

Mirco Ravanelli

PhD Candidate

Fondazione Bruno Kessler (FBK)

University of Trento

"Imagination is more important than knowledge"

Albert Einstein



CONTACTS

mrvanelli@fbk.eu

+393408174323

Via della Canova 35
38121 Trento- ITALY



26/02/1986

PHD TOPIC

Most state-of-the-art speech recognizers provide a satisfactory performance only in close-talking scenarios, forcing the users to speak very close to a microphone-equipped device.

Considering the growing interest towards speech recognition and the progressive use of this technology in everyday lives, it is easy to predict that in the future users will prefer to relax the constraint of handling or wearing any device to access speech recognition services, requiring technologies able to cope with distant-talking interactions also in challenging acoustic environments characterized by non-stationary noises and reverberation.

A prominent limitation of most of the existing distant-talking solutions is that such systems are still based on several cascade building blocks (e.g., speech enhancement, speech activity detection and acoustic models) which are typically rather independent with a poor matching and communication between them.

To overcome such a limitation, a promising research direction which I'm investigating concerns the adoption of a "network of deep neural networks", where all the basic elements of the system are implemented with deep neural networks, are fully connected, and cooperate together to achieve a single global goal.

My efforts are thus focused not only to study proper neural network architectures, but also on devising novel learning algorithms and jointly training strategies, which can be more suitable for the addressed scenario.

RESEARCH INTERESTS

- Distant-Talking Speech Recognition in Noisy and Reverberant Environments
- Deep Neural Networks and Deep Learning
- Multi-Microphone Signal Processing based on Distributed Microphone Networks and Microphone Arrays
- Robust Acoustic Scene Analysis in Adverse Conditions
- Methods for Impulse Response Measurement and Multi-Microphones Data Contamination
- Acoustic Event Detection for Large Scale Video Event Classification

PUBLICATIONS

- M. Ravanelli**, L. Cristoforetti, R. Gretter, M. Pellin, A. Sosi, M. Omologo, "The DIRHA-English corpus and related tasks for distant-speech recognition in domestic environments", accepted at ASRU 2015.
- M. Ravanelli**, B.Elizalde, G. Friedland, J. Bernd, "Insights into Audio-Based Multimedia Event Classification with Neural Networks", accepted at ACM@MMCOMMONS 2015.
- M. Ravanelli**, M. Omologo, "Contaminated speech training methods for robust DNN-HMM distant speech recognition", in Proceedings of INTERSPEECH 2015, Dresden.
- E. Zwysig, **M. Ravanelli**, P. Svaizer, M. Omologo, "A multi-channel corpus for distant-speech interaction in presence of known Interferences", in Proceedings of ICASSP 2015, Brisbane, Australia.
- M. Ravanelli**, M. Omologo, "On the selection of the impulse responses for distant-speech recognition based on contaminated speech training", in Proceedings of INTERSPEECH 2014, Singapore.
- M. Matassoni, R. Astudillo, A. Katsamanis, **M. Ravanelli**, "The DIRHA-GRID corpus: baseline and tools for multi-room distant speech recognition using distributed microphones", in Proceedings of INTERSPEECH 2014, Singapore.
- A. Brutti, **M. Ravanelli**, M. Omologo, "SASLODOM: Speech Activity detection and Speaker Localization in DOMestic environments", Proceedings of Evalita 2014, Pisa, Italy.
- M. Ravanelli**, V.H. Do, A. Janin, "TANDEM-Bottleneck Feature Combination using Hierarchical Deep Neural Networks", in Proceedings of ISCSLP 2014, Singapore.
- M. Ravanelli**, B. Elizalde, K. Ni, G. Friedland, "Audio Concept Classification with Hierarchical Deep Neural Networks", in Proceeding of the European Signal Processing Conference, EUSIPCO 2014, Lisbon, Portugal.
- B. Elizalde, **M. Ravanelli**, K. Ni, D. Borth, G. Friedland, "Audio-Concept Features and Hidden Markov Models for Multimedia Event Detection", in Proceedings SLAM 2014, Penang, Malaysia.
- A. Brutti, **M. Ravanelli**, P. Svaizer, M. Omologo, "A speech event detection and localization task for multiroom environments", in Proceedings HSCMA 2014, Nancy, France.
- L. Cristoforetti, **M. Ravanelli**, M. Omologo, A. Sosi, A. Abad, M. Hagmueller, P. Maragos, "The DIRHA simulated corpus", in Proceedings of LREC 2014, Reykjavik, Iceland.
- A. Sosi, F. Brugnara, L. Cristoforetti, M. Matassoni, **M. Ravanelli**, M. Omologo, "Embedding speech recognition to control lights", in Proceedings of INTERSPEECH 2013, Lion, France.
- B. Elizalde, **M. Ravanelli**, G. Friedland, "Audio Concept Ranking for Video Event Detection on User-Generated Content", in Proceedings of SLAM 2013, Marseille, France.
- M. Ravanelli**, A. Sosi, M. Matassoni, M. Omologo, M. Benetti, G. Pedrotti "Distant Talking Speech Recognition in Surgery Room : the DOMHOS project", in Proceedings of AISV 2013, Venice, Italy. (Awarded)
- M. Ravanelli**, A. Sosi, P. Svaizer, M.Omologo, "Impulse response estimation for robust speech recognition in a reverberant environment", in Proceeding of the European Signal Processing Conference, EUSIPCO 2012, Bucharest, Romania.

CAREER HISTORY

PhD STUDENT



Fondazione Bruno Kessler (FBK)
University of Trento
November 2013 – Present



After two years of research at the Bruno Kessler Foundation (FBK) and after an experience at the International Computer Science Institute (ICSI-Berkeley), I have decided to undertake a PhD on Distant-Talking Speech Recognition with Deep Neural Networks (DNNs). In the context of my PhD I'm also collaborating with ICSI to develop a video information retrieval system based on deep learning techniques.

Advisor: Maurizio Omologo (*omologo@fbk.eu*)

VISITING RESEARCHER



I.C.S.I (University of California, Berkeley)
January 2013 – April 2013
www.icsi.berkeley.edu

During my experience in ICSI, I had the opportunity to directly work with prof. Nelson Morgan, a pioneer of the Hybrid approaches for ARS. The topic of my work was the exploration of Pre-Trained Deep Neural Networks (DNNs) for Large Vocabulary Speech Recognition in the context of Babel-Swordfish project. In particular, I develop a multi-stream Hierarchical Deep Neural Network which is able to significantly outperform traditional speech recognition baselines.

Advisor: Nelson Morgan (*morgan@icsi.berkeley.edu*)

RESEARCHER



Fondazione Bruno Kessler (FBK)
November 2011 – November 2013
www.fbk.eu

I was involved in the European Project DIRHA (<http://dirha.fbk.eu/>) which aimed to study and develop a multi-microphone distant-talking speech recognition system in a domestic environment. In particular my research was focused on the development of robust acoustic models effective in challenging environment characterized by the presence of both non-stationary noises and reverberation. I was also involved in the DomHos project developing a distant-talking ASR system in an operating theater.

Advisor: Maurizio Omologo (*omologo@fbk.eu*)

AWARD



AISV – Italian Conference of Speech Science

January 2014

www.aisv.it

I won the “Franco Ferrero” award for the best paper at *AISV 2013* with the following paper: “Distant Talking Speech Recognition in Surgery Room : the DOMHOS project” . The award ceremony has been held in Turin on January 23 during the AISV 2014 conference.

EDUCATION

Master Degree in Telecommunications



University of Trento

October 2006 – February 2011

www.unitn.it

Mark: Full marks and Honors (110/110 cum laude)

Thesis Title: “A Multi-Microphone Front-End for Speech Transcription in Court Environment”

Advisors: M. Omologo (FBK) , F. Brugnara (FBK), F. De Natale (UNITN)

Description: The thesis engages with the transcription of the speech in court environment, a challenging context for the state of the art speech recognition technologies. In particular, the work proposes the introduction of a distributed microphone network based on triangular-arrays located near each speaker able to perform an appropriate acoustic scene analysis as well as a robust speech enhancement.

SKILLS

- *Programming Languages:* C/C++ , Perl, Shell Scripting, Python, XML
- *Software/Libraries:* MATLAB/OCTAVE (Advanced knowledge), Kaldi, HTK, TNET, Caffe, Quicknet, CNTK, CuDNN, Theano.
- *Languages:* Italian (Native), English (Fluent)
- *Personal Abilities:* good team worker, problem solving oriented, critical thinking
- *Hobbies:* soccer, cycling, swimming, trekking, reading, pet lover