

スライド発表における癖検出機構に基づく発表練習支援システム

松浦 辰雄[†] 大園 忠親[†] 新谷 虎松[†]

[†]名古屋工業大学大学院情報工学専攻

1 はじめに

学会発表などで自分の研究結果をわかりやすく伝えるために、スライドを用いた発表を行う機会は多い。しかし、スライドを用いた発表の際に、緊張や不安から発表者は声の大きさが小さくなったり、口癖が出てしまうなど発表者の癖が出てしまうことがある。そのため発表者は、事前に発表練習を行うが、発表者は無意識な癖により自身の癖に気づけなかったり、発表中に客観的に自分の発表を評価するのは困難である。

本研究では、PC 内蔵マイク、カメラおよび Kinect を用いて発表者がスライドで発表している様子を発表履歴として記録し、ユーザの発表の癖を検出する発表練習支援システムを開発した。ユーザは、過去の発表履歴を閲覧し、癖の出ている箇所や回数を可視化することで客観的に評価し、発表練習を効果的に行うことができる。本稿では、発表中の発表者の視線、声量、早口および口癖を検出する癖検出機構について述べる。

2 癖検出機構による癖検出

本節では、発表中の発表者の癖を検出する癖検出機構について述べる。スライド発表において、声量不足や早口などで聴講者は発表内容を聞き取り辛かったり、口癖により発表内容に集中できないことがある。また、無意識に出てしまう癖は、ユーザ自身気づけなかったりするため発表の客観的な評価が困難である。そこで、本システムでは、PC 内蔵マイク、カメラおよび Kinect を用いて、発表履歴として動画情報、音声情報、骨格情報およびスライド情報を記録し、発表履歴から癖検出機構でユーザの癖を検出し可視化を行う。

発表中の非言語である身体の動きに注目し、身体の動きを検出しフィードバックを行い、発表の向上を図る手法もあり [1]、本研究では、非言語的な癖としてユーザの視線を検出する。癖検出機構で検出する癖の種類

A Presentation Practice Support System based on a Habit Detection Mechanism in a Slide Presentation
Tatsuo MATSUURA[†], Tadachika OZONO[†] and Toramatsu SHINTANI[†]

[†]Department of Computer Science, Graduate School of Engineering, Nagoya Institute of Technology.

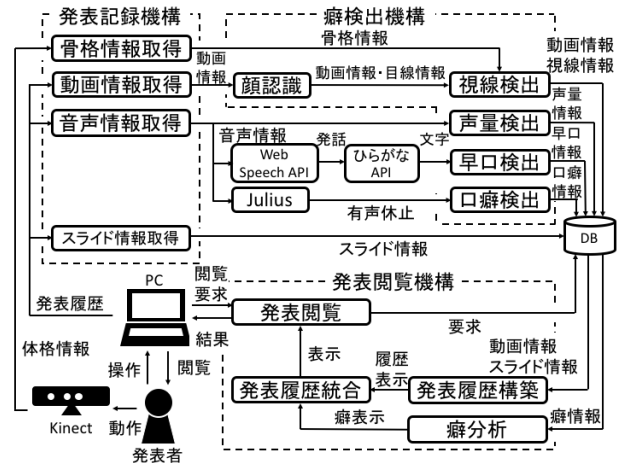


図 1: システム構成図

は、視線、声量、早口および口癖である。視線検出は、動画情報から顔認識を行い、発表者の目の位置を検出することで PC 画面に視線が向いているか検出する。さらに、Kinect の骨格情報から体の位置および傾きを検出し、スクリーンに向いているかを判定する。視線検出を行うことで発表者が聴講者の方向に視線が向いているかを判定する。声量の検出には、PC 内蔵マイクで声量 (dB) を計測し、声量が小さくなっている箇所を検出する。早口の検出には、PC 内蔵マイクにより音声認識である Web Speech API¹を用いて発話を記録する。記録した発話に対して、goo ラボ API のひらがな API²を用いて平仮名変換し、1 分あたりの発話の文字数の計測を行い判定する。口癖の検出には、「えー」と「あー」などの有声休止に対して検出を行う。音声情報に対して大語彙連続音声認識エンジンである Julius³を用いて音声認識を行う。有声休止用の Julius の辞書を作成し、有声休止を検出する。

3 発表練習支援システム

本節では、発表練習支援システムについて述べる。本システムは、発表者の発表の様子を記録および癖を検

¹https://developer.mozilla.org/ja/docs/Web/API/Web_Speech_API

²<https://labs.goo.ne.jp/api/jp/hiragana-translation/>

³<http://julius.sourceforge.jp>



図 2: 発表履歴の閲覧

出し、記録した発表および検出された癖を見直すことで発表の練習を支援する。本システムは発表記録機構、癖検出機構および発表閲覧機構で構成されており、図 1 を用いて説明する。

発表記録機構は PC 内蔵マイク、カメラおよび Kinect を用いて発表履歴を記録する。発表履歴としては、発表者の動画情報、音声情報、骨格情報およびスライド情報を取得する。記録された発表履歴は癖検出機構に渡され、癖検出を行う。また、動画情報およびスライド情報はデータベースに保存される。

癖検出機構では、視線検出、声量検出、早口検出および口癖検出を行う。視線検出では、動画情報から顔認識を行い目線情報と骨格情報から発表者の体の向きを取得し、発表者の視線が聴講者に向いているか判定する。声量検出では、音声情報から声量 (dB) を測定し、声量が小さい箇所を特定する。早口検出では、音声情報から Web Speech API で発話認識を行い、発話認識結果をひらがな API で平仮名変換をして文字数を計測する。口癖検出は、音声情報に対して Julius で音声認識を行い、Julius の辞書に登録されている有声休止を検出する。癖検出機構で検出された癖の情報はデータベースに保存される。

発表閲覧機構では、発表履歴および癖検出機構で検出された癖を閲覧し、過去の発表を閲覧し確認することができる。ユーザは、閲覧したい過去の発表を選択すると、データベースから、動画情報およびスライド情報を取得し、発表履歴として構築する。さらに、検出された癖情報を用いて癖分析を行い、癖が出ている部分を表示する。発表履歴と癖情報を統合し、ユーザが過去の発表履歴を確認することで発表練習の支援を行うことができる。

3.1 発表履歴の閲覧

本節では、ユーザが発表履歴を閲覧している様子を示している図 2 の実行例について説明する。図 2 の (A) では、発表ごとのスライドの 1 枚目のサムネイル画像が表示されている。図 2 の (B) では、図 2 の (A) で選択した発表のスライドのサムネイル画像が 1 枚ずつ表示されている。図 2 の (C) では、図 2 の (B) で選択したスライドの詳細情報として、発表している様子の動画と癖が検出された部分と回数が表示されている。表示されている癖の種類は、視線、早口、声量および口癖である。声量のグラフが表示されており、声量が小さくなっている部分が赤く表示される。視線、早口および口癖が検出されている部分も赤く表示され、声量のグラフと連動している。

3.2 考察

図 2 の本実行例では、スライドごとの発表している様子の動画および癖が検出されている部分を詳細に閲覧することができた。発表履歴に含まれる動画と癖が出てくる部分を表示することで、ユーザは癖が出ているタイミングや傾向を確認し、癖に注意しながら効果的に発表を行うことができた。しかし、口癖は有声休止のみだけではなく、意味のある言葉ではあるが繰り返す同じ言葉を用いることで聴講者に不快を与えてしまうことがある。本システムでは、口癖に関しては有声休止にのみ検出を行なったが、他にも、ユーザが多用してしまう独特な言い回しを検出する必要がある。また、視線以外にも非言語的な身体の動きの癖を検出し、ユーザに示すことでさらなる発表の向上が見込まれる。

4 おわりに

本稿では、スライド発表をしている様子を発表履歴として記録を行い、ユーザの癖を検出する発表練習支援システムを提案した。本システムを用いることで、ユーザは過去に行った発表の閲覧および癖を確認し、効果的に発表練習を行うことができた。

参考文献

- [1] Jan Schneider, Dirk Borner, Peter van Rosmalen and Marcus Specht. "Presentation Trainer, your Public Speaking Multimodal Coach", Proceedings of the 2015 ACM on International Conference on Multimodal Interaction, pp.539-546, 2015.