

ホームネットワークにおける音声を利用した位置情報システム*

2K-06

東京電機大学 理工学部 情報システム工学科†

中村 祐介 梅島 慎吾 高橋 将 桧垣 博章‡ § 横山 智紀 花崎 泉¶ ||

1 背景と目的

一般家庭では、様々な家庭電化製品 (家電) が利用されており、その多くがマイコンなどのコンピュータによって制御されている。これまで独立に制御されていたこれらの家電をネットワークで相互接続することによって、遠隔での制御や管理を行ったり、協調動作によって効率的、統合的なサービスをユーザに提供することができる。このような、家電 (情報家電とよぶ) や情報機器を制御するコンピュータからなるネットワークはホームネットワークとよばれ、家庭の情報化を実現するための技術として注目されている。ホームネットワークにおいて、ユーザ要求を満足するように適切に情報家電を制御するためには、時々刻々変化するユーザの位置を獲得することが重要である。すなわち、家庭内においては、どのユーザがどの部屋に存在するかをリアルタイムに獲得する位置情報システムの実現が求められる。本論文では、ホームネットワークを対象として、音声 (話者認識) と無線 LAN プロトコルを用いた、位置情報の獲得率が高く、導入が容易である分散協調型の位置情報システムを提案する。

2 従来手法

ホームネットワークのサービスのようにより、ユーザの位置がサービスの提供に対して重要な役割を果たす場合のために、様々な位置情報獲得手法が提案されている。これらはユーザ検出の方法によって以下の 2 種類に大きく分類することができる。

- デバイス携帯型: ユーザが固有のデバイスを常時身に付けることによって、ユーザの位置情報を獲得する手法である。GPS や PHS を用いるシステム [2, 8]、Active Badge [5] などがある。
- 環境発見型: ユーザは固有のデバイスを携帯しない。音声、映像、体温などを手がかりとして、環境に設置したデバイス (センサ) がユーザを検出する手法である [1, 4, 7]。

家庭で生活するユーザに対して、固有のデバイスを携帯し続けることを要求することは困難であることから、ホームネットワークにおける位置情報システムでは、環境発見型のユーザ検出方法を用いることが望ましいと考えられる。

3 協調型話者認識による位置情報獲得

従来提案されているユーザ位置獲得手法は、複雑な装置を用いた高度な手法によって高いユーザ認識率や細粒度の位置情報の獲得を実現している。例えば [7] では、部屋の天井に設置した 3 台のカメラによって得られた動画画像を用いており、[4] では超音波、画像、音声の 3 つの情報からユーザ位置を特定している。しかし、一般家庭

においてこのような装置を導入することは、コストの面からも困難である。したがって、設置と利用に高度な技術を必要としない簡易な装置であることが求められる。装置が獲得したユーザの位置情報は、情報家電を統合的に制御するために、あるいは情報家電そのものが協調動作してサービスを提供するために利用されることから、ネットワークへの接続が不可欠である。IEEE1394 のようなケーブルフリーな状態が望まれる家庭内においては、無線通信の利用が適当である。本論文では、ユーザの音声情報を利用し、話者認識技術によってユーザ位置を獲得する手法を提案する。2 章で述べたように、環境発見型の位置情報システムでは、画像を用いる様々な方法 [1, 4, 7] が提案されている。しかし、この方法では、カメラの位置から見通すことができるユーザのみが検出可能となる (図 1)。これに対して、音声を用いた検出では、家具等によって見通すことができない場合にも正しく検出することが可能である (図 2)。

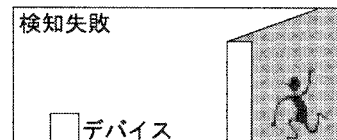


図 1: 画像によるユーザ認識

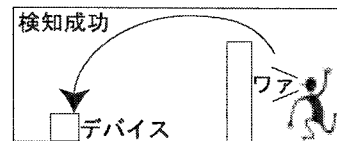


図 2: 音声によるユーザ認識

ただし、以下に挙げる問題点がある。

- (1) ユーザが常に音声を発するとは限らない。
- (2) 家庭内には様々な雑音 (ノイズ) がある。
- (3) 話者認識の成功確率は必ずしも高くない。
- (4) 音声を取得したデバイス (センサ) の位置とユーザの位置が必ずしも一致しない。

これらの問題を解決するために、提案手法では、建物内に複数のデバイスを設置する。設置の密度は、音声の到達距離と獲得するユーザ位置の粒度に依存するが、すべての場所を見通すために設置するカメラの数よりは少ないことが一般的に成り立つ。

(1) について、ユーザが音声を発さない場合、これを利用した話者認識は不可能である。しかし、人間は生活の中で様々な音 (生活音) を発生することから、これを利用することが可能である。例えば、ロックをせずに部屋に入ったことを足音などの情報から得たならば、部屋の所有者がそこにいと検出 (推定) できる。

また、音声を発さないためにユーザが特定できないが、生活音からそこにいずれかのユーザが存在すること

*Location Information System for Home-Networks

†Tokyo Denki University

‡Yusuke Nakamura, Shingo Umeshima, Susumu Takahashi and Hiroaki Higaki

§{yusuke, shin5, susumu, hig}@higlab.k.dendai.ac.jp

¶Tomoki Yokoyama and Izumi Hanasaki

||{yoko, hana}@k.dendai.ac.jp

が分かる。ホームネットワークは、このような匿名ユーザに対して、様々なサービス(照明を点灯するなど)を提供することが可能である。これは、(3)の問題についても同様であり、話者認識の成功確率が低いためにユーザの特定に失敗した場合でも、そこに匿名ユーザが存在するという情報を獲得してサービスを行なうことは可能である。また、建物内のレイアウト等の情報を利用することもできる。例えば、玄関で匿名ユーザを検出した場合には、来客をむかえるためのサービスを行なえばよい。一方、屋外との出入り口が窓のみである部屋で突然匿名ユーザが検出された場合は、侵入者である可能性を考慮したサービスを行なうことができる。

位置情報獲得デバイス(センサ)は、その情報に基づいて情報家電を制御、管理することから、通信能力を備えている。この通信をユーザ位置情報の獲得に積極的に活用することで、より高い認識率を得ることができる。例えば、部屋 R' に隣接する部屋は R のみであり、 R にはユーザ X のみが存在する。このとき、 R' で匿名ユーザが検出され、 R にはユーザが存在しない(音声と生活音が一定時間取得できない)ならば、 R' に存在するのは X であると推定できる。このように、単体では必ずしも高い認識率を実現できない、あるいは、音声ではなく、生活音のみしか取得できない場合でも、複数の分散配置された認識デバイスの協調によって、位置情報を獲得することができる。

また、(4)は音声を用いる場合特有の(映像を用いる場合は発生しない)問題であると言える。室内位置情報システムの場合、これは重大な問題となることがある。しかし、ホームネットワークで提供されるサービスの大部分は、ユーザ位置について部屋単位の粒度があれば十分であると考えられる。協調型話者認識の実現のために複数(少なくとも1部屋に1つ)のデバイスを装備することから、複数のデバイスで取得された音声によって認識されたユーザの位置は、その音声強度(大きさ)が最大と観測されたデバイスの周辺に存在すると考えられる。さらに、建物内の部屋や家具のレイアウト情報の組み合わせることによって、より正確な認識を可能とすることができる。

(2)に関して、ラジオやテレビの音声、情報家電が発する駆動音や音声ガイダンスなどをどのように扱うかが問題である。これらは、今後のフィールド実験等によって閾値を定めることで解決できると考えられる。

4 評価

図3のシステム構成に基づいて作成したプロトタイプを用いて、ユーザ位置情報の獲得実験を行なった。論文[3]で述べたように、位置情報獲得装置(デバイス)間の通信には無線LANを用いる。デバイスのプロトタイプには、マイクロホンを接続したパーソナルコンピュータを用いた。デバイス間の通信には、Bluetooth [6]を用いている。図4において、部屋 R_1 、 R_2 間を(1)、(2)、(3)の順に移動した場合の結果を表1に示す。

ホームネットワークアプリケーション	
協調型話者識別	SNMP
話者識別	TCP/IP
マイクロホン	マイクロセル無線LAN

図3: ソフトウェア構成

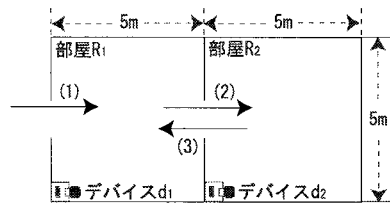


図4: シミュレーション環境

表1: シミュレーション結果(%)

	(1)	(2)	(3)
ユーザ認識	66.7	75.0	66.7
匿名ユーザ認識	20.8	4.2	20.8
誤認識	8.3	8.3	4.2
音声取得失敗	4.2	12.5	8.3

(2)において、(1)よりも認識率が高かったのは、 d_1 で得た情報を d_2 で利用したためである。

5 まとめと今後の課題

本論文では、ホームネットワークにおいて環境発見型の手法によってユーザ位置情報を獲得する簡易デバイスについて述べた。ここでは、複数のデバイスによる協調型話者認識を行なう。また、匿名ユーザとしての認識についてホームネットワークにおけるホームネットワークサービスにおける有効性について述べた。今後は、情報家電制御も含めたシステム全体についての検討、実験を行なう。

参考文献

- [1] Kusumoto, A., Nakazawa, J., Tobe, Y., and Tokuda, H., "A Location-Adaptive Virtual Networked Appliance," Proc. of the 21st IEEE ICDCS Workshops, pp. 214-219(2001).
- [2] 加藤誠巳, 酒井真哉, "インターネット上の3次元上智大学キャンパス案内システム 3D Walk Navi," 情報処理学会第54回全国大会, No.4, pp. 397-398 (1997).
- [3] 中村祐介, 梅島慎吾, 高橋将, 横山智紀, 花崎泉, 絵垣博章, "ホームネットワークにおける音声を利用した位置情報システム," 信学技報 (to appear).
- [4] 渡辺正規, 松尾直司, 野田拓也, 村瀬健太郎, 渡辺一宏, 長田茂美, 森田修三, "超音波, 画像, 音声を用いたマルチモーダルインターフェース," 情報処理学会第62回全国大会, 6J-07, CD-ROM (2001).
- [5] "The Active Badge System," <http://www.cam-ork.co.uk/ab.html>.
- [6] "The Official Bluetooth Wireless Info Site," <http://www.bluetooth.com>.
- [7] "Microsoft Easy Living," <http://research.microsoft.com/easyliving/> (2000).
- [8] "PHS 位置情報システム," <http://www.eva.hi-ho.ne.jp/mobile/phs.navi.html>.