Demo of Idlak Tangle, An Open Source DNN-Based Parametric Speech Synthesiser

Blaise Potard[1], *Matthew P. Aylett*[1,2], *David A. Baude*[1]

[1]CereProc Ltd., United Kingdom
[2]The Centre for Speech Technology Research, University of Edinburgh, United Kingdom
{blaise,matthewa,dave}@cereproc.com

Abstract

We present a live demo of Idlak Tangle, a TTS extension to the ASR toolkit Kaldi [1]. Tangle combines the Idlak front-end and newly released MLSA vocoder, with two DNNs modelling respectively the units duration and acoustic parameters, providing a fully functional end-to-end TTS system. The system has none of the licensing restrictions of currently available HMM style systems, such as the HTS toolkit, and can be used free of charge for any type of applications.

Experimental results using the freely available SLT speaker from CMU ARCTIC, reveal that the speech output is rated in a MUSHRA test as significantly more natural than the output of HTS-demo.

The tools, audio database and recipe required to reproduce the results presented are fully available online at `https://github.com/bpotard/idlak`.

The live demo will allow participants to measure the quality of TTS output on several ARCTIC voices, and on voices created from commercial-grade recordings.

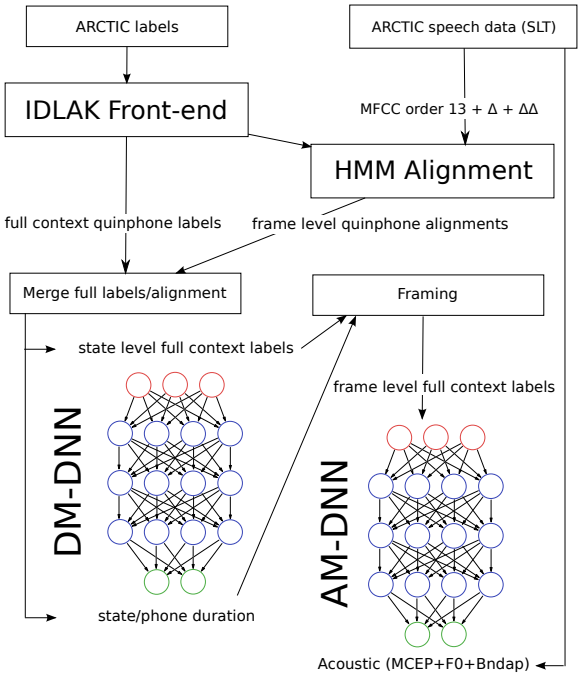

Figure 1: *Tangle DNN training architecture*

1. References

[1] D. Povey, A. Ghoshal, G. Boulianne, L. Burget, O. Glembek, N. Goel, M. Hannemann, P. Motlíček, Y. Qian, P. Schwarz, J. Silovský, G. Stemmer, and K. Veselý, "The Kaldi speech recognition toolkit," *Proc. IEEE ASRU*, 2011.