

[研究助成 (B)]

歌唱による嚥下機能保持・向上のための人工知能と
脳科学の融合システム基盤の開発

2221905

研究代表者
(助成金受領者)
共同研究者兵庫県立大学
先端医療工学研究所
関西医科大学
リハビリテーション学部准教授 八 木 直 美
准教授 前 澤 仁 志

[研究の目的]

肺炎は、超高齢化社会で大きな死亡原因になっている。肺炎の原因である誤嚥を歌唱によって防止したい。従来、医療従事者が咽頭運動の評価を行うため多大なコストを要す。また、嚥下機能保持・向上の方法もない。本研究では、喉周辺の AI による動画像処理で新規特徴量を発見し、嚥下機能評価の新指標・方法を解明することを目的とする。次に、脳磁図計測により嚥下運動の中枢制御機構から誤嚥誘発のメカニズムを解明して、歌唱による咽頭筋強化が中枢に与える可塑性変化を明確にし、歌唱が有効である医学的根拠を確立する。このように、歌唱による喉・脳の動態を機械により見える化して、嚥下機能の保持・向上システムを研究する。これが人間（歌唱）と機械の調和の促進である。

嚥下障害の評価・診断は、言語聴覚士や耳鼻咽喉科医、リハビリテーション科医等の主観的評価にゆだねられている。先行研究では、歌唱者は声帯健康状態が良いと報告されているが、嚥下能力への影響は一定の見解が得られていない。以前に、新型コロナウイルスの蔓延で、日本嚥下医学会は嚥下障害診断指針の中で詳細な感染予防策を作成しているが、解決可能な遠隔診断システムは存在しない。更に、いくつかの動態解析が試みられているが、患者と医療従事者双方の身体、心理的負担を軽減されたものは、未だ临床上広く使用されている現状にはない。

一方、歌と声のトレーニングは、咀嚼、嚥下、呼吸機能に大きな影響を与えている。歌唱者が嚥下反射の惹起性において統計的に優れているという医学的根拠立証することを研究デザインし、これを検証することを目指す。

[研究の内容、成果]

1. 動態解析

歌唱経験者と非歌唱経験を分類し、反復唾液嚥下テスト (repetitive saliva swallowing test: RSST) を評価基準として、喉頭部の動態解析および動画像診断が嚥下評価に有効であることを解明する。次に、脳機能計測による大脳皮質運動野の可塑性を評価し、これによって歌唱が嚥下機能を保持・向上させることの根拠を調査する。

1.1 計測方法

歌唱経験と嚥下機能の関係を解明し、歌唱による影響を RSST で定量化する。誤嚥のスクリーニングとして最も簡便な方法である RSST を実施する。言語聴覚士が被験者の喉仏・舌骨に人差指と中指の腹を軽くあてた状態で 30 秒間唾液を飲み込む様子を確認する。1 種類のデバイス (可視光カメラ GoPro) を用いて、被験者の正面と側面から嚥下動作を計測する。

さらに、歌唱者と非歌唱者の臨床データを用いて、遠隔 AI 動画診断の有効性を評価する。ゴクンと飲みこむとき、甲状軟骨 (喉仏) は

前上方に移動する。近年、モーションキャプチャ技術の進歩により、身体にマーカを装着することなく、人間の行動を正確に計測することが可能である。しかしながら、例えば喉頭（皮膚表面）にマーカを装着したとしても、甲状軟骨の動態解析をすることは不可能である。そこで、本研究では動画画像からマーカーレスで位置推定、トラッキングすることができる人工知能技術 DeepLabCut を基本として、喉仏の動きを自動追跡する（図 1）。



（黄丸：喉仏の位置）

図 1 動画自動計測方法

嚥下時は喉仏が前上方に移動することが解剖学的に知られている。各対象者について、教師用データとして 20 フレーム程度のラベリングし、学習する。高齢者になるほど喉の動きが不明瞭で、正確に自動追跡ができないため、最新の AI、画像処理、信号処理等の方法を用いて、全ての動態に対して喉の状態を解析できるロバストな手法を研究開発する。言語聴覚士が実際に実施する RSST では、嚥下動態の様子、回数等を観察している。

本研究では、嚥下回数に加えて、喉頭の空間的な座標の変化、スピードなどの特徴量を取得する。嚥下評価の指標として新たな特徴量を発見し、嚥下機能評価の新しい指標・方法を解明する。

1.2 対象者および解析方法

40 歳以上の健常者を対象に臨床試験参加者を募集する。事前に RSST やスクリーニングテストである簡易嚥下評価ツール EAT-10 等を実施し、嚥下機能に大きな問題がないことを確認済である。対象者に、30 秒間にできるだけ多くの唾液嚥下をしてもらうように指示し、

2 回の実験を実施する。1 回目は言語聴覚士が触診で検査を行い、2 回目は可視光カメラ GoPro（GoPro 社）でその様子を撮影する。画像サイズは、1920×1440 pixels（内 6 データ 1920×1080 pixels）、フレームレート 60 fps（内 2 データ 30 fps）とする。

対象者は、それぞれの歌唱経験の程度により 3 群（非歌唱経験者グループ、一般歌唱経験者グループ、プロ歌唱経験者グループ）に分類する。一般歌唱経験者グループは歌唱レッスン等を受けて日常的に歌唱している人、プロ歌唱経験者グループは声楽家としてプロ活動を行っている人と定義する。非歌唱経験者グループおよび一般歌唱経験者グループの 2 群においては、統計解析ソフトウェア SPSS v29（IBM 社：International Business Machines Corporation）を用いて、傾向スコアマッチングにより各群の対象者の背景因子を揃える。対象者の詳細を表 1 に示す。

表 1 対象者の詳細

歌唱経験 グループ	年齢 [歳]	歌唱経験 年数 [年]	対象者数 [名]	嚥下回数 [回]
非歌唱経験	60.6±14.0	0	27	105
一般歌唱経験	58.7±9.8	9.7±12.9	27	136
プロ歌唱経験	58.2±9.8	42.8±10.3	6	33
合 計	—	—	60	274

本研究は、兵庫県立大学先端医療工学研究所倫理審査（承認番号 22-001）等の承認を得て実施している。

本研究には、i7 64-bit CPU（Windows 10 Home ベース）、GeForce RTX 2060/2080 Ti の GPU を搭載したコンピュータを使用して機械学習を行う。信号処理は、MATLAB R2024a（Signal Processing Toolbox, Statistics and Machine Learning Toolbox 等、MathWorks 社、USA）数値解析ソフトウェアを用いる。

1.3 解析結果

計測した嚥下動画に対して、マーカーレストラッキング技術 DeepLabcut を適用し、嚥下時の動態波形を抽出後、外れ値処理、固定部減算、

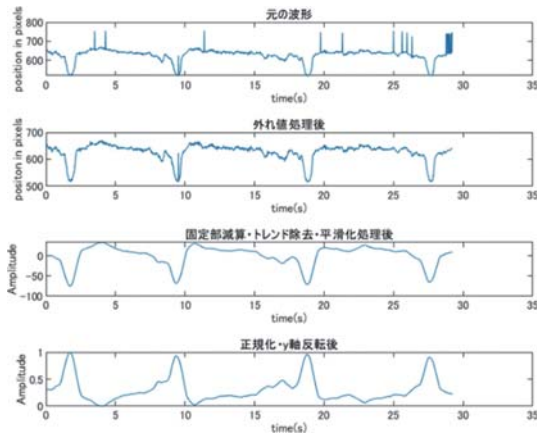


図2 嚥下波形の信号処理結果の一例

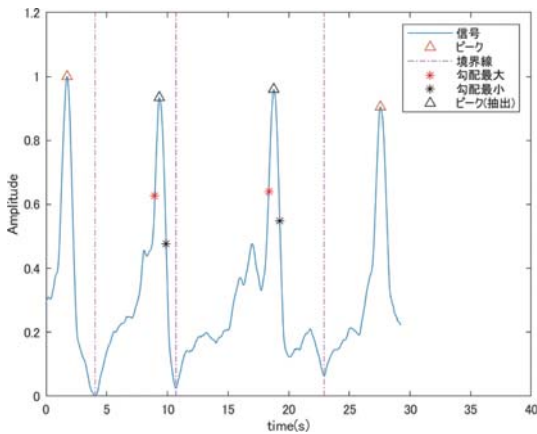


図3 嚥下波形解析の一例

トレンド除去，平滑化，正規化等の信号処理および特徴量抽出を行った。解析結果の一例を図2および3にそれぞれ示す。

3群全体において，RSST との相関を調査した結果，言語聴覚士の目視および動画解析において，それぞれ相関係数 0.752 および 0.749 と高い相関が得られた。この結果より，言語聴覚士の目視検査および動画解析は有用であると考えられる。RSST との相関関係および歌唱経験グループによる3群比較を行なった結果を表2に示す。非歌唱経験者グループに対して，一般

表2 嚥下回数の比較結果

計測方法	非歌唱経験者	一般歌唱経験者	プロ歌唱経験者
RSST [回]	6.1±1.9	7.5±2.0*	6.5±1.9
言語聴覚士の目視 [回]	5.8±1.3	6.7±2.1	7.7±3.1
動画解析 [回]	5.9±1.4	7.0±1.9*	7.5±2.1

* $p < 0.05$ (vs 非歌唱経験者)

歌唱経験者は RSST および動画解析結果に有意な差が認められた。

嚥下特徴量時間は，「① 1 回の嚥下波形の傾きが最大の時点から喉頭が最大位になる時点」，「② 1 回の嚥下波形の傾きが最大の時点から波形の傾きが最小の時点」，「③ 1 回の嚥下波形の傾きが最大の時点から嚥下終了時点」，「④ 1 回の嚥下開始時点から喉頭が最大位になる時点」とそれぞれ定義した。RMS (Root Mean Square：振幅を二乗した値の平均値の平方根) の値が大きいことは，喉仏を高い位置で保っていることを示唆している。波形の形状を表している。また，プロミネンス (他のピークに対して相対的に内部的な高さによってどの程度顕著であるかを示す) において，非歌唱経験者と比較して一般歌唱経験者の方が優位に高値を示し，嚥下がスムーズに行われていることが考えられた。嚥下計測の特徴量の比較結果を表3に示す。

表3 嚥下計測の特徴量の比較結果

計測特徴量項目	非歌唱経験者	一般歌唱経験者	プロ歌唱経験者
嚥下継続時間 [s]	3.64±1.83	3.15±1.84*	3.07±1.59
嚥下特徴量時間① [s]	0.55±0.54	0.42±0.44	0.42±0.22
嚥下特徴量時間② [s]	1.00±0.69	0.76±0.52 [#]	0.80±0.33*
嚥下特徴量時間③ [s]	1.96±1.21	1.84±1.31	1.50±0.92*
嚥下特徴量時間④ [s]	2.22±1.15	1.75±0.96 [#]	1.99±1.16
RMS	0.47±0.11	0.45±0.10	0.56±0.10 [#]
プロミネンス	0.61±0.17	0.67±0.18**	0.57±0.22

* $p < 0.05$, ** $p < 0.01$, [#] $p < 0.005$, ^{##} $p < 0.001$ (vs 非歌唱経験者)

2. 嚥下造影検査 (Video Fluorography: VF) との比較検証

嚥下機能評価においてゴールドスタンダードとされている検査は VF である。VF 検査時に，可視光カメラ GoPro で嚥下の様子を同時撮影し，VF 動画と可視光カメラ動画の解析結果を比較することで，可視光カメラでの解析の有効性を検証する。具体的には，GoPro と VF の同時撮影した動画の喉頭部の動きを DeepLabcut により追跡する。抽出した波形データを同期させて解析結果を確認する。さらに，数値計算ソ

フトウェア Matlab を使用して、嚥下波形のピーク、嚥下区間、喉頭の最大挙上のタイミング等を特定する。1 回の嚥下区間に対して、4 種類の時間特徴量を抽出し、範囲の相関関係を確認する。共同研究機関において、対象者 6 名（男性 3 名、女性 3 名）の臨床試験の結果、特徴量毎に高い相関が認められ、可視光カメラを用いた嚥下評価は有効な方法であることが示唆された。

3. 脳機能計測

圧センサーを用いて圧変動に伴うリズムと脳反応との脳反応 —— 圧コヒーレンス解析 (Cortico-force coherence, CFC) システムを構築した。従来の脳反応 —— 運動コヒーレンス解析 (Cortico-kinematic coherence, CKC) では、加速度計測装置 (ACC) を対象とする身体部位に設置し運動のリズムを評価している。しかし、臨床では、大きな反復運動を行うことが困難な患者も多く、微弱な力でも評価できる手法が求められている。そこで、われわれは、微弱なリズムミカルな圧変化を圧センサーで検出し、脳信号と圧とのコヒーレンス解析を行った。

3.1 対象者および解析方法

健康成人 14 名（女性 7 名、年齢 22~40 歳）が参加した。被験者は、右人差し指または右つま先を用いて、能動的な条件（随意運動）と受動的な条件（検者による受動運動）の 2 つの異なる条件下で、リズムミカルな押圧課題を行うよう指示された。押圧は圧センサーで計測し、運動のダイナミクスを捉えるために、表面筋電図と加速度計による記録も追加された。脳波 (EEG) は国際 10-20 システムに準拠した 32 チャンネルキャップで記録された。コヒーレンス解析は脳波信号と対応する力、筋電図、または運動信号 (ACC) との間で施行した。

3.2 解析結果

すべての被験者および条件において、コヒーレンススペクトルは、被験者の運動リズムと一致する周波数において、有意なピークを示した。

これらのピークは、EEG-圧力 (CFC) だけでなく、EEG-ACC (CKC) および EEG-EMG でも観察された。また、能動的な押圧条件と受動的な押圧条件において CFC ピークの振幅に統計的に有意な差が認められなかった。さらに、コヒーレンスの強度は指とつま先を押す課題間で差がなかったことから、CFC は身体部位を問わず臨床応用が可能なことが示唆された。親指と足の運動と脳反応とのコヒーレンスは左半球（対側半球）に認められた。本 CFC 手法は、大きな運動や強い力を必要とせず、四肢や口腔領域の感覚運動障害の患者に広く適応することが可能という利点がある。CFC は、従来の加速度計を用いた CKC 計測とは異なり大きな運動を行う必要がなく、微弱な圧での計測が可能のため、計測部位（脳）と距離の近い口腔領域の脳機能解析には特に有利である。今後は、CFC 計測を口腔領域に適応し、歌唱者のコヒーレンス値を解析することで、歌唱の口腔運動制御における影響を明らかにしたい。

4. まとめ

本研究では、深層学習を用いた動画解析による、簡易かつ定量的な嚥下の動態評価システムを構築した。ゴールドスタンダードとされる VF との同時計測および解析を行うことで、評価結果の同等性を示すことができた。さらに、脳機能計測において、脳信号と圧センサーで検出した圧変化とのコヒーレンス解析により、特定の周波数で有意な特徴量を捉えることができ、口腔領域の脳機能解析の有効性が示唆された。今後も嚥下リハビリテーション促進に貢献することを目指す。

〔謝 辞〕

本研究を実施するにあたり、多大なる御支援を頂きました立石科学技術振興財団に厚く御礼申し上げます。

[成果の発表, 論文など]

- [1] 八木直美, “人工知能による嚥下ヘルスケアシステムの開発 (生体信号計測と解析),” 日本バイオリオロジー学会誌 (B&R), 37(2), 9-14, 2024. 4. 雑誌
- [2] N. Yagi, S. Nagami, H. Maezawa, Y. Hata, “Medical trends and developments in system of systems engineering for dysphagia,” Studies in Systems, Decision and Control, Springer Nature (in press) 書籍
- [3] N. Yagi, “Medical support system with bioinformation,” and “Healthcare monitoring by AI and signal processing,” Taichung Veterans General Hospital, 2023. 9. 16. 招待講演
- [4] 八木直美, “人工知能で嚥下を科学する,” 第44回バイオリオロジー・リサーチフォーラム, 2023. 6. 3. 招待講演
- [5] 八木直美, “人工知能によるリハビリテーション評価システムの開発,” はりま産学交流会 6 月創造例会, 2023. 6. 16. 招待講演
- [6] A. Yoshida, N. Yagi, Y. Fujii, H. Shibutani, Y. Kobayashi, Y. Saji, Y. Sakai, Y. Hata, “A Visible Camera Approach to Motion Tracking based Swallowing Evaluation,” ICMLC & ICWAPR 2024, 2024.
- [7] T. Ueyama, N. Yagi, Y. Fujii, H. Shibutani, Y. Kobayashi, Y. Saji, Y. Sakai, Y. Hata, “Aspiration Detection System with Video Fluoroscopic Examination of Swallowing Testing,” ICMLC & ICWAPR 2023, pp. 376-380, 2023.
- [8] 八木直美, 中村朱里, 川村直子, 前澤仁志, 酒井良忠, 柏岡秀紀, 平田雅之, 柳田敏雄, 畑豊, “DeepLabcut を用いた嚥下動態解析方法の検討,” 人工知能学会第二種研究会, 2024.
- [9] H. Maezawa, M. Wakida, M. Matsuhashi, “Cortico Force Coherence of the Finger and Toe with Slight Rhythmic Pressure on Force Sensors using Electroencephalography,” Scientific Reports, pp. 1-12, 2025.