KAILASH PATIL
Center for Language and Speech Processing,
Johns Hopkins University, Baltimore, MD
kailash@jhu.edu, www.clsp.jhu.edu/people/kailash

CURRENT POSITION

Ph.D. Candidate, Department of Electrical and Computer Engineering, Sept. 2008 – present
Johns Hopkins University, Baltimore, MD

- Currently pursuing my Ph.D. under Dr. Mounya Elhilali, at the Center for Language and Speech Processing.

FIELDS OF INTERST

Auditory Scene Analysis, Speech Processing, Automated Speaker/Speech Recognition, Machine Learning, Acoustic Modeling, Large Vocabulary Continuous Speech Recognition.

PH.D. THESIS TOPIC

Neuro-Computational Basis of Sound Object Recognition

- Developed novel feature extraction methods which are extremely robust to noise for both speech and speaker identification tasks.
  - These features capture speech specific regions in the modulation domain maximizing reliability.
  - Multi-stream approach to further divide this region into subparts performs better in various noise conditions.
- Successfully demonstrated models for timbre which capture the perceptual space of musical instruments.
  - Attentional mechanisms in this perceptual space have been developed which can further boost the representation of any given target object.
  - Developed methods to adapt feature extraction and modeling stages to out-of-domain data.

RESEARCH EXPERIENCE

Research Assistant, Center for Language and Speech Processing Sept. 2008 - present
Johns Hopkins University, Baltimore, MD

- **Derived-STRF(Spectro-Temporal Receptive Field) contours for Speech recognition**
  - Developed an algorithm to learn STRFs from speech data to give sustained response
  - Successfully used the resulting contour to derive robust features which show improved performance in noisy conditions
- **Phoneme recognition framework using STRFs**
  - Developed a mechanism to automatically select STRF features for each broad phoneme class
  - Combined posteriors from multi-layered perceptrons trained on these features give improved performance.
- **Multi-resolution Analysis for Lung sounds**
  - Successfully extracted multidimensional features from lung sounds that were able to predict presence of abnormalities.
- **Speech based filter banks**
  - Derived filter banks from average spectrum of speech that were compared with perception based mel-like filter banks.

Graduate Researcher, CLSP Summer Workshop June - July 2009
Johns Hopkins University, Baltimore, MD
Project Guide: Dr. Dekang Lin, Google

- Collaborated as a part of the N-gram team on extracting similar phrases based on a context.
- Worked in a Hadoop environment to parallelize algorithms for Google N-Gram Data
- Developed a web interface for users to search similar phrases.

Graduate Researcher, Telluride Neuromorphic Cognition Engineering Workshop, July 2011
Telluride, CO
Project Guides: Dr. Mounya Elhilali, JHU; Dr. Malcom Slaney, Microsoft

- Collaborated as a team with the Attention Driven Scene Analysis group to understand and model the role of attention in scene understanding.
- Developed a new technique to apply a top-down attention weight on the features by back-propagation in multi-layered perceptron.
- Achieved improved performance in recognizing the target phoneme class.

Intern, Indian Institute of Science May – June 2006
Bangalore, India
Project Guides: Prof. A.G.Ramakrishnan, IISc, Bangalore; Dr. S.R.M Prasanna , IIT Guwahati.

- Implemented a text independent speaker identification system using MFCC features and VQ & GMM for speaker modeling

Undergraduate Researcher, Indian Institute of Technology, Augt. 2007 – April 2008
Guwahati, India
Advisor: Dr. S. R. M Prasanna, IIT Guwahati.
Topic: Exploration of New Feature extraction methods for speaker recognition

- Demonstrated usefulness of amplitude modulation features and low level acoustic features in encoding speaker information and showed the complementarity to MFCC features.

Teaching Assistant, Department of Electrical and Computer Engineering,
Johns Hopkins University, Baltimore, MD

- Taught classes, graded coursework for "Introduction to Speech and Audio Processing" during Fall 2011 and Fall 2012.
- Taught classes, graded coursework for "Signals & Systems I" during Spring 2012.

Intern, ENST-Bretagne, May – June 2007
Rennes Campus, France
Project Guide: Prof. Laurent Toutain**,** ENST-Bretagne, France.

- Worked in a team on the DisPairSe project.
- Worked on making a NAS device IPv6 ready so that it could be used in p2p protocols
- Helped build a p2p protocol for such devices using freepastry.

SKILLS

- **Languages:** C, C++, Java, UNIX shell scripting
- **Environments:** Windows, Linux, Unix
- **Software:** MATLAB, Quicknet Neural Network Toolbox, HTK, Kaldi, Open Fst, SRI Language Modeling Toolkit.
- **IDE:** Eclipse
- Microsoft Office, Excel, PowerPoint, LaTeX

EDUCATION
Master of Science in Engineering, Department of Electrical and Computer Engineering May 2011
Johns Hopkins University, Baltimore, MD

Bachelor of Technology, Department of Electronics and Communication Engineering July 2008
Indian Institute of Technology, Guwahati, India
Departmental Rank: 2
CPI(Cumulative Point Index): 9.13/10

REFEREED JOURNAL AND CONFERENCE PAPERS

- K. Patil, M. Elhilali, "Top-down attentional mechanisms for improving scene recognition".(Under Preparation)
- K. Patil, M. Elhilali, "Robust timbre recognition framework for cross-database testing".(Under Preparation)
- S. K. Nemala, K. Patil, M. Elhilali, "Robust speech recognition by humans and machines: A role of the modulation transfer function", *Computer Speech and Language*.(Under Review)
- K.Patil, M. Elhilali, "Task Driven Attentional Mechanisms for Auditory Scene Recognition", *Proceedings of 38th International Conference on Acoustics, Speech, and Signal Processing,* Vancouver, Canada, May 2013.
- S. K. Nemala, K. Patil, M. Elhilali, "A multistream framework based on modulation filtering for robust speech recognition", *Transactions on Audio Speech and Language Processing*, pp. 416-426, Vol. 21(2) Feb 2013.
- S. K. Nemala, K. Patil, M. Elhilali "Recognizing the message and the messenger: Biomimetic Spectral Analysis for Robust Speech and Speaker Recognition", *International Journal of Speech Technology,* pp. 1-10, Dec 2012.
- K. Patil, D. Pressnitzer, S. Shamma, M. Elhilali, "Music in our ears: The biological bases of musical timbre perception", *PLOS Computational Biology*, Vol. 8(11), Nov 2012.
- K.Patil, M. Elhilali, "Goal Oriented Auditory Scene Recognition", *Proceedings of 13th Conference of International Speech Communication Association*, Portland, USA, Sep 2012.
- M. Carlin, K. Patil, S. K. Nemala, M. Elhilali, "Robust phoneme recognition based on biomimetic speech contours", *Proceedings of 13th Conference of International Speech Communication Association*, Portland, USA, Sep 2012.
- D. Emmanouilidou, K. Patil, J. West, M. Elhilali, "A multiresolution analysis for detection of abnormal lung sounds", *Proceeding of 34th Conference of the IEEE Engineering in Medicine and Biology Society*, San Diego, USA, 2012.
- S. K. Nemala, K. Patil, M. Elhilali, "Multistream bandpass modulation features for robust speech recognition", *Proceedings of 12th Conference of International Speech Communication Association*, Florence, Italy, Aug 2011.
- S. Thomas, K.Patil, S. Ganapathy, N. Mesgarani, H. Hermansky, "A Phoneme Recognition Framework based on Auditory Spectro-Temporal Receptive Fields", *Proceedings of 11th Conference of International Speech Communication Association*, Makuhari, Japan, Sept 2010.
- D. Lin, K. Church, H. Ji, S. Sekine,D. Yarowsky, S. Bergsma, K. Patil, E. Pitler, R. Lathbury, V. Rao, K. Dalwani, S. Narsale, "New tools for web-scale N-grams", *Proceedings of 7th International Conference on Language Resources and Evaluation*, Valletta, Malta, May 2010.
- K. Patil, H. S. Dhillon and A. Mitra, "A Telephone Based Wireless Remote Controller for Home Appliances", *Proceedings of 14th National Conference on Communications* (NCC-2008), Bombay, Feb. 2008.

PRESENTATIONS AND ABSTRACTS

- K.Patil, M. Elhilali, "Attentional Mechanisms for Recognizing Acoustic Scenes", *36th Annual MidWinter Meeting*, ARO, Baltimore, MD, USA, 2013.
- K.Patil, M. Elhilali, "Neuro-Computational Basis of Sound Object Recognition", *Thesis Proposal at Department of Electrical and Computer Engineering*, Johns Hopkins University, Baltimore, USA, March 2012.
- K. Patil, "Goal-Oriented Auditory Scene Recognition", *Center for Language and Speech Processing*, Baltimore, USA, April 2011.
- K. Patil, M. Elhilali, M. Slaney, "High Level Saliency and Features", *Telluride Neuromorphic Cognition Engineering Workshop 2011,* Telluride, USA, July, 2011.
- K. Patil, "Robust Recognition of the message and the messenger in Speech", *Center for Language and Speech Processing*, Baltimore, USA, Oct 2010.
- K. Patil, M. Elhilali, "A biomimetic multi-resolution spectrotemporal model for musical timbre recognition", *Joint 159th ASA Meeting and Noise-Con 2010*, Baltimore, USA, Mar 2010.
- K. Patil, V. Rao, D. Lin, "Distributional Features and Clustering for Ngrams", *Center for Language and Speech Processing Summer Workshop 2009,* Baltimore, USA, July 2009.

PATENTS

- M. Elhilali, S. K. Nemala, K. Patil, "Noise robust multistream speech signal analysis/encoding".(Pending)
- M. Elhilali, S. K. Nemala, K. Patil, "Multi-resolution cortical features as front-end for robust speech and speaker recognition" .(Pending)

HONORS/AWARDS

- Recipient of the ICSA student grant for Interspeech, Portland, 2012.
- Awarded Merit Scholarship at IIT Guwahati for the period 2005-2006 for being the topper in Electrical and Computer Engineering department.
- Placed in the National top 1 percent in Indian Physics Olympiad in the year 2003-4.
- Recipient of the National Talent Search Exam Scholarship by NCERT

References will be provided on request.