

生活支援ロボットのためのクラウドロボティクス

○山内悠嗣, 山下隆義, 藤吉弘亘 (中部大学)

1. はじめに

少子高齢化に伴い、労働者や介護者の不足が懸念されており、障がい者や高齢者の自立支援にロボットの活用が期待されている [1]. 人の生活の質 (QoL: Quality of Life) の向上を目的とした生活支援ロボットのタスクとして、我々は以下の2つの自立支援・見守りが重要であると考え、

1. 短時間における自立支援・見守り:
物(日用品)を拾い、渡す, 対話コミュニケーション等
2. 長時間における自立支援・見守り:
病気の早期発見, 安否確認, 服薬支援等

上記のタスクを実現するには、膨大な情報を実時間で処理することが要求され、ロボットに多くのリソースを持つ計算機を搭載する必要がある。これはロボットの大型化、高価格化に繋がるため、生活支援ロボットの実用化を妨げる要因となる。この問題を解決する方法として、2013年に Kuffner 等が提案したクラウドロボティクス [2] がある。クラウドロボティクスとは、ロボットとクラウド上の認識エンジンや情報処理技術を連携させ、大規模な計算やデータベースを要求するような情報処理をクラウドで担うという考え方である。特に、近年ではネットワークの高速化、ネットワーク機器が急激に増加し、センサネットワークや Machine to Machine(M2M), Internet of Things(IoT) 社会の到来も、クラウドロボティクスの発展に拍車を掛けている。また、音声認識・合成の分野では、クラウド型音声コミュニケーションツールキット rospeech [3] が公開され、容易にロボットに組み込むことが可能である。本稿では、我々が取り組んでいる生活支援ロボットのための人を観る、物を観るクラウドロボティクスとそのプロトタイプシステムについて述べる。

2. 生活支援ロボットのためのクラウドロボティクス

人とロボットが長期間にわたり共に暮らす共生社会では、タスクを完了するまでの時間という観点において、ロボットに要求されるタスクは短時間的な支援と長時間的な支援に分けられる。本章では、生活支援に必要なタスクと技術について整理し、提案するシステムの特徴について述べる。

2.1 生活支援に要求されるタスク

短時間的なタスクと長時間的なタスクを整理し、その目的と必要となる技術について述べる。

短時間的な自立支援・見守り

短時間的な自立支援・見守りのタスクとしては、物を拾い渡すような“物を対象としたタスク”、人とロボットの対話コミュニケーションのような“人を対象とし

たタスク”がある。“物を対象としたタスク”では、まず画像中の指定された対象物をロボットが認識する必要がある。次に、認識した物を拾うためには、物体周囲の空間の認識し、物体の姿勢の推定及び物を掴むための把持点を検出する必要がある。認識する対象物体を日用品に限定しても、膨大な種類の物体が認識の対象となる。

“人を対象としたタスク”では、例えば人とロボットの対話コミュニケーションにおいて、ロボットが人物の特定及び状態を理解することが重要である。コミュニケーションに必要不可欠な要素は、対話する相手が誰であるかを理解することであり、ロボットは人物を正確に照合する必要がある。また、初対面もしくは面識がある場合においても人の状態を理解することも重要となる。例えば、顔の表情から心理状態を推定し、その状態に合わせてロボットが振る舞うことで、人に対してストレスを与えないコミュニケーションを実現できる。このようなタスクを実現するためには、時間にわたり安定して認識する必要があるため、高精度な顔画像解析が要求される。

長時間的な自立支援・見守り

長時間的な自立支援・見守りのタスクとしては、病気の早期発見、安否確認、服薬支援等を対象とする。長時間にわたり人の行動解析及び経年変化を分析することで、健康の維持・増進に繋がる見守り技術を実現できる。特に、人の顔は状態を把握するうえで重要な情報発信源であるとされ、顔の部位の変化から臨床症状及び、疾患が随伴する可能性を推測することができる [4, 5]. これをロボットで実現するためには、顔の器官点を検出、検出した器官点を解析することで、顔画像からの健康状態の自動診断が期待できる。また、人と共生するロボットは長時間にわたり人を観察することが可能であるため、日常的に変化する顔画像やコミュニケーションの取り方を総合的に判断することで、早期の臨床症状の発見が期待できる。

上記の各タスクでは、ロボットが観る対象物により必要となる技術が大きく異なる。観る対象で必要となる技術を表1に整理する。これらの高度な画像処理技術は、多くの計算量とデータベースを必要とする。ロボット単体での情報処理、データベースを保持する場合には、ロボットの大型化、高電力化、高価格化の問題が発生する。これらの問題を一挙に解決するためにもクラウドロボティクスによる情報処理が適している。

2.2 本システムの特徴

我々は、2.1で挙げた観る技術を実現するクラウドロボティクスのシステムを開発している。現在、開発しているクラウドロボティクスのシステムの特長を下記に示す。

- リソース負担の軽減

表1 ロボットに要求されるタスクと必要技術の例

観る対象	必要な技術
物(日用品)	物体認識, 姿勢推定, 空間認識, 把持点検出
人(本人, 他者)	顔照合, 表情認識, 器官点検出, 年齢推定, 性別推定, 人種推定

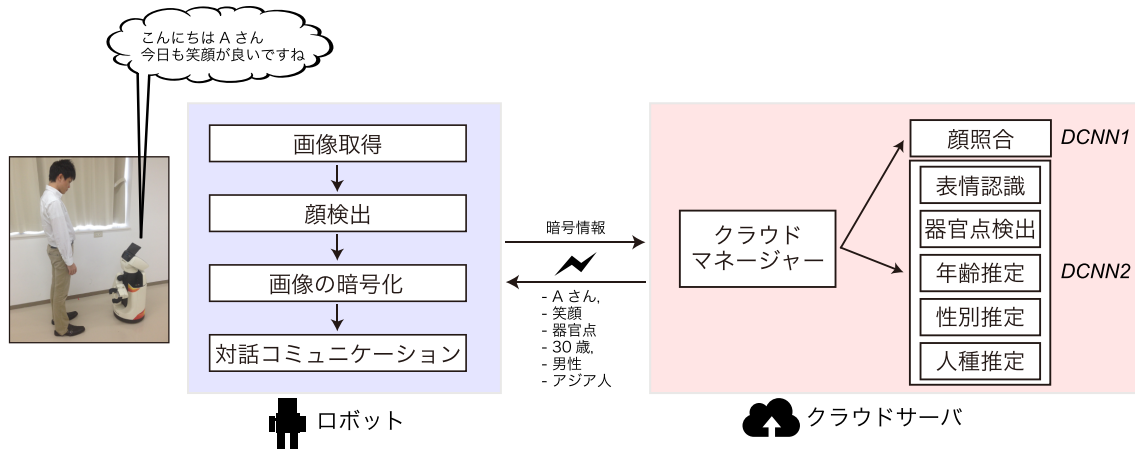


図1 クラウドロボティクスによる顔画像解析を用いた対話コミュニケーションの例

高度な画像認識は、一般的に多くの計算量が必要となるため、ロボット側の情報処理能力を占有する問題がある。また、画像認識には膨大なデータベースや学習したモデルを必要とするため、ロボット側で全ての情報を保持することが難しい。クラウドロボティクスを用いることでロボット側のリソース負担を軽減できる。

● 画像認識エンジンの高精度化

ロボットから送信される画像をクラウドサーバに蓄積し、蓄積した画像を用いて逐次的に学習することで、画像認識エンジンを高精度化することができる。また、様々なロボットから送信される画像を追加学習に用いることで、画像の質、撮影環境等の違いを吸収できる認識エンジンを学習できる。

● ROS への対応

Robot Operating System (ROS) に対応したクラウドロボティクスを構成する。これにより、ROSに対応したロボットであれば本システムを使用できる構成とする。

3. 観る技術を実現する認識エンジン

本章では、クラウドロボティクスによる観る技術について述べる。2章にて述べた生活支援ロボットに要求されるタスクは、人を対象にしたタスクと物を対象にしたタスクの2つに分けられる。本章では、人を観る技術と物を観る技術によるクラウドロボティクスの詳細について述べる。

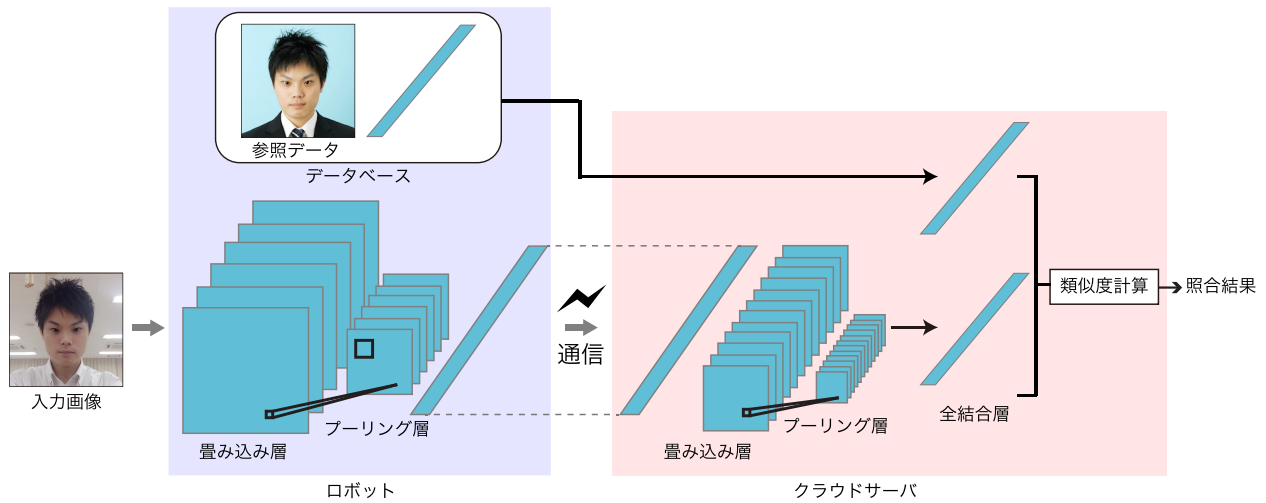
3.1 人を観る技術

図1に人を観る技術による対話コミュニケーションの例を示す。まず、ロボットが取得した画像から顔を検出し、後述で述べる方法により暗号化する。そして、ロボットからクラウドサーバへ暗号化した情報を送り、顔画像解析を行い、その結果をロボットに送信する。以

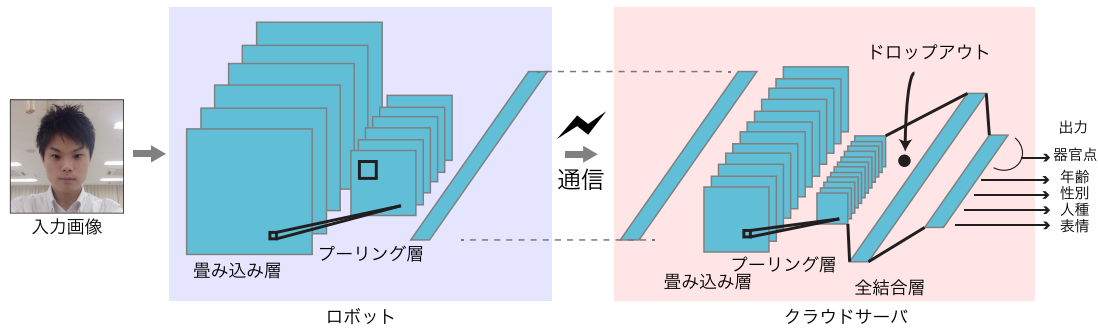
上の処理は、半リアルタイムで処理可能であるため、ロボットは顔画像解析の結果を用いて円滑な対話コミュニケーションを実現することが可能となる。

我々は特徴量を自動的に設計、学習する Deep Convolutional Neural Network(DCNN)[6, 7]による顔器官点検出、顔属性推定、顔照合の3つの顔画像解析を用いる。顔照合では、ロボットが取得した画像より検出した顔画像と、予めロボットが持つ顔画像が同一人物であるかを照合する。図2(a)にDCNNによる顔照合の流れを示す。ロボット側では、撮影した画像から顔画像を検出し、畳み込み及びプーリングを行う。そして、プーリング後の出力をクラウドサーバに送信する。顔画像照合及び属性認識を行うために、顔画像をクラウドサーバに直接送信することも考えられるが、通信時及びサーバ送信後に画像が流出する恐れがある。提案する方法では、プーリング後の出力を送るため、情報が暗号化されることになる。暗号化されたデータには個人を特定するような情報を含まないため、個人情報という観点において安心できるようなシステムとなっている。また、画像の暗号化にDCNNの処理の一部を利用することにより、計算効率の低下を防いでいる。

従来の顔属性認識の研究では、各属性(器官点、人種、年齢、性別、表情等)を認識するためにそれぞれ識別器を用意する必要があった。従って、推定する属性の数に比例して計算量が増加する問題を抱えていた。そこで、我々は1つのDCNNで複数の属性を推定する方法を提案している。図2(b)にDCNNによる属性推定の流れを示す。本手法では、1つのDCNNを用いるだけで効率良く複数の属性を同時に認識することができるため、半リアルタイムで属性を推定することが可能である。本システムを用いて属性を推定した結果を図3に示す。現在のプロトタイプでは、Robot Operating System (ROS)を基盤としたシステムに実装している。



(a) 照合



(b) 属性推定

図2 DCNNによる属性の推定と照合

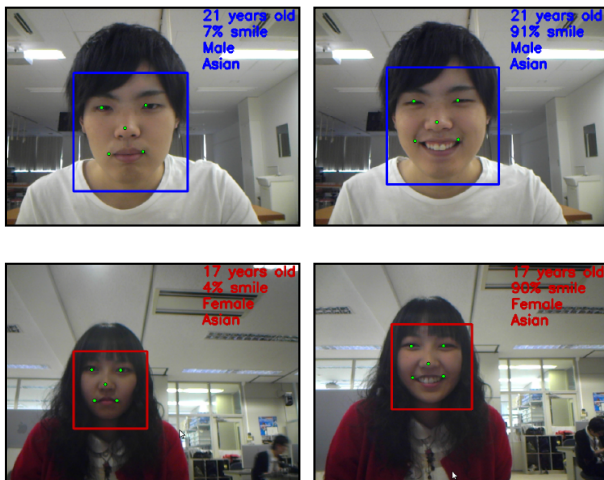


図3 顔画像解析の結果

3.2 物を観る技術

図4に物を観る技術を用いた把持の例を示す。本システムでは、まず人が渡して欲しい日用品をタブレット上で指定し、画像や指定した画像座標等の情報がロボットに送信される。送られた情報から物体名や姿勢の推定と把持可能な点を検出し、ロボットに送信する。ロボットは、送られた情報に基づき把持することで、物体に適した方法で把持できる。

物(日用品)を観る技術として、Weighted Hough

Forest[8]を用いる。Weighted Hough Forestは、複数の決定木を用いた機械学習手法の一種であるRandom Forestをベースとした物体検出法である。Weighted Hough Forestでは、局所パッチに対してクラス(検出対象、もしくは非検出対象)尤度を表す学習サンプルの重みを学習に導入している。決定木の階層が深くなるにつれ、検出対象物と非検出対象物において類似しない特徴的なサンプルの重みが大きくなる。これにより、検出時に背景や類似物体領域への投票値を小さくすることができるため、誤検出を低減することが可能である。図6にWeighted Hough Forestによる物体検出の例を示す。

Weighted Hough Forestを用いて物体を検出した際に、物体同士が密着するなどして対象物体の把持が不可能な場合がある。このような場合、対象物体の検出に加えて対象物体が把持可能であるか判定する必要がある。Weighted Hough Forestを用いて把持判定をするためには、把持可能な事例と把持不可能な事例を学習することで把持判定の基準を獲得する必要がある。しかし、大量の把持不可能な事例を収集するには大きな労働コストが必要となる。そこで、追加学習を用いてオンラインで把持判定の基準を獲得する手法を提案している。図5にWeighted Hough Forestによる物体把持判定とオンライン追加学習の概要を示す。Weighted Hough Forestは木構造を用いてサンプルを識別するため、追加学習の際に把持不可能サンプルが辿り着いた

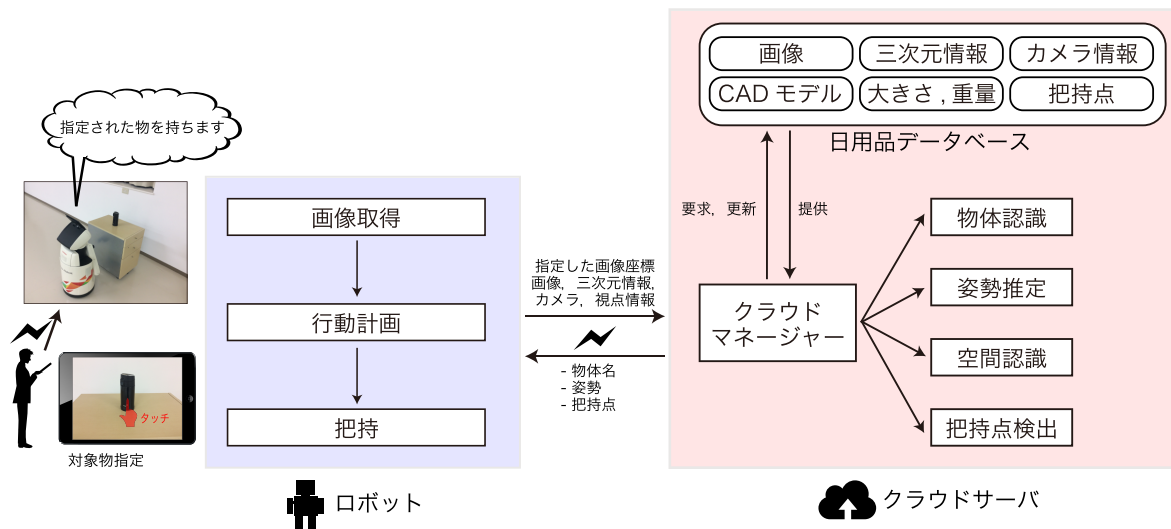


図4 クラウドロボティクスによる画像認識に基づく物体把持の例

末端ノードのみを更新することで、識別器の変更を最小限に抑えながらも把持不可能サンプルに適応できる。

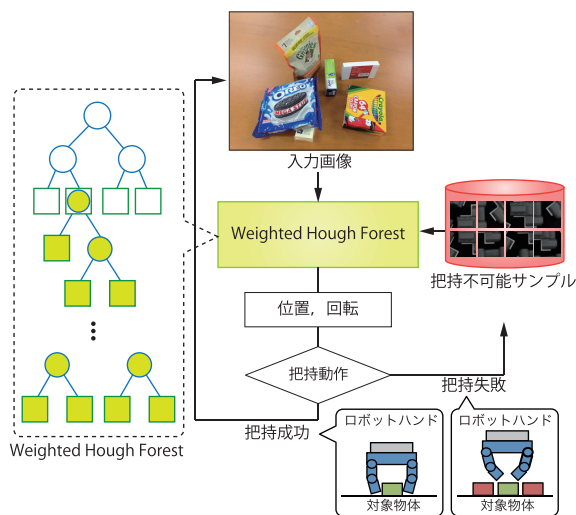


図5 Weighted Hough Forest による物体把持判定とオンライン追加学習の概要

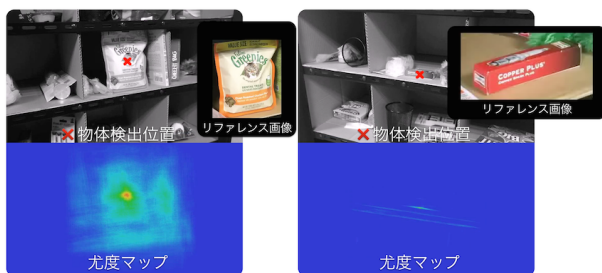


図6 Weighted Hough Forest による物体検出の例

4. おわりに

本稿では、生活支援ロボットのためのクラウドロボティクスとそのプロトタイプシステムについて述べた。生活支援ロボットに期待されているタスクと必要技術

について整理し、そのタスクを実現するための人を観る、物を観る技術について述べた。現在は各認識エンジンを構築し、ROSを基盤としたクラウドシステムに実装している。人を観る技術については既にクラウド上のサーバに実装し、人の状態に合わせた簡易的なコミュニケーションを実現した段階にある。

今後は、提案するクラウドロボティクスのシステムの開発を進めると共に、ROSに対応したロボットが本システムを容易に利用できるようなツールキットを開発する予定である。

参考文献

- [1] 山本貴史, 齋藤史論, 橋本国松, 池田幸一, “生活支援ロボット HSR の開発”, 日本ロボット学会学術講演会, 2012.
- [2] B. Kehoe, A. Matsukawa, S. Candido, J. Kuffner, and K. Goldberg, “Cloud-Based Robot Grasping with the Google Object Recognition Engine”, IEEE International Conference on Robotics and Automation, 2013.
- [3] K. Sugiura and K. Zettsu, “Rospeex, A Cloud Robotics Platform for Human-Robot Spoken Dialogues”, IEEE/RSJ International Conference on Intelligent Robots and Systems, 2015.
- [4] 北岡哲子, 工藤千秋, “Clinical Facial Expression Scale(C-Face) による早期認知症患者の表情特徴抽出 — FACS の併用の有用性 —”, 日本早期認知症学会論文誌, Vol. 6, No. 2, 2013.
- [5] 北岡哲子, 岸太一, “臨床的顔表情評価法 C-Face(Clinical Facial Expression Scale) の開発 — 認知症患者における検証 —” 日本早期認知症学会論文誌, Vol. 6, No. 1, pp. 84-89, 2013.
- [6] 木村真稔, 福井宏, 山下隆義, 山内悠嗣, 藤吉弘亘, “Convolutional Neural Network による顔の向きに頑健な顔器官点検出”, パターン認識・メディア理解研究会, 2015.
- [7] 山下隆義, 木村真稔, 福井宏, 山内悠嗣, 藤吉弘亘, “Deep Convolutional Neural Network による顔器官点検出における最適なミニバッチ作成方法”, 画像センシングシンポジウム, 2015.
- [8] Y. Murai, Y. Yamauchi, T. Yamashita, and H. Fujiyoshi, “Weighted Hough Forest for Object Detection”, IAPR International Conference on Machine Vision Applications, 2015.