

深度画像から生成された頸部シルエットの画素数計測を用いた嚥下時刻推定 Estimation of swallowing time by counting pixels in the neck silhouette generated by depth images

浦西雄大*
Yudai Uranishi

小澤恵子*
Keiko Kozawa

酒井道*
Osamu Sakai

宮城茂幸*
Shigeyuki Miyagi

1. はじめに

近年、日本人の肺炎による死亡率が増加傾向にあり、現在では脳卒中に代わって日本における死因第3位となっている[1]。肺炎の原因の1つは高齢者の嚥下機能低下によって起こる誤嚥であり、肺炎の約70%は誤嚥性肺炎であると言われている[2][3]。したがって嚥下機能の評価により誤嚥の危険性を早期発見することが求められている。

病院で用いられている検査手法として、嚥下内視鏡検査(VE:Videoendoscopic examination of swallowing)や嚥下造影検査(VF:Videofluoroscopic examination of swallowing)がある。これらの手法はいずれも高価な医療機器を必要とし、検査を行える場所が限られている。またVFはX線照射による被爆を伴うため侵襲性がある。

非侵襲な検査手法としては、嚥下時の喉頭挙上運動を触診により判定する方法や、頸部聴診法を用いた嚥下音による評価手法も使われているが、医療知識を持つ熟練者でないと嚥下機能の評価が難しい。嚥下音を用いた評価手法の研究[4]も行われているが、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。

非侵襲、非接触かつ手軽に行える手法として、ビデオカメラにより嚥下の様子を外部から撮影し、嚥下に伴う喉頭隆起の挙上運動を検出し、その挙上量と挙上時間によって嚥下評価を行った研究[5][6]がある。しかし、この手法では喉頭隆起の不明瞭な被験者に対しての嚥下検出および嚥下評価が行えていない。

そこで本研究では嚥下に伴う頸部表面の形状変化に着目し、誤嚥の早期発見を補助するシステムの構築において必要となる嚥下動作の検出およびその時刻を推定する手法を提案する。健康者を対象として深度カメラにより外部から撮影を行い、嚥下に伴う頸部表面の形状変化を検出した。またその形状変化量から嚥下時刻の推定を行った。

2. 嚥下時刻推定のアルゴリズム

ここでは、撮影した深度画像から被験者頸部表面の形状変化を検出し推定嚥下時刻を求める手順について述べる。

2.1. アルゴリズムの概要

図1に、アルゴリズムの概要を示す。深度動画画像各フレーム上の赤枠で示した部分は頭頸部領域を表す。まず初期フレーム上で頭頸部領域のみ抽出する。その頭頸部領域から嚥下に伴う変化を観測するための計測領域を切り出す。計測領域画像を2値化し、次のフレイ

ムにおいて同様に処理された計測領域の2値化画像と排他的論理和を計算する。これにより嚥下に伴う頸部の形状変化が発生した部分のみが抽出される。この部分の面積を求め、時系列的に変化を観測することにより嚥下時刻を推定する。

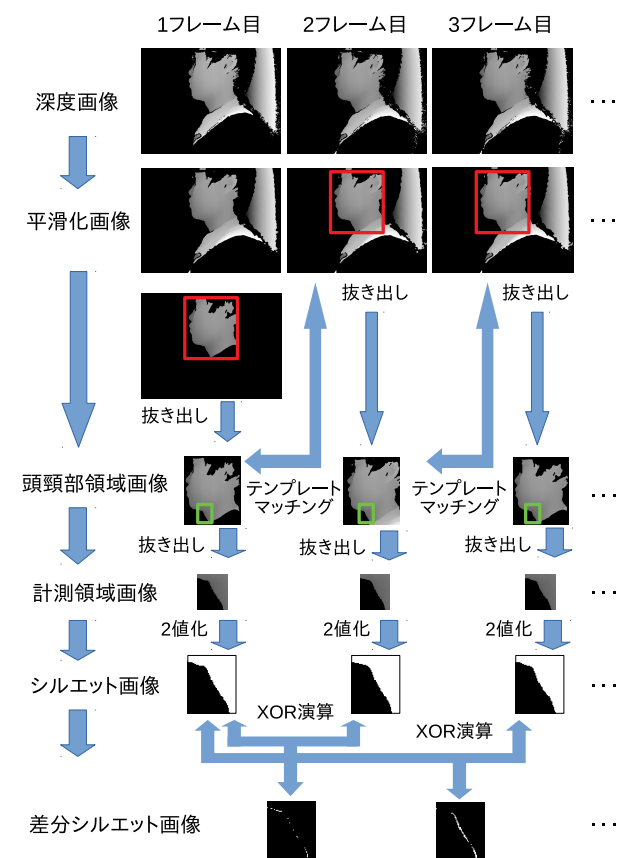


図 1: 嚥下時刻推定アルゴリズムの概要図

2.1.1. 頭頸部領域の抽出

カメラと被験者までの距離がほぼ一定であると仮定すると、深度画像中の頭頸部領域の深度値は、ある深度範囲に収まる。そこでこの条件を用いて深度画像内の不要な背景を除去する。深度画像を $z(x, y)$ とすると、頭頸部領域が含まれる深度画像 $z_h(x, y)$ は

$$z_h(x, y) = \begin{cases} z(x, y) & z_{\min} \leq z(x, y) \leq z_{\max} \\ 0 & \text{otherwise} \end{cases}$$

である。 $z_h(x, y)$ には、被験者の頭頸部と同じ深度範囲にある物体は画像中に残る。そこで画像内に残る物

*滋賀県立大学, The University of Shiga Prefecture

体の中で、頭頸部輪郭線が最も長いと仮定し、境界線追跡により得られた輪郭線の長さが最も長い物体が被験者の頭頸部であるとみなし、頭頸部領域画像 $z_H(x, y)$ を抽出した。

嚥下時に、嚥下がしやすいように姿勢を変化させる、または頭部を上下に振るなどの動作、すなわちグローバルモーションを伴う場合がある。そこで、2 フレーム目以降では、前のフレームから切り出した頭頸部領域画像 $z_H(x, y)$ をテンプレートとし、深度画像 $z(x, y)$ とのテンプレートマッチングにより頭頸部領域の位置を決定する。この際、エッジ部分の欠損の影響でマッチング結果にずれが発生する可能性があるため、あらかじめ $z(x, y)$ にメディアンフィルタを適用し、エッジ部分の平滑化処理を行った後にテンプレートマッチングを行う。マッチング率が最大となる位置を原点として、前のフレームから切り出した頭頸部領域画像 $z_H(x, y)$ を同じサイズの画像を切り出すことにより新たに頭頸部領域画像を得る。

2.1.2. 計測領域画像と嚥下時刻推定

1 フレーム目の $z_H(x, y)$ において、手動で指定した矩形領域を計測領域とし、計測領域画像 $z_T(x, y)$ を切り出す。2 フレーム目以降は $z_H(x, y)$ に対して相対的に同じ位置を計測領域とし順次切り出す。切り出した $z_T(x, y)$ を

$$s(x, y) = \begin{cases} 0 & z(x, y) = 0 \\ 1 & \text{otherwise} \end{cases}$$

と2値化処理し、シルエット画像 $s(x, y)$ を得る。1 フレーム目のシルエット画像 $s_1(x, y)$ と2 フレーム目以降のシルエット画像 $s_n(x, y)$ $n = 2, 3, 4, \dots$ と排他的論理和を計算することにより差分シルエット画像が得られる。嚥下によって計測領域の形状変化があった場合、差分シルエット画像中の画素数が増加すると考えられるので、この変化を観測することにより嚥下時刻が推定できる。後述の実験では、観測区間における画素数の最大値に対して90%の値をしきい値と設定し、しきい値を超えた区間を推定嚥下区間とした。

3. 実験方法

今回撮影に使用したカメラは Intel 社の Intel RealSense F200 である。認識範囲は20cm から120cm までで、解像度 640×480 、フレームレート 30fps と設定している。

図2に撮影環境を示す。20代の男性健常者5名を対象とし、嚥下食としてスプーン1杯分の水を用意した。

被験者には椅子に座ってもらい、被験者の頸部側面から距離 d の位置に三脚台に取り付けた RealSense を配置した。RealSense の仰角は 0° とし、頸部表面の映り方の違いによる嚥下時刻の推定結果を比較するために撮影距離 $d=200\text{mm}, 300\text{mm}$ の2通りに設定した。撮影開始前、被験者には水を口に含んだ状態で壁に頭を付けるように指示した。また撮影距離が目的の距離となるように RealSense の位置調整を行った。そして撮影

開始後、撮影者の合図によって嚥下を行ってもらい、撮影を終了した。1回の撮影につき嚥下は1回のみとし、前後の嚥下に影響が無いよう適度な時間を空けてから撮影を行った。また嚥下の前後でのグローバルモーションをなるべく小さくして撮影を行うために、被験者には頭部を壁に付けた状態で嚥下を行ってもらうように指示した。器具などを使用しての頭部の固定は、被験者が自然な嚥下を行えなくなることが想定されるため避けている。得られた深度画像に対して $z_{min}=30\text{mm}$ 、 $z_{max}=50\text{mm}$ の範囲の深度値を取得し、第2章で述べたアルゴリズムを適用した。

次に嚥下時刻から嚥下評価を行う方法について述べる。今回使用した RealSense では深度画像と RGB 画像を同時取得することが可能である。そこで取得した RGB 画像から嚥下の開始フレームおよび終了フレームを目視で確認し、そのフレーム番号を記録した。この嚥下開始から嚥下終了までの区間を正解嚥下区間とした。正解嚥下区間内に推定嚥下区間が含まれる場合に嚥下検出成功とみなし、検出に成功した嚥下区間数を記録した。

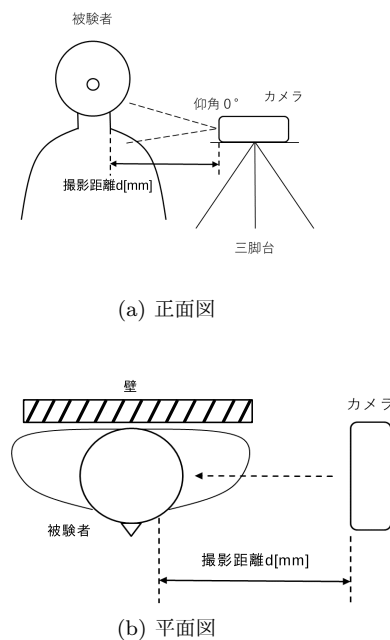
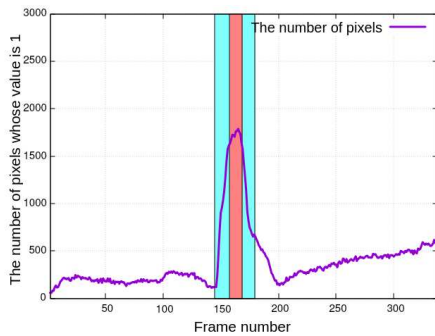


図 2: 撮影環境

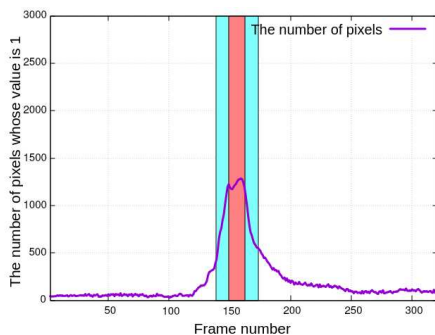
4. 嚥下時刻の推定結果

図3から図5は3名の被験者について撮影した深度画像に本手法を適用し、得られた計測領域シルエット画像の画素値が1である画素数を全フレームについて計測した結果である。ただしグラフ中の青で示す区間は正解嚥下区間、赤で示す区間は推定嚥下区間をそれぞれ表している。また被験者全5名の実験結果から推定嚥下区間数、正解嚥下区間数および検出に成功した正解嚥下区間数を算出した結果は以下の表1のようになった。

推定した区間はそのほとんどが正解区間内に含まれていたが、図4のように1つの正解嚙下区間内に複数の推定嚙下区間を検出するケースが見られた。また嚙下の最中やその直前、直後にグローバルモーションを含む被験者については嚙下検出に失敗するケースが見られた(図5)。



(a) 撮影距離 200mm



(b) 撮影距離 300mm

図3: 被験者Aの計測結果

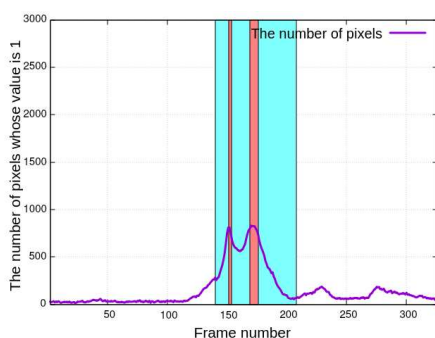


図4: 被験者Bの計測結果(撮影距離 300mm)

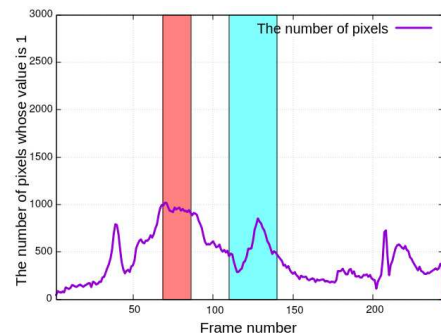


図5: 被験者Cの計測結果(撮影距離 200mm)

表1: 嚙下区間数の計測結果

撮影距離	200mm	300mm
推定嚙下区間数	13	14
正解嚙下区間数	10	9
検出した正解嚙下区間数	10	9
再現率	1.000	1.000
適合率	0.769	0.643

5. 考察

5.1. グローバルモーションの除去手法について

本手法ではテンプレートマッチングを用いて嚙下と同時に起こるグローバルモーションの除去処理を行った。この処理により平行移動の動作は打ち消す事ができている。しかし除去処理の対象としている嚙下時の姿勢変化や頭部の上下動作は、平行移動だけでなく回転移動も合わさった動作となっている。そのため頭部の固定を行わない条件下で本手法を適用する場合、完全にグローバルモーションを打ち消すことはできないと考えられる。計測領域の位置は取得した頭頸部画像に対して相対位置で決定しているため、頭頸部領域画像が回転したまま取得すると測定領域が正しく取得できず、嚙下の誤検出の原因となる(図6(c))。図4のように正解嚙下区間内で複数の推定嚙下区間がみられたのはこの事が原因と考えられる。したがって回転移動にも対応したアルゴリズムを取り入れ、グローバルモーションの除去手法を見直す必要がある(図6(d))。

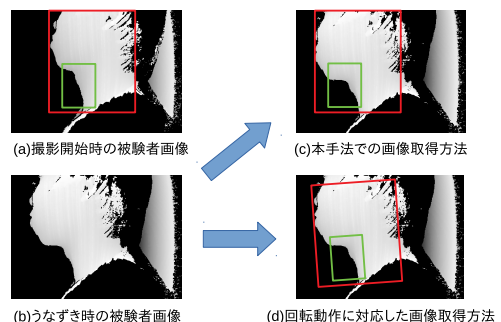


図6: 回転動作の除去処理

5.2. 頭頸部追跡失敗時の処理について

本手法ではテンプレートマッチングによるグローバルモーションの除去を行っているが、この処理は隣りあうフレームについて、前のフレームで取得したテンプレートを更新しながらのマッチングによって行っている。そのため、あるフレームで頭頸部の追跡に失敗し頭頸部と異なる領域をテンプレートに更新した場合、それ以降のフレームについては誤ったテンプレートでのマッチングが行われることとなる。

今回使用した画像データでは頭部を壁に付けた状態から撮影を開始し、不要な背景を除去した後にマッチングを行っているため、誤ったマッチング結果となるケースは見られなかった。しかし今後頭部固定を行わない条件下での適用を想定し、マッチング失敗時の補正処理も加える必要があると考えられる。

5.3. 追跡可能範囲について

今回使用したアルゴリズムでは、 n フレーム目の頭頸部領域画像をテンプレート画像として $(n+1)$ フレーム目の画像にマッチングを行い、マッチングにより得られた頭頸部領域画像をテンプレート画像に更新して $(n+2)$ フレーム目の画像にマッチングを行っている。テンプレート画像の大きさは撮影開始フレームの画像から目視で決定し、撮影終了フレームまで不変としている。2 フレーム目以降のテンプレート画像は、マッチング率が最大となる点 (x, y) を原点としてテンプレート画像と同じ大きさの矩形領域画像を抜き出すことで取得している。

この時マッチング点 (x, y) を取得することのできる座標範囲は、探索用画像の領域内に限られている。そのため頭頸部領域が画角から外れた場合には、頭頸部領域を正しく取得することができない。例として撮影距離を 200mm に設定した場合、被験者の頭頸部が一部画角から外れた状態から撮影が開始されるため、マッチングに失敗し頭頸部領域を正しく取得できないケースが見られた。したがって撮影時に被験者の頭部全体が画角内に収まるような適切な撮影距離に設定する、または頭頸部領域が画角から外れた場合の追跡手法について見直す必要があると考えられる。

追跡手法の改善案として、取得するマッチング点を 2 点に増やす方法が考えられる。本手法ではマッチング点を 1 点のみの取得とし、取得する画像の大きさも不変としていた。しかしマッチング点を左上と右下の 2 箇所取得し、その 2 点で囲まれる矩形領域画像を取得すれば、被験者が画角から外れた場合でも追跡が可能になると考えられる。ただしこの方法で取得する頭頸部画像の大きさはフレームごとに変化すると考えられるので、頭頸部画像から計測領域画像を取得する際には取得する位置や大きさの観点で考慮する必要がある。

6. 結論

本研究では、深度カメラより取得した深度画像から嚥下による頸部表面の形状変化を検出することにより、嚥下時に頭部を固定している条件下での嚥下動作時刻の推定が可能であることが確認できた。今後は頭部の固定を無くした撮影データについても検討し、より複

雑なグローバルモーションに対応したアルゴリズムを構築していく。また嚥下評価を画像のみで行うことは困難であるため、本手法により得られた嚥下時刻に対応する嚥下音の解析を行い、喉頭隆起の不明瞭な被験者にも適用可能な嚥下評価手法を検討していく。

謝辞

本研究の一部は JSPS 科研費 JP18K45678 の助成を受け行われた。また、滋賀県立大学工学部の地域ひと・モノ・未来情報研究センターの研究活動として行われ、内閣府の地方創生推進交付金による支援を受けた。

参考文献

- [1] 厚生労働省, "我が国の人口動態平成 30 年," pp.17-18, ISSN 1345-5222, 2018.
- [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 -健康成人を対象として-, " *昭和歯学会雑誌* 26(2), pp.163-174, 2006.
- [5] 高橋明子, 清水良昭, 北邦宏, 仲筋宣子, 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.
- [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*, 43(1), pp.92-97 2015.