

遺伝的アルゴリズムによる物体ラベリングの一手法と ロボットビジョンへの応用

橋本 学 鷺見和彦
三菱電機(株) 産業システム研究所
〒661 尼崎市塚口本町 8-1-1

email : hasimoto@fas.sdl.melco.co.jp

あらまし 視野内に複数の同一形状の物体が存在するシーンを認識する手法を提案する。画像から抽出した誤ったものを多数含む物体の候補群に対して、各々の候補の存在有無をラベルづけし、「存在する候補」の集合として画像を認識する。ラベルづけ問題をラベル組合せの最適化ととらえ、遺伝的アルゴリズムを適用した手法を提案する。近傍の物体同士の干渉度合いと、候補各々のモデルフィット率を用いた評価関数を適応度に組み込んだ。解をコーディングした染色体は、遺伝的操作によって物体間の局所的干渉を避けながら徐々に全体として望ましい画像解釈を生成するようにふるまう。本手法を用いて、ランダムに積まれた多数のダンボール箱を認識するロボットのためのビジョンシステムを構成し、実験により実用的な性能であることを示した。

和文キーワード ラベルづけ / 遺伝的アルゴリズム / ハフ変換 / デパレット化ロボット

A labeling method using genetic algorithm and its application to robot vision

Manabu Hashimoto Kazuhiko Sumi
Industrial Electronics and Systems Laboratory
Mitsubishi Electric Corporation
1-1, Tsukaguchi Honmachi 8-chome, Amagasaki, 661 Japan

Abstract We propose a method to recognize a scene, which consists of same shaped objects. Many object candidates including false ones are extracted, and each candidate is labeled as "exist" or "not-exist". The scene is recognized using only "exist" object. Genetic algorithm(GA) was applied to the labeling as a combinatorial optimization problem. Fitness function represents positional relation of neighboring objects and the model-fitting ratio of each candidates. Chromosomes behave as they generate a global solution avoiding local inconsistency by genetic operations. We developed a vision system for robot which can recognize stacked carton boxes, and proved its practicality.

英文 key words labeling / genetic algorithm / Hough transform / depalletizing robot

1 はじめに

工場や自動倉庫の出入荷場などの物流拠点では、コンベヤーを流れる荷物をパレットに積み作業(パレタイズ)、あるいは反対にパレット上の積荷をひとつづつ荷下ろしする作業(デパレタイズ)が頻繁に行なわれるが、これらは多大な労力を必要とするので自動化が強く望まれていた。ロボットによる自動化を考えた場合、パレタイズではコンベヤーの機械的位置決めを利用できる。しかしデパレタイズでは、輸送中の荷くずれや作業者の積載ミスにより積荷の位置が不確定であるため、完全な自動化のためには積荷の位置を認識するセンサ系が不可欠である。

この課題は一般的には3次元認識を必要とするため、従来からレンジファインダを利用した方法^{[1][2]}が研究されてきた。しかし、物流システムにおける管理情報、例えば積荷の形状や寸法データ、個数、積み段数などを利用すれば問題を単純化できる。即ち、問題を(1)山積みされた荷物の上方に設置されたカメラによる画像から物体の2次元的位置と相対的な上下関係を認識する問題と、(2)必要に応じて物体の絶対高さを計測する問題の2つに分解し、結果を統合させてロボット動作に必要な情報を生成する。前者の問題に対するアプローチとして、テンプレートマッチングは汎用的な手法と言われている。しかし、物体の回転ずれを認識するためには、少しずつ回転させたテンプレートが必要とし、高速化アルゴリズム^[3]を用いても処理時間は実用的でない。

本報告では、簡単な方法を使って1枚の画像から対象物体らしいものをたくさん抽出して物体候補群とし、別途用意した評価基準を用いて各候補に真偽のラベルをつけ、正しい物体候補のみで画像を再構成することで画像全体の認識を行なう方法を提案する。ラベルづけ問題は可能なラベルの組合せ最適化と考え、遺伝的アルゴリズム(GA)^{[4][5]}を用いた手法を提案する。GAのbuilding-block仮説を有効に利用することで、効率のよい探索が可能である。

まず我々の扱う認識問題の性質(前提)を明らかにする。次いで提案した認識アルゴリズムについて、前段の物体の直線輪郭に注目した物体候補の抽出処理、および後段の遺伝的アルゴリズムを用いたラベ

ルづけ処理をそれぞれ説明する。最後に本手法に関して染色体群のふるまい、他の手法との処理速度比較、および開発した実システムにおける認識性能について実験結果を明らかにする。

2 物体認識アルゴリズム

ここでは、我々の扱う認識問題の性質と提案する認識アルゴリズムの概要を述べる。認識対象に関する前提と、利用可能な情報は次のとおりである。

1. 認識すべき物体は寸法既知の箱状物体である。
2. パレット上の物体は1種類である。
3. 積荷を上方から見た場合、画像には高さ2段分の物体しか存在しない。但し個数は未知である。
4. パレット上の積荷の段数は既知である。

上記のような情報は、物流システムにおいては通常利用可能なものである。我々のゴールは積まれた物体の画像を処理し、物体の2次元的位置と、相対的高さ(上下どちらか)を知ることである。ワークの絶対的高さは、上記の情報を組み合わせることで得られる。ロボットの動作タクトへの影響が少ないことから、カメラは対象物の上方に設置する。図1に示したのは、原画像の一例である。物体の輪郭は必ずしも明確でなく、梱包テープや印刷文字が画像を複雑にしている。



図1: 積まれたダンボール箱の画像例

認識アルゴリズムの流れを図2に示す。前段では、物体の部分情報の組合せとして物体候補が多数抽出される。部分情報としては、扱う物体が箱状であることから直線輪郭を利用する。物体候補は直線輪郭に囲まれた矩形パターンである。部分情報の抽出は

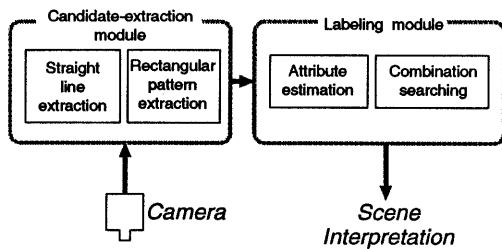


図 2: 輪郭情報とラベルづけに基づいた箱状物体の認識アルゴリズム

高速にできるが、その反面、画像上の外乱成分の影響を受けやすく、誤った物体も候補として抽出されることがある。後段では候補それぞれのもつ属性を使って、正しい候補と誤った候補、また候補の長さ属性をラベルづけにより分類する。最終的に、誤った候補を除外して正しい候補のみで画像を構成し、認識結果とする。このアルゴリズムはボトムアップ手順であり、各モジュールの役割が明確で処理が簡単であり、モジュール間の情報フローも単純なので高速な処理系を作りやすいという利点がある。

このような、認識要素の抽出とそれらのラベルづけを組み合わせた認識ストラテジーにおいて、認識要素としてどのようなものを用いるかは重要な問題である。認識要素として、一つ一つの画素、あるいは直線やコーナーパターンといった形状プリミティブを用い、それらの組合せとして徐々に高次の表現を作り上げていくアプローチは興味深い。しかし、ラベルづけ過程で扱うデータ量が大きく、最適な組合せを求めるのに膨大な処理時間がかかることから、工業用ビジョンシステムには不向きであると考えられる。我々はひとつの物体の外形を表す矩形パターンを認識要素とし、それらを数多く抽出することで候補に冗長さを持たせつつ、データ量のある程度抑制する。これによってラベルづけ段階での組合せを計算する時間が短縮され、実用的な時間で近似解を得ることができる点を重視した。

3 直線輪郭に着目した物体候補の抽出

原画像からエッジを検出し、ハフ変換^[6]で直線を抽出する。画像上の不要なエッジ点はハフ空間のS/N

比を劣化させるので、簡単なエッジ追跡によって短い線セグメントを求め、線セグメント単位でハフ空間への投票を行なう。投票値は線セグメントの長さとする。投票座標 (ρ_s, θ_s) は以下の式で求められる。

$$\rho_s = x_k \cdot \cos \theta_s + y_k \cdot \sin \theta_s, \quad k = 1, 2 \dots N \quad (1)$$

ここで、 (x_k, y_k) は線セグメントのエッジ点列であり、 N はその長さである。投票座標 (ρ_s, θ_s) は、検出点列 (x_k, y_k) と直線方程式の距離の二乗和が最小になる方程式に対応するハフ空間上の座標として明示的に計算される。

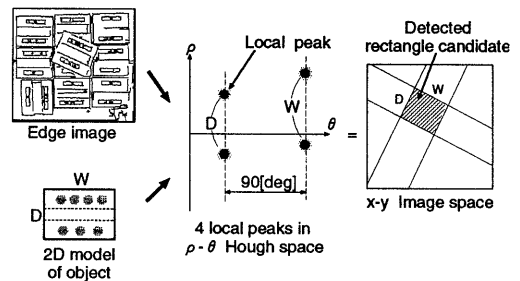


図 3: ハフ変換を用いた物体候補の抽出

図 3は矩形パターンとしての物体候補抽出の様子を示す模式図である。システムは抽出したい矩形パターンのモデルを持っており、ハフ空間上での局所ピーク 4 点の組合せがモデルの寸法 (D, W) に合致する場合に、候補が検出されたとして登録される。

エッジ画像は上段の物体だけでなく下段の物体も含んでいるので、大きさの異なる 2 種類のモデルを用意して両方も抽出する。物体の厚みが大きい時は矩形パターンの大きさによる上下段の区別が容易であるが、厚みが小さい時はこの区別はいまいなものになる。この場合は、同一位置で上下段両方に対応する候補が発生することを許す。またエッジ検出の段階で、矩形パターンを構成する 4 つの輪郭が必ずしもすべて明確に検出できないこともあるので、内部モデルの 3 辺だけで対応づけるなど柔軟に対応させる。その結果、梱包用のテープや箱表面の図柄による誤った候補の抽出もあるが、正しい候補の見逃しは少なくなる。

4 遺伝的アルゴリズムによるラベルづけ

4.1 ラベルづけによる画像認識

本手法では、画像の記述は多数の物体候補をラベルづけすることによって得られる。ここでは物体の相対の高さを示す情報は既に候補自身が持っているため、可能なラベルは「存在する」(正しい候補)あるいは「存在しない」(偽の候補)の2種類である。図4のように、存在するというラベルが付けられた候補群によってシーンが構成され、画像が認識される。

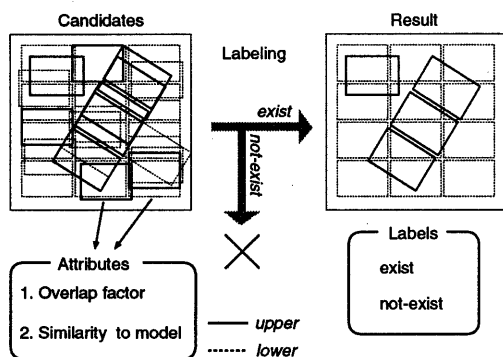


図4: 物体候補のラベルづけ

全ての候補に関し、重なり度係数とモデルフィット率の2つの属性があらかじめ計算される。重なり度係数は、ある候補とその近傍にあるいくつかの候補との相対的な位置関係をもとに計算され、近傍の物体との相互干渉が大きいほど係数は大きくなる。但し、異なる高さにある物体同士は2次的に重なりがあっても差し支えない。また、物体候補各々のモデルとの類似性を示すモデルフィット率は、候補に対して輪郭モデル^[7]をあてはめることで計算される。この方法は、濃淡パターンを直接用いる相互相関に比べて信頼性の点で劣るが、高速に計算できるという利点がある。

候補をラベルづけするにあたっては、画像の局所的な部分に対しては相互干渉すなわち重なり度係数の小さいことが重要である。これは解の部分的無矛盾性を維持するという意味である。また、一方では認識結果に含まれる物体のモデルフィット率は、できるだけ高い方が好ましい。このような評価基準のもとで、全ての物体候補それぞれに対して上記2種

類のラベルのどちらかを与えることによって、好ましい解が得られる。

4.2 染色体のコーディング方法と適応度

遺伝的アルゴリズム(GA)^{[4][5]}は、生物の進化の過程を表現した最適化のための計算手順の一つである。またこれは、多点探索とランダム探索の両方の側面を持つが、複数の部分解の組合せによってより好ましい解を生成するという building-block 仮説を反映しており、一般には大域探索に優れて大局解への収束は速いといわれている。

GAでは、問題をどのように染色体として表現するかは重要である。我々の提案する方法は、図5のようなものである。物体候補それぞれを染色体上のビットに対応させ、各ビットはラベル「存在する」の場合は値1を、「存在しない」の場合は値0をとる。したがって、染色体の長さは候補数と同じになる。1本の染色体は、ビットの値1の候補のみからなる画像として解を与える。

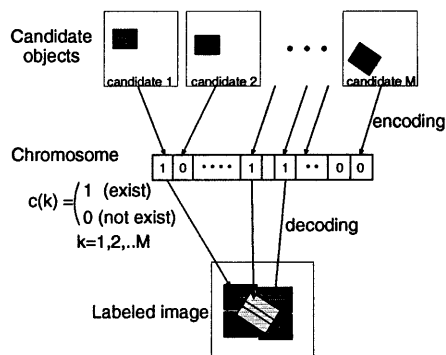


図5: 染色体のコーディング方法

適応度関数は個体集団のふるまいに大きな影響を与える。個体集団中の i 番目の個体(ここでは一つの個体は一つの染色体から構成される)の持つ染色体の適応度 $F(i)$ を、次式のように定義する。

$$F(i) = w_O \cdot F_O(i) + w_R \cdot F_R(i) + w_N \cdot F_N(i) \quad (2)$$

ここで、 F_O は重なり度合いに関する係数、 F_R はモデルとのマッチングによって得られた類似度から計算される候補のモデルフィット率に関する係数、

F_N は画像に含まれる候補数が多くなるほど大きくなる係数である。 w_O 、 w_R 、 w_N は重み係数であり、実験的に決められる。適応度の計算を容易にするために、これら3つの項は線形結合されている。また、 F_O 、 F_R 、 F_N はそれぞれ以下の式で定義され、0から1までの値をとるように正規化されている。

$$F_O(i) = 1 - \frac{1}{\sum_{k=1}^M c_i(k)} \sum_{k=1}^M \{c_i(k) \sum_{l=1, l \neq k}^M \{c_i(l) \frac{S_{k,l}}{A_k}\}\} \quad (3)$$

$$F_R(i) = \frac{1}{\sum_{k=1}^M c_i(k)} \sum_{k=1}^M \{c_i(k) \frac{\text{Score of } O_k}{\text{Full score of } O_k}\} \quad (4)$$

$$F_N(i) = \frac{\sum_{k=1}^M c_i(k)}{M} \quad (5)$$

ここで、 $c_i(k)$ は i 番目の染色体の k ビット目の値であり、1か0のどちらかをとる。 M は染色体の長さである。 O_k は k 番目の候補を表す。また、 $S_{k,l}$ は候補 O_k と O_l との画像上での重なり部分の面積、 A_k は候補 O_k の矩形パターン面積である。

(2) 式において、第1項は画像の局所的な部分においては、物体間の重なりのような矛盾を起こしてはいけないことを表現している。第2項は画像全体が信頼性の高い候補から構成されるべきであることを表現している。この2つの項は、画像の部分的、および全体的それぞれにおける解の好ましさを表しているが、解に含まれる物体の個数という点では同じようなふるまいをする可能性がある。つまり、信頼性の高いわずかな候補の作る局所解に陥りやすい。そこで、それらの項と相反する項として第3項により候補数が多いほど適応度が高くなるようにし、解の動きを多様にした。

4.3 遺伝的操作による解の探索

GA では、世代交代の過程で、効果的な解の探索のために遺伝的操作を確率的に行なう。図6と図7にそれぞれ用いた交叉と突然変異を示す。

一様交叉は、ランダムに設定されたマスクビット列を用いる。子 a は、マスクの値が1のビットについては親 A から、0のビットについては親 B から値を受け継ぐ。子 B に関しては反対の親から受け継ぐ。染色体上でのビット並びは、必ずしも画像上での位

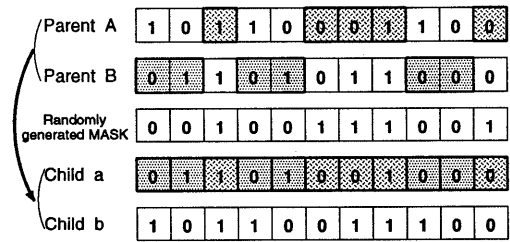


図6: 一様交叉

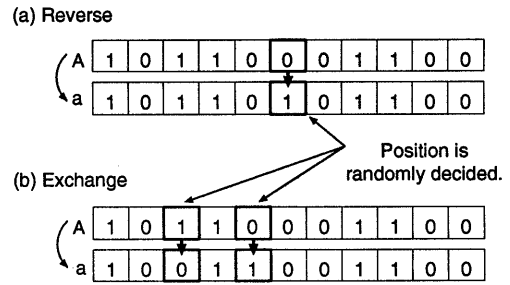


図7: 2種類の突然変異

置関係を反映しているわけではないので、ビット並びに依存しない一様交叉を用いた。

突然変異としては、反転タイプと交換タイプの2つを用いた。反転タイプはランダムに選ばれた位置の値を反転させるもの、交換タイプは染色体上の値が1のビットと0のビットをそれぞれ1箇所づつランダムに選び、それらの値を交換するものである。交換による突然変異は、画像中の「存在する」候補の数を変化させずに新しい解を生成する操作に相当する。

重要な点は、我々の提案した染色体のコーディング方法と遺伝的操作の組合せは、building-blockの考え方に沿っているということである。画像認識に関しては、画像の部分的な無矛盾(あるいは矛盾の少ない)の解の組合せとして、画像全体としても好ましい解を生成する必要がある。GAでは高い適応度を持つ解の形質を受け継いだ次世代の集団が、より高い適応度を持つことを期待するのである。この点において、GAによるラベルづけは例えば単純なランダムサーチに比べて高効率であると考えられる。

5 実験と考察

5.1 列挙法との比較

本手法の処理速度を他の手法と比較する実験をおこなった。比較の対象としては、単純ツリー探索による列挙法を選んだ。この方法は、与えられた全ての組合せをツリーで記述して順次試行するが、あるノードで一定しきい値以上の評価値が得られなかった場合にはそのノードから先のツリーは探索しないで次の枝に移るようにした。

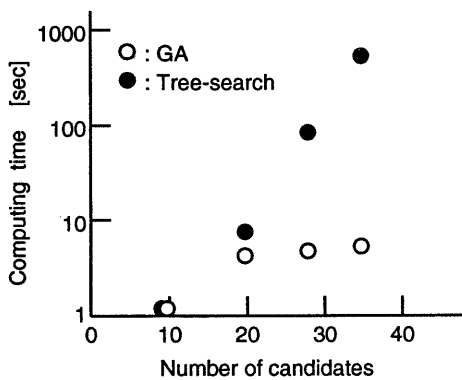
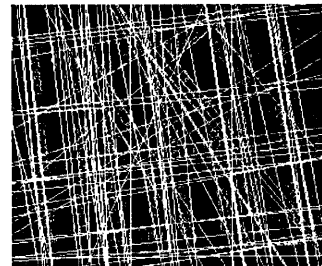


図 8: GA とツリー探索との処理時間の比較

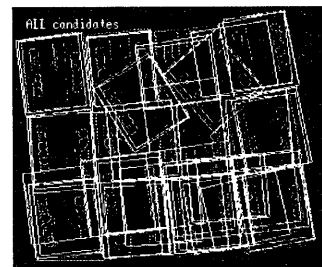
結果を図 8 に示す。実験には 12 個の段ボール箱からなる画像を用い、35MIPS の計算機によるソフトウェアで実行し、処理時間を実測した。どちらの手法も物体候補は同じものを用い、ラベルづけにおける評価関数なども全く同一のものを用いた。GA では、個体集団の平均適応度の変化量が 300 世代にわたってしきい値以下ならば世代交代を停止させた。GA によるラベルづけは候補数が多くなるほど上記の列挙法に対する高速性が顕著になり、候補数が 35 くらいになると 100 倍程度高速であることがわかる。また、両手法において認識結果の正しさに大きな差はなかったが、ツリー探索のほうが最終解の評価値がやや高いということは起こった。GA による探索では、正しい解に接近するのは速いが、その近傍を細かく探索する能力は不十分であることを示している。

5.2 世代交代における染色体のふるまい

次に、本手法によるラベルづけの過程で、時間とともに染色体がどのようにふるまっているかを調べた。図 9 は物体候補を抽出する際の処理画像である。同図 (a) にハフ変換で得られた直線群、(b) に直線の組合せとして得られた物体候補群を示した。なお、実験に用いた原画像は図 1 と同じものである。また、初期交叉率は 0.9、初期突然変異率は 0.5 とし、世代交代につれて徐々に下げた。個体数は 46 であった。



(a) ハフ変換によって検出された直線群
(画像サイズ = 362 × 254)



(b) 抽出された 46 個の物体候補群
(偽の候補を含む)

図 9: 検出された直線と候補の例

図 10 に世代交代における適応度の履歴を示す。集団の中で最も適応度の高い個体は階段状に適応度を上げており、遺伝的操作によって偶発的に高い適応度を持つ個体が発生したことがわかる。一方集団中の平均適応度は徐々に増加し、集団中に適応度の高い個体が徐々に広がっていることがわかる。

図 11 は、集団中で最大適応度を持つ個体の表現する解を、世代交代の途中で画像にデコードしたもの

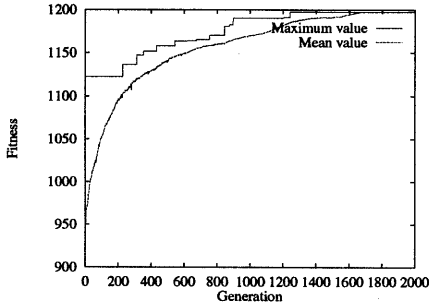


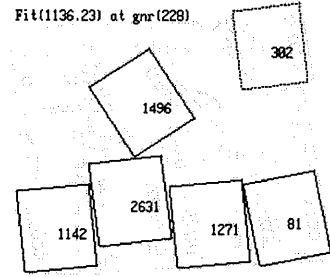
図 10: 適応度履歴曲線 (個体数 46)

である。画像 (a)(b)(c)(d) は、それぞれ世代数 228、353、674、1241 のものである。図では上段の物体は実線、下段の物体は点線で描かれており、物体についた番号は候補につけられた ID である。(a) では、6 個の物体が“存在”しており、それらは互いに 2 次元の干渉を起こしていない。同図 (b) では候補 1493 が候補 302 と入れ代わっている。その理由は、候補 1493 は 302 よりも高いモデルフィット率を持つからである。この変化は交換タイプの突然変異によって引き起こされた。(c) では、候補 2673 と候補 2679 が交叉によって追加される。また、この時点では上段の候補 1496 が削除されて下段の候補 1 が追加されているが、これは誤りである。(d) は平均適応度が一定期間安定しているため、最終解と認められたものである。候補 1496 も再生して、先ほどの誤りはなくなっている。このように、遺伝的操作により世代交代につれて徐々に望ましい解に収束している。

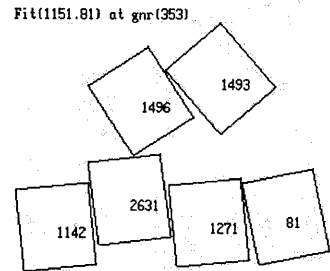
5.3 実システムの評価

図 12 に本手法を応用したデパレタイズシステムの構成を示す。図でロボットハンドについての奥行きセンサは、認識に失敗した時に補助的に用いられるもので、システム全体として高い信頼性を確保するためのものである。

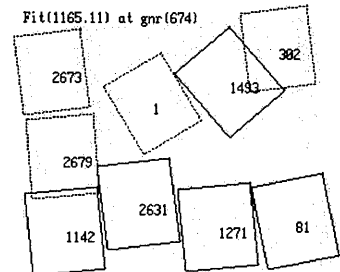
表 1、表 2 はそれぞれシステムの信頼性と処理時間の評価結果である。認識信頼性は約 98% であり、実システムとして十分である。処理時間については、エッジ検出も含め候補抽出に 3.2 秒、ラベルづけは約 1.8 秒であった。合計 5 秒の認識時間は円滑なロボット動作にとって十分と思われる。



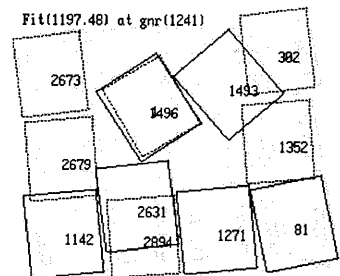
(a) 世代数 = 228.



(b) 世代数 = 353.



(c) 世代数 = 674.



(d) 世代数 = 1241.

図 11: 世代交代過程における集団中の最良解

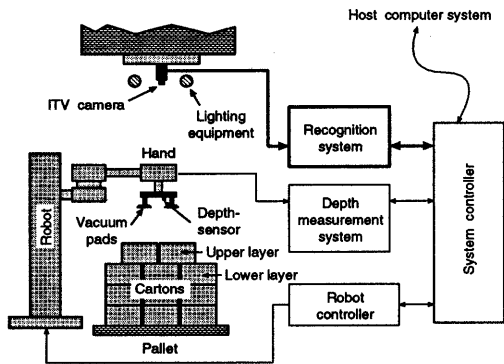


図 12: デパレットシステム構成

表 1: 認識信頼性評価実験の結果

正解	53 枚
見逃し	1 枚
虚報	0 枚
合計	54 枚

表 2: 処理時間計測結果

(計算機性能 = 35MIPS、染色体長 = 28)

エッジ検出	1850 msec
ハフ変換	644 msec
物体候補列挙	754 msec
候補属性計算	1477 msec
GA によるラベルづけ	366 msec
合計	5091 msec

6 まとめ

物体候補抽出と遺伝的アルゴリズムによるラベルづけを組み込んだ物体認識アルゴリズムを提案した。GA によるラベルづけの主な特徴は以下の通りである。

- 染色体の定義として、各ビットが候補の存在有無を表現する。
- 一様交叉と反転・交換突然変異により、部分解の組合せで大局解を得ることが可能である。

GA によるラベルづけによって高速な解の探索が可能である。GA を画像認識の高次処理に適用することでリアルタイム動作を実現した。ラベルづけの性

能は、ツリー探索より 100 倍以上高速で、実システムとしても認識率 98%、処理時間 5 秒と実用的であることがわかった。また、GA の各世代における染色体のふるまいを調べ、交叉や突然変異によって解の探索が効率よくおこなわれていることを確認した。

今後の課題としては、染色体が長い場合の処理時間の抑制がある。染色体コーディング方法の改良が必要であろう。また、形状が不安定で形状モデルを利用しにくい物体認識への拡張についても検討していきたい。

参考文献

- [1] S. Inokuchi and K. Sato, et al., "Vision System for Depalletizer in Automated Storage", Proceedings of JAPAN-U.S.A. Symposium on Flexible Automation, 1986.
- [2] H. Suzuki and T. Tanaka et al., "Vision System for Depalletizer Robot", Proceedings of JAPAN-U.S.A. Symposium on Flexible Automation, 1994.
- [3] D.I.Barnea and H.F.Silverman, "A Class of Algorithms for Fast Digital Image Registration", IEEE Trans. on Comput., C-21, 2, pp.179-186, 1972.
- [4] D.E. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning", Addison Wesley, 1989.
- [5] 北野宏明, "遺伝的アルゴリズム", 産業図書, 1993.
- [6] R.O. Duda and P.E. Hart, "Use of the Hough Transformation to Detect Lines and Curves in Pictures", Comm. ACM, Vol.15, No.1, pp.11-15, 1972.
- [7] 橋本, 鷺見他, "輪郭点情報を用いた高速テンプレートマッチングアルゴリズム", 信学論 J74-D-II, 10, 1991