

II. 物体認識の研究

辻 三郎, 谷内田正考
(阪大基礎工学部 制御工学科)

辻研究室では、人工知能研究の一環として物体認識の研究を行い、同時にその応用を探索している。以下に、実験システムの構成とその問題点を説明した後、最近行なった研究の目的・内容について概説する。

1. システム構成

本研究室における画像処理関係の研究設備を図1に、またそのシステム構成のブロック図を図2に示す。本システムの主な特徴と問題点を以下に述べる。

(a) 高速パターン・プロセッサ図形処理の一つの問題点として、前処理、特徴抽出に時間がかかりすぎるといふ点があげられる。そこで、図に示す高速パターンプロセッサを試作し[1, 2], PDP 8/E がこれを制御しながら処理をするシステム構成を使用している。パターン・プロセッサは、1対の画像メモリの一方の指定された窓内の画像データを、パターン演算ユニットで処理し、その結果を他方に書込む。窓の大きさ、移動、パターン演算の内容は、すべて制御メモリで決定されるので、PDP 8/E から前処理・特徴抽出の内容を自由に選択できるのが、フィードバックを持つ処理に適している。また本システムには、2種のディスプレイを自由に使い分けて処理結果を表示しつつ、対話形式でプログラムを開発できるように会話形モニタが用意されている。

(b) Multi-Terminal System 画像処理の研究では、特に、計算機をオン・ラインで使って、処理結果を見つつプログラムを開発するのが能率的であるが、一方、ひとりの研究者が計算機を独占して使用するため効率が悪い。この点を解決するため、YHP-21MXには効端末から何人かのユーザが同時に計算機を使用できる Multi-Terminal System を採用している。現在の所、YHP-21MXには一台の端末しか接続してないが、増設の予定である。

(c) コンピュータ・リンクーYHP-21MX, PDP 8/E の2台の計算機を使用しているが、それぞれの計算機に接続されているデバイス、あるいはソフトウェアを互いに有効に使用できるように、TTYインタフェイスを介して2台の計算機を

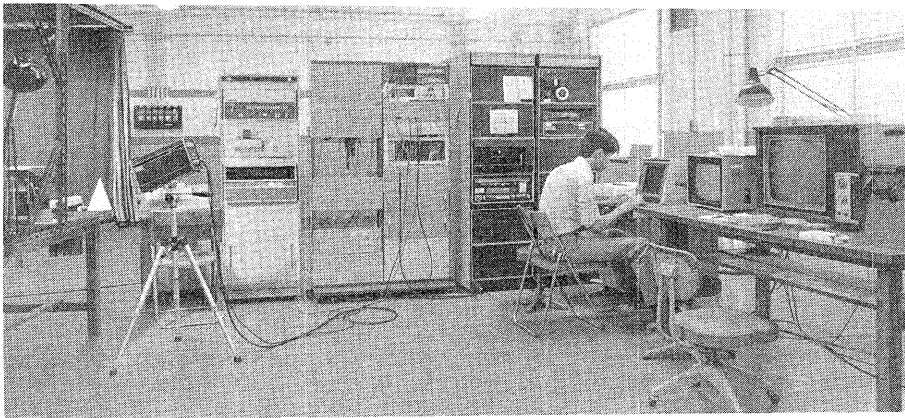


図1. 画像処理システムの写真

接続している。互いの計算機、デバイス間のデータ転送は完了し、現在YHPの Multi-Terminal Systemの下にPDPを管理するシステムを実験中である。また2台の計算機をDMAで結合し、多量のデータ転送も高速に行う予定である。

(d) 画像ファイル管理・転送プログラム—各研究者が開発した種々のプログラムに互換性を与えるため、任意のサイズの各種画像ファイルへのアクセスを管理する画像ファイル管理プログラムが用意されている。画像データは計算機のコアに入り切らないので、このプログラムではディスクにストアされたファイルから必要な部分のデータを計算機のバッファに自動的にR/Wする virtual memory 方式をとっている。また任意のデバイス間の画像データの転送（例えばT.カメラからデータを入力してディスクにストアする、あるいはディスクにストアされたファイルをディスプレイに表示する）を管理するプログラムとして画像ファイル転送プログラムがある。もち論、異なった計算機のデバイス間の転送も可能である。

(e) Virtual LISP—画像処理に、problem solving などの高次の機能を導入するために、LISPを活用している。本LISPはPDP 8/Eに接続している32k bytesのコアを8kw (1w=36bits) として、データセルをストアするのに使用している。しかし、データはコア内に入り切らないので、画像データの場合と同様に、ディスクから必要なセル領域をコア内に自動的にR/Wする Virtual Memory 方式を採用している。

2. テキスチャの解析 [3, 4, 5, 6]

テキスト領域を検出するため、画像の局部領域の統計的性質を算出する非線形フィルタの研究と、テキストを構成する基本粒子の性質の一様性を調べる構造解析の研究を行った。非線形フィルタの目的は、同じテキスト領域内の点は同じ出力の値を持ち、しかも領域の境界はぼけないように画像の平均化を行う。

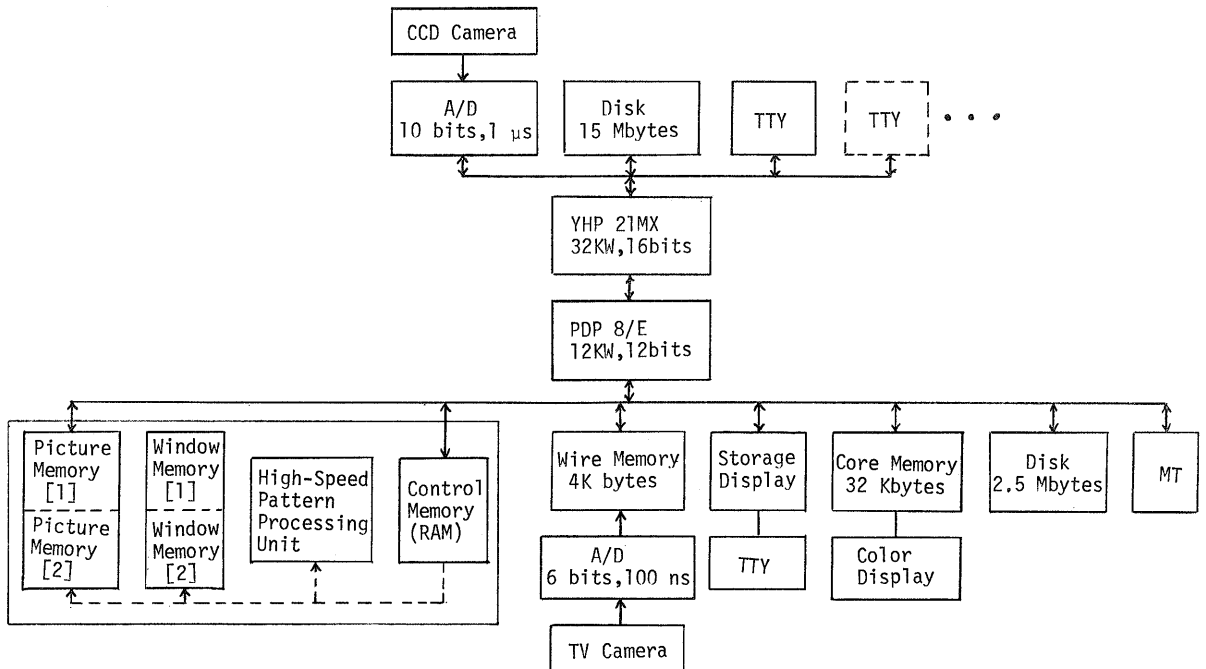


図2. システム構成

従って、画像内の各点について最良の大きさで、境界を含まない最適位置の平均化領域を発見しなければならぬ。図3(a)に示すように、注目の近傍を4象限に分割すると、少なくとも1象限は、たゞ一種類のテクスチャ領域しか含まない。そこでまず、各象限の正方形近傍を、その平均濃淡レベルの変動が小さい値以下になるまで拡大し平均化領域の最良の大きさを求める。次に最良の象限を決定する。例えば、図3(b)のように、才1象限と才3象限を比較すると、最良の大きさの半分の領域を2つ接続した領域の平均値は、境界を含まない才1象限の平均値により近い。このようは比較を各象限間について行い、最も差の小さい象限を最良とする。この最良の象限内で最良の大きさの領域の平均値を注目点の出力とする。図4(a)の入力画像から(b)のフィルタの出力を用いてregion法により(c)の領域を得る。

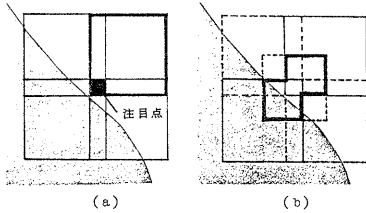


図3. 非線形フィルタ

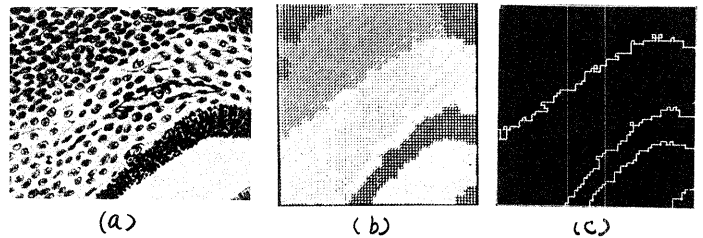


図4. 実験結果の1例

構造レベルでは、テクスチャの基本粒子の性質・配列などの分布を調べて解析を行おう。まず領域法により基本粒子を抽出し、次に各粒子の属性として、大きさ(面積, 周囲長), 形状(5種のモーメント), 密度などを計算し、各属性のヒストグラムから明らかに分類可能な属性を用いて領域分割を行おう。例えば図5の入力画像の粒子の属性の中から、図6に示す周囲長のヒストグラムにより粒子が分類され、粒子間の接続性などがテストされ、背景と図7のような立方体領域が得られる。次に各領域ごとに粒子の属性のヒストグラムを作り、分類可能な属性が発見されると、再分割を実行して行く。図7の粒子のモーメント M_{11} のヒストグラムを図8に示す。これを用いて、領域 R_1 を図9に示す3つのテクスチャ領域に分割できる。

3. 曲面体の認識 [7, 8]

画面の特徴を記号で表わすと、それらの記号を文法を用いて処理することが出来る。このような文法的手法を用いて、入力画像から平面と曲面を分離し認識する方法を開発した。図10のような入力画像に対して、文法的手法を用いて、平面と曲面に分離した結果を図11に示す。

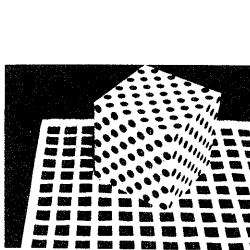


図5. 入力画像

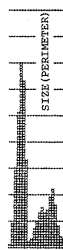


図6. 周囲長

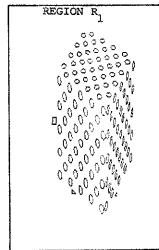


図7. 周囲長による分割



図8. M_{11}

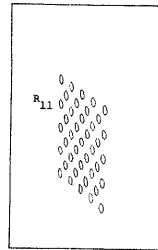
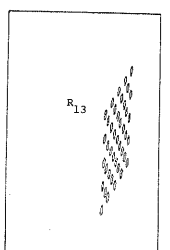
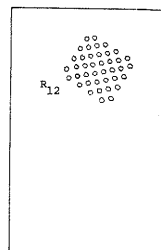


図9. R_1 の再分割



4. 機械部品の学習認識 [9, 10, 11, 12]

複雑な物体の例として、あるいは物体認識技術の一つの応用例として、図12に示すような機械部品を学習・認識するシステムの研究を行なった。このような複雑な物体から主要な特徴を信頼性良く抽出するにはBottom-upな方法では困難があり、対象の構造に関する知識を用いたTop-downな方法を用いる必要がある。しかし、従来のTop-down

な方法は、対象が顔写真というように限定されている場合に、その部分構造(鼻など)を同定するには良いが、機械部品のように対象の種類が多い場合に、あるいは入力物体が“何であるか”を認識するには適用できない。ここでは、図13に示すように、入力物体に関して現在得られている特徴と各物体のモデル(知識)とを比較して、次にどのような特徴を抽出するのが最も有効かというstrategyを決定する問題解決機能をscene analyzerに与えることにより解決している。まず、認識への糸口となる主要な特徴として、物体の大きさ、外形などの特徴が抽出され、これを基に、いくつかの候補モデルが決定される。次に、scene analyzerはこれらの候補モデルの構造を調べて、次にどのような特徴を調べるのが認識に最も有効であるかをheuristicに予測し、次にとるべきstrategy(その特徴を抽出するのに使うspecialistsのタイプとその特徴の予測位置)を決定する。提案された特徴は、特徴抽出ルーチン(hole finder, line finder, texture detector等)の集まりである

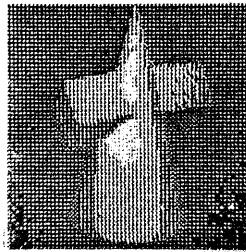


図10. 入力画像

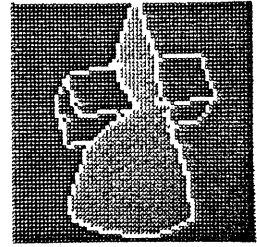


図11. 処理結果

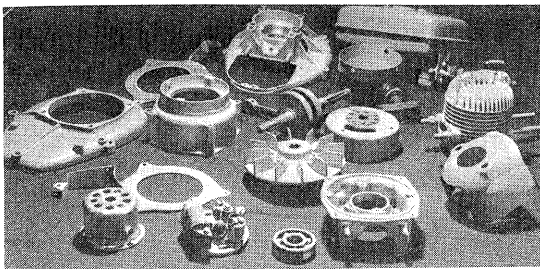
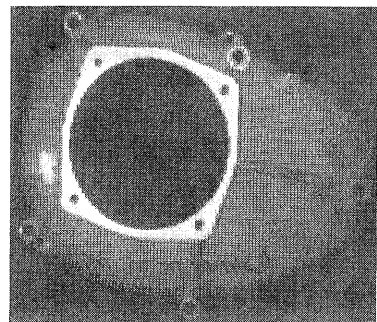


図12. 機械部品の例



(a) 入力物体

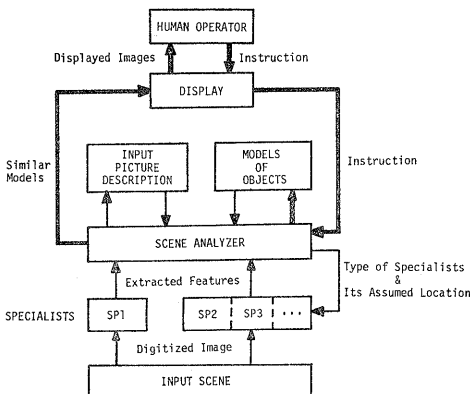
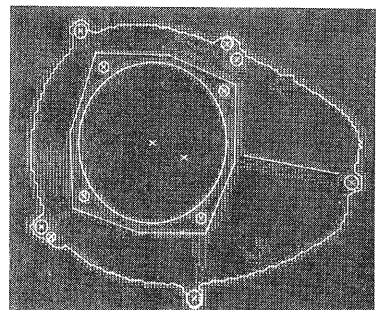


図13. 学習・認識システム



(b) 処理結果

図14. 実験結果の1例

specialists により、入力画像から抽出される。抽出された特徴は scene analyzer に
より評価され、再び次に調べるべき特徴が決定される。図14に実験結果の一例を
示す。入力物体は部品番号20の上面と認識され、同時にその位置、方向と紐立て
に必要なボツ穴の位置も計測される。また、本システムには実例物体を機械に提
示し、その主要な特徴をディスプレイを介して interactive に教示するだけで、物
体の構造モデルが自動的に作成されるという学習機能が備わっているので、新し
い機械部品に対しても容易に適応できる。

5. 積み重なった部品の認識 [13, 14]

上で述べた学習認識システムでは、実用性という点を考慮して物体を上方から
見て、しかもそれらが重なり合っていないという制限をおいた。しかし、積み重
なった部品を任意の方向から見て認識するのは、人工知能的な観点からも興味あ
る問題であるので、積み重なった部品の内、隠されていない部品を認識する研究
を行なった。この場合、上方から見た場合とは異なり、見る方向により物体の形
が異なる、また重なり合った物体からいかに物体を分離するかという問題が生じ
る。これらの点を解決する方法として、まず入力画像から楕円などの見られた特
徴を見つけ、それを手がかりにして物体の原型像に逆変換する。次に、変換され
た画像中から、円弧、穴などの強い特徴を見つけ、モデルとの対応を調べ、欠け
ている特徴を推測して、弱い特徴を見つけ認識する方法をとっている。図16は四
角の入力物体から原型像に逆変換し、穴、円弧などの特徴を抽出した結果であり
、図17はそれらからモデルを使って弱い特徴も見つけた結果である。

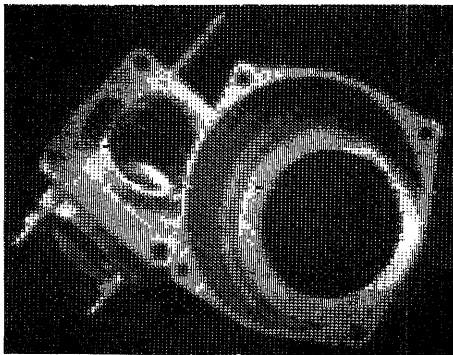


図14. 入力画像

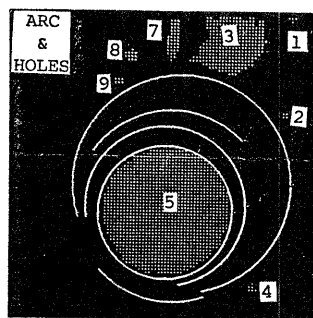


図15. 穴と円弧

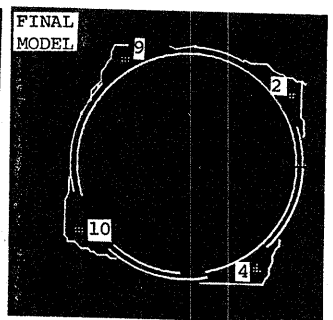


図16 処理結果

6. 回転台上の物体のパラメータ測定 [15]

一方向から見た画像だけで、隠されている物体を認識するのは難かしい。また
その3次元的位置を計測するのも困難である。このような場合、多方向から見た
結果を総合すると、一画面では解決できないあいまいな点も解決でき、3次元の
位置も求めることができる。その基礎的研究として、計算機で制御される回転台
上に置かれた多面体を2方向から見て、その3次元的位置を計測する研究を行な
った。この方法の特徴は、一般のステレオ法とは異なり、T.V.カメラが一台で良く、
また複数画面間の視差を大きくとれるので、測定精度が高いという長所がある。
その反面、図17に示すように、視差を大きくとれば画面間の特徴点を対応させる
過程が複雑になるという問題がある。ここでは、回転前後における頂点形状、稜
の傾斜角などの遷移にいくつかの法則が存在することを利用して、2画面の特徴
点の対応づけを解決している。

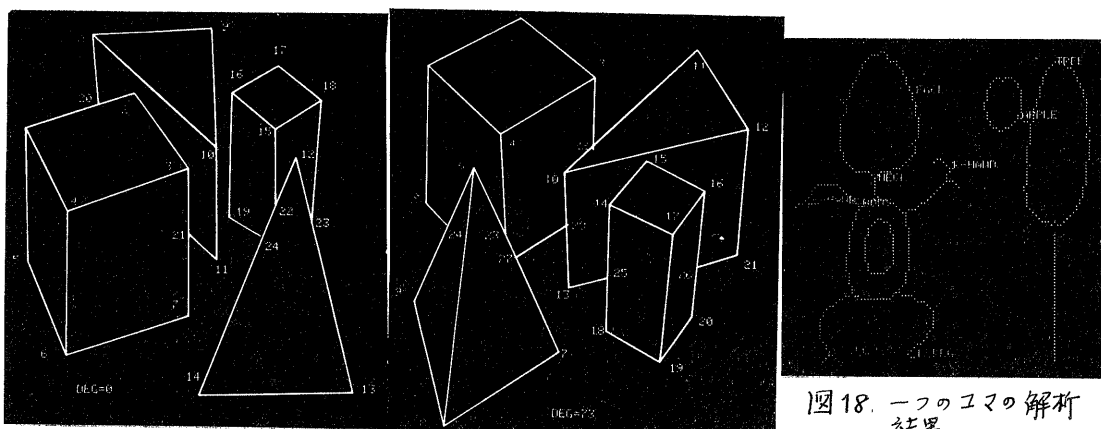


図17 (a) 回転前

(b) 回転後(73°)

図18. 一つのコマの解析結果

7. 動画像の質問回答システム

動き回る物体を解析するには、静止物体とは異なり時間、行動などの新しい概念が必要と becoming。その一例として、動画から主人公の行動を理解し、動画の内容に関する質問に対して回答するシステムの研究を行った。動画の場合、情報量が圧倒的に多いという問題があるが、一コマずつの動きは少ないという性質を利用すれば、以前のコマの解析結果を利用して現在のコマを処理できる。また一画面ではあいまのない点も前後のコマとの関連から解釈できるという利点がある。各コマ毎の解釈が終了後、コマ間の物体の動きから直接的に導出できる動作を調べる。これらの primitive な動作から、より高次の物体間の相互作用を行動に関する知識を用いて推論する。こうして導出された行動・動きに関する記述を用いれば、動画に関する質問、例えば "What did Besi?" に対して "Besi caught an apple." と回答できる。図18にあるコマの解析結果を、図19にいくつかの連続したコマから動作を推論していく過程を示す。

8. CCDカメラの学習制御

動き回る物体を解析するには、その視対象の動きに応じて、素早くカメラを制御し、視対象をカメラに捕える必要がある。このため直径6cmの球にレンズ系およびCCDを内蔵した小型カメラとカメラを連応性良く学習制御するシステム

MOVIE FRAMES	1st F	2nd F	3rd F	4th F	5th F	6th F
BESI 's center	(45, 89)	(50, 89)	(52, 89)	(55, 90)	(55, 85)	(55, 79)...
TREE 's center	(102, 56)	(104, 57)	(105, 56)	(104, 57)	(103, 57)	(103, 55)...
APPLE's center	(85, 43)	(86, 43)	(86, 42)	(86, 42)	(86, 42)	(85, 41)...
	(M BESI RIGHT)	"	"		(M BESI UP)	"
SCENE GROUP	SG=1				SG=2	...
PRIMITIVE	(M BESI RIGHT)				(M BESI UP)	
ACT & DESCRIP	(N BESI TREE)				(N BESI APPLE)	
DEDUCED ACT	(WALK BESI TREE)				(JUMP BESI APPLE)	

図19. 連続したコマからの動作の推論過程

を試作した。カメラは図21に示されているスピーカ型ソレノイドにより駆動されるが、その制御信号はカメラを動かしている間に準最適値が学習され、視対象をより早く捕えられるようになっていく。カメラの回転速度は約 $1000^\circ/\text{sec}$ である。

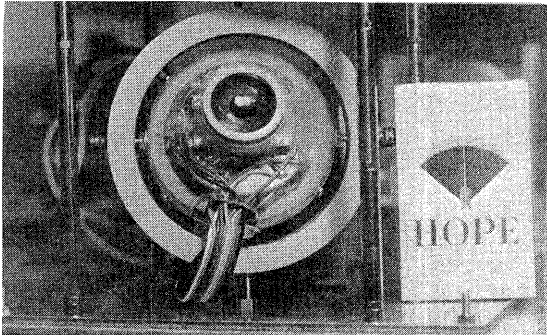


図20. CCDカメラのクローズアップ

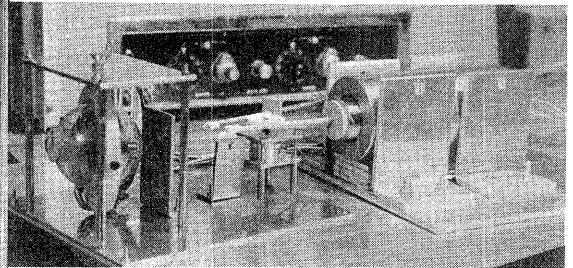


図21. 眼球運動の制御システム

謝辞 以上の研究には文部省科学研究費(一般, 特定, 試験)の補助を受けた。
[参考文献]

- 1) 谷内田, 辻: パターン情報処理装置について, 信学会パターン認識と学習研 誌 72-69 (1972)
- 2) 谷内田, 富田, 辻: 3次元物体の高速パターン処理システム, IEE Trans. 58D, 4 (1975)
- 3) 富田, 谷内田, 辻: Textureの判別, 信学会パターン認識と学習研 誌 72-69 (1972)
- 4) Tomita, Yachida, Tsuji: Detection of Homogeneous Regions by Structural Analysis, Proc. of Int. Joint Conf. on Artificial Intelligence (IJCAI), (1973)
- 5) Tsuji, Tomita: A Structural Analyzer for a Class of Textures, Computer Graphics & Image Processing, 2, 3. (1973)
- 6) 富田, 谷内田, 辻: テキスチ識別のための非線形フィルタ, IEE Trans. 58D, 2 (1975)
- 7) 藤原, 辻: 曲面体の認識, 信学会パターン認識と学習研, PRL 73-95 (1974)
- 8) Tsuji, Fujiwara: Linguistic Segmentation of Scenes into Regions, Proc. of Int. Joint Conf. on Pattern Recognition (1974)
- 9) 谷内田, 辻: 機構部品の学習認識システム, 信学会パターン認識と学習研, PRL 74-52 (1975)
- 10) Yachida, Tsuji: Machine Vision for Complex Industrial Parts with Learning Capabilities, 4th IJCAI (1975)
- 11) 谷内田, 辻: 機構部品の認識, 情報処理学会イメージプロセッシング研 2 (1975)
- 12) 谷内田, 辻: 機構部品の学習認識システム, IEE Trans. (1976)
- 13) 中村, 辻: 積み重なった部品のシーンアナリシス, 信学会パターン認識と学習研 PRL 74-51 (1975)
- 14) Tsuji, Nakamura: Recognition of an Object in a Stack of Industrial Parts, 4th IJCAI (1975)
- 15) 菊池, 辻: 回転台上の多面体パラメータ測定, 第4回SICE学術講演会 (1975)