

File : README

AISHELL-2 Speech Data

北京希尔贝壳科技有限公司
Beijing Shell Shell Technology Co.,Ltd
03/10/2018

1. AISHELL-2 Speech Data

- Sampling Rate : 16kHz
- Sample Format : 16bit
- Environment : Indoor
- Speech Data Type : PCM
- Channel Number : 1
- Recording Equipment : iOS
- Time : 1000 hours of speech data (around 1 million utterances)
- Speaker : 1991 speakers (845 male and 1146 female)

2. Data Structure

Directory tree	
Corpus Catalog	
AISHELL2-Data-Specification .docx	Corpus Information
ChangeLog	Change Information
└─ data	Data File
─wav.scp	Audio list
─trans.txt	Transcript Content
─spk_info.txt	Speaker Information
─wav	Audio Data File
─D2166	Speaker ID File
ID2166W0001.wav	Audio

3. Evaluation Data:

Currently we open-sourced AISHELL2-2018A-EVAL

- dev: 2500 utterances from 5 speaker
- test: 5000 utterances from 10 speakers

You can download above evaluation set from:

<https://github.com/kaldi-asr/kaldi/tree/master/egs/aishell2>

4. RECIPE

Based on Kaldi standard system, AISHELL-2 provides a self-contained Mandarin ASR recipe

- a word segmentation module, which is a must-have component for Chinese ASR systems
- an open-sourced Mandarin lexicon(DaCiDian)
- a simplified GMM training recipe
- 80-dim FBank without pitch as NN input feature
- slightly different data augmentation setup(tempo perturbation, max-volume perturbation)
- acoustic channel adaptation recipe(AM fine-tuning)
- a real-time streaming ASR demo via laptop microphone (tested on MacOS)

You can download recipe set from: http://www.aishelltech.com/aishell_2