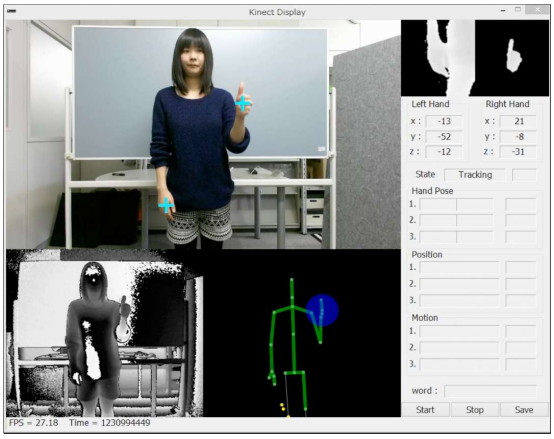
くらし
における喜び・驚き

L-02 聴覚障害者とのコミュニケーションを支援する自動手話単語認識技術

情報工学専攻　准教授　酒向 慎司

概　要

音声処理技術を活用して人の動作を認識



- 音声認識の要素技術を取り入れた手話の自動翻訳技術を構築
- 聴覚障害者とのコミュニケーションを支援

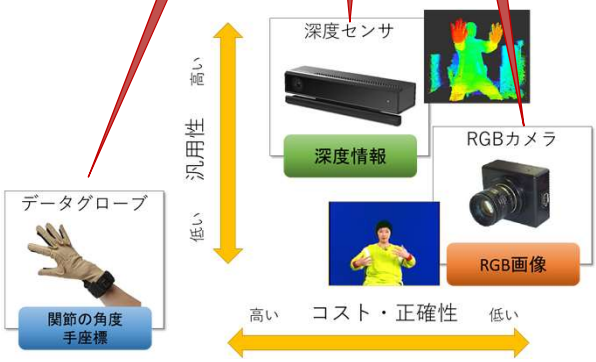
特　長

手話の音韻構造に基づいて分類し、その組み合せから語彙（単語）を認識する（サブユニットモデル）

- 語彙の拡張性に優れている
- 多量の手話データによる事前学習を必要としない
- 特定のセンサに依存しない汎用性

	認識精度	拡張性（汎用性）	データ収集量
サブユニットモデル（本技術）	△	○	○ 共通パターンから類推可能
単語毎の個別認識モデル（従来技術）	△	×	× 多量の学習データが必要

精度やコストの違いに応じた様々なセンサを選択可能！



実用化イメージ と もたらされる喜び・驚き

- 障害の有無によって分け隔てられることなく、
  すべての情報が平等に行き渡る超スマート社会の実現
- 手話分野を超えた、ジェスチャや身体動作のパターン認識への幅広い応用
（工場内など騒音環境の意思疎通、熟練工の技能抽出など）

今後の課題

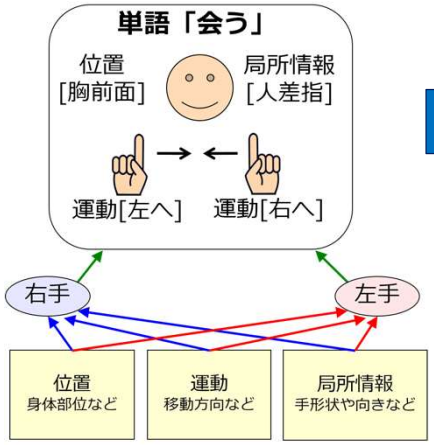
- 日常生活レベルの語彙数（2000語彙以上）への対応
- 手話者の「個人差（微妙な表現の違い）」に対する頑健性への対応
- 手話対話キオスク端末の開発

求める連携先 と メッセージ

- 情報保障や支援技術の他、ジェスチャや身体動作のパターン認識の活用に関心をお持ちの企業との連携を期待します
- 音声研究（音声認識・音声合成・音声対話など）の技術を視覚言語（手話）に応用した研究を進めています

特長が発揮される仕組み

- 手話の音韻構造：「3種の要素」が「両手」で表現され、空間内で“同時に”表出される
- 複雑な組み合わせで構成されるので、語彙（単語）ごとにモデルを作成するのは膨大で困難
- 音韻要素の共通性を利用した効率的な単語モデル表現：**手話の音韻構造に基づいたサブユニットモデル**



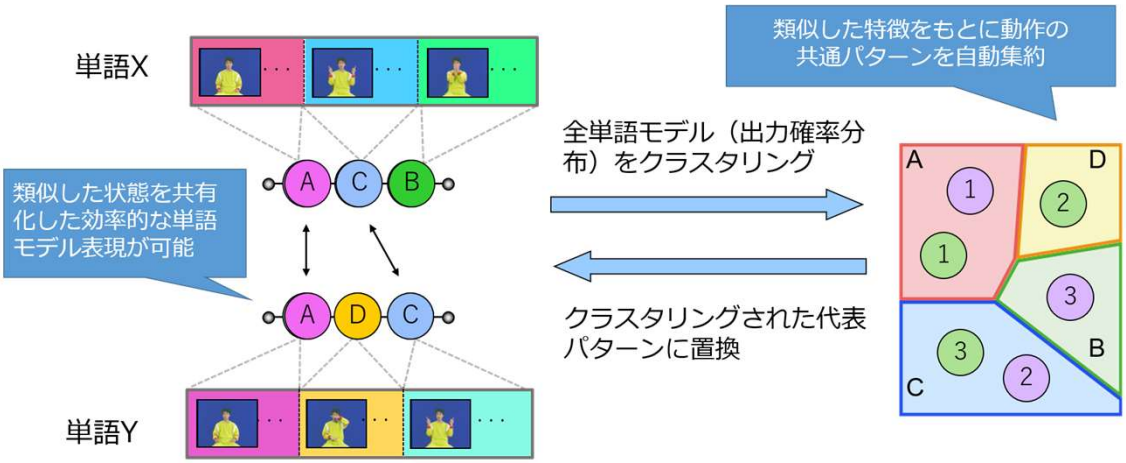
手話を構成する3要素

試作したリアルタイム単語認識システム（デモ有）

サブユニットモデル生成の概念と効果

- サブユニットモデルの効果：従来技術（単語単体のモデル）より高い単語認識率
- 音韻構造の導入の効果：サブユニット化の効果をさらに高める効果

⇒ **従来技術よりも小さなフットプリント（モデルデータのサイズ）で単語認識率を約１０％向上**



試作品の状況

提示可

研究フェーズ

基礎固め　実用性評価

1　2　3　4　5

原理検証　開発研究　技術移転可

文献・特許の情報

- B. Kwolek, and S. Sako, “Learning Siamese Features for Finger Spelling Recognition”, Advanced Concepts for Intelligent Vision Systems Sep. 2017.
- H. Hosoe, S. Sako, and B. Kwolek, "Recognition of JSL Finger Spelling Using Convolutional Neural Networks", 15th IAPR International Conference on Machine Vision Applications, 04-12, pp.72-75, May. 2017.

【お問合せ】名古屋工業大学　産学官金連携機構

〒466-8555　名古屋市昭和区御器所町字木市29番

TEL:052-735-5627　FAX:052-735-5542

E-mail: nitfair@adm.nitech.ac.jp　URL: http://technofair.web.nitech.ac.jp/