

環境音認識のための最尤状態数の検討

岡本 亜紗子^{1,a)} 林田 亘平¹ 中山 雅人² 西浦 敬信²

概要: 近年, 安心・安全な生活環境を実現するために, 防犯システムが広く用いられている. 特に住居への不正侵入や異常事態の検出を目的として, 監視カメラが使用されている. しかし, 監視カメラは暗所や死角においての異常検出が困難であるという問題が存在する. この問題を解決するために, 日常生活において発生する環境音を利用した異常音検出システムが提案されている. 環境音を認識し, 音事象を特定することで, 異常音を検出することができるため, 即座に住居の管理会社や警備員に異常事態を報知することができる. 従来, 環境音認識法として, Hidden Markov Model(HMM)を用いた手法が広く研究されてきた. HMMは時間変化に伴うスペクトルの変動を統計的にモデル化したものであり, HMMの状態数は信号列の状態遷移に関わる重要なパラメータである. 環境音は短時間に減衰する単発音や長時間持続する音, 定常な音や非定常な音など, 環境音によって継続長や定常性などの音響的特徴が大幅に異なるため, HMMの最尤状態数は環境音によって異なると考えられる. しかし, 従来法では, 各音響モデルは全て同じ状態数で構築されていたため, 認識性能が劣化する問題があった. そこで本稿では, 環境音の認識性能向上を目指し, 環境音の音響的特徴に基づき最尤状態数を推定し, 各環境音で異なる状態数の音響モデルを用いて環境音を高精度に認識する手法を提案する. 評価実験の結果, 提案法の有効性を確認した.

1. はじめに

近年, 安心・安全な生活環境を実現するために, 防犯システムが広く用いられている [1]. 特に住居への不正侵入や異常事態の検出を目的として, 監視カメラが使用されている [2]. 監視カメラの映像を人手で監視し, 異常時に住居の管理会社や警備員に異常事態を報知することで, 必要に応じて現場に人を派遣することが出来る. しかし, 監視カメラは暗所や死角での異常検出が困難であるという問題が存在する. この問題を解決するために, 環境音を利用した異常音検出システム [3] が提案されている. 環境音とは日常生活において発生する音であり, 扉の開閉音や電話の呼出音など様々な音が存在する. これらの環境音を認識し, 音事象を特定することによって, 異常事態の検出が可能となり, 監視カメラでは検出できない暗所や死角における異常事態も検出可能となる.

従来, 環境音認識法として, Hidden Markov Model (HMM) を用いた手法が広く研究されてきた [4]. HMMは時間変化に伴うスペクトルの変動を統計的にモデル化したものであり, HMMの状態数は信号列の状態遷移に関わる

重要なパラメータである. 環境音は短時間に減衰する単発音や長時間持続する音, 定常な音や非定常な音など, 環境音によって継続長や定常性などの音響的特徴が大幅に異なるため, HMMの最尤状態数は環境音によって異なると考えられる.

しかし, 従来の環境音認識法 [5] では, 各音響モデルを全て同じ状態数で構築して認識を行っていた. そのため, 認識性能が劣化する問題があった. そこで本稿では, 環境音の時間変化に伴うスペクトルの変動を表す特徴に基づき最尤状態数を推定し, 各環境音で異なる状態数の音響モデルを構築することにより, 環境音の認識性能向上を検討する.

2. 従来の環境音認識法とその問題点

幅広い環境音の認識を試みた従来研究として, 三木らによる HMM を用いた環境音認識の検討がある [5]. 三木らは, 様々な環境音が存在する実環境下において, 音声認識を行うために, HMMによる環境音の認識と, 環境音と音声の識別を行った. 三木らは, モデル学習のため環境音のラベル付けを行い, それぞれの環境音を個々にモデル化した. 従来法の概念図を図 1 に示す.

従来研究では, HMM を用いて環境音をモデル化し, 音響モデルと入力環境音との尤度を計算することで, 認識を行う. この際, 各環境音の HMM は全て同じ状態数で構築していた. しかし, 環境音は手を叩く音とブザー音のよう

¹ 立命館大学大学院 情報理工学研究科 〒525-8577 滋賀県草津市野路東 1-1-1

² 立命館大学 情報理工学部 〒525-8577 滋賀県草津市野路東 1-1-1

a) is0009sv@ed.ritsumei.ac.jp

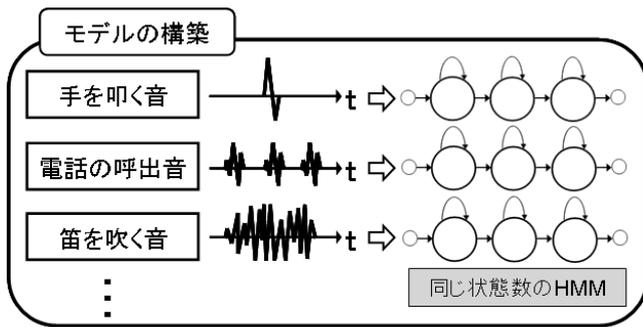


図 1 環境音認識における従来法概念図

Fig. 1 The outline of the environmental sounds recognition in the conventional method.

に単発音と連続音や、空調の音と紙を破る音のように定常な音と非定常な音が存在し、環境音によって HMM の最尤状態数は異なると考えられる。そのため、従来研究では、各環境音に対して HMM の状態数を全て同一で構築することで、認識性能が劣化する問題があった。

そこで、本稿では、各環境音に対して同じ状態数の音響モデルを構築するのではなく、各環境音の音響的特徴に応じて最尤状態数を推定し、各環境音で異なる状態数の音響モデルを構築することで、環境音の認識性能向上を図る。

3. HMM の最尤状態数推定に基づく環境音認識法の提案

本稿では、環境音の音響的特徴に基づく HMM の最尤状態数の推定による環境音認識法を提案する。最尤状態数の推定に用いる環境音の音響的特徴として、各環境音の時間波形に対する Δ パワーの変曲点数に着目する。 Δ パワーの変曲点数は環境音のパワーの変動を表しており、環境音の継続長や定常性を表すことができる。継続時間が長い音や非定常な音は、 Δ パワーの変曲点数が多いことが予想でき、単発音など継続時間が短い音や継続時間が長くても定常な音は、 Δ パワーの変曲点数が少ないことが予想できる。以上のような特徴から、各環境音の最尤状態数推定に Δ パワーの変曲点数を用いる。図 2 に提案法である最尤状態数推定に基づく環境音認識法概念図を示す。まず、構築する音響モデルの各音源に対する Δ パワーの変曲点数を算出する。そして、算出した変曲点数を基に、各環境音の最尤状態数を推定する。この最尤状態数推定式算出のために、あらかじめ予備実験を実施し、 Δ パワーの変曲点数と最尤状態数の関係を調査する。そして、その結果を基に最尤状態数推定式を算出する。本稿では、この最尤状態数推定式の算出において、対数近似と線形近似をそれぞれ用いた 2 つの手法を検討する。算出した最尤状態数推定式を基に推定した状態数の音響モデルを構築し、認識を行うことで、環境音の認識性能向上を目指す。また、対数近似と線形近似をそれぞれを用いた 2 つの最尤状態数推定法の認識性能

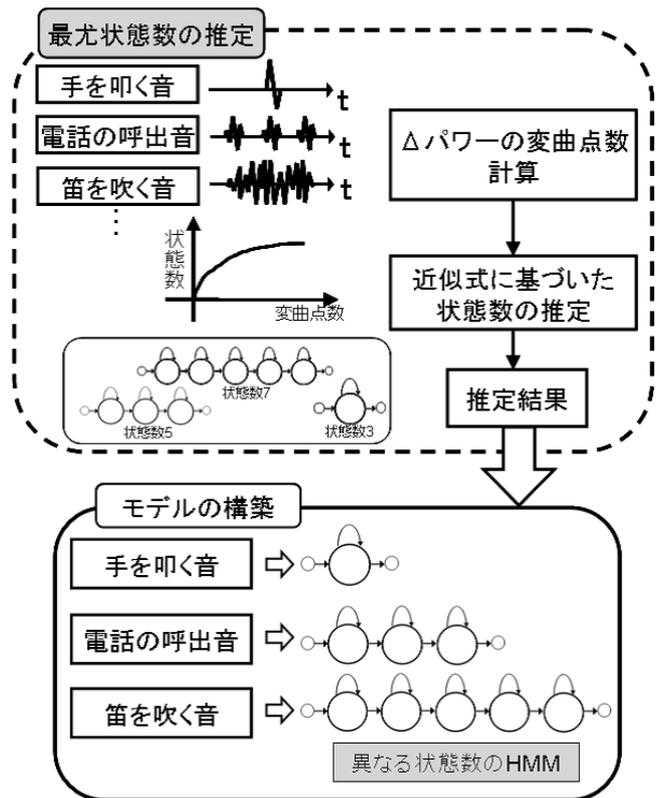


図 2 環境音認識における提案法概念図

Fig. 2 The outline of the environmental sounds recognition in the proposed method.

を比較する。

3.1 環境音の時間波形における Δ パワーの変曲点数の算出

本稿では、最尤状態数を推定するための環境音の特徴として、環境音の時間変化に伴うスペクトルの変動を考慮するために Δ パワーの変曲点数に着目する。環境音のパワーの変動を表す Δ パワーは式 (1) によって算出する。

$$d_t = \frac{\sum_{\theta=1}^{\Theta} \theta \times (\text{power}_{t+\theta} - \text{power}_{t-\theta})}{2 \times \sum_{\theta=1}^{\Theta} \theta^2}, \quad (1)$$

ここで、 d_t は時刻 t における Δ パワー、 power_t は t 番目のフレームのパワーであり、 Θ は Δ パワーを算出する際、考慮する前後のフレーム数である。本稿において、 Θ は 2 に設定したため、時刻 t における Δ パワーは、 $(t+2)$ 番目のフレームから $(t-2)$ 番目のフレームまで、計 5 フレームにおけるパワーの差の加重平均となる。

つまり Δ パワーの変動が大きいと、環境音のパワーが時間によって大きく変動することを表し、 Δ パワーの変動が小さいとパワーの変動が小さいことを示す。この Δ パワーの変動を表す値として、変曲点数を利用した。ここで、変曲点数はゼロ交差法 [6] により算出する。 Δ パワーの変曲点数は、環境音のパワーの時間変動を表す指標となる。例えば、木で板を叩くような単発音はブザー音のような連続音と比べ Δ パワーの変曲点数が少ないことが予想できる。

表 1 予備実験の実験条件

Table 1 Preliminary experimental conditions.

| | |
|---------|---|
| データベース | RWCP-DB (92 種類, 9,700 サンプル) |
| 標本化周波数 | 16 kHz |
| 量子化ビット数 | 16 bit |
| モデル数 | 38 種類 |
| 特徴量 | MFCC 12 次元 + パワー 1 次元 + Δ MFCC 12 次元 + Δ パワー 1 次元 + $\Delta \Delta$ MFCC 12 次元 + $\Delta \Delta$ パワー 1 次元 |
| 状態数 | 3~15 状態 (left-to-right) |
| 学習データ | 2,660 サンプル (38 種類 \times 70 サンプル) |
| テストデータ | 1,140 サンプル (38 種類 \times 30 サンプル) |

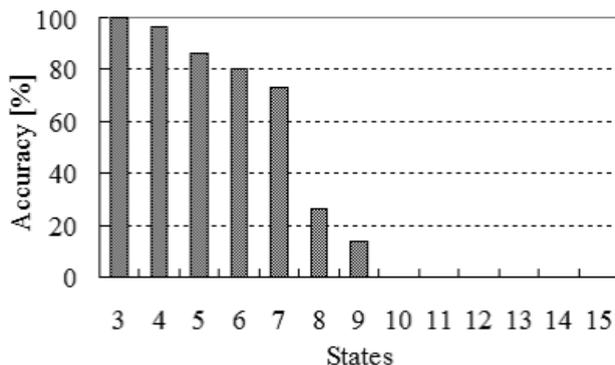


図 3 各状態数における手を叩く音の認識結果

Fig. 3 Recognition accuracy of hand clap on each state.

また、紙を破る音のような非定常な音は、空調の音のような定常な音と比べ、 Δ パワーの変曲点数が多くなることが予想できる。以上のような特徴から、 Δ パワーの変曲点数は各環境音の最尤状態数と相関があると考えられるため、最尤状態数推定のための特徴として Δ パワーの変曲点数を用いる。

3.2 最尤状態数推定式算出のための予備実験

各環境音の最尤状態数と Δ パワーの変曲点数の関係を調査するために、予備実験を行った。各環境音に対して3~15状態の音響モデル13種類を構築し、環境音ごとに認識を行った。そして、各状態数の音響モデルでの認識率を算出することで、各環境音の最尤状態数を調査した。音響モデルの学習にはRWCP-DB[7]の環境音38種類各70サンプルを用いて、テストには各30サンプルを用いた。

認識した38種類の環境音から、手を叩く音の各状態数の認識結果を図3に示す。横軸は状態数、縦軸は認識率を表す。本稿において、各環境音の最尤状態数は、最も高い認識率を得た状態数とした。また、最も高い認識率を得た状態数が複数存在する場合は、それら状態数の平均値とした。これは、状態数を推定する際に、多少の推定誤差が発生しても認識率を大幅に低下させないためである。例えば、図3の手を叩く音においては最尤状態数は3となる。

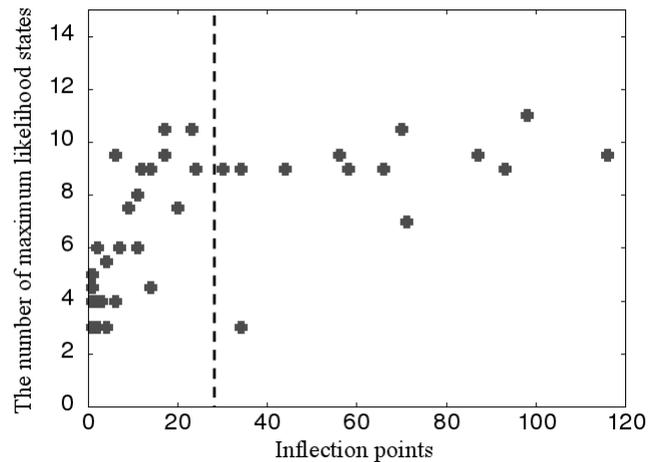


図 4 各環境音における変曲点数と最尤状態数

Fig. 4 The number of maximum likelihood states and inflection points on each environmental sound.

以上の結果に基づき、環境音38種類の最尤状態数と Δ パワーの変曲点数の関係を調査した。最尤状態数と Δ パワーの変曲点数の関係を図4に示す。各環境音の Δ パワーの変曲点数は、各環境音の学習データである各70サンプルの平均値とする。横軸は Δ パワーの変曲点数、縦軸は最尤状態数を表す。また、各点が各環境音の最尤状態数と Δ パワーの変曲点数の対応を表す。図4から Δ パワーの変曲点数が少ない環境音においては最尤状態数が少なく、 Δ パワーの変曲点数が増加するにつれて最尤状態数も増加し、図4の点線に示すように、変曲点数が30あたりから最尤状態数が飽和する傾向を確認できた。したがって、この傾向を利用することで、 Δ パワーの変曲点数を基に最尤状態数を推定できると考えられる。

3.3 Δ パワーの変曲点数に基づく最尤状態数推定式の算出

Δ パワーの変曲点数と最尤状態数の関係を基に、音響モデルの最尤状態数を推定し、構築することで認識性能の向上を目指す。本稿では、最尤状態数推定式の算出において、2つの手法を検討する。まず1つ目が、 Δ パワーの変曲点数と最尤状態数の関係から、対数近似を用いて近似曲線を算出する手法である。この手法では、予備実験で使用した全環境音を対象に近似曲線を算出する。もう一方が、線形近似を用いて近似直線を算出する手法である。この手法では変曲点数が少ない環境音を対象に近似直線を算出する。両手法を用いて算出した近似式を基に、各環境音の状態数を推定し、評価実験を行う。

3.3.1 対数近似を用いた最尤状態数の推定法

予備実験において算出した各音源の最尤状態数と各音源の Δ パワーの変曲点数との近似曲線を、式(2)の対数近似式より算出する。予備実験に使用した全環境音を対象に近似曲線を算出する。

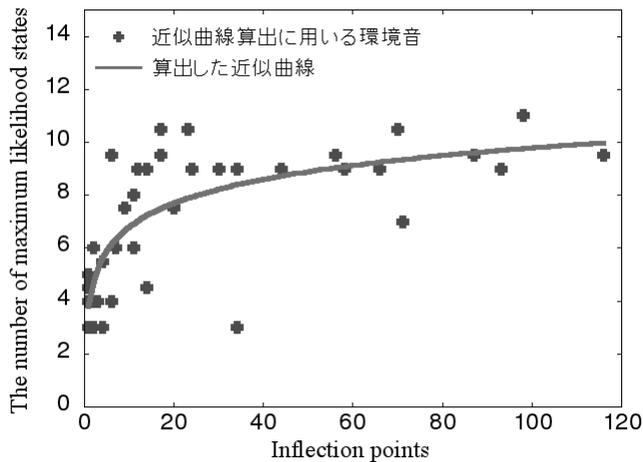


図 5 変曲点数と最尤状態数の近似曲線 (変曲点数 0 から 120 まで)
Fig. 5 Regression curve of inflection points and the number of maximum likelihood states.

$$y = (a \times \log x) + b, \quad (2)$$

ここで、 y は状態数、 x は Δ パワーの変曲点数を示す。近似曲線を算出した結果、 $a = 1.3$ 、 $b = 3.8$ の値を得られた。算出した近似曲線を図 5 に示す。横軸は Δ パワーの変曲点数、縦軸は最尤状態数を表す。図 5 中の●印は近似曲線算出に用いる環境音を示す。算出した近似曲線を基に、 Δ パワーの変曲点数から最尤状態数を算出し、各環境音に対する最尤状態数の音響モデルをそれぞれ構築することで、認識性能の向上が期待できる。

3.3.2 線形近似を用いた最尤状態数の推定法

図 4 より、点線で示す変曲点数以降の状態数が飽和している環境音は、状態数による認識率の変化がないことが確認できる。そのため、それらの環境音における最尤状態数の推定精度が低下した場合でも、認識性能への影響は少ないと考えられる。そこで、本研究では最尤状態数が飽和している環境音は考慮せず、変曲点数が少ない環境音を用いて最尤状態数推定式を算出する場合も検討した。実験的に決定した閾値 30 を基に、変曲点数が 30 までの環境音を最尤状態数推定式の算出に使用した。変曲点数が 30 以下の場合、 Δ パワーの変曲点数と最尤状態数が線形の関係を持つ傾向が確認できるため、式 (3) の線形近似式を用いて近似直線を算出する。

$$y = (a \times x) + b, \quad (3)$$

ここで、 y は状態数、 x は Δ パワーの変曲点数を示す。近似直線を算出した結果、 $a = 0.23$ 、 $b = 4.26$ の値を得られた。算出した近似直線を図 6 に示す。横軸は Δ パワーの変曲点数、縦軸は最尤状態数を表す。図 6 中の●印は近似直線算出に用いる環境音を示す。変曲点数が少ない環境音を対象に近似直線を算出することで、それら環境音に対する最尤状態数の推定精度向上が期待できる。

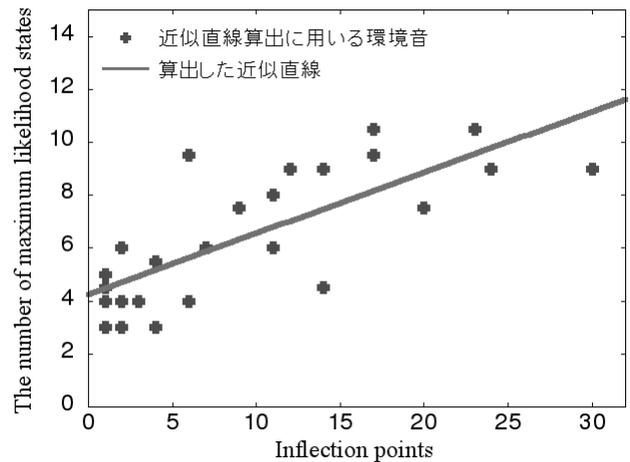


図 6 変曲点数と最尤状態数の近似直線 (変曲点数 0 から 30 まで)
Fig. 6 Regression line of inflection points and the number of maximum likelihood states.

4. 評価実験

提案法の有効性を確認するために、従来法と提案法を用いて認識性能比較実験を行い、認識性能を評価する。

4.1 実験条件

提案法により算出した近似式を用いた最尤状態数の推定法が有効であるかを確認するために、最尤状態数推定のための予備実験に使用した 38 種類の環境音に対するクロードテストを実施した。また、提案法が近似式の学習に使用した 38 種類の環境音に依存していないかを確認するために、予備実験に使用した環境音を除く他種類の環境音に対するオープンテストを実施した。両テストにおいて、対数近似を用いて算出した近似曲線と線形近似を用いて算出した近似直線を基に、各環境音の最尤状態数を推定する。それぞれで推定した状態数の音響モデルを用いて認識を行い、認識結果を比較する。

音響モデルの学習とテストには、中村らが構築した環境音データベースである RWCP-DB[7] の環境音を用いた。環境音各 70 サンプルを学習に使用し、テストには各 30 サンプルを使用した。クロードテストには RWCP-DB 内から予備実験に使用した 38 種類、オープンテストには予備実験に使用していない 18 種類の環境音を用いて認識を行った。従来法として全ての音響モデルを 5 状態で構築して認識を行い、提案法の認識率と比較する。実験条件を表 2 に示す。認識結果から、認識率を式 (4)、(5) より算出する。

$$A = 100 \times \frac{1}{C} \sum_{i=1}^C G(i), \quad (4)$$

$$G(i) = \begin{cases} 1(\text{correct}) \\ 0(\text{incorrect}), \end{cases} \quad (5)$$

表 2 実験条件

Table 2 Experimental conditions.

| | |
|---------|---|
| データベース | RWCP-DB (92 種類, 9,700 サンプル) |
| 標準化周波数 | 16 kHz |
| 量子化ビット数 | 16 bit |
| モデル数 | Closed test: 38 種類 Open test: 18 種類 |
| 特徴量 | MFCC 12 次元+パワー 1 次元 + Δ MFCC 12 次元 + Δ パワー 1 次元 + $\Delta \Delta$ MFCC 12 次元 + $\Delta \Delta$ パワー 1 次元 |
| 状態数 | Closed test: 5 状態 (left-to-right) Open test: 3~15 状態 (left-to-right) |
| 学習データ | Closed test: 2,660 サンプル (38 種類× 70 サンプル) Open test: 1,260 サンプル (18 種類× 70 サンプル) |
| テストデータ | Closed test: 1,140 サンプル (38 種類× 30 サンプル) Open test: 540 サンプル (18 種類× 30 サンプル) |

ここで, A は認識率, i は環境音番号, C は環境音の総数を示す. $G(i)$ ($i = 1, \dots, C$) は環境音の認識結果の正誤により, 1 か 0 の値を返す関数を示す.

4.2 実験結果

クローズドテストとオープンテストにおける従来法と提案法の認識数と認識率を表 3 に示す. 実験結果より, 従来法と比較して近似曲線を用いた推定法では認識率がクローズドテストにおいて 1.2 %, オープンテストにおいて 2.4 % 向上しており, 近似直線を用いた推定法では認識率がクローズドテストにおいて 2.3 %, オープンテストにおいて 2.8 % 向上していることが確認できた. このことより, 近似式を用いた最尤状態数推定による環境音認識法の有効性を確認した.

また, 近似曲線を用いた推定法と比較して, 近似直線を用いた推定法では認識率がクローズドテストにおいて 1.1 %, オープンテストにおいて 0.4 % 向上していることが確認できた. このことより, 近似直線を用いた推定法がより有効であることを確認した.

4.3 考察

近似曲線を用いた推定法と比較して, 近似直線を用いた推定法において認識率の向上が確認できた. 表 4 に, 変曲点数が 30 以下の環境音における, 対数近似を用いた近似曲線と線形近似を用いた近似直線の, 最尤状態数の真値との相関係数を示す. 線形近似による近似直線に比べて, 対数近似による近似曲線の相関係数が低いことが確認できる. そのため, 近似曲線を用いた推定法では近似直線を用

表 3 従来法と提案法の認識結果

Table 3 Experimental results with the conventional and proposed methods.

| | 認識手法 | 認識数 (認識率 [%]) |
|-------------|------------|--------------------|
| Closed test | 従来法 | 1,008/1,140 (88.4) |
| | 提案法 (対数近似) | 1,022/1,140 (89.6) |
| | 提案法 (線形近似) | 1,034/1,140 (90.7) |
| Open test | 従来法 | 404/540 (74.8) |
| | 提案法 (対数近似) | 417/540 (77.2) |
| | 提案法 (線形近似) | 419/540 (77.6) |

いた推定法と比べ, 変曲点数が少ない環境音において, 最尤状態数の推定誤差が大きくなると考えられる. 一方, 近似直線を用いた推定法では, 変曲点数が少ない環境音を用いて近似直線を算出することで, それらの環境音における最尤状態数の推定精度が向上し, 認識率が上昇したと考えられる. そして, 変曲点数が多い環境音については状態数の変化に伴う認識率への影響が少ないため, 推定誤差による認識率低下が発生せず, 全体の認識率が上昇したと考えられる.

また, 従来法と比較して提案法において陶器をスプーンで叩く音の認識率が低下していることを確認した. 予備実験における陶器をスプーンで叩く音の各状態数での認識率を図 7 に示す. 横軸は状態数, 縦軸は認識率を表す. 図 7 より陶器をスプーンで叩く音の最尤状態数は 3~6 であることが確認できる. しかし, 提案法において推定した最尤状態数は 7 であった. 図 7 から確認できるように, 状態数が 7 以降において認識率は大幅に低下している. そのため, 最尤状態数の推定誤差により陶器をスプーンで叩く音の認識率が低下したことがわかる.

ここで, 算出した最尤状態数推定式による推定精度を確認するために, 推定式算出に使用した 38 種類の環境音の最尤状態数の真値を用いて認識を行い, クローズドテストの結果と比較する. 比較結果を図 8 に示す. 横軸は手法, 縦軸は認識率を表す. 図 8 から, クローズドテストにおける近似直線を用いた推定法の結果と比べ, 最尤状態数の真値を用いた場合の方が, 認識率が約 3 % 高いことが確認できる. このことから, 最尤状態数の推定精度をさらに向上できる余地があると考えられる.

また, 継続長が異なる環境音同士の誤認識率を調査するために, 近似直線を用いた推定法において, 誤認識した各環境音の継続長を基に 3 グループに分類し, 誤認識した環境音の中で他グループの環境音に誤識別された割合を算出した. この割合が高いと継続長が大きく異なる環境音同士の誤認識が多いことを示す. その結果を表 5 に示す. 従来法に比べて提案法では他グループへの誤識別率が低下していることが確認できる. このことから Δ パワーの変曲点数によって各環境音ごとに最尤状態数の音響モデルを構築したことで, 継続長が大きく異なる環境音同士の誤認識が減

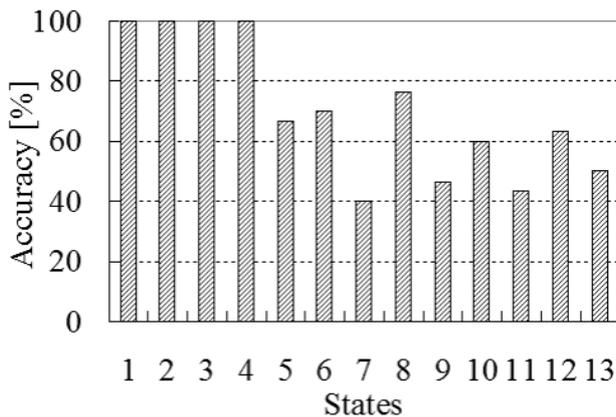


図 7 陶器をスプーンで叩く音の各状態数での認識率

Fig. 7 Recognition accuracy of sound to beat ceramics with a spoon on each state.

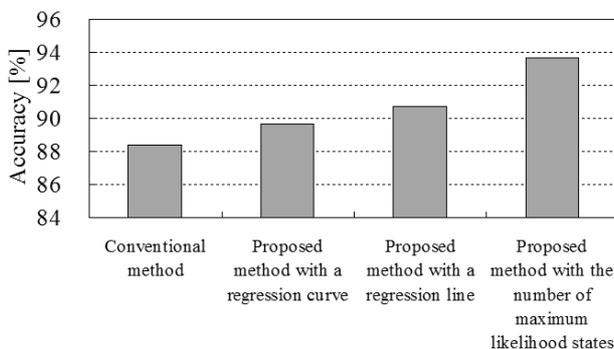


図 8 各手法と最尤状態数の真値を用いた場合の認識率

Fig. 8 Recognition accuracy of environmental sounds with the number of maximum likelihood states and each method.

表 4 最尤状態数と各手法を用いて推定した状態数の相関係数

Table 4 Correlation coefficient between the number of maximum likelihood states and the number of states estimated with each method.

| 近似手法 | 相関係数 |
|-------------|------|
| 対数近似による近似曲線 | 0.70 |
| 線形近似による近似直線 | 0.77 |

表 5 継続長が異なる環境音同士の誤識別率

Table 5 Identification error rate with 3 groups for recognition error sounds.

| | 認識手法 | 誤識別数 (誤識別率 [%]) |
|-------------|------------|-----------------|
| Closed test | 従来法 | 36/132 (27.3) |
| | 提案法 (線形近似) | 20/106 (18.9) |
| Open test | 従来法 | 44/136 (32.4) |
| | 提案法 (線形近似) | 28/121 (23.1) |

少したことがわかる。

今後の課題として、最尤状態数の推定精度向上が挙げられる。これには、最尤状態数の真値を用いた場合、認識率はさらに向上することがわかっているため、近似式の算出についての再検討が有効であると考えられる。

5. おわりに

本稿では、環境音の認識性能向上を目指し、環境音の特徴によって最尤状態数を推定し、各環境音で異なる状態数の音響モデルを構築する環境音認識法を提案した。最尤状態数の推定に用いる環境音の特徴として、環境音のパワーの変動を表す Δ パワーの変曲点数に着目し、 Δ パワーの変曲点数と状態数との関係を調査した。その関係を基に、 Δ パワーの変曲点数から最尤状態数を求める近似式を算出した。近似式から推定した状態数の音響モデルを構築し、認識を行うことで、環境音の認識性能向上を試みた。

従来法と提案法の認識性能比較実験を実施し、近似曲線を用いた最尤状態数推定法と近似直線を用いた最尤状態数推定法による認識率の向上を確認した。また、全環境音を対象に算出した近似曲線による推定法と比較し、変曲点数が少ない環境音を対象に算出した近似直線による推定法を用いることで認識率の向上を実現した。このことから、変曲点数が少ない環境音に対する状態数の推定精度が重要であるといえる。

今後の課題として、さらなる環境音認識精度向上を目指し、最尤状態数の推定を高度化する必要がある。また、継続長が同等な環境音同士の誤認識を減少させるために、 Δ パワーの変曲点数以外の特徴の利用を検討する。

謝辞 本研究の一部は科研費の助成を受けた。

参考文献

- [1] 浅野浩幸, 中野潔:安全安心まちづくりと情報通信技, 情報処理学会研究報告, Vol. 2005-EIP-27, No. 32, pp. 9-16, 2005
- [2] 桐山有美, 平山洋介:公共空間の監視と統制について:その1監視カメラの普及メカニズム(都市計画), 日本建築学会近畿支部研究報告集, No. 45, pp. 437-440, 2005
- [3] 河本満, 浅野太, 車谷浩一:マイクロフォンアレイを用いた音環境の見守りによる非日常音と危険状態の検出システム, 電子情報通信学会技術研究報告, USN, Vol. 108, No. 138, pp. 19-26, 2008
- [4] R. J. Elliott, L. Aggoun and J. B. Moore: Hidden Markov Models Estimation and Control, Springer, 1994
- [5] 三木一浩, 西浦敬信, 中村哲, 鹿野清宏:HMMを用いた環境音識別の検討, 電子情報通信学会技術研究報告, SP, Vol. 99, No. 525, pp. 79-84, 1999
- [6] 傳田遊亀, 田中貴雅, 溝口遊, 中山雅人, 西浦敬信, 山下洋一:話者方位推定を利用した動的時間領域処理に基づく遠隔発話区間検出, 電子情報通信学会論文誌, D, Vol. J92, No. 1, pp. 112-122, 2009
- [7] 中村哲, 比屋根一雄, 浅野太, 遠藤隆:実環境における音響シーンデータベースの構築, 日本音響学会秋季研究発表会, pp. 137-138, 1998