

表情・音声を用いた褒め方の上手さを評価する システムの基礎検討

山内 愛里沙¹ 大西 俊輝¹ 呉 健朗² 武藤 佑太¹ 石井 亮³ 青野 裕司³ 宮田 章裕^{1,a)}

概要: 対話において褒めることは重要であり、相手を上手く褒めるためには、言語的・非言語的行動を適切に用いる必要があると考えられる。しかし、発話内容、声色、表情、姿勢、ジェスチャといった多様なモダリティをどのように用いればよいか明らかにされていない。我々は、対話において相手を上手く褒めるためには、人間がどのような言語的・非言語的行動を用いているかを定量的な分析から明らかにし、対話における褒め方の上手さのスコアリングを行い、褒め方のスキルを向上させるためのアドバイスを提示するシステムの構築を目指している。本研究では、対話における自身の褒め方の上手さを自身が認識できるようにするために、対話コーパスの作成、褒め方の上手さの推定モデルの構築、褒め方の上手さを評価するシステムの構築を行った。対話コーパスの作成では、褒め方の上手さの評価と各モダリティの記録を含む2者対話を収録し、アノテーションを行った。褒め方の上手さの推定モデルの構築では、対話において褒める人の表情特徴量、音声特徴量の抽出を行い、表情特徴量、音声特徴量を用いた回帰モデルの構築を行った。褒め方の上手さを評価するシステムの構築では、提示された人物の発話シーンに対して、ユーザの褒めるシーンの映像と音声を入力とし、そのシーンにおけるユーザの褒め方の上手さを推定し、スコアとしてユーザに提示するシステムの構築を行った。構築したシステムによって、対話における褒め方の上手さを推定・評価することを初めて実現した。

1. はじめに

対話において、褒めることは重要である。例えば、初対面の相手と話す場合、相手の見た目やエピソードを積極的に褒めることで相手が心を開き、話が弾み、その後の人間関係が円滑になることが期待される。あるいは、同僚や部下の行為を適切に褒めることで彼らのパフォーマンスを最大化することが、組織マネジメントの観点からも重要であるとの認識が広がっている。本研究で対象となる褒める行為は、承認・称賛を表現すること、対象の良い側面を示唆する言動的または身体的な行動、対象の行動や性格に向けられた称賛を示す言語的・非言語的な行為、相手の良さ、優れている点、長所に気づき、価値を正しく認めて伝える行為と考えられている [1][2][3][4][5]。これらの既存知見より、相手を上手く褒めるためには、言語的・非言語的行動を適切に用いる必要があると考えられる。しかし、対話において相手を上手く褒めるためには、発話内容、声色、表情、姿勢、ジェスチャといった多様なモダリティをどのように用いれば良いかは明らかにされていない。このため、

自身の褒め方の上手さを認識することができず、何を意識すれば褒め方のスキルが向上するのか分からないという問題が存在する。

このような背景を踏まえ、我々は、上手く褒めるために人間がどのような言語的・非言語的行動を用いているかを定量的な分析から明らかにし、対話における褒め方の上手さのスコアリングを行い、褒め方のスキルを向上させるためのアドバイスを提示するシステムの構築を目指している。先行研究において、対話において相手を上手く褒めるためには、表情が欠かせない要素の1つであること、さらに、対話における褒め方の上手さは、褒める人の頭部の縦軸周りの回転角度、視線の横方向の角度、顎を上げる動き、目を見開く（瞼に力を入れる）動きが、褒め方の上手さに関係していることを明らかにした [6][7]。本研究では、先行研究 [6], [7] で扱った表情に加えて、未だ扱われていなかった音声にも着目する。具体的には、表情と音声を用いた褒め方の上手さのスコアを算出するシステムを構築し、対話におけるユーザ自身の褒め方の上手さをユーザ自身が認識できるようにすることを研究課題とする。本研究の貢献は、次のとおりである。

- 先行研究では扱われていなかった音声情報を初めて扱い、表情、音声を用いた褒め方の上手さの推定モデル

¹ 日本大学文理学部

² 日本大学大学院総合基礎科学研究科

³ 日本電信電話株式会社 NTT メディアインテリジェンス研究所

a) miyata.akihiro@acm.org

を構築した。

- 対話における褒め方の上手さのスコアを算出するシステムを構築した。

2. 関連研究

本研究は、発話内容、声色、表情、姿勢、ジェスチャといった多様なモダリティを用いた褒め方の上手さを評価するシステムの構築を目指すものであり、対話に関するスキルの推定や評価を行う研究に関連している。

[8] は、発話言語から得られる様々な種類の自然言語解析情報を用い、人間と同様に適切なタイミングで発話を伴う、エージェントの全身運動を自動生成するモデルを提案している。頷き、頭部の姿勢、表情、ハンドジェスチャ、上半身の姿勢を自動生成している。提案するモデルによって、エージェントの体の動き、動きの自然さに関する印象、発話と動きの一致度、人間らしさ、好ましさは改善された。[9] は、人間と同様に適切なタイミングで発話を伴う全身運動を自動生成するモデルを用い、簡単に動きを生成することができる API を構築している。音声言語のみから仮想エージェントを自動的に制御することができるデモシステムを構築している。[10] は、新たに形成された小グループ内で、どのようにリーダーの出現が認識されるか分析している。グループ内におけるディスカッションの音声や映像から抽出される非言語情報を用いて、リーダーの出現を自動的に推測している。グループコーパスから音響特徴量・視線特徴量を抽出して組み合わせ、ルールベースで出現したリーダーかそうでないかを分類するアプローチをとっている。[11] は、言語的行動と非言語的行動の相関分析を行い、言語的行動における言語特徴量、頭部のジェスチャなどの非言語的行動が親密な自己開示と関連していることを明らかにした。言語的・非言語的行動を用いて、自己開示の親密度を自動推定することができるマルチモーダルディープニューラルネットワークを開発した。[12] は、人前で話す能力について、専門家の意見に関する視聴覚非言語的行動を自動評価している。仮想の聴衆が、人前で話すトレーニングに必須なフィードバックを提供することを可能にしている。[13] は、参加者の対話からの特徴量抽出、外部の評価者による参加者のコミュニケーションスキルの評価、ディスカッションタイプに関する状況の変動の影響の分析を行い、グループにおける参加者の個々のコミュニケーションスキルを解析している。[14] は、ベテランの高校教員によって行われたヒューマンアセスメントに基づく、教員のパフォーマンス評価のための自動評価モデルを提案している。[15] は、プレゼンテーションデータから得られたマルチモーダル情報から、プレゼンテーションスキルの推定モデルの構築・評価を行っている。[16] は、発話、顔、体の動きなどのマルチモーダル情報から得た特徴量を、プレゼンテーションのスコアリングに適用する方法の分析を行って

いる。

3. 研究課題

対話におけるスキルの推定や評価を行う研究は数多く行われているが、本研究が目指す褒め方の上手さの推定や評価とスコアの算出とは異なる点が存在する。[8], [9], [10], [11] は、表情や音声等のマルチモーダル情報を用いて、人間と同様な全身運動を自動生成するモデルの提案やグループ内における対話参加者の役割推定、自己開示における親密度の自動推定を行っている。本研究においても表情や音声等のマルチモーダル情報を用いて研究を行うが、既存研究とは対話コーパスの内容や褒め方の上手さを推定するという点で異なる。[12], [13], [14], [15], [16] は、マルチモーダル情報を抽出し、公衆の面前で話す能力の自動評価やコミュニケーションスキルの評価と分析、プレゼンテーションスキルの評価やスコアリングへの適用方法の分析を行っている。本研究においても、マルチモーダル情報を抽出して研究を行うが、既存研究とは褒め方の上手さを評価するという点で異なる。

以上より、既存研究において、褒め方の上手さの推定や評価、対話における自身の褒め方の上手さの認識を可能にする取り組みは行われていない。このことから、本研究では、対話における褒め方の上手さのスコアリングを行い、ユーザに対して褒め方の上手さのスコアと褒め方のスキルを向上させるためのアドバイスを提示するシステムの構築を目指す。先行研究において、対話において相手を上手く褒めるためには、表情が欠かせない要素の1つであること、さらに、対話における褒め方の上手さは、褒める人の頭部の縦軸周りの回転角度、視線の横方向の角度、顎を上げる動き、目を見開く（瞼に力を入れる）動きが、褒め方の上手さに関係していることを明らかにした [6][7]。しかし、褒め方の上手さに関係すると考えられる音声は扱っていない。本研究では、先行研究 [6], [7] で扱った表情に加えて、未だ扱われていなかった音声情報を用いて、表情、音声で褒め方の上手さの推定に有用であるかを検証すること、さらに、褒め方の上手さのスコアを算出するシステムを構築し、対話におけるユーザ自身の褒め方の上手さをユーザ自身が認識できるようにすることを研究課題とする。

4. 提案手法

3章の研究課題を達成する手法として、対話において褒めている人の映像と音声を入力し、褒め方の上手さのスコアを算出するシステムを提案する。褒め方の上手さをスコアとして算出することにより、ユーザは算出されたスコアから自身の褒め方の上手さを認識することができると考えられる。本研究は、下記に示すステップに分けて推進する。

- **Step1** : 対話コーパスの作成
対話における褒め方の上手さの評価と各モダリティの

記録を含む対話コーパスの作成を行う。

- **Step2**: 褒め方の上手さの推定モデルの構築
表情特徴量と音声特徴量を抽出し、褒め方の上手さを推定するモデルを構築する。
- **Step3**: 褒め方の上手さを評価するシステムの構築
褒め方の上手さの推定モデルを用いて、対話における褒め方の上手さを評価するシステムを構築する。

5. 対話コーパス

5.1 2者対話の収録

2者対話を対象とした対話における褒め方の上手さの評価と各モダリティの記録を含む対話コーパスの作成を行った。2者対話の参加者は、20代の大学生20名（男性19名、女性1名）であり、2名1組の組み合わせを10組構成した。10組のうち、初対面が8組、顔見知り1組、友人同士が1組であった。対話の収録を始めるにあたり、参加者に対話材料を準備させることを意図して、アンケート（表1）に回答してもらった。参加者は、図1のように互いに

表1 アンケートの質問事項と回答方法

項目	質問事項	回答方法
Q1	年齢	単一選択
Q2	性別	単一選択
Q3	相手との関係	単一選択
Q4	今まで頑張ってきたこと（2つ以上）	自由記述

向き合って着座する。対話の収録は、各参加者の様子と2



図1 2者対話の様子

者対話全体の様子を撮影するためのビデオカメラ、各参加者の声を録音するためのマイク、各参加者のハンドジェスチャなどの動きを計測するための深度センサを用いて行った。各組の参加者達（参加者A、参加者B）は、実験者の合図に従い以下の（1）～（3）を行った。

- （1）自己紹介（5分間）
- （2）参加者Aが主に話し手となり今まで頑張ってきたことを話し、参加者Bが主に聞き手となり随時意見を述べ

る対話（5分間）

- （3）参加者Bが主に話し手となり今まで頑張ってきたことを話し、参加者Aが主に聞き手となり随時意見を述べる対話（5分間）

（2）、（3）を10組分、計100分間の2者対話を収録する。対話において話し手となる参加者には、アンケートで回答した今まで頑張ってきたことについて話すように指示した。また、アンケートで回答しなかった今まで頑張ってきたことについて話すことを許可した。対話において聞き手となる参加者には、話し手の発話内容について褒めるように指示した。加えて、自由に質問したりリアクションしたりすることを許可した。なお、（1）は各組の多くが初対面同士であり、参加者の緊張をほぐす目的で行っているため、今回の分析では使用しない。

5.2 アノテーション

音声データや映像データに対して注釈付けを行うツールであるELAN[17]を用いて、収録した音声データと映像データに対して表2に示すアノテーション項目を人手で付与した。発話シーンは、沈黙時間が200ミリ秒未満の連続した

表2 アノテーションの項目

項目	説明
発話シーン	沈黙時間が200ミリ秒未満の連続した音声区間さらに発話が1つの文として成り立つ区間
発話内容	参加者の発話内容をテキストに書き起こしたものと聞き手の参加者が相手の発話に対して褒める発話を行っていることを判断したシーン

した音声区間であり、さらに、発話が1つの文として成り立つ区間である。発話内容は、参加者の発話内容を書き起こしたものである。Praiseシーンは、聞き手の参加者が相手の発話に対して褒める発話を行っていることを判断したシーンである。判断基準が揺るがないよう、1名のアノテータが全てのPraiseシーンの付与を行った。以上のアノテーションで抽出した情報を表3に示す。参加者20名の話し

表3 アノテーションで抽出したデータの情報

項目	件数	継続時間の最小値	継続時間の最大値
話し手の発話シーン	2,963	0.02秒	13.02秒
聞き手の発話シーン	1,973	0.04秒	11.39秒
Praiseシーン	227	0.29秒	7.80秒

手の発話シーンが2,963件、聞き手の発話シーンが1,973件、Praiseシーンが227件であった。

5.3 褒め方の上手さの評価

5.2節で付与したPraiseシーンについて、当該シーンで話し手であった参加者による褒め方の上手さの評価を行った。2者対話の収録後、当該シーンで話し手であった参加者に映像データを参照してもらい、当該シーンで相手に褒

められて気分が良くなったかについて、7段階（1：気分が良くなったと思わない～7：気分が良くなったと思う）で評価してもらった。

6. モデル構築

本研究では、表情特徴量を用いた推定モデル、音声特徴量を用いた推定モデル、表現特徴量と音響特徴量の2つを用いた推定モデルを構築した。具体的には、特徴量抽出とモデル構築を行った。

6.1 特徴量抽出

6.1.1 顔・視線方向と表情特徴量の抽出

顔画像処理ツールである OpenFace[18] を用いて、参加者の正面に設置されたビデオカメラで撮影した映像データから、顔の向きと表情特徴量の抽出を行った。具体的に、OpenFace から抽出可能な頭部、視線、Action Units[19] を抽出した。頭部に関する特徴量としては、ビデオカメラ側から顔を見て横方向を x 軸、縦方向を y 軸、奥行き方向を z 軸とした際の、頭部の x 軸、y 軸、z 軸方向の回転角度 (pose.Rx, pose.Ry, pose.Rz) の Praise シーンの前後1秒ずつを含む区間における中央値 (_med), 分散 (_var), 10 パーセンタイル値 (_p10), 90 パーセンタイル値 (_p90) を用いた。Action Units に関する特徴量としては、各 Action Units の強度 (AU01, AU02, AU04, AU05, AU06, AU07, AU09, AU10, AU12, AU14, AU15, AU17, AU20, AU23, AU25, AU26, AU45) の Praise シーンの前後1秒ずつを含む区間における中央値, 分散, 10 パーセンタイル値 (_p10), 90 パーセンタイル値 (_p90) を用いた。各 Action Units の内容について表4に示す。全 Praise シーンに対して上記の

表4 Action Units の内容

項目	内容	項目	内容
AU01	眉の内側を上げる	AU14	笑窪を作る
AU02	眉の外側を上げる	AU15	唇の両端を下げる
AU04	眉を下げる	AU17	顎を上げる
AU05	上瞼を上げる	AU20	唇の両端を横に引く
AU06	頬を持ち上げる	AU23	唇を固く閉じる
AU07	瞼を緊張させる	AU25	顎を下げて唇を開く
AU09	鼻に皺を寄せる	AU26	顎を下げて唇を開く
AU10	上唇を上げる	AU45	瞬きをする
AU12	唇の両端を引き上げる		

特徴量を抽出し、各特徴量を平均値0, 分散1になるように正規化を行った。

6.1.2 音声特徴量の抽出

音声情報処理ツールである openSMILE[22] を用いて、聞き手である参加者が装着したマイクで録音された音声データから、音声の韻律の代表的な特徴量の抽出を行った。抽出可能な特徴量のうち、表5に示す音声特徴量を、利用する特徴量とした。表5に示す各音声特徴量に対し、表6に

表5 音声特徴量の内容

特徴量	内容
pcm_RMSenergy_sma	音のエネルギーの二乗平均平方根
pcm_ffftMag_mfcc_sma[1]-[12]	1~12 次のメルケプストラム係数
pcm_zcr_sma	ゼロ交差率
voiceProb_sma	声である確率
F0_sma	基本周波数

表6 抽出可能な統計量

統計量	内容	統計量	内容
_max	最大値	_linregc1	線形近似の勾配
_min	最小値	_linregc2	線形近似のオフセット
_range	最大値と最小値の差	_linregerrQ	線形近似の二乗誤差
_maxPos	最大値の絶対位置	_stddev	標準偏差
_minPos	最小値の絶対位置	_skewness	歪度
_amean	平均値	_kurtosis	尖度

示す統計量が算出される。本研究では、表5に示す16個の各特徴量につき表6に示す12個の統計量を算出した、計192個の特徴量をモデル構築で用いる音声特徴量とした。以上の特徴量に加えて、Praise シーンにおける聞き手の発話の長さ(秒)を用いた。6.1.1節と同様に、全 Praise シーンに対して上記の特徴量を抽出し、各特徴量を平均値0, 分散1になるように正規化を行った。

6.2 モデル構築

Praise スコアを目的変数、6.1.1 項と 6.1.2 項の表情特徴量と音声特徴量を説明変数とする回帰モデルを構築した。Praise スコアは、5.3節で行った、聞き手の褒め方の上手さに対する話し手による7段階評価である。モデル構築には、構造化データに対して高精度なモデルを構築することができる Gradient boosting[20] を用いた。学習率や木の深さなどのハイパーパラメータは、Hyperopt[21] でチューニングした。データセットを80%の訓練データ、20%の検証データに無作為に分け、訓練データで構築したモデルを用いて検証データの Praise スコアを推定するタスクを100回行ったところ、各指標の平均値は表7のようになった。MSE は平均二乗誤差 (Mean Square Error) を表し、 R^2 は

表7 各指標の平均値 (N=100)

指標	平均値
MES(train)	0.131
MSE(test)	1.391
R^2 (train)	0.928
R^2 (test)	0.206

決定係数を表している。

7. システム構築

7.1 システム構成

ユーザの映像と音声を入力すると、入力されたシーンにおけるユーザの褒め方の上手さを推定し、スコアとしてユーザに提示するシステムを構築する。システムの構成を

図2に示す。入力は、ユーザの映像と音声とする。入力さ

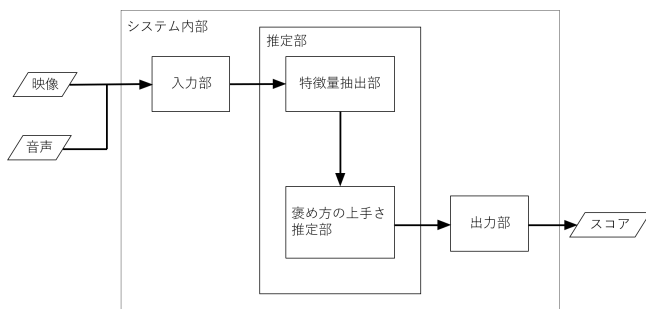


図2 システムの構成図

れたデータは入力部へ渡され、推定部で特徴量抽出と褒め方の上手さの推定を行う。推定部では、6.2節で構築した推定モデルを用いる。褒め方の上手さの推定を行った後、出力部でディスプレイへの出力処理を行う。以降、入力部、推定部、出力部に分けて述べる。

7.2 入力部

ユーザの顔部をビデオカメラで撮影した映像と、マイクから録音した音声に対して入力処理を行う。入力処理を行った後、映像データ、音声データを推定部へ渡す。

7.3 推定部

7.2節の入力データに対する特徴量抽出と褒め方の上手さの推定を行う。特徴量抽出では、7.2節の映像データと音声データから、6.1.1項と6.1.2項で示した表情特徴量と音声特徴量を抽出し、抽出された表情特徴量と音声特徴量のマージを行う。褒め方の上手さの推定では、マージされた特徴量と6.2節で構築した褒め方の上手さの推定モデルを用いて、褒め方の上手さの推定とスコアの算出を行う。具体的には、6.2節で用いたデータセットを訓練データ、マージされた特徴量を検証データとし、6.2節で構築した表情・音声を特徴量としたモデルを用いて、Praiseスコアの推定を行う。推定されたPraiseスコアを出力部へ渡す。

7.4 出力部

褒め方の上手さのスコアをディスプレイに出力する処理を行う。具体的には、7.3節で算出された1~7の数値を、1~100にスケールし、100点を満点としたスコアをユーザに提示する。図3に、スコアの出力例を示す。

8. おわりに

対話における自身の褒め方の上手さを自身が認識できるようにするために、対話コーパスの作成、褒め方の上手さの推定モデルの構築、褒め方の上手さを評価するシステムの構築を行った。対話コーパスの作成では、褒め方の上手さの評価と各モダリティの記録を含む2者対話を収録し、



図3 スコアの出力例

アノテーションを行った。収録した映像データ、音声データから、対話において褒める人の表情特徴量、本研究で初めて用いた音声特徴量の抽出を行い、これらの表情特徴量、音声特徴量を用いた回帰モデルの構築を行った。さらに、構築した褒め方の上手さの推定モデルを用いて、ユーザの映像と音声を入力すると、入力されたシーンにおけるユーザの褒め方の上手さを推定し、スコアとしてユーザに提示するシステムの構築を行った。

本稿では、システムの構築によって、対話における褒め方の上手さを推定・評価することは可能となったが、いくつかの課題が挙げられる。本システムは、褒め方の上手さのスコアリングを行うものであるが、モデルやスコアの精度には向上の余地がある。推定アルゴリズムの改良や、利用する音声特徴量の種類の追加を行い、精度の向上を行っていききたい。さらに、表情や音声に加えて、これまで扱っていないユーザの発話内容、ジェスチャ等の特徴量を利用していきたいと考えている。ユーザへのフィードバックの観点においては、本システムが出力するスコアから、ユーザが自身の褒め方の上手さを認識できたか否かの調査を行う必要がある。フィードバック方法も、スコアの提示のみでなく、ユーザの褒め方のスキルを向上させるためのアドバイスを提示することを目指したい。

参考文献

- [1] Brophy, J.: Teacher praise: A full functional analysis. Review of Educational Research, Vol.51, No.1, pp.5-32 (1981).
- [2] Kalis, T. M., Vannest, K. J. and Parker, R.: Praise Counts: Using Self-Monitoring to Increase Effective Teaching Practices. Preventing School Failure. Alternative Education for Children and Youth, Vol.51, No.3, pp.20-27 (2007).
- [3] Jenkins, L. N., Floress, M. T. and Reinke, W.: Rates and Types of Teacher Praise: A Review and Future Directions. Psychology in the Schools, Vol.52, No.5, pp.463-476 (2015).
- [4] 永崎一則: 元気と幸せを生み出す ほめかたの研究。早稲田教育出版 (2015)。
- [5] 西村貴好: 結果を引き出す 大人のほめ言葉。同文館出版株式会社 (2017)。
- [6] 大西俊輝, 柴田万里那, 呉健朗, 石井亮, 富田準二, 宮田章裕: 対話における上手い褒め方のモデリングの基礎検討。情報処理学会シンポジウム論文集, マルチメディア

- ア, 分散, 協調とモバイル (DICOMO 2019), Vol.2019, pp.656–662 (2019).
- [7] 大西俊輝, 柴田万里那, 山内愛里沙, 呉健朗, 石井亮, 富田準二, 宮田章裕: 褒め方の上手さの推定における頭部・顔部の効果, 情報処理学会グループウェアとネットワークサービスワークショップ 2019 論文集, Vol.2019, pp.1–6 (2019).
- [8] Ishii, R., Katayama, T., Higashinaka, R., and Tomita, J.: Generating Body Motions using Spoken Language in Dialogue. Proceedings of the 18th ACM International Conference on Intelligent Virtual Agents (IVA '18), pp.87–92 (2018).
- [9] Ishii, R., Katayama, T., Higashinaka, R. and Tomita, J.: Automatic Generation System of Virtual Agent's Motion using Natural Language. Proceedings of the 18th International Conference on Intelligent Virtual Agents (IVA '18), pp.357–358 (2018).
- [10] Sanchez-Cortes, D., Aran, O., Mast, M. S. and Gatica-Perez, D.: A Nonverbal Behavior Approach to Identify Emergent Leaders in Small Groups. IEEE Transactions on Multimedia, Vol.14, No.3, pp.816–832 (2012).
- [11] Soleymani, M., Stefanov, K., Kang, S. H., Ondras, J., Gratch, J.: Multimodal Analysis and Estimation of Intimate Self-Disclosure. Proceedings of the 2019 ACM on International Conference on Multimodal Interaction (ICMI '19), pp.59–68 (2019).
- [12] Wörtwein, T., Chollet, M., Schauerte, B., Morency, L. P., Stiefelhagen, R. and Scherer, S.: Multimodal Public Speaking Performance Assessment. Proceedings of the 2015 ACM on International Conference on Multimodal Interaction (ICMI '15), pp.43–50 (2015).
- [13] Okada, S., Nguyen, L. S., Aran, O. and Gatica-Perez, D.: Estimating communication skills using dialogue acts and nonverbal features in multiple discussion datasets. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), Vol.15, No.1, pp.13:1–13:30 (2019).
- [14] Fukuda, M., Huang, H. H., Kuwabara, K., Nishida, T.: Multimodal Assessment on Teaching Skills in a Virtual Rehearsal Environment. Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents (IVA '19), pp.191–193 (2019).
- [15] 八木 悠太郎, 岡田 将吾, 塩原 翔太, 杉村 聡太: ビジネスプレゼンテーションにおける言語・非言語的能力の自動推定. 人工知能学会全国大会論文集, pp.1–3 (2019).
- [16] Ramanarayanan, V., Leong, C. W., Chen, L., Feng, G. and Suendermann-Oeft, D.: Evaluating Speech, Face, Emotion and Body Movement Time-series Features for Automated Multimodal Presentation Scoring. Proceedings of the 2015 ACM on International Conference on Multimodal Interaction (ICMI '15), pp.23–30 (2015).
- [17] Brugman, H. and Russel, A.: Annotating Multimedia Multi-modal resources with ELAN. Proceedings of the 4th International Conference on Language Resources and Language Evaluation (LREC '04), pp.2065–2068 (2004).
- [18] Baltrušaitis, T., Zadeh, A., Lim, Y. C. and Morency, L. P.: OpenFace 2.0: Facial Behavior Analysis Toolkit. 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG '18), pp.59–66 (2018).
- [19] Ekman, P. and Friesen, W. V.: Manual for the Facial Action Coding System. Palo Alto: Consulting Psychologists Press (1977).
- [20] Friedman, J. H.: Greedy Function Approximation: A Gradient Boosting Machine. The Annals of Statistics Vol.29, No.5, pp.1189–1232 (2001).
- [21] Bergstra, J., Yamins, D. and Cox, D. D.: Hyperopt: A Python Library for Optimizing the Hyperparameters of Machine Learning Algorithms. Proceedings of the 12th Python in Science Conferences (SciPy '13), pp.13–20 (2013).
- [22] Eyben, F., Wöllmer, M. and Schuller, B.: openSMILE – The Munich Versatile and Fast Open-Source Audio Feature Extractor. Proceedings of ACM Multimedia, pp. 1459–1462 (2010).