

解説

日常生活支援ヒューマノイドの環境認識・行動制御

Environment Recognition and Behavior Control of Humanoid for Daily Tasks

岡田 慧* *東京大学大学院情報理工学系研究科

Kei Okada* *Graduate School of Information Science and Technology, University of Tokyo

1. はじめに

少子高齢社会の進行を背景として、日常生活支援を目的としたロボット研究が盛んになってきている。我々も等身大ヒューマノイド HRP2 を利用し、図 1 に示すように掃除や炊事を中心として様々な日常生活支援タスクに取り組んできている。

このようなロボット知能は、環境の空間的な広がりや三次元的にとらえ自身の立ち位置、自分の手先、その手が持っている物体、その物体の先が接触しようとしている環境と、それぞれの位置関係を把握しながら注意と行動を制御していく必要がある。そこでは我々は環境の三次元幾何情報を基盤として環境認識・行動制御に必要な知識表現を明らかにしながら、これらの知識表現を記憶にとらえたロボットシステムを構築してきている。また、次の段階としてこれらの記憶をどのようにしてロボット自身が獲得したらよいのかという問題に取り組みつづける。

本稿では、ヒューマノイドによる日常生活支援タスク実現のための記憶システム構成法と、これを用いた動作生成、環境認識、行動制御例、またロボット自身による記憶形成のアプローチを紹介する。

2. 日常生活支援ヒューマノイドの記憶システム

日常生活支援ヒューマノイドのための、認識行動共有記憶システムを図 2 に示す。この中の三次元 CG 画像が、実際のロボットの記憶系でありシステムの中核である。ここで各物体は形状だけではなく、様々な意味情報を有している点に特徴がある。すなわち、各物体は CG 用の形状・色を持つ単なるデータとしてソフトウェア実装されているのではなく、一つ一つがインスタンスとして扱われており、物体が有する情報（記憶）の問い合わせや操作が可能になっている [1]。

システム構成上の特徴として、認識行動系の密な統合を



図 1 等身大ヒューマノイドによる日常生活支援タスク例

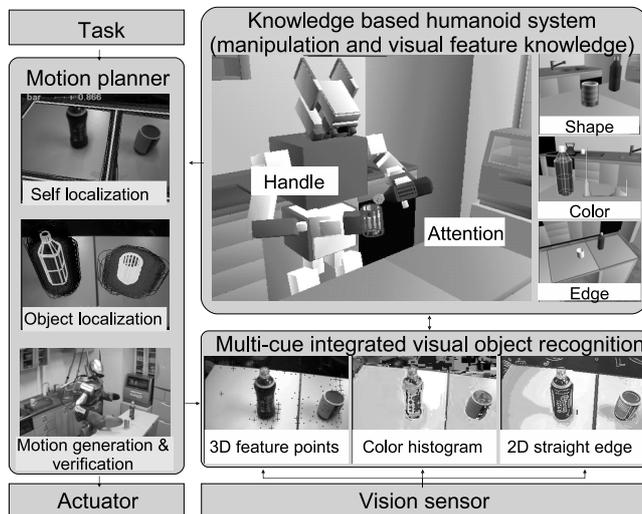


図 2 日常生活支援ヒューマノイドの行動認識記憶システム

目的とし、動作生成と視覚認識の双方のサブシステムがロボットの記憶系を共有している点にある。これにより効率的な動作生成と頑強な物体認識が実現され、これらを組み合わせることで動作生成時の環境認識処理や行動実行時の行為検証処理の簡潔な記述による行動制御を実現している [2]。

2.1 道具操作知識を用いた動作生成

道具操作に必要な知識として、以下の三つを導入した。タ

原稿受付 2008 年 2 月 14 日

キーワード: Humanoid, Recognition, Behavior Control

*〒113-8656 文京区本郷 7-3-1

*Bukyo-ku, Tokyo

スク遂行のための基準位置座標を示す **spot**, 物体を把持する際の位置と姿勢の拘束条件を示す **handle**, 物体を操作する際の基準座標並びに視覚検証のための注視点情報を示す **attention** である。

handle を導入することでリーチング動作の目的条件を決定でき, **attention** を導入することでツール座標系での動作記述によるロボットの全身行動の生成が可能になる. 一般にロボットの全身の自由度に比べツール座標系の自由度は少ないため, 効率のよい動作記述となる [3].

2.2 視覚特徴知識を用いた環境認識

視覚による環境認識のための視覚特徴知識として以下の三つを導入した. 三次元特徴点を用いた認識のための三次元表面形状 (**shape**), テクスチャ情報を用いた認識のための色情報 (**color**), 見かけエッジ比較を用いた認識のための三次元エッジ情報 (**edges**) である.

パーティクルフィルタを用いた複数視覚特徴の統合による物体認識を行う. 認識対象物体の三次元空間中での位置姿勢を表す確率分布をパーティクル (x) の集合で表し, 各パーティクルの尤度 ($p(x|Z)$) は複数 (M 個) の視覚特徴 ($z_{i=0...M}^i$) を用いて $p(x|Z) = \sum_{i=0}^M \prod_i p_i(x|z^i)$ とした.

視覚から得られる三次元特徴点と三次元表面形状 (**shape**) の類似度は, 三次元表面形状のうちロボットの視点から見える面の集合と三次元特徴点との最短距離の和を用いて計算する. 図3の左図に三次元表面形状のうちロボットの視点から見える面の集合を示し, 右図は認識結果をスーパーインポーズしたものである. 中図は三次元特徴点を点で表し, 直線の位置で各パーティクルの位置を, 直線の長さでそれぞれの尤度を表している. 色情報, エッジ情報を用いた計算の詳細は文献[4]を参照されたい.

2.3 認識動作共有知識ベースシステムを用いた行動制御

認識動作共有知識ベースシステムでは, 動作生成と視覚認識の双方のサブシステムによりロボットの記憶系が共有され認識行動系が密に統合されている構成に特徴がある. 本稿では, この知識が認識と行動に共有されることで実現可能な認識動作融合型の行動制御法を図4に示すお茶注ぎ, コップ洗い動作を例に示す.

2.3.1 自己位置同定と対象物体認識

各認識対象が環境に固定されているか否かの知識を導入し, 自己位置同定と操作対象物認識の両処理を行う. 自己位

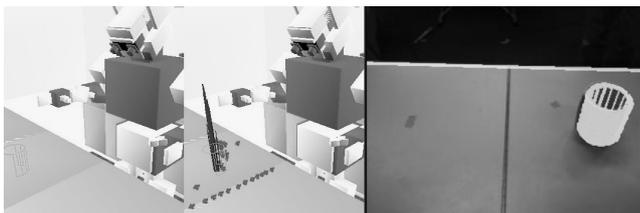


図3 三次元特徴点と三次元形状知識を用いたコップの視覚認識

置同定では環境に固定されている対象物体を認識しロボットの現在位置を更新する. 一方, 操作対象物は環境に固定されておらず, 認識結果を用いて記憶内の物体の現在位置の更新あるいは消滅を行う. 図5は, カウンターバー上のコップとペットボトルを認識したところである. コップの認識には **shape** 知識を用い, ペットボトルの認識は **shape** と **color** を用いた. ロボットはこれらをつかみ, お茶を注ぎ, 人が飲み終わったコップをつかんで後方のシンクの位置へと移動する. 移動先の座標は **spot** で示されている. 移動後は移動誤差が含まれるため, 図6のようにシンクを認識することでロボットの現在位置を同定し必要な修正量の移動動作を行い, ロボットを **spot** に誘導する.

2.3.2 動作知識を用いた視覚探索範囲の注視誘導

動作生成時の知識を利用することで, 一般には認識困難な対象でも認識すべき対象物体候補とその探索範囲を限定することで認識が可能になる. 例えば図7に示す例では, 視覚探索範囲を蛇口の存在可能範囲の下側と限定し, 三次元特徴点と円筒状の水流モデル (**shape**) との距離比較することで水流を認識した. このようにして, 認識した水流の位置にコップを持った手を誘導し, コップ洗い動作を実現している.

2.3.3 視覚検証埋め込み型動作生成処理

動作生成に必要な情報は, 対象物体とその操作基準座標



図4 日常タスク行動実験 (お茶注ぎ, コップ片付け)

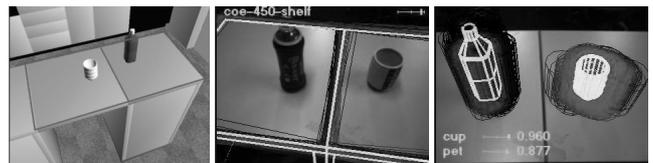


図5 視覚によるカウンター, コップとペットボトルの認識



図6 シンクの視覚認識と位置情報を用いた自己位置同定



図7 三次元特徴点と環境知識を用いた水流の認識

で表された動作系列である。一方、視覚検証を行う場合に必要必要な情報は対象物体の視覚特徴知識である。したがって、動作生成関数の中で対象物体の視覚特徴知識を参照した環境・物体認識関数を呼び出すことで、ユーザは動作生成と同様の引数を持つ関数で視覚検証埋め込み型の動作生成を実現できる。詳細は文献 [2] を参照されたい。

2.4 高次行動タスクプランナとの接続

これまで示してきたような環境のモデルを有し、把持、移動など抽象度の高い動作記述で実ロボットの認識行動を実現できるシステムでは、高次の行動タスクプランナと接続することで複雑な行動でも手順の推論を行いながら実現できるようになる。Fabienらは文献 [5] において料理タスクを例題とし、高次行動タスクプランナを用いて目的とする行動を自動的に分解し、不足している情報はユーザに問い合わせながらタスクを遂行する行動制御法を示している。

3. 認識行動記憶のオンサイト形成

日常生活環境で行動するロボットでは、環境や動作の記憶をあらかじめ与えるのではなく、その場で獲得していく手法が課題となる。本章では、認識行動記憶のオンサイト形成に関する研究例を紹介する。

3.1 身体直接教示に基づく感覚動作獲得

ロボットの身体を直接接触してその場で動作を教示・修正することで、新しいタスクや状況での行動獲得が可能になる。

小倉らは、人がロボットの手先を持って誘導することで行動を教示し再現可能なシステムを構築した [6]。図 8 では、ユーザがロボットの手を引っ張ることで移動行動と把持行動を誘導している。教示フェーズでは移動行動の目標座標情報を *spot* として記憶し、把持行動の把持位置を対象物体（ヤカン）に対する相対座標情報である *handle* として記憶する。動作再現時には教示フェーズで記憶した *spot* まで移動し、ヤカンを視覚で認識し、記憶した *handle* へリーチングし把持動作を行っている。それぞれ目標座標だけを記憶し、現在状態から目標状態までの動作系列の生成は動作プランナの機能を援用し実現している。

3.2 対象物の能動観察に基づく視覚認識記憶の獲得

ロボットの操作対象物もまたオンサイト獲得し、記憶に

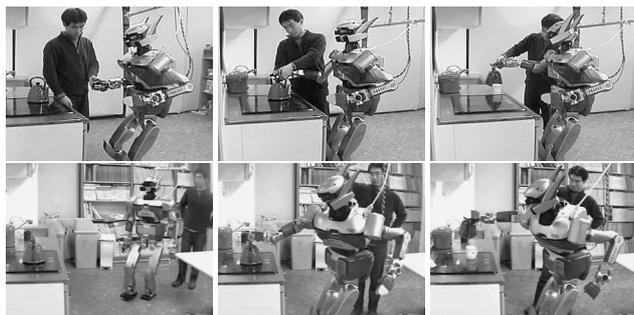


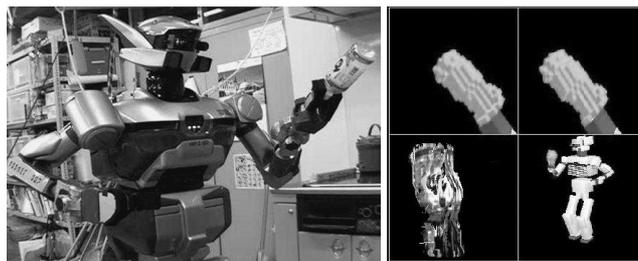
図 8 物体把持行動の教示（上：教示，下：再現）

追加できることが望ましい。小島らはヒューマノイドが人から手渡された物体を把持し（図 9 (a) 左）、これを 1 周させることで全周のテクスチャ情報（color）と形状情報（shape）を取得した（図 9 (a) 右） [7]。形状情報はシルエット法を用い、そのための背景差分処理を図 9 (b) に示す。ロボットが人から手渡された物体の視覚特徴を獲得することで、(1) 実際にロボットの視野から認識しやすい視覚特徴を判断しながら、取得すべき視覚特徴を選択できる、(2) ロボットのハンドに物体が拘束されているため、任意の視点から物体を観測できる、というメリットがある。

3.3 実演の観察に基づく感覚動作獲得

道具利用動作の獲得では道具の操作と、道具により操作される対象物の関係が重要になる。佐川らの研究ではロボットは人の実演を観察し、箒掃き動作において対象物（この場合はゴミ）が操作主体である人に対してどのように動かすかを認識し記憶した [8]。

図 10 はこの記憶に基づく感覚動作の生成例を示している。図 10 (a) 左でロボットはゴミを認識することで、図 10 (a) 中、図 10 (a) 右のようにロボット相対でどこにゴミが存在

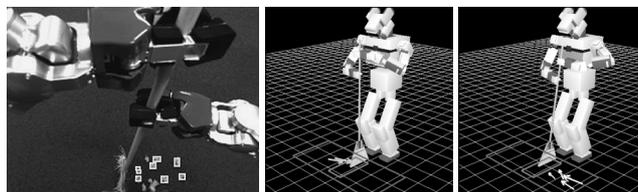


(a) 能動観察動作の様子と得られた三次元形状テクスチャモデル



(b) 背景差分処理による把持物体の発見

図 9 ロボット自身の操作に基づく視覚認識記憶の獲得



(a) ゴミの状態観察とロボットの行動生成



(b) 箒掃き動作の様子

図 10 実演の観察に基づく感覚動作獲得

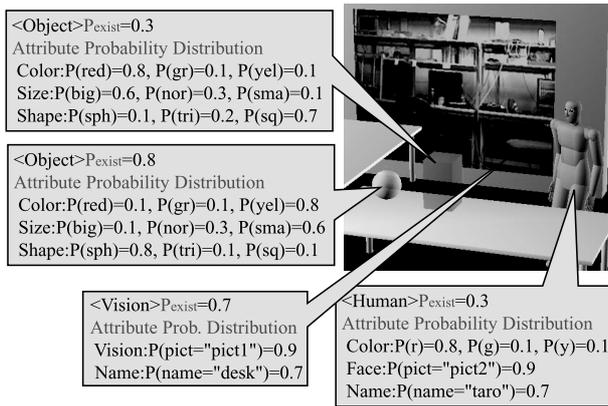


図 11 確率的空間記憶モデル

するかが分かる。本タスクにおいて、ロボットの取るべき動作はゴミの存在する場所に応じてゴミが中央に集まるよう記憶されており、図 10 (b) のようにゴミの左右の配置に応じて適切な動作を生成適用しゴミを中央に集める。

3.4 確率的空間記憶と対話に基づく知識生成

ダイナミックな環境でロボットが自身で記憶を形成していく過程では、センシングの不完全さから物体の存在は確率的に表現せざるを得ない。稲邑らは、環境中の知識に関する確率的な表現を導入した確率的空間記憶モデルを提案した [9]。ここでロボットは、状況と文脈に基づいて環境の確率的記憶を更新しながら人間からの指示を対話的に理解し、タスクの実現を行っている。

確率的空間記憶モデルは、(1) 存在確率、(2) 物体属性(色、形、質量など)の確率分布、(3) 三次元幾何情報を保持している (図 11 参照)。これらの情報は視覚認識の結果を反映し、また、記憶を減衰させる係数を用いて存在確率の確信度が時間の経過とともに減少するよう制御されている。

確率的空間記憶モデルを用いることで人から「それ取って」というような曖昧な依頼がなされた際に「それ」が示す物体を確率的推論モデルを用いて判断し、必要があれば適切な質問を生成することが可能になっている。

4. おわりに

人間と姿形が似ている等身大ヒューマノイドを見ていると、日々の雑多な作業を手伝ってほしいと思うのは筆者だけではないだろう。しかしながら、我々が普段行っている何事でもないタスクをロボットで実現しようとする、そこにある動作や認識がいかに複雑なものであるかを改めて確認する。ここではそのような複雑な問題に対し、三次元形状モデルと動作・認識知識という形でトップダウンに記憶を与えた認識行動システムの構成法と、より動的で未知の状況に対応するためのオンサイト記憶獲得法や確率的記憶表現に関する研究例を紹介した。後者に関しては個別のシステムのプロトタイプは存在するものの、それらが統合

され常に働いて必要なときに呼び出されるような構成法を明らかにしていくのが大きな課題となっている。

他の研究グループに目を向けると、産総研自律行動制御 G やカールスルーエ大でも記憶ベースの認識と動作計画を核に、それぞれ遠隔操作、模倣という指向性を持って日常生活支援タスクの実現に取り組んでいる [10] [11]。一方 MIT では、これらとは逆にボトムアップ・ビヘイビアベースドアプローチで取り組んでいる [12]。

参考文献

- [1] 稲葉雅幸, 岡田慧, 水内郁夫, 稲邑哲也: “Euslisp によるヒューマノイドロボットのシステム実現”, コンピューターソフトウェア, vol.23, no.2, pp.45–61, 2006.
- [2] K. Okada, M. Kojima, Y. Sagawa, T. Ichino, K. Sato and M. Inaba: “Vision based behavior verification system of humanoid robot for daily environment tasks,” Humanoids 2006, pp.7–12, 2006.
- [3] K. Okada, T. Ogura, A. Haneda, J. Fujimoto, F. Gravot and M. Inaba: “Humanoid Motion Generation System on HRP2-JSK for Daily Life Environment,” ICMA 2005, pp.1772–1777, 2005.
- [4] K. Okada, M. Kojima, S. Tokutsu, T. Maki, Y. Mori and M. Inaba: “Multi-cue 3D Object Recognition in Knowledge-based Vision-guided Humanoid Robot System,” IROS 2007, pp.1505–1506, 2007.
- [5] F. Gravot, A. Haneda, K. Okada and M. Inaba: “Cooking for humanoid robot, a task that needs symbolic and geometric reasonings,” ICRA 2006, pp.462–467, 2006.
- [6] 小倉崇, 羽根田淳, 岡田慧, 稲葉雅幸: “自律行動の誘導に基づく等身大ヒューマノイドの全身動作のオンライン教示システム”, 第 23 回日本ロボット学会学術講演会予稿集 CD-ROM, 1F25, 2005.
- [7] 小島光晴, 岡田慧, 稲邑哲也, 稲葉雅幸: “手渡された回転対称物体を観察して見かけモデルを作るヒューマノイド”, 日本機械学会ロボティクス・メカトロニクス講演会'06 講演論文集, pp.2A1–D27, 2006.
- [8] 佐川裕一, 岡田慧, 稲邑哲也, 稲葉雅幸: “ヒューマノイドにおける道具操作の認識と模倣行動の実現に関する研究”, 日本機械学会ロボティクス・メカトロニクス講演会'06 講演論文集, pp.2A1–D22, 2006.
- [9] 稲邑哲也, 園田朋之, 川路友博, 稲葉雅幸: “確率的空間記憶モデルに基づく人間ロボット間の協調的タスク遂行システム”, 第 20 回人工知能学会全国大会論文予稿集, pp.3F3–1, 2006.
- [10] E.S. Neo, T. Sakaguchi, K. Yokoi, Y. Kawai and K. Maruyama: “A Behavior Level Operation System for Humanoid Robots,” Humanoids 2006, pp.327–332, 2006.
- [11] T. Asfour, K. Regenstein, P. Azad, J. Schroder, A. Bierbaum, N. Vahrenkamp and R. Dillmann: “ARMAR-III: An Integrated Humanoid Platform for Sensory-Motor Control,” Humanoids 2006, pp.169–175, 2006.
- [12] A. Edsinger and C. Kemp: “Manipulation in Human Environments,” Humanoids 2006, pp.102–109, 2006.



岡田 慧 (Kei Okada)

1997 年京都大学工学部情報工学科卒業。2002 年東京大学大学院工学系研究科情報工学専攻博士課程修了。博士 (工学)。2002 年東京大学大学院情報理工学系研究科科学技術振興特任教員, 2006 年同特任講師となり現在に至る。ヒューマノイドシステム, 実時間三次元ビジョンの研究に従事。1999 年第 4 回ロボティクスシンポジウム優秀論文賞, 2006 年 IEEE-RAS International Conference on Humanoid Robots Best Paper Award。情報処理学会, IEEE の各会員。(日本ロボット学会正会員)