

平成十八年度
電気通信大学 大学院 修士論文

足音を用いた個人識別に関する研究

電気通信大学 大学院
情報システム学研究科
情報ネットワーク学専攻

0551008 江守 健太郎

指導教員

阪口 豊

出澤 正徳

長岡 浩司

提出年月日

平成十九年一月三十日

修士論文の和文要旨

大学院 情報システム学研究科 博士前期課程 情報ネットワーク学 専攻		
氏名	江守 健太郎	学籍番号 0551008
論文題目	足音を用いた個人識別に関する研究	
要旨	<p>我々は足音を聞いてその足音の主が誰であるか識別することができる。本研究では人間の持つこのような能力を工学的に実現することを目指し、人はいかなるような要素を聞き分けて識別を行っているのか、という問題意識の下で、足音を識別するのに有効な特徴量は何であるかについて調査を行った。</p> <p>足音の識別に関する研究報告は少なく、どのような特徴量が識別に有効であるかわかっていない。よって、本研究では人が足音を聞いた場合に感じる聴感上の特徴を手がかりに、足音の大きさを表す「足音全体のエネルギー」、歩き方を表す「時間軸のエネルギー分布」、音の構成要素を表す「周波数軸のエネルギー分布」を表現できるものを特徴量として用いた。</p> <p>本研究では11人分、それぞれ約100歩の足音を録音し、サンプルデータとして用いた。簡単のために、全ての特徴量の分布を正規分布だと仮定し、このサンプルデータから歩行者ごとに分布を推定する。それから識別したい足音一步一步について下で述べるような特徴量を抽出し、学習した分布との距離からその歩行者で観測される条件付き確率を求める。最後にベイズ推定により全ての特徴量を統合して事後確率を求め、最大となる歩行者を識別結果として与える。また、一つ抜き法により識別率を計算している。</p> <p>特徴量として以下のものを用いた。まず、足音全体のエネルギーを表すものとして足音のエネルギーを用いた。また、時間軸のエネルギー分布を表すものとして歩行の時間間隔、1次のモーメント、エネルギー比、及び分位点を用いた。次に、周波数軸上のエネルギー分布を表すものとしてパワースペクトル、1次のモーメント、ラウドネス、及びシャープネスを用いた。</p> <p>結果、有効な特徴量を組み合わせて得られた識別率の最大値は56.73%である。しかし、11人中2名で識別率が5%以下となった。この結果は個人を識別するシステムとして十分な数値とは言えない。しかし、最も有効な特徴量であるパワースペクトルの識別率の平均値と比べると、複数の特徴量をベイズ推定の手法により統合する提案システムでは約5%識別率が向上し、提案システムが有効であることを示せた。</p> <p>識別率があまり高くない理由として、特徴量の外れ値が多く、分布が正規分布に従っていない可能性がある。また、全ての歩行者に有効な特徴量はないので、特徴量同士で効果を打ち消し合っていることも考えられる。</p> <p>今後、特徴量の分布の精密な表現方法や、新たな特徴量を提案することによって、さらなる識別率の向上が望まれる。</p>	

目次

第1章	序論	1
第2章	足音の性質	2
2.1	歩行の性質	2
2.2	足音の信号的性質	3
第3章	識別手順	4
3.1	サンプルデータの収集	4
3.2	特徴量の抽出	4
3.3	特徴量の学習	4
3.4	足音の識別	5
第4章	サンプルデータの収集	6
4.1	足音の録音	6
4.1.1	録音機器	6
4.1.2	録音環境	8
4.1.3	録音方法	8
4.1.4	録音結果	8
4.2	足音の切り出し	10
4.2.1	イベント検出アルゴリズム	10
4.2.2	自動切り出しの評価と手動による補正	11
第5章	特徴量の抽出	14
5.1	先行研究で使用された特徴量	14
5.2	聴感上の足音の特徴	14
5.3	足音全体のエネルギーを表現する特徴量	15
5.3.1	足音のエネルギー	15
5.4	時間軸のエネルギー分布を表現する特徴量	15
5.4.1	歩行間隔	15
5.4.2	1次のモーメント	15
5.4.3	エネルギー比	15
5.4.4	分位点	16
5.5	周波数軸のエネルギー分布を表現する特徴量	17

5.5.1	パワースペクトル	17
5.5.2	1次のモーメント	18
5.5.3	ラウドネス	18
5.5.4	シャープネス	20
第6章	ベイズ法による特徴量の合成	21
6.1	特徴量の学習	21
6.2	ベイズの定理	23
6.2.1	特徴量が単一の場合	23
6.2.2	特徴量が複数の場合	23
第7章	特徴量の考察	24
7.1	特徴量の評価	24
7.2	音のエネルギーを表現する特徴量	25
7.2.1	足音のエネルギー	25
7.3	時間軸のエネルギー分布を表現する特徴量	28
7.3.1	足音の間隔	28
7.3.2	1次のモーメント	29
7.3.3	エネルギー比	31
7.3.4	分位点	32
7.4	周波数軸のエネルギー分布を表現する特徴量	34
7.4.1	パワースペクトル	34
7.4.2	1次のモーメント	35
7.4.3	ラウドネス	36
7.4.4	シャープネス	37
第8章	提案システムによる識別結果	38
第9章	結論	40
第10章	謝辞	41
	参考文献	41

第1章 序論

我々は日常的に、足音を聞いてその主が誰であるかを識別することができる。

個人的な経験だが、実家の自分の部屋は2階の階段のすぐ前の部屋であった。そして、その階段を登る音で父か母か、はたまた弟なのかわかることができた。

本研究では人間の持つ、このような能力を工学的に実現することを目指し、人はいかなるような要素を聞き分けて識別を行っているのか、という問題意識の下で、特に足音を識別するのに有効な特徴量は何であるかについて調査を行う。

これまでに報告された足音の識別に関する先行研究は多くない。数人程度の識別に成功した例 [1, 2] や、環境音の中から足音を抜き出すことを目標としたもの [3] が報告されている。

以上のように、足音の識別についての研究は進んでおらず、どのような特徴量が識別に有効であるかわかっていない。よって、本研究では聴感上の特徴を手がかりに、使用する特徴量の選定を行う。

本研究では、まず足音の録音を行い、次に録音した信号から足音の検出を行った。この検出された足音を用いて識別を行い、人間が使用しているであろう足音の特徴について考察を行った。

第2章 足音の性質

足音の検出、及び識別を行うために、まずは足音の性質を知らなくてはならない。本章では歩行、及び足音の性質について、先行研究による知見と研究を行って感じて自分の所見を交えて述べる。

2.1 歩行の性質

足音を生じさせる歩行の性質について述べる。歩行は次のような段階を経て行われる [4]。

- 踵接地 踵が床に接地
- 足底接地 足底全体が床に接地
- 踵離地 踵が地面から離れる
- 爪先離地 足部全部が地面から離れる

はじめに、踵接地時には地面と踵が激しくぶつかり合い、インパルス状の音が発生する。

次の足底接地時は歩き方によって大きく変わってくる。一般的には踵が床に着いたあと、足の裏の後部から順番に床に接地していく。しかし、足を擦る癖のある人は、踵接地時にそのまま踵を擦ったり、足の前部から床に接地することもある。また、人によっては、足の指の付け根部分が床に当たることによって第2のピークが出来ることもある。

最後の踵離地、及び爪先離地では大きな音を生じることはない。

このように足の複数の部分によって生じる音によって構成され、また、歩く、走るなど動き方の違いによって音の構成が変化する。

2.2 足音の信号的性質

図 2.1 に本研究において実際に録音し、サンプルデータとして使用した足音の例を示す。

この信号の前半部分と後半部分では振幅が同じくらいなので、背景雑音の影響を受けているといえる。よってどの区間が足音の鳴っている領域であるかは、この図からだけでは明確に判断することはできない。しかし、耳で聞くと確かにこの瞬間に足音が現れていることがわかる。また足音には、この図からも読み取れるように、踵が床にぶつかる瞬間に一番大きな音を発し、その後、0.1~0.3秒程度の継続時間を経て減衰していく性質がある。よって、足音の終点を検出することは難しい。

また、自分で測定した限りでは、一般的な歩行速度の場合、足音の間隔は0.5秒から0.7秒の間くらいである。

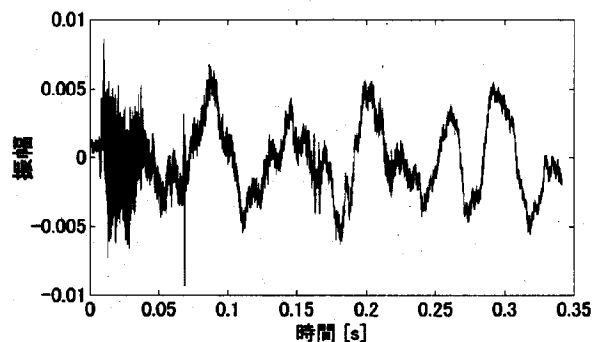


図 2.1: 足音の例

第3章 識別手順

本章では足音を識別するまでの手順について簡単な説明を行う。各項目の詳しい条件やアルゴリズムについては後述する。

3.1 サンプルデータの収集

識別を行うためには、識別を行いたい歩行者のサンプルデータを収集する必要がある。識別を行いたい歩行者ごとに足音を含む音響信号を録音し、足音を切り出す。本研究では、この切り出した足音に歩行者名と足音の開始時刻のラベルを付けた物をサンプルデータと呼ぶ。また、歩行者ごとのサンプルデータの集まりをクラスと呼ぶ。

3.2 特徴量の抽出

歩行者間の違いを表現するために、歩行者間の違いを際立たせるような情報を抽出する。このような情報を特徴量と呼ぶ。

識別率の高低は有効な特徴量を抽出できるかどうかにかかっている。この有効な特徴量とは、クラス間のばらつきが大きく、クラス内のばらつきが小さいようなものである。また、歩行者の特徴は様々に表現できるため、通常は複数の特徴量を組み合わせて使用する。

3.3 特徴量の学習

全ての特徴量についてクラスごとの分布を推定する。本研究では簡単のため、全ての特徴量は正規分布に従うと仮定したので、サンプルデータに基づいて特徴量の平均値と分散から分布を推定する。

分布を推定することによって、あるデータが入力された場合に分布の平均値からの距離で、そのクラスで観測される確率を計算することができる。

3.4 足音の識別

ここでは、複数の特徴量を統合する一つの方法としてベイズ推定を用いる。

まず、クラスごとの観測確率を条件付き確率として、ベイズの定理より入力された足音がどの歩行者のものであるかを表す事後確率を計算する。この事後確率が最大となる歩行者を識別結果として与える。このとき入力したデータのラベルと、識別結果が同じ歩行者ならば正解とする。

識別率は同じラベルを持つ入力された足音の数のうち、いくつかの足音で正解を得られたかで表す。

第4章 サンプルデータの収集

本章ではサンプルデータを収集するために必要な事柄について述べる。本研究で扱うサンプルデータとは、歩行者名と音響信号中での出現時間のラベルを持った足音の信号のことである。

足音を録音するにあたって「録音機器」、「録音環境」、「録音方法」の3つの条件について考えなければならない。よって本章では、これら3つの条件について順番に説明を行う。

4.1 足音の録音

4.1.1 録音機器

音を録音するのに必要な機械は、音を録るためのマイクロフォン、マイクロフォンで録った信号を増幅するためのアンプ、録った音を保存するためのレコーダの3種類がある。大抵のレコーダにはアンプが付いているため、本稿ではアンプについて言及せず、マイクロフォンとレコーダについてのみ選定を行う。

マイクロフォンの選定

マイクロフォンにはダイナミックマイクとコンデンサマイクの2種類がある。

ダイナミックマイクは物理的に耐久性が高く、高入力に耐えることができる。しかし一部の機種には、ボーカルマイクとして使用した場合のハンドグリップによるノイズを除去するためにハイパスフィルタが入っているものがある。このような機種では低い周波数帯域の信号を録音できないため選定の対象外となってしまう。

一方コンデンサマイクはダイナミックマイクに比べて周波数特性がよく、低域から高域までフラットである。また感度が高いため細かい音の起伏も録音できる反面、空調などの不要な音も拾いやすい。

足音は音声に比べて小さな音なので、感度が良いに越したことはない。また、足音はどのような周波数帯域に特徴を持つかわからないので、周波数特性の良いマイクを使いたい。以上のことから、足音の録音にはコンデンサマイクが適当であると判断し、図4.1の audio-technica 社のコンデンサマイク AT2020 を実験に用いた。

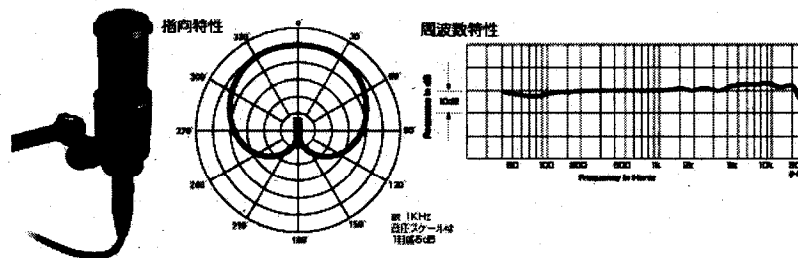


図 4.1: コンデンサマイク AT2020

レコーダの選定

録音した音響信号はPCによって信号処理される。よってPCに直接入力して録音しても良いが、パソコンそれ自体が強いノイズの発生源なので録音先としては不適である。そのため、単体のオーディオレコーダを使うことが適当であると考えた。結局、利便性の点から ROLAND R-4 をレコーダとして採用した。

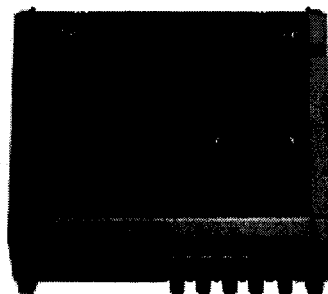


図 4.2: ポータブルレコーダ R-4

4.1.2 録音環境

表 4.1 の条件に従って、被験者 11 人について約 100 歩ずつ足音の録音を行った。録音機器は前述の通りである。この録音機材を図 4.3 に示す。また、履物や床面の性質の違いによる影響を取り除くため、履物はスニーカーに限定し、同じ部屋の同じ床において録音を行った。

表 4.1: 足音の録音条件

被験者数	11 人
足音の数	約 100 歩
履物	スニーカー
床面	研究室の床
録音形式	16bit, 48kHz, wav フォーマット

4.1.3 録音方法

足音は、音声などと比べると大きな音ではないうえに、音源である被験者が移動してマイクから離れていくと足音の信号の振幅は急激に減衰してしまう。そのため、サンプルデータの収集においては、マイクから一定距離を保った状態で足音を録音するために図 4.4 のようにマイクスタンドの周りを歩いてもらうことにした。

図 4.1 より、使用したマイクは斜め 45 度くらいまで録音感度が一定であると思われるので、マイクを下向きにして地面から 30cm の高さに固定した。被験者にはマイクスタンドの周りでできるだけ中心に近い位置をスタンドの足やケーブルを踏まないように歩いてもらった。

4.1.4 録音結果

実際に録音した信号の例を図 4.5 に示す。

この図を見ると、背景雑音の録音レベルが高くどこで足音が開始するか目で見てもわからない。中には良く聞くと車の走行音が入っている物もある。したがって、本研究で収録したサンプルデータは、学習用データとしては望ましいものとは言えない。

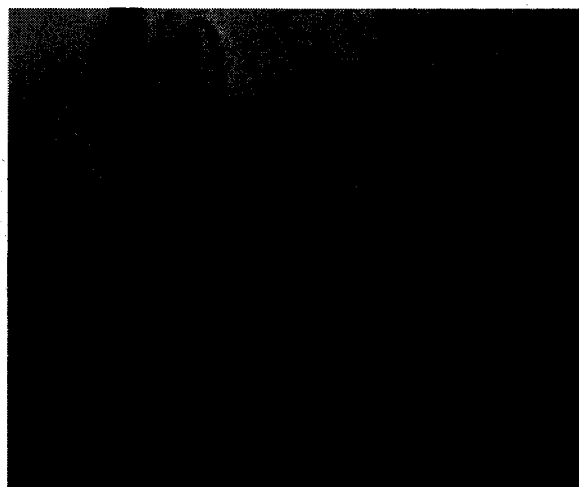


図 4.3: 実験装置

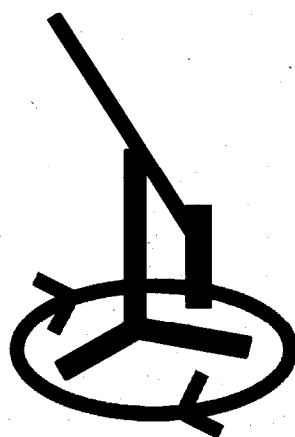


図 4.4: 歩行の仕方

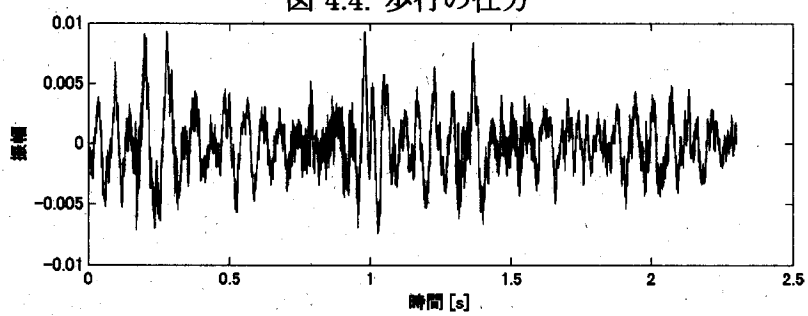


図 4.5: 録音した足音の例

4.2 足音の切り出し

録音した足音は、1つの長いwavファイル中のどこかに存在している。識別を行うためには、この一連の音響信号中から足音を検出しなければならない。

第2章では、足音はインパルス状の信号で急激に信号が立ち上がる性質を持っていること、また足音の終了点は明確でないことについて述べた。そこで音響信号中より信号の立ち上がり調べ、足音の開始点のみ検出することを考えた。しかし、この方法では足音以外の物音にも反応してしまう。検出後に足音とそれ以外を区別する方法があるならばよいが、それは非常に困難である。そこで、足音とそれ以外をすべて含めてイベントと呼び検出を行った。また、それらのイベントが発生する時刻をラベルと呼ぶ。

本手法では検出されたイベントを足音候補点とし、実際には耳で足音を聞いて真のラベルを付けるための手がかりとした。よって本手法の目指すところは過検出は許容し、誤検出をなくすことである。ここでいう過検出とは、足音挿入位置以外をイベント開始点と判断することで、誤検出とは足音挿入位置を見逃してしまうことである。

4.2.1 イベント検出アルゴリズム

音響エネルギー

すでに述べたように、足音の開始点で急にエネルギーが立ち上がる性質を利用してイベントの検出を行う。

まず、音響信号を時間方向に粗く分割し、各時間区間におけるエネルギーを求め、ここでは0.01秒ごとに信号を分割しエネルギーを求めた。この音響信号を f と表現し、0.01秒間のデータ数を L とする。この区間のエネルギー E は次のように計算することができる。

$$E = \sum_{i=t+1}^{t+L} f^2(i) \quad (4.1)$$

図4.6(a)に足音の含まれる音響信号の例を示す。この音響信号に式(4.1)を適用したのが図4.6(b)である。

次に E を式(4.2)のような5点の移動平均法により平滑化を行った。平滑化されたエネルギー \bar{E} は図4.6(c)のようになった。

$$\bar{E}(t) = \sum_{i=-2}^2 E(t+i)/5 \quad (4.2)$$

また図4.6(d)は、耳で聞いて調べた足音の開始位置を灰色の縦線として図4.6(c)上に書き込んだものである。足音の開始点の直後を見ると、どの場合でもエネルギーの立ち上がりの部分に相当することがわかる。

エネルギーのエッジ検出

エネルギーの立ち上がりを検出するために、平滑化された信号のエネルギー \bar{E} を微分する。この結果を E' として図 4.7(a) に示した。

$$E'(t) = \frac{d\bar{E}(t)}{dt} \quad (4.3)$$

ここではエネルギーの立ち上がりのみ知りたいので、微分 E' の正の成分だけが必要になる。そこで式 (4.4) の 5 点の移動平均法を使って E' を平滑化し、0 以上の値だけをプロットしたものが図 4.7(b) となる。また、この図 4.7(b) 中に足音の開始位置を書き入れたものが図 4.7(c) である。

$$\bar{E}'(t) = \sum_{i=-2}^2 E'(t+i)/5 \quad (4.4)$$

最初の足音以外はすべて本手法によってイベントを検出できていることがわかる。そもそも、この開始位置は自分の耳で聞いてラベルをつけた物であり、当然誤差が含まれる。耳で聞いた場合の誤差を 0.05 秒程度と考えると、最初の足音も許容範囲内とすることができる。

4.2.2 自動切り出しの評価と手動による補正

一度本手法で足音候補点を抽出し、それに従い手動で足音のラベルをつけたその後、再度本手法で最適な閾値の場合でラベルを付けて、その場合の検出率を調べた。

表 4.2 はその結果である。「足音の数」は手動でラベルを付けて調べた正解の足音の数。対して「検出数」は検出されたラベルのうち正解の物の数。「検出率」は $100 \times (\text{検出数}) / (\text{足音の数})$ で表す。また「過検出数」は検出されたラベルのうち不正解と判断された物で、「過検出率」は $100 \times (\text{過検出数}) / (\text{足音の数})$ で表す。

足音を検出できたかどうかの判定は、手動で付けたラベルの前後 0.05 秒以内に検出されたラベルがあればそのラベルは正解とする方法を用いた。

足音の検出率は 98% 以上、過検出率は 4 倍弱という結果になった。検出率が 100% でないのは足音が小さすぎる、もしくは前後に突発的なノイズが入ったなどの理由で上手く検出できなかった場合である。

また、過検出数を実際の足音の数の 5 倍まで限定できている。この数字は実際に耳で聞いてラベルを付ける段に、十分小さい数だと言える。

以上のことより、本手法は音響信号中からイベントを検出する際に有効な補助を行えることを示せた。

表 4.2: 足音の検出率

足音の数	検出数	検出率 [%]	過検出数	過検出率 [%]
1159	1144	98.7	4476	386.2

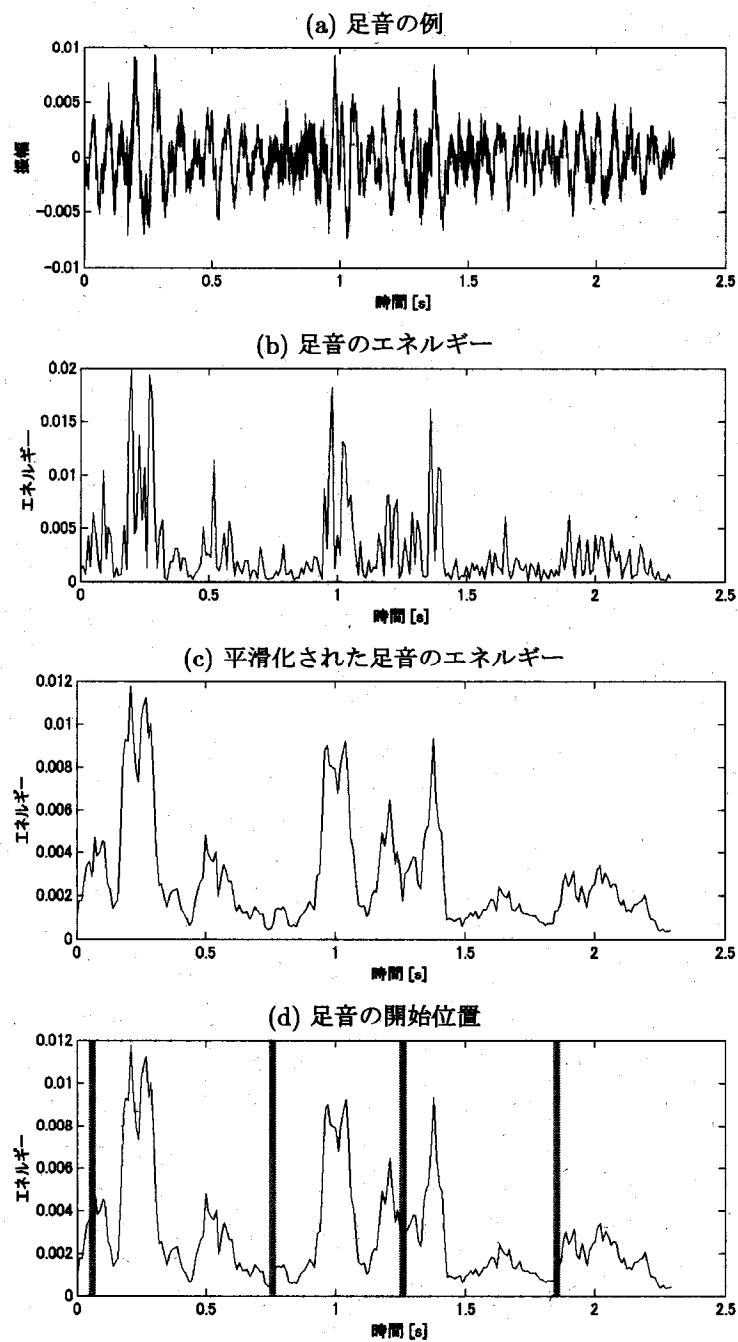


図 4.6: 足音のエネルギー

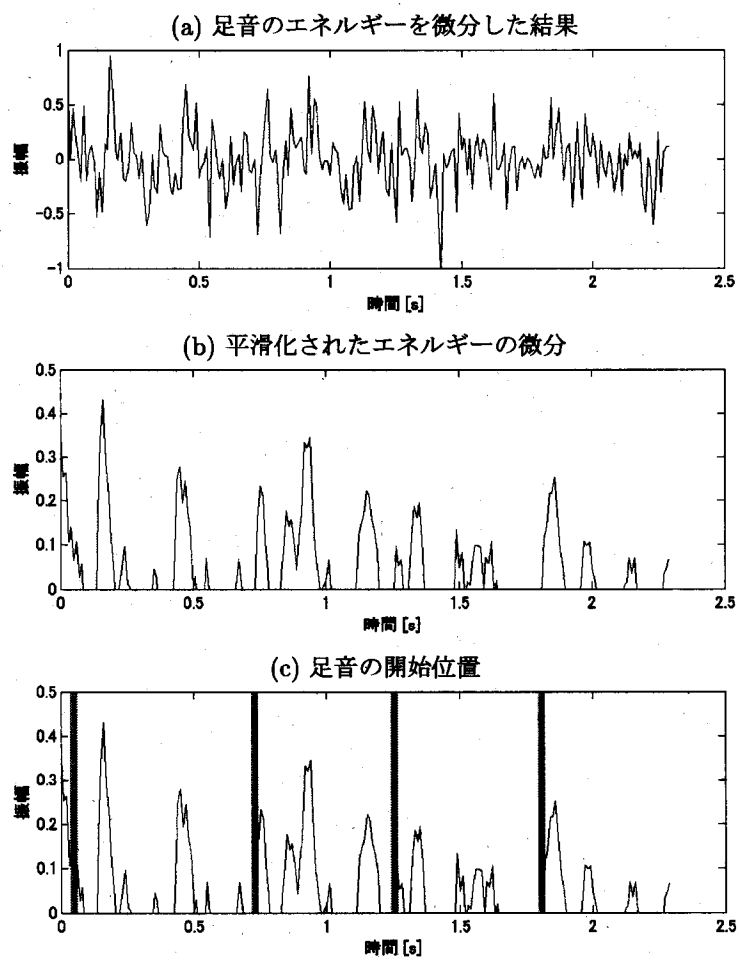


図 4.7: 足音のエネルギーを微分した結果

第5章 特徴量の抽出

5.1 先行研究で使用された特徴量

足音の識別に関する先行研究は数少ない。文献 [1] では、足音の周波数パワースペクトルを求め、移動平均処理による包絡線の相関を特徴量としている。また、文献 [2] ではパワースペクトルの包絡線の類似度、メルケプストラム分析、歩行間隔、それから心理音響パラメータ [7] としてラウドネス、シャープネス、変動強度、ラフネスを使用している。

これら2つの先行研究では3人や5人の被験者を約90%の確率で識別できることを示した。本研究でもこれらと同じように周波数領域のパラメータや心理音響パラメータを利用している。

5.2 聴感上の足音の特徴

実際に人間が耳で聞いて感じる足音の特徴として、次の3つを挙げることができる。

- 音の大きさ
- 足を擦るかどうか
- 重い音、軽い音

まず、音の大きさは足音のエネルギーの大きさとして表すことができる。よって足音のエネルギーを特徴量として使うことが考えられる。

次に、被験者が足を擦るような癖を持っているかどうか注目する。足を擦らない人は踵と床がぶつかって生じるインパルス状の信号を発するが、足を擦る人は足の裏と床が擦れて、擦らない場合に比べて長い間音が出続ける。したがって、足を擦る人は足を擦らない人と比べて時間軸方向のエネルギーの分布が異なる。つまり、エネルギーの時間軸方向の分布を特徴量として使うことが考えられる。

また重い音、軽い音という表現があるが、一般的に重い音は低い周波数成分を含む音で、軽い音は低い周波数成分はあまり含まない音とすることができる。つまり足音の周波数を特徴量として使うことが考えられる。

以上のことから「足音全体のエネルギー」、「時間軸のエネルギー分布」、「周波数軸のエネルギー分布」を表現できるような特徴量を使用する。

5.3 足音全体のエネルギーを表現する特徴量

5.3.1 足音のエネルギー

足音の大きさを表すことができる足音のエネルギーを特徴量として用いる。

0.3秒間のデータを時間軸方向に N 等分して、各区間ごとに足音のエネルギー E を求める。ここで、第 i 区間のデータの集合を I_i とすると、その区間のエネルギー E_i は以下のように計算することができる。ただし、 $f(t)$ は時刻 t での音の信号値である。

$$E_i = \sum_{t \in I_i} f^2(t) \quad (5.1)$$

5.4 時間軸のエネルギー分布を表現する特徴量

5.4.1 歩行間隔

足音の時間間隔 Δt を特徴量として用いる。次式のように、サンプルデータの足音開始時間 t_i と、一つ前のデータの足音開始時間 t_{i-1} の差を使用する。この特徴量のみ複数の足音から抽出される。

$$\Delta t = t_i - t_{i-1} \quad (5.2)$$

5.4.2 1次のモーメント

エネルギーの1次のモーメント M_i を特徴量に用いる。足音の信号 f の長さを L で表し、次式のように計算する。1次のモーメントは重心を表すので時間軸方向のエネルギーの偏りを表すことができる。

$$M_i = \sum_{t=1}^L t f^2(t) \quad (5.3)$$

5.4.3 エネルギー比

時間軸を N 等分し、第 i 区間に含まれる全体に対するエネルギーの比 $R(i)$ を特徴量に用いる。ここで第 i 区間のデータの集合を I_i と置くと、次のように表すことができる。

$$R(i) = \frac{\sum_{s \in I_i} f^2(s)}{\sum_{t=1}^L f^2(t)} \quad (5.4)$$

5.4.4 分位点

足音全体の長さ L に対して I の位置でのエネルギーの割合を分位点 P_t と呼び、次式のように表す。任意の位置でのエネルギーの割合を表すので、この特徴量は時間軸方向のエネルギーの分布を表すことができる。

$$P_t(N) = \frac{\sum_{s=1}^I f^2(s)}{\sum_{t=1}^L f^2(t)} \quad (5.5)$$

5.5 周波数軸のエネルギー分布を表現する特徴量

5.5.1 パワースペクトル

時間方向に N 分割し、各区間ごとに式 (5.6) のハミング窓 w を作用させて式 (5.7) のフーリエ変換を行う。ハミング窓を使うので、分割の仕方は図 5.1 のように半分ずつオーバーラップするものである。次に、式 (5.8) よりパワースペクトル P を得る。

最後に、後述する臨界帯域幅に従って、周波数帯域を表 5.1 のように 24 個に分割したパワースペクトルを特徴量として用いる。

$$w(t) = 0.54 - 0.46 \cos \frac{2\pi t}{L-1} \quad (5.6)$$

$$F(\omega) = \sum_{t=-\infty}^{\infty} w(t)f(t)e^{-j\omega t} \quad (5.7)$$

$$P(\omega) = |F(\omega)|^2 \quad (5.8)$$

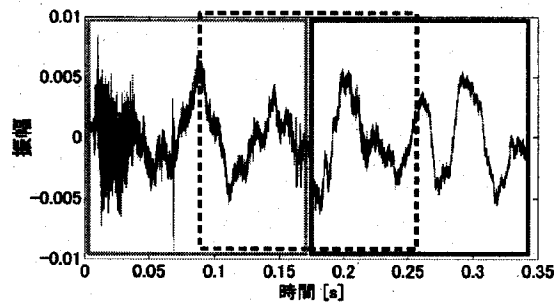


図 5.1: パワースペクトルの分割方法

表 5.1: 分割する周波数帯域

開始点 [Hz]	0	100	200	300	400	510	630	770
終了点 [Hz]	100	200	300	400	510	630	770	920
開始点 [Hz]	920	1080	1270	1480	1720	2000	2320	2700
終了点 [Hz]	1080	1270	1480	1720	2000	2320	2700	3150
開始点 [Hz]	3150	3700	4400	5300	6400	7700	9500	12000
終了点 [Hz]	3700	4400	5300	6400	7700	9500	12000	15500

5.5.2 1次のモーメント

また、得られたパワースペクトルに対して次式の計算を行なうことにより、周波数軸 ω でのエネルギーの偏りを1次のモーメント M_f から求めて特徴量に用いる。ここで W' は角周波数の最大値を表す。

$$M_f = \sum_{\omega=0}^{W'} \omega P(\omega) \quad (5.9)$$

5.5.3 ラウドネス

心理音響パラメータとして、ラウドネスを特徴量に使用する。

音の大きさを表す音圧は、同じ数字であっても周波数によって聴感上の大きさが異なる。ラウドネスとは、このような人間の耳が持つ周波数特性を考慮した音の大きさを表す指標である。まず臨界帯域幅ごとに図 5.2 の等ラウドネス曲線に沿って音圧を求め、次に、マスキングの影響を考慮することでラウドネスを求めることができる。

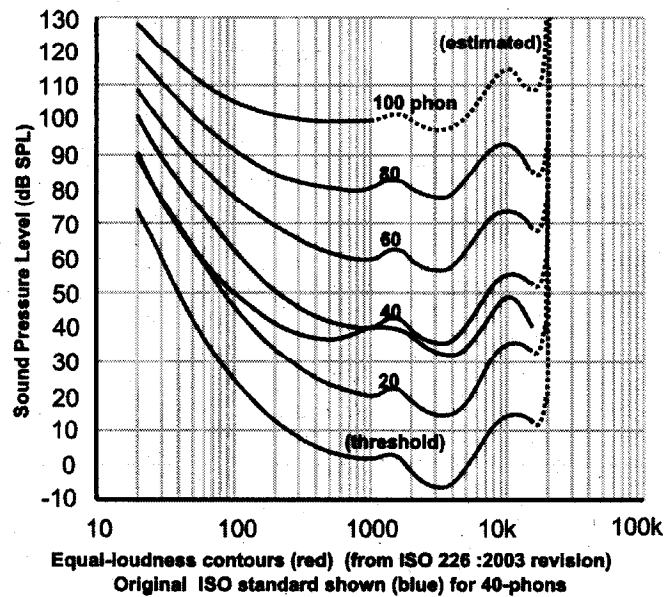


図 5.2: 等ラウドネス曲線

はじめに臨界帯域幅について説明する。臨界帯域幅とは、耳の周波数分解能を表すための指標である。離れた周波数を持つ、同じ音圧の2つの純音が同時に鳴ると音圧は2倍になるが、2つの周波数が近い場合には聴感上の音圧は2倍よりも小さくなる。このように、片方の音によってもう片方の音が聞こえなくなる現象のことを

マスキングという。この、互いにマスキングしあわなくなる周波数帯域のことを臨
界帯域幅という [8]。

低い周波数帯域から順番に臨界帯域幅を区切っていくと表 5.2 のようになる。こ
の区間ごとに、図 5.2 の等ラウドネス曲線に沿って音圧を求め、マスキングの影響
を反映させる。図 5.3 に 1000Hz の純音についてラウドネスを計算した結果を示す。
特に周波数の大きい側でマスキングの影響が強いことがわかる。

表 5.2: 臨界帯域幅

開始点 [Hz]	0	100	200	300	400	510	630	770
終了点 [Hz]	100	200	300	400	510	630	770	920
開始点 [Hz]	920	1080	1270	1480	1720	2000	2320	2700
終了点 [Hz]	1080	1270	1480	1720	2000	2320	2700	3150
開始点 [Hz]	3150	3700	4400	5300	6400	7700	9500	12000
終了点 [Hz]	3700	4400	5300	6400	7700	9500	12000	15500

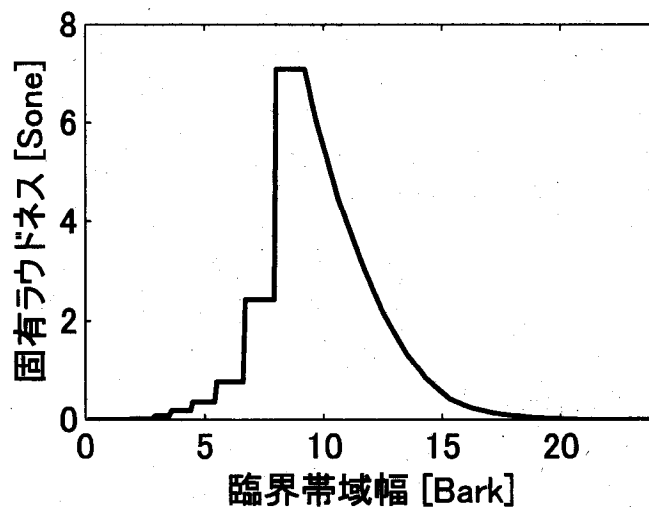


図 5.3: 1000Hz の純音のラウドネス

ラウドネスは 24 個の臨界帯域幅ごとに図 5.3 のような固有なラウドネスを求め、
全てについて論理和をとったものが得られるラウドネスとなる。図 5.4 に足音のラ
ウドネスを求めた例を示す。

図 5.3 より、マスキングは滑らかに減衰するので、本研究では臨界帯域幅間を 10
倍にサンプリングし、全部で 240 点のデータを得られるようにしている。

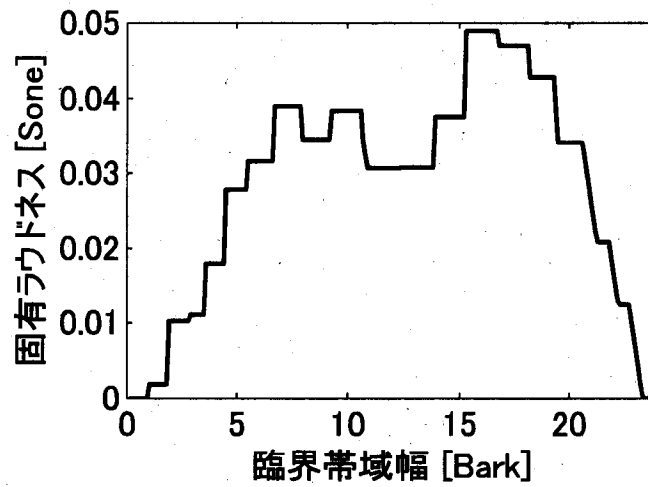


図 5.4: 足音のラウドネスの例

5.5.4 シャープネス

特徴量に心理音響パラメータの一つであるシャープネスを用いる。この特徴量は音の甲高さを表すことができる。この特徴量は、ラウドネスを N と表すとこれに窓関数 w をかけて1次のモーメントをとり、ラウドネスの総和で正規化することにより得られる。この窓関数は高い周波数に重みが付けられているので、値が大きいと高い周波数の成分が多く含まれているとわかる。

$$s = \frac{\sum_{z=1}^{24} zw(z)N(z)}{\sum_{i=1}^{24} N(i)} \quad (5.10)$$

第6章 ベイズ法による特徴量の合成

6.1 特徴量の学習

本章では、全ての特徴量について、歩行者ごとの分布を推定する方法について説明する。本研究では簡単のため、全ての特徴量は正規分布に従うと仮定し、収集したサンプルデータに基づいて特徴量の平均値と分散から分布を推定する。この推定に使用するサンプルデータを学習データと呼び、学習データに使用しなかったサンプルデータをテストデータと呼ぶ。

ここでは、テストデータから抽出された特徴量とクラスの平均値間の距離をマハラノビス距離により求め、観測確率を計算する。マハラノビス距離とはユークリッド距離などと同じように、2点間の距離を表現する方法の一つである [5]。

簡単のため図 6.1 左のような2次元の概念図を考える。図上の黒で描かれた楕円は、点Oを重心に持つ2次元正規分布の例を表す。点Oを中心とする灰色の円上の2点、点Aと点Bのどちらがよりこのクラスで観測される確率が高いかを考える。

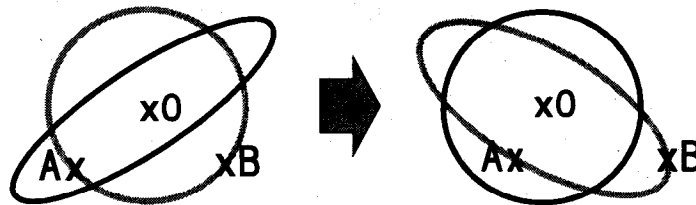


図 6.1: マハラノビス距離の概念図

点Oを中心とする灰色の真円上に点Aと点Bがあることから、ユークリッド空間では点AC間の距離 \overline{AC} と点BC間の距離 \overline{BC} は同じである。しかし点Aはクラスの境界の範囲内にあるが、点Bは範囲外にある。よって、クラスの分布の偏りを考慮すると、点Aの方が点Bよりも観測確率が高くならなければならない

そこで、図 6.1 右のように対象とするクラスの分布の偏りをなくして超球となるような空間にデータを写像すると、点Aの方が点Bより重心に近くなる。このように、クラスのデータの偏り具合を考慮したものがマハラノビス距離である。

クラス分布の共分散行列を Σ で表し、対象とするクラスとの距離を比較するデータを \mathbf{x} 、クラスの重心を μ とすると、マハラノビス距離 D^2 は以下のように表せる。

$$D^2 = (\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu}) \quad (6.1)$$

マハラノビス距離 D^2 はデータの次元数と同じ自由度 ϕ を持つのカイ 2 乗分布に従う。この性質より、そのデータ \mathbf{x} が対象とするクラスで観測される確率を式 (6.2) の、カイ 2 乗分布の確率密度関数の上側確率より計算することができる。

$$f_{\phi}(x^2) = \frac{(x^2)^{\phi/2-1} e^{-x^2/2}}{2^{\phi/2} \Gamma(\phi/2)} \quad (6.2)$$

式 (6.2) の x^2 にはマハラノビス距離 D^2 が入る。また Γ は、式 (6.3) のガンマ関数のことである。自由度 1~5 までのカイ 2 乗分布の確率密度関数を図 6.2 に示す。

$$\Gamma(\lambda) = \int_0^{\infty} x^{\lambda-1} e^{-x} dx \quad (6.3)$$

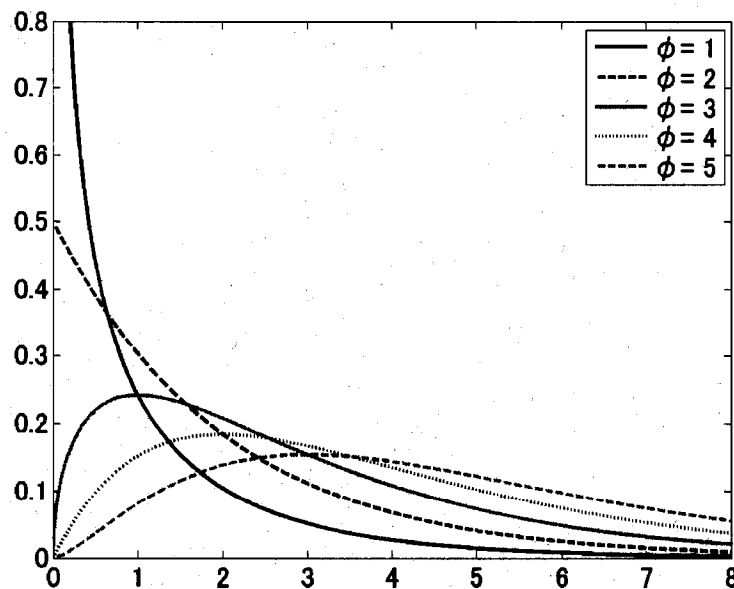


図 6.2: カイ 2 乗分布の確率密度関数

6.2 ベイズの定理

6.2.1 特徴量が単一の場合

マハラノビス距離を用いることにより、あるテストデータの特徴量 \mathbf{x} が歩行者 c で観測される条件付き確率 $p(\mathbf{x}|c)$ を計算することが出来た。次にこの条件付き確率を使って、ベイズの定理より事後確率の計算を行う [6]。

以下のようなパラメータを使用して、ベイズの定理を計算する。

- $p(\mathbf{x})$ 特徴量 \mathbf{x} の全体の確率分布
- $p(c)$ 事前確率
- $p(\mathbf{x}|c)$ 歩行者 c で特徴量 \mathbf{x} を観測する条件付き確率
- $p(c|\mathbf{x})$ 事後確率

全歩行者の数を K とすると、ある歩行者 c について何も観測していない状態での確率、つまり事前確率は $p(c) = 1/K$ で、すべての歩行者について共通である。

また、特徴量 \mathbf{x} 全体の確率分布は次のように条件付き確率 $p(\mathbf{x}|c)$ に事前確率 $p(c)$ をかけたものの和で表せる。

$$p(\mathbf{x}) = \sum_{c=1}^K p(\mathbf{x}|c)p(c) \quad (6.4)$$

以上のことから、特徴量 \mathbf{x} が入力された場合の事後確率 $p(c|\mathbf{x})$ は、ベイズの定理を利用して次のように書くことができる。

$$p(c|\mathbf{x}) = \frac{p(\mathbf{x}|c)p(c)}{p(\mathbf{x})} = \frac{p(\mathbf{x}|c)p(c)}{\sum_{c=1}^K p(\mathbf{x}|c)p(c)} \quad (6.5)$$

6.2.2 特徴量が複数の場合

まず、同じテストデータから抽出された $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ という n 個の特徴量があると考える。これら n 個の特徴量の観測確率からクラス c での事後確率を求める。

i 番目の特徴量 \mathbf{x}_i について、クラス c で観測される確率は $p(\mathbf{x}_i|c)$ と表せる。また事前確率 $p(c)$ は特徴量によらず $p(c) = 1/K$ で一定である。よって、複数の特徴量からあるクラス c についての事後確率 $p(c|\mathbf{x})$ を計算するためには、式 (6.5) の n 回分積を取ればよい。これを表したのが式 (6.6) である。

$$p(c|\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n) = p(c) \prod_{i=1}^n \frac{p(\mathbf{x}_i|c)}{p(\mathbf{x}_i)} \quad (6.6)$$

第7章 特徴量の考察

7.1 特徴量の評価

第5章では複数の特徴量について提案を行った。しかし、それらの特徴量全てが識別に有効であるとは限らない。また、有効であっても最も識別率が良くなるようにパラメータを設定するべきである。そのためには評価したい特徴量単独で識別率を求め、結果がよいものだけを用いる。

以下では識別率を求めるまでのアルゴリズムについて述べる。第6章ではベイズ推定を用いて図7.1のように歩行者ごとの事後確率を求めた。この事後確率が最大になる歩行者をこのテストデータの識別結果とする。

サンプルデータの中から一つ抜き法を用いて学習データとテストデータを振り分ける。一つ抜き法とは、全サンプルデータのうち一つだけをテストデータに選び、残り全てを学習データに用いる。そして全てのサンプルデータを1回ずつテストデータに選び、識別を行う方法である。一つ抜き法を使うことにより、式(7.1)のように識別率は求められる。また、式(7.1)の分子を識別結果が歩行者 d となる数とすると、歩行者 c のサンプルデータが歩行者 d に誤識別された割合を表すことができる。

$$(\text{識別率}) = \frac{(\text{正解数})}{(\text{歩行者 } c \text{ のサンプルデータの数})} \quad (7.1)$$

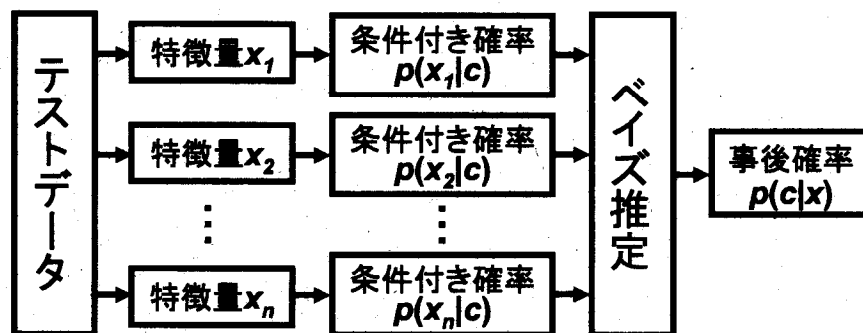


図 7.1: 事後確率を計算するまでのフローチャート

7.2 音のエネルギーを表現する特徴量

7.2.1 足音のエネルギー

足音のエネルギーを特徴量に用いる。

時間方向に1~10等分に分割して、区間ごとに足音のエネルギーを求めて識別を行った。図7.1に識別結果を示す。この表は行が分割数を表し、列が歩行者を表す。また、セル内の数字は識別率を表している。この結果から、足音のエネルギーは歩行者00,05,07に有効であることがわかる。しかし、それ以外の歩行者では極端に識別率が悪いことがわかる。また分割数を増やした場合、歩行者00の識別率は向上し、歩行者07の識別率は悪くなるがわかる。

表 7.1: 足音のエネルギーによる識別率

分割数	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10	平均値
1	21.21	0.99	11.54	8.51	0.00	26.85	5.98	65.49	4.59	0.00	3.54	13.52
2	7.07	0.99	4.81	4.26	0.88	53.70	14.53	46.90	7.34	0.00	12.39	13.90
3	33.33	1.98	1.92	4.26	1.75	55.56	13.68	38.05	7.34	0.00	2.65	14.59
4	36.36	0.00	1.92	3.19	14.91	57.41	4.27	32.74	15.60	0.00	0.00	15.13
5	46.46	0.00	0.96	10.64	11.40	56.48	1.71	20.35	17.43	2.08	0.88	15.31
6	44.44	0.99	0.00	9.57	9.65	64.81	4.27	15.93	12.84	1.04	0.88	14.95
7	38.38	0.00	0.00	10.64	7.89	70.37	1.71	10.82	9.17	0.00	0.00	13.53
8	47.47	0.99	0.00	11.70	21.93	62.96	0.85	7.96	15.60	0.00	0.00	15.41
9	60.61	0.00	0.00	8.51	14.91	56.48	5.13	6.19	10.09	0.00	0.00	14.72
10	62.63	0.00	0.00	7.45	18.42	50.93	5.13	4.42	11.93	1.04	0.00	14.72

次に、識別率の平均値が一番大きくなる5等分の場合の識別率を表7.2に示す。行が入力されるテストデータの主を表し、列は分類される歩行者を表すので、歩行者08のテストデータのうち5.50%が歩行者03に誤識別されたと表現することができる。この表を見ると、歩行者00では識別率が高く、かつ誤識別されることが少ないので歩行者00にとっては非常に有効な特徴量であるといえる。しかし、歩行者05,08では識別率は高いものの、誤識別される確率が高いことがわかる。これらの違いを見るために、分割しない場合のエネルギーを散布図として図7.2に示した。この図を見ると外れ値が非常に多く、平均値が大きい方に引っ張られていることがわかる。また、クラス間の分散に比べてクラス内の分散が大きい。特に、歩行者04,05,08は分散が大きく、他の歩行者によるクラスと重なっているため、誤識別が多いとわかる。

表 7.2: 5等分にした場合の識別率と誤識別率

	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10
歩行者00		2.02	1.01	9.09	8.08	27.27	0.00	0.00	6.06	0.00	0.00
歩行者01	13.86		2.97	5.94	10.89	30.69	0.99	18.81	14.85	0.99	0.00
歩行者02	5.77	0.96		4.81	0.96	47.12	1.92	12.50	24.04	0.96	0.00
歩行者03	14.88	2.13	2.13		8.51	41.49	2.13	8.51	7.45	2.13	0.00
歩行者04	27.19	3.51	0.88	10.53		36.84	0.88	3.51	3.51	0.88	0.88
歩行者05	7.41	4.63	2.78	7.41	5.56		0.00	1.85	11.11	0.93	1.85
歩行者06	11.97	2.56	1.71	11.11	5.13	35.90		11.97	16.24	0.00	1.71
歩行者07	3.54	2.65	2.65	7.08	1.77	34.51	5.31		21.24	0.88	0.00
歩行者08	9.17	0.92	1.83	5.50	4.59	45.87	2.75	9.17		0.92	1.83
歩行者09	6.25	1.04	4.17	6.25	5.21	47.92	0.00	10.42	16.67		0.00
歩行者10	15.04	6.19	0.88	11.50	4.42	37.17	0.88	5.31	16.81	0.88	

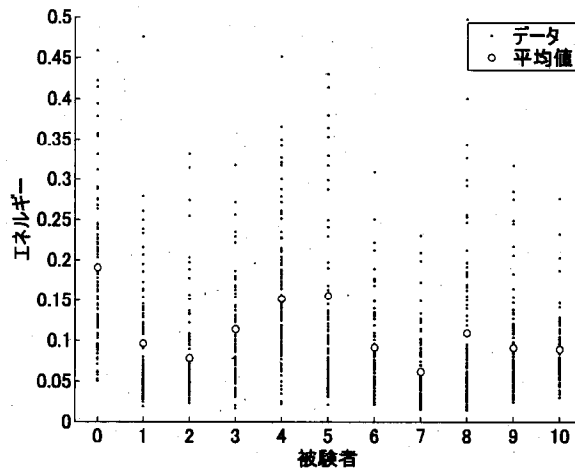


図 7.2: 足音のエネルギーの散布図

エネルギーは0以下にはならないので平均値より小さい値ではばらつきが少ないが、大きい値の方ではばらつきが大きい。つまり分布が正規分布に従っていない可能性がある。クラスごとの条件付き確率を求める場合に特徴量の分布が正規分布に従っていると仮定しているため、実際の場合と矛盾しており正しい識別率が得られていない可能性がある。

最後に、1~10まで時間方向に等分割して初めの n 個を識別に用いた場合について識別率を求める。その結果を表7.2に示した。1列目の分割数に従って時間方向に分割し、2列目のデータ数分だけ初めから n 個使用して識別を行う。歩行者ごとに一番識別率が高くなったセルに色を付けている。この表から、歩行者ごとに識別率が最大になる分割数とデータ数は異なることがわかる。また、平均値は15%弱くらいで安定してしまっている。これはある歩行者の識別率が上がったなら、その分他の歩行者の識別率が下がっていることを表す。よって、足音のエネルギーでは全ての歩行者を一度に識別できることはないといえる。

結果、足音のエネルギーは値のばらつきは大きく、パラメータによって特定の歩行者を識別することができるが、全ての歩行者を同時に識別できるような特徴量ではないといえる。

表 7.3: 足音のエネルギーによる識別率

分類数	データ数	実行者00	実行者01	実行者02	実行者03	実行者04	実行者05	実行者06	実行者07	実行者08	実行者09	実行者10	平均値
1	1	21.21	0.99	11.54	8.51	0.00	26.85	5.98	61.06	4.59	0.00	3.54	13.52
2	1	9.09	0.00	6.73	4.26	4.39	28.70	0.00	15.60	6.25	2.65	12.61	
2	2	7.07	0.99	4.81	4.26	0.88	53.70	14.53	46.90	7.34	0.00		13.90
3	1	14.14	1.98	4.81	3.19	26.32	21.30	5.13	61.06	5.50	3.13	0.88	13.40
3	2	5.05	0.99	6.73	7.45	7.89	48.15	14.53	41.59	8.26	2.08	3.54	13.30
3	3	33.33	1.98	1.92	4.26	1.75	55.56	13.88	38.05	7.34	0.00	2.65	14.59
4	1	19.19	0.00	5.77	9.57	36.84	2.78	8.55	57.52	11.01	3.13	0.88	14.11
4	2	20.20	0.00	2.88	0.00	25.44	26.85	5.98	49.56	27.52	1.04	4.42	14.90
4	3	20.20	0.00	0.96	1.06	19.30	57.41	4.27	36.28	13.78	1.04	1.77	14.19
4	4	36.36	0.00	1.92	3.19	14.91	57.41	4.27	32.74	15.60	0.00	0.00	15.13
5	1	29.29	1.98	0.96	9.57	23.68	4.63	4.27	53.98	8.26	4.17	1.77	12.96
5	2	26.26	0.00	0.96	2.13	21.05	25.93	8.55	35.40	22.94	11.46	7.96	14.78
5	3	29.29	0.00	0.96	11.70	9.65	37.96	4.27	25.66	30.28	4.17	2.65	14.24
5	4	38.38	0.00	0.00	6.38	13.16	56.48	2.56	23.01	11.93	1.04	5.31	14.39
5	5	46.46	0.00	0.96	10.64	11.40	56.48	1.71	20.35	17.43	2.08	0.88	15.31
6	1	16.16	0.00	4.81	8.51	28.07	3.70	5.13	44.25	24.77	3.13	1.77	12.75
6	2	19.19	0.99	7.69	1.06	24.56	28.70	10.26	28.32	19.27	10.42	0.00	13.68
6	3	29.29	1.98	0.00	4.26	13.16	42.59	6.84	21.24	22.02	3.13	0.88	13.22
6	4	30.30	0.00	0.00	5.32	8.77	56.48	5.13	7.96	22.94	1.04	2.65	12.78
6	5	40.40	0.00	0.00	4.26	9.65	66.67	3.42	17.70	7.34	2.08	2.65	14.02
6	6	44.44	0.99	0.00	9.57	9.65	64.81	4.27	15.93	12.84	1.04	0.88	14.95
7	1	19.19	0.00	0.00	7.45	26.32	1.85	5.13	29.20	35.78	3.13	0.00	11.64
7	2	9.09	0.00		8.51	26.32	28.70	11.11	13.27	24.77	8.33	0.00	12.96
7	3	11.11	0.00	5.77	3.19	24.56	42.59	3.42	12.39	30.28	3.13	1.77	12.56
7	4	21.21	0.00	1.92	7.45	14.04	52.78	2.56	10.62	35.78	1.04	0.00	13.40
7	5	30.30	0.00	0.96	9.57	9.65	58.33	1.71	9.73	19.27	0.00	0.00	12.68
7	6	43.43	0.00	0.00	10.64	7.02	68.52	0.85	11.50	12.84	0.00	0.00	14.07
7	7	38.38	0.00	0.00	10.64	7.89		1.71	10.62	9.17	0.00	0.00	13.53
8	1	15.15	0.00	0.00	3.19	25.44	2.78	1.71	41.59	26.61	7.29	8.85	12.06
8	2	3.03	0.99	11.54	14.89	34.21	14.81	11.97	25.66	21.10	2.08	1.77	12.91
8	3	6.06	0.99	2.88	11.70		27.78	3.42	14.16		0.00	6.19	14.47
8	4	21.21	0.99	0.00	8.51	29.82	41.67	4.27	10.62	31.19	0.00	1.77	13.84
8	5	27.27	1.98	0.00	11.70	21.05	57.41	1.71	2.65	27.52	0.00	0.00	13.75
8	6	34.34	0.00	0.00	12.77	23.68	62.96	1.71	9.73	18.35	0.00	0.88	14.95
8	7	43.43	0.00	0.00	12.77	23.68	64.81	2.56	9.73	17.43	0.00	0.88	14.95
8	8	47.47	0.99	0.00	11.70	21.93	62.96	0.85	7.96	15.60	0.00	0.00	15.41
9	1	9.09	0.00	4.81	1.06	24.56	3.70	2.56	48.67	19.27	1.04	5.31	10.92
9	2	26.26	3.96	2.88	12.77	21.93	7.41		22.12	24.77	1.04	2.65	12.91
9	3	32.32		1.92	15.96	22.81	23.15	5.13	15.04	32.11	3.13	1.77	14.39
9	4	40.40	0.99	0.00	13.83	16.67	32.41	4.27	8.85	33.94	1.04	1.77	14.02
9	5	38.38	0.00	0.00	14.89	8.77	37.96	2.56	5.31	37.61	0.00	0.00	13.23
9	6	39.39	0.00	0.00	10.64	7.89	53.70	4.27	2.65	25.69	0.00	0.00	13.11
9	7	46.46	0.00	0.00	8.51	13.16	56.48	3.42	8.85	12.84	0.00	0.00	13.61
9	8	60.61	0.99	0.00	7.45	9.65	51.85	5.98	7.08	14.68	0.00	0.00	14.38
9	9	60.61	0.00	0.00	8.51	14.91	56.48	5.13	6.19	10.09	0.00	0.00	14.72
10	1	5.05	0.99	0.00	3.19	28.95	5.56	3.42	41.59	4.59		7.08	11.12
10	2	30.30	3.96	0.96		24.56	5.56	3.42	23.01	11.01	15.83	3.54	13.21
10	3	22.22	1.98	3.85	19.15	32.46	22.22	2.56	15.93	33.94	12.50	3.54	15.49
10	4	31.31	0.00	1.92	12.77	28.95	29.63	5.98	8.85	34.86	4.17	1.77	14.56
10	5	31.31	1.98	1.92	13.83	23.68	37.96	3.42	7.08	36.70	2.08	0.00	14.54
10	6	30.30	1.98	0.96	15.96	15.79	48.15	1.71	3.54	27.52	1.04	0.00	13.36
10	7	47.47	0.00	0.96	11.70	22.81	53.70	2.56	4.42	23.85	5.21	0.88	15.78
10	8	56.57	0.00	0.96	9.57	20.18	55.56	3.42	5.31	16.51	3.13	0.00	15.56
10	9	60.61	0.00	0.00	4.26	16.67	49.07	5.13	4.42	20.18	2.08	0.00	14.77
10	10		0.00	0.00	7.45	18.42	50.93	5.13	4.42	11.93	1.04	0.00	14.72

7.3 時間軸のエネルギー分布を表現する特徴量

7.3.1 足音の時間間隔

足音のラベルの時間差を特徴量として用いる。この特徴量のみ、複数の足音のからの情報を利用している。図 7.3 にその値を示す。この図から歩行者 10 の歩行間隔が短く、歩行者 04 と 05 は歩行間隔が長いとわかる。よって、この 3 名には有効な特徴量だと予想できる。また、足音のエネルギーの場合と同様にクラス間の分散よりもクラス内分散の方が大きい。

次に、この特徴量のみを使用して識別を行う。その結果は表 7.4 に示す。予想通り、歩行間隔が短い歩行者 10 の識別率が著しく高くなった。また、歩行者 04 も高い識別率となり、歩行者 05 も他の歩行者に比べれば高い識別率となった。

以上のことから、足音の間隔は足音の識別に有効な特徴量であるといえる。しかし、足音間の時間差を利用するので足音に 1 対 1 対応できず、またサンプルデータの数が減ってしまうという欠点もある。

図 7.3: 足音の時間間隔の散布図

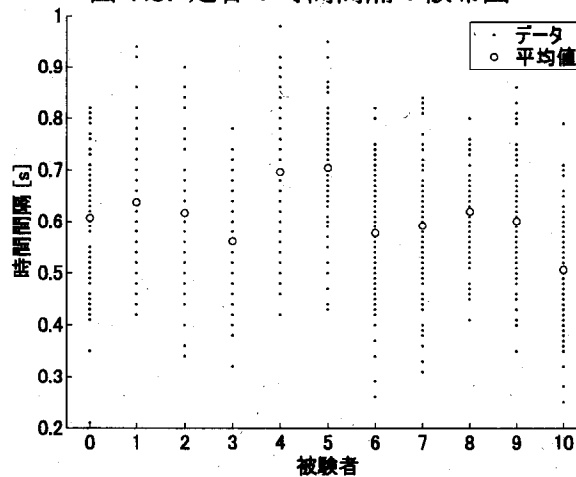


表 7.4: 足音の時間間隔による識別率と誤識別率

	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10
歩行者00		13.33	0.00	6.67	18.89	10.00	2.22	6.67	2.22	6.67	24.44
歩行者01	0.00		0.00	8.70	26.09	10.87	6.52	0.00	11.96	8.78	14.13
歩行者02	0.00	8.51		10.64	29.79	2.13	12.77	0.00	5.32	8.51	22.34
歩行者03	0.00	14.12	0.00		7.06	2.35	10.59	0.00	4.71	8.24	37.65
歩行者04	0.00	14.42	0.00	4.81		18.27	1.92	0.00	5.77	8.65	6.73
歩行者05	4.04	20.20	0.00	1.01	51.52		1.01	2.02	0.00	1.01	5.05
歩行者06	3.74	11.21	0.00	13.08	13.08	9.35		3.74	4.67	3.74	33.64
歩行者07	1.89	12.26	0.00	8.49	16.04	7.55	8.49		4.72	6.60	29.25
歩行者08	6.00	23.00	0.00	11.00	15.00	13.00	6.00	7.00		4.00	11.00
歩行者09	2.30	9.20	0.00	12.64	12.64	17.24	8.05	1.15	5.75		26.44
歩行者10	0.97	3.88	0.00	12.62	2.91	0.97	11.65	3.88	0.97	1.94	

7.3.2 1次のモーメント

エネルギーの1次のモーメントを特徴量に使用した。この特徴量では、エネルギーの時間軸方向での重心がわかる。

表 7.5 に時間軸方向に1~10等分した場合の識別率を示す。また、図 7.4 に、時間方向に分割しない場合の1次のモーメントの値を歩行者ごとに示す。図 7.4 より、クラス間の分散よりもクラス内の分散の方が圧倒的に大きいことがわかる。また、歩行者の中で平均値が一番大きい歩行者 00 と、一番小さい歩行者 07 の識別率が大きくなっている。つまり、歩行者全体の中で一番平均値が大きい人と一番小さい人のみ有効だとわかる。このデータでは歩行者 00 と 07 に有効であるが、図 7.4 よりばらつきが極端に大きいことがわかるので別のデータではこの2者に有効であるとは限らない。

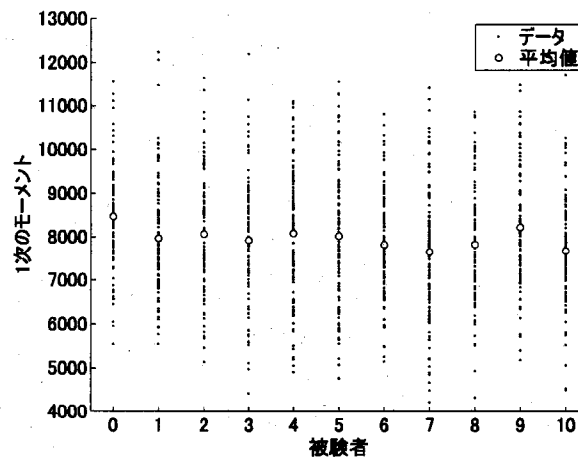


図 7.4: 分割しない場合の1次のモーメントの散布図

表 7.5: 1次のモーメントによる識別率

分割数	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10	平均値
1	42.42	0.99	0.96	3.19	14.04	1.85	1.71	46.90	4.59	5.21	3.54	11.40
2	28.28	6.93	4.81	0.00	0.00	45.37	2.56	23.89	19.27	10.42	1.77	13.03
3	16.16	0.00	4.81	0.00	0.88	48.30	0.00	16.81	4.59	25.00	0.00	10.41
4	14.14	8.91	8.65	0.00	8.77	44.44	0.00	24.78	2.75	5.21	4.42	11.10
5	14.14	0.00	0.00	3.19	15.79	32.41	0.00	18.58	13.78	28.13	10.62	12.42
6	8.08	0.00	0.96	3.19	2.63	51.85	3.42	19.47	6.42	28.13	9.73	12.17
7	24.24	0.00	3.85	22.34	3.51	26.85	0.85	11.50	16.51	23.96	11.50	13.19
8	14.14	0.00	2.88	3.19	13.16	35.19	13.68	17.70	7.34	15.63	6.19	11.74
9	20.20	4.95	1.92	14.89	8.77	28.70	11.11	18.58	12.84	27.08	5.31	14.03
10	13.13	0.99	0.96	4.26	7.89	31.48	15.38	13.27	11.93	16.67	4.42	10.94

さらに、足音のエネルギーの場合と同様に分割数を増やしその前半部分のデータを用いて識別を行った場合の結果を表 7.6 に示す。この結果から、パラメータを少し変えるだけで急激に識別率が変わってしまうことがわかる。よって、分割数を増やしてもクラス間の距離は開かず、図 7.4 のようにクラスが互いに重なり合っている様子が予測できる。

よって、1 次のモーメントは足音の識別に有効な特徴量であるとは言えない。

表 7.6: 1 次のモーメントによる識別率

分割数	データ数	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10	平均値
1	1	42.42	0.99	0.96	3.19	14.04	1.85	1.71		4.59	5.21	3.54	11.40
2	1		1.98	0.96	1.06	1.75	50.93	4.27	4.42	2.75	3.13	0.88	11.61
2	2	28.28	6.93	4.81	0.00	0.00	45.37	2.56	23.89	19.27	10.42	1.77	13.03
3	1	51.52	0.00	1.92	1.06	0.68	58.28	0.85	3.54	0.92	3.13	3.54	11.51
3	2	31.31	0.00	0.00	0.00	4.39	42.59	0.00	0.00	9.17	27.08	0.88	10.49
3	3	16.16	0.00	4.81	0.00	0.88	46.30	0.00	18.81	4.59	25.00	0.00	10.41
4	1	1.01	1.98	18.23	1.06	6.14	38.81	0.00	43.36	0.92	0.00	3.54	10.84
4	2	2.02	3.96		0.00	3.51	33.33	0.00	38.94	2.75	0.00	0.00	10.74
4	3	7.07	2.97	9.82	1.06	7.02	53.70	0.00	28.32	0.92	1.04	0.88	10.24
4	4	14.14		8.65	0.00	8.77	44.44	0.00	24.78	2.75	5.21	4.42	11.10
5	1	0.00	0.99	1.92	1.06	1.75	58.28	5.13	19.47	2.75	2.08	28.32	11.16
5	2	14.14	0.99	0.00	0.00	2.63	44.44	0.00	19.47	0.92		30.97	13.92
5	3	19.19	0.00	0.00	5.32	2.83	35.19	0.00	15.04		31.25	19.47	13.98
5	4	17.17	0.00	0.00	8.38	12.28	32.41	0.00	20.35	21.10	29.17	10.62	13.59
5	5	14.14	0.00	0.00	3.19	15.79	32.41	0.00	18.58	13.76	28.13	10.82	12.42
6	1	5.05	0.99	9.82	1.06	0.88	62.04	3.42	4.42	1.83	2.08		12.17
6	2	23.23	0.99	1.92	1.06	0.00		1.71	10.62	0.00	25.00	30.97	14.66
6	3	12.12	0.00	10.58	0.00	0.00		0.00	7.96	3.67	9.38	24.78	12.20
6	4	9.09	0.00	0.96	7.45	0.00	82.04	0.85	5.31	4.59	25.00	17.70	12.09
6	5	15.15	0.00	0.00	6.38	0.88	59.28	3.42	6.19	3.67	22.92	15.04	12.08
6	6	8.08	0.00	0.96	3.19	2.83	51.85	3.42	19.47	6.42	28.13	9.73	12.17
7	1	3.03	2.97	3.85	0.00	0.00	57.41	0.85	2.65	0.92	5.21	38.94	10.53
7	2	8.08	0.99	9.82	0.00	0.88	64.81	4.27	15.93	0.00	6.25	25.66	12.41
7	3	5.05	0.00	1.92	4.26	7.02	48.15	1.71	12.39	1.83	12.50	30.09	11.36
7	4	19.19	0.00	0.96	1.06	10.53	30.56	0.00	9.73	22.94	22.92	15.93	12.17
7	5	26.26	0.00	0.96	2.13	3.51	29.63	2.56	8.85	19.27	31.25	12.39	12.44
7	6	24.24	0.00	0.00	6.38	0.88	25.83	2.56	11.50	17.43	25.00	15.04	11.72
7	7	24.24	0.00	3.85		3.51	28.85	0.85	11.50	16.51	23.96	11.50	13.19
8	1	1.01	0.99	1.92	4.26	4.39	56.48		7.08	4.59	6.25	6.19	12.04
8	2	4.04	1.98	9.82	3.19		40.74	31.82	20.35	0.00	0.00	7.08	12.70
8	3	4.04	0.99	3.85	1.06	20.18	37.96	29.06	14.16	0.92	15.63	15.04	12.98
8	4	4.04	1.98	6.73	1.06	16.67	46.30	23.08	13.27	4.59	7.29	11.50	12.41
8	5	9.09	0.00	1.92	2.13	11.40	41.67	15.38	13.27	6.42	14.58	11.50	11.58
8	6	22.22	0.00	3.85	0.00	19.30	41.67	12.92	14.16	7.34	6.25	7.96	12.32
8	7	19.19	0.00	3.85	1.06	15.79	36.11	23.08	12.39	6.42	8.33	6.19	12.04
8	8	14.14	0.00	2.88	3.19	13.16	35.19	13.88	17.70	7.34	15.63	6.19	11.74
9	1	22.22	2.97	6.73	4.26	0.88	32.41	38.46	6.19	3.67	3.13	3.54	11.31
9	2	2.02	5.94	2.88	2.13	14.04	36.11	27.35	24.78	0.92	5.21	6.19	11.60
9	3	22.22	0.99	0.00	5.32	11.40	45.37	23.93	20.35	0.00	12.50	12.39	14.04
9	4	13.13	2.97	1.92	4.26		43.52	18.80	17.70	0.00	9.38	15.04	13.43
9	5	24.24	7.92	0.00	2.13	7.89	31.48	17.95	15.93	15.60	12.50	15.04	13.70
9	6	17.17	3.96	0.00	4.26	11.40	27.78	17.09	11.50	15.60	28.13	16.81	13.97
9	7	13.13		0.96	6.38	12.28	27.78	14.53	12.39	12.84	26.17	12.39	13.71
9	8	24.24	6.93	0.96	5.32	11.40	28.70	19.66	11.50	11.01	23.96	8.85	13.87
9	9	20.20	4.95	1.92	14.89	8.77	28.70	11.11	18.58	12.84	27.08	5.31	14.03
10	1	22.22	4.95	2.88	5.32	0.00	3.70	41.88	2.65	2.75	31.25	1.77	10.85
10	2	17.17	6.93	1.92	0.00	8.77	25.93	33.33	14.16	1.83	11.46	6.19	11.61
10	3	22.22	0.99	1.92	3.19	14.04	40.74	23.08	15.04	1.85	9.38	8.85	12.84
10	4	16.16	0.99	0.00	3.19	12.28	27.78	26.50	14.16	11.01	32.29	11.50	
10	5	13.13	0.00	3.85	6.38	17.54	37.04	23.93	13.27	10.09	16.67	10.62	13.87
10	6	11.11	0.99	0.96	6.38	15.79	33.33	23.08	11.50	20.18	17.71	8.85	13.63
10	7	14.14	0.00	0.96	5.32	12.28	28.70	23.93	11.50	14.68	20.83	6.19	12.60
10	8	9.09	0.99	1.92	6.38	10.53	32.41	20.51	11.50	11.01	19.79	6.19	11.85
10	9	12.12	2.97	0.96	3.19	10.53	35.19	19.66	12.39	11.01	16.67	4.42	11.74
10	10	13.13	0.99	0.96	4.26	7.89	31.48	15.38	13.27	11.93	16.67	4.42	10.94

7.3.3 エネルギー比

足音を2~10等分して、それぞれの区間の比を特徴量として用いる。この特徴量ではエネルギーの分布の違いを表現することができる。

表7.7にエネルギー比のみ使って識別を行った場合の識別率を示す。傾向は1次のモーメントの場合と同じで歩行者00と07の識別率が高い。次に、2分割した場合に足音の前半部分のエネルギー比を散布図にしたのが図7.5である。この図から歩行者00は前半部のエネルギーの平均値が最も小さく、歩行者07は最も大きいことがわかる。この結果は1次のモーメントの場合と全く同じである。また、分割数を2以上にすると歩行者05の識別率が增大する点も同一である。

以上のことから、エネルギー比は足音の識別に有効ではなく、1次のモーメントと同じ特徴を表現していると考えられる。

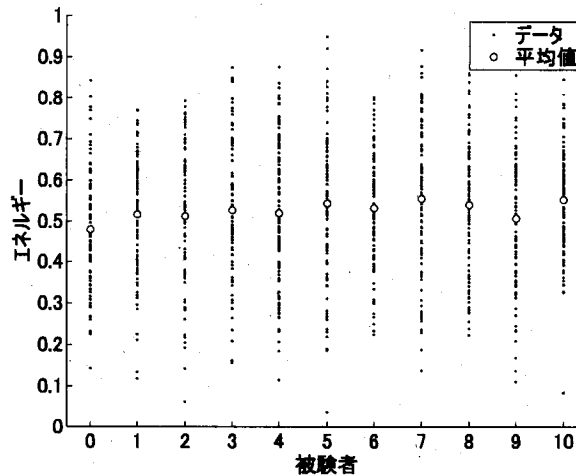


図 7.5: 2分割した足音前半のエネルギーの散布図

表 7.7: エネルギー比による識別率

分割数	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10	平均値
2	37.37	0.00	0.96	2.13	0.88	0.93	0.00	38.94	1.83	15.63	0.88	9.05
3	28.28	0.00	0.00	1.06	6.14	29.63	0.00	33.63	13.76	4.17	1.77	10.77
4	8.08	0.00	0.00	5.32	0.00	50.00	0.00	28.32	3.67	20.83	7.08	11.21
5	23.23	0.00	1.92	14.89	0.00	51.85	0.00	14.16	9.17	10.42	7.08	12.07
6	14.14	0.99	0.96	10.84	1.75	48.15	2.56	17.70	6.42	21.88	6.19	11.94
7	9.09	4.95	4.81	10.84	3.51	49.07	5.13	16.81	12.84	18.75	9.73	13.21
8	16.16	4.95	1.92	9.57	14.91	39.81	8.55	16.81	11.93	11.46	4.42	12.77
9	21.21	3.96	3.85	6.38	16.67	36.11	14.53	18.58	6.42	15.63	9.73	13.92
10	25.25	0.00	3.85	8.51	11.40	37.04	16.24	13.27	11.93	19.79	9.73	14.27

7.3.4 分位点

開始点からの累積エネルギーが指定した割合になる時刻を特徴量として用いる。

表 7.8 に分位点を 1~10 までは 1 ずつ、それ以降は 50 まで 4 ずつ変化させた場合の分離度を示す。この結果より、歩行者ごとに有効な分位点が異なることがわかる。分位点が小さい場合には歩行者 01,05,07 に有効で、分位点が大きい場合には歩行者 00,05,07,08、中央あたりが有効なのは歩行者 09,10 である。

これら全てを使用して識別を行うと、結果は表 7.9 のようになる。行がテストデータの持つラベルで、列が識別されたクラスを表す。対角線上の灰色のセルが正解での識別率を表す。この表から単体で用いた場合よりも悪くなっていることがわかる。また、事前確率は 9% 位なので、最大で 19% 程度と結果もあまり良くない。

そこで、分位点が 1,30,50,98% の場合について値の分布を図 7.6~7.9 に示した。この結果を見ると、分位点が極端に大きいか小さい場合にはばらつきが小さいが、その分足音のエネルギーの場合と同様に分布が正規分布に従っていない可能性がある。また、30% や 50% ではクラス間の分散よりもクラス内の分散の方が極端に大きくなっている。以上のことから、識別率がよい場合でも、平均値が一番大きい値と小さい値になる歩行者で識別率が大きくなっている可能性がある。実際に同じ分位点では、2 人の歩行者の識別率のみ高くなっている。

よって、分位点は足音の識別に有効であるとは言えない。

表 7.8: 分位点による識別率

分位点	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10	平均値
1	5.05	68.34	9.82	0.00	0.88	55.58	2.56	34.51	0.00	0.00	0.00	15.86
2	17.17	70.30	0.00	0.00	8.14	47.22	1.71	15.93	0.00	0.00	8.85	15.21
3	5.05	60.40	3.85	0.00	0.00	45.37	0.00	26.55	7.34	0.00	2.65	13.75
4	0.00	60.40	0.96	2.13	0.00	39.81	0.00	0.88	5.50	5.21	16.81	11.97
5	0.00	65.35	4.81	3.19	13.16	42.59	0.85	5.31	8.26	0.00	4.42	13.45
6	2.02	63.37	3.85	3.19	8.77	37.96	0.85	10.82	0.00	7.29	5.31	13.02
7	5.05	57.43	0.96	2.13	0.88	42.59	8.55	2.85	13.76	1.04	0.88	12.36
8	4.04	65.35	2.88	7.45	2.83	46.30	3.42	1.77	0.00	8.33	11.50	13.97
9	7.07	61.39	0.00	0.00	0.00	44.44	7.69	9.73	2.75	12.50	7.96	13.96
10	0.00	57.43	4.81	4.26	7.99	45.37	11.97	6.19	0.82	7.29	7.96	14.01
14	8.08	42.57	3.85	1.06	1.75	42.59	18.80	5.31	0.00	9.38	3.54	12.45
18	1.01	36.63	2.88	3.19	2.63	22.22	5.98	3.54	0.82	14.58	18.58	10.20
22	0.00	22.77	0.96	3.19	0.88	3.70	17.95	2.65	0.82	40.63	14.16	9.80
26	0.00	0.00	0.96	5.32	1.75	0.93	0.00	20.35	3.67	42.71	33.63	9.94
30	16.16	0.00	1.92	29.79	0.88	4.63	4.27	1.77	4.59	37.50	36.28	12.53
34	37.37	0.99	2.88	2.13	0.00	0.00	1.71	8.85	1.83	14.58	38.05	9.86
38	32.32	2.97	1.92	4.26	0.00	2.78	3.42	30.97	0.00	16.67	27.43	11.16
42	28.28	1.98	5.77	2.13	0.88	0.93	1.71	44.25	2.75	18.79	15.93	11.31
46	37.37	4.95	0.96	4.26	0.00	0.93	3.42	51.33	0.82	13.54	0.00	10.70
50	49.49	0.00	1.92	1.06	0.88	0.93	1.71	48.67	0.00	6.25	7.96	10.81
54	49.49	1.98	1.92	2.13	0.88	0.93	0.00	53.10	0.82	6.25	1.77	10.85
58	50.51	0.00	0.00	1.06	3.51	5.56	0.85	53.10	3.67	7.29	0.00	11.41
62	51.52	0.99	1.92	0.00	6.14	2.78	1.71	51.33	0.82	7.29	0.88	11.41
66	49.49	0.99	0.00	1.06	13.16	0.93	4.27	46.02	3.67	9.38	3.54	12.05
70	52.53	4.95	0.00	2.13	1.75	0.00	0.00	46.02	1.83	1.04	7.08	10.67
74	50.51	2.97	0.00	1.06	5.26	3.70	0.00	47.79	0.82	5.21	0.88	10.75
78	36.36	5.94	0.00	2.13	16.67	6.48	0.00	38.94	4.59	0.00	1.77	10.26
82	53.54	7.92	0.96	1.06	0.88	6.48	0.00	18.58	18.35	1.04	6.19	10.46
86	57.58	7.92	0.00	3.19	0.00	27.78	0.00	4.42	24.77	1.04	0.88	11.60
90	54.55	0.00	1.92	2.13	0.00	37.04	7.69	0.00	8.26	0.00	0.00	10.14
94	42.42	9.90	11.54	0.00	0.00	53.70	0.00	3.54	0.00	1.04	0.00	11.10
98	36.36	0.00	8.65	0.00	0.00	57.41	0.00	0.88	0.82	0.00	2.65	9.72

表 7.9: 全ての分位点を使用した場合の識別率

	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10
歩行者00		2.02	3.03	13.13	15.15	10.10	3.03	4.04	5.05	16.16	9.09
歩行者01	22.77		4.95	9.90	10.89	14.85	3.96	4.95	2.97	13.86	8.91
歩行者02	11.54	2.88		8.65	18.27	15.38	0.96	2.88	-3.85	18.27	5.77
歩行者03	11.70	3.19	7.45		19.15	14.89	3.19	5.32	6.38	10.64	8.51
歩行者04	11.40	0.88	7.02	13.16		26.32	1.75	7.02	2.63	15.79	8.77
歩行者05	16.67	0.93	0.93	10.19	13.89		0.00	7.41	2.78	7.41	2.78
歩行者06	19.66	0.00	9.40	10.26	16.24	13.68		3.42	1.71	9.40	8.55
歩行者07	11.50	1.77	2.65	7.96	15.04	18.58	1.77		7.08	11.50	14.16
歩行者08	11.01	0.00	3.67	13.76	13.76	21.10	0.82	5.50		19.27	6.42
歩行者09	14.58	0.00	7.29	12.50	15.63	15.63	4.17	3.13	6.25		4.17
歩行者10	22.12	0.00	4.42	9.73	8.85	15.04	3.54	8.85	1.77	13.27	

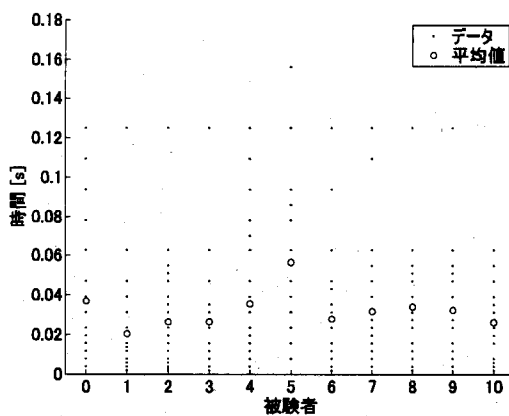


図 7.6: 分位点 1% の散布図

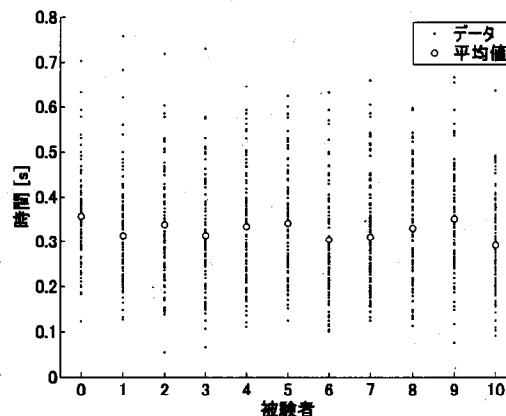


図 7.7: 分位点 30% の散布図

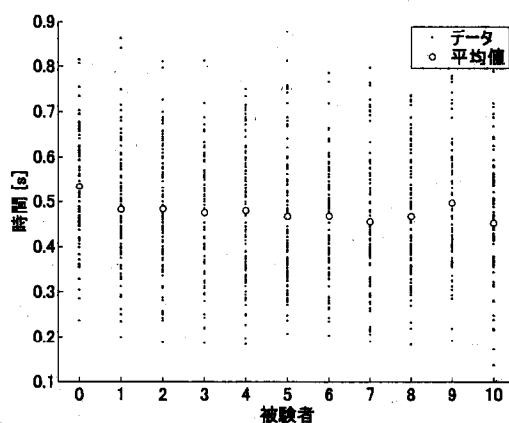


図 7.8: 分位点 50% の散布図

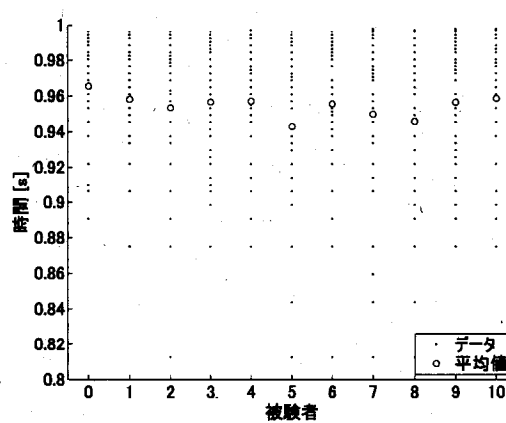


図 7.9: 分位点 98% の散布図

7.4 周波数軸のエネルギー分布を表現する特徴量

7.4.1 パワースペクトル

足音を N 個に分割し、各区間ごとにパワースペクトルを求め特徴量に用いる。分割した区間にはハミング窓をかけるので、図 5.1 のように区間を半分ずつオーバーラップして計算を行う。また、周波数帯域は表 5.1 のように 24 個に分割する。

まずは、パワースペクトルの形を比較するために足音の信号のエネルギーを 1 に正規化して、分割数を 1 から 19 まで変化させた場合の分離度を表 7.10 に示す。また、正規化せずにパワースペクトルを求めて識別率を得たものを表 7.11 に示す。ただし、データ数よりも特徴量の次元数が多くなるとマハラノビス距離が計算できないので、特徴量の次元が 80 次元よりも大きくなる場合は主成分分析により 80 次元まで圧縮している。

さて、この表 7.10 と表 7.11 を見ると、どちらの場合も歩行者 00 と 06 を除き時間方向に分割しない場合に識別率が一番高くなっている。また、正規化した場合としない場合では、正規化しない場合の方が識別率がよい。どちらも識別率の良い歩行者には同じ傾向があるので、エネルギーの違いも表現できる正規化しない場合のパワースペクトルのみ特徴量として用いるべきだろう。

表 7.10: 正規化したパワースペクトルによる識別率

分割数	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10	平均値
1	2.02	63.37	39.42	0.00	6.14	13.89	16.24	0.00	22.94	3.13	0.00	15.19
3	25.25	20.79	33.65	0.00	3.51	3.70	6.84	1.77	7.34	0.00	5.31	9.83
5	35.35	9.90	24.04	0.00	5.26	3.70	18.24	0.00	9.17	0.00	0.00	9.42
7	34.34	8.91	24.04	0.00	6.14	4.63	17.95	1.77	11.93	0.00	4.42	10.38
9	30.30	12.87	25.96	0.00	5.26	7.41	23.93	0.88	9.17	2.08	6.19	11.28
11	33.33	9.90	20.19	0.00	7.89	8.33	30.77	1.77	12.84	2.08	7.08	12.20
13	34.34	10.89	16.35	0.00	6.14	6.48	27.35	4.42	9.17	2.08	11.50	11.70
15	34.34	7.92	15.38	0.00	8.77	12.96	33.33	5.31	8.26	0.00	4.42	11.88
17	32.32	8.91	18.27	0.00	7.02	7.41	28.21	4.42	9.17	0.00	5.31	11.00
19	34.34	10.89	14.42	0.00	10.53	6.48	29.91	7.96	14.68	1.04	7.08	12.49

表 7.11: 正規化しないパワースペクトルによる識別率

分割数	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10	平均値
1	55.56	48.51	29.81	0.00	42.11	25.93	12.82	4.42	18.35	2.08	2.65	22.02
3	83.84	18.81	13.48	0.00	31.58	8.33	7.69	0.00	3.67	1.04	3.54	15.63
5	87.88	7.92	9.62	0.00	29.82	3.70	11.11	0.00	5.50	0.00	4.42	14.54
7	89.90	7.92	8.65	0.00	35.09	6.48	13.68	0.00	3.67	0.00	6.19	15.60
9	92.93	1.98	5.77	0.00	29.82	13.89	14.53	0.00	7.34	0.00	0.88	15.20
11	86.87	9.90	6.73	0.00	29.82	10.19	18.80	0.00	8.26	0.00	1.77	15.67
13	87.88	11.88	4.81	0.00	28.32	12.96	17.09	0.00	6.42	0.00	0.88	15.30
15	88.89	2.97	2.88	0.00	28.07	16.67	8.55	0.00	5.50	0.00	0.00	13.96
17	85.86	5.94	2.88	0.00	18.42	12.04	11.11	0.00	1.83	0.00	0.88	12.63
19	83.84	9.90	5.77	0.00	28.07	20.37	14.53	0.00	5.50	0.00	0.00	15.27

7.4.2 1次のモーメント

パワースペクトルの場合と同じように、1~19個まで信号を分割し、信号のエネルギーを1に正規化する。そして、それぞれの区間においてパワースペクトルを求め、さらに1次のモーメントを計算し特徴量に用いる。この特徴量は周波数領域でのエネルギーの偏りを表現することができる。

周波数領域における1次のモーメントによる識別結果を表7.12に示す。この結果より、歩行者01,02,04,05と複数の歩行者に有効な特徴量だとわかる。また、ほとんどの歩行者で識別率が高くなるので、時間方向は3つに分割するべきである。

次に、時間方向に分割しない場合について、歩行者ごとに特徴量の値を図7.10にプロットする。この結果から時間軸方向のエネルギー分布を表す特徴量よりもクラス内のばらつきが少ないことがわかる。

よって周波数領域における1次のモーメントは足音の識別に有効な特徴量であるといえることができる。

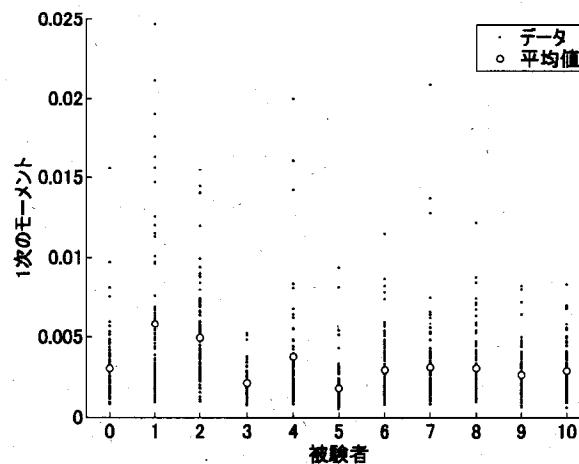


図 7.10: パワースペクトルによる1次のモーメントの散布図

表 7.12: パワースペクトルによる1次のモーメントの分離度

分割数	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10	平均値
1	1.01	19.80	8.65	14.89	42.98	22.22	2.56	2.65	0.92	10.42	4.42	11.87
3	1.01	43.56	38.46	0.00	36.84	60.19	0.00	8.85	2.75	4.17	0.00	17.80
5	0.00	53.47	44.23	1.06	23.68	37.04	0.85	22.12	9.17	1.04	2.65	17.76
7	1.01	57.43	50.96	0.00	18.42	25.00	2.56	7.08	8.26	1.04	3.54	15.94
9	0.00	59.41	47.12	1.06	13.16	19.44	5.13	12.39	10.09	3.13	3.54	15.86
11	1.01	62.38	45.19	1.06	16.67	15.74	4.27	8.85	6.42	1.04	0.88	14.87
13	0.00	66.34	45.19	0.00	11.40	16.67	5.13	6.19	3.67	3.13	2.65	14.58
15	2.02	67.33	50.00	0.00	6.14	6.48	4.27	8.85	2.75	3.13	5.31	14.21
17	0.00	68.32	46.15	0.00	5.26	6.48	4.27	11.50	6.42	2.08	0.88	13.76
19	0.00	57.43	53.85	0.00	4.39	3.70	2.56	8.85	3.67	1.04	0.88	12.40

7.4.3 ラウドネス

LOUDNESSは人間の耳に聞こえる音の大きさとそのエネルギーを表現する心理音響パラメータの一つである。パワースペクトルの場合と同じように信号を分割し、各区間のLOUDNESSを求め、特徴量として用いる。

LOUDNESSは1つの区間につき240個の特徴量を得られるが、特徴量の次元数がサンプルデータの数を上回ってしまうのでマハラノビス距離を計算することができない。よって、主成分分析を行いLOUDNESSの第80成分までを識別に使用する。

足音の信号のエネルギーを1に正規化した場合のLOUDNESSによる識別率を表7.13に示す。また、正規化しない場合のLOUDNESSによる識別率を表7.14に示す。まず両者をパワースペクトルの識別結果と見比べると、歩行者01,02はパワースペクトルの方が識別率が高いが、歩行者06,10ではLOUDNESSの方が識別率が高いなど違いが見られる。次に正規化した場合としてない場合でLOUDNESSの識別率を比べる。傾向としては、どちらも同じ歩行者で識別率が高くなり、識別率も同程度である。

表 7.13: 正規化した場合のLOUDNESSによる識別率

分割数	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10	平均値
1	42.42	1.98	4.81	0.00	20.18	4.63	30.77	15.04	4.59	2.08	18.58	13.19
3	38.38	3.96	6.73	1.06	16.67	3.70	29.91	16.81	12.84	1.04	26.55	14.33
5	35.35	2.97	5.77	0.00	12.28	1.85	31.62	8.85	8.26	0.00	19.47	11.49
7	31.31	1.98	9.62	0.00	1.75	5.56	28.21	5.31	4.59	1.04	4.42	8.53
9	28.28	0.00	13.46	0.00	0.88	5.58	22.22	5.31	8.26	0.00	1.77	7.79
11	31.31	0.00	5.77	0.00	0.88	7.41	20.51	5.31	8.26	1.04	3.54	7.64
13	30.30	3.96	5.77	0.00	2.63	3.70	17.95	4.42	8.26	0.00	7.96	7.72
15	40.40	0.99	16.35	0.00	0.00	0.93	17.09	0.00	16.51	0.00	1.77	8.55
17	36.38	4.95	13.46	0.00	0.00	2.78	18.80	0.00	20.18	0.00	1.77	8.94
19	36.36	1.98	11.54	0.00	0.00	4.63	18.80	0.00	11.01	1.04	0.88	7.84

表 7.14: 正規化しない場合のLOUDNESSによる識別率

分割数	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10	平均値
1	82.83	0.99	3.85	0.00	29.82	9.26	23.08	4.42	9.17	1.04	20.35	16.80
3	77.78	1.98	3.85	0.00	34.21	3.70	6.84	0.00	3.67	1.04	11.50	13.14
5	83.84	1.98	0.00	0.00	28.95	1.85	11.11	0.00	0.92	0.00	10.62	12.66
7	81.82	1.98	0.00	0.00	16.67	13.89	4.27	0.00	0.00	1.04	0.00	10.88
9	83.84	0.99	0.96	0.00	15.79	12.96	5.13	0.00	0.00	0.00	0.88	10.96
11	85.86	0.99	1.92	0.00	21.93	14.81	3.42	0.00	0.00	0.00	0.00	11.72
13	85.86	0.00	0.00	0.00	16.67	13.89	5.98	0.00	0.00	0.00	0.00	11.13

7.4.4 シャープネス

心理音響パラメータの一つであるシャープネスを特徴量に用いる。この特徴量は音の甲高さを表現することができる。

足音の信号を図 5.1 のように 1~5 個まで分割し、各区分についてラウドネスを計算してシャープネスを計算する。

表 7.15 にシャープネスだけ用いて求めた識別率を示す。また、図 7.11 には時間方向に分割しない場合のシャープネスを散布図として示す。図 7.11 より、最も特徴量の値が小さい歩行者 00 と、最も値が大きい歩行者 07 の識別率が大きくなっていることがわかる。このことが分割数を増やすと急激に識別率が下がる要因となっている。しかし、時間領域のエネルギー分布を表現する特徴量に比べればクラス内の分散は明らかに小さい。よって、シャープネスは足音の識別自体には有用であると考えられる。

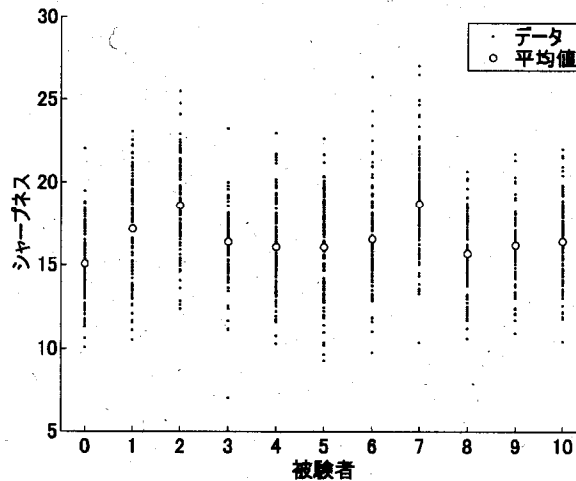


図 7.11: シャープネスの散布図

表 7.15: シャープネスの識別率

分割数	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10	平均値
1	52.53	8.91	12.50	1.05	0.88	7.41	5.13	50.44	7.34	3.13	0.88	13.65
3	5.05	4.95	14.42	12.77	26.32	41.67	24.79	5.31	18.35	10.42	0.00	14.91
5	2.02	1.98	8.65	14.89	15.79	46.30	35.90	1.77	19.27	12.50	0.00	14.46

第8章 提案システムによる識別結果

以下の特徴量を用いて識別を行う。

- 足音の間隔
- 時間方向を分割しない場合のパワースペクトル（正規化なし）
- 時間方向に3分割した場合のパワースペクトルの1次のモーメント
- 時間方向に3分割した場合のラウドネス（正規化なし）
- 時間方向に3分割した場合のシャープネス

表 8.1 に本研究により提案されたシステムによる識別率を示す。1 列目はテストデータの持つラベルを表し、1 行目は識別されるクラスを表す。よって、被験者 04 のテストデータを識別した際に、クラス 01 に誤識別してしまう確率は 4.81% である。また、対角線上の灰色のセルは正解のクラスでの識別率である。

表 8.1 より、11 人中 8 人で正解クラスが最大の識別率になった。しかし、識別率が 50% を超えたのは 1 名だけで、識別率が 5% 以下と極端に低い歩行者も 2 名いる。この結果は、個人識別を行うシステムとしては十分な精度であるとは言えない。

表 8.1: 提案システムによる識別率

	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10
歩行者00		8.89	20.00	0.00	23.33	0.00	17.78	1.11	2.22	1.11	2.22
歩行者01	9.78		23.91	0.00	8.70	2.17	15.22	2.17	4.35	1.09	1.09
歩行者02	8.51	19.15		1.06	5.32	1.06	10.64	2.13	4.26	1.06	1.06
歩行者03	3.53	12.94	22.35		12.94	0.00	21.18	4.71	10.59	1.18	5.88
歩行者04	4.81	10.58	9.82	0.00		9.82	0.00	5.77	0.00	1.92	1.92
歩行者05	5.05	5.05	20.20	0.00	13.13		17.17	0.00	10.10	2.02	2.02
歩行者06	9.35	10.28	19.63	0.00	9.35	4.67		0.93	2.80	0.00	3.74
歩行者07	1.89	10.38	16.04	1.89	14.15	3.77	22.64		4.72	1.89	3.77
歩行者08	7.00	13.00	15.00	0.00	19.00	3.00	17.00	1.00		2.00	1.00
歩行者09	8.05	13.79	20.69	0.00	18.39	2.30	12.64	0.00	17.24		3.45
歩行者10	6.80	10.68	15.53	0.97	10.68	0.00	17.48	0.00	5.83	0.97	

使用した特徴量の中で、最も有効だったパワースペクトルのみ使用した場合の識別率を表 8.2 に示す。歩行者 00 と 01 は提案システムよりも高い識別率だが、それ以外の被験者では同等かそれ以上の識別率になった。このことから、本研究で提案したような複数の特徴量をベイズ推定で統合する方法が有効であることを確認できる。これまで何度か触れてきたが、特に時間軸のエネルギー分布を表す特徴量においてばらつきが大きい。クラス間分散よりもクラス内分散が大きく、上手くクラス同士を分離できないばかりか、特徴量の分布を正規分布としてモデル化できないよ

表 8.2: 特徴量にパワースペクトルのみ使用した場合の識別率

	歩行者00	歩行者01	歩行者02	歩行者03	歩行者04	歩行者05	歩行者06	歩行者07	歩行者08	歩行者09	歩行者10
歩行者00		25.25	10.10	0.00	7.07	0.00	1.01	0.00	1.01	0.00	0.00
歩行者01	27.72		13.86	0.00	8.91	0.99	0.00	0.00	0.00	0.00	0.00
歩行者02	13.46	42.31		0.00	9.62	1.92	2.88	0.00	0.00	0.00	0.00
歩行者03	13.83	55.32	9.57		6.38	3.19	6.38	0.00	4.26	1.06	0.00
歩行者04	27.19	21.93	6.14	0.00		0.00	0.00	0.00	1.75	0.88	0.00
歩行者05	15.74	19.44	24.07	0.00	2.78		8.33	0.00	3.70	0.00	0.00
歩行者06	11.97	38.46	22.22	0.00	9.40	0.00		0.00	5.13	0.00	0.00
歩行者07	8.85	52.21	11.50	0.00	0.00	5.31	14.16		2.65	0.88	0.00
歩行者08	25.69	29.36	10.09	0.92	6.42	0.00	9.17	0.00		0.00	0.00
歩行者09	19.79	51.04	12.50	0.00	3.13	0.00	5.21	0.00	6.25		0.00
歩行者10	16.81	62.83	7.96	0.00	7.08	0.88	0.88	0.00	0.00	0.88	

うなケースも多い。また、ある特徴量が全ての被験者について有効であることは少ない。つまり、ベイズ推定で組み合わせる特徴量によって、識別率の平均値は低くとも、特定の歩行者のみ識別できるようなケースも有り得る。

以上のことをまとめると、識別率があまり高くない理由としては外れ値が多いことと、特徴量の分布が正規分布に従っていないことの2つの理由を挙げられる。したがって、改善方法として特徴量のクラス分布をより正確に表現できる確率モデルを用いることや、歩行者ごとに使用する特徴量を変えることが考えられる。

第9章 結論

本研究では、複数の特徴量を組み合わせて足音の識別を行うシステムの構築を試みた。その結果、パワースペクトルが特に足音の識別に有効であるとわかった。また、それ以外の特徴量は識別率を多少向上させるに留まったが、提案した複数の特徴量をベイズ推定により統合するシステムが有用であることは示せた。

今後、特徴量の分布の精密な表現方法や、新たな特徴量の提案によって、さらなる識別率の向上が望まれる。

第10章 謝辞

本研究を進めるにあたり、ご指導及びご助言を頂いた阪口豊助教授、出澤正徳教授、島井博行助手、石田文彦助手に心よりお礼を申し上げます。ならびに、助言や激励を頂いたヒューマンインターフェース学講座の学生、研究生、修了生の皆様、及び大島さんに心より感謝いたします。

参考文献

- [1] 田中元志, 井上浩, “足音スペクトルの比較による木造家屋内の歩行認識に関する一検討,” 電学論 C, vol.122-C, No.3, pp.525-526, 2002.
- [2] 高須賀崇, 小路康弘, 安川博, “心理音響パラメータによる足音の識別に関する一検討,” 信学技報, CST2004-37, pp.13-18, 2004.
- [3] Bin She, “Framework of footstep detection in in-door environment,” ICA2004, Mo.P2.1, pp.715-718, 2004.
- [4] 土屋和夫 監修 臨床歩行分析懇談会 編, “臨床歩行分析入門,” 医歯薬出版株式会社, 1989.
- [5] 田口玄一, “品質工学の数理,” 日本規格協会, 1999.
- [6] 麻生英樹, 津田宏治, 村田昇, “統計科学のフロンティア 6 パターン認識と学習の統計学 新しい概念と手法,” 岩波書店, 2003.
- [7] E.Zwicker and H.Fastl, “Psychoacoustics Second Edition,” Springer, 1999.
- [8] 三浦種敏 監修 電子通信学会 編, “新版 聴覚と音声,” コロナ社, 1980.