

VCC 2020 baseline design

Sian-Yi Chen

Advisor : Tay-Jyi Lin and Chingwei Yeh

Outline

- Action item

- 進度太慢，盡速做出結果

- Status report

- 在 VCC2020 比賽中，語音轉換分為兩個任務

- 本周進度 (已重現任務一)

Task1：同種語言內轉換 (semiparallel VC)

- 在訓練語料中，僅存有少量共同語句，其餘不同 (訓練語料 source & target 各 70 句，共同語料 20 句)
- 轉換完成後，source speaker's voice 完全被轉換為 target speaker's voice (下一頁有轉換完成前後音檔示例)
- **Example sentence**：Moroccan agriculture enjoys special treatment when exporting to Europe.



Train source voice



Train target voice

- 下周規劃

Task2：跨語言轉換 (cross-lingual VC)

- 訓練語料完全不同
- 希望轉換完成後聲學特徵相同，且保持語料內容不變
- ASR 模型與 task1 相同，輸出英文文本後需轉換成 phoneme，並使用英文與目標語言 (中文) 訓練 TTS model

Task1 conversion process

ASR

- model : Transformer-based
- dataset : LibriSpeech

TTS

- model : multi-speaker, x-vector Transformer-TTS
- dataset : Libritts

Vocoder

- model : Parallel WaveGAN

*ASR result case conversion :

在 TTS 訓練資料中， Libritts 提供的是 normalized text

Ex : Tom, the Piper's Son

但 ASR 輸出結果皆為大寫，並且沒有標點符號

Ex : IN REALITY THE EUROPEAN PARLIAMENT IS PRACTISING DIALECTICS

因此作者將 ASR 結果全轉換為小寫，雖然會有部分不匹配，但作者認為是可接受的

Example sentence : In reality, the European Parliament is practising dialectics.

輸出入音檔：各 25 個音檔
(取其中一個音檔作為例子)

