

研究室配属の1つの考え方

研究室選びは最初の就職活動

- 研究室選択権は学生にあり (cf. 就職とは違う)
- マッチする学生さんに来てほしい。
- 教員・在籍学生との相性が一番重要
- 【情報収集】研究室のHP, 年の離れていない先輩の話, 研究室就職状況。
- 参考: <http://winnie.kuis.kyoto-u.ac.jp/>

2013年度の奥乃研究室メンバー

教授 奥乃 博 (2001年4月採用 東大・企業 26.5yrs・独法・東京理科大)

准教授・講師 公募中

助教 糸山 克寿 (2011年4月採用, 京大D)

特定助教 西出 俊 (2011年4月採用(次世代研究者育成センター), 阪大・京大D)

秘書 岡崎 裕美 (2010年5月採用)

博士課程 D3 3名 (学振特別研究員1名+国費留学生1+私費留学生1) 

D2 2名 (国費留学生1+社会人:東芝) 

D1 1名 (社会人:日立)

修士課程 M2 6名 (計算機科学コース5名)

M1 4名 (計算機科学コース3名, 内飛び級入学1)

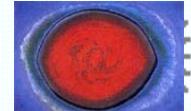
4回生 4~5名(予定)

ロボット HRP-2, HIRO, Keepon, SIG-2, Robovie-IIIs, Robovie-R2, パロ,

アクトロイド, Nao, Robovie-X, Replie, AIBO, Latte, ネコロ, ドラえもん, 個人所有Nao (正価200万円!!!)



2013年度 奥乃研究室紹介

- 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一メディア分野 
- 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」
 - 「二次創作CGMのための音楽音響処理」
 - 「音によるHuman Robot Interaction」
- 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Problem Based Learning
 - 日本・世界の研究を牽引

2013年度の居室・実験室

1. 総合研究7号館4階412号室: 糸山助教, 学生居室

2. 総合研究7号館4階407号室: 奥乃研実験室 

3. 総合研究7号館4階408号室: 奥乃教授室

4. 総合研究7号館4階410号室: 秘書室

5. 総合研究7号館4階412号室: 准教授室・倉庫

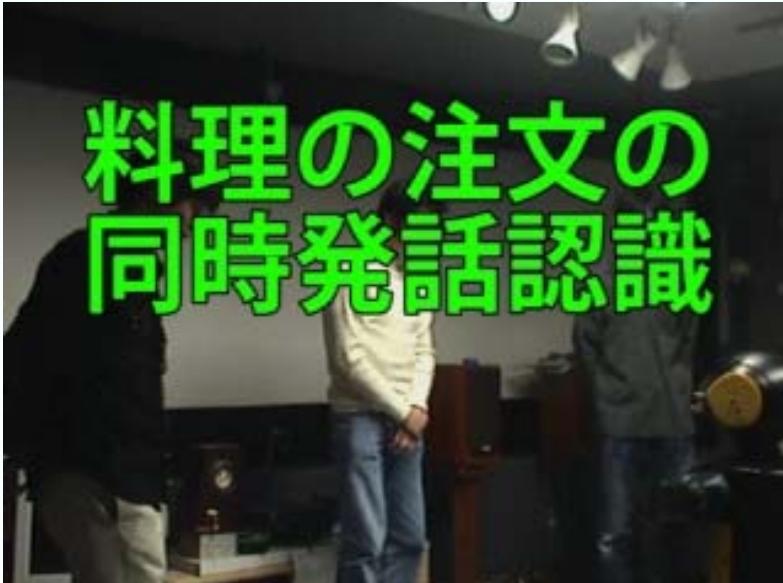
6. 坂記念館(7号館正面2階) 無響室, 移動ロボット, 空中ロボット実験室

7. 総合研究1号館3階(白眉): 西出助教室, 学生居室

私たちの夢： 世界中のロボットを聖徳太子に



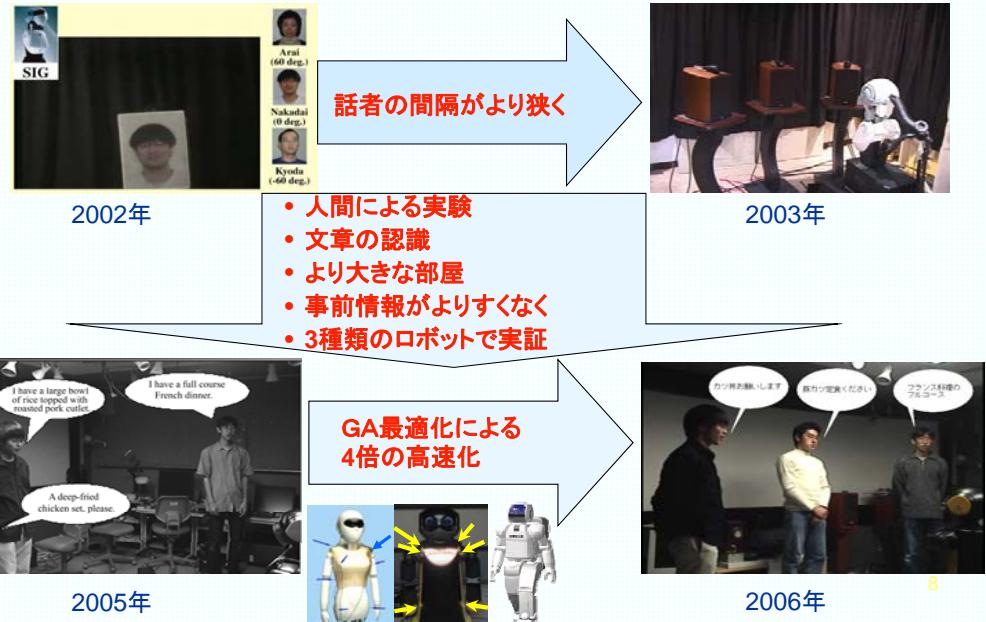
ロボット聴覚システムHARKの応用



3人の注文
終了後、
ロボットの
応答時間
に注意！

Before: 7.9秒
After: 1.9秒

三話者同時発話認識技術の進展



HAKK メンバ Willow Garageに招待



<http://www.willowgarage.com/>

無人飛行ロボット・ヘビ型に耳を

東日本大震災を目のあたりにして世の中の役立つことをしたい



Active Scope Camera



日経新聞2013年3月24日

東日本大震災を目のあたりにして 世の中の役立つことをしたい

- ・災害用ロボットでなぜ耳は使われていないのか
- ・サイバーフィジカルによる安全安心へのアプローチ
- ・Human-in-the-loopによる人とロボットとの協調

新メディア展開にも積極的に貢献

待集 >>

情報処理学会誌5月号(4/15発売)

CGMの現在と未来

初音ミク、ニコニコ動画、ピアプロの切り拓いた世界

編集にあたって

後藤真孝 (産業技術総合研究所)
奥乃 博 (京都大学)

日本の技術・社会・文化の強みの相乗効果。一般の人々が活発な創作活動を繰り広げ、「Consumer Generated Media」、消費者生成の未来を切り拓く画期的な現象が起きている現象の中核に焦点を当てた特別セッション「CgMの現在と未来：初音ミク、ニコニコ動画、ピアプロの切り拓いた世界」を、情報処理学会創立50周年記念全国大会のイベント企画として、2010年3月10日（水）に東京大学で開催した。本イベントは開催前に2件、開催後に8件報道されるなど、注目を集め、700名の会場がほぼ満席となった。インターネットでの無料ニコ動画内蔵映像「アメニコ」がニコニコ動画

- CGM: Consumer Generated Media (消費者生成メディア)
- N次創作
- 能動的音楽鑑賞
- 映像・音響の融合



音楽パートの置換

1. 混合音から置換する楽器パートを抽出するための音源分離



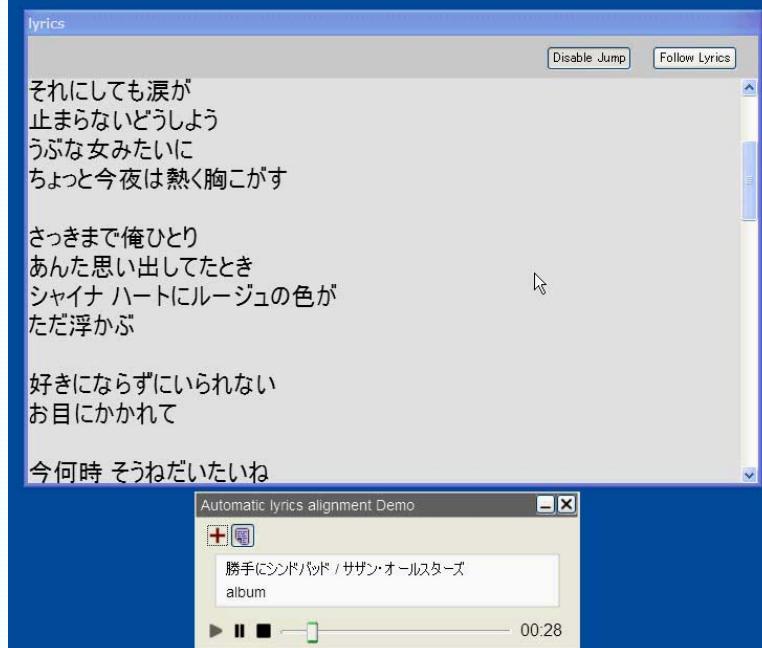
2. 別の楽器音から楽器パートの構成に必要な音の音合成



音の高さ、長さを変えてはもちろん、
発音時の印象を変えるetc.
様々な加工を施した音で音色置換可能

ピアノ音を フォルテ奏法に操作
ピアノ奏法に操作
※音量差がなくなるよう調節済み

CD音楽に歌詞を時間的対応付け



1. 文字ごとに音楽と対応付け

2. 歌詞をクリックすればそこから演奏

原理:ボーカルパートを抽出

応用:電子辞書の対話データ(シャープ・パピルスで商用)

松下幸之助人間力のページ(ドコモ, AU)

カエルの合唱機構の測定による解明

- アマガエルの発声行動の数理モデリング[合原08]

- 2匹・3匹のアマガエルが同期して鳴く現象
- より多いカエル(多体系)への拡張

→ 水田で鳴き声を収録(H20年6~7月)

- フィールド収録音の解析は困難

1. 無数の雑音源

- 別種カエル・コウモリ・自動車

→ 雜音の位置が仮定できない



アマガエル

2. 広い空間

- 大量のマイクアレイが必要

→ 設置・回収が困難

・音光変換デバイスの開発



京大農学部水田
最小でも一辺5~10[m]

水田でのカエルホタルによる観測



90 [deg]



We can see
where the sound is.
that two frogs are calling alternately.

Okuno Lab. Kyoto University

身体性の獲得

音による Human Robot Interaction

音楽音響信号処理によるCGM

認知発達
ロボティクス

マルチパーティ対話

音楽合奏ロボット

楽器演奏音
分析合成補助関数法
による分析

表情認識生成統一モデル

多様な音を聞き分ける

自分の音は抑制

音を見せる

バフリングによる模倣

ロボット聴覚

HARK

Kinect

音のLifelog

歌声・話声弁別

擬音語認識

複数話者同時
発話認識実時間視聴
覚情報統合

音楽情報処理

楽譜追跡

道具身体化

能動的両耳聴

複数話者同時
発話認識

音源3D定位・追跡・分離

調波非調波統合モデル

音によるField観測

神経力学

移動ロボット
AV-SLAMマルチチャネル統
計的音響信号処理Non-Parametric
ベイス推定

力エルホタル

合唱の数理モデル

Recurrent
Neural
Network

人工知能 (AI)

計算機科学 (CS)



2013年度の卒業特別研究の流れ

- 4/10 研究室配属, 4/11 第1回ミーティング・グループ仮決定
- Matlab による信号処理トレーニング (4月集中, 木月, 4/12開始)
- 音声認識システム修得 (5月)
- グループでの演習 (パターン認識, PRML)
- 6月~7月 テーマ決定・大学院願書出願
- 8月6~7日 情報学研究科大学院試験
- 卒業特別研究を 育々と 進める
- 1月初旬:情報処理学会全国大会論文投稿 (2頁)
- 1月末:卒業特別研究報告提出
- 2月中旬:卒業特別研究諮問会
- 3月6~8日:情報処理学会全国大会発表 (名工大)
- 3月下旬:国際会議投稿 (英文4~6頁)
ロボット: IROS2014(Chicago), Humanoids 2014 (?)
音声対話: Interspeech2014(Singapore), 音楽: ISMIR2014



2013年度の研究プロジェクト

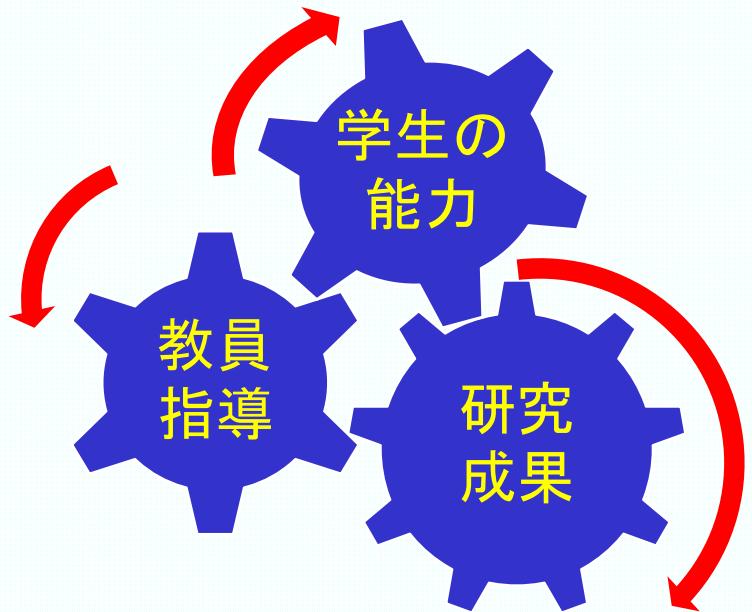
- 科研費基盤研究(S)「ロボット聴覚の実環境理解に向けた多面的展開」
 - 屋内ロボット, 屋外ロボット, 飛行ロボットにロボット聴覚機能の搭載
 - カエルの合唱機構解明のための音光変換デバイス「カエルホタル」の開発
- 同若手研究「Non-Parametric Bayesによる多重奏音楽情報処理」
- JST 日仏研究交流「ヒューマノイドロボットのための能動的両耳聴」
- 共同研究:
 - ホンダ・リサーチ・インスティチュート・ジャパン
 - NTTコミュニケーション科学基礎研究所
- H24年度終了: 科研費萌芽研究「カエルの合唱解明と音楽共演ロボット」
- H23年度終了: 科研費基盤研究(S)「音環境理解からロボット聴覚の構築」
- H23年度終了: グローバルCOE「知識循環社会」「音環境理解」
- H22年度終了: 科研費萌芽研究「カエルの合唱の解明」
- H20年度終了: 科研費萌芽研究「数独の難易度定義と自動問題作成」

研究室配属学生への教育サービス

- 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視野に、1年間だけの教育はご勘弁を。
- 専門教育 (過去全員が信号処理は初めて)
 - 信号処理プログラミング: Matlab (C, C++, Python)
 - 音声認識システムの修得
- 発表指導・論文・申請書執筆指導
 - 学士: 情報処理学会全国大会・国際会議投稿
 - 修士: 国際会議発表、ジャーナル論文投稿
- 毎週水曜日 (10:30~18:00) 昼食会・研究会
- 他にグループ会合 (週末は完全休息のはず)

奥乃研の指導原理



研究指導

Before

ate abbreviations. However, general morphological analysis tools can not segment proper nouns correctly.

2. Reducing recognition errors caused by added abbreviations

Adding generated abbreviations often causes recognition errors, because the vocabulary size increases. Abbreviations whose phonemes is similar to other words' in the original dictionary.

To solve the former problem, we segment words by using morphological analysis and conjunction probability among characters. For the latter problem, we change probabilities of occurrence of generated abbreviations according to the phonemic similarity between abbreviations and words in original dictionary.

We evaluated our method by using 1494 utterances of real users of Kyoto bus information system. Recall of generating abbreviation by our method is 94.3%. On the other hand, that by base line method is 85.7%. This shows effectiveness of our segment approach using conjunction probability. We also evaluate ASR accuracy by using this data. To utterances which contain abbreviations, our method increase ASR accuracy by 24.2 point from baseline. To utterances which contain words which baseline can recognize, our method increase ASR errors by 0.1 point from baseline. The method which is only added generated abbreviations increased ASR errors by 15.1 point from baseline. These two results show that our approach of changing probabilities of generated abbreviations reduced ASR errors caused by adding words into vocabulary. This shows that our method can recognize user's abbreviations by expanding vocabulary with reduced recognition errors.

We evaluated our method by using 1494 utterances of real users of Kyoto bus information system. Recall of generating abbreviation by our method was 94.3%. On the other hand, that by base line method is 85.7%. This shows effectiveness of our segment approach using conjunction probability. We also evaluate ASR accuracy of this data. To utterances which contain abbreviations, our method increase ASR accuracy by 24.2 point from baseline. To utterances which contain words which baseline can recognize, our method increase ASR errors by 0.1 point from baseline. The method which is only adding generated abbreviations increased ASR errors by 15.1 point from baseline. These two results show that our approach of changing probabilities of generated abbreviations reduced ASR errors caused by adding words into vocabulary. This shows that our method can recognize user's abbreviations by expanding vocabulary with reduced recognition errors.

was sharply added to the ASR dict. dictionary
These results

研究指導・論文添削

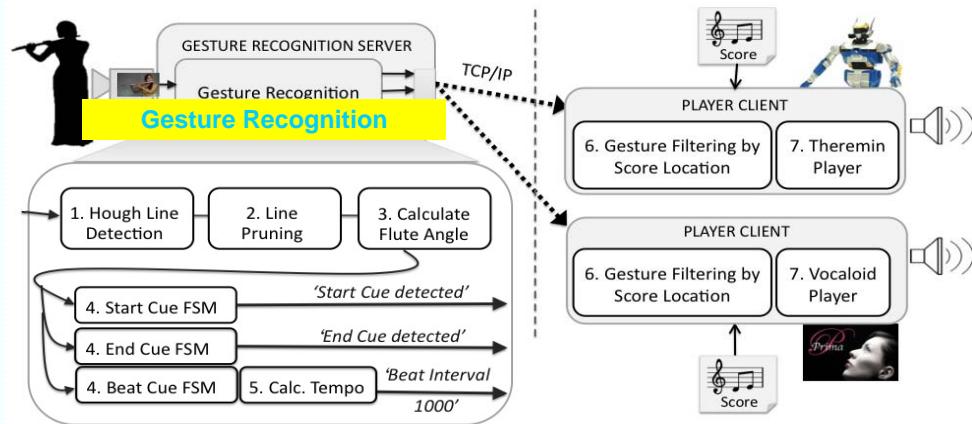
Before

Afte

奥乃研からのB4・B3の受賞

- IEEE/RSJ IROS2006 Best Paper Nomination Finalist
(知能ロボットのトップレベル国際会議) (4/950)
- IEEE/RSJ IROS2008 Award for Entertainment
Robots and Systems (NTF Award) Nomination
Finalist (4/649)
- 情報処理学会大会優秀賞(全発表から10件程度, 1名)
- 同 大会奨励賞(大学卒業10年以内で年10件程度, 3名)
- 同 学生奨励賞(多数, H24春は2名, H23春は3名)
- 財団海外研究発表助成
- 学生支援機構奨学金全額・半額返済免除(大学院)
- 学生の発表論文・受賞とも京大でも全国でもダントツ!

ロボットアンサンブルシステムの概要



The robot can play the theremin and sing while:

- 1) Starting and stopping in synchrony
- 2) Adapting to human's tempo mid-song

35

共演者としての音楽ロボット



音楽ロボット（カルテット演奏）

HARK

Quartet Ensemble Greensleeves

Two humans:
Guitarist (T. Itohara) and Flutist (A. Lim)
Two robots:
Thereminist and Dancing singer

Graduate School of Informatics, Kyoto University
Honda Research Institute Japan

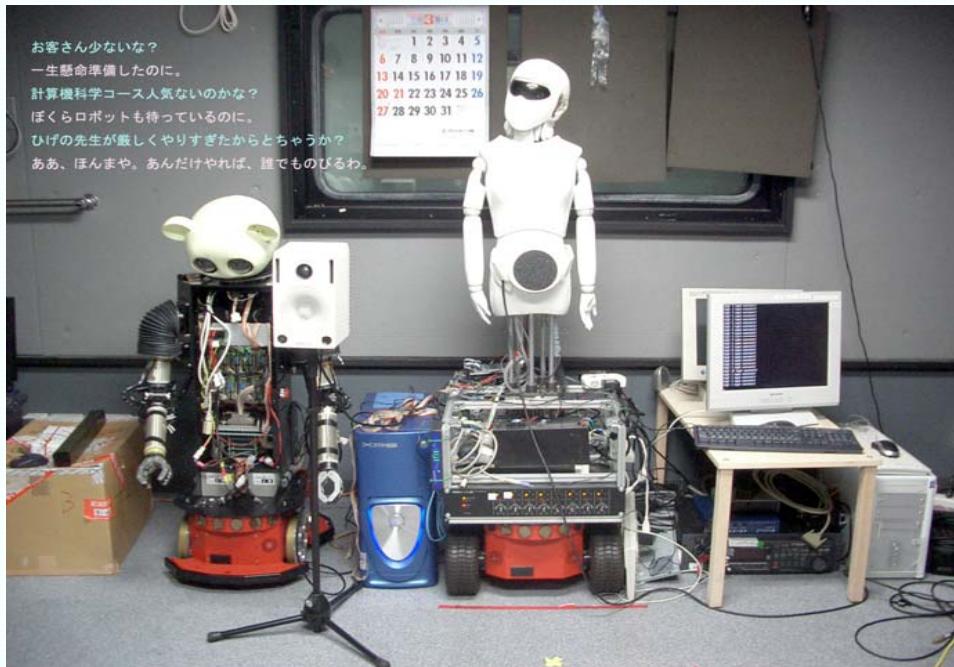
- ビート追跡と視聴覚統合による人との協調演奏
[IEEE Humanoids 2012]

DESIREモデルによる感情

Converting emotional voice
to motion for robot
telepresence

Angelica Lim, Tetsuya Ogata and Hiroshi G. Okuno
Kyoto University

見学会（後）で質問をしてください。



Take Home Messages

奥乃研究室に来れば、投稿論文が1本書けるように指導します。それをして論理的思考(logical thinking)の一端を学べるので、企業に就職するときの武器を授けられます。

あなたの手にする研究室希望調書は

奥乃研究室への

招待状です。