

科学研究費助成事業 研究成果報告書

平成 28 年 5 月 19 日現在

機関番号：17102

研究種目：挑戦的萌芽研究

研究期間：2014～2015

課題番号：26630099

研究課題名(和文) 知能化空間とウェアラブルセンサによる第4人称センシング

研究課題名(英文) Fourth-person sensing using informationally structured environment and wearable sensors

研究代表者

倉爪 亮 (Kurazume, Ryo)

九州大学・システム情報科学研究科(研究院・教授)

研究者番号：70272672

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：本研究では、サービスロボットの普及に向けた新たな計測システムとして第4人称センシングを提案した。第4人称センシングとは、ロボットに搭載されたセンサ(2人称センサ)、環境に分散配置されたセンサ(3人称センサ)と、ウェアラブルセンサ(1人称センサ)を組み合わせた計測システムである。それぞれの利点を最大限に活用し、かつ欠点を補うことで、曖昧な事象に対しても高精度で信頼性の高い認識が実現できる。曖昧な物品取り寄せ指示に焦点を当て、1人称視点映像により認識したユーザ行動と環境固定センサで計測された物品情報を基に、ユーザが意図した物品を特定するシステムを構築し、意図の正確な理解に有効であることを示した。

研究成果の概要(英文)：In this research, we proposed a new concept of "fourth-person sensing" for service robots, which combines wearable cameras (the first-person viewpoint), sensors mounted on robots (the second-person viewpoint), and sensors embedded in the informationally structured environment (the third-person viewpoint) to recognize the surroundings of the service robot correctly and efficiently. Each sensor has its advantage and disadvantage, while the proposed concept can compensate the disadvantages by combining the advantages of all sensors. This technique is quite useful for accurate understanding of a user's intention and context of the scene. As one of applications of the proposed concept, we developed a HCI system which combines the first-person sensing and the third-person sensing, and showed the effectiveness of the proposed concepts through experiments.

研究分野：ロボット工学

キーワード：知能ロボティクス 環境知能化 ウェアラブルセンサ サービスロボット 一人称画像

1. 研究開始当初の背景

高齢化に伴い、介護現場における労働力不足が深刻化しており、人との共生を目指したサービスロボットの開発が進められている。一方で、サービスロボットが実際に生活支援サービスを計画・提供するためには、複雑に変動する生活空間の中で多くの環境情報を取得し、それらを実時間で処理する必要がある。そのため、センサの可搬能力や処理能力に限界のあるサービスロボットが、ロボット単体が全てを実行することは困難である。

この問題に対し、我々はこれまで、ヒト、モノの動きや環境の幾何構造など、サービスロボットの実現に必要な様々な情報を、レーザやカメラ、RFID タグ、圧電センサなど、周囲環境に埋め込まれたセンサにより取得し、ロボットからの要求に応じて伝達する空間知能化（環境情報構造化）の研究を推進してきた。これは、多くのセンサを搭載した All-in-one 型のロボットの実現を目指した従来型のロボット開発とは一線を隔すものであり、予め環境に様々なセンサを設置し、それをロボットの一部として取り込むことで、ロボットの可搬性能や処理性能の制約を受けず、仮想的にロボットのセンシング能力の拡大を実現する試みである。

この研究の一環として、我々は分散センサネットワークの構築に必要な環境情報構造化アーキテクチャ Town Management System (TMS)の開発を進めてきた。TMSでは、環境全体に分散配置したセンサにより空間内の人やロボット、物品の位置や状態といった情報を取得し、クラウド型データベースで統合管理する。サービスロボットは、作業を行う際にこれらの環境情報を利用することで、仮想的に拡大したセンシング能力を得ることができる。また、システムのみドルウェアに Robot Operating System (ROS)を導入し、ロボットやセンサ、機能の追加に柔軟なアーキテクチャ ROS-TMSとして開発を行っている。

従来の ROS-TMS で管理される環境情報を生活支援を受けるユーザの視点（1 人称）から整理すると、サービスロボットに搭載するセンサから得られる情報を 2 人称、環境全体に固定したセンサから得られる情報を 3 人称とすることができる。これら 2 人称・3 人称視点の情報は、環境全体を計測することができる反面、ユーザに近い環境に対しては、解像度や死角の存在などの問題が起きやすく、ユーザの指示や要求を信頼性高く認識することが困難な場合がある。

一方、近年、Google Glass などのウェアラブルカメラを用いた First Person Vision の研究が、コンピュータビジョンの分野で盛んになりつつある。これは当事者の見ている 1 人称視点の画像から、ライフログの作成や視野内への情報提示を行う試みである。これをロボットサービスに展開し、2 人称、3 人称センサから得られるモノの位置、種類などの

知識と 1 人称情報を組み合わせることで、サービス受益者の興味や要求をより早く、より正確に推定できる可能性がある。

2. 研究の目的

1 人称視点は、興味や要求などの内的要因（創発）に拠る身体反応の計測に適している。しかし計測範囲は限定的であり、局所的、断片的な情報になりがちである。また 2 人称視点は対象と環境を同時に計測できるが、ロボットに搭載可能なセンサ数や処理能力には制限がある。さらに 3 人称視点は、対象と環境、ロボットを俯瞰的に観測でき、指示対象候補の検出など外的要因（誘発）の把握に適している。しかし対象から離れた計測となり、精度、解像度、死角が問題となる。

そこで本研究では、ロボットおよび環境センサによる従来の 2 人称、3 人称情報と、ウェアラブルカメラからの 1 人称画像を同時に用い、3 者を相補的に組み合わせることで、サービス受益者の指示をより正確に理解するとともに、対象者自身の興味や行動をトリガにプロアクティブ（先読み）なロボットサービスを実現する「第 4 人称センシング」を提案する。また、第 4 人称センシングの適用例として、曖昧性を含むサービスロボットへの物品取り寄せ指示に焦点を当て、1 人称視点映像により認識したユーザ行動と TMS の 3 人称センサで計測された物品情報を基に、物品特定を行うシステムを構築する。さらに、構築したシステムを用いた実験を行い、第 4 人称センシングが曖昧な指示に対する正確な理解に有効であることを示す。

3. 研究の方法

(1) 第 4 人称センシングの概念

ここで述べる第 4 人称という言葉は、1 人称・2 人称・3 人称の 3 者の状態を客観的な立場から理解し、独自の解釈や分析を行う視点を指す。小説を例に挙げると、主人公を始めとした登場人物らが展開する世界を、物語として読み取る「読者」の視点に相当する。読者は、物語を読み進めていく中で、その世界とは完全に独立した視点から、通常では知り得ない主人公（1 人称）、相手（2 人称）、それを取り巻く人々（3 人称）の心の動きを把握し、独自の予測を立てることができる。第 4 人称による環境計測が目指す究極の目標は、3 つの人称視点を以って環境を分析することで、ユーザの心理状態からコンテキスト、環境の状態に至るまで包括的な空間の理解を行うことである。

一方で、各人称で得られる情報には、それぞれ長所と短所がある。

1 人称センサは、ウェアラブルカメラ装着者の行動を認識したり、細かな変化からユーザの意図や興味を推定することができるが、計測範囲が狭く、局所的・断片的な情報になりがちである。

2 人称センサは、サービスロボット自体が

生活空間内を移動できることから、環境に固定されるセンサに比べて計測の自由度が高く、実際にサービスを受ける人とその周囲環境を計測するのに適している。一方で、可搬能力や処理能力に制約を受けるため、多くのセンサを搭載することはできず、生活支援に十分な情報を得ることができない。

3人称センサは、対象・ロボット・環境を俯瞰的に計測することができるが、計測対象から離れた位置に固定されていたり、何らかの計測のみに特化した配置になっていることが多いため、死角や解像度といった問題が起きやすく、空間内の人の要求や指示を高精度に理解することは困難である。

一方、これら3者を相補的に組み合わせることで、サービスロボットへの指示に関連して次のことが期待できる。

1つ目に、より正確な指示理解である。システムに対するサービス要請の手段としては、音声が多く利用される。音声による指示はユーザから自発的に明示されるため、サービスのトリガとしては有用である。しかし、自然な音声指示の中で、ユーザの意図や要求が十分に表現される場合は少ない。一方、ウェアラブルカメラによって得られる1人称視点には、装着者が何を見ているか、何をしているかといった情報が含まれている。1人称視点映像の見えや動きの特徴を分析すれば、これらを行動情報や注視情報として抽出することができ、音声指示が曖昧な場合でも、指示内容を明確にできる可能性がある。

2つ目に予見的なサービスの開始である。1人称視点からは、2人称・3人称センサでは捉えることのできない細かな変化を計測することができる。これらには装着者の意図や興味といった心理的要因に依る動作も含まれており、直近で明示的な指示が行われる可能性が高い。それら特徴的な動作を検出した時点でサービスを開始すれば、従来の2人称・3人称によるシステムよりも早く生活支援を提供することができる。以下の節では、1人称、2人称、3人称の各種センサについて説明する。

(2) 1人称センシング

近年、高性能なウェアラブルカメラが手軽に入手できるようになった。なかでも、一般的に smart glasses と呼ばれるものの多くは、それ自身が Android OS を搭載しており、ウェアラブルカメラとしての側面だけでなく、可搬性の高いコンピュータとして幅広い用途に利用できる。また、マイクやスピーカー、ディスプレイといったユーザインターフェースを内蔵しているため、Human-Computer Interaction (HCI) を担うデバイスとして、TMS アーキテクチャに導入することもできる。本研究では、Epson 社の Moverio BT-200AV (図1) を利用した。



図1 1人称センサ Epson Moverio

(3) 2人称センシング

サービスロボットは、生活支援サービスを提供する側の立場にあり、自身に搭載するセンサから、生活支援対象の周囲環境を計測したり、作業に必要な環境情報を取得する。そのため、ロボットの視点から計測される情報を2人称情報と表現することができる。現在TMSで稼働するサービスロボット SmartPal V (図2) は、頭部に LRF と RGB-D センサ、胴体には同様の LRF とカメラを搭載している。

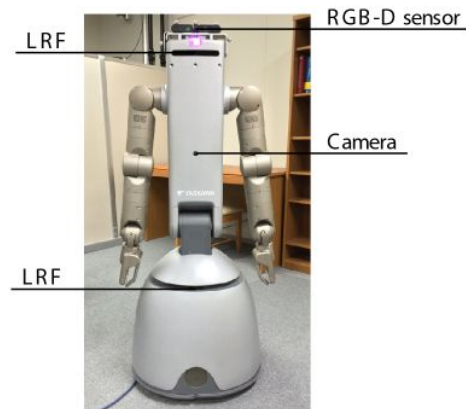


図2 2人称センサ SmartPal V

(4) 3人称センシング

我々は、環境全体に分散センサネットワークを構築する ROS-TMS の開発を行っており、生活空間で計測された環境情報はクラウド型データベースで管理される。分散センサとしては、LRF や RFID タグリーダー、Load cell、RGB-D センサなどが挙げられ、空間全体の物体位置や状態を計測している。本研究では、生活支援を受けるユーザの1人称情報、生活支援を提供するロボットの2人称情報に対して、環境全体の分散センサから取得する情報を3人称情報と呼ぶ。

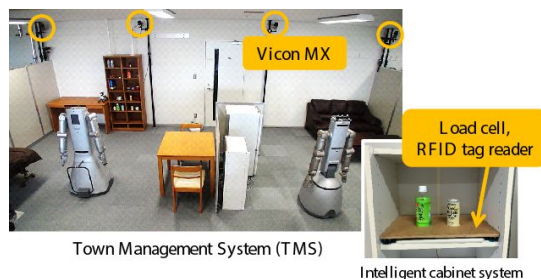


図3 3人称センサ ROS-TMS

4. 研究成果

平成 26 年度

曖昧性を含むサービスロボットへの物品取り寄せ指示に焦点を当て、ユーザの視線方向を利用した指示物品の提示手法を開発した。具体的には、1人称画像が計測できるウェアラブルカメラと、これまでに構築した3人称情報を取得する知能化空間を組み合わせ、知能化空間に登録された候補物品と1人称画像内の物品画像の対応を決定するシステム、および対応関係が一意に定まらない場合には、装着者の視野内に3人称情報の候補物品を提示することで、指示の正確性を向上するシステムを構築した。については、シースルー型ウェアラブルカメラ (Moverio, Epson) から、装着者の見ている画像を取り込み、その画像と位置計測システム (Bonita, Vicon) により得られる装着者の視線方向、および空間データベース ROS-TMS に登録された候補物品から、装着者が意図する物品を決定する仕組みを開発した。またについては、シースルー型ウェアラブルカメラ (Moverio, Epson) 上に、3人称情報である知能化空間に登録した候補物品を提示し、装着者に信頼性高く選択させるシステムを開発した。

評価実験の結果、第4人称センシングがあいまいな指示に対する正確な理解に有効であることが確認された。

平成 27 年度

第4人称センシングの適用例として、知能化空間内で想定される曖昧な物品取り寄せ指示に着目し、実際に音声と画像により適切な行動を選択、実行するシステムを構築した。例えば、ユーザが水の取り寄せを指示した場合、飲料水や園芸のための水など、水に関連した物品は様々な存在するため、従来のシステムが適切なものを判断することは難しい。一方、1人称視点からユーザ行動を認識することができれば、曖昧な指示を理解できる。例えば、ユーザが食事中であると分かれば、候補から飲料水を選択することができる。以上のシナリオを想定し、1人称視点の行動情報と3人称視点の物品情報を組み合わせた物品取り寄せシステムを構築した。システムは、ウェアラブルカメラと処理サーバにより構成される。ウェアラブルカメラでは、一人称視点画像と音声指示の取得及び送信を行う。処理サーバでは、受信画像列からユーザ行動を認識し、データベースの物品情報と併せてサービスロボットへ提示物品を指示する。実験では、ウェアラブルカメラを装着したユーザが行う5つの行動カテゴリ (読書, 食事, 植木を注視, ロボットを注視, 辺りを見回す) に対し、最初の3つの行動を行いながら同時に水の取り寄せを要求した (図4~6)。

図4は、読書している場合の様子である。1人称視点から得た動画像のカテゴリは、「読書」とであると識別された。また、水の取り寄せを指示したところ、サービスロボット

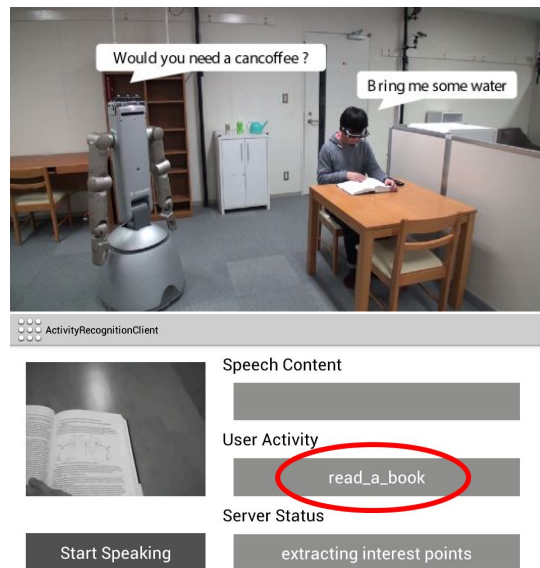


図4 読書の認識

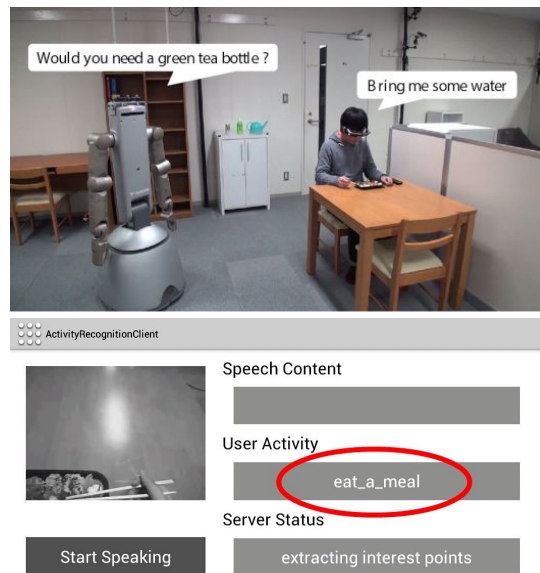


図5 食事の認識

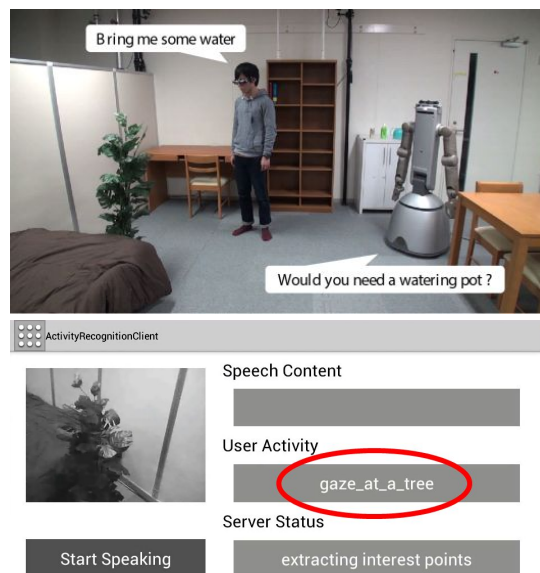


図6 植木の注視の認識

から「cancoffee」の提示を受けた。図5は、食事している場合の様子である。1人称視点から得た動画像のカテゴリは、「食事」であると識別された。また、水の取り寄せを指示したところ、サービスロボットから「green tea bottle」の提示を受けた。図6は、植木を注視している場合の様子である。1人称視点から得た動画像のカテゴリは、「植木を注視」であると識別された。また、水の取り寄せを指示したところ、「watering pot」の提示を受けた。これらの実験を通して、共通の曖昧な音声指示に対しても、その時点の行動情報によって適切な物品を特定できていることを確認した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計4件)

Yumi Iwashita, Kazuto Nakashima, Yoonseok Pyo, Ryo Kurazume, Fourth-person sensing for pro-active services, Fifth International Conference on Emerging Security Technologies (EST-2014), 査読有, 113-117, 2014

Yumi Iwashita, Asamichi Takamine, Ryo Kurazume, Michael S. Ryoo, First-Person Animal Activity Recognition from Egocentric Videos, 22nd International Conference on Pattern Recognition (ICPR 2014), 査読有, 4310-4315, 2014

Yumi Iwashita, Asamichi Takamine, Ryo Kurazume, Michael S. Ryoo, First-Person Activity Recognition from Animal Videos, 3rd Workshop on Egocentric (First-person) Vision in CVPR 2014, 査読有, 2014

Kazuto Nakashima, Yumi Iwashita, Yoonseok Pyo, Asamichi Takamine, Ryo Kurazume, Fourth-Person Sensing for a Service Robot, Proc. of IEEE International Conference on Sensors, 査読有, 1110-1113, 2015

[学会発表](計3件)

Asamichi Takamine, Yumi Iwashita, Ryo Kurazume, M. S. Ryoo, First-person activity recognition from DogCentric videos, The Tenth Joint Workshop on Machine Perception and Robotics (MPR14), K-P-06, 2014.10.17, 北京

中嶋 一斗, 岩下 友美, ピョ ユンソク, 高嶺 朝理, 倉爪 亮, サービスロボット

のための第4人称センシングの提案, 画像の認識理解シンポジウム (MIRU2015), SS4-11, 2015.7.30, 大阪

中嶋 一斗, 岩下 友美, ピョ ユンソク, 高嶺 朝理, 倉爪 亮, サービスロボットのための第4人称センシングの提案, 日本機械学会ロボティクスメカトロニクス講演会, 1A1-003, 2015.5.18, 京都

[その他]

ホームページ等

<http://robotics.ait.kyushu-u.ac.jp/~kurazume/research-j.php?content=sv#s2>

6. 研究組織

(1)研究代表者

倉爪 亮 (Ryo Kurazume)

九州大学・大学院システム情報科学研究
院・教授

研究者番号: 70272672

(2)研究分担者

岩下 友美 (Yumi Iwashita)

九州大学・大学院システム情報科学研究
院・准教授

研究者番号: 70467877