

生体音と環境音の同時収録による 高齢者の行動および身体状態認識に関する検討*

○西村雅史，小林悠一，桐山伸也，峰野博史（静岡大学大学院・情報）

1 はじめに

高齢者に適切な生活支援や質の高い介護サービスを行うため，個々の高齢者の日常行動や心身の状況に関する情報の重要性が高まっている。我々は高齢者への負担が小さく，かつ簡便なシステムで実現が可能な方法として，「生活行動に伴って生じる音」から日常行動，身体状態を認識・分析する手法を検討している。特に，通常集音マイクによって収録される環境音に加えて咽喉マイクによって収録される生体音を新たな情報源とすることで，日常会話などの知的活動の検出や，摂食・嚥下等の身体状態の認識を行なう方法を検討したので報告する。

2 行動音データの収録方法

2.1 行動音データ収集の問題点

ライフログの研究などではセンサー情報の一つとして周囲の環境音や音声を収録することが多いが，その情報が十分に活用されてきたとは言いがたい。特に通常マイクでは対象者に関して音声以外の生体音情報を得ることは出来ないし，その音声も多くの場合口元から遠く離れた位置で周囲の人の発話や環境音と共に収録されることになる。結果としてノイズや反射等の影響を強く受けるため，音声認識技術などの適用は極めて困難であり，十分な情報を収集することが出来なかった。

2.2 咽喉マイクによる収録

本研究では上記のような問題点を考慮し，周囲ノイズの影響を受けにくいマイク構成として，咽喉マイクと通常集音用ピンマイクの併用を検討した（Fig.1）。咽喉マイクは高騒音下の会話用に開発された接触型のマイクで，咽喉付近の皮膚の振動を音として収集するものである。このため収録音の帯域はおおよそ 4KHz 以下となり，こもった音に感じら

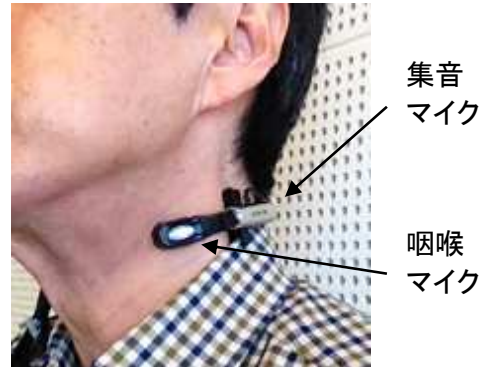


Fig.1 咽喉マイクと集音マイクの同時装着の様子
A man wearing throat and lavalier microphones

れるが，発話内容はおおむね問題なく理解できる。また，Fig.1 から分かるようにネックバンド型の形状をしており，装着は容易で，外れにくい。装着者の負担も比較的小さいので，就寝時を含む高齢者の長時間の装着にも適している。また，唾や飲食物を飲み込む際に咽喉付近で発生する嚥下音（=喉頭蓋の開閉音など）は高齢者の身体状態を知る上で重要な情報であることが知られており，医療分野でも咽喉マイクによる計測が行われている。このように，周囲ノイズの影響を受けにくい音声データの収集に加えて，嚥下音を中心とした生体音を収集出来るという特徴を持つ。

同様の目的に使用されるマイクとしては骨伝導マイクが知られているが，通常，イヤホンかヘッドセットの形状をしているため，長時間の装着が困難な場合がある。また，意図せず外れてしまうことも多いため，我々の目的には必ずしも適していない。

2.3 2 マイクによる同時収録の意義

集音マイクと咽喉マイクの両者を用いることで，嚥下などの生体音が収集出来ることに加え，特に音声について以下の効果が期待される。1) 本人を交えた会話か，他人の一方的な発話か，本人の独語かの識別精度の向上。2) 咽喉マイクの信号を用いた集音マイク側

* A study on combined use of throat and lavalier microphones for recognition of daily activity and physical condition of aged people, by M. Nishimura, Y. Kobayashi, S. Kiriya and H. Mineno (Shizuoka University). E-mail: nisimura@inf.shizuoka.ac.jp

の信号からの雑音除去^[1]。特に1)は高齢者の知的活動を検出するために大変重要である。つまり、従来の集音マイクのデータでは音声を検出できたとしても、それが認知症の特徴でもある“独語”なのか、テレビ視聴などの状況なのかを識別することは不可能、もしくは大変困難であった。一方、咽喉マイクを併用することで、それぞれのマイクで検出された音声区間の重なりによって会話とこれらの状況を従来よりも簡単に区別することが出来る。Fig. 2 に、実際の会話音声を集音マイクと咽喉マイクによってステレオ録音した波形と、咽喉マイク側データによる発話区間検出 (VAD) 結果の例を示す。

3 咽喉マイク収録音による行動認識

咽喉マイクによる収録音で、どのような身体状態の認識が可能かを調べるため、予備的な検討を行った。

3.1 対象となる行動

事前検証で検出の可能性が考えられた「深呼吸」「咀嚼」「咳 (咳払い)」「嚥下」「発話」の5つの行動を実験の対象とした。なお、「嚥下」には食物摂取、水分摂取に加えて唾を飲み込む行為(「空嚥下」)を含む。

3.2 実験データ

被験者6名(成人男性)が咽喉マイクを装着して上記の5種類の行動を繰り返し行い、その際に発生する音を収録した。これをVADによって自動分割した後、人手によってそれぞれの検出区間に対して正解行動ラベルを付与した。正解行動ラベルを付与したイベントの総数は1,622である。なお、被験者によっては深呼吸の収録音が微弱で、そもそもVADによって検出されないケースがあったが、これらは認識対象に含めていない。

3.3 実験結果

識別にはSVM(ガウシアンカーネル)を用い、各被験者の1名の全データを評価、残りの5名の全データを学習用とし、交差検証で評価した。なお、特徴量としては20次元のメルフィルターバンク出力を用いている。結果をTable 1に示す。また、被験者ごとの各行動に対するF-値をFig.3に示す。発話や咀嚼の識別については比較的安定して高いF-値が得られたが、嚥下、咳については被験者による性能のばらつきが大きい。特に嚥下の識別

においては被験者Bの結果が悪く、嚥下と発話の誤判別が高頻度で起きていた。また、深呼吸についてはVADでの検出に加えて識別も困難であることが分かった。

4 おわりに

日常行動、身体状態を簡便に観測するための手段として咽喉マイクと集音マイクを併用する方法について検討を行った。今後は高齢者の実際の日常行動に伴う音の収集・分析を進めたい。

参考文献

- [1] Zhang, Z., et al. Multi-sensory microphones for robust speech detection, enhancement and recognition. ICASSP2004. p. 781-784.

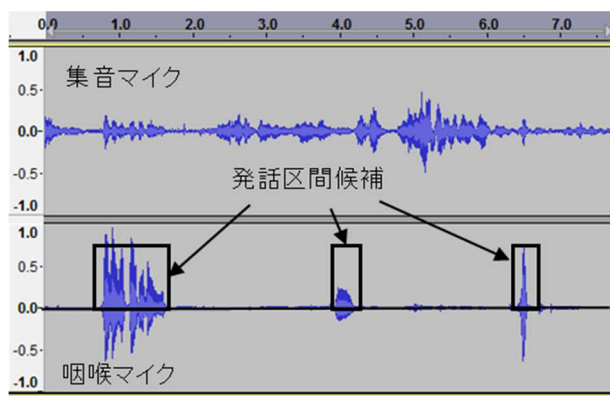


Fig.2 2チャンネル収録による発話検出の例
An example of 2 mic stereo-recording and VAD

Table 1 行動認識結果 (Recognition results)

行動の種類	Precision	Recall	F-measure
深呼吸	0.37	0.39	0.38
咀嚼	0.76	0.84	0.80
咳	0.74	0.75	0.74
嚥下	0.73	0.65	0.69
発話	0.90	0.92	0.91

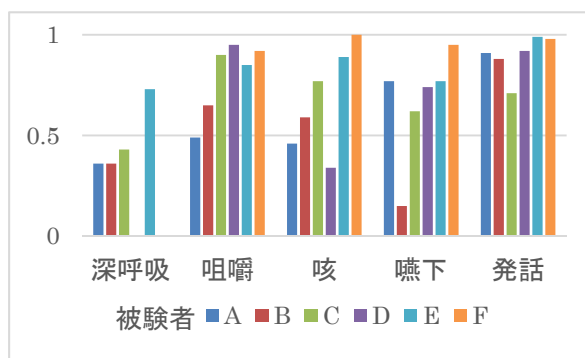


Fig.3 被験者別および行動別のF-値
F-measures for each subject and each activity