

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

1 はじめに

近年、大学教育の「質」に対する社会的要請が高まってきている。文部科学省は「学士力」を定義し、大学の組織的なファカルティ・ディベロップメント (FD) 活動、教員の自己研鑽を求めている。この背景には、大学全入時代を迎えての大学入学者の質的变化のみならず、大学教員の教授能力に疑問符がつけられていることがある。

本研究では、大学教員の教授力向上のために教員のスピーチ能力向上を目標とした「音声認識技術を活用した自習可能な口話評価システム」を開発することである。

2 評価システムの概要

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1)。音声認識エンジンの認識率で明瞭さを測定する。また、波形データによる音声特性分析では、声の大小、声の高低、間合い、話す速さの4要素を抽出する。予め用意されたお手本音声データ (以下 GE data) と比較し、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る作成する顔型ユーザーインターフェースで表示し改善を促す。

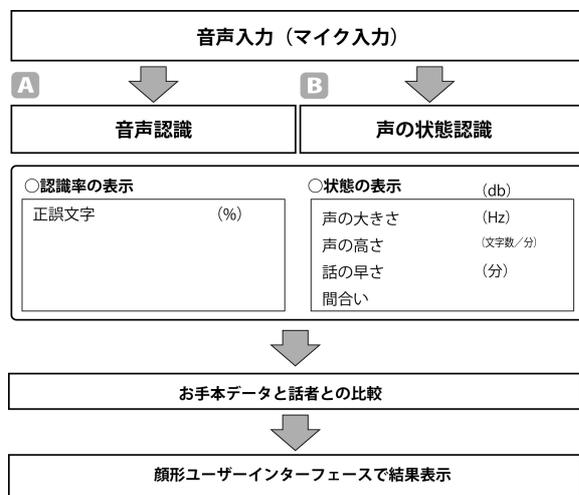


図1 評価システムの概要

3 顔型ユーザーインターフェースの構築

顔型グラフは図2の様に大小、高低などが人相の変化として現れる。バランスの良い話し方をすると人相のバランスは整い、良くない話し方をすると崩れる。

また、顔型グラフは GE Data と話し方が近くなると笑顔になり、遠ざかると困った顔になる。実験者は顔型グラフが整った笑顔になるよう話すことを目指す。

顔型グラフの読み取りに順応することで、実験者は顔グラフから自分の発声状態を理解できるようになるが、慣れないうちは難しい。そこで GE data との差異を数値的に理解できるレーダーチャート等のグラフと組み合わせ、これを顔型ユーザーインターフェースとする。

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2 顔型グラフ変化のルール

4 顔型グラフ生成アルゴリズム

上記のルールに基づき、実際に顔型グラフを用いたユーザーインターフェースを作成し評価システムを構築する。

音声信号はマイクから入力され、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が行われる (図3)。

顔型グラフを作成するために、音声信号から種々のパラメータを抽出する。まず、周波数解

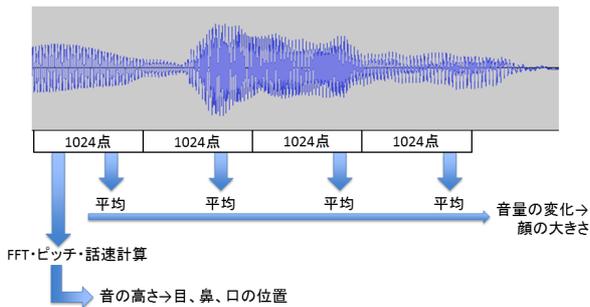


図3 音声信号からの音声特性抽出

析によってピッチを検出する。その後、ピッチが検出された区間を音声が入力された区間だとみなし、この区間での音量や話速など他のパラメータを求める。これらのパラメータを用い、以下のようにして顔グラフにおける各パーツの形状と位置を決定する。

まず、事前にGE Dataの各パラメータの平均値を求めておく。例えば、音声の平均を V_{ave} 、ピッチの平均を P_{ave} とする。

次に、直前の処理単位（1024点分）のデータにおける音量の平均と V_{ave} を比較し、差がある範囲 V_1 以内にあるときを2点、ある範囲 V_2 以内（ただし、 $V_1 < V_2$ ）にある時を1点、そうでない場合を0点とする。また、差が範囲 V_1 より大きいときは、 V_{ave} と比較した音量の大小により、図2のルールに基づき顔のパーツ位置を決定する。ピッチやその他のパラメータについても、それぞれのパラメータに応じた範囲を定め、同様に処理する。

最後に、各パラメータに対し、上記の処理で求めた点数の合計を求め、合計値により顔パーツの形状を決める。合計値が高いほど目や眉が笑顔の形状になり、低いほど悲しい顔の形状になる。

5 検証とまとめ

本研究では、「教員の為のスピーチ評価システム」を考案し、顔形ユーザーインターフェースのルール作りの検討とそれに従ったシステム構築を行った。結果として、変更のルールは4で定めた通り音声特性が反映され、話者のスピーチ能力を判定した。

今回は音声信号から4つの特徴を抽出し、リアルタイムで顔形グラフを生成するアルゴリズムを構築し、プロトタイプシステムを実装した。プロトタイプシステムは問題なく動作したが、

今後は音声認識による明瞭さの抽出アルゴリズムと、それらをリアルタイムに顔形グラフに反映させるシステム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調整し適切な値をさぐる予定である。

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, 「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, 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, 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)