

### 3次元認識によるビンピッキングシステムの実用化\*

中原 智治<sup>†</sup>・顧 海松<sup>†</sup>・荒木 秀和<sup>†</sup>・藤井 裕之<sup>†</sup>・廣田 雅之<sup>†</sup>

#### A Practical Bin-Picking System Using 3D Object Recognition\*

Tomoharu NAKAHARA<sup>†</sup>, Haisong GU<sup>†</sup>, Hidekazu ARAKI<sup>†</sup>, Hiroyuki FUJII<sup>†</sup>  
and Masayuki HIROTA<sup>†</sup>

We present a practical bin-picking system that uses stereo vision from multiple viewpoints and the three-dimensional recognition using salient features and uncertainty evaluation. First, the salient features of objects in images are extracted. Under the guidance of the features, the two-dimensional object candidates are detected. And the feature correspondence among stereo images is made based on the global features of an object model. Second, the three-dimensional features are reconstructed using the stereo vision results. Next, with the object features and the three-dimensional measurements, the uncertainty is evaluated. Then the results from the different viewpoints are integrated to determine the optimal results for each object by using the uncertainty and the three-dimensional position of object candidates. Based on this practical three-dimensional vision, an intelligent robot system is realized to pick up objects in a bin.

#### 1. はじめに

生産工程において、多品種中少量生産のものは一貫自動化ラインとなりにくく、コンテナよりばら積みされた部品を取り出してラインに供給する作業が数多く存在している。多くの場合、これらの部品供給作業にはパーツフィーダが用いられているが、振動・騒音・部品の破損・部品詰まりによるトラブル停止等の多くの問題が発生する。そのため、ばら積み状態から部品を取り出し部品供給を自動化する、ビンピッキングシステムの開発が必要とされている [1]。

ビンピッキングについての研究は 1970 年代中頃より多数の研究が行われてきている。センシング機能として触覚だけを用いるもの [2]、2次元認識と触覚を用いるもの [3]、3次元認識によるもの [4,5] と様々な方法が提案されてきているが、実験室レベルのものが多く、実際の生産に使用されたものはほとんど無かった。最近、実用例として、2眼ステレオ計測による3次元認識を使うもの [6]、画像処理とアクティブ3次元計測を組み合わせたもの [7] が報告されている。

本システムでは、多視点からのステレオ3次元計測と、顕著な特徴に基づくモデルベースの3次元認識を使うビンピッキングシステムを提案する。

ステレオ視は Marr ほか [8] により定式化されて以来、左右画像間の特徴対応付け問題や、実時間の距離画像獲得などについて多数研究されてきたが [9]、実用的なステレオ視の構築は依然難しい課題である。実用的な3次元認識には精度の高い3次元計測と安定した3次元認識とが必要とされる。

本文では、精度の高い3次元計測を実現するための物体の顕著な特徴に着目するセグメントベースステレオ計測、安定した認識を行うための計測信頼性を考慮するモデルベースマッチングによる3次元認識 [10]、ばら積みのためランダムな方向を向く部品を精度良く計測するための多視点からの計測方法 [11]、ラインの生産スピードに対応させるための並列分散処理による高速化 [12]、許容範囲内の計測誤差に対応するためのロボットハンド、を用いたビンピッキングシステムの実用化について説明する。

#### 2. 3次元認識

##### 2.1 顕著な特徴による物体認識

人間が物を観察する時、二つの場合があると考えられる。一つは、物の位置が既知で、対象を直接注視する場

\* 原稿受付 2000年6月28日

<sup>†</sup> 松下電工(株) Matsushita Electric Works, Ltd.; 1048, Kadoma, Osaka 571-8686, JAPAN

**Key Words:** 3D object recognition, salient feature, stereo vision, uncertainty evaluation, bin-picking.

合、もう一つは、物の存在位置が未知で、目立つ物から順次探していく場合である。目立つ物とは、色の鮮やかな物・急に動き出す物・鋭い輝度変化のある物などが考えられる。これらの物体が画像（あるいは画像列）に写る時に顕著な特徴として現われる [13]。これまでも、認識処理の効率向上のために、物体の同定や追跡および画像記述などの分野において顕著な特徴が利用されてきている。本項では、この考えを応用した顕著な特徴による物体認識について述べる。

山積み状態に置かれている物体のシーンでは同じ物体が多数存在しているため、どの特徴に着目すべきか、どの物体を先に認識すべきかが全体の認識結果を大きく左右する。Fig. 1 に金属部品のばら積み状態の一例を、Fig. 2 にこの物体のモデルを示す。このモデルは、表向き時の長辺対、裏向き時の長辺対、および短辺対により構成される。この部品は従来のピンピッキングの対象と異なり、面の要素が少なく、かつ大きな開口部がある。このため下の部品が上の部品により隠されないの、認識がより一層難しくなる。そこで、安定した認識を実現するために顕著な画像特徴を利用する手法を開発した。

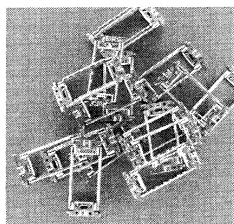


Fig. 1 Randomly stacked parts

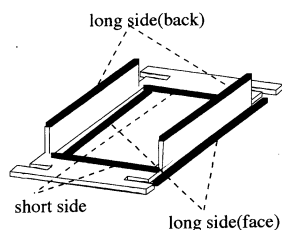


Fig. 2 Model of a part

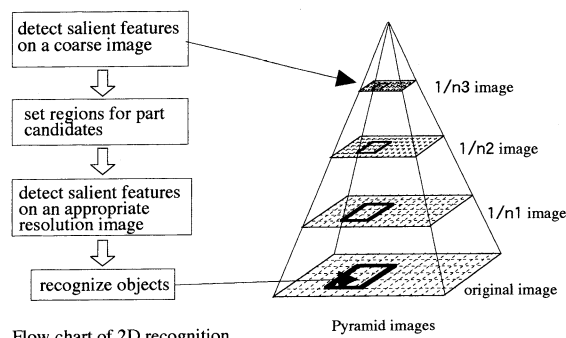
ピンピッキング用の3次元認識のための顕著な特徴として以下の特性が必要と考えられる。

- (1) 画像分解能に不変。
- (2) 物体の3次元姿勢変化に影響されにくい。
- (3) 物体の3次元姿勢を決定できる。

また、複雑な形状を持つ部品でも、通常複数の基本形状要素から構成されている。ここで基本形状要素とは、円・楕円・同心円・同心楕円、平行線・平行線群・矩形・平行四辺形などである。手法の汎用性の観点から、顕著な特徴が基本形状要素を含む必要がある。本文では、Fig. 1 の金属棒部品を対象に Fig. 2 の部品モデルの長辺において、金属光沢により表・裏向きともに安定して現われるハイライトを利用し、これらのハイライトによる平行線

群を顕著な特徴として検出し、これらを認識の手掛かりとする。

顕著な特徴の位置は画像の分解能に影響されないため、このような特徴抽出においては Scale-Space filtering 手法がよく使われている。本文では処理速度も考慮し、ピラミッド画像を利用して、最初に粗い分解能の画像からハフ変換により概略の特徴位置を検出する。そして、計測精度と検出する特徴の寸法および画像分解能との関係により、物体の他の特徴を適切な分解能の画像上で抽出する。処理の流れを Fig. 3 に示す。まず、粗い分解能の画像上で長辺対を検出する。つぎに、長辺対に囲まれている帯状領域内で更に短辺抽出を行い、物体の全体形状を認識する。



Flow chart of 2D recognition

Fig. 3 Two dimensional processing on pyramid images

## 2.2 ステレオ対応付け

ピンピッキングが対象とするシーン中では、同じ形状を持つ物体が多数存在するため、従来のように点やセグメントのような局所的な特徴のみを用いてステレオ対応付けを行うと、誤対応が多く発生する。そこで、以下に示す大局的特徴を用いてステレオ対応付けを行う必要がある。

- (1) 物体全体に関する形状特徴（円形・矩形）
- (2) 物体全体の代表点の位置と2次元での向き
- (3) 2次元形状構成要素の特徴（各直線の傾き・直線同士とのなす角度・直線長等）

Fig. 4 にステレオ入力画像を示す。今回の対象物体では、物体全体に関する形状特徴として矩形を、物体全体の代表点の位置と2次元での向きとして画像上での矩形の向きを、2次元形状構成要素の特徴として矩形を構成する長辺対・短辺対を使用する。したがって、対象物体

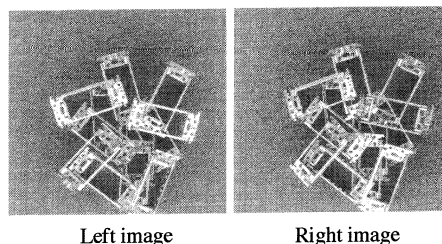


Fig. 4 Stereo image

に対して、エピソード拘束を基に、左右画像上で似た矩形同士の中心位置を用いて対応する物体同士の候補を求める。つぎに、各矩形を構成する長辺対・短辺対を用いて信頼度を評価し、対応候補から間違っただけの候補を削除した後ステレオ対応付けを行って3次元距離を計測する。

### 2.3 認識の信頼度評価

表面光沢・隠れ・画像ノイズ・外乱光などの影響で、ばら積みされた物体に対して、3次元位置姿勢を高い精度で求めることは極めて困難である。しかし、ピンピッキングでは対象物体の3次元形状が既知であるため、これをモデルデータとして利用し、3次元計測結果から一番精度の良いものを選択することが可能である。本文では、以下の二つの目的で、認識信頼性の評価を行う。

- (1) 認識結果の信頼性の定量化。
- (2) 各認識段階ごとの評価による認識効率の向上。

画像上での矩形検出段階と、3次元計測後の3次元認識段階に使用する評価式を(1)式と(2)式に示す。

$$\begin{aligned}
 U_{2d} = & w_1(\text{短辺の平行度}) \\
 & + w_2(\text{画像上での短辺間距離の一致度}) \\
 & + w_3(\text{短辺付近濃度勾配方向の一致度}) \\
 & + w_4(\text{長辺の平行度}) \\
 & + w_5(\text{画像上での長辺間距離の一致度}) \\
 & + w_6(\text{長辺付近濃度勾配方向の一致度}) \quad (1)
 \end{aligned}$$

$$\begin{aligned}
 U_{3d} = & w_1(\text{短辺の平行度}) \\
 & + w_2(\text{短辺間距離の一致度}) \\
 & + w_3(\text{短辺平面の平面度}) \\
 & + w_4(\text{長辺の平行度}) \\
 & + w_5(\text{長辺間距離の一致度}) \\
 & + w_6(\text{長辺平面の平面度}) \\
 & + w_7(\text{短辺平面と長辺平面間の距離の一致度}) \\
 & + w_8(\text{短辺平面と長辺平面の平行度}) \\
 & + w_9(\text{短辺平面と長辺平面の法線方向の一致度}) \\
 & + w_{10}(\text{短辺平面主軸方向と長辺平面主軸方向との一致度}) \quad (2)
 \end{aligned}$$

上記(1)式、(2)式において、距離に関する項は、モデルと計測値とを比較するものである。

評価方法は、各画像において検出された矩形に対して、(1)式の画像上での信頼度  $U_{2d}$  の評価を行い評価の低いものを切り捨てる。つぎに残された矩形に対して、前項の大局的観点に立ってステレオ対応付けを行い3次元計測を行う。つぎに、3次元計測結果を用いて(2)式の3次元認識の信頼度  $U_{3d}$  を計算し、3次元空間中でのモデルとの一致性を求める。

評価項目の重みに関しては、それぞれの評価項目の重要性和計測値の安定性に基づき値を決定する。

### 2.4 3次元認識結果

認識実験では、カメラと物体間距離：1.5[m]、視野：300×300[mm]と設定した。ばら積み状態の金属部品(Fig.4)に対して、顕著な特徴に基づき画像上での認識を行い、認識された物体候補の中心点をFig.5に示す。同じ形状の物体が互いに重なって存在しているため、最初の画像上での認識の候補には偽の結果が含まれている。つぎに、この画像上での概略認識で得られた各候補に対して、候補ごとに処理領域を限定し、限定された領域内において詳細に3次元認識を行う。Fig.6に原画像分解能で一つの候補に対して限定された物体の存在領域を示す。各物体候補ごとにこの原画像上で限定された存在領域内で、線分特徴の検出・計測の信頼性の評価・最適な長辺と短辺の検出を行い、これらの検出された長辺と短辺より物体の3次元空間での位置・姿勢を決定する。Fig.7の認識結果画像上に示された数字は、認識された結果の高さの順位を示したものである。信頼度を考慮して最終的に選択された結果は#1である。検出された短辺と長辺を白線で示す。認識に成功した割合を示す認識率は99.5%である。

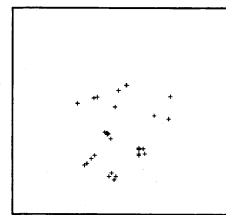


Fig. 5 Center points of part candidates

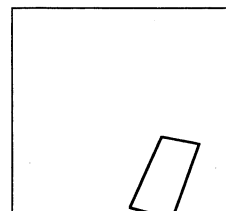


Fig. 6 A region for detail processing

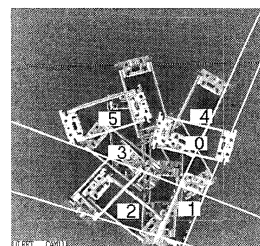


Fig. 7 Result

## 3. 多視点からの計測

### 3.1 多視点ステレオ計測

カメラを使って計測を行う場合、Fig.8に示すように、カメラの視線方向と物体の面・線とのなす角  $\phi$  が計測精

度に大きく影響を及ぼす。角 $\phi$ が直角に近い程精度の高い計測を行うことができる。また、セグメントベースのステレオ計測においては、エピ極線方向と平行なセグメントの計測は不可能であるので、Fig. 9に示すように、画像上において、エピ極線と特徴とのなす角 $\theta$ が計測精度に大きく影響を及ぼす。この場合も、角 $\theta$ が直角に近い程精度の高い計測を行うことができる。ピンピッキングの場合、対象物体がランダムな方向を向くため、上記視線方向と物体特徴とのなす角、すなわち、画像上においてエピ極線と特徴とのなす角による計測精度の低下に対策をしておく必要がある。これらの問題の影響を緩和するために、Fig. 10のように複数の基線方向を有する多視点計測ステレオ法を提案する。任意の二つのカメラを選択することで、様々な基線方向での計測が行えるため、誤差の少ない計測が可能となる。また、複数の方向からの計測となるので、各セグメントの観測に適した方向から計測することができ、見え方の違いにより生じる計測誤差を減少させることが可能となる。

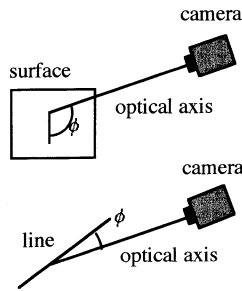


Fig. 8 Optical axis and feature directions

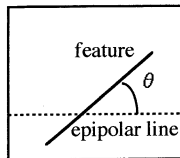


Fig. 9 Epipolar line and feature direction

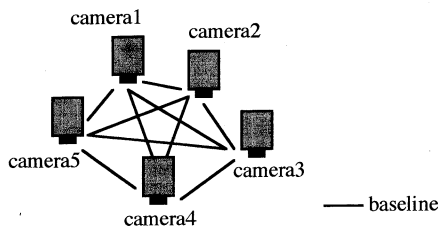


Fig. 10 Multi-view stereo vision

### 3.2 計測結果の統合

計測の信頼性は、2.3で示したモデルと計測値との誤差による要素と、上記3.1のエピ極線とセグメントとの位置関係より決まる要素により表される。したがって、計測信頼度 $C$ は(3)式で表現できる。

$$C = \sum f(M, S_i) \cdot g(\theta_i) \quad (3)$$

ここで、 $f$ は(2)式に示すモデルと計測値とのマッチングを表す評価関数、 $M$ はモデルデータ、 $S_i$ は $i$ 番目の矩形に対する計測値、 $g$ はエピ極線と矩形との位置関係を表す評価関数、 $\theta_i$ は $i$ 番目の矩形とエピ極線との成す角である。信頼性が高いのは、あらかじめ登録されたモデル形状と計測された形状との誤差が小さく、セグメントとエピ極線とのなす角が直交に近い場合となる。

つぎに、 $n$ 台のカメラを用いた多視点ステレオでは、 $nC_2$ 通りのステレオペアが選択でき、一つの部品に対して最大 $nC_2$ 個の計測結果が得られる。そのため、一部品に対して得られた複数の計測結果を一つにまとめる必要がある。その指標として、(3)式で示した計測信頼度 $C$ を用いる。つまり、複数の結果の中で、最も計測信頼性が高い結果を選択する。

### 3.3 統合性能確認実験結果

上記方法の有効性を確認するため、次の実験を行った。Fig. 11に示すように、およそ正三角形の頂点に配置した三台のカメラで、回転テーブル上の部品を順次回転させて計測する。テーブル回転量と部品の姿勢計測値との相対誤差と、最終結果を出力したステレオペアとの関係をFig. 12に示す。Fig. 12中において、実線で結ばれた結果が最終的に選択された結果の姿勢誤差を、黒丸が最終結果を出力したステレオペアを、斜線を施した領域が特徴とエピ極線との関係より計測精度が悪くなるため選択されるべきではない領域を示す。最終的に選択された結果の姿勢誤差はおおむね2度以内と良好である。信頼性の評価式(3)におけるモデルとのマッチング項の影響で、選択されたステレオペアが多少ばらついてはいるが、計測

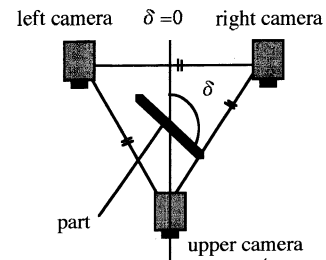


Fig. 11 Arrangement of cameras and a part

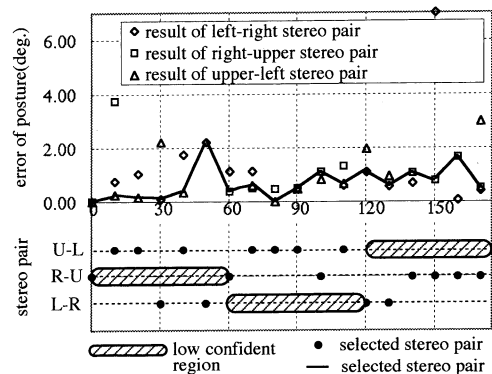


Fig. 12 Selected stereo pair

精度の悪い領域からは結果が選択されず良い動作を示している。

## 4. ピックアップ戦略

### 4.1 計測信頼性に基づくピックアップ戦略

ピックアップ動作前に、複数の計測結果中からどの計測結果を把持対象とするかを決定する必要がある。本項では、計測信頼性を利用した把持対象の決定方法を提案する。まず、計測結果間での干渉の有無を調べる。他の計測結果と干渉せずに孤立している計測結果の中で、計測信頼性がしきい値以上の計測結果があれば、その計測結果を把持対象とする。つぎに、高い位置にある計測結果から順に計測信頼性を調べ、しきい値以上の計測結果があれば、その計測結果を把持対象とする。ただし、上下に重なり下の計測結果の信頼度が高い場合については、例外処理として、上の計測結果を選択するようにしている。つまり、計測結果相互の位置関係と計測結果の信頼性とを同時に利用することによって、確実にピッキング作業を行うことができる計測結果を選択する。

### 4.2 ピックアップ対象認識結果

Fig. 13に示すシーンに対して、上記提案手法を適用した。その時の部品のばら積み状態を Fig. 14に、得られた計測結果を Table 1に示す。Table 1において、Zは物体を置いたテーブル上から物体中心までの高さを、Cは計測信頼度を示す。計測信頼度は値が小さい程信頼性が高いことを意味する。部品#1についての計測結果は計測信頼性が低く、部品#2についての計測結果は信頼性が高いため、#1が一番高い位置にあるが二番目に高い#2をピックアップ対象として選ぶことになる。部品#1の計測信頼性が低いのは、姿勢が大きく傾いているためである。

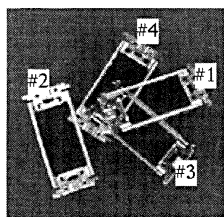


Fig. 13 Scene

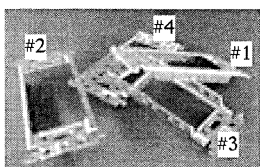


Fig. 14 Oblique view of stacked parts

## 5. ビンピッキングシステム

本節では、前節までに説明してきた3次元認識方法を用いるビンピッキングシステムの構築について述べる。

Table 1 Results

No.	Z*	C
1	22.5	5155
2	13.8	1052
3	13.0	1079
4	9.4	4079

\*Z indicates z position of a part from the table surface

### 5.1 構成

本ビンピッキングシステムの概略構成を Fig. 15に示す。システムは部品供給部、三組のステレオカメラ、画像処理装置、ロボットコントローラと6軸多関節ロボットで構成される。まず、部品供給部でばら積み部品をロボット動作位置に供給する。つぎにばら積み部品を計測・認識し、把持対象を決定してロボットにより部品を把持する。ランダムな姿勢をとる部品を安定して認識するために、異なる三つの基線を持つ三組のステレオ構成としている。

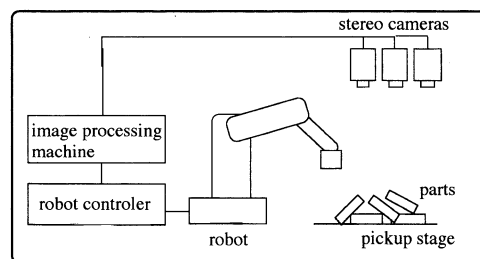


Fig. 15 Configuration of the bin-picking system

### 5.2 並列分散処理

ステレオベースのマシンビジョンでは認識処理の計算量が非常に多くなる。したがって、実用的なシステムを構成するうえでは認識処理部の高速化が大きな課題となる。そこで、本システムでは実時間での認識処理を行うために、粗密画像を用いて多段階認識処理を行うことにより、効率的な認識を行っている。さらに、これらの処理を並列して行うことができるアーキテクチャを開発し、当社で独自に開発した密結合型並列画像処理装置と市販のリアルタイム OSにより高速処理を実現した。認識処理の並列処理アーキテクチャを Fig. 16に示す。この装置は画像専用バスを持ち、一枚のマスタボードと複数のスレーブボードから構成されている。3眼ステレオの場合、各カメラから入力された画像はいったん各スレーブボードで粗画像に変換された後、概略認識・概略ステレオ対応付けが行われる。つぎに詳細認識部では、概略認識で候補になったもののみを、密画像を用いて候補数に応じた  $n$  枚のスレーブボードで並列に処理する。そして、マスタボードで認識結果の併合・部品間の干渉認識処理を行うことにより把持対象部品を決定する。この並列アーキテクチャにより、処理時間においてボトルネックとな



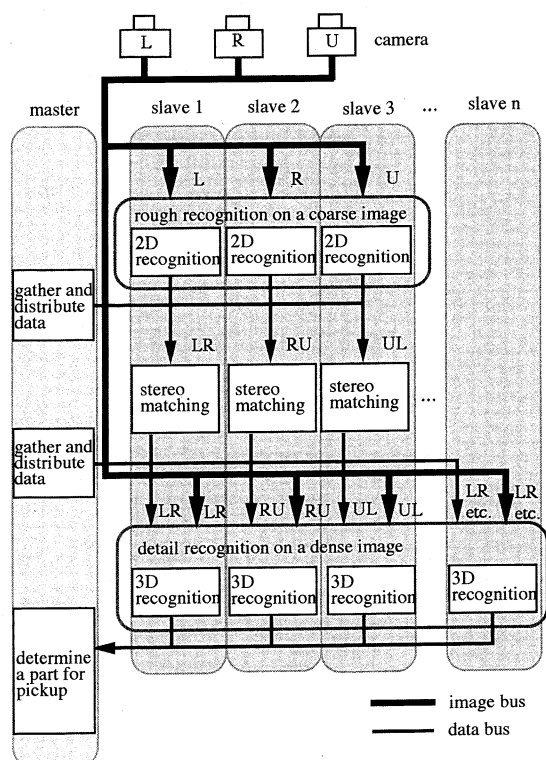


Fig. 16 Parallel processing architecture

る認識処理を、スレーブボードの増減で所望の時間で処理させ、実用的な時間でビンピッキングを行うことが可能となった。

### 5.3 安定したピックアップのための把持機構

三組のステレオによる高精度な計測、計測信頼度を用いた認識を行っているが、認識結果に誤差が含まれることは避けられない。このような誤差を含む計測値を使い、かつ、重なりのある複雑なシーンを正確に認識することは難しい。他の部品とハンドとの干渉を防ぎ安全にピックアップするために、部品の法線方向からアプローチし電磁吸着により把持する。吸着部にフロート機構を付加することにより安定したピックアップを実現した。ピックアップ時のハンドの様子を Fig. 17 に示す。

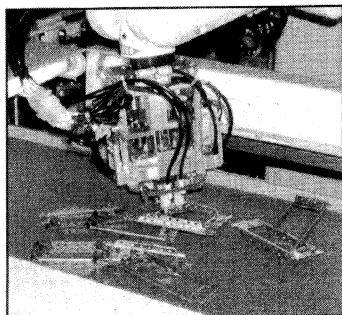


Fig. 17 The robot hand and stacked parts

### 5.4 ビンピッキング結果

ばら積み状態の金属部品に対して、ビンピッキングを行った結果を Table 2 に示す。位置・姿勢精度はハンド

にフロート機構を持つため、ピックアップ動作にはこの表に示される値で十分な精度である。認識処理とロボットの動作は並列に行われるため、本システムではこの処理時間で生産ラインの生産速度を満足する。今回、ビンピッキングの性能を示す指標として、すべての部品を取りきるまでにリトライする割合を示すリトライ率を定義した。リトライ率は0%が望ましいが、現在の結果で実用性は満足している。

Table 2 Performance of the bin-picking robot

position accuracy	±3 mm
posture accuracy	±5°
processing time/part	2s
retry ratio*	12%

$$*\text{retry ratio} = \frac{\text{failures}}{\text{all trials}}$$

## 6. おわりに

ビンピッキングの実用化を目的に、物体の顕著な特徴に着目するセグメントベースステレオ計測、計測信頼性を考慮するモデルベースマッチングによる3次元認識、多視点からの3次元計測方法、並列分散処理、計測誤差を吸収するロボットハンドを開発し、ビンピッキングシステムへと適用しその実用性を確認した。

今後、より多くの部品に対してビンピッキングシステムを適用し、汎用システムへと開発を進めていく予定である。また、今回の装置は部品供給のみを行うが、組立ロボットと3次元視覚システムを組み合わせることにより、さらに付加価値の高いフレキシブルなロボットシステムを実現していきたいと考えている。

## 参考文献

- [1] 井口, 佐藤: 3次元画像計測, 昭晃堂, pp. 142-147 (1990)
- [2] C. Rosen, D. Nitzan, G. Agin, A. Bavarsky and G. Gleason: Research applied to industrial automation; 8th Report (prepared for the National Science Foundation), Stanford Research Institute, Menlo Park, CA (1978)
- [3] J. Birk, R. Kelley, J. Duncan, H. Martins and R. Tella: A robot system which feeds workpieces stored in bins; *Proc. 9th Intl. Symposium on Industrial Robots*, pp. 339-355 (1979)
- [4] 池内, 長田, B. K. P. ホーン, H. K. ニシハラ: 照度差ステレオとプリズムステレオを用いたビンピッキングタスク中のグリッパの姿勢決定; 信学論, Vol. J65-D, No. 4, pp. 546-553 (1985)
- [5] T. Fujita, K. Sato and S. Inokuchi: Range image processing for bin-picking of curved object; *Proc. IAPR Workshop on CV*, pp. 451-454 (1988)
- [6] 井倉, 恩田, 藤原, 阿部, 森: 鋳造部品の加工機へのロー

- ディングのための視覚ベースピンピッキング; 日本機械学会ロボティクス・メカトロニクス講演会'99 講演論文集, 2A1-19-007 (1-2) (1999)
- [7] 榊原: 生産システムに急速な変革を迫る知能ロボット; 第17回日本ロボット学会学術講演会, pp. 69-70 (1999)
- [8] D. Marr: Cooperative computation of stereo disparity; *Science*, 194, Oct. 15, pp. 283-287 (1976)
- [9] 松山, 久野, 井宮: コンピュータビジョン 技術評論と将来展望; 新技術コミュニケーションズ, pp. 123-137 (1998)
- [10] 顧, 中原, 荒木, 藤井: 物体の顕著な特徴に基づく3次元認識; 第17回日本ロボット学会学術講演会, pp. 801-802 (1999)
- [11] 荒木, 中原, 顧, 藤井: ピンピッキングのための3次元計測とピックアップ戦略; 第17回日本ロボット学会学術講演会, pp. 803-804 (1999)
- [12] 藤井, 中原, 荒木, 顧, 廣田: 実用的ピンピッキングシステム; 第17回日本ロボット学会学術講演会, pp. 805-806 (1999)
- [13] L. Itti, C. Koch and E. Niebur: A model of saliency-based visual attention for rapid scene analysis; *IEEE PAMI*, Vol. 20, No. 11, pp. 1254-1259 (1998)