

# Introduction to Audio Signal Processing

## *Notes on Speech and Audio Processing*

Chia-Ping Chen

Department of Computer Science and Engineering

National Sun Yat-Sen University

Kaohsiung, Taiwan ROC

# Where Do We Need Audio Processing

---

- Speech recognition: to communicate with machines with voice interface, we need to process the speech that a machine can recognize.
- Speech synthesis: to use computers to help us “speak”, we need to know how to produce the desired “sounds”.
- Speech coding: to save disk space or transmission bandwidth, we need to know how to encode speech efficiently without sacrificing intelligibility.

# Speech Recognition Is Hard

---

- There is no clear boundary between the words in speech.
- Speech varies in many ways.
- There is often corruption of speech by environmental noises.
- There are constraints in computational resource, such as for a mobile device.
- There is a big gap between the performance level of speech recognition by human and machines.

# Task Parameters of ASR

---

- speaker-dependent vs speaker-independent
- small vocabulary vs large vocabulary
- isolated speech vs continuous speech
- read speech vs spontaneous speech
- clean speech vs noisy speech

# Historical Notes

---

- 1920s: Radio Rex
- 1950s: IBM digit recognition system
- 1960s: spectral features, linear prediction, dynamic time warping for decoding, and HMM for acoustic modeling
- 1970s: task-oriented projects with definite objectives
- 1980s: collections of large speech databases (TIMIT, WSJ)
- 1990s: more efforts on both front-end and back-end, such as vocal tract normalization, speaker adaptations, gender-dependent models, etc.

# Recent Works

---

- More corpora
  - Broadcast News
  - Switchboard
  - Call Home
- More research on core technology
  - noise-robustness
  - pronunciation models
- More applications
  - speaker verification and identification
  - language identification

# The Handyman Robot

---

- **Robot:** I am handyman robot. Tell me a task, and I will do it for \$5 per hour. This money will be applied to further research in artificial intelligence.
- **Human:** \$5 dollar an hour? Sounds great! Can you paint?
- **Robot:** My painting is of the highest quality.
- **Human:** OK. See that paint brush and bucket of paint? Take them out and paint the porch.
- **Robot:** Your request will be fulfilled.

# The Handyman Robot

---

(The robot trundles off to do his job, and returns in an hour.)

- **Robot:** The task is completed. Please deposit \$5 to aid in further research.
- **Human:** (Handing over the cash) This was a great deal! Come back again!
- **Robot:** (While leaving) Oh, by the way, it wasn't a Porsche. It was a BMW.