

# Suwon Shon

Research Scientist  
Computer Science and Artificial Intelligent Laboratory  
Massachusetts Institute of Technology  
<http://people.csail.mit.edu/swshon>

32 Vassar Street, 32-G436  
Cambridge, MA, USA  
(+1) 617-435-0584  
swshon@csail.mit.edu

## EDUCATION

- 2017 **Integrated Ph. D of Electrical Engineering**, Korea University Seoul, Korea  
Thesis: Compensating Resource Imbalance in Speech Signal Processing: Application to Speaker Recognition System  
Advisor: Hanseok Ko
- 2010 **B.S of Electrical Engineering**, Korea University Seoul, Korea

## PROFESSIONAL APPOINTMENT

- Apr.2019 - present **Research Scientist**  
Computer Science and Artificial Intelligence Laboratory (CSAIL), MIT Cambridge, MA, USA  
Faculty host : James Glass
- Jul.2017 - Mar.2019 **Postdoctoral Associate**  
Computer Science and Artificial Intelligence Laboratory (CSAIL), MIT Cambridge, MA, USA  
Faculty host : James Glass
- Mar.2017 - Jun.2017 **Research Associate**  
School of Electrical Engineering, Korea University Seoul, Korea  
Faculty host : Hanseok Ko

## RESEARCH ACTIVITIES

- Jul.2017 - present **Research Staff**  
Spoken Language Systems (SLS) group at CSAIL, MIT  
- Lead projects on data-driven supervised and unsupervised learning method to learn a person identity and spoken language from speech, text and face
- ASRU 2019 **Special Session Organizer** [arabicspeech.org/mgb5](http://arabicspeech.org/mgb5)  
The Fifth Multi Genre Broadcast Challenge (MGB-5)  
- Evaluation of automatic speech recognition and dialect ID systems for the multi-genre speech  
- Organized with Ahmed Ali, James Glass, Steve Renals  
- Published 17 Arabic dialect identification dataset that has over 3,000 hours speech collected from YouTube
- ICDM 2018 **Data Challenge Organizer** [mce.csail.mit.edu](http://mce.csail.mit.edu)  
Multi-target Challenge Evaluation 2018 (MCE 2018)  
- Evaluation of Multi-target (blacklist) detection and identification for speaker recognition system  
- Organized with Najim Dehak, Douglas Reynolds and James Glass  
- Total 21 submissions from 12 teams
- COLING 2018 **Workshop Organizer** [alt.qcri.org/vardial2018](http://alt.qcri.org/vardial2018)  
Varieties and Dialects 2018 (VarDial 2018)  
- Arabic dialect identification challenge in broadcast news domain using acoustic and linguistic (text) inputs  
- Organized with Preslav Nakov, Ahmed Ali and James Glass
- NIST SRE 2018 **Member of JHU-MIT team** [sre.nist.gov](http://sre.nist.gov)  
- Applying speech enhancement and speaker diarization on the speaker recognition task  
- Collaborators : Jesús Villalba, Fred Richardson and David Snyder and so on.

- MGB-3, 2017 **Leader of MIT-QCRI team** [www.mgb-challenge.org](http://www.mgb-challenge.org)  
*Ranked #1* - Built dialect ID system by fusion of fine-tuned i-vector linguistic feature  
 - Collaborators : Ahmed Ali and James Glass
- NIST SRE 2016 **Leader of Korea University team** [sre.nist.gov](http://sre.nist.gov)  
 - Domain adaptation for language mismatches between training and test utterance  
 - Collaborators : Hanseok Ko
- Jan.2016 - Feb.2016 **Visiting Graduate Student**  
 Center for Robust Speech Systems, University of Texas at Dallas Richardson, TX, USA  
 - Faculty host : John H.L. Hansen
- Mar.2010 - Feb.2017 **Graduate Student Researcher**  
 Intelligent Signal Processing Laboratory, Korea University Seoul, Korea  
 - Led projects related speaker recognition, acoustic scene analysis and sound source localization  
 - Funding : Samsung Electronics, LG Electronics, Bell Lab. Seoul, Office of Naval Research
- Mar.2009 - Feb.2010 **Undergraduate Student Internship**  
 Samsung Techwin-Korea University Collaboration Center Seoul, Korea  
 - Develop gesture recognition for hand signal interface between human operator and unmanned ground vehicles  
 - Host : Jounghoon Beh

#### PUBLICATIONS

- [1] **S. Shon**, T.-H. Oh, and J. Glass, "Noise-tolerant Audio-visual Online Person Verification using an Attention-based Neural Network Fusion," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019
- [2] **S. Shon**, A. Ali, and J. Glass, "Domain Attentive Fusion for End-to-end Dialect Identification with Unknown Target Domain," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019
- [3] S. Mun and **S. Shon**, "Domain Mismatch Robust Acoustic Scene Classification using Channel Information Conversion," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2019
- [4] **S. Shon**, H. Tang, and J. Glass, "Frame-level Speaker Embeddings for Text-independent Speaker Recognition and Analysis of End-to-end Model," in *IEEE Spoken Language Technology Workshop (SLT)*, 2018
- [5] **S. Shon**, W.-N. Hsu, and J. Glass, "Unsupervised Representation Learning of Speech for Dialect Identification," in *IEEE Spoken Language Technology Workshop (SLT)*, 2018
- [6] M. Zampieri, S. Malmasi, P. Nakov, A. Ali, **S. Shon**, J. Glass, Y. Scherrer, T. Samardžić, N. Ljubešić, and J. Tiedemann, "Language Identification and Morphosyntactic Tagging: The Second VarDial Evaluation Campaign," in *Proceedings of the Fifth Workshop on NLP for Similar Languages, Varieties and Dialects (VarDial 2018)*, pp. 1–17, 2018
- [7] **S. Shon**, A. Ali, and J. Glass, "Convolutional Neural Networks and Language Embeddings for End-to-End Dialect Recognition," in *Proc. Odyssey 2018 The Speaker and Language Recognition Workshop*, pp. 98–104, 2018
- [8] M. Najafian, S. Khurana, **S. Shon**, A. Ali, and J. Glass, "Exploiting convolutional neural networks for phonotactic based dialect identification," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 5174–5178, 2018
- [9] **S. Shon**, A. Ali, and J. Glass, "MIT-QCRI Arabic Dialect Identification System for the 2017 Multi-genre Broadcast Challenge," in *IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU)*, pp. 374–380, 2017
- [10] **S. Shon**, S. Mun, W. Kim, and H. Ko, "Autoencoder based Domain Adaptation for Speaker Recognition under Insufficient Channel Information," in *Interspeech*, pp. 1014–1018, 2017

- [11] **S. Shon**, S. Mun, and H. Ko, “Recursive Whitening Transformation for Speaker Recognition on Language Mismatched Condition,” in *Interspeech*, pp. 2869–2873, 2017
- [12] S. Mun, **S. Shon**, W. Kim, D. K. Han, and H. Ko, “Deep Neural Network based Learning and Transferring Mid-level Audio Features for Acoustic Scene Classification,” in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 796–800, 2017
- [13] S. Mun, **S. Shon**, W. Kim, and H. Ko, “Deep Neural Network Bottleneck Features for Acoustic Event Recognition,” in *Interspeech*, pp. 2954–2957, 2016
- [14] **S. Shon.**, D. Han, S. Mun, and H. Ko, “Non-negative matrix factorisation-based subband decomposition for acoustic source localisation,” *Electronics Letters*, vol. 51, no. 22, pp. 1723–1724, 2015
- [15] **S. Shon**, S. Mun, D. K. Han, and H. Ko, “Maximum likelihood Linear Dimension Reduction of heteroscedastic feature for robust Speaker Recognition,” in *IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 1–5, IEEE, 2015
- [16] **S. Shon**, D. K. Han, and H. Ko, “Abnormal acoustic event localization based on selective frequency bin in high noise environment for audio surveillance,” in *IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, pp. 87–92, 2013
- [17] **S. Shon**, D. K. Han, J. Beh, and H. Ko, “Full azimuth multiple sound source localization with 3-channel microphone array,” *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, vol. E95-A, no. 4, pp. 745–750, 2012
- [18] **S. Shon**, E. Kim, J. Yoon, and H. Ko, “Sudden noise source localization system for intelligent automobile application with acoustic sensors,” in *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 233–234, 2012
- [19] **S. Shon**, J. Beh, C. Yang, D. Han, and H. Ko, “Motion primitives for designing flexible gesture set in Human-Robot Interface,” in *International Conference on Control, Automation and Systems (ICCAS)*, pp. 1501–1504, 2011

#### TEACHING EXPERIENCES

Teaching Assistance	Statistical Signal Processing at Korea University Signals and Systems at Korea University
Advising	2 undergraduate student as mentor for Undergraduate Research Opportunity Program at MIT Gabrielle C. Rivera, Master of Engineering student at MIT

#### TALKS

<i>Robust Speaker Recognition influenced by noise and face</i>	
Apr.2019	- ETRI <span style="float: right;">Daejeon, Korea</span>
Apr.2019	- KAIST <span style="float: right;">Seoul, Korea</span>
Apr.2019	- Naver <span style="float: right;">Gyeonggi, Korea</span>
<i>Analyzing Hidden Representation of End-to-end Speaker Recognition System</i>	
Jul.2018	- KAIST <span style="float: right;">Daejeon, Korea</span>
Jul.2018	- Kookmin University <span style="float: right;">Seoul, Korea</span>
Jul.2018	- Korea University <span style="float: right;">Seoul, Korea</span>
Jul.2018	- NC Soft <span style="float: right;">Gyeonggi, Korea</span>
Jul.2018	- Naver <span style="float: right;">Gyeonggi, Korea</span>
<i>Recent Speaker Recognition Progress</i>	
Apr.2018	- Philips Visit Day <span style="float: right;">Cambridge, MA, USA</span>
<i>Speaker and Dialect Recognition under Limited Resources</i>	
Nov.2017	Qatar Computing Research Institute <span style="float: right;">Doha, Qatar</span>

## REFEREE SERVICE

ISCA Speech Communications  
ISCA Computer Speech and Language  
ISCA Interspeech  
IEEE/ACM Transaction on Audio, Speech and Language Processing  
IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)  
IET Electronics Letter  
IET Science, Measurement and Technology

## REFEREES (persons who I may ask reference letters)

- Postdoc advisor **James Glass**, Senior Research Scientist, MIT CSAIL  
E-mail: glass@mit.edu
- Ph.D advisor **Hanseok Ko**, Professor, Korea University  
E-mail : hsko@korea.ac.kr
- Ph.D. mentor **David Han**, Senior Scientist, U. S. Army Research Laboratory  
E-mail: ctmkhan@gmail.com
- Collaborator **Ahmed Ali**, Principal Engineer, Qatar Computing Research Institute  
E-mail: amali@hbku.edu.qa
- Collaborator **Douglas Reynolds**, Senior Technical Staff, MIT Lincoln Laboratory  
E-mail: dar@ll.mit.edu
- Collaborator **Najim Dehak**, Professor, Johns Hopkins University  
E-mail: ndehak3@jhu.edu
- Collaborator **Taesu Kim**, CEO, Neosapience, Inc.  
E-mail: taesu@neosapience.com