



ウェアラブルコンピューティングのための状況依存 音声情報提示手法

矢高, 真一
田中, 宏平
寺田, 努
塚本, 昌彦
西尾, 章治郎

(Citation)

情報処理学会論文誌, 51(12):2384-2395

(Issue Date)

2010-12-15

(Resource Type)

journal article

(Version)

Version of Record

(URL)

<https://hdl.handle.net/20.500.14094/90002099>



ウェアラブルコンピューティングのための 状況依存音声情報提示手法

矢 高 真 一^{†1} 田 中 宏 平^{†2,*1} 寺 田 努^{†1,†3}
塚 本 昌 彦^{†1} 西 尾 章 治 郎^{†2}

音声を用いた情報提示は、イヤホンなどの小型デバイスでハンズフリーで利用でき、視覚情報と比較して他の作業への影響が小さいことからウェアラブルコンピューティングにおける情報提示方式として期待されている。しかし騒音などの周辺状況や会話中などの活動状況の影響を受けやすく、ユーザが提示情報を聞き取れない場合が多い。また、不必要に大きな音量で情報提示を行うと、ユーザのストレスは大きくなる。そこで本研究では、ウェアラブルコンピューティング環境におけるユーザ状況を考慮した音声情報提示手法を提案する。提案手法では、マイクやセンサを用いてユーザ状況を取得し、予備実験の結果をもとに適切な方法で音声情報を提示することで確実に情報をユーザに伝える。また、本研究では評価実験により、提案手法の有効性を確認した。

A Context-aware Audio Information Presentation Method for Wearable Computing

SHINICHI YATAKA,^{†1} KOHEI TANAKA,^{†2,*1}
TSUTOMU TERADA,^{†1,†3} MASAHICO TSUKAMOTO^{†1}
and SYOJIROU NISHIO^{†2}

In wearable computing, the audio information is suitable for information presentation because it can be used without hands by wearable small earphones, and it has little influence on other tasks compared with other information presenting methods. However, since the presented sound is frequently affected by surrounding situations such as background noise and conversational voices, the user cannot catch the presented information while too loud sound annoys him/her. Therefore, we propose an audio presenting method considering user contexts for wearable computing. Our method acquires the user context from wearable sensors and microphone, and outputs audio information controlled to be surely transmitted to the user. The evaluation results confirmed the effectiveness of our method.

1. はじめに

近年の計算機の小型化・軽量化にともない、コンピュータを身に付けて利用するウェアラブルコンピューティングへの期待が高まっている。ウェアラブルコンピューティング環境において、ユーザは移動中や他の作業をしている途中などさまざまな状況で情報を閲覧する。したがって、従来のデスクトップコンピューティング環境のような視覚を主に用いた情報提示ではつねに多くの情報を快適に閲覧できるとはいえず、ユーザ状況を考慮し、聴覚や触覚などのさまざまなメディアを組み合わせた情報提示が必要である。たとえば、ユーザが作業に集中しているときに頭部装着型ディスプレイ（HMD：Head Mounted Display）に情報を提示されてもそれに気がつかないことがあるため、振動などの触覚情報でユーザの注意を引くことが有効である。また、歩行中や運転中はHMDを用いた情報閲覧は危険がともなうため、音声など歩行や運転に影響を与えにくい情報提示方法が適しているといえる。

ここで、音声情報提示はイヤホンのような一般的なデバイスを用い、また受動的に情報を得られるためウェアラブルコンピューティングに適した情報提示方式として特に期待されている。津村らによるウェアラブルコンピュータに適した情報提示方法の評価¹⁾では、タスク遂行中に視覚情報提示と音声情報提示の認識度や快適度などを調べており、静かな環境下であれば音声メディアがテキスト情報の提示に最も適していると述べている。

しかし、音声情報提示においても、ユーザ状況によって得られる情報量が変化するという問題点がある。たとえば、提示音量が小さいと騒音下では提示された内容が聞き取れないことがあり、集中が必要なタスクを行っているときには提示内容を聞き逃すことも考えられる。一方、提示音量をあらかじめ大きく設定しておいたり、提示前にチャイム音を鳴らしユーザの注意を引いてから提示したりすれば聞き逃すことは少なくなるが、ウェアラブルコンピューティング環境ではつねに情報提示デバイスを身に付けているため、静かなところで

^{†1} 神戸大学大学院工学研究科
Graduate School of Engineering, Kobe University

^{†2} 大阪大学大学院情報科学研究科
Graduate School of Information Science and Technology, Osaka University

^{†3} 科学技術振興機構さきがけ
PRESTO, Japan Science and Technology Agency

*1 現在、三菱電機先端技術総合研究所
Presently with Advanced Technology R&D Center, Mitsubishi Electric Corporation

不必要に大きい音量で提示されたり、チャイム音が鳴ったりすることでユーザはストレスを感じる。そこで本稿ではユーザ状況と提示された音声情報の認識率の関係を調べ、動的に変化するユーザの状況を考慮した音声情報提示手法を提案する。ここで、本稿では「活動状況」は歩行や会話などのユーザ自身の状況、「周辺状況」は周辺音量などのユーザの周囲の状況、「ユーザ状況」はこれらの活動状況、周辺状況を合わせたものと定義している。

本稿は以下のように構成されている。2章でウェアラブルコンピューティング環境における情報提示手法と音声情報を用いたアプリケーションの関連研究について記述し、3章では本研究の想定環境について詳しく説明する。4章では提案手法について述べ、5章で評価実験と考察を行う。最後に6章でまとめと今後の課題を述べる。

2. 関連研究

ウェアラブルコンピューティング環境では、さまざまな状況下で他の作業と並行してコンピュータを操作する。そのため、従来のようにただディスプレイに情報を映すという単純な情報提示ではユーザ状況の影響で得られる情報が減少してしまうという問題がある。

この問題を解決するため、これまでにユーザ状況を考慮した情報提示手法が提案されてきた。田中らによるウェアラブルコンピュータのための知覚影響度に基づく情報提示手法²⁾は、ユーザの利用できる情報提示デバイスの特徴とサービスが提示させたい情報の特徴から提示可能なデバイスを選択し、最適な提示方法で情報を提示する機構である。たとえば、ユーザが車の運転中などディスプレイに目をやれない場合にテキスト情報を提示する場合、テキスト情報を音声情報に変換し提示することでユーザに正しく情報を伝える。この研究は情報提示デバイスを切り替える枠組みの提案であり、各提示デバイスがどのような場合に提示に不適切になるのかという点には言及しておらず、各提示デバイスごとに情報提示の適切さに関する評価関数があらかじめ与えられていることを想定している。したがって、本研究の結果をこの機構の音声提示部の評価部分として動かすことで、汎用的なデバイスの切替えにも対応できる。また、田中らによるシースルー型頭部装着ディスプレイのための背景を考慮したオブジェクト配置手法³⁾は背景が透過できるシースルー型HMDにおいて、HMDの背景の明暗などの影響で画面の一部が見えにくくなった場合に表示する情報の配置を動的に変更し、情報の視認性を高めている。このように、ユーザ状況を考慮することで情報の伝達性能を高めることができる。

振動などの触覚情報を用いて情報提示を行う手法として、沖野らによる触覚デスクトップ⁴⁾や塚田らによるActive Belt⁵⁾がある。触覚デスクトップでは触覚提示デバイスを用い

てHMDでは判断しにくい細かな操作を補助することを目的としている。Active Beltはベルトに方位センサ、GPS、アクチュエータを取り付け、目的地へたどり着くにはどちらに向かえばよいのかをアクチュエータの振動で知らせる。どちらの研究も他の作業との並行性は高いが、提示できる情報量は音声と比較すると少ない。また、状況に応じたユーザへの伝達性能の変化については考慮されていない。

モバイル環境において、メール受信などの通知をユーザ状況を考慮して提示する手法についても多く提案されている。Kernらは、加速度センサやマイクによってユーザの状況をユーザの中断可能性、周囲の中断可能性という2軸上にマッピングし、そのマッピングされた位置によりピープ音のみで通知するか、画像と音声を併用して通知するかなどを決定し提示するシステムを提案している⁶⁾。SawhneyらによるNomadic Radio⁷⁾はマイクによりユーザが会話しているかどうかを認識し、提示する情報の優先度と過去の履歴から、アラート音や読みあげなどに提示方法を変更する。この研究ではユーザ状況により得られる情報量に変化するが、本研究では提示音量などを変更することにより、情報量を変えずユーザに確実に情報を伝達する手法を目指す。また、これらは通知という情報量の小さい情報提示に関する取り組みであり、ニュースやナビゲーションといった長文の再生は考慮していない。

ウェアラブルコンピューティングにおける音声情報についても、多くのアプリケーションが提案されている。庄司らによるSoundWeb⁸⁾は、音声データにハイパーリンクの概念を付与し、音声データを相互に関連付ける。別の音声データへのリンクがある部分の再生時に同時に信号音を流すことでユーザにリンクの存在を知らせ、音声のみで構造化された音声情報を閲覧することを可能にしている。WilsonらによるSWAN⁹⁾は進行方向や近くのパス停やコンビニなどの存在を信号音などの非言語音を用いて提示する。広田らによるシステム¹⁰⁾は、聴覚定位を利用してユーザの周囲に仮想的にアイコンを配置し、音声の聞こえた方向をポインティングすることでコンピュータを操作できる。在塚らによるシステム¹¹⁾ではGUI画面のレイアウトを空間的な音の配置で効果的に表現し、直観的に操作することを可能にしている。RoyらによるNewsComm¹²⁾はあらかじめ登録しておいた自分の趣向に応じて、ニュースなどをネットワーク上から自動で取得して提示する。これらの音声提示アプリケーションでは、ユーザ状況は考慮されておらず、ユーザが正しく、かつ快適に情報を取得できない場合がある。したがって、これらのアプリケーションと提案手法を組み合わせることでより有効なシステムが構築できる。

3. 想定環境

提案手法では、ユーザがコンピュータを身に付けて生活するウェアラブルコンピューティング環境を想定している。ウェアラブルコンピューティングにはハンズフリーで常時利用できるという特徴がある^{13),14)}。このような特徴を利用して、腕や腰に付けたセンサからさまざまな生体データを得ることで健康を保つ支援をするシステム¹⁵⁾や、ウェアラブルコンピュータと中央制御用コンピュータを高速無線通信でつなげることで、作業しながら電子マニュアルを確認したり、制御室と音声や画像の通信をしたりするシステム¹⁶⁾などが実現されている。このようなアプリケーションでは健康状態の異常や遠隔地からの指示など、音声情報がよく用いられている。そのため、本研究では音声情報の提示を想定する。

また本研究では、ユーザが音声情報提示デバイスとして骨伝導イヤホンを着用していることを想定する。一般のヘッドホンのような耳をふさぐデバイスは、ウェアラブルコンピューティング環境では周囲の音が聞こえにくくなるため危険をとまなう。骨伝導イヤホンは振動を鼓膜を介さずに伝えるため図1に示すように耳をふさがず、周辺音も同時に聞き取れる。

さらに、ユーザは活動状況や周辺状況を取得するためのセンサやマイクを着用しているとする。これらのデバイスは小型で、装着していても日常生活の邪魔にならず、ナビゲーションサービスや健康管理システムなどさまざまなアプリケーションでの利用が想定されているため、ウェアラブルコンピューティング環境での利用は一般的である。



図1 骨伝導イヤホンの装着例

Fig. 1 A snapshot of a user wearing a bone-conduction earphone.

4. 提案手法

本研究では、ウェアラブルコンピューティング環境においてユーザが提示された情報を認識でき、かつ不必要にユーザにストレスを与えない方法で音声情報を取得したいという要求に対し、ユーザ状況を考慮して音声情報を提示する手法を提案する。具体的にはニュースの読み上げやナビゲーションサービスの音声案内などで音声情報を出力する際、ユーザの活動状況や周辺状況を考慮し、ユーザが音声情報を認識しやすい方法で提示する。また、提案手法ではユーザの希望する聞き取り精度や情報の持つ緊急性などの性質により提示方法を変更する。たとえば作業支援システムにおいて電子マニュアルの読み上げを行う場合、ユーザの聞き逃しは作業ミスや危険につながる。一方、ニュースの読み上げなどにおいて、ユーザがおおよその文意が伝わればよいと設定した場合、精度は低くてもかまわないため作業の邪魔にならないことを優先する。

本章ではまず、ユーザ状況と提示された音声情報の認識率の関係を明らかにするための予備実験について述べ、次に実験結果に基づき提示方法を制御する手法について述べる。

4.1 予備実験

音声情報を提示する際、小さすぎる提示音量ではユーザは情報を認識できない。一方で不必要に大きい音量で提示した場合はユーザはストレスを感じる（この点については5章で評価を行う）。そこで、ユーザ状況に適した音量を調べるために次に示す方法でユーザの活動状況や周辺音と、提示された音声情報の認識率の関係を調査する予備実験を行った。

4.1.1 実験方法

実験では、図2に示すように腰に小型PC、首元にマイクを装着した被験者に、アルファベット4文字を成年男性の声（400～450 [Hz]）で順に読み上げた音声情報を、平均1分とするポアソン到着で提示した。被験者は聞こえたアルファベット4文字を記録した。また実験環境に設置したスピーカからBGM（日本語の歌詞あり）を流して周辺音量を制御し、録音した。ここで、活動状況において、「体の動きの大きさ」、「他のタスクへの集中力」、「会話の有無」が音声情報の認識率に影響を与える要因であるという仮説を立てた。「会話の有無」は、同時に2人以上の話を聞くのは困難であるという筆者らの経験や文献17)で述べられている知見から、他の目や指を使った作業とは異なる影響を音声情報の認識に与えると予想したため要因として含めた。このような仮説に基づき、図3のようにマッピングされる以下の5つの活動状況について調査した。

- タスクなし



図 2 実験の様子

Fig. 2 A snapshot of a user in the experiment.

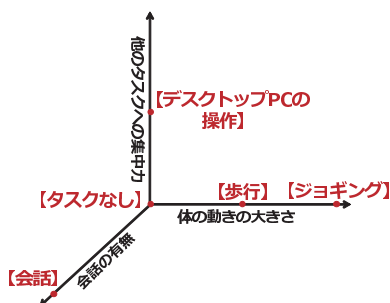


図 3 ユーザの活動状況

Fig. 3 User contexts for the evaluation.

- 歩行
- ジョギング
- デスクトップ PC の操作
- 会話

歩行、ジョギングは「体の動きの大きさ」、デスクトップ PC の操作は「他のタスクへの集中力」、会話は「会話の有無」がそれぞれどのように音声情報の認識に影響を与えるかを調べるために選択した。デスクトップ PC の操作は、文書作成やウェブブラウジングといった比較的集中力の小さい作業を行った。これは現在多くの人々が仕事やプライベートにおいて頻繁に行っており、ウェアラブルコンピューティングにおいても同様のタスクは多くなると考えた

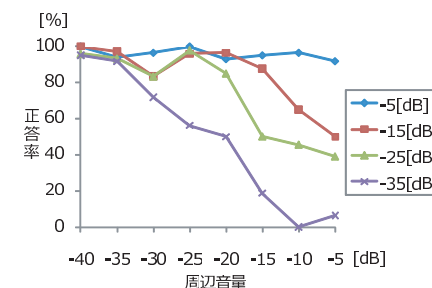


図 4 タスクなしのときの正答率

Fig. 4 Accuracy with no task.

ためである。会話時における実験は、被験者を含む 21～22 歳の学生 3～4 人があるテーマについて話しをし、被験者が会話中に音声情報が提示されるように、第三者が制御した。実験にはウェアラブルコンピュータとしてソニー社の VAIO UX90S (OS: Microsoft Windows XP Professional, CPU: インテル Core Solo プロセッサ U1400 (1.20 GHz), メモリ: 1 GB), 骨伝導イヤホンに TEMCO 社の HG40SAN-TBT, マイクにはパナソニック社の RP-VC201 を用いた。提示した音声の音量は減衰なしの音量と比較して -5 dB, -15 dB, -25 dB, -35 dB で提示した。なお、被験者は 21～22 歳の男性 5 人である。

4.1.2 実験結果と考察

タスクなしのときのアルファベット 1 文字ごとの正答率を図 4 に示す。結果から、今回の実験の最大音量である -5 dB で提示したときの正答率は安定して高く、平均正答率は 97% であった。このような状況での誤答は b[bi:] と d[di:] などの母音が同じで子音のみが異なるアルファベットがほとんどであったため、アルファベットを聞き分けるというタスクの難易度だといえる。一方で -15 dB 以下の提示音量では全体として、周辺音が大きくなるにつれて正答率が下がる傾向がみられることから、周辺音量が音声情報の認識に影響を与えていることが分かる。

次に各活動状況の違いに注目する。提示音量ごとの周辺音量に対する正答率を図 5 に示す。まず会話時以外について述べると、周辺音量が大きくなるにつれて、また提示音量が小さくなるにつれて、正答率が下がる傾向がみられる。一方、会話時を除く活動状況間の正答率の違いはあまりみられない。したがって、身体動作や他のタスクへの集中力は音声情報の認識にほとんど影響を与えないことが分かる。ただし、タイムトライアルやリズムゲームのようなより集中力の必要な作業では音声情報の認識に影響が出ると予想されるため、今後検

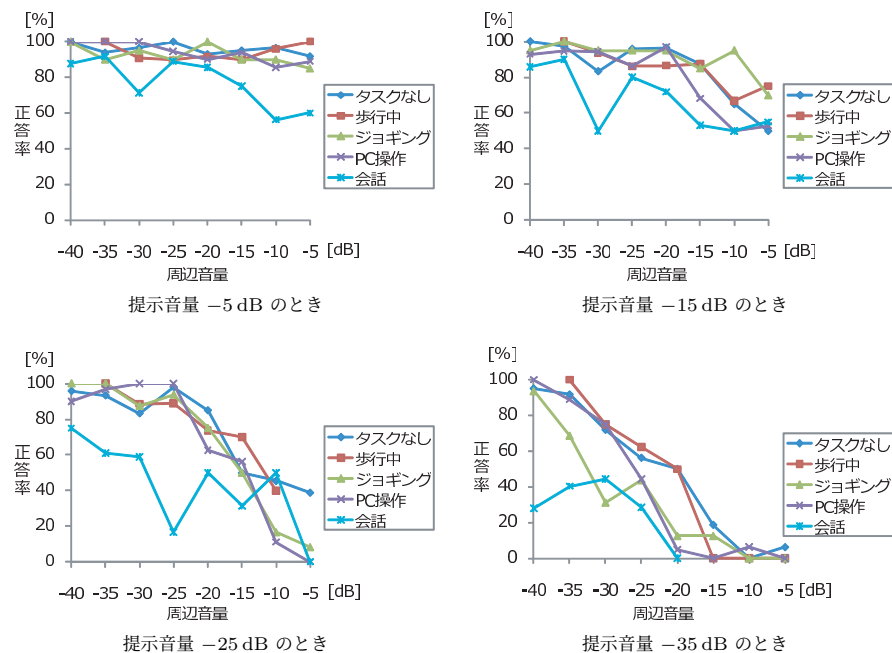


図 5 提示音量 -5 dB, -15 dB, -25 dB, -35 dB のときの正答率

Fig. 5 Accuracy with -5 dB, -15 dB, -25 dB and -35 dB presentation sound volume.

討が必要である。次に会話時について述べると、周辺音量にかかわらず正答率が上下した。これは集中する対象が話すことや人の話を聞くことであり、会話中に集中の度合いが頻繁に変化し、会話と同時に音声聞き取ることが困難だった場合があったためだと考えられる。

4.1.3 音声情報の認識のために必要な提示音量

予備実験の結果から、各活動状況における周辺音量と必要な提示音量の関係を明らかにする。各活動状況において正答率が一定値（閾値 T ）を超えたときの周辺音量と提示音量の関係を表す散布図を図 6、図 7、図 8、図 9、図 10 に示す。ここで、ユーザが確実に音声情報を聞き取れる閾値として $T = 94\%$ 、ある程度聞き逃しはあるがおおよその文意は伝わる閾値として $T = 85\%$ と設定する。前者の閾値は、周辺音量が -30 dB 以下の環境でタスクなしのときに、 -5 dB で提示した際の平均正答率 97% に対して、信頼区間が 3% であったことによる。後者の閾値はおおよそ 10 文字に 1 文字聞き間違ってもよい状況を想定したこ

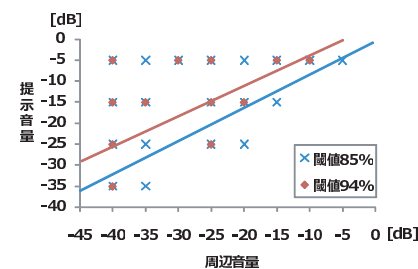


図 6 タスクなしのときの散布図

Fig. 6 A point diagram for no task.

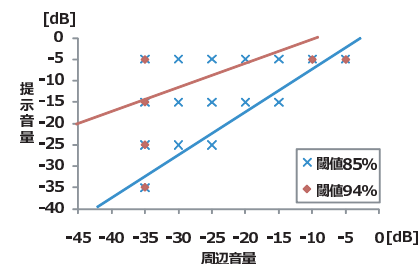


図 7 歩行時の散布図

Fig. 7 A point diagram for walking.

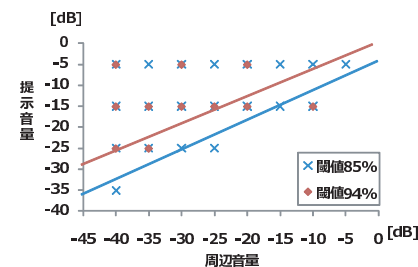


図 8 ジョギング時の散布図

Fig. 8 A point diagram for jogging.

とによる。結果より、会話以外の状況では周辺音量と認識に必要な最小の提示音量には dB スケールにおいて比例関係がみられることが分かる。この傾向から、会話以外の各活動状況において、各周辺音量に対して正答率が閾値 T を超える最小の提示音量について、最小二

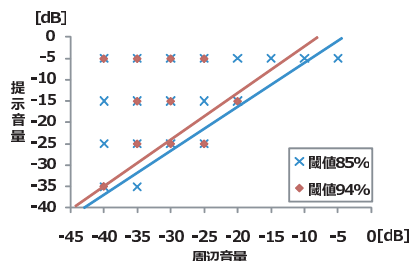


図 9 PC 操作時の散布図

Fig. 9 A point diagram for using a PC.

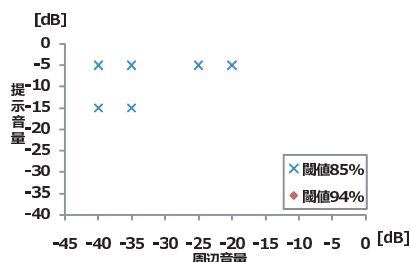


図 10 会話時の散布図

Fig. 10 A point diagram for conversation.

乗法を用いて近似直線を求めた。求めた直線を図 6～図 9 に示す。本研究の提案する手法ではユーザが会話中でなければ、この導出した直線を用いて、周辺音量、活動状況から最適な提示音量を算出する。

ただし、周辺音量が大きすぎる場合、導出された提示音量が正の値となり提示できない。また、会話中の場合、周辺音量と認識に必要な提示音量に比例関係はみられず、提示音量を変更しても十分な正答率は得られない。このような場合は、提示前にアラート音を鳴らしユーザの注意を引いたり、提示可能となってから提示したりするといった手法を用いる必要がある。

4.2 音声情報提示手法

提案手法の動作フローを図 11 に示す。提案手法は以下のステップで動作する。

- (1) 活動状況、周辺音、提示する程度の取得

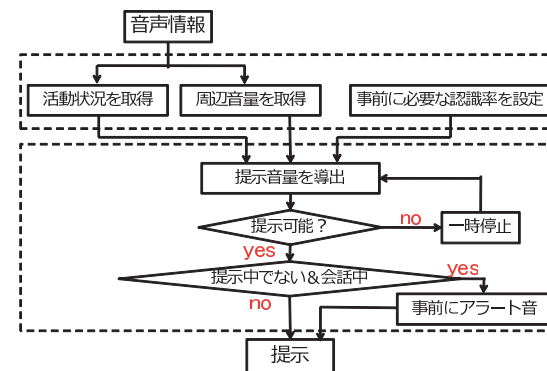


図 11 提案手法の動作チャート

Fig. 11 A flowchart of proposed method.

提案手法ではあらかじめ希望する聞き取り精度をユーザが入力する。そして音声情報を出力する際とそれ以降 n 秒間隔（本研究では $n = 2$ とした）で、周辺音量、活動状況を取得する。周辺音量はマイクを使用して測定し、活動状況については 3 軸加速度センサやマイクを用いて認識する。活動状況の取得方法に関しては現在さかんに研究が行われており、本研究ではタスクなし、歩行、ジョギング、PC 操作については文献 18), 19) で示されているような、会話に関しては文献 7) で示されているような手法を利用することを想定している。

- (2) 取得情報から最適な提示方法の判断

活動状況が会話以外であれば、4.1.3 項で求めた近似直線の式を用いて提示音量を決定する。その際、導出された音量が大きすぎて提示が不可能な場合は一時停止し、提示可能となった後に提示する。ユーザが会話中であれば、ユーザの注意を引くために提示前にアラート音を鳴らし、タスクなしの場合の関係式を用いて提示音量を決定し提示する。ただし、すでに情報提示中の場合はアラート音の再生は行わない。

4.3 実装

提案システムのプロトタイプシステムを実装した。開発は Windows XP 上で Microsoft Visual C# 2005 を使用し、音声情報制御のために DirectX の DirectSound を使用した。実装したシステムは、3 軸加速度センサとマイクを使用して活動状況を取得する。具体的には、まず「タスクなし」、「歩行」、「ジョギング」、「PC 操作」については右腕、腰、右足に装着

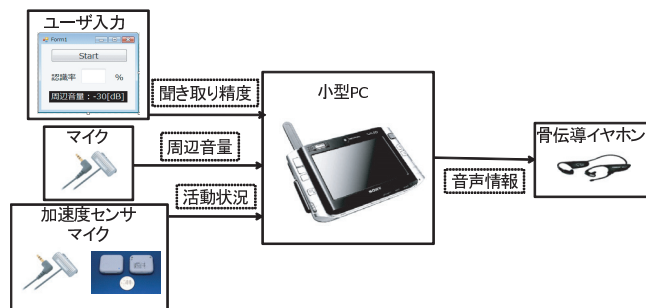


図 12 システム構成
Fig. 12 System structure.

した 3 つの 3 軸加速度を用いて認識を行い、センシングデータを 0.5 秒間の平均値および分散値を特徴量として、SVM²⁰⁾ により学習および認識を行う。「会話」については首元に装着したマイクを用い、取得した音に言語が含まれているかどうかを音声認識器 Julius²¹⁾ を用いて調べ、言語が含まれていればそれ以降 10 秒間を「会話中」とした。システム構成を図 12 に示す。またアプリケーションとして、twitter²²⁾ のタイムラインの更新内容を読み上げる twitter 音声クライアントと現在地付近のグルメ情報などを音声で配信する地域情報サービスを実装し、動作の確認を行った。

5. 評価実験

提案手法の有効性をプロトタイプシステムを用いて評価する。評価では、ユーザ状況による提示音量の変更、提示前のアラート音、提示不可能な場合の一時停止の 3 つの手法が、それぞれどのような場合について効果的であるかを明らかにする。

5.1 提示音量変更の評価

状況によって提示音量を変更する手法の有効性を評価した。実験は予備実験と同様に、被験者は腰に小型 PC、首元にマイク、耳元に骨伝導イヤホンを装着し、平均 1 分のポアソン到着で提示されたアルファベット 4 文字を記録する。また、同時に提示音量の適切さを 5 段階（不必要に音量が大きく不快であるとき 5、聞こえにくいとき 1、適切なときは 3）で主観評価させた。比較対象としては、-5 dB、-35 dB の一定の提示音量で提示したものを用いた。実験した活動状況は音量変更が有効と考えられるタスクなしのとき、歩行時、PC 操作時とした。ただし、活動状況はセンサを用いて認識したのではなく、あらかじめ固定して

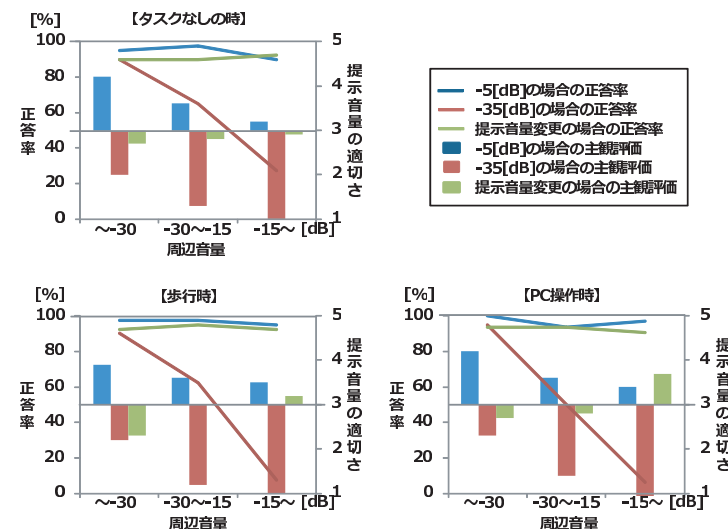


図 13 提示音量変更の評価実験結果
Fig. 13 The evaluation result in changing presentation sound volume.

行った。周辺音量の制御には実験環境に設置したスピーカから流す BGM の音量を変更することで行い、-30 dB 以下、-30 dB から -15 dB、-15 dB 以上の 3 つについて行った。実験では、おおよその文意が伝わればよいと考えられるニュース速報や天気予報を読み上げるアプリケーションを想定し、閾値 T は 85% とした。なお、被験者は 21~23 歳の男性 5 人であった。

活動状況ごとの周辺音量に対する正答率と主観評価のグラフを図 13 に示す。まず正答率に注目すると、提示音量を動的に変更した場合の正答率は活動状況や周辺音量にかかわらず高く、今回の閾値である 85% を超える値となった。提案手法と提示音量を -5 dB、-35 dB に固定した場合の正答率について、それぞれ有意水準 0.05 の Fisher の正確確率検定を用いて有意差を判別した。-35 dB で提示した場合と提案手法を比較すると、すべての活動状況において周辺音量が -30 dB 以上になると、提示音量を動的に変更する場合の正答率が有意に高くなった。一方、-5 dB で提示した場合と提案手法ではすべての周辺音量と活動状況において有意差はなかった。結果から、提示音量を大きな値に固定しなくても、活動状況や周辺音量によって動的に音量を変更することで高い正答率が得られることが分かる。

主観評価については図 13 より、提示音量を -35 dB に固定した場合の評価は、周囲が静かなときは 3 に近いが、周辺音量が大きくなるにつれて 1 に近づいている。つまり周囲が静かなときは適切な音量で提示できているが、周辺音量が大きくなると被験者は提示内容を聞きとりにくいと感じたことが分かる。これは正答率の結果からも明らかである。また提示音量を -5 dB の一定の提示音量で提示したときには、周囲が静かであれば評価は 3 を大きく超え、周辺音量が大きくなるにつれて 3 に近づく。つまり、周りが騒がしければ提示音量は適切であるが、静かであれば不必要に大きい提示音量なため、被験者は不快に感じたことが分かる。一方、提案手法を用いて提示音量を動的に変更したときの評価は周辺音量にかかわらず 3 に近い値が得られており、適切な提示音量であったといえる。

また実験後、被験者に「ウェアラブルコンピューティング環境において音声情報を提示されることについてどう思うか」、「提示音量が変更されることに違和感を感じるか」という 2 つの問いからなるアンケートを行った。前者の問いに対する回答としては、「メインの情報提示手段としては少し物足りないが、補助的な利用は便利である」「メールが来るたびにその内容を読み上げられることは煩わしいが、送信者、件名などが音声で提示されたら助かる」「目の不自由な人には便利である」「ナビゲーションなどのアプリケーションで進行方向から音がすれば便利である」といった意見が得られた。後者の問いに対する回答としては、「周囲の音量によって聞こえ方が違うから自動で音量調整してくれるのは助かる」「自動調整と手動調整を使い分ければ便利である」「自分が想定している音量より大きい音量で急に提示されると不快である」といった回答が得られた。

前者の質問に対する回答から、ウェアラブルコンピューティング環境において音声情報のみで大きな情報量を提示するのはユーザのストレスを大きくすると思われる。しかし、用途によって音声情報を視覚情報などと使い分けて用いれば効果的に情報提示が行えるといえる。また、後者の質問に対する答えから、ユーザ状況を考慮して提示音量を変更することは有効だと考えられるが、急に大きく音量を変更することはユーザのストレスを大きくすることにつながるので適切に変化させる必要があることが分かる。

5.2 アラート音の評価

音声情報の提示前にアラート音を鳴らすことでユーザの注意を引く手法の有効性を評価した。実験方法は提示音量変更の評価実験と同様で、アラートありの場合とアラートなしの場合を比較した。ただし、主観評価として提示の煩わしさを 5 段階で評価させ、被験者が煩わしくないと感じると評価は 1 に近づき、煩わしいと感じると評価は 5 に近づくようにした。また、提示音量はユーザ状況によって動的に変更した。実験した活動状況はタスクなし

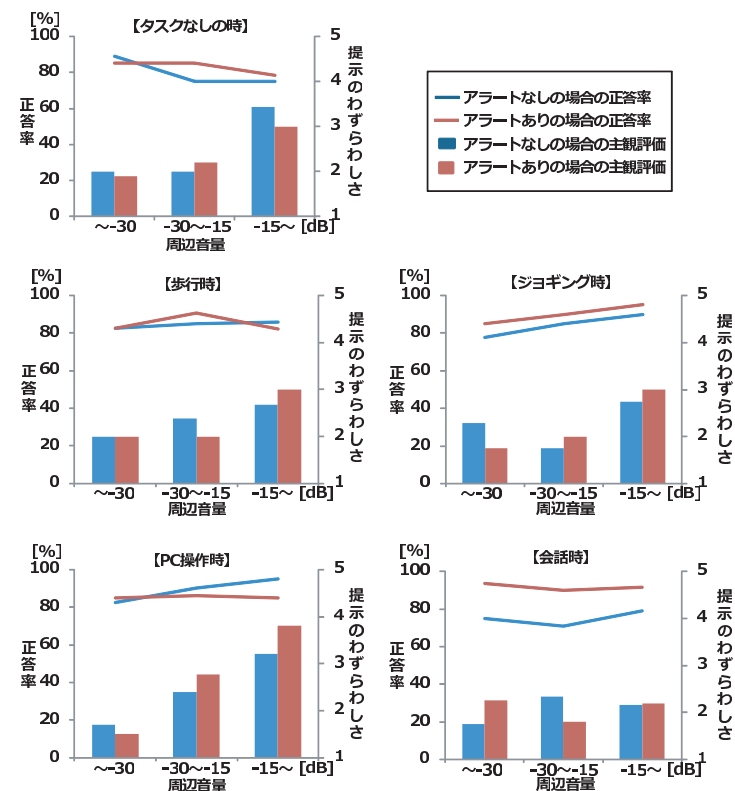


図 14 アラート音の評価実験結果
Fig. 14 The evaluation result on the effect of alert.

のとき、歩行時、ジョギング時、PC 操作時、会話時である。

活動状況ごとの周辺音量に対する正答率と主観評価のグラフを図 14 に示す。まず正答率に注目すると、会話以外のすべての活動状況において、アラートの有無にかかわらず正答率は閾値 T の 85% に近い値となった。アラートの有無による正答率の変化を有意水準 0.05 の Fisher の正確確率検定を用いて判別した結果、有意差はみられず、会話以外のすべての活動状況であればアラート音を鳴らすことは正答率に影響を与えないといえる。一方、会話時には周辺音量が -30 dB 以下、 -30 dB から -15 dB の場合、アラート有無によ

る正答率の変化に有意差があり、アラートありの場合の正答率が高くなった。周辺音量が-15 dB の場合は有意差はみられなかったが、他のタスクと比較してアラートの有無による正答率の違いは大きくなった。これはアラート音を鳴らすことで、話すことや人の話を聞くことに集中している被験者の注意を向けることができたためだと思われる。

次に主観評価については、図 14 より、アラートの有無による違いはあまりみられないことが分かる。アラートの有無による主観評価の違いを有意水準 0.05 の Wilcoxon の符号付き順位検定を用いて判別した結果、すべての活動状況、周辺音量において、有意差はみられなかった。したがって、音声情報提示前のアラート音の有無はユーザが感じる提示の煩わしさに影響を与えないことが分かる。ここで会話時に注目すると、提示前にアラートを鳴らすことで正答率が上がり、アラートの有無は提示の煩わしさに影響しない。つまり会話時には提示前にアラートを鳴らすことで効果的な情報提示が実現できるといえる。また、周辺音量が大きくなるにつれてアラートの有無にかかわらず提示の煩わしさが大きくなる傾向がみられる。一方、周辺音量によって提示音量を変更しているため、正答率は周辺音量によってあまり変化がみられない。つまり、周辺音量が大きくなると、提示音量を大きくすることでユーザに音声情報を伝えることはできるが、その提示にユーザは煩わしさを感じるという。

5.3 一時停止の評価実験

導出された音量が大きすぎて音声情報の提示が不可能な場合に、提示可能になるまで一時停止した後提示する手法の有効性を評価する実験を行った。実験方法は提示音量変更の評価実験と同様であり、一時停止する場合と、一時停止せず、その場で提示できる最大の提示音量で提示する場合を比較した。ただし、主観評価として提示の煩わしさを 5 段階で評価させ、被験者が煩わしくないと感じると評価は 1 に近づき、煩わしいと感じると評価は 5 に近づくとした。周辺音量は、提示不可能な提示音量を導出する大きな音量を実験環境に設置したスピーカから流した。実験した活動状況は歩行時、ジョギング時、PC 操作時である。タスクなしの場合にはすべての周辺音量に対して提示可能な提示音量を導出できるため実験を行わなかった。また、周辺音量が非常に大きい中で会話を行うのは困難であるため、会話時についても実験を行わなかった。

活動状況ごとの周辺音量に対する正答率と主観評価のグラフを図 15 に示す。まず正答率に注目すると、実験したすべての活動状況においてにおいて一時停止ありの場合の正答率が高くなる傾向がみられ、有意水準 0.05 の Fisher の正確確率検定で判別すると、ジョギング時と PC 操作時において、有意差がみられた。つまり、周辺音量が大きくて導出された提示

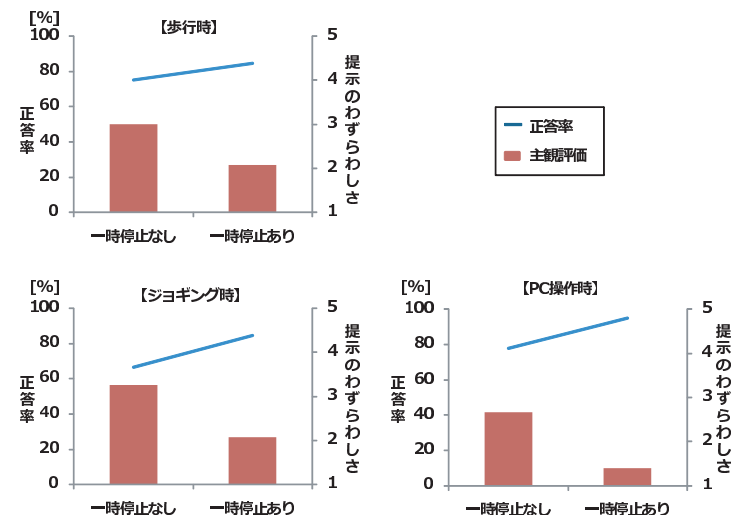


図 15 一時停止の評価実験結果

Fig. 15 The evaluation result on the effect of pausing.

音量が提示不可能であるときは、提示可能になるまで一時停止後に提示することでユーザに確実に音声情報を伝えることができるといえる。

次に主観評価については一時停止あり場合の提示の煩わしさはなしの場合と比較して小さくなる傾向がみられ、有意水準 0.05 の Wilcoxon の符号付き順位検定で判別すると、ジョギング時と PC 操作時において、有意差がみられた。これは大きな提示音量で提示されることに被験者が煩わしさを感じるためだと思われる。つまり、導出された音量が大きすぎて音声情報の提示が不可能な場合には、提示可能になるまで一時停止することで、音声情報の認識率も上がり、提示の煩わしさも小さくなるという。

5.4 アプリケーションの評価

本節では、プロトタイプを用いた音声提示アプリケーションの有用性を主観評価する。具体的には、普段から Twitter を使用している 5 人の学生に Twitter 音声クライアントを使用させて感想を得た。ここで、提案手法を用いて提示方法を決定する場合に対して、-5 dB, -35 dB の一定の提示音量で提示した場合を比較した。提示手法の違いに関するコメントとしては、「提案手法で提示した場合は煩わしさもなく、聞き取りやすかった」「音量が大きい場合 (-5 dB) は音声は聞き取れたが、煩わしい」「提案手法で提示した場合と音量が小さ

い場合(−35 dB)にあまり差がないように思え、ともにちょうどよい音量だと思った」といったものが得られた。上記の意見から、提案手法を用いて提示することで、煩わしくなく、かつ聞き取れる音量で提示できたといえる。提案手法と音量が小さい場合(−35 dB)に差がみられないという意見は、使用した環境が静かであり、提案手法における提示音量が−35 dBに近かったためだと考えられる。次に、全体を通してのコメント(Twitterの更新情報などの一般的な情報を提示されること、音声合成の精度について)としては、「わざわざブラウザを開いて更新情報を確認するという手間が省けて楽だった」「更新情報を音声で提示されることはうれしい」「全提示手法を通して、作業に集中している場合は少し煩わしかった」「ユーザ名が聞き取りにくい」「URLなどを音声で提示されても分からない」といったものがあげられる。これらの意見から、音声で一般的な情報を提示することは有用性があると考えられる。これは、ブラウザを開くなどの能動的な行動に対し、音声で受動的に情報を得ることがウェアラブルコンピューティング環境に適しているためだと考えられる。一方、作業中に提示されることが煩わしいという意見もみられた。今回はTwitterの更新情報を提示したが、地震速報やスケジュール情報など緊急性やリアルタイム性を持つ情報は作業の中断となっても提示する必要があると考えられるため、情報の持つ性質と提示方法の関係を評価し、重要性の低い情報は作業中には提示しないといった処理を加える必要性が示唆された。また、ユーザ名やURLなどのアルファベットの読み上げが聞き取りにくいという意見は、今回日本語の音声合成を使用したため、アルファベットをそのまま読み上げた(「yes」を「ワイイエス」と読む)ためだと考えられる。これらはアルファベットと読み方を対応付ける辞書データなどで対応が可能である。ただし、URLなどを音声で提示しても理解できないため、提示する内容によっては音声での提示は不適切だと考えられる。

6. まとめ

本稿では、ウェアラブルコンピューティング環境において、活動状況や周辺音の音量に合わせて適切な音声情報提示を行う手法を提案した。本研究では予備実験により、ユーザの活動状態や周辺音の音量と音声情報の認識率の関係を明らかにし、提示情報の音量や出力タイミングを適切に変更して提示するシステムを実装した。評価実験から、提案手法を用いることで、精度良く快適に音声情報が閲覧できることが分かった。また、会話中であれば事前にアラート音を鳴らすことで認識率を上げることができ、周辺音量が大きすぎる場合は一時停止することで十分な認識精度が得られるなど提案手法の有効性が明らかになった。

今後の課題としては、多様な状況で実験を行うことで、音声情報の認識に影響を与える要

因を一般化する必要がある。音声提示中の音量変更を適切に行うことで、提示中の環境変化に対応する手法を提案する予定である。

謝辞 本研究の一部は、科学技術振興機構戦略的創造研究推進事業(さきがけ)および文部科学省科学研究費補助金基盤研究(A)(20240009)、特定領域研究(21013034)によるものである。ここに記して謝意を表す。

参考文献

- 1) 津村弘輔, 井上亮文, 加藤淳也, 住谷哲夫, 重野 寛, 岡田謙一: ウェアラブルコンピュータに適した最適な情報提示方法の評価, 情報処理学会研究報告(グルーブウェアとネットワークサービス), Vol.2005, No.49, pp.13–18 (2005).
- 2) 田中宏平, 寺田 努, 西尾章治郎: ウェアラブルコンピュータのためのユーザ状況を考慮した知覚影響度に基づく情報提示手法, 情報処理学会マルチメディア, 分散, 協調とモバイルシンポジウム論文集, Vol.2008, No.1, pp.1479–1486 (2008).
- 3) 田中宏平, 岸野泰恵, 宮前雅一, 寺田 努, 西尾章治郎: 光学式シースルー型HMDのための読みとりやすさを考慮した情報提示手法, 情報処理学会論文誌, Vol.48, No.4, pp.1847–1858 (2007).
- 4) 沖野将司, 寺田 努, 塚本昌彦, 義久智樹: ユーザの状況を考慮した触覚情報によるウェアラブル入出力支援システム, 情報処理学会研究報告(モバイルコンピューティングとユビキタス通信研究会研究報告), Vol.2008, No.18, pp.17–24 (2008).
- 5) 塚田浩二, 安村通晃: Active Belt: 触覚情報を用いたベルト型ナビゲーション機構, 情報処理学会論文誌, Vol.44, No.11, pp.2649–2658 (2003).
- 6) Kern, N. and Schiele, B.: Context-Aware Notification for Wearable Computing, *Proc. 7th IEEE International Symposium on Wearable Computing (ISWC'03)*, pp.263–274 (2003).
- 7) Sawhney, N. and Schmandt, C.: Nomadic Radio: Speech and Audio Interaction for Contextual Messaging in Nomadic Environments, *ACM Trans. Computer-Human Interaction*, Vol.7, No.3, pp.353–383 (2000).
- 8) 庄司 武, 中村聡史, 塚本昌彦, 西尾章治郎: ウェアラブル環境における音声データのハイパーリンク型ナビゲーションシステム, 情報処理学会論文誌, Vol.44, No.5, pp.1380–1391 (2003).
- 9) Wilson, J., Walker, N.B., Lindsay, N., Cambias, C. and Dellaert, F.: SWAN: System for Wearable Audio Navigati on Wearable Computers, *Proc. 11th IEEE International Symposium on Wearable Computing (ISWC'07)*, pp.91–98 (2007).
- 10) 広田光一, 廣瀬通孝: ウェアラブルのための聴覚定位によるインタラクション手法, 情報処理学会論文誌, Vol.44, No.1, pp.156–165 (2003).
- 11) 在塚俊之, 畑岡信夫: 視覚障害者向け三次元音場利用情報提示システム, 電子情報通信学会誌, Vol.84, No.5, pp.333–336 (2001).

- 12) Roy, D.K. and Schmandt, C.: NewsComm: A Hand-Held interface for Interactive Access to Structured Audio, *Proc. SIGCHI Conference on Human Factors in Computing Systems (CHI'96)*, pp.173-180 (1996).
- 13) 藤原礼征: ウェアラブルとモバイルの大きな違い, チームつかもと会報誌「ウェアラブルでいこう」, Vol.1, p.1 (2004).
- 14) 塚本昌彦: ウェアラブルコンピューティングと GIS, 空間情報科学研究センター第 5 回シンポジウム, pp.1-19 (2002).
- 15) 沖電気工業株式会社. <http://www.oki.com/jp/profile/info/>
- 16) 平岩 明, 武藤伸洋, 三樹 聡, 奥平雅士, 柿崎隆夫: ウェアラブル PC による現場作業支援システムのコンセプトと常装着型ディスプレイの検討, ロボティクス・メカトロニクス講演会講演概要集, Vol.2000, p.98 (2000).
- 17) 柏野牧夫, 平原達也: 一度に何人の声を聞き分けられるか?, 日本音響学会研究発表会講演論文集, Vol.1996, No.1, pp.467-468 (1996).
- 18) Murao, K., Terada, T. and Nishio, S.: Toward Construction of Wearable Sensing Environments, Book Chapter, *Wireless Sensor Network Technologies for Information Explosion Era (Book Series: Studies in Computational Intelligence)*, Hara, T., Zadorozhny, V.I. and Buchmann, E. (Eds.), Vol.278, pp.207-230, Springer-Verlag (2010).
- 19) 佐藤 誠, 森田千絵, 土井美和子: 生体データと加速度データを用いた行動認識, 情報処理学会第 65 回全国大会論文集, Vol.5, pp.211-214 (2003).
- 20) Vapnik, V.: *The nature of statistical learning theory*, Springer (1995).
- 21) 音声認識システムの開発・研究のためのオープンソースの高性能な汎用大語彙連続音声認識エンジン julius. <http://julius.sourceforge.jp/>
- 22) twitter. <http://twitter.com/>

(平成 22 年 3 月 31 日受付)

(平成 22 年 9 月 17 日採録)



矢高 真一

2009 年神戸大学工学部電気電子工学科卒業。エンタテインメントコンピューティング・ウェアラブルコンピューティングに興味を持つ。



田中 宏平 (正会員)

2005 年大阪大学工学部卒業。2007 年同大学院情報科学研究科博士前期課程修了。2010 年同研究科博士後期課程修了。同年三菱電機株式会社先端技術総合研究所入社。博士 (情報科学)。在学時, ウェアラブルコンピューティング・ユビキタスコンピューティングの研究に従事。



寺田 努 (正会員)

1997 年大阪大学工学部情報システム工学科卒業。1999 年同大学院工学研究科博士前期課程修了。2000 年同大学院工学研究科博士後期課程退学。同年より大阪大学サイバーメディアセンター助手。2005 年より同講師。2007 年神戸大学大学院工学研究科准教授。現在に至る。2004 年より特定非営利活動法人ウェアラブルコンピュータ研究開発機構理事を兼務。博士 (工学)。アクティブデータベース, ウェアラブルコンピューティング, ユビキタスコンピューティングの研究に従事。IEEE 等 5 学会の会員。



塚本 昌彦 (正会員)

1987 年京都大学工学部数理工学科卒業。1989 年同大学院工学研究科修士課程修了。同年シャープ (株) 入社。1995 年大阪大学大学院工学研究科情報システム工学専攻講師, 1996 年同専攻助教授, 2002 年同大学院情報科学研究科マルチメディア工学専攻助教授, 2004 年神戸大学電気電子工学科教授となり, 現在に至る。2004 年より特定非営利活動法人ウェアラブルコンピュータ研究開発機構理事長を兼務。工学博士。ウェアラブルコンピューティングとユビキタスコンピューティングの研究に従事。ACM, IEEE 等 8 学会の会員。



西尾章治郎（フェロー）

1975 年京都大学工学部数理工学科卒業。1980 年同大学院工学研究科博士後期課程修了。工学博士。京都大学工学部助手，大阪大学基礎工学部および情報処理教育センター助教授，大阪大学大学院工学研究科情報システム工学専攻教授を経て，2002 年より大阪大学大学院情報科学研究科マルチメディア工学専攻教授となり，現在に至る。2000 年より大阪大学サイバーメディアセンター長，2003 年より大阪大学大学院情報科学研究科長，その後 2007 年より大阪大学理事・副学長に就任。この間，カナダ・ウォータールー大学，ビクトリア大学客員。データベース，マルチメディアシステムの研究に従事。現在，Data & Knowledge Engineering 等の論文誌編集委員。本会理事を歴任。本会論文賞を受賞。電子情報通信学会フェローを含め，ACM，IEEE 等 8 学会の各会員。
