

3次元視覚によるプログラミングレス マニピュレーションシステムの開発

富田 文明[†] 松下 俊夫[†] 河井 良浩[†]

産総研では、多分野・多目的に利用できる高機能3次元視覚システム VVV (Versatile Volumetric Vision) を開発している。3次元視覚として、距離計測、形状計測、物体認識、運動追跡が主要な処理であるが、VVVは、多様な状況で任意の形状の物体を対象として、これらの処理を一貫的に実時間で高精度に実行することができるソフトウェアシステムである。VVVの応用システムの1つとして、各種物体の各種操作の基本動作を体系化、モジュール化することにより、専門家によるプログラミングを必要としないプログラミングレスハンドアイシステムの実現を目指している。本稿では、3次元視覚を用いた物体操作の実験システムとして、人が操作する物体の移動方法を見て覚える物真似システムのほか、自動システムとして、空き缶等のピンピッキングシステム、円柱部品のペグインホールシステム、ブロックの組み立てシステム、作業移動型システムとしてヒューマノイドによるパネル運搬作業等を紹介する。

R&D on Programming-less Manipulation Systems Using Three-dimensional Vision

FUMIAKI TOMITA,[†] TOSHIO MATSUSHITA[†] and YOSHIHIRO KAWAI[†]

We are developing an advanced 3-D vision system VVV (Versatile Volumetric Vision) which can be used for many purposes in many fields. Range sensing, shape description, object recognition and motion tracking are the principal functions of 3-D vision. The VVV is a software system which can implement these functions consistently and accurately in real time. Using VVV, we are also developing programming-less hand-eye systems which need not any professional robot programmer by systematically have a set of various elementary manipulation modules for various kinds of objects. In this paper, we introduce such experimental automatic hand-eye systems as a bin-picking system which selects, for example, empty cans among other objects, a peg-in-hole system which inserts cylindrical parts into moving holes, an assembly system which combines scattered blocks into a structure, and a humanoid which works to carry panels as a system of integrated locomotion and manipulation, as well as interactive systems which teach robots how to move objects by showing.

1. ま え が き

産業用ロボットの第1世代はプレイバック方式で、スプレー塗装のように、危険だが正確さをあまり要求されない作業に応用された。しかし、大規模なシステムに取り込むことは困難であった。第2世代はその動作をプログラムすることができるロボットで、VALのようなロボット言語が開発された。この第2世代のロボットは、精度と融通性に非常な改良をもたらし、現在溶接や組み立て等の作業に広く利用されている。しかし、実際にはプログラマが起こりうるすべての状況

を想定しなければならないので、ロボットは注意深く設定された環境でしか作業できていない。現在目標とされているロボットは、いろいろな環境の中で、自分で判断して自律的に行動する知能ロボットである。このような知能ロボットを実現するために必要な機能がロボットの眼 - 視覚である。

産学ともに従来から開発されている大多数の視覚(画像処理)システムは、カメラ1台(単眼視)で対象を平面的に扱う2次元視覚であり、半導体等高速な処理が必要な大量生産型システムには有効であったが、立体的な対象を扱う場合は、原理的に情報不足で不良設定問題となり、環境や対象に専用的な条件付けが必要であることから、現在の変種変量生産型システムには機能的にも経済的にも対応できなくなっている。

[†] 独立行政法人産業技術総合研究所
National Institute of Advanced Industrial Science and
Technology (AIST)

結果として、生産ラインの全自動化ができず、自動化のシンボルであったはずの産業用ロボットが、人間にとってかわられる例も出てきている。ひいては、人件費削減のため、工場を低賃金の海外へ移転する大きな理由の1つとなっている。来たる高齢化、少子化時代の労働者不足に対処するためにも、事は緊急であると考えられる。この問題を解決できるのが、立体を立体として処理することにより、相対的に情報充足、良設定問題となり、個別的な条件付けが不要で、結果として汎用的、共通的に利用できる3次元視覚技術であり、実際、多くの現場からその早期実現が求められている¹⁾。

3次元視覚技術を使った産業用システムの例もあるが、対象物(ワーク)の形状に制約があり、ワークにある丸穴(円)のような単純な平面的特徴を扱う程度にとどまっております²⁾、ワークが変わった場合は設定・調整に熟練を要する、あるいは対応できないのが現状である。さらに(社)日本ロボット工業会の市場予測では、今後のロボットは、その活躍の場をFA分野と同等に、家庭や病院等の生活分野へと広がり、今まで以上に、自律性や汎用性が要求されることになる。

産総研では(旧電総研時代から)10年以上にわたり、多分野・多目的に利用できる高機能3次元視覚システムVVV(Versatile Volumetric Vision)を開発している³⁾。3次元視覚として、距離計測、形状計測、物体認識、運動追跡が主要な処理であるが、VVVは、これらの処理の対象として、観測条件を限定することなく任意の形状の物体を扱うことができる点(汎用性)、各処理を実時間で高精度に実行することができる点(高性能)、各処理をつながりよく1つのシステムの中にまとめている点(一貫性)等、他に競合するシステムがない機能の高さと適用範囲の広さを有するソフトウェアシステムである。VVVは、人間の眼が必要とされる多くの作業や機械に共通的に利用でき、その自動化を促進することが期待される。

VVVの応用システムの1つとして、専門家によるプログラミングを必要としないプログラミングレスマニピュレーションシステムTORCSの実現を目指している¹⁰⁾。このシステムでは、種々の物体に対する多様な操作に対して基本となる操作を抽出・体系化し、それらの機能を実行するプログラムをモジュール化しておくことで、各種の作業プログラムの半自動的な生成を行う。

本システム開発の基本方針として、当初は過度に複雑な軌道計画を必要としない環境で、付録にあるできるだけ多くの種類の基本的な動作⁴⁾ができる実験システムを実現しながら、プラグマティックに必要な知識

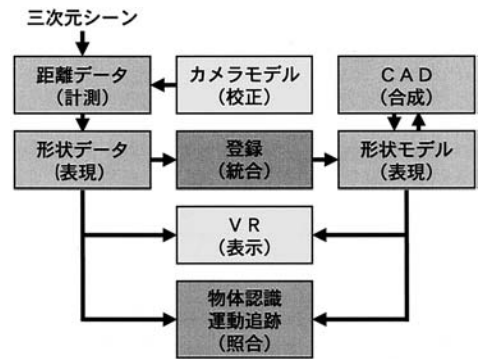


図1 3次元視覚プロセス
Fig. 1 3-D vision process.

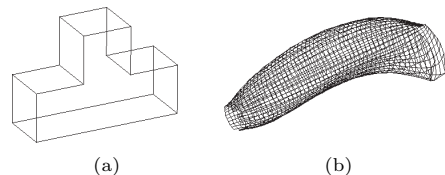


図2 形状表現: (a) ブロックの境界表現, (b) バナナの曲線ネット表現

Fig. 2 Geometrical shape models: (a) Boundary representation of a block, (b) Curved-line net representation of a banana.

と機能を蓄積することから始めている。本稿では、これまで開発した実験システムを紹介する。

2. 3次元視覚システム

3次元視覚の基本プロセスを図1に示す。まず、3次元視覚センサによって、観測している環境(以下、シーン)の3次元距離を計測する。次に、この距離データの構造的解析からシーンに存在する物体の面や境界線の3次元形状を計測する。この形状データは、CADやVRに利用できる表現である。さらに、対象とする物体の形状データ(モデル)をあらかじめ用意しておけば、形状モデルと形状データとを照合することにより、その物体の存在の有無と、存在する場合はその3次元の位置と姿勢を決定する。いわゆる物体認識である。また、物体が動いている場合は、照合処理を継続的に繰り返すことによって、その3次元運動を追跡する。物体のモデルの登録には、3次元CADによって合成する方法と、自然物等CADでは合成が難しい場合には、実物の3次元形状を計測し、その形状データを形状モデルとして利用することができる。その基本データ構造は、CAD分野で標準的な境界表現(b-rep)で、輝度情報も付加される。曲面の距離データが得られている場合には、曲線ネットによって曲面

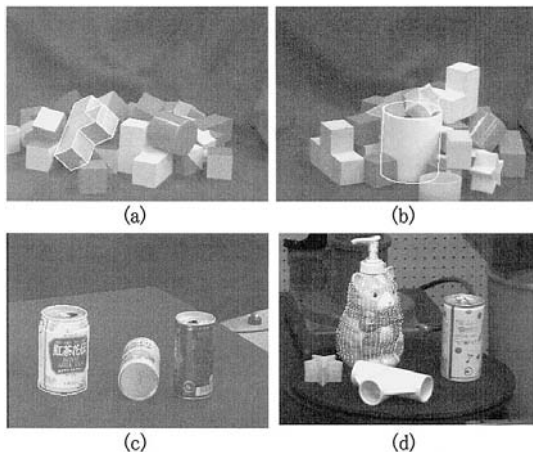


図3 物体認識：(a) 他の物体と混在するブロック，(b) 部分的な隠れのあるコップ，(c) 模様のある空き缶，(d) 自由曲面体の陶器（白線が物体のモデルを表し，その物体の3次元の位置と姿勢が認識されたことを示す）

Fig. 3 Object recognition: (a) Building block mixed with other objects, (b) Cup that is partly occluded, (c) Empty can with various patterns, (d) Ceramic product with free forms (White lines represent object models and indicate the 3D positions and attitudes have been recognized).

形状を表現する（図2）。

3次元視覚センサとして，複数台のカメラを用いる受動的なステレオビジョンと，レーザ光等を物体に投影する能動的なレンジファインダのいずれも利用可能であるが，標準的にはステレオビジョンを基本とし，レンジファインダは確実な距離測定が要求される場合の最後の手段と位置付けている．これは，人間の視覚がそうであるように，ステレオビジョンでたいていのはできるはずであり，工学的にできるようにすることが筆者らの目標であることと，実用的にも簡便性，融通性，汎用性，安全性等の観点から優位にあると考えるが，一般的には用途に応じて使い分ければよい．

物体認識は，通常，対象物の境界線（曲面の見かけの輪郭線を含む）の照合だけで十分であり，境界線の頂点（分岐点，屈曲点，変曲点，遷移点）を局所的な初期照合に用いることにより，多様な状況で各種の物体の認識が可能となる⁵⁾．具体的には，任意の3次元形状（多面体から自由曲面体まで）でよい，表面に模様，光沢があってもよい，任意の3次元位置・姿勢（6自由度）でよい，不特定の背景に他の不特定の物体と混在してもよい，部分的に隠れて見えなくてもよい等の汎用性の条件の基本はクリアしている（図3）。

本システムは，標準的なUNIX（Linux）コンピュータに実装でき，処理時間は，シーンや物体の複雑さに依存するが，PC（Pentium III 866 MHz）上で，標

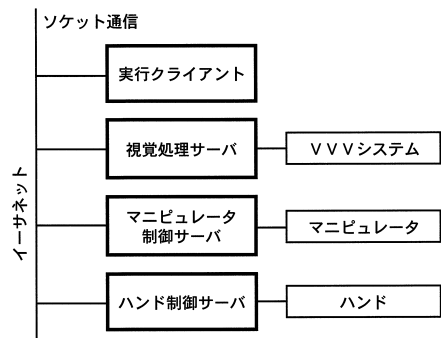


図4 実験システムの構成図

Fig. 4 Hand-eye system configuration.

準オプションでの距離・形状計測と物体認識が1～数秒，運動追跡が数ミリ秒のオーダーである．測定精度は，使用するカメラの分解能，レンズの焦点距離，カメラ間距離，そして，対象までの距離によって定まる．したがって，要求仕様を満足するそれらの組合せを事前に選ぶことができる．参考として，市販のカメラ（ソニーXC-75）に焦点距離16mmのレンズを付け，カメラ間距離を30cmとした場合，2m離れた物体に対して，位置精度が0.5mm，姿勢精度が0.5度程度である．

本システムをロボットの眼として利用することにより，整備されていない環境，刻々変化する環境においてもロボットによる自律的物体操作が可能となる：

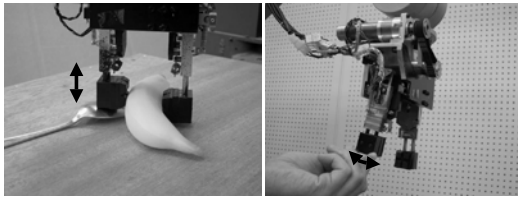
- ロボットの作業環境の3次元幾何モデルが自動的に随時得られることから，ロボットの軌道計画ができる．
- 任意の対象物のモデルを用意するだけで，環境中の対象物の存在とその3次元の位置と姿勢が得られることから，任意の場所に置かれた対象物の把持等の操作ができる．
- 移動する対象物の3次元運動を実時間（ビデオレート）で追跡できることから，ビジュアルフィードバックやエラーリカバリーができる．

3. ハンドアイシステム

ハンドアイシステムの基本構成を図4に示す．

3.1 マニピュレータ

ハンドアイシステムの開発の基本方針として，異機種のマニピュレータを共通の作業プログラムから利用できるように，たとえば，手先の位置とハンドの姿勢角を与えてハンドを移動する等，作業プログラムで必要な基本的マニピュレータ制御命令を抽出して，中間命令セットを作成している．中間命令は，作業プログラムのコンパイル時（あるいは実行時）にマニピュ



(a) (b)

図5 衝突吸収ハンド

Fig. 5 Collision absorption hand.

レータに固有の命令に変換される。

制御サーバとマニピュレータコントローラ間の通信は、RS232Cを用いるもの、独自のネットワーク方式を用いるもの等、機種ごとに異なる。異なる機種のマニピュレータを使用する環境でも、これらの差異も吸収し、ネットワーク上のどの計算機からも任意のマニピュレータが同一命令で制御できるように、ユーザ、あるいは、作業プログラムがクライアントとして発行する中間命令を、マニピュレータの制御サーバで翻訳する、クライアント・サーバ方式のマニピュレータ制御を行うようにしている。サーバとクライアント間はソケット通信を用いている。

3.2 ハンド

本稿で紹介するハンドアイシステムで標準的に利用するハンドとして、平行2指ハンドを選択しており、またハンドには触覚あるいは力覚センサを装着していない。これは、大部分の作業が2指ハンドで実行できると考えられることと、視覚だけで実行可能な作業の機能整備を優先するという方針による。ただし、視覚では検知できないミクロ的な衝突を回避するために、衝突吸収ハンドを開発し、利用している。

このハンドは、各指に開閉方向と垂直な2軸の方向の衝突を吸収するバネ機構を取り付け、衝突時にハンド（指）と抵抗物の双方に無理な力が加わらない仕組みとなっている（図5）。各方向に1cm程度可動であり、精度が1mm程度の標準的な視覚系には十分な余裕がある。このような簡単な機構だけで、多少の衝突が許容されることによって物体操作の自由度が増し、従来のように衝突回避のために触覚（力覚）センサを用いて複雑な制御系を構成する必要がなくなるメリットは大きい。さらに、逆に衝突を利用する作業も可能となる。たとえば、押印作業⁶⁾を視覚センサだけで容易に実現することができている。

3.3 視覚系

3次元視覚センサは、3台のカメラを1セットとするステレオカメラシステム—ステレオビジョンが基本である。カメラの種類や配置は用途に応じて自由に

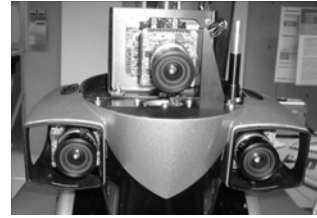


図6 3眼ステレオカメラシステム

Fig. 6 Trinocular stereo camera system.

設定できる。たとえば、図6は後述するヒューマノイド HRP-2¹⁹⁾の頭部に搭載されている3眼ステレオカメラシステムである。3眼である最大の理由は、2眼では不可避なエピソード線と平行な部分の対応の曖昧性を除去するためである。

ステレオカメラとマニピュレータの配置は、目的や状況に応じて自由に設定できる。本稿で紹介する主な実験は対向型の配置で実施している（図8）。その他、ヒューマノイドのように人間と同様な並列型や、アームに視覚系を搭載したアイハンド型でもよい。

ステレオビジョンによる距離計測の準備として、カメラパラメータを測定するカメラキャリブレーションがあり⁷⁾、その精度が距離精度に大きく影響する。標準的なピンホールカメラモデルに基づくカメラパラメータには、カメラの位置と姿勢の外部パラメータ（6自由度）とカメラの焦点距離、分解能等の内部パラメータ（6自由度）がある。これらのカメラパラメータは、その3次元位置が既知の指標（たとえば、×印）とその画像上の2次元位置の対応付け（6点以上）から計算することができる。ハンドアイシステムの場合、正確な位置制御ができるマニピュレータを利用して、ハンドに指標を持たせるか、アームに指標を貼ってマニピュレータを移動し、その画像上の位置を計測する方法が便利である。カメラキャリブレーションと同時に、ハンドアイシステムにおけるカメラの座標系とロボットの座標系のコーディネーションが同時に行える利点がある。

4. 作業の記述

実行クライアントでの作業の記述として、作業フレーム、把持フレーム、監視フレームについて述べる。

4.1 作業フレーム

作業計画と実行コントロール部はオブジェクトのクラスの集合として作成している。たとえば、すべての作業の基本である Pick & Place の場合、対象物の検出と初期計測の後、それぞれの対象物の移動は、次のような基本サイクルを繰り返すことで実行される。

把持計画 把持位置へのハンドの移動
 把持 持ち上げ ハンド姿勢変更
 ゴールまたはサブゴール位置への運搬
 把持物体を解放 ハンドを元に戻す

これらの基本作業要素に、各動作の実行前後の確認行動要素、異常を検出した場合の修正行動要素を加えたものが作業の構成単位となる。この状態遷移に関する規則と、作業に固有の環境変数（テーブルの高さ、許容位置誤差、移動経路等）および物体モデルが作業記述を表す。

4.2 把持フレーム

物体モデルとして、幾何モデルに加えて、物体の把持位置とハンドの把持姿勢を、物体座標系に対する座標変換として、あらかじめ物体ごとに記述する。たとえば、円柱の場合、直立状態ではZ軸は上向き、X軸はロボットベース座標中心を向く方向とし、水平状態ではZ軸はロボットベース座標中心を向き、X軸は水平とする。

4.3 監視フレーム

対象物の検出と初期計測は、認識機能で行い、実行時の監視は追跡機能^(8),9)を用いる。追跡機能は、動きのある操作対象や環境を監視してロボットに確実に作業を行わせるために有用な技術である。単に低レベルの制御指標として追跡データを使用するのみならず、実行時に作業の流れを修正するような高レベルの制御にも追跡機能を利用する。作業課題に容易に適應できる柔軟性と自律性をシステムに与えるためには、何のような状態で追跡しているかをシステムがつねに知る必要がある。物体記述および作業記述に基づき作業の各局面で追跡機能を利用することで、一連の物体操作作業をロバストに実行するシステムを構築している。認識機能と関連付けられた追跡機能の用途には以下のようなものが考えられる。

- 状況の事前確認：一度認識された物体に対して、その位置で追跡を試みることにより状況を確認する。
- 実行中のモニタ：操作中の物体を追跡し、予測された位置・姿勢にあることを監視する。
- 動きのあるものの監視：移動物体を追跡し、特定位置や姿勢への到達を監視する。
- 動きのある対象への追従：移動物体を捕捉したり、移動対象へのはめあい等の作業に使用する。

作業はオブジェクト表現された作業モデルのメソッドを呼び出すことで実行される。初期認識の後、追跡を用いた作業の実行管理は、選択された各作業ステップに対して以下のように行われる。

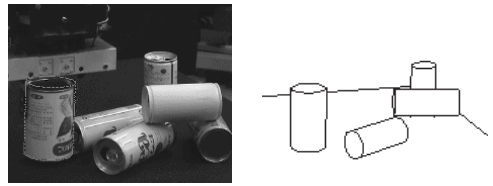


図7 空き缶の認識

Fig.7 Recognition of empty cans.

- (1) アーム動作 + 追跡
- (2) 物体の位置・姿勢の修正
- (3) 追跡異常が発生した場合
 - (3-1) アームモデルから物体位置を推定して再追跡
 - (3-1-1) 追跡成功 動作継続
 - (3-1-2) 追跡異常 初期認識に戻る
- (4) イベント検出 動作終了

5. ピンピッキング

異種、同種を問わず、散在するあるいは山積みされた多くの物の中から特定の対象物を拾い上げるピンピッキングは、生産工程で自動化が要求される典型的な作業である。実験システムとして、空き缶を拾ってゴミ箱に捨てる「掃除システム」を作成した。

認識処理は、登録された複数種類の空き缶のモデルを順次適用し、存在する（閾値以上の照合度を有する）空き缶の位置姿勢を（照合度順にソートして）出力する（図7）。拾い上げ（Pick）動作では、ソート順に空き缶を選択し（結果として上側にある空き缶を優先し）、各空き缶の最上位の把持点の真上から垂直方向にアプローチした。拾い上げた後は、定位置にあるゴミ箱の上まで移動し、ハンドを開放して空き缶をゴミ箱に捨てる。

本作業では、空き缶を拾い上げる際に接触等で他の空き缶の位置が変わってしまう場合があり、各空き缶のPick動作開始前にその存在を確認するために、追跡機能を使用した。移動が小さい場合は、追跡機能で位置を修正することができるが（フィードバック）、大きな移動で見失った場合は、認識処理から再実行する（エラーリカバリ）。

また、認識した空き缶をすべて拾った後、認識処理を再実行し、以前は隠れて見えなかった空き缶を認識すれば、同作業を繰り返す。

本ピンピッキング作業は、ハンドで把持できる大きさと重さであれば、モデルを入れ替えるだけで、任意

ビジュアルフィードバックの一種と見なせるが、追跡機能を利用して動作の停止や開始を行う。

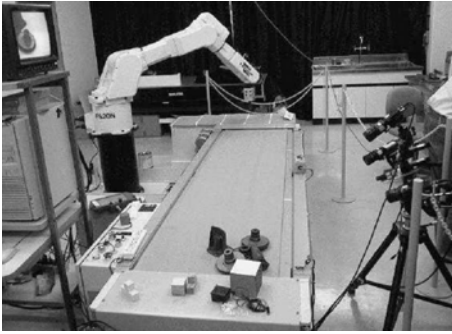


図 8 鋳物選別システム

Fig. 8 System to select castings.

の対象物に適用できる。たとえば、図 2 (b) のバナナのモデルを使えば、図 5 (a) のようにバナナを認識して拾うことができ、用意された皿に置く「給仕システム」も作成した。

このほか、鋳造業の生産工程の模擬実験 (図 8) として、コンベア上を移動するランダムに置かれた鋳造品をコンベアの速度に合わせてマニピュレータを移動しながら動的に Pick する「選別システム」を作成した¹⁰⁾。また、空き缶拾いシステムでは、特定の空き缶の幾何モデルを用いるのではなく、空き缶の共通モデル (Generic Model) として「円筒」を用いることにより、任意の空き缶を認識し拾うことができるシステムも開発した¹¹⁾。

6. ベグインホール

2つの対象物の相対的位置関係を認識する必要がある基本動作の1つとして、部品を部品の断面と同形の穴に挿入するベグインホールがある。本章では、直径49mmの円筒を直径50mmの丸穴に挿入する実験例について述べる。

6.1 基本操作

- (1) 円筒の把持は5章と同様に行う。
- (2) 穴の位置を認識処理により計測し、穴近傍へハンドを移動する。
- (3) ハンドと把持物体の相対的位置関係は(1)の把持時に多少ずれるので、ハンドと把持物体との相対関係を得るため、円筒の位置を再計測し、円筒の位置の微調整を行う。
- (4) 最後に、穴への接近と挿入動作を行う。

視覚系の距離計測精度を0.5mmに設定しているので、穴上空からの直接の挿入が可能であった。

追跡物体やハンドの位置・姿勢によって追跡対象のカメラから可視な部分はずねに変化する。追跡処理は、かなりのオクルージョンが生じても部分輪郭を手が

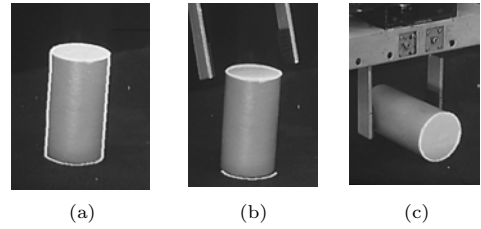


図 9 モデルの選択

Fig. 9 Selective models.

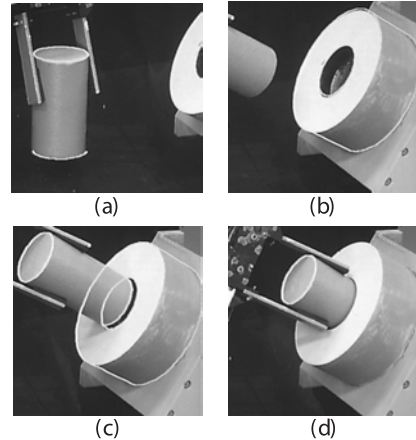


図 10 ベグインホールの作業監視

Fig. 10 Monitor of a peg-in-hole task.

りに追跡が可能であるが、高速化のために認識処理より局所的な処理に基づいているため、他の物体の輪郭線と密着していると動作が安定しない場合がある。そこで本システムでは、物体の属性と作業記述の両方で、物体の姿勢に応じて適用すべき追跡用形状モデルを指定するようにした。図 9 は、白線部は追跡用モデルを入力画像にオーバーレイしたもので、(a) は初期認識で使用されるモデルで追跡をかけた状態を、(b)、(c) では誤動作を与える可能性のある輪郭部を除いたモデルを適用している状況を示す。

6.2 動的な操作

図 10 (a) は初期位置からの持ち上げ動作時の確認、(b) は移動時における目標物体の存在状況監視を行っている様子を示す。当視覚系は、複数の物体の移動を同時に追跡することができるので、実験では、挿入前に把持物体と穴のある物体の双方を追跡対象として最終位置合わせを行っている (図 10 (c))。したがって、この状態で穴を動かしても円柱を正しい姿勢で穴に追従させることができる。また、イベント検出の例として、図 10 (d) では、所定の位置まで確実に挿入動作を行うため、円柱上面の穴の入り口からの距離を監視している。

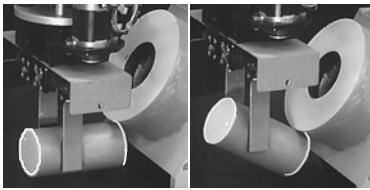


図 11 把持変更の監視
Fig. 11 Monitor of regrasping.

6.3 把持変更

イベント検出の他の応用として、ハンドに把持された物体の姿勢変更を行った。マニピュレータの可動範囲の制限から、指定された把持姿勢では挿入開始点へ移動できないことがある。この場合、挿入が可能になるように途中でハンドと物体の相対姿勢を変えるのが、1つの解決法である。図 11 は、穴物体の下部で、円筒の一端を穴物体に接触させながら持ち上げて握り方を変えている様子を示す。円筒の傾きがある範囲に達すると握り修正動作は停止して、挿入動作に復帰する。

7. 組み立て

組み立て (Assembly) は、基本的動作の組合せ、マクロな (複合) 動作 - 作業である。

実験システムでは、図 12 のような 3 種のブロックの組合せを作業例とした。Pick 動作と Place 動作の繰り返し操作である。本作業では「持ち替え」を主要課題としている。一般的に、マニピュレータとハンドの機構の自由度によっては、Pick 時の把持状態から Place 時の把持状態に直接移行できない場合があり、間に持ち替え動作が必要となる。

ハンドによる把持物体の持ち替えについては、多指ハンドによる物体の操りや、指先の回転機構等を利用するもの¹²⁾等の研究がある。これらは、ハンド指と物体との力学的つり合いの解析や、機構の提案を主としている。視覚の利用に関しては、多指ハンドの操り時に画像を利用したもの¹³⁾や作業目的を考慮した視覚認識計画の試み¹⁴⁾等があるが、本システムでは、視覚を用いることにより持ち替え操作を含む作業全体を安定に行う方法の実現を目的とした¹⁵⁾。

実験では、図 13 のように、3 種のブロックの組合せの順序、接近方向は既知とする。また、図 14 のように、それぞれのブロックの安定姿勢と、最終姿勢に至る姿勢遷移、およびそれぞれの姿勢においてブロックのどの部分を握るかについても、あらかじめ与えるものとする。

7.1 把持計画

(1) 観測された姿勢が最終姿勢であれば、その姿勢

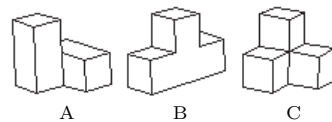


図 12 使用ブロック
Fig. 12 Building blocks used.

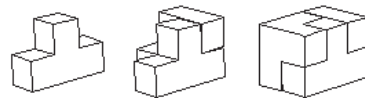


図 13 組み立て手順
Fig. 13 Order of assemble.

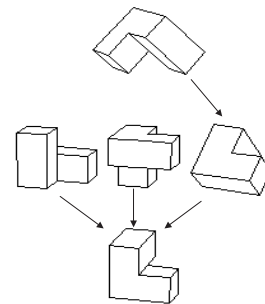


図 14 姿勢遷移
Fig. 14 Transitions of attitudes.

での移動 (Pick & Place) 計画を行う。すなわち、物体に指定された把持部分に対して、初期状態と最終状態 (移動目標点) との双方で可動かつ環境との衝突を生じないハンド把持姿勢を探索する。探索は基準となる把持姿勢の回りで、一定角度ずつハンドの向きを変えて行う。

(2) 観測された姿勢が最終姿勢でなければ、その場で姿勢遷移図の次の姿勢に置き直す移動計画を行う。すなわち、物体の初期姿勢と次の姿勢との双方で可能なハンドの把持姿勢を (1) と同様にして探索する。

(3) (2) で解が見つからないときには、視野内の空いた作業領域に一度物体を移動し、その状態から持ち替えが可能かどうかを調べる。このため、作業領域を格子点地図で表し、各格子点について可能なハンド把持姿勢を探索する。

7.2 作業実行

把持物体をハンドが同じ姿勢で把持して初期状態からゴールへ移動できる場合は、1 回の運搬で作業は終

ハンドの幾何モデルと環境の幾何モデルのいわゆる干渉チェックで、多面体近似で実行している。

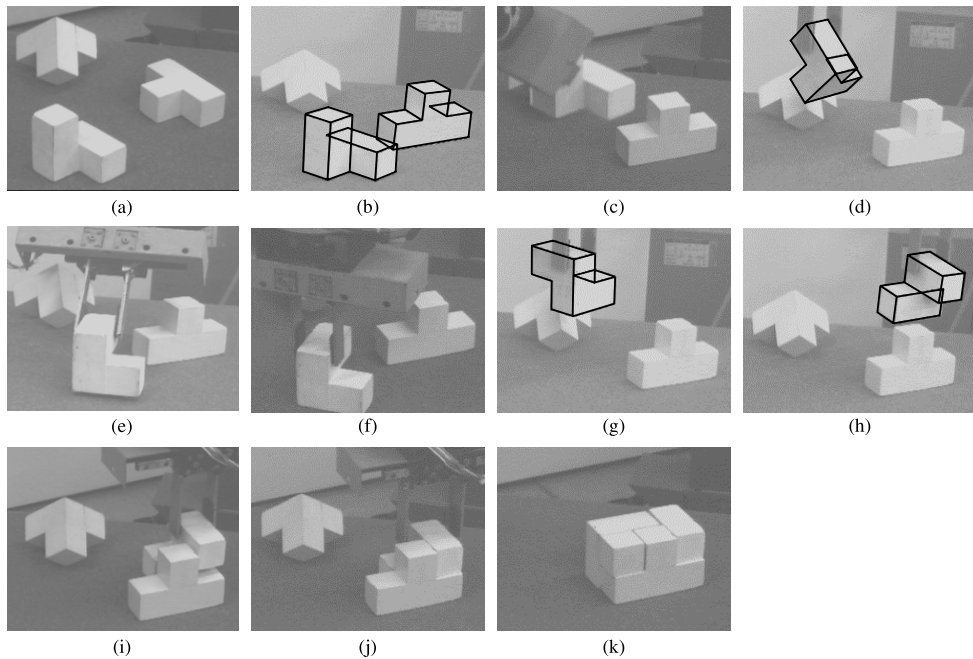


図 15 ブロックの組合せ作業状況
Fig. 15 Task of assembling building blocks.

了する。持ち替えが生じた場合や把持や床への設置時に積み木の姿勢が変わってしまった場合には、その位置で計測されたブロックの姿勢から再度把持計画を行い、ゴールに到達するまでこの過程を繰り返す。

7.3 作業監視

対象物体の検出と初期計測は、認識機能で行い、実行時の監視は追跡機能を用いる。本作業では、視覚システムの効果は以下の点で顕著となる。

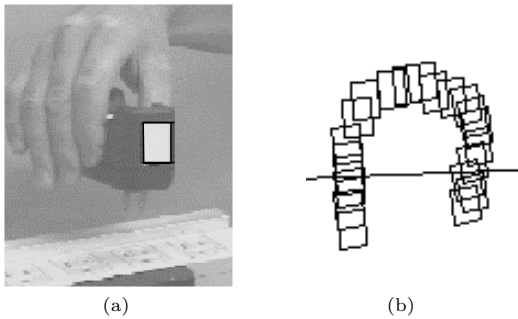
- ブロックの姿勢を変えて床に置くとときブロックが倒れてしまうような状況を追跡機能で迅速に検出することができる。
- 持ち替えるたびに、ハンドとブロックの位置関係がずれるが、視覚計測でそれを修正することができる。
- 認識結果を地図に反映させることにより、置き換え操作に必要な空いた地点を知ることができる。

7.4 実験結果

ブロックは、カメラ視野内の平面上に互いに重ならない任意の位置・姿勢で置かれた。図 15 に実行状況の一例を示す。(a) は初期状態、(b) ~ (j) は「ブロック A」の操作過程、(k) は最終作業結果である。ブロックに黒い縁どりが見えるのは、追跡結果をカメラ画像に重畳したものである。(d)、(g) は移動途中における把持状態の確認・計測状況を示している。

8. 作業の教示

作業によってはロボットの自律性に任せずにユーザが動作を教示したい場合がある。これまでも、ロボットの作業教示の負荷を軽減するために、視覚を用いた観察による教示 (Teaching by Showing) 法が提案されている。しかし、人の手と物体間の相対的な動きを観測するために、小さな窓を用いて局所的な特徴点を追跡する方法¹⁶⁾は、不変な特徴点を確保するために手の動きを制限しなければならないし、物体の種類や移動も限定される等の限界がある。また、アイインハンドシステムで、カメラを取り外して人手で動かしてその映像を記録した後、カメラをハンドに搭載して記録した映像と合うようにマニピュレータを制御する方法¹⁸⁾は、2次元視覚による対象物の初期位置の認識や追跡する特徴点の選択等の問題があるほか、物体の移動には利用できない等、用途が限定される。このほか、カメラで移動作業の区切りを検出した後、レンジファインダに切り替えて、そのときの物体間の拘束関係を調べる方法¹⁷⁾は、多面体を対象とした行動理解が主題であり、移動追跡機能はない。いずれにしても、多様な作業対象に対して簡便に精度の高い実時間計測を行うことができれば、観察による作業教示の適用範囲はさらに拡大し、また作業理解のレベルも向上すると期待される。そこで、当視覚系の高精度な認識と追跡機



(a) (b)

図 16 (a) 経路教示, (b) 移動軌跡

Fig. 16 (a) Teaching of motion, (b) Trajectory of motion.

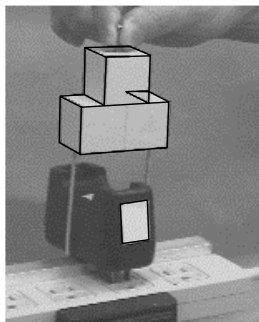


図 17 把持教示

Fig. 17 Teaching of grasping.

能を、対象物の把持と移動の作業教示へ応用することを試みた。

8.1 移動経路

精度の高い経路情報が取得できることを示すため、図 16(a) のような電池充電器をコンセントに差し込む作業を実験例とした。まず目標位置に物体を置き、最終位置・姿勢を認識処理を用いて計測した。次に移動物体を手で移動させながら目標位置までの経路を教示した。このときデータのサンプル間隔は、移動物の位置・姿勢の変化量を閾値としてコントローラブルである。サンプル例を図 16(b) に示す。

8.2 把持方法

物体に対するハンドの把持位置・姿勢教示には図 17 に示すようなハンドの模型を用い、その形状モデルを認識、追跡の対象とした。追跡を行いながら適切な把持位置まで移動し、その地点でのデータを取り込むことで教示を行った。

8.3 再生

再生時には対象物がつねに同じ場所のあるとは限らないので、まず、対象物の現在の位置を認識処理で計測し、Pick & Place 動作により、対象物を先に教示した初期位置へ置き換える。

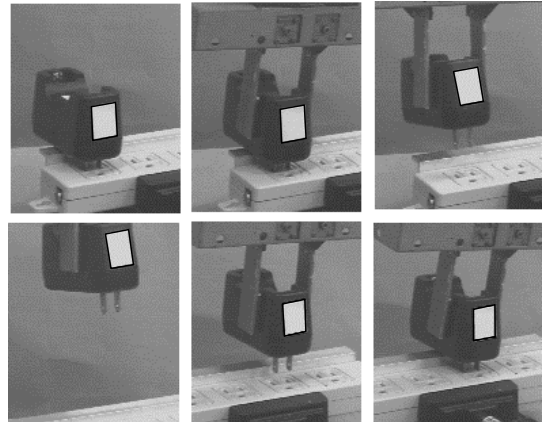


図 18 プレイバック

Fig. 18 Playback of task.

次に、教示した把持点データを使って物体への接近、把持を行い、経路データに従ってハンドを目標位置の近傍まで移動させる(図 18)。

最終位置への位置合せは、追跡処理を用いて把持物体の位置・姿勢を計測しながらビジュアルフィードバックを行った。これは、教示時に使用した輪郭へ移動時に観測される同じ輪郭を重ね合わせる操作を行っていることになる。

9. 作業移動型システム

車両等の移動体にマニピュレータを搭載して、自律移動と物体操作を組み合わせた作業移動型システムの開発も行っており、工場での自律搬送やオフィス業務への適用を図っている。産総研が中心になって開発しているヒューマノイド HRP-2(プロメテ)もその 1 つであり、筆者らがその視覚を担当している¹⁹⁾。搬送デモシナリオの中で、ロボットは視覚を用いて、地形形状計測を行い、自分が運搬作業を行うパネル山を探索し、その位置へ近づき、最終的にはパネルの把持位置の計測を行う。

9.1 視覚システムの構成

HRP-2 には 3 眼ステレオカメラが搭載されている(図 6)。首の左右(カメラ間距離 140mm)とその中央上部 70mm の 3 か所に白黒ボードカメラが配置されている。また首部にはパン・チルト機能が備わっている。ロボット固有の問題として、顔面シールドの歪補正を行っている。被写体距離が 0.5m(パネル把持位置検出時)から 3~4m(パネル板置き場検出時)の範囲をカバーできるように、画角約 33°のレンズを用いている。シャッタースピードは制御可能であり、照明環境の変化に自動対応できる。

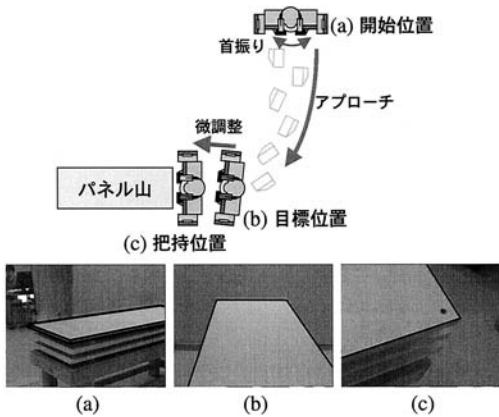


図 19 (a) パネル山検出, (b) 立ち位置検出, (c) 把持点検出
 Fig. 19 (a) Detection of mount of panels, (b) Detection of standing position, (c) Detection of grasping points.



図 20 パネル把持
 Fig. 20 Grasping of a panel.

9.2 パネル位置姿勢検出

パネル山を検出してアプローチし、パネルを把持する作業において、数回のパネルの位置姿勢検出を行い、歩行誤差・計測誤差を修正しながら歩行経路を自動生成している。カメラ画像から対象物体の輪郭線の3次元形状復元を行い、6.1節の場合と同様に、あらかじめ想定したパネルモデルと照合しその位置姿勢を求めている(図19)。最も精度が要求されるパネル把持位置は2mm以内の精度で検出でき、安定して把持することができる(図20)。

10. あとがき

3次元視覚を用いた物体操作の実験システムとして、人が操作する物体の移動方法を見て覚える物真似システムのほか、自動システムとして、空き缶等のピンピックシステム、円柱部品のペグインホールシステム、ブロックの組み立てシステム、作業移動型システムとしてヒューマノイドによるパネル運搬作業等を紹介した。実験例で扱っている操作対象は、ニキータンブロック、空き缶、バナナなど比較的単純な形状の物体であるが、本稿では操作機能を主題としているので、VVVでは一応任意の形状の物体を扱うことがで

きるという記述にとどめている。視覚機能の観点からは複雑な形状の物体ほど(特徴点が多くなるので)認識はむしろやさしくなる。しかし、同じ操作でも物体の形状が操作に影響する場合があるかどうか等の検証は今後の課題である。また、任意の物体の把持点と安定姿勢を幾何モデルから自動生成するシステムを開発する予定でもある。現在、まだ最も基本的な動作しか実現できていないが、使われている機能は、開発を予定している液体注ぎ等の他の動作でも共通的に利用されるものである。

今後も、つねに成長するシステムとして、不足している機能を増強するとともに、多様な実用例に適用することにより、その有効性の実証と実用に耐えるロバスト化を図る。また、マクロ的動作に有効な視覚とミクロ的動作に有効な触覚の協調システムは別途並行的に開発する予定である。

ハンドアイシステムは古い課題であり、解決済みの過去の問題として誤解される場合もある。しかし、実際には、簡単な作業ですぐに使えるシステムがないのが現状であり、それが筆者らのシステム開発の動機の1つでもある。技術的には、視覚がネックになっていることが現場では周知の事実である。人間が利用する情報の大半(80%以上といわれている)が視覚情報に基づいていることから容易に理解できる。したがって、人工の知能システム(AI, ロボット)においても、視覚は、本来、知能システムの情報源であり、知能システムに使えるようなものでなければならない。また、逆に、知能システムが、視覚を使わないのは無理がある。現在、AI, ロボット(ソフト)がともに壁にぶつかっているといわれているが、視覚を使えないでいること、使える視覚がないことが大きな原因である。使える視覚ができれば、他の多くの分野でも技術の高度化が図れることになる。

情報分野では、技術的にできている、できていないにかかわらず、表向きのシナリオや概念はほぼ出尽くしていると思われる。したがって、ことばの新規性に捕らわれすぎることなく、技術レベルによってできることのレベルが違うことに事の本質があることに注意する必要がある。ホンダの二足歩行ロボットがよい例で、過去と同じテーマでもアプローチが違えば新しい結果が生まれ、より高いステップへの芽となると考える。

本研究は、来るべきユビキタス情報環境社会を構成する多数の情報処理ユニット(人間の行う作業を支援または代行するエージェント)が「共通的」に利用可能とすべき知的機能として、CV(視覚), AI(自然言

語, 推論), ロボット(操作, 移動), ヒューマンインタフェース(対話, VR), 知識データベース(表現, 検索)を統合する知能情報処理システムの具体例の1つとして開発している。不特定ユーザの様々な知識と目的に応じてそのシステム構成が可変なように, 知識は階層構造的であり, 機能はモジュール的である。そして, 標準的なタスク(機能モジュールの組合せ)は体系的に整備され, また, 随時容易に追加, 編集できるものとする。そして, ユーザは, プログラミングすることなく, 対話的に知識をチューニングするだけで, 効率的に作業を実行するタスク指向システムを生成できるものとする。

参 考 文 献

- 1) 加藤頭剛: 1秒で物体の位置と姿勢が認識できる眼がほしい, 日本ロボット学会誌, Vol.12, No.1, pp.15-16 (1993).
- 2) 恩田寿和, 藤原伸行, 阿部清秀, 森 宣仁: 三次元検出による部品位置決めと事前のハンド干渉チェックにより実現した視覚ベースピッキングシステム, 日本ロボット学会誌, Vol.18, No.7, pp.995-1002 (2000).
- 3) 富田文明: 高機能3次元視覚システム VVV, 情報処理学会誌, Vol.42, No.4, pp.370-375 (2001).
- 4) 高瀬国克: トルク制御ロボットマニピュレータの機構と制御に関する研究, 電総研研報, 第876号 (1986).
- 5) Sumi, Y., Kawai, Y., Yoshimi, T. and Tomita, F.: 3D Object Recognition in Cluttered Environments by Segment-Based, *Int. J. of Computer Vision*, Vol.46, No.1, pp.5-23 (2002).
- 6) 北垣高成, 末廣尚士, 小笠原司, 劉 雲輝: 並列処理型センサベーストマニピュレーションシステム: 匠, 日本ロボット学会誌, Vol.15, No.3, pp.363-372 (1997).
- 7) 高橋裕信, 富田文明: ロボットビジョンのためのカメラキャリブレーション, 日本ロボット学会誌, Vol.10, No.2, pp.177-184 (1992).
- 8) 石山 豊, 角 保志, 富田文明: ステレオビジョンによる三次元物体の三次元運動追跡, 日本ロボット学会誌, Vol.18, No.2, pp.213-220 (2000).
- 9) 角 保志, 石山 豊, 富田文明: ステレオビジョンシステムのためのモデルベーストアプローチによる自由曲面体の実時間運動追跡, 信学論 D-II, Vol.J84-D-II, No.8, pp.1693-1700 (2001).
- 10) 平成10年度地域コンソーシアム研究開発事業「三次元視覚システム VVV を用いたタスク指向型・ロボット制御システム TORCS の研究開発」成果報告書, 新エネルギー・産業技術総合開発機構(NEDO) (2000).
- 11) Takahashi, H., Hirono, H., Terasaki, H. and Tomita, F.: Primitive Based Stereo for the Can-Picking Robot, *Proc. MVA '92*, pp.607-610 (1992).
- 12) 寺崎 肇, 長谷川勉: 指先に回転機構を有する平行2指ハンドによる知的物体操作のための行動計画, 日本ロボット学会誌, Vol.13, No.7, pp.1044-1052 (1995).
- 13) 横小路泰義, 坂本守行, 吉川恒夫: カメラ画像を併用したソフトフィンガー型の手による物体の操り, 第16回日本ロボット学会学術講演会予稿集, pp.501-502 (1998).
- 14) 三浦 純, 池内克史: 作業の目的を考慮した視覚認識戦略の生成, 日本ロボット学会誌, Vol.14, No.4, pp.574-585 (1996).
- 15) Matsushita, T., Sumi, Y., Ishiyama, Y. and Tomita, F.: A Tracking Based Manipulation System Built on Stereo Vision, *Proc. IROS'98*, pp.185-190 (1998).
- 16) 國吉康夫, 井上博允, 稲葉雅幸: 人間が実演して見せる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).
- 17) 池内克史, 末廣尚士: 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).
- 18) 中村幸博, 荒川賢一: マルチモーダルコミュニケーションによる教示, 日本ロボット学会誌, Vol.17, No.2, pp.166-169 (1999).
- 19) Kawai, Y., Fukase, Y., Ikeno, R., Ishiyama, Y. and Tomita, F.: Development of the Vision System for the HRP-2P Humanoid Robot, *Proc. MVA 2002*, pp.257-260 (2002).

付 録

動作の分類

- (1) 単なる動き, 加力動作, 指の動作
Pick, Release, Move, Include, Pull, Lift, Turn, Twist, Push, Support, Shake, Vibrate, Swing, Impact, Strike, Fit, Slide, Grasp
- (2) 物の移動, 他の物体との結合状態を変える動作
Transfer, Insert, Throw, Place, Put-on, Arrange, Lean, Hang, Combine, Screw, Unscrew, Separate, Extract, Attach, Set, Lock, Unlock, Pack, Unpack
- (3) 液体や粉体のような集合体を扱う動作
Pour, Paint, Plaster, Spray, Distribute, Sift, Fill, Lubricate, Stuff, Mix, Wipe, Gather, Draw, Scoop, Ladle, Clean, Write
- (4) 加工を加える動作

Punch , Drill , Saw , Cut-off , Whet , Sharpen ,
Shave , Plane , Whittle , Polish , Grind , Weld ,
File , Squeeze , Powder , Scratch , Drive-Nail ,
Unnail , Dig

(5) 柔軟体を操作する動作

Wash , Squeeze , Wind , Tie , Wire , Spread ,
Bend , Wrap , Saw , Knead , Tear-off

(平成 15 年 5 月 19 日受付)

(平成 15 年 9 月 9 日採録)

(担当編集委員 加藤 晃市)



富田 文明 (正会員)

昭和 53 年大阪大学大学院基礎工
学研究科物理系博士課程修了。工学
博士。同年通産省電子技術総合研
究所入所。現在、産業技術総合研
究所 3 次元視覚シ
ステム研究グループ長。筑波大学大学院システム情報
工学研究科機能工学系教授(兼任)。東京理科大学大学
院基礎工学研究科電子応用工学専攻客員教授。昭和 58
年～昭和 59 年カーネギーメロン大学計算機科学科客
員研究員。画像処理、コンピュータビジョン、知能口
ボットの研究に従事。平成 2 年情報処理学会創立 30 周
年記念論文賞。ICRA2003 Best Video Award。著書
「Computer Analysis of Visual Textures」(Kluwer
Academic Publishers)等。電子情報通信学会、日本
ロボット学会各会員。



松下 俊夫

昭和 47 年大阪大学基礎工学部生
物工学科卒業。昭和 49 年同大学大学
院基礎工学研究科修士課程修了。同
年電子技術総合研究所入所。現在、
産業技術総合研究所 知能システム研
究部門主任研究員。画像処理、ハンドアイシステム
等に関する研究に従事。電気学会、日本ロボット学会
会員。



河井 良浩 (正会員)

平成元年名古屋大学大学院博士課
程前期課程情報工学専攻修了。同年
通産省工業技術院電子技術総合研
究所入所。平成 13 年改組により独立
行政法人産業技術総合研究所 知能シ
ステム研究部門主任研究員、現在に至る。ICRA2003
Best Video Award。コンピュータビジョン、視覚障
害者用福祉機器開発の研究に従事。電子情報通信学会
会員。