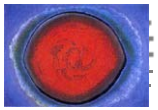
2008年度　奥乃・尾形研究室紹介

- 所属　大学院情報学研究科　知能情報学専攻
  知能メディア講座音　メディア分野



- 研究テーマ
  - 音とロボットに関する知能の解明
  - 「聞き分ける」,「音を見せる」
  - 「音声対話」
  - 「音・動作を模倣する」
- 方針
  - 学生の能力の発見と着実な成果
  - 挑戦的な研究テーマで新たな地平を切り拓く
  - 日本・世界の研究を牽引

2008年度の研究室メンバー

- 教　授　　奥乃 博　（2001年4月採用，企業・機関・東京理科大）
- 准教授　　尾形 哲也　（2003年10月採用, 早稲田大・理研）
- 助　教　　駒谷 和範　（2002年12月採用）
- GCOE助教　高橋 徹　（2008年2月採用, 和歌山大）
- 秘　書　　西井 美季
- 学振PD　　白松　俊
- 博士課程　D3　2名　（学振特別研究員1名、国費留学生1名）
  D1　2名　（学振特別研究員1名）
- 修士課程　M2　5名（留学生1名）
  M1　4名（留学生1名）
- 4回生　　1＋4名の予定
- ロボット　HRP-2, SIG, SIG-2, Robovie-IIs, Robovie-R2, Keepon, Replie, Latte, AIBO, ネコロ, パロ, ドラえもん

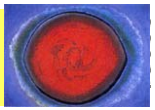
究極の目標：聖徳太子コンピュータ

- 人と共生するには混合音の「聞き分け」が重要
- あかちゃんのような発達論的なコミュニケーション機能の獲得
- 音響信号から記号の獲得
- 身体性の活用
- 人の耳はなぜ2つか

研究のテーマを分かりやすく言うと

1. 聞き分ける
   - 事前知識最小の「聞き分け」機能
   - ラピッドプロトタイピングによる「聞き分け」
   - 市販CD音楽，実環境下での「聞き分け」
   - 実時間処理を目指す高速「聞き分け」
2. 身体性
   - 身体の音響的特徴を活用した「聞き分け」
   - 自己生成音のモデル化・抑制
3. 音を使った発達論的コミュニケーション
   - 対話能力の自己獲得の工学的モデル化
   - シンボル・グラウディングによる言語獲得

2008年度参加研究プロジェクト



1. **グローバルCOE 「知識循環社会」**
   - 原初知識モデル教育研究コア 「音環境理解」(奥乃+ all)
2. **科研費基盤研究(S)「ロボット聴覚」**(奥乃+ all)
3. **科研費特定領域研究「情報爆発IT基盤」**
   - 「音の情報爆発」(奥乃)
   - 「音声対話技術」(駒谷)
4. **JST CrestMUSE**「AIアプローチによる音楽転写」(片寄関学大教授・奥乃)
5. **科研費若手研究(B) 「音声対話のヘルプ」**(駒谷)
6. **科研費基盤研究(A) 「聴覚障害者支援」**(河原和歌山大教授・ 奥乃)
7. 申請中 (尾形)
8. **共同研究**:
   - ホンダ・リサーチ・インスティチュート・ジャパン(奥乃 + all)
   - 理化学研究所(尾形)
   - 企業1(奥乃), 企業2(駒谷, 奥乃), 企業3(奥乃, 高橋) (人気大手企業)
9. **財団資金**: SCAT(駒谷)

研究室配属学生への教育サービス

1. **修士課程への進学は前提(3年間一貫教育)**
   博士課程への進学も考えてほしい
2. **専門教育(過去全員が信号処理は初めて)**
3. **プログラミング指導: Matlab (C、C++、Java)**
4. **音声認識システムの修得**
5. **発表指導・論文・申請書執筆指導**
6. **学士:情報処理学会全国大会発表**
7. **修士:国際会議発表、ジャーナル論文投稿**
8. **毎週水(火?)曜日: 昼食会・研究会**
9. グループ会合(週末は完全休息)

2008年度の卒業特別研究の流れ

1. **研究室配属, グループ仮決定**
2. **Matlab による信号処理トレーニング(4月集中, 木金)**
3. **音声認識システム修得(5月)**
4. **グループでの演習・英文輪講(パターン認識)**
5. **テーマ決定・大学院願書出願**
6. **8月6~7日 情報学研究科大学院試験**
7. **卒業特別研究を進める**
8. **1月初旬:情報処理学会全国大会論文投稿(2頁)**
9. **2月初旬:卒業特別研究報告提出**
10. **2月中旬:卒業特別研究諮問会**
11. **3月10~12日:情報処理学会全国大会発表(立命館大)**
12. **3月下旬:国際会議投稿(英文4~6頁)**
    IROS2009(米), RO-MAN2009(富山), Interspeech2009(英), ISMIR2009(神戸)

奥乃研からの受賞者(B3でも受賞!)

- (B3):情報処理学会
- (M2, D3):渡航助成
- (D1):IEEE RAS日
- (D1):**学振特別研究**
- (D3): SI2006優秀
- (D修):情報処理学会
- (D修):FIT論文賞, 財
- (D修):SI2004ベスト
- (M修):情報処理学会
- (M修):情報処理学会
- (M修):人工知能学会
- (M修):情報処理学会
- (M修):情報処理学会
- (B卒):情報処理学会
- (D修): FITヤングリ
- (D修):人工知
- (東京理科大
- 詳しくはHPを。

nalist, 渡航助成

), テレコムシステ
研究員(DC2),

ブ発表賞, Best in Class Award (Drum),
技術学生賞
部学生研究奨励
助成

京理科大学長賞、

技術・実績が積める
- 国際会議発表年1回以上
- 国内は好きなだけ発表
- ジャーナル論文も

名誉も
- うまくいけば種々の賞も

経済的なサポートも
- TA, RA等の学内アルバイト
- 奨学金全額返済免除

将来もバラ色
- 希望企業に就職
- OBとのネットワーク
- 行事等の手配が上手に

学生の発表論文・受賞とも京大でもダントツ!

特別研究のテーマの候補（案・応談）

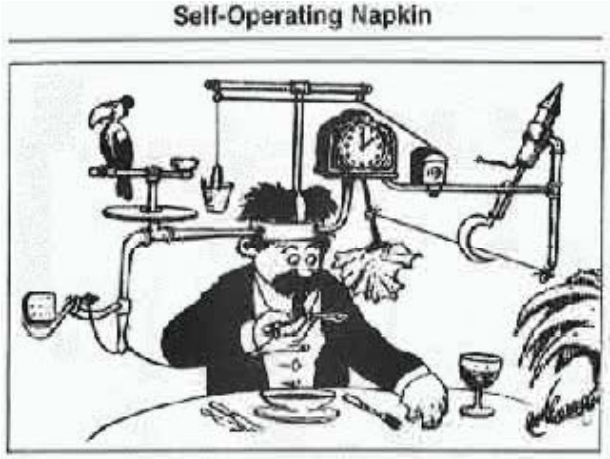
音やロボットに関連したあなたの夢，やりたい研究

- ロボット聴覚・音環境理解
  - ソフトマスクを用いた雑音頑健な音声認識システム
  - 環境音認識システム
- 音楽情報処理
  - 音楽映像情報検索システム
  - 音楽に合わせて歌うロボット
  - メロディー抽出
- 音声対話システム
  - ロボットを使った複数人との音声対話
  - 数独解法チュータシステム
  - 擬音語によるヒューマンコンピュータインタラクション
- ロボットのための音声模倣発達
  - 共感覚によるロボットの動作・音声の生成
  - あかちゃんのような模倣機能を有するロボット

今年は源氏物語誕生1000年

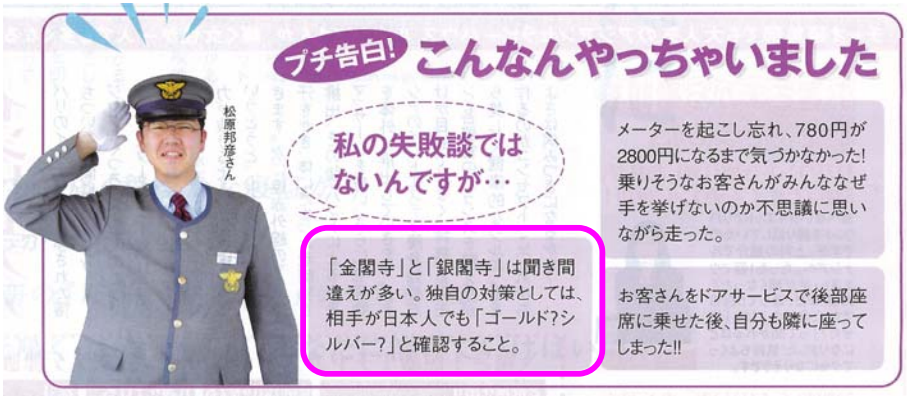
源氏物語千年紀
紫のゆかり、ふたたび

身体性を使った音研究：
定性的な処理だけでは不十分

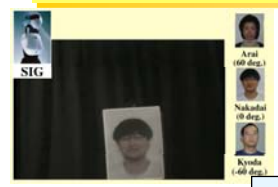


音声対話システムの課題

- システムが聞き間違う
- 人が聞き間違う



三話者同時発話認識技術の進展



2002年

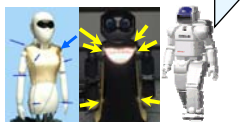
話者の間隔がより狭く

2003年

- 人間による実験
- 文章の認識
- より大きな部屋
- 事前情報がより少なく
- 3種類のロボットで実証

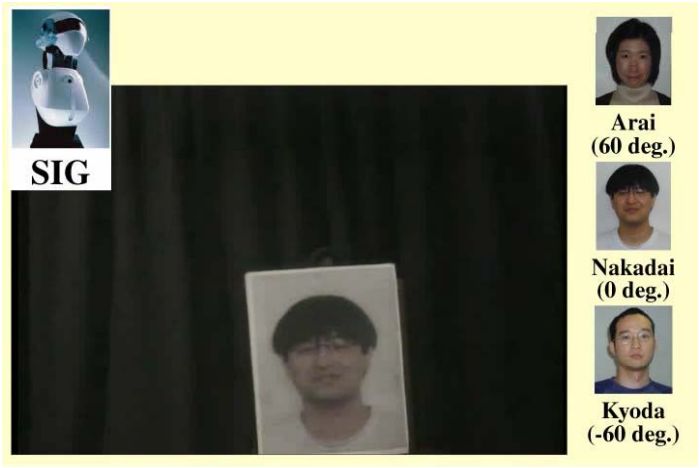
2005年

GA最適化による
4倍の高速化



2006年

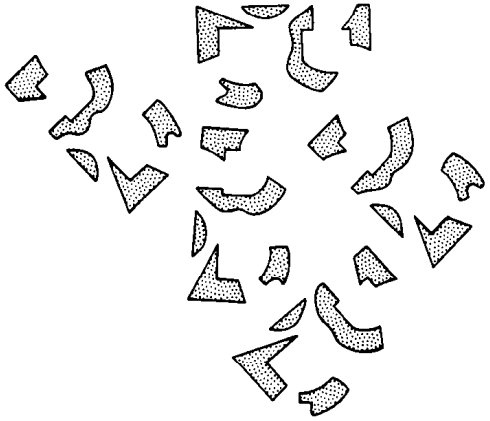
2002: 60° 間隔の3話者



Please notice you hear SIG's motor noises.

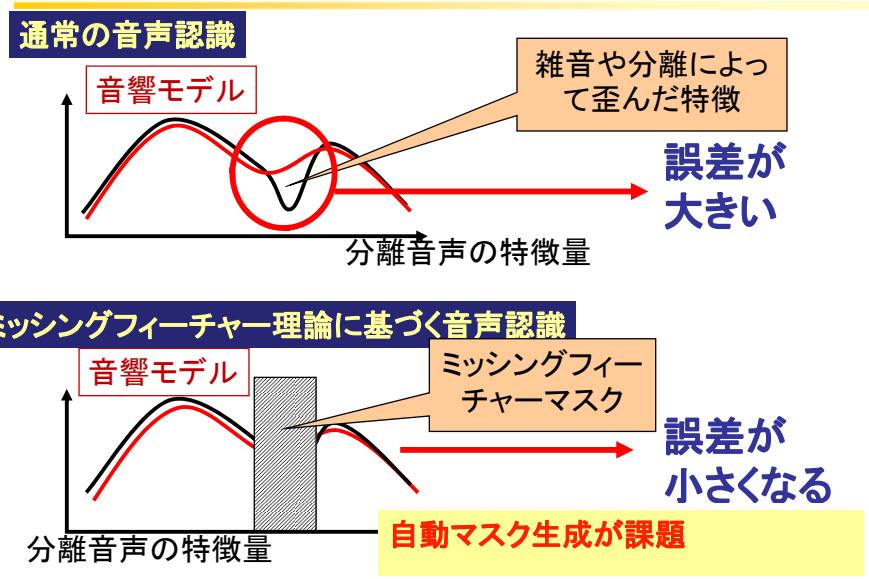
2003: 30° 間隔の3話者

雑音や残差には意味がある



この文字は？　Occlusion（見え隠れ）？

ミッシングフィーチャー理論 (MFT) とは

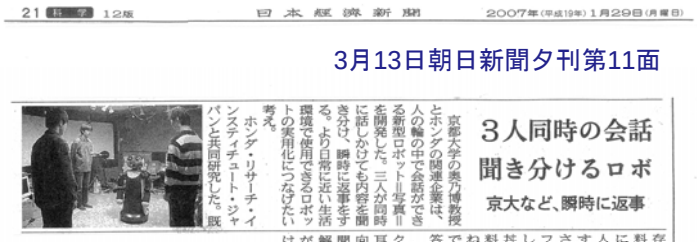


2005: 3人実話者が料理を注文

応答までの時間に注目を

2006: 4倍の高速化

応答までの時間に注目を



3月13日朝日新聞夕刊第11面

3人同時の会話
聞き分けるロボ
京大など、瞬時に返事

京都大学の奥乃博教授とホンダの関連企業は、人の輪の中で会話ができる新型ロボット「写真」を開発した。三人が同時に話しかけても内容を聞き分け、瞬時に返事をする。より日常に近い生活環境で使用できるロボットの実用化につなげたい考え。

ホンダ・リサーチ・インスティチュート・ジャパンと共同研究した。既存のロボットを改良し、料理の注文専用ロボットに応用した。実験で、三人が「カツ丼お願いします」「トンカツ定食ください」「フランス料理のフルコース」と同時注文した。ロボットは「カツ丼とトンカツとフランス料理のフルコースですね」と答え、合計金額まで含めて返事をした。返答までの時間は二秒。

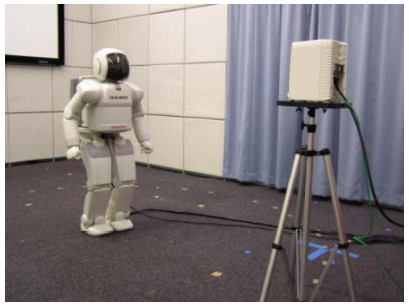
ロボットに八個のマイクロホンをつけた。人の耳と同じように、どの方向でどんな音がしたかを聞き分け、会話の意味を解析する。理論上は八人が一斉に話しても聞き分けられる。

2008年3月26日
フジテレビ「偉大なる未来図鑑」で聖徳太子ロボット放映
山本・武田・水本君が登場

2007年2月21日
テレビ東京
WBSで放映
吉井君が登場

ロボットが音楽に合わせて足踏み

1. スピーカから流れてくる音楽を聴き，ビートトラッキングを行う.
2. 得たビートから次のビートを予測し，足踏み
3. ビートが変わると追従



（Honda Research Institute Japan との共同研究）

ロボットが音楽に合わせて足踏み

（Honda Research Institute Japan との共同研究）

共感覚 （Synaesthesia）

視聴覚情報統合

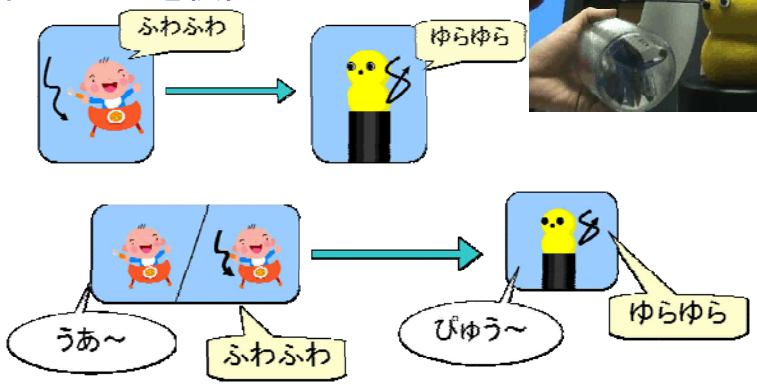
どちらが Kiki でどちらが Bouba ？

Figure 1 (Ramachandran & Hubbard, 2001)

音響と画像とのモダリティ変換

- 音イベント＝音源の動き＋音響信号

- 音イベントを模倣



見学会でのデモンストレーション
工学部10号館4階EV前に集合
2グループに分かれて見学

1. マルチドメイン音声対話システム（414）
2. ロボットの経験による挙動獲得（424）
3. 音楽情報処理（地下MM・前室）
4. ロボット聴覚・音環境の可視化（地下MM ）
5. 面談希望者： 夕方3講義で打ち上げ.
   4/3（木）午後、4/4（金）～12時.
   夜は何時でも可.　午前中はメイルで予約を.

見学会（後）で質問をしてください。