

唇検出による英語発音指導教材への適応の可能性

Development of English pronunciation materials
utilizing shape detection of learners' lips伊藤 敏^{*1}, 大石 晴美^{*2}, 鷺野 嘉映^{*3}Satoshi ITOU^{*1}, Harumi OISHI^{*2}, Kaei WASHINO^{*3}^{*1} 岐阜聖徳学園大学経済情報学部^{*1} Department of Economics and Information, Gifu Shotoku Gakuen University^{*2} 岐阜聖徳学園大学教育学部^{*2} Department of Education, Gifu Shotoku Gakuen University^{*3} 岐阜聖徳学園大学看護学部^{*3} Department of Nursing, Gifu Shotoku Gakuen University

Email: itous@gifu.shotoku.ac.jp

あらまし：口唇の形状は英語発音時の重要な要素の一つである。そこで、音読時の口唇の形状を取得し、学習者へフィードバックする教材を目指す。口唇形状の検出には、顔の特徴点検出が可能である RealSense SDK を用いた。口唇形状の把握には口唇の位置座標を指標として用い、その指標が教材適応に有効であるかの検討を行った。結果、単一の単語発音レベルでは位置座標の指標が有効であることを確認した。

キーワード：唇形状検出、英単語発音、RealSense SDK、行動解析

1. はじめに

学習者の負荷時の行動を観測し、その結果を学習者へフィードバックすることが学習効果の向上につながると考えられる。我々は学習者にセンサを装着して、学習者の行動を観測する試みを行った¹⁾。課題中の頭部の動きや、脈波、呼吸数を取得することで負荷の状態などを推測した。しかし、検出装置を装着することによる学習者への負担の問題がある。よって、非接触で学習負荷時の行動を記録・解析する方法の検討も重要であると考えられる。

学習者の行動情報を非接触で取得する手段として、Kinect²⁾と Leap Motion³⁾が知られている。これらのデバイスは、学習者に装置を装着するのではなく、非接触で行動を観測可能である。また、ゲームなどの入力デバイスとしても盛んに利用されている。Kinect は全身骨格の位置や形状や動き検出が可能であり、スキル学習支援やその他多方面への適応研究が行われている。Leap Motion は手や指の位置や形状や動きを検出可能である。本研究で用いる Intel の RealSense SDK⁴⁾は Kinect と Leap Motion の中間距離領域の検出範囲を持ち、座位学習者の観測に適した検出範囲を持つ。観測可能な量は顔の検出、顔の特徴点座標、傾き、表情などである。

今回我々は、RealSense SDK を用いた正面からの録画画像から得られる口唇の形状把握に着目した。学習者の英単語発音時の口唇の形状を検出し、その結果を提示し、正しい発音に必要な口唇形状との差を提示することで、発音の修正に寄与可能な教材開発を目指す。そのための第一歩として、口唇の形状を検出する方法の開発と検証を行う。

英語発音学習では、母音と子音の特徴を把握することが必要である。音声学の基準では、母音は、唇の形、舌の位置と形、顎の開閉度で分類される⁵⁾。唇が丸い場合を円唇母音、そうでない場合を非円唇母音と言う。口の中で、舌の上がり方で舌の頂上が前にあるものを前舌母音、中位にあるものを中舌母音、後方にあるものを後舌母音と呼ぶ。

また、子音は、舌、歯、唇、声門を使用して口の中で空気の流れを調節することによって発する音である。舌の位置(両唇音、歯茎音、硬口蓋音、軟口蓋音)と調音法(閉鎖音、破裂音、摩擦音、破擦音、鼻音、側音)によって、音が決められる。

本研究では、母音発音時の唇の形と顎の開閉度に限定した。唇の開き具合と舌の位置を模式的に表した International Phonetic Alphabet (IPA)を参考に、円唇母音、非円唇音母音、顎の開き度合いで、学習者が発する音の判別をする。

本稿では、2章で口唇検出の方法開発と検証を行い、3章で英語音読時の発音指導教材としての可能性を検討し、4章でまとめる。

2. 方法

2.1 RealSense SDK の唇検出への適用

RealSense SDK は、カメラによる顔検出機能を持ち、顔の特徴点として顔の輪郭や口唇の位置など 78 点を抽出することが可能な仕様である⁴⁾。カメラ SR300 を用いた口唇の検出において、図 1 に示すように唇の位置に 20 個の特徴点表示が可能であり、仕様通りに唇位置座標が検出可能であることを確かめた。さらに、音読中の学習者の顔面を録画し、その

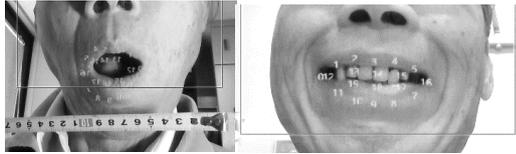


図 1. RealSense SDK による唇の index 表示

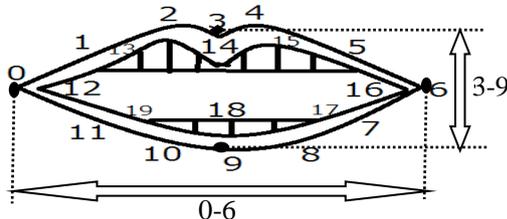


図 2. 唇に割り振られた index

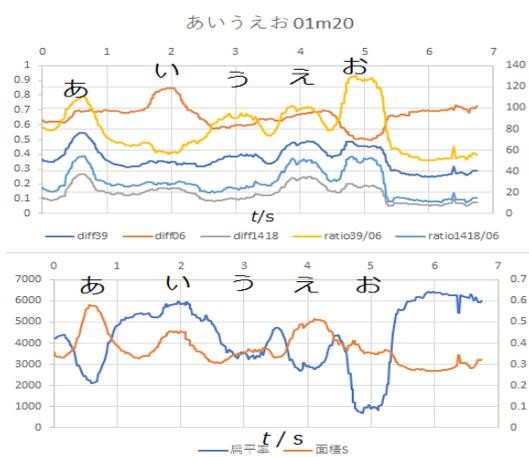


図 3. 「あいうえお」の音読による index 座標と比

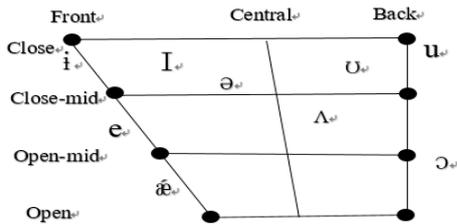


図 4. International Phonetic Alphabet

動画から同様の唇座標検出が可能であることを確かめた。

図 2 に示すように唇の位置を 20 個の index に割振り、各 index の位置座標と、口唇の横幅を示す index 0-6 から、以下の指標を解析に用いた。diff39, diff06, diff1418 は、それぞれ二点の位置座標の差であり、ratio39/06 と ratio1418/06 は、二点の位置座標の差を口唇の横幅に対応する dif06 で除した指標である。Area は口唇の面積、oblate は口唇の扁平率を示す。

2.2 日本語発音への適用

「あいうえお」を区切りながら音読した結果を、

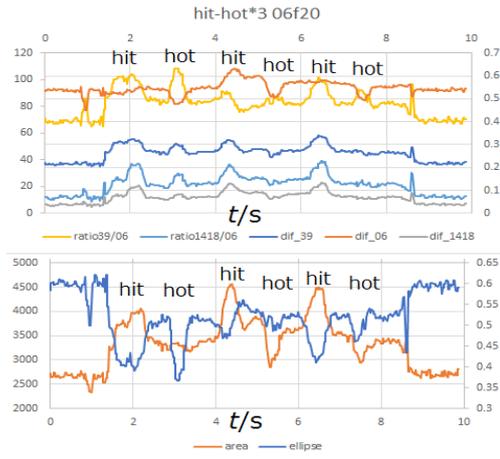


図 5. hit-hot を発音時の口唇の形状

経時変化として図 3 に示す。「あ」では縦横ともに大きく広がり、「い」は横へ広がるなど、発音と検出した唇の形状は対応する結果が得られた。実験した 8 名に同じ傾向がみられた。一文字ずつ区切りながら発音をする方法では、日本語に関しては発音と唇の形状は対応関係を持つことが示された。

3. 英単語発音への適用の可能性

英単語を発音し、母音部分の唇形状を調べた。音読する単語の選定は図 4 に示す IPA の唇の開き具合（縦軸）を参考に、唇形状が遠、中、近距離の次の単語群を選定した。唇形状に近い母音ほど発音区別難易度が上がる。（遠距離：hit [hit] - hot [hɒt]，中距離：hit [hit] - hat [hæt]，近距離：hat [hæt] - hut [hʌt]）

円唇と非円唇母音を含む hit-hot を区切りながら 3 回発音をした結果を図 5 に示す。hit と hot 発音時の口唇形状の違いが検出できている。

4. まとめ

本方法により音読中の口唇の位置座標を非接触で検出可能であり、日本語及び英語の発音する単語により口唇座標が異なることを確かめた。英語の発音指導教材に活用できる可能性を見出した。

本研究の一部は科研費 (16K01083) の助成を受けたものである。

参考文献

- (1) 伊藤敏, 王琳琳, 鷺野嘉映, 井上祥史: “慣性センサを用いた行動検出試行”, 教育システム情報学会 2016 年度第 2 回研究会, pp9-13 (2016)
- (2) <https://developer.microsoft.com/ja-jp/windows/kinect/develop> 2017 年 5 月 23 日確認
- (3) <https://developer.leapmotion.com/> 2017 年 5 月 23 日確認
- (4) <https://software.intel.com/en-us/intel-realsense-sdk> 2017 年 5 月 23 日確認
- (5) Peter Rosh: "English Phonetics and Phonology: A Practical Course", 4th ed. (2010) http://assets.cambridge.org/9780521717403/excerpt/9780521717403_excerpt.pdf