

ActivityAnalyzer:携帯電話搭載センサによる リアルタイム生活行動認識システム

大内 一成[†] 土井 美和子[†]

社会の高齢化に伴い、高齢者・障がい者など社会生活弱者の家庭内の生活状態を見守ることの重要性が増してきている。我々は、環境側へ機器を設置せず、携帯電話などの小型情報端末にすでに搭載されている加速度センサとマイクのみを活用した生活状態見守りを検討している。小型端末上での処理を念頭に置き、計算量を抑えた認識手法として、まず加速度センサを活用してユーザの動作状態を大まかに推定し、推定した状態に応じてマイクを起動して環境音の分析を行う方式を提案する。胸ポケットに格納した試作デバイスと PC の構成でリアルタイムに生活行動を認識する「ActivityAnalyzer」を開発し、実際の家庭生活のデータを収集して評価した結果、加速度センサのみで歩行、作業、安静の 3 状態を概ね 95% 以上の精度で分類でき、環境音の分析により、「歯磨き」、「電気シェーバーによる髭剃り」、「ドライヤーの使用」、「トイレ水洗/手洗い」、「掃除機がけ」、「皿洗い」、「アイロンがけ」の各作業を 75.8% の精度で、本人の学習データを用いた場合は 85.9% の精度で分類できた。

A real-time living activity recognition system by using sensors on a mobile phone

KAZUSHIGE OUCHI[†] and MIWAKO DOI[†]

In aging societies, such as that of Japan, there is a growing need to watch over elderly people. We propose a low-throughput recognition technique for in-home context awareness using only off-the-shelf sensors, namely, an accelerometer and a microphone, which are commonly applied in mobile phones. The proposed technique firstly estimates a user's movement condition roughly by acceleration sensing. Secondly, it classifies the working condition in detail by acoustic sensing based on the estimated condition. As a result of our experiment, three movement conditions are classified with more than 95% accuracy by acceleration sensing: quiet, walking, and working. Moreover, it classified working into brushing teeth, shaving, drying the hair with a blower, flushing the toilet, vacuuming, washing the dishes, and ironing with 75.8% accuracy by acoustic sensing and improved the accuracy to 85.9% by training subject's own data.

[†] 株式会社東芝 研究開発センター
Corporate Research and Development Center, Toshiba Corporation

1. はじめに

加速度センサ、GPS (Global Positioning System) など、ユーザの状況、位置などの把握に活用可能なセンサが、携帯電話など小型の情報端末へ搭載され、ユーザの現在の状況に基づいて適切なサービスを提供する状況依存型サービスの普及が期待されている。一方、社会の高齢化に伴い、高齢者・障がい者など社会生活弱者の家庭内の生活状態を見守ることの重要性が今後更に増してくると考えられる。家庭内の生活状態見守りのためには、部屋間の移動、食事、掃除、炊事、洗濯、入浴、排泄など、様々な生活行動の認識が必要となるが、屋内では GPS による測位が使えず、また単一の加速度センサでこれらの状態を把握することは困難であるため、上述の小型情報端末上の状況依存型サービスとは別のアプローチがこれまで試みられてきた。例えば、家庭内の至る所に各種センサを配置し、複数のセンサ情報を統合してユーザの行動を判断する取り組みがある[1]-[3]。専門の施設等ではこのような方式による実運用の可能性も考えられるが、一般の家庭に持ち込むためには、設置コスト、運用コストが大きな問題となる。これらに対し、我々は、携帯電話などの一般的な小型情報端末にすでに搭載されている加速度センサとマイクによる家庭内の生活状態見守りを検討している。

本論文では、まず、小型情報端末を用いた状況認識に関する関連研究を概観し、我々の目指す方向性を明らかにする。次に、加速度センサの情報を活用してユーザの動作状態を大まかに分類し、推定した状態に応じて環境音を分析して詳細な作業状態を分類する提案手法について説明する。続いて、加速度データ、音データを同期記録可能な試作デバイスと PC の構成でリアルタイムに生活行動を認識する「ActivityAnalyzer」を開発し、実際の家庭生活のデータを収集して、提案手法の基本性能を確認した結果を述べ、最後に考察する。

2. 関連研究

身体に装着した加速度センサによりユーザの動作状態を認識する研究は多い。例えば、加速度センサを左右の肩、肘、手首、腰、膝、足首に計 12 個装着した研究[4]や、上腕、手首、腰、太腿、足首に計 5 個装着した研究[5]などがあり、多いものでは 20 種類の動作状態の推定を行っている。複数の加速度センサを装着することで、詳細な動作状態の推定が可能となるが、日常生活でこれらのように多数の加速度センサを常時身につけることは、拘束性が高く受け入れがたい。

これに対し、近年、携帯電話への加速度センサの搭載が進んでいる。このような背景のもと、携帯電話内蔵の単一の加速度センサを活用した動作状態認識の研究も行われている。例えば、携帯電話内蔵の 3 軸加速度センサのデータを基に、歩行、速歩、階段上り、階段下り、走行の 5 動作状態を約 80% の精度で認識する研究[6]がある。他の研究[7][8]でも歩行、走行を含む 4 種類程度の動作状態推定を行っている。我々がタ

ターゲットとしている家庭内の生活状態見守りのためには、歩行、走行などの動作状態だけでなく、家事などの作業状態も認識対象とする必要があるが、単一の加速度センサだけで作業状態を推定することは困難である。

一方、携帯電話には通話用のマイクがあるため、これをセンサとして活用することも可能である。環境音の分析による状況推定に関する研究としては、音特微量のみで屋内外 17 種類の生活シーンを平均 68.4%の精度で分類する取り組み[9]などがある。しかし、加速度センサのデータは、例えば 10bit 20Hz 程度でサンプリングすれば良いが、文献[10]では 16bit 48kHz で音データをサンプリングして分析している。環境音分析では、加速度の分析に比べて扱うデータ量が 3 桁以上増え、また、周波数成分の抽出のためには FFT (Fast Fourier Transform) を実行するなど、ある程度の計算量を要する処理が必要となり、環境音分析を常時行うこと（音声認識を常時動かし続けることに相当する）はできれば避けたい。また、音声認識では、入力意図のある発話だけ切り出して認識することが音声認識精度向上に有効である[10]ことから、環境音分析を行う場合においても、分析すべき区間を切り出して認識させる手法が有効であると考えられる。そこで、我々は加速度センサでユーザの動作状態を大まかに推定し、推定した状態に応じて環境音分析を実施する手法を提案する。

携帯電話に搭載された加速度センサとマイクを活用した状況認識に関する研究としては、加速度センサ、マイク、GPS を用いてユーザの移動状態を推定する報告がある[11]。歩行、歩行、停止に加えて自転車、電車、バス、自動車などの移動状態を推定する場合、加速度センサだけでは精度が低下するため、マイクからの環境音と GPS の位置情報とあわせて分析することで精度向上を図っている。具体的には、加速度センサで移動状態が電車、バス、自動車のいずれかであると判断した場合にマイクを使って自動車であるかそれ以外であるかを判断する。その後、電車とバスの判断には GPS から得られる位置情報の時間変化を使っている。他にも携帯電話に複数のセンサを搭載してユーザの状態を推定する試み[12][13]が行われているが、文献[12]では携帯電話のマナーモードへの自動切り替えなどを行うためにユーザの状態を Normal / Idle / High Activity / Uninterruptible の 4 つに分類し、文献[13]では動作状態としては歩行、走行、着席の 3 状態を推定しているなど、家庭内生活状態の推定は行われていない。これに対し、我々のターゲットは家庭内の生活状態見守りであるため、推定対象は移動状態だけでなく家事などの生活状態も含む点異なる。

また、携帯電話ではなくウェアラブルセンサを用いた行動認識に関する研究もある[14][15]。加速度センサ、マイクに加えて、カメラが搭載されたウェアラブルセンサを手首に装着して、歯磨き、掃除機がけ、皿洗い、コーヒーを作る、ココアを作るなど、15 種類程度の生活状態の認識を行っている。カメラで操作対象物の画像の特徴を捉えることで、コーヒーを作る、ココアを作るなど、加速度と音だけでは違いの判断が困難な行動の識別も可能になる点が特徴である。これらの取り組みに対し、我々は、特

別なデバイスを必要とせず、またカメラを使わず低消費電力で、日常的に使われている携帯電話単体による家庭内の主要な生活状態の認識を目指している点が異なる。

3. 加速度と音による家庭内ユーザ状況認識手法

3.1 認識対象とする生活行動

一口に生活行動といってもその内容は多岐にわたる。このため、食事、着替え、移動、排泄、整容、入浴など生活を営む上で欠かせない基本的な活動を指す日常生活動作 (ADL: Activities of Daily Living) と、炊事、掃除、洗濯などの家事全般や、金銭管理、服薬管理、外出など、ADL より複雑で高次の活動を指す手段的日常生活動作 (IADL: Instrumental Activities of Daily Living) による評価が高齢者介護やリハビリテーションの分野で用いられている。高齢者・障がい者の生活自立度評価の際には、バーセルインデックス^{*}に代表される ADL の評価だけでは不十分で、IADL も重要な指標であるとされている[16][17]。よって、我々が目指す小型情報端末による家庭内生活状態見守りでは、ADL、IADL の両方を認識対象とする。

3.2 想定するサービスイメージ

上述の通り、日常生活において負担なく ADL、IADL を継続的にモニタリングすることは、特に高齢者、障がい者、認知症患者などにとって重要であるが、対象者によってモニタすべき ADL、IADL は異なることが考えられる。例えば、特定の部位に身体的な障がいを持っているだけのひとと、認知症のひとでは、モニタすべき ADL、IADL は自ずと異なる。また、家電機器 (掃除機、電気シェーバーなど) の動作音、キッチンやトイレの水洗音などは家庭によって異なり、歯磨きに一般的な歯ブラシを使用するか/電動歯ブラシを使用するかなど、同じ生活行動において発生する音は対象者や環境によって異なる。あらゆる環境に適用可能な汎用性の高い生活行動認識の実現が最終的には理想であるが、本研究では、次のようなサービスイメージを想定した場合に有用であると考える手法を提案する。

まず、独居の高齢者や、障がい者、あるいは認知症患者 (以下、対象者とする) などの生活状態をモニタリングするため、その家族あるいはケア関係者が、対象者に該当する ADL、IADL のデータを収集する。つまり、立ち会いの下で認識対象の各生活行動を行ってもらい、その際に収集したデータをシステムに学習させ、各生活行動を認識させる。その際、学習に必要な時間が長いと双方にとって負担になるため、学習はできるだけ短い時間で完了することが必要である。モニタする本人の実際の生活行動のデータを学習するため、家庭による音の違い、使用する機器による音の違いがなく、短い学習時間であっても、精度良い認識が可能になることが期待できる。認

^{*} Barthel Index: 食事、車椅子からベッドへの移動、整容、トイレ動作、入浴、歩行、階段昇降、更衣、排便、排尿の ADL10 項目を 2~4 段階で評価

識結果は、リアルタイムにサーバへアップロードして家族あるいはケア関係者がいつでもモニタリングできるようにする、あるいは携帯電話内に記録して後から参照するなど、用途によって必要なリアルタイム性に応じた構成とする。

上記サービスを実現するため、本論文ではまず、ADLの移動の基本行動としての「歩行」、整容に含まれる「歯磨き」、「電気シェーバーによる髭剃り」、「ドライヤーの使用」、排泄と関連する「トイレ水洗/手洗い」、および IADL の家事に含まれる「掃除機がけ」、「皿洗い」、「アイロンがけ」を対象として、提案手法による家庭内ユーザ状況認識の適用可能性を検証する。

3.3 処理概要

本論文で提案する家庭内ユーザ状況認識手法の処理概要を図1に示す。加速度センサ内蔵携帯電話では、端末の向きによって画面の切り替えを行うなど、ユーザの動作に基づいた操作を常に行うことができるように加速度センサは常時 ON になっている。本手法では、この常時 ON の加速度センサのデータにより、まず動作状態を「歩行」「作業」「安静」の3つに分類する。「作業」とは、歩行以外に、身体的動きを伴う何らかの生活行動をしていると想定される状態のこととし、「安静」とは「歩行」も「作業」もしていない、身体を動かしていない状態のこととする。尚、本論文では「階段昇降」も「歩行」とすることとした。家庭内の生活状態としては部屋間の移動がわかれば、生活状態の切り替わりを検出する手がかりとなるため、今回はこれらをすべて「歩行」とすることとした。

次に、「作業」と分類された場合にはマイクを起動し、環境音を分析して作業内容を推定する。以下、加速度による動作状態分類、音による作業状態分類について、詳しく説明する。

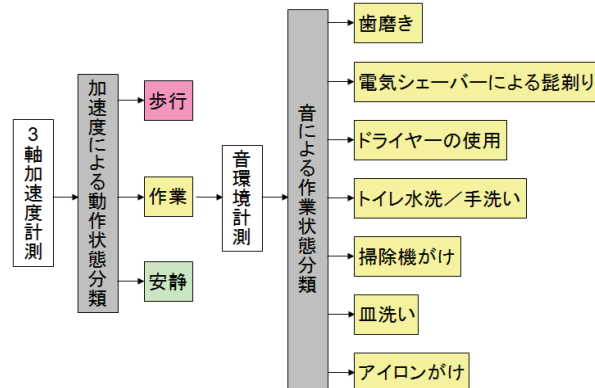


図1. 提案手法の処理の流れ

3.4 加速度による動作状態分類

加速度による動作状態分類は、上述の通り常時動作させ続ける想定であるため、できるだけ計算負荷の低い処理にする必要がある。そこで、統計量として1秒間の分散のみを用いる簡便な手法を用いることとした。処理の流れを図2に示す。

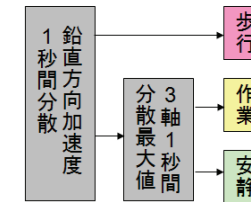


図2. 加速度による動作状態分類

ユーザの胸ポケットに入れたデバイス（本研究では携帯電話を想定している）に内蔵した3軸加速度センサの測定データを用いる。文献[9]にあるように一定の時間幅内（2秒）の平均加速度方向が重力加速度方向と推定し、重力加速度方向の加速度の1秒間の分散を算出する。これは歩行時に最も顕著に加速度変化が現れるのが重力加速度方向であるため、この分散の大きさによって「歩行」か否かを判断する。ここでは、閾値を実験的に $0.05[G^2]$ と定めた。

「歩行」以外の場合には、3軸の各加速度の1秒間の分散を算出し、最大の分散を持つ軸の値の大きさによって、「作業」か「安静」かを弁別する。その閾値は $0.0001[G^2]$ とした。これは、平方根を取って標準偏差とすると $0.01[G]$ となり、睡眠・覚醒判別の際に体動の有無を検出する加速度の大きさとして広く使われている基準値と同一である[18]。以上のように、3軸加速度の1秒間の分散のみを用いる簡便な手法により、「歩行」、「作業」、「安静」の3種類の動作状態进行分类する。

3.5 音による作業状態分類

動作状態が「作業」と分類された場合に、マイクを起動し音による作業状態分類を行う。処理の流れを図3に示す。音による作業状態分類では、事前に分類対象の作業状態の音データを学習しておく。分類器には、高い汎化性能を持つサポートベクターマシン（SVM: Support Vector Machine）を用いることにした。

音データから抽出する特徴量としては、広く音声認識に使われているメル周波数ケプストラム係数（MFCC: Mel Frequency Cepstral Coefficients）に加え、算出に際して計算量の比較的小さい二乗平均平方根（RMS: Root Mean Square）、ゼロ交差率（ZCR: Zero-Crossing Rate）の使用をあわせて検討した。MFCCは人間の聴覚上重要な周波数成分を強調した特徴量で、文献[9][14][15]など音を活用した状況認識に関する関連研究でも使われており、我々もMFCCを使用することとした。また、RMSは主に音の

強度を表し、ZCR は当該区間の音高（ピッチ）に相当する特徴量で、これらは MFCC に含まれない音の特徴を表すため、これらを MFCC に追加することによる性能向上を期待した。今回用いる音データは 16bit 16kHz サンプリングのデータを用いることとし、各特徴量算出の際のウィンドウ幅は 512、ウィンドウオーバーラップは 0 とし、1 秒ごとにそれぞれの平均値を出力するようにした。

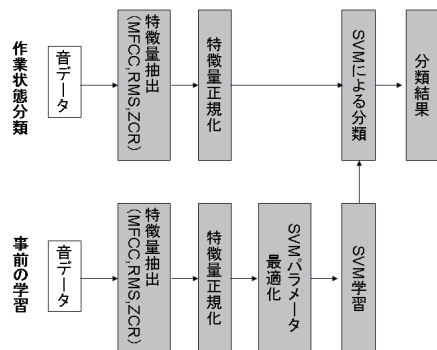


図 3. 音による作業状態分類

MFCC の算出方法自体の説明はここでは割愛するが、FFT、メル周波数フィルタ、log、離散コサイン変換（DCT）などの処理が行われるため、ある程度の計算コストが必要となる。よって、本手法では環境音分析による状況推定を常時実施せず、計算コストが少ない加速度による動作状態分類をまず行い、その結果「作業」と分類された場合のみ音による作業状態分類を行う。実運用時にはさらにある程度の尤度を持って作業状態を推定できた場合は、音による作業状態分類を一旦終了させて、その作業状態が継続しているものとし、作業状態から一旦別の動作状態（安静、歩行）が発生して、その後再度作業状態に移行した場合に改めて音による作業状態分類を行うようにするなど、さらなる計算量、消費電力削減の工夫が必要になってくると考えられる。尚、MFCC の次数は 13 とした。

N 個の音データ $\{a_1, a_2, \dots, a_N\}$ の RMS (a_{rms}) は、

$$a_{rms} = \sqrt{\frac{1}{N} \sum_{i=1}^N a_i^2}$$

で算出する。同じく ZCR (a_{zcr}) は、

$$a_{zcr} = \frac{1}{N-1} \sum_{i=2}^N F\{a_i a_{i-1} < 0\}$$

で算出する。ここで、関数 $F\{A\}$ は、A が TRUE の場合に 1 となり、そうでない場合に 0 となる関数である。このように、RMS, ZCR は、MFCC に比べて計算量が少なく、MFCC に含まれない特徴を表す特徴量である。抽出した合計 15 次元（MFCC : 13 次元, RMS : 1 次元, ZCR : 1 次元）の特徴量は、それぞれでスケールが異なるため、 ± 1 で正規化を行う。

4. 評価実験

4.1 リアルタイム生活行動認識システム

前章で提案した手法の妥当性を評価するためには、実際の生活状態の加速度と音を同期して記録し、分析する必要がある。最終的には携帯電話などの小型機器単体での状況推定を目指しているが、そのための基礎的なデータ収集を効率的に行うことを目的に、3 軸加速度センサとマイクを搭載し、両センサのデータを microSD カードに同期記録できる評価用デバイスを試作した。デバイスの外観と概略仕様を図 4 に示す。



3 軸加速度センサ	LIS344ALH (STMicroelectronics) ±2G/6G 10bit 20Hz サンプリング
MEMS マイク	SPM0408HE5H (Knowles) 8/16bit 8/16/32kHz サンプリング
記録メディア	microSD カード
オーディオ/ファイル制御用 CPU	ARM Cortex-M3
システム制御用 CPU	PIC18LF4685 (Microchip)
オーディオ LSI	ML2308 (OKI)
通信機能	Bluetooth SPP / A2DP

図 4. 試作した評価用デバイスと概略仕様

データ収集時には、本デバイス単体で動作して microSD カードに加速度センサとマイクのデータを同期記録する。また、Bluetooth モジュールを搭載しており、PC などの外部機器に加速度データは SPP (Serial Port Profile) で、音データは A2DP (Advanced Audio Distribution Profile) で PC に送信し、PC で 3 章に示した一連の処理を実行してリアルタイムに生活行動を認識する「ActivityAnalyzer」を開発した。認識結果表示の一例を図 5 に示す。

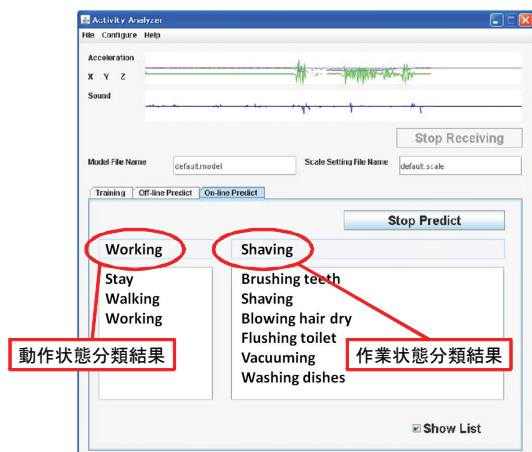


図 5. リアルタイム生活行動認識システム「ActivityAnalyzer」

4.2 データ収集実験

試作した評価用デバイスを胸ポケットに入れ、一般の家庭における日常生活のデータ収集を実施した。被験者 4 名（60 代男女、30 代男女、それぞれ 1 名ずつ）に対して行動リスト（被験者毎に順番を入れ替えた）を提示し、それに従って一通り行動してもらい実験を、日を変えて 2 日分実施し、被験者 1 名あたり 2 データセット、計 8 データセットを収集した。分類対象とした生活状態と、それぞれの平均収集時間は表 1 の通りである。連続して収集したデータから、対象の生活状態に該当する部分のデータを切り出し、分析対象とした。尚、「歯磨き」は電動歯ブラシではなく一般の歯ブラシによる歯磨きのデータを収集した。「電気シェーバーによる髭剃り」は男性被験者 2 名のみ実施した。また、排泄中は「安静」と判別される可能性が高く、排泄後にトイレの水洗と手洗いをする際に「作業」となることが予想されることから、今回のデータ収集では、被験者に便座に座ってもらい、そこから立ち上がってトイレを水洗し、手を洗ってもらうように指示し、その際のデータを収集した。

3.2 節にも述べた通り、本研究では、できるだけ短時間の学習データで、精度良い生活行動認識を実現する点を目的としている。よって、「作業」の 7 種類の分類性能の評価方法は、収集した全データセットのうち 1 データセットをテスト用とし、残りを学習用として、全データセットが 1 回ずつテスト用となるように評価を実施する Leave-one-out Cross-validation とは異なる評価方法を用いることとした。つまり、収集した全 8 データセットの生活状態データのうち、1 データセットから無作為に抽出した 10 秒分のデータを学習用とし、それ以外の 7 データセットをテスト用として分類す

る。次に別の 1 データセットから再度任意の 10 秒分を新規の学習用データとして抽出して同様のテストを行う。これを全データセットから 1 回ずつ学習用データを抽出するようにして評価を行った。

最適な学習時間については、実運用時の観点も含めた検討が今後必要であるが、3.2 節で述べた、我々が想定するサービスイメージのためには、極力短い時間で学習できるようにすべきであると考え、それぞれ 10 秒分ずつという少ない学習データで分類を試みた。分類器に汎化性能の高い SVM を選択したのはこのためでもある。

表 1. データ収集した生活状態とその平均時間

生活状態		平均時間 (分)
歩行 (階段昇降も含む)		1.15
作業	歯磨き	1.14
	電気シェーバーによる髭剃り	0.88
	ドライヤーの使用	1.48
	トイレ水洗/手洗い	0.45
	掃除機がけ	3.18
	皿洗い	1.55
	アイロンがけ	1.67
安静 (TV 鑑賞)		3.24

4.3 結果

4.3.1 加速度による動作状態分類性能

加速度による動作状態分類（「歩行」、「作業」、「安静」のいずれかに分類する）の結果を表 2 に示す。数値は、分析対象の生活状態の全データのうち、正しくその動作状態（「歩行」、「作業」、「安静」）に分類されたデータの割合、すなわち適合率を示す。「歯磨き」、「電気シェーバーによる髭剃り」など作業状態に含まれる各生活状態は、ここでは「作業」に分類されることが正解となる。過去 1 秒間の分散を 1 サンプル (50ms) ごとに算出し、その都度動作状態分類を実施した。加速度の 1 秒間の分散のみを用いた分類法で、概ね 95%以上の精度で「歩行」、「作業」、「安静」を分類できていることがわかる。尚、今回は 4 名の被験者全員で 3.4 節に示した同一の閾値を適用して表 2 の結果を得たが、対象者の身体状態（年齢、障がいの有無）等によっては、学習時に適切な閾値設定が必要になる可能性がある。

「掃除機がけ」の際には掃除機をかけながら移動する際の歩行も発生するが、掃除機をかけながらの歩行の大半は「作業」と分類された。3.4 節で説明した手法で、掃除中において姿勢がかがめたりしながらの移動は、部屋移動のための歩行と区別されることがわかった。

表 2. 加速度による動作状態分類性能

生活状態		動作状態分類性能 (%)	
歩行 (階段昇降も含む)		95.4	
作業	歯磨き	98.6	97.4
	電気シェーバーによる髭剃り	99.5	
	ドライヤーの使用	92.9	
	トイレ水洗/手洗い	95.5	
	掃除機がけ	97.6	
	皿洗い	98.5	
	アイロンがけ	99.5	
安静 (TV鑑賞)		97.5	

4.3.2 音による作業状態分類性能

音による作業状態分類の結果を表 3 に示す。MFCC, RMS, ZCR の各特徴量を 1 秒毎に算出し、それを用いて 1 秒単位で作業状態分類を実施し、正しくその作業状態に分類されたデータの割合 (適合率) を示す。作業状態ごとに多少の性能のバラツキはあるが、任意の 10 秒間の学習データで平均 75.8% の分類性能があることが確認できた。

「電気シェーバーによる髭剃り」、「ドライヤーの使用」については、90%以上の精度で分類できているが、その他の生活状態については、いずれも 70%程度の精度であった。原因としては、「歯磨き」については、電動ではない歯ブラシによる歯磨きのデータを収集したが、歯磨きで発生する音の特徴が被験者によって異なるためであると考えられる。「トイレ水洗/手洗い」、「皿洗い」は、それぞれ水洗時の水量を指定せずに被験者任せとしたため、被験者によってバラツキが出たと考えられる。「掃除機がけ」についても同様に動作モードは被験者任せとした。極力自然な生活状態のデータを収集することを主眼に置いて実施したことによる影響ではあるが、逆に言えば、水量や、動作モードが異なるデータを学習しても、約 70%程度の精度で同一の作業状態として分類できることがわかった。

しかし、本研究では、3.2 節で述べた通り、対象者本人のデータを学習データとして用いることを当面のターゲットとして考えているため、今回収集した各被験者の 2 データセットのうち、1 データセットから任意の 10 秒間を抽出して学習データとし、別の 1 データセットをテスト用として評価し、次に逆の組み合わせで同様のテストを実施した場合とあわせて分類性能を評価した。表 4 にその結果 (全被験者の平均) を示す。本人の学習データを用いることによって、全体として約 10 ポイント、性能が向上できていることがわかった。想定したサービスイメージに対する手法として、提案手法が有用であることが確認できた。

表 3. 音による作業状態分類性能 (全被験者データによる評価)

生活状態 (作業)	作業状態分類性能 (%)
歯磨き	68.3
電気シェーバーによる髭剃り	91.4
ドライヤーの使用	91.6
トイレ水洗/手洗い	71.7
掃除機がけ	70.2
皿洗い	67.2
アイロンがけ	70.3
平均	75.8

表 4. 音による作業状態分類性能 (本人データの学習による評価)

生活状態 (作業)	作業状態分類性能 (%)
歯磨き	73.0
電気シェーバーによる髭剃り	91.9
ドライヤーの使用	96.1
トイレ水洗/手洗い	94.5
掃除機がけ	87.0
皿洗い	81.9
アイロンがけ	77.2
平均	85.9

5. 考察

5.1 音のみによる分類との比較

提案した加速度と音を組み合わせた家庭内ユーザ状況認識手法が有用であるかは、従来の加速度のみ、あるいは音のみでの状況認識手法と比較して検証する必要がある。文献[7]-[9]などの関連研究で単一の加速度センサで認識対象とする動作は、現状、歩行、走行など、移動を中心とした数種類の動作状態だけで、本論文で対象とするその他の作業状態を認識することは困難であるため、加速度センサだけによる作業状態分類についてはここでは比較の対象としない。従って、ここでは音だけで同様の生活状態分類を行う場合との比較検証を行う。

提案手法では、加速度センサで「作業」と分類した動作状態のみ、音によりさらに細かく作業状態分類を行っている。作業状態の分類は音だけを用いているので、比較

すべきは、加速度センサで分類している「歩行」、「安静」となる。つまり「歩行」、「安静」の分類を音で行った時と、加速度で行った時とを比較すればよい。

3.4節で述べた音による作業状態分類法と同じ方法で「歩行」「安静」を認識できるか確認した。一般家庭の廊下を歩行している時(歩行)と、TVを鑑賞している時(安静)の音データについて、任意の10秒分を切り出し、同一の音特徴量(15次元)データをそれぞれ「歩行」「安静」クラスとして他の作業状態に加えて学習させた。テストデータは学習データと同じ場所(一般家庭の廊下)を歩行した際、および同一のTV番組(スポーツ中継)を視聴している際の音データ(歩行:1分、安静:3分)とした。場所が異なったり、番組が異なったりすると、環境音の変化により、著しく性能が劣化することが考えられるため、今回の比較では同じ場所の歩行、同一TV番組のデータを使うことにした。加速度による「歩行」認識、「安静」認識との比較を表5、表6に示す。また、加速度の場合も、音の場合も、分類の際に処理する生データは1秒間分であるため、それぞれ処理対象である1秒間の生データのサイズも算出した。

表5. 「歩行」分類性能と対象データサイズの比較

分類に使用するデータ	「歩行」分類性能 (%)	1秒間データサイズ (Byte)
加速度 (10bit 20Hz)	95.4	75
音 (16bit 16kHz)	86.7	32,000

表6. 「安静」分類性能と対象データサイズの比較

分類に使用するデータ	「安静」分類性能 (%)	1秒間データサイズ (Byte)
加速度 (10bit 20Hz)	97.5	75
音 (16bit 16kHz)	88.8	32,000

音だけを使用して86.7%の精度で「歩行」を分類でき、88.8%の精度で「安静(TV鑑賞)」を分類できることが確認できたが、加速度だけによる分類性能よりはそれぞれ劣る。また、分類のために必要な1秒間のデータサイズは、加速度の場合は75バイトであるのに対し、音の場合は32,000バイトと3桁大きいデータの処理が必要となる。このことから、「歩行」、「安静」を認識するためには、性能とデータ量(計算量)双方の観点から、加速度による動作状態分類を行うことが望ましいと言える。

また、「安静」時にユーザ自身から発せられる音は、呼吸音など非常に小さいレベルの音のみで、周囲の環境音の影響を大きく受けることが想定され、今回は同一のTV番組を視聴中のデータで評価したが、(番組内容に依存しない)TV視聴、読書、音楽鑑賞、睡眠など、音のみによる複数の「安静」検出はさらに困難であると考えられる。

以上の観点から、加速度による動作状態分類で「歩行」、「作業」、「安静」を分類し、

「作業」の場合に環境音分析を行って詳細な作業状態分類を行う提案手法の妥当性を確認することができる。

5.2 検出すべき生活状態について

今回の実験では、高齢者や障がい者の家庭内生活状態見守りのために検出すべき作業状態として、ADL、IADLの中から家事、排泄管理、整容の代表的なものについてデータ収集を行い、提案手法によるユーザ状況認識の基本的な適用可能性を確認できた。しかし、今回対象とした作業状態以外にも、家事であれば、炊事、洗濯など、排泄以外の基本的な生活行動としては食事、睡眠、入浴、着替えなど、整容であれば洗顔、化粧など、ADL、IADLには他にも分類対象として検討すべき生活行動が多く存在する。

また、総務省統計局の平成18年(2006年)の調査結果[19]によれば、平成13年(2001年)に比べて特に高齢者は男女ともに家事関連の生活時間が増加している。また、身だしなみ、化粧、入浴、トイレなど身の回りの用事的生活時間は、高齢になるほど長くなり、全年齢で10年前に比べて顕著に増加している。このことから、高齢者の家事や身の回りの用事に関する生活状態を見守ることは今後更に重要性が増すと考えられるため、分類対象とする生活状態の数を増やす方向で検討を進めていきたい。

ただし、今後、様々なユースケースを検討していく上で、そのユースケースにとって分類が必要な作業状態のうち、どうしても音に基づいた作業状態分類が困難な場合があるかもしれない。例えば、食事、睡眠など、人間の基本的な生活状態の取得は、高齢者・障がい者に限らず生活習慣の管理には重要である。しかし、本論文で提案した手法では、食事については、動作状態が「作業」だけでなく「安静」と判断される可能性があり、また、食事に伴う音(咀嚼音、食器の音など)に加えて、複数人で食事をしている場合は、他の人が発する音や会話の音、場合によってはTVやステレオの音なども含まれるため、本手法では分類が難しいと想定される。また、睡眠時はほとんどの時間が「安静」と判断されることになり、時刻と合わせることである程度は推測できるが、直接的には分類が困難である。

これらについては、ターゲットとしたユースケースにおいて許容される拘束性に依存するが、腕時計型の生体センサを装着して、脈波、手の動き(加速度)、皮膚温度、皮膚電気反射(GSR: Galvanic Skin Response)などを計測することによる食事検出[20]、脈波、体動を計測することによる睡眠状態計測[21]などが可能であり、これらの計測手法と本提案手法とを組み合わせた利用形態も考えられる。

6. まとめと今後の課題

本論文では、環境側へ機器を設置せず、携帯電話などの小型情報端末にすでに搭載されている加速度センサとマイクのみを活用した生活状態見守りのための家庭内ユー

ザ状況認識手法を提案した。まず加速度センサでユーザの動作状態を大まかに推定し、推定した状態に応じてマイクを起動して環境音の分析を行うことで、計算量を抑えた家庭内ユーザ状況認識が可能となる。

胸ポケットに格納した試作デバイスで実際の家庭生活のデータを収集し、各生活状態につき任意の10秒間ずつの音特徴量データをSVMに学習させて評価した結果、加速度センサのみで「歩行」、「作業」、「安静」の3状態を概ね95%以上の精度で分類でき、「作業」時には環境音の分析を行うことで、「歯磨き」、「電気シェーバーによる髭剃り」、「ドライヤーの使用」、「トイレ水洗/手洗い」、「掃除機がけ」、「皿洗い」、「アイロンがけ」の各作業状態を平均75.8%の精度で、本人の学習データを用いた場合は平均85.9%の精度で分類できた。このことにより、提案手法が我々の想定する家庭内の生活状態見守りに活用可能な基本性能を持つことを確認した。

分類器にSVMを用いることで、各生活状態につき任意の10秒間分という少ない学習データでも良好な分類性能と応答性が得られたが、目的とする生活状態によっては学習させるデータの区間を適切に選択する必要がある。その区間の効果的な選び方については今後検討が必要である。また、運用前の事前学習が必要な前提で議論してきたが、大規模なデータ収集を実施することによって各作業状態の標準的な学習データを作成することにより、事前学習を不要にすることができる可能性がある。また、運用開始後の自動的な再学習の枠組みも検討したい。

今後は、携帯電話上に提案手法を搭載し、より実用に近い形で、これらの課題に関する検討と実用化に向けた評価を行っていく。

謝辞

本研究の一部は総務省の研究委託により実施したものである。

参考文献

- 1) 美濃導彦：家庭におけるユビキタス環境の構築－ゆかりプロジェクト－，電子情報通信学会2004総合大会，No.A-16-8，p.317 (2004).
- 2) 森武俊，野口博史，佐藤知正：センシングルーム－部屋型日常行動計測蓄積環境 第2世代ロボティックルーム－，日本ロボット学会誌 Vol.23 No.6，pp.665-669 (2005).
- 3) 松岡克典：宅内ネットワークを用いた生活見守り技術，第49回自動制御連合講演会論文集SU1-2-2，pp.1-2 (2006).
- 4) N. Kern, B. Schiele, and A. Schmidt: Multi-sensor Activity Context Detection for Wearable Computing, Proceedings of the First European Symposium on Ambient Intelligence (EUSAI 2003), LNCS 2875, pp.220-232 (2003).
- 5) L. Bao and S. S. Intille: Activity Recognition from User-Annotated Acceleration Data, Proceedings of the Second International Conference on Pervasive Computing (PERVASIVE 2004), LNCS 3001, pp.1-17 (2004).
- 6) T. Iso and K. Yamazaki: Gait Analyzer based on a Cell Phone with a Single Three-axis Accelerometer, Proceedings of the 8th conference on Human-computer interaction with mobile devices and services (MobileHCI2006), pp.141-144 (2006).
- 7) 倉沢央，川原圭博，森川博之，青山友紀：センサ装着場所を考慮した3軸加速度センサを用いた姿勢推定手法，情報処理学会 研究報告，2006-UBI-11 (3)，pp.15-22 (2006).
- 8) 池谷直紀，菊池匡晃，長健太，服部正典：3軸加速度センサを用いた移動状況推定方式，情報処理学会 研究報告，2008-UBI-19 (14)，pp.75-80 (2008).
- 9) V. Peltonen, J. Tuomi, A. Klapuri, J. Huopaniemi, and T. Sorsa: Computational auditory scene recognition. Proceedings of Acoustics, Speech, and Signal Processing (ICASSP 2002), pp.1941-1944 (2002).
- 10) 大内一成，土井美和子：センサ駆動ハンドヘルド型音声認識入力方法の提案－センサを用いた発話動作検出手法，情報処理学会論文誌 Vol.51 No.2，pp.324-333 (2010).
- 11) 小林亜令，岩本健嗣，西山智：釈迦：携帯電話を用いたユーザ移動状態推定方式，情報処理学会論文誌 Vol.50 No.1，pp.193-208 (2009).
- 12) D. Siewiorek, A. Smailagic, J. Furukawa, A. Krause, N. Moraveji, K. Reiger, J. Shaffer, and F. Wong: SenSay: A Context-Aware Mobile Phone, Poster of 7th IEEE International Symposium on Wearable Computers (ISWC 2003) (2003).
- 13) T. Iso, N. Kawasaki, and S. Kurakake: Personal context extractor with multiple sensor on a cell phone, 7th IFIP International Conference on Mobile and Wireless Communications Networks (MWCN2005), D. 2 C200525 (2005).
- 14) 前川卓也，柳沢豊，岸野泰恵，石黒勝彦，亀井剛次，櫻井保志，岡留剛：ウェアラブルセンサによるモノを用いた行動の認識について，情報処理学会 研究報告，2010-UBI-25 (57)，pp.1-8 (2010).
- 15) T. Maekawa, Y. Yanagisawa, Y. Kishino, K. Ishiguro, K. Kamei, Y. Sakurai and T. Okadome: Object-Based Activity Recognition with Heterogeneous Sensors on Wrist, Proceedings of the Eighth International Conference on Pervasive Computing (PERVASIVE 2010), LNCS 6030/2010, pp.246-264 (2010).
- 16) Katz S.: Assessing self-maintenance: activities of daily living, mobility, and instrumental activities of daily living, Journal of the American Geriatrics Society 31(12), pp.721-727 (1983).
- 17) Fillenbaum GG: Screening the elderly. A brief instrumental activities of daily living measure, Journal of the American Geriatrics Society 33(10), pp.698-706 (1985).
- 18) Cole RJ, et al.: Automatic sleep/wake identification from wrist actigraphy, Sleep 1992, 15 461-469 (1992).
- 19) 総務省統計局 平成18年社会生活基本調査，<http://www.stat.go.jp/data/shakai/2006/index.htm>
- 20) K. Ouchi, T. Suzuki, and M. Doi: LifeMinder: A wearable healthcare support system with timely instruction based on the user's context, IEICE Transactions on Information and Systems, Vol.E87-D, No.6, pp.1361-1369 (2004).
- 21) T. Suzuki, K. Ouchi, K. Kameyama, and M. Takahashi: Development of a Sleep Monitoring System with Wearable Vital Sensor for Home Use, Proceedings of BIOSTEC BIODEVICES 2009, pp.326-331 (2009).