Chanwoo Kim

Corporate Executive Vice President at Samsung Research.
Head of Language and Voice Team
56 Seongchon-Gil, Seoul R & D campus, Seocho-gu, Seoul, 06765, South Korea
Cell: +82-10-4043-3996, E-mail: chanw.com@samsung.com

Google Scholar page : Link to the page (h-index :24, citations: 2677)

Semantic Scholar page : `https://www.semanticscholar.org/author/Chanwoo-Kim/14073033`

DBLP : `https://dblp.org/pid/62/79.html`

Personal homepage: `https://chanwcom.github.io/`

Previous CMU personal homepage(No longer updated from 2017): `http://www.cs.cmu.edu/~chanwook`

Education

Ph. D.
Language Technologies Institute, School of Computer Science
Carnegie Mellon University, Pittsburgh, PA — Sept. 2005 - Dec. 2010
Advisor: Prof. Richard Stern
THESIS - Signal Processing for Robust Speech Recognition Motivated by Auditory Processing

M. S.
School of Electrical Engineering
Seoul National University, Seoul, Korea — March 1998 - Feb. 2001
Advisor: Prof. Wonyong Sung
THESIS - Implementation of an Intonation and Pronunciation Checking System for Embedded Systems

B. S.
School of Electrical Engineering
Seoul National University, Seoul, Korea — March 1994 - Feb. 1998
Thesis advisor: Prof. Min-Koo Han

High School
Seoul Science High School, Seoul, Korea — March 1991 - Feb. 1994

Award

- 2019 IEEE Signal Processing Society Best Paper Award, Dec. 2019 for Power-Normalized Cepstral Coefficients (PNCC) for Robust Speech Recognition.
(https://signalprocessingsociety.org/newsletter/2020/01/2019-ieee-signal-processing-society-awardees)

- Google spot bonus award, Acoustic modeling for Google Home and its commercialization. Dec. 2016.
- 17-th Samsung Humantech Thesis Bronze Prize, C. Kim and R. M. Stern, "Power-Normalized cepstral coefficients for robust speech recognition", Feb. 2011.
- 16-th Samsung Hamantech Thesis Honour Prize, C. Kim and R. M. Stern, "Small power boosting and spectral subtraction for robust speech recognition", Feb. 2010
- Research fellowship for Ph. D. study 2005 - 2010, Carnegie Mellon University, Jan. 2005.
- Scholarship for studying abroad 2005 - 2006, Korea Institute of Energy Evaluation and Planning, July, 2005.
- Graduate cum laude with a bachelors degree, Seoul National University, Feb. 1998.

Research Interests

- Automatic speech recognition: end-to-end speech recognition and noise robustness
- Text-to-speech: end-to-end text-to-speech and speech-to-translated text
- Machine learning, deep-learning, and data-augmentation
- Human auditory processing
- Natural language understanding

Employment

Samsung Research — Feb. 2018 - Current
56 Seongchon-Gil Seoul R & D Campus, Seoul, 06765, Korea
Corporate Vice President, Head of Speech Processing Lab. and Natural Language Processing Lab.

- On-device and server neural end-to-end speech recognition
  Initiated and have been leading the end-to-end speech recognition project. Designed the overall architecture and implemented the feature extraction, data augmentation, and the data pipeline parts. The on-device end-to-end speech recognition system was commercialized for high-end Samsung phones in Nov. 2019.
- Neural Text-to-speech engine development
  The end-to-end Neural TTS was commercialized for Samsung Bixby Service in Oct. 2019.
- Wake-up engine
  Have been leading the wake-up engine project. This will be commercialized for Samsung AI speakers in Feb. 2020.
- Far-field processing
  Have led been leading multi-microphone far-field processing project. This will be commercialized for Samsung AI speakers in Feb. 2020.

Google Google Speech — Feb. 2013 - Feb. 2018
1600 Amphitheatre Pkwy Mountain View, CA, 94043
Senior Software Engineer

- Room simulation system for acoustic model simulation
  Designed and implemented the room acoustics simulation system using the image method for various kinds of real room environments. Using deep learning techniques, acoustic models were trained using data generated by this room acoustics simulator. This system has reduced Word Error Rate to half [C. Kim. et. al., Interspeech 2017 (160+ citations)]. This system is used for commercialized Google speech recognition systems.

- Developing next-generation voice-assistant devices for Google Have participated and lead efforts for the future voice assistant devices at Google.

- Robust speech recognition for Google Home
  Have participated in the Google Home project by generating simulated training sets and performing training and evaluation. This system was launched and commercialized.

- Noise robustness for speech recognition using deep learning
  Developed a system which learns the mapping from noisy feature to clean feature using deep learning techniques for feature enhancement and noise robustness of speech recognition systems.

- Multi-style training for speech recognition systems
  Trained acoustic models for speech recognition systems using various machine learning techniques and the aforementioned room simulator. Using this Multi-style TRaining (MTR) approach, we obtained remarkable improvement under noisy environments.

- Development of a Signal-to-Noise Ratio (SNR) estimation algorithm
  Developed a modified Waveform Amplitude Distribution Analysis (WADA) algorithm as a Mapper in MapReduce at Google. This algorithm has been used extensively at Google speech (130+ citations).

Microsoft Microsoft Speech Jan. 2011 - Jan. 2013
One Microsoft Way, Redmond, WA, 98052
Speech Scientist

- Development of the SMAD algorithm
  Developed a sound source separation algorithm named Statistical Modeling of Angle Distribution (SMAD) for multi-microphone sound source separation

- Distributed processing for speech recognition
  Participating in developing the infrastructure for processing big data for speech recognition using Cosmos.

- Post-processing of Gaussian Mixture Model (GMM)
  Introducing an approach to balance deletion/insertion errors by modifying the mixture components of GMM.

LG Electronics June 2003 - Aug. 2005
128, Yeoui-daero Yeongdeungpo-gu, Seoul, Korea, 150-721
Senior Research Engineer , VoIP Phone Conference Call Tech Leader

- VoIP handset project Manging the conference-call part using the Session Initiation Protocol (SIP).

- Video On Demand (VOD) system project Participating in the VOD project working on streaming protocols such as Real-Time Protocol (RTP), Real-Time Control Protocol (RTCP), and Real-Time Streaming Protocol (RTSP).

- Development of an audio/video synchronization algorithm Developing an audio/video synchronization algorithm for Video Telephony (VT) using Network Time Protocol (NTP).

- Testing and evaluation of multimedia codes H.263, H.264, AAC, AAC+, and H.323

- Multimedia platform survey Texas Instruments (TI) Open Multimedia Application Platform (OMAP), Neomagic MiMagic6, Renesas SH-Mobile V. etc

SK Teletech — June 2003 - Aug. 2005
Research Engineer

Edumediatek — June 2000 - July 2002
Research Engineer
A research venture company established by Prof. Wonyong Sung at Seoul National University

- Computer Aided Language Learning (CALL) System Intonation and vowel checking system for embedded systems

- Development of a robust formant extraction algorithm Combination of the spectral peak picking and root solving

- Development of a robust pitch extraction algorithm

Graduate research

- Power Normalized Central Coefficients (PNCC) Motivated by human auditory processing, this new feature set incorporates modulation frequency, temporal masking, and rate-nonlinearity concepts. The features require low computation and on-line implementation is possible.

- Small Power Boostin (SPB) Algorithm Developed the SPB algorithm motivated by the human rate-intensity as well as temporal and spectral smoothing. This algorithm works especially well for non-stationary noise such as music noise.

- Dual-microphone speech enhancement systems Developed the Phase Difference Channel Weighting (PDCW) algorithm which performs sound source separation without a priori knowledge of source locations.

- Single-microphone speech enhancement system: Developed the Power-function-based Power Distribution Normalization (PPDN) algorithm which enhances speech.

- Automatic speech recognition with low computational cost Focused on developing online algorithms, which are robust against noise with low computational complexity.

- SNR Estimation algorithm for speech: Developed a new algorithm which estimates Signal-to-Noise Ratio (SNR) more accurately than the standard NIST STNR algorithm.

- Automatic ITD threshold detection algorithm: Developed a new algorithm which can obtain the optimal ITD threshold for spatial separation of signals.

- Dereverberation algorithm Developed Suppression of Slowly-varying and Falling edges (SSF): a simple algorithm for robust speech recognition that is highly effective in reverberant environments.

- HLab C++ automatic speech recognition system Developed a C++ HMM (Hidden Markov Model) speech recognition system from scratch.

- Power Normalized Voice Activity Detection (PN-VAD) Developed a new voice activity detection algorithm that is very accurate and computationally efficient.

- DARPA GALE project Robust front end for the DARPA GALE project. Applied a new feature set that we developed for the 180-hr training set and the large vocabulary Broadcast News test set.

- NSF auditory signal processing project Developed features motivated by auditory processing, binaural processing, and poly-aural processing.

- Samsung speech-to-speech translator project Developed low-complexity online algorithms for embedded processors. Developed an on-line noise-robustness feature extraction algorithm, a binaural source-separation algorithm, an online VAD, an online MVN, and supporting technologies. Applied the algorithm to a 64,000-word database with a 500-hour training set.

- Voice activity detector project Designed a robust VAD which requires very low computation.

Invited Talks

- "Introduction to contemporary speech recognition systems and recent advances from deep-learning and human auditory processing techniques," Korea Aerospace University, Goyang, Korea, May. 2015.
- "Research direction for robust speech recognition," Samsung Advanced Institute of Technology, Suwon, Korea, Dec. 2014.
- "Speech recognition research using neural networks," Seoul National University, Seoul Korea, Sept. 2014.
- "Speech Noise Robustness for Intelligent Cars and Home Appliances," Yonsei University, Wonju, Korea, Sept. 2014.
- "TMT, SSF, and PAINT algorithms for far-field speech recognition," Sogang University, July, 2014.
- "Speech Recognition Infrastructure and Noise Robustness Efforts," Seoul National University. July, 2014.
- "Signal Processing for Robust Speech Recognition Motivated by Auditory Processing," Sogang University, Sept. 2012.

Teaching Experience

- Teaching Assistant, 18-792 Advanced Digital Signal Processing, Electrical and Computer Engineering Department, Carnegie Mellon University, Spring, 2010.
- Teaching Assistant, including weekly recitations, 15-211 fundamental data structures and algorithms, School of Computer Science, Carnegie Mellon University, Fall, 2009.

,

International Jounral Papers

[1] T. Sainath, R. J. Weiss, K. W. Wilson, B. Li, A. Narayanan, E. Variani, M. Bacchiani, I. Shafran, A. Senior, K. Chin, A. Misra, and **C. Kim**, "Multichannel signal processing with deep neural networks for automatic speech recognition," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, Feb. 2017.

[2] **C. Kim** and R. M. Stern, "Power-normalized cepstral coefficients (pncc) for robust speech recognition," *IEEE/ACM Trans. Audio, Speech, Lang. Process.*, pp. 1315–1329, July 2016, (IEEE Signal Processing Society (SPS) Best Paper Award - 2019).

[3] B. Cho, H. Kwon, J-W. Cho, **C. Kim**, R. M. Stern, and H. Park, "A Subband-Based Stationary-Component Suppression Method Using Harmonics and Power Ratio for Reverberant Speech Recognition," *IEEE Signal Process. Lett.*, pp. 780–784, June 2016.

[4] **C. Kim**, K. Seo, and W. Sung, "A robust formant extraction algorithm combining spectral peak-picking and roots polishing," *Eurasip Journ. on Applied Signal Processing*, vol. 2006, pp. Article ID 67 960, 16 pages, 2006.

[5] **C. Kim**, K. Seo, and W. Sung, "Efficient media synchronization method for video telephony system," *IEICE Trans. Information and Systems*, vol. E89-D, no. 6, pp. 1901–1905, June 2006.

[6] **C. Kim** and K. Seo, "Robust DTW-based recognition algorithm for hand-held consumer devices," *IEEE Trans. Consumer Electronics*, vol. 51, no. 2, pp. 699–709, May 2005.

Book Chapter

[1] T. Sainath, R. J. Weiss, K. W. Wilson, B. Li, A. Narayanan, E. Variani, M. Bacchiani, I. Shafran, A. Senior, K. Chin, A. Misra, and **C. Kim**, "Raw Multichannel Processing Using Deep Neural Networks," in *New Era for Robust Speech Recognition: Exploiting Deep Learning*, S. Watanabe, M. Delcroix, F. Metze, and J. R. Hershey, Ed. Springer, Oct. 2017.

International Conference Papers: Peer-Reviwed

[1] J. Kim, M. Kumar, D. Gowda, A. Garg, and **C. Kim**, "Semi-supervised transfer learning for language expansion of end-to-end speech recognition models to low-resource languages," in *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2021 (accepted).

[2] J. Kim, M. Kumar, D. Gowda, A. Garg, and **C. Kim**, "A comparison of streaming models and data augmentation methods for robust speech recognition," in *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2021 (accepted).

[3] D. Gowda, A. Garg, J. Kim, M. Kumar, S. Singh, A. Gupta, A. Kumar, N. Dawalatabad, A. Maghan, S. Singh, and **C. Kim**, "hitnet: byte-to-bpe hierarchical transcription network for end-to-end speech recognition," in *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2021 (accepted).

[4] N. Dawalatabad, T. Vatsal, A. Gupta, S. Kim, S. Singh, D. Gowda, and **C. Kim**, "Two-pass end-to-end asr model compression," in *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2021 (accepted).

[5] S. Singh, A. Gupta, A. Maghan, D. Gowda, S. Singh, and **C. Kim**, "Streaming end-to-end speech recognition with jointly trained neural feature enhancement," in *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2021 (accepted).

[6] A. Gupta, A. Jayasimha, A. Maghan, S. Singh, D. Gowda, and **C. Kim**, "Voice to action : spoken language understanding for memory-constrained systems," in *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, 2021 (accepted).

[7] **C. Kim**, A. Garg, D. Gowda, S. Mun, and C. Han, "Streaming end-to-end speech recognition with jointly trained neural feature enhancement," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 6773–6777.

[8] A. Gupta, A. Kumar, D. Gowda, K. Kim, S. Singh, S. Singh, and **C. Kim**, "Neural utterance confidence measure for rnn-transducers and two pass models," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 6398–6402.

[9] S. Indurthi, M. A. Zaidi, N. Kumar Lakumarapu, B. Lee, H. Han, S. Ahn, S. Kim, **C. Kim**, and I. Hwang, "Task aware multi-task learning for speech to text tasks," in *ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2021, pp. 7723–7727.

[10] J. Park, **C. Kim**, and W. Sung, "Convolution-based attention model with positional encoding for streaming speech recognition on embedded devices," in *2021 IEEE Spoken Language Technology Workshop (SLT)*, 2021, pp. 30–37.

[11] **C. Kim**, D. Gowda, D. Lee, J. Kim, A. Kumar, S. Kim, A. Garg, and C. Han, "A review of on-device fully neural end-to-end automatic speech recognition algorithms," in *2020 54th Asilomar Conference on Signals, Systems, and Computers*, Nov. 2020.

[12] A. Garg, G. P. Vadisetti, D. Gowda, S. Jin, A. Jayasimha, Y. Han, J. Kim, J. Park, K. Kim, S. Kim, Y. yoon Lee, K. Min, and **C. Kim**, "Streaming On-Device End-to-End ASR System for Privacy-Sensitive Voice-Typing," in *Proc. Interspeech 2020*, 2020, pp. 3371–3375. [Online]. Available: http://dx.doi.org/10.21437/Interspeech.2020-3172

[13] A. Garg, A. Gupta, D. Gowda, S. Singh, and **C. Kim**, "Hierarchical Multi-Stage Word-to-Grapheme Named Entity Corrector for Automatic Speech Recognition," in *Proc. Interspeech 2020*, 2020, pp. 1793–1797. [Online]. Available: http://dx.doi.org/10.21437/Interspeech.2020-3174

[14] A. Kumar, S. Singh, D. Gowda, A. Garg, S. Singh, and **C. Kim**, "Utterance confidence measure for end-to-end speech recognition with applications to distributed speech recognition scenarios," in *INTERSPEECH-2020*, Oct. 2020.

[15] D. Gowda, A. Kumar, K. Kim, H. Yang, A. Garg, S. Singh, J. Kim, M. Kumar, S. Jin, S. Singh, and **C. Kim**, "Utterance Invariant Training for Hybrid Two-Pass End-to-End Speech Recognition," in *Proc. Interspeech 2020*, 2020, pp. 2827–2831. [Online]. Available: http://dx.doi.org/10.21437/Interspeech.2020-3230

[16] **C. Kim**, K. Kim, and S. Indurthi, "Small energy masking for improved neural network training for end-to-end speech recognition," in *2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2020, pp. 7684–7688.

[17] S. Indurthi, H. Han, N. K. Lakumarapu, B. Lee, I. Chung, S. Kim, and **C. Kim**, "End-end speech-to-text translation with modality agnostic meta-learning," in *2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2020, pp. 7904–7908.

[18] **C. Kim**, M. Kumar, K. Kim, and D. Gowda, "Power-law nonlinearity with maximally uniform distribution criterion for improved neural network training in automatic speech recognition," in *2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, Dec. 2019, pp. 988–995.

[19] **C. Kim**, S. Kim, K. Kim, M. Kumar, J. Kim, K. Lee, C. Han, A. Garg, E. Kim, M. Shin, S. Singh, L. Heck, and D. Gowda, "End-to-end training of a large vocabulary end-to-end speech recognition system," in *2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, Dec. 2019, pp. 562–569.

[20] A. Garg, D. Gowda, A. Kumar, K. Kim, M. Kumar, and **C. Kim**, "Improved multi-stage training of online attention-based encoder-decoder models," in *2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, Dec. 2019, pp. 70–77.

[21] K. Kim, K. Lee, D. Gowda, J. Park, S. Kim, S. Jin, Y.-Y. Lee, J. Yeo, D. Kim, S. Jung, J. Lee, M. Han, and **C. Kim**, "Attention based on-device streaming speech recognition with large speech corpus," in *2019 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, Dec. 2019, pp. 956–963.

[22] **C. Kim**, M. Shin, A. Garg, and D. Gowda, "Improved vocal tract length perturbation for a state-of-the-art end-to-end speech recognition system," in *INTERSPEECH-2019*, Graz, Austria, Sept. 2019, pp. 739–743. [Online]. Available: http://dx.doi.org/10.21437/Interspeech.2019-3227

[23] D. Gowda, A. Garg, K. Kim, M. Kumar, and **C. Kim**, "Multi-task multi-resolution char-to-bpe cross-attention decoder for end-to-end speech recognition," in *INTERSPEECH-2019*, Graz, Austria, Sept. 2019, pp. 2783–2787. [Online]. Available: http://dx.doi.org/10.21437/Interspeech.2019-3216

[24] A. Menon, **C. Kim**, and R. M. Stern, "Robust recognition of reverberant and noisy speech using coherence-based processing," in *ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, May 2019, pp. 6775–6779.

[25] **C. Kim**, E. Variani, A. Narayanan, and M. Bacchiani, "Efficient implementation of the room simulator for training deep neural network acoustic models," in *INTERSPEECH-2018*, Sept 2018, pp. 3028–3032. [Online]. Available: http://dx.doi.org/10.21437/Interspeech.2018-2566

[26] X. Zhou, C. Kwan, B. Ayhan, **C. Kim**, K. Kumar, and R. Stern, "A comparative study of spatial speech separation techniques to improve speech recognition," in *International Symposium on Neural Networks*, June 2018, pp. 494–502.

[27] **C. Kim**, T. Sainath, A. Narayanan, A. Misra, R. Nongpiur, and M. Bacchiani, "Spectral distortion model for training phase-sensitive deep-neural networks for far-field speech recognition," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2018, pp. 5729–5733.

[28] **C. Kim**, A. Menon, M. Bacchiani, and R. Stern, "Sound source separation using phase difference and reliable mask selection selection," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, April 2018, pp. 5559–5563.

[29] A. Menon, **C. Kim**, U. Kurokawa, and R. M. Stern, "Binaural processing for robust recognition of degraded speech," in *2017 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, Dec 2017, pp. 24–31.

[30] **C. Kim**, A. Misra, K.K. Chin, T. Hughes, A. Narayanan, T. Sainath, and M. Bacchiani, "Generation of simulated utterances in virtual rooms to train deep-neural networks for far-field speech recognition in Google Home," in *INTERSPEECH-2017*, Aug. 2017, pp. 379–383.

[31] B. Li, T. Sainath, A. Narayanan, J. Caroselli, M. Bacchiani, A. Misra, I. Shafran, H. Sak, G. Pundak, K. Chin, K-C Sim, R. Weiss, K. Wilson, E. Variani, **C. Kim**, O. Siohan, M. Weintraub, E. McDermott, R. Rose, and M. Shannon, "Acoustic modeling for Google Home," in *INTERSPEECH-2017*, Aug. 2017, pp. 399–403.

[32] A. Menon, **C. Kim**, and R. M. Stern, "Robust Speech Recognition Based on Binaural Auditory Processing," in *INTERSPEECH-2017*, Aug. 2017, pp. 3872–3876.

[33] **C. Kim** and K. Chin, "Sound source separation algorithm using phase difference and angle distribution modeling near the target," in *INTERSPEECH-2015*, Sept. 2015, pp. 751–755.

[34] **C. Kim**, K. Chin, M. Bacchiani, and R. M. Stern, "Robust speech recognition using temporal masking and thresholding algorithm," in *INTERSPEECH-2014*, Sept. 2014, pp. 2734–2738.

[35] H. Park, M. Maciejewski, **C. Kim**, and R. M. Stern, "Robust speech recognition in reverberant environments using subband-based steady-state monaural and binaural suppression," in *INTERSPEECH-2014*, Sept. 2014, pp. 2715–2718.

[36] **C. Kim**, C. Khawand, and R. M. Stern, "Two-microphone source separation algorithm based on statistical modeling of angle distributions," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, March 2012, pp. 4629–4632.

[37] **C. Kim** and R. M. Stern, "Power-normalized cepstral coefficients (pncc) for robust speech recognition," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, March 2012, pp. 4101–4104, (17–th Samsung Humantech Thesis Bronze Prize).

[38] K. Kumar, **C. Kim**, and R. M. Stern, "Delta-spectral cepstral coefficients for robust speech recognition," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2011, pp. 4784–4787.

[39] **C. Kim**, K. Kumar, and R. M. Stern, "Binaural sound source separation motivated by auditory processing," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2011, pp. 5072–5075.

[40] **C. Kim** and R. M. Stern, "Feature extraction for robust speech recognition based on maximizing the sharpness of the power distribution and on power flooring," in *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*, March 2010, pp. 4574–4577.

[41] **C. Kim** and R. M. Stern, "Nonlinear enhancement of onset for robust speech recognition," in *INTERSPEECH-2010*, Sept. 2010, pp. 2058–2061.

[42] **C. Kim**, K. Eom, J. Lee, and R. M. Stern, "Automatic selection of thresholds for signal separation algorithms based on interaural delay," in *INTERSPEECH-2010*, Sept. 2010, pp. 729–732.

[43] **C. Kim**, K. Kumar and R. M. Stern, "Robust speech recognition using small power boosting algorithm," in *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, Dec. 2009, pp. 243–248, (16–th Samsung Humantech Thesis Honour Prize).

[44] **C. Kim** and R. M. Stern, "Power function-based power distribution normalization algorithm for robust speech recognition," in *IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*, Dec. 2009, pp. 188–193.

[45] **C. Kim** and R. M. Stern, "Feature extraction for robust speech recognition using a power-law nonlinearity and power-bias subtraction," in *INTERSPEECH-2009*, Sept. 2009, pp. 28–31.

[46] **C. Kim**, K. Kumar, B. Raj, and R. M. Stern, "Signal separation for robust speech recognition based on phase difference information obtained in the frequency domain," in *INTERSPEECH-2009*, Sept. 2009, pp. 2495–2498.

[47] **C. Kim** and R. M. Stern, "Robust signal-to-noise ratio estimation based on waveform amplitude distribution analysis," in *INTERSPEECH-2008*, Sept. 2008, pp. 2598–2601.

[48] R. M. Stern, E. Gouvea, **C. Kim**, K. Kumar, and H.-M. Park, "Binaural and multiple-microphone signal processing motivated by auditory perception," in *Hands-Free Speech Communication and Microphone Arrays (HSCMA)*, May. 2008, pp. 98–103.

[49] **C. Kim**, Y.-H. Chiu, and R. M. Stern, "Physiologically-motivated synchrony-based processing for robust automatic speech recognition," in *INTERSPEECH-2006*, Sept. 2006, pp. 1975–1978.

[50] **C. Kim**, K. Seo, W. Sung, and S. Jung, "Efficient audio/video synchronization method for video telephony system in consumer cellular phones," in *IEEE Int. Conf. on Consumer Electronics*, Jan. 2006, pp. 137–138.

[51] **C. Kim** and K. Seo, "Robust dtw-based recognition algorithm for hand-held consumer device," in *IEEE Int. Conf. on Consumer Electronics*, Jan. 2005, pp. 433–434.

[52] **C. Kim** and W. Sung, "Implementation of intonational quality assessment system," in *INTERSPEECH-2002*, Sept. 2002, pp. 1225–1228.

[53] **C. Kim** and W. Sung, "Vowel pronunciation accuracy checking system based on phoneme segmentation and formants extraction," in *Int. Conf. Speech Processing*, Aug. 2001, pp. 447–452.

International Conference Papers: Non-Peer-Reviwed

[1] R. M. Stern, **C. Kim**, A. R. Moghimi, and A. Menon, "Binaural technology and automatic speech recognition," in *International Congress on Acoustics, ICA1016-633*, Buenos Aires, Argentina, Sept. 2016, pp. 1–10. [Online]. Available: http://www.ica2016.org.ar/ica2016proceedings/ica2016/ICA2016-0633.pdf

Domestic Conference Papers

[1] **C. Kim** S. Park and K. Seo, "Efficient audio/video synchronization method for video mobile communication terminals (in korean)," in *Korea Computer Congress*, July 2005, pp. 355–357.

Theses

[1] **C. Kim**, "Signal processing for robust speech recognition motivated by auditory processing," Ph.D. dissertation, Carnegie Mellon University, Pittsburgh, PA USA, Dec. 2010.

[2] **C. Kim**, "Implementation of an Intonation and Pronunciation Checking System for Embedded Systems," Master's thesis, Seoul National University, Seoul, Korea, Feb. 2001.

US Patents Issued

[1] **C. Kim**, R. C. Nongpiur, and T. Sainath, “Device-specific multi-channel data compression neural network,” Patent US 10,490,198, Nov. 26, 2019.

[2] R. Nongpiur, A. Misra, and **C. Kim**, “Auralization for multi-microphone devices,” Patent US 10,412,489, Sept. 10, 2019.

[3] **C. Kim**, R. C. Nongpiur, V. Peddinti, and M. Bacchiani, “Speech recognition using acoustic features in conjunction with distance information,” Patent US 10,339,929, July 2, 2019.

[4] **C. Kim**, R. C. Nongpiur, and M. Bacchiani, “query endpointing based on lip detection,” Patent US 10,332,515, June 25, 2019.

[5] **C. Kim**, R. C. Nongpiur, and A. Narayanan, “Sound source estimation using neural networks,” Patent US 10,063,965, Aug. 28, 2018.

[6] **C. Kim**, R. C. Nongpiur, and A. Misra, “Auralization for multi-microphone devices,” Patent US 9,992,570, June 5, 2018.

[7] **C. Kim**, R. C. Nongpiur, and T. Sainath, “Device specific multi-channel data compression,” Patent US 9,875,747, Jan. 23, 2018.

[8] C. Khwand and **C. Kim**, “Target object angle determination using multiple cameras, United States Patents,” Patent US 9,269,146, Feb. 23, 2016.

[9] **C. Kim** and C. Khwand, “Multi-microphone Audio Source Separation Based On Combined Statistical Angle Distributions,” Patent US 9,131,295, Sept. 8, 2015.

[10] **C. Kim**, K. Eom, J. Lee, and R. M. Stern, “Signal separation system and method for automatically selecting threshold to separate sound sources,” Patent US 8,718,293, May 6, 2014.

[11] **C. Kim**, “Formants extracting method combining spectral peak picking and roots extraction,” Patent US 8,000,959, Aug. 16, 2011.

[12] **C. Kim**, “Speech distinction method,” Patent US 7,761,294, July 20, 2010.

[13] K. Seo and **C. Kim**, “Synchronizing video/audio data of mobile communication terminal,” Patent US 7,710,943, May 4, 2010.

[14] **C. Kim**, “Method of filtering speech signals to enhance quality of speech and apparatus thereof,” Patent US 7,590,524, Sept. 15, 2009.

[15] **C. Kim**, “Baseband modem for speech recognition and mobile communication terminal using the same,” Patent US 7,593,853, Sept. 22, 2009.

[16] **C. Kim**, “Mobile device and method for preventing undesired key depression in the same,” Patent US 7,602,377, Oct. 13, 2009.

[17] **C. Kim**, “Speech coding apparatus with perceptual weighting and method therefor,” Patent US 7,603,271, Oct. 13, 2009.

[18] **C. Kim**, “Telephone number retrieval system and method,” Patent US 7,356,356, Apr. 8, 2008.

US Patents Application

[1] **C. Kim**, R. C. Nongpiur, and M. Bacchiani, "Query endpointing based on lip detection," Patent Application, Oct. 31, 2019.

[2] A. Narayanan, **C. Kim** and R. Nongpiur, "Sound Source Estimation Using Neural Networks," Patent Application, March, 2016.

[3] A. Misra, **C. Kim** and R. Nongpiur, "Room Auralization For Multi-Microphone Devices," Patent Application, March, 2016.

[4] **C. Kim**, C. Khawand, and J. Moon, "Automatically optimizing capture of images of one or more subjects," Patent Application US 2012/0 300 092, Nov. 29, 2012. [Online]. Available: http://www.lens.org/lens/patent/US_2012_0300092_A1

[5] **C. Kim**, "Voice coding/decoding method and apparatus," Patent Application US 2006/0 015 330, Jan. 19, 2006. [Online]. Available: http://www.lens.org/lens/patent/US_2006_0015330_A1

[6] **C. Kim**, "Voice coding apparatus and method using PLP in mobile communications terminal," Patent Application US 2006/0 025 991, Feb 2, 2006. [Online]. Available: http://www.lens.org/lens/patent/US_2006_0025991_A1

[7] **C. Kim**, "Mobile terminal having support power pack," Patent Application US 2006/0 111 155, May 25, 2006. [Online]. Available: http://www.lens.org/lens/patent/US_2006_0111155_A1

[8] **C. Kim**, "Method for extracting feature vectors for speech recognition," Patent Application US 2006/0 129 392, June 15, 2006. [Online]. Available: http://www.lens.org/lens/patent/US_2006_0129392_A1

[9] **C. Kim**, "Apparatus and method for reducing power consumption in a mobile communication terminal," Patent Application US 2005/0 057 548, March 17, 2005. [Online]. Available: http://www.lens.org/lens/patent/US_2005_0057548_A1

[10] **C. Kim**, "Voice recognition method," Patent Application US 2005/0 131 693, June 16, 2005. [Online]. Available: http://www.lens.org/lens/patent/US_2005_0131693_A1

Korea Patents Issued

[1] **C. Kim**, "A mobile terminal having a support power pack," Patent Korea 10-0 677 397, Jan. 26, 2007.

[2] **C. Kim**, "Separable mobile terminal," Patent Korea 10-0 677 347, Jan. 26, 2007.

[3] **C. Kim**, "Voice coding/decoding method, and apparatus for the same," Patent Korea 10-0 672 355, Jan. 16, 2007.

[4] **C. Kim**, "Mobile phone," Patent Korea 10-0 677 304, Jan., 2007.

[5] **C. Kim**, "A multi-party system and method for requires reducedcomputational amount," Patent Korea 10-0 733 713, June 23, 2007.

[6] **C. Kim**, "Apparatus for removing noise by using hands-free mike of mobile terminal," Patent Korea 10-0 739 178, July 6, 2007.

[7] **C. Kim**, "Slide type mobile communication terminal applying subdisplay device," Patent Korea 10-0 560 919, March 7, 2006.

[8] **C. Kim** and K. Seo, "A method and a apparatus of synchronization videosignal with audio signal for mobile phone," Patent Korea 10-0 565 333, March, 2006.

[9] **C. Kim**, "Guidance method and apparatus for telephone number," Patent Korea 10-0 595 610, June, 2006.

[10] **C. Kim**, "Key pushing prevention method for portable apparatus," Patent Korea 10-0 595 614, June 23, 2006.

[11] **C. Kim**, "A method and a apparatus of advanced low bit rate linear prediction coding with plp coefficient for mobile phone," Patent Korea 10-0 619 893, Aug. 28, 2006.

[12] **C. Kim**, "Method and apparatus for enhancing quality of speech," Patent Korea 10-0 640 865, Oct. 25, 2006.

[13] **C. Kim**, "Baseband modem and mobile terminal for voice recognition," Patent Korea 10-0 640 893, Oct. 2525, 2006.

[14] **C. Kim**, "Formant frequency detecting method of voice signal," Patent Koarea 10-0 511 316, Aug. 23, 2005.

Technical skills

- Machine Learning Toolkits: Tensorflow, Keras, etc.
- Programming Language : C, C++, Python, Perl, Java, Matlab, Bash, etc.
- Programming Environments : Win32 API, Microsoft Foundation Class (MFC), NumPy/SciPy, MapReduce, and various Linux environments.
- Speech Recognition Engine : Sphinx, Hidden markov model Tool Kit (HTK), returnNN, Janus, etc.
- Development Tool : Various Linux tools including gcc. Microsoft visual studio, Arm Developer Suite (ADS), Code Composer Studio (CCS), etc.
- Real-time OS : Rex, Nucleus

References

- Prof. Richard M. Stern (Ph. D. study advisor 2005-2010)
  Department of Electrical and Computer Engineering, Carnegie Mellon University,
  B24 Porter Hall
  5000 Forbes Ave, Pittsburgh, PA 15213, USA (Tel: +1-412-268-2535)
  E-mail: rms@cs.cmu.edu

- Prof. Wonyong Sung (M. S. study advisor 1999-2001)
  Department of Electrical and Computer Engineering, Seoul National Univ.
  1 Gwanangno, Gwanak-gu, Seoul, Korea, 151-744 (Tel: +82-2-880-1816)
  E-mail: wysung@dsp.snu.ac.kr

- Prof. Gary Geunbae Lee (Supervisor at Samsung Research between Feb. 2018 - Dec. 2019)
  Department of Computer Science,
  Postech Intelligent Software Lab., PIRL 354, POSTECH, 77, Cheongam-ro, Nam-gu, Pohang-Si, Gyeongsangbuk-do, 37673, Rep. of Korea
  E-mail: gblee@postech.ac.kr

- Prof. Yoonsuck Choe (Worked together at Samsung Research AI Center between Feb. 2018-Dec. 2019)
  Department of Computer Science and Engineering Texas A&M University
  3112 TAMU College Station, TX 77843-3112, USA (Tel: +1-979-845-5466)
  E-mail: choe@tamu.edu

Updated Feb., 2021