

## 複数の聴診箇所の聴診音から選択した学習データを用いての異常肺音検出 Abnormal lung-sound detection using training data selected among respiration samples recorded from multiple auscultation points

梅野 直幹<sup>†</sup>  
Naoki Umeno

松永 昭一<sup>†</sup>  
Shoichi Matsunaga

高田 寛之<sup>†</sup>  
Hiroyuki Takada

### 1. はじめに

肺疾患の検査は医療従事者のみが行うことができ、病院へ行くことが困難な人は疾患の発見が遅れる。我々は、肺疾患を自動で識別し、早期発見につなげることを目的とし、肺疾患に起因して発生する異常肺音の自動検出を目指して研究を行っている。

これまで、呼吸音の音響分析に基づき、異常音を特定するための研究が行われてきた[1-2]。その中で我々は聴取した肺音データについて、隠れマルコフモデル (HMM) を利用した音響モデルを学習し、最尤法に基づく異常肺音の識別を行った[3-4]。肺音データは複数の聴診箇所から聴取しているが、聴診箇所ごとに肺音の音響特徴が異なるため、評価する肺音データと同じ聴診箇所のみデータを用いて音響モデルを学習していた。また、より多くの肺音データを収集するには、医師や患者の協力が必要となるため大規模なデータベースを整備することは困難である。そのため少ないデータ数で学習を行うこととなり、認識精度が低い原因になることが考えられる。

本稿では、評価箇所と異なる箇所の肺音データにも学習に有効なデータが存在すると考え、複数の聴診箇所のデータを利用して学習データを拡張する二つの手法を提案した。学習データの拡張を行い、より多くのデータを用いて学習を行うことで、異常音の識別精度向上を目指す。

### 2. 肺音データ

肺音データは、電子聴診器を使用して肺患者と健常者について図 1 の 13 の聴診箇所から録音したものを利用した。各被験者のそれぞれの聴診箇所毎に 1 つの肺音データが記録されている。肺患者から記録したデータ数は 605 で、健常者から記録したデータ数は 837 であり、1 つのデータにつき平均で約 10 回分の呼吸吸気が含まれている。

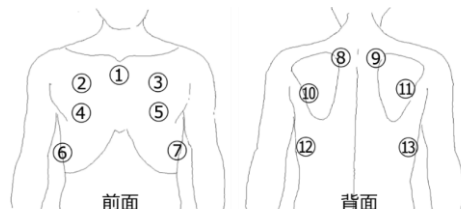


図 1 肺音データを収録した聴診箇所

### 3. 異常肺音識別の流れ

学習データ拡張手順を含む、異常肺音識別の流れを図 2 に示す。この節では既存手法である異常肺音識別の流れについての説明を行い、次節で学習データ拡張についての説明を行う。

識別は学習と評価の 2 つの過程で行う。学習の過程で、

<sup>†</sup>長崎大学大学院工学研究科 School of Engineering, Graduate School of Engineering Nagasaki Univ.

あらかじめ肺音データのラベル付けを行った学習データをもとに音響特徴を抽出し、HMMを用いた音響モデルを作成する。評価の過程で、肺音データから特徴抽出を行った評価データについて異常音と正常音の識別を行う。作成したモデルを用いて正常音としての尤度と異常音としての尤度を算出し、これらの尤度の高い方を識別結果として出力する。また、この一連の流れについて既存手法 (Baseline I)[4-6] では、ある特定の聴診箇所のみについて学習と評価を行っており、その他の聴診箇所の肺音データは利用していない。

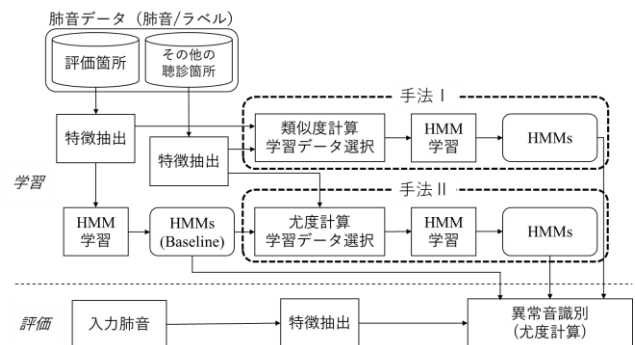


図 2 学習データ拡張の手順と異常肺音識別の流れ

### 4. 学習データ拡張

高精度な音響モデルを作成するために、評価箇所とは異なる複数の聴診箇所のデータを利用して学習データを拡張する手法を提案する。それらの箇所のデータには、評価箇所の肺音と音響特徴が大きく異なるデータも存在するために、高精度な音響モデルを作成する際の妨げになる。そのため、学習データとして適切なデータを選択する必要があると考えられる。本稿では二つの選択手法を提案する。

#### 4.1 選択手法 I

一つ目は肺音データにおける音響特徴量の類似性を考慮した選択手法である。これは評価箇所の肺音データと音響特徴量が類似している肺音データを学習に用いることで識別精度が向上するという考えに基づく。肺音データの音響特徴量としてメルケプストラム係数 (MFCC) とパワーを利用し、評価箇所の肺音データ分布と評価箇所と異なる箇所の肺音データとの類似性を求める。類似性の計算には、評価箇所の肺音データの分布と他の聴診箇所の肺音データとの分布間の距離を求めるためにバタチャリヤ距離[5]を用いた。この距離が小さいほど同様の音響特徴がある。類似性があらかじめ設定した閾値  $\alpha$  以上であれば、評価箇所のデータと類似しているものと考え、そのデータを学習データとして採用する。閾値  $\alpha$  が小さいほどより高い類似度が必要となる。

## 4.2 選択手法 II

二つ目は候補データの正常と異常の尤度差による信頼性を考慮した選択手法である。これは尤度差が大きいほど正常音と異常音を明確に区別することができ、このようなデータを学習に用いることで識別精度が向上するという考えに基づいている。

まず、評価箇所と異なる聴診箇所の肺音データである候補データについて、一つ一つの呼吸ごとに正常と異常の尤度を計算し、その二つの尤度差  $d$  を計算する。この尤度差  $d$  について事前に定義した閾値  $D_T$  よりも大きい場合、候補データを学習データとして選択する。 $D_T$  が大きいほどより高い信頼性が必要となる。

## 5. 評価実験

### 5.1 実験条件

音響モデルは、3 状態 2 混合の HMM を利用し、音響特徴量には MFCC (6 次元) とパワー (1 次元) を利用した。評価には、図 1 の④の異常肺音 254 個、正常肺音 245 個を用いた。学習にはそのデータセットに加え、他の聴診箇所から学習データに選択した肺音データを用いた。評価には leave one out 法を用いており、評価対象のデータは学習に用いていない。

提案手法の比較対象として 2 種類の Baseline を準備した。一つは図 2 に示した通り評価箇所と同一箇所みの肺音データを利用し学習を行った音響モデル(Baseline I)。もう一つは、全ての聴診箇所からの全ての肺音データを利用し学習を行った音響モデルである(Baseline II)。また、4.1、4.2 の手法に従い作成した 2 種類の音響モデルを準備した。この 4 種類の音響モデルを用いて異常肺音識別を行い、それぞれについての比較を行う。

### 5.2 実験結果

4 つの音響モデルを利用したときの識別率と、既存手法から拡張したデータ数を呼吸数単位でそれぞれ表 1 に示す。手法 I と手法 II は識別率が最も高かった場合の結果を表示した。これらは、閾値  $\alpha$  及び  $D_T$  を変化させることにより取得した。

Baseline II の識別率は 86.2% であり、Baseline I の 88.6% から 2.4 ポイント低下した ( $p = 0.048$ )。この結果から、他の聴診箇所の肺音データを全て利用しての学習データ数の拡張は識別精度向上につながらないことが分かった。

一方、手法 I の 91.1% と手法 II の 90.7% の識別率は、Baseline I の識別率から向上した。(それぞれ  $p = 0.10$  及び  $p = 0.12$ )。

次に、手法 I と手法 II についての比較を行う。それぞれの拡張数における識別率を図 3 に示す。二つの手法で同じ拡張数を確保するのは困難なため、横軸は手法 I についての拡張数を表示している。聴診箇所④についての評価を行う際、手法 I では 621 の呼吸数を拡張して 91.1% の識別率が得られ、手法 II では 518 の呼吸数を拡張して 90.7% の識別率が得られた。この結果より、手法 I の方が手法 II より高精度な識別ができる可能性があることを示した。しかし、手法 II において、正常と異常を正しく分類することのできた全ての呼吸を拡張データとして利用した場合 ( $D_T = 0$  で拡張数約 4700) でも識別率が 90.6% と、高い精度を維持することができたのに対して、手法 I では拡張数が 4624

のとき識別率が 81.3% と大幅に低下した。この結果から、手法 II の方が、閾値の変動に対して頑健であることがわかる。また、拡張数が多い場合の手法 I の識別率が Baseline II を下回った原因として、異常肺音と正常肺音の学習データ数に偏りがあったことが考えられる。異常肺音は患者の一部の呼吸にしか含まれていないが、手法 I は類似度が高いデータから被験者毎の全ての呼吸を学習データとして選択している。そのため異常肺音に比べ正常肺音のデータ数が著しく多くなり、識別率が低下したことが考えられる。

表 1 異常肺音の識別率とそれぞれの学習データ拡張数

手法	正常[%]	異常[%]	平均[%]	拡張数
Baseline I	86.6	90.6	88.6	0
Baseline II	86.6	85.8	86.2	5460
手法 I	87.8	94.5	91.1	621
手法 II	87.8	93.7	90.7	520 ~ 4720

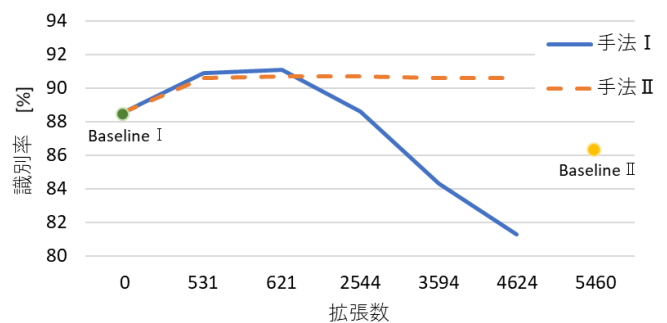


図 3 それぞれの拡張数における識別率

## 6. まとめ

本稿では、評価箇所と異なる聴診箇所の肺音データから学習に適切な肺音データを選択することによる、2 種類の学習データの拡張手法を示した。1 つは音響特徴量の分布の類似性に基づき、もう 1 つは正しく分類された正常音と異常音の尤度差による信頼性に基づき選択を行った。2 種類の提案手法は Baseline よりも高精度な結果を得ることができ、学習データの拡張が有効であることを示した。また、音響特徴量の類似性を考慮した手法 I は手法 II よりも高い識別精度を得ることができたが、拡張数によって識別率の変動が小さかった点から、閾値の変動により頑健であるのは手法 II であることが分かった。事前に設定する必要のある閾値に対してよりロバストである手法 II は手法 I より実用的であると考えられる。

### 参考文献

- [1] Y. P. Kahya, S. Yere and O. Cerid, "A wavelet-based instrument for detection of crackles in pulmonary sounds," *Proc. of IEEE EMBS*, pp. 3175-3178, 2001.
- [2] M. Bahoura and X. Lu, "Separation of crackles from vesicular sounds using wavelet packet transform," *Proc. of IEEE ICASSP*, II, pp. 1076-1079, 2006.
- [3] S. Matsunaga, et al., "Classification between normal and abnormal respiratory sounds based on maximum likelihood approach," *Proc. of IEEE ICASSP*, pp. 517-520, 2009.
- [4] M. Yamashita, et al., "Discrimination between healthy subjects and patients with pulmonary emphysema by detection of abnormal respiration," *Proc. of ICASSP*, pp.693-696, 2011.
- [5] Guy B. Coleman, Harry C. Andrews, "Image Segmentation by Clustering," *Proc IEEE*, Vol.67, No.5, 1979, pp. 773-785.