Laxmi Pandey

1061 W Yosemite Avenue #8
Merced CA - 95348
Email-id : **lpandey@ucmerced.edu**, **pandey.laxmi21@gmail.com**
Mobile No.: **+16464794737**

ACADEMIC DETAILS

Examination	Institute	Year	CPI/%
Ph.D. (EECS)	UC Merced	2018	-
Master of Science (Electrical Engineering)	IIT kanpur	2017	7.5
Bachelors of Technology (Electronics and Communication Engg)	UPTU, Lucknow	2012	75.7%
Higher Secondary (CBSE)	KV IIT Kanpur	2007	63%
Secondary (CBSE)	KV IIT Kanpur	2005	75%

FIELDS OF INTEREST

Audio-Visual Speech Recognition and Perception, Human Computer Interaction, Machine Learning, Speech Denoising and Audio Source Separation.

PUBLICATIONS

- Aditya Raikar , Saurya Basu, Laxmi Pandey and Rajesh M. Hegde **"Multi-Channel Joint speech dereverberation and denoising using Deep Priors"**, INDICON 2018.
- Laxmi Pandey and Rajesh M. Hegde **"LSTM Based Attentive Fusion of Spectral and Prosodic Information for Keyword Spotting"**, Interspeech 2018.
- Laxmi Pandey, Anurendra Kumar and Vinay Namboodiri **"Monaural Audio Source Separation Using Variational Autoencoders"** , Interspeech 2018.
- Laxmi Pandey, Nitish Divakar, Krishna D N, Anuroop Iyengar **"Deep Clean: GPU Powered Speech Denoising using Adversarial Learning"** , GTC 2018.
- Laxmi Pandey and Rajesh M. Hegde **"Keyword Spotting in Continuous Speech using Spectral and Prosodic Information Fusion"** in CSSP, Springer,2018.
- Laxmi Pandey and Rajesh M. Hegde **"Fusion of Spectral and Prosodic Information using Combined Error Optimization for Keyword Spotting"**, National Conference on Communications, NCC 2017.
- L. Pandey, K. Nathwani, S. Kaur, I. Hussain, R. Pathak, G. Singh, S. Tiwari and Rajesh M. Hegde **"Domain Specific Audio Indexing Using Linguistic Information"**, Proceedings of IEEE Symposium on Signal Processing and Information Technology (ISSPIT), Noida, December 2014.

WORK EXPERIENCE (3+ years)

- Worked as a Research Associate at MIPS Lab, IIT Kanpur (July 2012-July 2015)
  - Developed an automatic speech recognition systems based on Hidden Markov Models using sphinx3 and Kaldi toolkit .
  - Developed a speech based search engine to search a spoken keyword with in an audio archive for the retrieval of keyword related information.
  - Developed a tool for real time goal detection tool for automatic generation of cheer sound effects at every goal hit in the soccer match.

RESEARCH AND ACADEMIC PROJECTS

- **Domain Specific Speech based Search Engine** *(Jan 2013 - Dec 2014), DIT*: Developed a novel methodology for indexing domain specific audio archives using linguistic information present in the speech signal. The audio indexing system is phone based and can work under limited training data conditions. A training data set that captures the linguistic information within Hindi language at the syllable level is first developed. A reduced phone set is then derived from the super syllabic set of the Hindi language. The system is then bootstrapped at the phone level with domain specific data. The audio indexing itself is then performed using a novel sliding phone protocol technique. The performance of such a audio indexing system is then evaluated for Indian parliament speech and read news. The proposed bootstrapping method with sliding phone search provides reasonable improvements in phone recognition accuracy and in terms of search retrieval efficiency when compared to conventional methods.

- **Development of Prosodically Guided Phonetic Engine for Hindi** *(July 2012 - Dec 2014), DIT*: Developed a phonetic engine using prosodic and phonetic information for Hindi language. Phonetic engine is a machine which represents all the information present in the speech so that speech can be exactly reproducible. It is thought of different tiers, viz., phonetic transcription, syllabification, pitch marking (pitch index) and break index (break marking). To demonstrate this work, a telephone based audio indexing system for searching through audio archive in Hindi is developed.
- **Build and demonstrate a real time continuous speech recognition system in Hindi** *(Jan 2016 - Apr 2016), Course project for EE627, Speech Signal processing* :Developed and Demonstrated a system for spotting keywords in continuous speech utterances. The system consists of a front-end and a back-end processing. The front-end consists of the word boundary identifier i.e. segmentation of the speech utterance into lexical items (words). The back-end is a language model with an associated dictionary i.e. matching the items with the units present in the memory and performing an action depending upon the semantics of the unit matched.
- **License Plate Detection and Recognition in Complex Scenes Using Mathematical Morphology and Support Vector Machines** *(Aug 2015 - Nov 2015), Term Paper for EE604, Image Signal processing* : A highly reliable license plate detection and recognition approach using mathematical morphology and support vector machines (SVM) is implemented. The approach is composed of three main stages including license plate detection, character segmentation and recognition. A preprocessing step is applied to improve the performance of license plate localization and character segmentation in case of severe imaging conditions. The first and second stages utilize edge detection, mathematical morphology followed by connected component analysis. While SVM is employed in the last stage to construct a classifier to categorize the input numbers of the license plate into one of 9 classes. The algorithm has been applied on 208 car images with different backgrounds, license plate angles, distances, lightning conditions, and colors. The average accuracy of the license plate localization is 97.60%, 90.74% for license plate identification, and 97.89% for number recognition.
- **Indoor Positioning Using UWB-IR Signals in the Presence of Dense Multipath with Path Overlapping** *(Aug 2015 - Nov 2015), Term Paper for EE602, Statistical Signal processing*: A method for positioning using ultra-wideband impulse radio (UWB-IR) signals that is robust in indoor environments characterized by dense multipath channel with path overlapping is implemented. Path overlapping effects arising from multipath in dense cluttered environments decrease the direct path resolution in time domain, and hence induce time-of arrival (TOA) and angle-of-arrival (AOA) based positioning inaccuracy. To mitigate this problem, system design that is capable of resolving the closely-spaced multipath at low cost is of value. The method yields the least-squares estimation of joint TOA and AOA with low computational cost. It is based on the spectral observation of beam forming, in which the path overlapping effect is mitigated using multipath-aided acquisition. The computational cost is reduced using in-band power spectrum and accurate initial estimations.

INDUSTRIAL COLLABORATION

- **Real Time Goal Detection in Soccer Match using Hidden Markov Model** *(Feb 2016 - Oct 2016), LG Soft India*: Developed a real time goal detection tool that will generate an automatic cheering effects at every goal hits in the soccer match video. An approaches used for the keywords spotter implementation is to consider individual models for the keywords i.e. actual model and to represent non keywords "background" or "garbage" models are used. Classifier is modelled using hidden Markov models (HMM). Conditional probability of test signal given actual model i.e. P(test | actual_model) and conditional probability of test signal given background model i.e. P(test | background_model) is compared with the theshold set to make final decision.
- **Development of SMS Compression Techniques for Indian Languages** *(Sept 2015 - Jan 2016), CDAC*: Developed a standard loss less compression algorithm that can allow more number of characters of Hindi language in SMS text. A novel encoding scheme is proposed along with several modifications to standard schemes making them efficient for transmission of Hindi and multilingual text. The encoding schemes allow the transmission of around 160 characters for pure Hindi, and multilingual text. The efficiency of the proposed schemes is evaluated by conducting experiments on a multilingual database specially collected from twitter using dictionary learning. Performance evaluation shows that these

encoding schemes allow nearly 160 characters per SMS for messages in both Hindi and multilingual text.

- **Video Analyzer using Command Line Video Quality Metric** *(March 2015 - Feb 2016), LG Soft India*: Designed and developed a video analyzer tool which is used for scoring the similarity or difference between two videos on the basis of dropped frames, broken macro blocks, etc. Tool performs automated processing on a pair of video files. One contains an original video sequence (e.g., straight from the reference device) and the other contains a processed video sequence (e.g., after coding and transmission and decoding) and performs video calibration and video quality estimation.
- **Audio Analyzer using Dynamic Time Warping Algorithm** *(Jan 2014 - Feb 2015), LG Soft India*: Designed and developed an audio analyzer tool which is used for scoring the similarity and difference in audio effects between reference and test device. The proposed algorithm is capable of preprocessing the test music or speech signal for silence removal and extraction of meaningful parameters (also called feature extraction) for template matching using Dynamic Time Warping (DTW) method and Frame Matching(FM) using Sliding window method.
- **Digital Mandi for the Indian Kissan** *(July 2012 - Mar 2013), BSNL*: Digital Mandi Application for Indian Kisan is a unique web and cell phone based multimodal agriculture commodity price retrieval system. It has been developed by BITCOE (BSNL IIT-Kanpur Centre of Excellence) at IIT-Kanpur. My key role was to develop multi-lingual automatic speech recognition models, based on Hidden Markov Model, as a root for information retrieval system. Currently, the application has been deployed on the BSNL national network in Orissa and Haryana. The application service was formally inaugurated by Honorable Minister of Communications and IT on 29th August 2011 and is operational since then.

MASTER'S THESIS

Fusion of Spectral and Prosodic Information for Keyword Spotting in Continuous Speech
A methodology for information fusion of prosodic and spectral information in the syllabic search space at model level and feature level is proposed. In order to fuse information at feature level, I have concatenated the spectral features (MFCCs) and the prosodic features, which includes statistical features describing energy and pitch contour. In order to enhance the learning performance of HMM models for different prosodic variations of a particular syllabic unit, I proposed a linear transformation method to estimate the missing prosodic variations using the available ones in the dataset. Further, HMMs for different prosodies are learnt using the thus estimated feature set. Finally the decision from the HMMs of different prosodies is combined to reduce the word-error rate of the overall system. To perform all the experimentations, a huge and diverse database of publicly available audios has been collected. Variations in prosody in the acquired database has been studied extensively, which supports the notion of added discriminative nature with prosodic information in context of speech recognition.

TECHNICAL SKILLS

- **Programming Language:** C, C++, Microcontroller 8051 and Microprocessor 8085 (Assembly language)
- **Database:** MySQL (WAMP, LAMP)
- **Scripting Language:** Matlab, Octave, Shell, PHP, PRAAT, Python (tensorflow)

RELEVANT COURSES

- Statistical Signal Processing, Speech Signal Processing, Image Signal Processing, Introduction to Signal Analysis, Photonics, Networks and Switching.

REFERENCES

- Prof. Rajesh M. Hegde - *rhegde@iitk.ac.in*
  Professor, Electrical Engineering, IIT Kanpur, India
- Dr. Ahmed Arif - *asarif@ucmerced.edu*
  Assistant Professor, Electrical and Computer Science Engineering, UC Merced, USA
- Dr. Vinay Namboodri - *vinaypn@iitk.ac.in*
  Assistant Professor, Computer Science and Engineering, IIT Kanpur, India
- Dr. Karan Nathwani - *karan.nathwani@iitjammu.ac.in*
  Assistant Professor, Electrical Engineering, IIT Jammu, India