

赤外線センサにHMMを適用した物体の認識手法

宇田川 健 山崎 裕二 庄境 誠

旭化成株式会社 情報技術研究所

〒243-0021 神奈川県厚木市岡田 3050 厚木アクスト・メインタワー 22 階
E-mail: {udagawa.kb, yamazaki.yb, shozakai.mb}@om.asahi-kasei.co.jp

概要 赤外線センサは人を検知するセンサで、防犯や自動照明などに広く使われている。中でもパッシブセンサとして使われている焦電センサは、人体が発する赤外線の量を電圧レベルに変換し、閾値判定論理により、人の存在を ON/OFF で出力する。本論は、この焦電センサの電圧レベル（アナログ出力）に対してパターン認識を適用することで、人の存在だけでなく、さらに詳しい物体の情報を取り出す技術を提案する。今回使用したパターン認識手法は音声認識で広く使用されている隠れマルコフモデル（HMM）で、認識対象は、人の移動方向・少人数での人の識別・人と人以外のカテゴリ認識である。シミュレーションベースでの性能評価においては、いずれも 9 割以上の認識結果が得られており、新しい焦電センサの使用法として大いに期待できる技術である。

キーワード 焦電センサ、赤外線、パターン認識、HMM、信号処理

Infrared Sensing using HMM

Ken UDAGAWA, Yuji YAMAZAKI and Makoto SHOZAKAI

Information Technology Laboratory, Asahi Kasei Corporation

Atsugi AXT Maintower 22F, 3050 Okada, Atsugi, Kanagawa 243-0021 JAPAN
E-mail: {udagawa.kb, yamazaki.yb, shozakai.mb}@om.asahi-kasei.co.jp

Abstract An infrared sensor, widely used for security systems, lighting automation, etc, senses the presence of human. A pyroelectric sensor, the most typical one as a passive type, changes the infrared rays that a human body emits into an analog voltage level. The output of this sensor is usually digital, high means human is present and low means absent, created by plain comparator. Now we suggest the new sensing scheme applying pattern recognition into this analog voltage signal and getting not only human's presence and absence but also the more detailed information about human. In this paper, we describe three experimental results of new sensing scheme using Hidden Markov Model (HMM) currently used for speech recognition, walking direction detection, human identification and human/non-human classification. We get more than 90 percent accuracies in each simulation, and expect to develop the new usage of pyroelectric sensor.

Key words pyroelectric sensor, infrared rays, pattern recognition, HMM, signal processing

1. はじめに

来るべきユビキタス社会では、様々な現実世界の情報がセンサにより取得され、それらがネットワークで結合されることで特定の状況に最適なサービスが供給できるようになると言われている。この時のセンサは、物理量を電気信号に変換するだけでなく、それ自身が状況認識の機能を持つようなインテリジェントなものも用いられる。我々は各種センサの出力にパターン認識技術を応用することで、今まで埋もれていた情報を引き出すことができると考え研究を進めているが、中でも赤外線センサは、物体が自発的に放射している赤外線を検出するため、興味深い素材である。すなわち赤外線センサの出力には、赤外線放射源となっている物体の動き、大きさ、形状などの情報が含まれており、信号処理の手法を確立すれば、ユビキタスネットワークに対して、人の状況を伝達するためのインタフェースとして欠かせない存在になることが予想される。

本論では赤外線センサのアナログ出力に、音声認識で広く用いられている隠れマルコフモデル（HMM）を応用する実験を行い、いくつかの状況認識が可能であることを報告する。

第二章では実験に使用した焦電型赤外線センサの原理に触れ、続く第三章では赤外線センサの出力信号には赤外線放射体の動きにより固有のパターンが観測されることを示す。第四章では今回の信号処理の中心である HMM について解説し、その後の第五章で認識実験の結果を示す。最後に第六章では現在進めている作業と今後の課題について触れる。

2. 焦電型赤外線センサ

2.1 原理

赤外線センサは赤外線エネルギーの変化を電気信号の変化に変換する装置のことで、大きく量子型と熱型に分類できる。熱型赤外線センサは、常温で動作することや、波長依存性が少ないことなどの特徴を持つが、中でも焦電型赤外線センサは比較的感度が高く、また応答速度も速いため、物体検出などの用途に用いられている。

図 1 には焦電型赤外線センサの動作原理を示す。焦電効果を持つ結晶は自発分極しており、熱的に安定な状態では、表面に吸着した電荷と打ち消しあって平衡状態にある。ここに赤外線エネルギーが加えられると温度が上昇し、分極の状態が変化することから、電荷の平衡状態が崩れ、表面に顕在化した電荷を電圧の変化として観測することができる。なお、電荷は時間がたつと、漏洩・吸着により新たな平衡状態に達するため、観測する信号は、赤外線の変化に対する微分型となる。

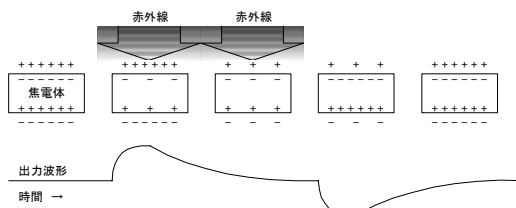


図 1 焦電型赤外線センサの動作原理

焦電型赤外線センサを物体検出用途に使用する場合、同一素子を 2 つまたは 4 つ組み合わせ、電気的に逆極性に直列接続する。これにより、周辺温度のゆらぎの影響を打ち消したり、外乱光の影響を押さえると共に、物体のより細かい動きを検出できる。

2.2 フレネルレンズ

センサとフレネルレンズを組み合わせることで赤外線の検出領域をコントロールすることができる。これはカメラに例えるとセンサ素子がフィルムであり、フィルムに像を結ぶ視野範囲が赤外線の検出領域となる。視野はレンズの焦点距離とセンサ素子の大きさで決まる。

一般に、物体検出用途ではセンサ素子に対して複数のレンズを組み合わせ使用し、空間内に複数の赤外線検出領域を分散させる（図 2）。この領域のいずれかを赤外線放射体が通過する場合、その動きの微分に相当する信号が出力される。

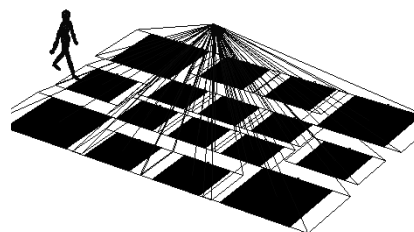


図 2 感度領域の分布例

2.3 出力信号に含まれる情報

従来、焦電型赤外線センサを侵入者の検出や自動スイッチ等に用いる場合は、センサ出力信号を増幅した後、閾値で二値化して機器制御用の ON/OFF 信号を作り出していた。しかし、センサのアナログ出力の中にはもっと複雑で豊富な情報が含まれている。例えば、図 3 はフレネルレンズにより分散された、床上の赤外線検出領域の模式図であるが、この領域を A、B、および C のように赤外線放射体が移動した場合、それぞれの経路で通過する赤外線検出領域の個数や大きさが異なるため、それぞれ異なる出力信号が観測

されるはずである。

高さ方向で見ると、図 4 に示すように、センサに近い部分では複数のレンズにより分散された赤外線検出範囲が密に詰まっているため、同じ移動速度であっても、センサに近い部分を通過したものは高い周波数を出力する。

また、同じ空間を縦長の物体が通過する場合は、まず、図 5 の A のように床に近い部分の赤外線検出領域から先に反応するため、移動開始直後は低周波成分を多く含む信号が出力される。続いて B のようにセンサの真下に近づくにつれて、高周波成分が増えて行き、センサから遠ざかると、再び低周波成分を含むようになる(図 5)。すなわち、周波数成分の時間的な変化を調べることにより、赤外線放射体の形状をある程度推定できる。

ここまで述べたように、焦電型赤外線センサの出力信号には赤外線放射体の動きや形状に関する情報が含まれている。これらを解析的に求めることは簡単ではないが、信号処理を工夫することで一部の情報を抽出することができると考えられる。

3. 物体とセンサ出力パターン

以上のような推測のもとに、物体及び物体の動作によって実際の焦電型赤外線センサのアナログ出力がどのように変化するかを調査した。調査にあたっては、同一条件のセンサ出力を複数収集し、統計的な手法を用いて規則性・パターンを確認する方法を用いた。その結果、焦電型赤外線センサのアナログ出力は、歩行経路や各個人に対して、ある固有パターンが存在することが判明した。以下に、データ収集系、解析に使用したツール、そして解析結果についての詳細を述べる。

3.1 データ収集

まず、焦電型赤外線センサのアナログ出力を A/D コンバータでデジタル信号に変換し、パソコンに取り込む実験系を作成した。今回使用したセンサはアナログ出力型のモーションセンサ NaPiOn 標準タイプ(松下電工製)^[1]で、概観とその検出範囲は図 6 に示す通りである。使用した A/D コンバータのサンプリングレートは 10Hz、解像度は 16 ビットである。

このセンサを床からの高さ 2.5m の天井に取り付け、センサの検出範囲を被験者が歩行したときのセンサ出力を収集する実験を行った。

収集にあたっては、被験者の歩行経路を図 6 のようにセンサ検出範囲に対して、検出範囲外からセンサの真下(検出範囲の中心)を通り、検出範囲外へ抜けるという直線に限定した。歩く直線は破線矢印のような 8 方向で、試行回数は方向毎に各 50 回である。これらの歩行時のセンサ出力

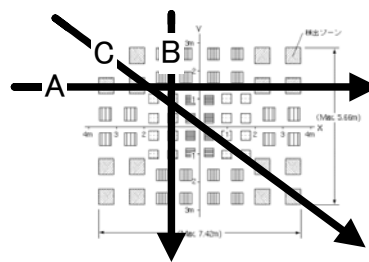


図 3 検出領域分布と移動経路

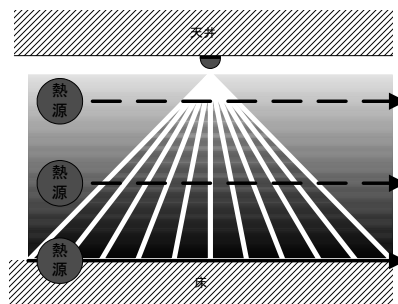


図 4 検出領域分布と熱源の高さ

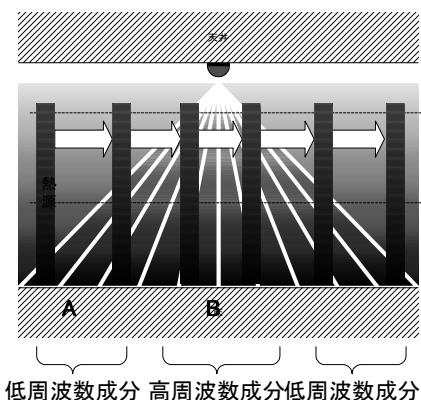


図 5 検出領域分布と熱源の形状

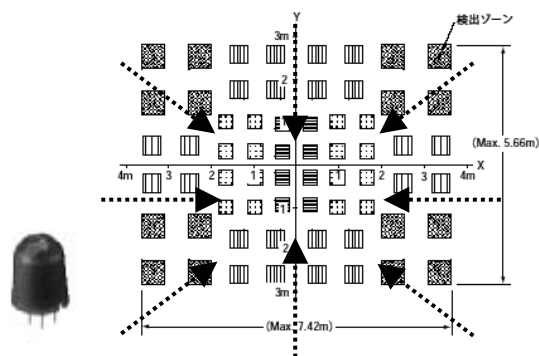


図 6 NaPiOn 標準タイプとその検出範囲

表 1 被験者のカテゴリ

カテゴリ		人(匹)数
人	幼稚園児(身長 100cm 前後)	6名
	小中学生(身長 140cm 前後)	5名
	20~30代	男女各5名
	40~50代	男女各5名
	60歳以上	男女各5名
犬	大型犬	5匹
	小型犬	5匹

をパソコンに取り込み、被験者1人当たり 400 サンプルのデータを収集した。

被験者は人 41 名と犬 10 匹で、詳細なカテゴリと人数を表 1 に示す。人に関しては、普段着・脱帽で、ポケットに手を入れることなく、普段通りの自分のペース・歩き方で歩行してもらった。収集時の気温は常温 24~27 度である。

3.2 解析手法とツール

収集したセンサ信号の例を図 7 に示す。これは 20 代女性 A が方向 ① を歩行したときのセンサ信号である。この信号の周波数スペクトルの時間変化を見るために、信号に対して音声のスペクトル解析で使用されている短時間フーリエスペクトル解析を適用した。これは、信号をフレームと呼ばれる短時間で区切り、各フレーム単位でスペクトル解析を行う手法である。

今回は、フレーム内における周波数解像度とフレーム分割による時間変化の度合いを考慮し、フレーム長 1.6 秒(16 サンプル)、フレーム間隔 0.4 秒(4 サンプル)のフレームに分割した。フレーム処理では、ハミング窓、16 ポイント FFT、パワー計算を行い、各フレームを 8 ポイントのスペクトルに変換した。

図 7 の信号を上記の手法によりスペクトルに変換した図を図 8 の左上に示す。これを見ると、低周波領域に、スペクトルの時間変化が存在することが判る。さらに、同じ女性が同じ方向 ① を二回目に歩いたときのスペクトルを右上に、方向 ① を一回目に歩いたときのスペクトルを左下に、方向 ② を二回目に歩いたときのスペクトルを右下に示す。これらを比較すると、歩いた方向毎にスペクトルの固有パターンが存在することが判る。

このように、センサ信号のスペクトルにはなんらかの規則性が存在するが、すべてのデータに対してスペクトルを目で比較し、その規則性を発見することは困難である。そこで、自社で開発した COSMOS^{[2][3]}と呼ばれるツールを使用することで、スペクトルパターンの規則性を解析した。

COSMOS は、複数の多次元ベクトル間の相互

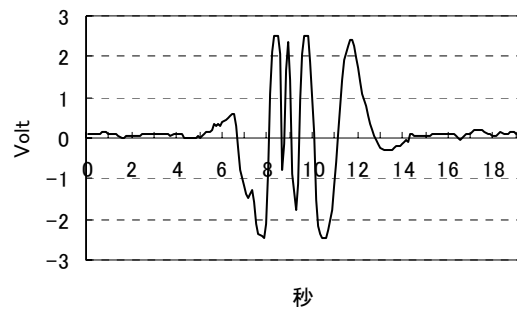


図 7 収集したセンサ信号例

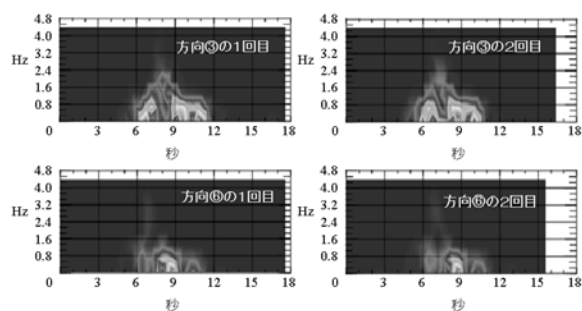


図 8 スペクトル例

距離を計算し、Sammon 法^[4]により多次元空間を二次元空間にマッピングするツールである。これにより、距離の小さい多次元ベクトル同士、つまり類似している多次元ベクトル同士は近い点に、距離の大きい多次元ベクトル同士は遠い点に配置された二次元散布図が作成でき、視覚的に多次元ベクトルの類似度を判断することが可能となる。

今回の信号スペクトルはデータ毎にフレーム数が異なるため、長さの異なるベクトル同士の距離計算を行う必要がある。従って、距離計算の手法として Euclid 距離を用いることができない。そこで今回は、パターン認識で用いられる時間軸伸縮マッチング (Dynamic Time Warping: DTW)^[5]を用いスペクトル間の距離を計算し、COSMOS に表示した。

3.3 センサ出力の規則性

以上のような手法により、人と犬のデータの一部をスペクトル変換し二次元上にマッピングした COSMOS 図を図 9 に示す。図中の × 印が犬、□印が大人、○印が子供の信号を表している。ただし、大人は表 1 の小中学生以上、子供は幼稚園児としている。図 9 より、人のスペクトルと犬のスペクトルはそれぞれ群れをなしていることが判る。これは、人と犬のスペクトルは異なり、

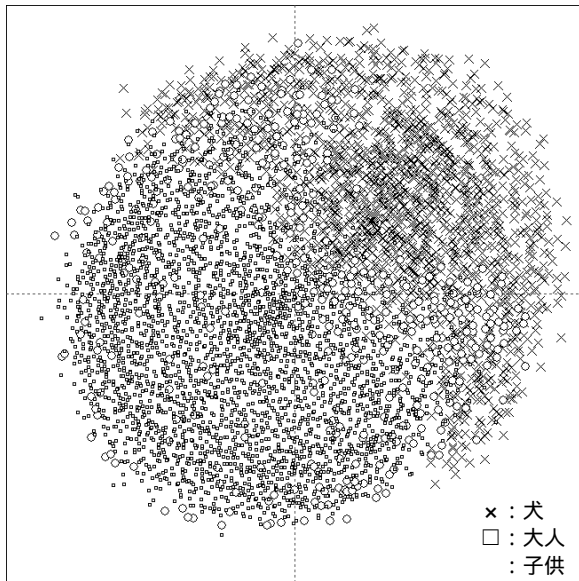


図 9 人と犬

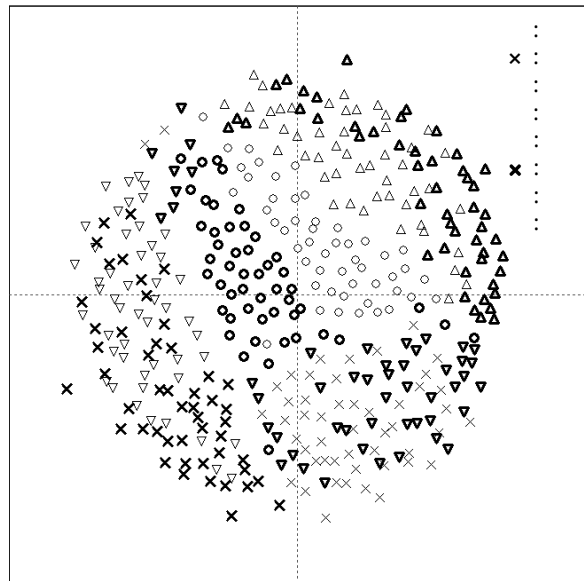


図 10 人の動線

それぞれ固有のスペクトル変化を持っていることを表している。従って、このスペクトル変化に対してパターン認識を適用することによって、この焦電型赤外線センサの出力で人と犬の区別ができる可能性があると言える。

次に、特定の人 1 名 (20 代女性) のスペクトルを二次元上にマッピングした COSMOS 図を図 10 に示す。図中の △印が方向 A、○印が方向 B、x 印が方向 C、□印が方向 D の信号を表している。歩いた方向によって群れが出来ていることから、方向毎にスペクトルが異なっていることがわかり、これもパターン認識を適用することで人間の歩いた動線を判別できる可能性があることを示している。

さらに、同一方向を歩いた特定の 3 名 (20 代女性 A,B,C) のスペクトルを二次元上にマッピングした COSMOS 図を図 11 に示す。図中の △印が A、x 印が B、○印が C の歩行時の信号を表している。図 11 から、各個人によってもスペクトルに差があり、それらは固有のパターンを持っていることがわかる。これは、歩く動線を限定した場合に、赤外線センサの出力を用いて個人の識別ができる可能性があることを示している。

以上のように、焦電型赤外線センサのアナログ出力は、赤外線を発する物体やその動きに対して、ある固有のスペクトルを出力していることが判明した。従って、同じようなスペクトルが出る物体もしくは動きでクラスタリングし、そのスペクトルに固有のモデルを作成し、パターン認識を適用することで、人の在不在だけでなく、物体の状

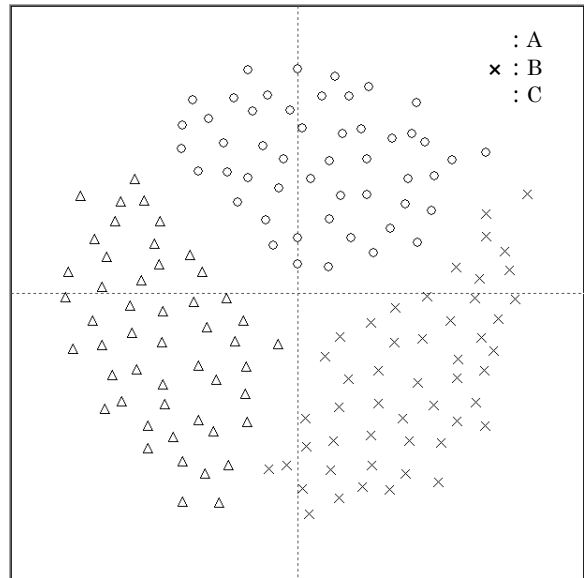


図 11 個人

態や性質などの情報が取得可能であると思われる。

4. HMM による赤外線センシング

これまでの検討をもとに、焦電型赤外線センサのアナログ出力から求められたスペクトルにパターン認識を適用し、新しい情報の認識を試みた。今回用いたパターン認識手法は HMM (Hidden Markov Model) [6] で、認識対象は、1. 人の移動方向、2. 人と人以外、3. 個人、の 3 つである。

4.1 HMM

HMM は主に音声認識で使われているパター

ン認識の手法で、時系列信号に対する確率モデルである。定常的な信号を出力する状態を複数用意し、その状態遷移モデルにより、非定常な信号をモデル化する。出力される信号からその状態遷移を直接観測することはできず、状態遷移が隠れているため、隠れマルコフモデルと言われている。

具体的には、各モデルは状態毎に状態を遷移する遷移確率分布と、状態から出力される信号の出力確率分布を持っている。この二つの分布は、大量のデータを用いた学習によって算出される。

認識を行う際には、各モデルごとに信号(フレーム)が入力される度に、入力信号から計算される出力確率分布と遷移確率分布を積算しながら状態遷移を行い、全フレーム入力後に可能なすべての状態遷移の中から最も確率尤度の高い遷移を行うモデルを選び、認識結果として出力するという処理を行う。

今回の赤外線センシングに適用するにあたっては、1回の歩行を1モデルとしたため、音声認識で例えるならば、単語認識のようなものである。学習から得られた各パラメータの単一ガウス分布から出力確率を計算し、同じく学習から得られた遷移確率と合わせて確率を積算し、最終的にはモデル同士で積算確率を比較し、最も尤度の高いモデルを選択するという処理である。

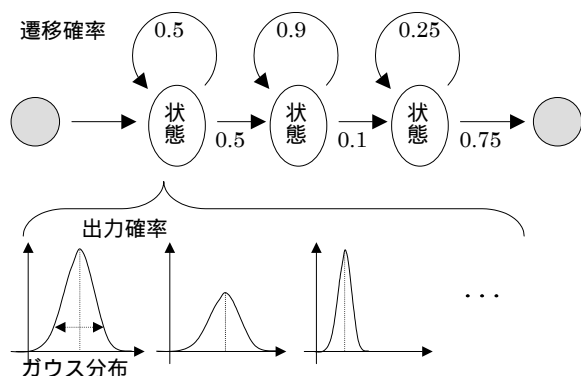


図 12 3 状態 HMM の例

4.2 人の移動方向識別

3.1 のデータ収集系を用いて収集した人 17 名のセンサ信号を用いて、移動方向の識別実験を行った。一人当たりのデータは、図 6 に示す 8 方向の移動方向に対して各方向 5 回分、計 40 サンプルの歩行データである。

これらのデータを 3.2 で説明したフレームに分割し、フレーム処理によってスペクトル分析に基づいた特徴パラメータに変換した。この特徴パラメータに変換されたセンサ信号を各移動方向別に分け、8 方向の移動方向単位でモデルを作成した。作成した HMM は内部状態 7 状態で、自

己遷移と一つ隣の状態への遷移のみを持つ Left-to-Right モデル(図 12)である。学習には 40 サンプル×17 名すべてのデータを用いたため、各モデル当りの学習データ量は 85 サンプルである。

今回の評価では学習データと同じデータを用い、認識率を求めた。

図 13 に認識結果を示す。黒の棒グラフで表されている認識率が、入力したセンサ信号から人の移動方向を正しく認識した確率で、灰色の部分で動線は同じだが方向を反対に認識した確率である(例えば ⑤ の方向のデータを与えたときに、⑥ であると認識したもの)。

学習データと評価データが同一であるクローズな評価であるとはいえ、平均で 90% を超える認識ができており、かなり高い認識率であると言える。方向別に見ると、方向 ② と ⑤ の認識率が特に悪い。これは、使用したセンサのフレネルレンズの配置が原因で、この方向の信号レベルが小さく不安定になり易いためであり、レンズの配置を変えることにより、安定した認識ができるように改良が可能であると思われる。

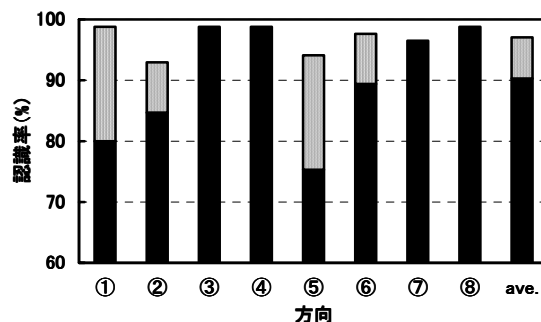


図 13 移動方向認識結果

4.3 人と人以外の識別

移動方向識別と同じような手法にて人と人以外を識別する実験を行った。今回の人以外とは犬で、使用したデータは 3.1 で収集した全データである。

まず、すべてのデータを移動方向識別と同じように特徴パラメータに変換し、大人・子供・大型犬・小型犬の 4 つのカテゴリに分類した。大人は小中学生以上で、子供が幼稚園児である。これは人の歩行データの中で、幼稚園児のセンサ波形が他の年齢のセンサ波形と少し異なるためである。

これら 4 つのカテゴリ内で、移動方向毎に内部状態 7 状態の HMM を作成し、計 32 個の HMM を作成した。学習データは各カテゴリの各移動方向別にランダムに選択したデータで、全データの

約 1/5 を用いた。

この認識の評価にあたっては、学習データとして用いなかった残りのデータ、大人 11200 サンプル、子供 1920 サンプル、大型犬 1920 サンプル、小型犬 1920 サンプルを評価データとして用いた。

図 14 に認識結果を示す。今回の識別は、人と人以外の識別であるため、大人の認識率は、大人が歩行したときのセンサ信号を入力した場合に、大人と認識した回数と子供と認識した回数合わせて算出した認識率になっている。同様に子供の認識率は大人と認識した回数と子供と認識した回数の加算、大型犬の認識率は大型犬と小型犬の加算、小型犬の認識率は小型犬と大型犬の加算でそれぞれ算出した認識率となっている。

この結果を見ると、学習データと評価データが異なるオープンな評価結果であるにも関わらず、平均 97% といった非常に高い認識率が得られており、人へのみ反応して欲しいようなアプリケーションへの応用が期待できる。追記すると、移動方向別にモデルを作成しているため、必ずセンサの真下を通る直線であれば、侵入角度によらず同様な認識が可能である。

4.4 個人の識別

同様な手法にて、3.1 で収集した特定の 3 名 (A,B,C) のデータを用い、個人の識別を試みた。HMM は、各個人単位で移動方向別に 8 つ、計 24 個のモデルを作成した。認識率の評価は、学習データと同じデータを用いて行った。

図 15 に認識結果を示す。今回の識別は個人の識別であるため、A が歩行したときのセンサ出力を入力した場合に、移動方向の正解は問わず、A のモデルが選択された合計で認識率を算出している。これを見ると、平均で 97% 近い認識率が出ており、少人数での個人識別が可能であることを示している。

さらに、図 16 に個人 + 移動方向の識別結果を示す。これは誰がどこを歩いたのか？ といった個人と移動方向の識別結果である。クローズドな評価であるとはいえ、平均で 90% を超える非常に高い認識率が得られており、興味深い結果である。

5. 終わりに

焦電型赤外線センサのアナログ出力には、検知物体の大きさや動きの情報が含まれており、パターン認識手法を適用することで、これらの情報が少なからず得られることがわかった。今回の実験では、移動方向・人と人以外のカテゴリ・個人の識別を行ったが、移動速度やジェスチャー、さらには人間の行動自体を識別できる可能性もある。これらの認識に対しては、今現在画像による認識が主流であるが、赤外線センサでこのような認識

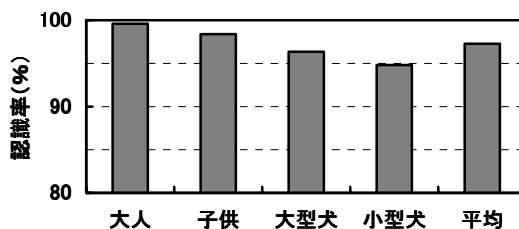


図 14 人と人以外の識別結果

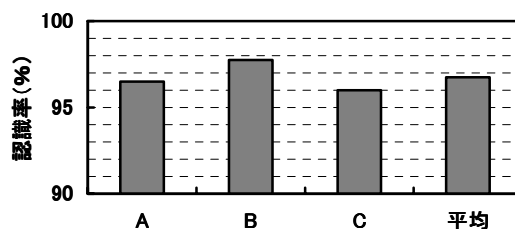


図 15 個人の識別結果

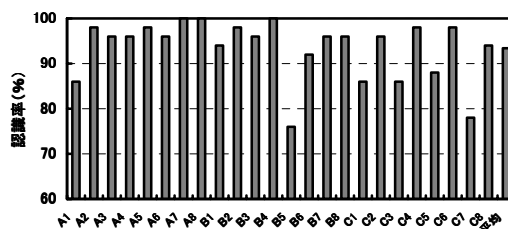


図 16 個人 + 移動方向の識別結果

が可能となれば、コストの削減やプライバシーの確保などメリットは大きい。

今回の結果は、環境の良い条件の下で、歩行経路まで限定したデータによるもので、実用化にはまだまだ多くの問題が残されている。外気温の変化による振幅レベルの変動や、検出範囲境界付近の認識など難しい問題が数多くある。

このような実用化に向けた問題を解決するために、現在、パソコン上で動作するリアルタイム認識システムを作成し、実環境下でのテストを行っている。近々、FPGA を用いたハードウェアの作成も予定しており、最終的には小型マイコン 1 個での実現を目標としている。

参考文献

[1] 松下電工 (株) カタログ, 「焦電型 MP モーシ

ヨンセンサ NaPiOn」, 2004.

[2] M. Shozakai et al., “Analysis of Speaking Styles by Two-Dimensional Visualization of Aggregate of Acoustic Models,” Proc. ICSLP-04, vol.1, pp.717-720, 2004.

[3] G. Nagino et al., “Design of Ready-Made Acoustic Model Library by Two-Dimensional Visualization of Acoustic Space,” Proc. ICSLP-04, vol.4, pp.2965-2968, 2004.

[4] J. W. Sammon, “A nonlinear mapping for data structure analysis,” IEEE Trans. Computers, vol.C-18, no.5, pp.401-409, May 1969.

[5] 古井貞熙, “デジタル音声処理”, 東海大学出版会, 第8章, 1985.

[6] 鹿野清広, 伊藤克巨, 河原達也, 武田一哉, 山本幹雄, “音声認識システム”, オーム社, 第2章, 2003.