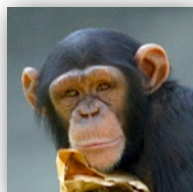


指導教員	峯松信明教授	研究場所	本郷・工学部2号館10階
------	--------	------	--------------

～ヒトに学び、ヒトを越える次世代音声情報処理の実現に向けて～

峯松研究室では、**音声**を主たる媒体とした**コミュニケーション**に関する**基礎研究・応用研究**を行なっています。音声は音、即ち、空気粒子の振動現象でしかありません。ですが、この振動現象を鼓膜が捉えると、あら不思議、その振動から様々な情報を人は抽出します。誰が、何を、どのように、喋った・歌ったのか、といった情報を抽出します。日頃、当たり前のように行なっている空気振動を媒体とした情報抽出・生成、実は、分かっていないことが沢山あります。

チンパンジーとヒトは遺伝子で言えば2%弱の違いしかありませんが、情報処理という観点から見ると何が違うのでしょうか？霊長類研の方々に聞くと「視覚の世界は両者で共通している。でも、聴覚の世界は別世界」との答えも返ってきます。サルとヒトの共通の祖先に対して、どのような情報処理能力が付与されると音声言語や音楽が生まれるのでしょうか？研究テーマの一つは、このような観点からヒトの行なう音声情報処理を計算機の上に実装することに有ります。その一方で、計算機には計算機のやり取り、クラウドコンピューティング、スマートな方法論が好きな学生な学生もいるでしょう。基礎研究、の興味とやる気に合わせてテーマ設



方、流儀もあります。ビッグデータ機械学習に代表される力技がそれでも、ガンガン計算しまえ！が好き
応用研究どちらが好きですか？各々定が可能です。幾つか紹介します。

■ 観測対象に内在する情報の不変的構造表象とその数理的モデリング ～アフォードンスを数理的に定義する～

環境から受け取る信号（視覚・聴覚・触覚・・・）は様々な変形を被ります。しかし、我々はこれらの変形前後の刺激に対して同一性を感じることができます（知覚の恒常性）。この問題に対する一般解を数学（位相幾何学・トポロジー）的に導出し、様々なメディア情報処理への応用を検討しています。変形不変の情報表象を峯松研では提案しています。

■ 話者の違い、年齢の違い、環境の違いに頑健な音声認識技術の構築 ～これ、サルには難しいんだな～

音声ストリームを対象として不変構造抽出を計算機実装すれば、話者の違いや環境の違いに対して不変な音パターン抽出技術が構築されます。逆に、雑音が混入された音声を逐一クリーン音声に戻す処理を実装することでも頑健な音声認識は可能です。様々な観点・方法論から、環境の違いに頑健な音声認識技術の構築とその応用を検討しています。

■ 音声からの年齢推定・女性度推定・感情推定 ～音声の中の様々な情報を推定してみよう～

音声認識は、音声から単語列を見つける作業を言いますが、音声には他にも様々な情報が含まれており、その情報を推定する技術の構築を行なっています。年齢、女性度、感情などなど。女性度の推定は、性同一性障害者（MtF話者）のボイストレーニングに応用されています。貴方は音声からどんな情報をとり出したいですか？

■ 外国語の発音評価システムの構築と教育現場での実践的活用 ～世界の人々を繋ぐ外国語教育支援～

先生の声と生徒の声は音としては大きく異なります。年齢・性別・体格などに影響を受けない、発音分析手法を提案し、世界中の英語発音を話者を単位として分類しています。最近注目されているシャドーイング発声に対する評価技術の構築や、また、世界中の日本語教育の現場で使われているアクセント辞書（OJAD）の開発も行なっています。

■ 話者変換／メディア変換／何でもかんでも変換しちゃえ ～手から声を出す技術の構築／音色テルミンの構築～

貴方の声を別話者の声に変換する技術を構築しています。多様なキャラクタ声を出す声優の声を使って、貴方の声を色々なキャラ声に変換します。声空間と声空間の写像を設計する訳ですが、声空間と手の運動空間の写像を設計すれば、手の運動からダイレクトに声が生成されます。貴方の手に喋らせることだってできちゃいます（構音障害者支援）。

■ 最後に ～学生さんへのメッセージ 「さあ、はじけちまおうぜい！」～

峯松研は、よく学び、よく語り、よく遊び、よく食い、よく飲む、そういう学生を求めています。先輩には、研究科長賞をとった方もいます。問うて考え、基盤技術を構築し、最後は社会貢献する。これが我々のモットーです。

