

# 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

吉田 淳<sup>1,a)</sup> 朝本 崇康<sup>2</sup> 小澤 恵子<sup>1</sup> 森谷 季吉<sup>3</sup> 坂本 眞一<sup>1</sup> 酒井 道<sup>1</sup> 宮城 茂幸<sup>1</sup>

**概要:** 近年高齢化社会の進展に伴い、誤嚥性肺炎による死亡率が増加傾向にある。このため誤嚥の危険性を早期発見するための嚥下評価が必要とされている。深度画像を用いた、非侵襲、非接触かつ安価な嚥下評価手法が着目されている。深度画像中の喉頭隆起の動きを利用し、喉頭挙上の追跡を行うが、喉頭隆起が不明瞭な場合は追跡に失敗する可能性がある。そこで本研究では頸部正面から撮影された RGB 画像と深度画像を併用し、喉頭隆起が不明瞭な被験者においても嚥下を検出する手法を提案する。輪状軟骨および両鎖骨骨頭の3点にマーカを貼付し、これらの3点で形成される三角形の変形量の変化から嚥下の検出を行う。女性を含む複数の健常者に提案手法を適用し、検出した嚥下区間に喉頭挙上が最大となる時刻が含まれているかどうかを判定した。その結果、再現率と適合率はそれぞれ0.631, 0.582であった。また内視鏡検査におけるホワイトアウトの時刻が含まれているかどうかで判定した場合は、それぞれ0.658, 0.546であった。いずれの場合も、喉頭隆起が不明瞭な被験者における嚥下検出の可能性が確認できた。

## 1. 序論

2015年の日本における死因の第3位は増加傾向にある肺炎であり、2011年に脳血管疾患を抜いている[6]。また、高齢者の肺炎の約70%が誤嚥性肺炎によるものと言われている[15][10]。誤嚥性肺炎は加齢や脳の障害による嚥下機能の低下が原因となり、食物を上手く飲み込めないことにより引き起こされる。高齢化の進む日本では高齢者の誤嚥が増加傾向にあるため、嚥下機能の評価をおこない、誤嚥する危険性を早期発見する必要がある。

現在、病院で行われている嚥下評価手法として、嚥下内視鏡検査(VE:Videoendoscopic examination of swallowing)と嚥下造影検査(VF:Videofluoroscopic examination of swallowing)が行われている。VEは内視鏡を鼻から挿入し、水やとろみのついた固形物(以降、とろみ)、ゼリーといった嚥下食を与え、嚥下の様子を内視鏡で観察して評価する手法である。一方でVFはバリウムを含む嚥下食を与え、X線照射の元で嚥下食の動きを観察することで評価する手法である。どちらの手法とも患者に負担をかける。さらに、VFにおいてはX線による被曝を伴い、長時間の

評価は行えず、侵襲性がある。また、これらの手法で用いられる装置は非常に高価で、大きく、評価のできる場所が病院など限られる。そのため、非侵襲で安価な嚥下評価手法が求められている。

非侵襲で安価な嚥下評価手法として、嚥下音から嚥下障害の判定を行う研究が行われている[9]。また、反射型フォトセンサを用いて嚥下障害の判定を行う研究もなされている[4]。しかし、これらの手法は頸部に装置を取り付け評価を行うため、非接触ではない。

非侵襲で非接触、安価な嚥下評価手法として、デジタルビデオカメラを用いた輪郭抽出による喉頭運動の検出を行った研究がある[13][12]。別の手法として非接触型のセンサを用いた研究も行われており、筋電図の出力結果と同期が確認された[7]。そのほかにカラー(RGB)画像に加えて深度画像も撮影可能なMicrosoft社のKinectカメラを使用した、喉頭隆起部のトラッキングによる特徴量検出の研究も行われている[8]。同様にKinectカメラを使用し、嚥下時の喉頭部の前後運動の振幅を測定している研究がある[14]。これらの手法には以下の2点の問題点が挙げられる。

- カメラ位置が固定され、かつ頭頸部のぶれがないことを前提とする。
- 喉頭隆起が不明瞭な被験者では喉頭部の追跡が十分でない。

そこで、本研究では正面から健常者の頸部の撮影を行い、喉頭隆起が明瞭、不明瞭に関わらず嚥下を検出可能な方法

<sup>1</sup> 滋賀県立大学  
The University of Shiga Prefecture  
<sup>2</sup> 滋賀県立大学院  
Graduate school of The University of Shiga Prefecture  
<sup>3</sup> 草津総合病院  
Kusatsu General Hospital  
a) of23jyoshida@ec.usp.ac.jp



図 1 頸部 RGB 画像

を提案する。

カメラを用いて正面から撮影した RGB 画像を図 1 に示す。RGB 画像だけでは頸部の色の変化が乏しく、喉頭の変化が目立たない。そこで、RGB-D カメラを使用し、喉頭部の深度情報の取得も行う。また、輪状軟骨と両鎖骨骨頭に注目し、この 3 点がつくる三角形の形状変化をみることで嚥下を検出する。なお、本研究では RGB-D カメラに安価で、持ち運びも容易な Intel 社の RealSense を用いる。

## 2. 嚥下検出のアルゴリズム

本章ではまず嚥下のメカニズム [11] についてまとめ、嚥下動作に伴う頸部の動きから嚥下検出を行う手法について説明する。

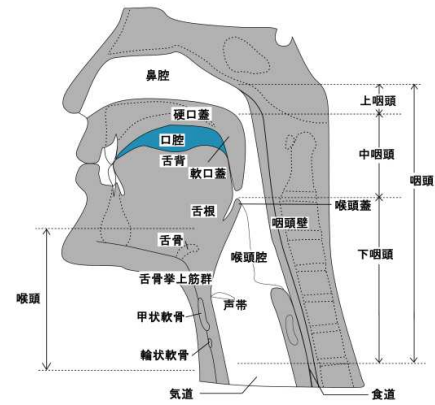
### 2.1 嚥下のメカニズム

嚥下は食塊を口控から胃へ送り込む一連の動作のことをいい、口控期、咽頭期、食道期の 3 段階に分けられる。

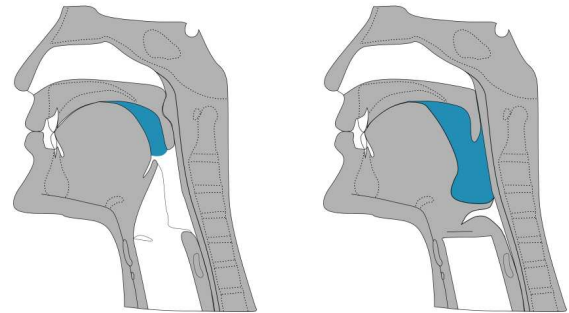
口控期は咀嚼などにより食塊の準備が出来た後に始まる。図 2(a) に食塊が口控内にある状態を、食塊を青色として示す。この時、舌前部が硬口蓋に押し付けられながら、食塊を後方に押す。また、舌根と軟口蓋が密着しており、食塊が口控内に保持される。咽頭期が始まると舌により食塊が咽頭に圧出される(図 2(b))。この時、軟口蓋が挙上し、収縮した咽頭壁と接し、上咽頭が閉鎖される。また、舌背が軟口蓋に密着し、口控が遮断される。その後、舌と舌骨が前上方へ動き始める。喉頭も挙上し、喉頭蓋が水平になり、気道が閉鎖する(図 2(c))。舌骨が最大限前上方に達し、喉頭が舌骨に近接する時、喉頭蓋が下方に傾く。また、喉頭の前上方への移動により、食道入口部は拡大する(図 2(d))。喉頭の収縮が下咽頭まで進み、中咽頭が咽頭壁、舌骨、軟口蓋の密着によって閉鎖される。最後に、食塊が頸部食道に送られ、舌と舌根、喉頭などの器官が安静位に戻り、気道が再開通する(図 2(e))。

### 2.2 座標系の定義

透視投影を用いた RGB-D カメラの座標系を図 3 に示す。ワールド座標系  $X, Y, Z$  における点の座標を  $(x, y, z)$  と表す。RGB カメラ座標系は RGB カメラ中心を原点とした  $X_c, Y_c, Z_c$  軸で表され、その座標を  $(x_c, y_c, z_c)$  とする。ま

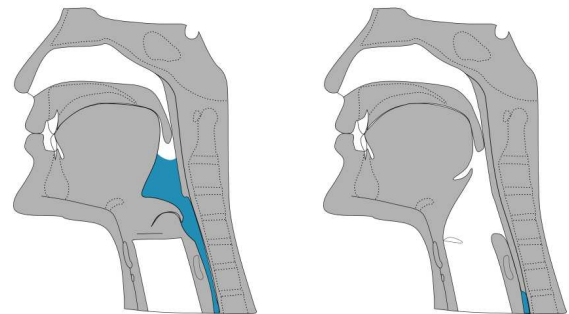


(a) 食塊が口控 (口控期)



(b) 食塊が口控から咽頭へ (咽頭期)

(c) 食塊が咽頭 (咽頭期)



(d) 食塊が咽頭から食道へ (咽頭期)

(e) 食塊が食道 (食道期)

図 2 嚥下のメカニズム (青: 食塊)

た、RGB カメラに対する画像座標系は  $U_c, V_c$  軸で表現され、その座標を  $(u_c, v_c)$  とし、1 画素は B(青), G(緑), R(赤) で構成される。深度カメラ座標系は深度カメラ中心を原点とした  $X_d, Y_d, Z_d$  軸で表され、その座標を  $(x_d, y_d, z_d)$  とする。また、深度カメラに対する画像座標系は  $U_d, V_d$  軸で表現され、その座標を  $(u_d, v_d)$  とし、1 画素は  $w$ (深度) のみで構成される。

RGB 画像座標と深度画像座標はずれているので、位置合せを行い深度画像と RGB 画像間において各画素の対応関係を決定する必要がある。位置合せをした深度画像の画像座標系では 1 画素に RGB 画像座標系の座標  $u'_c, v'_c$  が付加され、画素値として  $(u'_c, v'_c, w)$  をもつ。

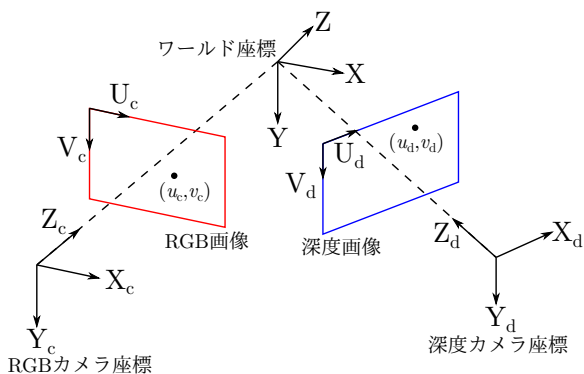


図 3 透視投影モデル



図 4 深度画像

実際に撮影をして取得した RGB 画像の頸部を図 1 に、深度画像の頸部を図 4 に示す。ただし、深度画像は 300mm を中心に 1 階調 1mm の 256 階調のグレースケール画像としている。また、図 1 に示すように安静時の輪状軟骨部と基準点とする両鎖骨骨頭に黄色マーカを貼付し、部位を認識できるようにしている。

### 2.3 検出アルゴリズムの概要

輪状軟骨部と両鎖骨骨頭で形成される三角形を図 5(a) に示す。頸部を撮影した画像の各フレーム毎にこの三角形の見かけの形状が変化する (図 5(b))。三角形の見かけの形状変化の要因としては、嚥下以外の体動 (以降、グローバルモーション) と嚥下動作による輪状軟骨部の前後上下運動の影響が考えられる。グローバルモーションによる三角形の見かけの形状変化は、カメラと頸部との相対的な角度や位置変により引き起こされるので、三角形の形状そのものは変化しない。一方で、嚥下動作による三角形の見かけの形状変化では、輪状軟骨部の前後上下運動があるために三角形の形状そのものが変化する。そこで、安静時の参照フレームにおける三角形と、それ以降の補正フレームにおける三角形を比較し、その変化量から嚥下動作を検出する。変化量を取得するために補正フレームの三角形の回転・並進により、参照フレームの両鎖骨骨頭への位置合わせを行い、輪状軟骨部の座標のずれを変化量として求める (図 5(c))。また、検出結果の統合を行い、最終的な嚥下区間を推定する。

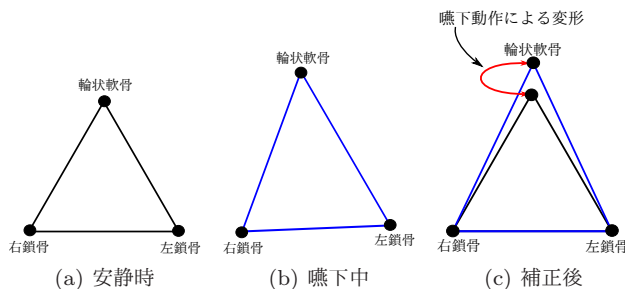


図 5 回転と並進処理による三角形の補正

表 1 2 値化の際に使用したしきい値

	色相 H	彩度 S	明度 V
上限	35	255	255
下限	18	200	50

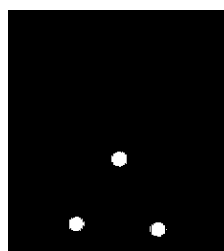


図 6 2 値化画像

### 2.4 マーカ座標の取得

マーカ 3 点の RGB カメラ座標を求める方法を述べる。使用したマーカは黄色であるため、それに対応した色領域を求め、ラベリング処理によりマーカ重心の RGB 画像座標  $(u_c, v_c)$  が導出できる。RGB 画像座標  $(u_c, v_c)$  に対応する位置合せをした深度画像での座標  $(u_d, v_d)$  を求め、その画素値  $(u'_c, v'_c, w)$  を実座標変換することによりマーカ 3 点の RGB カメラ座標を取得する。

初めに、色領域を求める際、外乱光による影響を軽減するために RGB 画像から HSV 画像へ変換する。変換した HSV 画像に対し黄色マーカを検出できるように表 1 のしきい値で 2 値化を行う。図 1 の RGB 画像に対し HSV 変換後、2 値化した画像を図 6 に示す。

2 値化した画像に対しラベリング処理を行い、各ラベリングされた領域の面積と重心を取得する。上位 3 つの面積が大きい領域をマーカにより生成された領域であるとし、その重心をマーカ点  $P_{cn}(u_{cn}, v_{cn}), (n = 0, 1, 2)$  とする。他の領域はノイズより生成された領域であるため以降では考慮しない。また、この処理の際にマーカ欠損によりマーカを 3 点検出できなかったフレームは嚥下検出の評価対象としない。

次に、RGB 画像上でのマーカ点  $P_{cn}$  に対応する深度画像上でのマーカ点を取得する。求めた RGB 画像上でのマーカ点  $P_{cn}$  に最も距離の近い深度画像上の点  $P_{dn}(u_{dn}, v_{dn}), (n = 0, 1, 2)$  をその画素値  $(u'_c, v'_c, w)$  を用いて探索し、深度画像上でのマーカ点とする。つまり、

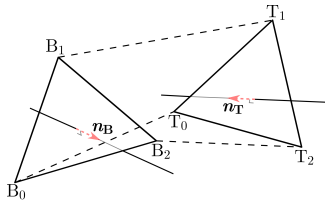


図 7 回転・並進処理前の点の位置関係

$$d^2 = (u_c - u'_c)^2 + (v_c - v'_c)^2, \quad (1)$$

$$(u'_c, v'_c) = \arg \min_{(u'_c, v'_c)} d^2.$$

その後、撮影時に生じる深度  $w$  の欠損による影響を防ぐため、マーカー点  $P_{dn}$  に対して、マーカー点を中心とする 1 辺 5 ピクセルの正方形範囲の深度値  $w$  のメディアンフィルタをかける。ただし、メディアンフィルタは欠損点 ( $w = 0$ ) である点を除いた点群の中央値をとる。

深度画像における求めたマーカー点  $P_{dn}$  それぞれに対し式 (2) を用いて RGB カメラ座標へと座標変換をする。

$$\begin{aligned} x_c &= \frac{u'_c - c_x}{f_x} w, \\ y_c &= \frac{v'_c - c_y}{f_y} w, \\ z_c &= w. \end{aligned} \quad (2)$$

ここで、 $f_x$  は X 軸方向、 $f_y$  は Y 軸方向に対応するピクセル単位で表された焦点距離である。また、 $(c_x, c_y)$  は光軸中心である。これらのカメラパラメータは事前に決定しておく。

## 2.5 回転・並進

各フレームで取得したマーカー 3 点で形成される三角形の位置合わせを行うために回転・並進による補正を行う。

撮影した画像から被験者が安静状態となったフレームを基準フレームと定義し、基準フレーム以降のフレームを回転と並進によるマーカー 3 点の位置合わせをする補正フレームと定義する。また、基準フレーム以前のフレームは以降では考慮しない。基準フレームのマーカーの座標を右鎖骨、輪状軟骨、左鎖骨の順に  $B_0, B_1, B_2$  とする。また、補正フレームのマーカーの座標を右鎖骨、輪状軟骨、左鎖骨の順に  $T_0, T_1, T_2$  とする。これらの座標の定義を図 7 に示す。グローバルモーションを補正するためには  $T_0, T_1, T_2$  をそれぞれ  $B_0, B_1, B_2$  へと一致させるように回転と並進を行う。

基準フレームと補正フレームにおいて、マーカー 3 点がつくる平面の単位法線ベクトルをそれぞれ  $n_B, n_T$  とし、求めたそれぞれの平面の単位法線ベクトル  $n_B, n_T$  を平行にするような回転行列  $R_0$  を導出する。回転行列  $R_0$  は四元数 [5] を用いて導出した。回転後のマーカー 3 点の座標を右鎖骨、輪状軟骨、左鎖骨の順に  $T'_0, T'_1, T'_2$  とする。回転行列  $R_0$  による回転後の状態を図 8 に示す。

しかし、ここまでの処理では法線方向は平行となるが、

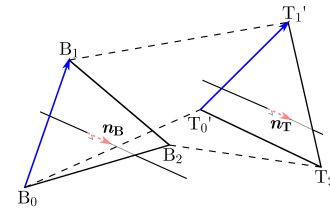


図 8 回転行列  $R_0$  による回転後の点の位置関係

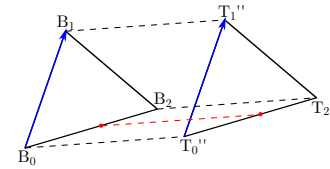


図 9 回転行列  $R_1$  による回転後の点の位置関係

平面上のベクトルは平行にならない。そのため、平面上のベクトル  $\overrightarrow{B_0B_1}$  と  $\overrightarrow{T'_0T'_1}$  を平行にするような回転行列  $R_1$  を四元数 [5] を使用して導出した。2 度目の回転後のマーカー 3 点の座標を右鎖骨、輪状軟骨、左鎖骨の順に  $T''_0, T''_1, T''_2$  とする。回転行列  $R_1$  による回転後の状態を図 9 に示す。

最後に並進補正をかけ、マーカー 3 点を一致させるように動かす。嚙下時の輪状軟骨部のマーカーの座標変化の影響を小さくするため、並進ベクトル  $t$  は鎖骨 2 点のマーカーの中心座標の差とする。RGB カメラ座標系の原点を  $O_c$  とすると、

$$t = \frac{\overrightarrow{O_cB_0} + \overrightarrow{O_cB_2}}{2} - \frac{\overrightarrow{O_cT''_0} + \overrightarrow{O_cT''_2}}{2},$$

で与えられる。並進補正後の輪状軟骨のマーカー座標を  $T''_1$  と表す。

## 2.6 変化量

点  $B_1$  の座標を  $(x_{cB}, y_{cB}, z_{cB})$ 、点  $T_1$  の座標を  $(x_{cT}, y_{cT}, z_{cT})$  とする。この時、輪状軟骨部のマーカーの変化量  $e$  を式 (3) のように定義する。

$$e = (x_{cB} - x_{cT})^2 + (y_{cB} - y_{cT})^2 + (z_{cB} - z_{cT})^2. \quad (3)$$

この変化量  $e$  について回転・並進による補正前の  $T_1$  を用いた変化量を  $e_b$  とし、補正後の  $T''_1$  を用いた変化量を  $e_a$  とする。図 10 に変化量  $e_b, e_a$  のマーカー座標の比較場所を示す。変化量  $e_b, e_a$  を導出した結果からグローバルモーション補正の評価を行う。また、変化量  $e_a$  から嚙下の検出を行う。これまでの処理をかけたデータの変化量  $e_b, e_a$  を図 11 に示す。

## 2.7 しきい値処理と投票による嚙下検出

求めた変化量  $e_a$  の時間間隔は均一ではないため、一次線形補間を用いてリサンプリングする。また、ノイズやスプリットの影響を除去するためにカットオフ正規化角周波数  $\pi/20$  のローパスフィルタを適用する。ローパスフィルタ後



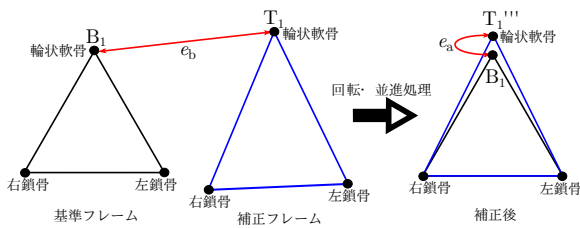


図 10 基準と補正フレーム間の輪状軟骨部マーカ座標変化量

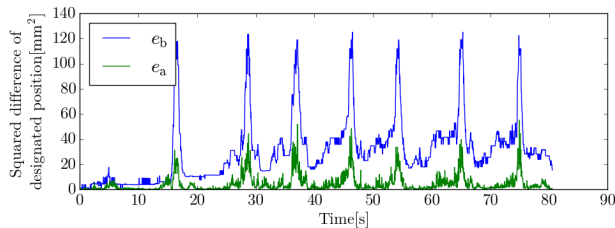


図 11 回転・並進補正前後の変化量グラフ

の変化量に対し、しきい値処理により嚙下が発生しているかどうかを判定する。その後、嚙下動作の連続性を考慮するために多数決投票を行い最終的な嚙下区間判定を行う。

ローパスフィルタ後の変化量の最大値は、撮影した画像毎に異なるのでしきい値を変化させる必要がある。今回は最大値の40%をしきい値と定義し、変化量がしきい値を越えているならば嚙下が発生していると判定する。

しきい値処理のみでは嚙下区間の誤判定をまねく恐れがある。例えば、ローパスフィルタ後の変化量において変化量の最大値付近でスプリットが発生することがある。しきい値処理ではこれを複数回の嚙下と判定する可能性がある。しかし、実際の嚙下は約1秒で行われる[2]のでそれ以下の極めて短い時間での嚙下は発生しえない。そこで、現在の点とその前後7点ずつの計15点で嚙下であるか非嚙下であるか多数決投票をすることにより、嚙下動作の連続性を考慮する。多数決投票において、各点の判定結果は前述のしきい値による結果を利用し、嚙下と判定している点が過半数を占めた時、現在の点を連続性を考慮した嚙下動作をしている点とする。

図11のデータの補正後変化量  $e_a$  に対して、ローパスフィルタをかけた後の変化量と投票による嚙下検出を行った結果を図12に示す。また、図12のデータに対して、連続して嚙下動作をしている区間を嚙下区間とし、その開始時刻と終了時刻を表2に示す。嚙下検出評価は各データで表2のような嚙下区間を導出し、正解嚙下時刻がその区間に含まれているかで判定する。

### 3. 撮影実験

本章では、初めに使用したRGB-Dカメラの仕様についてまとめ、RGB-DカメラによるRGB画像と深度画像の取得方法および記録方法、位置合せ処理を述べる。その後、

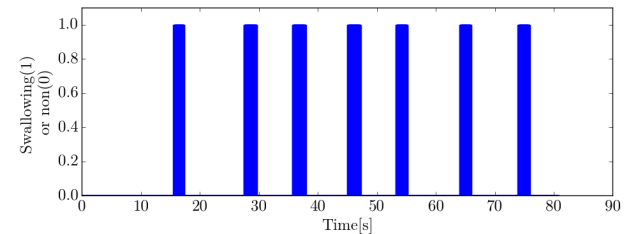
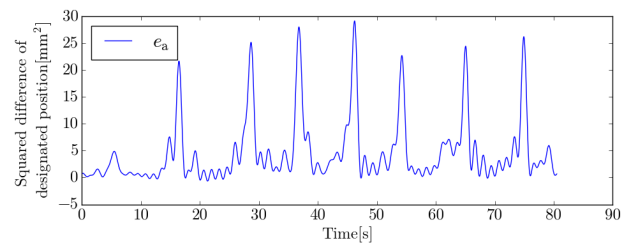


図 12 ローパスフィルタをかけた後の変化量グラフ (上) と嚙下検出結果 (下)

表 2 嚙下区間検出結果

嚙下区間 No.	嚙下開始時刻 [s]	嚙下終了時刻 [s]
1	15.9775	16.9114
2	27.9189	29.2197
3	36.1244	37.4920
4	45.4640	46.7983
5	53.6696	54.7703
6	64.5102	65.5443
7	74.3836	75.4510

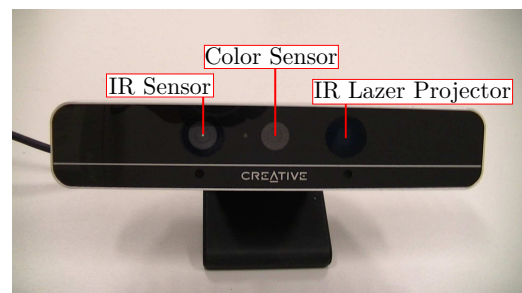


図 13 Intel RealSense F200

実際の撮影条件と撮影方法についてまとめる。さらに、嚙下検出の成否を判定、またその評価をする時に使用する正解嚙下時刻の取得方法も述べる。

### 3.1 RGB-D カメラ

本研究では、Intel社のRealSense F200を使用し、RGB画像と深度画像を同時に取得する。図13にRealSenseの外観を示す。RealSenseはstructured light方式[3]の赤外線センサであり、IR Lazer Projectorから照射されるパターンをIR Sensorで取得することで距離を求める。また、近距離での撮影が可能である。本研究では、参考文献[1]と同一のRealSenseを用いており、各種カメラパラメータの推定や妥当性の確認は行われているため、表3のカメラパラメータ値を用いる。

表 3 カメラパラメータ [1]

X 軸方向焦点距離 $f_x$ [pixel]	610.95
Y 軸方向焦点距離 $f_y$ [pixel]	611.04
X 軸方向光軸中心 $c_x$ [pixel]	324.14
Y 軸方向光軸中心 $c_y$ [pixel]	247.92

### 3.2 撮影と記録

本節では RGB カメラと深度カメラの解像度とフレームレートの設定、および撮影した画像の記録方法について述べる。

深度画像のみでは認識する部位の検出が難しく、その部位の深度値の取得が困難となる。そこで、認識する部位を RGB 画像で検出し、深度画像からその部位の深度値を取得する。よって、画像の撮影時は RGB 画像および深度画像の同時取得を行う。RGB 画像と深度画像の画像解像度は共に  $640 \times 480$  に設定する。また、30fps を RGB カメラと深度カメラの両方のフレームレートとした。RGB 画像と深度画像はそれぞれ別のバイナリファイルへ書き込み、画像の撮影タイミングのフレームタイムも同ファイルに書き出す。フレームタイムは高分解能カウンタの QueryPerformanceCounter 関数を用いて現在のカウント値を取得し、カウンタの周波数を取得できる QueryPerformanceFrequency 関数で割ることにより、高精度の時刻を取得する。

### 3.3 位置合せ処理

RGB カメラと深度カメラからそれぞれの画像を共に  $640 \times 480$  の解像度で取得している。しかし、RGB カメラと深度カメラには視野角に違いがある。また、図 13 より 2 つのカメラの設置位置が違うため、RGB 画像と深度画像は一致せずずれが生じる。このずれを取り除くため、RGB 画像と深度画像の位置合せ処理をする。深度画像の画素位置が RGB 画像の座標系においてどこに対応するかを取得し、深度画像にその座標  $(u'_c, v'_c)$  を付加する。この処理には Intel RealSense SDK に用意されている MapDepthToColor 関数を使用した。

### 3.4 撮影方法と撮影条件

本研究を行うにあたり、本学の倫理委員会の承認を得た。また、社会医療法人誠光会草津総合病院頭頸部外科の共同のもと研究を行っている。

撮影環境を図 14 に、撮影条件を表 4 に示す。

検査椅子に座した被験者の輪状軟骨と両鎖骨骨頭に黄色のマーカを貼付した。貼付した様子を図 1 に示す。頭はヘッドレストに軽く付けてもらう程度として、自然な嚥下を行えるようにした。被験者の喉頭と RealSense の距離は 300mm に固定し、RealSense が被験者の正面に位置するように移動させた。また、RealSense の仰角を表 4 の選択した撮影条件の仰角に設定した。表 4 の選択した撮影条件に

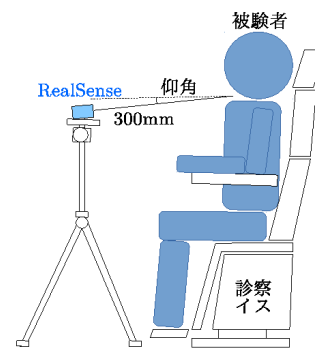


図 14 撮影環境

表 4 撮影条件

条件番号	内視鏡の撮影	仰角 [deg]
1	無	5
2	無	20
3	有	5
4	有	20

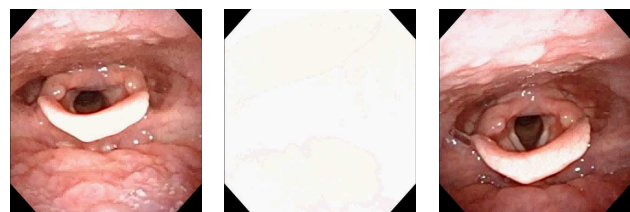


図 15 嚥下時の咽頭

において、内視鏡の撮影が有りの場合は、RealSense の撮影と同時に嚥下内視鏡画像の撮影を行った。撮影の際は合図を RealSense と内視鏡の両方で撮影し、時刻換算ができるようにした。嚥下食は実際の嚥下評価に使用されている水、とろみ、ゼリーを使用し、それぞれ 2 回ずつ被験者に与えた。それぞれの嚥下食の間は時間をとり、前後の嚥下で影響を与えないようにした。1 人の被験者に対して、表 4 に示す全ての撮影条件で撮影を行った。

### 3.5 正解嚥下時刻の取得

本節では正解嚥下時刻の取得方法について述べる。正解嚥下時刻は嚥下検出の成否とその評価をする際に使用する。

正解嚥下時刻は以下の 2 種類用意した。

**ホワイトアウト時間** 嚥下の様子を内視鏡で撮影している際に、嚥下ピーク時に起こる、映像が白くなるタイミング (図 15)。内視鏡の映像から該当する時間を取り出す。表 4 に示す内視鏡の撮影をしているデータで利用できる。

**最大挙上時間** 喉頭挙上運動で喉頭部が最大挙上しているタイミング。RealSense で撮影した RGB 画像から目視で全データ作成する。

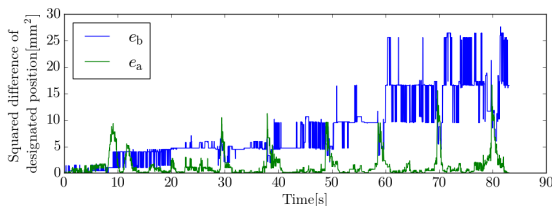


図 16 マーカ欠損がないデータの回転・並進補正前後の変化量

## 4. 結果

初めに回転と並進による補正の前後での結果を示す。その後、嚥下検出を評価する手法とその結果を示す。

### 4.1 回転・並進補正結果

撮影したデータそれぞれに対して、嚥下検出のアルゴリズムを適用した。回転・並進補正前後の変化量のグラフを図 16 に示す。図 16 では、補正前の変化量  $e_b$  は時間が経つにつれ、基準からの変化量が増加傾向にあるが、補正後の変化量  $e_a$  では時間が経つにつれ、変化量が大きくなる事はない。よって、グローバルモーションの補正ができていていると考えられる。また、補正後の変化量  $e_a$  では局所的に変化量が大きくなる区間が複数回みられる。この区間が嚥下によって変化量が増加する区間であると考えられる。

### 4.2 嚥下区間検出結果

喉頭隆起が不明瞭な被験者で撮影した2つのデータに対して、ローパスフィルタをかけた後の変化量  $e_a$  と嚥下区間検出結果を図 17 と図 18 に示す。図 17 は6回の嚥下全てを検出できたデータであり、各区間に各嚥下が対応している。一方で、図 18 は全ての嚥下が検出できなかったデータである。検出した区間をRGB画像で確認したところ、大きく口を開ける動作をしており、その動作に伴う輪状軟骨部の皮膚の移動を嚥下として検出している。

### 4.3 嚥下検出評価

アルゴリズムにより嚥下区間を検出した結果に対して、正解嚥下時刻により評価を行う。評価の際に使用する指標を次のように定義する。

#### True Positive(TP)

検出した嚥下区間に正解嚥下時刻を含む区間数

#### False Positive(FP)

検出した嚥下区間に正解嚥下時刻を含まない区間数

#### False Negative(FN)

検出できていない正解嚥下の回数

また、これらの指標を用いて再現率 (Recall) と適合率 (Precision) を式 (4) で定義する。再現率は総正解嚥下回数のうち正しく検出できた割合を示し、適合率は検出した嚥

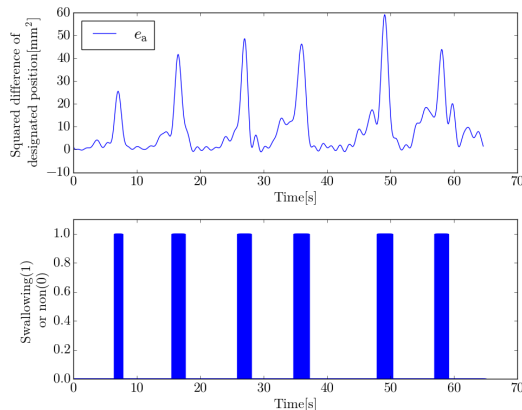


図 17 嚥下検出に成功したデータの輪状軟骨部の変化量と嚥下区間の検出結果

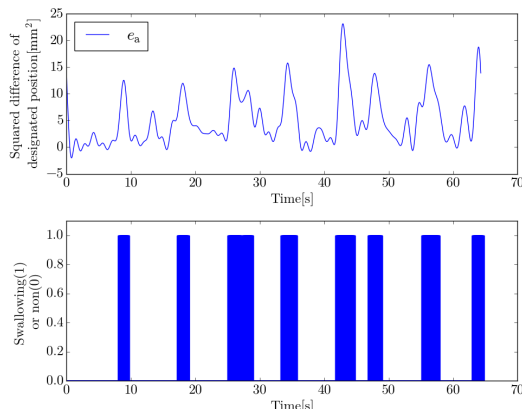


図 18 嚥下検出に失敗したデータの輪状軟骨部の変化量と嚥下区間の検出結果

下区間の総数のうち正しく検出できている割合を示す。

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (4)$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

正解嚥下時刻としてホワイトアウト時間と最大挙上時間を使用し、それぞれで評価を行う。本研究では7名の被験者に対し、表4に示す4つの条件で撮影をしたデータを用いた。アルゴリズムによる嚥下検出を評価した結果を表5に示す。表5より喉頭隆起の明瞭さに関係なく嚥下は検出できているが、喉頭隆起が明瞭な被験者に比べて不明瞭な被験者は評価値が低い。

再現率の評価値が低い理由としては、式(4)より嚥下を正しく検出できていないためである。これは喉頭隆起が不明瞭な被験者の嚥下時の頸部動作は輪状軟骨部だけでは変化量が小さく目立たないことが原因である。目視で嚥下動作を認識する際は頸部全体の動きに注目しているため、検出時は輪状軟骨だけでなく頸部全体に注目する必要があると考えられる。別の原因として、図19のように他動作を誤検出した際に最大変化量が大きくなり、嚥下を検出できないデータも挙げられる。

適合率の評価値が低い理由としては、式(4)より2つの



表 5 被験者の評価の平均結果

	人数	ホワイトアウト時間による評価				最大挙上時間による評価			
		データ数	嚥下回数	再現率	適合率	データ数	嚥下回数	再現率	適合率
喉頭隆起が明瞭	2	3	19	0.944	0.944	6	41	0.944	0.944
喉頭隆起が不明瞭	5	9	54	0.551	0.397	18	126	0.527	0.461
全被験者	7	12	73	0.658	0.546	24	167	0.631	0.582

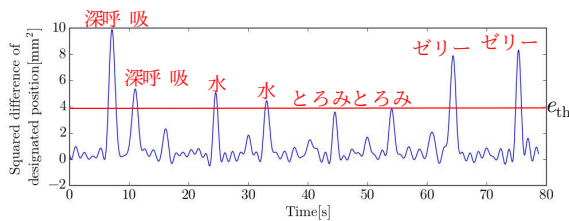


図 19 深呼吸動作による嚥下を検出できないデータの変化量グラフ ( $e_{th}$  は嚥下フレームを検出する際のしきい値)

理由を挙げられる。1 つ目は再現率と同様の理由で嚥下を検出できていないためである。2 つ目は検出区間を実嚥下数より多く検出しているためである。これは他動作を嚥下として検出しているためである。他動作と嚥下動作を区別することができれば、評価値の向上が見込める。

## 5. 結論

本研究では RealSense を用いて撮影した嚥下時の頸部の RGB 画像と深度画像から、喉頭隆起が明瞭、不明瞭に関係なく、健常者の嚥下を検出する手法を提案した。検出結果をホワイトアウト時間で評価した結果、再現率の平均は 0.658 となり、適合率の平均は 0.546 となった。最大挙上時間で評価した結果は再現率の平均が 0.631 となり、適合率の平均が 0.582 となった。また、喉頭隆起の明瞭さに関係なく、嚥下検出の可能性が確認できた。

## 参考文献

[1] 朝本崇康, 小澤恵子, 森谷季吉, 坂本真一, 酒井 道, 宮城茂幸: 深度画像による健常者嚥下時の喉頭挙上量計測の試み, 2017 年電子情報通信学会総合大会 情報・システム講演論文集 1, p. 71 (2017).

[2] 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-, 日本耳鼻咽喉科学学会会報, Vol. 95, No. 7, pp. 169-181 (1984).

[3] Hansard, M., Lee, S., Choi, O. and Horaud, R. P.: Time of Flight Cameras: Principles, Methods, and Applications, Springer Science & Business Media (2012).

[4] 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, Vol. 113, No. 103, pp. 33-38 (2013).

[5] 金谷健一: 幾何学と代数系 Geometric Algebra, 森北出版株式会社 (2014).

[6] 厚生労働省: 平成 29 年版厚生労働白書 資料編.

[7] 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会ポスターセッション予稿集 ISS-SP-219 (2015).

[8] 増山裕斗, 杉本千佳, 根本明宜: 嚥下機能評価のための喉頭領域抽出と PSO によるトラッキングを用いた喉頭挙上計測法, 2017 年電子情報通信学会総合大会 情報・システム講演論文集 1, p. 69 (2017).

[9] 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, Vol. 26, No. 2, pp. 163-174 (2006).

[10] 大類 孝: 超高齢化社会における誤嚥性肺炎の現状, 日本老年医学会雑誌, Vol. 56, No. 4, pp. 458-460 (2013).

[11] 才籾栄一, 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 (2007).

[12] 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, Vol. 43, No. 1, pp. 34-43 (2014).

[13] 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, Vol. 44, No. 1, pp. 92-97 (2015).

[14] 竹谷剛生, 古閑公治, 久保高明, 大塚裕一, 宮本恵美, 船越和美, 本木 実: Kinect を用いた健常者における非侵襲・非接触型嚥下機能評価法の研究, 保健科学研究誌, Vol. 14, pp. 103-113 (2017).

[15] Teramoto, S., Fukuchi, Y., Sasaki, H., Sato, K., Sekizawa, K. and 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*, Vol. 56, No. 3, pp. 577-579 (2008).