

令和 2 年 7 月 5 日現在

機関番号：24201

研究種目：基盤研究(C) (一般)

研究期間：2017～2019

課題番号：17K01571

研究課題名(和文) 嚥下音と深度画像を組み合わせた嚥下機能評価システムの構築

研究課題名(英文) Constructing a system for evaluating swallowing function by combination of swallowing sounds with neck depth images

研究代表者

宮城 茂幸 (Miyagi, Shigeyuki)

滋賀県立大学・工学部・准教授

研究者番号：20273469

交付決定額(研究期間全体)：(直接経費) 3,500,000円

研究成果の概要(和文)：食塊の飲み込みに伴う嚥下音を、臨床医による嚥下状態の評価結果を教師データとして機械学習の一種であるSVMを用いて学習し、判別器を作成することができた。実際に判別をさせ、スクリーニング用としては78%の精度が得られることが確認できた。また、嚥下状態の程度に応じた判別問題としては46%の精度となることが確認できた。

深度画像を用いた非接触センシングによる嚥下時間推定では、頸部深度画像から嚥下時間推定を行うためのアルゴリズムを開発できた。本アルゴリズムを健康者の頸部深度画像に適用したところ、男女問わず嚥下時間推定でき、精度が90%であることが確認できた。

研究成果の学術的意義や社会的意義

言語聴覚士による嚥下評価では触診や聴診を行うが、今回開発した深度画像による嚥下時間推定手法を用いると、嚥下の行われたタイミングを触診無しで行い得ることができる。嚥下状態の遠隔診断を確立するために必要な一つの要素技術として、嚥下のタイミングを検知することは重要であるが、そのための一つの手法となり得る。一方、嚥下音の解析は精度の点ではまだ十分ではないが、スクリーニングだけでなく、嚥下の状態に応じたクラス識別についての可能性は少なくとも確認できた。このことから、遠隔診断だけでなく、在宅療養やケアハウスといった看護や介護の現場における簡易型嚥下評価システム構築の可能性を明らかにすることができた。

研究成果の概要(英文)：A SVM based classifier for evaluating the level of swallowing disorder could be constructed by using the swallowing sounds annotated of the VE score by clinicians as training data that was collected from healthy subjects and patients. The performance achieved to 78% accuracy in the two class problems for screening, and 46% accuracy in the four class problems in which swallowing status was categorized into four levels.

A noble algorithm could be successfully constructed as a non-contact technique for estimating swallowing time periods by using depth images of neck part. Applying the developed algorithm to the neck part depth image from both male and female healthy subjects yielded the 90% accuracy of the estimated swallowing periods.

研究分野：信号処理

キーワード：嚥下評価 嚥下音 深度画像 機械学習 SVM 頸部

1 研究開発当初の背景

厚生労働省の人口動態統計によると日本人の死亡原因の第3位は肺炎である。その中で7割が高齢者の誤嚥性肺炎であるといわれている。高齢化社会の進展にともない、介護サービス付き住宅や訪問看護といった様々な形態での介護が行われることが予想される。そのような介護現場において、高齢者の嚥下機能が正常であるかどうか、あるいはどの程度嚥下機能が低下しているかといった評価が、特殊な機器を使用せずに行えるようになれば誤嚥性肺炎のリスクを減少させられるであろう。医療機関において嚥下機能を評価するために嚥下造影検査(VF)や嚥下内視鏡検査(VE)が行われている。VFはX線による侵襲をともなうが嚥下時の咽頭の動きを正確に把握することができるので、術後の経過を診断するためによく用いられる。またVEは侵襲をともなわないので、診察室での嚥下機能障がいスクリーニングによく利用される。一方言語聴覚士は頸部触診とともに頸部聴診法を用いて嚥下が行われているかどうかの判断を行う。VFやVEに比べ特殊な機器を必要としないが、熟練を要する。

これまで頸部聴診法の考え方を元に、嚥下音や嚥下時の咽頭の動きを加速度センサやマイクロフォンでとらえ嚥下障がいのスクリーニングを行う方法が提案されている。これらの研究において、嚥下音の産出部位、時間、あるいは男女差といった嚥下音に関する基本的な性質は徐々に明らかになってきている。そのため嚥下音によるスクリーニングは可能になりつつあるが、嚥下音と嚥下障がいの程度との関連性や、障がいの形態との関連性などはわかっていない。そこで実際の嚥下障がいのスコアと嚥下音の違いを詳細に検討することにより、スクリーニングからさらに進んだ評価が可能になると考えられる。

2 研究の目的

死亡原因の一つとして高齢者の誤嚥性肺炎があげられる。高齢化社会の進展にともない、介護現場では誤嚥につながる嚥下障害を早期に発見するためのシステムが必要とされている。これまで嚥下音の解析と、深度画像による嚥下時の喉頭挙動の計測に取り組んできた。その結果嚥下音の特徴と医師の付与した嚥下評価スコアとの関連性を示すことができた。また深度画像による喉頭の前後運動がX線により計測された挙動と対応し、相対的な移動量が一致するという知見を得た。そこでこれらの結果を用い嚥下音と深度画像を組み合わせることで、ベッドサイドやチェアサイドで利用可能な簡易型嚥下評価システムの開発を行う。

3 研究の方法

(1) 嚥下音の特徴探索

嚥下音の収集には咽頭マイクを用いるが、マイクの周波数特性や雑音により信号の性質に変化をおよぼす可能性がある。スペクトル減算法による雑音除去を行うとともに、マイクの周波数特性を事前に計測し、周波数特性が一定となる補正を行う。その後嚥下音にフーリエ変換あるいは短時間フーリエ変換を適用し、周波数領域および時間周波数領域における得られる中心周波数、周波数分散、周波数分布といった特徴量を求める。これらの特徴量と嚥下機能評価スコアとの相関分析を行い、相関の高い特徴量を決定する。特徴量を複数組み合わせ、機械学習を行うい、臨床医の診断により嚥下機能評価スコアを用いて群分けされた複数のグループを教師データとする。学習結果を用いて特徴量を用いたグループわけがどの程度正確に行えるかを検証する。

(2) 頸部深度画像の変位解析

嚥下前後の頸部深度画像のフレーム間差分を計算し、頸部の表面変動を求める。表面変動をしきい値処理し、変動の大きさから嚥下時間を推定するアルゴリズムを開発する。頸部深度画像から推定された嚥下時間とVE画像より得られた嚥下時間とを比較し、推定時間の正確さを検証する。

4 研究成果

(1) SVM による嚙下音分類

男性 17 名、女性 10 名 (年齢範囲:2147, 平均: 22.4) からなる健常な被験者および、男性 78 名、女性 65 名からなる入院患者 (年齢範囲:25102, 平均: 83.3) より嚙下音を採集した。南豆社製咽頭マイク SH-12iK および PC 上のソフトウェア Audacity を用い、サンプリング周波数 8000 Hz, Audacity 上の利得を 0.7 に固定し録音した。嚙下音のピーク音量を含む 800 ms の区間を切り出し、雑音除去および感度補正を施した嚙下音を以下の実験で使用した。これと同時に臨床医により兵頭スコア [1] による被験者の嚙下評価が行われ、それらを教師データとして利用した。SVM の学習時には、各クラスごとのデータ数が同一になるようにランダムにデータを抽出した。

特徴量として、エネルギー、スペクトルについての最大値、平均、標準偏差、中央値、四分位数の比、ピーク周波数、スペクトログラムについての最大値、ピーク位置、ピーク位置の相対比、総エネルギー、帯域ごとの四分位数の比、2 値化スペクトログラムについての面積比、平均位置、平均ピーク位置等を用いた。総特徴量数は 83 となるり、特徴量の組み合わせ ${}_{83}C_n$ が膨大となることから、あらかじめ予備実験により教師データと相関の強い特徴量 9 個に制限し、さらに SVM の学習が飽和する特徴数として 2 クラス問題の場合 4 に、4 クラス問題の場合 6 に制限し判別に有効な特徴量の組み合わせを調べた。それらの結果を表 1 および表 2 に示す。なお、2 クラス問題では健常者 (兵頭スコアが 0) とそれ以外の被験者の嚙下音に分類した。また 4 クラス問題では兵頭スコアにしたがって障害の程度を A、B、C、D の 4 つのクラスに分類した。これらの結果から、スクリーニングは精度 78% 行い得るが、単一の SVM だけでは障害の程度に応じた分類精度が 46% 程度と低く、嚙下障害の程度を推定することには不十分であることがわかった。

表 1: 嚙下音の判別を 2 クラス問題として考えたときの、各特徴量による分類の精度、適合率、再現率および F 値。各データは適合率が高い順に並べている。

Comb. #	Used Features				Accuracy	Precision	Recall	F-Measure
8	\bar{D}_v	\bar{D}	A_w	Q_2	0.780	0.787	0.790	0.781
18	\bar{D}_v	\bar{D}	Q_2	$Q_{8,2}$	0.780	0.804	0.750	0.771
54	\bar{D}_v	Q_2	\bar{D}_h	$Q_{8,2}$	0.780	0.831	0.710	0.760
48	\bar{D}_v	A_{\max}	Q_2	Q_{r1}	0.775	0.811	0.720	0.762
113	A_w	A_{\max}	Q_2	Q_{r1}	0.770	0.763	0.790	0.773
34	\bar{D}_v	\bar{f}	\bar{D}_h	Q_{r1}	0.770	0.752	0.820	0.781
46	\bar{D}_v	A_w	Q_{r1}	$Q_{8,2}$	0.770	0.737	0.870	0.789
53	\bar{D}_v	Q_2	\bar{D}_h	Q_{r1}	0.770	0.750	0.810	0.779
19	\bar{D}_v	\bar{D}	\bar{D}_h	Q_{r1}	0.755	0.717	0.870	0.780
17	\bar{D}_v	\bar{D}	Q_2	Q_{r1}	0.750	0.808	0.670	0.722

(2) 深度画像による嚙下時間推定

嚙下時には喉頭の挙上運動や頸部の筋肉の運動が生じる。それらの運動は頸部の形状変化としてあらわれる。図 1 に疑似カラーで色付けした深度画像を 2 枚示す。図 1(a) は嚙下前、図 1(b) は嚙下中の頸部の様子を示している。これらの 2 枚の画像を比較すると、矩形内のカラーパターンが変化していることがわかる。つまり、嚙下時には深度に変化が生じる。深度が大きく変化していれば、嚙下をしている時間であると推定できる。この原理から以下に示す 4 つの処理手順で嚙下時間の推定を行うことができる。

1. 深度画像から点群の取得する。
2. 点群系列間の深度差を頸部変化量として計算する。
3. 頸部変化量の統計量からしきい値処理により嚙下時刻を検出する。
4. 推定した嚙下時刻からマージ処理をして、嚙下時間を推定する。

表 2: 嚙下音の判別を 4 クラス問題として考えたときの、各特徴量の組み合わせと、それらを用いたときの全体の適合率および各クラスに対する F 値。各データは適合率が高い順に並べている。

Comb. #	Used Features					
42	\bar{D}_v	\bar{f}	A_w	Q_2	\bar{D}_h	Q_{r1}
3	\bar{D}_v	\bar{D}	\bar{f}	A_w	A_{\max}	Q_{r1}
29	\bar{D}_v	\bar{D}	A_w	Q_2	Q_{r1}	$Q_{8,2}$
17	\bar{D}_v	\bar{D}	\bar{f}	Q_2	\bar{D}_h	Q_{r1}
6	\bar{D}_v	\bar{D}	\bar{f}	A_w	Q_2	Q_{r1}
46	\bar{D}_v	\bar{f}	A_{\max}	Q_2	\bar{D}_h	Q_{r1}
9	\bar{D}_v	\bar{D}	\bar{f}	A_w	\bar{D}_h	$Q_{8,2}$
45	\bar{D}_v	\bar{f}	A_w	\bar{D}_h	Q_{r1}	$Q_{8,2}$
11	\bar{D}_v	\bar{D}	\bar{f}	A_{\max}	Q_2	\bar{D}_h
75	\bar{D}	A_w	A_{\max}	\bar{D}_h	Q_{r1}	$Q_{8,2}$

Comb. #	Accuracy	F-Measure (A)	F-Measure (B)	F-Measure(C)	F-Measure(D)
42	0.460	0.602	0.468	0.511	0.145
3	0.420	0.495	0.427	0.500	0.133
29	0.420	0.579	0.407	0.313	0.338
17	0.415	0.521	0.457	0.456	0.095
6	0.410	0.598	0.481	0.279	0.174
46	0.410	0.625	0.360	0.333	0.208
9	0.405	0.571	0.492	0.259	0.141
45	0.400	0.562	0.256	0.400	0.256
11	0.390	0.626	0.288	0.396	0.143
75	0.385	0.385	0.431	0.355	0.356

時刻 t_n における深度画像中の参照領域内の深度値を $z_k(t_n)$ $k = 1, 2, \dots, M$ と書くとする、隣接深度画像間の差分は $e_k(t_n) = z_k(t_{n-1}) - z_k(t_n)$ である。なお、深度は X, Y 座標と合わせて点群として得られるが、時刻 t_{n-1} と t_n における点群の位置 $(x_k(t_{n-1}), y_k(t_{n-1}))$ および $(x_k(t_n), y_k(t_n))$ は必ずしも一致しないので、実際には $e_k(t_n)$ を求めるときに深度値の補間が必要である。参照領域内における点ごとの差分 $e_k(t_n)$ が求まれば、それらを用い、平均、平均の絶対値の対数、平均電力、分散といった統計量の時間変動を計算し、しきい値処理から嚙下時刻の推定を行い、それらの結果を統合することにより嚙下時間推定を行う。統合処理では、推定時刻間隔が小さい場合の重畳処理や、統計量の時間変動における大小の位置関係による推定区間調整が行われるその処理結果の例を図 2 に示す。

RealSense F200 (Intel Corporation) を用い喉頭隆起が不明瞭な女性 3 名を含む 20 代の被験者 7 名の嚙下時の画像を撮影し、それらに対し提案手法を適用した結果、表 3 に示す嚙下時間推定の精度を得ることができた。このことから、非接触で嚙下が発生している時間を推定できることがわかった。

表 3: 性別と仰角ごとに評価した嚙下時間推定の精度

データ分類	総嚙下回数 (回)	再現率 R_{period}	適合率 P_{period}	F 値 F_{period}
男性	117	0.957	0.903	0.929
女性	78	0.821	0.853	0.837
仰角 5 度	100	0.910	0.892	0.901
仰角 20 度	95	0.895	0.876	0.885
総計	195	0.903	0.884	0.893

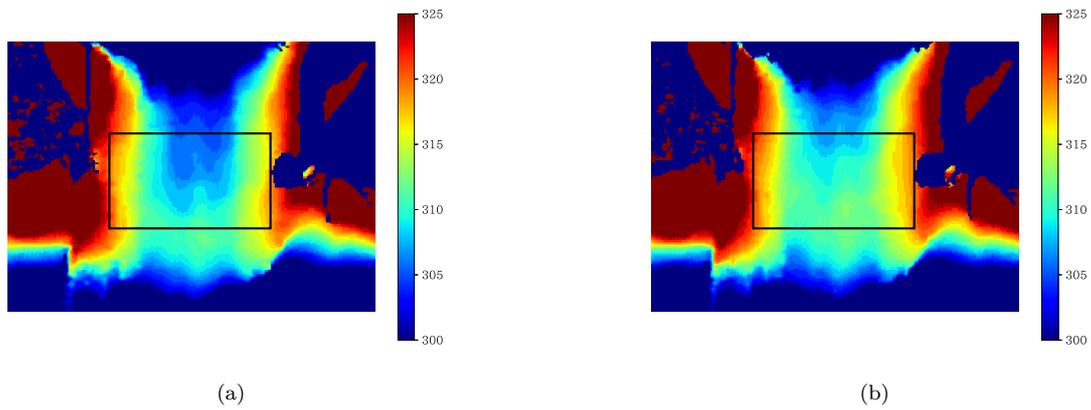


図 1: 深度を疑似カラーで色付けした頸部の深度画像。(a) の深度画像は嚥下前を撮影した。(b) の深度画像は嚥下中を撮影した。それぞれの画像内の矩形はアルゴリズムでの基準範囲を示している。

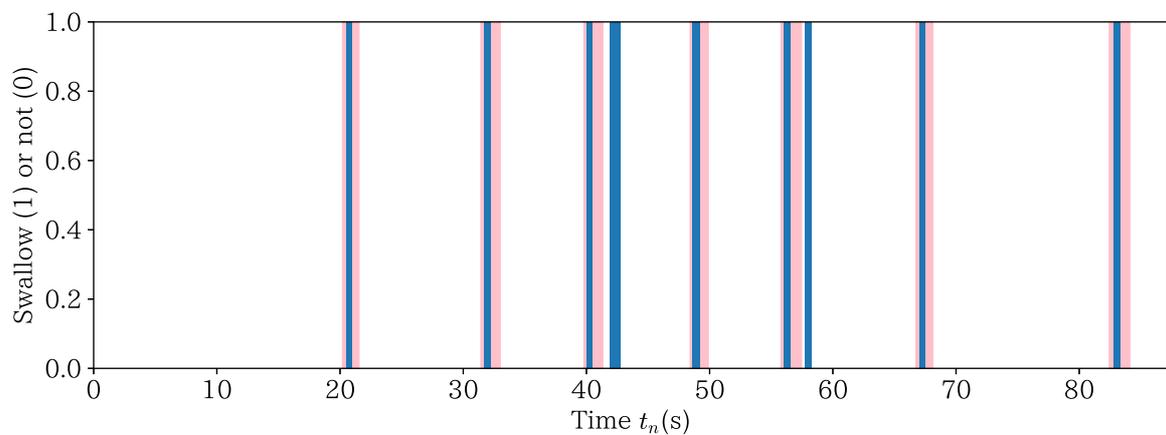


図 2: 嚥下時刻推定結果を統合し嚥下時間を推定した結果。青色の棒は推定した嚥下時間を示す。ピンク色の背景で正解嚥下時間を示す。

参考文献

- [1] 兵頭 政光, 西窪 加緒里, 弘瀬 かほり, “嚥下内視鏡検査におけるスコア評価基準 (試案) の作成とその臨床的意義,” 日本耳鼻咽喉科学会会報, 113(8), pp. 670-678, 2010.
- [2] Intel Corporation, “Intel(R) RealSense™ Camera F200 Product Specifications,” <https://ark.intel.com/content/www/us/en/ark/products/92255/intel-realsense-camera-f200.html>, (2020-01-12).
- [3] 中村薫, 前本和志, 斉藤裕佑, 谷口直嗣, 初音玲, “Intel RealSense SDK センサープログラミング,” 株式会社 翔泳社, p. 9, 2015.

5. 主な発表論文等

〔雑誌論文〕 計3件（うち査読付論文 3件／うち国際共著 0件／うちオープンアクセス 1件）

1. 著者名 宮城茂幸、小澤恵子、森谷季吉、坂本眞一、酒井道	4. 巻 20
2. 論文標題 簡易型嚥下評価システムの構築に向けて ---嚥下音と頸部深度画像を併用した嚥下評価手法の開発---	5. 発行年 2018年
3. 雑誌名 月刊地域ケアリング	6. 最初と最後の頁 73-77
掲載論文のDOI（デジタルオブジェクト識別子） なし	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 -

1. 著者名 宮城茂幸、小澤恵子、森谷季吉、坂本眞一、酒井道	4. 巻 21
2. 論文標題 簡易型嚥下評価システムの構築に向けて ---嚥下音と頸部深度画像を併用した嚥下評価手法の開発---	5. 発行年 2019年
3. 雑誌名 月刊地域ケアリング	6. 最初と最後の頁 75-79
掲載論文のDOI（デジタルオブジェクト識別子） なし	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 -

1. 著者名 Miyagi Shigeyuki, Sugiyama Syo, Kozawa Keiko, Moritani Sueyoshi, Sakamoto Shin-ichi, Sakai Osamu	4. 巻 8
2. 論文標題 Classifying Dysphagic Swallowing Sounds with Support Vector Machines	5. 発行年 2020年
3. 雑誌名 Healthcare	6. 最初と最後の頁 103 ~ 103
掲載論文のDOI（デジタルオブジェクト識別子） https://doi.org/10.3390/healthcare8020103	査読の有無 有
オープンアクセス オープンアクセスとしている（また、その予定である）	国際共著 -

〔学会発表〕 計7件（うち招待講演 0件／うち国際学会 1件）

1. 発表者名 吉田 淳、朝本 崇康、小澤 恵子、森谷 季吉、坂本 眞一、酒井 道、宮城 茂幸
2. 発表標題 RGB-Dカメラにより取得されたマーカ座標間の相対位置変化を用いた健常者の嚥下検出
3. 学会等名 情報処理学会研究報告
4. 発表年 2018年

1. 発表者名 橋本瀬菜、向樹里子、小澤恵子、宮城茂幸、森谷季吉
2. 発表標題 頸部聴診音による嚥下評価指標の検討-健常高齢者との比較-
3. 学会等名 第3回滋賀県多職種連携学会研究大会
4. 発表年 2018年

1. 発表者名 吉田淳, 小澤恵子, 森谷季吉, 坂本眞一, 酒井道, 宮城茂幸
2. 発表標題 深度画像の隣接フレーム間差分による嚥下時刻推定
3. 学会等名 ViEW2018 ビジョン技術の実利用ワークショップ2018
4. 発表年 2018年

1. 発表者名 小澤恵子, 宮城茂幸, 森谷季吉
2. 発表標題 在宅における嚥下障害早期発見に向けて --頸部聴診音による評価指標の確立の検討--
3. 学会等名 第23回日本摂食嚥下リハビリテーション学会学術大会
4. 発表年 2017年

1. 発表者名 藤戸遥, 土田侑奈, 小澤恵子, 宮城茂幸, 森谷季吉
2. 発表標題 在宅における嚥下障害早期発見に向けて --頸部聴診音による評価指標の確立の検討(中間報告)--
3. 学会等名 第2回滋賀県多職種連携学会研究大会
4. 発表年 2017年

1. 発表者名 朝本崇康, 吉田 淳, 小澤恵子, 森谷季吉, 坂本眞一, 酒井道, 宮城茂幸
2. 発表標題 投票により決定された局所深度勾配情報を用いた喉頭挙上運動の追跡
3. 学会等名 2018年電子情報通信学会総合大会
4. 発表年 2018年

1. 発表者名 Jun Yoshida, Keiko Kozawa, Sueyoshi Moritani, Shin-ichi Sakamoto, Osamu Sakai, Shigeyuki Miyagi
2. 発表標題 "Detection of Swallowing Times Using a Commercial RGB-D Camera
3. 学会等名 2019 IEEE 8th Global Conference on Consumer Electronics (GCCE) (国際学会)
4. 発表年 2019年

〔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	小澤 恵子 (Kozawa Keiko) (90747429)	滋賀県立大学・人間文化学部・准教授 (24201)	