



语音合成：第七章作业讲评



主讲人 谢启聪



- 第一部分：数据准备
- 第二部分：自适应（声学模型+声码器）

➤ 第一部分：数据准备

➤ 第二部分：自适应（声学模型+声码器）

作业

-
- 从之前课程提供的标贝科技开源数据中的label中随机挑选50句左右的文本进行录音。
 - 声学模型部分和声码器部分的数据准备请参照之前的作业流程。

➤ 第一部分：数据准备

➤ 第二部分：自适应（声学模型+声码器）

作业

● 声学模型自适应

- 修改声学模型部分，
加入模型 *restore* 的相关代码。

```
56 global_saver = tf.train.Saver(tf.global_variables(), max_to_keep=hparams.keep_checkpoint_max)
57
58 log("Model training set to a maximum of {} steps".format(
59     hparams.total_training_steps))
60
61 # Config TensorFlow GPU options
62 config = tf.ConfigProto(allow_soft_placement=True)
63 config.gpu_options.allow_growth = True
64
65 restore_path_v1 = "log_dir/example/facotron"
66
67 init_ops = [tf.global_variables_initializer(),
68             tf.local_variables_initializer()]
69
70 with tf.Session(config=config) as sess:
71     try:
72         summary_writer = tf.summary.FileWriter(
73             tensorboard_dir, sess.graph)
74
75         if args.restore:
76             try:
77                 checkpoint_state = tf.train.get_checkpoint_state(restore_path_v1)
78                 if (checkpoint_state and checkpoint_state.model_checkpoint_path):
79                     global_saver.restore(sess, checkpoint_state.model_checkpoint_path)
80             except tf.errors.NotFoundError as e:
81                 assert False, e
82         else:
83             global_saver.save(sess, checkpoint_path, global_step=global_step)
84
85     trainreader.start_threads()
86     sess.run(init_ops)
```

- 在之前课程中训练好的声学模型的基础上，固定 *encoder* 部分参数，尝试对 *attention/decoder* 的部分参数或全部参数进行训练。

```
123
124 def add_optimizer(self, global_step):
125     def is_frozen(name, frozen_list):
126         return any([frozen in name for frozen in frozen_list])
127     if self.hparams.retrain_decoder_only:
128         frozen_list = ["phone_embedding", "EncoderCBHG", "EncoderPreNet"]
129     else:
130         frozen_list = []
131     self.vars = [var for var in tf.trainable_variables() if not is_frozen(var.name, frozen_list)]
132
```

● 声码器自适应

- 在之前课程中训练好的声码器的基础上，尝试对整体参数/部分参数进行`refine`，看看能否在保持声码器稳定的基础上改善音色的相似度。
- 和声学模型对接，验证自适应效果。



感谢各位聆听 !

Thanks for Listening

