

生体データと加速度データを用いた行動認識

佐藤誠[†] 森田千絵[†] 土井美和子[†]

(株)東芝 研究開発センター[†]

1. はじめに

計算機やセンサ、記録デバイスの小型化や無線技術の発展により、人体に装着可能な計算機(ウェアラブルコンピュータ)に様々なセンサを組み込み、人体の動作や生体活動に関する情報を長期間収集することが可能になった。短期間のセンサ情報から、装着者がどのような行動を行っているか推定する技術は、収集されたデータを利用した、治療支援、健康管理支援、装着者の早期異常検知などの様々なサービス実現の基礎となる。

本研究では、日常生活において表れる、「歩く」、「食事する」等の比較的大まかな行動を、腕時計タイプのセンサから収集した加速度データや脈波データ、体温、GSR などを用いて推定することを目的とする。センサ情報に適切な前処理を施し、データマイニング技術の一つである決定木を用いたモデル構築を行うことで、各行動を高精度に識別できることを確認した。

2. データ収集装置

生体情報収集のために人間の身体に装着するセンサには、情報量と装着者の(精神的・肉体的な)負担の間のトレードオフが存在する。身体の様々な部位にセンサを装着すれば、細かい動作を推定するための多くの情報が得られるが、センサの量が多くなればなるほど、装着が許容される状況は限られる。健康管理や在宅医療など日常生活のほとんどの時間ウェアラブル PC を装着しなければならない状況では、腕時計や指輪など通常の生活で身につけている物と同程度の重量や装着感であることが望ましい。

本研究では、図 1 に示す腕時計型の 2 次元加速度センサおよび生体センサ(LifeMinderR)を対象とする[1][6]。この装置を手首および指に装着することにより、前腕と平行な軸(便宜上以下 y 軸と呼ぶ)方向の加速度、手の甲に平行で前腕と垂直な軸(x 軸と呼ぶ)方向の加速度、脈波、GSR(皮膚電気反射)、および皮膚温度を測定することができる(データは 0~1023 の値をとる)。

また、この装置は無線通信機能を有しているので、測定されたデータを一定の時間間隔ごとに PC や PDA 端末などに送信し、装着者の行動に制約を与えず、長時間に及ぶデータ計測を行うことができる。

以下に、現バージョンのセンサから得られるデータの基本的な性質をまとめる：

- ・脈波、GSR、皮膚温度の安静時のベースラインには個人差がある。
- ・脈波、GSR を得るためには、元データを数値処理する必要があり、数値処理はチューニングを要するパラメータを含む。また、加速度、皮膚温度と比較して計測ノイズが生じやすい。

Title: Human action recognition using physiological and acceleration data

[†]TOSHIBA Corp. Research & Development Center



図 1: LifeMinder®の構成。X 軸加速度(AccX)、Y 軸加速度(AccY)、皮膚温度(Temp)、皮膚電気反射(GSR)、脈波(Heart)が計測可能。Bluetooth 通信により PDA 等でデータが送受信可能。

3. 基本行動の認識

前節で述べた装置を用いて、装着者がどのような行動を行っているのか推定することを考える。本研究で扱う行動は、「歩く」、「走る」等の時間的にある程度の連続性があり、かつ日常生活において頻繁に生じる大まかなものであり(以下、基本行動と呼ぶ)、「頭を掻いたか鼻を掻いたか」などといった瞬間的で細かい動作の識別は行わない。また、ある時刻 t の行動を推定する場合、 t より過去のデータのみを使う場合と、 t より未来のデータも使う場合があるが、リアルタイム推定を必要とする応用のため、前者のアプローチを採用する。

本論文の実験では、CART 型の 2 進分類木[4]で行動識別モデルの構築を行った。ノード分割尺度は Gini 統計量を用い、交叉検定を使った postpruning で枝刈りした[5]。全データの 7 割を訓練データとしてモデル化に使い、残り 3 割をテストデータとして予測精度を評価した。

本節では、計測ノイズが生じにくく、複雑な数値処理を必要としないという理由から、加速度データのみを用いて基本行動の識別を行った結果を示す。まず、様々な状況下での識別結果を示し、次に、行動を連続して行った場合の識別結果を示す。

3.1. 多様な状況下での基本行動識別

20~40 代の男女 6 人(男性 5 人、女性 1 人)に対して、以下の状況における加速度データを収集した[3]：

- 「歩く」：平地、平地+装着側の手に(大きめの)鞆、平地+装着反対側の手に鞆、平地+装着側の手に(軽い)ファイル、平地+装着反対側の手にファイル、階段、階段+好きな側に鞆、階段+好きな側にファイル、の 8 種。
- 「走る」：平地、平地+装着側の手に鞆、平地+装着反対側の手に鞆、の 3 種。
- 「立ち止まる」：「走る」と同じ 3 種。
- 「デスクワーク」：PC 作業、書類を読む、書類を書く、自由行動、という 4 種。

これらの状況は、デスクワーカーの平日に行う可能性のある行動を中心にリストアップした。収集されたデータを、20点ずつ区切り(1秒分)、上述の4行動をクラス変数として決定木を構築した。属性変数として、1)生のデータ(40点)、2)フーリエ変換のパワースペクトル係数(20個)、3)以下に述べる10個の簡単な統計量を特徴抽出した場合の3種類を用いた。

ここで、一定長に区切られた時系列からの特徴抽出を簡潔に説明するため、時刻 t を最終時点とした時系列 $P=\{p(t), p(t-1), p(t-2), \dots\}$ の平均を表す関数と標準偏差を表す関数をそれぞれ、 $Ave(p(t))$, $StdDiv(p(t))$ と表す。また、計算対象時系列の長さを明示的に示したい場合には、 $Ave(p(t), N)$, $StdDiv(p(t), N)$ などと表す。上記の3)で属性変数として用いた特徴量は以下の10種類である。

- $Ave(AccX(t))$, $StdDiv(AccX(t))$,
- $Ave(AccY(t))$, $StdDiv(AccY(t))$,
- $Ave(|AccX(t)-AccX(t-1)|)$, $StdDiv(|AccX(t)-AccX(t-1)|)$,
- $Ave(|AccY(t)-AccY(t-1)|)$, $StdDiv(|AccY(t)-AccY(t-1)|)$,
- $Ave(\sqrt{(AccX(t)-AccX(t-1))^2+(AccY(t)-AccY(t-1))^2})$,
- $StdDiv(\sqrt{(AccX(t)-AccX(t-1))^2+(AccY(t)-AccY(t-1))^2})$.

表 1: 前処理による識別精度の比較(5 試行)

	1)	2)	3)
予測精度(%)	76.4	95.9	96.1
標準偏差	0.94	0.27	0.44

```

Ave(AccY(t)) > Th1 [D 45/1165]
| Ave(|AccY(t)-AccY(t-1)|) > Th2 [B 3/30]
| Ave(|AccY(t)-AccY(t-1)|) <= Th2 [D 18/1135]
Ave(AccY(t)) <= Th1 [A 661/1176]
| Ave(|AccY(t)-AccY(t-1)|) > Th3 [A 243/739]
| | ** > Th4 [B 48/258]
| | ** <= Th4 [A 32/481]
| Ave(|AccY(t)-AccY(t-1)|) <= Th3 [C 36/437]

A: 歩く, B: 走る, C: 立ち止まる, D: デスクワーク
** : Ave(sqrt((AccX(t)-AccX(t-1))^2
+ (AccY(t)-AccY(t-1))^2))

```

図 2: 単純化した決定木

モデル化用データと評価用データの分割を異なる乱数を用いて行った5試行の実験結果を表1に示す。この結果から、2)と3)の特徴抽出を行った場合の決定木が同程度の識別精度を示していた。計算量が少ないこと、および、データ欠損などの計測ノイズに頑強な特徴量であることから、現状の識別問題では3)で十分と考えられる。以下では、3)の前処理を用いた場合の典型的な決定木を詳しく調査した結果を示す。

- 図2は構築された決定木の典型例をさらに枝刈りして単純化したものである。図中の各ノードに記した[]の中のアλファベットは、そのノードに対応する行動コードを表し、数値(#1/#2)は、そのノードに分類されたテストデータの数が#2であり、そのうち識別誤りが#1個であることを表す。この単純化された木における識

別プロセスは、 $Ave(AccY(t)) (=Ave(AccY(t,20)))$ によりデスクワークとその他の行動を識別し、 $Ave(|AccY(t)-AccY(t-1)|) (=Ave(|AccY(t)-AccY(t-1)|,19))$ により立ち止まるを識別し、最後に歩くと走るを識別している。

- オリジナルの決定木の葉の数は11であり、各葉ノードの行動の内訳は、{歩く=4, 走る=4, 立ち止まる=1, デスクワーク=2}であった。
- 状況差に関しては、歩くの葉の3/4は階段におけるデータが70%以上あり、走るの葉の2/4は鞆を装着側の手に持った場合のデータが80%以上あるという偏りが見られた。これらの葉に分類されていたデータの割合は、歩く=7%, 走る=24%であった。
- 個人差に関しては、歩くの葉の2/4は特定被験者のデータが70%以上あり、走るの葉の1/4は特定被験者のデータが80%以上あるという偏りが見られた。なお、個人差が生じた葉は状況差のときに挙げられたものにすべて含まれた。これらの葉に分類されていたデータの割合は、歩く=4%, 走る=3%であった。

これらの結果を考察する：

- 腕を机と平衡に保つデスクワークと他の行動は $AccY$ の水準によって区別でき、 $AccY$ の変化の少ない立ち止まるは、歩く・走ると区別できる。
- 階段を歩いている場合と鞆を装着側の手に持って走っている場合のデータが、他の大部分のデータと異なる特徴を示す。ただし、歩くの場合は個人差による影響が強いと考えられる。
- 個人差が強く出ている葉に含まれるデータは共に5%以下になっているので、個人差への過適合を避けるため決定木からこれらの葉を取り除いても、識別性能の損失は小さいと思われる。

3.2. 連続行動識別

前節では、単一の行動を行っている際のデータを用いて、注目時刻の一定範囲の最近データを用いた識別モデルの精度を評価したが、現実には装着者の行動はある程度連続して行われ、その後に変化する。そのため、連続行動の推定では、行動が変化する途中の状態が存在するという困難さがある一方で、誤りの連続が十分小さければ、推定された前後の行動と異なる認識結果をノイズとみなせるという容易さがある。後者を生かすことを考えると、識別誤り率が同じでも時間的に誤りが連続しないモデルのほうが望ましいことになる。そこで、識別モデルの誤りの時間連続性を評価する。

前節の実験で特徴空間の分割が後になった、歩く(A)、走る(B)、立ち止まる(C)、という行動を、

A B C B A C A

という順序で行ったデータを収集した。このタスクでは、A-B-C間のすべての行動遷移が生じるように設定されている。なお、3人(全員男性)の被験者が、平地で手に荷物を持たずに実験を行った。また、各行動は30秒~1分程度継続させ、自然な動作の変化となるように被験者の自発的なタイミングで行動を切り替えさせた。データの延べ時間は14分30秒程度となった。データと時刻を同期させたビデオ映像から手入力で教師信号を生成した。

前節と同様に、収集された加速度データを20点ずつ区

切り(1 秒分), 上述の 3 行動をクラス変数とした. 本節のデータは状況差がないので, 前節のデータよりも性質が単純と考えられるため, 属性変数としては, $\{Ave(AccX(t)), Ave(|AccX(t) - AccX(t-1)|), Ave(AccY(t)), Ave(|AccY(t) - AccY(t-1)|)\}$ の 4 つの特徴量のみを用いて決定木を構築した. 得られた木の典型例を単純化した決定木を図 3 に示す.

```
Ave(AccY(t)) > Th1 [B 316/651]
| Ave(|AccY(t)-AccY(t-1)|) > Th2 [B 4/306]
| Ave(|AccY(t)-AccY(t-1)|) <= Th2 [C 13/345]
Ave(AccY(t)) <= Th1 [A 16/397]
| Ave(|AccX(t)-AccX(t-1)|) > Th3 [B 1/8]
| Ave(|AccX(t)-AccX(t-1)|) <= Th3 [A 9/389]
```

A: 歩く, B: 走る, C: 立ち止まる.

図 3: 単純化した決定木

得られた決定木の識別精度は 97% であり, 行動変化中のデータが混在する場合でも高精度で識別が可能であることが分かった.

次に, 得られたモデルで連続識別した場合の時間的にひとまとまりの誤りを連続誤り数と呼び, 連続誤り数の頻度を評価する. 例えば, 図 4 において, 行動の教師信号を T, T を連続識別した結果を R とすると, 連続誤り数 N の頻度 F(N) は, F(1)=3, F(2)=2, F(3)=1 である.

T: AAAAAAAAAABBBBBBBBBBCCCCCCCC
R: AAAACAACBBBBAABBBBCCBBBCCAC

図 4: 連続識別の例(太字が誤り)

図 5 は図 3 の決定木で連続識別した場合の連続誤り数 (N=N 秒に対応) の頻度である. この結果から, 連続する誤りをひとつの誤りと考えれば, 多くの誤りは 2 秒程度しか連続しないので, 単純なフィルタによって取り除くことができると考えられる.

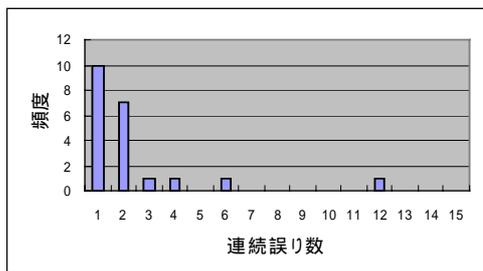


図 5: 連続誤り数の頻度 (行動識別)

4. 食事行動の認識実験

モバイル生体センサとリアルタイムの行動推定技術を組み合わせた健康管理や, 投薬管理, ダイエット支援などの応用を考えると, 「食事」という基本行動の認識は特に重要となる. 実用上最も重要なのは食事を始めた時点や食事を終了した時点の推定と考えられるが, これらは, 短時間のセンサデータから食事行動とそれ以外の行動(以下「食事外」行動とよぶ)を識別する統計モデルを

利用して実現することができる. 例えば, 食事外行動が一定時間推定されている状況で食事行動が認識され, その後も一定時間食事行動の認識が継続されれば, その地点が食事開始時点であると推定でき, 同様の方法で食事終了時点も推定できる.

そこで, 前節までと同様に, 短時間のデータセグメントを特徴抽出した属性変数を用いて食事行動を認識するモデル(決定木)を構築することを考える. 食事行動の認識のため, 生理学上の考察から脈波, GSR, 皮膚温度を用いることにした. また, 行動変化に対する応答の速さを利用するために加速度データも用いることにした(詳しい議論は[2]を参照). 本節では, 生体データ{脈波, GSR, 皮膚温度}のみを用いた場合と, 生体データと加速度データをともに用いた場合を, 識別精度と連続誤り率に関して比較する.

データ収集実験では, 食事タスクを 8 食分(被験者 3 人)行ったのに加え, 食事以外の様々なタスク(以下, 食事外タスクと呼ぶ)も行った. 精神状態と身体の動きがなるべく多様になるように, {クレペリン検査(簡単な一桁の足し算を繰り返すテスト)を行う, 音楽を干渉する, 新聞を読む, PC 作業を行う, 安静にする, お茶を飲む, 電話をかける, 歯を磨く, 読書をする, タバコをすう, 食事のふり(体の動きだけ食事をしているが実際は食べない)}, という 11 種類の食事外タスクを行った.

食事タスクの生体データを観察したところ, 食事前, 食事中, 食後の各フェーズで異なる傾向があったため, 「食後」という行動(状態)を新たに導入し, {食事, 食後, 食事外}の 3 クラス問題としてモデル構築を行うことにした. データセグメントへの教師信号としては, 食事タスクの食事中のものに「食事」を, 食事タスクの食後のものに「食後」を, 食事タスクの食前と食事外タスクのものに「食事外」を付加した.

さらに, 食事タスク実行中の被験者の身体の動きを観察したところ, 手の甲に垂直な成分の動作が比較的多く見られたので, 以下の計算式を用いて Z 軸加速度の推定を行った(詳しい議論は[2]を参照).

$$Z'(t) = 256^2 - (Ave(AccX(t), 80) - 512)^2 - (Ave(AccY(t), 80) - 512)^2$$

$$\text{If } Z'(t) > 0 \text{ Then } Z(t) = \sqrt{Z'(t)}$$

$$\text{Else } Z(t) = 0.$$

そして, Z(t) から特徴抽出した属性変数 $\{Ave(Z(t), 100), StdDev(Z(t), 100), Ave(|Z(t)-Z(t-1)|, 99), StdDev(|Z(t)-Z(t-1)|, 99)\}$ を用いて, ユークリッド距離を距離尺度とした k-means クラスタリングを行い, 得られた各クラスタのクラス分布({食事, 食後, 食事のふり, 食事外(食事のふり以外)})の分布を観察した. その結果, 完全なクラス分類ではないものの, {食事, 食事のふり}行動の比率が全データの比率より明らかに大きいクラスタが存在した. そこで, それらのクラスタに含まれるデータセグメントに 1 を, それ以外のクラスタに含まれるデータセグメントに 0 を付加し, 新たな属性変数(Cluster(Z(t)))として利用することにした.

モデル化の際に用いた生体データからの特徴量は:

- Ave(Temp(t)), StdDev(Temp(t)),
- Temp(t)-Temp(t-L+1),
- StdDev(Heart(t)),

- Ave(|Heart(t)-Heart(t-1)|),
- HeartRate(Heart(t), 1),
- GSRFrec(GSR(t), 2),

ここで、L はセグメント長、HeartRate(*, 1)はパラメータ 1 を用いた数値処理の結果得られた心拍数、GSRFrec(*, 2)はパラメータ 2 を用いた数値処理の結果得られた GSR 頻度である。

加速度データからの特徴量は：

- Ave(AccX(t)), StdDev(AccX(t)),
- Ave(AccY(t)), StdDev(AccY(t)),
- Ave(Z(t)),
- Ave(Cluster(Z(t))).

なお、すべての属性変数は連続変数として扱った。

表 2：食事識別決定木の精度

識別精度 (%)		食事	食後	食事外	全体
加速度なし	訓練データ	88.5	72.6	95.0	93.7
	テストデータ	73.8	42.9	91.4	88.1
加速度あり	訓練データ	88.2	91.8	97.2	96.0
	テストデータ	79.1	74.0	95.0	92.3

表 2 は得られた決定木の精度である。実験結果では、加速度を用いた場合の予測精度が優れていた。特に、食後行動(状態)の認識において、加速度を用いたほうが良いことが分かった。これは、食事を終了した影響が体温などには比較的ゆっくりと表れるが、加速度パターンは即座に変化するためと思われる。ただし、今回のデータは食事外クラスの割合が多いため、全体の識別精度の差は 4%程度となっている。

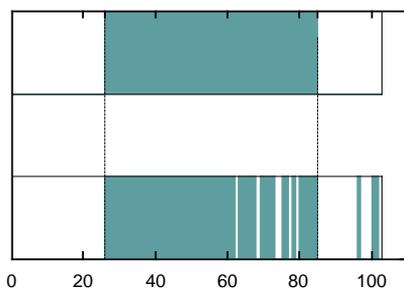


図 6：連続識別の典型例(加速度ありの決定木)。横軸がセグメントでひとつが 7 秒相当。上が教師信号で下が識別結果。塗りつぶされた領域が食事行動。

得られた決定木を用いて食事タスクにおけるデータを連続識別した。以下では、食後行動も食事外行動とみなして評価する。図 6 は、食事タスクの実験結果のひとつに対して、加速度あり決定木を用いて連続識別を行った結果を示している。識別誤りはある程度あるものの、全体としては良好な識別が行われており、食事開始時点や終了時点の推定も容易に行うことができると思われる。

次に、加速度あり決定木と加速度なし決定木の典型例を用いて、全ての食事タスクのデータに対して連続識別を行った場合の連続誤り数の頻度比較結果を図 7 に示す。図 7 の結果からも、加速度あり決定木のほうが誤りの連続性に関して好ましい性質を示している。

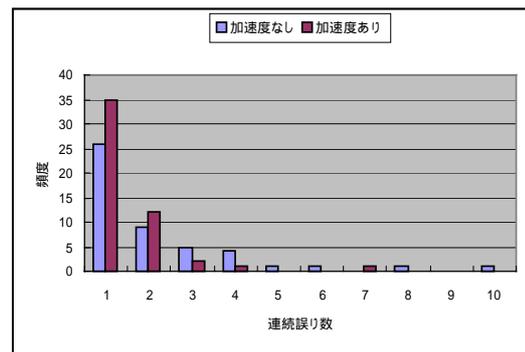


図 7：連続誤り数の頻度(食事認識)。左側が加速度なし。右側が加速度あり。

5. おわりに

腕時計型の生体情報センサーである LifeMinder®から収集されたデータを用いて、日常生活における「歩く」、「走る」、「立ち止まる」、「デスクワーク」、「食事する」という基本行動を識別する決定木を構築し、高精度の行動推定が可能であることを示した。

多くの被験者や状況からのデータを扱い、モデルの頑健性を示すことや、装着者の異常など稀で突発的な状態をすばやく検出できる認識モデルの構築が今後の課題である。

謝辞

本研究の一部は文部科学省の平成 14 年度科学技術振興調整費「人間支援のための分散リアルタイムネットワーク基盤技術の研究」の一環として行われた。

参考文献

- [1] 鈴木 琢治, 大内 一成, 土井 美和子: LifeMinder: ウェアラブル健康管理システム, 電子情報通信学会技術研究報告ヒューマン情報処理, Vol.101, No.699, pp.33-38 (2002).
- [2] 鈴木 琢治他: LifeMinder: ウェアラブル健康管理システム, 第 65 回情報処理学会全国大会(2002) (To appear).
- [3] 森田 千絵, 佐藤 誠, 土井 美和子: 加速度データを用いた行動認識, 情報科学技術フォーラム(FIT2002), pp.415-416 (2002).
- [4] Breiman, L., Friedman, J.H., Olshen, R.A. and Stone, C.J.: *Classification and Regression Trees*, Wadsworth Int. Group, Belmont, California, USA(1984).
- [5] Duda, R.O., Hard, P., E., and Stork, D.G.: *Pattern Classification*, John Wiley & Sons, New York (2000).
- [6] Ouchi, K., Suzuki, T. and Doi, M.: LifeMinder: A wearable healthcare support system using user's context, Proceedings. of ICDCS2002 Workshops (IWSAWC2002), pp.791-792, (2002).