

教員のためのスピーチ評価システム

新井浩*, 南保英孝**, 瀬戸就一*, 下村有子***, 川辺弘之***

*金城大学短期大学部, **金沢大学, ***金城大学

arahiro@kinjo.ac.jp

概要: 大学教員は、教育者としての指導法や教授法について専門的教育を受けていない。また平均年齢が高く、客観的評価を受ける機会が少ないため、評価されることを嫌う傾向にある。そこで我々は、自らのスピーチを評価するシステムを構築している。音声データから声の大小・高低・速遅・間合い・明瞭さ等のパラメータを抽出し、お手本となるスピーチデータと比較している。話者はマイクに向かって話すだけで、システムが顔グラフで評価を与える。

1 はじめに

我々は聴覚障害学生支援を目的とした教員の講義の内容や講義の雰囲気伝えるため、講義内容を聴覚障害者に正確に伝える技術についての調査・研究を行ってきた[1]。授業における非言語的な情報である「臨場感」や「感情」という教室内のドラマ性が、学生の講義理解を豊かなものとするため欠かせないものであると主張し、聴覚障害者が豊かな情報を感じられる字幕システムの構築するなかで「感情フォント」の提唱を行っている [2] [3] [4]。

「感情フォント」システムの構築にあたっては、音声認識技術を用いて、音声から言語情報と非言語情報を抽出することを試みている。これら2つの情報は音声特性データとみなすことができ、教員教育研究のための基礎データとして活用できる。

そこで我々は大学教員の教授力向上のために教員の話し方、特に「言語的・非言語的コミュニケーション能力」を成長させることを目標とした



図1 教員のためのスピーチ自習システム

「教員の為のスピーチ自習システム」(図1参照)を開発する。また本研究において、話し方の良否といった複合的な判定を表示する為に、チャーノフが創案した顔形グラフを利用し、より直感的に理解できる顔形ユーザーインターフェースの構築を目指している[5]。

2 自習システムの概要

マイクより入力されたユーザ音声は、音声認識と付加情報抽出し、音声特性分析が行われる(図2参照)。音声認識においては認識された文字と認識できた割合でスピーチの評価を行う。

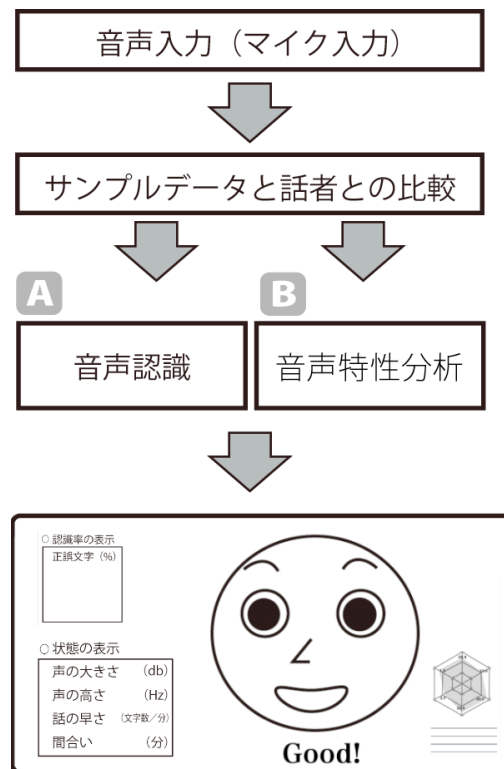


図2 教員のためのスピーチ自習システム概要

一方、音声特性分析では、声の大小、声の高低、間合い、話す速さの4要素を抽出する。その結果から、話者のスピーチの欠点をレーダーチャートグラフや顔型グラフを組み合わせて作成する顔形ユーザーインターフェースで表示し、改善を促す。

3 チャーノフの顔形グラフ

学習するにあたり学習者の学習意欲を保つため親しみやすく直感的なインターフェースの採用が必要である。

多くの数値データを使用者に効果的に伝えるため、グラフを用いて視覚的明快さを付加する。これまで使用されてきたグラフ表現では、1変量または2変量のデータでは棒グラフや折れ線グラフ、散布図が使用される。一方、音声解析データのような3変量以上の複雑な構造を持っているデータは、レーダーチャートやクモの巣図、3次元の棒グラフなどで表現されてきた。こういった3変量以上を扱うグラフ表現の中にチャーノフが創案した顔形グラフ(図3参照)がある。これは多数のデータを目、鼻、口、眉など顔のパーツ形状に対応させ、一つのデータを人間の顔が作る表情として示すものである[6][7]。

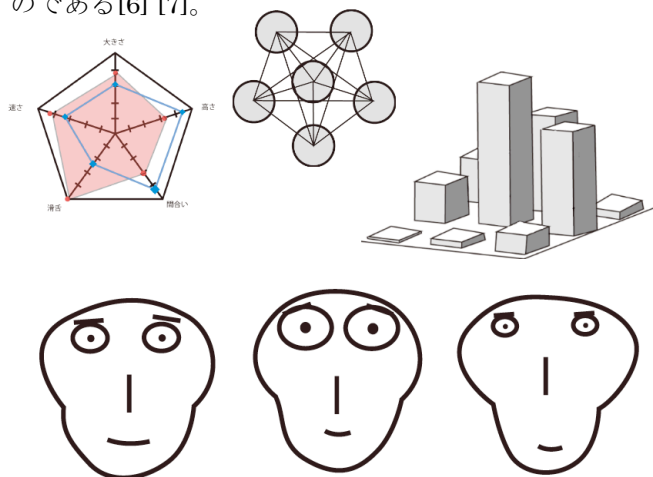


図3 多次元グラフの例と顔形グラフの変化

顔形グラフではデータ量を表すだけでなく、表情から読み取れる印象に良否の判断も付け加えることができ、また表情の移り変わりで時間的な変化を表す事もできる。

評価と顔形グラフの表情が一致するよう調整

することで、あたかも聴講する学生が眼前にいるような感覚で表示させる事ができ、話者は直感的に評価結果を感じる事ができる。

使用例の中にはデータ分布と表情の連動が適切に調整されていない例も見られ、顔形要素へ十分な吟味をせず変数を割り当てると、表情の印象より結果に誤解をまねきやすい。

4 顔形グラフのルールづくり

2つの段階を経て話し方の評価を表示する。第1段階は声の大小、声の高低、間合い、話す速さ、明瞭さの5つパラメータを顔のパーツへ反映し表示する。眉、目の開き具合、口角の角度などは変化する部位から除外し、表情の喜怒哀楽に反映しないようにした。また、声の大きさは目の大きさに、声の高さは顔パーツの表示位置の高さ、など変化する部位と連動している音声特性を想像しやすいように調整した(図4参照)。

音声特性	変化する部位	小	大
大・小	目の大きさ (輪郭)		
高・低	パーツの位置 (上下)		
速・遅	輪郭		
間合い	パーツの位置 (左右)		
明瞭さ	口の大きさ		

図4 顔形グラフの変化例(第1段階)

第2段階は話し方の良否の判定である。事前に5人の評価者がループリック評価を用いて選定した手本音声データを基準値とし、この音声データとの類似性で評価する(図5参照)。評価が高いほどに黒目が大きくなり力のある笑顔になり、評価が下がると目が小さく閉じぎみになり、力のない困った顔になる(図6参照)[8][9]。この様に音声特性を顔要素の配置で表示し、良否

判定は表情としてその後付け加える。それぞれのパラメータの最大値に達する事が、必ずしも良い話し方に繋がるのではなく、適切な音量、速さ、高さ等が評価につながり、この様な状態になった時に顔グラフは最も安定した顔の形を表示する。

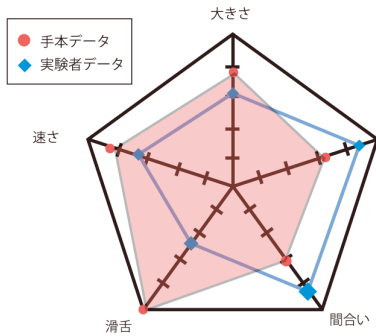
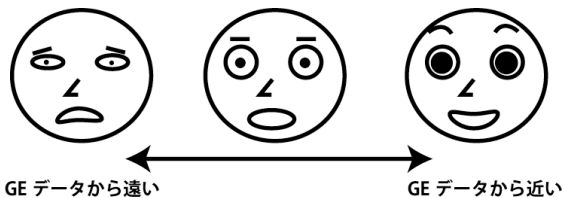


図5 手本データと比較した良否判定



変化する部位	GE データに遠い	GE データから近い
黒目の大きさ	小	大
目の形状	開く	閉じる
口角	下がる	上がる
眉の角度	下がる	上がる
眉と目の距離	せばまる	ひらく
眉の形状	直線	山なり

図6 顔形グラフの変化例 (第2段階)

5 顔形グラフ生成アルゴリズム

上記のルールに基づき、実際に顔形グラフを用いたユーザーインターフェースの作成に入った。今回は音の大小・高低という2つの要素にリアルタイムで反応する顔形グラフを作成した。

はじめは音声信号の取得である。はじめに音声信号は、マイクまたはWAVファイルから入力される。入力される順に沿って、時系列順に1024点のデータが一単位として処理される。

まず、音量の平均値V_aveを求める。これは、各データ(16bit -32768~32767 の値)の絶対値の

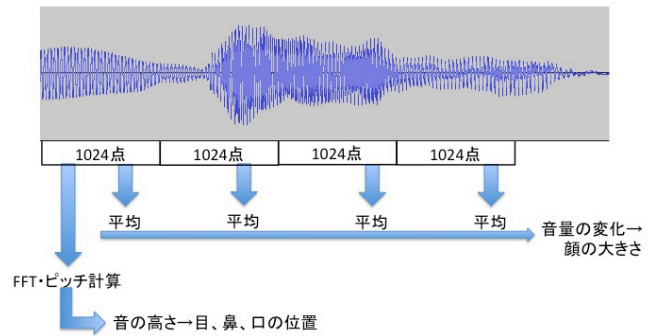


図7 顔形グラフ生成のアルゴリズム

和を1024で割ったものである。

次に音量の計算をする。1024点のデータからピッチを計算し、ピッチの平均値P_aveを求める。なおピッチは1024点全ての区間で求められるわけではない。求められない箇所もあるので、P_aveは求められたピッチの和を求められた区間の数で割ったものになる(図7参照)。

V_aveとP_aveを用いて顔グラフにおける各パツの位置を決定する。

これまでの10単位(過去10240点分)のデータの音量の平均とV_aveを比較し、二つの差が100を超えた場合に顔の大きさを変化させることとする。V_aveが大きくなった場合は顔形状を拡大し、V_aveが小さくなった場合は顔形状を縮小する。

また、P_aveの値が600Hzより大きくなった場合には音程が高い、200Hzより小さくなった場合には音程が低い、それ以外の場合には音程は普通と判断し、それぞれの結果に応じて目・口の表示位置を変更する。変更のルールは4で定めた通り音量の大小は顔の大小に反映され、音の高低は顔つきに反映される(図8参照)

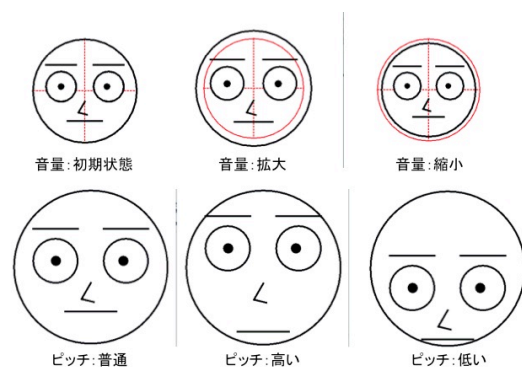


図8 顔形グラフの変更ルール例

6 まとめ

本研究では、「教員の為のスピーチ評価システム」で使われる、直感的に理解できる顔形ユーザーインターフェースのルール作りを提案し、そのルールに基づき顔形グラフを用いたユーザーインターフェースの作成を行った。

今回はプロトタイプとして2つの要素(大小・高低)をリアルタイムで反映することができた。今後は残り3要素(速遅・間合い・明瞭さ)を反映できるよう顔形グラフの構築をすすめていく。その後実地テストを重ねながら顔形グラフに反映される値を微調整し適切な値をさぐる。

謝辞

本研究は文部科学省平成 23 年度科研費<挑戦的萌芽研究:課題番号 23650556>の支援を受けた。

参考文献

- [1] Kawabe, H., Sugimori, K., Shimomura, Y. and Seto, S., Computer Simulation of Text Reproduction Based upon “Quantity-Rather-Than-Quality” Concept, Proceeding of the 40th International Conference on Computers and Industrial Engineering, 2010
- [2] Seto, S., Arai, H., Sugimori, S., Shimomura, Y. and Kawabe, H., Subtitle system visualizing non-verbal expressions in voice for hearing impaired - Ambient Font -, Proceeding of the 10th Asia-Pacific Industrial Engineering and Management Systems, 2010
- [3] 瀬戸就一, 南保英孝, 新井浩, 川辺弘之, 杉森公一, 下村有子, 「聴覚障害学生に教員の口調と授業の雰囲気伝えるシステム」, 情報処理学会第 75 回全国大会(2013)
- [4] 下村有子, 瀬戸就一, 南保英孝, 新井浩, 川辺弘之, 杉森公一, 「聴覚障害学生に授業の雰囲気伝えるシステムの構築」, 情報処理学会第 74 回全国大会(2012)
- [5] 杉森公一, 新井浩, 川辺弘之, 下村有子, 瀬戸就一, 「自ら学習・訓練する大学教授のスピーチ評価システム」, ヒューマンインタフェースシンポジウム (2012)
- [6] 田中豊, 脇本和昌 著「多変量統計解析法」現代数学社 (1983)

- [7] 脇本和昌「チャーノフの顔形グラフの変形のこころみ-体型グラフ-」行動計量学 4 巻 2 号 (1977)
- [8] Gary Faigin 著 みつじまちこ 訳 「表情顔の微妙な表情を描く」マール社 (2005)
- [9] P・エクマン, W・V・フリーセン 著 工藤力 訳編 「表情分析入門 表情に隠された意味をさぐる」誠信書房 (1987)