

# イラストで学ぶ 音声認識



荒木雅弘 著  
講談社 2015

# 1. はじめに

1.1 音声認識とは

1.2 音声認識の歴史

1.3 どこで役立つのか

1.4 何が難しいのか

1.5 本書のねらい

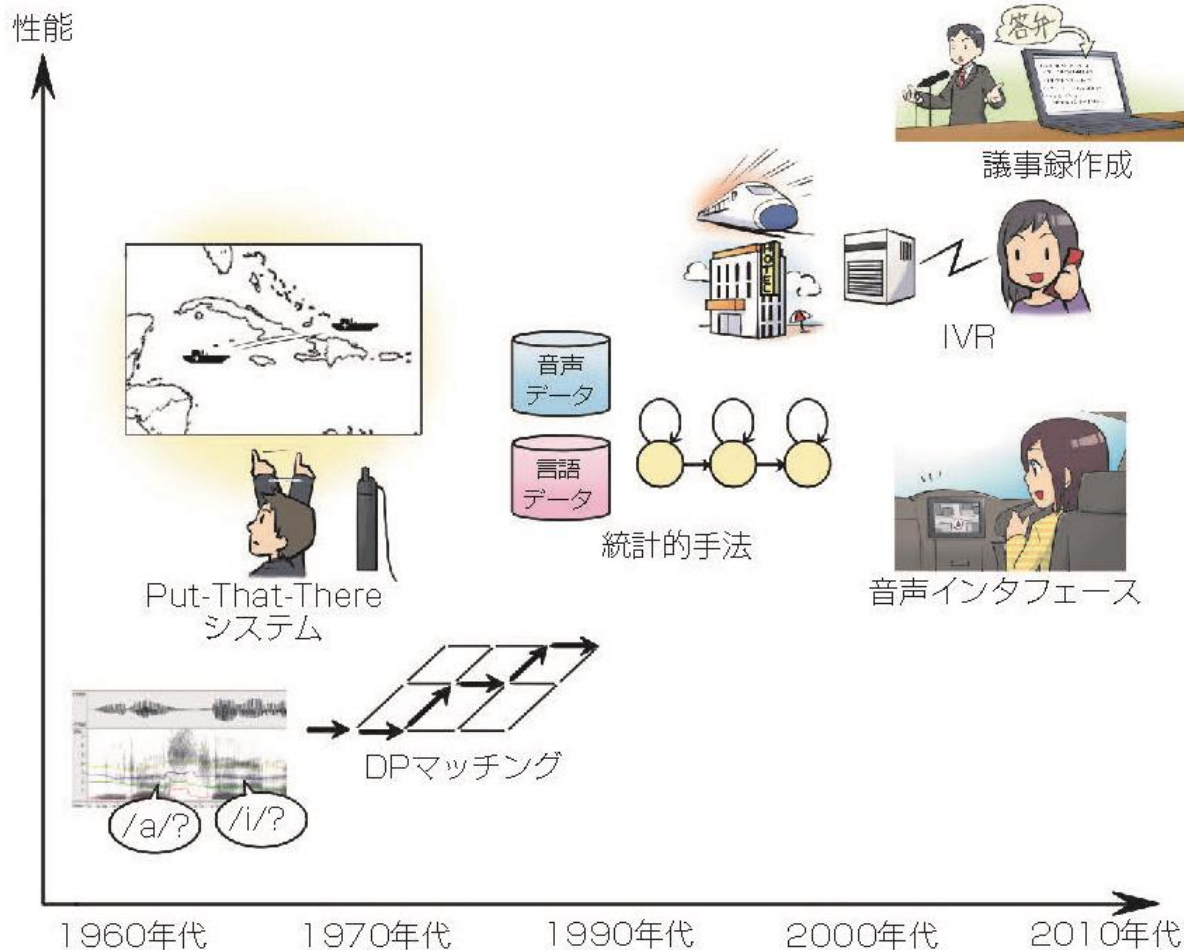
# 1.1 音声認識とは

- 音声認識の定義
  - 人間が声で話す言葉を文字に変換する技術
  - 機械可読な意味表現に変換するところまでを含む場合もある
- 音声認識における問題設定
  - 使用環境
  - 利用者
  - 単語数
  - 出力



図 1.1 音声認識の問題設定

# 1.2 音声認識の歴史



- 2015年以降
- ・ ロボットの音声  
インタフェース
  - ・ スマートスピーカー

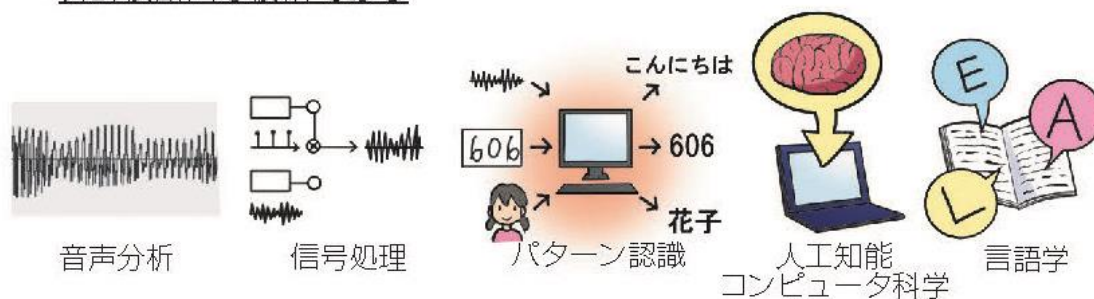
図 1.2 音声認識の歴史

# 1.3 どこで役立つのか

- ヒューマンインタフェースとして
  - 携帯端末（スマートフォン、タブレット）や家電・スマートスピーカーなどへの入力手段
  - カーナビなど、ハンズフリーが要求される状況
- 情報蓄積・検索
  - 会議の議事録作成
  - テレビ放送・動画への字幕付与
  - 音声検索

# 1.4 何が難しいのか

- ・ 分野横断的な技術である



- ・ パターン認識問題の中でも最も難しいクラスの問題である



- ・ 実用化のためには多くのノウハウが必要である



図 1.5 音声認識の難しさ

# 1.5 本書のねらい

- 音声認識技術の難しさへの対処

1. 分野横断的な技術である

- ➡ イラストによる直観的な理解

2. パターン認識問題の中でも最も難しいクラスの問題

- ➡ WFSTによる解決法の解説

3. 実用化の際にノウハウが必要

- ➡ ノウハウを学べるフリーソフトの紹介



# 本書の構成

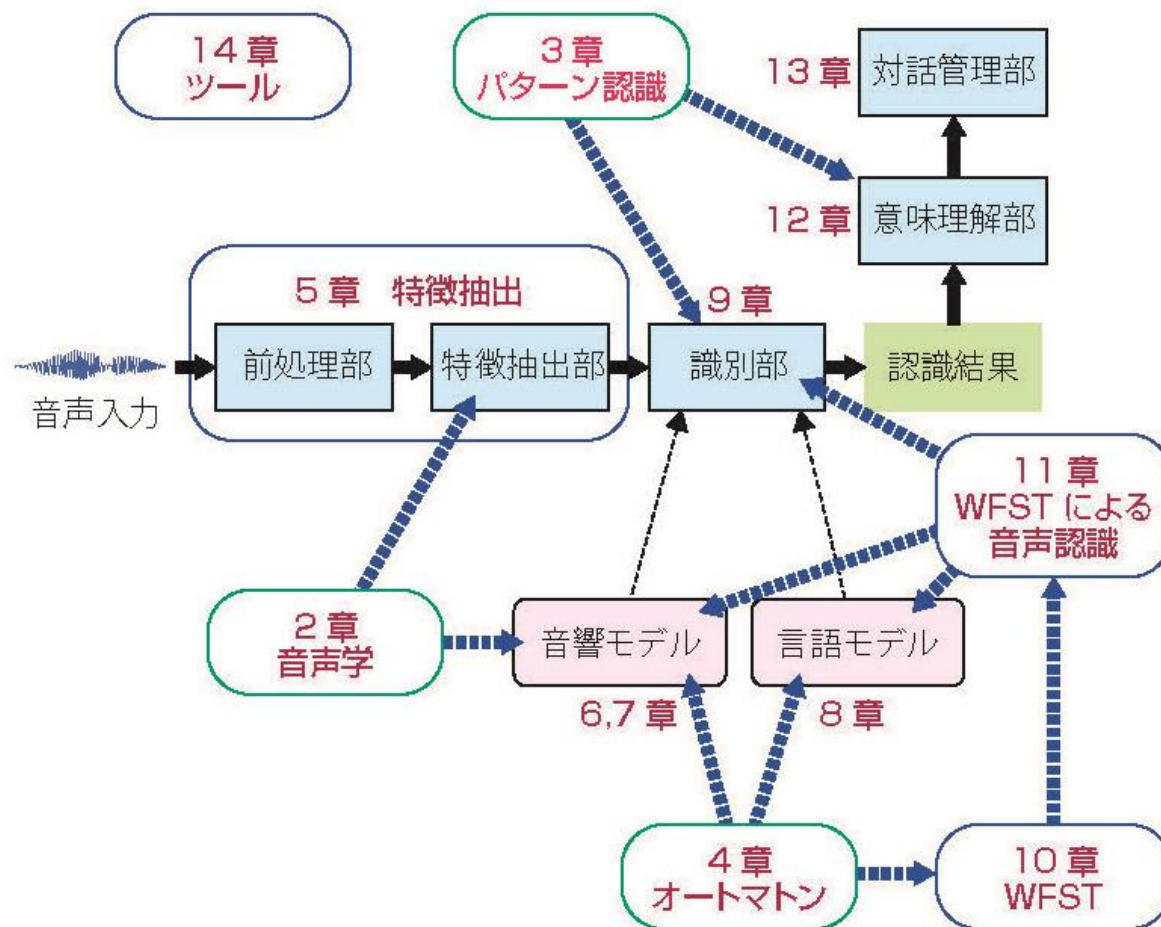


図 1.7 本書の章構成