

人間の知覚、動作特性の測定、評価に関する基礎的研究

Fundamental research on measurement and evaluation of characteristics of human perception and movement

代表研究者 東京農工大学工学部教授 本多庸悟
Prof., Faculty of Eng., Tokyo Univ. of Agri. and Technol.
Tsunenori HONDA

協同研究者 東京農工大学工学部助教授 北嶋克寛
Assoc. Prof., Faculty of Eng., Tokyo Univ. of Agri. and Technol.
Katsuhiro KITAJIMA

東京農工大学工学部助手 金子俊一
Assist., Faculty of Eng., Tokyo Univ. of Agri. and Technol.
Shun-ichi KANEKO

東京農工大学工学部助手 五十嵐智
Assist., Faculty of Eng., Tokyo Univ. of Agri. and Technol.
Satoshi IGARASHI

With the recent development of information processing and automated systems, the new types of technological and social circumstances are being constructed. To make them more human-adaptable, human mental and physical characteristics against the engineering or artificial circumstances should be reviewed and examined from the new viewpoints, and the results should be effectively utilized for constructing the coming information processing and automated systems. Miscellaneous functional configurations with diverse objects and aspects such as FA (factory automation), OA (office automation), HA (home automation), etc. are being formed, and some new problems are appearing. For example, in FA, methods for I/O, preprocessings preceding to automatic processes, certification of efficiency and safety for robot controls, etc. are now problems to be solved.

Based on the recognition of the states abovementioned, we intend to collect and analyse basic data on human perception and motion, and to get basic materials to construct more effective and human-adaptable information processing and automated systems. The items to be studied in the research are as follows:

- (1) Recognition of 3D objects,
 - (a) Recognition of polyhedral objects based on multiple views,
 - (b) 3D shape recognition of metal products,
 - (c) Autonomic robot arm with visual recognition,
- (2) Recognition of free-hand drawings,
- (3) Robot language based on verbs for operations, and
- (4) Visual recognition and positioning movement.

研究目的

近年の高度情報化に伴って、我々の周囲には新しい形態の技術的、社会的環境が形成されていく

ものと考えられる。これを我々人間にとて、真に好ましいものとしなければならない。そのためには、我々自身の情報処理に関わる精神的、肉体

的特性を新しい観点から調べ、今後の情報システムや自動システムの構築に生かしていくことが必要であると思われる。既に、FA, OA, HA などさまざまな対象と形態を持つ情報的環境が形成されつつあり、多くの問題が提起されつつある。たとえば FA に関するものを見ると、システムへの入力方法の拡大、自動化に先だって行なわれるべき多くの準備作業の処理の効率化、ロボットのような高度に自動的な機械の制御の効率や安全性の確保、などがある。

人間と、情報・自動システムのような人工システムの融合を考えるとき、それらの間に存在する次のような解決すべき問題を挙げることができる。

(1) 我々にはほとんど何の苦もなくできる動作で、機械化はまだむずかしいもの、たとえば、3 次元物体を視覚的に識別し、つかみあげること。

(2) 我々の考えを肉体的機能によって表現し、我々人間どうしのコミュニケーションに用いている様式を、そのままのかたちで情報システムに受け入れさせること、たとえば、フリーハンドで書いた図をそのまま計算機に読み込ませること。

(3) 我々が日常使っている言葉と知覚の関係をどのように人工システムに持ち込めるのか、たとえば、その 1 個のなかに多くのシーケンスを含んでいることの多い動詞を機械に対する指示に用いるとしても、単なる固定的シーケンスでは不確定な事態に対応できないので、なんらかの知覚との統合的処理が必要である。

本研究はこのような観点に基づき、計算機を用いて自動認識や制御のためのプロトタイプを構築し、どの程度のレベルの人工システムが実現できるのかを試みるとともに、人間の知覚、動作に関する基礎的データを収集、解析し、より有効な情報処理あるいは自動システムを構築するための基礎資料を得ようとするもので、具体的に次のような項目を取り上げる。

(1) 3 次元物体の視覚認識

(a) 多視野の画像に基づく多面物体の認識

(b) 金属物体の形状認識

(c) 視覚認識を利用した自律的ロボットアーム

(2) フリーハンド手書き機械図面の認識

(3) 作業動詞を用いるロボット言語

(4) 視覚認識と位置決め動作

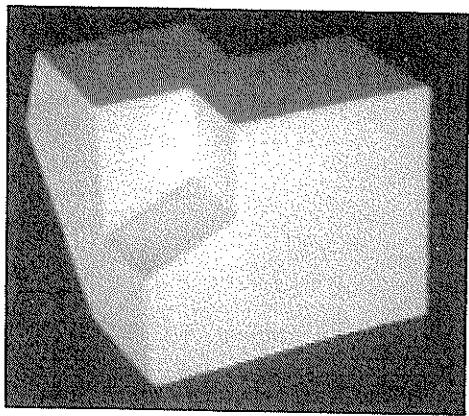
研究経過および成果

(1) 3 次元物体の認識

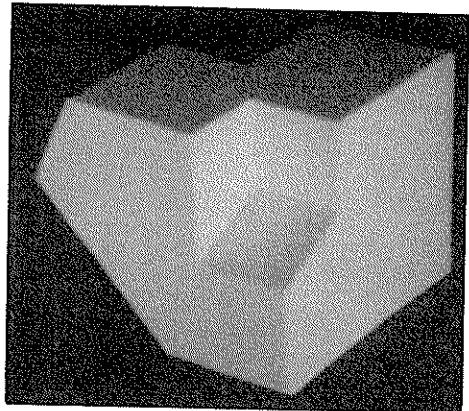
(a) 多視野の画像を利用した多面体物体の認識

3 次元多面体の「全体形状」を再構成する体系的な手法の試みについて述べる。従来のステレオ撮像を利用した位置計算法に比べて、この手法の特色は、構造的照明などを用いない通常の撮像環境で得られる濃淡画像を用い、かつなるべく少ない枚数の画像データにより、効率良く 3 次元対象物の全体形状を再構成することである。すなわち、対象物を載せた回転台の回転角度を操作することにより、一つの視野を代表するステレオ画像対を、再構成するために十分で、かつなるべく冗長でない（共通部分を少なくする）いくつかの視野に関して撮像し、それらを併合して全体形状を得る。Fig. 1 にステレオ原画像対、および画像処理の例を示す。各画像対ごとに頂点の対応付けを行るために、次の三つの手法を用いている。①幾何学的に決定できる拘束直線上を探索する。一意に決定不可能な頂点（2 重の対応など）については、②対応付け可能な面の情報から、一意に対応が決定できるものもある。さらに、③面と視線との交点として計算できる 3 次元位置を参考にすることも可能である。各視野を代表する部分形状を表すグラフを、共通の特徴（共通頂点、および共通稜線）などをより併合する。この例の場合には、4 枚の画像で再構成が完了した。Fig. 2 は再構成したワイヤーフレーム表示である。Fig. 3 に頂点位置の計算精度を示す。使用した CCD カメラ、およびフレームメモリ装置は通常のものである。画像内の頂点を目視で検出した場合と同等の精度が得られていることが分かる。平均位置誤差 6.1 mm (撮像距離の 0.76%)、標準偏差 2.5 mm 以内で頂点を位置決めすることができた。

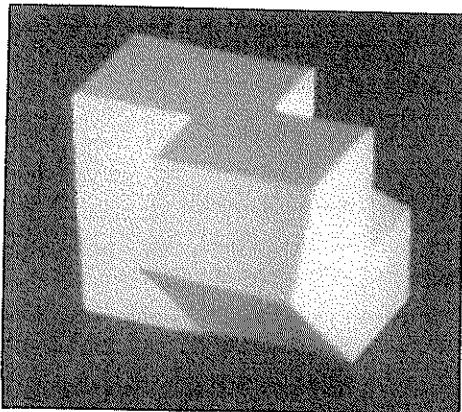
(b) 金属物体の形状認識



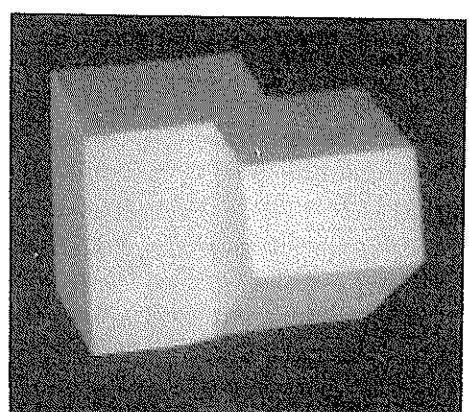
(a) Rotation angle: 0°



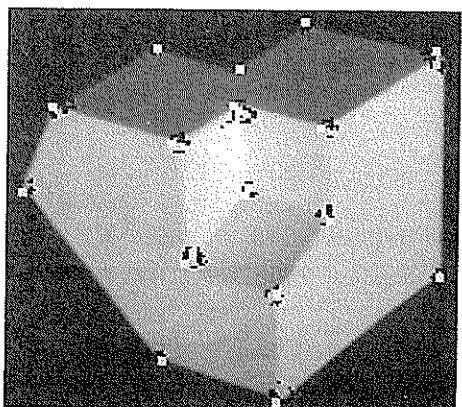
(b) Rotation angle: 30°



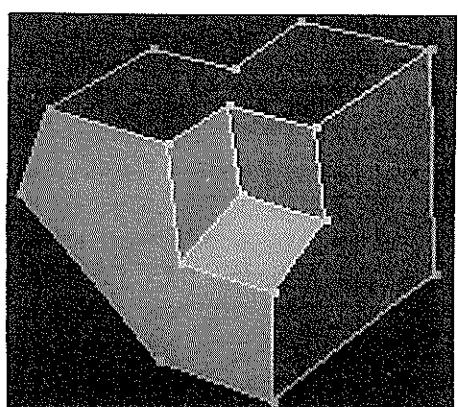
(c) Rotation angle: 130°



(d) Rotation angle: 170°



(e) Extraction of vertices



(f) Extraction of surfaces

Fig. 1 Original images and picture processings.

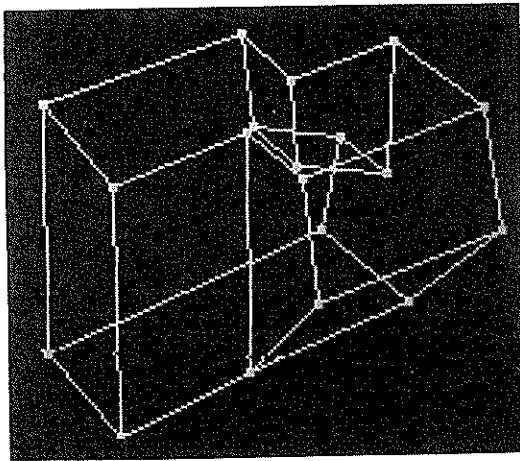


Fig. 2 Wire-frame display.
(Rotation angle: 180°)

従来 3 次元物体の認識手法が多く提案されているが、そこで扱われた対象物は、主につや消し面をもつ多面体物体であった。そこで開発された画像処理技術がある程度の汎用性をもつことは確かであるが、光学的特性の大きく異なる種類の対象物の処理も併せて検討する必要がある。そこで、2 種類の金属物体を取り上げ、その形状認識を指向した画像処理手法に関する基礎実験を行った。

Fig. 4 に、①石膏模型、②やすりがけされたアルミニウムブロック、および③鋸びた鋳鉄ブロックの3種類の対象物の濃淡原画像および微分画像を示す。このうち、①石膏模型はつや消し面に近いのでこれまでと同様に処理可能である。②やすりがけしたアルミニウム面で構成される直方体では、微分画像を見ても分かるように、稜線上では連結している画素が多いが、面内に多くのハイライトが観測されている。③鋸びた鋳鉄においては、稜線も面内の区別なく孤立ノイズ状の特徴点が現れている。これら3種類の多くの表面画像サンプルについて、平均値、分散などの統計量を算出してみても、顕著な違いが見られる。たとえば、アルミニウム面などでは拡散反射とともに、かなりの鏡面反射をすることなどが分かっている。

このようなかなり性質の異なる表面をもつ多面体対象物の形状を再構成するために、各面の傾き

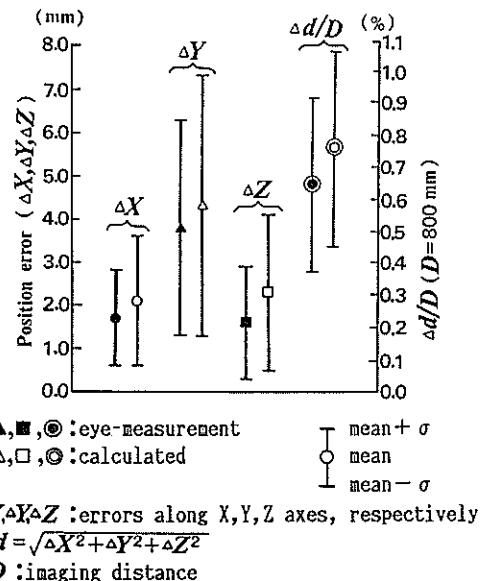


Fig. 3 Position errors of vertices.

を決定するための特徴点抽出法として、次の二つのアプローチについて検討した。①稜線部分が比較的取り出しやすいつや消し面やアルミニウム面は、微分処理、2 値化に続いて小さな孤立領域を除去し、残ったものを特徴点（主に稜線部分となる）とする方法、および②鋸びた鋳鉄表面では稜線部分を単純な処理で取り出すことがむずかしいので、微分処理、2 値化の結果をそのまま特徴点（主に面内の点が多い）とする方法である。Fig. 5 に再構成した面の傾きを示す。パラメタの値によって多少違いが見えるが、ほぼ安定に再構成が可能であることが分かる。そこにおいては、面法線の角度誤差に換算して、②のアルミニウムの直方体の面は平均角度誤差 3.57°、③の鋳鉄の直方体は平均角度誤差 6.70° で再構成できている。

(c) 視覚認識を利用した自律的ロボットアーム

ロボットアームの位置決め制御のために視覚情報を利用する試みを行なった。多視野画像を利用した3次元物体の視覚認識などの先行の研究で得た知見を基礎として、より実際的応用のレベルに近付くために、ロボットアームの把握操作を自律的に、すなわち、あらかじめの教示なしに制御

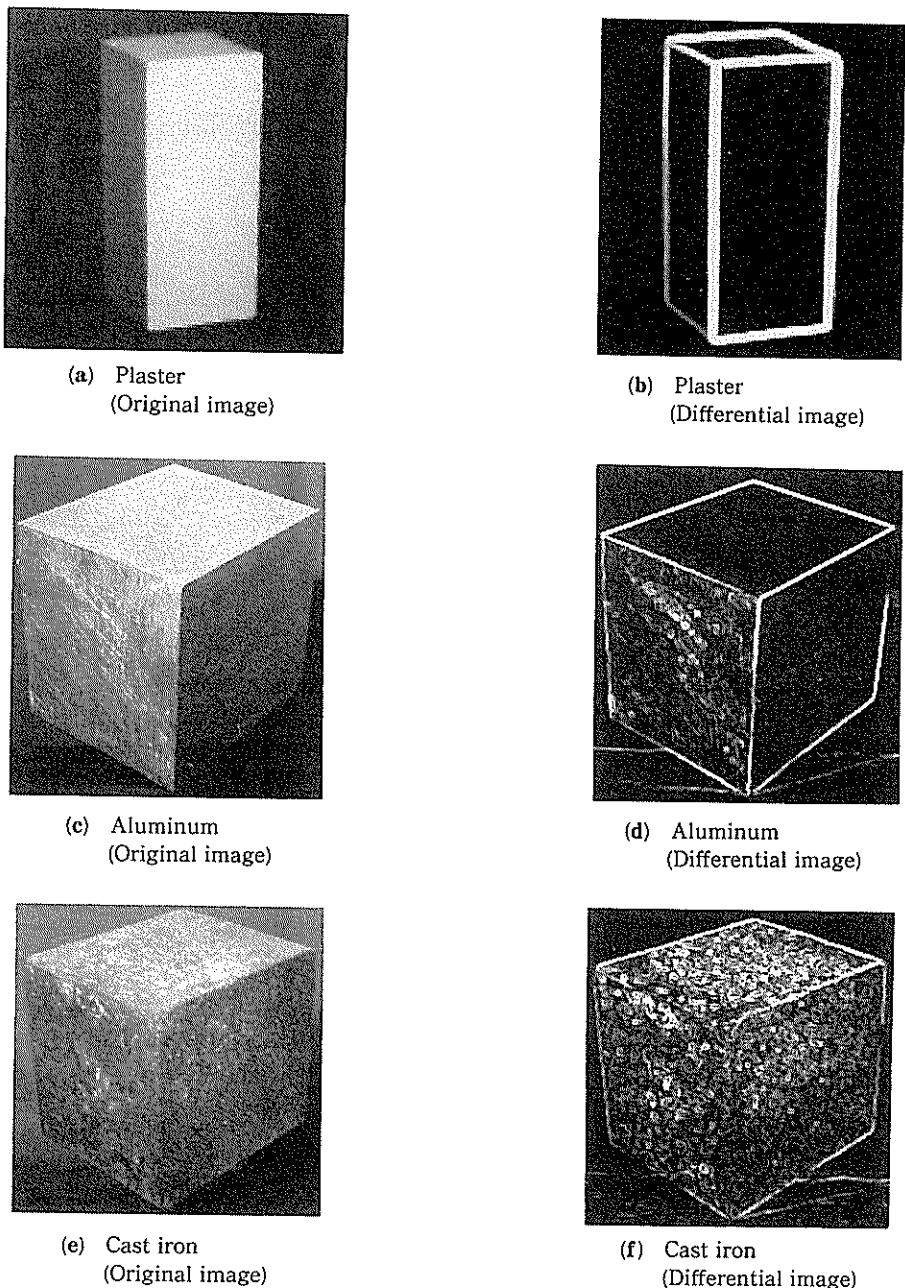
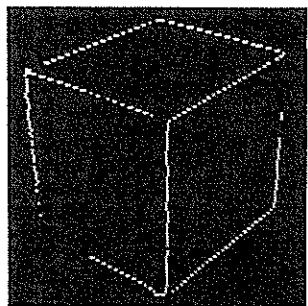


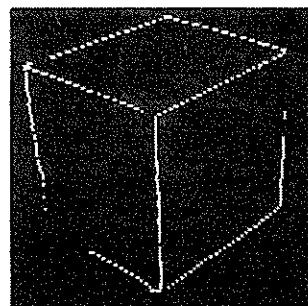
Fig. 4 Results of picture processings.

するために視覚情報を利用する方法について検討した。実験システムの構成は次のようにある：①視覚的捕捉サブシステム：ロボットアームに装着したCCDカメラを移動させながら、対象物の像

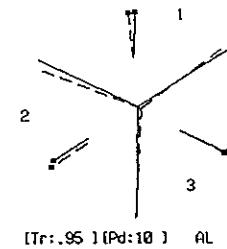
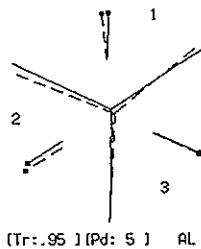
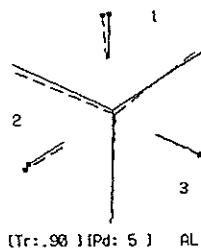
を画面中央に捉える。②視覚的認識サブシステム：対象物の2次元的特徴量、およびステレオ撮像による位置計算法によって求めた3次元的特徴量を用いて対象物の識別を行う。③把握制御サ



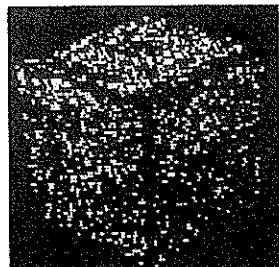
(a) Aluminum (right)



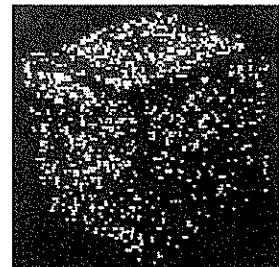
(b) Aluminum (left)



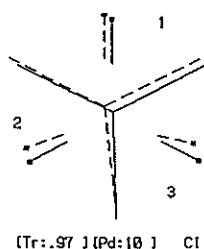
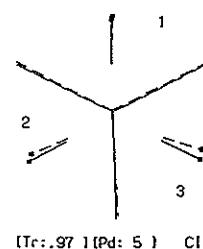
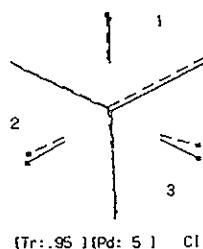
(c) Reconstruction of surface geometry



(d) Cast iron (right)

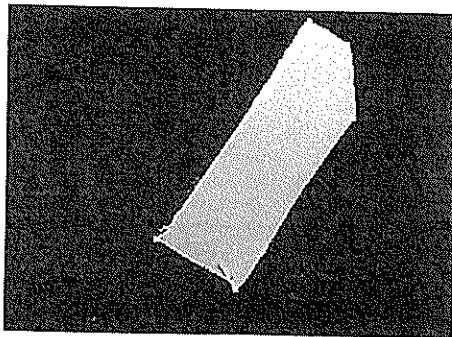


(e) Cast iron (left)

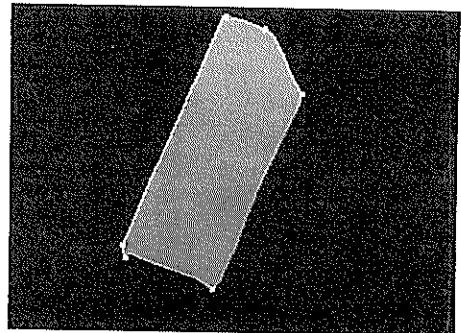


(f) Reconstruction of surface geometry

Fig. 5 Surface reconstruction.



(a) Right image (first)



(b) Left image (second)

Fig. 6 Contour extraction.
(Stereo-pair, convergence angle: 5°)

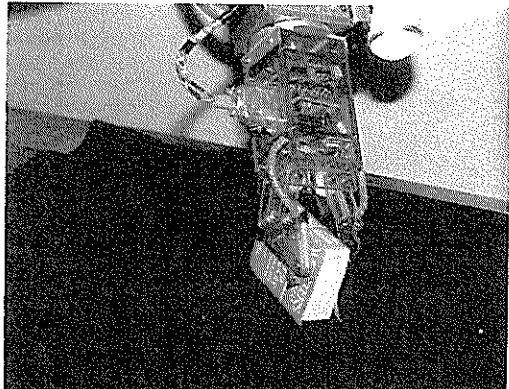


Fig. 7 Visual positioning and automatic grasping.
(Object: book)

システム：ステレオ位置計算法に含まれる位置誤差に対処するために、ハンドの指先に装着した接触センサを利用した把握制御を行う。これら三つのサブシステムはこの順序で起動され処理を進める。その中で、特に次の三つの項目について、新しい方法を提示した。

- ①教示なしに、対象物を視覚的に捕捉する手法について検討する。
- ②凹多面体をも扱うために、2次元輪郭形状を凸包によって近似的に記述し、それに基づいて識別用特徴を抽出する。
- ③対象物の3次元空間における姿勢を判定するため、輪郭頂点座標の共分散行列の固有

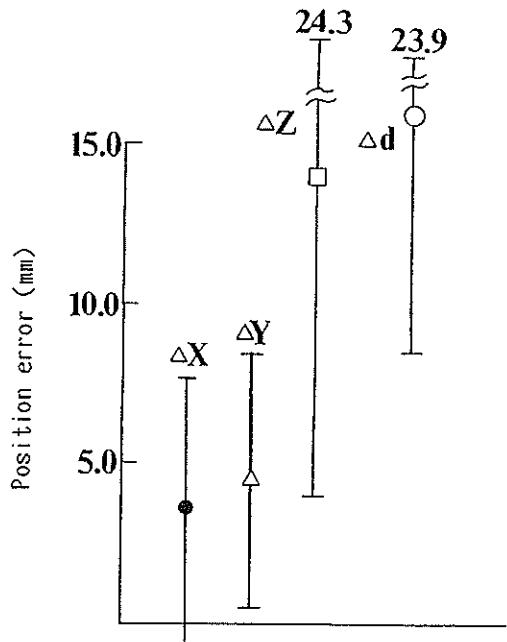


Fig. 8 Position errors of vertices.

ベクトルの組を利用する。

まず、視覚的捕捉では、まず探索領域をあらかじめ設定した順序で走査し、対象物を検出して、推定した対象物の中心へセンタリングを行う。Fig. 6 にステレオ画像における輪郭抽出の例を示す。実験に用いた対象物はつや消し面の多面体が6種類、光沢面の積木、カウンタ（プラスチック製）、小型辞書の計9種類で、それら輪郭形状の

Table 1 Precision of Free-hand drawings

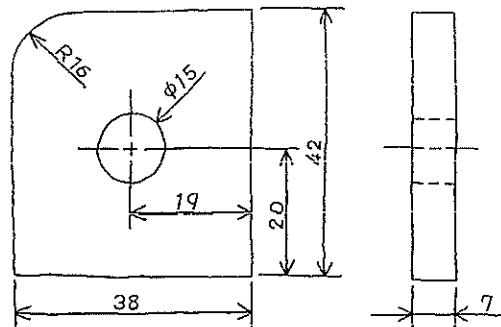
Subject	Straight line (300 mm)		Circle (100 mm diameter)	
	Absolute error	Linearity	Absolute error.	Circularity
H	4.3%	2.70 mm	19.6%	3.30 mm
K	18.9%	1.92 mm	36.5%	6.30 mm
I	5.6%	2.46 mm	6.3%	5.10 mm

Note: Each value is the mean of five trials.

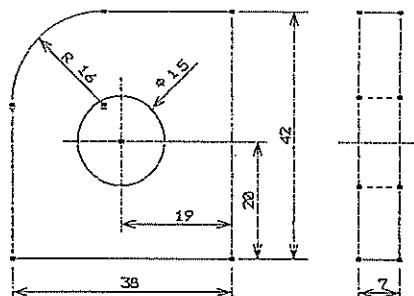
特徴量が安定に検出できることを確認した。Fig. 7 はロボットアームによる実対象物の把握の様子を示す。Fig. 8 は位置計算の結果を示す。計算誤差は、平均値 16 mm、標準偏差が 7.5 mm であり、3 次元の姿勢に関する角度誤差は、平均値 13°、標準偏差 5° であった。認識率は 90%、最終的な把握成功率は 60% 程度となった。

(2) フリーハンド手書き機械図面の認識

フリーハンドで書かれた機械図面をスキャナなどの一括入力方式で読み取り、図面内の情報を認識するための基本的検討と、そのための方法を検討した。フリーハンド手書きの質に関して独自に行った実験によると、Table 1 のような結果が得られている。すなわち、フリーハンドで目のこで直線を引くとき、長さで平均約 4~20% の誤差があり、真直度で全長の 10% 近くである。円を書くと直径は実に 10~40% 近くの誤差である。フリーハンド図の線図形そのものからは、定量的情報はもとより、線の種類ですら機械に認識することは不可能であろう。このような実験結果を参考にして、認識システム内の各種閾値などのパラメータを設計し、文字情報を重視する方針を立てた。まず、入力・変換された図面を原始要素と呼ぶ微小矩形により近似しデータ圧縮を行う。ある程度小さな連結領域を文字・記号として分離・抽出したものを文字・記号データとし文字認識を行う。残りを図形部として特徴点・線抽出の処理を行う。文字認識は輪郭の構造を発見的に認識する構造解析法とフーリエ解析法を併用している。図形部は、まず、ベクトル化され、端点を含め、大きな変曲点を特徴点として抽出する。次に、矢先パターンの認識の後、寸法線および寸法補助線を検出し、先に認識されている文字・記号情報と



(a) Free-hand drawing



(b) Reconstructed drawing

Fig. 9 Recognition of free-hand drawings.

結合することにより、外形線の寸法付けを行う。側面図と正面図との 2 次元対応付けを行った後、寸法の明記されていない外形線部分について寸法付けがなされる。Fig. 9 に入力図面と、連結構造の認識および寸法付けが正しくなされた出力清書図面の例を示す。

(3) 作業動詞を用いるロボット言語

ロボットへの動作指令を、逐一記述しなければならない動作レベル言語ではなく、より高度な作業レベルの記述が可能であるようなロボット言語

の実現を試みた。Fig. 10 は対象とする作業環境を示す。このような環境下の作業に関する動詞を抽出するため、まず 38 人の被験者にそのような動詞を挙げてもらった結果が Table 2 である。さらに、検討を加えて、18 動詞に絞り、それらを Fig. 11 に示すような階層構造として整理した。この言語においては、作業を表す作業動詞に三つのレベルがあり、上位レベルの動詞は、より下位レベルの動詞をモジュールとして組み合わせて実現される。Fig. 12 はロボットアームの動作の作業の様子である。

このように高度に自動化されたシステムでは、不確定的な事態（たとえば、対象物が倒れた）に対応する機能が必要である。その一つとして、言語から視覚情報が利用できる機能を実現した。作業環境上部に CCD カメラを固定し、画像処理プロ

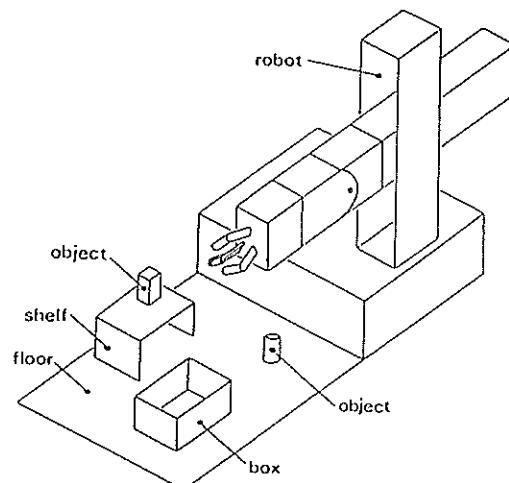


Fig. 10 Circumstances and objects.

セッサを結合して、「VISION」というコマンドを用いて、プログラム内に視覚情報を取り込むことを可能にした。Fig. 13 は画像処理例およびメインコンソール上の表示であるが、2 種類の対象物を 2 次元的に識別し、その位置座標などを算出している。ロボット言語のユーザは、このように算出した位置に基づいて対象物の操作ができる。Table 3 は位置測定の精度を示す。

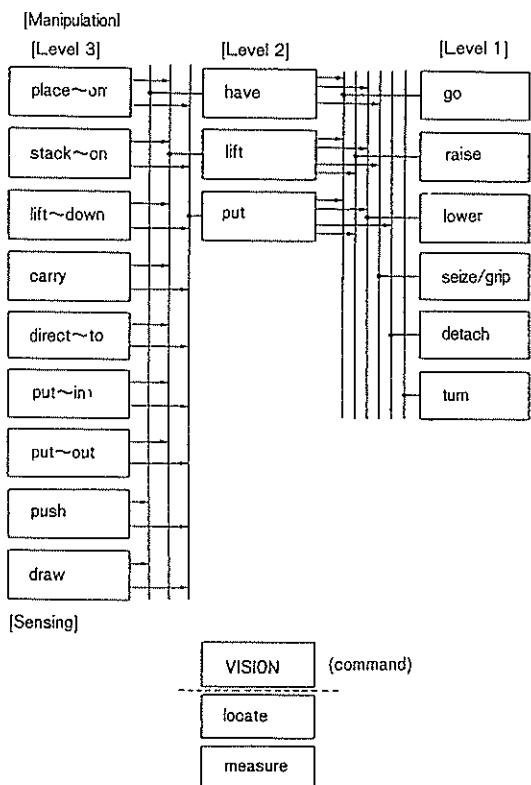
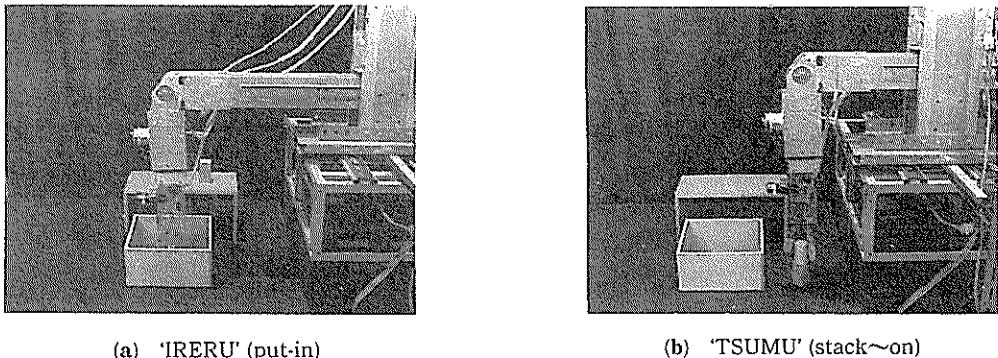


Fig. 11 Hierarchical structure of verbs.

Table 2 Verbs selected to describe manipulation on workbench (by thirty eight subjects)

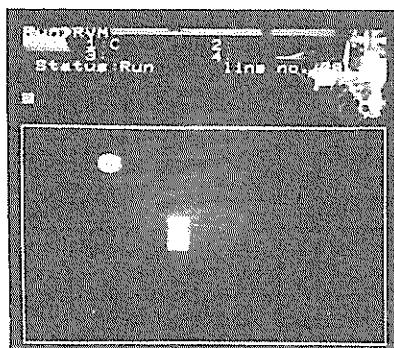
Verbs	Freq.	Verbs	Freq.	Verbs	Freq.
put-in	29	lift, lift-down	19	detach	7
direct-to	27	have, seize	19	approach	5
stack-on, place-on	26	arrange	18	support	3
put	24	touch, hit	15	memory	1
stand, set-upright	23	push, draw	14	destroy	1
release	22	drop	14	exist	1
go	21	throw-away	13	detach (slang)	1
lay, throw-down	20	lift-down	10		
put-out	19	carry	8		



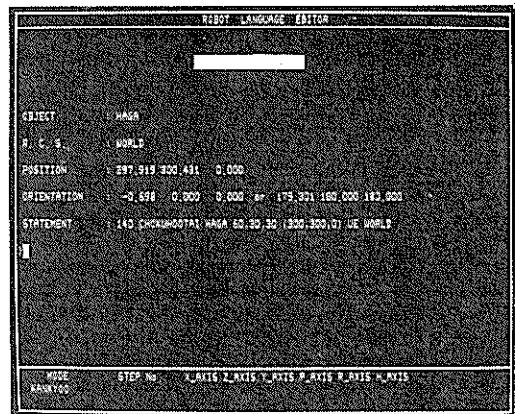
(a) 'IRERU' (put-in)

(b) 'TSUMU' (stack~on)

Fig. 12 Examples of operation.



(a) Original image



(b) Result of 'VISION' command

Fig. 13 Execution of 'VISION' command.

Table 3 Measurement errors of vertices (mm)

	Center		Peripheral	
	X	Y	X	Y
Mean	1.22	1.33	1.15	0.73
Standard deviation	0.10	0.17	0.08	0.44

Note: Each value is the mean of ten samples.

(4) 視覚認識と位置決め動作

既に述べたように、通常の照明下での視覚認識による位置決めシステムの精度はミリメートルのオーダであり、動作範囲の相対的精度で1%程度であった。これは決して高精度とはいえない。しかし、人間の行っていることも、少なくとも部分

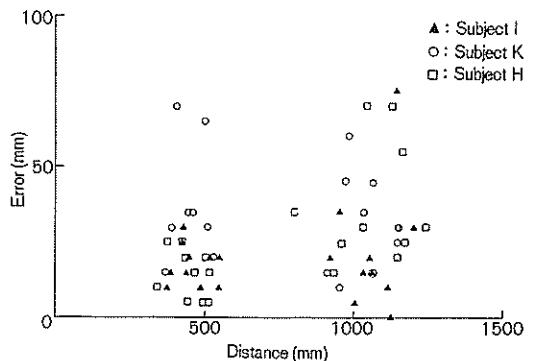


Fig. 14 Error of human positioning.
(after watching of one second)

的には高精度ではないことは、フリーハンド手書きに関する例にも見られたとおりである。

前述の(1)項の(c)に示されたロボットと同様のことを人間が行ってみるとどのようになるであろうか。幅700mm、奥行き1200mmの机の幅側に被験者を座らせ、5mm四方の白片の中央に直径1mmの黒点を打ったマークを面上に置き、被験者は1秒間だけこれを見た後、目を閉じてこのマークを指先で押さえるよう動作する。

Fig. 14は3人の被験者に対し座ったままの範囲と身をのりだす範囲の二通りについて試みた例である。誤差はマークからの絶対距離であるが、平均値で20mmから40mm、相対誤差で2%から7%に及ぶ。実際の場合には連続的な視覚フィードバックによって（あるいは他の感覚も併用して）、高速で正確な動作が遂行されるのである。人間の場合、個々の要素の性能を巧妙に統合する仕組みがあり、個々の性能はそれに必要な範囲内に抑えられているということを再認識せられる。その程度の性能の要素をどのように組み合わせて高い性能を実現しているかという統合の仕組みの問題は今後の課題である。

まとめおよび今後の課題

(1) 本研究によって、人間の知覚、動作特性をとくに視覚と上腕に関わる動作を中心に調べ、視覚認識と物体操作、フリーハンド図の認識、および作業動詞によるロボット言語システムのプロトタイプ・システムを動作させることができた。

(2) 人間のある要素的動作（たとえば、対象位置を視覚的に確認したのち、その位置を指示する；指示された图形をフリーハンドで書く、など）の性能は、個人差はあるが、大まかに位置誤差で数%前後に達する。人間だけが行動しているときはそれら知覚、動作の要素を種々、かつ連続的に統合させて目的を達していると思われるが、人間と人工システムとの融合（あるいはインターフェース）を考える場合には、このような人間のもつ部分的には意外に粗い性能のことも考慮せねばならない。

(3) 今後、(1)項で述べた個々のプロトタイプシステムについては性能の向上をはかり、工業的

成果へと導く努力をしたい。

(4) 本研究により、人間の知覚と動作の統合について興味ある手がかりが得られた。今後、知覚では、色覚、力覚、触覚など、動作では手指、下肢なども含めた統合的なシステムとして考察し、人工システムとのインターフェース、人工システムのもつべき新しいタイプの性能（たとえば、直観に類するようなもの）などを考究したいと考えている。

謝 辞

本研究は昭和61年度から63年度までの日産科学振興財團の研究助成によるものである。この助成により、専用の実験用機器を設置して充実した研究を遂行することができ、またその成果によって国内外での学術交流の実もあがるなど、大きく前進することができた。ここに深甚の謝意を表する次第である。また、この期間に研究室に在籍して本研究に寄与され、研究発表などを共にした諸氏に感謝するものである。

研究発表

- 1) T. Honda: CAM and Computer Vision, The 1st International Conference on Computer Aided Production Engineering, (Invited Paper) (1986).
- 2) T. Honda, S. Kaneko and H. Inazumi: 3-D Shape Recognition of Metal Products, Proceedings of the 2nd International Conference on Computer Aided Production Engineering, pp. 9-15 (1987).
- 3) T. Honda, S. Kaneko, S. Igarashi and Y. Fujii: A Robot Language for Parts-handling Based on Hierarchical Structure of Verbs, Proceedings of International Conference on the Robotics, pp. 1-4 (1987).
- 4) T. Honda, S. Kaneko, S. Igarashi, K. Murata and Z. Z. Cui: A Method for Visually Autonomous Robot Arm to Manipulate Polyhedral Parts, Proceedings of the 6th International Conference on Production Engineering, pp. 772-777 (1987).
- 5) T. Honda, S. Kaneko, Y. Fujii and S. Haga: A Flexible Robot Control System with Both Task-level Language and Visual Monitoring, *Annals of the CIRP*, 37, (1), 29-32 (1988).
- 6) T. Honda, S. Kaneko, Y. Takahashi and A. Watanabe: Recognition of Free-hand Machine Drawings, Proceedings of the 3rd Internation-

- al Conference on Computer Aided Production Engineering, pp. 104-114 (1988).
- 7) S. Kaneko and T. Honda: Visually Controlled Autonomic Robot Arm System for Manipulation of Polyhedral Parts, Proceedings of the 4th International Conference on Computer Aided Production Engineering, pp. 121-126 (1988).
- 8) 金子俊一, 本多庸悟, 池田恵一: 多視野のステレオ画像対を利用した多面体物体の認識, 精密工学会誌, 55, 107-114 (1989).