

認識行動共有知識ベースシステムにおける 複数視覚特徴統合による物体認識

岡田 慧* 小島 光晴* 稲葉 雅幸*

Object Recognition with Multi Visual Cue Integration for Shared Knowledge-based Action Recognition System

Kei Okada*, Mitsuharu Kojima* and Masayuki Inaba*

We have been tackled action recognition system design for a real world robot in a constructive approach. We have proposed the integrated system that shares knowledge-base between action and recognition. In this paper, we focus on a knowledge based object recognition system. Our system enables navigate attention to a visual search area, view prediction based recognition, robust recognition with multi-cue integration and time-series stochastic filter. Developed object recognition system is able to recognize both self location and object location by means of knowledge-base information, so that it is able to apply to humanoid tasks that contain both manipulation and locomotion. Finally, we showed a daily environment tasks consists of tea pouring, cup carrying and wash dishing in order to evaluate an availability of proposed shared knowledge-based action recognition system and object recognition method.

Key Words: Humanoid, Object Recognition, Multi-cue Integration, Knowledge-base

1. はじめに

ロボットによる人間の日常生活の支援は、少子高齢社会対応や次世代基幹産業など様々な側面からの期待もあり、世界中の研究機関や企業で研究が進められている。そこではヒューマノイドプラットフォームの研究 [1] が進み高信頼性と高安定性を有した高性能ロボットが幅広く利用できるようになりつつあり、ロボットが扱う物体や行動する環境を理解する認識処理研究や、行動と認識を結びつける知的制御研究など認識行動システムの研究が重要になりつつある。例えば ASIMO による知能化研究 [2]、欧州でのロボットの認識・判断・対話・学習研究プロジェクト [3]、独カールスルーエ大でのサービスヒューマノイド研究 [4] 等において、ヒューマノイドによる日常生活支援タスクの実現を目的とした認識・行動・知能システム研究がなされてきている。

これまで我々は日常生活環境での人間支援タスクの実現を目的としたヒューマノイドの認識行動システムの構成法として、タスクに関連する環境・物体知識を埋め込んだ認識行動共有知識ベースシステムを提案し実装してきた [5]。このシステムの特徴は、認識システムと行動システムの密な統合を利用して注視誘導・視覚検証処理を行動生成へ埋め込み、簡潔な記述で複雑

な認識行動を実現できる点である。そのために、物体の三次元形状モデルに意味情報を付加できるソフトウェア環境を利用し、視覚認識と動作生成の双方に共有される知識ベースを構築した。

本論文では本システムにおける物体認識サブシステムの構成法に関して述べたものである。具体的には知識ベースを利用することで、行動知識を用いた視覚探索範囲の注視誘導、見え方の予測に基づく高精度認識処理、複数視覚特徴の時系列確率フィルタリング統合による頑強な物体認識が実現可能であることを示す。さらに、提案する物体認識システムは自己位置同定と操作物体認識の双方に適用可能であり、移動と操作を含むヒューマノイドタスクの実現できる。また、提案するシステムならびに物体認識法の有用性を示すため、実ヒューマノイドを用いた日常生活でのお茶そそぎ、コップ運搬、水洗いを含む複雑タスクを実現した。

このように行動系と認識系で同じ表現を参照できるような統合システムを構築し、感覚により行為の検証を行う従来研究としてハンドアイシステムを対象とした研究例 [6-8] がある。本稿で対象としている日常生活支援ロボットでは、さまざまなタスク実現が求められ、簡潔な記述の導入作業で頑強な認識を伴うさまざまな行動が可能なシステム構成法が重要になる。本稿で示した物体認識システムでは視覚特徴知識はソフトウェア操作可能な三次元形状モデルに対して付与することで、複数視覚特徴による頑強な認識処理を容易に記述可能である。また、ヒューマノイドシステムでは必要となる移動行動と操作行動の双方が

原稿受付

*東京大学大学院情報理工学系研究科

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

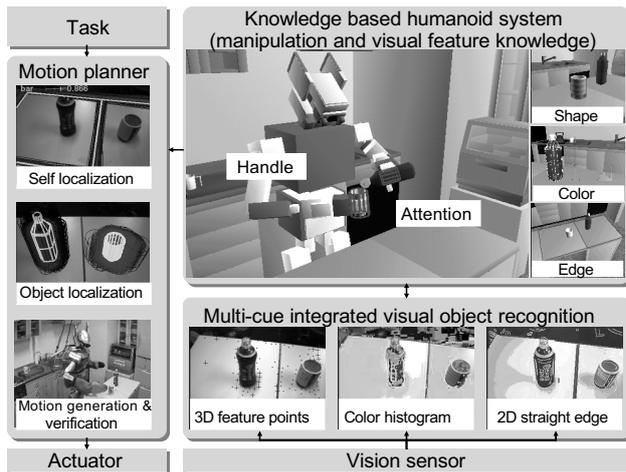


Fig. 1 Knowledge-based recognition system

可能な知識ベースを提案し、実際にこれらの知識を用いて移動と操作を含む日常生活支援タスクが実現できることを実証した。

2. 認識行動共有知識ベースシステム

Fig.1に本システムの概要を示す。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知識ベースである。ここで各物体は形状だけでなく様々な意味情報を有している点に特徴がある。システム内で各物体はCG用の形状・色を持つ単なるデータではなく、各々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[9]。知識ベースには環境・物体の形状情報に加え、タ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ならびに視覚検証のための注視点情報を示すattentionを行動生成知識として持つ[10]。また視覚特徴知識として三次元特徴点を用いた認識のための三次元表面形状を示すshape, 色ヒストグラム情報を示す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を示すedgesを持つ[11]。

本システムの特徴は、行動生成時に視覚検証のための注視誘導処理を埋め込むことができるため、認識行動の簡潔な記述が可能である点にある。これは認識システムと行動システムが同一の知識ベースを共有しており、相互の知識の参照が可能のために実現できている。視覚検証を行わない行動生成関数の引数は、対象物体への参照とロボットの動作系列であり、関数の中では対象物体が有する物体の操作知識を参照し動作を生成する。一方、視覚検証を含む動作生成関数の引数は視覚なしの動作生成関数と同じであるが、関数の中では対象物体が有する物体の視覚特徴知識もまた参照され、視覚認識による知識ベース内の物体位置更新に基づいた動作生成や、動作生成後の物体認識を通じた視覚検証を実現する。

3. 認識行動共有知識を用いた物体認識システム

物体認識システムではシステム内の物体に付与された知識を用いたモデルベース物体認識を行う。ここでの知識は行動生成と

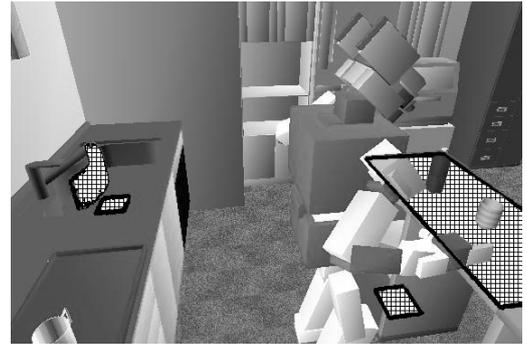


Fig. 2 Search area knowledge in the knowledge based system

共有されるため次のような特徴を有する。(3)(4)については次章で述べる。

- (1) 行動知識を用いた視覚探索範囲の注視誘導
- (2) 見え方の予測に基づく高精度の物体認識
- (3) 複数視覚特徴の時系列確率統合による頑強な物体認識法
- (4) 自己位置と操作物体の双方の認識可能なシステム

3.1 行動知識を用いた視覚探索範囲への注視誘導

視覚の役割はロボットの動作生成時の環境情報の確認や更新、あるいは動作終了時の行為検証である。視覚認識と動作生成の密な結合がなされている本システムでは、視覚認識時にそれを呼び出した動作とそれが参照している知識ベースを用いることで、視覚が注意を向けるべき対象を知ることができ、認識すべき対象物体候補とその探索範囲を限定した効率的な視覚探索が可能である。

Fig.2にシステム内に知識として記述されている視覚探索範囲を編目で示した。右側のカウンター上の矩形はカウンター上の物体を認識する際に利用する視覚探索範囲であり、カウンター下面およびキッチンシンク内の矩形はカウンターおよびキッチンシンクの認識に利用する。前者の探索範囲は床面上の二次元探索であり、後者はこれに鉛直軸周りの回転を追加した三次元探索である。キッチンの蛇口の下側にある探索範囲は水流と蛇口の角度を認識するためのものである。この領域は蛇口の幾何構造に拘束された一次元探索となっている。

これによりロボットは実際の視野内のすべての画像情報を処理する必要がなく、視覚探索範囲内に限定した注視処理を適用すればよい。これにより効率的な視覚認識処理と、例えば水流の検出のように一般的には認識困難な対象の認識が実現できる。

3.2 見え方の予測に基づく高精度認識処理

三次元物体の認識ではロボットの視野から見えない隠れ部分への対応が重要になる。本システムの認識システムは行動生成システムと結合しており、認識時に現在のロボットの身体情報を獲得可能である。そこでロボットの視点を仮想的に生成しシステムが有する物体情報からその時点でロボットから見えるであろう物体情報を予測し、この情報を用いた認識処理を行うことで隠れ部分の影響のない高精度の視覚認識を行う。

Fig.3上に見え方予測に基づかない場合(A)を、下に見え方予測に基づく場合(B)を示す。それぞれ左列の画像中のカウンター上に物体認識に利用するモデルを表示した。見え方予測に

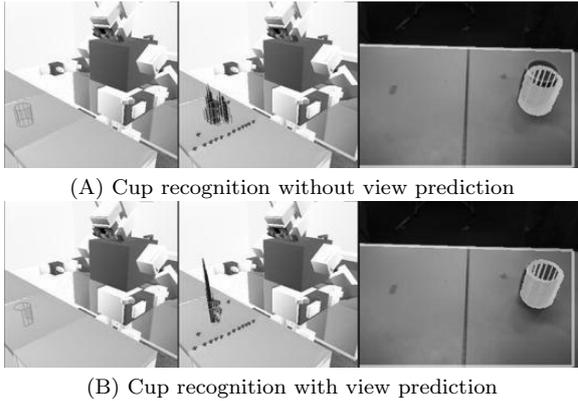


Fig. 3 Comparison of accuracy in object recognition with respect to view prediction

基づかない場合はすべての面が見えているが、見え方予測に基づく場合はロボットの視点からみえる面だけが表示されている。物体認識では三次元特徴点と認識モデルの面の距離計算（詳細は次章で説明）を評価関数として利用するが、見え方予測に基づく場合は基づかない場合に比べてロボットの視点からは見えない面も評価に含まれるため誤差が大きい。

また、中列のカウンター上の鉛直線で認識結果の確率分布を示している（直線の位置がパーティクルの位置に対応し、直線の長さが重みに対応する）。見え方予測に基づく場合(B)は認識結果が一点に集中しピークを形成しているが見え方予測に基づくかない場合(A)ではピークが分散していることがわかる。左列に視野画像に認識結果をスーパーインポーズしたところであり、見え方予測に基づくかない場合は1cm程度の誤差を確認できる。

本節では三次元特徴点を用いた物体認識を対象に説明したが、後述する色ヒストグラム、二次元エッジについてもロボットの視点からの見え方を予測生成し認識処理に利用している。

4. 複数視覚特徴の統合による三次元物体認識法

複数視覚特徴を利用することで頑強な物体認識が可能となる。本システムでは三次元特徴点、色ヒストグラム特徴、二次元エッジ特徴の三種類を用い、この視覚特徴に対応する知識として三次元表面形状、色ヒストグラム情報、三次元エッジ情報を知識ベース内に導入した。Fig.4の上列が実際のロボットの視野画像とそこでの視覚特徴を、下列に視覚特徴知識を示している。視覚特徴知識はこの図に示すような三次元形状モデルに対して直接付与されるため、視覚的に確認しながら視覚特徴知識を容易に導入することができる。

三次元特徴点にはペットボトルや模様のついたコップなどテクスチャが多い物体の認識に効果がある。エッジ特徴はテーブルやキッチン等の直線が多く含まれる人工物の認識に効果がある。知識として有する三次元エッジ情報は、ロボットの視点情報を用いて二次元エッジ情報に変換され、視覚処理で得られた二次元エッジ特徴と比較される。これは日常生活タスクではロボットが目の前のテーブルなどに正対し認識する状況が多くなり、2眼ステレオカメラで精度よい三次元エッジ特徴の検出が困難な場面が多くある点を考慮した。

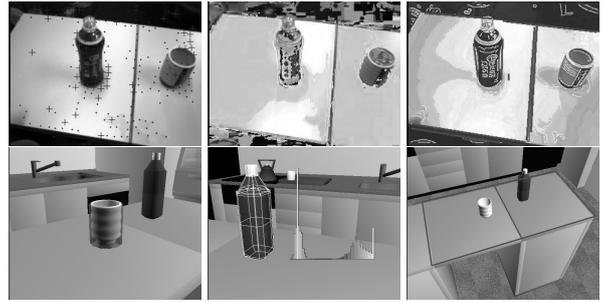


Fig. 4 Visual cues for object recognition (Left : 3D feature point, Middle : Color Histogram, Right : Bottom 2D edge segment)

4.1 パーティクルフィルタを用いた時系列物体認識

実世界で行動するロボットでは時系列画像に対して物体認識を行うことで過去の認識結果を考慮しつつ最新の視覚情報を用いた物体の認識を行うことができる。

本稿ではパーティクルフィルタ [12] を用いた時系列物体認識を行った。パーティクルフィルタは状態を示す確率分布を多数のパーティクル（粒子）の集合で表現した時系列フィルタである。

時刻 t における認識対象物体の三次元空間中での位置姿勢を \mathbf{x}_t （状態変数）、視覚センサによる観測を \mathbf{Z}_t （観測変数）としたとき、ある観測が得られた際の認識対象物体の位置姿勢の確率分布（事後確率） $p(\mathbf{x}_t|\mathbf{Z}_t)$ を状態量 \mathbf{x}_t の仮説群 $s_t^{(1)}, \dots, s_t^{(N)}$ と各仮説に対応する重み w_t^1, \dots, w_t^N を用いて以下のように離散的に近似表現する。

$$p(\mathbf{x}_t|\mathbf{Z}_t) \approx \sum_{i=1}^N s_t^{(i)} w_t^{(i)} \quad (1)$$

w_t は尤度 $p(\mathbf{z}_t|\mathbf{x}_t)$ を用いて以下のように求める [13].

$$w_t = w_{t-1} \frac{p(\mathbf{z}_t|\mathbf{x}_t) p(\mathbf{x}_t|\mathbf{x}_{t-1})}{q(\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{Z}_t)} \approx w_{t-1} p(\mathbf{z}_t|\mathbf{x}_t) \quad (2)$$

状態変数 \mathbf{x} の自由度は位置姿勢の6自由度となるが、実際の処理では3.1節で説明した視覚探索範囲に関する知識を用いて自由度を拘束し、効率と頑強性の向上を図っている。

また、確率分布 $p(\mathbf{x}_t|\mathbf{Z}_t)$ の各要素の分散と閾値より低く、尤度 $p(\mathbf{z}_t|\mathbf{x}_t)$ が閾値より大きい場合において、対象物体が存在し、かつ物体の位置姿勢が認識できたと判断した。

4.2 複数視覚特徴の統合による物体認識

視覚センサから複数の視覚特徴 ($z^i (i = 0..M)$) を抽出し、式(2)の各パーティクルの尤度を、各視覚特徴を用いて計算された尤度の重み和として以下の式のように計算する。

$$p(\mathbf{z}_t|\mathbf{x}_t) = \sum_{i=0}^M \varpi_i p_i(z_t^i|\mathbf{x}_t) \quad (3)$$

ここで $p(z^i|\mathbf{x})$ は各視覚特徴を用いた尤度であり、 ϖ_i は $\sum_{i=0}^M \varpi_i = 1$ となる各視覚特徴の重みを示す。

各視覚特徴を用いた尤度の積ではなく重み和としたのは視覚特徴の適応的な選択 [14] への展開を考えているためである。

4.3 視覚特徴知識を用いた尤度計算

4.3.1 三次元特徴点

三次元特徴点を用いた尤度計算を以下に示す。

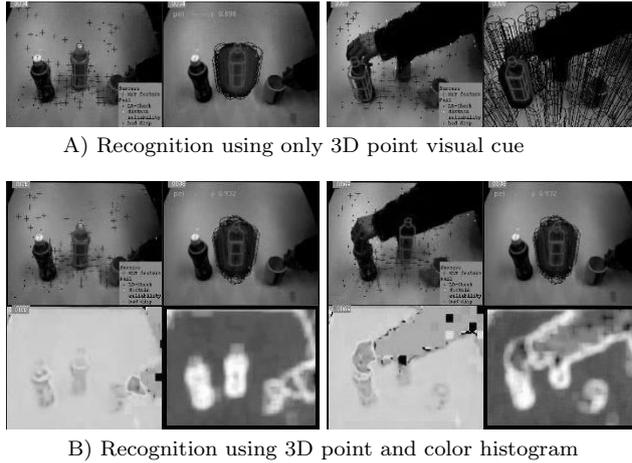


Fig. 5 Comparison of single visual cue object recognition and that of multi-cue integration

$$p_{point} \propto \exp[-\alpha_{point}(\frac{1}{|P|} \sum_{p \in P} D_{point}(p, F_{ref}^{visible}))^2] \quad (4)$$

$F_{ref}^{visible}$ は 3.2 節に示した物体モデルの面情報 F_{ref}^{all} のうちロボットの視点から見える面の集合であり、 P は視覚処理で得られる全ての三次元特徴点 P_{all} のうち最近傍の面 $F_{ref}^{visible}$ との距離が閾値以下のものである。 $D_{point}(p, F)$ は面と点の距離を示し、 $|P|$ は P に含まれる三次元特徴点 p の数である。

三次元特徴点は片目の画像から KLT 特徴点 [15] を計算し、この点に対応する点をもう一方の目の画像中から探索する。探索には SAD 評価を利用した。ステレオカメラは歪み補正、レクタフィケーションがなされているため探索はエピポーラライン上に拘束した。

4.3.2 色空間特徴

色空間特徴を用いた尤度計算は、視野上の各点における小領域の HSI 色ヒストグラム h_B と、あらかじめ物体に記憶させた色特徴情報 $h_{B,ref}$ を用いて以下のように計算した [16]。

$$p_{color} \propto \exp[-\alpha_{color}(B^2(h_B^H, h_{B,ref}^H) + B^2(h_B^S, h_{B,ref}^S) + B^2(h_B^V, h_{B,ref}^V))] \quad (5)$$

$B(h_1, h_2) = [1 - \sum_{b=1}^{Nb} \sqrt{h_1^b h_2^b}]^{1/2}$ は Bhattacharyya 距離 [17] である。3.2 節に示した三次元特徴点の場合と同様、色ヒストグラム処理においても見え方の予測を導入し、ロボットの視点から見える物体の形状を入力視野画像にマスクし色ヒストグラムの計算を適用する。

4.3.3 二次元エッジ特徴

二次元エッジ特徴を用いた尤度計算は以下である。

$$p_{edge} \propto \exp[-\alpha_{edge}(\frac{1}{|E^{2D}|} \sum D_{edge}(E^{2D}, E_{ref}^{2D}))^2] \quad (6)$$

入力画像に対しエッジ抽出処理を行い、エッジ点を追跡し直線性の評価により抽出した直線エッジセグメントを E_{all}^{2D} とした。 E_{ref}^{2D} は E_{ref}^{2D} と E_{ref}^{2D} との距離が閾値以下のものである。

E_{ref}^{2D} は 3.2 節に示す見え方の予測により、物体モデルが保持している三次元エッジ知識 E_{ref}^{3D} を現在のロボットの視点の位置姿勢情報に基づいて視野画面上に投影したものである。

距離の計算 $D_{edge}(E^{2D}, E_{ref}^{2D})$ は以下ようになる。まず E_{ref}^{2D} を固定長の長さのエッジ ($e_{ref}^{(1)} \dots e_{ref}^{(M)}$) として分割し、それぞれのエッジ ($e_{ref}^{(m)}$) に対して、 $E^{2D}(= e^{(l)})$ のなかからもっとも近いものを選択する。近さの評価は $value(e_{ref}^{(m)}, e^{(l)}) = d(2-a)$ とした。 $d = distance(e_1, e_2)$ は e_2 の両端点から e_1 との距離の和であり、 $a = angle(e_1, e_2)$ は 2 つのエッジのなす角度を示す。それぞれ閾値以内でない場合は 0 を返すようにした。

得られた最近傍エッジ対 $e_{ref}^{(m)}, e^{(min)}$ に対し、その評価値が非零の値を持つエッジの数を $|value|$ としたとき、エッジ対 $D_{edge}(E^{2D}, E_{ref}^{2D})$ の距離を $\frac{\sum value(e_{ref}^{(m)}, e^{(min)})}{|value|}$ として計算した。

4.4 複数視覚特徴統合による物体認識システムの評価

複数視覚特徴統合により頑強な物体認識が可能であることを確認した。Fig.5 は机の上に 2 つのペットボトルがあり人が手を伸ばし隠れが発生する状況である。左列は手を伸ばす前の画像、右列は手を伸ばした後の画像である。

A) は三次元特徴点のみを利用した場合である。左の画像は三次元特徴点を示し、右は位置姿勢を状態とする各仮説（パーティクル）にある物体の三次元エッジと結果がスーパーインポーズされている。人が手を伸ばす前は右のペットボトルを認識していたが人が手を伸ばすと隠れが生じパーティクルが収束しなくなり認識に失敗していることがわかる。

一方 B) は三次元特徴と色ヒストグラムの両視覚特徴を統合した場合である。左下画像は Hue 画像を示し右下画像は視覚特徴知識として持っているペットボトルの色ヒストグラムを用いて計算した各ピクセルの尤度を示している。これにより右上画像に示すように隠れが生じても右のペットボトルを認識し続けることができている。

4.5 知識ベース物体認識システムを用いた自己位置同定と操作対象物認識

各認識対象が環境に固定されているか否かを示す物体の可動性知識を用いることで、前節までに示してきた物体認識システムを用いて、自己位置同定と操作対象物認識の両処理を行うことができる。自己位置同定では環境に固定されている認識対象物体のロボットからの相対位置を認識することでシステム内のロボットの現在位置を更新し自己位置同定を行う。一方、環境に固定されていない物体はシステム内のオブジェクトの現在位置を更新、またはオブジェクトを生成消滅させることで操作対象物の認識とする。

4.5.1 自己位置同定

本システムではロボットの初期位置はあらかじめ与えており、歩行動作による移動誤差を視覚認識により計測し修正することで自己位置同定を行う。Fig.6, Fig.7 に自己位置同定の様子を示す。それぞれ左上がロボットが自己位置同定をしている様子を示し、右上は視覚処理で得られた二次元直線エッジセグメントである。左下は物体認識処理中の様子を示し、細線は認識処理における各仮説（パーティクル）に対応する物体の位置姿勢を示し、線の濃さが仮説の尤度に反比例している。右下で認識結果を太線で示した。

シンク、カウンターの位置は環境に固定されているという知

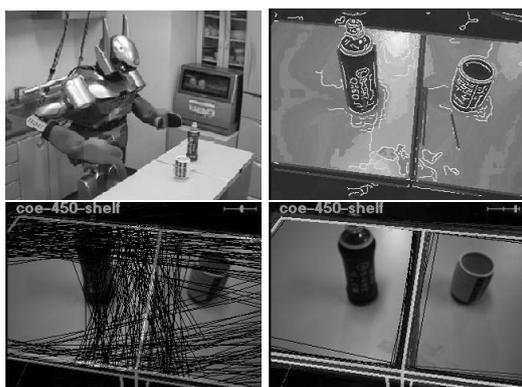


Fig. 6 Vision based self localization using counter knowledge.

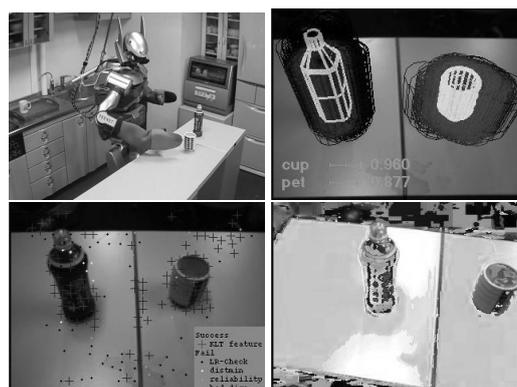


Fig. 8 Vision based cup and plastic bottle recognition



Fig. 7 Vision based self localization using sink knowledge.



Fig. 9 Water flow recognition using tap knowledge and 3D features points

識を利用し、物体認識結果からロボットの現在自己位置を同定し、この情報を用いてロボットの位置修正することで知識内の目標とする位置にロボットを誘導する。

4.5.2 操作対象物認識

Fig.8の左上画像はロボットがカウンター上のペットボトルとカップを認識している様子を示し、右上は認識結果を示す。細線が各仮説の位置姿勢にある三次元エッジを示し、尤度が高いものが薄く表示されている。太線で描かれた物体が認識結果である。下段はこの認識のために利用した視覚特徴を示している。左側は三次元視覚特徴点、右側は色ヒストグラムに利用したHue特徴である。

Fig.9では水流の位置を認識し蛇口の回転角度を推定している。環境知識には蛇口から出ている水流を三次元形状としてっており、水流の位置は蛇口の回転ジョイント構造の制約をうけている。探索パラメータは3.1節に示した視覚探索範囲知識を用いて蛇口の回転角度の自由度で水流の位置を拘束した。認識結果により知識ベース内の蛇口の回転角度を更新し、視覚により実世界の状態を反映させた情報に基づいて、ロボットの水洗い動作を生成する。また、水流の認識がなされなかった場合は蛇口から水が流れていないと判断できる。

4.6 物体認識システムの計算時間評価

物体認識処理においては100個のパーティクルを用い同一画像フレームに対してN回の処理を行った。また、4.1節で示した条件で物体が存在すると判断された状態がMフレーム連続

	M	N	Search Area
Counter (Fig.6)	2	5	200[mm]×200[mm]×40[deg]
Sink (Fig.7)	2	3	150[mm]×150[mm]×40[deg]
Cup (Fig.8)	5	5	915[mm]×450[mm]
Bottle (Fig.8)	5	3	915[mm]×450[mm]
Water (Fig.9)	2	5	120[deg]

Table 1 Parameters used for the object recognition

した場合に処理を中断し、その状態を知識ベースに反映させた。それぞれの繰り返し回数の具体的な数値をTable 1に示す。Fig.6では $N = 2, M = 5$, Fig.7では $N = 2, M = 3$, Fig.8のコップは $N = 5, M = 5$, ペットボトルは $N = 5, M = 3$, Fig.9では $N = 2, M = 5$ であった。数値は実験を通じて経験的に決定した。

また、現状のシステムにおける計算時間は100パーティクルの観測、重み計算、リサンプル、更新の手順を合わせて、三次元特徴点を用いた場合は7msec程度、色空間特徴を用いた場合は110msec程度、三次元エッジ特徴は370msec程度の時間の計算時間がかかっている。計算時間の差はそれぞれの尤度の計算処理の差に起因するが、最適化を施すことで大幅(100倍程度)な高速化が可能であることを確認している。

また、画像処理自体は三次元特徴点計算を用いた場合は45msec程度、色空間特徴を用いた場合は58msec程度、三次元エッジ特徴は90msec程度の計算時間がかかっている。計算は3.0GhzのXeonCPUを用いた。結果として2,3秒で物体の認識を行う

Behaviors	Visual controls	Object	Knowledge
Behaviors with self localization		Cup Bottle Counter Sink	Shape Histogram, Shape Edge Edge
Move to counter	Recog. counter		
Move to kitchen	Recog. sink		
Behaviors with object localization		Search area	Ttarget
Grasp a cup	Recog. cup	On counter Counter foot Sink foot	Cup, Bottle, Dish Counter Sink
Grasp a bottle	Recog. bottle		
Place a cup	Recog. cup		
Place a bottle	Recog. bottle		
Behaviors with visual verification		Event	Knowledge
Pour tea	Recog. tea	Recog. tea	Color histogram
Open tap	Recog. water	Recog. water	Water flow model
Close tap	Recog. water		
Wash cup	—		

Table 2 Knowledge description in the kitchen experiment.

システムとなっている。

5. 日常生活支援タスク行動例

開発した知識ベース認識行動システムを利用した日常生活支援タスク例としてキッチンでペットボトルからお茶をコップに注ぐ、また人が利用したコップを把持しキッチンシンクまで運び水洗いするというタスクを実現した。

5.1 等身大ヒューマノイドロボットシステム HRP2JSK

本実験で利用したロボットは等身大ヒューマノイドロボットプラットフォーム HRP2 [18] をベースに、手首の自由度追加 (7 自由度腕)、ヘッド部へのステレオカメラセンサの取り付けのハードウェア改造を施した。また、ソフトウェアシステムは OpenHRP 上に構築されたプラグイン機構を持った HRP2 ロボットシステム [19] の実時間ループ層とロボットシステム環境で記述された感覚・知能制御層 [9] から構成した。

5.2 日常支援タスクのための知識記述

このタスクを実現するのに必要となる知識を Table 2 に示した。ここでは右上の表 (Object) にあるように 4 種類の物体認識用の知識を記述した (Cup, Bottle, Counter, Sink)。また 3 種類の探索領域 (On counter, Counter foot, Sink foot) を右中の表 (Search area) に示した。右下の表は視覚を用いたイベント認識のための知識を示す。水流の認識は 4.5.1 節を用いて、またコップのなかのお茶の認識は文献 [20] で示した手法を用いた。

これらの知識を用いて 10 種類の行動を記述した。移動行動にはカウンターへの移動とキッチンへの移動の 2 種類がある (Move to counter, Move to kitchen)。それぞれカウンターとシンクの認識に基づいた自己位置同定を行う。

物体認識結果に基づいた行動にはコップの把持、設置、ペットボトルの把持、設置の 4 種類を記述した (Grasp a cup, Place a cup, Grasp a bottle, Place a bottle)。また視覚を用いた行為検証を行う行動としてお茶を注ぐ、水道の蛇口をあけるを記述した (Pour tea, Open tap, Close tap)。また認識を用いない行動としてコップ洗い行動を記述した (Wash cup)。

5.3 等身大ヒューマノイドによる日常生活支援行動実験

実験の様子を Fig.10 に示す。ロボットはまずコップの認識と把持 (1,2) を行い、次にペットボトルの認識と把持 (3,4) を行っている。水注ぎ動作 (5) を行いコップとペットボトルの設置 (6,7) を行う。

次に人がコップの中の飲料を飲み空いたコップをカウンターに置き (8,9)、コップの存在と位置を認識 (10) し、知識ベース内の物体情報を更新している。これを把持した後カウンター側に移動する (11,12)。移動後は歩行制御の誤差があるためキッチンシンク認識 (13) により自己位置を同定した。

これによりキッチンに対してロボットの位置を知ることができるとキッチン上の蛇口の位置情報がわかる。この情報をつかって蛇口をあけ (14)、水認識 (15) により蛇口の角度の認識と水が流れていることの確認をおこない、洗い動作 (16) を実行する。

5.4 認識行動共有ベースシステムの有用性評価

東大 COE 実世界情報プロジェクト [21] における最終デモンストレーションを含め、さまざまなメディア対応・研究室見学の場におけるヒューマノイドの生活支援タスク実験を通じ、本システムの評価を行った。撮影に伴う証明条件の変化などにも関わらず認識に失敗する割合は非常に稀であり頑強な物体認識システムを実現できているといえる。認識失敗の理由としては、正しい結果を得ているが閾値を超えない場合 (特に確率分布の分散の項) や、パーティクルがローカルミニマムに陥る場合等である。

また、新規タスクへの対応として Fig.11 に示すような 4 台のヒューマノイドを用いたサービスタスクに取り組んだ。そのために物体記述として皿と机、視覚探索領域として机の上と机の足元、動作として皿を把持・設置を追加する必要があったが、数人/日の作業で対応が可能となった。これにより本システムでは少量の記述の追加により異なるタスクの実現が可能であり、システムが十分な拡張性を有することを確認できた。

6. おわりに

本稿では日常生活支援ロボットのための認識行動共有知識ベースの物体認識システム構成法を示した。日常生活支援タスクの実現には、さまざまな環境と物体に対して簡潔な記述で頑強な認識行動が可能なシステム構成法が求められる。提案したシステムの特徴は、知識ベース内の三次元形状モデルに行動・認識に関わる知識を付与する構成法である。これにより簡単な作業で知識を導入することで、視覚探索範囲の注視誘導、見え方の予測に基づく高精度の物体認識、複数視覚特徴の時系列確率統合による頑強な物体認識を実現できる。また、開発した物体認識システムが移動と操作が可能なヒューマノイドでは重要となる自己位置同定と操作物体認識の双方の適用可能である。

また、本システムの評価を目的とし、ペットボトルとコップを用いたお茶注ぎ動作、コップを把持して歩行する把持移動動作、水流の認識に基づくコップの水洗い動作を含んだ等身大ヒューマノイドの日常生活支援行動タスクを実現し、デモンストレーション等の機会をとらえ検証してきた。その結果、物体認識に基づいた動作が頑強に実現できており、システムが有効に機能していることを確認した。

本稿で示すような知識ベースシステムに対する一般的な議論として、未知の物体・状態への対応法やシステムの拡張性と記述量が挙げられる。本稿で提案したシステムでは物体がタスク記述に含まれていれば、その物体が存在しなかったり配置されて



Fig. 10 daily life support experiments using knowledge based on recognition system.

いる場所が異なっている場合でも視覚処理によりその状況を判断し適切な対応をとることができる。一方、システムの記述量は5.2節に示すようにタスクの複雑さに比較して多くなく、また、Fig.11に示すような新規タスクへの対応も容易であり、十分な拡張性を有していると考えている。

今後の展開としては(1)その場で知識を獲得する方法 [22]、(2)モデル化困難な事象・対象への対応、(3)行動生成・物体認識の学習発達機構、が考えられる。既存の知識ベースシステムを高次の生得的な自律型認識行動記憶能力として捕らえた学習戦略採用することで認識・学習・行動の融合日常生活支援行動のような複雑なタスクを対象とした研究が可能になると考えている。

参考文献

- [1] H. Inoue, S. Tachi, K. Tanie, K. Yokoi, S. Hirai, H. Hirukawa, K. Hirai, S. Nakayama, K. Sawada, T. Nishiyama, O. Miki, T. Itoko, H. Inaba, and M. Sudo. HRP: Humanoid Robotics Project of MITI. In *Proceedings of the First IEEE-RAS International Conference on Humanoid Robots (Humanoids 2000)*, 2000.
- [2] Y. Sakagami, R. Watanabe, C. Aoyama, S. Matsunaga, N. Higaki, and K. Fujimura. The intelligent ASIMO: System overview and integration. In *Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'02)*, pp. 2478-2483, 2002.
- [3] Raja Chatila. The cognitive robot companion and the european 'beyond robotics initiative'. In *6th EAJ International*

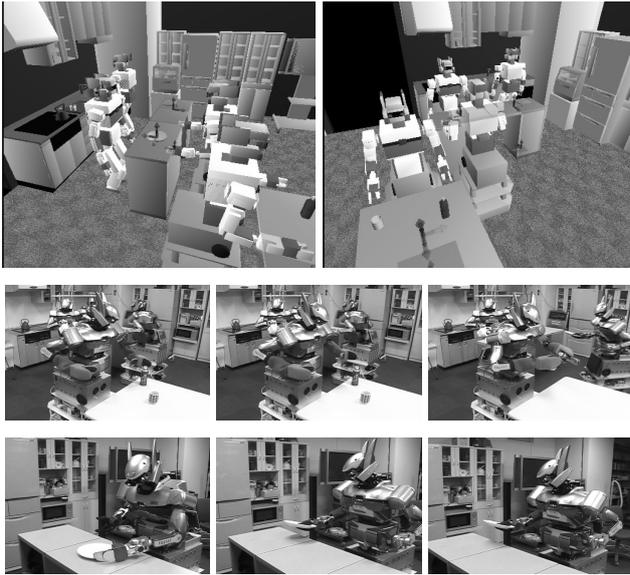


Fig. 11 Multi humanoid experiment in kitchen service task

- Symposium "Living with Robots"*, 2004.
- [4] 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. In *Proceedings in 2006 International Conference on Humanoid Robots (Humanoids 2006)*, 2006.
- [5] Kei Okada, Mitsuharu Kojima, Yuichi Sagawa, Toshiyuki Ichino, Kenji Sato, and Masayuki Inaba. Vision based behavior verification system of humanoid robot for daily environment tasks. In *2006 6th IEEE-RAS International Conference on Humanoid Robots (Humanoids 2006)*, pp. 7–12, 2006. 2006 6th IEEE-RAS International Conference on Humanoid Robots Best Paper Award.
- [6] R.C.Bolles. Verification Vision for Programmable Assembly. In *Proceedings of 5th International joint Conference on Artificial Intelligence*, pp. 579–575, 1977.
- [7] 白井良明. モデルベースト・ビジョン. 日本ロボット学会誌, Vol. 2, No. 6, pp. 593–598, 1984.
- [8] 長谷川勉. ロボット作業のためのモデルベーストビジョン. 日本ロボット学会誌, Vol. 10, No. 6, pp. 153–158, 1992.
- [9] 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.
- [10] Kei Okada, Takashi Ogura, Atsushi Haneda, Junya Fujimoto, Fabien Gravot, and Masayuki Inaba. Humanoid Motion Generation System on HRP2-JSK for Daily Life Environment. In *2005 IEEE International Conference on Mechatronics and Automation (ICMA05)*, pp. 1772–1777, 2005.
- [11] Kei Okada, Mitsuharu Kojima, Satoru Tokutsu, Toshiaki Maki, Yuto Mori, and Masayuki Inaba. Multi-cue 3D Object Recognition in Knowledge-based Vision-guided Humanoid Robot System. In *Proceedings of the 2007 IEEE/RSJ Intl. Conference on Intelligent Robots and Systems (IROS'07)*, pp. 1505–1506, 2007.
- [12] Genshiro Kitagawa. Monte Carlo filter and smoother for non-Gaussian nonlinear state space models. *Journal of Computational and Graphical Statistics*, Vol. 5, No. 1, pp. 1–25, March 1996.
- [13] Michael Isard and Andrew Blake. Condensation – conditional density propagation for visual tracking. *International Journal of Computer Vision*, Vol. 29, No. 1, pp. 5–28, 1998.
- [14] Jochen Triesch and Christoph von der Malsburg. Democratic

integration: Self-organized integration of adaptive cues. *Neural Computation*, Vol. 13, No. 9, pp. 2049–2074, 2001.

- [15] Jianbo Shi and Carlo Tomasi. Good features to track. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*, pp. 593–600, Seattle, June 1994.
- [16] Patrick Perez, Carine Hue, Jaco Vermaak, and Michel Gangnet. Color-based probabilistic tracking. In *ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part I*, pp. 661–675, London, UK, 2002. Springer-Verlag.
- [17] T. Kailath. The Divergence and Bhattacharyya Distance Measures in Signal Selection. *IEEE Transaction on Communication Technology*, Vol. 15, pp. 52–60, 1967.
- [18] H. Hirukawa, F. Kanehiro, K. Kaneko, S. Kajita, K. Fujiwara, Y. Kawai, F. Tomita, S. Hirai, K. Tanie, T. Isozumi, K. Akechi, T. Kawasaki, S. Ota, K. Yokoyama, H. Honda, Y. Fukase, J. Maeda, Y. Nakamura, S. Tachi, and H. Inoue. Humanoid Robotics Platforms developed in HRP. In *Proceedings of the 2003 IEEE-RAS International Conference on Humanoid Robots (Humanoids 2003)*, 2003.
- [19] 金広文男, 藤原清司, 原田研介, 金子健二, 梶田秀司, 横井一仁, 比留川博久. HRP-2 の運動制御システム –プラグイン群とその実行制御–. 第 21 回ロボット学会学術講演会予稿集, p. 3A33, 2003.
- [20] 小島光晴, 岡田慧, 稲葉雅幸. 日常生活支援ヒューマノイドによるお茶注ぎ行動における双腕協調と状態認識. 第 25 回ロボット学会学術講演会予稿集, p. 3L12, 2007.
- [21] 佐藤知正編著. 東京大学 21 世紀 COE 実世界情報プロジェクト監修, 『人と共存するコンピュータ・ロボット学実世界情報システム』. オーム社, 2004.
- [22] 小島光晴, 岡田慧, 稲葉雅幸. 能動的に物を手にとって視覚特徴を覚えるヒューマノイド行動の実現. 第 25 回ロボット学会学術講演会予稿集, p. 3H22, 2007.

岡田 慧 (Kei OKADA)

1974 年生. 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. ヒューマノイド統合システム, ヒューマノイドプランナ, 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(Humanoids2006) Best Paper Award. 日本ロボット学会, 情報処理学会, IEEE の各会員.} (日本ロボット学会正会員)

小島 光晴 (Mitsuharu KOJIMA)

1983 年生. 2006 年東京大学工学部機械情報工学科卒業. 同年東京大学大学院情報理工学系研究科知能機械情報学専攻入学. 現在に至る. ヒューマノイドロボットの視覚認識処理と環境物体モデリングの研究に従事. (日本ロボット学会学生会員)

稲葉 雅幸 (Masayuki INABA)

1958 年生. 1981 年東京大学工学部機械工学科卒業, 1986 年東京大学大学院工学系研究科情報工学専門課程博士課程修了. 工学博士. 1986 年東京大学講師, 1989 年助教授, 2000 年教授. 現在, 大学院情報理工学系研究科創造情報学専攻所属. 日本機械学会, 計測自動制御学会, 情報処理学会, 人工知能学会, ソフトウェア科学会等会員. (日本ロボット学会正会員)