

京都大学 大学院情報学研究科
知能情報学専攻

音声メディア分野

講師：吉井 和佳
yoshii@kuis.kyoto-u.ac.jp

統計的機械学習

ノンパラメトリックベイズ
ディープラーニング
最適化 (MCMC, VB)

聴覚機能の 計算モデル

音声認識・自動採譜

信号処理

音源分離 (NMF, ICA)
音源識別 (HMM, GMM)
残響除去 (AR/MA)

記号処理

フレーズ解析 (N-gram)
構文解析 (PCFG)
意味解析 (LDA)

研究室選びは最初の就職活動

教員・在籍学生との相性が一番重要

マッチする学生さんに来てほしい！

研究室選択権は学生にあり (≠実際の就職)

研究室に関する情報収集が大事

<http://winnie.kuis.kyoto-u.ac.jp>

研究発表のアクティビティの高さ

受賞・助成・奨学金返済免除

研究室OB (修士・博士) の就職状況

年齢が近い先輩の話

研究の進め方

音とロボットに関する知能の解明

キーワード：聴き分ける・音を見せる

音楽音響処理：自動採譜・音源分離・CGMのための音楽解析

ロボット聴覚：屋内/屋外環境での音環境理解・インタラクション

学生の得意な面を伸ばします

新しい問題を見つける・挑戦する

学生の自由な発想を尊重

着実な研究成果 (国内会議 → 国際会議 → 論文誌)

世界の一線で戦う (英語化必須)

2015年度 研究室メンバー

今が旬の若手の教員・学生が世界を舞台に活躍中

この1年で旧奥乃研の成果を土台にさらに飛躍！

教授	選考中	H27年5月以降	計算機科学は30代の若手が牽引
講師	吉井 和佳	H26年2月採用・京大D修了→産総研	
助教	糸山 克寿	H23年4月採用・京大D修了	
秘書	岡崎 裕美	H22年5月採用	 
修士課程	M2 3名	うち計算機科学コース1名	
	M1 4名	うち計算機科学コース4名	
学部4回生	4～5名 (予定)	M1&M2に女性も在籍	
ロボット	Nao 3台・移動ロボット(地上・飛行)など多数所有		

2014年度 受賞一覧

この一年で11件の受賞歴あり (研究室中最多?)

- (1) M2池宮由楽, 優秀研究賞, 京都大学第9回ICTイノベーション, 2015年3月
- (2) M1 土橋彩香, 情報処理学会 第77回全国大会, 2015年3月
- (3-7) M2 池宮由楽, M2 坂東宣昭, B4 大喜多美里, B4 板倉光佑, B4 関口航平
- (8) M2 Yukara Ikemiya, Winner of Singing Voice Separation Track at MIREX 2015, 2014年10月
- (9) M2 Yoshiaki Bando, Best Student Paper Award, 12th International Symposium on Safety, Security, and Rescue Robotics (SSRR 2014), 2014年10月
- (10) M2 池宮由楽, ベストプレゼンテーション賞
情報処理学会 第104回音楽情報科学研究会 (夏のシンポジウム), 2014年8月
- (11) M2 望月敬太, 大会優秀賞, 情報処理学会 第776回全国大会, 2014年6月

B4が大活躍！すでに3名とも国際会議に投稿済み (残り1名も投稿準備中)

2015年度 研究スペース

総合研究7号館

408号室：教授室

412号室：吉井講師室

410号室：岡崎秘書室

417号室：糸山助教・学生室

407号室：実験室 (ロボット・3Dプリンタ)

坂記念館 (7号館正面2階)

無響室・実験室 (クアドロコプター)

見学・相談は随時受け付けていますのでお気軽にどうぞ
yoshii@kuis.kyoto-u.ac.jp

音を聴き分ける

「音声」を聴き分ける

聖徳太子：複数同時発話の内容を認識

マイクロホンアレイ処理：音源分離・反響除去

「音楽」を聴き分ける

音楽理解：音源分離・自動採譜など

歌声解析：歌唱スタイルの認識・転写

演奏支援：人間との合奏・伴奏

「物音・鳴き声」を聴き分ける

上空から/災害現場での対象の発見

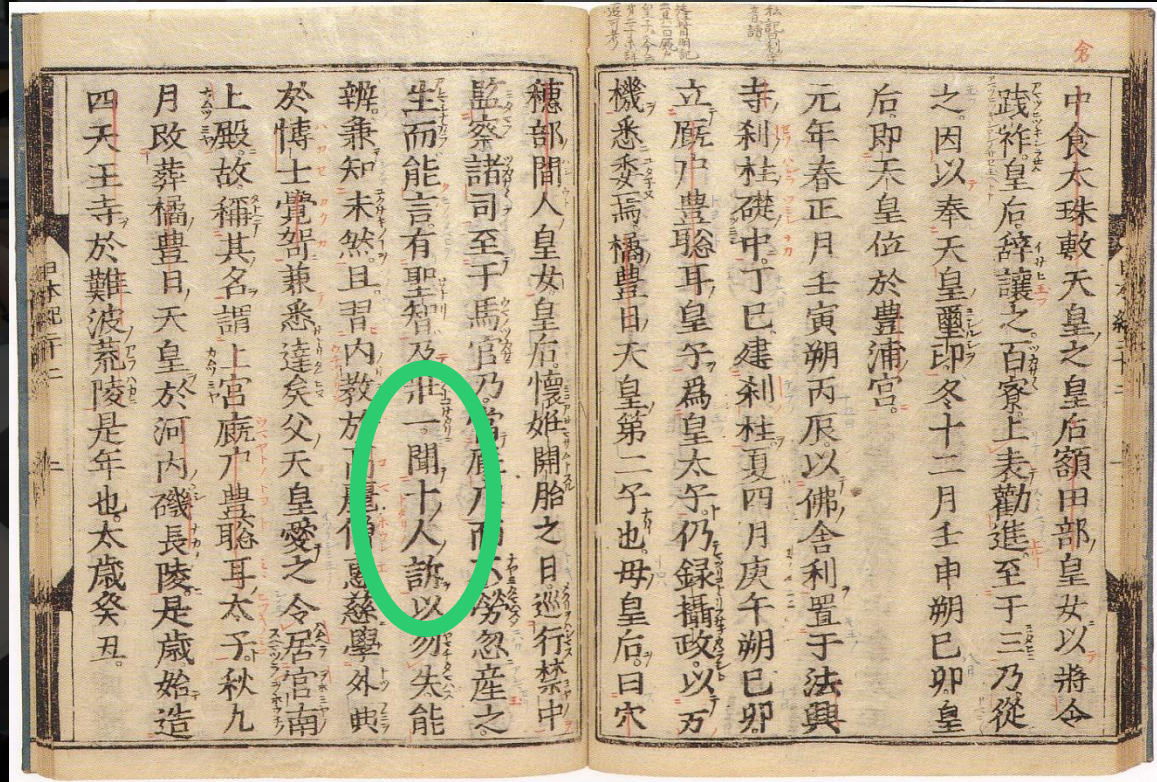
カエルや野鳥の鳴き声の分析



「音声」を聴き分ける

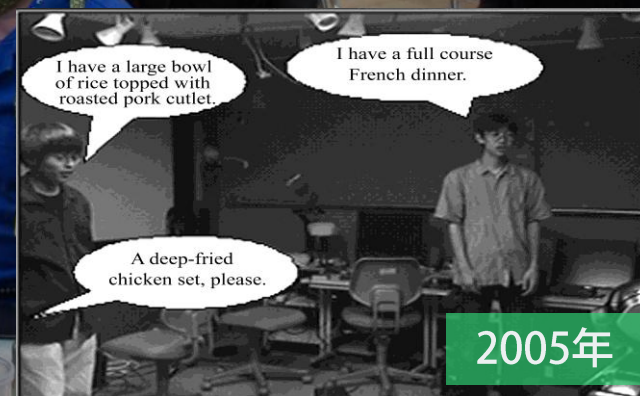
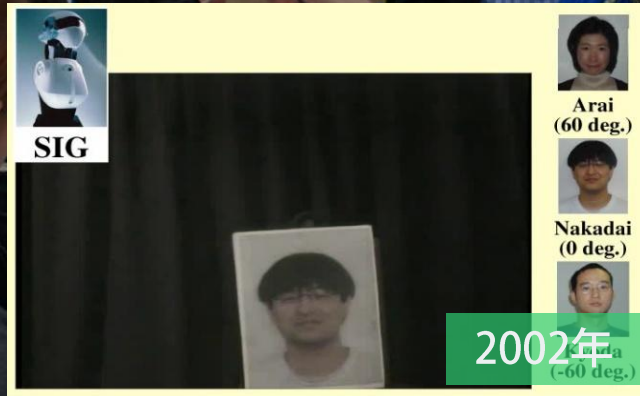
聖徳太子ロボット

複数人の同時発話を聴き取れるか？

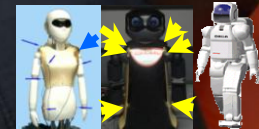


三話者同時発話認識の発展

話者
間隔
を狭く



- ・ 人間による実験・単語 → 文章の認識・4倍の高速化
- ・ 大部屋・少ない事前知識・3種類のロボットで実証



ロボット聴覚システム#ARK

料理の注文の 同時発話認識

自然発話に対する応答時間に着目

聖徳太子ロボット2012

耳がたくさんあれば人間を超える！？

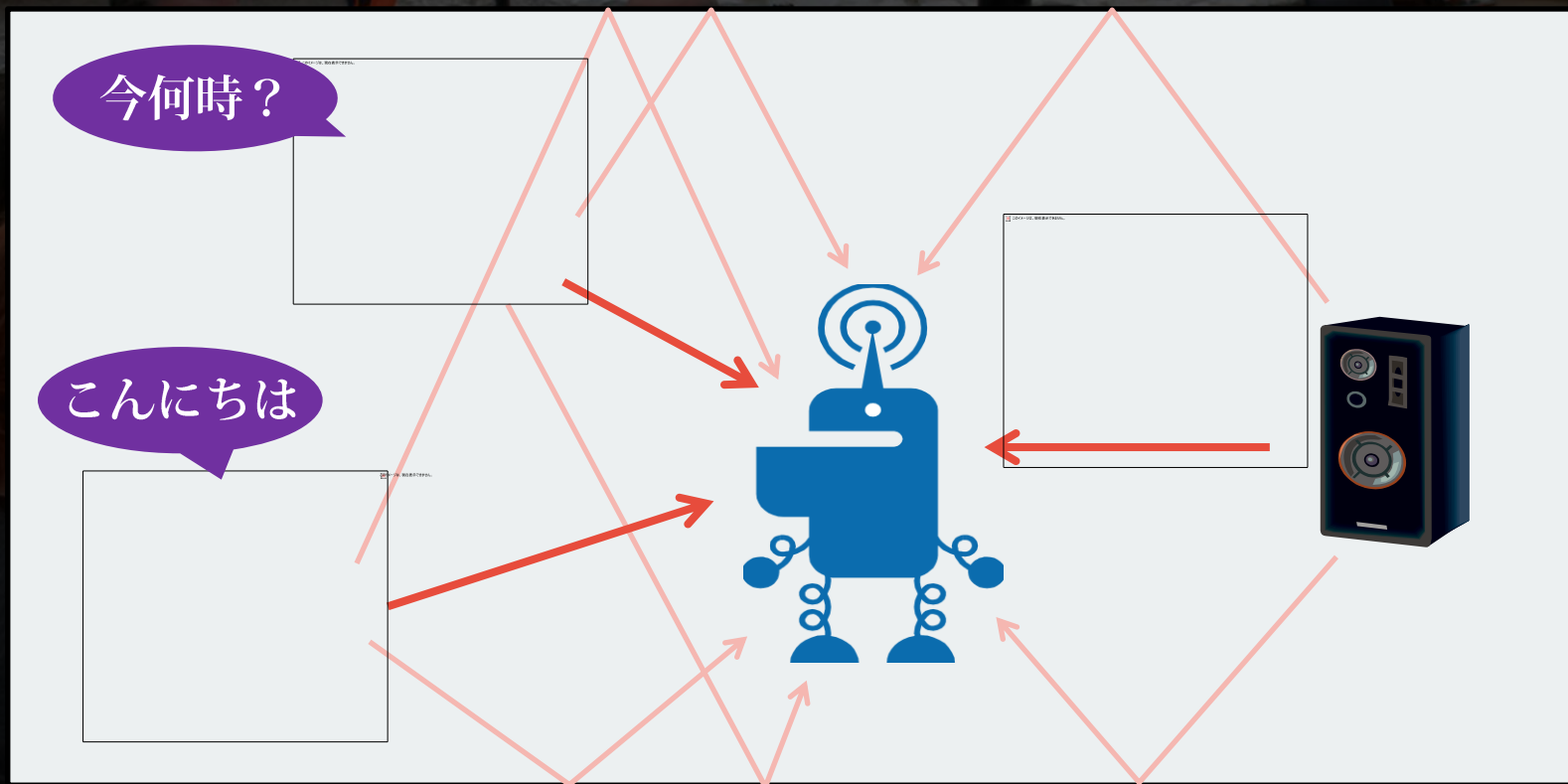
Simultaneous Speech Recognition

～ Meal Order Taking ～

16chマイクロホンアレイ・音源方向は所与

マイクロホンアレイ処理

音源数・残響時間が未知で聴き取れるか？



ベイズモデルによる確率的統合

鶏と卵の関係を一挙に取り扱う方法論

定位 → 分離 → 残響除去による誤差の累積を回避可能

研究のトレンド
信号処理
+
機械学習
+
実問題への応用

残響除去モデル

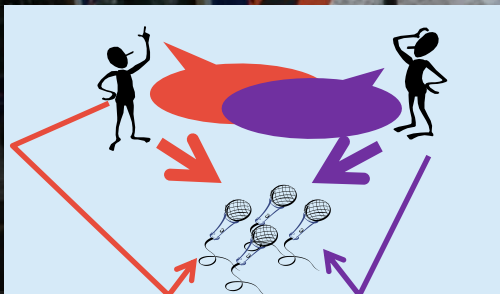
統一的
ノンパラメトリック
ベイズモデル

音源分離モデル

音源定位モデル

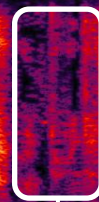
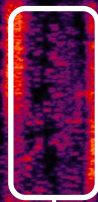
「人工知能」
「信号処理」
両分野で最難関
国際会議AAAI・
論文誌IEEEに採択

定位＋分離＋残響除去

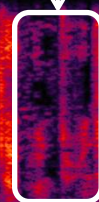
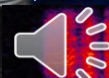
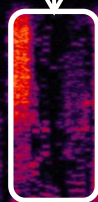


2話者同時発音

残響除去なし



残響除去あり



ポスター会場で実稼働

時計台国際交流ホール

マイクロホンアレイ



観測した混合音



分離音



周囲の人の話し声

背景雑音



「音楽」を聴き分ける

人間とロボットの協奏

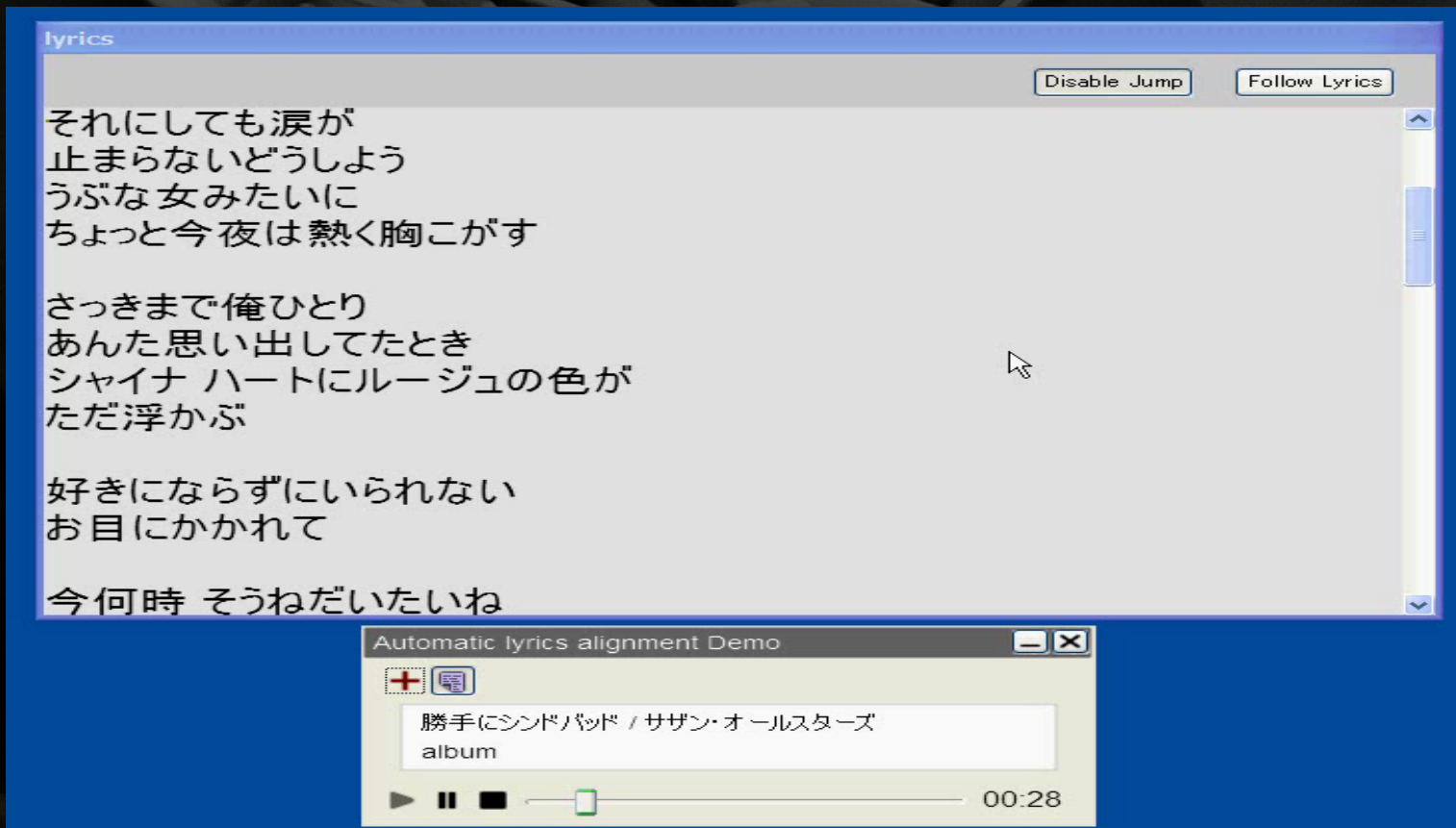
ギター演奏の楽譜位置をリアルタイムで追跡

ロボットの耳で「相手の出す音」を聴き分ける
テンポ変化にも追従



音響信号と歌詞の自動同期

歌詞で簡単に頭出しが可能 (電子辞書で商用化)

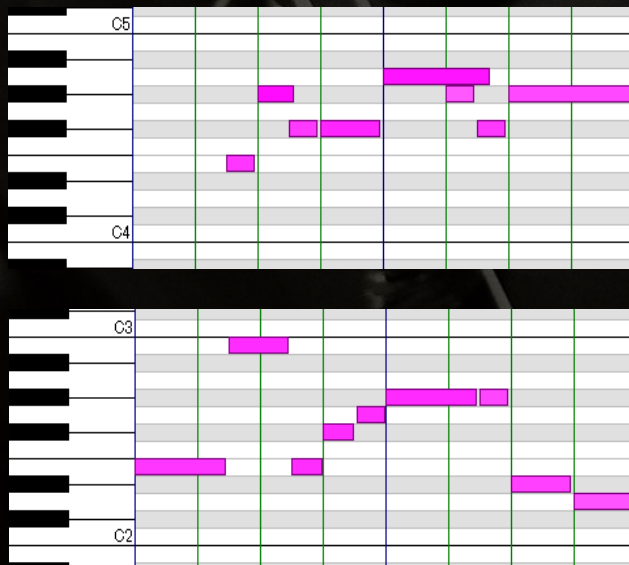


自動採譜・音源分離

音楽音響信号を「パーツ」に分解したい

無限個の [音高] と [音色] を取り扱えるノンパラメトリックベイズモデル

音高 (ソース)



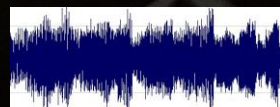
音色 (フィルタ)



組み合わせて
様々な
楽器音を生成



重ね合わせ



音楽音響信号

音楽信号の楽器パート分離

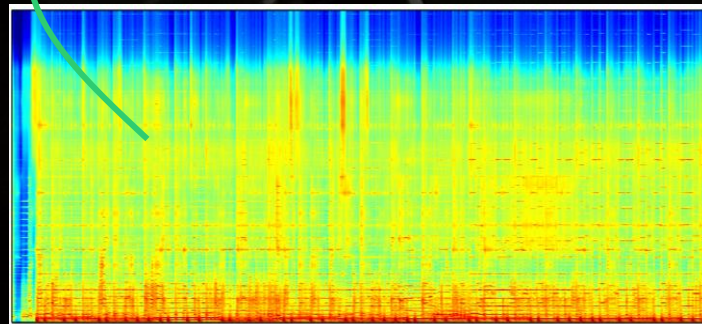
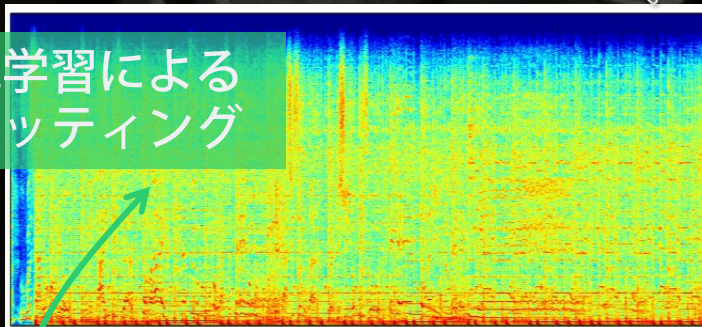
混合音を音色ごとにある程度分離できる

観測スペクトログラム

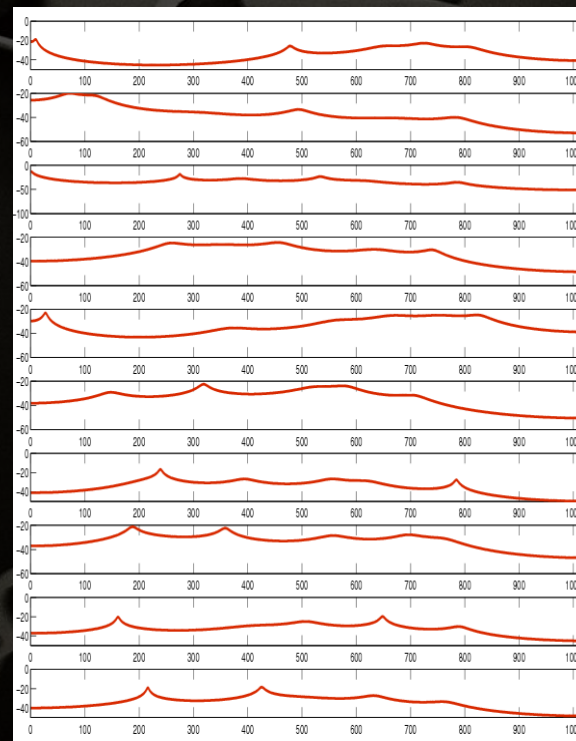


機械学習による
フィッティング

ノンパラメトリックベイズモデル



学習された周波数応答 (音色)



CGM文化への積極的なコミット

一般ユーザとクリエイータの境界が曖昧に

特集>>

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

産総研でのボス

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により、一般のユーザが参加するCGM（Consumer generated media、消費者生成メディア）の未来を切り拓く画期的な現象が起きている。その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「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

- 音楽・イラスト等を合法的に再利用した創作を可能にするコンテンツ投稿サイト「ピアプロ」の生みの親で、上記の「VOCALOID」に基づく歌声合成ソフトウェア「初音ミク」販売元の伊藤博之氏（クリプトン・フューチャー・メディア（株））
- 上記の「初音ミク」等を用いたCGM動画が活発に投稿され、ユーザによる創作を促している動画コミュニティサービス「ニコニコ動画」の生みの親の懸塚昭彦氏（（株）ドワンゴ）
- 上記の「ニコニコ動画」の先進性を考察する上で重要な「擬似同期型アーキテクチャ」や「N次創作」等の概念の生みの親の濱野智史氏（（株）日本芸芸）という4名の第一人者の皆様に記事の執筆をお願いした。また、本イベントを企画し、司会を担当した後藤真孝（産業技術総合研究所）も記事を執筆した。

情報処理学会誌2013年5月号（初の完売！）



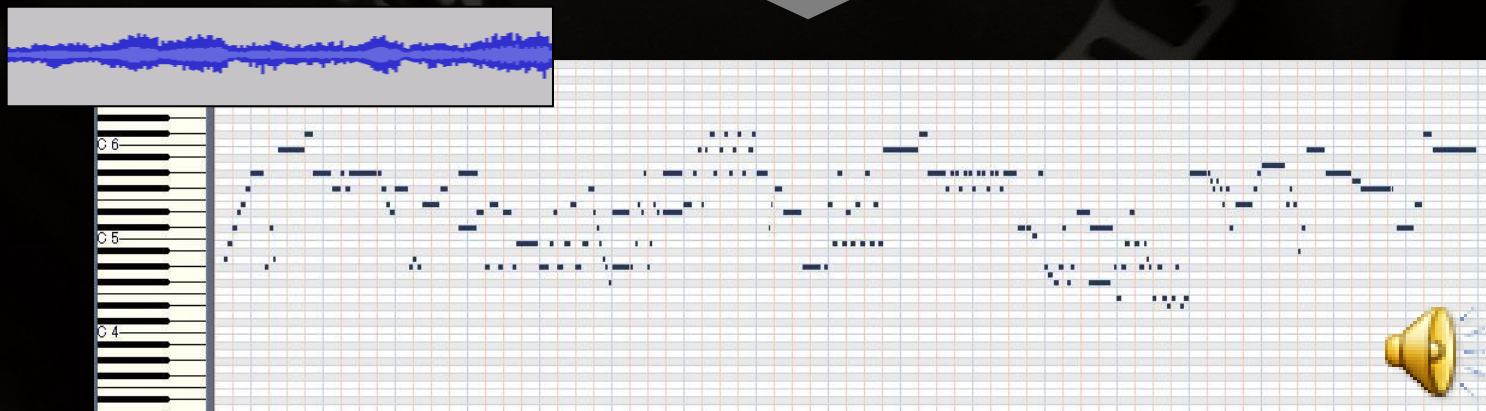
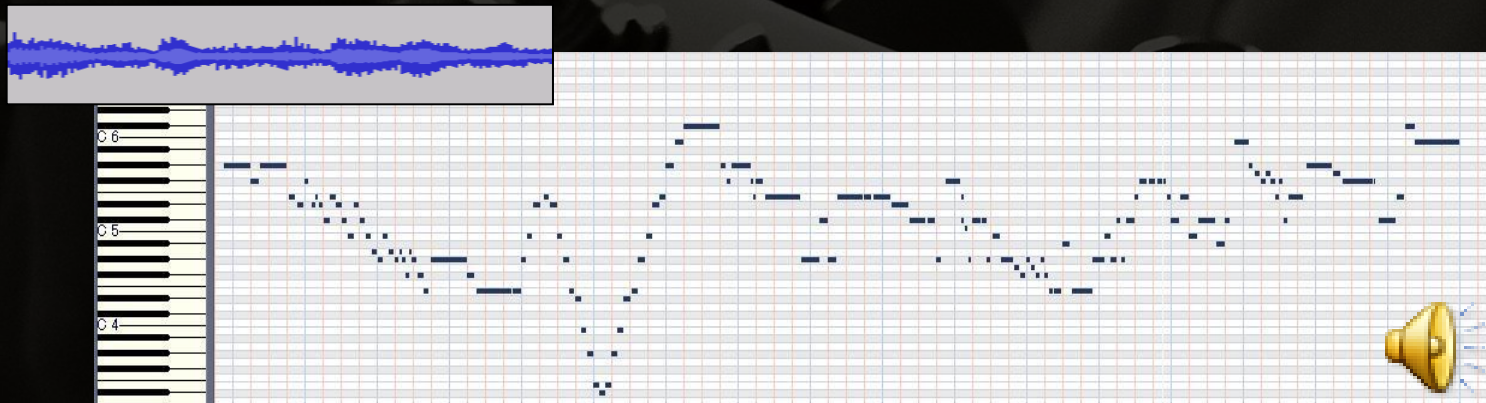
表紙

混合音中のドラムパートの置き換え

Drumix: An audio player with
a function for re-arranging
drum parts in real time

Kazuyoshi Yoshii
Masataka Goto
Kazunori Komatani
Tetsuya Ogata
Hiroshi G. Okuno

混合音中のメロディの置き換え



歌声と伴奏の分離

オリジナル



歌声



音高



伴奏



歌声と伴奏の分離

オリジナル

歌声

伴奏

そばかす /
JUDY AND
MARY



I'm in the mood for
dancing / The Nolans



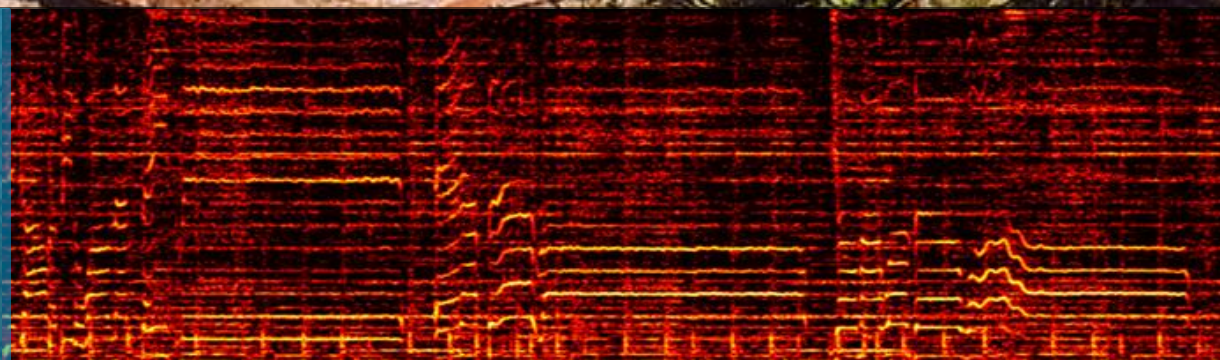
Morning Glory /
Oasis



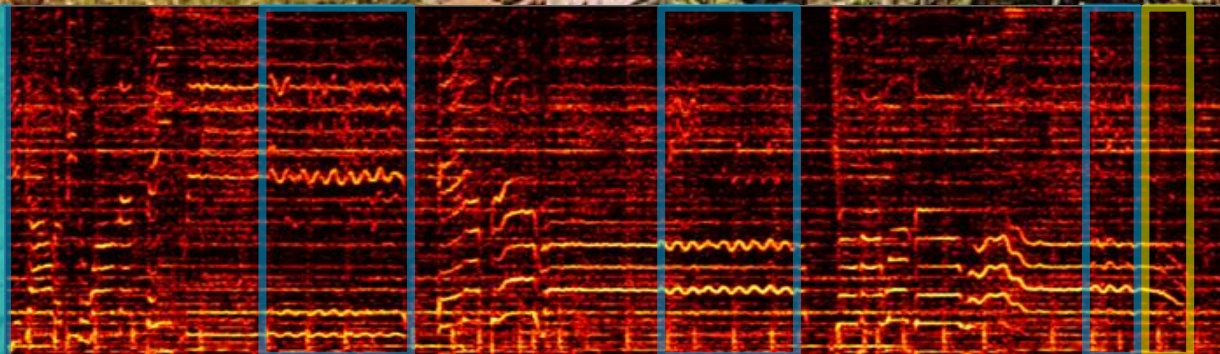
混合音中の歌声編集

「Wonderwall / Oasis」 にビブラートを付加する

編集前



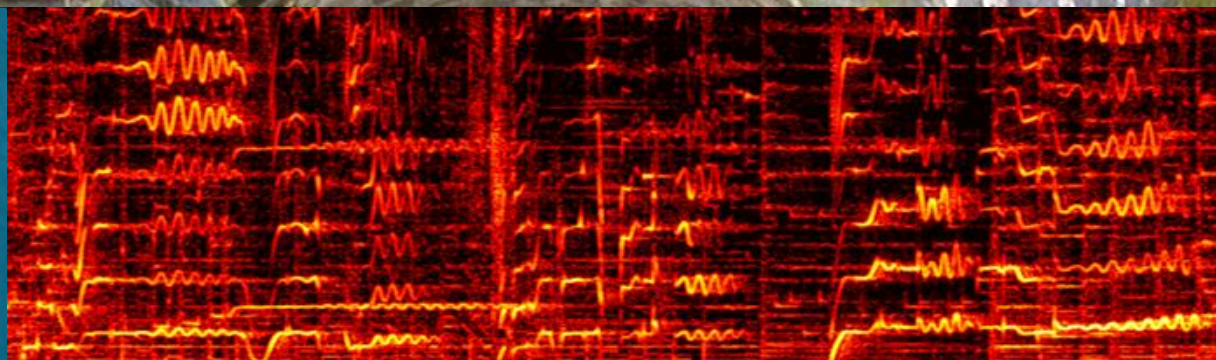
編集後



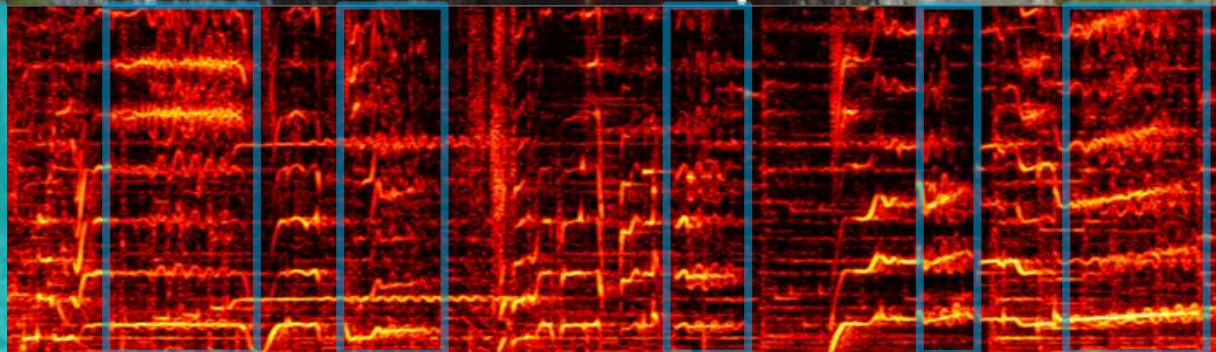
混合音中の歌声編集

「天城越え / 石川さゆり」からビブラートを取り除く

編集前

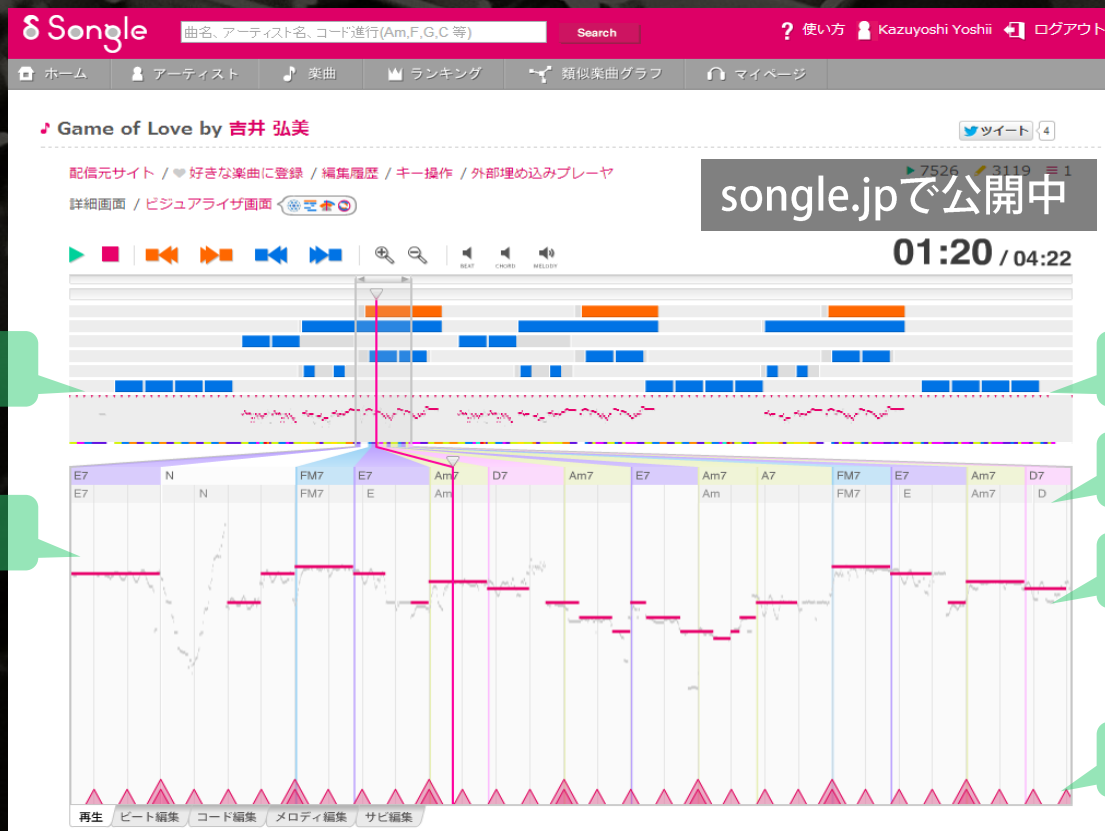


編集後



Songle: 能動的音楽鑑賞WEBサービス

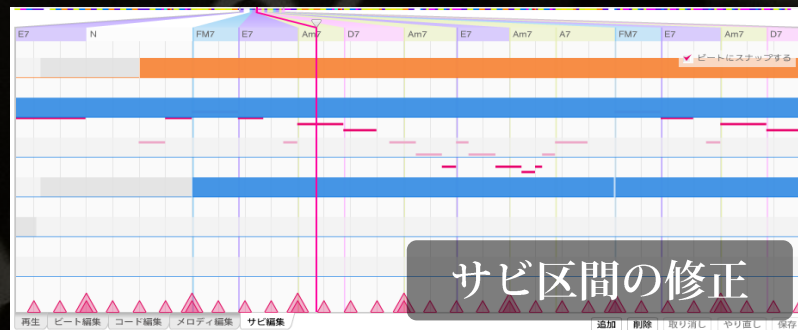
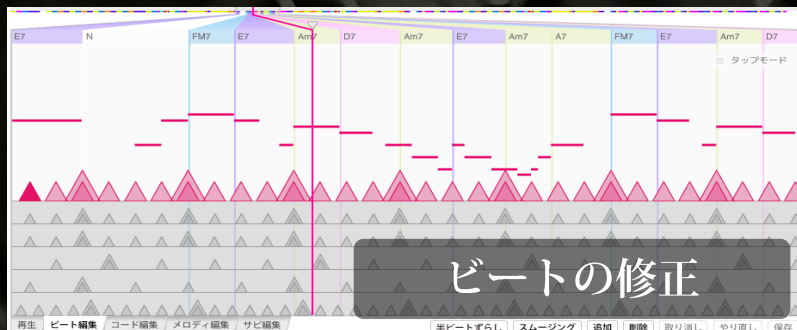
解析結果を目で見ながら音楽を楽しむ



学術研究/発表の新しいかたち

機械学習技術でユーザ貢献を増幅する！

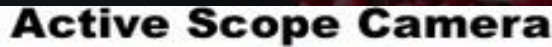
ユーザが修正 → システムが再学習 → 認識率が向上 → 全ユーザが喜ぶ



A kingfisher bird with a bright yellow-orange body and a dark blue head and back is perched on a branch. The bird is facing left, and its wings are slightly spread. The branch is covered with clusters of small, round, red berries. The background is a soft, out-of-focus green, suggesting a natural, outdoor setting. A semi-transparent dark blue horizontal band runs across the middle of the image, containing white Japanese text.

「物音・鳴き声」を聴き分ける

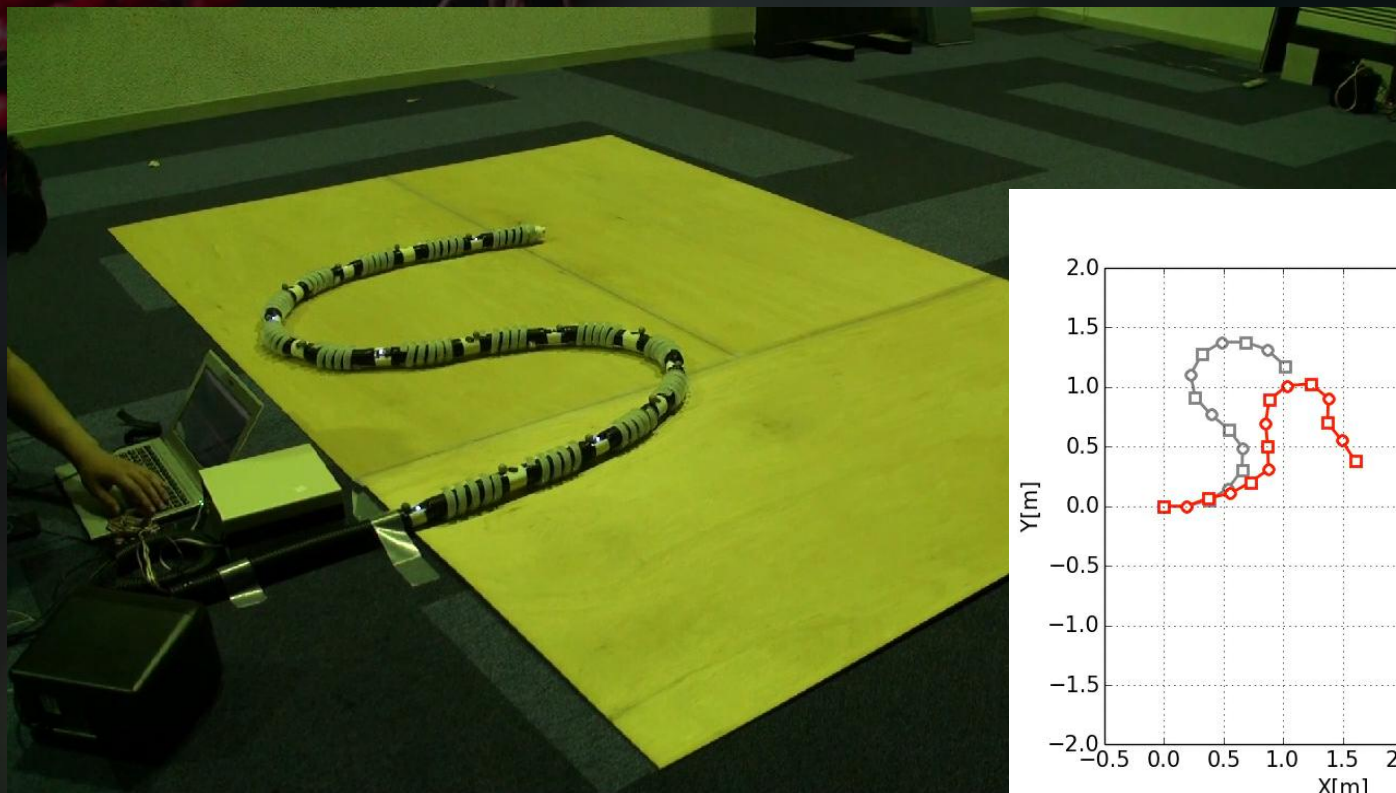
災害時/極限環境で役立つ音の研究を



日経新聞2013年3月24日

ホース型ロボットによる音源探索

極限環境化における姿勢推定 & 音源探索



飛行ロボットによる音源探索

移動型マイクロホンアレイを搭載



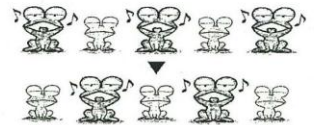
カエルの鳴き声の可視化

カエルの声を聴き分けるデバイスの開発



©読売新聞（大阪） 2014年01月28日 朝刊◇34面
※無断複製転載禁止

カエルの歌 輪唱してた



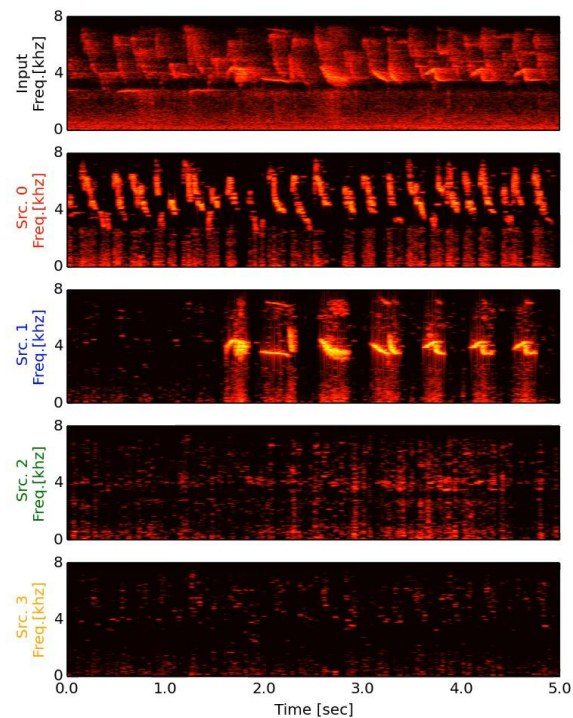
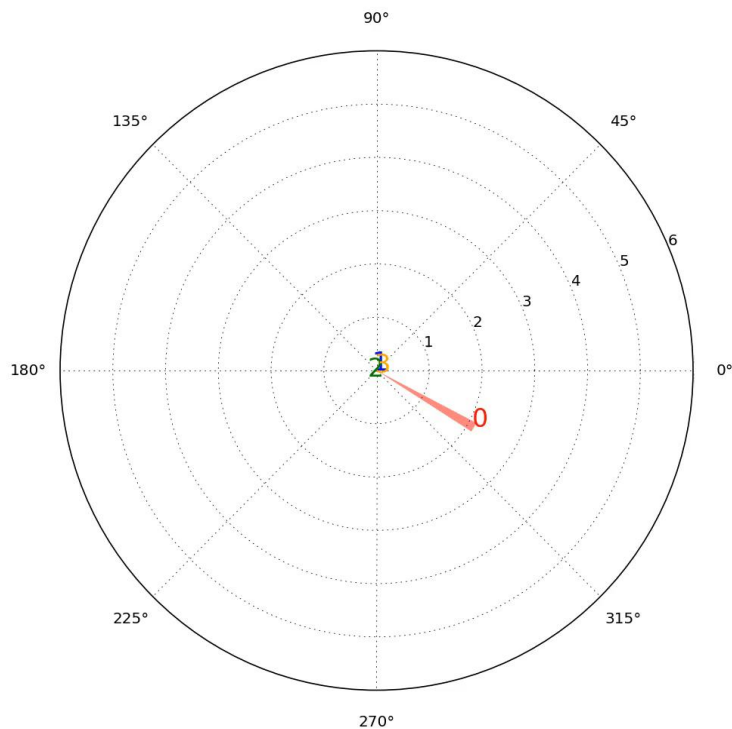
1匹おきタイミング合わせ「ゲツ」
多くのアマカエルがリズムを合わせて鳴き交わす合唱には、知られていなかった法則があったと、理化学研究所脳科学総合研究センターと京都大などの研究チームが27日、科学誌サイエンス・フィジック・レビューに発表した。
理研・京大など発表
カエルは求愛や縄張りを示すために声を合わせて鳴くと考えられるが、鳴き声を聞き分けて記録するのが難しく、具体的な鳴き方はよくわかっていなかった。理研の合原一孝・基礎科学特別研究員と奥乃博・京大教授らは、半径30センチ以内で

カエルが鳴くと光る装置を開発。京都市や臨岐（島根県・屋久島・鹿兒島県）などで、水田のあぜ道に40〜50センチ間隔で40〜60個を並べて観測した。その結果、カエルの雄は0.5〜2センチ間隔で並び、1匹おきの2グループに分かれて、輪唱のように鳴くことがわかった。はじめのグループがタイミングを合わせて「ゲツ」と鳴くと、次は隣のグループが一斉に「ゲツ」と鳴く動作を、交互に繰り返していた。

研究チームは「自分の縄張りをはっきりと主張しやすいように、鳴くタイミングをずらしているのではないか」とみている。

野鳥の鳴き声の解析

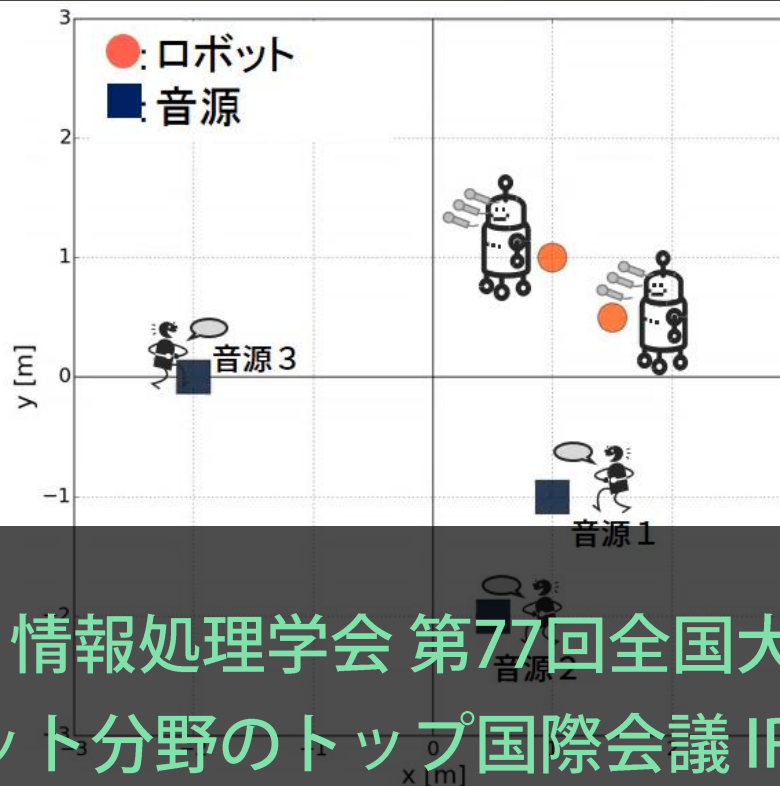
ロボット聴覚システムが実環境で動作！



B4関口君：複数移動ロボット

38

複数ロボットが音源分離しやすい位置へ自律移動



- 音源1
あらゆる現実を
すべて自分のほうへ
ねじ曲げたのだ
- 音源2
一週間ばかり
ニューヨークを
取材した
- 音源3
立春が過ぎても
日々が続く

情報処理学会 第77回全国大会 学生奨励賞
ロボット分野のトップ国際会議 IROS 2015 に投稿済み

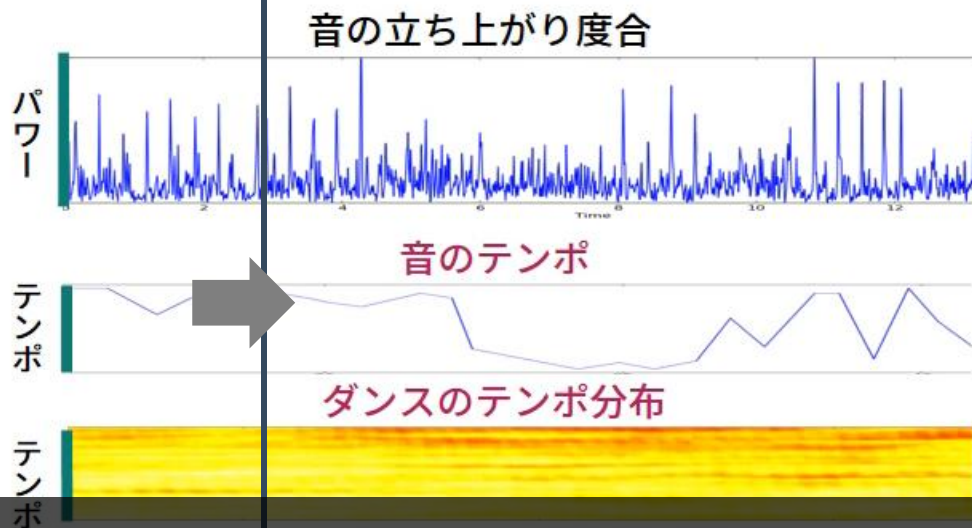
B4大喜多さん：ダンス共演ロボット

39

音楽とダンスから拍に合わせて踊るロボットを開発



骨格情報



情報処理学会 第77回全国大会 学生奨励賞
ロボット分野のトップ国際会議 IROS 2015 に投稿済み

B4板倉君：同時発話音声認識

音源分離と音声認識をベイズモデルで統合

人間のように分離音を作らずに混合音を直接認識できる機構を初めて実現



モンテカルロ積分&マルコフ連鎖モンテカルロ法による近似

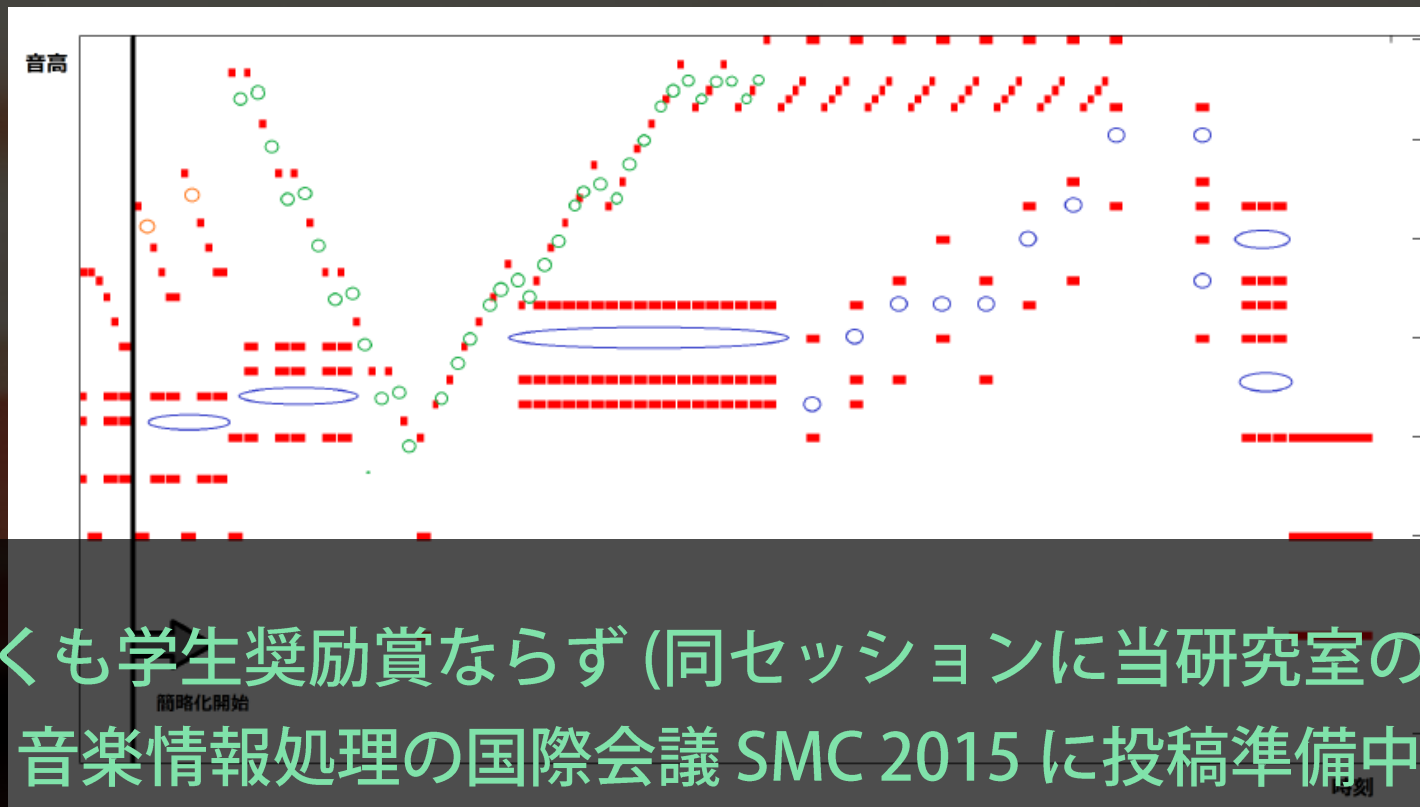
$$p(Z|X) = \int p(Z|S)p(S|X)dS$$

情報処理学会 第77回全国大会 学生奨励賞

音声分野のトップ国際会議 Interspeech 2015 に投稿済み

B4福田君：音楽演奏支援

ピアノ演奏で間違った箇所を簡略化



音声メディア研究室の実情



2014年度卒業/修了生の進路

修士課程修了者



SONY

M2坂東：国際会議4件・英文ジャーナル採録

M2池宮：国際会議4件・英文ジャーナル投稿準備中

M2西牟田：国際会議2件・英文ジャーナル投稿中(推薦)

M2古川：国際会議1件・和文ジャーナル採録(推薦)

2013年度卒業/修了生の進路

博士課程修了者 (学振DC1)



修士課程修了者



全員が円満に就職先を見つけています

よい就職のためには研究能力/プレゼン能力≫就職活動

2012年度卒業/修了生の進路

博士課程修了者 (学振DC1)



Honda Research Institute

修士課程修了者



MITSUBISHI

Changes for the Better

KORG

株式会社コルグ

mixi

過去の博士課程在籍者は**ほぼ全員が学振特別研究員**に採用 (月額20万円の支給)

多数の学生が日本学生支援機構**奨学金の全額/半額免除**を認められています

2011年度 修士課程修了者



2010年度 博士課程修了者 (学振DC1 & DC2)



HITACHI
Inspire the Next

2010年度 修士課程修了者



Panasonic
ideas for life



AsahiKASEI



2014年度 卒業特別研究の流れ

4/14 (火) 研究室配属決定

4/15 (水) 第1回ミーティング・グループ決定・新B4歓迎会

4月：信号処理トレーニング (4/9開始 4月毎週木金)

5月：グループでの演習 (PRML, BRMLなど)

6月～7月：テーマ決定・大学院願書出願

8月：情報学研究科 大学院試験

卒業研究を適宜議論しながら進める

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

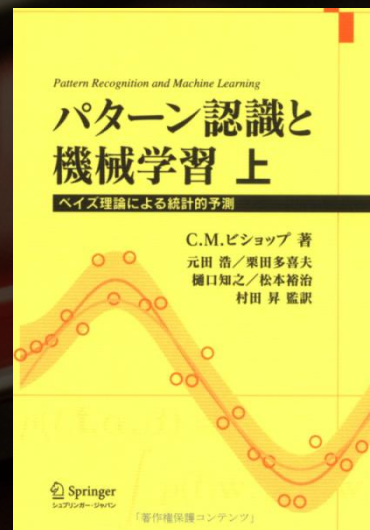
2月中旬：卒業特別研究諮問会

3月：情報処理学会全国大会発表

3月～5月：国際会議投稿 (英文4～6項)

ロボット：IROS 2016 (Daejeon) 音声：Interspeech 2016 (San Francisco)

音楽：ISMIR 2016 (New York)



当研究室の教育サービス

修士課程への進学を前提とした専門教育

3年間一貫教育・博士課程への進学も視野に

初心者到手厚い知識習得カリキュラム

統計的信号処理・機械学習の勉強

信号処理プログラミング (Matlab/C++/Python)

論文/申請書執筆指導・発表練習 (極めて重要)

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

体系的かつ健康的な研究体制

毎週水曜日10:30～18:00 昼食会・研究会

他にグループミーティング (週末は完全休息)

当研究室の駆動原理



論文執筆・プレゼン指導

論文のクオリティをあげきる経験が重要

Before

音声対話システムにおける固有名詞の簡略化に対処する語彙拡張
尾形 哲也¹ 奥乃 博²
¹ 京都大学 工学部情報学科 ² 京都大学大学院 情報学研究科 知能情報学専攻

1. はじめに

音声対話システムにおける音声認識辞書内の語は、正式名称で登録されていることが多い。なぜならそれは一般に、正式名称で構成されたデータベースをもとに構築されるからである。一方でシステムに慣れていない初心者ユーザは、システムが受理可能な語彙がわからない。そのため、正式名称を簡略化した表現で発話することが多い[1]。このような発話は音声認識誤りを引き起こすため、システムの音声認識辞書に簡略表現を追加する必要がある。

この問題への対応は、音声認識辞書内の語には固有名詞が多いことである。なぜなら固有名詞の語の多くは、Webなどの外部の知識から簡略表現を取得することが困難だからである。そこで我々は、固有名詞の簡略表現を得るために、データベースから構築された音声認識辞書(既存辞書)の語の部分列を省略することで対応した。簡略表現の生成・追加のそれぞれの段階で以下の課題がある。

1. 簡略表現生成のための固有名詞分割
簡略表現生成には、既存辞書の単語を省略する要素ごとに分割しなければならない。この際、形態素解析部による分割では、固有名詞を正確に分割できない可能性がある。

2. 生成した簡略表現に起因する既存辞書の語に対する悪影響の抑制
生成した簡略表現を単純に音声認識辞書に加えるのは、語彙の増加により認識率悪化を招いてしまう。特に、既存辞書の語は簡略前の正式名称であるため、正しく音声認識できない可能性がある。それらと音素列が似ている簡略表現は音声認識上混同されやすいため対応が必要である。

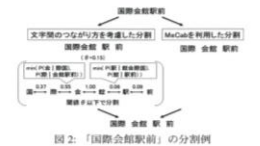
本研究での簡略表現追加までの流れを図1に示す。まず、既存辞書の語の分割と省略により簡略表現を生成する。それらの生成確率を、既存辞書の語との類似度に応じて補正する。

簡略表現を辞書に追加する従来研究[2,3]では、音声認識時に混同される度合いを考慮していない。そのため語彙の増加により、既存辞書の単語に対しての音声認識が悪化する危険性がある。

2. 固有名詞の分割と簡略表現生成

まず、辞書内の単語 w を要素列 $c_1, c_2, \dots, c_n$ に分割する。固有名詞を分割するために、形態素解析部 McCab [4] だけでなく、辞書固有の文字のつながり方を考慮した、一般的な意味の関切りを McCab を用いて抽出し、加えて辞書固有の文字のつながり方を考慮した、二つの抽出結果どちらから区

Vocabulary Expansion to Cope with User's Abbreviated Expression in Spoken Dialogue Systems: Masaki Katsurama, Kanami Katsurama, Tetsuya Ogino, and Hiroshi G. Okuno (Kyoto Univ.)



切られていけば、その位置で単語を切り分け結果とする。辞書固有の文字のつながり方による分割は次のように行なう。ある単語 w は文字列 $c_1c_2\cdots c_n$ で表わされる。単語の分割の尺度として、 c_1 と c_1 との連続度を音声認識辞書内の文字 N -gram 確率を用いて

$$\min(P(c_1|c_2, \dots, c_n), P(c_1|c_2, \dots, c_n)) \quad (1)$$

と定義した。つまり c_1 への前からの N -gram 確率と、 c_1 への後からの N -gram 確率の小さい方の値を、文字 c_1 と c_2 との連続度とした。 c_1 と c_2 との連続度が閾値 θ 以下の場合、 w を c_1 と c_2 の間で区切る。本手法により「国際会館」を分割した例を図2に示す。McCab による分割では「国際」を分割できないが、文字の連続度をもとめることで「国」(と「際」)に分割することができる。

次に、要素列 $c_1, c_2, \dots, c_n$ から任意個の要素の連続性によって、(2 n -1)個の簡略表現を生成した。生成した簡略表現の読みは、要素とこの読みを McCab による読み付与と簡略前の読みとのマッチングを用いて付与した。

3. 既存辞書に悪影響を与える簡略表現の抑制

既存辞書の語を音声認識する上で、既存辞書の語の部分列と音韻的に類似した簡略表現は悪影響を与えやすい。そこで音韻的距離を定義し、その値が近い簡略表現は生成確率を減少させた。

After

音声対話システムにおけるユーザの固有名詞の簡略化に対処する語彙拡張
奥乃 博¹ 尾形 哲也² 奥乃 博²
¹ 京都大学 工学部情報学科 ² 京都大学大学院 情報学研究科 知能情報学専攻

1. はじめに

音声対話システムにおける音声認識辞書内の語は、正式名称で登録されていることが多い。なぜならそれは一般に、正式名称で構成されたデータベースをもとに構築されるからである。一方でシステムに慣れていない初心者ユーザは、システムが受理可能な語彙がわからない。そのため、正式名称を簡略化した表現で発話することが多い[1]。このような発話は音声認識誤りを引き起こすため、システムの音声認識辞書に簡略表現を追加する必要がある。

その問題への対応は、音声認識辞書内の語には固有名詞が多いことである。なぜなら固有名詞の語の多くは、Webなどの外部の知識から簡略表現を取得することが困難だからである。そこで我々は、固有名詞の簡略表現を得るために、データベースから構築された音声認識辞書(既存辞書)の語の部分列を省略することで対応した。簡略表現の生成・追加のそれぞれの段階で以下の課題がある。

1. 簡略表現生成のための固有名詞分割
簡略表現生成には、既存辞書の単語を省略する要素ごとに分割しなければならない。この際、形態素解析部による分割では、固有名詞を正確に分割できない可能性がある。

2. 生成した簡略表現に起因する既存辞書の語に対する悪影響の抑制
生成した簡略表現を単純に音声認識辞書に加えるのは、語彙の増加により認識率悪化を招いてしまう。特に、既存辞書の語は簡略前の正式名称であるため、正しく音声認識できない可能性がある。それらと音素列が似ている簡略表現は音声認識上混同されやすいため対応が必要である。

本研究での簡略表現追加までの流れを図1に示す。まず、既存辞書の語の分割と省略により簡略表現を生成する。それらの生成確率を、既存辞書の語との類似度に応じて補正する。

簡略表現を辞書に追加する従来研究[2,3]では、音声認識時に混同される度合いを考慮していない。そのため語彙の増加により、既存辞書の単語に対しての音声認識が悪化する危険性がある。

2. 固有名詞の分割と簡略表現生成
まず、辞書内の単語 w を要素列 $c_1, c_2, \dots, c_n$ に分割する。固有名詞を分割するために、形態素解析部 McCab [4] だけでなく、辞書固有の文字のつながり方を考慮した、一般的な意味の関切りを McCab を用いて抽出し、加えて辞書固有の文字のつながり方を考慮した、二つの抽出結果どちらから区

3. 既存辞書に悪影響を与える簡略表現の抑制

既存辞書の語を音声認識する上で、既存辞書の語の部分列と音韻的に類似した簡略表現は悪影響を与えやすい。そこで音韻的距離を定義し、その値が近い簡略表現は生成確率を減少させた。



切られていけば、その位置で単語を切り分け結果とする。辞書固有の文字のつながり方による分割は次のように行なう。ある単語 w は文字列 $c_1c_2\cdots c_n$ で表わされる。単語の分割の尺度として、 c_1 と c_1 との連続度を音声認識辞書内の文字 N -gram 確率を用いて

$$\min(P(c_1|c_2, \dots, c_n), P(c_1|c_2, \dots, c_n)) \quad (1)$$

と定義した。つまり c_1 への前からの N -gram 確率と、 c_1 への後からの N -gram 確率の小さい方の値を、文字 c_1 と c_2 との連続度とした。 c_1 と c_2 との連続度が閾値 θ 以下の場合、 w を c_1 と c_2 の間で区切る。本手法により「国際会館」を分割した例を図2に示す。McCab による分割では「国際」を分割できないが、文字の連続度をもとめることで「国」(と「際」)に分割することができる。

次に、要素列 $c_1, c_2, \dots, c_n$ から任意個の要素の連続性によって、(2 n -1)個の簡略表現を生成した。生成した簡略表現の読みは、要素とこの読みを McCab による読み付与と簡略前の読みとのマッチングを用いて付与した。

3. 既存辞書に悪影響を与える簡略表現の抑制

既存辞書の語を音声認識する上で、既存辞書の語の部分列と音韻的に類似した簡略表現は悪影響を与えやすい。そこで音韻的距離を定義し、その値が近い簡略表現は生成確率を減少させた。

「国際会館」の簡略表現の生成に必要で必要な要素。→「国際」は必要で必要な要素。→「国際」は必要で必要な要素。

論文執筆指導

卒論を英語化して国際会議に行こう！

◎ 限定用法の主格関係代名詞には、which ではなく that を使う。
(本格的には which も OK だが、慣例 & disambiguation ため)

④ baseline method は何なの？ 4月

ing ad words zero
ate abbreviations. However, general morphological analysis tools can not correctly
segment proper nouns correctly.
2. Reducing recognition errors caused by added abbreviations ing ad words into ASR dictionary
Adding generated abbreviations often reduces recognition errors, because
ok the vocabulary size increases. Abbreviations whose phonemes is similar to those of
other words in the original dictionary might be misrecognized as the word
in the original dictionary. are close
To solve the former problem, we segment words by using morphological analyzer
analysis and conjunction probability among characters. For the latter problem, issue
we change occurrence probabilities of generated abbreviations according to the
phonological similarity between abbreviations and words in original dictionary. ed words
We evaluated our method by using 1494 utterances of real users of Kyoto
city bus information system. Recall of generating abbreviation by our method
is 94.3%. On the other hand, that of base line method is 85.7%. This shows ed
the superiority of our segment approach using conjunction probability. We also evalu-
ated ASR accuracy of this data. For utterances which contain abbreviations,
our method increases ASR accuracy by 24.2 point from baseline. For utterances
which contain words which baseline can recognize, our method increases ASR
accuracy by 0.1 point from baseline. The method which only adding
abbreviations increased ASR accuracy by 15.7 points. These results
show that our approach of changing probabilities of generated abbreviations
reduced ASR error caused by added abbreviations. results
that our method can recognize user's abbreviations with reduced recognition errors.
was simply added to the ASR dictionary
utterance data of novices who used the system only by
without increasing degradation of ASR accuracy

ロボット聴覚
音楽情報処理
においては
日本が世界を牽引！

日本から投稿される情報系の論文は、主要な国際会議にほとんど通りません。国際コミュニティにおける日本人のプレゼンスは非常に低いです。ほとんどの論文が十分に完成度が上がらないまま投稿され、不採択になっています。一般的な国際会議では通ることがあっても、グレードの高い国際会議に完成度が低い論文が通ることはまずありません。完成度が上がらないこと、これが日本の情報系の論文が海外で通用しない最大の理由です。

松尾豊先生@東大のHPより抜粋

http://ymatsuo.com/japanese/ronbun_jpn.html

http://ymatsuo.com/japanese/ronbun_eng.html

当研究室B4/B3の受賞歴

IEEE/RSJ IROS (ロボットのトップカンファレンス)

2006 Best Paper Nomination Finalist (4/950)

2008 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

大会奨励賞 (大学卒業10年以内で年10件程度) 3名

学生奨励賞 H26:3名 / H25:1名 / H24:2名 / H23:3名 など多数
財団海外研究発表助成・奨学金全額/半額返済免除 (大学院)

発表論文・受賞とも京大でも全国でもダントツ！

教員の受賞歴

教員は分野を代表するアクティブな研究者

The 14th Int. Society for Music Information Retrieval Conference (ISMIR 2013) Best Oral Presentation Award, 2013/11/7.

電子情報通信学会 第15回情報論的学習理論ワークショップ (IBIS 2012) Honorable Mention, 2012/11/8.

情報処理学会 インタラクション 2012 ベストペーパー, 2012/03/15.

電子情報通信学会 第13回情報論的学習理論ワークショップ (IBIS 2010) ポスター奨励賞, 2010/11/5.

船井情報科学振興財団 船井研究奨励賞, 2010/4/17.

情報処理学会 音楽情報科学研究会 夏のシンポジウム ベストプレゼンテーション賞, 2009/07/31.

Award for Entertainment Robots and Systems (IROS 2008 NTF Award) Nomination Finalist, 2008/9/24.

Award for Entertainment Robots and Systems (IROS 2008 NTF Award) Nomination Finalist, 2008/9/24.

IPSJ Digital Courier 船井若手奨励賞, 2008/4/19.

電気通信普及財団 テレコムシステム技術学生賞, 2007/3/19.

情報処理学会 山下記念研究賞, 2007/3/6.

IEEE関西支部 学生研究奨励賞, 2007/2/6.

情報処理学会 音楽情報科学研究会 夏のシンポジウム ベストプレゼンテーション賞, 2006/08/08.

情報処理学会 インタラクション 2006 ベストインタラクティブ発表賞, 2006/03/02.

The 1st Annual Music Information Retrieval Evaluation eXchange (MIREX 2005) Best-in-Class Award, 2005/9/14.

情報処理学会 全国大会 学生奨励賞, 2005/03/03.

第3回情報科学技術フォーラム (FIT 2004) 論文賞, 2004/09/07.

アウトリーチ活動

研究成果を一般に向けて積極的に発信

NEWS SCAN

ROBOTICS

Playing It by Ear

A machine-listening system that understands three speakers at once BY TIM HORNYAK

Prince Shotoku was a seventh-century politician attributed with authorship of Japan's first constitution. Famed as a nation builder, he is said to have been able to listen to many people simultaneously, hearing the petitions of up to 10 supplicants at once and then handing down judgments or advice.

Inspired by the legendary prince, Japanese researchers have spent five years developing a humanoid robot system that can understand and respond to simultaneous speakers. They posit a restaurant scenario in which the robot is a waiter. When three people stand before the robot and simultaneously order pork cutlet meals or French dinners, the robot understands at about 70 percent comprehension, responding by repeating each order and giving the total price. This process takes less than two seconds and, crucially, requires no prior voice training.

Such a system could be used to help mental health professionals to understand how to communicate with patients. It could also be used to help people with hearing impairments to understand what others are saying. It could also be used to help people with hearing impairments to understand what others are saying.

Sounds from the other speakers are audible. The result is a robust robot listener that is closer to the human brain's auditory powers than other systems. Okuno says it might handle as many as six talkers depending on their relative angles and the number of microphones utilized (currently eight). The robot can move and orient toward speakers, too, thereby enhancing performance.

"Okuno's project for robots that can understand overlapping voices does a really nice job of combining the best ideas in multichannel source localization with the powerful technique of missing-feature speech recognition," remarks Dan Ellis, head of Columbia University's Laboratory for the Recognition and Organization of Speech and Audio. "What makes his work stand out from anything else is the commitment to solving all the practical problems that come up in a real-world deployment ... and making something that ... can enable a

Now hear these: Hiroshi G. Okuno poses with Robovie R2 and SIG2, machines that

BS JAPAN
ECONOMY NEWS
INFO/CULTURE
ENTERTAINMENT
MOVIE/DRAMA

世の中は変わる
毎週金曜 夜10時24分

放送内容
■第123回[11/02/04]
■『盲者の音を聞き分ける聖徳太子のようなロボットが登場！』

「日本全国、進歩の足跡」。今回は聖徳太子の地、京都で誕生した最新の技術を紹介。京都大学奥の研究室では、なんと聖徳太子のようロボットが開発されたという。複数の人の声を聞き分けたり、聖徳太子のよう、同時に複数の人の声を聞き分けたりできるロボットなのだ。さらに奥の研究室では、音の聞き分けに関する様々な研究を行っている。音を聞き分けるシステムも開発し、様々な分野で役立てようとする「聖徳太子コンピュータ」の全貌に迫る。

聖徳太子ロボット テルミン演奏ロボット

放送 ニッポン放送 >> オールナイトニッポン >> SuonoDolce NOW PLAYING 2A

KYOCERA 近未来story

中川翔子のサイエンス!

FRIDAY 24:30-25:00 ON AIR!!

ON AIR > KBS京都 Radio FRIDAY 24:30-25:00 > ラジオ大阪 QRC TUESDAY 23:30-24:00 > KBS

今日の研究内容

- 09.03.27 九州大学大学院「証言を聞く」
- 09.03.28 東京理科大学「アポロロケット」
- 09.03.13 京都大学大学院「プラナリアの宇宙旅行」
- 09.03.09 筑波大学大学院「後世の歴史メディア」

最近の研究内容

研究内容

参加ゼミ

2008.07.25 放送
京都大学大学院 情報学専攻
知能情報学専攻
奥乃 研究室

聖徳太子は10人の訴えを同時に聞き分けたという。奥乃研究室では、「聞き分ける」をキーワードに、「聖徳太子」ロボットの開発、多重音源からの発音やボーカルの聞き分け、音楽を聴いて口ずさむ音楽ロボット、システムが発話中にユーザが発話する「バースキン」が可能な音声対話システムなどの研究を行っています。

本日

聖徳太子ロボット

聖徳太子ってなんでも、しなやかな、今なら、テレビで、京都大学、そこは、研究、続けている。

神田 高
水本 武志
藤原 真樹
の3人が、神田さん、気がつけば、

久慈 隆雄 監修 / 藤井 大 制作

収録 藤井 大 / 藤井 大

01 02 03 04 05

PREV

京都大学奥の研究室
さま
Hoko
またん
ギガ
Gサイエンス

KYOCERA 近未来story

中川翔子のサイエンス!

JOLF AM1242

制作 久慈ニッポン放送

聖徳太子ロボット テルミン演奏ロボット

研究テーマ案 (応相談)

各自の自由な発想を最も重視します！

ロボット聴覚

聖徳太子ロボットによる複数人との音声対話
移動ロボットと飛行ロボットの協調型音環境理解
音楽をリアルタイムで認識するロボット

音楽情報処理

音源分離・自動採譜・残響処理・楽譜追跡
音高・音長・音色の操作による二次創作
推薦・検索・可視化のためのインタフェース

統計的機械学習

ノンパラメトリックベイズ
ディープラーニング

趣味としての研究テーマ

習得した知識を生かした副業のススメ

研究室内ハッカソンの不定期開催

〇〇と××を組み合わせてみたら？

新しいWEB APIが出たから早速使ってみよう
おもしろアプリ大歓迎！（深い研究の芽生えかも）

パズル解法

16ヒントの数独問題の不存在証明

ZBDDによる数独問題難易度の定義

ルール適用だけで数独を解く

動物音響学

カエル・鳥の鳴き声認識（農学・総人等から）

研究室見学会の内容

ロボット聴覚グループ

音楽ロボット・同時発話認識

飛行ロボット・ホース型ロボット

音楽情報処理グループ

音楽音響信号の解析・加工・編集

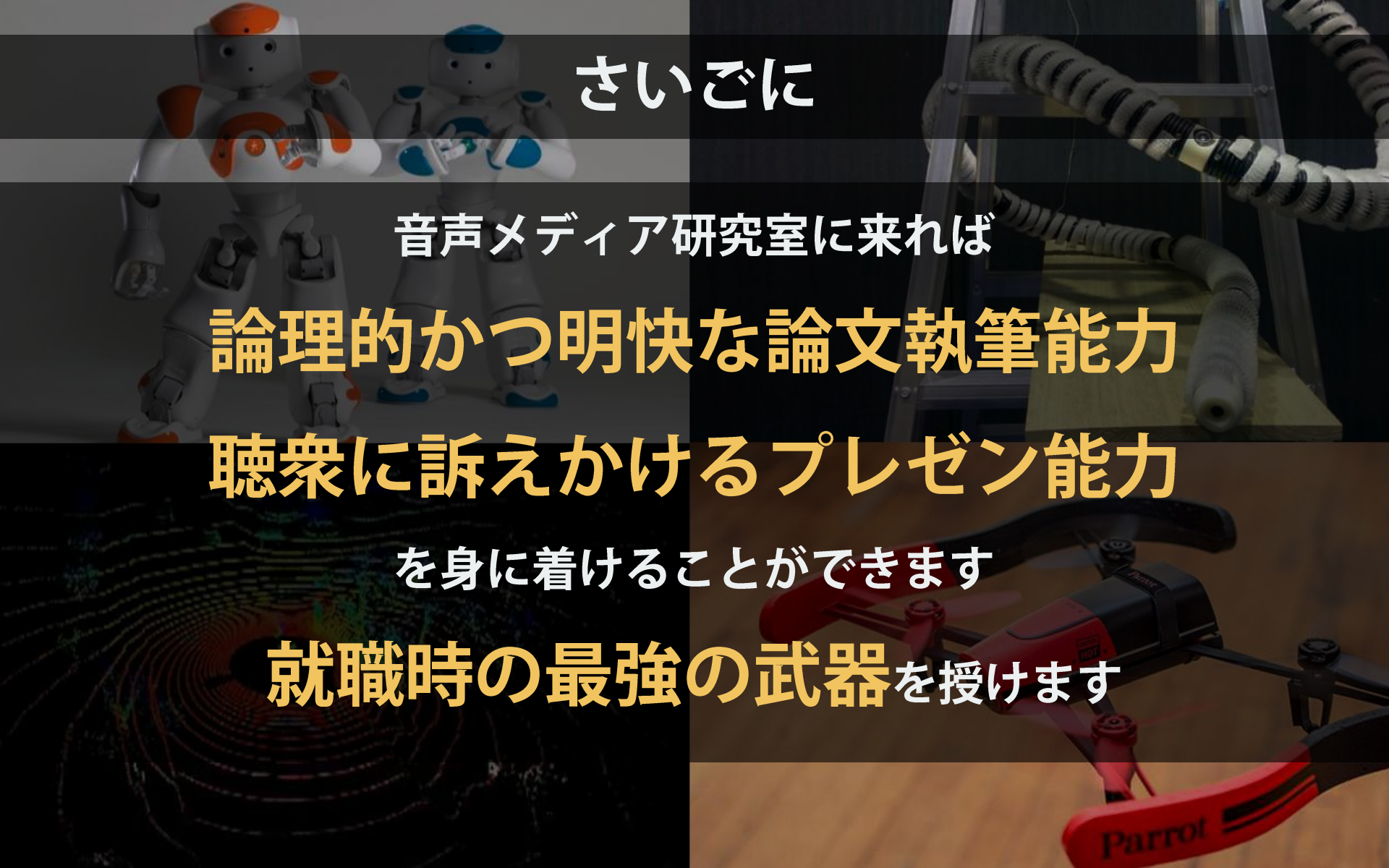
歌い方の解析・転写

懇談会：総合研究7号館 4階 談話室 本日18:00開始

飛び入り大歓迎・無料 昨年は10名以上の参加あり！

研究室メンバーの生の声が聞けます

個別相談は随時受け付けています



さいごに

音声メディア研究室に來れば

論理的かつ明快な論文執筆能力

聴衆に訴えかけるプレゼン能力

を身に着けることができます

就職時の最強の武器を授けます



さいごに

あなたの手にする
研究室希望調書は

当研究室への招待状

若返って勢いのある研究室
だからこそ面白い！

<http://winnie.kuis.kyoto-u.ac.jp/>

研究のデモ
ビデオが
いっぱい

研究室
メンバー
情報が充実



B4の
卒論発表の
ビデオ

奥乃研
旧ページも
参考に

B4 & M2 計算機コース謝恩会



当研究室主催BBQ@八瀬

