

人間システム工学概論

人間とコンピュータの 音声対話システム

関西学院大学 理工学部
教授 川端 豪

人間システム工学概論

第8回

「人間とコンピュータの音声対話システム」

授業日程:

【情報科学科/人間システム工学科】

5/29/2018(Tue) 13:30-15:00

授業の進め方

- 授業時間内に講義と演習(試験)を交互に進める
- 開始時刻に教室を施錠し、解答用紙を配布する。
途中退室する場合は解答用紙を返却すること
- 「序論を講義(約20分)」
「講義(約15分)+演習(5分)」×3回
- 演習は試験に準ずる。答案は独自に作成すること。
会話厳禁。違反者には退室を命じることがある

Copyright © 2006-2018 by Takeshi Kawabata

授業内容

序論 音声対話システム

1. 音声認識
2. 音声合成
3. いろいろな応用

Copyright © 2006-2018 by Takeshi Kawabata

音声対話システム (1)

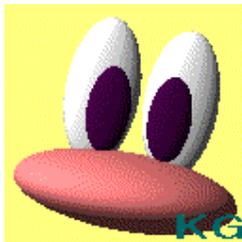
目標

■人間と音声で自然に会話するコンピュータを創り出す

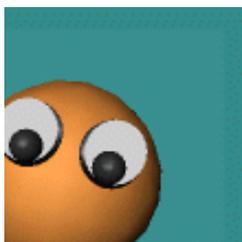
Copyright © 2006-2018 by Takeshi Kawabata

音声対話システム (2)

もう、ここまでできる



Mr. PINK



日経サテライトニュース (8/15/98)

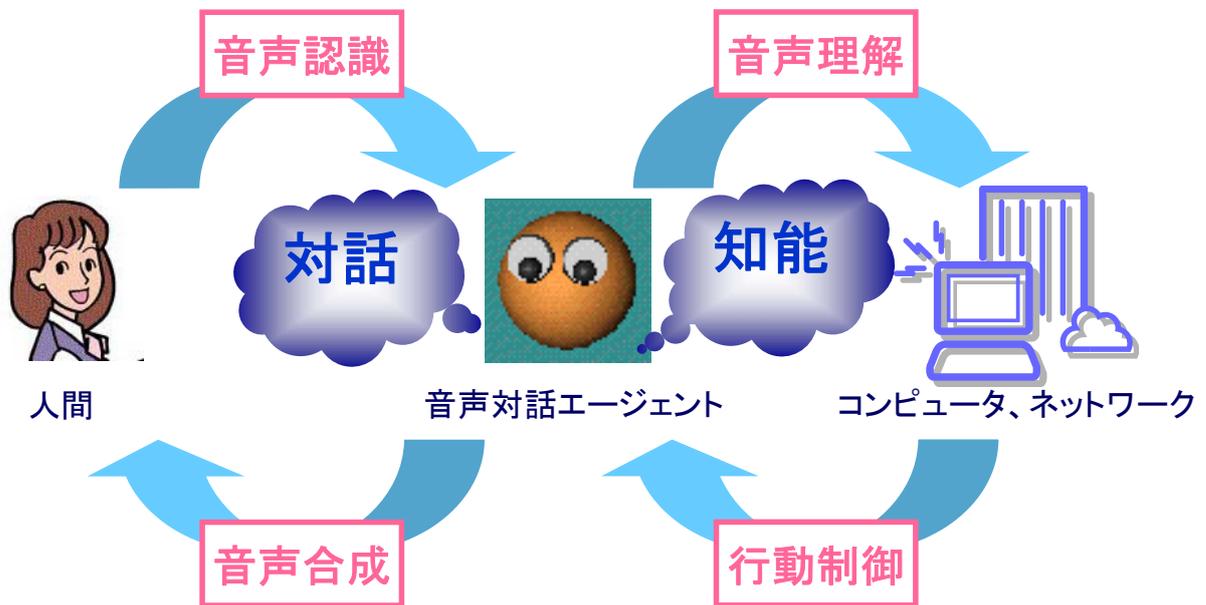


Sherry とおしゃべり



Copyright © 2006-2018 by Takeshi Kawabata

秘密は？⇒最先端技術の複合



Copyright © 2006-2018 by Takeshi Kawabata

1. 音声認識 (1)

■ 音声認識とは？

- コンピュータに人間の音声を「聞き取らせる」技術

■ 認識の単位

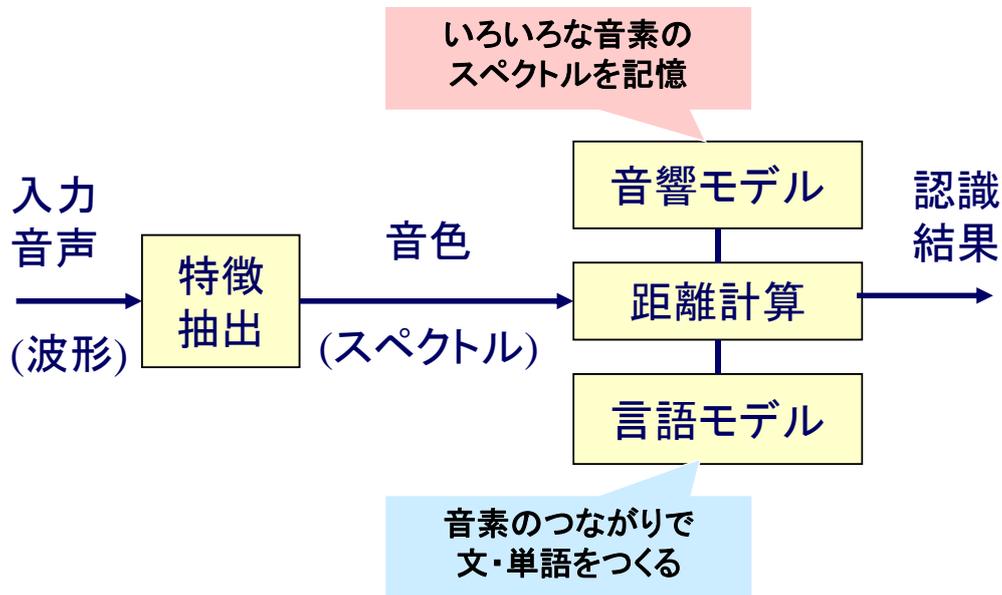
- 単語: 「あさひ」
- 音節: 「あ さ ひ」
- 音素: 「a s a h i」

音節、音素
が主流

Copyright © 2006-2018 by Takeshi Kawabata

1. 音声認識 (2)

音声認識のしくみ



Copyright © 2006-2018 by Takeshi Kawabata

1. 音声認識 (3)

■ 特徴抽出

- 入力された音声を
特徴ベクトルの時系列に変換する
⇒ 「スペクトル系列」

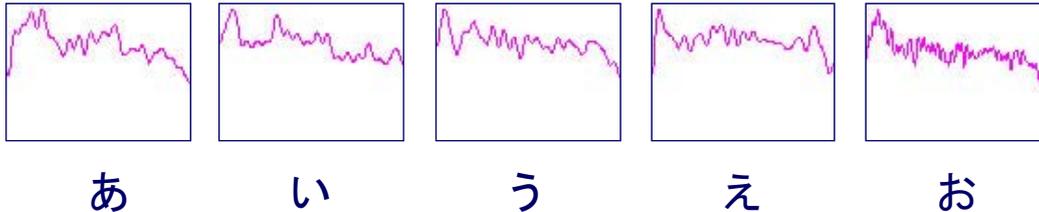


Copyright © 2006-2018 by Takeshi Kawabata

1. 音声認識 (4)

■音響モデル

– いろいろな音素のスペクトルを記憶



Copyright © 2006-2018 by Takeshi Kawabata

1. 音声認識 (5)

■音響モデルの役割

–

–



Copyright © 2006-2018 by Takeshi Kawabata

1. 音声認識 (6)

■ 言語モデル

- 音素の並び方が似ていても、まるで違う意味のことがある。

/k a w a r a n a i /

/k a o a r a w a n a i /

誤認識しやすい

Copyright © 2006-2018 by Takeshi Kawabata

1. 音声認識 (7)

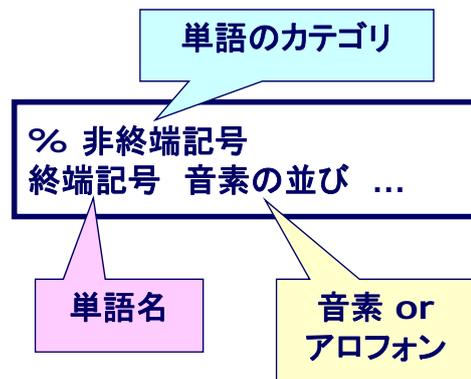
■ 言語モデル

- 状況に合わせて「語彙」と「文型」を制限
- 語彙の制限 ⇒ 辞書
- 文型の制限 ⇒ 文法、確率モデル

Copyright © 2006-2018 by Takeshi Kawabata

1. 音声認識 (8)

辞書による語彙の制限



% WORD	
一の関	ichinoseki
宇都宮	utsunomiya
浦佐	urasa
越後湯沢	echigoyuzawa
燕三条	tsubamesanjo
u	
岡山	okayama
角館	kakunodate
掛川	kakegawa
岐阜羽島	gifuhashima
京都	kyoto
熊谷	kumagai
郡山	kooriyama
軽井沢	karuizawa
古川	furukawa

...

Copyright © 2006-2018 by Takeshi Kawabata

1. 音声認識 (9)

文法による文型の制限

S ← NP
← ADJ NOUN
← 赤い 帽子
← akai boushi

S ← NP
NP ← NOUN
NP ← ADJ NOUN

% NOUN
帽子 boushi
服 fuku
...
% ADJ
赤い akai
青い aoi
...

Copyright © 2006-2018 by Takeshi Kawabata

1. 音声認識 (10)

■ 言語モデルの役割

–



Copyright © 2006-2018 by Takeshi Kawabata

2. 音声合成 (1)

■ 音声合成とは？

– コンピュータに人間の音声を「話させる」技術

■ 音声合成手法の大分類

–

–

Copyright © 2006-2018 by Takeshi Kawabata

2. 音声合成 (2)

■録音編集 (Pre-recorded voice)

- 前もって録音した音声波形を切り貼り
- 駅の構内アナウンス、自販機



Copyright © 2006-2018 by Takeshi Kawabata

2. 音声合成 (3)

■テキスト音声合成 (TTS: Text To Speech)

- 任意の文章を音声に変換する技術
- 電話自動応答、Web の読み上げなど



Copyright © 2006-2018 by Takeshi Kawabata

3. いろいろな応用 (1)

■川端研 卒研究生のとりくみ

- 仮想空間、お話し散歩 (Virtual KSC)
- ロボットと音声対話
- 話せばわかるカーナビ

Copyright © 2006-2018 by Takeshi Kawabata

3. いろいろな応用 (2)

■Virtual KSC

- 仮想空間中のKSCを、CGキャラクター Susie の案内で散歩
- 将来 Web 上で動作するようになれば、居ながらにして、関学キャンパスの美しさを体験



Copyright © 2006-2018 by Takeshi Kawabata

3. いろいろな応用 (3)

■ ロボットと音声対話

- 近い将来、ロボットは社会の一員に
- コミュニケーションはやはり「音声」
- 音で周囲の環境を把握する



Copyright © 2006-2018 by Takeshi Kawabata

3. いろいろな応用 (4)

■ 話せばわかるカーナビ

- 運転中の情報機器操作の安全確保
 - ⇒ 音声対話インタフェースが有望
- 会話による情報検索・意思決定



Copyright © 2006-2018 by Takeshi Kawabata

3. いろいろな応用 (5)

目標: 人間と音声で自然に会話するコンピュータを創り出す

仮想空間、お話し散歩
(Virtual KSC)



(市販のトイロボットへの実装写真)

ロボットと
音声対話

Hello!



イマジ
ネーショ
ンが決
めてだ

話せばわか
るカーナビ



Copyright © 2006-2018 by Takeshi Kawabata

3. いろいろな応用 (6)

■ 応用を考えるポイント

- 利用シーンを思い浮かべる
「イマジネーション」が決め手
- 誰の役に立つのか？
- 逆に困る人はいないか？

演習

Copyright © 2006-2018 by Takeshi Kawabata