

令和 3 年 6 月 14 日現在

機関番号：13701

研究種目：基盤研究(C) (一般)

研究期間：2018～2020

課題番号：18K11371

研究課題名(和文)音響処理と画像処理の協調的統合による嚥下タイミング計測の研究

研究課題名(英文)Swallowing timing estimation by coordinated integration of acoustic signal and image processing

研究代表者

速水 悟 (Hayamizu, Satoru)

岐阜大学・工学部・教授

研究者番号：90345794

交付決定額(研究期間全体)：(直接経費) 3,400,000円

研究成果の概要(和文)：本研究は、嚥下機能を非侵襲かつ定量的に計測するために、嚥下時の音と甲状軟骨の動きのデータをモデル化した。音源食品を飲み込ませ、その音を録音したデータから、音響信号処理及びニューラルネットワークによって音源食品が食道に送られるタイミングを推定した。また、3-IR照度差ステレオ法によって、喉周辺の動画像から甲状軟骨の動きを得ることで、これと同期して動く喉頭蓋の動きを推定した。最後に、これら二つを比較することにより、嚥下機能を計測した。嚥下機能が正常である被験者に対する実験を行い、音データと動画データから得られた二つの時刻を比較することにより、嚥下タイミングの推定とその評価を実施した。

研究成果の学術的意義や社会的意義

嚥下障害者の誤嚥防止のためには、嚥下能力の把握は最も重要であり、簡便な計測装置は病状的確な進行具合の把握、治療に役立つ。現状ではX線による嚥下造影検査や内視鏡検査しか検査方法がなく、侵襲性が非常に高い。一方、非侵襲に嚥下を計測する研究も広く行われているが、嚥下物がいつ飲み込まれたかを簡便に計測する方法はなかった。本研究では、音を発する音源食品を嚥下することにより、この問題の解決を目指す。具体的には、耳に装着した骨伝導マイクから得られた音源食品の音を音響処理により解析することにより、嚥下物が嚥下され食道に入った時刻を検出する方法の研究を行った。

研究成果の概要(英文)：In this study, we modeled the data of swallowing sound and thyroid cartilage movement in order to measure swallowing function non-invasively and quantitatively. From the recorded data of the sound source food swallowed, the timing at which the sound source food was sent to the esophagus was estimated by acoustic signal processing and a neural network. In addition, the movement of the epiglottis was estimated by obtaining the movement of the thyroid cartilage from the video image around the throat by the 3-IR illuminance difference stereo method. Finally, the swallowing function was measured by comparing these two. An experiment was conducted on subjects with normal swallowing function, and a method for estimating swallowing timing and its evaluation were performed by comparing the two times obtained from sound data and video image data.

研究分野：メディア情報学

キーワード：嚥下機能計測 音響信号処理 照度差ステレオ法 ニューラルネットワーク

## 1. 研究開始当初の背景

飲食物を飲み込み、食道、胃へと送る機能(嚥下機能)は、加齢に伴い低下する。これにより、飲食物を正常に嚥下することが困難になる嚥下障害を引き起こす可能性が高まる。そのため、近年の高齢化に伴い、嚥下障害の患者は増加してきている[1]。嚥下障害は、窒息や肺炎の原因となり、特に嚥下障害が原因で起こる誤嚥性肺炎は、高齢者の肺炎の7割以上を占め、深刻な社会問題となっている[2]。そのため、嚥下障害の早期発見や予防において、嚥下機能を計測することが非常に重要である。

誤嚥を起こすのは、咽頭蓋の閉鎖により気管が閉じたタイミングと、口腔内にある食塊が食道に送られるタイミングがずれるために起こる。各種方面から嚥下評価の研究が行われているが、接触式である、あるいは患者を拘束する必要があるので課題が多い。また、外部からの計測では、口腔内にある食塊の位置の同定が難しいため、咽頭蓋が閉鎖するタイミングと、食塊が飲み込まれるタイミングを計測できる方法は提案されていなかった。

## 2. 研究の目的

本研究では、骨伝導マイクを耳に装着し、カメラ装置の前で音源食品を嚥下することで非接触、非侵襲で、かつ簡単にこのタイミングを計測し、嚥下機能計測を行う方法を提案した。

嚥下機能が健常である人は、飲食物を咽頭から食道へ送る際に、喉頭蓋という器官が反射的に気道を塞ぐように倒れる(図1)。これにより、飲食物の気道への流入を防いでいる。このとき、食道に飲食物が送られるタイミングと、喉頭蓋が気道を塞ぐタイミングは一致している。しかし、嚥下障害をもつ人は、老化などによる喉周辺の筋肉の衰えにより、喉頭蓋の動きが鈍化している場合が多い。それにより、これらのタイミングが一致しなくなり、誤嚥を引き起こすリスクが高まる[3]。そのため、食道に飲食物が送られるタイミングと、喉頭蓋が気道を塞ぐタイミングの二つを検出し、それらが一致しているかどうかを調べることで、嚥下障害を患っているか否かを判断することができる。

本研究では、音源食品を嚥下する際の音から、食道に音源食品が送られるタイミングを推定し、そのときの喉頭周辺の動画像から、甲状軟骨の動きを追跡することで喉頭蓋が気道を塞ぐタイミングを得る。最終的に両者を比較することによって、嚥下機能障害であるかを判断する。

図2にシステムの概要を示す。撮影システムの外観を図3、音の取得方法を図4、録音システムの外観を図5に示す。

本研究では、図3の装置の前に座り、音源食品を嚥下したときの音を耳に装着した骨伝導マイクで録音し、同時に喉周辺の動画像を図3の装置で撮影する。耳装着型の骨伝導マイクは、装着が容易で、嚥下時の喉の動きへの影響がないため用いた。被験者は装置の前で音源食品を嚥下するだけでよく、非侵襲な嚥下機能計測を行うことが可能である。また、音とフラッシュが同時に出るアプリを作成し、これをインストールしたスマートフォンを図5のようにカメラに写るように設置し、音とフラッシュが発生した時刻を開始時刻とすることで、音と動画像の同期をとる。

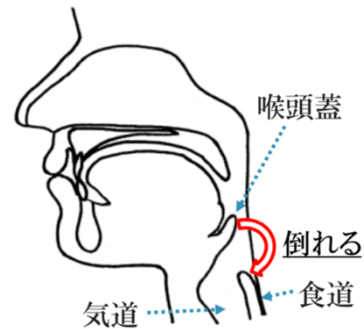


図1 嚥下の仕組み

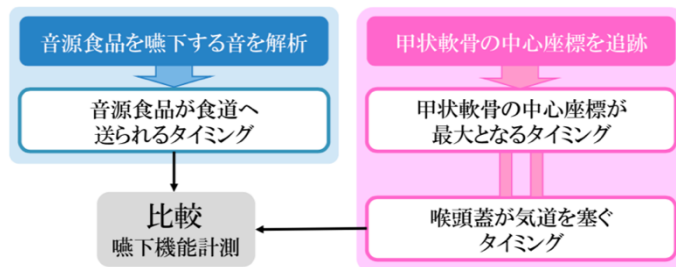


図2 提案システムの概要

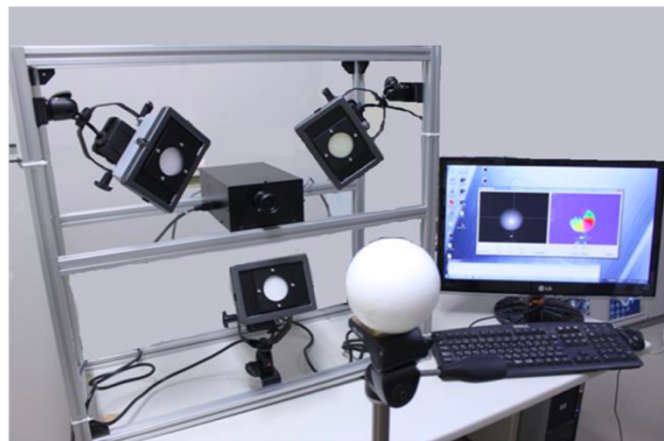


図3 撮影システムの外観

### 3. 研究の方法

音源食品を嚥下する際の音を、耳に装着した骨伝導マイクで録音し、その音から、音源食品が食道へ送られたタイミングを求める。ここで、音源食品から発生する音を“食塊音”と呼ぶこととする。音源には、高压ガス封入キャンデー(明治産業製のはじけるキャンデー)に水を加えることにより発生するランダムな発泡音を利用し、このキャンデーを水とともに封入した医薬品用のカプセルを音源食品として用いる。

音源食品が食道へ送られると、骨伝導マイクから遠ざかり、食塊音を収集できなくなることを利用し、食道に音源食品が送られるタイミングを推定する。つまり本研究では、嚥下音を分析し評価を行う従来の方法とは異なり、音源食品から発生する食塊音を抽出し分析を行うことで嚥下タイミングの推定を行う。

しかし、収集される音には、食塊音だけでなく、嚥下音も含まれる。食道に音源食品が送られるタイミングを得るには、混ざりあった音の中から、食塊音を含む時刻を選ぶ必要がある。そのため、食塊音と嚥下音をそれぞれ別々に録音したデータと、二つを合成した音(合成音)を、ニューラルネットワークで学習させる。本研究では、食塊音の発生しているフレームの検出を行モデルとして、深層学習の中で最も基本的な全結合ニューラルネットワーク(Feed Forward Neural Network: FFNN)を用いた。そして、有音区間に対して、各時刻(フレーム)が食塊音を含んでいるかどうか判定する。この流れを図6に示す。

### 4. 研究成果

本システムを用いて、嚥下運動の計測精度を調べた。本システムでは、甲状軟骨の追跡において、甲状軟骨がはっきりと確認できる被験者であれば、動きの追跡が可能である。本研究では、20代男性5名に常温の水を飲ませ、その様子を撮影した。得られた動画像の開始フレームでの甲状軟骨の位置を0として、テンプレートマッチングによって推定された甲状軟骨の位置の推移を嚥下運動の計測結果とした。目視で1フレームずつ甲状軟骨の位置を追ったものを理想値とし、計測結果と理想値を比較し、その差を計測精度として算出した。

学習したモデルを用いて、音源食品を嚥下する音に対して、各フレームが食塊音を含んでいるかどうか判断する。また、学習データと同様、実験に用いるデータも、音の発生しているフレームを抽出し、各フレームとその前後1フレームの正規化したメル周波数スペクトルを求めたものを特徴量として、入力に用いた。同時に撮影した喉頭の動きのデータと比較した結果を図7に示す。図の1段目は、各フレームの振幅スペクトルの和を時間軸上に並べたものと、食塊音が含まれていると判定されたフレームに縦線を引いたものである。2段目では、追跡した喉頭の動

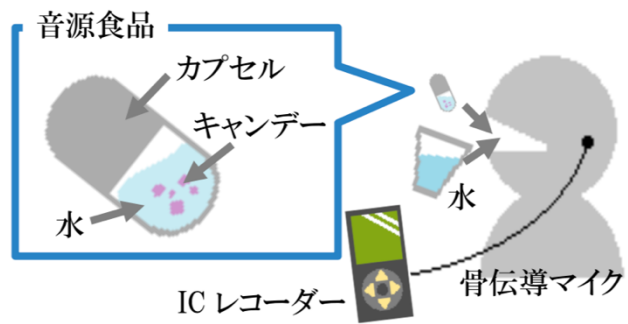


図4 音の取得方法



図5 録音システムの外観

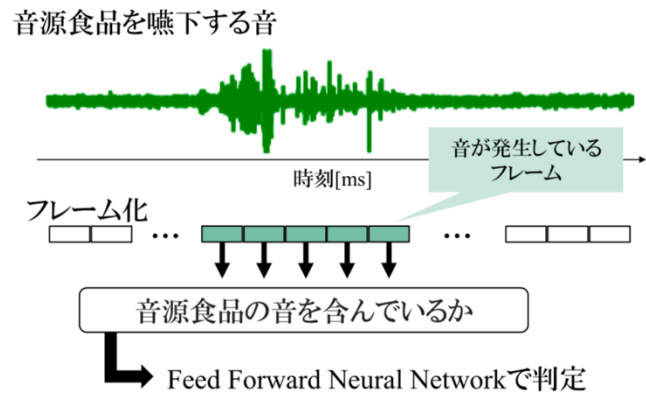


図6 食塊音を含むフレームの選択方法

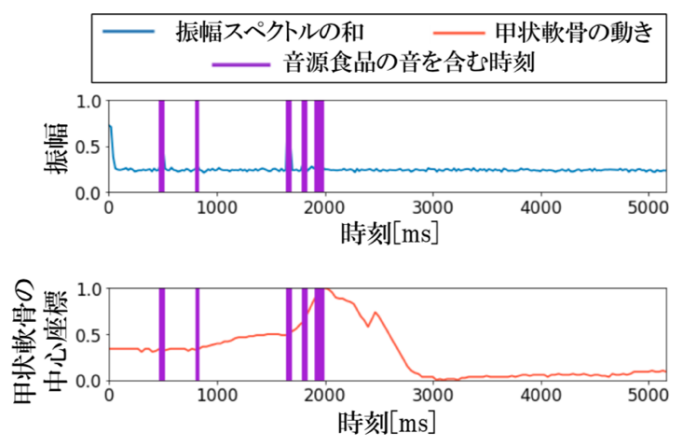


図7 甲状軟骨の中心座標と音源食品の音の推定結果

き(甲状軟骨の中心座標)と食塊音が含まれるフレームに縦線を引いたものである。喉頭の動きは、人により差があるため、データごとに0~1の範囲に正規化してからグラフ化している。喉頭の動きは、値が大きくなるほど喉頭蓋が気道を塞いでいく。そのため、値が最大となる地点で気道を完全に閉鎖するタイミングとなる。つまり、その地点で食塊音が観測されなくなれば、正常な嚥下を行えていることになる。図7を見ると、喉頭蓋が気道を閉鎖するタイミングで食塊音が観測されなくなっていることから、この被験者は正常な嚥下が行えていることがわかる。

嚥下タイミングの推定の精度向上に向けた検討を行った。音源食品が食道に送られるタイミングを推定する上で、音源食品から発生する音の頻度が少ない場合や、音にばらつきがある場合の精度向上とノイズ除去が課題となっていた。これらの課題について、音声・音響を対象とした関連分野の研究動向を調査し、時間領域の特徴をそのまま利用することを検討した。具体的には、時間波形に畳み込みニューラルネットワークを適用して、時間領域の特徴を抽出するためのモデル構造と学習方法などについて、基礎的な検討を行った。

一般に深層学習のモデルは、個人差やノイズ環境などの様々な条件下で収集した大量のデータの存在が前提となっている。時間領域の特徴をそのまま利用する方法は、パラメータ数は少ないもののかなりのデータ量が必要である。このため、特徴量の見直しとモデル適応の手法を調査した。これまでの研究においては、あらかじめクラス分類モデルを学習させ、フレーム単位で音の種類を識別することで、嚥下タイミングの推定を行ってきた。これをさらに発展させる方法として、異常検知モデルの適用と、時間方向と周波数方向で特徴的な領域を抽出する方法を検討した。具体的には複数フレームを入力として、注意機構付きのモデルを学習させる。再構成誤差を可視化することで、周波数方向の特徴的な領域が抽出できる。また注意機構によって得られた重み係数を可視化することで、フレーム方向の重要度を推定することができる。機械装置の異常音データを用いて動作を確認した。

#### <参考文献>

- [1] 小口和代, 才藤英一, “機能的嚥下障害スクリーニングテスト「反復唾液嚥下テスト」(RSST)の検討(1)正常値の検討,” リハビリテーション医学, no. 37, pp. 375-382, 2000.
- [2] 厚生労働省, “5 疾病・5 事業について,” 第二回医療計画の見直し等に関する検討会, 資料 2, 2016.
- [3] 小口和代, 才藤英一, “機能的嚥下障害スクリーニングテスト「反復唾液嚥下テスト」(RSST)の検討(2)妥当性の検討,” リハビリテーション医学, no. 37, pp. 383-388, 2000.
- [4] 中原智喜, 古閑公治, 永谷正巳, 伊賀崎伴彦, 村山伸樹, “マイクロフォン波形を用いた嚥下運動機能の解析(1),” 電気関係学会九州支部連合大会, 06-2P-14, 2004.
- [5] 八木直美, 永見慎輔, 上野博司, 家邊 徹, 尾家慶彦, 越久仁敬, “音声信号解析を用いた嚥下抽出システムの開発,” 第 57 回自動制御連合講演会, 3A05-2, 2014.
- [6] 株式会社 ハッピーリス, <https://www.happyris.jp/swallowingsound-check/>
- [7] 林 豊彦, “飲み込みを科学する,” バイオメカニズム学会誌, vol. 40, no. 1, pp. 13-19, 2016.
- [8] 山本 伸, 青木義満, “喉表面の 2 領域解析による嚥下機能計測システムの開発,” 社団法人映像情報メディア学会技術報告, vol. 33, no. 11, pp. 1-4, 2009.
- [9] 藤野将弘, 加藤邦人, 村 絵美, 永井 元, “3-IR 照度差ステレオ法を用いた嚥下運動計測システムの提案,” 第 20 回画像センシングシンポジウム, pp. IS3-07-1-IS3-07-8, 2014.
- [10] R. J. Woodham, “Photometric method for determining surface orientation from multiple images,” Opt. Eng., vol. 19, no. 1, pp. 139-144, 1980.
- [11] 池内克史, “反射率地図に基づき, 二次元濃淡画像より三次元形状を再構成する 2 手法,” 信学論(D), vol. J65-D, no. 7, pp. 842-849, July 1982.
- [12] 金子俊一, 富田 安, 本多庸悟, “3 色光の同時照射による照度差ステレオ法,” 信学論(D-II), vol. J76-D-II, no. 10, pp. 2243-2246, Oct. 1993.
- [13] M. Fujino, K. Kato, E. Mura, and H. Nagai, “Real time swallowing measurement system by using photometric stereo,” Proc. International Conference on Quality Control by Artificial Vision (QCAV 2015), pp. 95341 A-1-8, 2015.
- [14] 児玉千紗, 加藤邦人, 田村哲嗣, 速水 悟, “音響信号処理による嚥下タイミング推定手法,” 計測自動制御学会ライフエンジニアリング部門シンポジウム 2017, 3B1-1, 2017.
- [15] C. Kodama, K. Kato, S. Tamura, and S. Hayamizu, “Swallowing function evaluation using deep-learning-based acoustic signal processing,” Proc. APSIPA ASC 2017, TA-P3. 15, 2017.
- [16] 児玉千紗, 加藤邦人, 田村哲嗣, 速水悟, “音響信号処理と 3-IR 照度差ステレオ法による嚥下機能計測”, 電子情報通信学会論文誌 D, Vol. J102-D, No. 3, pp. 173-184, 2019

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 1件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 児玉千紗、加藤邦人、田村哲嗣、速水悟	4. 巻 J102D-3
2. 論文標題 音響信号処理と3-IR 照度差ステレオ法による嚙下機能計測	5. 発行年 2019年
3. 雑誌名 電子情報通信学会論文誌	6. 最初と最後の頁 pp.173-184
掲載論文のDOI（デジタルオブジェクト識別子） 10.14923/transinfj.2018PDP0001	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計0件

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

	氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考
研究分担者	加藤 邦人 (Kato Kunihito) (70283281)	岐阜大学・工学部・准教授  (13701)	
研究分担者	田村 哲嗣 (Tamura Satoshi) (10402215)	岐阜大学・工学部・准教授  (13701)	
研究分担者	木村 暁夫 (Kimura Akio) (00362161)	岐阜大学・大学院医学系研究科・准教授  (13701)	
研究分担者	西津 貴久 (Nishizu Takahisa) (40228193)	岐阜大学・応用生物科学部・教授  (13701)	

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8 . 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関
---------	---------