ZHUO CHEN

414, W. 120 street Apt 301, New York, NY,10027
(646)-465-1394, zc2204@columbia.edu,
www.ee.columbia.edu/~zhuo

HIGHLIGHTS

- Star Researcher with 6+ years of deep learning and machine learning on speech/audio processing and recognition with double-digit publications and patents.
- Proven track-record on winning competitions and deliver products
  - State-of-the-art far-field and text-independent speaker recognition as Microsoft Intern
  - 1st and 2nd best performance, MIREX, task of singing voice separation, 2016
  - 2nd best performance, the 3rd CHiME challenge, 2015
  - Winner, IARPA-ASpIRE challenge, 2015
- Revolutionary single-microphone speech separation algorithm for the "cocktail party" problem
- Python master and AI evangelist

PROFESSIONAL EXPERIENCE

MICROSOFT, BELLEVUE, WA
Research intern, June 2016 – September 2016

- Delivered a far-field speech recognition and text-independent speaker recognition system

JELENIK SPEECH & LANGUAGE TECHNOLOGIES WORKSHOP, SEATTLE, WA
Far-field speech recognition team member, July 2015 – August 2015

- Developer and Researcher focusing on the speech recognition in far-field and mis-matched condition.

MITSUBISHI ELECTRIC RESEARCH LABORATORIES, CAMBRIDGE, MA
Research intern, November 2014 – June 2015

- Invented the state-of-the-art Semi-supervised audio source separation algorithm for a single microphone

- Designed and implemented a single-microphone robust automatic speech recognition and enhancement system

- Deep learning with Long Short-term Memory neural network

INTERNATIONAL COMPUTER SCIENCE INSTIDUTE, BERKELEY, CA
Research assistant, June 2013 – June 2014

- Researcher of robust front end for multi-language automatic speech recognition and speech enhancement

COLUMBIA UNIVERSITY, NEW YORK, NY
Research assistant, September 2012 – present

- Research & develop a series of audio signal processing algorithms and systems, which lead to patents and publications in top conferences & journals.

EDUCATION

COLUMBIA UNIVERSITY, NEW YORK, NY
PhD in electrical engineering, September 2012 – December 2016(expected)

COLUMBIA UNIVERSITY, NEW YORK, NY
MS in electrical engineering, September 2010 – December 2011

XIAN JIAOTONG UNIVERSITY, XIAN, CHINA
BS in electrical engineering, September 2006 – July 2010

AWARDS AND HORNOR

- 1st and 2nd best performance, MIREX, task of singing voice separation, 2016
- 2nd best performance, the 3rd CHiME challenge, 2015
- Winner, IARPA-ASpIRE challenge, 2015
- Second Award, Chinese National Electronics Design Contest, 2009
- Third Award, Applications of Texas Instruments, 2009
- Siyuan Scholarship, 2008
- Nanrui Jibao Scholarship, 2007
- Freshman Scholarship in Xian Jiaotong University, 2007
- Excellent student in Electrical Engineering of Xian Jiaotong University, 2007

PUBLICATION

1. Yi Luo, Zhuo Chen, Jonathan Le Roux, John Hershey, Nima Mesgarani, *"Deep Clustering and Conventional Networks for Music Separation: Stronger Together"*, (submitted to ICASSP 2017)

2. Zhuo Chen, Yan Huang, Jinyu Li, Yifan Gong, *"Improving mask learning based speech enhancement system with restoration layers and residual connection"*, (submitted to ICASSP 2017)

3. Zhuo Chen, Yi luo, Nima Mesgarani, *"Deep attractor network for single-microphone speaker separation"*, (submitted to ICASSP 2017)

4. Yi Luo, Zhuo Chen, Daniel P.W Ellis, *"Deep Clustering For Singing Voice Separation"*, MIREX, task of Singing Voice Separation, 2016 (1st and 2nd performance)

5. Shi-Xiong Zhang, Zhuo Chen, Yong Zhao, Jinyu Li, Yifan Gong, *"End-to-End Attention based Text-Dependent Speaker Verification"*, in 2016 IEEE Workshop on Spoken Language Technology, San Diego, CA

6. Yusuf Isik, Jonathan Le Roux, Zhuo Chen, Shinji Watanabe, John R. Hershey, *"Single-Channel Multi-Speaker Separation Using Deep Clustering"*, in Proc. Interspeech, San Francisco, Sep 2016

7. Z. Chen, J. O'Sullivan, S. Sheth, G. Mckann, A. D. Mehta, N. Mesgarani, *"Neural decoding of attentional selection in multi-speaker environments without access to separated sources"*, in Neuroscience 2016, November, San Diego, CA

8. Tasha Negamine,Zhuo Chen, Nima Mesgarani, , *"Adaptation of Neural Networks Constrained by Prior Statistics of Node Co-Activations"*, in Proc. Interspeech, San Francisco, Sep 2016

9. John R. Hershey, Zhuo Chen, Jonathan Le Roux, Shinji Watanabe, Yusuf Isik , *"Deep clustering: Discriminative embeddings for segmentation and separation"*, in Proc. ICASSP, Shanghai, April 2016

10. T. Hori, Z. Chen, H. Erdogan, J. Hershey, J. Roux, V. Mitra, S. Watanabe, *"The Merl/sri System For The 3rd Chime Challenge Using Beamforming, Robust Feature Extraction, And Advanced Speech Recognition"*, in Proc. ASRU, Arizona, Dec 2015

11. Roger Hsiao, Jeff Ma, William Hartmann, Martin Karafiat, Frantisek Grezl, Lukas Burget, Igor Szoke, Jan Honza Cernocky, Shinji Watanabe, Zhuo Chen, Sri Harish Mallidi, Hynek Hermansky, Stavros Tsakalidis, Richard Schwartz, *"Robust Speech Recognition in Unknown Reverberant and Noisy Conditions"*, in Proc. ASRU, Arizona, Dec 2015

12. Z. Chen, S. Watanabe H. Erdogan, J. Hershey, *"Speech enhancement and recognition using multi-task learning of long short-term memory recurrent neural networks"*, in Proc. Interspeech, Dresden, Sep 2015

13. Z. Chen, B. McFee, D. Ellis , *"Speech enhancement and recognition using multi-task learning of long short-term memory recurrent neural networks"*, in Proc. Interspeech, Dresden, Sep 2015

14. D. Ellis and H. Satoh and Z. Chen, *"Detecting proximity from personal audio recordings"*, in Proc. Interspeech, Singapore, Sep 2014

15. Z. Chen, H. Papadopoulos, D. Ellis, *"Content-adaptive speech enhancement by a sparsely-activated dictionary plus low rank decomposition"*, in Proc. HSCMA, Nancy, May 2014

16. Z. Chen, D. Ellis, *"Speech enhancement by sparse, low-rank, and dictionary spectrogram decomposition***",** in Proc. 2013, Workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, 2013

17. Z. Chen, G. Grindlay, D. Ellis, *"Transcribing multi-instrument polyphonic music with transformed eigeninstrument whole-note templates", MIREX, task of Multiple Fundamental Frequency Estimation Tracking, 2012*

PATENT

1. Mesgarani, N., O'Sullivan, J., Chen, Zhuo, *"Neural decoding of attentional selection in multi-speaker environments without access to separated sources", Provisional Patent filed June 2016*

2. John Hershey, Jonathan Le Roux, Shinji Watanabe, Zhuo Chen, *"Method for distinguishing components of an acoustic signal",* US patent No. 9368110, 2016

INVITED TALKS

1. *Speech segmentation and separation with deep clustering,* Microsoft Corporation, Redmond, Aug. 2016

2. *Speech separation with neural network,* Ecole Normale Superieure*, Paris, Jun. 2016*

3. *Deep clustering: Discriminative embeddings for segmentation and separation,* Columbia neural network research seminar series, Nov. 2015

4. *Using and modifying CURRENNT for LSTM neural network,* Columbia neural network research seminar series, Jul. 2015

5. *Segmentation and separation with discriminative model using recurrent deep neural network*, Mitsubishi electric research laboratories seminar series, Jun. 2015

6. *Speech enhancement by low-rank and convolutive dictionary spectrogram decomposition*, Interspeech 2014, the 15th annual conference of the international speech communication association, Singapore, Sep 2014

7. *Detecting proximity from personal audio recordings*, Interspeech 2014, the 15th annual conference of the international speech communication association, Singapore, Sep 2014

8. *Content-adaptive speech enhancement by a sparsely-activated dictionary plus low rank decomposition*, The 4th joint workshop on Hands-free Speech Communication and Microphone Arrays , Nancy, May 2014

9. *Speech enhancement by sparse, low-rank, and dictionary spectrogram decomposition*, The 17th workshop on Applications of Signal Processing to Audio and Acoustics, New Paltz, NY, 2013