

観察に基づく手作業の獲得における視覚の利用

小川原 光一[†] 高松 淳^{††}
木村 浩^{†††} 池内 克史^{††}

筆者らは、日常の手作業をロボットに自動的に獲得・再生させることを目的とし、主に視覚を用いて人間のお手本動作を観察することによって、再利用可能な抽象化された作業のモデルを構築し実ロボットによる動作の再現を可能にするための手法について研究を行っている。本稿では、物体どうしの相互作用の連続形として表現可能な作業を対象とし、獲得から行動再生までの手順を、(1) 観察に基づく物体および把持の認識、(2) 教示動作からの本質的な相互作用の抽出と視覚の誤差修正による運動軌跡の補正、(3) 行動再生時の視覚フィードバックによる外乱補償の3つに分けて概略を述べる。

Vision Based Acquisition of Manipulation Tasks

KOICHI OGAWARA,[†] JUN TAKAMATSU,^{††} HIROSHI KIMURA^{†††}
and KATSUSHI IKEUCHI^{††}

The purpose of our research is to establish a method for acquiring and reproducing everyday manipulation tasks automatically by a robot system. A demonstration of a task is observed mainly by vision and a reusable abstract task model is constructed. In this paper, a process for acquiring and reproducing a manipulation task which can be represented as a sequence of interactions between objects is discussed. The process is made up of three parts; namely, (1) Observation of demonstration: identification and recognition of objects and grasps, (2) Task analysis: extraction of essential interactions and correction of trajectory obtained from vision, (3) Reproduction: generation of robot behavior which compensates collapse of grasps by visual feedback mechanism.

1. はじめに

近年、家庭やオフィスなどの人間共存社会で行動することを目的とした共生型ロボットの開発が積極的に行われているが、このような多様かつ動的に変化する環境で使用されるロボットには、環境もしくはユーザのニーズに応じて新規の動作を適応的に獲得する能力が求められる。この種のロボットの役割の1つは人間に代わって日常における種々の作業を代行することであり、このとき新たな作業内容をいかにロボットに伝達し、またいかに再利用性のある形で習得させるのが課題となる。

ロボットの行動を生成する手段として、従来プログラミング言語や専用のティーチングデバイスを使用す

ることが一般的であったが、これらはロボットに関する専門的な知識を必要とするため、一般的なユーザのための行動教示手法としてはそぐわない。そこで、初心者が熟練者のお手本を、もしくは子供が大人のお手本を見ることで作業を覚えていく過程との類似性から、作業の実演教示を観察させることによってロボットに作業を自動獲得させる方法が脚光をあびており、筆者らも観察に基づく日常の手作業の獲得を目的として研究を進めている。

以下、2章では、観察に基づく動作獲得に関するこれまでの研究例を紹介し、観察・解析・再生の3段階からなる筆者らの取り組みについて概略を述べる。3章では、実演教示の観察から得られる視覚情報を抽象化するための手法として、剛体の位置・姿勢の同定手法である3D Template Matching(3DTM)法を利用した環境物体の同定および時系列での3次元追跡手法について述べる。また3DTMを関節物体の同定に拡張した手法を紹介し、これを利用して把持を行う手の形状を推定する手法について述べる。この抽象化された情報を解析し作業の再生に必要な情報を切り出す手

[†] 科学技術振興機構

Japan Science Technology Agency

^{††} 東京大学

The University of Tokyo

^{†††} 電気通信大学

University of Electro-Communications

法として、4章では、まず複数回の教示動作から作業の遂行に必須である相互作用を多次元 DP Matching を用いて推定する手法について述べる。次に、組立て作業のように作業が接触状態遷移の形で記述できる場合に、各接触要素の状態を連立線形方程式に帰着させることによって軌跡の視覚誤差を修正する方法について述べる。5章では、切り出された情報を利用した人間型ロボットによる動作の再生について述べる。その際作業の失敗の認識と回復動作の一環として、外乱などの影響によるロボットの把持の崩れを視覚フィードバックを用いて補正する手法について述べる。最後に、6章でまとめを述べる。

2. 観察に基づく動作獲得

教示動作の観察に基づきロボットを用いて動作の理解と獲得を試みる研究例を概観すると、大きく分けて、(1) 脳の情報処理機能の工学的な実現に基づきボトムアップ的に動作の理解と獲得を試みるアプローチと、(2) ある限定された作業領域の中での汎用性のある作業の習得を目的としてトップダウン的に動作の理解と獲得の仕組みを考えるアプローチの2つに分類される。

2.1 脳の情報処理機能の工学的な実現

1つは人間の脳機能を理解するという目的の下で、脳の神経生理の知見に基づき計算機上に脳もしくは認知のモデルを構築し、そのモデルを検証する意味でロボットを用い人間の脳機能の仕組みを探るアプローチである^{1)~3)}。

古くは Meltzoff らによる新生児における模倣の実験⁴⁾ などにより、生得的な模倣メカニズムの存在が主張されており、これを生理学的に説明する知見として、たとえばサルで発見されたミラーニューロン (mirror neuron) の存在が報告されている。このニューロンは、自身が行う意味を持った特定の動作に反応して活動するだけでなく、他者が行う同じ意味の動作を見た場合にも同様に活動する⁵⁾ といった特徴があり、霊長類における見まねの基礎と考えられる²⁾。このような生得的に備わっていると考えられるいわば認知のための要素機能を工学的にモデル化し検証することで、脳機能の解明を図る研究が行われて^{3),6)}。

一方動作生成に関しては、Kawato らと Miyamoto らによって、けん玉やテニスサーブといった腕の動的な動きを必要とする動作を獲得するために、観察によって動作軌跡の最適な経路点を推定し^{7),8)}、経路点間をスムーズに遷移する腕の軌跡を順と逆のモデルを用いた順逆緩和計算による双方向神経モデル¹⁾を用い

て高速に計算する方法が提案されている。順逆緩和計算は、小脳で計算されているとされる腕の運動軌道生成の計算モデルを参考に提案された多関節腕の運動指令を決定する計算アルゴリズムであり、この実験結果は人間の動作の計測結果とよく一致する。

また、設計者の構築した身体運動や物体操作の物理としての理解の体系を明示的にロボットに埋め込むのではなく、外界との相互作用の中から外界に対する理解を創発的に獲得するメカニズムの実現を目指す「認知ロボティクス」という枠組みが提唱されている^{9),10)}、現状は方法論自体を模索している段階である。

これらの研究例では、基本的には脳の処理系の末端の要素機能に着目してその機能を同定することから全体を押し量ろうとするボトムアップ的な取り組みがなされているが、要素機能を組み合わせたときの総体としての挙動の解明にまでは踏み込んでいないのが現状である。そのため、比較的単純な全身運動の獲得・生成は実現されているものの、環境や物体との複雑な相互作用を必要とする動作の獲得法に関しては手探りの状態である。

2.2 特定の作業領域内での汎用的な動作獲得

もう一方のアプローチは、ある設定された作業領域 (たとえば凸物体どうしの組立て作業など) の枠内で実行される任意の作業を再現することを目的とし、そのために必要十分な情報を取得するための方法および作業の記述方法を工学的に設計する方法である。

たとえば研削作業のような定型動作は、ロボットアームの水平方向の位置と垂直方向の力の時系列情報を教示デバイスから得ることによって、直接作業の記述を獲得することができる¹¹⁾。

単純な Pick & Place 動作においても、指先と物体の双方を視覚により常時観察することにより、環境の状態遷移の形で作業の抽象表現を得ることができる¹²⁾。

一方、実行時のロボット動作のわずかな誤差が結果に重大な影響をおよぼす作業においては、一連の作業行程の中のどの箇所が実行に注意を要し、どの箇所がそうではないのかを区別し、重要な箇所については特別な動作生成規範を用いる方法が一般的である。掃引動作については、複数の教示動作ごとの軌跡のずれから局所的な許容誤差を推定する方法が提案されており¹³⁾、また組立て作業のように物体どうしの接触状態遷移列として記述可能な作業については、実行時には基本的には教示された軌跡をなぞりつつ、状態遷移が生じる箇所では所望の状態遷移が生じるように視覚・力フィードバック制御を行うことによって作業を遂行する方法が提案されている¹⁴⁾。

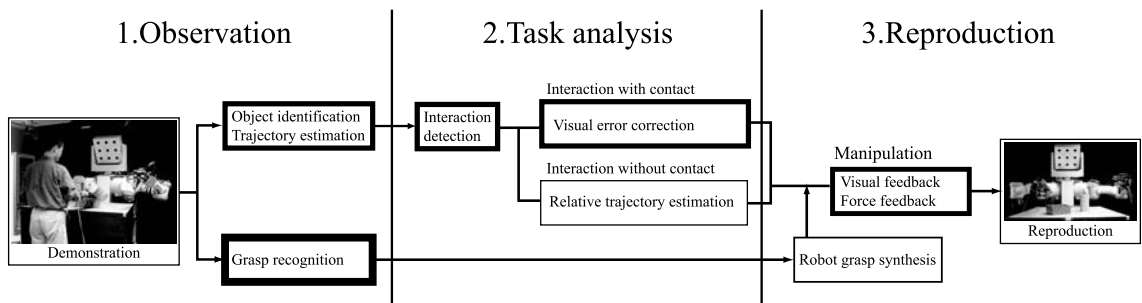


図 1 作業獲得の流れ

Fig.1 Overview of task acquisition.

また、教示対象の挙動を表す物理方程式が分かっている場合には、そのパラメータを推定することが作業を獲得することになり、たとえば観察による初期解から数回の試行を経て振り子の倒立動作を学習する研究がある¹⁵⁾。しかし、一般にパラメトリックな物理モデルを自動推定することは非常に困難であり、一方ノンパラメトリックの場合には単純な動作にしか適用できないという欠点がある。日常の手作業を簡単な物理モデルで記述できる可能性を示唆する報告¹⁶⁾はあるものの、一般的な作業を記述するための物理モデルの構成法についてはまったく分かっていない。

2.3 本研究における取り組み

本研究では、日常行われる手を使った作業を視覚などによる観察手段によって獲得し、これから何らかの抽象化された作業記述を導出し、これを用いてロボットにより同じ作業を再生することを目的とする。本稿では「日常の手作業」のなかでも、物体と物体との連続的な相互作用（接触・非接触動作を含む）で表現される作業を対象を絞り、この条件に該当する任意の作業を再生するための作業記述の構築方法、およびそのための観察手段・再生方法について述べる。上で述べた2つのアプローチでいえば後者にあたる。

図1に作業獲得の流れを示す。

2.3.1 観 察

まず、システムは実演教示を視覚により観察する。本稿で想定する手作業は物体と物体との連続的な相互作用であるが、この時点では、ある瞬間にどの物体とどの物体とが相互作用を行っているのかが明らかではないため、環境中に存在するすべての物体について時系列上で位置と姿勢の3次元追跡を行う（上段）。

またこのとき、教示者がどのように物体を把持しているのかという情報は、再生の際に重要となる。把持は、物体の安定を目的とした圧力把握と指先による操作を目的とした精密把握の2つ¹⁷⁾、もしくはその中間の把握を加えた3つに分類される¹⁸⁾。つまり、操作

物体・もしくは被操作物体の軌跡が分かったとしても、これをロボットが再生する際に、圧力把握によって主に腕全体を用いて操作するのか、それとも精密把握によって指による操りを行うのが異なり、これらは、たとえば作業に必要なとされるコンプライアンスや位置決め正確さなどに影響を与える。そのため、静的な把持形態の時系列変化の認識を同時に行う（下段）。

2.3.2 作業の解析

ここで想定している作業は、物体と物体との相互作用を、その作用が行われている期間相対運動を保った状態で再生することにより再現可能である。そのため、相互作用が行われていない期間の運動、もしくは相互作用と関係のない物体の運動を無視することにより、教示時とは物体の配置などが異なる環境においても動作の再生が可能になる。そこで、観察によって得られた物体の軌跡情報から、作業の遂行に必要な相互作用の切り出しを行う。

切り出された相互作用が接触をとまうものである場合には、さらに観察による誤差を取り除き妥当な接触関係を保つ軌跡に修正することが、組立て作業におけるはめあいのような厳しい作業の再生にとって重要である。そのため、妥当な接触状態を保つように視覚誤差を補正し軌跡の修正を行う。

2.3.3 再 生

再生は作業解析で求められた相互作用を連続的に実現することによってなされる。このとき、ロボットは各相互作用ごとに、教示者が行った把持と同等の形態によって物体の把持を行う。

また、把持物体の軌跡を生成する際には、相互作用が接触をとまう場合には、必要に応じて視・触覚フィードバックにより要求される接触状態遷移を実現し、接触をとまわない場合にも、把持の崩れや外乱による不正な相互作用を補償するための視覚フィードバックにより適宜動作の補正を行う。

以降の章では、図1の太線で囲まれた内容について

概要を述べる．

3. 観 察

教示動作を解釈する場合には、まず場面で何が行われているのかを認識する必要がある．本研究の対象は手作業であるため、操作物体・被操作物体の動きの情報、および操作時の教示者の把持形態に関する情報の2種類を観察によって獲得する．

本章では、剛体物体の同定方法、剛体物体の3次元軌跡を獲得する方法、および把持物体によって部分的に隠蔽された関節物体である手の形態を推定する手法について述べる．いずれも、対象の3次元幾何モデルは既知とする．

3.1 3D Template Matching による同定手法

対象の3次元幾何モデル(3角パッチモデル)が既知であり、また位置や姿勢が不明な対象の実際の3次元形状が観測できる場合に、対象の位置・姿勢 $\langle q, t \rangle$ を適切な初期位置 $\langle q_0, t_0 \rangle$ を元にロバストに推定することが可能な3D Template Matching(3DTM)法¹⁹⁾について説明する．ここで、 t は位置を表す3次元ベクトルであり、 q は姿勢を表す4次元のquaternionベクトルである．

幾何モデルの i 番目の頂点のモデル座標系での3次元位置を mp_i とし、観測された3次元データ中の対応する3次元点を rp_i とすると、両者の関係は変換式 $p = \langle q, t \rangle$ によって誤差を含んだ式(1)の形で表される．ただし、演算 $q \cdot mp_i$ は、 mp_i を4次元ベクトルであるquaternion q に従って回転させることを意味する．

$$rp_i = q \cdot mp_i + t + \beta \quad (1)$$

ここで、 β はランダムな3次元誤差であり、 β がガウス分布に従うならば、 p は最小二乗法により式(2)を最小化することによって解くことができる．

$$E(p) = \sum \| q \cdot mp_i + t - rp_i \|^2 \quad (2)$$

しかし、実際の誤差分布は通常ガウス分布に従わないため、最小二乗法の場合外れ値の影響により同定処理が不安定になる．そのため、WheelerらはM推定法を誤差分布の評価に用いた物体の同定手法を提案した¹⁹⁾．M推定法は最小二乗法の一般形であり、 $E = \sum_i \rho(z_i)$ と定式化される．この関数を、共役勾配法を用いた反復計算によって漸的に局所最小値に落とすことにより p を推定する．ただし、 $\rho(z_i)$ は誤差 z_i の任意の関数である．式(3)で表されるLorentzian関数を $\rho(z)$ として用いることで、外れ値の影響を低減し同定性能を向上させることができる．これにより、対象物体の一部が隠蔽され不可視の場合にも、メッシュ

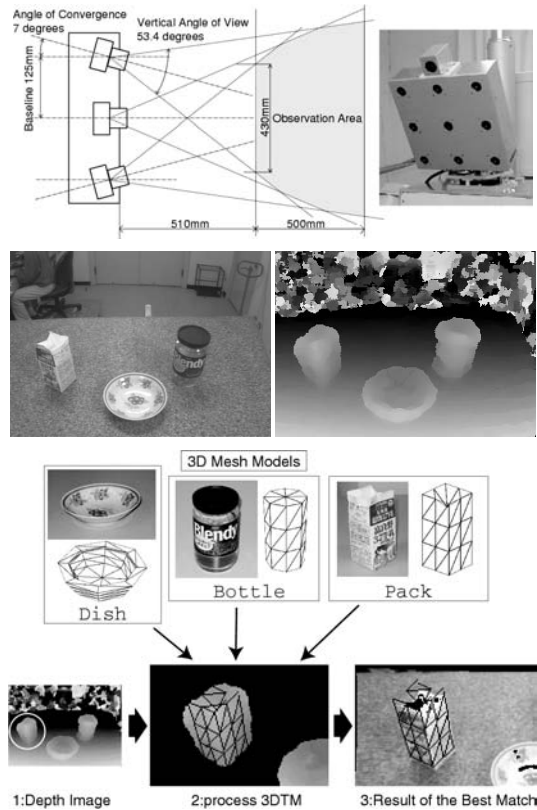


図2 多眼ステレオを利用した物体の同定
Fig. 2 Object identification.

モデル上で正しく対応する距離データがない箇所は外れ値として共役勾配法への寄与が低減されるため、正確な同定が可能となる．

$$\psi(z) = \frac{\frac{z}{\sigma}}{1 + \frac{1}{2}(\frac{z}{\sigma})^2} \cdots \text{Lorentzian weight function.} \quad (3)$$

ただし、 σ は重みである．

3.2 物体認識²⁰⁾

図2の多眼マルチベースラインステレオ装置²¹⁾を利用し、生成される色画像および視差画像より物体の抽出を行う．本研究ではテーブル上での手作業を想定しているため、テーブル上が空の状態での背景画像をあらかじめ記録しておき、距離データに対する背景差分法によってまず候補となる物体領域を抽出する．

次に、この各物体領域に対して、その重心を初期座標とし、あらかじめ用意された複数の3次元モデルについて上記の3DTMによって位置合せを行う．初期姿勢については、異なる4方向の姿勢を与え、同定結果の最も良かったものを選択する．このとき、ステレオ視によって得られる距離データは $2\frac{1}{2}$ 次元の情報であるため、同定の際も幾何モデルの3角パッチのうち法線が視点側を向いており、なおかつ自身の他の3角



図 3 操作物体の 3 次元追跡
Fig. 3 3D tracking of the grasped object.

パッチによって隠蔽されていないものだけを使う必要がある。位置合せが収束した後に、 M 推定を用いない純粋な二乗誤差の平均を出し、誤差の最も小さなモデルを対象物体とする。また、物体表面色の色相ヒストグラムをモデルとして持っておき、カラー画像上の候補領域の色相ヒストグラムとの差が閾値以上の場合には、同定を行う前に候補モデルを棄却することで同定性能の向上を図っている。

3.3 3次元軌跡推定²²⁾

本ステレオ視覚装置により、最大 15 fps の周期で連続するカラー画像と距離データを得ることが可能である。対象物体の時刻 0 での位置は上記の認識手法によって既知であるため、教示者が操作する物体の軌跡 $\langle q(t), t(t) \rangle$ を獲得するためには、 $\langle q(t-1), t(t-1) \rangle$ を初期値として、各物体ごとに上述の 3DTM を連続的に適用することによって図 3 のように求められる。図において、白のワイヤフレームモデルが追跡結果である。

3.4 把持形態の認識²³⁾

現在我々は、データグローブと隠れマルコフモデルを用いて把持形態を 14 種類の把持プリミティブ¹⁸⁾ に分類²⁴⁾、これをロボットハンドの把持にマッピング²⁵⁾ を行う研究を進めている。しかし、データグローブを含む接触型のセンサの場合、位置や姿勢・力覚といった情報を精度良く取得することが可能である反面^{20), 26)}、煩雑なセットアップやデバイスによって教示動作そのものが拘束されることが教示者の負担となる。

また手作業においては、必然的に物体を把持する局面が多くなるが、従来提案されてきた視覚に基づく手の形状認識手法は、そのほとんどがジェスチャ認識などを念頭に開発されているため、物体の把持を考慮していないものが大半であった。把持をともなう場合に問題となるのは、手の姿勢によってはカメラ画像上で手が把持物体により部分的に隠蔽されるため、この場合欠損データから把持形態を推定することになり推定精度が低くなる点である。

そこで、筆者らは指を積極的に動かす「操り」をともなわない把持を対象とし、把持期間中手の形状は変化しないが手の姿勢は変化するという仮定の下で、連

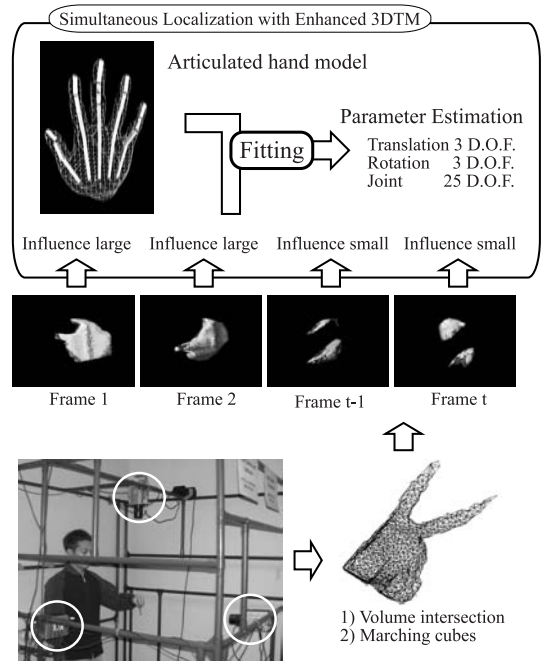


図 4 把持形態の推定
Fig. 4 Estimation of hand shape.

続する時系列データから N 枚のサンプルを取り出し、このうち欠損データからの影響をうまく低減することによって、結果的に欠損の少ないデータのみを利用して手の形状を推定する手法を提案する。

具体的には、光軸が直交するように配置した 3 台の赤外線カメラを用い、静的な把持を行っている手について姿勢が時間軸上連続して変化する画像列を獲得する。次に、時系列データの中から一樣な時間間隔で N セットの画像を抽出し、各セットについて図 4 下部に示すように同時刻の 3 枚の画像から視体積交差法によって手のポリウム表現を復元する。一方手の幾何モデルについては、図 4 上部に示すように手の表面形状を表すメッシュモデル(皮膚)を 25 関節のリンク構造(骨)にアタッチし、関節角を指定することによって自由変形する手の形状モデルを作成する。次に、式(4)に示すように、3DTM を関節物体の同定に適用可能なように拡張した手法(Enhanced 3DTM)を用いて、手の形状モデルの復元データに対するフィッティングを行い、姿勢と関節角の推定を行う。

$$E(p, \theta) = \sum_i \rho(|Rm_i(\theta) + t - r_j|^2) \quad (4)$$

$$\begin{pmatrix} m_i(\theta) \\ 1 \end{pmatrix} = \prod_l T_l(\theta) \begin{pmatrix} m_i \\ 1 \end{pmatrix}$$

ここで、 T_l は l 番目の関節角を変数にとり、モデル

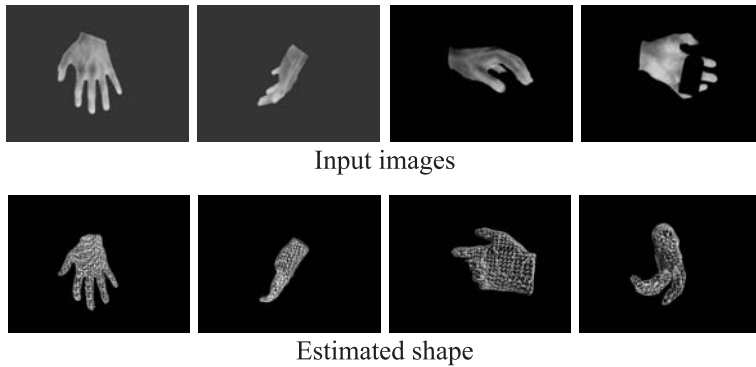


図 5 実環境における把持形態推定結果
Fig. 5 Estimation of hand shape in real environment.

のローカル座標系でメッシュの頂点 m_i の位置を変換する 4×4 の同次変換行列を表す。

このフィッティングの際に、欠損画像によって復元されたボリュームデータは大半が誤対応を引き起こす。そこで、 N 個の復元データすべてを同時に用いて関節角の収束計算を行うことで、 M 推定により誤対応の寄与を正しい対応からの寄与よりも相対的に低くすることができるため、結果的に欠損データを棄却することになり手の形状推定が可能になる。

図 4 の場合は、円筒形の物体を把持した状態で手を回転させた時系列データであるが、時刻 $t-1$ と t の復元データはオクルージョンによって指の大部分が欠損している。この場合、モデルの指に正しく対応する復元データがなく誤対応となるため、 M 推定によって影響が軽減される (Influence small)。シミュレーションによる定量的な評価では、関節角度の推定誤差は約 15 度であった。これは、復元データの条件が良い場合でも不完全な Visual Hull にしかならないことを考えると十分な結果であり、またこの程度の誤差であれば前記の 14 種類の把持プリミティブへのマッピングのための指標になりうる。また実環境においても、図 5 に示すように良い推定値が得られることを確認している。

4. 作業の解析

本章では、本質的には同一の作業を表す複数の異なる教示動作を相互比較することによって作業の遂行に必須の相互作用を抽出する手法と、抽出された相互作用が接触をともなう場合に、視覚の誤差を修正することによって妥当な接触状態を保った物体軌跡を推定する手法について述べる。

4.1 必須相互作用の推定^{22),27)}

本手法では、作業内容や対象物体に関する事前の知

識がない場合に、環境中の物体の見かけの運動のみから作業の遂行に必須の相互作用の推定を行う。ここで相互作用とは、(1) 相互作用を行う物体の組、(2) 相互作用が行われる期間、(3) 物体どうしの相対運動の 3 つの情報からなり、作業中に離散的に現れる必須相互作用の数とそのパラメータの推定を行う。

このとき、従来のように 1 回の観察結果のみを利用して解析を行う場合、(1) どの相互作用が作業の遂行に必須であるのか、(2) 把持物体とどの環境物体が相互作用を行っているのかの 2 点についてあいまい性があり、1 回の観察結果からでは事前の知識なしに必須相互作用を確定することはできない。

そこで、本質的には同一の作業を示す N 個の異なる教示動作の観察結果を統合し、すべての教示動作に共通する相互作用を抽出することで作業に必須の動作と相互作用を行う物体の組合せを同時に推定する手法を提案する。

まず、本質的には同一の作業を表す教示動作を複数 (N) 回実行しこれを記録する。ここで本質的に同一とは、作業の目的を達成するために必須の把持物体・環境物体間の相互作用が決まった形態・順序で現れるものと定義する。このとき、必須ではない相互作用が任意の順番で現れることは許す。すると、各教示動作は必須相互作用を内包した相互作用の連続形として表現することができる。

前述の理由から単一の観察結果には解析するうえであいまい性が存在するため、そこから必須相互作用のみを抽出することは困難である。そこで、教示動作を複数回行う際に、必須相互作用に影響を与えずに環境もしくは教示動作自体を毎回変化させる。すると、必須相互作用は同じであるが、それ以外の相互作用については教示動作によって異なる観察結果を得ることができる。

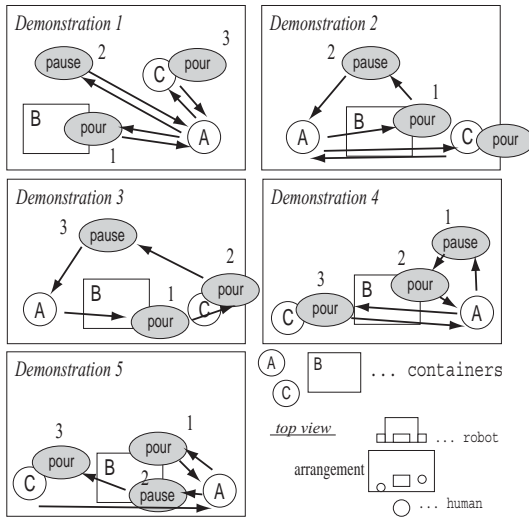


図 6 複数の教示動作

Fig. 6 Multiple demonstrations.

図 6 は、対象作業として「注ぎ動作」を選択し、5 回教示を行った例である。この例では、本質的な動作は、(1) 容器 A の中身を容器 B に注ぎ、(2) 次に容器 A の中身を容器 C に注ぐことである。個々の教示動作の違いは物体の配置や必須動作と関係のない手の動きにある。

次に、複数の観察結果を相互比較し、すべての観察結果に共通する相互作用を抽出することで必須の相互作用を推定する。相互比較を行う方法として、多次元 Dynamic Programming (DP) Matching²⁸⁾ を用いる。

各観測結果を要素配列 $Seq_i (1 \leq i \leq N)$ の形で表現したときに、DP Matching を適用するには任意の要素の組合せ $(Seq_1[index_1], \dots, Seq_N[index_N])$ の類似度を評価する関数 $E(index_1, \dots, index_N)$ を作成する必要がある。ここで、要素は検出された相互作用候補を指す。

$G(index_1, \dots, index_N)$ を、その要素の組合せに至るまでの評価値の重み付き和とすると、以下の漸化式によって最適パスつまり最適対応を計算することができる。すなわち、計算された最適パスが経由する要素の組合せが推定された必須相互作用である。

$$G(index_1, \dots, index_N) = E(index_1, \dots, index_N) \cdot \text{path}(ii, \text{index})/2 + \min\{\cup[G(ii_1, \dots, ii_N) + E(ii_1, \dots, ii_N) \cdot \text{path}(ii, \text{index})/2] \mid \text{index}_1 - S \leq ii_1 < \text{index}_1, \dots, \text{index}_N - S \leq ii_N < \text{index}_N\}.$$

ただし、 $\text{path}(ii, \text{index})$ は現在の格子点 index から探索

対象の格子点 ii までの距離 $(\sqrt{\sum_{k=1}^N (\text{index}_k - ii_k)^2})$ を表し、探索範囲内の連続した必須相互作用が縮退して 1 つの必須相互作用として扱われることを防ぐ役割を持つ。また、(1) ある要素が 2 回以上他の配列の要素と一致することを許さない、(2) 探索範囲を N^S の格子とする、との 2 つの拘束条件を設定する。前者は、必須相互作用が他の必須相互作用と重なって検出されないための条件である。後者の S は、必須相互作用間に入りうる必須ではない相互作用の数を規定するものである。 $S - 1$ が連続する必須ではない相互作用の数よりも少ないと、必須ではない相互作用が必須であると誤検出され、 S を大きくした場合には計算コストが増大する。各教示動作を見たときに最大の相互作用候補数に合わせるのが最も安全である。

本稿では、把持物体と環境物体との相互作用の検出を目的としており、ある要素の組合せにおいて把持物体とある環境物体との相対関係がすべての要素にわたって一致しているときに、その環境物体と把持物体が必須の相互作用をしている (評価関数 $E(\text{index}_1, \dots, \text{index}_N)$ が低い値を返す) と見なす。

ここで相対関係とは、図 7 に示すように把持物体 $GObj$ からある環境物体 Obj_j までの距離 $\text{dist}_i(GObj, Obj_j)$ および、把持物体からその環境物体へ向けた正規化ベクトル $\text{vec}_i(GObj, Obj_j)$ によって規定され、両者の標準偏差の重み付和の最小値をもって評価値とする。

評価関数 E は式 (5) のように表される。

$$E = \min_j \{ \text{dist_dev}_j + w \cdot \text{vec_dev}_j \}. \quad (5)$$

ただし w は両者の重みを示す定数であり、本研究では評価関数がより適切な値を出力するように経験的に定めた。

図 7 の例では、把持物体と Obj_3 との関係 $(GObj, Obj_3)$ はすべての教示動作を通してほぼ同一である。そのため、評価関数 E ではこの関係 $(GObj, Obj_3)$ を表す重み付き分散 dev_j が他の関係のものに比べて最小となるため、この値を評価値として返す。この関係 $(GObj, Obj_3)$ が相互作用候補となり、この評価値が小さいほどこれが必須動作である確率が高くなるが、最終的には DP マッチングの結果によってどの相互作用候補が必須であるのかが決定される。逆に、もし評価関数 E が小さい値を返さない場合には、 $(GObj, Obj_i) (1 \leq i \leq 3)$ はいずれも必須相互作用ではないと予想することができる。

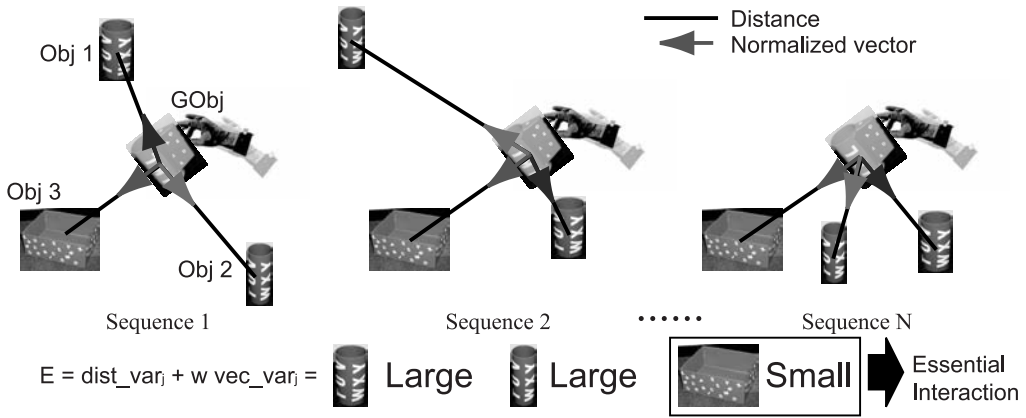


図 7 必須相互作用の評価
Fig. 7 Evaluation of essential interaction.

4.2 接触状態誤差の修正²⁹⁾

3章で述べた方法を用いて操作物体の3次元軌跡や環境物体の位置を計算した場合、これらは一般に誤差を含んでいるため、接触をとまなう作業においては物体どうしの埋め込みや離脱が生じてしまう。そのため、組立て作業のように拘束が増える方向に進む作業の接触状態を解析する場合、このままのデータでは正確な接触状態を推定することができない。

このとき、対象とする物体の幾何モデルが分かっている場合には、操作物体の位置・姿勢の誤差のみを考慮すればよいため、次の2段階の方法によって誤差を解消することができる。

- (1) 接触が生じている個所の候補を検出する。
- (2) すべての接触候補を利用して、妥当な接触状態を推定する。

第1段階では、接触要素として Vertex-Face, Face-Vertex, Edge-Edge の3種類を考え、画像より推定された位置の情報から、操作物体と環境物体との間のすべての接触要素の組合せについて、接触要素の距離がある閾値より小さい場合を接触候補とする。多面体の場合には、図8に示すように9種類の接触関係がありうるが、これらはこの3つの接触要素の組合せで記述できるため、この3つが必要十分である。

第2段階では、物体の幾何形状および位置・姿勢のパラメータから、各接触要素ごとに誤差量(接触要素の距離)を解析的に求めることが可能であるが、この解析式をテイラー展開によって線形形式に変換すると、いずれの接触要素についても式(6)の形で表されることが分かる。

$$\Delta e = c - [a^T b^T] \begin{bmatrix} T \\ r \end{bmatrix} \quad (6)$$

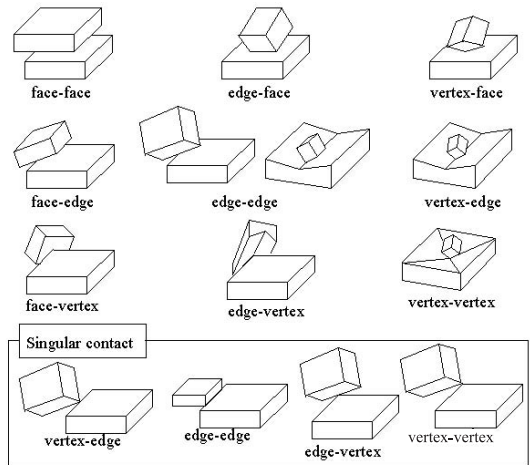


図 8 9種の接触要素表
Fig. 8 Nine types of contact elements.

ここで、 a, b, c は、元の解析式に応じて各接触要素ごとに決まる定数である、 T と r はそれぞれ操作物体の並進ベクトルと回転を表すオイラー角のベクトルである。

次に、最小二乗法を用いて式(7)を解くことによって、誤差を最小にする操作物体の位置姿勢を求めることができる。

$$\frac{\partial}{\partial q} \sum (\Delta e)^2 = 0 \quad (7)$$

5. 再生

5.1 ロボットによる動作の再現

図9に、接触をとまわらない作業を再生したときのロボット動作の連続画像を示す。再生の際には、物体の配置は教示のときの配置とは異なっているが、把持物体の軌跡が対象物座標系で教示のときと同一となる

ようにロボットアーム，もしくは指を制御することによって同じ作業を再現することが可能になる．必須相互作用間を結ぶ動作は，1つ前の相互作用の終了位置と次の相互作用の開始位置とを直線で結ぶ軌道によって生成される．

5.2 把持の崩れの補正による失敗の検出と回復³⁰⁾

ロボットによる作業再現時に注意すべきこととして，何らかの外乱や予期しない接触によって作業が失敗したときに，そのことを検出し必要なら回復動作をとることがあげられる．

本節では，たとえば物体操作中に環境にぶつけるなどして把持が崩れた場合に，それを適切に回復するための手法について説明する．本研究で対象とする作業



図9 作業の再現
Fig. 9 Reproduction from the task model.

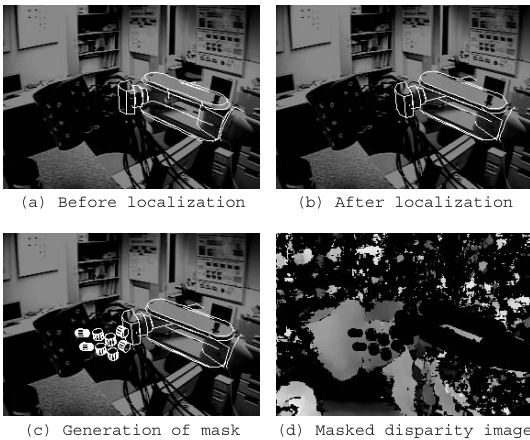


図10 把持物体を隠蔽するロボット指の影響を低減するためのマスク生成
Fig. 10 Generation of mask of robot figure.

においては，各相互作用に対応する相対軌跡が保たれれば作業は成功しているとして扱う．そのため，各相互作用の先頭で，以下の方法によって把持物体の姿勢を検証し，必要なら姿勢の補正を行う．

まずロボットは，作業モデルから想定されるカメラ座標系における把持物体の位置・姿勢と，ステレオ視覚装置より得られる色画像・距離データ中での位置・姿勢との比較を行う．このとき，把持物体はロボットハンドによって部分的に隠蔽されていると想定されるため，次の手順によって距離画像のマスキングを行い同定精度を上げる．

ロボットは，ロボットアームおよびハンドの既知の関節幾何モデルと現在の各関節角より，実行時のアームの幾何形状を生成する．これを，距離データから構築される3次元空間中の想定位置に投影し(図10(a))，3DTMによってアームの同定を行う(図10(b))．次に，カメラの視線方向とロボット指および予想される把持物体との位置関係から，把持物体を隠蔽する可能性のあるロボット指についてのみ上記のアームの幾何モデルに追加し，先ほどの同定で得られた個所に再投影する(図10(c))．これをマスクとし，この領域は後述の把持物体の推定の際には使用しないものとする(図10(d))．

次に，把持物体の想定位置に把持物体の幾何モデルを投影する(図11(a))．把持の崩れがある場合，これと実際の位置との間にはずれが生じているため，ここで3DTMを再び用い把持物体の同定を行うことによってずれの度合いの推定を行う(図11(b))．次に，このずれを最小にするようにロボットアームの逆運動学を解き，把持物体を想定位置に移動させる(図11(c))．この手順を，誤差が十分少なくなるまで繰り返す(図11(d))．ずれが比較的小さな場合は，この推定されたずれをオフセットとして本来の相対軌跡に付加することによって，作業を中断することなく外乱を補正することが可能になる．もしずれが大きい場合には，いったん把持物体をテーブル上に置き，持ち直すなどの操作が必要になる．

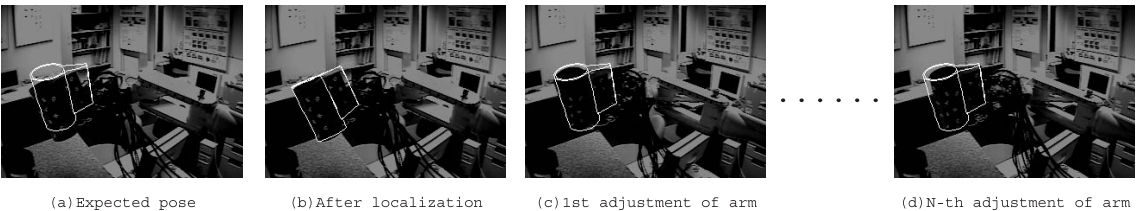


図11 把持物体のずれの推定
Fig. 11 Estimation of displacement of grasped object.

6. ま と め

本稿では，人間の行う手作業を観察に基づきロボットが習得し再生するための手法の一環として，物体の相対運動に着目し，主に視覚情報を利用して，教示時とは異なる環境下で人間のお手本動作をいかに忠実に再現するのに重点を置いた手法について述べた．

しかし，現状は動作の意味を考慮しないお手本をなぞるだけの原始的な模倣である．たとえば，容器の中身を注ぐという動作の本質は，容器自体を観察するだけでは理解することは困難であり，また成功例だけを観察していても分からないと思われる．今後の課題として，物体どうしの相互作用を観察の中心としながらも，その周辺で生じる事象が成功時と失敗時でどのように異なってくるのかを定量的に評価し，それから動作の成功を判定することがあげられる．また，高度な操りにおいて，5章で示したような静的な視触覚フィードバックを初期解としながら，徐々にフィードフォワード制御へ移行することによって操作への熟達化を実現することもあげられる．

謝辞 なお本研究は，文科省科研費補助金特定領域研究（C）課題番号 15017222 および，科学技術振興事業団・池内 CREST プロジェクトの補助を受けている．

参 考 文 献

- 1) 川人光男：脳の計算理論，産業図書（1996）.
- 2) Schaal, S.: Is imitation learning the route to humanoid robots?, *Trends in Cognitive Sciences*, Vol.3, No.6, pp.233–242 (1999).
- 3) Mataric, M.: *Sensory-Motor Primitives as a Basis for Imitation: Linking Perception to Action and Biology to Robotics* (2000).
- 4) Meltzoff, A. and Moore, M.: Imitation of facial and manual gestures by human neonates, *Science*, Vol.198, pp.75–78 (1977).
- 5) di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. and Rizzolatti, G.: Understanding motor events: a neurophysiological study, *Experimental Brain Research*, pp.176–180 (1992).
- 6) Inamura, T., Nakamura, Y., Ezaki, H. and Toshima, I.: Imitation and Primitive Symbol Acquisition of Humanoids by the Integrated Mimesis Loop, *Int. Conf. on Robotics and Automation*, pp.4208–4213 (2001).
- 7) Kawato, M., Gandolfo, F., Gomi, H. and Wada, Y.: Teaching by showing in kendama based on optimization principle, *Int. Conf. on Artificial Neural Networks*, Vol.1, pp.601–606 (1994).
- 8) Miyamoto, H., Schaal, S., Gandolfo, F., Gomi, H., Koike, Y., Osu, R., Nakano, E., Wada, Y. and Kawato, M.: A kendama learning robot based on bi-directional theory, *Neural Networks*, Vol.9, No.8 (1996).
- 9) 浅田 稔, 石黒 浩, 國吉康夫：認知ロボティクスの目指すもの，日本ロボット学会誌，Vol.17, No.1, pp.2–6 (1999).
- 10) 國吉康夫，ベルトウーズリエク：身体性に基づく相互作用の創発に向けて，日本ロボット学会誌，Vol.17, No.1, pp.29–33 (1999).
- 11) Asada, H. and Asari, Y.: The direct teaching of tool manipulation skills via the impedance identification of human motions Learning Tasks From A Single Demonstration, *Int. Conf. on Robotics and Automation*, pp.1269–1274 (1988).
- 12) Kuniyoshi, Y., Inaba, M. and Inoue, H.: Learning by watching, *IEEE Trans. Robotics and Automation*, Vol.10, No.6, pp.799–822 (1994).
- 13) Delson, N. and West, H.: Robot Programming by Human Demonstration: Adaptation and Inconsistency In Constrained Motion, *Int. Conf. on Robotics and Automation*, pp.30–36 (1996).
- 14) Ikeuchi, K. and Suehiro, T.: Toward an Assembly Plan from Observation Part I: Task Recognition With Polyhedral Objects, *IEEE Trans. Robotics and Automation*, Vol.10, No.3, pp.368–384 (1994).
- 15) Atkeson, C.G. and Schaal, S.: Learning Tasks From A Single Demonstration., *Int. Conf. on Robotics and Automation*, pp.1706–1712 (1997).
- 16) Arimoto, S.: Robotics research toward explication of everyday physics, *Int. Journal of Robotics Research*, Vol.18, No.11, pp.1056–1063 (1999).
- 17) Napier, J.: The prehensile movements of the human hand, *Journal of Bone and Joint Surgery*, Vol.38B, No.4, pp.902–913 (1956).
- 18) 鎌倉矩子：手のかたち手のうごき，医歯薬出版株式会社（1989）.
- 19) Wheeler, M.D. and Ikeuchi, K.: Sensor Modeling, Probabilistic Hypothesis Generation, and Robust Localization for Object Recognition, *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol.17, No.3, pp.252–265 (1995).
- 20) Ogawara, K., Iba, S., Tanuki, T., Kimura, H. and Ikeuchi, K.: Recognition of Human Task by Attention Point Analysis, *Int. Conference on Intelligent Robot and Systems*, Vol.3, pp.2121–2126 (2000).

- 21) <http://www.komatsu.co.jp/research/study56.htm>
- 22) 小川原光一, 高松 淳, 木村 浩, 池内克史: 複数教示動作の時系列上での統合に基づく人間作業のモデル化手法, 情報処理学会論文誌: コンピュータビジョンとイメージメディア, Vol.43, No.SIG4, pp.117-126 (2002).
- 23) 橋本謙多郎, 小川原光一, 高松 淳, 木村 浩, 池内克史: 赤外線画像を用いた視体積交差法による把持形体の認識, 情報処理学会研究報告(CVIM 135) (2002).
- 24) Bernardin, K., Ogawara, K., Ikeuchi, K. and Dillmann, R.: A Hidden Markov Model Based Sensor Fusion Approach for Recognizing Continuous Human Grasping Sequences, *The International Conference on Humanoid Robots* (2003).
- 25) 園田展人, 小川原光一, 木村 浩, 池内克史: 日常動作に置ける把持の認識とロボットによる再現, 情報処理学会研究報告(CVIM 137), pp.61-68 (2003).
- 26) 津田雅之, 高橋友一, 小方博之: 実演教示による組み立て作業モデルの生成, 日本ロボット学会誌, Vol.18, No.4, pp.73-82 (2000).
- 27) Ogawara, K., Takamatsu, J., Kimura, H. and Ikeuchi, K.: Extraction of essential interactinos through multiple observations of human demonstrations, *IEEE TIE*, Vol.50, No.4, pp.667-675 (2003).
- 28) Fuellen, G.: A Gentle Guide to Multiple Alignment, *Complexity International*, Vol.4, pp.1-56 (1997). <http://www.csu.edu.au/ci/vol04/mulali/mulali.html>
- 29) Takamatsu, J., Ogawara, K., Kimura, H. and Ikeuchi, K.: Correcting Observation Errors for Assembly Task Recognition, *Int. Conf. on Intelligent Robots and Systems*, pp.232-237 (2002).
- 30) Ogawara, K., Kimura, H. and Ikeuchi, K.: Refining hand-action models through repeated observations of human and robot behavior by combined template matching, *Int. Conference on Intteligent Robot and Systems* (2001).

(平成 15 年 5 月 6 日受付)

(平成 15 年 9 月 9 日採録)

(担当編集委員 角 保志)



小川原光一

1975 年生. 1997 年東京大学工学部機械情報工学科卒業. 2002 年東京大学大学院工学系研究科電子情報工学専攻博士課程修了. 工学博士. 2002 年より科学技術振興機構博士研究員. 知能ロボットの動作獲得に関する研究に従事. 日本ロボット学会, IEEE 各会員.



高松 淳

1974 年生. 2001 年東京大学大学院理学系研究科情報科学専攻博士前期課程修了. 現在同大学院情報理工学系研究科コンピュータ科学専攻博士後期課程在籍. 観察に基づくロボットプログラムの自動生成に関する研究に従事.



木村 浩(正会員)

1961 年 2 月 5 日生. 1983 年東京大学工学部機械工学科卒業. 1988 年東京大学大学院工学系研究科機械工学専攻博士課程修了. 工学博士. 東北大学工学部助手, 電気通信大学講師を経て, 現在, 電気通信大学情報システム学研究科助教授. 1995 年~1996 年カーネギーメロン大学客員研究員. 歩行ロボット, 知能ロボットの研究に従事. 1989 年度日本ロボット学会論文賞受賞. 日本ロボット学会, 日本機械学会, 計測自動制御学会, IEEE 各会員.



池内 克史(正会員)

1973 年京都大学工学部機械工学科卒業. 1978 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. MIT 人工知能研究所, 電総研, CMU 計算機科学科を経て, 1996 年より東京大学生産技術研究所教授. 人間の視覚機能, 明るさ解析, 物体認識, 人間による組立て作業の自動認識等の研究に従事. 工学博士. D.Marr 賞(ICCV : 1990 年), IEEE 優秀論文賞(CVPR : 1991 年), 最多引用論文賞(AI Journal : 1992 年) 等受賞. IEEE Fellow.