

Dr. Premkumar (Prem) Natarajan

47 Hickory Road
Sudbury, MA 01776

Cell: (617) 512-4359
pnataraj@gmail.com

EXECUTIVE VICE PRESIDENT, Raytheon BBN Technologies

At BBN, my salient contributions include creating several new lines of research and technology practice; developing and nurturing relationships with new domestic and international customer and partner organizations; and establishing an extensive network of partnerships with top University research groups. Managing innovation and interacting with clients is an important part of my daily life at BBN. I routinely initiate and direct large, interdisciplinary and multi-organization research and development efforts; negotiate partnerships and commercial deals; and develop and execute integrated technology and business development strategy.

My work has allowed me to gain a deep understanding of US Government research and technology procurements and extensive experience in operating within Government and Commercial markets. I work closely with a diverse set of customers including DARPA, IARPA, Army, the Intelligence Community, commercial organizations, and foreign governments. I hold a TS/SCI clearance.

In recent years, operationalizing our technological and IP assets in speech, language and multimedia through the development of a portfolio of products and solutions, and leveraging those assets to enter new interdisciplinary research areas – such as predictive analytics and human social cultural behavior modeling (HSCB) – has been an important emphasis of my work. I continue to be active in my areas of technical interest: document image processing, handwriting recognition, computer vision, machine learning, speech recognition, speech-to-speech translation, topic classification, and multimedia processing. Finally, I contribute to the greater professional community by organizing conferences and workshops, serving on program committees, contributing to papers and book chapters, and serving as guest editor.

PROFESSIONAL EXPERIENCE

Raytheon BBN Technologies – Cambridge, MA **1996 – Present**
Executive Vice President, 2011 – Present; Vice President, 2006 – 2011
Manager of Speech, Language and Multimedia Business Unit, 2006 – Present
Elected Principal Scientist, 2010

- Serve as a member of BBN's executive leadership team.
- Develop and implement long-term technical, business, and operational strategies – focused on sustainable innovation and growth – which consolidate BBN's market leadership in its areas of operations.
- Provide long-term technical and business vision, leadership in program execution and operational management of the Speech, Language and Multimedia Business Unit; driving growth momentum by leading strategic business development and technology thrusts; and achieving short and long-term goals, while meeting or exceeding budget objectives.
- Result: Business Unit has grown by a factor of two in revenue during 2008-2012 with a corresponding increase in staff size.

Selected Accomplishments:

- In 2006, led the development of a 5-year growth plan for the Business Unit with minimum and stretch targets. In 2010, we exceeded the *stretch* target.
- Developed scalable organizational structure to support and manage rapid growth. Identified key new positions and recruited the talent to fill the positions.
- Participated in development and refinement of corporate business strategy as a member of BBN's Business Management Council. Secured buy-in for strategically significant initiatives and pursuits. Established valuable Industrial and University partnerships.
- Created a new deployment-line-of-business with a hybrid business model combining product licensing revenue with professional services. Deployments are now an important driver of growth in operating income for the Business Unit.
- Conceptualized, pursued, and captured first-ever DARPA program in document image processing and recognition. Served as capture manager during every phase, from conceptualizing the technical research area, pitching the idea to DARPA, obtaining seedling funding, then leading a community effort to define the value of the proposed research, and finally winning a 5-year \$35+ million R&D program.
- Established new Computer Vision research area at BBN. Identified and recruited key talent, developed University partnerships, and captured beachhead projects including IARPA ALADDIN.
- Shaped and led effort to capture a multi-million dollar, multi-year open-source monitoring technology deployment contract with a foreign government.
- Established human social-cultural behavior research and applications area at BBN, attracting several million dollars of Government research funding within the first two years of operation. Developed strong relationship with clients and helped form a new partnership between the research sponsor and an important Government user group.
- Led successful capture of BBN's first predictive analytics program from IARPA under IARPA's OSI program.
- Led the growth of speech-to-speech translation business to cumulative revenues of more than \$55M since inception.
- Shaped and led BBN's successful capture effort for the Army MFLTS (Machine Foreign Language Translation System) Program of Record.
- Conceptualized a commercial, hosted, speech recognition SaaS (Software-as-a-Service) offering. Negotiated subscription deal with a commercial transcription start-up. Within 12 months, the business model has been validated and a growing and profitable revenue stream has been established.

Deputy Manager, Speech and Language Processing Department, 2004 – 2006**Elected Lead Scientist, 2006**

Led key programs and projects. Developed technical and business strategy for optical character recognition (OCR) and speech-to-speech translation business areas. Participated in day-to-day management of overall department.

Selected Accomplishments:

- Proposed an integrated research-to-solutions operations model as a framework for strategic planning and investments.

- Led the effort to develop laptop and PDA-based speech-to-speech translation technology. Within 18 months, BBN progressed from a new entrant in the field to being the top-ranked system in DARPA Speech-to-Speech Translation evaluations. In 2010, BBN's was the *only* system selected by the Government for operational field testing (on the Android smartphone platform).
- Expanded OCR capability to include new domains (handwritten data, video data), new platforms, and new customers. Principal author of BBN's winning IARPA VACE Phase II and Phase III proposals for OCR of videotext resulting in total funding of \$6M. The VACE project initiated video processing (i.e., computer vision) work at BBN.
- Initiated pursuit with an important foreign government. The relationship with that government has since grown to several million dollars of revenue.

Manager, Speech Engineering, and Senior Scientist, 2001-2004

Technical lead for the advancement and deployment of several large-volume Directory Assistance Automation and Call Center Solutions. Managed day-to-day interaction with telecom client organization and represented BBN at client site over extended periods. Served as technical lead for the OCR development effort. Participated in business development activities focused on commercial and US Government clients. Played a key role in capturing and executing BBN's first speech-to-speech translation research effort.

Selected Accomplishments:

- Led the speech engineering team on the first large-scale Directory Assistance Automation effort in North America for a major telecom customer.
- Initiated, obtained customer funding, and served as project and technical lead for porting BBN's Byblos OCR system to the Windows environment.
- Contributed to and led winning proposal development efforts for innovative research and engineering work from several US Government customers including BBN's first speech-to-speech translation effort.

Scientist, BBN, 1999-2001

Staff Scientist, BBN, 1996-1999

Performed research in OCR of Arabic using hidden Markov model (HMM) technology – the first successful use of HMMs for script-independent OCR. This effort formed the cornerstone of a series of projects in OCR and document processing that continue to this day. Made several innovative technical contributions to core recognition technology and was a key developer of BBN's first prototype client-server OCR system that accepted scanned images and produced text output. Individually responsible for a 100x speed-up of the OCR engine with marginal loss of accuracy – the speed-up made BBN's OCR engine usable in operational systems.

EDUCATION

Graduate of Greater Boston Executive Program, Sloan School of Management, MIT.
 Ph.D. (Electrical Engineering), Tufts University.
 M.S. (Electrical Engineering), Tufts University.
 B.E. (Electrical Engineering), University of Poona, India, (*with Distinction*).

SOCIETY MEMBERSHIPS AND PROFESSIONAL ACTIVITIES

Member, *Advisory Board, Computer Science Department, Lehigh University*, 2009-present
General Co-Chair, *International Conference on Document Analysis and Recognition (ICDAR)*, Washington DC, 2013
Guest Editor (with Shrikanth Narayanan, Panayiotis Georgiou, Tanya Schultz, and Alan Black), *Special Issue on Speech-to-Speech Translation*, Computer Speech and Language (CSL), 2013
Program Co-Chair, *International Workshop on Document Analysis Systems*, June 2010;
Multilingual OCR Workshop at ICDAR Barcelona, July 2009; *Multilingual OCR Workshop at ICDAR Beijing* 2011
Area Chair for *Spoken Language Systems and Applications, NAACL-HLT*, Denver, May 2009
Program Committee Member, ICDAR 2007; ICDAR 2009; ICDAR 2011
BBN Representative at the *ISCA-Industry Roundtable* held in conjunction with InterSpeech, the international conference on spoken language processing, 2006-Present
Program Committee, *Workshop on Spoken Language Understanding for Conversational Systems*, 2004
Senior Member, *Institute of Electrical and Electronics Engineers*
Member, *Technical Committee on Statistical Techniques in Pattern Recognition, IAPR*
Tau Beta Pi; Eta Kappa Nu

HONORS

A Best Paper Award for: S. Ananthkrishnan, R. Prasad, and **Prem Natarajan**, “Phrase Alignment Confidence for Statistical Machine Translation,” *Proceedings of Interspeech*, pp. 2878-2881, Makuhari, Japan, September 2010.

“Special technology achievement award in recognition of excellence and outstanding contribution,” in recognition of the highly successful deployment of Automated Directory Automation at a major telecom company, from CTO of Verizon in 2002 (BBN was part of Verizon at the time).

Scholarship from J. N. Tata Endowment for Higher Studies of Indians Abroad, India.

National Merit Scholar, 1988-1992, Govt. of India.

PATENTS

“Multi-frame videotext recognition,” *United States Patent 8,290,273*, October 2012.

“Method and Apparatus for Training an Automated Speech Recognition-based System,” *United States Patent 7,346,507*, March 2008.

“Unsupervised Training in Natural Language Call Routing,” *United States Patent 7,092,888*, September 2006.

“Blind adaptive equalization using cost function that measures dissimilarity between the probability distributions of source and equalized signals,” *United States Patent 6,049,574*, 2000.

“Segmental Rescoring in Text Recognition,” Filed April 2009.

“Unsupervised Categorization of Videos in the Wild,” in preparation.

SALIENT INVITED TALKS

Invited Panelist and Speaker at Translation Automation User Society (TAUS), Seattle, Washington, October 2012 (“*Human Language Project Panel*” and “*Speech-to-speech Translation*”)

Invited Speaker at Translation Automation User Society (TAUS), Portland, Oregon, October 2010 (“*Translation Lessons from the Intelligence Space*”)

Keynote Speaker at International Workshop on Spoken Language Technology, Tokyo, Japan, December 2009 (“*Robust Speech-to-Speech Translation*”)

“HMMs for Speech, Language, and OCR,” Invited talk for Distinguished Lecture Series at Charles River Analytics (CRA, Inc.), November 2003, Cambridge, Massachusetts.

“Examining the Potential of Voice Solutions for Implementing Automated Directory Assistance,” Invited Speaker, *Using Voice-Enabled Wireless and Multimodal Applications to Drive Business Revenues*, May 12 – 14 2003, Chicago, Illinois.

“Directory Assistance Automation,” Invited Speaker/Panelist, *SpeechTek 2002*, New York, NY.

MEDIA MENTIONS

Fox News (online edition), September 2012 (in “*Military Tech Can Monitor World Media at the Press of a Button*”)

MIT Technology Review magazine, August 2012 (in “*Securing Your Voice*”)

MIT Technology Review magazine, June 2012 (in “*Wiping Away Your Siri Fingerprint*”)

Fox News (online edition), May 2012 (in “*Star Trek Universal Translator to Join War on Terror*”)

News.cnet.com, July 2010 (in “*Working to Make Combat Soldiers’ Lives Easier*”)

The Christian Science Monitor, June 2010 (in “*Technology that Translates, and Unites*”)

Medical News Today, September 2009 (in “*English/Spanish Speech-To-Speech Translation System to Help In Health Care Situations, Funded By NSF*”)

Government Computer News, June 2009 (in “*DARPA Project Seeks to Teach Machines Better Reading Skills*”)

Washington Post, September 2008 (in “*Even Spies Go to Trade Conferences*”)

Federal Computer Week, June 2008 (in “*A New Life for Talk-to-Text*”)

MIT Technology Review magazine, November 2007 (in “*Searching Video Lectures*”)

Tactical Weapons Magazine, September 2007

Defense Technology Magazine, July 2007

MIT Technology Review magazine, March 2004 (in “*Translation in the Age of Terror*”)

PUBLICATIONS

- E. Boschee, **Prem Natarajan**, R. Weischedel, “Automatic Extraction of Events from Open Source Text for Predictive Forecasting,” *Handbook of Computational Approaches to Counterterrorism*, Springer, December 2012 (in publication).
- H. Cao, K. Subramanian, X. Peng, J. Chen, R. Prasad, and **Prem Natarajan**, “Extracting Information from Handwritten Content in Census Forms,” *21st International Conference on Pattern Recognition*, Tsukuba, Japan, November 2012.
- H. Cao, J. Chen, J. Devlin, R. Prasad, and **Prem Natarajan**, “Document Recognition and Translation System for Unconstrained Arabic Documents,” *21st International Conference on Pattern Recognition*, Tsukuba, Japan, November 2012.
- S. Vitaladevuni, F. Choi, R. Prasad, and **Prem Natarajan**, “Detecting Near-Duplicate Document Images Using Interest Point Matching,” *21st International Conference on Pattern Recognition*, Tsukuba, Japan, November 2012.
- R. Kumar, R. Prasad, S. Ananthakrishnan, A. N. Vembu, D. Stallard, S. Tsakalidis, **Prem Natarajan**, “Detecting OOV Named-Entities in Conversational Speech,” *InterSpeech 2012*, Portland, September 2012.
- S. Tsakalidis, X. Zhuang, R. Hsiao, S. Wu, P. Natarajan, R. Prasad, and **Prem Natarajan**, “Robust Event Detection From Spoken Content In Consumer Domain Videos,” *InterSpeech 2012*, Portland, September 2012.
- X. Zhuang, S. Tsakalidis, S. Wu, P. Natarajan, R. Prasad, and **Prem Natarajan**, “Compact Audio Representation for Event Detection in Consumer Media,” *InterSpeech 2012*, Portland, September 2012.
- P. Natarajan, S. Wu, S. Vitaladevuni, X. Zhuang, S. Tsakalidis, U. Park, R. Prasad, and Prem Natarajan, “Multimodal Feature Fusion for Robust Event Detection in Web Videos,” *Computer Vision and Pattern Recognition (CVPR) 2012*, Providence, June 2012.
- L. Kang, D. Doermann, H. Cao, R. Prasad and **Prem Natarajan**, “Local Segmentation of Touching Characters using Contour based Shape Decomposition,” *10th IAPR International Workshop on Document Analysis Systems*, Gold Coast, Australia, March 2012.
- X. Peng, H. Cao, K. Subramanian, E. Boschee, R. Prasad, and **Prem Natarajan** “Information Extraction from Historical Semi-Structured Handwritten Documents,” *Family History Technology Workshop, RootsTech*, February 2012.
- V. Manohar, S. Tsakalidis, P. Natarajan, R. Prasad, and **Prem Natarajan**, “Audio-Visual Fusion Using Bayesian Model Combination for Web Video Retrieval,” *ACM Multimedia 2011*, Scottsdale, December 2011.
- J. Kumar, R. Prasad, H. Cao, W. Abd-Almageed, D. S. Doermann, and **Prem Natarajan**, “Shape Codebook based Handwritten and Machine Printed Text Zone Extraction,” *Proceedings of Document Recognition and Retrieval (DRR) XVIII*, pp. 1-10, San Francisco, December 2011.
- Pradeep Natarajan, **Prem Natarajan**, V. Manohar, S. Wu, S. Tsakalidis, S. N. Vitaladevuni, X. Zhuang, R. Prasad, G. Ye, David Liu, I-H Jhuo, S. Chang, H. Izadinia, I. Saleemi, M. Shah, B. White, T. Yeh, and L. Davis, “BBN VISER TRECVID 2011 Multimedia Event Detection System,” *TRECVID 2011 Multimedia Event Detection Workshop (MED’11)*, November 2011.
- S. Ananthakrishnan, S. Vitaladevuni, R. Prasad, and **Prem Natarajan**, “Source Error-Projection for Sample Selection in Phrase-Based SMT for Resource-Poor Languages,” *Proceedings of IJCNLP*, pp. 819-827, Chiang Mai, Thailand, November 2011.

- S. N. Vitaladevuni, Pradeep Natarajan, R. Prasad and **Prem Natarajan**, "Efficient Orthogonal Matching Pursuit using Sparse Random Projections for Scene and Video Classification," *Proceedings of IEEE International Conference on Computer Vision (ICCV)*, Barcelona, Spain, November 2011.
- R. Prasad, **Prem Natarajan**, D. Stallard, S. Saleem, S. Ananthakrishnan, S. Tsakalidis, C. Kao, F. Choi, R. Meermeier, M. Rawls, J. Devlin, K. Krstovski, and A. Challenner, "BBN TransTalk: Robust Multilingual Two-Way Speech-to-Speech Translation for Mobile Platforms," *Computer Speech and Language: Special Issue on Speech-to-Speech Translation*, November 2011.
- S. Ananthakrishnan, R. Prasad, D. Stallard, and **Prem Natarajan**, "Batch-Mode Semi-Supervised Active Learning for Statistical Machine Translation," *Computer Speech and Language: Special Issue on Speech-to-Speech Translation*, November 2011.
- X. Peng, H. Cao, R. Prasad, and **Prem Natarajan**, "Text Extraction from Video Using Conditional Random Fields," *Proceedings of the Eleventh International Conference on Document Analysis and Recognition*, pp. 1029-1033, Beijing, China, September 2011.
- Pradeep Natarajan, D. Belanger, R. Prasad, M. Kamali, K. Subramanian, and **Prem Natarajan**, "Baseline Dependent Percentile Features for Offline Arabic Handwriting Recognition," *Proceedings of the Eleventh International Conference on Document Analysis and Recognition*, pp. 329-333, Beijing, China, September 2011.
- V. Manohar, S. Vitaladevuni, H. Cao, R. Prasad, and **Prem Natarajan**, "Graph Clustering-Based Ensemble Method for Handwritten Text Line Segmentation," *Proceedings of the Eleventh International Conference on Document Analysis and Recognition*, pp. 574-578, Beijing, China, September 2011.
- H. Cao, R. Prasad, and **Prem Natarajan**, "OCR-Driven Writer Identification and Adaptation in an HMM Handwriting Recognition System," *Proceedings of the Eleventh International Conference on Document Analysis and Recognition*, pp. 739-743, Beijing, China, September 2011.
- H. Cao, R. Prasad, and **Prem Natarajan**, "Handwritten and Typewritten Text Identification and Recognition Using Hidden Markov Models," *Proceedings of the Eleventh International Conference on Document Analysis and Recognition*, pp. 744-748, Beijing, China, September 2011.
- S. N. Vitaladevuni, H. Cao, D. Belanger, K. Subramanian, R. Prasad and **Prem Natarajan**, "Sparsity-based super-resolution for offline handwriting recognition," *Proceedings of Joint Workshop on Multilingual OCR and Analytics for Noisy Unstructured Text Data (J-MOCR-AND)*, pp. 1-5, Beijing, September 2011.
- X. Peng, H. Cao, K. Subramanian, R. Prasad, and **Prem Natarajan**, "Automated Image Quality Assessment for Camera-Captured OCR," *Proceedings of the International Conference on Image Processing*, pp. 2621-2624, Brussels, Belgium, September 2011.
- Pradeep Natarajan, S. Tsakalidis, V. Manohar, R. Prasad, and **Prem Natarajan**, "Unsupervised Audio Analysis for Categorizing Heterogeneous Consumer Domain Videos," *InterSpeech*, pp. 313-316, Florence, Italy, August 2011.
- S. Ananthakrishnan, S. Tsakalidis, R. Prasad, and **Prem Natarajan**, "On-line Language Model Biasing for Multi-Pass Automatic Speech Recognition," *Proceedings of Interspeech*, pp. 621-624, Florence, Italy, August 2011.
- S. Ananthakrishnan, R. Prasad, and **Prem Natarajan**, "On-line Language Model Biasing for Statistical Machine Translation," *Proceedings of ACL-HLT*, pp. 445-449, Portland (OR), U.S.A., June 2011.
- V. Manohar, S. Tsakalidis, Pradeep Natarajan, R. Prasad, and **Prem Natarajan**, "Audio-Visual Fusion using Bayesian Model Combination for Web Video Retrieval," *Proceedings of the 19th ACM international conference on Multimedia (MM '11)*, New York, USA, 1537-1540, 2011.

- Pradeep Natarajan, Y. Wu, R. Prasad, S. Saleem, E. MacRostie, F. Bernardin, and **Prem Natarajan**, “Large-scale, Real-time Logo Recognition in Broadcast Videos,” *Proceedings of International Conference on Multimedia and Expo (ICME)*, pp. 1-4, Barcelona, Spain, July 2011.
- Y. Wu, **P. Natarajan**, J. Noonan, R. Prasad and **Prem Natarajan**, “Wavelet Band-Pass Filters for Matching Multiple Templates in Real Time,” *Proceedings of British Machine Vision Conference (BMVC)*, pp. 1-12, Dundee, U.K., August 2011.
- K. Subramanian, R. Prasad, and **Prem Natarajan**, “Robust named entity detection from optical character recognition output,” *International Journal on Document Analysis and Recognition*, vol. 14, no.2, pp. 189-200, 2011.
- E. Boschee, **Prem Natarajan**, R. Weischedel, “An Evaluation of Alternative Approaches for Identifying, Coding, and Forecasting Events Data,” *Proceedings of the 52nd Annual Convention of the International Studies Association*, Montreal, Canada, March 2011.
- S. Ananthakrishnan, R. Prasad, and **Prem Natarajan**, “An Unsupervised Boosting Technique for Refining Word Alignment,” *Proceedings of the IEEE Workshop on Spoken Language Technology (SLT)*, Berkeley, CA, December 2010.
- H.\. Cao, R. Prasad, and **Premkumar Natarajan**, “Improvements in HMM Adaptation for Handwriting Recognition Using Writer Identification and Duration Adaptation,” *Proceedings of the International Conference on Frontiers in Handwriting Recognition (ICFHR)*, Kolkata, India, 2010.
- S. Ananthakrishnan, R. Prasad, D. Stallard, and **Prem Natarajan**, “Discriminative Sample Selection for Statistical Machine Translation,” *Proceedings of EMNLP: Conference on Empirical Methods in Natural Language Processing*, pp. 626-635, Cambridge, MA, October 2010.
- M. Raugus, V. Sridhar, R. Prasad, and **Prem Natarajan**, Multi-Resolution Discriminative Models for Subvocalic Speech Recognition, *Proceedings of Interspeech*, pp. 2878-2881, Makuhari, Japan, September 2010.
- V. Sridhar, R. Prasad, **Prem Natarajan**, “Mutual Information Analysis for Feature and Sensor Subset Selection in Surface Electromyography based Speech Recognition,” *Proceedings of Interspeech*, pp. 2878-2881, Makuhari, Japan, September 2010.
- S. Ananthakrishnan, R. Prasad, and **Prem Natarajan**, “Phrase Alignment Confidence for Statistical Machine Translation,” *Proceedings of Interspeech*, pp. 2878-2881, Makuhari, Japan, September 2010
[Received a best paper award]
- R. Prasad, A. Bhardwaj, K. Subramanian, H. Cao, and **Prem Natarajan**, “Stochastic Segment Model Adaptation for Offline Handwriting Recognition,” *Proceedings of the Twentieth International Conference on Pattern Recognition (ICPR)*, pp. 1993-1996, Istanbul, Turkey, 2010.
- R. Prasad, M. Kamali, D. Belanger, A-V. I. Rosti, S. Matsoukas, and **Prem Natarajan**, “Consensus Network Based Hypotheses Combination for Arabic Offline Handwriting Recognition,” *Proceedings of the Twentieth International Conference on Pattern Recognition (ICPR)*, pp. 2861-2864, Istanbul, Turkey, 2010.
- S. Ananthakrishnan, R. Prasad, D. Stallard, and **Prem Natarajan**, “A Semi-Supervised Batch-Mode Active Learning Strategy for Improved Statistical Machine Translation,” *Proceedings of the 14th Conference on Computational Natural Language Learning (CoNLL)*, pp. 126-134, Uppsala, Sweden, July 2010
- J. Chen, H. Cao, R. Prasad, A. Bhardwaj, and **Prem Natarajan**, “Gabor Features for Offline Arabic Handwriting Recognition,” *Proceedings of the International Workshop on Document Analysis Systems (DAS)*, pp. 53-58, Boston, MA, 2010.

R. Prasad, S. Tsakalidis, I. Bulyko, C. Kao, and **P. Natarajan**, "Pashto Speech Recognition with Limited Pronunciation Lexicon," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, USA, March 2010.

D. Stallard, R. Prasad, S. Ananthkrishnan, F. Choi, S. Saleem, and **P. Natarajan**, "Evaluation of Different Confirmation Strategies for Speech-to-Speech Translation Systems," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Dallas, USA, March 2010.

S. Ananthkrishnan, R. Prasad, and **P. Natarajan**, "Alignment Entropy as an Automated Predictor of Bibtex Fidelity for Statistical Machine Translation," *7th International Conference on Natural Language Processing (ICON'09)*, Hyderabad, India, December 2009.

H. Cao, R. Prasad, **P. Natarajan**, and V. Govindaraju, "Nested State Indexing in Pairwise Markov Networks for Fast Handwritten Document Image Rule-line Removal," *IEEE Intl. Conference on Image Processing*, Egypt, November 2009.

D. Stallard, R. Prasad, and **P. Natarajan**, "Development and Internal Evaluation of Speech-to-Speech Translation Technology at BBN," *NIST Performance Metrics Intelligent Systems Workshop*, Gaithersburg, MD, September 2009.

K. Subramanian, R. Prasad, and **P. Natarajan**, "Robust Named Entity Detection using an Arabic Offline Handwriting Recognition System," *Third Workshop on Analysis of Noisy Unstructured Text Data (AND)*, Barcelona, Spain, July 2009.

H. Cao, R. Prasad, and **P. Natarajan**, "A Stroke Regeneration Method for Cleaning Rule-lines in Handwritten Document Images," *First International Workshop on Multilingual OCR*, Barcelona, Spain, July 2009.

P. Natarajan, K. Subramanian, A. Bhardwaj, and R. Prasad, "Stochastic Segment Modeling for Offline Handwriting Recognition," *International Conference on Document Analysis and Recognition (ICDAR)*, Barcelona, Spain, July 2009.

S. Saleem, H. Cao, K. Subramanian, M. Kamali, R. Prasad, and **P. Natarajan**, "Improvements in BBN's HMM-based Offline Handwriting Recognition System," *ICDAR*, Barcelona, Spain, July 2009.

H. Cao, R. Prasad, S. Saleem, **P. Natarajan**, "Unsupervised HMM Adaptation using Page Style Clustering," *ICDAR*, Barcelona, Spain, July 2009.

S. Tsakalidis, R. Prasad, and **P. Natarajan**, "Context-dependent Pronunciation Modeling for Iraqi ASR," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Taiwan, April 2009.

R. Prasad, C. Moran, F. Choi, S. Saleem, R. Meermeier, S. Saleem, C.-L. Kao, D. Stallard, and **P. Natarajan**, "Name Aware Speech-to-Speech Translation for English/Iraqi," *IEEE/ACL Workshop on Spoken Language Technology (SLT 2008)*, Goa, India, December 2008.

F. Choi, S. Tsakalidis, S. Saleem, K. Krstovski, C.-L. Kao, R. Meermeier, C. Moran, K. Subramanian, R. Prasad, **P. Natarajan**, and D. Stallard, "Recent Improvements in BBN's English/Iraqi Speech-to-Speech Translation System," *IEEE/ACL Workshop on Spoken Language Technology (SLT 2008)*, Goa, India, December 2008.

R. Prasad, S. Saleem, M. Kamali, R. Meermeier, and **P. Natarajan**, "Improvements in Hidden Markov Model Based Arabic OCR," *International Conference on Pattern Recognition*, Tampa, FL, December 2008.

C.-L. Kao, S. Saleem, R. Prasad, F. Choi, **P. Natarajan**, D. Stallard, K. Krstovski, and M. Kamali, "Rapid Development of an English/Farsi S2S Translation System," *International Workshop on Spoken Language Translation (IWSLT 2008)*, Hawaii, October 2008.

- K. Krstovski, E. MacRostie, R. Prasad, and **P. Natarajan**, "End-to-End Trainable Thai OCR System Using Hidden Markov Models," *Eighth IAPR Workshop on Document Analysis Systems (DAS 2008)*, Nara, Japan, September 2008.
- K. Krstovski, M. Decerbo, R. Prasad, D. Stallard, S. Saleem, and **P. Natarajan**, "A Wearable Headset Speech-to-Speech Translation System," *ACL-08: HLT Workshop on Mobile Language Processing*, pp. 10-12, Columbus, OH, USA, June 2008.
- R. Prasad, S. Saleem, E. MacRostie, **P. Natarajan**, and M. Decerbo, "Multi-Frame Combination for Robust Videotext Recognition," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Las Vegas, NV, USA, March 2008.
- K. Subramanian, R. Prasad, E. MacRostie, and **P. Natarajan**, "Robust Named Entity Detection in Videotext using Character Lattices," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Las Vegas, NV, USA, March 2008.
- D. Liu, S. McVeety, R. Prasad, and **P. Natarajan**, "Semi-Supervised Topic Classification for Low Resource Languages," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Las Vegas, NV, USA, March 2008.
- D. Stallard, C.-L. Kao, K. Krstovski, D. Liu, **P. Natarajan**, F. Choi, R. Prasad, S. Saleem, and K. Subramanian, "Recent Improvements and Performance Analysis of ASR and MT in a Speech-to-Speech Translation System," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Las Vegas, NV, USA, March 2008.
- P. Natarajan**, R. Prasad, K. Subramanian, S. Saleem, and R. Schwartz, "Finding Structure in Noisy Text: Topic Classification and Unsupervised Clustering," *Special Issue of Springer International Journal on Document Analysis and Recognition (IJDAR) on Noisy Text Analytics*, ISSN: 1433-2833 (Print) 1433-2825 (Online), Vol. 10, No. 3-4, pp. 187-198, December 2007.
- K. Subramanian, D. Stallard, R. Prasad, S. Saleem, and **P. Natarajan**, "Semantic Translation Error Rate for Evaluating Translation Systems," *Proceedings of 2007 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU'07)*, Kyoto, Japan, December 9-13, 2007.
- H. Cao, R. Prasad, **P. Natarajan**, E. MacRostie, V. Govindraj, "Page Segmentation Based on Smearing and Error Correction Unifying Top-down and Bottom-up Approaches," *International Conference on Document Analysis and Recognition (ICDAR)*, Brazil, September 2007.
- D. Stallard, F. Choi, C. Kao, K. Krstovski, **P. Natarajan**, R. Prasad, S. Saleem, and K. Subramanian, "The BBN 2007 Displayless English/Iraqi Speech-to-Speech Translation System," *Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech 2007)*, pp. 2817-2821, Antwerp, Belgium, August 27-31, 2007.
- K. Subramanian, R. Prasad, and **P. Natarajan**, "Optimal Estimation of Rejection Thresholds for Topic Spotting," *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, Honolulu, Hawaii, USA, April 2007.
- R. Prasad, K. Krstovski, F. Choi, S. Saleem, **P. Natarajan**, M. Decerbo, and D. Stallard, "Real-time speech-to-speech translation for PDAs," *IEEE Conference on Portable Devices (Portable'07)*, Orlando, FL, March 2007.
- R. Prasad, **P. Natarajan**, K. Subramanian, S. Saleem, and R. Schwartz, "Finding Structure in Noisy Text: Topic Classification and Unsupervised Clustering," *IJCAI-2007 Workshop on Analytics for Noisy Unstructured Text Data*, Hyderabad India, Jan 2007.
- D. Stallard, F. Choi, K. Krstovski, **P. Natarajan**, R. Prasad, S. Saleem, and R. Suleiman, "Design and Evaluation of the 2006 BBN English/Iraqi Two-way Speech Translation System," *IEEE/ACL 2006 Workshop on Spoken Language Translation*, Aruba, December, 2006.

- P. Natarajan**, S. Saleem, R. Prasad, E. MacRostie, K. Subramanian, “Multilingual Offline Handwriting Recognition,” *Symposium of Arabic and Chinese Handwriting*, Maryland, 2006.
- S. Saleem, K. Subramanian, R. Prasad, D. Stallard, C. Kao, **P. Natarajan**, and R. Suleiman “Improvements in Machine Translation for English/Iraqi Speech Translation,” *Proceedings of the 8th Annual Conference of the International Speech Communication Association (Interspeech 2007)*, pp. 2445-2448, Antwerp, Belgium, August 27-31, 2007.
- S. Saleem, R. Prasad, and **P. Natarajan**, “Colloquial Iraqi ASR for Speech-to-Speech Translation,” *Int. Conf. Spoken Language Processing*, Pittsburg, PA, Sept. 2006.
- D. Stallard, F. Choi, K. Krstovski, **P. Natarajan**, R. Prasad, and S. Saleem, “A Hybrid Phrase-based/Statistical Speech Translation System,” *Int. Conf. Spoken Language Processing*, Pittsburg, PA, Sept. 2006.
- E. MacRostie and **P. Natarajan**, “Simplified Training System for BBN Byblos OCR System,” *Symposium on Document and Image Understanding Technologies*, November 2005, Baltimore, Maryland.
- P. Natarajan**, R. Sundaram, R. Prasad, and E. MacRostie, “Character Duration Modeling for Speed Improvements in the BBN Byblos OCR System”, *Int. Conference on Document Analysis and Recognition*, Seoul, Korea, Aug. 29-Sep. 1, 2005.
- M. Decerbo, **P. Natarajan**, R. Prasad, E. MacRostie, and A. Ravindran, “Performance Improvements to the BBN Byblos OCR System”, *Int. Conference on Document Analysis and Recognition*, Seoul, Korea, Aug. 29-Sep. 1, 2005.
- E. MacRostie, **P. Natarajan**, M. Decerbo, and R. Prasad, “The BBN Byblos Japanese Optical Character Recognition (OCR) System,” *Int. Conference on Pattern Recognition*, Cambridge, UK, Aug. 23-24, 2004.
- D. Stallard, J. Makhoul, F. Choi, E. Macrostie, **P. Natarajan**, R. Schwartz, B. Zawaydeh, “Design and Evaluation of a Limited Two-way Speech Translator,” *Eurospeech 2003*, Geneva, Switzerland.
- P. Natarajan**, M. Decerbo, T. Keller, R. Schwartz, and J. Makhoul, “Porting the BBN BYBLOS OCR System to New Languages,” *Proc. Symposium on Document and Image Understanding Technology*, April 9 – 11 2003, Greenbelt, Maryland.
- P. Natarajan**, R. Prasad, B. Suhm, and D. McCarthy, “Speech-Enabled Natural Language Call Routing: BBN Call Director,” *Int. Conf. Spoken Language Processing*, Denver, CO, pp.1161-1164, Sept. 16-20, 2002.
- D. Stallard, **P. Natarajan**, M. Noamany, R. Schwartz, and J. Makhoul, “Design for a Speech-to-Speech Translator for Field Use,” *Proc. International Conference on Spoken Language Processing 2002*, Denver, Colorado.
- P. Natarajan**, R. Prasad, R. Schwartz, and J. Makhoul, “A Scalable Architecture for Directory Assistance Automation,” *IEEE Int. Conf. Acoustics Speech and Signal Proc.*, Orlando, FL, pp. 3273-3276, May 13-17, 2002.
- P. Natarajan**, B. Elmieh, R. Schwartz, and J. Makhoul, “Video-text OCR Using Hidden Markov Models,” *Proc. International Conference on Document Analysis and Recognition*, September 2001, Seattle, Washington.
- P. Natarajan**, R. Schwartz, J. Makhoul, “OCR of Low-resolution Text Images from Diverse Sources,” *Proc. Symposium on Document and Image Understanding Technology*, April 2001, Greenbelt, Maryland.
- P. Natarajan**, Z. Lu, R. Schwartz, I. Bazzi, J. Makhoul, “Multilingual Machine Printed OCR,” Chapter in *Hidden Markov Models – Applications in Computer Vision*, Series in Machine Perception and Artificial Intelligence, Editors. Horst Bunke and Terry Caelli, Vol. 45, World Scientific Publishing Company.

- P. Natarajan**, Z. Lu, R. Schwartz, I. Bazzi, J. Makhoul, "Multilingual Machine Printed OCR," *International Journal of Pattern Recognition and Artificial Intelligence*, Vol. 15, No. 1 (2001) 43-63, World Scientific Publishing Company.
- P. Natarajan**, I. Bazzi, Z. Lu, R. Schwartz, J. Makhoul, "Robust OCR of Degraded Documents," *Proc. International Conference on Document Analysis and Recognition*, September 19-22, 1999, Bangalore, India.
- R. Schwartz, Z. Lu, **P. Natarajan**, I. Bazzi, A. Kornai, J. Makhoul, "Advances in the BBN BYBLOS OCR System," *Proc. Symposium on Document Image Understanding Technology*, , April 1999, Annapolis, Maryland.
- I. Bazzi, **P. Natarajan**, R. Schwartz, A. Kornai, Z. Lu, J. Makhoul, "OCR of Degraded Documents using HMM-Based Techniques," *Proc. Symposium on Document Image Understanding Technology*, April 1999, Annapolis, Maryland.
- Z. Lu, I. Bazzi, A. Kornai, J. Makhoul, **P. Natarajan**, and R. Schwartz "Robust Language-Independent OCR System," *Proc. SPIE Vol. 3584*, p. 96-104, 27th AIPR Workshop: Advances in Computer-Assisted Recognition, Edited by R. J. Mericsko, January 1999.
- P. Natarajan**, J. P. Noonan, S. S. Agaian, "Logical Wavelets," *Proc. SPIE Vol. 3164*, p. 78-89, *Applications of Digital Image Processing XX*, Edited by A. G. Tescher, Oct. 1997.
- J. Makhoul, R. Schwartz, C. LaPre, I. Bazzi, Z. Lu, **P. Natarajan**, "A Language-Independent Methodology for OCR," *Proc. Symposium on Document Image Understanding Technology*, Annapolis, MD, pp. 99-107, May 1997.
- J.P. Noonan and **P. Natarajan**, "A Class of Iterative Signal Restoration Algorithms," *Proc. IEEE International Symposium on Information Theory*, September 1995, Whistler, Canada.
-
-