

## タッチログ：爪装着型触音計測に基づく 触対象推定システム

牧野 泰 才<sup>†1</sup> 村尾 将 和<sup>†2</sup> 前野 隆 司<sup>†2</sup>

本研究で我々は、指下で生じる振動を取得する方法を提案する。爪に圧電素子を貼付し、接触や操作により生じた振動を検出する。このようにして得られた振動は、指の姿勢や外来ノイズに影響されることなく、接触対象や動作によって異なる波形を示すことを確認した。対象により接触音が異なることから、何を触ったのかを検出可能である。ヒトは行動を起こす際、多くの場合何かに触れる。したがって、この触対象が何かという情報は、新しいライフログのデータとして利用可能である。実験により提案手法の対象識別精度を確認した結果、12種類の日常動作を94.4%の精度で識別できることを確認した。また、ライフログの利用価値についても考察し、本手法が有効である用途の検討も行った。

### Touch Log: Estimation of Touch Objects by Measuring Vibration on Fingernail

YASUTOSHI MAKINO,<sup>†1</sup> MASAKAZU MURAO<sup>†2</sup>  
and TAKASHI MAENO<sup>†2</sup>

In this paper, we propose a new method that detects tactual vibration at a fingernail. A piezoelectric device is attached on the fingernail to detect sounds of touch and manipulation. We call the recorded sound "Touch log." We experimentally show that the touch log is robustly detected against external audio noise and finger motions. Since the touch logs depend on objects, we can record what we touched in our daily lives by this method. This information can be used as a new source for life log. We show our experimental result that our prototype system recognizes 12 different actions with 94.4% accuracy. We also discuss what types of life log needs the touch log related information.

### 1. はじめに

本研究で我々は、ライフログとして利用可能な「タッチログ」を提案する。人が何か行動を起こす際、洋服を着る、テレビをつける、扉を開けるといったように、多くの場合、手指が何らかの物体に接触する。したがって、接触対象が何であるかという情報は、その人の行動履歴と関連が高い。この情報は、ライフログとしての利用が可能である。

ライフログとは、人の生活を計測および記録したものである。これまでも人は、日記、写真、議事録といったように様々な方法で生活の一部を記録してきた。これに対して、近年のデジタル化やセンシング技術の発達により、ライフログは従来のような「記録するもの」から、機器によって自動的に「記録されるもの」へと変わりつつある。このようなライフログのあり方の変容にともない、これらの新しい個人情報を利用した新たなサービスの提案および実用化がさかに行われている<sup>1),2)</sup>。

近年、ネットワーク技術の進展により、従来個人で用いてきたライフログが、他者と共有するものへと変化をとげている。たとえば Google で検索する場合、自身が入力した検索ワードに対して、一般的に頻繁に同時に検索される単語も提示される。すなわち、他者の検索履歴というライフログを、ネットワークを介して共有し活用していることが分かる。

このような、検索履歴や使用したファイルの履歴といった、主に PC 上で記録可能な電子化された情報は、ネットワークの普及により、近年その活用機会が増えている。一方、扉を開ける、食事をとるといった電子機器を利用しない行動情報（以下、「非電子化情報」と呼ぶ）は、電子化情報と比べて取得するのが難しい。このような情報も取得することができれば、ライフログの利用の幅は大きく広がると考えられる。

非電子化情報を取得する手段として、大きく2つの方法が考えられる。1つはユーザにセンサを搭載し、ユーザの行動履歴を記録する方法である。簡便な手段としては、カメラやマイクロフォンを利用する方法が提案されている<sup>3),4)</sup>。ヒトの動作を映像や音声で記録し、画像処理や音声認識技術を利用して動作を推定する。また、これらを複合的に組み合わせ、手首に配置したカメラやマイクロフォン、加速度センサなどを利用し、人の動作を推定する手法も提案されている<sup>5)</sup>。

<sup>†1</sup> 慶應義塾大学環境共生・安全システムデザイン教育研究センター

Center for Education and Research of Symbiotic, Safe and Secure System Design, Keio University

<sup>†2</sup> 慶應義塾大学大学院システムデザイン・マネジメント研究科

Graduate School of System Design and Management, Keio University

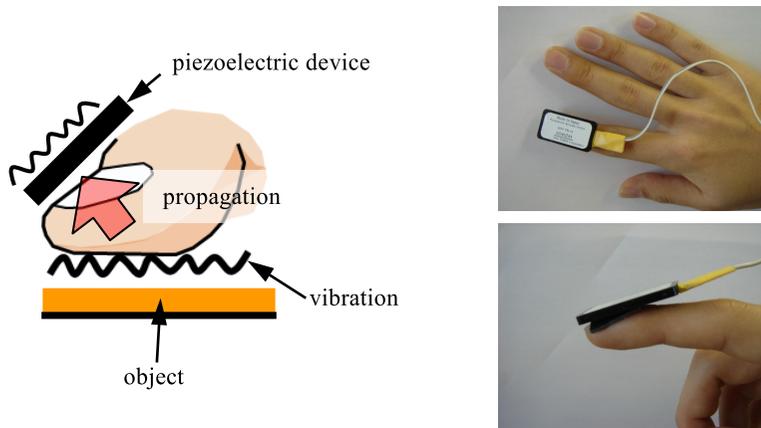


図1 触音検出システムのプロトタイプ  
Fig. 1 Prototype tactile sound detection system.

もう1つの方法は、ヒトが動く空間にセンサを配置し、動作履歴を取得する方法である。こちらでもカメラやマイクロフォンを利用できるが、それ以外に、センサネットワークを利用した方法も提案されている<sup>6),7)</sup>。たとえば、室内の床に圧力センサなどを配置することで、人の動きを記録することが可能となる。

本稿で提案する「タッチログ」は、ユーザにセンサを搭載するタイプの、新しいログ取得法である。図1に提案手法の概略を示す。爪に圧電素子を貼付し、指腹で生じた振動を検出する。本稿では以降、この爪上で観測される振動を「触音」と呼ぶ。触音は接触対象や操作に依存してその振動パターンが異なるため、音声認識の手法を用いることで接触対象が何であるかを知ることが可能となる。

提案するタッチログ計測手法は、主に以下の2つの特長を持つ。

- 1) 指自体の運動は計測されないため、指の姿勢などの影響を受けにくい。
- 2) 光学的、音響的な外来ノイズの影響を受けにくい。

これらの特長から、指が触対象に触れ、何らかの操作をしている「意味のある信号区間」を抽出することが容易になる。また、カメラやマイクロフォンにより取得されるデータにノイズが大きく重畳されるような環境下（たとえば暗闇や騒音環境）であっても利用可能である。したがって、従来手法に新たに情報を付与し、補完的に利用することができる。たとえば、取得された映像の中から、画像処理を必要とする区間を抽出するためのキーとしての利

用や、照明環境が動的に変化するようなシーンでの、安定した行動認識などができると期待される。

本稿では、提案する爪上での振動計測手法について2章で説明し、従来の測定手法と比較した場合の利点を提示する。3章ではタッチログをライフログとして利用する場合の利点などを、一般的なライフログの利用方法を概観することで比較する。4章では実際に得られた波形を用いて触対象の識別実験を行った結果を示す。5章、および6章で、考察、まとめを行う。

## 2. 爪上での振動計測による触音取得

### 2.1 先行研究

これまでに提案されてきた触覚センサは、大きく3つに分類できる。1つはロボットに搭載し、指先や全身などで、ヒトと同じような触覚を検出するためのセンサである<sup>8),9)</sup>。2つめは、タッチパネルのように、ヒトが触覚を介して情報を入力する際に利用されるインタフェースとしてのセンサである<sup>10)</sup>。

これらに対して、近年3つめの触覚センサとして、ヒトの指の変形を間接的に計測することで、指腹に生じている触現象を推定するというタイプの、新しいセンサが多く提案されている。これは、ヒト自身をセンサの一部として利用するという点で、上記2つのタイプとは異なる。指腹にはデバイスを装着しないため、ヒトの動作時の触覚を変えないことなく、触現象の推定が行えるという利点がある。

Mascaroらは、爪部の色変化に着目し、爪に光学センサを配置することで、指腹部での接触状態を推定した<sup>11)</sup>。安藤らも同様の手法に独立成分分析を用い、接触力の他に屈曲も検出している<sup>12)</sup>。牧野らは、指を透過する光に着目し、タッチパネル上の光源パターンの透過光を爪上の光学センサで計測し、押下力や指の姿勢を推定するシステムを提案している<sup>13)</sup>。湯村らは、爪上に加速度センサを配置し、物体をなぞった際に生じる指腹の振動を検出し、触覚の伝送に利用している<sup>14)</sup>。福本らは、手首部分に圧電衝撃センサを設置し、指先のタップ音を取得するという方法を提案している<sup>15)</sup>。Iwamotoらは、チューブを利用した指輪状のインタフェースによりチューブ内の空気圧変化を計測し、新しい入力デバイスとしての利用を提案している<sup>16)</sup>。そのほかにも、指先にひずみゲージを付けて接触力を検出する方法や<sup>17),18)</sup>、マイクロフォンをつけて接触時の音を検出する方法<sup>19)</sup>など、多様な手法が提案されている。

## 2.2 提案手法

本稿では、ヒトの指を利用した触覚センサの新しい実現方法として、図1に示すように、爪に圧電素子を貼付し、指腹に生じた振動を検出するシステムを提案する。両面テープにより、爪上に圧電素子（NEC トーキン KDS-UM-01）を貼付し、指腹に生じた振動を検出する。圧電素子から出力される信号は、直流成分をカットする1次のローパスフィルタ（カットオフ周波数：2 Hz）を介した後に、マイクロフォン端子からPCへ直接取り込まれるようになっている。

多くの場合、人は対象の操作に拇指と示指を利用する。したがって、操作時の指腹の振動を取得するのであれば、どちらかの指、あるいは両方にセンサを装着するのが望ましい。今回、これ以降の実験では、示指にセンサを装着している。マウスのクリックなど、通常示指以外では行わない作業があったため、示指に装着することを選んだ。

図2に、4種類の動作時に実際に取得された触音波形の例を示す。上段左がクリック操作

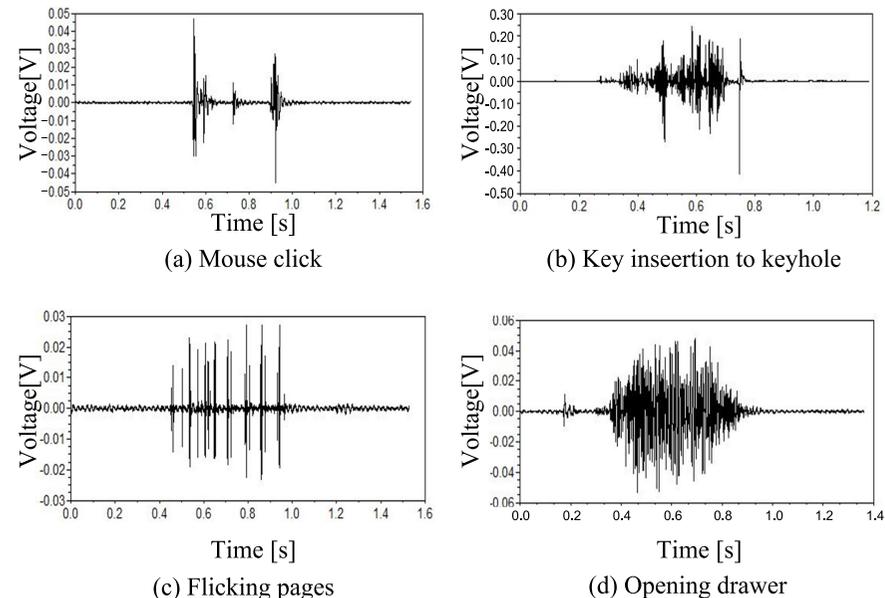


図2 観測された触音  
Fig. 2 Observed tactile sounds.

時の触音、上段右が鍵穴に鍵を差し入れた際の触音、下段左が本のページをめくった際の触音、下段右が引き出しを引いた際の触音である。本のページめくり動作は通常拇指を利用するが、今回は示指により行った。これらの触音は、それぞれ特徴的な波形を示すことが確認された。以下でそれぞれの特徴について述べる。

クリック時の触音は押下時および除荷時それぞれに振動が生じる。つまり小さな振動が2つ生じることがこれの特徴である。鍵穴に鍵を差し込んだ際の触音は徐々に振動が大きくなるという特徴を有することが分かった。これは鍵の差し込み量が増加するほど、指と振動が生じる点が近くなるからである。本のページをめくった際の触音はこれがめくられるたびに振動が検出される。この触音に対してFFTを行うと、2 kHzという比較的高い周波数成分を含むという特徴がみられた。引き出しを引いた場合は、比較的、定常な振動が得られた。始動する際に小さな振動が生じることもこれの特徴である。

### 2.3 既存手法との比較実験

本手法では圧電素子を利用している。このことにより、加速度センサを利用した場合とは異なり、指の姿勢変化の影響を受けないという特徴を持つ。指の姿勢や運動によらず、指腹に生じた振動のみを検出することができる。また、今回利用した圧電素子は、音響インピーダンスが空気と大きく異なるため、空気中を伝搬する環境音を取得することもなく、音響的な外来ノイズの影響を受けにくい。指腹から指内部を伝搬し爪へと至る振動のみを検出する。

以上の理由より、先述のように、

- 1) 指自体の運動は計測されないため、指の姿勢などの影響を受けにくい、
  - 2) 光学的、音響的な外来ノイズの影響を受けにくい、
- という2つの特長を有するセンサとなっている。以下本節では、この2点を実際に既存手法と比較し評価する。

#### 2.3.1 加速度センサとの比較

本手法の利点を確認するための比較実験を行った。まず、これまでに提案されている加速度センサによる手法<sup>13)</sup>と比較した。加速度センサ、圧電素子ともに右手示指に両面テープにより貼付して実験を行った。なお、加速度センサとしては、AMI601を利用し、圧電素子と同程度のサイズ、質量となるように実装を行った。実験の手順としては、最初に手のひらを下にした状態で机の上に静止し、10 cm 横にあるマウスに手を移動させ、1 クリックした後、元的位置へ戻すという動作を行った。

図3(a)に圧電素子で検出された触音を、図3(b)に加速度センサの出力をそれぞれ示す。加速度情報は、センサから得られた3軸のデータのうち、爪表面に対して垂直な方向の加速

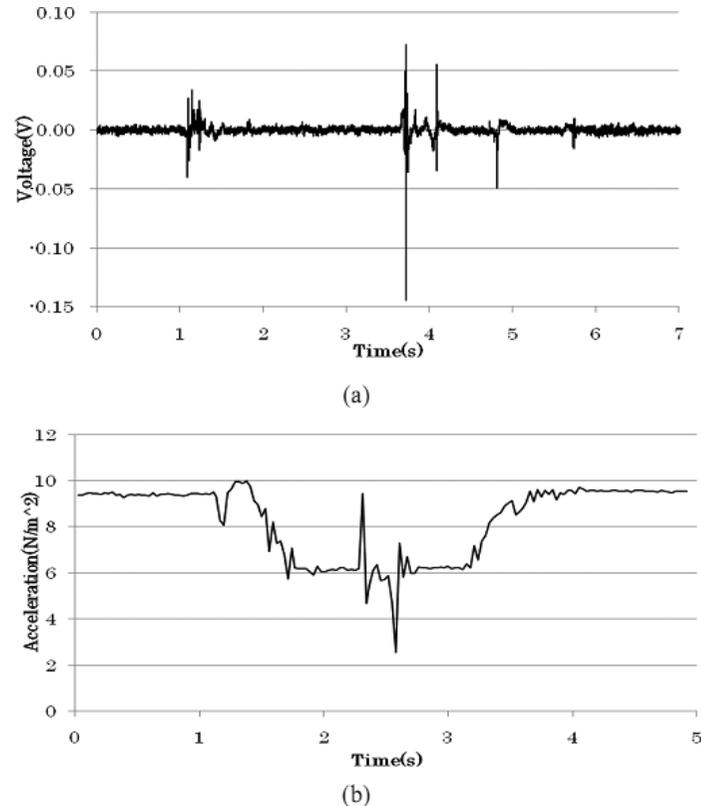


図3 マウスクリックにともなう観測信号：(a) 触音，(b) 加速度

Fig. 3 Observed mouse click related signals: (a) tactile sound and (b) acceleration.

度のみを示す．一般的に，クリック時に生じる振動は，指腹に対して垂直な方向の振動であるため，主にこの成分に振動情報が取得される．

図3(a)の触音の結果を実際の動作と比較すると，1秒付近の振動は指先がマウスに触れた瞬間，4秒前後の2つのピークが押下時と除荷時のクリック動作に対応する振動，5秒あたりの振動がマウスから指が離れた瞬間，6秒あたりの振動が机に戻した瞬間である．提案手法の場合，クリック時の振動，その他関連する振動が，十分なSN比で検出されていることが分かる．

一方，図3(b)の加速度センサの場合，計測開始時は手のひらを下にした状態で静止しているため，ほぼ重力加速度程度の値が得られている．1.5秒付近より手をマウス上へと移動させているが，そのときの加速度変化，およびマウスのボタンに指を沿わせることによる指の姿勢変化の影響が出力されている．マウスをクリックする前後では，ボタンに沿わせる形に指が傾いたことにより，加速度センサの出力が低下している．クリック動作によって生じた振動は，2～3秒付近で押下時と除荷時ともに検出された．

加速度センサの場合，他の2軸のセンサ出力などをもとに，指の動作とクリックによる振動とを切り分けることは可能ではある．しかし提案手法の方が，操作に関する振動のみを検出できていることが確認された．このことは，有意な振動区間を切り出すことが容易であることを意味する．

### 2.3.2 マイクロフォンとの比較

次に，音響的な外来ノイズの影響を受けにくいという利点を示すために，指先にマイクロフォンを配置して同様に接触音を取得する場合<sup>18)</sup>と比較した．マイクロフォンはヘッドセットなどで用いられる市販のものを利用し，指先にマイクロフォンの集音面が向くように，爪部にテープで固定した．

加速度センサのときと同様，机上に置いた手をマウス上に移動させ，クリックして戻すという動作を行った．外来ノイズとしてはホワイトノイズを用い，単一のスピーカにより再生した．このとき，マウスのボタン上におけるノイズの音圧は，約73dBであった．

結果を図4に示す．(b)のマイクロフォンの場合4秒付近でクリックが行われている．しかし，図からも明らかなように，ノイズレベルが大きく，クリックが生じている区間を信号処理により安定に切り出すことは難しい．このときのSN比は2～3程度であることが確認された．

一方，提案手法ではノイズ環境下においても，その影響を受けることなく，十分なSN比でクリックにともなう振動が検出されることが確認された．ノイズレベルの平均振幅が0.002V程度であったのに対し，振動振幅は0.05V程度であり，SN比としては25程度である．ノイズ環境下においても，指先がマウスへ触れた瞬間(1～2秒の間)，マウスから離れた瞬間(4秒すぎ)，机においた瞬間(5秒あたり)がそれぞれ取得できることも確認した．

以上2つの比較実験より，本手法は人の指を利用した第3のタイプの触覚センサとして見た場合に，加速度センサやマイクロフォンと比べてノイズなどの影響を受けにくく，対象の振動のみを検出できる手法であることが確認された．これは，本手法が多様な環境下で利用可能な手法であることを示す結果である．

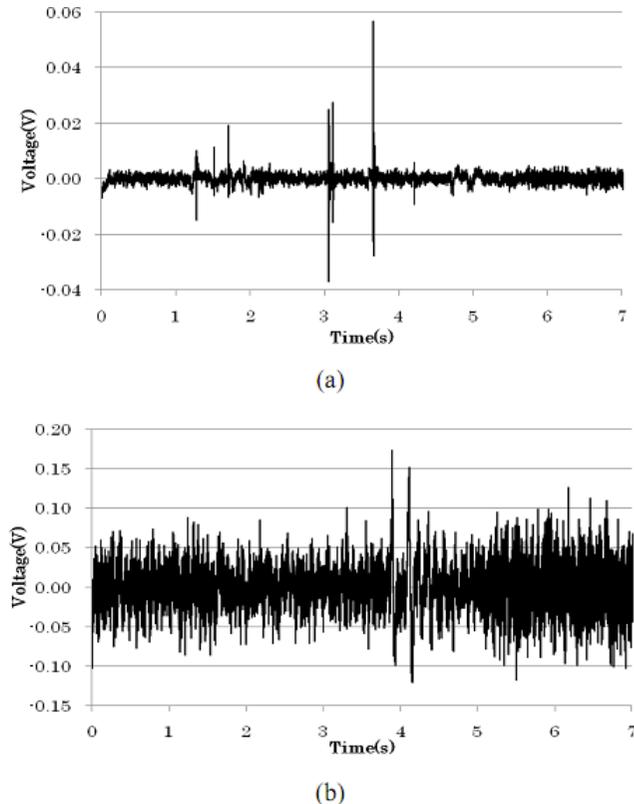


図4 音響的ホワイトノイズ環境下でのマウスクリックにともなう観測信号：(a) 触音，(b) 音響マイク  
 Fig.4 Observed mouse click related signals with acoustical white noise (a) tactile sound and (b) acoustical microphone.

### 3. 触音のライフログとしての利用

提案する圧電素子を利用した触音検出手法は、接触の瞬間や、把持対象を操作している瞬間を切り出すことが容易である。つまり、ヒトが何かしらの意志を持って対象に働きかけている瞬間を抽出できる。このとき、既存の画像や音声を用いたライフログと併用すると、それらデータの中の意味のある区間を容易に抽出することが可能となる。

また、接触対象が何であるか、あるいはどのような操作をしているかという情報も触音から識別できれば、ユーザの行動把握につながる。したがって、そのままタッチログ単体でライフログとして利用可能になると考えられる。このように、提案手法は触覚センサとしての利用だけでなく、ライフログとしての利用も可能である。

#### 3.1 ライフログの利用価値

タッチログをライフログとして利用する際に、新たに実現されるであろう利用法を議論するために、まず一般的なライフログの利用価値について述べる。

ライフログは、ヒトの記憶を様々な媒体に記録する技術であり、記憶の代替と考えることができる。我々はこの部分に着目し、先行研究<sup>20),21)</sup>を参考に、ヒトの記憶の機能を図5のように2つの軸で分類することとした。縦軸は記憶を利用する際に対象とする時間（過去-未来）を、横軸は記憶されたデータの利用範囲（個人-社会）を示す。このようにして得られた4つの象限の意味するところを以下に記述する。

- 個人的過去機能：個人内で閉じた状態で過去を振り返ることを目的とした機能。日記やアルバムなどのように過去を振り返ることによって、気分の調節や自己発見を行う。
- 個人的未来機能：個人内での目標設定や行動の動機づけを行う機能。メモやリマインダなどの備忘録、探し物を見つけることなどがあげられる。また、自己分析などの未来のための目標設定もこの機能に分類される。
- 社会的過去機能：他者とともに過去を振り返ることを目的とした機能。友人とアルバムを見ること、昔話で盛り上がることや、ブログを用いた体験の共有などがあげられる。これにより、個人間のつながりの強化やコミュニケーションの支援などを行う。また、国民調査などの集団の性質や動向を把握することもこの機能に分類される。
- 社会的未来機能：他者とともに目標設定や行動の動機づけを行う機能。経験に基づいて他者にアドバイスすること、マイルストーンなどの作業管理などがあげられる。また、視聴率調査のように、集団の性質や動向に基づいて行うマーケティングなどもこの機能に分類される。

これらの機能をもとに、現状提案されているライフログの機能をリストアップし、同じく図5に描画した。

1. 日常行動の管理：スポーツや読書といった趣味や日常行動に関する情報を記録する。これにより、目標設定や達成度などを確認することが可能になる。ライフログでは、これらの記録の検索やグラフ化を可能にし、場合によっては利用者の記録を自動化する。
2. 記憶の想起支援：探し物、電気の消し忘れ、鍵の閉め忘れといった日常生活において記

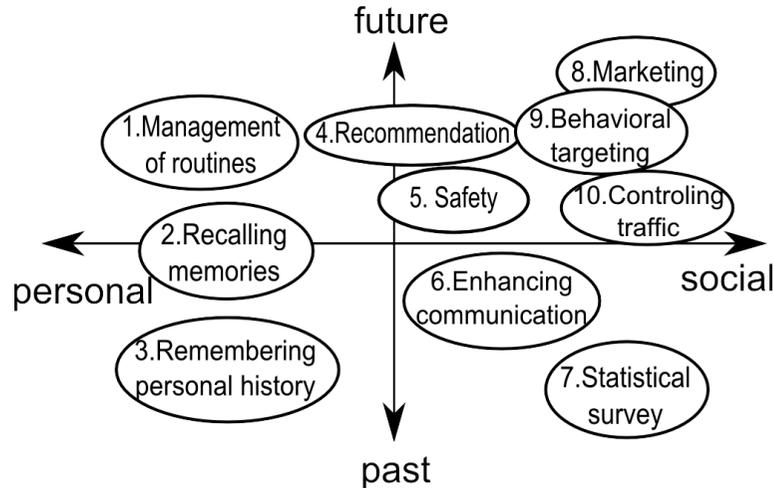


図5 ライフログの利用法の分類  
Fig. 5 Usages of life log.

憶の想起が必要とされるタスクを支援する。ライフログとしていつ何を使ったという情報を記録することにより、これらの問題は解決すると考えられている。

3. 個人の行動履歴を思い出すための支援：日記、写真、動画、検索履歴、GPS 情報などを記録し、それらを整理する。近年ではこれらを自動に時系列化し、検索を容易にするアプリケーションが開発されている。過去を振り返ることによる気分調節や他者とのコミュニケーションの支援といったメリットがある。
4. レコメンド：Web ページやレストランの場所といった情報を利用者にお勧め（レコメンド）する。レコメンドを行うためには利用者の趣味や嗜好を把握することが重要である。ライフログを用いることにより、利用者の嗜好により近いものを提供することができるようになる。主に Web アプリケーションなどで実用化されている。
5. 安全・安心：健康状態の変化を記録することにより、緊急事態の場合は病院や警察に自動的に連絡するなどの対応を行う。高齢者の一人暮らしにおける不安を解消するために、電化製品の使用履歴を記録することで、緊急時を検出する技術はすでに実用化されている。
6. コミュニケーションの向上：ライフログを用いることで、共通の趣味や経験を有するコ

ミュニティを発見することが想定されている。SNS などへの適用が期待されている。利用者にとって適したコミュニティを探すことが容易になると考えられる。

7. 生活行動調査：人々の趣味、嗜好や行動パターンを調査する際に利用する。特にアンケートでは調べることができない日常の行動を観察することが重要であり、ライフログのように自動的に行動を記録する手法は有効である。
8. マーケティング：ライフログを用いてマーケティングを行い、製品やサービスの開発に役立てる。生活行動調査の場合は、過去の動向を調べることに主眼が置かれるが、マーケティングはその情報をもとに、経営戦略を立てるなど、未来機能に特に焦点が当てられる。
9. 行動ターゲティング：レコメンドとは異なり、情報を伝える相手を絞り込む。主にリスティング広告といわれる検索連動型広告のために実用化されている。ある検索キーワードが入力された際に、それと関連する広告を表示する。効率的に情報を伝える相手を絞り込むためには、ライフログを用いることが有効である。
10. トラフィック制御：自動車や電車、通信網など、社会インフラのトラフィック制御を行う。たとえば自動車であれば、GPS 信号データに基づき、利用者の状況に応じて渋滞の緩和などが行える。

ここで注意が必要なのは、これら 4 象限の中でどのタイプの機能を持たせるかにより、ライフログに必要とされる要求仕様が異なる点である。たとえば社会的機能を実現する場合、特に不特定の他者とライフログを共有するような場合には、共有されるデータには個人を特定できるような情報が極力入らないことが望まれる。一方、個人的機能として、個人内でのみ自身のログを使い、行動を振り返るのであれば、個人情報が入っていることは、さして問題にはならない。

また、過去を思い出すような場合には、必ずしも精密な動作履歴が記録されている必要がない場合もある。たとえば、日記作成支援のような形でライフログを利用する場合、自身の過去の経験と、記録されたログとを照合することになる。そのため、整合性のとれない行動認識結果などは、自身で修正することができる。一方、過去の行動履歴から、何かをレコメンドするような場合には、過去の行動が正確に認識されていないと、適切な推薦は行えない。このように、目的とする利用方法により、要求される情報の質と精密さが異なることが分かる。

### 3.2 タッチログの利用価値・アプリケーション

本稿で提案するタッチログは、カメラやマイクロフォンと比較すると、取得されるデータの情報量は少ない。すなわち、情報の精密さが要求されるような用途には向かないことが考えられる。

一方、取得されるデータ内に他者、あるいは利用者自身の情報がほとんど入らないという利点を持つ。カメラやマイクロフォンでは、そこに介在している人物が誰であるかが、映像や音声で記録されてしまう。一方、触音として得られるのは、接触対象に依存した振動音であり、そこに操作者や関連他者の情報は入りにくい。

我々は事前にアンケート調査を行い、ライフログとして取得可能な情報である、When (いつ)、Who (誰が、誰と)、Where (どこで)、What (何を)、Why (なぜ)、How (どのように) の 5W1H の中で、最も取得されたくないと感じられる情報は Who であるという結果を得た。特に、自身の情報が他者により勝手に記録され、利用されることに対する抵抗感が大きいという傾向があった。したがって、カメラやマイクロフォンにより取得されたデータは、社会的機能として利用すること、あるいはそれ以前に勝手に記録されることに対して抵抗をいなくユーザが多いと予想される。一方、タッチログのように動作者や関連する他者の情報がほとんど含まれないデータは、社会的機能として共有することや、記録されることに対する抵抗感が低くなると期待される。

タッチログはまた、意味のある情報区間を切り出すことに優れている。画像や音声の場合、区間の切り出しのためには画像処理や音声認識を必要とする。一方、本手法では、対象に触れたり操作したりしていない時間は、ほとんど無音であるため、振動振幅に着目するだけで、有意な信号区間をピックアップできる。

このような特徴をもとに、タッチログでは以下のようなアプリケーションでの利用が想定される。

- 1) 玄関の扉を閉めた後に、車のエンジンをかけるログが出力されたような場合には、玄関の鍵を閉め忘れた可能性が高い。このように、通常の動作履歴と異なる場合を検出し、「し忘れ」を防止するのに有用である (1. 日常行動の管理, 2. 記憶の想起支援)。
- 2) Twitter や SNS などに、自身の行動履歴を自動でアップロードするよう支援できる。カメラで取得された画像や、音声情報を自動でアップロードすると、プライバシーの問題が大きい。触音をもとに識別された「触れたものリスト」に基づく行動履歴のアップロードであれば、ユーザの抵抗は少ない。また誤認識があったとしても、さして問題にはならない (6. コミュニケーションの向上)。
- 3) 本手法はティッシュの使用量といった非電子化情報による資源浪費行動の把握を可能にする。近年の環境問題の観点から消費した CO<sub>2</sub> や電気使用量の可視化が注目されている。従来困難であった、ティッシュの使用量といった非電子化情報による資源浪費行動の頻度、または冷蔵庫の開閉といった電気使用量と行動との因果関係を利用者に提示する

ことが想定される (7. 生活行動調査, 8. マーケティング)。

- 4) マウスのクリック回数や、キーボードの打鍵回数を記録することで、仕事の能率をログとして取得できる (1. 日常行動の管理)。
- 5) 子供や高齢者などの危険検出といった用途が想定される。包丁やハサミといった道具の使用法の誤りによる事故を未然に防ぐことが期待される (1. 日常行動の管理, 5. 安全・安心)。

このように、一般的なライフログの用途として提案されているものの多くに適用可能であることが分かる。もちろん、これらのアプリケーションをそれぞれ個別に実現するのであれば、タッチログを利用しなくても可能である。たとえば、ティッシュの消費量を検出するのであれば、ティッシュ箱に専用のセンサを配置するだけでよい。我々の提案するタッチログの利点は、これらを単一のセンサで横断的に取得できる点であり、また既存の手法と併用することでその精度を向上させられる点である。

## 4. 実 験

### 4.1 プロトタイプ

提案手法により取得される触音は、図 2 で見たように接触対象により異なる。そこで、既存の音声認識手法により、触音をどの程度識別可能であるかを検証した。このとき、取得される触音は各個人の皮膚の弾性的なフィルタの影響によっても、その波形が変わりうる。そこで、個人の皮膚の弾性的性質の差が、対象の認識にどの程度影響を与えるかも検証する。

プロトタイプシステムを試作し、複数人の触音をもとに平均的な音響モデルを構築した。もし個人差の影響が、対象の違いの影響よりも大きい場合には、このモデルでは十分な精度で認識が行えない可能性がある。その場合には、ユーザごとにモデルを作成しなければならない。一方、このモデルで高い認識率が得られた場合には、接触対象の違いによる触音の差が大きく、個人の弾性的性質の差は影響が少ないということを示すことができる。

本研究では日常生活で生じる触音に着目し、12 種類の接触対象を識別するシステムを構築した。図 6 に本研究で対象とする 12 種類の対象物 (動作) を示す。

本研究では従来の音声認識の手法を触音に適用した。構築したシステムは、

- 前処理部
- 特徴抽出部
- 音響モデル
- 識別演算部

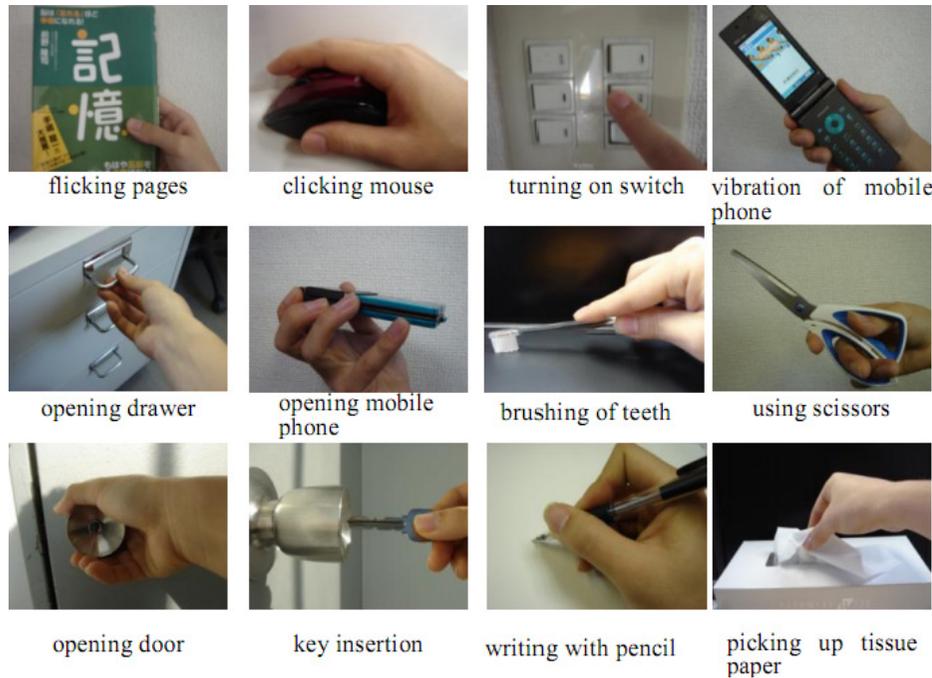


図 6 認識実験に用いた 12 種類の日常動作  
Fig. 6 Twelve daily actions used in recognition experiments.

の 4 つのサブシステムに分解される。以下にそれぞれのサブシステムの機能について述べる。

前処理部は入力である触音波のノイズ除去やオンライン認識において入力区間の決定を行う。

特徴抽出部はパターンの識別で利用する情報を特徴ベクトルとして抽出する。特徴量としては、音声認識の分野で利用される MFCC (メル周波数ケプストラム) を利用し、得られたピーク値を低次から順に 12 次元、これの 1 次差分を 12 次元、絶対値エネルギーの 1 次差分を 1 次元抽出し、計 25 次元のデータとして用いた。今回認識したい触音波形は、人の音声のようなスペクトル上の特徴を有するわけではない。したがって、音声認識で利用される特徴量をそのまま利用するのは最適ではない可能性がある。今回は実装のしやすさからこの手法を利用した。なお、特徴量抽出のパラメータに関しては、フレーム周期は 10 ms、ハ

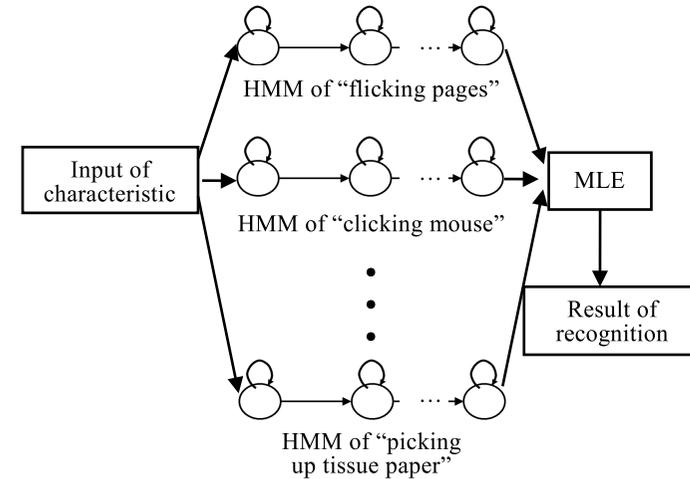


図 7 隠れマルコフモデル (HMM) による最尤推定の模式図  
Fig. 7 Maximum likelihood estimation of HMM.

ミング窓の幅は 25 ms, フィルタバンクのチャンネル数は 24, 高域協調係数は 0.97 とした。

音響モデルは、音声認識の分野で利用されている、HMM (Hidden Markov Model) を用いた学習手法を適用し、これを構築した。時系列データである触音を各状態とそれらの間の遷移によって表す。各状態は、特徴ベクトルの平均値と分散値を有し、確率密度関数を定義する。これにより、ある入力に対して各行動の尤度を求める。

識別演算部は音響モデルと入力の特徴ベクトルを比較し、出力を決定する。決定方法は、用意された音響モデルごとに尤度を算出し、最も尤度の高い音響モデルを出力とする。構築した 12 種類の触対象ごとの HMM を図 7 のように並列に並べることにより、入力の特徴量に対してビダリアルゴリズムを用いて MLE (Maximum likelihood Estimation) を行い、触れたものを識別できる。

これらのサブシステムのうち、前処理部、特徴量抽出部、識別演算部は音声認識ソフト Julius<sup>\*1</sup>を、音響モデルは HTK<sup>\*2</sup>を用いて構築した。Julius に本研究で構築した触音の音響モデルを適用し、タッチログのオンラインによる記録を行った。

\*1 大語彙連続音声認識エンジン Julius: <http://julius.sourceforge.jp/>

\*2 HTK Speech Recognition Toolkit: <http://htk.eng.cam.ac.uk/>

音響モデルの構築のために、7人の被験者（20～30代の男性6人、20代女性1人）に、12種類の触動作を各50回ずつ行わせ、その触音を記録した。したがって、各動作につき、7人分の350データを学習データとして収集した。この学習データを用いて Baum-Welch アルゴリズムで HMM の学習を行った。なお、マウスや携帯電話などは、同一のものを利用した。

触音収集時のサンプリング周波数は予備実験を行い決定した。予備実験で12種類の触音に含まれる周波数帯域を検証した結果、ほとんどの触音は、そのピークが1kHz以下であり、帯域としては2kHz程度までであった。また、これらのスペクトルには、触対象の特徴が現れていた。たとえば携帯のバイブレーションなどは、内部の偏心モータの振動周波数がほぼそのまま得られるため、200 Hz程度の低周波にピークを持つことを確認した。一方、歯磨き動作やページめくり動作などいくつかの動作では高周波振動が生じやすく、1kHz以上にピークを持っていた。ただしこの場合でも、有意な振幅を持つ帯域はただだか5kHz程度であったため、本稿ではサンプリング周波数を16kHzと決定し、以降の実験を行った。

#### 4.2 識別実験

2つの識別実験を行った。1つはオフラインでの識別性能の評価であり、もう1つはオンラインでの識別性能の評価である。

単純にログをとり、後で動作履歴を利用するような過去機能をメインとする利用においては、オフラインで十分である。一方、鍵の閉め忘れといったし忘れを指摘するような、現在から未来にかけての機能を想定する場合には、オンラインでの識別が求められる。そのため、両者の識別実験を行い、その性能の差を評価した。

オフラインの識別では、対象とする触音が生じている区間を、実験者が確認しながら切り出している。図3で示したように、触音ではマウスのボタンから指が離れた瞬間なども検出される。今回は、直接的な動作に関連する区間のみを手動で切り出した。一方、オンラインの場合には、Juliusにより、機械的に一定の振動振幅以上の領域を切り出す。したがって、今回の実験におけるオンラインとオフラインの差は、意味のある振動区間を手動で切り出すか、機械的に切り出すかの違いである。オンラインの場合、振動振幅をもとに有音区間が決定されるため、必要十分な区間が切り出されない可能性があり、それが認識率に影響すると考えられる。

まず、オフライン識別の性能評価およびオンライン認識のための HMM の状態数の選定を行った。その後、その HMM を用いてオンライン認識の性能評価を行った。また、触対象の多様性が本手法にどの程度影響するかを明らかにするために、異なるマウスを用いた認

識率の比較を行った。本研究では、HMM の学習にマウスや携帯電話などは同一のものを用いた。しかし、実際の運用時には利用者ごとにそれらの触対象は異なったものを使用する。このため、これらの触対象の多様性は本手法の認識率に大きな影響を与えると考えられる。本稿ではマウスを例にとり、異なったマウスを用いた場合、どの程度認識率が低下するのかを測定した。

##### 4.2.1 オフライン認識

オフラインでの認識率の算出には交差検証法を適用した。データサンプルを7人の被験者から10データずつ抽出し、70データずつの5つのグループに分けた。この5つのグループのうち4つのグループをモデル構築のためのトレーニングデータ、残りの1つを認識率算出のためのテストデータとして用いた。すべてのグループがテストデータとなるように、この操作を5回行い、すべての認識結果をもとに、平均認識率を算出した。

HMM の状態数を変化させながら、認識率を算出した結果を図8に示す。12種類の動作の認識率の平均値を示している。ここで、認識率  $A$  は正解数  $H$  および認識結果の総数  $N$  を用いて、

$$A = \frac{H}{N} \quad (1)$$

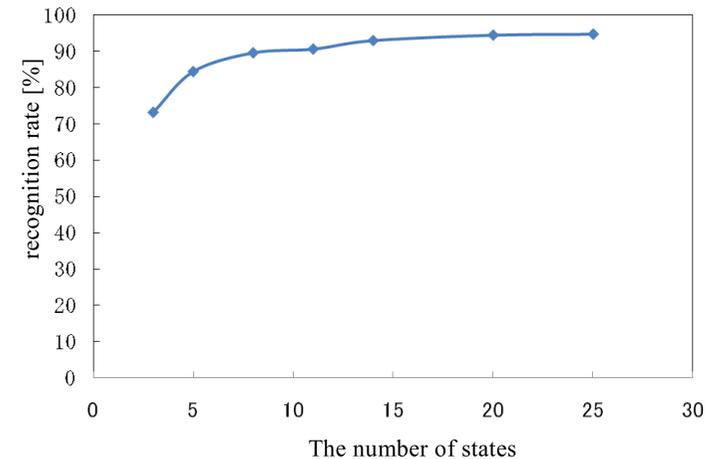


図8 状態数と平均認識率の関係

Fig. 8 Averaged recognition rates of changing the number of states.

と定義した．図より，状態数が増加するにともない認識率が上昇することが分かる．状態数が増加すると，その認識率の値が収束していくことも確認した．

オンライン認識においては，状態数の増加にともない計算時間も増加する．十分な認識率を保ちながら，実時間処理を可能とすることを考慮して，本稿では状態数を 20 として以降の実験を行った．

#### 4.2.2 オンライン認識

オンライン認識では，上記実験結果から選定した状態数が 20 の HMM を利用した．音響モデル作成時と同じ 7 人の被験者に，12 種類の行動を 20 回ずつ行ってもらい，その認識率を算出した．先述のように，オンライン認識の場合，有意な振動区間を振幅に応じて機械的に切り出して認識を行っており，その部分がオフライン認識とは異なる．今回開発したオンライン認識のプロトタイプでは，指定された動作を終えてから約 1 秒後に認識結果が表示される．これは，前述の未来機能的な目的を果たすには，十分な速度であると考えている．

なお対象の識別自体は，振動振幅が閾値を超えた場合には必ず行われ，尤度が最も大きくなる認識結果を出力する．したがって，登録されていないような動作をした場合には，その動作と最も類似していると推定された登録済み動作が出力される．今回の実験では，登録された動作のみを被験者に行わせることを想定したため，未登録動作の認識については考慮していない．また実際，そのような動作は実験中には含まれておらず，以下で示す認識率の結果には影響していない．

図 9 にオフラインとオンラインの識別結果を示す．オフライン識別では平均 94.4%，オンライン識別では平均 84.6%であった．また，12 種類の行動においてオンラインになることで極端に認識率が低下するものはみられなかった．

今回，個人を区別せずに 7 人のデータをまとめて利用し，平均的な触音モデルを構築した．これを利用した結果，男女問わず，十分な認識率が得られることが確認された．この結果は，

- 1) 触対象の違いによる触音の差が大きいこと
- 2) 個人間の皮膚の弾性的な性質の差による波形の違いが小さいこと

をそれぞれ示唆するものである．すなわち，平均的な音響モデルでも，利用者に依存せず，十分な精度で触対象が識別されることが示された．

#### 4.2.3 異なる触対象の識別

次に，オンライン認識において異なったマウスを利用した際の認識率の測定を行った．8 人の被験者（20～30 代男性 6 人，20 代女性 2 人）に，3 つのマウスを各 20 回クリックさ

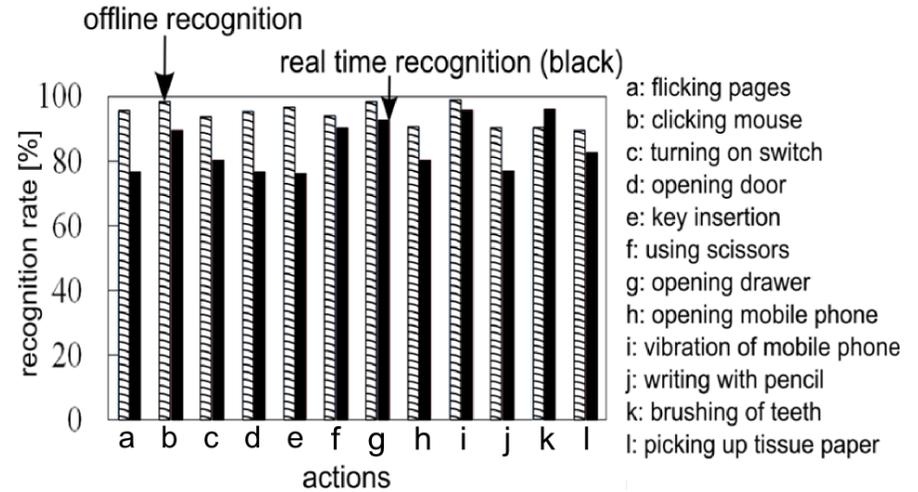


図 9 12 の異なる動作に対する平均認識率  
Fig.9 Averaged recognition rates of 12 different actions.

せ，その認識率を算出した．この結果を図 10 に示す．マウス A が音響モデルの構築の際に使用したものである．図 10 から分かるように，やはりモデル構築に用いたマウス A の認識率が高い．一方，マウス B およびマウス C においても 70%以上の認識率を得た．これは，同じカテゴリの触対象であれば，似た触音を検出できることを示唆する結果である．

以上の実験結果より，触音を利用した触対象の認識では，個人間の差は小さくなく，触音により対象を識別できることを確認した．また，同一カテゴリの触対象間では，モデルと異なるデバイスを利用した場合に認識率は低下するが，似た触音を検出されることを確認した．これらの結果は，提案手法がライフログとして利用可能であることを示すものである．

## 5. 考 察

本研究では，オフライン・オンラインともに，登録された 12 種類の動作のみを行わせた．つまり，4 章で示した認識率は，12 種類の動作の中からある 1 つの動作を行わせた際に，「他の 11 種類ではないと認識されるかどうか」を評価したものである．したがって，12 種類以外の動作を含む様々な日常動作をさせた際に，「正しく識別されるべき場面で認識されなかった」，あるいは「認識されるべきでない未登録の動作時に誤認識された」といった点

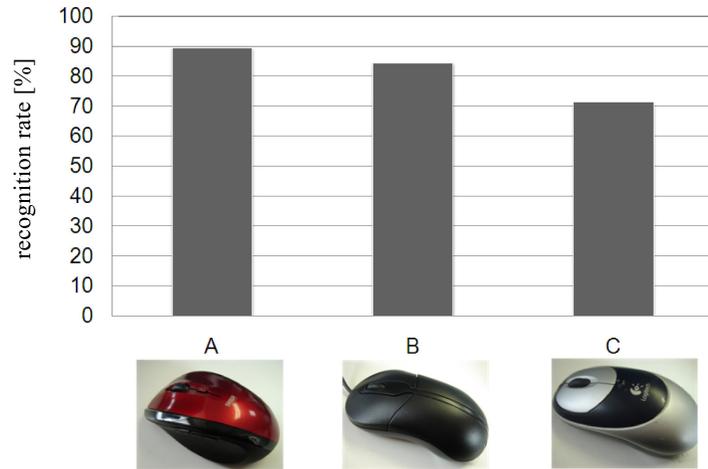


図 10 3 種類のマウスにおける平均認識率  
Fig. 10 Averaged recognition rates of 3 different mouse devices.

は考慮しておらず、今回の認識率の結果には反映されていない。その結果、比較的高い認識率が得られたと考えられる。

今後、実運用を考えた場合、登録されていない動作が入力されることは容易に想像できる。また日常動作の中で、意図せずに物に接触するなど、ノイズが入ってくることも想定される。このとき、現状のシステムでは、閾値以上の振動であれば、登録された動作のいずれかとして判断される。これについては当然、「登録されていない動作」であると認識されるのが望ましい。これは、尤度に注目することで解決可能と考えている。このような未登録の振動の場合には、算出される尤度が低いことが予想される。したがって、それに応じて結果を棄却し、登録されていないデータであることが判別できると考えられる。この実装は今後の課題である。

異なる触対象の識別として今回マウスの認識を行った。結果としては登録されていないマウスの識別率は、低いもので 70% 程度の結果であった。この結果は、同じカテゴリの触対象であれば、似た触音を検出できるということを示唆するものであるが、今回のように、候補が 12 種類の対象に限定された条件での識別結果としては、けっして高いものではない。なお、ここでの誤認識である 30% は、今回のシステムの仕様から分かるように、すべて「マウスクリック以外の動作」として識別されたものである。

今回、マウスは比較的デバイス間での差が少ないであろうと予想し利用したものである。ドアや引き出しの開閉など対象のバリエーションの広いものは、対象により大きな差が出るものと予想される。これを解消するためには、やはり各モデルを作成する段階において学習する対象の種類を増やし、1 つのカテゴリでカバーできる触音の範囲を広げる必要がある。この場合に、誤認識がどの程度生じるかの検証も、今後の課題である。

今回の実験結果をもとに、タッチログがどの程度ライフログに利用可能であるかを考察する。まず他のライフログと併用する場合、タッチログは行動が生じている区間を容易に検出できるため、現状の性能で十分であると考えられる。ただし、拇指にのみ振動が生じるような動作（携帯電話のボタン操作など）も多いため、実利用においては、示指と拇指の 2 本の指に装着するのが望ましい。また、大きなデバイスを爪に装着することは、操作の邪魔となるため、小型化や無線化などを進める必要がある。

併用ではなく、タッチログを単体で利用する場合、ライフログの用途により要求仕様が異なるため、今回の性能で十分かどうかは用途に依存する。日記作成支援のような過去機能としての利用を考えた場合、現状の認識率でも十分であることを我々はすでに確認している。タッチログを記録しながら被験者に 12 種類の様々な動作をさせ、数日後にその動作の順序を思い出させる実験を行ったところ、タッチログを提示した方が、記憶を頼りに思い出させた場合よりも、その正答率が向上した。今回の認識システムでは、意図せぬ接触などでもつねに 12 種類の動作のいずれかと認識されてしまうため、実験において被験者に提示したログは不完全なものであった。しかしその場合でも、被験者は自身の行動を正しく思い出せたことから、過去機能のように自身の記憶と照合するような用途においては、現状の認識率でも十分であると考えている。

一方、テレビの視聴率調査のような社会的未来機能として、マーケティングなどに利用しようとした場合には、誤認識は大きな問題となる。音声認識の精度を向上させると同時に、登録する触対象の数を増やす必要がある。

精度向上を実現するための 1 つの方法として、認識に用いる特徴量を変更することが考えられる。今回は実装の容易さを優先し、ヒトの音声認識で利用される特徴量をそのまま用いた。これらの特徴量は、ヒトの音声のパワースペクトルが持つ特徴をもとに決定されたものであり、触音には必ずしもそぐわない。触音として検出される波形のパワースペクトルが持つ特徴を考慮し、それに応じた特徴量を利用することで、認識精度が向上すると期待される。

音響モデルに関しては、場所に応じて複数に分けてモデルを作成するという方法も考えられる。たとえば部屋にいる場合と、電車に乗っている場合とでは、検出される触音の傾向が

異なる。目覚まし時計のアラームを押すというような触音は、電車内ではけっして検出されない。したがって、GPS データなど、異なるログ情報と比較することで、認識に利用するモデルを切り替えると、より認識率が向上すると考えられる。

## 6. 結 論

本研究では、ヒトが動作する際に生じる接触音や操作音（触音）を、指の振動を経由して爪上で計測するタッチログシステムを提案した。圧電素子を爪に貼付すると、空気中ではなく指先を介して伝搬された指下の振動のみが検出されるため、外来音響ノイズによらず触音を検出できる。また、加速度センサなどと異なり、指の動作や姿勢の影響は受けないため、触音のみを安定に検出できる。これは、触覚センサとして有用であるばかりでなく、触対象が何であるかを検出する際に有用である。

日常、我々が行動を起こす際、多くの場合において指に振動が生じる。よって、上記触音はライフログとして利用可能であると考えた。外来ノイズに強く、動作が生じているときのみ触音が検出されることから、既存のライフログシステムと組み合わせたときに、意味のある行動区間を容易に検出可能である。

また同時に、触音に基づいて接触対象が何であるかを識別できれば、タッチログ単体でもライフログとして利用できる。我々は既存の音声認識手法を適用し、触音に基づいて 12 種類の動作を識別可能であることを確認した。7 人の被験者により得られた触音から、平均的な音響モデルを構築して認識を行った結果、オフラインおよびオンラインの認識率はそれぞれ、94.4%、84.6%であった。これは、個人間の触音の違いよりも、接触対象ごとの触音の違いの方が支配的であることを示す。また、音響モデルの構築に使用していないマウスを用いた認識率の測定を行った結果、70%以上の認識率を得た。これは、同一カテゴリの触対象では、対象ごとの差が出にくいことを示唆するものである。

謝辞 本研究は文部科学省グローバル COE プログラム（環境共生・安全システムデザインの先導拠点）の支援を受けた。

## 参 考 文 献

- 1) Gemmell, J., Bell, G. and Lueder, R.: MyLifeBits: A Personal Database for Everything, *Comm. ACM*, Vol.49, No.1, pp.88–95 (2006).
- 2) 小塚宣秀, 森川大補: ケータイ・ライフログとしての実空間プロファイルと流通・管理技術, *情報処理学会誌*, Vol.50, No.7, pp.603–612 (2009).
- 3) Ueoka, T., Kawamura, T., Kono, Y. and Kidoe, M.: I'm Here!: A Wearable Object

Remembrance Support System, *Proc. 5th International Symposium on Human-Computer Interaction with Mobile Devices and Services*, pp.422–427 (2003).

- 4) Veuri, S., Schmandt, C., Bender, W., Tellex, S. and Lassey, B.: An Audio-Based Personal Memory Aid, *Proc. 6th International Conference on Ubiquitous Computing*, pp.400–417 (2004).
- 5) Maekawa, T., Yanagisawa, Y., Kishino, Y., Ishiguro, K., Kamei, K., Sakurai, Y. and Okadome, T.: Object-based Activity Recognition with Heterogeneous Sensors on Wrist, *Proc. International Conference on Pervasive Computing (Pervasive 2010)*, pp.246–264 (2010).
- 6) Mori, T., Hayama, N., Noguchi, H. and Sato, T.: Informational support in distributed sensor environment sensing room, *Proc. 13th IEEE International Workshop Robot and Human Interactive Communication (ROMAN 2004)*, pp.353–358 (2004).
- 7) De Silva, G.C., Yamasaki, T. and Aizawa, K.: An Interactive Multimedia Diary for the Home, *IEEE Computer*, Vol.40, Issue 5, pp.52–59 (2007).
- 8) 郡司大輔, 荒木拓真, 並木明夫, 明 愛国, 下条 誠: 触覚センサによる滑り検出に基づく多指ハンドの把持力制御, *日本ロボット学会誌*, Vol.25, No.6, pp.172–180 (2007).
- 9) Chigusa, H., Makino, Y. and Shinoda, H.: Large Area Sensor Skin Based on Two-Dimensional Signal Transmission Technology, *Proc. World Haptics2007*, pp.151–156 (2007).
- 10) Kamiyama, K., Vlack, K., Mizota, T., Kajimoto, H., Kawakami, N. and Tachi, S.: Vision-Based Sensor for Real-Time Measuring of Surface Traction Fields, *IEEE Computer Graphics and Applications Magazine*, Vol.25, No.1, pp.68–75 (2005).
- 11) Mascaro, S. and Asada, H.: Measurement of Finger Posture and Three-Axis Fingertip Touch Force Using Fingernail Sensors, *IEEE Trans. Robotics and Automation*, Vol.20, No.1, pp.26–35 (2004).
- 12) 安藤英由樹, 渡邊淳司, 杉本麻樹, 前田太郎: 独立成分分析による爪装着型センサの接触力と屈曲の分離, *日本バーチャルリアリティ学会論文誌*, Vol.8, No.4, pp.379–388 (2003).
- 13) 牧野泰才, 渡部陽一, 前野隆司: 爪上での透過光計測によるタッチパネル操作時の力推定, 第 27 回センシングフォーラム講演論文集, pp.157–160 (2010).
- 14) 湯村武士, 安藤英由樹, 前田太郎: 爪上装着型触覚提示技術を利用した触覚伝送手法の研究—凹凸感伝送のための加速度計測, *日本バーチャルリアリティ学会第 13 回大会論文集*, pp.471–474 (2008).
- 15) 福本雅朗, 外村佳伸: “指鉤”: 手首装着型コマンド入力機構, *情報処理学会論文誌*, Vol.40, No.2, pp.389–398 (1999).
- 16) Iwamoto, T. and Shinoda, H.: Finger Ring Tactile Interface Based on Propagating Elastic Waves on Human Fingers, *Proc. World Haptics*, pp.145–150 (2007).

- 17) 仲谷正史, 川副智行, 豊田成人, 木下智史, 和田 潤: 指の変形に注目したウェアラブル接触力センサシステムの検討, 日本機械学会ロボティクス・メカトロニクス講演会'10, 講演論文集 1A1-C22 (2010).
- 18) 田中由浩, 佐野明人, 伊藤真由実, 藤本英雄: 触覚ネイルチップ—デバイス設計と基礎実験, 日本機械学会ロボティクス・メカトロニクス講演会講演論文集, 2P1-N08 (2007).
- 19) 田中由浩, 佐野明人, 藤本英雄: 自己言及性と双方向性を考慮した触覚センシング, 日本機械学会ロボティクス・メカトロニクス講演会'10, 講演論文集 1A2-D17 (2010).
- 20) 佐藤浩一: 思い出の中の教師: 自伝的記憶の機能分析, 群馬大学教育学部紀要, 人文・社会科学編 49, pp.357-378 (2000).
- 21) 太田信夫: 記憶の心理学と現代社会, 有斐閣 (2006).

(平成 22 年 6 月 28 日受付)

(平成 23 年 1 月 14 日採録)



牧野 泰才 (正会員)

2002 年東京大学工学部計数工学科卒業。2004 年同大学大学院情報理工学系研究科システム情報学専攻修士課程修了。2007 年同専攻博士課程修了。同年 4 月より同専攻にて、学術振興会特別研究員。2008 年 4 月より特任研究員。2009 年 4 月より慶應大学にて、特別研究助教。触覚情報処理の研究に従事。



村尾 将和

2008 年慶應義塾大学工学部機械工学科卒業。2010 年同大学大学院システムデザイン・マネジメント研究科修士課程修了。



前野 隆司

1984 年東京工業大学工学部機械工学科卒業, 1986 年同大学大学院理工学研究科機械工学専攻修士課程修了, 同年キャノン株式会社入社。1993 年博士(工学)学位取得(東京工業大学)。1995 年慶應義塾大学専任講師, 同大学助教授を経て, 現在, 慶應義塾大学大学院システムデザイン・マネジメント研究科教授。