

視覚による動作認識に基づく人とロボットの協調

Vision Based Cooperation between the Human and Robot

正 木村 浩 (電通大) 正 池内 克史 (東大)

Hiroshi KIMURA, Univ. of Electro-Comm, 1-5-1 Chofu-ga-oka, Chofu, Tokyo 182

Katsushi Ikeuchi, Univ. of Tokyo, 7-22-1, Minato-ku, Tokyo 106

In order to assist the human, a robot must autonomously recognize the human's motions in real time and must plan and execute the needed assistant motion based on the task purpose and the context. We defined the abstract Task-Model, analyzed the human demonstration by using Event and State Buffer, and automatically generated the Task-Models (Task-Knowledge) needed in the assistance by the robot. The robot planned and executed the assistant motions based on the Task-Knowledge in the cooperation with the human by analyzing the human motions. We implemented the 3D object recognition system by using the appearance based and minute templates based matching and the sub-pixel stereo vision. The effectiveness of these methods was tested through an experiment in which the human and the robotic hand assembled toy parts in cooperation.

Key Words: Motion Recognition, Human-Robot Cooperation, Human Demonstration, Task Model

1 はじめに

将来の育児や介護への応用を目指して、ロボットが人と協調して作業を行う、特に、人の作業の補助を行うことを考えると、各時点で、ロボットは人により何が行われているかを正確に理解し、それに合った補助を行う必要がある。人間の動作の観察・認識において、最も重要な役割を担うのは視覚である。ロボットは視覚を用いて、現在の作業対象物や人間の把持形態を認識し、作業目的・動作文脈に基づいて、必要な補助作業を計画・実行しなければならない。

本研究では、人とロボットによるおもちゃパーツの組み立て (Fig.1) を対象とし、上記の問題の解決を試みる。ここで、おもちゃパーツは、組み合わせるための面、穴、軸、ネジ部などの機能を持っている。おもちゃパーツ自体や機能部の幾何学情報を CAD データなどから与えることは簡単であるが、それらへのハンドの操作を幾何推論のみから自動生成することは簡単ではない。このような問題点を解決するために、人の動作の観察・解析からロボットの動作系列^{[1][2]}や把持方法^[3]を自動生成する手法が提案されている。また、人との協調作業を考えると、人のどのような動作に対してどのような補助動作を行うべきかを、幾何推論のみから自動生成することはやはり簡単ではない^[4]。そこで、本研究では、作業モデルの枠組を定義した上で、人が単独で行う動作の観察・解析から、ロボットによる補助動作に必要な作業モデル (作業知識) を自動生成することを考える。

さらに、人の動作の認識とロボットによる協調の制御のために、オクルージョンがある場合にも高速・安定に物体の 3 次元認識を行うことができる視覚処理システムを構築した。この手法についても、簡単に述べる。

2 視覚処理

近年、見え方モデルと固有値空間法を組み合わせた 3 次元物体認識手法が、特徴を抽出する必要がない、記憶量・計

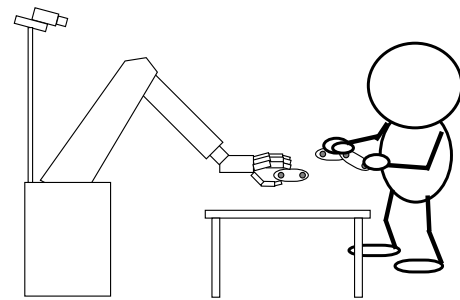


Fig.1 Cooperation between the human and robot

算量の点で優れているなどの理由から着目されている^[5]。特に、見え方モデルと微小テンプレートを用いた認識^[6]は、最初に物体のセグメンテーションを行う必要がない、オクルージョンに強いなどの利点がある。本稿では、^[6]を参考にした 2 次元画像上での 3 次元物体認識と、微小テンプレートのサブピクセル・ステレオ計測による 3 次元位置・姿勢測定を行った。ただし、固有値空間法を用いてはいない。

2.1 データ画像 対象物固有の座標系中心が回転台の回転中心上にあるように載せ、回転台を 10 度刻みで回転させながら 36 枚のステレオ画像を撮り、画像データとして保存する。

2.2 認識アルゴリズムの特徴 本稿で用いた 3 次元物体認識手法 (Fig.2 ~ 3) の特徴を以下に示す。

- (1) Trackability^[6]の高い微小領域についてのみ相関演算による Matching を行うので、高速・安定に結果を得ることができる。
- (2) データ画像と現在のシーン画像間で高い相関が得られた微小テンプレートを幾何学的拘束を用いてクラスタリングすることにより領域の Matching を行

う (Fig.3) ので、オクルージョンがある場合も部分の Matching が得られる。

- (3) 幾何学的拘束判定に幅を持たせることにより、奥行き変化による画像上での見かけの大きさ変化や、姿勢変化によるテンプレート間のゆがみにも、ある程度対応できる。
- (4) 相関演算に微小テンプレート間の Mean Distortion を用いているので、データ画像との照明条件の違い、陰影にある程度対応できる。
- (5) Trackability の高い微小領域についてのみステレオ計測を行うので、高速・安定に結果を得ることができる。

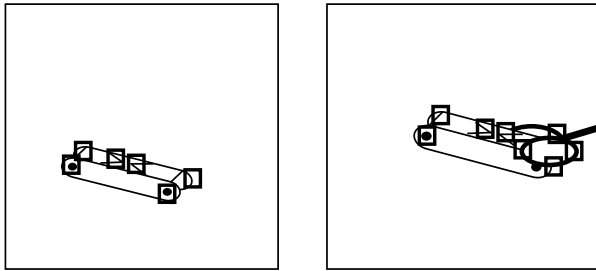


Fig.2 Minute templates on data image(left) and current scene(right) generated by considering trackability

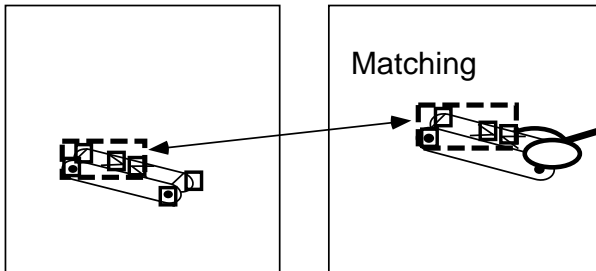


Fig.3 Templates cluster matched between data image(left) and current scene(right) by considering correlation and geometric constraints

この手法により、あるシーンの中でおもちゃパーツやドライバの 3 次元認識が可能となった。ただし、現在この手法には対象物の並進・回転について以下の短所がある。

- (a) 奥行き方向の大きな移動に対応できない
- (b) 2.1において、鉛直と直交する 2 軸の回転に対応するデータ画像を撮影すると全体として画像記憶量が膨大となる。これは、処理速度の低下をまねく。

2.3 人の手による把持の簡易認識 本研究では、人の手の形態の認識^[7]は実現されていないので、作業背景が一様であるとして、認識対象物まわりの高い Trackability を持つ連続した領域を人の手であるとし、これを直方体で近似した。対象物が人の手により把持されているかどうかは、3 次元空間での重なりと画像上でのオクルージョンの発生の有無で判断される。

3 人の動作の解析と作業知識の生成

3.1 作業モデルの枠組 本稿では、作業対象物について以下があらかじめ与えられているものとする。

- 回転台上で視覚系でキャリブレーションされた固有の座標系
- 固有の座標系内での穴、軸などの機能部 (Function) の座標系

ハンドによる機能部への操作 (Operation) を機能部座標系への操作と定義とした上で、「物体の機能部レベルの作業モデル」の枠組を以下の要素で定義する (Fig.4)。

- (1) 組となる各パーツと各機能部の組み合わせ (Combination of Functions)
- (2) 機能部への操作が行われる前提条件 (Precondition of Operations)
- (3) 機能部への操作 (Operations to Functions)

具体的には、(1) は軸合わせや面合わせのオス・メスの組、ネジとドライバなどの組み合わせを表す。(2) はネジ部を持つ軸の組み合わせのあとドライバによるネジ回し操作が行われることなどを表す。(3) は機能部座標の原点・軸の方向について組み合わせ時に一致すべき条件となる。

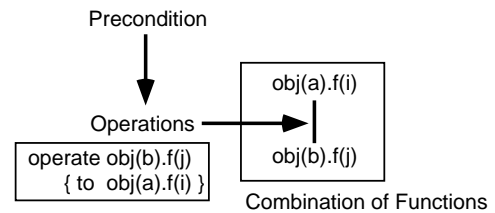


Fig.4 Abstract Task-Model: $\text{obj}(a).f(i)$ means the i -th function of object named $\#a$.

3.2 作業知識の自動生成 人の動作の各シーンに対して、2. で述べた視覚処理により、おもちゃパーツの 3 次元認識と把持の有無、パーツ機能部間の関係を検出することができる。ここでは、人によるパーツ把持の発生やパーツ機能部の組み付けの発生などを事象 (Event) と呼ぶ。

人が単独で行う動作の観察・解析から、「作業知識」を以下の方法で自動生成した (Fig.5)。

- 1) 現在のシーンを解析し、検出された Event を State Buffer と呼ばれるスタック型の作業領域に積む。
- 2) State Buffer の最上位にパーツやツール間の機能部の組み付け (機能部座標系原点・軸方向の一致) が検出されたとき、State Buffer に積まれている事象を取りだし下から順に Prediction, Operation1, Operation2 (Combination of Functions) とし (Fig.5-(b),(d)), Task-Model (Fig.6, 7) を生成する。

Fig.5-(b) においては、 $\#parts1$ と $\#parts2$ の組み付けが発生したので、二つの事象 ($\#parts1$ grasped と $\#parts2$ grasped) は $\#parts1-2$ grasped という状態と置換される。また、Fig.6, 7における Operation fix は、以下のような機能部座標の原点・軸の方向を一致させる命令に展開される。

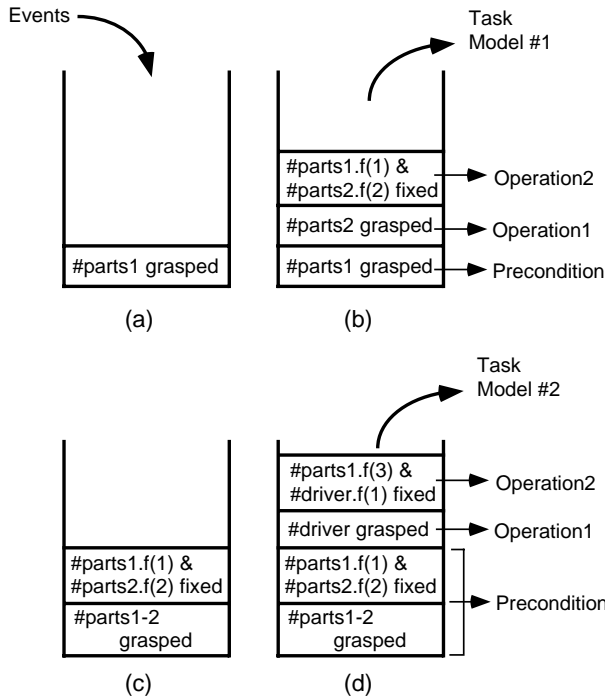


Fig.5 Events and state buffer

fix obj(a).f(i) with obj(b).f(j)
 coincide obj(a).f(i).origin with obj(b).f(j).origin
 align obj(a).f(i).axis(k) with obj(b).f(j).axis(l)

ただし、本稿では人の手の把持形態は認識されていないので、手のアプローチ・把持座標および把持点の自動抽出、さらに、それらのロボットハンドへのマッピング^[3]は今後の課題となっている。

4 人の動作の理解とロボットの協調

4.1 協調動作の計画・実行 人が片手または両手で行う動作に対して、3.2と同様のEventとState Bufferによる人の動作解析を行い、3.2で得られた「作業知識」に基づき、以下の方法でロボットハンドが協調動作を決定・実行した。

- 1) 現在のシーンを解析し検出されたEventを、State Bufferに積む。
- 2) システムに登録されている「作業モデル」の前提条件がState Bufferの最上位に現れたとき、Operationsが実行可能であれば、それらをロボットハンドが順次実行する。

ロボットハンドが行った操作により発生するEventは、視覚による確認(4.2)の上でState Bufferに積まれる。

4.2 協調作業の制御 ステレオ視覚系において、マニピュレータのキャリブレーションは比較的容易であるが、ハンドのそれは自由度の多い点や空気圧・ワイヤ駆動系の制御精度の点から容易ではない。従って、作業実行時に視覚を用いた補正が必要となる。ロボットハンドのような多指・多関節ロボットの動作認識は一般に難しい問

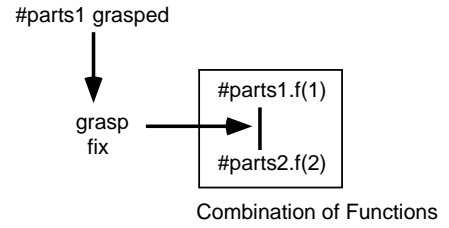


Fig.6 Task-Model #1 generated by analysis

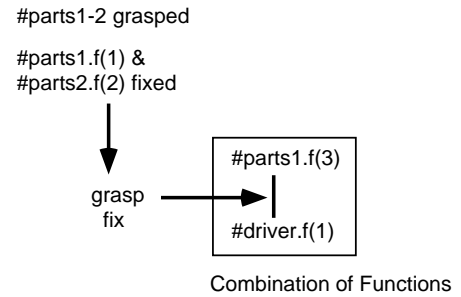


Fig.7 Task-Model #2 generated by analysis

題である^[8]が、本研究では自分自身が対象でありキャリブレーションされた視覚系のなかでだいたいどのような位置・姿勢にあるかは予測可能である。従って、ここでは、ロボットハンドの親指と人差し指を見つけるために2.で述べた手法を採用し、見え方モデルのデータ画像としてロボットハンドのOpenからCloseまでの変化について手首部を-50 ~ +50度の範囲で10度ごとに回転して撮影した画像を用い、各データ画像上で親指と人差し指の腹と指先の位置を登録した。協調作業時には、この対応するデータ画像を用いて親指と人差し指の指先を見つけ(Fig.8)、ステレオ計測することにより3次元位置を求め、必要な位置補正を行った。

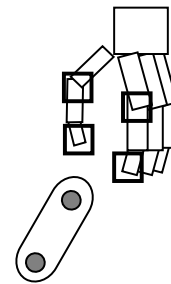


Fig.8 Recognition of robot fingers and their tips

また、組み合わせ協調時には、人の持つパーツだけでなく、ハンドで把持したパーツやツールの3次元位置・姿勢計測も連続的に行い、誤差を補正した。

4.3 実験 実験には、Puma560とUtah-MIT Handを用いた。ロボットの制御はChimeraと呼ばれるReal Time

OS を用いて行われ、画像処理とロボット動作計画は WS 上の SmallTalk-80 と C を用いて記述された。

人とロボットの協調組立ての基礎実験の様子を、以下に示す。Fig.9では、人が #parts1 を把持したとき、#parts2 との軸・半穴組み合わせ Task-Model#1(Fig.6) の前提条件が満たされ、Operations が実行可能であったので、ロボットハンドは #parts2 を把持し、#parts1 に組み付けた。Fig.10では、人が #parts1 と #parts2 をほぼ同時に把持したので、#parts1 と #driver 間ネジ・ドライバ組み合わせ Task-Model#2(Fig.7) の前提条件が満たされ、Operations が実行可能であったので、ロボットハンドは #driver を把持し、その先端を #parts1 のネジ部に挿入した。

いくつかの実験映像が WWW 上 (<http://www.kimura.is.uec.ac.jp>) で公開されている。

5 おわりに

本研究では、ロボットは視覚を用いて、現在の作業対象物や人間による把持の有無を認識し、作業目的・作業知識に基づいて、必要な補助作業を計画・実行した。具体的には、見え方モデル、微小テンプレートのクラスタリング、サブピクセル・ステレオ視覚による 3 次元物体認識システムを構築した。さらに、作業モデルの枠組を定義した上で、人が単独で行う動作に対して Event と State Buffer による解析を行い、ロボットによる補助動作に必要な作業モデル(作業知識)を自動生成した。そして、人とロボットの協調作業時に、同様な手法で人の動作解析を行い、ロボットは、作業知識に基づき協調動作を計画・実行した。以上を用い、人とロボットハンドによるおもちゃパーツの組み立て実験を行い、動作の認識に基づく人間・ロボット協調システムの可能性を示した。

なお、本研究は Carnegie Mellon University, School of Computer Science において行われた。研究にあたり御討論頂いた Task-Oriented Vision Lab. の皆様に感謝します。

参 考 文 献

- (1) 國吉, 井上, 稲葉, 人間が実演して見せる作業の実時間視覚認識とそのロボット教示への応用, 日本ロボット学会誌, Vol.9, No.3, pp.295-303, 1991
- (2) K.Ikeuchi, T.Suehiro, Towards an Assembly Plan from Observation Part I: Task Recognition With Polyhedral Objects, IEEE Trans. on RA, Vol.10, No.3, pp.368-385, 1994.
- (3) S.B.Kang, K.Ikeuchi, Grasp Recognition and Manipulative Motion Characterization from Human Hand Motion Sequences, Proc. of ICRA94, pp.1759-1764, 1994
- (4) 木村, 梶浦, 動作理解に基づく操縦ロボットと自律ロボットの協調作業, 日本ロボット学会誌, Vol.15, No.3, 1997 (掲載予定)
- (5) 村瀬, 古くて新しい画像認識法 - 固有空間法による画像認識 -, 情報処理, Vol.38, No.1, pp.54-60, 1997
- (6) K.Ohba and K.Ikeuchi, Recognition of the Multi Specularity Objects using the Eigen-Window, CMU-CS-96-105, 1996
- (7) Y.Cui and J.Weng, Hand Segmentation Using Learning-Based Prediction and Verification for Hand Sign Recognition, Proc. of CVPR96, pp.88-93, 1996
- (8) 片野, 木村, 視覚に基づくロボットの動作認識, ロボティクス・メカトロニクス講演会予稿集, 1997 (発表予定)
- (9) Y.Kuniyoshi, Behavior Matching by Observation for Multi-Robot Cooperation, Proc. of ISRR'95, 1995

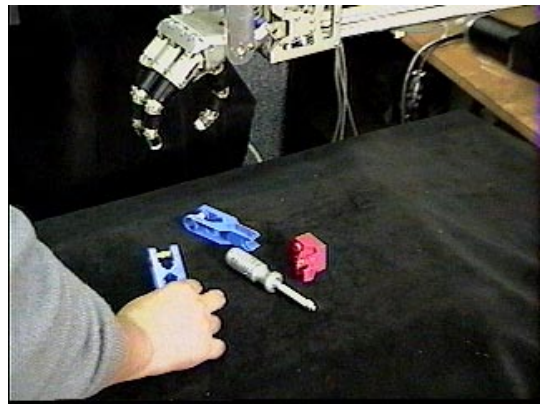


Fig.9 Experiment1: parts axis fixation

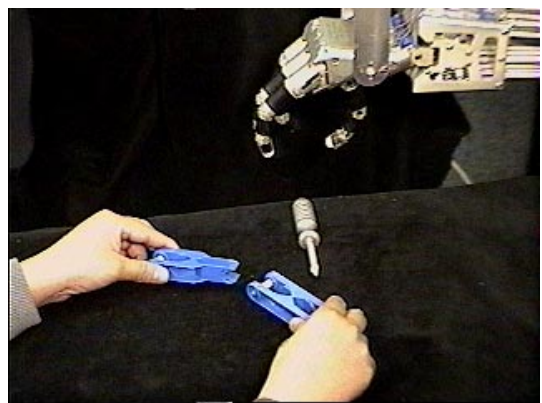


Fig.10 Experiment2: screw hole and driver tip fixation