognition (ASR), Probabilistic Linear Discriminant (PLDA), Support Vector Machine (SVM), Speaker Specific Feature Extractor (SSFE)

INTRODUCTION

The process of human learning often involves the introduction of a concept that is related to a previous assimilated concept which is relatively easy [1]. To facilitate the assimilation of new information, this mechanism of human learning is based on the knowledge already received, motivated a specific category of Linear Predictor (LP) based algorithms in machine learning. different types of training strategies based on LP has been explored to number of applications as: to improve the robustness of automatic speaker recognition (ASR) [2], to create deeper architecture for the Deep Neural Network (DNN) [3], get better estimates of the latent variable models[4]. In this study I propose a new framework to improve the percentage of identification accuracy of ASR system in the presence of severe noise and distortion inspired by a LP based machine learning algorithm. Advanced text-independent ASR systems use the i-vector framework with a Probabilistic Linear Discriminant Analysis (PLDA) back-end [5]. The i-vectors of a human utterance provide a compact representation of the specific characteristics of speakers limited to them in a very low dimension space.

In order to overcome the degradation of PIA performance in the presence of noise, several methods have been proposed which can be classified into two main categories: feature based methods that incorporate the use of strong functionality in PIA front-end or reduce the effect of post processing noise based on the model which better keep the presence of noise in mind. A method based on speaker-specific features has been proposed as a result of the distortion of tasks according to the time interval set for delivery of a specific goal [6]. Inspired by human hearing system Cochlear Cepstral Coefficients (CFCC) feature was used to improve PIA performance in noisy conditions [7].

On the basis of speech enhancement, improvement has also been proposed for various methods, which work at the functional level of ASR performance in noisy conditions. A DNN auto encoder was used to optimize speech for PIA improvement in ASR system [8]. De-noising technique based on DNN was proposed to estimate the speaker specific features for clean speech to train i-vector-based ASR system [9]. A speech enhancement method was used to extract a clean speech on the basis of DNN, which was then used to train the ASR system in [10]. Several methods to improve PIA have also been proposed based on speech enhancement for robust ASR systems in i-vector level: application of maximum a posteriori (MAP) estimation procedure assuming an additive model for noise a clean i-vectors were estimated from the noisy i-Vectors reported in [11]: application of Minimum Mean Square Error (MMSE) based estimator reported in to noisy i-vector [12].

In this study, I propose a novel model-based approach to improve the PIA performance of the ASR system, in which the presence of strong disturbances and channel deformities is driven by the LP, based Support Vector Machine (SVM) learning algorithms. LP based algorithm manages by listing training speech data in subsets based on an appropriate difficulty criteria.

Automatic Speaker Recognition System Using i-vector PLDA

Speaker specific feature extraction by i-vector

i-vector represent the expressions of speech in compact manner, preserving the speaker-specific features which are most important component for ASR system [5]. An i-vector extractor is a mechanism that maps the speaker-specific GMM mean super vector M in terms of speaker and channel independent supervector m and a total variability matrix T which is low rank. The vector w is represented as:

$$M = m + Tw \qquad (1)$$

Where w represents random vector in eqn.(1) with $N(0, I)$ standard normal distribution. The learned using large amount of training data is represented as matrix T. The subspace spanned by is obtained from a large collection of data representative of the task at hand by ML estimation [13].

Probabilistic Linear Discriminant Analysis (PLDA)

A PLDA learning with within and across-class of a very large labeled training data set variability by making use of Expectation Maximization (EM) algorithm [14]. Considering total training utterances of a speaker are J, a speaker i can be represented from entire collection of J i-vectors as $\{n_{i,j} : j = 1, 2, 3, \dots, J\}$. Using Gaussian PLDA (G-PLDA) formulation this collection $\{n_{i,j} : j = 1, 2, 3, \dots, J\}$ can be represented as,

$$n_{i,j} = m + \Phi\beta_i + \epsilon_{i,j} \qquad (2)$$

In eqn. (2) global offset is represented as m, basic speakers subspace is represented by column vector of Φ, coordinate in speakers subspace is represented by β_i, zero mean and covariance of Gaussian is represented by $\epsilon_{i,j}$. Large collection of speaker-labeled i-vector can be estimated by using EM algorithm and G-PLDA parameters can be represented as $\{m, \Phi, \Sigma\}$. Given a test utteranceof a speaker, the recognition rate can be calculated using a closed form solution with the G-PLDA model as presented in [15]. A single speaker test requires access to the mean of the speaker registration i-vector, G-PLDA model parameter $\{m, \Phi, \Sigma\}$ and test i-vector.

SVM and probabilistic linear discriminant analysis based training approached (SVM-PLDA)

The fundamental step in this approach is to divide the training data in the subset on the basis of a difficulty criterion and to increase the difficulty; the training algorithm has to gradually introduce more data. I recommend that the main power of the SVM-based learning algorithm is generated by gradual increase in the diversity of training samples, which leads to better normalization of projected model parameters and also provides guidance for better localization. The proposed technique is different from the approach based on data selection that is reported in [16, 17], I do not try to find a subset of data that can provide comparative or better performance than normal PDA. Instead, SVM-based technique provides training on all available data and learning algorithm is allowed. Training parameters are used sequentially to a difficulty criterion that allows better assessment of model parameters. The basic requirement of any SVM-based learning algorithm is metrics for difficulty in training examples, which are presented later. It has been assumed that SVM approaches in ASR can function similarly to continuation methods [18, 19], in which the start of the optimization procedure with a simpler and more uniform version of a non-convex function can drive the estimated parameters to a dominant best of the function. In the SVM-PLDA paradigm, this equates to better estimation of PLDA model parameters, allowing the learning algorithm to achieve better overall minima than a traditional multi-session PLDA training approach. In addition, it has also been hypothesized that SVM-based approaches may also have the effect of a regularized leading to better performance on the test data [18]. I believe that our proposed SVM-PLDA approach can also benefit from greater generalization by adapting an estimation procedure based on an SVM model.

To deepen the SVM-PLDA parameters, I first propose a treatment for EM G-PLDA training. Using the derivation of PLDA model parameters given in [14], I can calculate the order moments of β_i in a simplified form for speaker i with utterances J as:

$$E[\beta_i\beta_i^T] = (J\Phi^T\Sigma^{-1})^{-1} + E[\beta_i]E[\beta_i] \qquad (3)$$

Considering a speaker I can calculate the complete G-PLDA (m, Φ, Σ) model parameters as $m = {}^{1}/_{IJ} \sum_{i,j} \eta_{i,j}$, $\Phi = \left[\sum_{i,j}\left(\eta_{i,j} - m\right)\right] E[\beta_i^T]\left[\sum_{i,j} E[\beta_i\beta_i^T]\right]^{-1}$ and $\Sigma = {}^{1}/_{i,j} \sum_{i,j}\left[\left\{\eta_{i,j} - m\right\}\left\{\eta_{i,j} - m\right\}^T - \Phi E[\beta_i]\left(\eta_{i,j} - m\right)^T\right]$.

To take into account the difference in training duration and enrollment tests form separate PLDA and SVM-PLDA backbends for the different sets of enrollment tests with corresponding durations to those of the enrollment tests. The main approach in this work, UBM and the i-vector extractor are formed on whole utterances, while the i-vectors to form PLDA / SVM-PLDA models are extracted using mismatch duration.

The steps (S1,…., S13) involved to compute SVM-PLDA algorithm for noise robust ASR system is represented as follows:

S1 Partitioning of training data progressively difficult and distinct into K subsystems $d_1, d_2, \dots\dots, d_k$
S2 Randomly initializing current SVM-PLDA parameters Φ_{cu} and Σ_{cu}
S3 $d_{cu} \leftarrow \emptyset$ as a current empty dataset
S4 Select suitable value n for iterations to identify new dataset
S5 For k=1:K
S6 $d_{cu} \leftarrow d_{cu} \cup d_k$
S7 For l=i:n
S8 Update values as Φ_{up} and Σ_{up}
S9 $\Phi_{cu} \leftarrow \Phi_{up}$
S10 $\Sigma_{cu} \leftarrow \Sigma_{up}$
S11 END
S12 END
S13 SVM-PLDA updated parameters are Φ_{cu} and Σ_{cu}.

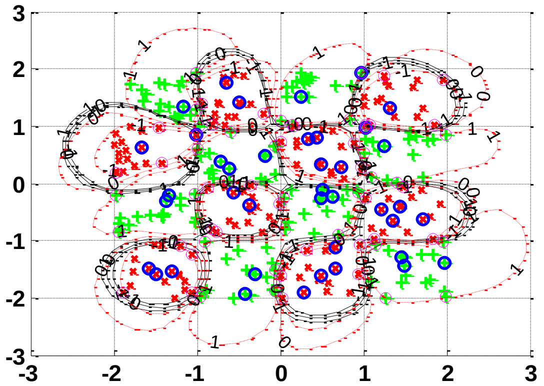
The training for the ISVM-PLDA (Inverse Support Vector Machine) algorithm differs from SVM-PLDA only in the organization of the learning data in the opposite direction. The organization of learning data in ISVM-PLDA may seem exactly opposite to the way humans actually learn. However, I hypothesize that the gradual change in the difficulty of the training data, starting from the most difficult data and then including relatively simpler data, can lead to a better regularization compared to the traditional PLDA to several sessions, with a consequent improvement in ASR performance. The steps (S1,…., S13) involved to compute ISVM-PLDA algorithm for noise robust ASR system is represented as follows:

S1 Partitioning of training data progressively less difficult and distinct into K subsystems $d_1, d_2, \ldots\ldots, d_k$
S2 Randomly initializing current ISVM-PLDA parameters Φ_{cu} and Σ_{cu}
S3 $d_{cu} \leftarrow \emptyset$ as a current empty dataset
S4 Select suitable value n for iterations to identify new dataset
S5 For k=1:K
S6 $d_{cu} \leftarrow d_{cu} \cup d_k$
S7 For l=i:n
S8 Update values as Φ_{up} and Σ_{up}
S9 $\Phi_{cu} \leftarrow \Phi_{up}$
S10 $\Sigma_{cu} \leftarrow \Sigma_{up}$
S11 END
S12 END
S13 ISVM-PLDA updated parameters are Φ_{cu} and Σ_{cu}.

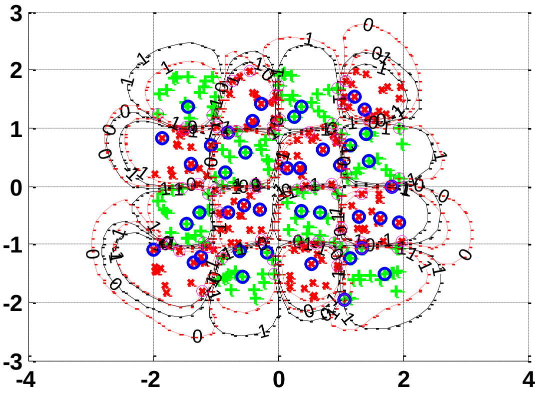
SVM and i-vector based speaker specific feature extractor (SVM-IVSSFE)

The SVM-IVSSFE algorithm estimates i-vector T extractor matrix using an inspired SVM approach. I distribute the training data in subsets based on a difficulty criterion and progressively using more difficult data to estimate i-vector extractor. The details of the estimation of the i-vector extractor table are available in [20, 21]. In SVM-IVSSFE, before the first iteration of the formation, the i-vector T extractor matrix is initialized using UBM. Once training has begins, T is estimated using a predetermined subset of data, d_{cu} to which the training data is added only sequentially in the increasing order of difficulty. Like the number of iterations increase, relatively more difficult data are included in the subset of training d_{cu}. The various steps (s1, ….,S12) involved to compute SVM-IVSSFE algorithm is represented as follows:

S1 Partitioning of training data progressively less difficult and distinct into K subsystems $d_1, d_2, \ldots\ldots, d_k$
S2 Randomly initializing current ISVM-PLDA parameters Φ_{cu} and Σ_{cu}
S3 $d_{cu} \leftarrow \emptyset$ as a current empty dataset
S4 Select suitable value n for iterations to identify new dataset
S5 For k=1:K
S6 $d_{cu} \leftarrow d_{cu} \cup d_k$
S7 For l=i:n
S8 Update values as Φ_{up} and Σ_{up}
S9 $\Phi_{cu} \leftarrow \Phi_{up}$
S10 $\Sigma_{cu} \leftarrow \Sigma_{up}$
S11 END
S12 END
S13 ISVM-PLDA updated parameters are Φ_{cu} and Σ_{cu}

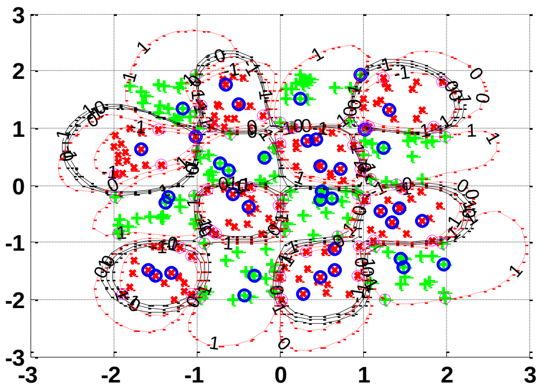


(a) SVM-PLDA algorithm for noise robust ASR

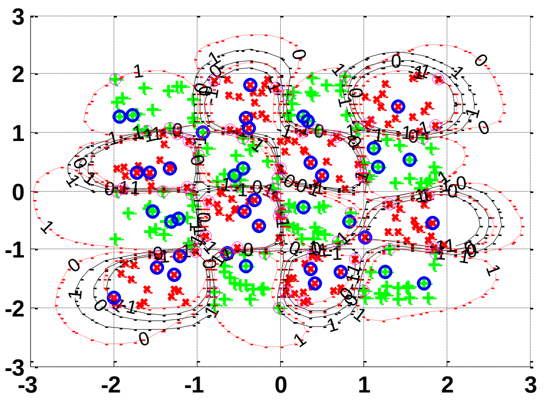


(b) ISVM-PLDA algorithm for noise robust ASR

Figure 1: Support Vector Machine- Probabilistic Linear Discriminant Analysis and Inverse Support Vector Machine-Probabilistic Linear Discriminant Analysis algorithm for noise robust ASR system.



(a) SVM-PLDA algorithm for noise robust ASR



(b) SVM-IVSSFE algorithm for noise robust ASR

Figure 2: Support Vector Machine- Probabilistic Linear Discriminant Analysis and i-vector Based Speaker Specific Feature Extractor algorithm for noise robust ASR system.

In Figure 1. (a) The hyper planes computed with SVM-PLDA algorithm for noise robust ASR and (b) the hyper planes computed with ISVM-PLDA algorithm for noise robust ASR for the hyper plane comparison. Figure 1. represents support vector machine- probabilistic linear discriminant analysis and inverse support vector machine- probabilistic linear discriminant analysis algorithm for noise robust ASR system.

In Figure 2.(a) The hyper planes computed with SVM-PLDA algorithm for noise robust ASR and (b) the hyper planes computed with SVM-IVSSFE algorithm for noise robust ASR. Figure 2. Represents support vector machine probabilistic linear discriminant analysis and i-vector based speaker specific feature extractor algorithm for noise robust ASR system.

Speaker recognition PIA experimental setup

Mel frequency cepstral coefficient (MFCC) with 25 ms frame size and 10 ms frame shift have been used in SVM-PLDA, ISVM-PLDA and SVM-IVSSFE based ASR systems. Steps involved extracting the feature (a) Divide test/train speech signal into short overlapping segments of 25 ms. (b) Apply Hamming and Hanning windowing techniques. (c) Take logarithm. (d) Consider 24 channel filter bank energy coefficients and apply Mel scale filtering analysis [22, 23]. (e) Apply DCT on the filter bank energy parameters and retaining the12 coefficients after DCT and discarding the remaining [24, 25, 26].

Noisy NIST 2010 data set for training and testing ASR system

The noisy version of NIST SRE 2010 created using the FaNT Toolkit, by adding noise to the PRISM for PIA analysis of ASR system. PRISM noise reduction involves real noise collected in crowded area like bars, canteens, offices and airports. I followed the following protocol to create a noisy NIST SRE 2010 C5 condition training version (PLDA) and testing ASR system. (i) A09 to A15 training data for PLDA was randomly divided into three approximate subsets and each subset has been added randaly 5 dB, 10 dB and 15 dB noise respectively using FaNT toolkit. (ii) B1 to B4 enrolment data has been divided into three subsets and each subset has been added randaly 5 dB, 10 dB and 15 dB noise respectively using FaNT toolkit.(iii) The test data was randomly divided into 3 approximate subsets. Then, within each of these subsets, test noise of PRISM set (B5 to B8) was added using the face toolkit respectively in 5 dB, 10 dB and 15 dB respectively.

Table 1: Noisy version of NIST SRE 2010 C5 Condition of Training, Enrollment and Test Data set

Data Set	Training	Enrollment	Test
Noise Level	5dB	5dB	5dB
	10dB	10dB	10dB
	15dB	15dB	15dB
Number of utterances	12000	4000	255
	12000	4000	255
	12612	3983	257

This is a true realistic scenario of handling noise, in which the enrollment and test data sets have different noise levels compared with the same noise level added to the enrollment and test data set at the same dB level. Table 1 shows the structure of training, enrollment and test set of data the noise version of NIST SRE 2010 C5 condition tests used in this research work.

Percentage of identification accuracy and discussion

For the analysis of PIA, I designed another system that includes the original clean training data in addition to the noisy training package. I made it so that SVM-PLDA could observe some clear statements, thus increasing the diversity of the training data set for the ASR system. It should be noted that access to your data corresponding to the noisy assembly is not a requirement of the SVM-PLDA algorithm. I hypothesized that including a subset of clean data with statements that were not used to create the learning set noise is also sufficient to increase the diversity of learning SVM-based PLDA training algorithms in the proposed ASR. The SVM-PLDA system started training with clean data followed by 20 dB, 15 dB, and 8 dB SNR respectively. To estimate the SVM-PLDA ASR system model parameters, I used 10 iterations for each new subset of learning for all except the last subsets. After including the data from the last subgroup of 8dB SNR, I used 5 training iterations to obtain the final parameter model to maintain the number of iterations comparable to the reference system.

For ISVM-PLDA experiments, the same 3 groups of instructions with SNR 5dB, 7dB, 10dB, and 15dB, but the formation proceeded in inverse mode starting from 15dB followed by 10dB, 7dB and 5dB respectively. For the ISVM-PLDA format system with an enhanced training package that also included its own data, the training included a further subset of data corresponding to the specific data set at the end. For any closed-set speaker recognition problem, speaker recognition accuracy is defined as follows.

$$PIA = \frac{No\ of\ utterances\ correctly\ identified}{Total\ No\ of\ utterences\ under\ test}$$

A total of 12000 utterances were put to test with SNR 5dB, 7dB, 10dB, and 15dB, Speaker recognition efficiency is compared with the performances of the SVM-PLDA, ISVM-PLDA and SVM-IVSSFE algorithms. Table II shows the PIA for SVM-PLDA, ISVM-PLDA and SVM-IVSSFE algorithms based ASR system. It can be observed from Table II that the SVM-IVSSFE algorithms show significant improvement up to 99.6% of speaker recognition efficiency for clean training and testing speech data. When ASR system trained with noisy and clean speech data and tested with clean speech data then the ASR system shows 99.4% of efficiency. The overall result represented in Table 2 shows some relevant observations as the PIA of SVM-IVSSFE combination is performing well compare to other combinations. SVM-IVSSFE ASR system shows 1.99% of improvements compare to SVM-PLDA ASR system as well as SVM-IVSSFE ASR system shows 1.5% of improvements compare to ISVM-PLDA ASR system.

Table 2: PIA for SVM-PLDA, ISVM-PLDA and SVM-IVSSFE algorithms based ASR system

ASR	ASR Training Data	ASR Testing Data	Percentage of EER	PIA
SVM-PLDA ASR System	C	C	02.39	97.8
	C	N	08.99	97.0
	N	C	05.97	97.1
	N+C	C	03.89	97.6
ISVM-PLDA ASR System	C	C	02.39	98.1
	C	N	08.99	97.2
	N	C	05.97	97.3
	N+C	C	03.89	97.9
SVM-IVSSFE ASR System	C	C	02.39	99.6
	C	N	08.99	97.7
	N	C	05.97	98.1
	N+C	C	03.89	99.4

Figure 3. shows DET plots for SVM-PLDA for the noisy NIST SRE 2010 speech data training sets that also include clean as well as noisy utterances. It is clear from these plots that the minimum false alarm probability 1% achieved.

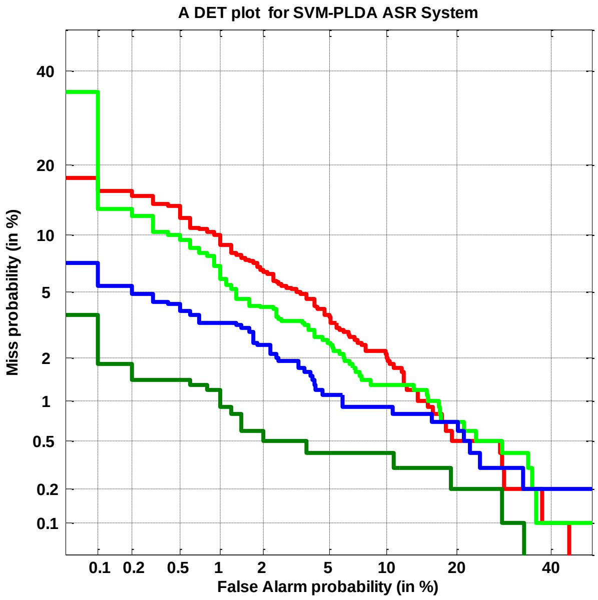


Figure 3: DET plots for SVM-PLDA for the noisy NIST SRE 2010 speech data.

Figure 4. shows DET plots for ISVM-PLDA for the noisy NIST SRE 2010 speech data training sets that also include clean as well as noisy utterances. It is clear from these plots that the minimum false alarm probability 1.1% achieved.

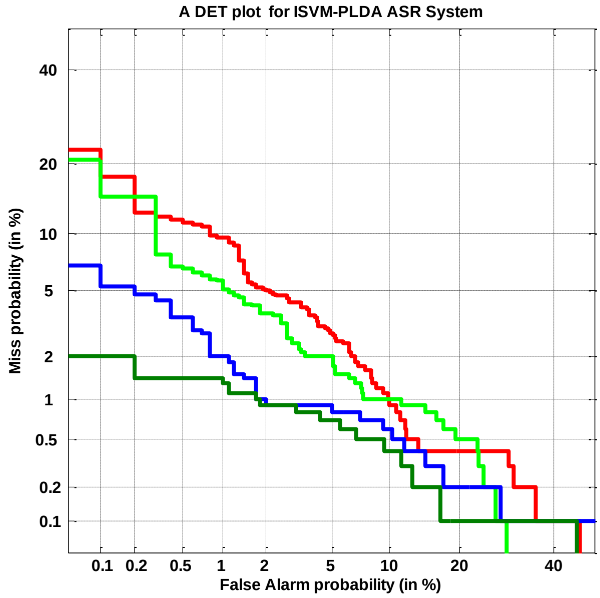


Figure 4: DET plots for ISVM-PLDA for the noisy NIST SRE 2010 speech data.

Figure 5. shows DET plots for SVM-IVSSFE for the noisy NIST SRE 2010 speech data training sets that also include clean as well as noisy utterances. It is clear from these plots that the minimum false alarm probability 0.8% achieved.

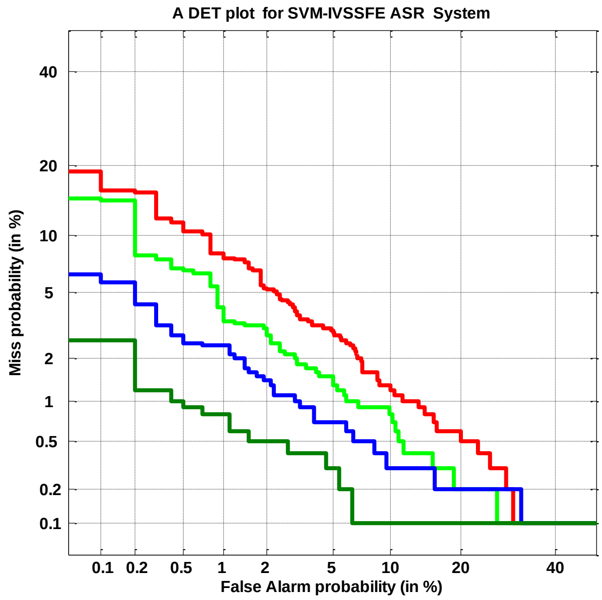


Figure 5: DET plots for SVM-IVSSFE for the noisy NIST SRE 2010 speech data.

CONCLUSION

The SVM-IVSSFE ASR algorithm operates by initializing the training with easier data and gradually includes difficult data as the training progresses. I hypothesize this approach of gradually increasing the diversity of the training leads to better estimation of the model parameters. For the severely noisy and degraded trials, proposed SVM-IVSSFE based ASR system achieves hieghest efficiency 99.6% and 99.4% compared to a regular baseline SVM-PLDA system. SVM-IVSSFE ASR system shows 1.99% of improvements compare to SVM-PLDA ASR system as well as SVM-IVSSFE ASR system shows 1.5% of improvements compare to ISVM-PLDA ASR system.

REFERENCES

[1] S Singh, "Forensic and Automatic Speaker Recognition System," International Journal of Electrical and Computer Engineering, vol. 8, no 5, 2018.

[2] S. Braun, D. Neil, and S.-C. Liu, "A curriculum learning method for improved noise robustness in automatic speech recognition," arXiv preprint arXiv:1606.06864, 2016.

[3] D. Amodei, R. Anubhai, E. Battenberg, C. Case, J. Casper, B. Catanzaro, J. Chen, M. Chrzanowski, A. Coates, G. Diamos et al., "Deep speech 2: End-to-end speech recognition in English and Mandarin," arXiv preprint arXiv:1512.02595, 2015.

[4] M. P. Kumar, B. Packer, and D. Koller, "Self-paced learning for latent variable models," NIPS, pp. 1189–1197, 2010.

[5] N. Dehak, P. Kenny, R. Dehak, P. Dumouchel, and P. Ouellet, "Frontend factor analysis for speaker verification," IEEE Trans. Audio, Speech, and Language Processing, vol. 19, no. 4, pp. 788–798, 2011.

[6] J. Pelecanos and S. Sridharan, "Feature warping for robust speaker verification," ISCA ODYSSEY 2001, 2001.

[7] S. Singh, Abhay Kumar, David Raju Kolluri, "Efficient Modelling Technique based Speaker Recognition under Limited Speech Data," International Journal of Image, Graphics and Signal Processing. vol.8(11) 2016, pp.41-48.

[8] O. Plchot, L. Burget, H. Aronowitz, and P. Mat¨ejka, "Audio enhancing with DNN autoencoder for speaker recognition," IEEE ICASSP 2016, pp. 5090–5094, 2016.

[9] F. Richardson, B. Nemsick, and D. Reynolds, "Channel compensation for speaker recognition using map adapted PLDA and denoising DNNs," ISCA ODYSSEY 2016, pp. 225–230, 2016.

[10] J. Chang and D. Wang, "Robust speaker recognition based on DNN/i- Vectors and speech separation," IEEE ICASSP 2017, pp. 5415–5419, 2017.

[11] W. B. Kheder, D. Matrouf, J.-F. Bonastre, M. Ajili, and P.-M. Bousquet, "Additive noise compensation in the i-Vector space for speaker recognition," IEEE ICASSP 2015, pp. 4190–4194, 2015.

[12] W. B. Kheder, D. Matrouf, M. Ajili, and J.-F. Bonastre, "Probabilistic approach using joint clean and noisy i-Vectors modeling for speaker recognition," ISCA INTERSPEECH 2016, pp. 3638–3642, 2016.

[13] P. Kenny, "Bayesian speaker verification with heavy-tailed priors." ISCA ODYSSEY 2010, 2010.

[14] S. J. Prince and J. H. Elder, "Probabilistic linear discriminant analysis for inferences about identity," IEEE ICCV-2007, pp. 1–8, 2007.

[15] D. Garcia-Romero and C. Y. Espy-Wilson, "Analysis of i-Vector length normalization in speaker recognition systems." ISCA INTERSPEECH 2011, pp. 249–252, 2011.

[16] S. Biswas, J. Rohdin, and K. Shinoda, "i-Vector selection for effective PLDA modeling in speaker recognition," ISCA ODYSSEY 2014, pp. 100– 105, 2014.

[17] L. M. Arslan and J. H. L. Hansen, "Selective training for hidden Markov models with applications to speech classification," IEEE Transactions on Speech and Audio Processing, vol. 7, no. 1, pp. 46–54, 1999.

[18] Y. Bengio, J. Louradour, R. Collobert, and J. Weston, "Curriculum learning," ICML, pp. 41–48, 2009.

[19] E. L. Allgower and K. Georg, "Numerical continuation methods: an introduction," vol. 13, 2012.

[20] P. Kenny, G. Boulianne, and P. Dumouchel, "Eigenvoice modeling with sparse training data," IEEE Trans. Speech and Audio Processing, vol. 13, no. 3, pp. 345–354, 2005.

[21] S.Singh, Mansour. H. Assaf and Abhay Kumar, "A Novel Algorithm of Sparse Representations for Speech Compression/Enhancement and Its Application in Speaker Recognition System," International Journal of Computational and Applied Mathematics, vol 11, no 1 , pp. 89-104, 2016.

[22] S.Singh and Ajeet Singh "Accuracy Comparison using Different Modeling Techniques under Limited Speech Data of Speaker Recognition Systems" Global Journal of Science Frontier Research: F Mathematics and Decision Sciences, vol. 16, Issue 2, pp.1-17, 2016.

[23] S.Singh and Mansour H. Assaf "A Perfect Balance of Sparsity and Acoustic hole in Speech Signal and Its Application in Speaker Recognition System" Middle-East Journal of Scientific Research. Vol. 24, No.11, pp. 3527-3541, 2016.

[24] S.Singh , Assaf Mansour H, Abhay Kumar and Nitin Agrawal, "Speaker Recognition System for Limited Speech Data Using High-Level Speaker Specific Features and Support Vector Machines," International Journal of Applied Engineering Research (IJAER), Vol.

12, Number 19, pp. 8026-8033, 2017.

[25] S.Singh and Dr. E.G. Rajan "MFCC VQ Based Speaker Recognition and Its Accuracy Affecting Factors" International Journal of Computer Application.vol 21(6), pp 1-6, 2011.

[26] S.Singh and Dr. E.G. Rajan "Application Of Different Filters In Mel Frequency Cepstral Coefficients Feature Extraction And Fuzzy Vector Quantization Approach In Speaker Recognition" International Journal of Engineering Research & Technology, vol. 2 Issue 6, 2013.