

Research highlights

New speech recognition model: Hidden Conditional Neural Fields

Toyohashi Tech researchers propose the Hidden Conditional Neural Fields (HCNF) model for continuous speech recognition. The model is a combination of the Hidden Conditional Random Fields (HCRF) and a Multi-Layer Perceptron (MLP), that is, an extension of Hidden Markov Model (HMM).

This new speech recognition model has the discriminative property for sequences from HCRF and the ability to extract non-linear features from an MLP. Furthermore, the HCNF can incorporate many types of features from non-linear features can be extracted, and is trained by sequential criteria.

In this paper, the researchers describe the formulation of HCNF and examine three methods to further improve automatic speech recognition using HCNF, which was an objective function that explicitly considered training errors, provided a hierarchical tandem-style feature, and included a deep non-linear feature extractor for the observation function.

HCRF can use a deep feed forward neural network (DNN) in the observation function, and therefore, a sophisticated pre-training algorithm such as the deep belief network (DBN) can be used to provide a deep observation function.

The research shows that HCNF can be trained realistically without any initial model and outperform the HCRF and triphone hidden Markov model trained by the minimum phone error (MPE) manner using experimental results for continuous English phoneme recognition on the TIMIT core test and Japanese phoneme recognition on the IPA 100 test set.

Reference:

Authors: Ysuhisa FUJII, Kazumasa YAMAMOTO, Seiichi NAKAGAWA

Title of original paper: Hidden Conditional Fields for Continuous Phoneme Speech Recognition

Journal, volume, pages and year: *IEICE Trans. Inf.&Sys.*, **E95-D**,2094-2104 (2012).(2012 IEICE best paper award)

Digital Object Identifier (DOI): 10.1587/transinf.E95.D.2094

Affiliations: Department of Computer Science and Engineering, Toyohashi University of Technology

Department <http://www.slp.cs.tut.ac.jp/index-e.html>



Seiichi Nakagawa

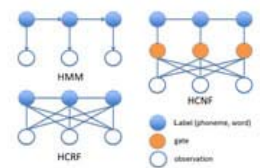


Figure caption: Model architectures of HMM, HCRF and HCNF