



- ▶ 卒研(就職希望者)のみ指導。2024年度の研究テーマは未定→卒研配属希望学生と相談して決める。
- ▶ 卒研生の要件：UNIX系OSとC言語を使いこなせること。工学に限らない幅広い見識を持ち、毎日平均して一人で努力を積み重ねることが得意な人。
- ▶ 配属を希望する可能性が有る方は、必ず面接を受けてください。 **面談必須**

機械による音声言語の認識 Spoken Language Recognition by machine

2018年度まで取り組んでいたこと

言語認識

「何語を話している？」

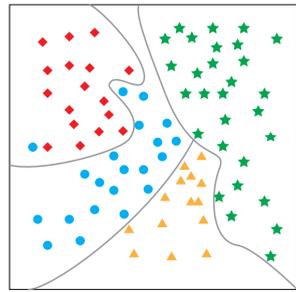
Language Recognition

Which language is being spoken?



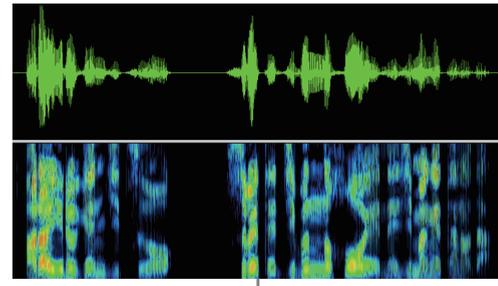
パターン認識

Pattern Recognition



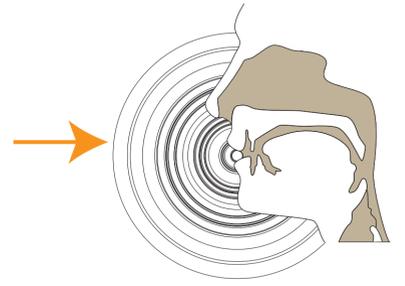
音声信号処理

Speech Signal Processing



音響音声学

Acoustic Phonetics



a i e o k s t n h m y r w
g d b p f t s ...

調音クラス articulatory class
母音 / 子音 摩擦 / 破裂 / 弾き / ...
円唇 / 非円唇 唇 / 歯 / 口蓋 / 声門 ...
vowel / consonant fricative / plosive / flap ...
round / unround, lips / dental / palatal / glottal ...

提案法 present method

スペクトル特徴を調音クラスに写像

mapping of spectrum featute to articulatory feature class

話者性に頑健、言語共通

roubust against speaker individuality, language independent

従来法 conventional method

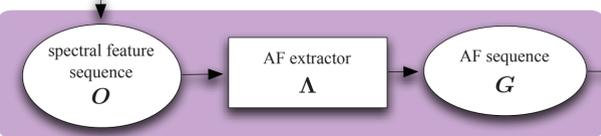
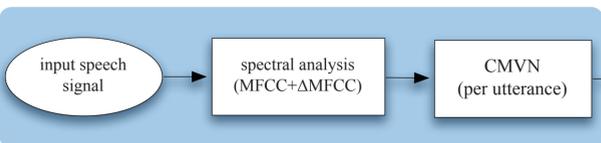
スペクトル特徴を言語音に写像

mapping of spectrum featute to language sound category

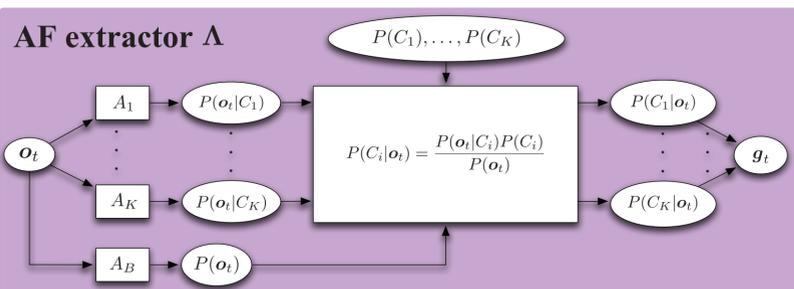
言語識別フローチャート

平田拓海, 高木一幸, "調音クラスの事後確率に基づく言語識別の検討 ~言語共通の調音特徴抽出器及び VQ 符号長について~," 電子情報通信学会技術研究報告, SP2014-169, pp.187--190, 2015年3月度音声研究会, 2015年3月3日.

signal processing



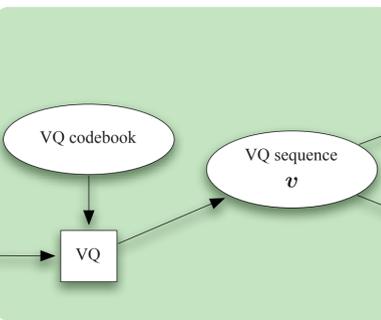
AF extraction



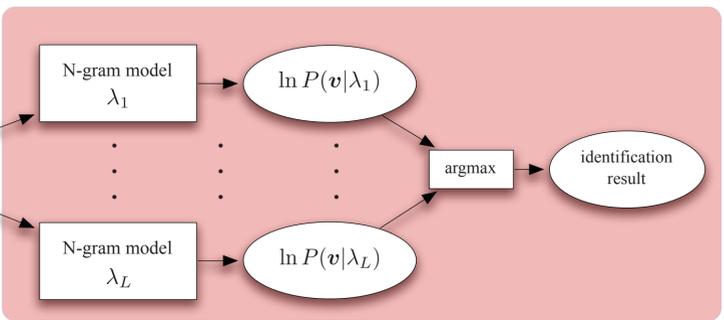
C_i ($i = 1, \dots, K$): 調音クラス
 $O = o_1 \dots o_T$: スペクトル特徴時系列
 T : フレーム数
 A_i : 調音クラス C_i の音響モデル
 A_B : 全学習データから学習した音響モデル (A_i, A_B として混合ガウス分布を用いた)

$P(o_t|C_i)$: 調音クラス C_i のクラス条件付き確率
 $P(C_i)$: 調音クラス C_i の事前確率
 $P(o_t)$: スペクトル特徴 o_t の生起確率
 $P(o_t|C_i)$: 調音クラス C_i の事後確率
 $g_t = (P(C_1|o_t), \dots, P(C_K|o_t))'$: AF ベクトル
 $G = g_1 \dots g_T$: AF 時系列

vector quantization



language identifaicaton



2023年度卒論

- ・アバターに合う音声を作るための音声分析合成を用いた感情不変リアルタイム音声合成システム

2021年度卒論

- ・歌詞専用の言語モデルと単語の読み類似度による分割した歌声の歌詞の特定
- ・音声特徴量の時系列情報を利用した音声感情認識の基礎検討

2017年度卒論

- ・調音クラスの事後確率に基づく言語識別法における多層順伝播ネットワークを用いた調音クラス認識率の改善