

サポートベクターマシンを用いたいびき音/非いびき音の識別

苫小牧工業高等専門学校 ○荒木 毅, 三上 剛, 小島洋一郎, 国立病院機構函館病院 米澤一也

要旨

本研究は、睡眠時に録音した音声データからいびき音を自動抽出する手法を提案することを目的とする。いびき音/非いびき音の各周波数成分に基づき 10 次元特徴ベクトルを作成し、サポートベクターマシンで識別した結果を報告する。

1. はじめに

睡眠時無呼吸症候群(Sleep Apnea Syndrome:SAS)は、睡眠時に一定の無呼吸(あるいは、低呼吸)を伴い、日中の過度の眠気や睡眠中の窒息感や喘ぎ、反復する中途覚醒等の症状を引き起こす病である。¹⁾

現在、SAS の診断には睡眠ポリグラフ検査が行われている。これは、患者の検査入院を必要とし患者に大きな負担をかける検査である。そのため、睡眠ポリグラフ検査よりも簡便にできる検査方法として、いびき音の音声解析による方法に注目が集まっている。²⁾しかし、これまでは後述する Sleep Sound データからいびき音を手作業で切り出し、解析を行っていたため膨大な労力が必要であり、比較的小規模なデータの分析しか行われてこなかった。

そこで、睡眠中のいびき音全体を解析するためには、いびき音を自動的に抽出するという前処理技術が必要になると考えた。

本研究では、いびき音を自動抽出するうえで必要とされるいびき音と非いびき音の識別を行った結果を報告する。

2. 使用するデータ

2.1. Sleep Sound データ

本稿では、睡眠中の人の周囲の音を録音した約 8 時間の音声データを Sleep Sound データと呼ぶ。

本研究で扱っている Sleep Sound データは、3 名の男性 OSAS 患者が国立病院機構函館病院に検査入院した際に録音したものである。録音は、携帯型リニア PCM レコーダーを用いてベッドの支柱にマイクを固定し、サンプリング周波数 11kHz、量子化 bit 数 16bit の音声データとして録音した。

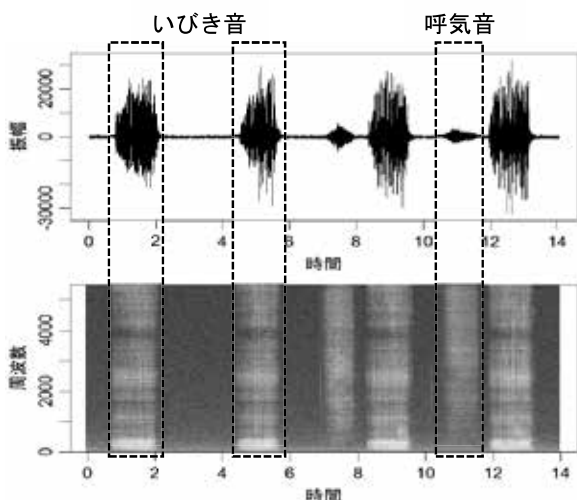


図1 Sleep Sound データの波形とスペクトrogram

また、本稿で頻繁に扱う言葉の定義について確認する。まず、“いびき音”と“非いびき音”とはそれぞれ一定以上の音量でいびき音、またはいびき音以外の音が録音されている信号区間のことを指す。また、“いびき候補音”とは一定以上の音量でいびき音または非いびき音が録音されている信号区間を指す。以降、本稿ではこれら 3 つの言葉をこのような意味で使用する。

図 1 の波形からわかるように、Sleep Sound データの中にはいびき音だけではなく、普通の呼吸音など非いびき音が含まれている。しかし、図 1 のスペクトrogramを見るといびき音と非いびき音の特徴が周波数成分の違い(スペクトrogramの濃淡)としてよく表れていることがわかる。

2.2. いびき音/非いびき音のデータ

まず、前述した Sleep Sound データに対し、3.1.にて説明するいびき候補信号の切り出し処理を行い 8150 個のいびき候補信号を用意した。次に、いびき候補信号の検聴および波形の観測により、いびき音が非いびき音かのラベル付を行い、いびき音 5430 個と非いびき音 2720 個に分別した。

本稿では、この分別したデータをトレーニングデータおよびテストデータとして扱い、10 分割交差検証を用いていびき音/非いびき音の識別性能の評価を行った。

3. 手法

3.1. いびき候補信号の切り出し

まず、Sleep Sound データからいびき音の候補となるいびき候補信号を次のような手順で抽出する。³⁾

(1) エネルギーの計算

Sleep Sound データから 100[ms]の区間毎にエネルギー E_k を計算する。エネルギー E_k は式(1)で計算し、区間は 50%オーバーラップさせる。

$$E_k = \sum_{i=0}^{N-1} s_k^2[i] \quad (1)$$

ここで、 k は区間の番号、 N は k 番目の区間内のサンプル数、 $s_k[i]$ は k 番目の区間内 i サンプル目の振幅を示す。

(2) しきい値を用いたいびき候補信号の切り出し

Sleep Sound データ全体からエネルギー E_k を計算した後、 E_k の最小値から 2 つのしきい値 T_{E1} 、 T_{E2} を用意する。

$$T_{E1} = a \times \min(E_k) \quad (2)$$

$$T_{E2} = b \times \min(E_k) \quad (3)$$

ここで、 a 、 b は経験的に定める定数を示す。($a < b$)
各区間毎のエネルギー E_k としきい値 T_{E1} を比較し、 E_k が T_{E1} より大きければ、その区間を候補区間とし区別する。その後、各候補区間におけるエネルギー E_k の最大値としきい値 T_{E2} の値を比較し、 E_k が T_{E2} より大きい場合にその候補区間をいびき候補信号として切り出す。

3.2. 特徴ベクトル作成

次に、切り出したいびき候補信号を表す特徴ベクトルを作成する。ここでは、従来の研究で扱っていた方法と今回新たに導入した方法の2つを紹介する。

(1) 方法1

切り出したいびき候補信号の0~5[kHz]に渡るスペクトログラムを求める。スペクトログラムを500[Hz]のサブバンドごとに10分割し、各サブバンドの周波数成分を要素にもつ10次元特徴ベクトル($\xi^k = [\xi_1^k \ \xi_2^k \ \dots \ \xi_{10}^k]^T$)を作成する。また、特徴ベクトルの各要素は、スペクトログラム全体の値で正規化したもので、式(4)で示す。

$$\xi_i^k = \frac{\sum_{j=1}^{N_k} \sum_{f=500(i-1)}^{500i} |y(j, f)|^2}{\sum_{j=1}^{N_k} \sum_{f=0}^{5000} |y(j, f)|^2} \quad i = 1, 2, \dots, 10 \quad (4)$$

ここで、 $y(j, f)$ はハンニング窓を用いた短時間フーリエ変換をいびき候補信号に行い求めたスペクトログラムの値、 N_k はk番目のいびき候補信号に含まれる候補区間の個数を示す。

作成した特徴ベクトルを3.3.にて紹介する識別方法でいびき音/非いびき音に識別し、各いびき候補信号の識別結果とする。

(2) 方法2

前述した方法1では、いびき候補信号1つに対し特徴ベクトル1つを作成していた。方法2では、いびき候補信号を100[ms]の区間毎に分割し、各区間のスペクトログラムの値から方法1と同様に、特徴ベクトルを作成する。この結果、いびき候補信号1つに対し複数の特徴ベクトルが作成される。

各いびき候補信号の識別は、一旦、1つのいびき候補信号を構成する複数の特徴ベクトルを3.3.にて紹介する識別方法でいびき音/非いびき音に識別した後、その複数の特徴ベクトルの識別結果に対し多数決を行うことで各いびき候補信号の最終的な識別結果とする。

3.3. いびき音/非いびき音の識別

いびき音/非いびき音の識別手法として、k近傍法(KNN)とSupport Vector Machine(SVM)を採用し識別を行う。また、識別結果の検証方法として、10分割交差検証を採用する。

4. 結果と考察

3.2.の特徴ベクトル作成の手順にて、方法1により特徴ベクトルの作成を行ったうえでk近傍法を用いた識別結果を図2に示す。図2は、k近傍法のkの値を1~29まで変化させたときのいびき音/非いびき音の識別率を表している。

方法1により特徴ベクトルを作成しSVMを用いて識別した結果を表1に示す。表2は、SVMのカーネル関数をrbfカーネル、コストパラメータであるCの値を0.001, 0.01, 0.1, 1, 10, 100, またカーネルのパラメータであるsigmaの値を0.01, 0.02, ..., 0.1, 0.2, ..., 1.0, 2.0, ..., 10, 20, ..., 100というように設定した際に得られた良好な結果を示したものである。方法2により特徴ベクトルの作成を行ったうえでSVMを用いて識別した結果を表2に示す。表記は、表1と同様である。

図2と表1の結果より、SVMを識別に用いたことで非いびき音の識別率の向上が見られる。しかし、表2からわかるように方法2を採用した場合は期待したような識別率の向上は見られなかった。

いびき音を解析することを視野に入れた際、いびき音を非いびき音に識別してしまう(解析対象を減らす)ことよりも、非いびき音をいびき音に識別してしまう(いびき音でないものを解析対象に入れる)ことのほうが望ましくないと考えられる。つまり、いびき音よりも非いびき音の識別率が高いほうが望ましいといえ、このような観点から今回のk近傍法とSVMの識別結果を比較すると、識別方法としてはSVMが有用であると考えられる。

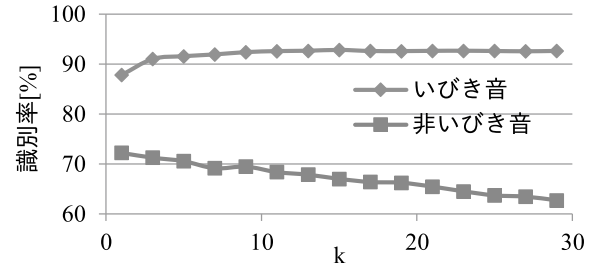


図2 k近傍法による識別結果

表1 SVMによる識別結果(方法1)

	C	sigma	識別率[%]	
			いびき音	非いびき音
1	100	2	89.9	79.0
2	100	1	89.9	74.0
3	100	0.9	90.2	73.6
4	100	0.8	90.5	73.0
5	100	0.7	90.6	72.1

表2 SVMによる識別結果(方法2)

	C	sigma	識別率[%]	
			いびき音	非いびき音
1	100	0.4	92.0	59.4
2	100	0.7	94.3	56.5
3	100	0.5	93.5	57.2
4	100	0.6	93.5	57.2
5	100	0.3	92.4	58.0

5. まとめ

本稿で示した結果としては、SVMを識別に採用したことで、非いびき音の識別率を向上させることができた。しかし、いびき候補信号を表す特徴ベクトルの数を増やし細分化して識別率を上げる試みでは、まだよい結果は得られていない。今後は、非いびき音の識別率の向上に重きを置き、新たな改良を行っていく。

謝辞

本研究は、日本学術振興会 科学研究費補助金 基盤研究(C)による研究助成(課題番号:26330338)によって行われたものである。

参考文献

- 1) 太田, いびきと睡眠障害, 東海大学出版会, pp.49-52, 2005
- 2) 榎本他, 雑音耐性に優れたいびきのホルマント周波数解析に基づく閉塞型睡眠時無呼吸症候群と単純いびき症との識別, 生体医工学, vol.48, no.1, pp.115-121, 2010
- 3) M. Cavusoglu, et al, "An efficient method for snore/nonsnore classification of sleep sounds", Physiol.Meas. vol.28, no.8, pp.841-853, 2007