

Dr. Nirmalya Sen

Email: nirmalyasen75@gmail.com



I am a **Machine Learning and Signal Processing** professional worked mainly on Audio signal. I have worked on various audio features, i.e., linear predictive coefficients, linear predictive cepstral coefficients, Mel frequency cepstral coefficients, inverted Mel frequency cepstral coefficients, prosodic features and GMM supervector. In classifier side, I have worked on vector quantization, Gaussian mixture model, Gaussian mixture model with universal background model, polynomial classifier and support vector machine. I have worked in detailed on Expectation Maximization (EM) algorithm and how EM is applied on Gaussian mixture model (GMM) and Hidden Markov model (HMM). I have exposure on image processing. I worked in number plate detection task. I applied Machine Learning techniques to classify/detect/categorize (clustering) as per the use cases.

CURRENT STATUS: Working as Associate Professor in the R. H. Sapat College of Engineering, Management Studies and Research, Nasik, Maharashtra.

INDUSTRY EXPERIENCE:

Research Engineer

Videonetics Technology Pvt. Limited (March, 2015 to August, 2016)

I was developing the **Audio Analytics** product for security system. I conceived the whole project from scratch. Audio Analytics products analyze the input audio signals which come from IP cameras and generate various alerts based on the results of analysis.

I was also involved in the project of **Automatic Vehicle Number Plate Recognition** task. Here, I was working on the feature extraction section to improve the recognition accuracy of the number plate.

Programming languages used: MATLAB, C

TEACHING EXPERIENCE:

Designation	Organization	Period		Nature of Job
		From	To	
Associate Professor	R. H. Sapat College of Engineering, Management Studies and Research, Nasik, Maharashtra	July 2017	Till date	Teaching DSP; Analog Communication
Assistant Professor	SGGSIE&T, Nanded, Maharashtra	August 2016	June 2017	Taught Signals & Systems; DSP
Assistant Professor	NITMAS, Diamond Harbour, West Bengal	July 2011	December 2012	Taught Signals & Systems; DSP
Scholar (Ph.D.)	IIT Kharagpur	July 2006	December 2010	Took tutorials & labs of DSP and digital speech processing

EDUCATION:

Examination	University	Year	% of marks	Subject
Ph.D.	IIT Kharagpur	2014	NA	Signal Processing and Machine Learning (Speaker Recognition)
M. Tech.	SGGSIE&T, Nanded, MS	2006	9.3 CGPA in a 10 point scale	Electronics Engineering
GATE	NA	2003 and 2004	92.77 percentile and 93.62 percentile	Electronics & Communication Engineering
B.E	Amravati University, MS	2002	69.3 %	Electronics & Telecommunication Engineering

COURSES ATTENDED AT THE TIME OF Ph.D.: Linear Algebra, Estimation of Signals and Systems, Matrix Algebra, Adaptive Signal Processing, Multirate Filter Banks & Wavelets.

Subjects of Interest:

I want to teach following subjects:

1. Signals and Systems.
2. Digital Signal Processing.

I want to teach following subjects in near future:

1. Linear Algebra for engineers.
2. Probability and Statistics for engineers.
3. Digital Image Processing.
4. Applied Multivariate Statistical Modeling.
5. Pattern Recognition.

List of Publications:

Book Chapters:

1. Sen, N., Patil, H. A., Mandal, S. Kr. D., & Rao, K. S., 2013. Importance of Utterance Partitioning in SVM Classifier with GMM Supervectors for Text-Independent Speaker Verification. Lecture Notes in Artificial Intelligence, vol. 8284, Springer International Publishing, Switzerland. pp. 780-789.
2. Sen, N., & Basu, T. K., 2012. A critical comparison between GMM classifier and polynomial classifier for text-independent speaker identification. In: Frontiers in Computer Education, Advances in Intelligent and Soft Computing, vol. 133, Springer Berlin Heidelberg. pp. 545-550.

3. Sen, N., & Basu, T. K., 2011. Features extracted using frequency-time analysis approach from Nyquist filter bank and Gaussian filter bank for text-independent speaker identification. Lecture Notes in Computer Science, vol. 6583, Springer Berlin Heidelberg. pp. 125-136.

Conference Publications:

1. Sen, N., Basu, T. K., & Chakroborty, S., January, 2011. Comparison of features extracted using time-frequency and frequency-time analysis approach for text-independent speaker identification. In: IEEE National Conference on Communications, NCC 2011. DOI: 10.1109/NCC.2011.5734720.
2. Sen, N., & Basu, T. K., January, 2011. Temporal energy and correlation features from Nyquist filter bank for text-independent speaker identification. In: IEEE Students' Technology Symposium, TechSym 2011. pp. 166-170.
3. Sen, N., Mandal, S. K. D., & Basu, T. K., February, 2011. Feature Set Extracted Using Frequency-Time Analysis Approach for Speaker Verification. In: IEEE International Conference on Devices and Communications, ICDeCom 2011. DOI: 10.1109/ICDECOM.2011.5738541.
4. Sen, N., Basu, T. K., & Chakroborty, S., February, 2011. Importance of using Log Function to Reduce the Correlation between Features in a Multidimensional Feature Space for Text-Independent Speaker Identification. In: IEEE International Conference on Devices and Communications, ICDeCom 2011. DOI: 10.1109/ICDECOM.2011.5738537.

5. Sen, N., Basu, T. K., & Patil, H. A., July, 2010. Significant improvement in the closed set text-independent speaker identification using features extracted from Nyquist filter bank. In: IEEE International Conference on Industrial and Information Systems, ICIIS 2010. pp. 303-308.
6. Sen, N., Gupta, R. D., & Chakroborty, S. July, 2010. Measurement of the correlation in a multidimensional feature space. In: IEEE International Conference on Industrial and Information Systems, ICIIS 2010. pp. 452-457.
7. Sen, N., Basu, T. K., & Patil, H. A., December, 2010. New features extracted from Nyquist filter bank for text-independent speaker identification. In: Annual IEEE India Conference INDICON 2010. DOI: 10.1109/INDCON.2010.5712689.
8. Sen, N., & Basu, T.K., December, 2009. A New Nyquist window with near optimal time-bandwidth product. In: Annual IEEE India Conference INDICON 2009. DOI: 10.1109/INDCON.2009.5409413.
9. Sen, N., Patil, H. A., & Basu, T. K., December, 2009. A New transform for robust Text-Independent speaker identification. In: Annual IEEE India Conference INDICON 2009. DOI: 10.1109/INDCON.2009.5409372.
10. Sen, N., Shashank, Adhikary, A., & Basu, T. K., 2009. A New Feature from Modified STFT Filter Bank for Robust Text-Independent Speaker Identification. In: IICAI 2009. pp. 1701-1716.

TITLE OF THE Ph.D. THESIS: Enhancement of Speaker Recognition Performance for Short Test Segments using GMM-SVM and Polynomial Classifiers.

ABSTRACT OF THE Ph.D. THESIS

This thesis investigates the effect of the length of the test segments on the speaker recognition performance in the context of GMM-UBM, GMM-SVM and Polynomial classifiers. Results reveal that for long test segments the GMM-SVM classifier performs better than classical GMM-UBM classifier but for short test segments the performance of GMM-SVM classifier is very poor compared to the classical GMM-UBM classifier. Results also indicate that, the performance of GMM-SVM classifier for short test segments decreases even further when the dimension of the GMM supervector increases. We have shown that, the cause for poor performance of GMM-SVM classifier for short test segments is the difference between the amount of MAP adaptations of training GMM supervectors and test GMM supervectors for the same speaker. To enhance the performance of GMM-SVM classifier for short test segments, we have proposed utterance partitioning for speaker class as well as impostor class. Results confirm that, partitioning of the training utterance can significantly reduce the problem of mismatch between the MAP adaptation of training GMM supervectors and test GMM supervectors. By choosing appropriate number of partitions of the training utterance, the performance of GMM-SVM classifier is far better than the classical GMM-UBM classifier for short test segments. We have also observed that, when the dimension of the GMM supervector increases, the overlap between the classes increases significantly. The intuitive reason for the greater overlap in higher dimensional GMM supervector space has been explained in this thesis.

This thesis has thoroughly investigated the necessity of utterance partitioning for speakers who belong to the impostor class for GMM-SVM classifier. We have shown that, if we do not perform utterance partitioning for the speakers of impostor class then, the average number of cohorts per speaker increases significantly and for short test segments, performance of GMM-SVM classifier becomes very poor. We have also shown that, the necessity of utterance partitioning for speakers belonging to the impostor class increases even more for higher dimensional GMM supervector space.

In this thesis, we have also investigated the effect of data imbalance problem in GMM-SVM classifier using minority (positive) class over-sampling approach. Results reveal that, performance of the GMM-SVM classifiers does not improve even after sufficient enhancement in the number of support vectors for the positive class of the SVM classifier. Thesis has given detailed intuitive reasoning for why the problem of data imbalance between speaker class and impostor class does not have adverse consequence on the performance of GMM-SVM classifier.

In addition to the over-sampling of the minority (positive) class data, we have also performed under-sampling of the majority (negative) class data (NCUS) by removing the support vectors of the majority class. Thesis demonstrates that after NCUS operation, the decision hyperplane of the SVM classifier shifts toward the ideal position however, the orientation of the decision

hyperplane is not satisfactory. When false alarm probability is high, the performance of the GMM-SVM classifier improves after NCUS operation. However, when false alarm probability is low, the performance of the GMM-SVM classifier decreases after NCUS operation due to the rejection of proper cohorts. We have also performed experiments using max rule speaker comparison with inner product discriminant functions (MRSCIPDF) technique. Thesis reveals that, similar to the GMM-SVM classifier, the partitioning of training utterance operation can improve the performance of MRSCIPDF technique significantly in recognition of short test segments and modestly in long test segments. Thesis comprehensively concludes that, the improvement in speaker verification performance after partitioning of the training utterances in GMM-SVM classifier is not due to the reduction in data imbalance problem. Reduction in the amount of mismatch of MAP adaptations between training GMM supervectors and test GMM supervectors is the reason for improvement in the performance of GMM-SVM classifier when partitioning of training utterance is performed.

Thesis highlights the problem of data imbalance and its adverse effects on the performance of polynomial classifier when trained with mean-square error criterion. To overcome the problem of data imbalance, we have performed the training of polynomial classifier using weighted mean-square error criterion. We have shown that, after training using weighted mean-square error criterion, polynomial classifier performs better than GMM classifier even for short test segments.

Key words: Speaker Recognition; GMM-UBM Classifier; GMM-SVM Classifier; Polynomial Classifier; Short Test Segments; GMM Supervector; Mismatch of MAP Adaptations; Utterance Partitioning; Data Imbalance Problem; NCUS; MRSCIPDF Technique; Mean-Square Error Criterion; Weighted Mean-Square Error Criterion.

REFERENCES:

Dr. Ajoy Kumar Ray
Professor, Department of Electronics and Electrical Communication Engineering, IIT Kharagpur
E-mail - ajoy_ray2004@yahoo.com, Ph: +919434004711

Dr. Tapan Kumar Basu
Retired Professor, Department of Electrical Engineering, IIT Kharagpur
Senior professor, Department of Electrical Engineering, Academy of Technology, Grand Trunk Road, Adisaptagram, Hooghly, West Bengal, Pin: 712121
Email: basutk06@rediffmail.com; Ph: +919748061419

Dr. Anup Kumar Ray
Retired Professor, Centre for Educational Technology, IIT Kharagpur
Email: anuptrg@gmail.com; Ph: +919434753375

Dr. K. Sreenivasa Rao
Professor, School of Information Technology, IIT Kharagpur
Email: ksrao@sit.iitkgp.ernet.in; Ph: +919933750853

Dr. P.K. Dutta
Professor, Department of Electrical Engineering, IIT Kharagpur
Email: pkd@ee.iitkgp.ernet.in; Ph: +919434006243