

# 社会的コミュニケーション能力 (スキル) の推定技術の現状と課題

岡田 将吾

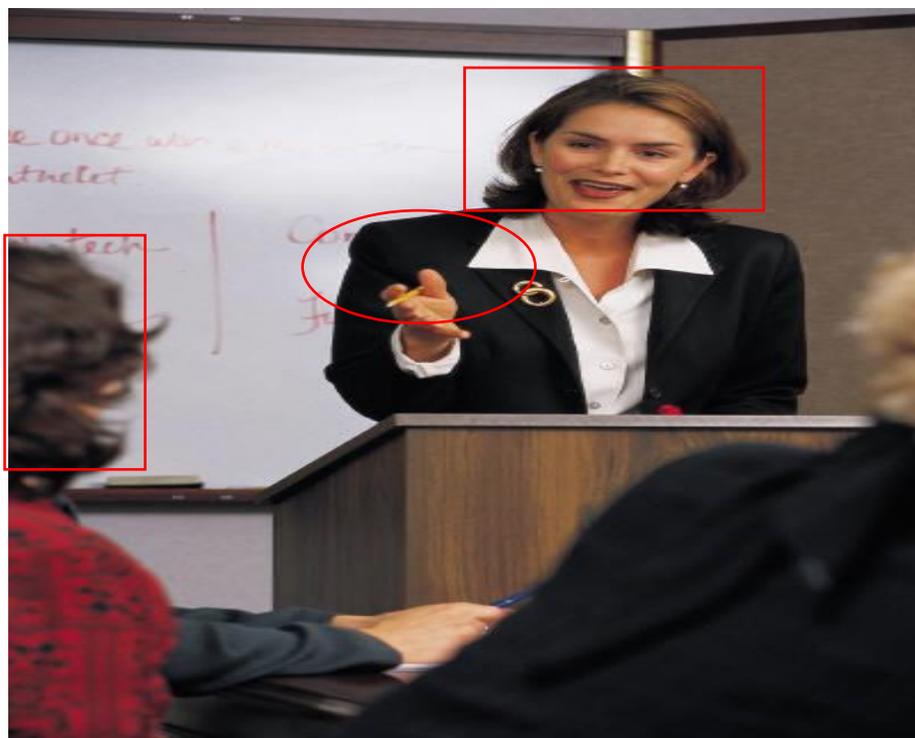
[北陸先端科学技術大学院大学 先端科学技術研究科 准教授]

# 目次

- 背景：社会的信号処理とコミュニケーションスキルの推定
- 研究紹介
  - ビジネスプレゼンスキル推定とフィードバック
  - その他のコミュニケーションスキル推定に関する研究成果
- 課題とまとめ

# 社会的信号処理 (Social Signal Processing [Vinciarelli et al. 2009])

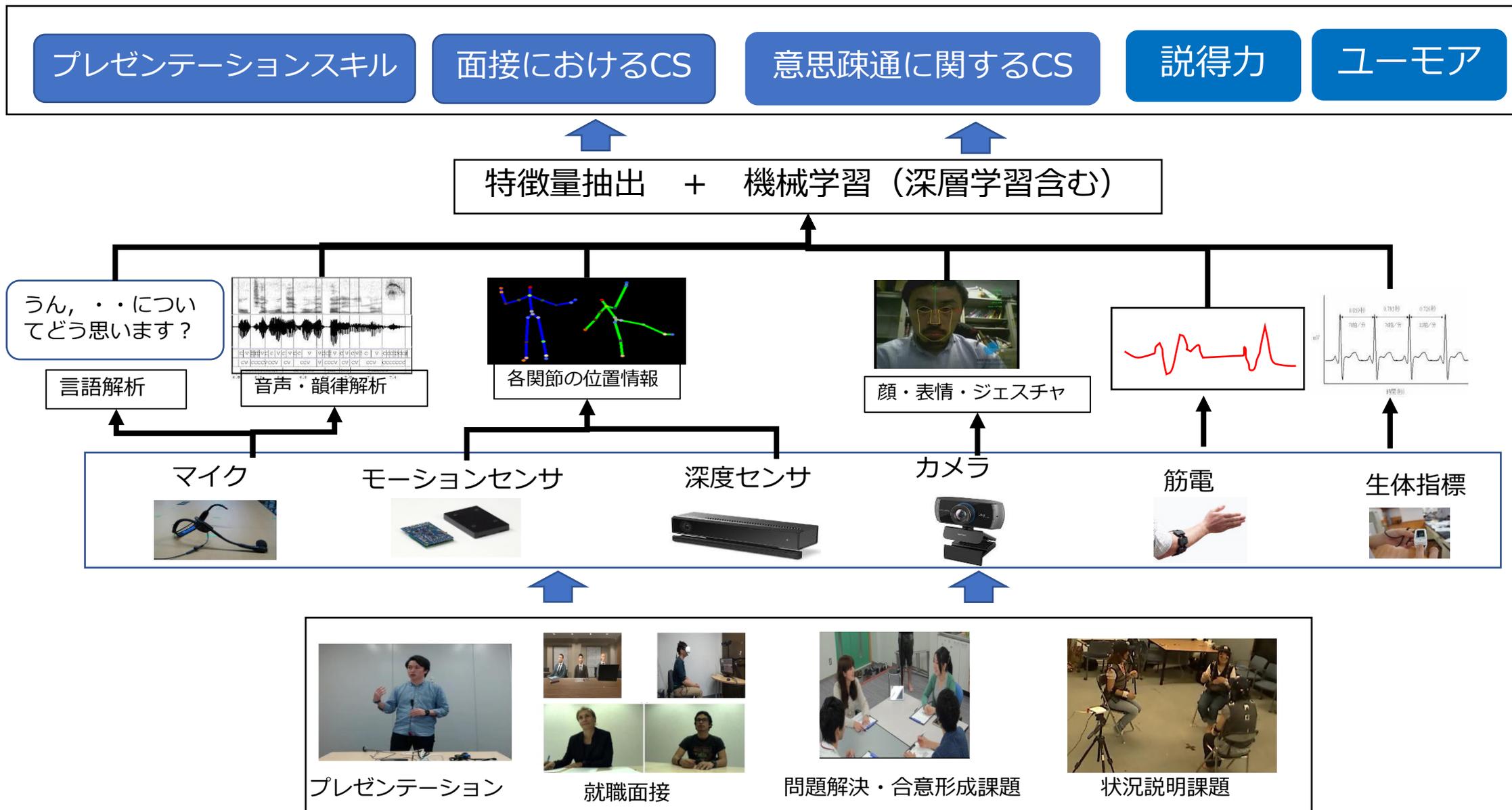
言語・非言語情報を通じて表出する**意図**, **態度**, **感情**, **スキル**などの**内面**の推定・生成



**入力変数**：音声，言語，視覚，生体情報

**目的変数**：同意，性格特性，リーダーシップ，傾聴，コミュニケーションスキル

# 社会的信号処理に基づくコミュニケーションスキル (CS)の推定



# コミュニケーションスキルの推定に関する関連研究

## • 対話タスクによる分類

- プレゼンテーション

[ICMI2020], [Vikram, ICMI2015], [Lei, ICMI2014]

- ソーシャルメディア [Sunghyun, ICMI2017]

- 面接 [Nguyen, IEEE2014]

- グループ対話

[Dairazalia, IEEE TMM2012], [Okada, ICMI2016], [Ishii, ICMI2018]

- エージェントとの対話 [Hoque,

UbiComp2013], [Tanaka, IUI2015], [Tanaka, ICMI2018]

## • スキルタイプによる分類

- 説得力 [Sunghyun, ICMI2017]

- リーダーシップ [Dairazalia, IEEE2012]

- 対人コミュニケーションスキル

[Okada, ICMI2016], [Nguyen, IEEE TMM2014], [Tanaka, IUI2015]

- 共感スキル [Ishii, ICMI2018]

- リスニング (傾聴)スキル [Tanaka, ICMI2018]

- スピーキングスキル

[Torsten, ICMI2015], [Vikram, ICMI2015], [Lei, ICMI2014]

**応用** : 教育・訓練のための支援 (e.g. SST) , HRテック

# 目次

- 背景：社会的信号処理とコミュニケーションスキルの推定
- 研究紹介
  - ビジネスプレゼンスキル推定とフィードバック
  - その他のコミュニケーションスキル推定に関する研究
- 課題とまとめ

# ビジネスプレゼンテーションスキルの推定と ユーザへのフィードバック生成

SoftBank University



マルチモーダル特徴量

- 音声情報
- 画像情報
- 言語情報



評価



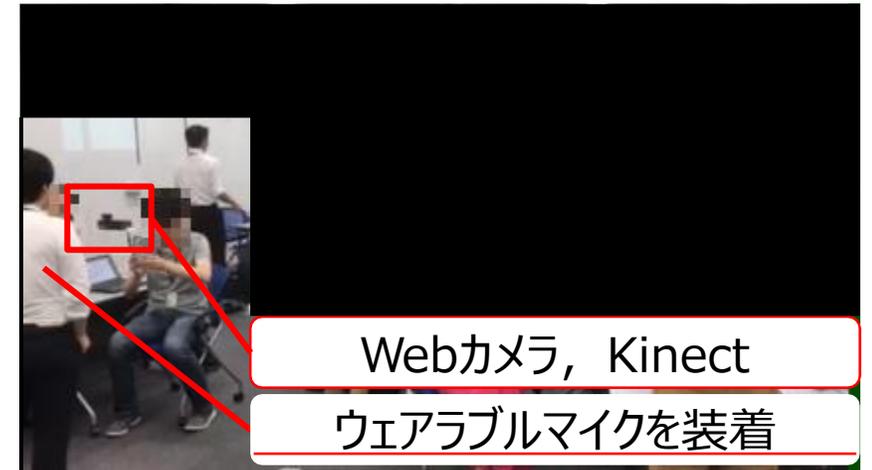
学習・評価

## 研究目的：ビジネスプレゼンテーションスキルのモデル化

- ① ビジネスプレゼンを対象とする研修プログラムでの実データの収集
- ② 言語・非言語特徴量を用いた多面的スキルの推定
- ③ マルチモーダル共変量適応型回帰モデルの有効性検証
- ④ スキル改善のためのフィードバックシステムの開発

# ビジネスプレゼンターデータセットの概要

- データ取得 ■ SoftBank University
  - Softbank University 「プレゼン話し方研修」にて取得（年7回）
  - 参加者にビジネスプレゼンテーションの基礎を教育する目的
  - 実験協力者はビジネスパーソン
- プレゼンテーションスキルの評定
  - 人事研修のエキスパート2名が評価を実施



# 評価項目の作成

## 総合的プレゼンテーションスキル

### 目的網羅性

- 誰に
- 何を
- どうしてほしい

### 内容ロジック

- 結論
- 根拠
- 相手の利益

### 効果的演出要素

- 強調
- 繰り返し
- 具体表現
- 双方向性

### 話し方基礎

- 抑揚
- 声量
- アイコンタクト
- ジェスチャ
- 表情

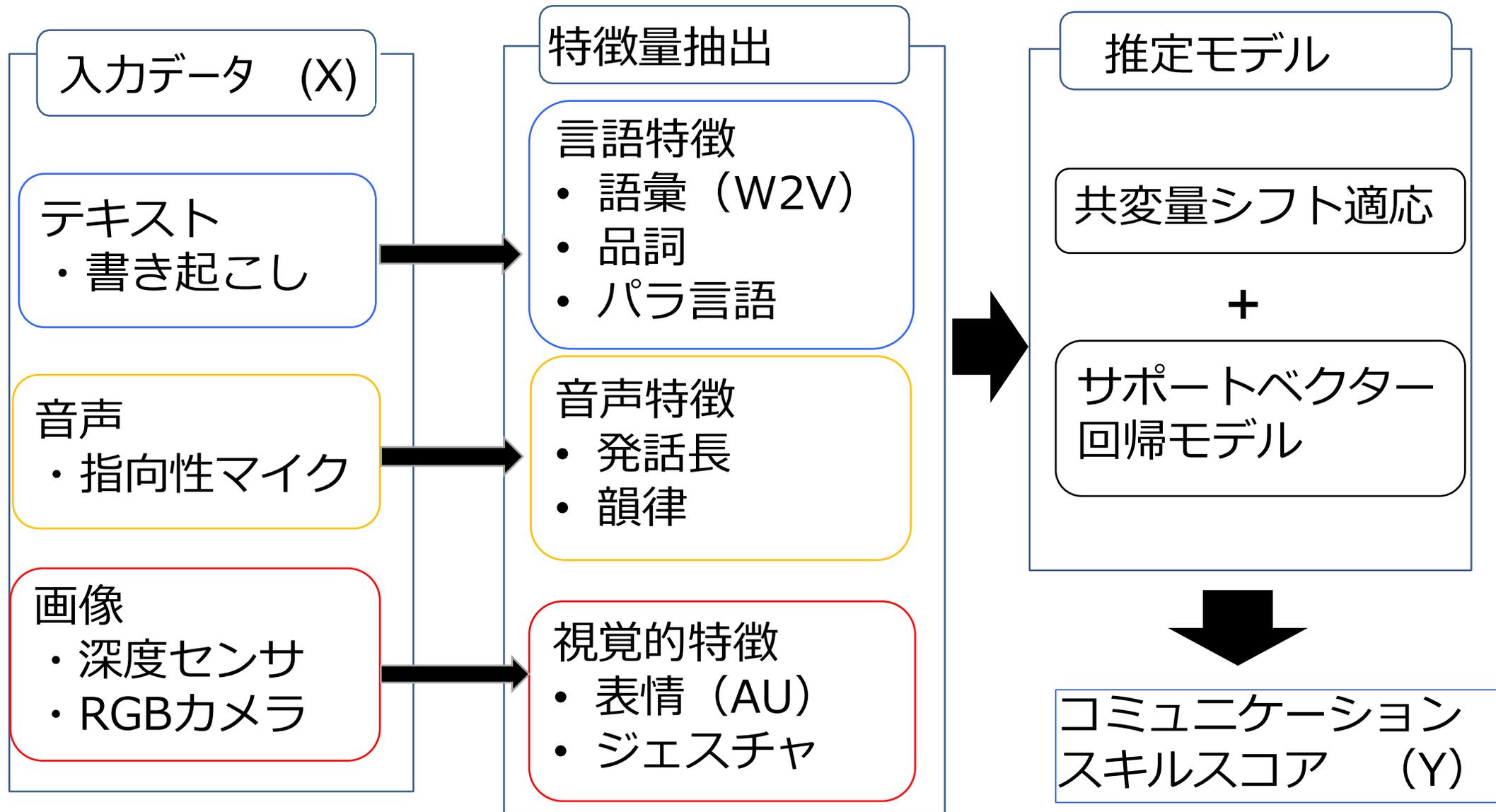
## アノテーション手順

- 人事採用経験者2名が各プレゼンターを評価
- 小項目を0,1,2の3段階で評価
- 2名の合計点 → 正解データ

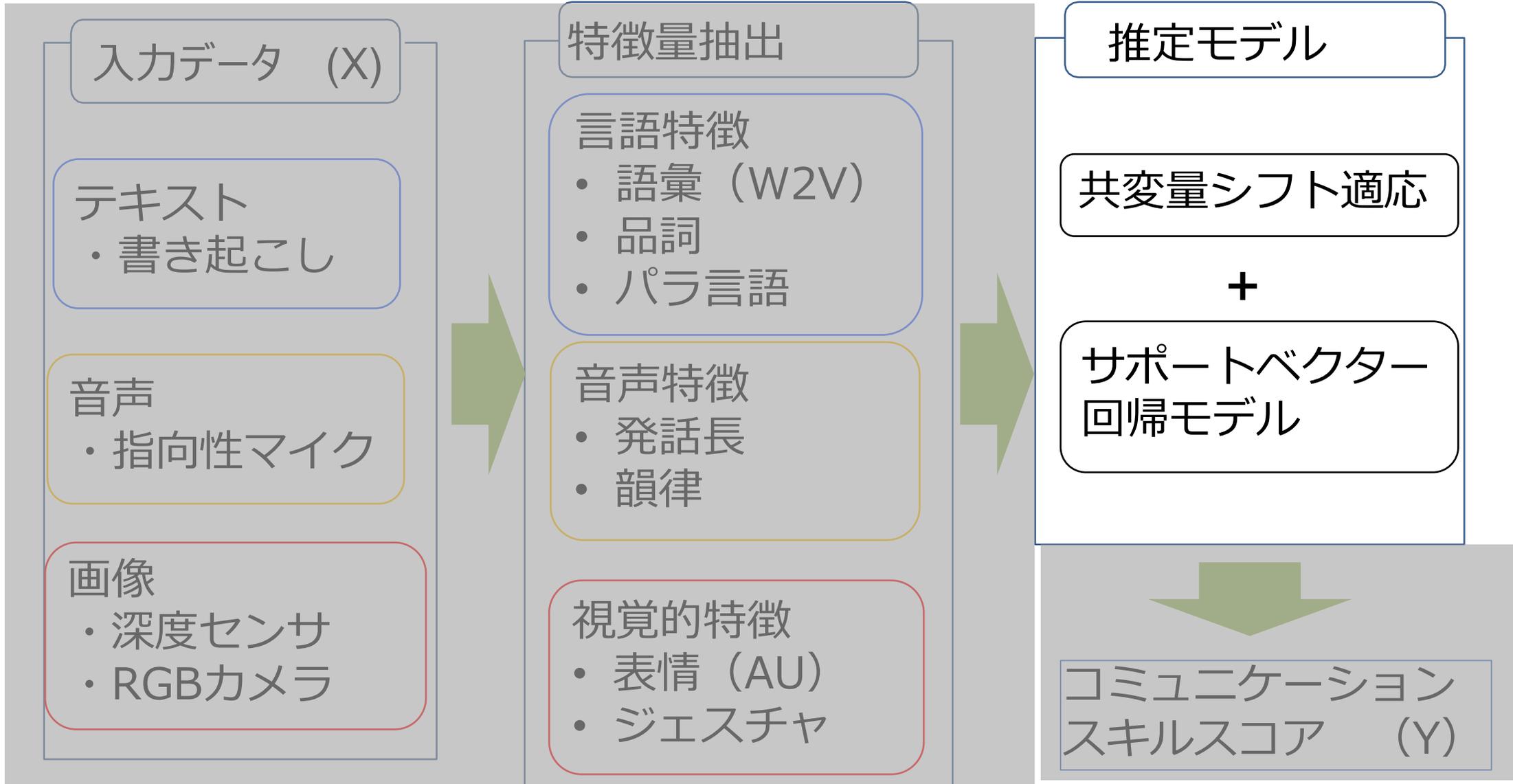
## 一致度算出

評価項目	重み付きカッパ係数
目的網羅性	0.45
内容ロジック	0.61
効果的演出要素	0.86
話し方基礎	0.72

# コミュニケーションスキルの推定手順



# コミュニケーションスキルの推定手順



# プレゼンテーションスキルの推定モデル

音声・言語・視覚情報

プレゼンテーションスキルスコア



- 発話中の韻律
- 言語内容
- ジェスチャー
- etc...

プレゼンテーションスキル **高**

## マルチモーダル情報に基づく推定における問題

- プレゼンターのマルチモーダル情報は多様かつ個人差を含む。
- 未知のプレゼンターのデータが訓練データの分布と一致するとは限らない。

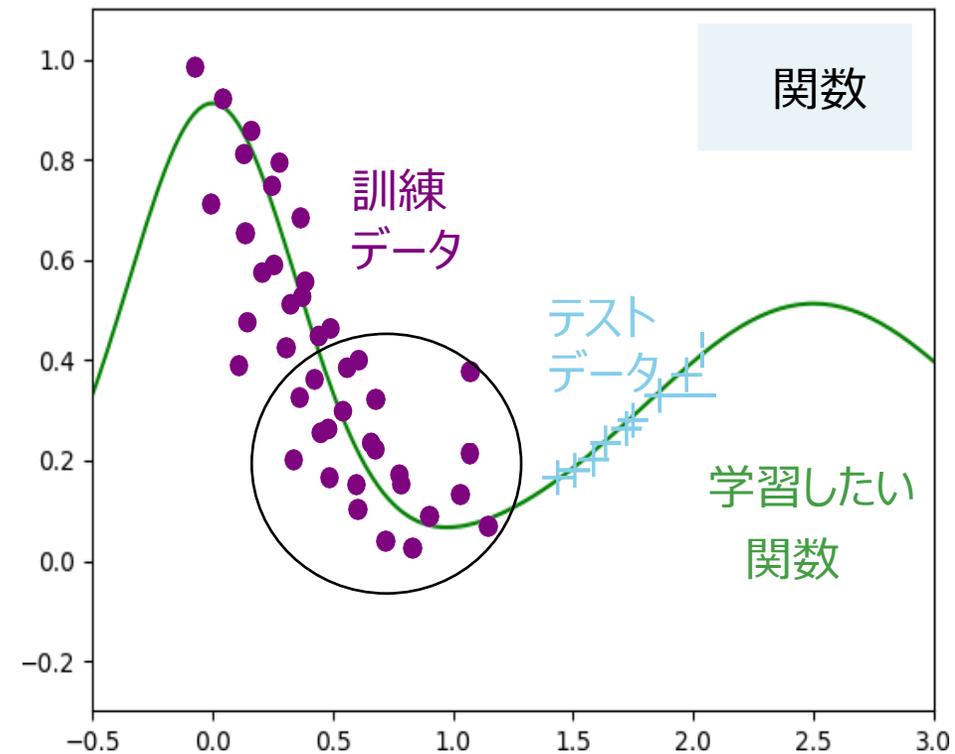
# 共変量シフト適応に基づくマルチモーダル機械学習

- 共変量シフト下での教師付き学習 [Shimodaira 2000] [Sugiyama 2007]
  - 重みの大きい（テストデータに類似する）学習データの誤差を考慮

- モデル：重要度重み付きサポートベクター回帰

$$\hat{\theta} = \operatorname{argmin}_{\theta} \left[ \frac{1}{N_{tr}} \sum_{i=1}^{N_{tr}} (\hat{w}(x_i^{tr})) | \hat{f}(x_i^{tr}; \theta) - y_i^{tr} | \epsilon \right]$$

- 重み $\hat{w}$ の推定にはKuLSIF [Kanamori+2009]を利用
- マルチモーダルフュージョンモデルとして定式化



# 回帰モデルの評価実験結果

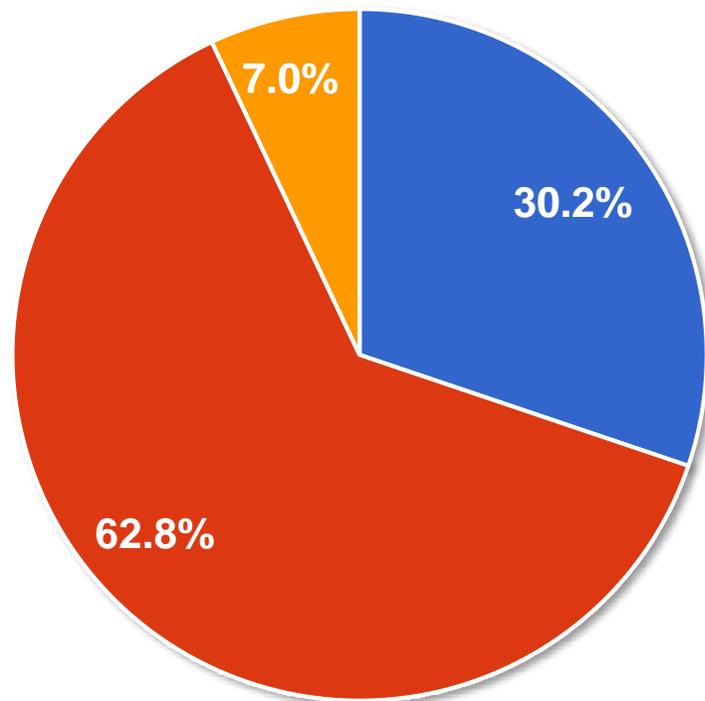
Corr. $\rho$	[A1]:目的網羅性			[A2]:内容ロジック			[A3]:効果的演出要素			[A4]話し方基礎		
	W/O IW	IW	diff	W/O IW	IW	diff	W/O IW	IW	diff	W/O IW	IW	diff
<i>A</i>	.18	.21	+.03	.00	.16	+.16	.06	-.07	-.13	.25	.02	-.23
<i>L</i>	-.75	-.71	+.04	.25	.15	-.10	.10	.26	+.16	.21	.33	+.12
<i>V</i>	.02	-.13	-.15	-.11	-.07	+.04	-.12	.02	+.14	-.70	-.44	+.26
<i>A + L</i>	.14	.17	+.03	-.08	<b>.27</b>	+.35	<b>.29</b>	.04	-.24	.20	.25	+.05
<i>A + V</i>	.21	<b>.25</b>	+.04	-.12	.05	+.18	.15	-.02	-.17	.19	.03	-.16
<i>L + V</i>	-.04	-.30	-.26	.03	.14	+.11	.16	.14	-.03	.17	<b>.36</b>	+.19
<i>All</i>	.13	.22	+.08	-.09	.26	+.35	.18	.10	-.08	.18	.24	+.06

精度は相関係数(Corr. $\rho$ )を用いた

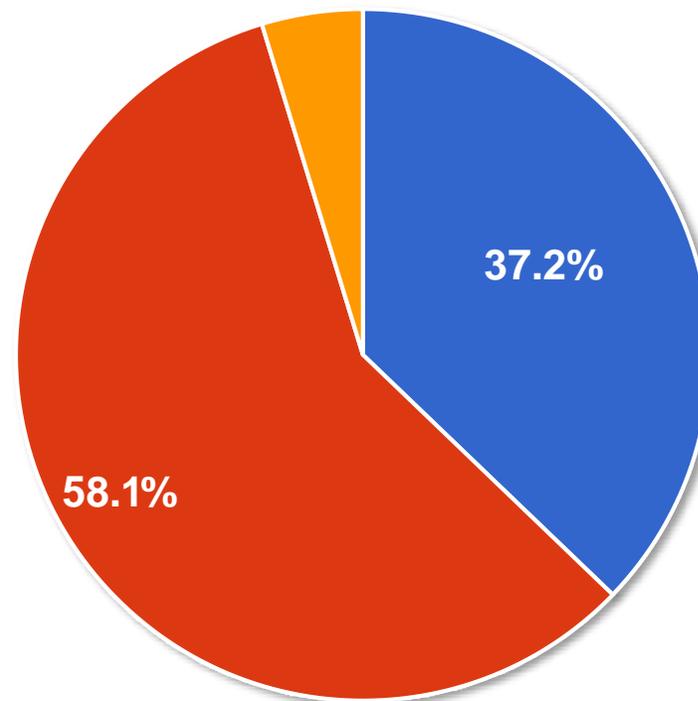
- [A1]~[A4]でIW適応後に精度が向上した
- 最大推定精度はマルチモーダルモデルで得られた
- [A1],[A2],[A4]ではIW適応モデルで最大推定精度が得られた

# 推定モデルに基づくフィードバックシステム評価

## AIの評価は妥当であったか？



## 講師の評価は妥当であったか？



- 妥当であった
- おおむね妥当であった
- おおむね妥当でなかった
- 妥当でなかった

自由回答例：

- AIにも改善フィードバックがあるといい.
- AI分析の基準が不明のため，納得感が不足.

# コミュニケーションスキルの推定に関する成果 1

## カウンセリング対話



**X**

### マルチモーダル特徴量

- 発話ターン
- 韻律
- 単語

**Y**

### 説得・納得場面の推定

## 推定精度

クライアントの  
納得場面を  
LSTMで約70%で推定

[Amari+ HCII2021]

## 就職面接



**X**

### マルチモーダル特徴量

- 発話ターン
- 韻律
- 頭部・全身動作
- 視線
- 参加者間のインタラクション

**Y**

### 面接官による就職適正・ CSの印象評定値

## 推定精度

就職適正度を  
二値分類で  
77%

[Okada+ ACM TOMM 2019]  
[Tian, Okada+ ACM HAI 2019]

# コミュニケーションスキルの推定に関する成果 2

## グループディスカッション (日本人4人一組)



**X**

### マルチモーダル特徴量

- 発話ターン
- 韻律
- 動作
- 発話内容の単語

**Y**

### コミュニケーション スキルの評定値

## 推定精度

CSの評定値の高低を  
L-SVM で93%で分類  
H-LSTMでR2 ~ 0.80

[Okada+ ACM ICMI2016]  
[Miura+ ACM ICMI2019]  
[Li+ ACM ICMI2019]  
[Okada+ JSAI論文誌2016]  
[Mawalim+2021 ACM TOMM 2021]

## グループディスカッション (欧州参加者4人一組)



**X**

### マルチモーダル特徴量

- 発話ターン
- 韻律
- 頭部・全身動作
- 視線
- 参加者間のインタラクション

**Y**

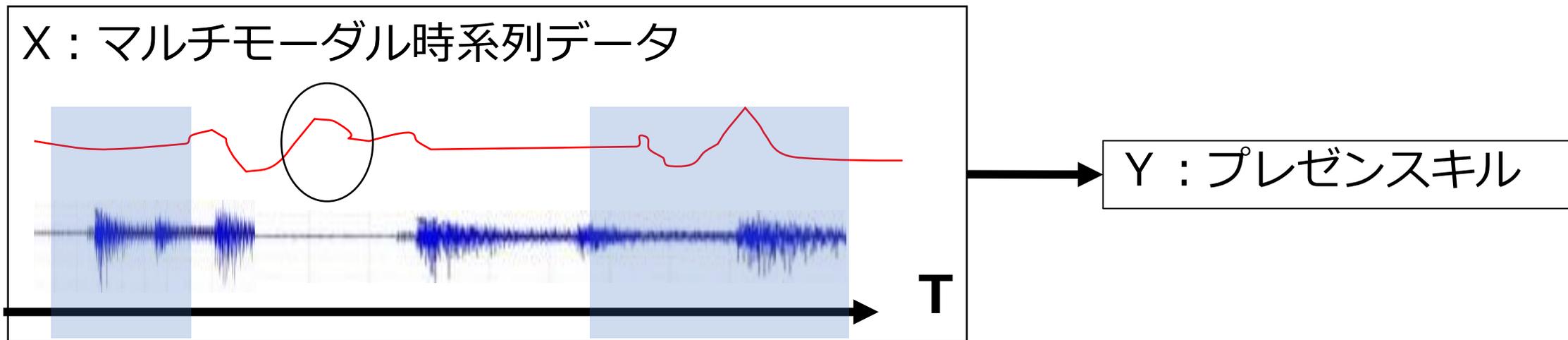
### リーダーシップ・ 性格 (Big Five) の印象評定値

## 推定精度

外向性の項目の  
高/低を73%  
リーダーシップの項目の  
高/低を77%  
の精度で分類

[Okada+ ACM ICMI 2015]  
[Okada+ ACM TOMM 2019]

# 社会的信号処理における機械学習の現状・課題



## 現状

- マルチモーダル情報の統合, 翻訳 (モーダルA→モーダルBの写像) , アライメント

## 課題

- 訓練サンプルの少量問題 (Many to One  $\Leftrightarrow$  時系列 to 1 ラベル)
- 学習に悪影響 (e.g. アノテーション一致率が低い) なサンプルの選定
- 時系列中の不要なフレーム/重要なフレームの選定や注視

## まとめと課題

- マルチモーダル情報から内面状態を推定する社会的信号処理技術
- コミュニケーションスキルの予測
  - 様々な状況・スキルタイプ
- ビジネスプレゼンススキル推定モデルとフィードバックシステムの研究開発
  - 重要度重み付き学習による個人差の問題の緩和と精度向上
  - フィードバックコメントに関して人間と概ね同等の評価
- 今後の課題
  - 推定モデル構築のための機械学習の精緻化
  - フィードバックシステムの運用・ユーザの行動変容分析
  - 内面状態の推定における倫理問題（プライバシー, 公平性, バイアス）