

Chanwoo Kim

Curriculum Vitae

Contact

5000 Forbes Avenue, Carnegie Mellon University, Pittsburgh, PA 15213
B45 Porter Hall

Office Phone: 412-268-2530 Cell Phone: 412-726-3996

E-mail: chanwook@cs.cmu.edu

Homepage: <http://www.cs.cmu.edu/~chanwook>

Research Interests

Automatic Speech Recognition, Speech Enhancement (Single Microphone, Dual Microphones, and Multiple Microphones), Machine Learning, Embedded Signal Processing, Signal Processing Systems, Human Auditory Processing

Education

Carnegie Mellon University, Pittsburgh, PA

Language Technologies Institute, School of Computer Science

Ph. D. in progress

Thesis : *Signal Processing for Robust Speech Recognition Motivated by Auditory Processing*

Advisor: Prof. Richard Stern

Thesis Committee : Prof. Richard Stern, Prof. Alex Rudnicky, Prof. Bhiksha Raj, Prof. Hynek Hermansky (Johns Hopkins University)

Seoul National University, Seoul, Korea

School of Electrical Engineering and Computer Science

M. S., Feb. 2001

Advisor: Prof. Wonyong Sung

Seoul National University, Seoul, Korea

School of Electrical Engineering

B. S., Graduate cum laude, Feb. 1998

Seoul Science High School, Seoul, Korea

Graduation, Feb 1994

Award

Honour prize, 16-th Samsung Humantech Thesis

Research fellowship for Ph. D. study, Carnegie Mellon University

Korea Power Electricity Engineer Fellowship

Six sigma black belt, LG Electronics

Graduate cum laude with a bachelor's degree, Seoul National University

Employment

- 5/2003–8/2005 **LG Electronics**
Senior Research Engineer
VoIP Phone Conference Call Development Part Leader
Seoul, Korea
- VoIP handset project - SIP protocol and conference call part
VOD (Video On Demand) system project - streaming protocols RTP, RTCP, and RTSP
VT (Video Telephony) project - Development of the audio and video synchronization systems
Multimedia Systems - H.263, H.264, AAC, AAC+, and H.323
Multimedia platform survey - TI OMAP, MiMagic6, SH-Mobile V. etc
Embedded System Programming - Porting codecs for ARM7 or ARM9-based systems
- 3/2003–5/2003 **SK Teletech**
Research Engineer
Handset UI Programmer
Seoul, Korea
- 6/2000–7/2002 **Edumediatic** (A research venture company established by Prof. Wonyong Sung at Seoul National University)
CALL (Computer Aided Language Learning System) Intonation and vowel checking system for embedded systems
Development of a robust formant extraction algorithm
Development of a robust pitch extraction algorithm

Publications

Peer-Reviewed International Journal Papers

- C. Kim and R. M. Stern. Signal separation in the frequency domain using the automatic selection of an ITD threshold. *IEEE Trans. Speech, Audio, Lang. Process.*, (in preparation).
- C. Kim, K. Kumar and R. M. Stern. Small power boosting and spectral subtraction for robust speech recognition. *IEEE Tran. Speech, Audio, Lang. Process.*, (in preparation).
- C. Kim and R. M. Stern. Power-Normalized Cepstral Coefficients for robust speech recognition. *IEEE Tran. Speech, Audio, Lang. Process.*, (in preparation).
- D. T. Chappell, R. M. Stern, Y. Chiu, and C. Kim. Voice Activity Detection: A Tutorial Review. *IEEE Signal Processing Magazine*, (in preparation).
- C. Kim, K. Seo, and W. Sung. Efficient media synchronization method for video telephony system. *IEICE Trans. Information and Systems*, E89-D(6):1901–1905, June 2006.
- 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*, 2006:Article ID 67960, 16 pages, 2006.
- C. Kim and K. Seo. Robust DTW-based recognition algorithm for hand-held consumer devices. *IEEE Trans. Consumer Electronics*, 51(2):699–709, May 2005.

Peer-Reviewed International Conference Papers

- C. Kim and R. M. Stern. Nonlinear enhancement of onset for robust speech recognition. In *INTERSPEECH-2010*, Sept. 2010 (accepted).

- 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 (accepted).
- 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*, pages 4574–4577, March 2010.
- C. Kim, K. Kumar and R. M. Stern. Robust speech recognition using small power boosting algorithm. In *IEEE Automatic Speech Recognition and Understanding Workshop*, pages 243–248, Dec. 2009.
- 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*, pages 188–193, Dec. 2009.
- C. Kim and R. M. Stern. Feature extraction for robust speech recognition using a power-law nonlinearity and power-bias subtraction. In *INTERSPEECH-2009*, pages 28–31, Sept. 2009.
- 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*, pages 2495–2498, Sept. 2009.
- C. Kim and R. M. Stern. Robust signal-to-noise ratio estimation based on waveform amplitude distribution analysis. In *INTERSPEECH-2008*, pages 2598–2601, Sept. 2008.
- R. M. Stern, E. Gouvea, C. Kim, K. Kumar, and H. Park. Binaural and multiple-microphone signal processing motivated by auditory perception. In *Hands-Free Speech Communication and Microphone Arrays, 2008*, pages 98–103, May. 2008.
- C. Kim, Y.-H. Chiu, and R. M. Stern. Physiologically-motivated synchrony-based processing for robust automatic speech recognition. In *INTERSPEECH-2006*, pages 1975–1978, Sept. 2006.
- 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. Consumer Elect.*, pages 137–138, Jan. 2006.
- C. Kim and K. Seo. Robust dtw-based recognition algorithm for hand-held consumer device. In *IEEE Int. Conf. Consumer Elect.*, pages 433–434, Jan. 2005.
- C. Kim and W. Sung. Implementation of intonational quality assessment system. In *INTERSPEECH-2002*, pages 1225–1228, Sept. 2002.
- C. Kim and W. Sung. Vowel pronunciation accuracy checking system based on phoneme segmentation and formants extraction. In *Int. Conf. Speech Processing*, pages 447–452, Aug. 2001.

Domestic Conference Papers (Korea)

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

Theses

- C. Kim. Signal Processing for Robust Speech Recognition Motivated by Auditory Processing. *Ph. D. thesis proposal*, May 2010. 122 pages.
- C. Kim. Implementation of an Intonation and Pronunciation Checking System for Embedded Systems. *M. S. thesis*, Feb. 2001. 71 pages.

US Patents Issued

- K. Seo and C. Kim. Synchronizing video/audio data of mobile communication terminal. *United States Patent*, (7,710,943), 2010.

- C. Kim. Method of filtering speech signals to enhance quality of speech and apparatus thereof. *United States Patent*, (7,590,524), Sept. 2009.
- C. Kim. Baseband modem for speech recognition and mobile communication terminal using the same. *United States Patent*, (7,593,853), Sept. 2009.
- C. Kim. Mobile device and method for preventing undesired key depression in the same. *United States Patent*, (7,602,377), Oct. 2009.
- C. Kim. Speech coding apparatus with perceptual weighting and method therefor. *United States Patent*, (7,603,271), Oct. 2009.
- C. Kim. Telephone number retrieval system and method. *United States Patent*, (7,356,356), Apr. 2008.

US Patent Applications Pending

Many patents are also applied in other countries (European patents, Japanese patents, Chinese patents, and Korea patents)

- C. Kim, R. M. Stern, K. Eom, and J. Lee. Automatic Interaural Time Delay Threshold Selection Method for Sound Source Separation. *United States Patent (Filed)*, 2010.
- C. Kim. Voice coding/decoding method and apparatus. *United States Patent Application*, (Application no. 20060015330).
- C. Kim. Voice coding apparatus and method using PLP in mobile communications terminal. *United States Patent Application*, (Application no.20060025991).
- C. Kim. Mobile terminal having support power pack. *United States Patent Application*, (Application no.20060111155).
- C. Kim. Speech distinction method. *United States Patent Application*, (Application no. 20060111900).
- C. Kim. Method for extracting feature vectors for speech recognition. *United States Patent Application*, (Application no. 20060129392).
- C. Kim. Apparatus and method for reducing power consumption in a mobile communication terminal. *United States Patent Application*, (20050057548).
- C. Kim. Formants extracting method. *United States Patent Application*, (20050075864).
- C. Kim. Voice recognition method. *United States Patent Application*, (20050131693).

Korea Patents Issued

- C. Kim. Formant frequency detecting method of voice signal. *Korea Patent*, (Application no. 10-2003-0069175 ,Registration no.10-0511316), Aug. 2005.
- C. Kim. Slide type mobile communication terminal applying subdisplay device. *Korea Patent*, (Application no. 10-2003-0071130, Registration no. 10-0560919), March 2006.
- C. Kim and K. Seo. A method and a apparatus of synchronization videosignal with audio signal for mobile phone. *Korea Patent*, (Application no. 10-2004-0046697, Registration no. 10-0565333), March 2006.
- C. Kim. Guidance method and apparatus for telephone number. *Korea Patent*, (Application no. 10-2003-0076089, Registration no. 10-0595610), June 2006.
- C. Kim. Key pushing prevention method for portable apparatus. *Korea Patent*, (Application no. 10-2003-0081627, Registration no. 10-0595614), June 2006.

- C. Kim. A method and a apparatus of advanced low bit ratelinear prediction coding with plp coefficient formobile phone. *Korea Patent*, (Application no. 10-2004-0057739, Registration no. 10-0619893), Aug. 2006.
- C. Kim. Speech distinction method. *Korea Patent*, (Application no. 10-2004-0097650, Registration no. 10-0631608), Sept. 2006.
- C. Kim. Method and apparatus for enhancing quality of speech. *Korea Patent*, (Application no. 10-2004-0071371, Registration no. 10-0640865), Oct. 2006.
- C. Kim. Baseband modem and mobile terminal for voice recognition. *Korea Patent*, (Application no. 10-2004-0071327, Registration no. 10-0640893), Oct. 2006.
- C. Kim. A mobile terminal having a support power pack. *Korea Patent*, (Application no. 10-2004-0095924, Registration no. 10-0677397), Jan. 2007.
- C. Kim. Separable mobile terminal. *Korea Patent*, (Application no. 10-2004-0067617, Registration no. 10-0677347), Jan. 2007.
- C. Kim. Voice coding/decoding method, and apparatus for the same. *Korea Patent*, (Application no. 10-2004-0055634, Registration no. 10-0672355), Jan. 2007.
- C. Kim. Mobile phone. *Korea Patent*, (Application no. 10-2004-0009331, Registration no. 10-0677304), Jan. 2007.
- C. Kim. A multi-party system and method for requires reducedcomputational amount. *Korea Patent*, (Application no. 10-2005-0075404, Registration no. 10-0733713), June 2007.
- C. Kim. Apparatus for removing noise by using hands-free mike of mobile terminal. *Korea Patent*, (Application no 10-2005-0072454. , Registration no. 10-0739178), July 2007.

Research Experience

Graduate Work

Power Normalized Central Coefficient (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 Boosting Algorithm (SPB): 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 the source locations.

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

Automatic Speech Recognition with low computational amount : F 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.

Automatic ITD threshold detection algorithm: Developed a new algorithm which can obtain the optimal ITD threshold for blind sound source separation.

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.

Major Graduate Projects

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 online 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.

Teaching Experience

Teaching Assistantship

18-792 Advanced Digital Signal Processing, Electrical and Computer Engineering Department, Carnegie Mellon University : Teaching assistant.

15-211 Fundamentals of Data Structures and Algorithms (using JAVA language), Computer Science Department, Carnegie Mellon University : Teaching assistant, including weekly recitations

Reviewer

Eusipco, Speech Communication

Membership

IEEE Student Member

Technical Skills

Proficient

Programming Language : C, C++ , Perl, Matlab, Shell Scripts

Programming Environment : Win32 API, MFC

Speech Recognition Engine : Sphinx, HTK

Development Tool : Microsoft Visual Studio

OS: Linux, Windows

Familiar

Programming Language: VHDL, JAVA

Document Tool: Latex

Real-time OS: Rex, Nucleus

Speech Recognition Engine: Janus

Development Tool : ARM Developer Suite, CCS for C55x and OMAP

References

Prof. Richard M. Stern

B24 Porter Hall, Electrical and Computer Engineering, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213 (rms @ cs.cmu.edu)

Prof. Bhiksha Raj

Gates-Hillman Center 6705, Language Technologies Institute, School of Computer Science, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213 (bhiksha @ cs.cmu.edu)

Prof. Alex Rudnický

Gates-Hillman Center 7221, Computer Science Department, School of Computer Science, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213 (air @ cs.cmu.edu)

Prof. Wonyong Sung

Room 301-805, School of Electrical Engineering, Seoul National University San 56-1, Shilim-Dong, Kwanak-Gu, Seoul 151-742, Korea (Tel: +82-2-880-1816) (wysung @ dsp.snu.ac.kr)

July 7, 2010