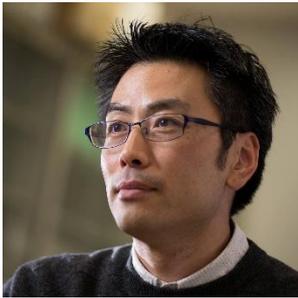


対話コミュニケーションの能力向上にむけた評価・フィードバックシステム

An Evaluation and Feedback System for Enhancing Communication Skills



山崎 俊彦 (Toshihiko YAMASAKI, Ph. D.)

東京大学大学院情報理工学系研究科電子情報学専攻
准教授

(Associate Professor, Department of Information and Communication Engineering, Graduate School of Information Science and Technology, The University of Tokyo)

電子情報通信学会 映像情報メディア学会 人工知能学会 情報処理学会 他

受賞：船井学術賞 (2016 年)、電子情報通信学会 IE 特別賞 (年間最優秀賞) (2020 年)、電子情報通信学会 HC 賞 (年間最優秀賞) (2011 年) 他

著書：『乳幼児の発達学』-眠る・食べる・遊ぶ- (共著, 朝倉書店 (2019 年)) 不動産テック (共著, 朝倉書店 (2020 年)) 他

研究専門分野：魅力工学 マルチメディア コンピュータビジョン

あらまし

本研究は、対話コミュニケーションの能力向上にむけたリアルタイム解析とフィードバックシステム構築を行う目的で行われた。これによって、これまで経験やセンスに依ると思われていた対話によるコミュニケーション能力を、客観的かつ定量的に評価・解析し、向上のための支援までをも行うシステムを研究開発した。対話コミュニケーション技術は対人関係を円滑にするだけでなく、学術・ビジネスの場においても非常に重要である。

1. 研究の目的

プレゼンテーションや会見など、対話による人と人とのコミュニケーションは非常に重要なスキルであるにも関わらず、系統立った学習パラダイムが存在していない。また、日本人のコミュニケーションスキルは、諸外国に比べて見劣りするとよく指摘される。このような現状を少しでも打ち破り、一人ひとりが自律的にコミュニケーションスキルを獲得できるような評価・フィードバックシステムをアプリケーションとして実現し、コミュニケーションスキル向上に資するこ

とが本研究の目的である。

2. 研究の背景

プレゼンテーションは、ビジネスの場・学術の場別け隔てなく必要不可欠なスキルの1つである。書店に行けばプレゼンテーションの指南書と呼べるような書籍を多数見つけることができる。しかし、書籍から学ぶだけだと実感が湧きにくく、書籍による学習だけで自分のプレゼンテーションを改善できる人は多くない。そのため、一般的には多くの時間と労力をかけて同僚などの前でリハーサルをして改善点を指摘してもらうほか、ひたすら場数を踏むことでプレゼンテーション・スキルを身につけていくのが一般的である。

しかし、リハーサルを行うには大変な手間と労力がかかる。パソコンやスマホの前でプレゼンテーションするだけでその良し悪しを自動的に評価し、ユーザにフィードバックしてくれれば、ちょうどカラオケの自動採点システムのように自力で練習が可能になるのではないかという着想で始めたのがこの研究である。プレゼンテーションを機械学習で自動的に評価するというシステムは当時（そして現在も）ほとんどなかったため、学術的にも非常に挑戦的かつ産業界からのニーズも高いものであった。

3. 提案システムの概要

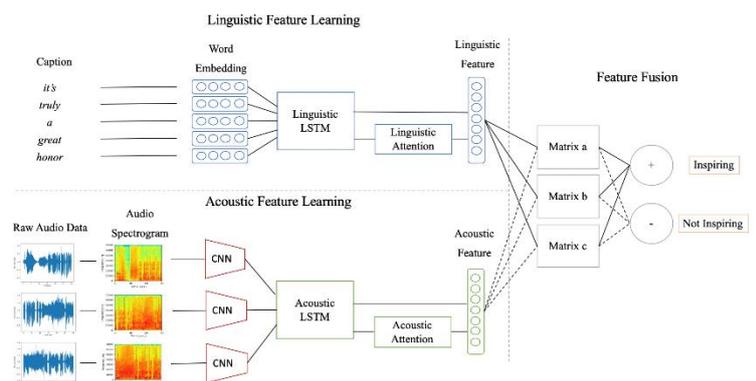


図1 プレゼンテーションを解析する深層学習モデル

現在のプレゼンテーションの印象解析のアーキテクチャを図1に示す。自然言語情報と音響情報を扱って

対話コミュニケーションの能力向上にむけた評価・フィードバックシステム

An Evaluation and Feedback System for Enhancing Communication Skills

いる。言語情報を扱う部分では言葉選びと論理の立て方に注目し、音響情報を扱う部分ではイントネーションやテンポなどの話し方分析に注目しており、それらを総合的に判断してプレゼンテーションが聴講者に与えるであろう印象を判断する [1]。一般的な自然言語解析と異なり、プレゼンテーションは単語数が多く時間も長いので、独自の Attention 機構(*1)を備えることで長時間のデータでも安定して分析できるようにしている。また、印象解析は良い・悪いの 1 軸ではなく、「説得力がある」、「印象的だ」など予め定義された 14 種類の印象語で評価を行っている。この 14 種類の印象以外の評価値が付与されているようなデータがある場合、例えば社内の研修用動画でわかりやすさ評価や受講者の成績などのデータがある場合は、それらを学習データとして追加的に加えることも可能であるという汎用性を持つ。

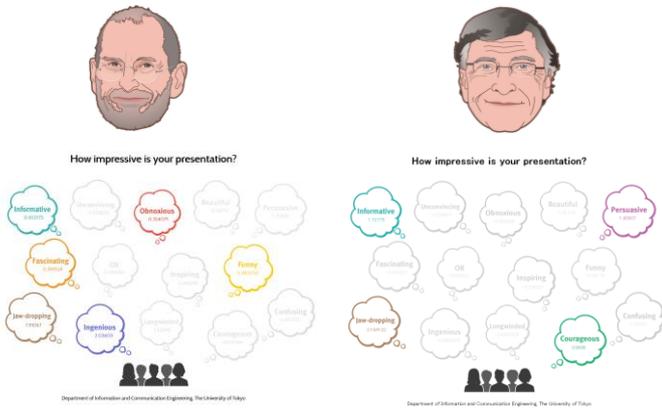


図2 2人のプレゼンテーションに対する印象予測結果

モデルを学習させたあと、初見となる2つの動画を入力して分析した結果を図2に示す。左側のプレゼンテーションの例では面白い、情報量豊か、魅惑的、天才的といった印象が予測されている。一方、右側のプレゼンではそのような華やかな印象は予測されない一方、説得力がある、勇敢だなどの印象が予測されていて、単に良し悪しだけではなくどの様に良いか悪いか分析されていることがわかる。ちなみに、この14種類の印象の予測の精度はTED Talksのデータを用いた実証実験で85%以上を実現している [1]。プレゼンテーション全体についての印象だけでなく、印象の経

時予測も行う事ができる。その様子を図3に示す。

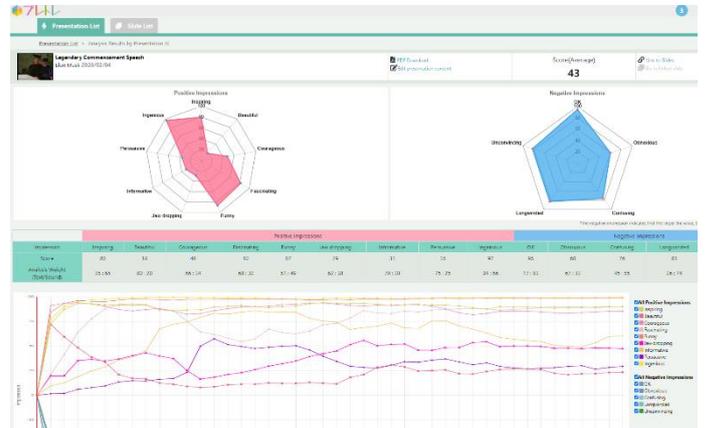


図3 会見の良し悪しを予測する深層学習モデル

これは、印象の予測モデルに対して「その時点でもしプレゼンテーションが終了したとしたら」という仮定で印象を予測させるものである。経時的な印象は正解データがないため精度を確認することはできないが、最終的な印象の予測精度が高いので、経時的な印象の予想もある程度高いことが期待される。これによりユーザには各時点でのプレゼンテーションの良し悪しも判断できる。すなわち、改善すべき点がどこにあるのかをポイントで示すことができる。さらには、長時間のプレゼンテーション動画のハイライト動画を作成する際にもこれらの情報を活用できる可能性がある。なお、どの様に改善すべきかという具体的な指示は現時点では生成しないようにしている。これは、具体的な指示だけに従い、それ以外の改善点に注意がいかなくなるのを防ぐ目的がある。

さらには、一方通行型コミュニケーションであるプレゼンテーションだけでなく、対話型コミュニケーションについても良し悪しを分析する技術を実現した。具体的には、記者会見における発表者と記者とのやり取りを対象に、チェック項目評価を実現した。YouTubeにアップロードされている日本の有名企業の謝罪会見動画合計7本(総計12時間)に対して、記者会見のコンサルティングを行っている株式会社プラップジャパンとの共同研究で11の項目について記者の質問に対して「うまく対応できている・対応がよくない・この評

対話コミュニケーションの能力向上にむけた評価・フィードバックシステム

An Evaluation and Feedback System for Enhancing Communication Skills

「評価項目は当てはまらない」の3値で正解ラベル付けを行い、それを予測する深層学習器を構築した。

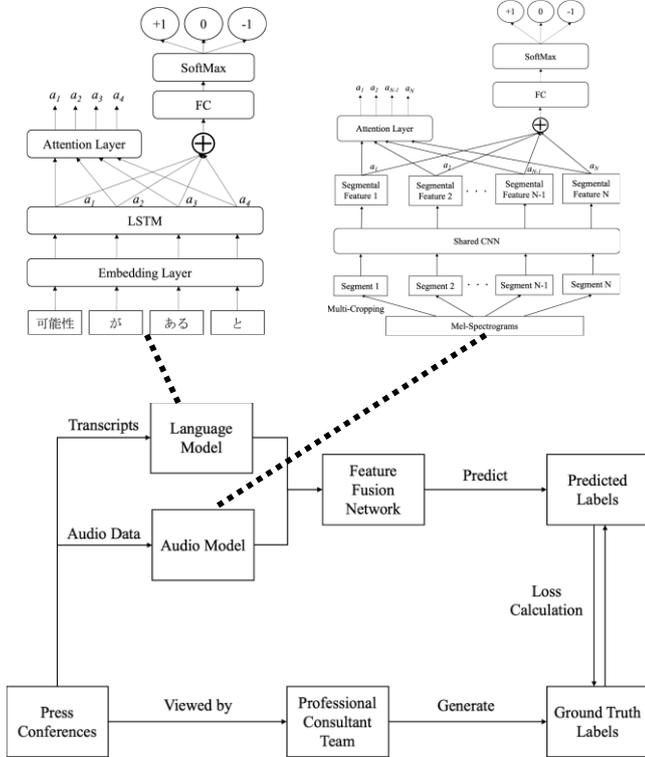


図4 会見の良し悪しを予測する深層学習モデル

テキスト化された会見内容について独自開発した深層学習アーキテクチャ(図4)を用いて、会見者の発言内容だけでなく記者の発言を受けてどのような回答をしているかを学習させた。その結果で57.6%の精度を実現し、ランダムに回答した場合の精度33%を大きく上回った[2][3]。この双方向コミュニケーションの解析は、プレゼンテーション以外にも例えば営業や販売活動における話し方改善などの業務支援のほか、チャットボットにおける相談者の感情推定や返答の良し悪しなどオンラインコミュニケーションにも応用可能である。

人と人のコミュニケーションを考える上では、口頭による言語コミュニケーションだけでなく、スライドなどの資料を用いた非言語コミュニケーションも重要であることが多い。そのため、スライドの良し悪しを判定し、改善すべき点を深層学習が自動判定して可視化するシステムを作成した(図5)[4][5]。

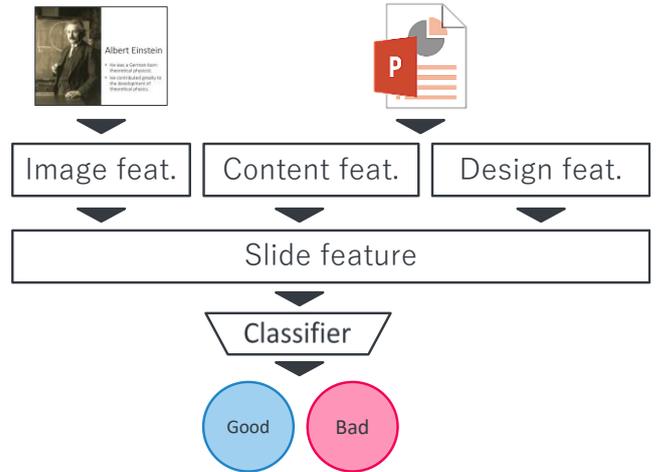


図5 スライドの良し悪しを分析する深層学習モデル

このシステムの実現のために、大規模なクラウドソーシングと独自の深層学習アーキテクチャの開発を行った。さらには株式会社Rubatoとの共同研究により、プロのスライド作成コンサルタントが特に注目する十項目について対応できているかを評価するシステムを実現した[5]。本システムの研究開発においては、スライドの良い例と悪い例の数をバランス良く揃えることが困難であるという学習データのアンバランス問題(*2)に直面し、それを解決する新しい方法を提案した。そのシステムを使ってスライドデザインを改善できた例を図6に示す。

研究開発したシステムの一部は株式会社ピーアンドアイにライセンス提供し、同社からプレトレ¹というサービスでビジネスパーソン向けに提供されており、継続的に機能のアップデートを行っている。本システムを使うことでパソコンやスマートホンの前でいつでもどこでも自分自身のプレゼンテーションの振り返りを行うことができる。国際会議AAAI2020で行ったデモ発表[6]では、デモシステムの前から人が絶えず、フライヤーもほとんどはけてしまうほどの盛況ぶりであった。

¹ <https://www.presentationtrainer.jp/>

対話コミュニケーションの能力向上にむけた評価・フィードバックシステム

An Evaluation and Feedback System for Enhancing Communication Skills

あると考えており、今後産学の共同研究などによって適用範囲を広げていきたい。

おわりに

人のコミュニケーションスキルを客観的かつ定量的に評価する研究や、それを実際にサービスとして実装した例は世界的に見ても少ない。これにより、これまで属人的なものとしてされてきたコミュニケーションスキルを誰でも簡単に改善・向上させることができればよいと願っている。また、このような研究はコミュニケーションスキルだけにとどまらず様々な学術・産業分野に応用展開が可能である。本プロジェクトを推敲する機会を与えてくださった(一財)テレコム先端技術研究支援センターの皆様には深く感謝の意を表す。

用語解説

- *1 Attention 機構：深層学習において、入力データ中の重要度の違いに応じて重みを変更する機構のこと。
- *2 アンバランス問題：一般的に、機械学習では全てのクラスの学習データの数がほぼ同数であるときに所望の性能を発揮する事が多い。仮にクラスによってデータの数に違いがあると、最頻出クラスを優先的に予測し続ける安易なモデルが作成されてしまうことがある。これがアンバランス問題である。

参考文献

[1] Toshihiko Yamasaki, Yusuke Fukushima, Ryosuke Furuta, L. Sun, Kiyoharu Aizawa, and D. Bollegala, "Prediction of User Ratings of Oral Presentations using Label Relations," ACMMM Workshop on Affect and Sentiment in Multimedia (ACMMM ASM), pp. 33-38, 2015.

[2] Shengzhou Yi, Koshiro Mochitomi, Isao Suzuki, Xueting Wang, and Toshihiko Yamasaki, "Attention-based LSTM for Automatic Evaluation of Press Conferences," IEEE International Conference on Multimedia Information Processing and

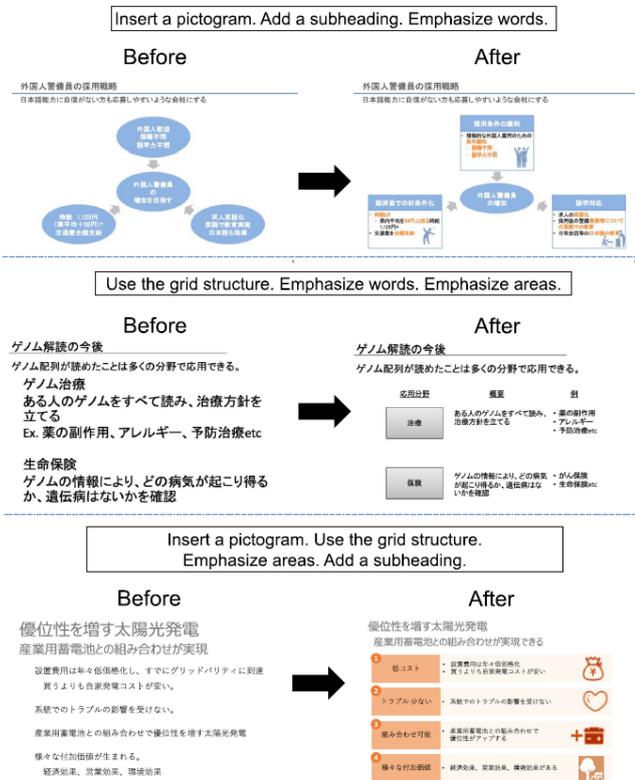


図6 スライドデザインの改善例

4. 将来展望

研究開発したシステムの一部は民間企業にソフトウェアライセンス提供という形で実サービスに結びつき、研究成果の社会還元に至っている。今後は利用者拡大を目指すとともに、評価・フィードバックモデルの拡充を行ってきたい。機械学習は、当然のことながら学習データに依存して出力が変化するのである。対話コミュニケーションには唯一無二の正解というのは想定し難く、様々なアプローチや解があるものだと考えている。例えば、産業界、学术界、政府や地方公共団体などによって理想とするコミュニケーションは異なるはずである。また、裏を返せばそれぞれの分野で蓄積されたデータを用いて学習すれば異なる複数のモデルが生成できることであり、今後の研究の発展が期待される。

本研究で実現した評価・フィードバックはプレゼンテーションや会見に限ったものではない。営業・店舗販売、Eコマース、さらには教育などにも発展可能で

対話コミュニケーションの能力向上にむけた評価・フィードバックシステム

An Evaluation and Feedback System for Enhancing Communication Skills

Retrieval (MIPR), Full paper, 2020.

[3] Shengzhou Yi, Koshiro Mochitomi, Isao Suzuki, Xueting Wang, and Toshihiko Yamasaki, “Attention-Based Multimodal Neural Network for Automatic Evaluation of Press Conferences,” *International Journal of Multimedia Data Engineering and Management (IJMDEM)*, Volume: 11, Issue: 3, Article: 1, Pages: 19, 2020.

[4] Shinji Oyama and Toshihiko Yamasaki, “Visual Clarity Analysis and Improvement Support for Presentation Slides” *The IEEE International Conference on Multimedia Big Data (BigMM)*, long, pp. 421-428, 2019.

[5] Shengzhou Yi, Junichiro Matsugami, and Toshihiko Yamasaki, “Assessment System of Presentation Slide Design using Visual and Structural Features,” *IEICE Transactions on Information and Systems*, Vol. E105-D, No. 3, Mar. 2021.

[6] Shengzhou Yi, Hiroshi Yumoto, Xueting Wang and Toshihiko Yamasaki, “PresentationTrainer: Oral Presentation Support System for Impression-related Feedback,” *The Thirty-Fourth AAAI Conference on Artificial Intelligence (AAAI), Demonstrations*, volume 34, no. 9, pp. 13644-13645, 2020.

この研究は、平成29年度SCAT30周年記念研究費助成の対象として採用され、平成30年度～令和2年度に実施されたものです。