

# 統計的音声区間検出法を用いた加速度センサによる 人間の動作検出

伊藤 真也<sup>†</sup> 浅野 翔<sup>‡</sup> 田村 哲嗣<sup>‡</sup> 速水 悟<sup>‡</sup>

<sup>†</sup> 岐阜大学大学院 工学研究科 応用情報学専攻

<sup>‡</sup> 岐阜大学 工学部 応用情報学科

## 1. はじめに

現在、ライフログに関する研究が大いに注目されている[1][2]. ライフログを取得するためには、人間の動作検出と認識を行う必要があり、不審者や不審行動の検出、ヒューマノイドのための動作認識の研究が行われている[3][4]. この際に、小型で人間の体につけることができ、取得したデータを無線で通信できる加速度センサを使う方法が広く用いられている. しかし加速度センサは、体の微妙な揺れなどから生じる動きも検出してしまいう問題がある. 一方、動作している領域と動作していない領域で、それぞれ別々の処理を施した方がよりよいライフログを構築できると考えられる.

それらの事から、加速度センサの情報から人間の動作区間を検出する必要がある. しかし、加速度センサの情報を用いた動作区間の検出は数少ないのが現状である. そこで本研究では、統計的音声区間検出法[5]を用いて、加速度センサから得られた情報の動作区間の検出を行う.

## 2. 加速度センサ

本研究では、ワイヤレステクノロジー社製の小型無線加速度センサを使用した. 加速度センサの情報は、10ms ごとに Bluetooth 通信により送られる. また、結果は 3 軸(x, y, z 軸)ごとに出力される. この結果は重力加速度の影響を含んでいる. また、加速度センサの情報には、設置した人間の姿勢の情報も含んでいる.

なお、加速度センサは消費エネルギーの測定などに利用され、身体活動測定計として製品化されている[6].

## 3. 統計的音声区間検出法

統計的音声区間検出は、Sohnらが提案した、雑音を含む音声信号から音声区間を統計的手法に基づき検出する手法である[5]. 具体的には、信号をある一定の長さに分割した解析フレームが音声区間 ( $H_1$ ) と非音声区間 ( $H_0$ ) であるかを尤度によって決定する. それぞれの尤度は、解析フレームの信

Human motion detection with 3D-accelerometers using statistical voice activity detection method.

<sup>†</sup>Shinya Ito, Graduate School of Engineering, Gifu University.

<sup>‡</sup>Shou Asano, Satoshi Tamura, Satoru Hayamizu, Faculty of Engineering, Gifu University.

号の DFT 係数の累積正規分布であると仮定し、信号の変化の違いを比較する. それぞれの仮定は次のように示される.

$$p(X|H_0) = \prod_{k=0}^{L-1} \frac{1}{\pi\lambda_w(k)} \exp\left\{-\frac{|X_k|^2}{\lambda_w(k)}\right\} \quad (1)$$

$$p(X|H_1) = \prod_{k=0}^{L-1} \frac{1}{\pi[\lambda_w(k) + \lambda_s(k)]} \exp\left\{-\frac{|X_k|^2}{\lambda_w(k) + \lambda_s(k)}\right\} \quad (2)$$

ここで  $X_k$  は観測信号の DFT 係数であり、 $\lambda_w(k)$  と  $\lambda_s(k)$  はそれぞれ雑音信号と音声信号の分散を表している. また  $L$  はフレームの長さである.

そして、音声区間と非音声区間の尤度比  $\Lambda$  を求め、閾値処理を行い、音声区間の検出を行う.

統計的音声区間検出では、雑音信号があらかじめ既知である必要がある. 加速度センサによる人間の動作検出では、雑音となる信号のパターンがある程度推測でき、統計的音声区間検出法は有効であると考えられる. また、加速度センサの値は重力加速度を含んでいるが、統計的音声区間検出法はその影響を受けないという利点がある. 以上の特性から、人間の動作検出を行う際に発生する体の揺れを背景雑音と考えると、統計的音声区間検出法を用いることによって軽減することができるのではないかと考えられる.

## 4. 実験

3 軸の情報を 1 つの結果に統合することを考える. まず、それぞれの軸毎に  $\Lambda$  を求める. 次に、3 軸分の  $\Lambda$  を足し合わせたものを新たな  $\Lambda$  として、それに対して閾値処理することによって、最終的な結果を算出することとする.

加速度センサは、腰(前太ももの上)に設置した. 検出動作には、以下に示す計 5 種類を用いた. 動作検出はそれぞれの動作に対して 20 回程度、一人の被験者に対して測定し、解析を行った.

- 歩く
- ベッドに寝る
- 椅子に座る
- ベッドから起きる
- 椅子から立つ

実際に測定した結果を時系列に表示したものを図 1 上図に示す. 縦軸は加速度の大きさである.

図 1 上図に示した測定結果は、停止している区間

は、出来る限り動かないように測定されている。そのため、本来の目的である体の微妙な揺れなどの検出したくない動作は含まれていない。よって、実際の状況で想定される揺れの動きを別に測定し、図1上図に示した測定データに揺れのデータを付加したものについて動作検出を行うこととする。18, 15, 12, 9dBのS/N比になるように揺れの動きを測定した結果に付加した。特に、15dBのときの結果を図1下図に示す。

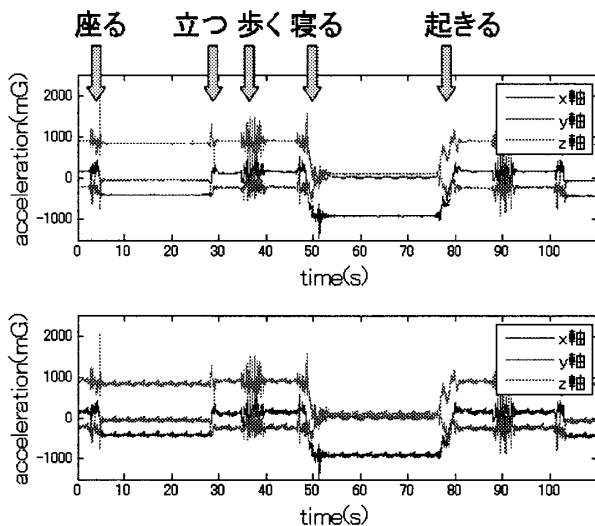


図1 各動作の測定結果

(上: 測定したもの, 下: 15dBの揺れを付加したもの)

また、動作検出の評価には音声区間検出で広く使われているFRR(False Rejection Rate)とFAR(False Acceptance Rate)によって評価する[7]。また、正解のフレームは、測定した波形と測定中に撮影した映像とを比べて判断した。FRRとFARは次のように表される。

$$FRR = \frac{\text{動作フレームが停止フレームと判断された数}}{\text{動作フレームの総数}} \times 100$$

$$FAR = \frac{\text{停止フレームが動作フレームと判断された数}}{\text{停止フレームの総数}} \times 100$$

実際に動作検出を行ったFRRとFARの関係を図2に示す。図2から分かるように、FRRとFARはトレードオフの関係になっている。そのため、本実験では、FRRとFARの合計が最小になるところを最適な結果とする。それぞれの結果を表1に示す。

## 5. 考察

4章の結果から、12dB以上では、ある程度正確に動作検出を行えていることが判明した。しかし、9dBのときは、12dB以上のときに比べて大幅に誤りが大きくなっていった。そのため、実際に動作検出する際には、一定以上の揺れを考慮して設計する必要があると考えられる。

また、本実験では揺れを後から付加することによ

って、実際の動作を再現したが、実際の状況で測定し解析することはしていない。そのため、実際の状況で実験を行う必要があると考えられる。

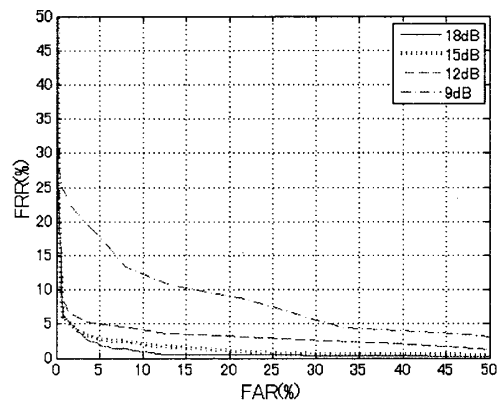


図2 FRRとFARの関係

表1 それぞれのレベルの最適な結果 (%)

	18dB	15dB	12dB	9dB
FRR	2.8	4.2	7.1	13.4
FAR	3.4	2.3	1.2	8.0
合計	6.2	6.5	8.3	21.4

## 6. まとめ

本研究では、統計的音声区間検出の手法を用いて加速度センサの情報から動作検出を試みた。実験の状況で想定される揺れの動きを付加して動作検出を行った。実験により、12dB以上の場合、誤り率10%未満の結果を得ることができた。今後は、動作検出した結果を用いて、動作認識や姿勢の認識を行う予定である。

## 参考文献

- [1] Microsoft: My LifeBits Project, <http://research.microsoft.com/barc/mediapresence/MyLifeBits.aspx>
- [2] KDDI: 情報家電プラットフォーム「いえもん」, [http://www.kddi.com/corporate/news\\_release/2005/1128/besshi.html](http://www.kddi.com/corporate/news_release/2005/1128/besshi.html)
- [3] 服部加奈子, 豊嶋伊知郎, 板倉豊和, 吉田琢史, 折原良平: 時系列の位置データを用いた人間行動分析, JSAI, 2007
- [4] 岡田慧, 鈴木義久, 國吉康夫, 稲葉雅幸, 井上博充: 視覚による人間動作認識と全身行動表現に基づくヒューマノイドの行動獲得, 記憶, 再現, 日本ロボット学会, 2001
- [5] Jongseo Sohn, Nam Soo Kim, Wonyong Sung: A Statistical Model-Based Voice Activity Detection, IEEE SIGNAL PROCESSING LETTERS, VOL. 6, No. 1, JANUARY, 1999
- [6] 松下電工株式会社: アクティマーカー, <http://www.mew.co.jp/metabo/index.html>
- [7] IPSJ SIG-SLP 雑音下音声認識評価WG, CNSREC-1-C 添付資料