

GMM を用いた反復唾液嚥下テスト自動評価システム

山下 大貴[†] 古川 大輔[‡] 村西 幸代[‡] 森野 智子[§]
 黒岩 眞吾[⊥] 綱川 隆司[†] 西田 昌史[†] 西村 雅史[†]

[†]静岡大学 大学院総合科学技術研究科 [‡]君津中央病院 リハビリテーション科

[§]静岡県立大学短期大学部 歯科衛生学科 [⊥]千葉大学 大学院工学研究院

1. はじめに

高齢者の肺炎の主な原因として誤嚥が挙げられる [1]. 誤嚥とは, 気管内に固形物や流動物など異物を誤って飲み込む, または吸い込むことであり, 嚥下機能が低下するほど誤嚥の可能性が高まる. 一方, 嚥下機能低下の早期発見は誤嚥の予防につながるため, 地域や病院などで定期的実施されている嚥下機能検診はその重要性を増している.

本稿では, 嚥下機能検査の 1 つである反復唾液嚥下テスト (RSST) の自動測定について, 性能改善を行ったのでその結果を報告する.

2. 先行研究とその課題

嚥下機能の検査方法としては, オーラルディアドコネシス (OD) や最大発声持続時間 (MPT), RSST など, 様々な方法が提案されている [2].

そこで, 我々は咽喉マイクロフォンとスマートフォンからなる簡単な装置で, OD, MPT 及び RSST の自動測定が可能な嚥下機能検査システムを開発した [3][4]. しかし, RSST では十分な性能を得ることができなかった. RSST とは, 30 秒間にできるだけ多く唾液を飲み込み, 飲み込んだ回数を測定する検査方法である.

なお, RSST については検診担当者が触診で得た嚥下回数を記録する装置 [5] が実用化されているものの, 自動測定ではない.

3. 検診データの詳細分析

今回, 先行研究 [4] で用いた 2 回の嚥下検診で収録されたデータ (69 名, イベント総数 402) について, 波形の視察と聴取により, 検診の収録データに含まれる音イベントの分析を行った. 2 回の嚥下検診の収録データに含まれた音イベントの内訳を表 1 に示す. 表 1 における喉の収縮は, 嚥下しようとするが, うまく飲み込めない際に発生する音である. 内訳を見ると, 嚥下の回数は全体の約 6 割を占めるが, それ以外の音も約 4 割含まれており, それらが先行研究における RSST の精度を下げる原因となっていた.

Automatic evaluation system for RSST using GMM
 D.Yamashita[†], D.Furukawa[‡], S.Muranishi[‡], T.Morino[§],
 S.Kuroiwa[⊥], T.Tsunakawa[†], M.Nishida[†], M.Nishimura[†],
[†]Graduate School of Integrated Science and Technology, Shizuoka University, [‡]Department of Rehabilitation, Kimitsu Chuo Hospital, [§]Department of Dental Hygiene, Shizuoka Prefecture University Junior College, [⊥]Graduate School of Engineering, Chiba University

表 1 2 回の嚥下検診収録データに含まれた音イベントの内訳

音イベント	嚥下	声	息
総数	245	52	51
音イベント	咳	喉の収縮	その他
総数	15	35	4

4. 提案手法

初めに, 検診データから, パワーに基づいた VAD (Voice Activity Detection) [4] を用いて音区間を検出する. その際, なるべく音イベントを漏らさないような設計を行う.

次に, VAD によって検出された音イベントから, 嚥下音のみを抽出するために, イベント単位で特徴量を抽出し, GMM (Gaussian Mixture Model) を用いて音イベントの識別を行う. 喉の収縮は, 嚥下音と音響的特徴が似ており, 人間が聞いても判断が難しい喉の収縮音が存在した. そのため, 喉の収縮音の学習モデルを構築すると, 逆に嚥下音の一部を誤って喉の収縮音と識別してしまい, 表 1 のように音イベントの約 6 割を嚥下音が占める場合においては精度が悪化する可能性がある. したがって, 嚥下, 声, 息, 咳, 喉の収縮音の 5 種類からなる学習モデル群と, 喉の収縮音を除いた 4 種類からなる学習モデル群を構築し, 性能の比較を行う. 図 1 に識別までのデータ処理の流れを示す.

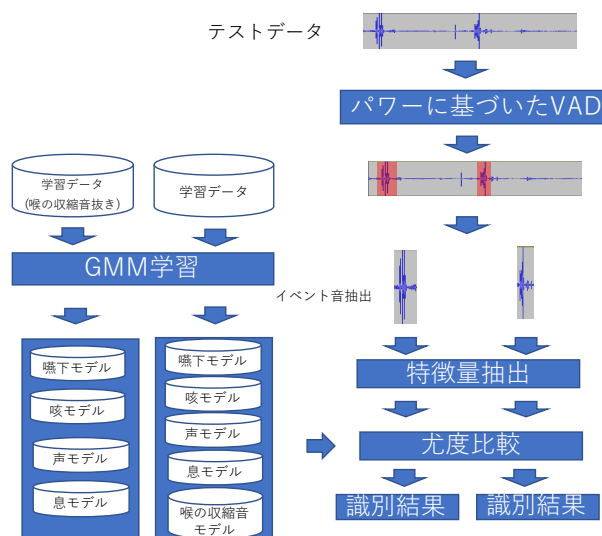


図 1 識別までのデータ処理の流れ

5. 識別実験

提案手法を基に識別実験を行った。今回、提案した2種類の学習モデル群の性能をそれぞれ評価し比較する。

5.1 実験データ

使用するデータは、2回の嚙下検診で収録されたデータ(データセットA)を用いる。ただし、これだけではデータ数が足りないため、嚙下音を中心に追加でデータ収集(20名)を行なった(データセットB)。収集したデータの内訳を表2に示す。

表2 データセットBのデータの内訳

音イベント	嚙下	声	息
総数	695	343	210
音イベント	咳	喉の収縮	
総数	384	227	

5.2 予備実験

識別実験を行う前段階として、音イベントを漏らさないことを重視したパワーに基づくVADで、音区間の検出を行った。その結果、表1のデータに対する嚙下音イベントの再現率は約96%となった。そして、このVADを用いて検出された音イベントに対し識別実験を行った。

5.3 実験方法

GMMを用いて学習モデルの構築、識別を行なった。今回、2つの学習モデル群を用意する。

- ・モデル群A 嚙下・声・息・咳
- ・モデル群B 嚙下・声・息・咳・喉の収縮

特徴量はMFCC12次元+パワー、またその1次差分、2次差分の計39次元をフレームごとに算出した後、イベント単位で平均した値を用いた。また、GMMの混合数は32とした。テストデータはデータセットAの被験者1名、学習データはテストデータを除くデータセットAとデータセットBを用いた。そして、データセットAに対して交差検証を行い、2つのモデル群に対して性能を評価した。

5.4 実験結果

実験結果を図2に示す。モデル群A、モデル群Bの再現率は、VADによる嚙下音の検出漏れを含めた結果となっている。また、VADのみの数値は先行研究[4]における結果である。まず、適合率はモデル群Aよりも喉の収縮音を含めたモデル群Bの方が約0.03高かった。次に、再現率はモデル群Aの方がモデル群Bに比べ約0.08高かった。そして、F値はモデル群Aの方がモデル群Bよりも約0.02高くなり、VADのみと比べると、約0.13向上していることが分かる。

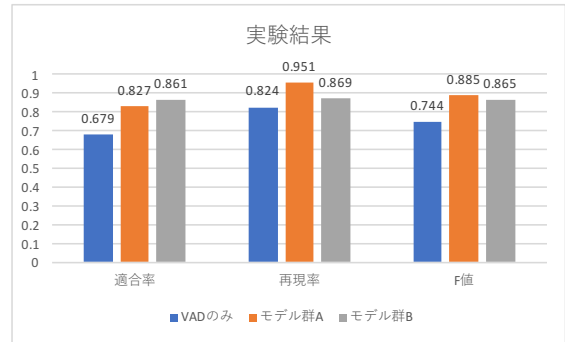


図2 実験結果

5.5 考察

まず、適合率においてモデル群Aがモデル群Bよりも低くなった理由は、喉の収縮音のモデルを構築しなかったことによって、喉の収縮音の約9割が嚙下音として識別されていたためである。喉の収縮音のモデルを含めたモデル群Bでは、喉の収縮音の約4割が正しく識別されていた。

次に、再現率について、モデル群Aでは、VADによる嚙下音の検出漏れの数を除いた場合、再現率は約0.99であり非常に高い結果となった。一方、モデル群Bでは、VADによる嚙下音の検出漏れの数を除いた場合、再現率が約0.91となった。識別結果を見ると、誤認識の多くが喉の収縮音として認識されていた。RSSTにおける音イベントのうち、多くが嚙下音であるため、誤認識の数がモデル群Aよりも多くなり、その結果モデル群Aの方がモデル群Bに比べF値が若干高くなったと考えられる。

6. おわりに

我々が開発した嚙下機能検査システムにおいて、精度の低かったRSSTの性能改善を行なった。その結果、先行研究[4]に比べF値が約0.13向上した。今後、現状では嚙下音として認識されている喉の収縮音を識別できるようにし、RSSTの更なる性能改善を図りたい。

謝辞

本研究の一部はJSPS科研費(16K13028)の助成を受けた。

参考文献

- [1] 大類孝, 海老原孝枝, 荒井啓行: 高齢者肺炎・誤嚥性肺炎, 日本内科学会 第99巻 第11号 pp 2746-2751 (2010)
- [2] 嚙下障害支援サイト スワロー -< <http://www.swallow-web.com/rehabilitation/evaluation.html>>(2018年1月6日取得).
- [3] 山下大貴・他: “咽喉マイクを用いた嚙下機能検査システムの開発”, 第14回情報学ワークショップ, B-37X(2016).
- [4] 山下大貴・他: “咽喉マイクを用いた嚙下機能検査システムの性能改善に関する検討”, 第79回全国大会講演論文集, 2017(1), pp577-578(2017)
- [5] 竹井機器工業株式会社 | 高齢者健康増進・介護予防関連 | 口腔機能:<<http://www.takei-si.co.jp/productinfo/detail/61.html>>(2018年1月10日取得)