

作業の目的を考慮した視覚認識戦略の生成

三浦 純^{*1} 池内克史^{*2}

Task-Oriented Generation of Visual Sensing Strategies

Jun Miura^{*1} and Katsuhi Ikeuchi^{*2}

In vision-guided robotic operations, vision is used for extracting necessary information for achieving the task. Since visual recognition is usually performed with limited resources, visual sensing strategies should be planned so that only necessary information is obtained efficiently. This paper describes a method of generating visual sensing strategies for assembly tasks based on knowledge of the task to be performed. The generation of the appropriate visual sensing strategy entails knowing *what* information to extract, *where* to get it and *how* to get it. This is facilitated by the knowledge of the task, which describes what objects are involved in the operation, and how they are assembled.

In the proposed method, using the task analysis based on face contact relations between objects, necessary information for the current operation is first extracted. Then, visual features to be observed are determined using the knowledge of the sensor, which describes the relationship between a visual feature and information to be obtained. Finally, feasible sensing strategies are evaluated based on the predicted success probability, and the best one is selected.

Our method has been implemented using a laser range finder as the sensor. Experimental results show the feasibility of the method, and point out the importance of task-oriented evaluation of sensing strategies.

Key Words: Task-oriented vision, Vision planning, Active vision, CAD-based vision

1. はじめに

視覚を利用するロボット作業においては、視覚は作業を適切に行うために必要な情報を作業環境から獲得するために用いられる。一般に、視覚認識は限られた資源の下で行われるため、必要な情報だけを効率よく獲得できるように視覚認識戦略を生成する必要がある。効率的な視覚認識戦略を生成するためには、作業に関する知識が不可欠である。作業に関する知識がなければ、観測すべき視覚特徴を適切に選ぶことは難しく、また、不要な視覚情報の処理に時間を費やしてしまう可能性がある。

以上のような見地から、タスク指向視覚、能動視覚、目的視覚に関する研究が活発に行われている [1][2][3]。作業に関する知識を用いることにより、高速かつロバストな視覚システムを構築することができるが、その開発はロボットの行う作業の内容に強く依存し、多くの労力を必要とする。したがって、作業に応じて効率的な視覚認識戦略を自動的に生成できる手法の開発が望まれている。

ロボット作業を適切に誘導するための視覚認識戦略の生成は、次の3つのステップに分解できる。

- どのような視覚情報が現在の作業に必要なかを決定する。
- どの視覚特徴を観測することにより、その情報を得るかを決定する。
- センサ配置などの具体的な認識戦略を決定する。

最初の2つのステップは有用な視覚特徴へ注視点を制御する問題であり、最後のステップは、視覚認識戦略候補の生成と評価にかかわるものである。

注視点制御の能力は効率的な視覚認識戦略生成のために重要である。視覚情報や事前情報に不確かさがある場合に、それを解消するために最適な視点を求める手法が提案されているが [4][5]、視覚特徴の組はあらかじめ与えられており、作業の記述から自動的に抽出されるわけではない。また、ロボットの作業目標や作業環境を詳細に解析することにより、効率的な視覚認識を実現しているものもあるが [6][7]、その開発はかなり各ロボット作業に依存したものになり、汎用性に欠ける。

3番目のステップ(具体的な認識戦略の決定)は、作業に適切な認識戦略の集合を生成するステップと、その中から最適なものを選ぶステップとに分けられる。前者は、視覚特徴の可視性や検出可能性などの条件を満たすべき観測条件を求めるものであ

原稿受付 1995年5月25日

^{*1}大阪大学工学部

^{*2}カーネギー・メロン大学計算機科学科

^{*1}Faculty of Engineering, Osaka University

^{*2}Computer Science Department, Carnegie Mellon University

る [8] [9]. 観測対象が与えられたときに, 自動的に条件を満たす観測条件を生成する機能は, すべての観測戦略プランナがサブルーチンとして備えるべき機能である.

最適な認識戦略を求める問題は, 視覚情報の「良さ」を測る評価関数を設定し, それを最適にする認識戦略を選ぶ, という手順で解くことができ, そこでは評価関数を適切に設定することが重要である. 多くの場合, 視覚による計測値の誤差を最小にする, という評価基準 [10] [11] が用いられる. 例えば, 計測値の誤差分布の共分散行列の行列式の値を最小にする, という評価値を考慮することができる. しかし, ある特定の作業に対して, 視覚情報のある部分が他の部分に比較して重要である (例えば, より正確な値を必要とする) 場合に, それを考慮した評価関数を生成する必要がある.

本論文では, 作業に関する知識を用いることにより, 各作業を遂行する上で最適な視覚認識戦略を自動的に生成することのできる手法について述べる. 対象とする作業は既知環境における組立作業であり, 物体の形, 大きさ, 概略の位置は CAD ベースの環境モデルから与えられるものとする. このような作業における視覚の役割は, 物体の位置を十分な精度で求め, 適切に組立が行われるようにすることである. 本論文では, 上で述べた視覚認識戦略生成過程のうち, (1) 各組立動作に必要な視覚情報をどのように決定するか, (2) 観測すべき視覚特徴の組をどのように選ぶか, (3) 各々の視覚認識戦略をどのように評価するか, について新たな手法を提案する.

以下, 第 2 節では, 面接触を基本とした組立動作の解析結果を用いて, 各組立動作に必要な視覚情報を決定する手法を述べる. 第 3 節では必要な視覚情報を得るために観測する視覚特徴の組を選ぶ方法を述べる. 第 4 節では, 作業の目的を考慮して認識戦略を評価する方法を述べる. 第 5 節では, ラインレンジファインダを用いた実装について述べる. 第 5.4 節では実験結果を述べ, 作業の目的を考慮した評価の重要性を示す. 第 6 節では提案する手法を基にした, 視覚認識戦略自動生成システムの構成について述べる.

2. 作業に必要な視覚情報の決定

ここでは, 組立作業を物体間の面接触に基づいて解析し [12], その結果を用いて, 各作業が視覚情報を必要とするか, 必要ならどのような情報が必要かを決定する手法について述べる.

この解析はまず, 多面体が並進運動を行う場合を扱い, 後でより一般的な場合を扱う. 作業の過程において, マニピュレータによって操作される物体は, 面接触状態間を遷移するものとし, 各組立作業の目的は, 目標とする面接触状態を実現することとする.

2.1 接触状態の表現と分類

操作される物体と環境との間の, ある 1 組の面接触による物体の運動に対する拘束は, 以下の不等式で表される.

$$N \cdot \Delta T \leq 0$$

ここで, ΔT は物体の可能な運動方向, N は環境側の接触面の法線方向である. 可能な運動方向はガウス球上の点の集合を用い

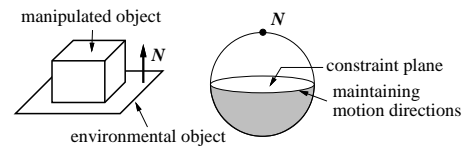


Fig. 1 Unidirectional constraint and its representation on the Gaussian sphere.

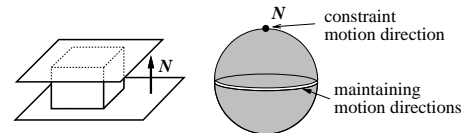


Fig. 2 Bidirectional constraint.

て表すことができる (Fig. 1 参照). 1 つの面接触によりガウス球面は 2 つの半球に分割される. 北極方向を法線 N の方向とすると, 赤道を含む北半球は物体の可能な運動方向, 南半球は不可能な運動方向を表す.

異なる方向の面接触が同時に生じるときには, 可能な運動方向は各面接触に対する複数の不等式を同時に満たす領域として計算でき, ガウス球上では複数の運動可能領域の論理積で表現される.

Fig. 1 では赤道に対応する方向は, 現在の接触状態を維持する運動の方向である. この運動を「維持動作」と呼び, その動作の自由度を「維持自由度 (maintaining DOF)」と呼ぶ. また, 赤道を含まない北半球内の方向への運動は現在の接触状態を解消する. この運動を「離脱動作」と呼ぶ. 離脱動作のうち, 維持動作の方向成分を全く含まないものを純粋「離脱動作」と呼び, その動作の自由度を「離脱自由度 (detaching DOF)」と呼ぶ.

Fig. 2 は, 2 つの接触面の方向が完全に向かい合っている場合を示す. 可能な運動方向は赤道上の維持動作に限られ, 離脱動作は存在しない. 自由空間の並進 3 自由度のうち 1 自由度は完全に拘束されている. このような自由度を「拘束自由度 (constraining DOF)」と呼ぶ.

ある接触状態は, (維持, 離脱, 拘束) の自由度の 3 つ組を用いて記述することができる. 例えば, Fig. 1 と Fig. 2 の状態は, それぞれ $(2, 1, 0)$, $(2, 0, 1)$ と表される. 各組立動作は 1 つの接触状態から別の接触状態への遷移であると見なすことができる. 可能な接触状態と可能な遷移とを調べることで, Fig. 3 に示す遷移グラフを得ることができる [13].

2.2 作業に必要な視覚情報の決定

すべての組立動作は, 現在ロボットによって操作されている物体のいくつかの自由度についての拘束を増加させる. 上で述べたように 3 種の自由度があることから, ある自由度の拘束の増加には, 維持自由度から離脱自由度へ, 離脱自由度から拘束自由度へ, および維持自由度から拘束自由度への, 3 通りの場合が考えられる. Fig. 4 に示す 3 つの場合において, 水平方向の自由度の変化がそれぞれの 3 通りの変化に対応している.

ここで, Fig. 4 の各動作がどのように実現できるかについて詳しく検討する. (a) の場合には, 水平自由度は維持自由度から離脱自由度へと変化する. 物体の移動方向は, 接触後の純粋離脱

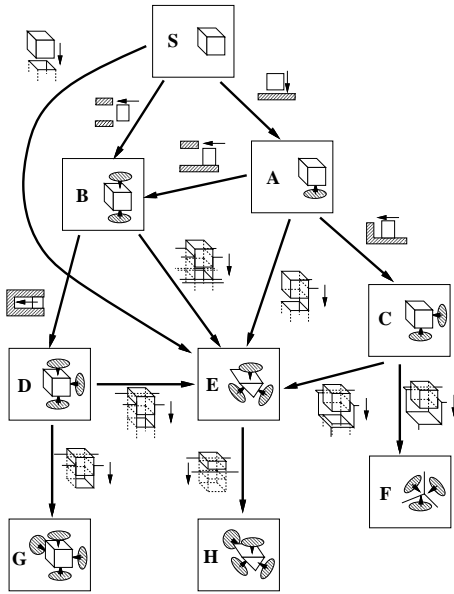


Fig. 3 Contact states and transitions represented as a directional graph [13].

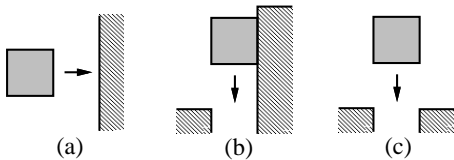


Fig. 4 Three typical cases of increase of constraint on a horizontal degree of freedom.

Type and transition of the triplet is as follows:

- (a): maintaining \rightarrow detaching $((3, 0, 0) \rightarrow (2, 1, 0))$.
- (b): detaching \rightarrow constraining $((2, 1, 0) \rightarrow (2, 0, 1))$.
- (c): maintaining \rightarrow constraining $((3, 0, 0) \rightarrow (2, 0, 1))$.

動作の方向，すなわち右側の壁の法線方向と一致しているので，この動作は接触が生じるまで物体を動かす，という戦略で実現される．したがって，視覚情報なしに力覚情報だけで実現できる．(b)の場合は，接触状態の遷移後には水平自由度は拘束自由度となるが，遷移に必要な物体の水平位置は，物体と右側の壁との接触を保ちながら動作することにより実現できるので，この場合も視覚情報は必要ない．それに対して，(c)の場合には，水平自由度は維持自由度から拘束自由度へと変化し，状態遷移の直前には視覚情報を用いて物体の水平位置を調節する必要がある．

ある自由度が維持自由度であるとき，その自由度に関しては物理的な接触は存在しないため，その自由度の位置決めには力覚情報を用いることはできない．したがって，その自由度が拘束自由度へと変化するときには，その自由度の位置決めを視覚情報を用いて行う必要がある．

以上をまとめると，ある自由度が維持自由度から拘束自由度へ変化するとき，その自由度を視覚によって観測する必要がある．現在の組立動作がそのような自由度の変化を含まないとき，その組立動作に視覚情報は必要ない．この理論を Fig. 3 に適用した結果，13 の接触状態遷移のうち4つ ($S \rightarrow B$ ， $S \rightarrow E$ ， $A \rightarrow E$ およ

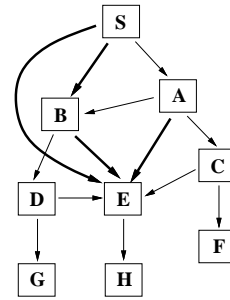


Fig. 5 Classification of state transitions. Bold lines indicate the transitions that require visual information. Thin lines indicate the transitions that do not require visual information.

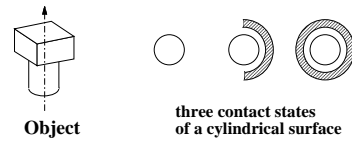


Fig. 6 The object considered in the extended contact state analysis.

び $B \rightarrow E$) が視覚情報を必要とすることがわかる (Fig. 5 参照) .

2.3 接触状態解析の拡張

ここでは，2.1節で述べた接触状態の解析を，物体が平面および円筒面からなり，かつ並進3自由度に加え回転1自由度を含む場合について拡張する．ほとんどの組立作業はこの4自由度を用いて行えるので，この解析は実用的には有用である．

Fig. 6は，平面と円筒面を両方持つような物体の典型的な例である．この物体を対象として解析を行った．円筒面の接触状態を図に示すような3つの場合，すなわち全く接触がない場合，半面だけ接触している場合，全面で接触している場合に限定した．これらはそれぞれ，円柱面の中心軸に垂直な方向の並進自由度が，維持自由度，離脱自由度，拘束自由度である場合に相当する．3つの場合以外の接触状況は，自由度の分類上は，すべてこれらの3つの場合のどれかであると見なすことができる．

この拡張した解析では各接触状態を並進3 + 回転3の自由度の6つ組で表現することができる．解析の結果，54の接触状態と85の接触状態間遷移が得られた．詳しい結果は文献[14]を参照されたい．

回転運動によって接触状態が遷移する場合，遷移直前の微小運動は並進運動とみなすことができる．したがって，上で述べた，各組立動作に必要な視覚情報を求める理論は，回転運動の場合にも適用可能である．各接触状態遷移の自由度の6つ組の変化にこの理論を適用した結果，19の状態遷移が視覚情報を必要とすることがわかる．さらに，視覚を用いて観測すべき自由度という点から，それらの動作を調べた結果，6つのグループに分類される．Fig. 7にそれぞれのグループの，典型的な動作例を示す．図中の黒い矢印が観測すべき自由度を示している．

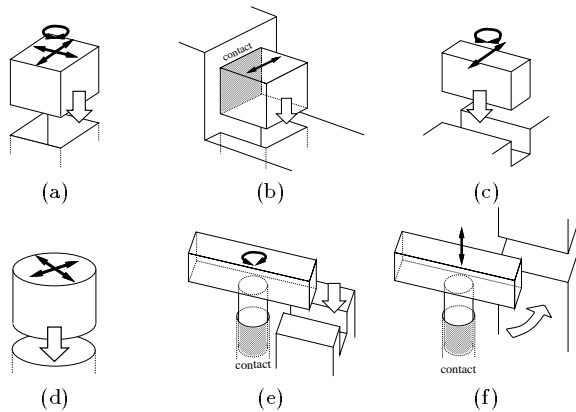


Fig. 7 Transition groups which need visual information. Thick arrows indicate the direction of movement. Thin arrows indicate degrees of freedom to be adjusted by use of visual information. The transition of sextuplet for each case is as follows:
 (a): $(3, 0, 0; 1, 0, 0) \rightarrow (1, 0, 2; 0, 0, 1)$.
 (b): $(2, 1, 0; 0, 0, 1) \rightarrow (1, 0, 2; 0, 0, 1)$.
 (c): $(3, 0, 0; 1, 0, 0) \rightarrow (2, 0, 1; 0, 0, 1)$.
 (d): $(3, 0, 0; 1, 0, 0) \rightarrow (1, 0, 2; 1, 0, 0)$.
 (e): $(1, 0, 2; 1, 0, 0) \rightarrow (1, 0, 2; 0, 0, 1)$.
 (f): $(1, 0, 2; 1, 0, 0) \rightarrow (0, 0, 3; 1, 0, 0)$.

3. 注視すべき視覚特徴の選択

視覚で観測すべき自由度が決定されたなら、次にその自由度が観測できるための、視覚特徴の組を選ばなければならない。ここでは、作業に関する知識を用いて、そのような視覚特徴の組を選ぶ手法について述べる。

3.1 センシング・プリミティブ

視覚特徴の選択問題を解くために、センシング・プリミティブという概念を導入する。センシング・プリミティブは抽象化された視覚認識手続きであり、ある視覚特徴とそれを観測することによって観測できる自由度との関係を記述する。多面体の辺や面といった、各々の視覚特徴に対して1つのセンシング・プリミティブを用意する。対象とする組立作業に現れる可能性があり、かつ使用する視覚センサで観測可能な視覚特徴を事前に数え上げることにより、必要なセンシング・プリミティブを用意しておく。センシング・プリミティブの例は次節で示す。

3.2 視覚特徴選択のプロセス

Fig. 8に視覚特徴選択のプロセスの概略を示す。作業についての記述から、前節で述べた理論を用いて、どの自由度を観測すべきかが決定される。一方、観測可能な視覚特徴の集合が、これも作業についての記述から得られる。それらの情報と用意されたセンシング・プリミティブとから、どの視覚特徴の組が現在の組立動作に十分な情報を与えるかを決定する。必要な幾何情報はCADベースの環境モデルを参照する。

この視覚特徴選択問題は次のように定式化できる。いま操作されている物体が6自由度を持つ一般的な場合について考える。並進自由度と回転自由度についての拘束は独立に考えることができるので、2つのガウス球、 G^T および G^R を考え、ある面接触に

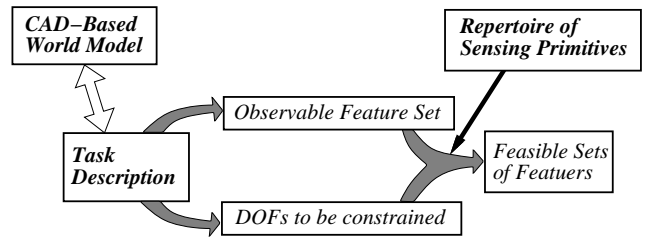


Fig. 8 Selection of features to be observed using task description and sensing primitives.

よる拘束の下で、 G^T 上で可能な並進運動方向、 G^R 上で可能な回転運動の軸方向を表すことにする。次の表記を用いる。

- $A_{goal}^T (A_{goal}^R)$: 現在の組立動作終了後に満たすべき拘束を表す $G^T (G^R)$ の球面上の点の集合。
- $A_{curr}^T (A_{curr}^R)$: 現在の (動作前の) 拘束を表す $G^T (G^R)$ の球面上の点の集合。Fig. 4(b)における物体と右側の壁との接触維持動作のような、ロボットの力覚を用いた制御によって実現可能な拘束も含む。
- $A_i^T (A_i^R)$: i 番目の視覚特徴を観測することによって得られる拘束を表す $G^T (G^R)$ の球面上の点の集合。この情報はセンシング・プリミティブとして記述されている。

n 個の異なる特徴 $\{feature_i | i = 1, \dots, n\}$ を観測すると、結果として以下の拘束 A_{obsd}^T, A_{obsd}^R が得られる。

$$A_{obsd}^T = A_{curr}^T \cap \left(\bigcap_{i=1}^n A_i^T \right),$$

$$A_{obsd}^R = A_{curr}^R \cap \left(\bigcap_{i=1}^n A_i^R \right).$$

この特徴の組が必要な拘束を与えるのに十分であるためには、以下の条件を満たす必要がある。

$$A_{goal}^T \supseteq A_{obsd}^T \text{ かつ } A_{goal}^R \supseteq A_{obsd}^R. \quad (1)$$

A_i^T と A_i^R のペアはセンシング・プリミティブを表現する。例えば、直線、平面、円弧、円筒面の4つの視覚特徴がある場合を考える。各特徴の3次元空間内の位置を観測することによって得られる拘束はFig. 9のように与えられる。例えば、ある物体の直線エッジを観測することにより、その物体の位置をそのエッジに沿った方向を除いて定めることができ、またその姿勢をそのエッジまわりの回転を除いて定めることができる。

視覚特徴選択の例題として、Fig. 10に示される2種のペグ挿入動作を考える。穴の位置をその辺を観測することによって定めるものとする。調節すべき自由度は挿入方向に垂直な平面上に限られるので、ガウス球の代わりにガウス円を用いて、拘束を表現する。また、回転自由度は少なくとも1つの辺を観測することによって拘束できるので、並進自由度に関する特徴選択だけを考える。

Fig. 10に示されるように、現在の拘束 A_{curr}^T について、(a)の場合には全円となり、(b)の場合には x 軸上の2点となる。 A_{goal}^T はどちらの場合も、可能な運動方向がないという状態である。次に、2つのエッジ e_1 と e_2 についての A_i^T は、Fig.

primitive feature	movable direction	possible rotation axes
line 	 two polar points	 two polar points
plane 	 a great circle	 two polar points
circular curve 	 no movable direction	 two polar points
cylindrical surface 	 two polar points	 two polar points

Fig. 9 Example sensing primitives represented by the Gaussian spheres. White regions indicate movable direction or possible rotational axes.

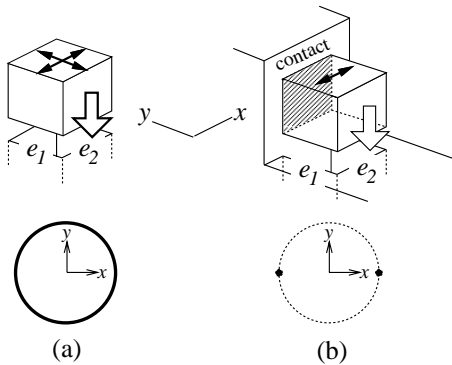


Fig. 10 Two insertion operations and movable directions.

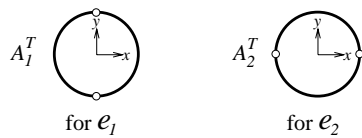


Fig. 11 Constraints obtained by observing two edges: A_1 for edge e_1 and A_2 for edge e_2 .

11に示すように、どちらもエッジの方向に対応したガウス円上の2点として表される。式(1)をこれらのデータに適用することにより、(a)の場合はエッジ e_1 を観測するだけで十分であるが、(b)の場合には e_1 と e_2 の両方を観測しなければならないことがわかる。

3.3 視覚センサ位置の候補の生成

観測すべき視覚特徴の組が選ばれたら、それらをすべて観測できるような視覚センサ位置の候補を生成する。視覚特徴の可視性

や検出可能性などの観測条件を考慮して、有効なセンサ配置を求める研究はすでに多く行われている[9]ので、それらを利用する。基本的には、センサの可動範囲を適当に量子化し、各量子化された位置において、(1) 必要な観測条件を満たすか、(2) センサと物体の衝突を起こさないか、をヒューリスティック規則(例えば、センサの方向と観測される面の法線方向の差がある範囲内に収まっていないといけない、といった規則)を用いて判定し、有効なセンサ位置の候補の組を生成する。

適切な視覚特徴の組が複数あるときには、それぞれの組に対し視覚センサ位置候補を生成する。最終的な視覚認識戦略は次節に述べる方法により決定する。

4. 作業の目的を考慮した視覚認識戦略の評価

4.1 視覚情報の不確かさに基づく評価

本論文で扱う組立作業では、視覚の目的は組み立てられる物体の、静的な環境に対する相対的な位置情報を得て、その情報をロボットにフィードバックして位置の修正を可能にすることにある。したがって、視覚認識戦略を、それによって得られるであろう物体位置の不確かさに基づいて評価することにする。

センサデータが3次元の計測点の集合から成り、各点はセンサのモデルから計算される不確かさを持つとする。このとき、物体の位置を決定する問題は次のように記述できる。 a を物体の位置を表すパラメータベクトルとする。ここで、 a は N 個の3次元計測点の集合 $\{x_i | i = 1, \dots, N\}$ から推定されるものとし、物体の幾何形状を表す次の方程式 f_i が各計測点について満たされるものとする。

$$f_i(x_i, a) = 0 \tag{2}$$

この方程式は一般に非線形であるが、概略の位置は作業の記述から与えられるので、拡張カルマンフィルタ[15]をこのパラメータ決定問題に適用することにより、パラメータベクトル a の不確かさを表す共分散行列 S は以下の式で与えられる。

$$S^{-1} = \sum_{i=1}^N M_i^T W^{-1} M_i,$$

$$M_i = \frac{\partial f_i}{\partial a}, \tag{3}$$

$$W_i = \frac{\partial f_i}{\partial x_i} \Lambda_i \frac{\partial f_i}{\partial x_i}^T.$$

ここで、 Λ_i は i 番目の計測データ x_i の不確かさを表す共分散行列である。

物体位置の見積りの確かからしさは共分散行列 S を用いて評価できる。ここで、 a の各要素の不確かさの、現在の組立動作の成功に対する相対的な影響が、扱われる物体の幾何形状に依存して変化することに注意しなければならない。相対的な影響度は、例えば S の各成分の重みつき和を用いて評価することも考えられるが、その重みも動作によって変化し、複雑な形状の物体に対しては適切な重みを計算することは容易ではない。そこで、物体位置のパラメータベクトルの誤差分布を組立動作に応じて適切に評価するための関数を、自動的に設定する方法を以下に提案する。

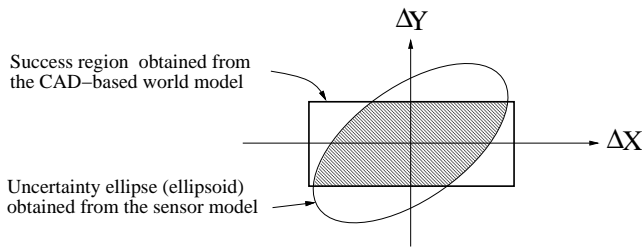


Fig. 12 Calculation of the predicted success probability. This figure shows the case where the position parameter of the object is two-dimensional, (X, Y) .

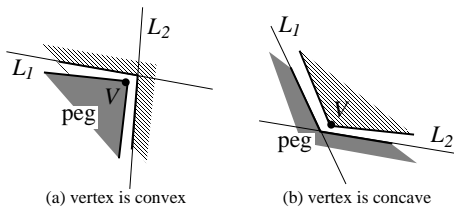


Fig. 13 Vertex (V) and edges (L_1 and L_2) used for generating inequalities in the case of insertion of a peg with a polygonal cross-section into a hole.

4.2 予測成功確率に基づいた評価

視覚情報で物体位置を補正して組立動作を行うとき、物体位置の計測精度がよければその動作は成功するが、計測精度が悪ければ失敗するであろう。したがって、ある視覚認識戦略の良さを測る直接的な尺度として、その戦略を用いて作業を行ったときに、その作業が成功するかどうか、というものが考えられる。そこで、我々は予測成功確率 (predicted success probability) を評価関数として用いることを提案する。

予測成功確率とは、ある視覚認識戦略を用いて組立動作を行ったときにその作業が成功する確率を、物体形状のモデルと視覚情報の不確かさのモデルを用いて計算したものである。その計算に当たっては、まず成功領域 (success region) と呼ばれる領域を、物体の位置のパラメータベクトルの誤差ベクトルの空間上で計算する。この領域は、誤差ベクトルがその領域内にあるときに、現在の動作が成功する領域として与えられる。次に、ある視覚認識戦略を採用したときのパラメータベクトルの誤差ベクトルの共分散行列 S を前節の方法により計算する。予測成功確率はパラメータベクトルの誤差ベクトルが、成功領域内に存在する確率として与えられる (Fig. 12参照)。この成功確率は、誤差ベクトル空間を量子化した後、成功領域内に存在する量子化された誤差が生じる確率の和として計算する。最終的な視覚センサ位置を求めるには、視覚センサの位置候補それぞれについてこの確率を計算し、予測成功確率が最高のもを選択する。

実際の応用に際しては、成功確率が 1.0 になるようなセンサ位置を選ぶことになる。そのようなセンサ位置の範囲が複数ある場合には、成功確率が 1.0 未満となるようなセンサ位置の範囲からできるだけ離れたセンサ位置を選ぶことにより、より信頼性のある視覚情報が得られると考える。

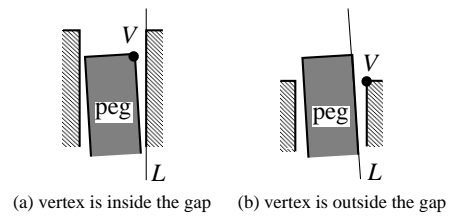


Fig. 14 Vertex (V) and edge (L) used for generating an inequality in the case of insertion of a peg with a polygonal cross-section into a parallel gap.

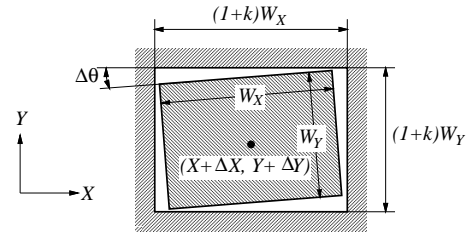


Fig. 15 Rectangular peg-in-hole operation.

4.3 成功領域を表す不等式の導出

成功領域は物体の CAD モデルから自動的に計算することができる。視覚情報を必要とする組立動作はすべて、ある種のペグ挿入動作とみなすことができる (Fig. 7参照)。また、本論文では、平面あるいは円筒面からなる物体の組立を対象としているので、成功領域の計算に当たっては次の 3 つの場合を考えれば十分である。

(1) 多角形断面を持つペグの穴への挿入 (Fig. 7の (a) と (b) の場合): ペグと穴の頂点の組それぞれに対して、2 つの不等式が得られる。ペグの頂点が凸なら、ペグの頂点と 2 つの穴の 2 辺との関係から、ペグの頂点が凹なら、穴の頂点とペグの 2 辺との関係から、それぞれ 2 つの不等式が得られる。Fig. 13は 2 つの場合を表す。 V, L_1, L_2 が計算に使われる頂点と辺である。

(2) 多角形断面を持つペグの溝への挿入 (Fig. 7の (c), (e) および (f) の場合): ペグと穴の頂点の組それぞれに対して、1 つの不等式が得られる。Fig. 13は 2 つの可能な場合を表す。 V と L が計算に使われる頂点と辺である。

(3) 円形断面を持つペグの穴への挿入 (Fig. 7の (d) の場合): 成功領域は挿入方向に垂直な平面上の円として与えられ、その半径は穴の半径とペグの半径の差である。

以上の計算は、動作前に拘束自由度のない、一般的な場合を扱っている。動作前に拘束自由度がある場合には、実際の成功領域は一般的な成功領域の、拘束された自由度での断面として与えられる。

4.4 成功領域の計算例

ここでは例として、長方形断面を持つペグを穴に挿入する動作について成功領域を計算する。Fig. 15はこの動作を上から見た図である。穴の辺の方向を X および Y 軸にとる。 W_x, W_y は辺の長さである。また、 k は穴のクリアランス比である。この動作では、ペグの位置と方向の 3 自由度、 (X, Y, θ) を調節しなければならない。

($\Delta X, \Delta Y, \Delta \theta$) を (X, Y, θ) の誤差ベクトルとすると、前節で述べた計算法により、8つの不等式の組が得られる。例えば、Fig. 15の右上の頂点については以下の2つの不等式が得られる。

$$\Delta X + \frac{W_X}{2} \cos \Delta \theta - \frac{W_Y}{2} \sin \Delta \theta \leq \frac{k+1}{2} W_X$$

$$\Delta Y + \frac{W_X}{2} \sin \Delta \theta + \frac{W_Y}{2} \cos \Delta \theta \leq \frac{k+1}{2} W_Y$$

得られた8個の不等式の組に対し、2組のデータセットについて実際の成功領域を計算した結果を Fig. 16に示す。(b)の場合を(a)の場合と比較すると、図からわかるように、 ΔX の許容範囲はより大きく、 $\Delta \theta$ の許容範囲はより小さい。もし、ペグの位置の誤差分布が両方の場合で等しければ、結果として出てくる成功の確率は異なるはずである。したがって、最適な視覚センサ位置を、成功領域と誤差分布を考慮した予測成功確率に基づいて決定することが必要である。

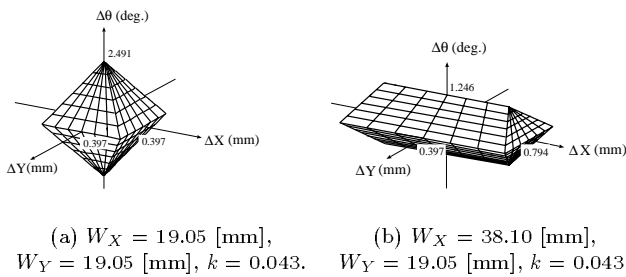


Fig. 16 The shape of the success region depends on the task.

5. ラインレンジファインダを用いた実装と実験結果

5.1 ラインレンジファインダと一般的な認識戦略

本論文で提案する視覚認識戦略生成手法を、ラインレンジファインダ[16]を用いて実現した。このレンジファインダはスリットレーザを照射し、物体表面の高輝度部分をTVカメラで検出することにより、線状の3次元位置データを得る(Fig. 17)。レンジファインダはRobotWorld[17]の可動プラットフォームに取り付けられ、3自由度の並進と垂直軸回りの回転を行うことができる。

視覚情報を必要とする組立動作はすべて、ある種のペグ挿入動作と見なすことができる。ペグの位置はその横の面を、穴の位置はその辺上で何点か(現在は5点)を観測することによって計測する。そこで、直線、円弧、平面、円筒面の4つの視覚特徴についてセンシング・プリミティブ(3.1節参照)を用意する。挿入方向に垂直な平面上での位置誤差が重要なので、各動作はFig. 18に示すように、レンジファインダを挿入方向に沿って移動させながら計測することにより観測する。穴を計測する際、各観測位置間の、挿入方向に沿った距離を一定にするように、レンジファインダを移動させた。また、計測対象がスリットレーザのある一定領域内に常に存在するように、レンジファインダの位置を制御し、その領域内では各計測点の位置の不確かさは一定であるとみなす。

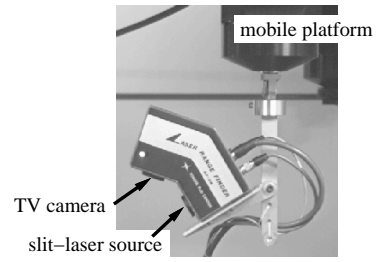


Fig. 17 Line laser range finder

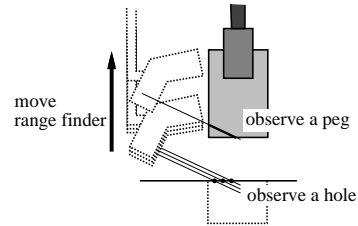


Fig. 18 A strategy for observing a peg and a hole.

5.2 視覚を用いた組立動作の誘導

実際の視覚を用いた組立動作の手順は以下になる。まず、ペグを穴の直前まで持ってくる。次に、レンジファインダを計画された位置に配置し、穴とペグの位置を計測する。穴とペグの相対位置の誤差が成功領域(4.3節参照)内に収まっていれば、そのままペグを挿入する。誤差が大きければペグの位置を修正し、ペグを再観測する。最後のステップを相対位置誤差が成功領域内に収まるまで繰り返し、そしてペグを挿入する。

5.3 レンジファインダの不確かさのモデル

我々の用いたレンジファインダは、奥行き方向の最大誤差が0.1[mm]以下、水平方向の最大誤差が0.3[mm]以下という、非常に精度のよい計測値を得ることができる。本論文では、視覚データの不確かさの、成功確率への影響を強調するために、計測値に比較的大きな正規ノイズを人工的に加えた。奥行き方向の計測値の分布の標準偏差を0.12[mm]、水平方向の計測値の分布の標準偏差を0.30[mm]とし、それら2つの分布は互いに独立になるようにした。Fig. 19は、このノイズを付加した状態で、3次元空間内の同じ点を500回計測した結果得られた計測値の分布を

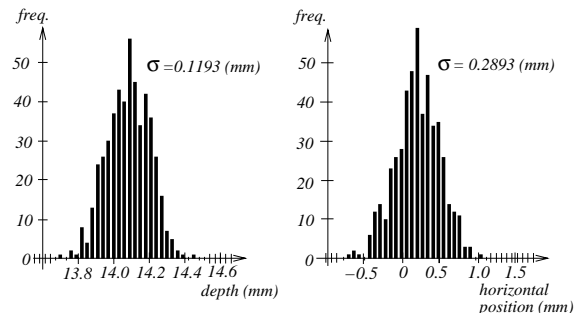


Fig. 19 Distribution of measurements and estimated standard deviation.

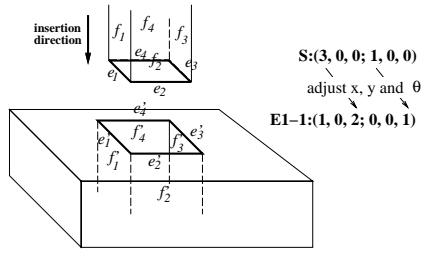


Fig. 20 Face contact analysis of the rectangular peg-in-hole operation. The sextuplet of DOFs (see Section 2.1) changes from (3, 0, 0; 1, 0, 0) to (1, 0, 2; 0, 0, 1).

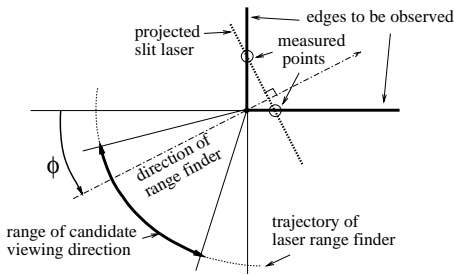


Fig. 21 Top view of candidate viewing directions for observing two edges and faces.

示す。得られた分布はほぼ適切な標準偏差を持つ正規分布に近くなっている。使用する視覚センサの不確かさのモデルが与えられれば、最適な視覚認識戦略が生成できることを検証することが本実験の目的であるので、このような人工的な不確かさのモデルを利用することにした。

5.4 実験結果

ペグ挿入動作、ねじ回しの先をボルトの頭の溝に挿入する動作、歯車のはめ合わせ動作の3種の動作について実験を行った。予測成功確率に基づく視覚認識戦略の評価法を有効性を確かめるために、各々の実験において、予測成功確率と、実際のロボットを用いて同じ動作を繰り返し行った結果得られた成功の割合とを比較する。

5.4.1 ペグ挿入動作

Fig. 20にペグ挿入動作の面接触状態遷移の解析結果を示す。この動作は Fig. 7の(a)に属し、並進2自由度および回転1自由度が拘束される。達成すべき面接触は $(f_1-f'_1)$, $(f_2-f'_2)$, $(f_3-f'_3)$ および $(f_4-f'_4)$ の4つである。また、視覚特徴の候補はペグについては f_1, f_2, f_3, f_4 であり、穴については e'_1, e'_2, e'_3, e'_4 のそれぞれ4つずつである。十分な位置決め情報を得るために、2つの隣り合った面あるいは辺を観測する。辺上の5点が必ず観測でき、レンジファインダとロボットとが衝突しないという条件から、可能な観測位置が Fig. 21に示すような、2辺の交点を中心とする円弧の一部に設定された。このとき、センサ位置を制御するパラメータは、スリットレーザの照射方向の水平成分と、穴の1つの辺とのなす角 ϕ のみである。Fig. 22に成功したペグ挿入動作を示す。

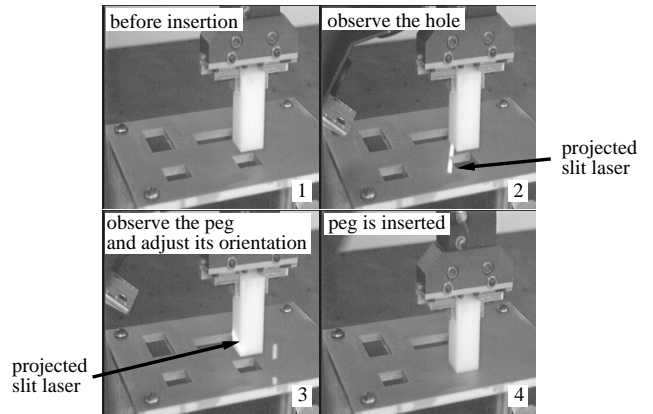


Fig. 22 A successful peg-in-hole operation.

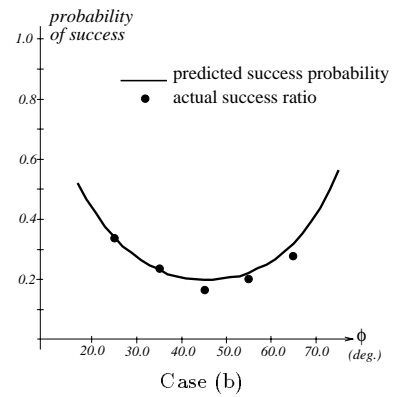
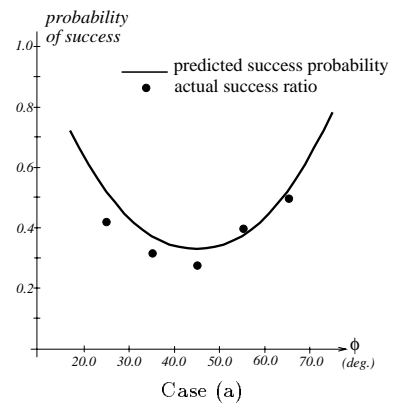


Fig. 23 Comparison of the predicted success probability with the actual success ratio.

5.4.2 ペグ挿入動作における予測成功確率と実際の成功割合の比較

次の2つのペグ挿入動作について、予測成功確率と実際の成功割合とを比較した。

Case (a): ペグの断面は $19.05[mm] \times 19.05[mm]$ の正方形。穴のクリアランス比は 0.043。この場合の成功領域は Fig. 16(a)に示されている。

Case (b): ペグの断面は $38.1[mm] \times 19.05[mm]$ の長方形。穴のクリアランス比は 0.043。この場合の成功領域は Fig. 16(b)に示されている。

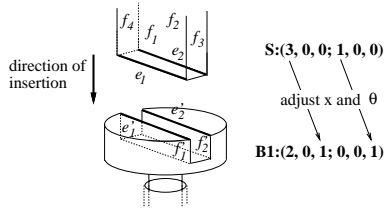


Fig. 24 Contact state analysis of putting a screwdriver on a bolt.

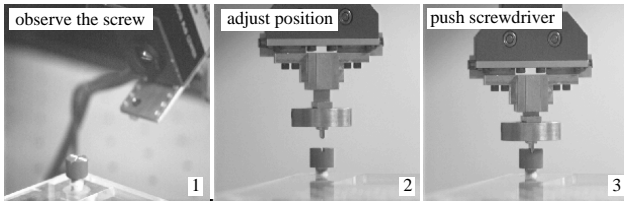


Fig. 25 The screwdriver was successfully inserted into the slot of the bolt head.

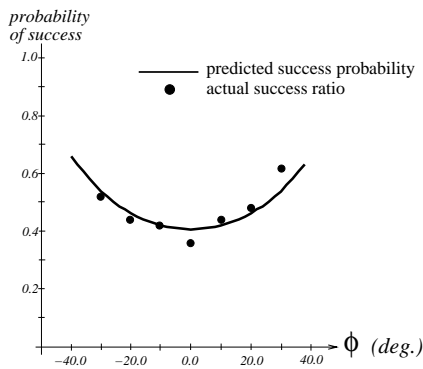


Fig. 26 Comparison of the predicted success probability with the actual success ratio in the screwdriver-bolt operation. The angle ϕ indicates the relative angle between the direction of the slot and the viewing direction of the range finder.

それぞれの場合で、いくつかの観測位置 (Fig. 21の ϕ) を選び、それぞれの観測位置について、レンジファインダでペグと穴の位置を計測して位置を修正し、実際のロボットで挿入するという作業を 50 回ずつ行った。うまくペグが挿入された場合を成功、挿入されなかった場合を失敗として、成功割合を計算した。

ペグの位置計測における測定点の数は、穴の位置計測の場合に比べて格段に多いので、ペグの位置の測定誤差は、穴の位置の測定誤差に比べてほとんど無視できる。そこで、予測成功確率の計算では穴の計測における誤差だけを考慮した。

Fig. 23に比較結果を示す。実線は計算された予測成功確率を、黒点は実験で得られた成功割合を示す。これら 2つが非常に良く合っていることがわかる。この結果は、ロボットの行う作業目的を考慮した認識戦略の評価が重要であることを示すものと考えられる。

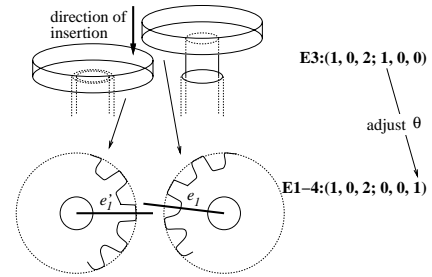


Fig. 27 Contact state analysis of gear mating.

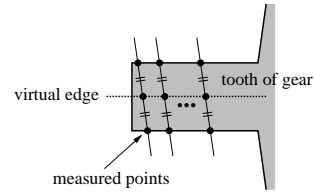


Fig. 28 Measuring the tooth position from edge positions.

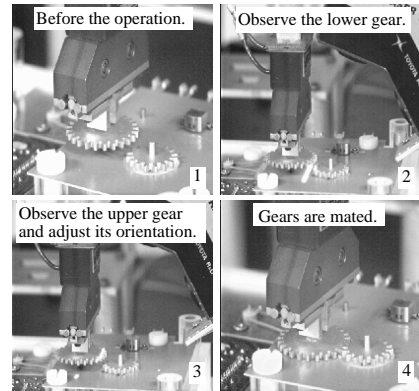


Fig. 29 The gears were successfully mated.

5.4.3 ねじ回しをボルトの溝に入れる動作

Fig. 24 にねじ回しの先をボルトの頭の溝に入れる動作の解析結果を示す。この動作は Fig. 7の (c) の動作に属し、並進 1 自由度と回転 1 自由度が拘束される。達成すべき面接触は $(f_1-f'_1)$ および $(f_2-f'_2)$ であり、視覚特徴の候補はねじ回しについては f_1, f_2, f_3, f_4 、ボルトについては e'_1, e'_2 である。

この動作では、ロボットとレンジファインダとの位置的制約から、ねじ回しの先とボルトの頭を同時に観測することはできなかった。そこで、ボルトだけを観測してねじ回しの位置と方向を修正した。Fig. 25に成功した動作を示す。また、Fig. 26に予測成功確率と実際の成功割合 (50 回試行) の比較結果を示す。この場合も、2つは非常に良く一致している。

5.4.4 歯車のはめ合い動作

Fig. 27に歯車のはめ合い動作の解析結果を示す。この動作は Fig. 7の (e) に属し、回転の 1 自由度が拘束される。歯車のはめあいでは歯の組合せが多数存在するので、以下のようにして方向を揃える。まず、Fig. 27に示すような 2つの仮想的なエッジ e_1, e'_1 を、2つの歯車の中心を結び直線にできるだけ近いとこ

ろに設定する．次にその2つのエッジが重なるように歯車の方向を調整する．

仮想的なエッジの位置は歯(あるいは歯と歯の間のギャップ)のエッジ上の点を計測することにより計算する(Fig. 28参照)．歯の形がほぼ長方形であるとみなし, エッジ点の組の中心点に直線フィッティングすることにより求める．

Fig. 29に成功した歯車のはめ合い動作を示す．Fig. 30は予測成功確率と実際の成功割合(50回試行)との比較結果である．この場合も, 実験結果は予測値と非常によく一致している．

6. 作業に関する知識を用いる視覚認識戦略生成システム

本節では, 本論文で提案する手法を基にした, 視覚認識戦略生成システムの構築法を示す．Fig. 31にシステムの概要を示す．対象とする組立作業についての知識と視覚センサのモデルから, 視覚認識作業モデル(sensing task model)と呼ばれる, 視覚認識戦略のテンプレートをあらかじめ生成しておくことにより, 計画生成時に実際の認識戦略を効率的に生成する．Figs. 22, 25, 29に示した視覚認識戦略はこのシステムによって生成された．

6.1 作業モデルとセンサモデル

各組立動作についての知識は作業モデル(task model) [12] [13] と呼ばれ, 各動作における接触状態の遷移, その遷移を実現する動作テンプレート, および動作実行に必要なパラメータから構成される．計画生成時には各パラメータに実際の幾何情報を代入し, ロボットの動作プログラムを生成する．必要な幾何情報は人間による作業の実行の観察結果から自動的に抽出される [12] か, あるいは, 作業プランナが作業の仕様記述から生成する．

センサモデル(sensor model) [18] は, センサによって検出できる特徴の種類, 距離や画角などの計測範囲, センサデータの不確かさといった, センサに関する知識を記述する．

6.2 視覚認識作業モデル

視覚認識作業モデルは以下の情報を持つ．

- 自由度の6つ組の遷移(2.3節参照)．
- 観測すべき自由度．
- 観測可能な視覚特徴の集合．

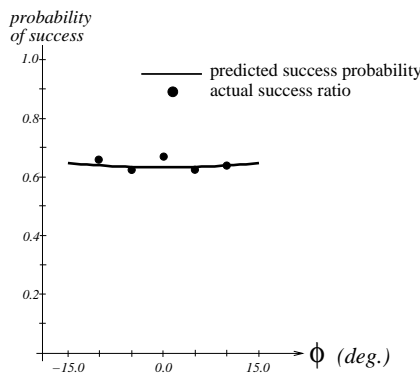


Fig. 30 Comparison of the predicted success probability and the actual success ratio in the gear-mating operation. The angle ϕ indicates the relative angle between the line connecting two gear centers and the viewing direction of the range finder.

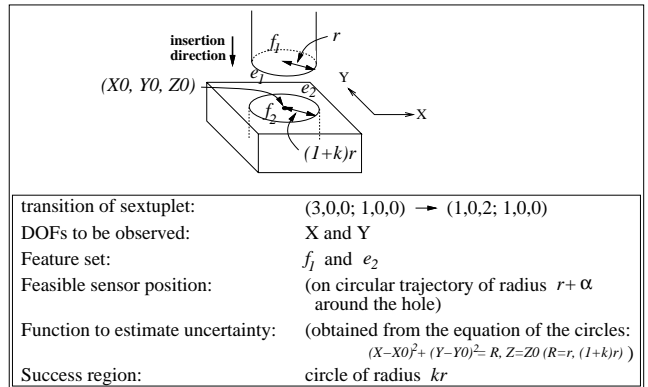


Fig. 32 A parameterized abstract sensing task model for the operation of inserting a peg with a circular cross-section into a hole. f_1 and f_2 indicate faces; e_1 and e_2 indicate edges. The radius of the cross-section of the peg is r ; α is determined so that there is no collision of the sensor; clearance ratio of this insertion is k .

- それらを観測できる視点の集合．
- 操作される物体の, 静止環境に対する位置の計測値の不確かさを見積もるための手続き．不確かさは位置パラメータの誤差ベクトルの共分散行列で表される(4.1節参照)．
- 認識戦略を評価するための成功領域(4.2節参照)．

上の情報の一部は, 物体の幾何形状に依存する．そこで, 効率的に視覚認識戦略を生成するために, 直方体や円柱といった典型的な形状を数え挙げ, それらに対して, 各辺の長さなどの幾何情報をパラメータとする形の視覚認識作業モデルを用意しておく．実行時にはパラメータに作業モデルから得られた実際の値を代入することにより, 効率的に視覚認識戦略を生成する．Fig. 32は円柱のペグ挿入動作についてのパラメータ化された視覚認識作業モデルの例である．

6.3 視覚認識作業モデルの生成手順

以下の手順で視覚認識作業モデルを生成する．

- (1) 面接触状態の遷移に基づく解析から, 視覚情報を必要とする動作を数え上げる(第2節参照)．
- (2) 各遷移について, 可能な物体形状を数え挙げ, 各形状に対しパラメータ化された視覚認識作業モデルを以下の手順で生成する．
 - (a) 必要な自由度を観測可能な視覚特徴の組を選び出す．
 - (b) 上の各組に対し, 視覚特徴の可視性およびセンサと観測される物体との衝突可能性を考慮して, 視覚センサ位置の候補を生成する．
 - (c) 物体の位置パラメータの不確かさを計算する関数を定義する．
 - (d) 成功領域を計算する．

6.4 視覚認識作業モデルを用いた視覚認識戦略の生成

計画生成時には, 視覚認識作業モデルから, 以下の手順で具体的な視覚認識戦略を生成する(Fig. 31参照)．

- (1) Instantiation: 現在の動作に適切な視覚認識作業モデルを選び出し, 幾何情報をパラメータに代入する．

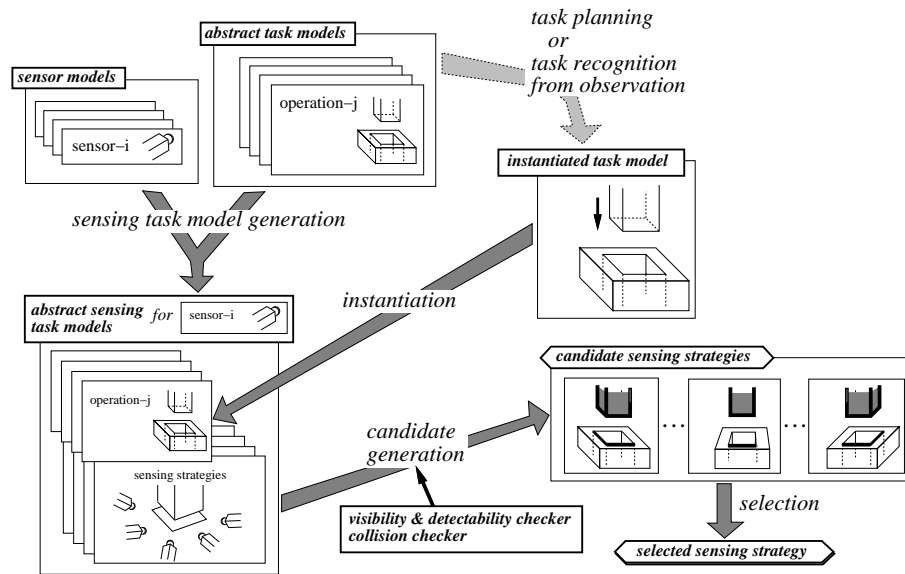


Fig. 31 Framework of task-oriented sensing strategy generation.

- (2) Candidate generation: 環境中のすべての物体を考慮して、視覚特徴の可視性やセンサの衝突可能性をチェックすることにより、実際に可能なセンサ位置の候補の集合を生成する。
- (3) Selection: 各視覚認識戦略候補を評価し、最適戦略を選ぶ。

7. おわりに

本論文では、作業に関する知識を用いることにより、最適な視覚認識戦略を自動生成する手法を提案した。組立動作の面接触状態の遷移に基づく解析結果を利用し、各動作を実行する際に視覚で観測すべき自由度を決定する。さらに、視覚特徴と観測可能な自由度に関する知識を用いて、観測すべき視覚特徴の組を生成する。最後に、予測成功確率に基づく評価を用いることにより、現在の動作を遂行する上で最適な視覚認識戦略を選び出す。

提案する手法をラインレンジファインダを対象センサとして実装した。実験結果から本手法の有効性、特に作業の目的を考慮して視覚認識戦略を決定することの必要性を示した。さらに、本手法を基にした、視覚認識戦略生成システムの構成について述べた。

第2章で述べたように、本論文で対象とする組立作業は、物体が面接触状態間を遷移するものに限られている。点接触や線接触を生じる場合については、接触状態グラフから接触状態の遷移計画を生成することにより[19]、力覚情報だけを用いて組立作業を行える場合がある。しかし、そのような場合においても、視覚情報を的確に利用することにより、例えば状態遷移数を減らすなどの利点が得られると考えられる。力覚情報と視覚情報を効果的に組み合わせて組立作業を的確に誘導する手法を、より広範囲の組立作業に対して開発することは今後の課題である。

また、ステレオ視覚など他の視覚センシング法に提案する手法を適用することも今後の課題である。

謝 辞

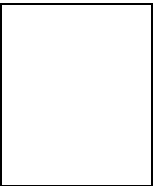
本研究は Carnegie Mellon University, The Robotics Institute において行った。実験装置について御協力頂いた川出雅人氏(現オムロン)、使用したレンジファインダについて御教示頂きました樋口和則氏(豊田中央研究所)に感謝します。また、研究にあたり御討論頂いた Task-Oriented Vision Laboratory の皆様に感謝します。

なお、第一著者の Carnegie Mellon University での滞在については(財)電気通信普及財団の援助を受けた。

参 考 文 献

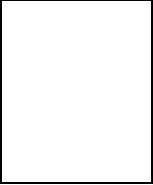
- [1] K. Ikeuchi and M. Hebert: "Task Oriented Vision", *Proc. of Int. Conf. on Intelligent Robots and Systems*, pp. 2187-2194, 1992.
- [2] R. Bajcsy: "Active Perception", *Proc. of IEEE*, Vol. 76, No. 8, 1988.
- [3] Y. Aloimonos, editor: "Active Perception", Lawrence Erlbaum Associates, Inc., New Jersey, 1993.
- [4] S.A. Hutchinson and A.C. Kak: "Planning Sensing Strategies in a Robot Work Cell with Multi-Sensor Capabilities", *IEEE Trans. on Robotics and Automat.*, Vol. RA-5, No. 6, pp.765-783, 1989.
- [5] R.D. Rimey: "Where to Look Next using a Bayes Net: An Overview", *Proc. of 1992 DARPA Image Understanding Workshop*, pp.927-932, 1992.
- [6] 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.
- [7] I. Horswill: "Polly: A Vision-Based Artificial Agent", *Proc. of AAAI-93*, pp. 824-829, 1993.
- [8] S. Sakane and T. Sato: "Automatic Planning of Light Source and Camera Placement for an Active Photometric Stereo System", *1991 IEEE Int'l Conf. on Robotics and Automat.*, pp.1080-1087, 1991.
- [9] K.A. Tarabanis, P.K. Allen, and R.Y. Tsai: "A Survey of Sensor Planning in Computer Vision", *IEEE Trans. on Robotics and Automat.*, Vol. 11, No. 1, pp.86-103, 1995.

- [10] H. Zhang: "Optimal Sensor Placement", *Proc. of 1992 IEEE Int. Conf. on Robotics and Automat.*, pp.1825-1830, 1992.
- [11] K. Kemmotsu and T. Kanade: "Sensor Placement Design for Object Pose Determination with Three Light-Stripe Range Finders", *Proc. of 1994 IEEE Int. Conf. on Robotics and Automat.*, pp.1357-1364, San Diego, CA, 1994.
- [12] 池内, 末広: "視覚による組立作業理解のための作業モデルとそれに基づく動作生成", *日本ロボット学会誌*, Vol. 11, No. 2, pp.281-290, 1993.
- [13] K. Ikeuchi and T. Suehiro: "Toward an Assembly Plan from Observation Part I: Task Recognition With Polyhedral Objects", *IEEE Trans. on Robotics and Automat.*, Vol. 10, No. 3, pp.368-385, 1994.
- [14] J. Miura and K. Ikeuchi: "Task-Oriented Generation of Visual Sensing Strategies in Assembly Tasks", Technical Report CMU-CS-95-116, School of Computer Science, Carnegie Mellon University, February 1995.
- [15] N. Ayache and O.D. Faugeras: "Maintaining Representations of the Environment of a Mobile Robot", *IEEE Trans. on Robotics and Automat.*, Vol. RA-5, No. 6, pp.804-819, 1989.
- [16] 樋口, 小関, 塚田, 山本, "形状計測用3次元視覚センサ", *精密工学会春季講演会論文集*, pp.949-950, 1989.
- [17] V. Scheinman: "A Multiple Robot Vision Guided Assembly System", in R. Bolles and B. Roth, editors, *Robotics Research 4*. The MIT Press, 1987.
- [18] K. Ikeuchi and T. Kanade: "Modeling Sensors: Toward Automatic Generation of Object Recognition Program", *Computer Vision, Graphics, and Image Processing*, Vol. 48, pp.50-79, 1989.
- [19] 横小路: "接触状態の分類と組立手順", *日本ロボット学会誌*, Vol.11, No.2, pp.185-191, 1993.



三浦 純 (Jun Miura)

1962年3月29日生。1984年東京大学工学部機械工学科卒業。1989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, 工学博士。同年, 大阪大学工学部電子制御機械工学科助手。知能ロボットの研究に従事。1994年~1995年, CMU 客員研究員。人工知能学会, 電子情報通信学会, 情報処理学会, IEEE, AAAI各会員。(日本ロボット学会正会員)



池内克史 (Katsushi Ikeuchi)

1949年5月29日生。1973年京都大学工学部機械工学科卒業。1978年東京大学大学院情報工学博士課程修了, 工学博士。MIT人工知能研究所, 電総研を経て現在カーネギーメロン大学計算機科学科, ロボティクス研究所, ヒューマンコンピュータインタラクティオン研究所兼任, 研究教授。明るさ解析, 物体認識, ハンドアイシステムの研究に従事。IROS-95 実行委員長, CVPR-96 プログラム委員長。IJCV, JOSA 等編集委員。マー賞等授賞。(日本ロボット学会正会員)