

Human Interface Laboratory



Masahide Sugiyama
Professor



Jie Huang
Senior Associate Professor



Konstantin Markov
Senior Associate Professor

Summary of Achievement

Refereed proceedings of an academic conference

[j-huang-304-012-01:2016] Y. Nagayama, A. Saji, and J. Huang. Distance factor for frontal sound localization with side loudspeakers. *sep*, page = 2016.

-

[j-huang-304-012-02:2016] R. Kaneta, A. Saji, and J. Huang. 3D sound localization system using two side loudspeaker matrices. *sep*, page = 2016.

-

[markov-304-012-01:2016] T.Matsui J.Yu, K.Markov. Articulatory and Spectrum Features Integration using Generalized Distillation Framework. In *IEEE Int. Workshop on Machine Learning for Signal Processing*, 2016.

It has been shown that by combining the acoustic and articulatory information significant performance improvements in automatic speech recognition (ASR) task can be achieved. In practice, however, articulatory information is not available during recognition and the general approach is to estimate it from the acoustic signal. In this paper, we propose a different approach based on the generalized distillation framework, where acoustic-articulatory inversion is not necessary. We trained two DNN models: one called 'teacher' learns from both acoustic and articulatory features and the other one called 'student' is trained on acoustic features only, but its training process is guided by the 'teacher' model and can reach a better performance that can't be obtained by regular training even without articulatory feature inputs during test time

[markov-304-012-02:2016] T.Matsui K.Markov. Robust Speech Recognition using Generalized Distillation Framework. In *Proc. Interspeech*, pages 2364–2368, 2016.

In this paper, we propose a noise robust speech recognition system built using generalized distillation framework. It is assumed that during training, in addition to the training data, some kind of 'privileged' information is available and can be used to guide the training process. This allows to obtain a system which at test time outperforms those built on regular training data alone. In the case of noisy speech recognition task, the privileged information is obtained from a model, called 'teacher', trained on clean speech only. The regular model, called 'student', is trained on noisy utterances and uses teacher's output for the corresponding clean utterances. Thus, for this framework a parallel clean/noisy

speech data are required. We experimented on the Aurora2 database which provides such kind of data. Our system uses hybrid DNN-HMM acoustic model where neural networks provide HMM state probabilities during decoding. The teacher DNN is trained on the clean data, while the student DNN is trained using multi-condition (various SNRs) data. The student DNN loss function combines the targets obtained from forced alignment of the training data and the outputs of the teacher DNN when fed with the corresponding clean features. Experimental results clearly show that distillation framework is effective and allows to achieve significant reduction in the word error rate

[markov-304-012-03:2016] Jianguo Yu Konstantin Markov. Deep learning based personality recognition from Facebook status updates. In *8th International Conference on Awareness Science and Technology*, 2017.

Major Conference

Advisor for undergraduate research and graduate research

[j-huang-304-012-03:2016] Yuzuru Nagayama. Master Thesis: Frontal Sound Localization with the Factor of Distance of Side Loudspeakers, University of Aizu, 2016.

Thesis Advisor: Huang, J.

[j-huang-304-012-04:2016] Ryo Kaneta. Master Thesis: A 3D Sound Localization System Using Two Side Loudspeaker Matrices, University of Aizu, 2016.

Thesis Advisor: Huang, J.

[j-huang-304-012-05:2016] Yuya Sato. Graduation Thesis: Frontal sound localization by two side loudspeakers, University of Aizu, 2016.

Thesis Advisor: Huang, J.

[j-huang-304-012-06:2016] Yuya Ohno. Graduation Thesis: Factors of frontal sound localization by headphones, University of Aizu, 2016.

Thesis Advisor: Huang, J.

[j-huang-304-012-07:2016] Chiaki Watanabe. Graduation Thesis: Elevation perception of side sound images by two horizontal side loudspeakers, University of Aizu, 2016.

Thesis Advisor: Huang, J.

Summary of Achievement

[j-huang-304-012-08:2016] Yoshihiko Sato. Graduation Thesis: A 3D sound system with two side loudspeaker matrices, University of Aizu, 2016.

Thesis Advisor: Huang, J.