

局所アスペクトテーブルによる定性的物体認識

楊子豊† 佐藤嘉伸†† 田村進一†† 屠錦榮*

†大阪大学 基礎工学部 情報工学科

††大阪大学 医学部 機能画像診断学研究所 565 吹田市山田丘 2-2

*上海華鐘コンピュータソフトウェア有限公司

あらまし 本論文では、局所アスペクトテーブルと呼ぶ認識に適した新しい表現形式を提案し、局所アスペクトテーブルの多面体モデルからの自動生成法、およびその表現に基づく柔軟かつ効率的な認識手法について述べる。局所アスペクトテーブルでは、物体モデルの局所アスペクト(局所的な見え方)の定性的特徴記述をキーとして、物体モデルと視点を検索でき、かつ、その近傍特徴の定性的記述を参照することができる。ここで、定性的特徴とは、画像線分の平行性やコーナーの見かけの角度の鈍角/鋭角などの特徴の組合せを表す。キーとなる定性的特徴には、物体と視点を識別する際の有効性がランク付けされている。これによって、認識に有効な局所的な画像特徴を画像から選び出し、効率よく信頼性の高い仮説生成を行うことができる。また、定性的特徴の利用によって、各アスペクトの定量的形状変化を吸収すると同時に、物体モデル自体の変形にも対処できるので剛体モデルだけでなく形状可変モデルを扱える柔軟な認識が可能になる。実験によって、認識の柔軟性と効率のよさを実証し、定性的記述による拘束がいかに強力であるかを示す。

Qualitative Object Recognition Using Local Aspect Table

Zifeng Yang†, Yoshinobu Sato††, Shinichi Tamura†† and Jin Rong Tu*

†Faculty of Engineering Science, Osaka University

††Division of Functional Diagnostic Imaging, Osaka University Medical School
Suita, Osaka, 565, Japan

*Shanghai Huazhong Computer Software Development Co. Ltd.

Abstract The goal of this work is to develop an efficient and flexible 3D-from-2D object recognition system. In this paper, we propose a novel object representation scheme, which is called the "local aspect table", and describe an automatic generation method of the local aspect table from polyhedral models and an efficient and flexible object recognition method using the local aspect table. From the local aspect table, given a qualitative description of a local aspect as a key, it is possible to retrieve corresponding object edges, a view point and qualitative descriptions of neighboring appearances. The "qualitative" descriptors, which are used as the keys and also used to describe the neighboring appearances, consist of connected three line segments. They are classified into 18 types based on convergent/parallel/divergent relations of appearances of two edges and acute/obtuse angles of appearances of corners. The qualitative descriptors are ranked based on effectiveness on discrimination of object edges and viewpoints for efficient recognition. Moreover, the qualitative descriptors cope effectively with both quantitative variations of each aspect and deformations of each object model itself for flexible recognition. We demonstrate powerfulness of the qualitative constraints by recognition experiments.

1 まえがき

物体認識とは、画像データを基に、どんな物体がどこに存在するのかを決定する問題であり、ロボットが環境中で知的な作業を行うために解決すべき最も基本的な問題の1つである。物体認識は、画像データと予め計算機に蓄えられた物体モデルを照合することによって行われる。3次元物体を効率よく認識するシステムを構築するには、多面体や一般化円筒などで表現された物体モデルから、予め認識に有効な情報を抽出し、それを認識段階においてうまく利用することが重要になる。認識に有効な情報の抽出と編集を、認識に先立つオフライン処理として自動的に行う処理系は、ビジョンコンパイラと呼ばれ、物体認識研究の1つの大きな流れとなっている[1]。これらの研究は、個々の物体モデルに対して認識手順自体を自動生成するアプローチ、および、個々の物体モデルから認識に適した表現を自動生成し、その表現に基づき共通の認識手順を用いて実行するアプローチ、の2つのアプローチに分類できる。前者のアプローチ[2],[3],[4]では、認識に有効な画像特徴を物体モデルから自動的に決定し、それらをランク付けすることにより、解釈木として認識手順を構成する。認識しようとする物体に特有な特徴を選択できれば、認識効率が格段に向上することが期待できる。認識手順の自動生成は魅力的な考え方ではあるが、手法の明確な定式化がなされるには至っていない。一方、後者のアプローチとしては、Geometric Hashing法[5]があげられる。まず、個々の物体モデルから、視点(あるいは、変換)に不変な表現形式であるHashテーブルをオフライン処理によって自動生成する。次に、このHashテーブルを利用して、物体モデルと画像データの間の変換パラメータを決定する。Geometric Hashing法は、明確に定式化された手法ではあるが、認識しようとする物体に特有な特徴の選択は考慮されていない。別の表現形式として、アスペクトグラフ[6]も効率的な認識を行うための表現として注目されている。Geometric Hashing法のHashテーブルが、モデル特徴点の画像における定量的な位置関係を記述しているのに対して、アスペクトグラフでは、画像特徴の定性的な(あるいは、トポロジカルな)関係が明確にされる。よって、アスペクトグラフを用いる照合は、定量的な関係よりも、定性的な関係を重視するものである。定性的な照合は、Geometric Hashing法が実行するような定量的な照合にくらべ、照合の柔軟性が高く、アスペクトグラフは潜在的には強力な表現形式になりうると期待できる。しかし、近年のアスペクトグラフ研究は、厳密なアスペクトグラフを生成する数学問題に焦点が当てられ[7],[8]、その認識への利用法に関する研究は極めて不十分である。

本研究は、アスペクトの定性的特徴を有効利用した“柔軟な認識”、および、認識しようとする物体に特有な特徴を利用した“効率的な認識”を行うための表現形式の自動生成法と、その表現形式に基づく明確に定式化された認識手法の確立を目的とする。“柔軟な認識”とは、形状の固定した剛体モデルだけではなく、形状の変容性を許した物体モデルを扱えることを意味する。剛体モデルしか扱えないと、あるカテゴリーの物体、例えば、“車”を認識するために、少しずつ形状の異なる多くの剛体モデルを蓄えておかねばならない。形状可変モデルを扱う場合、Geometric Hashing法のような定量的照合を用いると、モ

デルと画像の変換を記述するために多くのパラメータが必要になり、パラメータ推定が困難になる。本研究の1つの目的は、これまでの定量的拘束への過度の依存を反省し、定性的拘束の有効利用法を検討し、それらが、いかに強力であるか示すことである。本研究のもう1つの目的は、“効率的な認識”のためのアスペクト情報の組織化法を検討することである。従来の研究[7],[8]では、物体モデルからのアスペクトグラフの自動生成に重点が置かれていたが、各アスペクトから認識に有効な画像特徴を自動的に抽出し、それをいかに利用するかといった検討は行われていない。本研究では、認識効率向上の観点から、各アスペクト情報を基に認識すべき物体を効率よく検索できる方式を検討する。

本論文では、“柔軟”かつ“効率的”な認識を行うための局所アスペクトテーブルと呼ぶ新しい表現形式を提案し、局所アスペクトテーブルを多面体モデルから自動生成する方法とそれを利用した認識手法について述べる。本研究では、アスペクトを“同一の定性的特徴を有する形状群”と定義する¹。ここで、定性的特徴とは、2線分の収束/発散/平行性やコーナーの角度の鈍角/鋭角などの特徴の組合せを表す。局所アスペクトテーブルは、局所的な画像特徴の定性的記述をキーとして、その画像特徴に対応する物体モデルと視点を検索できるテーブルである。それぞれの画像特徴に対して、物体と視点を識別する際の有効性をランク付けすることにより、認識に有効な局所的な画像特徴を画像から選び出し、効率よく信頼性の高い仮説生成を行うことを目指す。さらに、定性的特徴の利用によって、各アスペクトの定量的形状変化を吸収すると同時に、物体モデル自体の変形にも対処することを目指す。

本論文の構成を以下に示す。まず、2章では、局所アスペクトテーブルを多面体モデルから自動生成する方法を述べる。次に、3章において、局所アスペクトテーブルを用いた認識手法を示す。4章では、実験によって、認識の柔軟性と効率のよさを実証する。5章で、本研究のまとめと今後の課題について述べる。

2 局所アスペクトテーブルの自動生成 — ビジョンコンパイラ

2.1 局所アスペクトテーブルの表現形式

物体認識を行う際には、画像が与えられ、それに投影されている物体が何かを推論する。画像中から、物体に起因する画像特徴を効率よく発見するためには、見かけの画像特徴が物体のどの部分に対応する可能性があるのかを、予め、表形式で記述しておくことと便利である。そのような表を用いることによって、画像特徴の記述をキーとして、物体の対応する部分を検索することができる。そのとき、ある画像特徴に対して、対応する物体部分が複数ある場合、その画像特徴は曖昧さが大きいといえる。逆に、少ない場合、曖昧さは小さいといえる。曖昧さの小さい画像特徴を優先的に選ぶことによって、認識の効率は向上すると考えられる。また、偶然に起こりそうにない規則的な画

¹Koenderinkによるオリジナルの定義[6]は、“見かけの形状のトポロジー的同一群”である。形状がトポロジー的に同一でも、角度の鋭角/鈍角が異なれば定性的特徴は同一ではなくなる。よって、本研究でのアスペクトの定義は、オリジナルの定義に比べ、アスペクトの分類はより細くなる。

像特徴も有用であると考えられる [9]。

局所アスペクトテーブルは、多面体モデルの認識に適した表現形式であり、物体の局所的な見え方、すなわち“局所アスペクト”をキーとして、物体の対応部分と視点を検索できる表である。その表現は、定量的な情報をいっさい含まず、定性的な記述のみから構成される。以下では、まず、局所アスペクトテーブル検索ためのキーとして用いる基本特徴について述べる。ここで述べる基本特徴は、検索ためのキーとしてだけでなく、以下に述べる近傍の画像特徴記述のプリミティブとしても用い、すべての認識処理はこの基本特徴を利用して行われる。次に、その基本特徴をキーとして検索されるデータ集合について述べる。このデータ集合には、対応する物体部分、視点、および近傍の画像特徴記述が含まれる。一般に、1つの画像特徴に対して対応する可能性のある物体部分は複数存在する。近傍の画像特徴記述は、そのような場合の曖昧さ解消のために用いる。

2.1.1 基本特徴 — コーナーペア

図1に本研究の実験で用いる2つの物体モデルを示す。共有頂点を持っている2つのエッジはコーナーを形成する。コーナーは、その2つの順番付けられたエッジで記述できる。例えば、図2(a)に示しているように、エッジe00からe01まで、両エッジの共有頂点の回りに回転させることにより形成されるコーナーを、e00e01で記述する。逆に、e01からe00までのコーナーをe01e00で記述する。また、回転の方向によって、コーナーの角度の正負を定義する。本研究では、時計まわりを正、反時計まわりを負としている。コーナーの画像上での見かけの角度も同様に記述する。その角度は、“正鋭角”、“正鈍角”、“負鋭角”、“負鈍角”の4つのタイプによって定性的に記述する。

本研究では、共有エッジを持つコーナーペア(以下、共有エッジを持つコーナーペアを単に“コーナーペア”と呼ぶ。)の画像投影を基本特徴要素とする(図2(b))。そのようなコーナーペアは、1つのエッジとそのエッジの両端で接続する2エッジを合わせた3つのエッジからなるとみなせる。コーナーペアの画像投影を基本特徴とする理由を以下に示す：

1. コーナーペアは比較的単純で、表現、および画像からの抽出が行いやすい。
2. コーナーに比べ、コーナーペアの見え方のバリエーションはかなり豊富であるので、検索のキーとして十分な識別性がある。

コーナーペアの見え方の定性的記述は、2つのコーナーの画像上での見かけの角度の鋭角/鈍角、および、共有エッジ以外の2つのエッジの関係に基づく。2つのエッジの投影の関係は、2つのエッジが共有エッジに対して、“同側”にある場合は、“平行”、“発散”、“収束”の3つのタイプによって、“異側”にある場合は、“平行”、“不平行”の2つのタイプによって定性的に記述する。その見え方は、表1に示す18種のタイプに分類できる。これら18種のタイプのコーナーペアの見え方を局所アスペクトテーブルのキーとして用いると同時に、曖昧さ解消のための近傍の画像特徴の記述にも用いる。

2.1.2 検索データ集合の記述

コーナーペアの見え方をキーとして検索されるデータ集合は以下の項目からなる：

1. 視点の記述。
2. コーナーペアの見え方の3つの線分に対応する物体モデルにおけるエッジ。
3. 物体モデル中の近傍エッジ、および近傍エッジを含む近傍コーナーペアの見え方の定性的記述。

視点の記述に関しては、同一のアスペクトを持つ視点方向群を1つの“視点”とみなす。本研究では、アスペクトを“同一の定性的特徴を有する形状群”と定義するので、3次元物体が、本研究で定義した定性的記述法において、同一の2次元形状としてみなせる視点方向群を1つの視点と定義する。

物体モデル中の近傍エッジに関しては、コーナーペアの見え方の3つの画像線分に対応する3つの物体エッジのいずれかと頂点を共有するエッジを近傍エッジとする(図3(a))。さらに、近傍エッジのいずれかと頂点を共有するエッジを次近傍エッジとする(図3(b))。この考え方を拡張していき、第 m 次近傍のエッジを定義できる。認識段階においては、特徴的なコーナーペアの見え方をキーとして局所アスペクトテーブルを検索し、さらに近傍コーナーペアの見え方の情報に基づき曖昧さの解消を行う。近傍エッジ、および次近傍エッジとそれらを含む近傍コーナーペアの見え方は、キーとしたコーナーペアとのキーとしたコーナーペアとの相対的位置関係が明確になるような形式で、表1の18種類のタイプにより記述される。例えば、近傍コーナーペアの場合は、近傍エッジがキーとしたコーナーペアのどの頂点を共有するかによって4種類の相対位置に分類でき、その分類が近傍コーナーペアの見え方のタイプと共に記述される。

2.2 ビジョンコンパイラ

ビジョンコンパイラは、物体モデルの入力データファイルを読み込んで、頂点種類の抽出、凹エッジの検出などの初期処理を行った後、局所アスペクトテーブルの生成処理を行う。局所アスペクトテーブルの生成は、2つの段階からなる。第1段階では、まず、入力データファイルの情報に基づき、幾何推論を行い、コーナーの見かけの角度の性質やエッジ間の関係を求め、アスペクトの分類を行う。次に、各アスペクトにおけるコーナーペアの見え方のタイプを求め、近傍特徴の最大次数の決定、およびその近傍特徴リストの生成を行い、各アスペクトの記述ファイルを生成する。以上の段階は、分割コンパイル段階と呼ぶ。第2段階は、各視点におけるアスペクトの記述を、コーナーペアの見え方のタイプ別に整理し、曖昧さの大きさによって、特徴的なコーナーペアの見え方を抽出し、探索優先順位を決定する。これによって、局所アスペクトテーブルを生成する。この段階をリンク段階と呼ぶ。その2つの段階は独立であるので、追加したい物体モデルがある場合、元の物体モデルをコンパイルし直す必要がなく、単に、追加したいモデルをコンパイルした後リンクすることによって、モデルベースを拡張できる。以下では、まず、ビジョンコンパイラの入力となる

モデルデータファイルについて述べ、次に、モデルデータファイルからアスペクトを分類する幾何推論システムも含め、ビジョンコンパイラの詳細について述べる。

2.2.1 モデルデータファイル

モデルデータファイルには、アスペクトの分類を行うための幾何推論システムの入力となる情報が書き込まれている。モデルデータファイルは、頂点記述部、コーナー記述部、エッジ記述部、および、エッジ関係記述部、可視面セット記述部の5つの部分からなる。

頂点記述部では、各頂点の隣接エッジ、隣接面を記述する。コーナー記述部では、対応頂点の各隣接コーナーの角度を記述する。本研究では、コーナーの見え方の定性的記述のみに基づく認識を目指しており、モデルのコーナーの角度も定性的に、“正直角”、“正鋭角”、“正鈍角”、“負直角”、“負鋭角”、および“負鈍角”の6タイプのいずれかを与える。エッジ記述部は、各エッジの2隣接頂点と2隣接面を記述する。

以上の記述に加え、物体モデルの関係記述の情報として、エッジ間の平行関係を記述する関係記述部を設ける。例えば、図1のような車の場合、コーナーの角度の定性的記述からエッジの平行関係を抽出することはある程度できるが、e02とe04のようなエッジの平行関係の抽出はできない。エッジ関係記述部により、そのようなエッジ間の関係情報が幾何推論システムに提供される。現在の段階では、エッジ関係記述部では平行なエッジ集合のみを記述しているが、より複雑な物体モデルを適切に表現するための拡張として、コーナー同士の角度の大小関係やその他の関係情報を記述することが考えられる。

さらに、同一の面の集合が観測できる視点を予め分類しておく、各視点から見える面の集合を、可視面セット記述部に記述する。実際の応用では、観察方向がある範囲内に限られているような場合がよくある。例えば、車の場合、車の底の方から車を見ることがなく、常に側方か上方の視点から車を認識する。その場合、選択的にいくつかの見える面の集合だけをモデルデータファイルに書いておくことによって、視点の範囲を選択できる。これによって、認識効率を向上することができる。

以上の記述は、定量的(数値的)な情報はまったく含んでいない。幾何推論システムでは、以上の記述を基に、可視面セット記述部の各視点を、さらに、“同一の定性的特徴を有する形状群”が観測される視点に細分類する。可視面セット記述部は、現在のところ人手で分類して入力しているが、図1の物体の場合も、“同一の定性的特徴を有する形状群”への分類は、人手による作業は非常に困難であり、自動推論の意義は大きい。可視面セットの自動分類は、アスペクトグラフ²の自動生成[7],[8]と類似の問題と考えられるが、本論文では扱わない。

2.2.2 コンパイルとリンク

まず、物体モデルのデータファイルの情報から、多面体投影の幾何推論によりアスペクト、すなわち“同一の定性的特徴を有する形状群”の分類を行う。1つのアスペクトが1つの視点に対応する。図4に幾何推論の手順を示す。この幾何推論シ

²この集合のアスペクトは“見かけの形状のトポロジー的同一群”である。

テムは一種のルールベースシステムとなっている。まず、1つの可視面セットを基に、可視のエッジ、および頂点を決定する。それらの頂点の見え方やエッジの見え方の関係に関する幾何推論を行う。図5にその処理過程の一例を示す。まず、コーナーの見え方に関する以下の基本的性質に基づき、いくつかのコーナーに関して鈍角か鋭角かを決定できる。

• 視線のコーナー面への投影がコーナー内あるいはその対角内の場合:

- (1) 直角コーナーの見かけの角度は鈍角になる。
- (2) 鋭角コーナーの見かけの角度は鋭角か鈍角になる。
- (3) 鈍角コーナーの見かけの角度は鈍角である。

• 視線のコーナー面への投影がコーナー外かつその対角外の場合:

- (1) 直角コーナーの見かけの角度は鋭角になる。
- (2) 鋭角コーナーの見かけの角度は鋭角である。
- (3) 鈍角コーナーの見かけの角度は鋭角か鈍角になる。

さらに、エッジの投影に関する以下の基本性質に基づき、エッジ間の見え方の関係を決定する。

• 共有エッジを持っているコーナーペアが共面であれば、共有エッジ以外の2エッジの関係を、“平行”、“同側発散”、“同側収束”、あるいは“異側”と定性的に記述することができる。その2エッジを画像へ直交投影した場合、その定性的関係は画像上でも変わらない。

エッジ間の見え方の関係は、性質が未決定のコーナーに関して、見え方の拘束を提供し、さらに、いくつかのコーナーに関して鈍角か鋭角かを決定できる。そのような拘束伝播の処理を新たに定性的性質が決定されるコーナーがなくなるまでくり返す。この処理によって、多くのコーナーに関して鈍角か鋭角かを決定できる。しかし、残りのコーナーに関しては、鋭角/鈍角の2つの可能性が残される。すべてのそれらのコーナーに関して、鋭角/鈍角の2つの可能性を考慮することによって、複数のアスペクトが生成される。このように、1つの可視面セットに対して、複数のアスペクトが生成され、それぞれが視点で定義する。ここで述べた幾何推論システムは、推論できなかったコーナーに関しては、すべての可能性を考慮するので、ありうるアスペクトの生成を推論しないということはないが、ありえないアスペクトを生成する可能性は残る。しかし、このことは、認識段階では、効率と信頼性の若干の低下につながるのみであり、致命的にはならない。

次に、生成された各アスペクト中のすべてのコーナーペアに対して、その見え方を表1に示す18のタイプに分類する。さらに、そのコーナーペアの近傍エッジリスト、および、近傍エッジを含む近傍コーナーペアの見え方である近傍特徴リストを生成する。このリストも表1に示す18のタイプで記述される。この処理を、次近傍、次々近傍と行っていく。ただし、近傍探索の回数 m は、第 m 次近傍エッジリストまでに物体モデルのエッジがほぼすべて含まれるように決定される。

各物体モデルに対して、以上の処理を行った記述ファイルを作成する。まず、コーナーペアの見え方のタイプ別に、各物

体モデルの検索データ集合のリストを統合する。次に、コーナーペアの見え方の各タイプに対して、認識段階での有効性に関する優先順位を以下の基準に従ってつける：

- その見え方タイプをキーとして検索されるデータリストの数が少ないほど曖昧さの少ない画像特徴、すなわち認識に有効な画像特徴である。
- コーナーの見え方において平行性や対称性の存在する見え方は、偶然によって形成される確率が少ないので認識に有効な画像特徴とする。

最後に、以上の処理結果を、コーナーペアの見え方をキーとして、視点、物体モデルエッジ、近傍エッジリスト、および近傍特徴リストを検索できるようにして、局所アスペクトテーブルを生成する。

3 物体認識

認識段階の処理手順は図 6 に示す。まず、画像の L 接続点ペアを画像特徴として抽出し、ビジョンコンパイラにより作成した局所アスペクトテーブルを検索し、投票による近傍特徴の支持率に基づき基本対応仮説を生成する。次に、近傍特徴の見え方を手がかりとして、近傍の探索を行いエッジ対応仮説を生成する。最後に、エッジ対応仮説の照合結果を判断する。

3.1 基本対応仮説の生成

画像から抽出された L 接続点ペアの中で、局所アスペクトテーブルにおいて最も高いランクのタイプに属する L 接続点ペアをキーとして、(視点, コーナーペア) を検索する。さらに、検索された各 (視点, コーナーペア) に対して、近傍エッジからの投票を行うことによって支持率 (local support) を計算する。支持率は、近傍エッジの中で、検索データ集合に記述されている近傍コーナーペアの見え方の定性的性質を満たす近傍エッジの個数とする。十分な (しきい値 T_1 以上の) 支持率を得られる (視点, コーナーペア) が無い場合は、次のランクの L 接続点ペアに対して支持率を計算していく。ただし、ある程度以下の低いランクの L 接続点ペアに対して検索処理を打ち切る。十分な支持率を得る L 接続点ペアと (視点, コーナーペア) の対応があった場合は、それを “基本仮説” とする。検索の打ち切りまでに、それが無い場合は、これまでで最も高い支持率のあった対応を基本仮説とする。

3.2 エッジ対応仮説の生成

エッジ対応仮説の生成手順は以下のようになっている：

1. 近傍エッジ対応仮説の生成。
2. 次近傍エッジ対応仮説の生成。
3. モデルから画像へ逆方向の照合による新しいエッジ対応仮説の成長。

近傍エッジ対応仮説は、基本対応仮説の L 接続点ペアとモデルの (視点, コーナーペア) の対応関係を基に、局所アスペクトテーブルの近傍特徴の見え方リストと整合のとれる近傍エッジ

と画像線分の対応をエッジ対応仮説として生成する。しかし、両立しない対応仮説が存在する場合は、それぞれの仮説を分離し、それぞれの仮説を含む 2 つの対応仮説群を生成する。次近傍以降に関しても同様の仮説生成を行う。この処理は、ビジョンコンパイラによって決定された第 m 次近傍まで行う。

以上の探索によって、モデルの各エッジへの照合がほぼ行われるが、残るエッジの照合は、モデルから画像へ逆方向に行う。逆方向の照合では、対応する視点におけるアスペクトの記述ファイルを読み込んで、未照合のエッジを含んでいるモデルコーナーペアを見出し、そして、そのコーナーペアの見え方、および前の段階で行ったエッジ対応仮説によって、画像から対応する L 接続点ペアを検出し、エッジ対応仮説を生成する。さらに、そのエッジの他の近傍の見え方への照合によって、仮説が正しいことを検証する。

3.3 照合状態の判断

照合状態を判断することによって、以後の探索の見通しに対する評価を以下の基準で行い、探索の中止、あるいは、探索のポイントの移動を行う。

1. 基本仮説の生成において、支持率がしきい値 T_2 以下の対応しか存在しない場合は、これ以上物体は存在しないと判断し、探索を中止する。
2. 近傍エッジ仮説と次近傍エッジ仮説の生成において、矛盾のない仮説からなる仮説群の中で、仮説群のエッジ対応仮説の数がしきい値以下であると仮説群を取り消し、探索ポイントを他の仮説群に移動する。

以上の基準によって、物体が存在しない場合、早い段階で判断することができる。また、見通しのよくない探索ルートから抜け出し、探索を効率よく制御することができる。認識されたかどうかの判断は、エッジ対応仮説の生成の後のエッジの数によって行う。

4 実験結果

本研究で提案した手法の性能を評価するため、Solbourne S4000DX (約 28MIPS) 上に局所アスペクトテーブルの自動生成システム、および物体認識システムをインプリメントして実験を行った。

4.1 局所アスペクトテーブルの自動生成

実験では、まず、車モデル (図 1(a)) のデータファイルをシステムに入力し、車の各視点におけるアスペクトの記述ファイルと局所アスペクトテーブルを生成した。実行時間は約 38 秒であった。次に、トラックを加え、2 つのモデルを持っているモデルベースへ拡張した。トラックモデル (図 1(b)) のデータファイルをシステムに入力しコンパイルした後、車の各視点におけるアスペクトの記述ファイルとリンクして、車とトラックのモデルベースの局所アスペクトテーブルを生成した。コンパイラの実行時間は約 23 秒で、リンクの実行時間は約 2 秒であった。

図7と図8は、幾何推論システムが推論した車とトラックの
 アスペクトである。図7は32、図8は16のアスペクトに分類さ
 れている。ここで、車とトラックの底面は通常の観測方向では
 見えないと仮定して、データファイルの可視面セットを設定し
 た。車のモデルからなる局所アスペクトテーブルでは、表1に
 示す、タイプ8, 9, 10, 11が優先順位1位、タイプ5, 6, 7, 14, 15
 が2位、タイプ3, 4が3位とランク付けされた。この11個の
 タイプを特徴的な局所アスペクトであるとみなし、検索キーと
 して用いた。

4.2 物体認識

ビジョンコンパイラにより作成した局所アスペクトテー
 ブルを利用して、物体認識の実験を行った。まず、抽出される画
 像特徴の欠落に対処するため、図9に示すように、画像線分の
 延長線が別のL接続点の頂点上を通る場合(type1)、2つのL
 接続点ペアにおいて2組の平行線分が存在する場合(type2)、
 および、2つの画像線分が共線条件を満たす場合(type3)、画
 像上のエッジ線分の補完を行う(実験では、エッジ線分の補完
 処理を行わない場合も示す)。なお、本実験での入力線分画像
 は、人間が描いた線画を用いた。本手法は、定性的拘束のみを
 用いているので、画像生成過程の厳密なシミュレートは不要で
 ある。また、鋭角/鈍角、および収束/平行/発散などの定性
 的性質の判断は、5度のマージンをとった。例えば、鋭角に関
 してはコーナー角95度以内とし、平行に関しては、エッジ線
 分の方向が±5度以内とした。基本仮説における十分な支持率
 のしきい値を $T_1 = 6$ 、探索を中止するための支持率のしきい値
 を $T_2 = 2$ と設定した。

まず、本システムがいかにかに認識を行うかを示すためエッジ線
 分の補完を行う場合と行わない場合での認識処理過程を順次
 追っていく。なお、この実験は、車の局所アスペクトテー
 ブルのみを用いて行った。

図10にエッジ線分の補完を行わない場合の認識処理過程を
 示す。図10(a)は入力線分画像であり、図10(b)に選ばれた基
 本対応仮説を示す。近傍エッジ対応仮説の生成(図10(c))、次
 近傍エッジ対応仮説の生成(図10(d))を行い、さらに、逆方向
 検索(図10(e))を行い最終的に車を検出した(図10(f))。実行
 時間は約2秒であった。

図11にエッジ線分の補完を行った場合の認識処理過程を示
 す。この場合、逆方向検索を行うことなく、その前の段階で車
 を検出した(図11(d))。実行時間は約3秒であった。エッジ線
 分の補完を行った場合は、一般に認識の信頼性は向上するが、
 いくつかの画像線分が加えられるので、L接続点ペアの数が増
 え、結局、認識時間が逆に長くなる場合がある。

以後では車とトラックのモデルベースをリンクした局所ア
 スペクトテーブルによって実験を行った。

図12に、形の違う2台の車が投影されている場合の認識例
 を示す。図12(b)はエッジ線分の補完を行った線分画像で、図
 12(c)はその認識結果である。このように、形状の異なる車を
 “車”という同一カテゴリーの物体とする柔軟性のある認識が
 行えた。実行時間は約6秒であった。

図13に、画像中に複数種類の認識可能な物体が投影されて
 いる場合の認識例を示す。車とトラックがそれぞれ検出された。
 実行時間は約6秒であった。

5 むすび

本研究は、アスペクトの定性的特徴を有効利用した“柔軟な
 認識”、および、認識しようとする物体に特有な特徴を利用し
 た“効率的な認識”を行うために、局所アスペクトテーブルと
 呼ぶ新しい表現形式を提案し、幾何推論システムによって、多
 面体モデルからその局所アスペクトテーブルを自動生成する
 方法を示した。そして、それに基づく明確に定式化された認識
 手法を確立し、実験では、本システムは、可変性を許した物体
 モデルを扱える“柔軟性”をもち、形の違う物体を同一カテ
 ゴリーの物体として認識できた。分割コンパイルによって、モ
 デルベースを簡単に拡張でき、複数種類の物体にもそれぞれ検出
 できた。また、雑音と隠れが存在する場合も対処できた。さら
 に、システムは数秒以内に物体を見出し、“効率的”な認識が
 行えることを実証した。

今後の課題としては、物体モデル中に平行や直角などの規
 則性が少ない場合あるいは画像計測の誤差が大きい場合にお
 いて、システムの認識能力に関しての実験的検討、およびモデ
 ルの数が増加した場合の効率的な局所アスペクトテーブルの組織
 化などがあげられる。

謝辞 本研究の推進にあたりご援助頂いたカシオ科学振興
 財団に感謝致します。

参考文献

- [1] 池内克史：物体認識と認識プログラムの自動生成、人工知能学会誌、Vol.4
 No.1, pp. 30-42 (1989).
- [2] Goad, C. : Special purpose automatic programming for 3d model-
 based vision, Proc. of DARPA Image Understanding Workshop, pp.
 94-104, DARPA (1983).
- [3] Ikeuchi, K. : Generating an interpretation tree from a cad model for
 3-d object recognition in bin-picking tasks, International Journal of
 Computer Vision, Vol.1, No.2, pp.145-165 (1987).
- [4] Kuno, Y., Okamoto, Y. and Okada, S. : Object Recognition Using
 a Feature Search Strategy Generated from a 3-D Model, Proc. of
 Third International Conference on Computer Vision, pp. 626-635
 (1990).
- [5] Lamdan, Y. and Wolfson, H. J. : Geometric Hashing : A Gen-
 eral and Efficient Model-Based Recognition Scheme, Proc. of 2nd
 International Conference on Computer Vision, pp. 238-249 (1988).
- [6] Koenderink, J.J. and Van Doorn, A.J. : The Internal Representa-
 tion of Solid Shape with Respect to Vision, Biological Cybernetics
 32, 211-216 (1979).
- [7] Stewman, J. H. and Bowyer, K. W. : Aspect Graphs for Convex
 Planar-Face Objects, Proc. of the IEEE Computer Society: Work-
 shop on Computer Vision, pp. 123-130 (1987).
- [8] Kriegman, D. J. and Ponce, J. : Computing Exact Aspect Graphs of
 Curved Objects: Solids of Revolution, Proc. of the IEEE Computer
 Society: Workshop on Interpretation of 3D Scenes, pp. 116-122
 (1989).
- [9] Lowe, D.G. : Visual Recognition from Spatial Correspondence and
 Perceptual Organization, Proc. of 9th International Joint Confer-
 ence of Artificial Intelligence, pp.953-958 (1985).

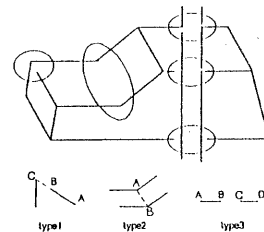
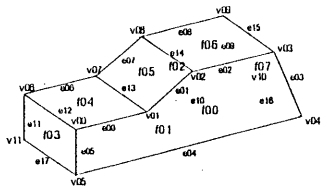
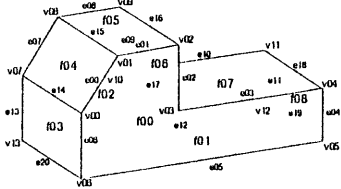


図9. エッジの補完仮説



(a) 車モデル



(b) トラックモデル

図1. 物体モデル

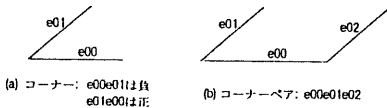


図2. コーナーとコーナーペアの記述

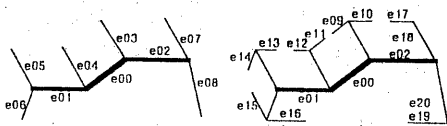


図3. コーナーペアe00e01e02の近傍

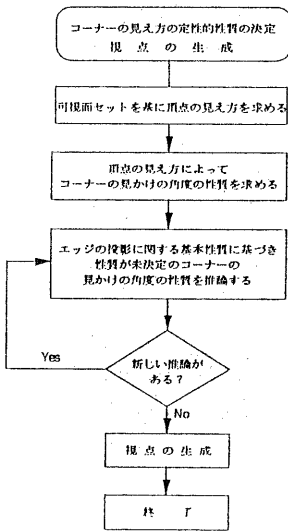


図4. コーナーの見かけの角度的な定性的性質の決定

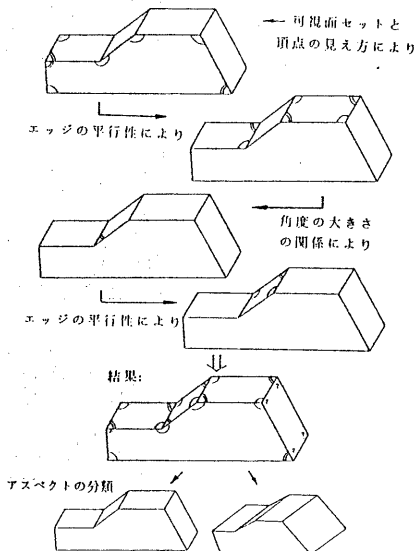


図5. コーナーの見かけの角度的な定性的性質を決定するための幾何推論

表1: コーナーペアの見え方のタイプとその記述

共有エッジ以外の2エッジの関係	2コーナーの見かけの角度	図例	タイプ (Pattern type)
同側・平行	正視-正視 cornerA - cornerB-100		0
	負視-正視 cornerB - cornerA-100		1
同側・垂直	負視-正視 cornerB - cornerA-100		2
	正視-負視 cornerA - cornerB-100		3
	負視-正視 cornerB - cornerA-100		4
同側・収束	負視-正視 cornerB - cornerA-100		5
	正視-負視 cornerA - cornerB-100		6
	負視-正視 cornerB - cornerA-100		7
異側・平行	正視-正視 cornerA - cornerB		8
	負視-負視 cornerA - cornerB		9
	正視-正視 cornerA - cornerB		10
	負視-負視 cornerA - cornerB		11
異側・不平行	正視-正視 cornerA - cornerB		12
	負視-負視 cornerA - cornerB		13
	正視-正視 cornerA - cornerB		14
	負視-負視 cornerA - cornerB		15
	正視-正視		16
	負視-負視		17

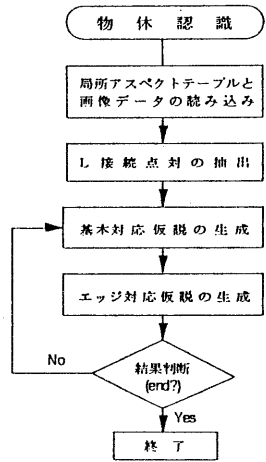


図6. 認識処理の流れ

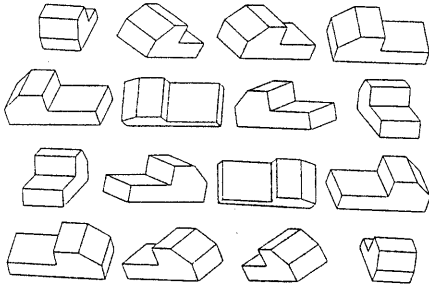


図8. トラックのアスペクト

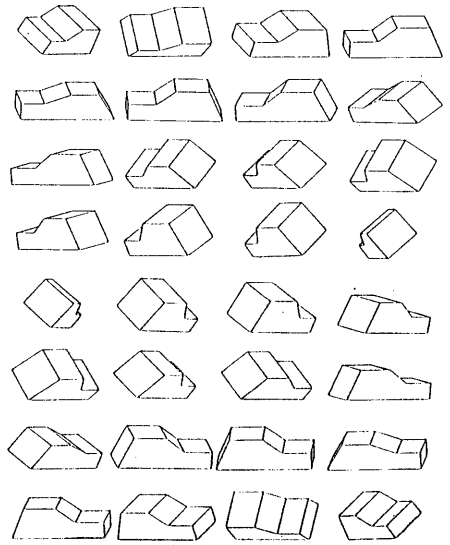


図7. 車のアスペクト

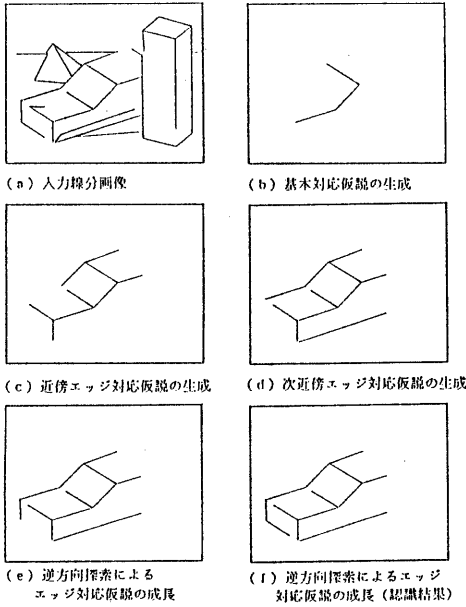


図10. 認識過程 (エッジ補完をしなかった場合)

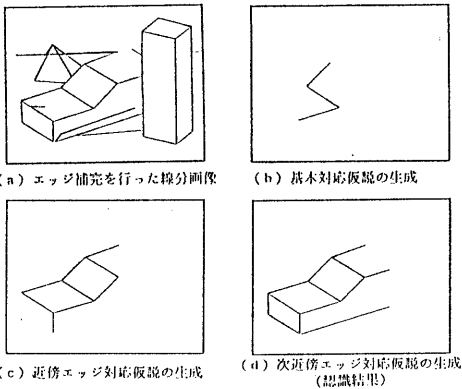


図11. 認識過程 (エッジ補完をした場合)

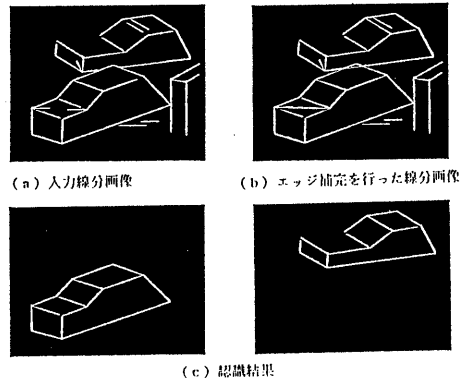


図12. 形の違う2台の車が存在する場合の処理結果

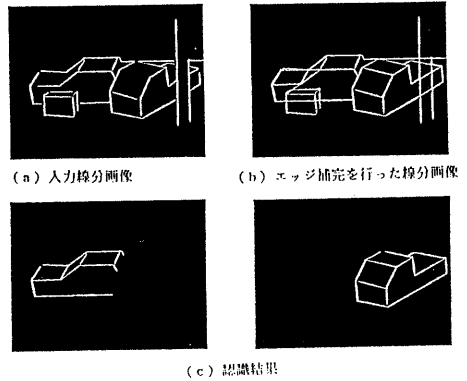


図13. 複数種類の認識可能な物体が存在する場合の処理結果