

英語プレゼンテーション学習支援用 AI の開発

Development of AI for English Presentation Learning Support

藤島 美空^{*1}, 上野 春毅^{*1}, 中山 誠一^{*2}, 小松川 浩^{*1}
Miku FUJISHIMA^{*1}, Haruki UENO^{*1}, Tomokazu NAKAYAMA^{*2}, Hiroshi KOMATSUGAWA^{*1}

^{*1} 公立千歳科学技術大学 理工学部

^{*1} Faculty of Science and Technology, Chitose Institute of Science and Technology

^{*2} 実践女子大学

^{*2} Jissen Women's University

Email: b2191970@photon.chitose.ac.jp

あらまし：本研究では、公開性の高い音声認識や画像認識といった AI 技術を用いることで、手元のデバイスのみだけで実行可能な英語プレゼンテーション動画の評価のプログラムを作成した。英語プレゼンテーションの評価観点として『発話内容』『無音となっている時間』『発表態度』『原稿を読んでいるか』『好印象を持つ表情をしているか』の5つの観点を定義し、それに基づいて評価を行った。

キーワード：AI, 学習支援

1. はじめに

近年、AI (人工知能) の技術が飛躍的に進歩し、教育分野における能力判断 (評価) の自動化という新たな社会的革新をもたらそうとしている⁽¹⁾。本研究では、大学初年次系の英語教育におけるプレゼンテーション能力養成を支援する AI 開発を目的とする。プレゼンテーションは発話や表情、視線といった要素が評価される。学生にとって上述の要素のトレーニングは、教員の評価を受けて改善を図るのが一般的である。しかし、教員の評価を常に受けられるとは限らず、自身を客観的に評価しながら自己トレーニングを行うことが求められる。一方で AI 技術の観点では、音声処理や画像処理などの精度が高まっており、学生の発話や表情、視線の情報を検出してプレゼンテーションを自動的に評価できる可能性がある。一例としてロボットを用いた研究⁽²⁾が行われており、身体的所作を Microsoft 社の Kinect で検出している。本研究チームでは、大学等の教育機関での運用を念頭に、特殊な機器を必要とせず広く普及しているデバイスであるスマートフォン等で利用できる公開性の高い技術を用いることを前提とする。

本研究では、教員の評価観点に沿ってプレゼンテーションの評価・学習支援を目指して、公開性の高い機械学習技術による自動的な評価の仕組み (AI) を開発する。本発表では、AI の肝要な部分である評価方法についてその設計と検証結果を報告する。

2. 概要

本研究で想定する授業の形態は、対面でのプレゼンテーションではなく、オンライン教育の普及により学生が手元のデバイスでプレゼンテーションを自ら撮影して動画として提出する形態とする。撮影されたプレゼンテーションの動画において、学生の『発話内容』『無音』『発表態度』『カンニング』『表情』

を自動で分析し、教員の評価基準に沿って評価を行う。

2.1 評価観点及び評価分類について

本研究では、撮影された英語プレゼンテーション動画 (大学初年次相当) を、機械学習技術を用いることで自動的な評価を行う。今回対象とする教員の評価観点を表 1 に示す 5 観点を定義した。

表 1 評価観点

種類	評価の観点	説明
発話	発話内容	キーフレーズを話している
	無音	言葉に詰まっている
振る舞い	発表態度	発表中、資料 (ポスター等) を使い説明をしている カメラを意識して発表している
	カンニング	カンニングペーパーを読んでいるような仕草が見られない
	表情	『笑顔』のような好印象を持つ表情をしている

また、上記の評価観点から判定される評価の段階は S ランク・A ランク・B ランク・C ランクの 4 段階とし、それぞれどの観点を満たしているとの評価になるかまとめたものを表 2 に示す。

表 2 評価分類

評価ランク	条件
S	表情が『笑顔』のような好印象 原稿を読んでいる カメラを意識しながら資料の活用を行っている
A	原稿を読んでいる カメラを意識しながら資料の活用を行っている
B	原稿を読んでいる又はカメラの意識・資料の活用をあまり行っていない キーフレーズをあまり話していない
C	原稿を読んでいる又はカメラの意識・資料の活用を行っていない キーフレーズをあまり話していない 言葉に詰まっている時間が一定以上

2.2 評価 5 観点のロジック

発表者の発話

評価観点 1 『発話内容』については、発話すべきフレーズ (以下キーフレーズ) を教員が設定し、それと発話の内容を比較し、適合数に応じて得点化し、

8割以上を1, 未満を0と評価した。そのため、まずはプレゼンテーション動画の音声を変換し、次に取得した文章をキーワードマッチングでキーフレーズの有無を判定して評価を行う。音声認識にはMicrosoft AzureのSpeech to Textを用いた。評価観点2『無音』については、言葉に詰まっているかどうか判定するため、音声が無音になっている時間を計測し、閾値に基づき無音を1, それ以外を0で評価した。

発表者の振る舞い

発表者の振る舞いについては、MediaPipeを用いることで顔のキーポイントを検出し、顔の向きや視線の推定を行うことで各観点の評価を行った。評価観点3『発表態度』については、推定した顔の向きから判定を行った。その判定は、推定した値と設定した閾値の差分の値に応じて0~3で評価した。閾値の設定は、教員による動画から模範となる顔の向きのデータを得て、そのデータが最大の評価となるようにした。評価観点4『カンニング』については推定した顔の向きと視線の値を用い、カンニングを1, それ以外を0と評価した。評価観点5『表情』については、好印象を持つ表情をしているか判定するためにFACS(Facial Action Coding System)で定義されている『AU(顔面の筋肉の動き)の組み合わせ』から『感情』へ変換する規則を用いた。今回はこの規則の中から『喜び(Happy)』に関するAUの検出を行い、そのAUの値から判定を行い、好印象を持つ表情をしているものを1, それ以外を0と評価した。

評価の算出

表1で定義した評価観点についてそれぞれ判定を行った結果から、表2で示した評価分類に基づいて評価の算出を行う。今回本研究で実装した評価アルゴリズムのフローチャートを図1に示す。

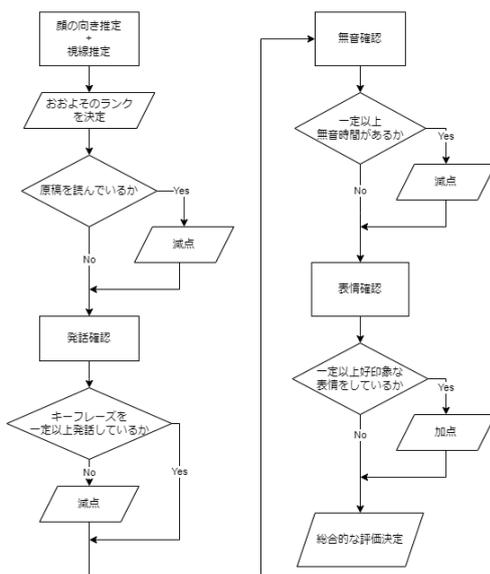


図1 評価フローチャート

3. 検証

本研究で開発した自動評価プログラムに、実際に評価を付けた学生の動画を適用し、プログラム上での評価と教員の評価を比較する。プログラム内で判定された総合評価ごとに評価観点を満たした動画の数をまとめたものを表4に示す。また、2つの評価の関係を示したものを表3に示す。

表3 プログラムの評価ごとの評価観点の平均点

プログラムの評価	動画の本数	発話内容	無音	発表態度	カンニング	表情
S	10	0.4	0	2.5	0	0.4
A	12	0.5	0	2.16667	0.08333	0.5
B	48	0.25	0.04167	1.20833	0.02083	0.25
C	47	0.06383	0	0.62766	0.08511	0.06383

表4 教員の評価と作成したプログラムの評価

		教員の評価			
		S	A	B	C
プログラムの評価	S	6	1	3	0
	A	7	3	2	0
	B	24	13	10	1
	C	18	16	13	0

4. おわりに

本研究では、公開性の高い技術を用いることで教員の評価観点に沿ってプレゼンテーションを評価・学習を支援できることを目指して、機械学習技術による自動的な評価の仕組み(AI)を開発して検証した。検証の結果、開発したAIで判定した評価は実際の評価よりも全体的に低く判定されていることがわかった。要因として、評価観点の判定方法や評価分類のアルゴリズムにおける評価観点の比重、そして検証データが評価指標を定める以前に評価された動画であることが考えられる。

今後の展望として、評価観点方法の見直しと実用性の高いアプリ化が挙げられる。今回設定した閾値の妥当性は検証できていないため、より柔軟に評価を行うことができるように検討を進める。また、今回作成した自動評価プログラムを用いてアプリ化をする際、アプリ利用者に評価を行った際の指摘箇所(原稿を読んでいるような仕草等)を、動画あるいは画像形式で視覚的に表示することで、より実用性のある学習支援アプリになると考える。

参考文献

- (1)宇都雅輝, “テスト理論と人工知能に基づくパフォーマンス評価の新技術”, 教育システム情報学会誌, Vol.37, No.1, pp.8-18 (2020)
- (2)後藤充裕, 石野達也, 稲澤佳祐, 松村成宗, 布引純史, 柏原昭博, “プレゼンタ動作を再現・再構成するロボットプレゼンテーションシステム”, JSiSE 研究会研究報告, Vol.32, No.5, pp.121-128 (2018)