

PRASANTA KUMAR GHOSH

Objective Research in human speech communication science, engineering models for speech science – speech production, processing and their links – and its application to speech technology development.

Research Interests

- Models of speech production and its relation to speech perception/processing,
- Automatic speech recognition especially inspired by speech production and perception links,
- Non-stationary signal processing methods (with applications to speech and music),
- Auditory motivated Signal Processing for speech and audio,
- Sampling theory, Nonuniform sampling and interpolation.

Education **University of Southern California (USC), Los Angeles, USA**
PhD in Signal Processing at Signal Analysis and Interpretation Laboratory (SAIL), August 2006 - August 2011.
(Supervisor: Prof. Shrikanth S. Narayanan)
CGPA (current) - 4/4

Indian Institute of Science (IISc), Bangalore, India
M.Sc.(Engg.) in Signal Processing, July 2006.
(Supervisor: Prof. Thippur V. Sreenivas)
CGPA - 7.5/8

Jadavpur University, Kolkata, India
B.E. in Electronics and Telecommunication Engineering, June 2003.
CGPA - 4.89/5 (2nd in the class)

Awards and Honors

1. *Honorable mention for the best paper award for the year 2010-11* in the Ming Hsieh Department of Electrical Engineering, University of Southern California (USC).
2. *Ming Hsieh Institute (MHI) Ph.D. scholar for the year 2010-11* in the Ming Hsieh Department of Electrical Engineering, University of Southern California (USC).
3. *Center for Excellence in Teaching's Award for Excellence in Teaching in the category of Electrical Engineering for the year 2010-11 in University of Southern California (USC).*
4. *Best TA awards for the year 2007-08 and 2008-09 for graduate class in the Ming Hsieh Department of Electrical Engineering, University of Southern California (USC).*
5. *Best M.Sc. (Engg) thesis award for the year 2006-07 in the Electrical Engineering Division at IISc;* thesis title: 'Non-uniform Sample Based Speech Analysis and Coding'.
6. *IEEE Bangalore chapter, First Prize in Mr. BRV Varadhan Post-graduate Student Paper Contest 2005;* paper title: 'ESTL: Novel Temporal Feature and Its Application to Speech Segmentation'.
7. *Ministry of Human Resources Development (MHRD), Govt. of India, scholarship Holder 2004-06.*

Courses Taken

At USC:

1. Probability Theory for Engineers, 2. Wavelets, 3. Mathematical Pattern Recognition, 4. Random Processes in Engineering, 5. Empirical Methods in Natural Language Processing, 6. Adaptive Signal Processing, 7. Statistics for Engineers, 8. Neural Computation with Artificial Neural Networks, 9. Selected Topics in Machine Learning.

At IISc:

1. Neural Networks, 2. Computer Vision, 3. Speech Information Processing, 4. Spectrum Analysis, 5. Mathematical Methods, 6. Advanced Digital Signal Processing.

Publications

Journals:

1. **P. K. Ghosh** and S. Narayanan, “Automatic Speech recognition using articulatory features from subject-independent acoustic-to-articulatory inversion”, *J. Acoust. Soc. Am. Express Letters (JASAEL)*, Volume 130, Issue 4, Sep 2011, pp EL251–EL257.
2. **P. K. Ghosh**, L. M. Goldstein, S. Narayanan, “Processing speech signal using auditory-like filterbank provides least uncertainty about articulatory gestures”, *J. Acoust. Soc. Am.*, Volume 129, Issue 6, Jun 2011, pp 4014–4022.
3. **P. K. Ghosh**, L. M. Goldstein, S. Narayanan, “Auditory-like filterbank: An optimal speech processor for efficient human speech communication”, *Springer Proceedings of Indian Academy of Sciences (Sadhana)*, Special Issue on Speech Processing. 36(5): 699–712, October 2011.
4. **P. K. Ghosh** and S. Narayanan, “A generalized smoothness criterion for acoustic-to-articulatory inversion”, *J. Acoust. Soc. Am.*, Volume 128, Issue 4, Oct 2010, pp 2162–2172.
5. **P. K. Ghosh** and S. Narayanan, “Joint source-filter optimization for robust glottal source estimation in the presence of shimmer and jitter”, *Speech Communication*, Elsevier, Volume 53, Issue 1, Jan 2011, pp 98–109.
6. **P. K. Ghosh**, A. Tsiartas, and S. Narayanan, “Robust voice activity detection using long-term signal variability”, *IEEE Trans. Audio, Speech and Language Processing*, Volume 19, Issue 3, March 2011, pp 600–613.
7. **P. K. Ghosh** and S. Narayanan, “Bark frequency transform using an arbitrary order allpass filter”, *IEEE Signal Processing Letters*, Volume 17, No. 6, June 2010, pp 543–546.
8. **P. K. Ghosh** and S. Narayanan, “Pitch contour stylization using an optimal piecewise polynomial approximation”, *IEEE Signal Processing Letters*, Volume 16, No. 9, September 2009, pp 810–813.
9. **P. K. Ghosh** and S. Narayanan, “Closure duration analysis of incomplete stop consonants due to stop-stop interaction”, *J. Acoust. Soc. Am. Express Letters*, Volume 126, Issue 1, July, 2009, pp. EL1–EL27.
10. **P. K. Ghosh** and T.V. Sreenivas, “Time-varying Filter Interpretation of Fourier Transform and its Variants”, *Signal Processing (Elsevier)*, Volume 86, Issue 11, November 2006, Pages 3258–3263.

Conferences:

1. **P. K. Ghosh** and S. Narayanan, “Analysis of inter-articulator correlation in acoustic-to-articulatory inversion using generalized smoothness criterion”, accepted in *Interspeech*, Florence, Italy, 2011.
2. **P. K. Ghosh** and S. Narayanan, “A subject-independent acoustic-to-articulatory inversion”, accepted in *ICASSP*, Prague, Czech Republic, 2011.

3. **P. K. Ghosh**, Andreas Tsiartas, Panayiotis G. Georgiou, and S. Narayanan, “Robust voice activity detection in stereo recording with crosstalk”, Proc. InterSpeech, 26–30 Sep, 2010, Makuhari, Japan.
4. **P. K. Ghosh**, S. Narayanan, Pierre Divenyi, Louis Goldstein, and Elliot Saltzman, “Estimation of articulatory gesture patterns from speech acoustics”, Proc. InterSpeech, 6-10 Sep, 2009, Brighton, UK, pp 2803–2806.
5. **P. K. Ghosh**, A. Ortega, and S. Narayanan, “Pitch Period Estimation using Multipulse Model and Wavelet Transform”, Proc. Interspeech 2007, Page(s): 2761–2764.
6. **P. K. Ghosh**, “Speech Segmentation using Extrema-Based Signal Track Length Measure”, ICASSP 2007, Volume 4, 15-20 April 2007 Page(s):IV-1065–IV-1068.
7. **P. K. Ghosh** and T.V. Sreenivas, “Dynamic Programming Based Optimum Non-Uniform Samples For Speech Reconstruction and Coding”, ICASSP 2006, Volume 1, Page(s):I–I.
8. **P. K. Ghosh** and T.V. Sreenivas, “Waveform Reconstruction from Non-uniform Samples with Application to Speech Coding”, IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing (NSIP), May 2005, Japan.
9. **P. K. Ghosh** and A. Konar, “Modification of the LMS Predictor to Reduce Signal Prediction Error in Linear Prediction”, International Conference on Communication, Devices and Intelligent Systems (CODIS 2004), Jan 2004, Kolkata, India.
10. **P. K. Ghosh** and T.V. Sreenivas, “Extrema based Unwarping for Time-varying Pitch Estimation”, accepted for publication in 12th National Conference on Communication (NCC) 2006.
11. S. Narayanan, E. Bresch, **P. K. Ghosh**, L. Goldstein, A. Katsamanis, Y. Kim, A. Lammert, M. Proctor, V. Ramanarayanan, and Y. Zhu, “A Multimodal Real-Time MRI Articulatory Corpus for Speech Research”, accepted in Interspeech, Florence, Italy, 2011.
12. B. Xiao, **P. K. Ghosh**, Panayiotis G. Georgiou, and S. Narayanan, “Overlapped speech detection using long-term spectro-temporal similarity in stereo recording”, accepted in ICASSP, Prague, Czech Republic, 2011.
13. A. Tsiartas, **P. K. Ghosh**, Panayiotis G. Georgiou, and S. Narayanan, “Bilingual audio-subtitle extraction using automatic segmentation of movie audio”, accepted in ICASSP, Prague, Czech Republic, 2011.
14. A. Tsiartas, **P. K. Ghosh**, and S. Narayanan, “Context-driven bilingual movie subtitle alignment”, Proc. InterSpeech, 6-10 Sep, 2009, Brighton, UK, pp 444–447.
15. A. Tsiartas, **P. K. Ghosh**, P. Georgiou, and S. Narayanan, “Robust word boundary detection in spontaneous speech using acoustic and lexical cues”, In Proceedings of ICASSP, Taipei, Taiwan, Apr 2009.
16. S. Ananthakrishnan, **P. K. Ghosh**, S. Narayanan, “Automatic classification of question turns in spontaneous speech using lexical and prosodic evidence”, In Proceedings of ICASSP, Las Vegas, Nevada, April 2008.
17. A. Das, M. Balwani, R. Thota, **P. K. Ghosh**, “Face Recognition from Images with High Pose Variations by Transform Vector Quantization”, ICVGIP 2006, Pages:674–685.
18. A. Das, **P. K. Ghosh**, “Audio-Visual Biometric Recognition by Vector Quantization”, IEEE Spoken Language Technology(SLT) Workshop, Dec 2006, Page(s):166–169.

- Talks/Posters**
1. International Seminar on Speech Production (ISSP), Montreal, Canada, June 2011.
 2. Ratheon BBN Technologies, Cambridge, Massachusetts, USA, May 2011.

3. CSAIL, Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, USA, May 2011.
4. IBM Watson Research Center, New York, USA, May 2011.
5. Haskins Laboratories, Yale University, New Haven, Connecticut, USA, May 2011.
6. Information Theory and Applications (ITA) Workshop, University of California, San Diego (UCSD), California, USA, Feb 2011.
7. Indian Statistical Institute (ISI), Kolkata, India, December, 2010.
8. EE Department Seminar, Indian Institute of Science (IISc), Bangalore, India, December, 2010.
9. Indian Institute of Technology, Madras (IITM), Chennai, India, December, 2010.
10. STCS Seminar, Tata Institute of Fundamental Research (TIFR), Mumbai, India, December, 2010.
11. Indian Institute of Technology, Kharagpur (IITKGP), West Bengal, India, December, 2010.
12. Acoustical Society of America (ASA) Meeting, Cancun, Mexico, November, 2010.
13. Information Theory School (ITS), University of Southern California, Los Angeles, USA, August, 2010.

Teaching Experiences

Teaching Assistant (TA), USC

- Introduction to Probability and Statistics for Electrical Engineering (Fall 2006)
- Probability Theory for Engineers (Spring 2007)
- Random Processes in Engineering (Fall 2007 and Spring 2008)

Professional Experiences

Researcher Oct 2011 – present
IBM India Research Lab, Delhi, India.

Research Intern March – July 2006
Microsoft Research India, Bangalore, India.
 Worked on *Speaker Identification and Audio-visual Biometric Recognition*.

Software Engineer July – December 2003
Ushacomm India Pvt. Ltd., Kolkata, India.

Summer Intern June – July 2002
Embedded DSP Group, TATA Consultancy Services, Kolkata, India.
 Worked on “*Study and Simulation of Orthogonal Frequency Division Multiplexing (OFDM)*”.

Professional Activities

- Reviewer for
1. IEEE Trans. Audio, Speech, and Language Processing (TASLP),
 2. IEEE Signal Processing Letters (SPL),
 3. Journal of the Acoustical Society of America (JASA),
 4. Speech Communication, Elsevier,
 5. EURASIP Journal on Audio, Speech, and Music Processing,
 6. ISA Transactions, ELSEVIER,
 7. ICASSP, INTERSPEECH, SPCOM, NCC.

Skills

1. Software Packages: Matlab, LATEX, SPSS.
2. Programming Languages: C, C++, Perl.

- Membership**
1. ASA member,
 2. IEEE member,
 3. ISCA member.