

深度画像による健常者嚥下時の喉頭前後運動の検出手法

Detection Method of Forward/backward Laryngeal Movement Using Depth Image during Swallowing in Healthy Subjects

朝本 崇康[†]
Takayasu Asamoto

小澤 恵子[‡]
Keiko Kozawa

森谷 季吉[§]
Sueyoshi Moritani
宮城 茂幸[‡]
Shigeyuki Miyagi

坂本 眞一[‡]
Shin-ichi Sakamoto

酒井 道[‡]
Osamu Sakai

1 はじめに

近年、肺炎は 2011 年に脳血管疾患を抜き、2014 年度、日本における死因第 3 位となっている [1]。そして、高齢者の肺炎の約 70% は誤嚥性肺炎によるものであると言われている [2][3]。

現在、病院で行われている嚥下障害の評価手法として、嚥下内視鏡検査 (Videoendoscopic examination of swallowing :VE) や嚥下造影検査 (Videofluoroscopic examination of swallowing :VF) などがある。いずれも、患者に負担をかけることになり、特に VF は X 線照射による被曝を伴い侵襲性がある。さらに、どちらも高価で大きな装置を必要とし、検査の行える場所が限られる。

非侵襲で評価を行う手法としては、嚥下音を用いる手法 [4] や光学センサを使用する手法 [5] などがある。嚥下音を用いる手法により、嚥下障害のスクリーニングができることはわかっているが、嚥下音の発生メカニズムや、嚥下障害診断との対象関係など不明な点もある。一方、光学センサを使用する手法では頸部にセンサアレイを装着する必要がある。

非侵襲かつ非接触で評価を行う手法として頭頸部を横方向から撮影した RGB 画像から輪郭の変化を追跡する手法 [6][7] が提案されている。しかし、実際の診察では必ずしも横方向からの撮影ができるわけではない。最近、安価な消費者向けの RGB-D カメラが普及し、これを用いて頸部を撮影した深度画像から頸部の複数点の深度変化と表面筋電図の出力結果との同期が確認された [8]。この結果は嚥下動作が頸部の深度変化として現れることを示唆している。我々は非接触・非侵襲な嚥下機能評価手法の確立を目指しており、RGB-D カメラにより撮影された深度画像を用いることにより嚥下時の喉頭運動に関する情報が得られるのではないかと考えている。

消費者向けの RGB-D カメラとして ASUS 社の Xtion や、Microsoft 社の Kinect などがあるが、これらは人体の骨格動作を検出するといった用途を想定しているため、撮影距離が 50cm 以上必要である。一方、嚥下機能評価のためには主に頸部、大きくても頭頸部までの範囲が撮影できればよい。特に頸部の詳細な深度画像が要求され

る。そこで我々は Intel RealSense F200 を用いることとした。RealSense は Intel 社が開発した structured light 方式の深度カメラであり、認識範囲が 20cm~120cm と比較的近距离の撮影に向いている (図 1)。

本研究では、嚥下障害評価を行う前段階として、嚥下時の喉頭前後運動の検出手法を提案する。健常者を対象とし、RealSense により外部から撮影を行い、喉頭前後運動の移動距離およびその動作に要する時間を測定した。また、その結果から速度の特徴量抽出を行った。喉頭前後運動の時間・速度の比率から、X 線学的解析 [10] と比較を行い、どの程度正確に喉頭前後運動が測定できるか、提案する検出手法に妥当性、信頼性があるか検討した。



図 1 Intel RealSense F200

2 実験方法

2.1 RGB 画像および深度画像の取得方法

RealSense には structured light 方式の深度カメラが搭載されている。IR レーザープロジェクタから照射されるパターンを赤外線カメラで撮影することにより距離を計測する。RGB カメラも搭載されているので、深度画像の取得と同時に RGB 画像の取得も行い、後に述べるように、RGB 画像上のマーカーから深度取得位置を定めた。実験では、RGB 画像、深度画像ともに解像度を 640 × 480 とした。喉頭挙上運動にかかる時間は極めて短く、その間できる限り多くの深度データを取得するために最大フレームレート 30 を利用した。

2.2 RGB 画像と深度画像のずれの修正

RGB カメラと深度カメラの間には視野角の違いがあり、さらに 2 つのカメラの位置が異なるため、取得される RGB 画像と深度画像にはずれが生じる。したがって、正確な喉頭前後運動を測定するためには、RGB 画像と深度画像の座標を対応させなければならない。このずれを修正することが必要となる。深度画像の方が RGB 画像に比べて画角が大きいため、深度画像の座標および深度値を RGB 画像にマッピングする必要があ

[†]滋賀県立大学大学院, Graduate school of The University of Shiga Prefecture

[‡]滋賀県立大学, The University of Shiga Prefecture

[§]草津総合病院, Kusatsu General Hospital

る。本研究では、Intel RealSense SDK に用意されている MapDepthToColor 関数を用いてマッピングを行った。

2.3 撮影方法

20 歳代の男性健常者 10 名を対象とし、測定に用いる嚥下食としては水を用いた。水量は改訂水飲み検査同様約 3ml とし、スプーンを用いて被験者に与えた。

図 2 に撮影環境を示す。撮影では、壁に対して平行に椅子を配置し、被験者を椅子に座らせた。嚥下時に被験者には顔をなるべく動かさないように指示したが、完全に動かないようにすることは不可能であり、かつ固定すると自然な嚥下を行うことができなくなる恐れがある。そこで、被験者には頭を軽く壁に付けるように指示した。1 回の撮影で被験者には水を 3 回与え、嚥下させた。

カメラ配置場所は被験者の正面とし、カメラの仰角がおよそ 5 度で、頸部が深度画像中心になるように設定する。頸部を大きく撮影するために、被験者の喉頭と RealSense までの距離は約 300mm とした。撮影を開始する前に、RealSense と喉頭までの距離ができる限り 300mm となるよう深度画像を確認しながら三脚の位置を調整した。

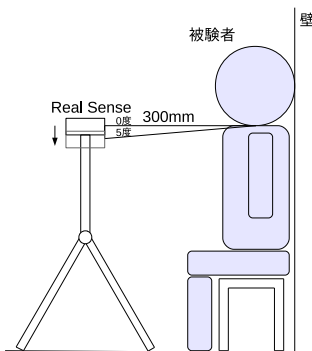


図 2 撮影環境

深度取得を容易にするために、被験者の頸部 3 箇所に、直径 8mm の単色のマーカーを貼付した。マーカーの位置は、図 3 のとおりで、頸部全体の動きがわかりやすいように舌骨、甲状軟骨、輪状軟骨とした。

2.4 マーカーの重心座標の深度値の取得

マーカーの重心座標に対応する深度値を取得するために、初めに RGB 画像からマーカーの重心位置を決定する。あらかじめ撮影した複数の RGB 画像から 3 点のマーカーの RGB 値を取得し、その RGB 値を HSV 値に変換した。複数の HSV 値から、どのフレームでもマーカーが検出できるように HSV 値の範囲を決めた。決定した HSV 値の範囲を用いて、RGB 画像を 2 値化処理する。次に、2 値化処理した画像に OpenCV 3.0 の connectedComponentsWithStats 関数を適用し、ラベリングを行い、マーカー 3 点に対応する領域を抽出し、それぞれの領域に対して重心座標を求める。深度画像において、得られた重心座標に対応する深度値をマーカー位置における深度値とする。

図 4 に、求めた 3 つのマーカーの重心座標における深度値の時間変化の 1 例を示す。撮影した RGB 画像から、嚥下していたと見られる区間は、9.9~11.5 秒、18.3~20.0 秒、26.7~28.1 秒と分かった。図 4 より、それらの区間で深度値が大きく変化していることから、嚥下に伴う喉頭前後運動が深度変化に対応していることが確認できる。

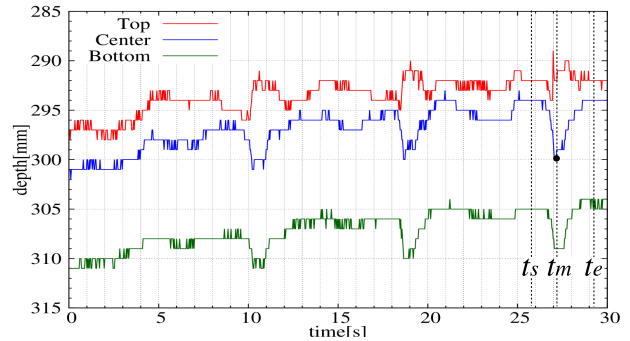


図 4 取得できたマーカーの重心座標の深度値の 1 例

3 特徴量抽出手法

ここでは、取得した深度値について、喉頭前後運動の特徴となる喉頭移動距離、喉頭前後運動時間、喉頭平均移動速度を求める。

3.1 深度取得位置の補正

嚥下時の喉頭挙上に伴い頸部の皮膚が上下するため、その影響によりマーカーの重心位置も上下する。図 4 のようなマーカー重心の深度値の追跡は頸部の固定位置の深度変化とは必ずしも一致しない。そこで、嚥下時にはほとんど頸部全体の移動が少ないことを利用し、ここでは以下の手順を提案する。

1. 甲状軟骨のマーカー重心座標の深度値変化から局所的に最大値を取る時刻 (t_m) を求める。例えば図 4 の t_m を参照のこと。
2. 時刻 t_m から 1.5 秒前の時刻 t_s における各マーカーの重心座標 (x_i, y_i) を求める。
3. 時刻 t_s から 3.5 秒後の時刻 t_e までは深度取得位置を (x_i, y_i) に固定し、その位置での深度変化を追跡する。

以上の手順を図 4 の 3 回目の喉頭前後運動に適用すると、図 5 となる。

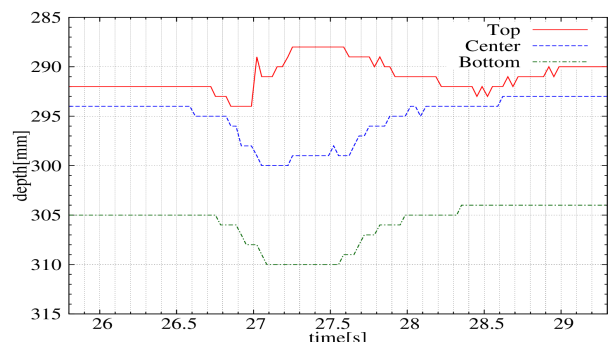


図 5 深度取得位置の補正後の深度値の 1 例

3.2 メディアンフィルタの適用

RGB 画像と深度画像を保存する際、使用機器の性能上、必ずしも同じ時間間隔で保存できるとは限らず、サンプリング位置にわずかなずれが発生する。そこで深度値変化に対し線形補間を用いてリサンプリングを行った。さらに深度画像からの計測結果にはスパイク状のノイズが重畳する。このノイズを低減するためにメディアンフィルタを適用した。メディアンフィルタは、変化によるエッジを保存しつつスパイク状のノイズを低減する効果が大きいことが知られている。今回はメディアンフィルタの区間長を 21 に設定した。以上の処理を図 5 に適用した結果の 1 例を図 6 に示す。

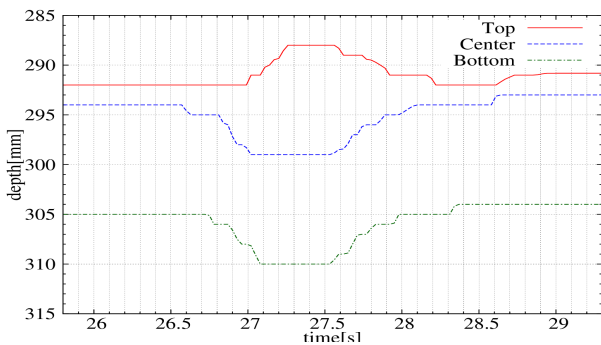


図 6 メディアンフィルタ適用後の深度値の 1 例

3.3 喉頭前後運動の特徴量抽出

喉頭前後運動の変位を求めるために、ピーク位置の深度値 Z_{max} および時刻 t を読み取る。ピーク位置を求めた後、喉頭前方運動が始まる前の深度値、つまり嚙下前の定常状態の深度値を求めるために、ピーク位置から約 0.8 秒前の時刻から約 0.15 秒前の深度値までの平均を算出する。また、同様に喉頭後方運動後の深度値はピーク位置から約 1.2 秒後の時刻から約 0.15 秒前の深度値までの平均を算出し、それぞれを嚙下を行う前後の定常状態の深度値とした。なお、約 0.8 秒、約 1.2 秒の値は撮影した複数のデータを参考に設定した。

その後、求めた平均値を超える点を立下り点、立ち上がり点とし、立下り点からピーク値までの時間を前方運動所要時間 t_1 、ピーク値から立ち上がり点までを後方運動所要時間 t_2 とした。さらに、立下り点と立ち上がり点を直線で結び、ピーク値からその直線に対して垂線を引き、その垂線の長さを喉頭が前後に移動した距離 d とした。求めた喉頭前後移動距離と前方運動所要時間および後方運動所要時間から平均速度 v_1, v_2 を求めた (図 7)。

4 特徴量抽出結果

健常者 10 名の実験結果から特徴量抽出をした結果、移動距離平均 \bar{d} は 2.9mm、所要時間平均 \bar{t}_1 が 0.61 ± 0.21 秒、 \bar{t}_2 が 0.87 ± 0.28 秒となった。速度平均 \bar{v}_1 は 4.8mm/sec となり、 \bar{v}_2 は 3.5mm/sec、前方運動と後方運動の時間および速度の比率を求めた結果、時間比率が 1:1.418、速度比率が 1:0.731 となった。なお、健常者 2 名において測定不能のものが 3 データあり、そのデータは除外した。

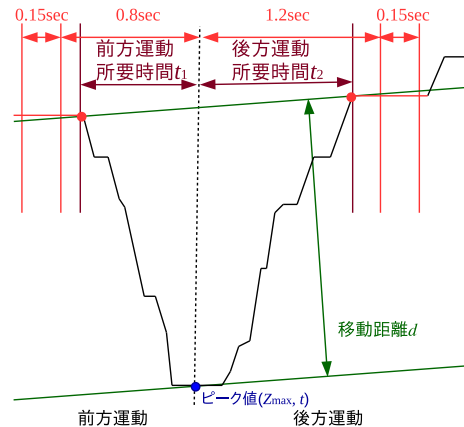


図 7 前後運動の各特徴量の抽出

5 考察

5.1 実験方法について

実験は、被験者になるべく頸部を動かさないように指示したが、完全に動かさないようにすることは不可能である。したがって、嚙下時の頸部移動を打ち消す補正が正確な深度値を取得するためには必要であると考え。また、正面 (仰角 5 度) での撮影に限定したが、必ずしもその角度で良いとは限らず、角度によって深度値の取得に影響があると考え。したがって、正面 (仰角 5 度) 以外に 10 度、15 度などあらゆる角度からの撮影を行い、最適な角度からの撮影が必要である。また、撮影距離は、RealSense の認識範囲が 200mm からであること、500mm での撮影時、ノイズによる影響が大きいため、被験者の頸部の移動も踏まえ、300mm での撮影とした。

5.2 特徴量抽出結果について

特徴量抽出する上で 1 次線形補間とメディアンフィルタを 21 区間で適応し、スパイク状のノイズの除去には成功したが、取得された嚙下時の深度値の変化が 1~3mm 小さくなる場合があった。したがって、他の手段を用いての処理を検討する必要があると考え。

特徴量抽出結果から、喉頭前後運動の移動距離は小さい人で 1~3mm 程度、大きい人で 5mm 程度となっており、同じ人であっても 3 回の嚙下でそれぞれ数 mm の変動があるものがあつた。これは、甲状軟骨が比較的確認しやすい人とそうでない人といった個人差があることが大きな要因となっていると考えられ、さらに、同じ人であっても、すべて同じように嚙下するとは限らないのではないかと考える。また、甲状軟骨が確認しやすい人の移動距離は大きくなる傾向にあつた。

次に、特徴量抽出の結果で得られたデータと X 線学的解析 [10] との比較を行う (表 1, 表 2)。X 線学的解析では、前方運動、後方運動にかかる平均時間は 0.42 ± 0.14 秒、 0.59 ± 0.16 秒であり、速度は 19.1mm/s、10.1mm/s となったと示されている。この結果と比較すると、前後ともに本研究結果の方が平均時間が少し大きく、速度は小さくなった。これは、外部から甲状軟骨の位置を正確

に求めることが難しいことと、頸部内部と外部から認識できる運動が異なるからであると推測される。しかし、喉頭前後運動の時間、速度比率を比較すると、X線学的解析では1:1.405, 1:0.529であり本研究結果に近い値を示した。したがって、絶対数は異なるものの喉頭前後運動全体の動きはほぼ正確に読み取ることができたと考え、取得されたデータに信頼性、妥当性があると考えられる。

表 1 各特徴量の比較

	t_1 [s]	t_2 [s]	\bar{v}_1 [mm/s]	\bar{v}_2 [mm/s]
本手法	0.61±0.21	0.87±0.28	4.8	3.5
X線学的解析	0.42±0.14	0.59±0.16	19.1	10.1

表 2 時間、速度比率の比較

	時間比率	速度比率
本手法	1 : 1.418	1 : 0.731
X線学的解析	1 : 1.405	1 : 0.529

5.3 嚥下区間の検出手法について

健常者 10 名の実験結果から、嚥下時に舌骨に貼付したマーカーの重心座標の深度値が大きく変化し、前方に移動していることがわかった。また、甲状軟骨の深度値は、舌骨の変化量に比べ数 mm 小さいが後方に移動していることがわかった。今回取得したデータのほとんどにこの変化が見られた。つまり、長い撮影時間の間に対象となる嚥下を検出するためには、この傾向を探索すれば良いことになる。

また、喉頭前後運動にかかる時間は、ばらつきはあるものの約 1.5 秒という結果が得られた。したがって、この傾向が約 1.5 秒の間で見られる区間を検出することで、嚥下に伴う喉頭前後運動を検出できると考える。

5.4 従来手法との比較

本研究によって、被験者に加速度トランスデューサー [4] や光学センサ [5] などのデバイスを取り付けることなく、RGB 画像のみで喉頭挙上を検出した例 [6][7] に比べて、周りの環境にも影響されず、非接触、非侵襲で喉頭前後運動の特徴量抽出を行うことができた。また、本研究によって、RealSense を用いて 300mm 離れた場所であっても嚥下にとまなう数 mm 程度の微小な変化を検出できることがわかり、様々な分野に応用できる可能性があることがわかった。

6 おわりに

本研究では、RealSense を用いて取得された RGB 画像および深度画像から特徴抽出を行い、嚥下に伴う数 mm 程度の喉頭前後運動を捉えることができた。X線学的解析と比べ、絶対数は異なるものの、前方運動と後方運動の比率はほぼ同じ結果となり、本検出手法に妥当性、信頼性が証明された。今後は、3 点だけでなく頸部全体の動きを検出し、より正確な嚥下時の頸部の運動を捉えるようにする。

謝辞

社会医療法人誠光会草津総合病院頭頸部外科センターの医師、看護師および関係者の皆様方にご協力いただいた。この場を借りて御礼申し上げる。また、実験に快く御協力頂いた皆様方に感謝の意を表す。

参考文献

- [1] 厚生労働省, 平成 27 年版厚生労働白書 資料編, p.10.
- [2] Teramoto S, Fukuchi Y, Sasaki H, Sato K, Sekizawa K, Matsuse T, "High incidence of aspiration pneumonia in community- and hospital-acquired pneumonia in hospitalized patients: a multicenter, prospective study in Japan", Journal of the American Geriatrics Society 56(3), pp.577-579, 2008.
- [3] 大類孝, "超高齢社会における誤嚥性肺炎の現状", 日本老年医学会雑誌 56(4), pp.458-460, 2013.
- [4] 中山裕司, 高橋浩二, 宇山理沙, 平野薫, 深澤美樹, 南雲正男, "嚥下音の産生部位と音響特性の検討 - 健常成人を対象として -", 昭和歯学会雑誌 26(2), pp.163-174, 2006.
- [5] 本間正寿, 林豊彦, 道見登, 谷口裕重, 井上誠, "嚥下機能評価システム SFN/3A における喉頭位置トラッキング法の改良", 電子情報通信学会技術研究報告, 113(103), pp.33-38, 2013.
- [6] 高橋明子, 清水良昭, 西條光雅, 高野梨沙, 小野大地, 滝田裕美, 安井利一, "頸部輪郭抽出法による喉頭運動の検出 - 食物物性の違いによる差の検出 -", 明海歯科医学会 44(1), pp.92-97, 2015.
- [7] 高橋明子, 清水良昭, 北邦宏, 中筋宣子, 小野大地, 滝田裕美, 安井利一, "頸部輪郭抽出法による喉頭運動の検出", 明海歯科医学会 43(1), pp.34-43, 2014.
- [8] 熊川瑛至, 本木実, 古閑公治, 非接触型センサを用いた嚥下動作解析法の一提案, 電子情報通信学会 2015 年総合大会 情報・システムソサイエティ特別企画 学生ポスターセッション予稿集, ISS-SP-219, 2015.
- [9] 才藤栄一, 向井美恵, 摂食・嚥下リハビリテーション 第 2 版, 医歯薬出版株式会社, pp. 130-136, 2007.
- [10] 古川浩三, "嚥下における喉頭運動の X 線学的解析 - 特に年齢変化について -", 日本耳鼻咽喉科学会会報 87(2), pp.169-181, 1984.
- [11] 中村薫, 前本和志, 齊藤祐佑, 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, 2015.
- [12] Intel Developer Zone, <https://software.intel.com/en-us/RealSense/Home>, 2016-02-05.