

# Curriculum Vitae of Chao Zhang

Machine Intelligence Laboratory  
Cambridge University Engineering Department

Email: [cz277@cam.ac.uk](mailto:cz277@cam.ac.uk)  
Web: <http://mi.eng.cam.ac.uk/~cz277>

## EDUCATION BACKGROUND

- PhD in Information Engineering** **2012.10 – 2017.1**
- Department of Engineering, Cambridge Univ., UK
  - Supervisor: Prof. Phil Woodland (FREng.)
- MSci in Computer Science and Technology (with honor)** **2009.9 – 2012.7**
- Tsinghua Univ., Beijing, China
  - GPA major: 92.2/100 (rank 12/180)
- BEng in Computer Science and Technology** **2005.9 – 2009.7**
- Tsinghua Univ., Beijing, China

## RESEARCH INTERESTS

Automatic speech recognition, artificial neural networks, natural language processing, pattern recognition, machine learning, data mining, quantitative finance, and cognitive science.

## RESEARCH EXPERIENCE

- Research Associate in Speech Recognition** **2017 – present**
- Machine Intelligence Lab., Cambridge Univ. Engineering Dept.
    - Worked on state-of-the-art ASR and HTK.
- Research Intern at Apple Siri** **2017**
- Siri Speech Team
    - Worked on core acoustic modelling for accented speech recognition.
- HTK V3.5 development** **2013 – present**
- Machine Intelligence Lab., Cambridge Univ. Engineering Dept.
    - Designed and developed the ANN extension to HTK with Prof. Phil Woodland.
- IEEE ASRU 2015 MGB challenge** **2015**
- Machine Intelligence Lab., Cambridge Univ. Engineering Dept.
    - Developed all HTK evaluation systems.
- DARPA BOLT Program 2014 evaluation** **2014**
- Machine Intelligence Lab., Cambridge Univ. Engineering Dept.
    - Developed all DNN tandem and hybrid systems and the HTK joint decoder.
- iARPA BABEL Program 2013 evaluation** **2013**
- Machine Intelligence Lab., Cambridge Univ. Engineering Dept.
    - Extended QuickNet for DNN and developed all DNN recipes.
- CICC stock index prediction system development** **2010 – 2011**
- Asset Management Dept., China International Capital Corporation
    - Created a CRF based index prediction system (25M RBM investment and 22% return rate).
- Research assistant in ASR and NLP** **2009 – 2012**
- CSLT, Tsinghua National Lab. for Information Science & Technology
    - Developed the ASR infrastructures and all ASR systems in the lab.

## **BOOKS & CHAPTERS**

1. P. Woodland, S. Renals, **C. Zhang**, L. Lu, Neural network acoustic modelling, In *Natural Speech Technology*, Cambridge Univ. Press, 2018. **(In Press)**
2. S. Young, G. Evermann, M. Gales, T. Hain, D. Kershaw, X. Liu, G. Moore, J. Odell, D. Ollason, D. Povey, A. Ragni, V. Valtchev, P. Woodland, **C. Zhang**, *The HTK Book (for HTK version 3.5)*, Cambridge Univ. Engineering Dept., 2015.

## **JOURNAL PUBLICATIONS**

1. **C. Zhang**, Y. Liu, Y.-Q. Xia, X. Wang, C.-H. Lee, "Reliable accent specific unit generation with discriminative dynamic Gaussian mixture selection for multi-accent Chinese speech recognition," *IEEE Trans. on ASLP*, Vol. 21(10), pp. 2073-2084, 2013.
2. **C. Zhang**, Y. Liu, T. F. Zheng, "Acoustic model reconstruction for multi-accent Chinese speech recognition," *Journal of Tsinghua Univ. (Sci. & Tech.)*, Vol. 51, pp. 1161-1166, 2011. **(best paper)**
3. C. Wingfield, L. Su, X. Liu, **C. Zhang**, P. Woodland, A. Thwaites, E. Fonteneau, W.D. Marslen-Wilson, "Relating dynamic brain states to dynamic machine states: Human and machine solutions to the speech recognition problem," *PLOS Computational Biology*, Vol. 13(9), pp. 1-25, 2017.

## **CONFERENCE PUBLICATIONS**

1. **C. Zhang**, P.C. Woodland, "High order recurrent neural networks for acoustic modelling", in *Proc. ICASSP*, 2018.
2. F.L. Kreyssig, **C. Zhang**, P.C. Woodland, "Improved TDNNs using deep kernels and frequency dependent grid-RNNs", in *Proc. ICASSP*, 2018.
3. **C. Zhang**, P.C. Woodland, "Joint optimisation of tandem systems using Gaussian mixture density neural network discriminative sequence training", in *Proc. ICASSP*, 2017.
4. P. Lanchantin, M.J.F. Gales, P. Karanasou, X. Liu, Y. Qian, L. Wang, P.C. Woodland, **C. Zhang**, "Selection of multi-genre broadcast data for the training of automatic speech recognition systems", in *Proc. Interspeech*, 2016.
5. C. Wingfield, L. Su, B. Devereux, X. Liu, **C. Zhang**, P. Woodland, E. Fonteneau, A. Thwaites, W. Marslen-Wilson, "Multi-level representations in speech processing in brain and machine: evidence from EMEG and RSA", in *Cambridge Language Science Symposium*, 2016. **(best poster shortlist)**
6. C. Wingfield, L. Su, X. Liu, **C. Zhang**, P. Woodland, A. Thwaites, E. Fonteneau, W. Marslen-Wilson, "Investigating human speech recognition: reverse-engineering the machine solution with EMEG and RSA", in *Proc. HBM*, Geneva, 2016.
7. P. Lanchantin, M.J.F. Gales, P. Karanasou, X. Liu, Y. Qian, L. Wang, P.C. Woodland, **C. Zhang**, "DNN speaker adaptation using parameterised sigmoid and ReLU hidden activation functions", in *Proc. ICASSP*, 2016.
8. **C. Zhang**, P.C. Woodland, "DNN speaker adaptation using parameterised sigmoid and ReLU hidden activation functions", in *Proc. ICASSP*, 2016.
9. J. Yang, **C. Zhang**, A. Ragni, M.J.F. Gales, P.C. Woodland, "System combination with log-linear models", in *Proc. ICASSP*, 2016.
10. L. Wang, **C. Zhang**, P.C. Woodland, M.J.F. Gales, P. Karanasou, P. Lauchantin, "Improved DNN-based segmentation for multi-genre broadcast audio", in *Proc. ICASSP*, 2016.
11. P.C. Woodland, X. Liu, Y. Qian, **C. Zhang**, M.J.F. Gales, P. Karanasou, P. Lanchantin, L. Wang, "Cambridge university transcription systems for the multi-genre broadcast challenge", in *Proc. ASRU*, 2015.

12. P. Lanchantin, M. Gales, P. Karanasou, X. Liu, Y. Qian, L. Wang, P. Woodland, **C. Zhang**, "The development of the Cambridge university alignment system for the multi-genre broadcast challenge", in *Proc. ASRU*, 2015.
13. P. Karanasou, M.J.F. Gales, P. Lanchantin, X. Liu, Y. Qian, L. Wang, P.C. Woodland, **C. Zhang**, "Speaker diarisation and longitudinal linking in multi-genre broadcast data", in *Proc. ASRU*, 2015.
14. R.C. van Dalen, J. Young, H. Wang, A. Ragni, **C. Zhang**, M.J.F. Gales, "Structured discriminative models using deep neural-network features", in *Proc. ASRU*, 2015. **(best paper candidate)**
15. X. Liu, F. Flego, L. Wang, **C. Zhang**, M. Gales, P. Woodland, "The Cambridge university 2014 BOLT conversational telephone Mandarin Chinese LVCSR system for speech translation", in *Proc. Interspeech*, 2015.
16. **C. Zhang**, P.C. Woodland, "Parameterised sigmoid and ReLU hidden activation functions for DNN acoustic modelling", in *Proc. Interspeech*, 2015.
17. **C. Zhang**, P.C. Woodland, "A general artificial neural network extension for HTK", in *Proc. Interspeech*, 2015.
18. H. Wang, A. Ragni, M.J.F. Gales, K.M. Knill, P.C. Woodland, **C. Zhang**, "Joint decoding of tandem and hybrid systems for improved keyword spotting on low resource languages", in *Proc. Interspeech*, 2015.
19. **C. Zhang**, P.C. Woodland, "Standalone training of context-dependent deep neural network acoustic models", in *Proc. ICASSP*, 2014. **(best paper)**
20. K.M. Knill, M.J.F. Gales, S.P. Rath, P.C. Woodland, **C. Zhang**, S.-X. Zhang, "Investigation of multi-lingual deep neural networks for spoken term detection," in *Proc. ASRU*, 2013.
21. **C. Zhang**, Y. Liu, Y. Xia, C.-H. Lee, "Discriminative dynamic Gaussian selection with enhanced robustness and performance for multi-accent speech recognition," in *Proc. ICASSP*, 2012.
22. **C. Zhang**, Y. Liu, C.-H. Lee, "Detection-based accented speech recognition using articulatory features," in *Proc. ASRU*, 2011. **(panel member paper)**
23. **C. Zhang**, Y. Liu, Y. Xia, T. F. Zheng, J. Olsen, J. Tian, "Reliable accent specific unit generation with dynamic Gaussian mixture selection for multi-accent speech recognition," in *Proc. ICME*, 2011. **(top 15% rated paper)**
24. **C. Zhang**, Y. Liu, T. F. Zheng, "Asymmetric acoustic model for accented speech recognition," in *Proc. APSIPA ASC*, 2011.
25. J. Hou, Y. Liu, **C. Zhang**, S. Huang, "An in-car Chinese noise database for speech recognition," in *Proc. IALP*, 2011.

## **RECENT AWARDS & GRANTS**

Research funding, Cambridge Univ. Engineering Dept. HTK Project	2017 - present
Research funding, EPSRC Natural Speech Technology Project	2013 - 2016
IEEE ASRU 2015, best paper candidate	2015
Research funding, DARPA BOLT Program	2015
IEEE ICASSP 2014, IBM spoken language processing best student paper award (about 5/600)	2014
Research funding, iARPA Babel Program	2013
Cambridge International Scholarship, Cambridge Overseas Trust (about 1/30)	2012
Excellent master graduate award, Dept. of Computer Sci. & Tech., Tsinghua Univ.	2011
Excellent academy award, State Key Lab. of Intelligent Tech. & Sys., Tsinghua Univ. (17/170)	2011
Google excellence scholarship (2/180)	2011
NCMMSC 2011, best student paper award (3/96)	2011
IEEE ASRU 2011, panel member paper award (19/97)	2011
IEEE ASRU 2011, student travel grant	2011
IEEE ICME 2011, top 15% rated paper	2011
IEEE ICME 2011, student travel grant	2011
Games of the XXIX Olympiad, outstanding volunteer (6/40)	2008

## **INVITED TALKS**

- State-of-the-art deep learning for acoustic modelling — an HTK perspective** 2018.2
- o AI Platform & Research, JD.com
- Tandem and hybrid speech recognition systems based on a generic ANN extension in HTK** 2016.3
- o Center for Speech and Language Technologies, Tsinghua Univ.
- HTK 3.5 based DNN system development and research** 2015.12
- o Microsoft Speech Recognition Team
- CRF models and their applications in stock market value prediction** 2011.4
- o Finance Engineering Group, Asset Management Department, CICC

## **SELECTED PROJECTS**

Deep Neural Networks for HTK (Cambridge Univ. Engineering Dept. funded)	2017 - present
Natural Speech Technology (EPSRC & BBC funded)	2013 - 2016
DARPA BOLT Program	2014
iARPA BABEL Program	2013
State Scholarship Fund of China Scholarship Council, 1010811231	2010 - 2012
Natural Science Foundation of China, 60975018	2009 - 2011
Telephony Bank 95533, China Construction Bank	2009 - 2010
Joint Research Grant of Nokia-Tsinghua Joint Funding	2008 - 2010

## **SELECTED SOFTWARE**

The hidden Markov model toolkit (HTK) V3.5, open source software	2015
Intelligent pronunciation to written character converter, software copyright (2010SR026836)	2010
iWallet - a remote financial management software, open source software	2008

## **SELECTED SERVICE**

Reviewer, IEEE/ACM Transactions on Audio, Speech, and Language Processing	2014 - present
Reviewer, ISCA Computer Speech & Language	2017 - present
Reviewer, APSIPA Transactions on Signal and Information Processing	2014
Sponsorship division officer, Chinese Students & Scholars Association in Cambridge	2012 - 2013
External reviewer, 11th Annual Conference of ISCA (Interspeech 2010)	2010
Reporters' standing seat assistant, Games of the XXIX Olympiad	2008
Reporters' standing seat assistant, The XXIX Olympic Synchronised Swimming Qualification	2008
TV broadcasting assistant, The 16th FINA Diving World Cup	2008

## **TEACHING EXPERIENCE**

Demonstration, Speech recognition (MLSALT2), Cambridge Univ.	2017
Demonstration, Speech recognition (MLSALT2), Cambridge Univ.	2016
Demonstration, Speech recognition (MLSALT2), Cambridge Univ.	2015
Teaching assistant, The principle of signal processing, Tsinghua Univ.	2010
Teaching assistant, C++ programming design and training, Tsinghua Univ.	2010

## **PROFESSIONAL SKILLS**

- o Assembly Language, C, C++, Java, C#, Scheme, Python, and Perl programming
- o Signal, Speech, Audio, and Language processing
- o Pattern Recognition, Machine Learning, and Data Mining applications