

# マルチエージェントによる多様な画像に対応した 物体認識システムの一構成法

柳井啓司† 出口光一郎†

実世界に対応した画像理解システムは、多種多様な画像に対して認識可能でなければならない。そのためには、従来のように認識対象となる画像の種類をあらかじめ想定してシステムを構成するべきではなく、多様な画像、認識対象に対して対応できる構成をしなければならない。そうしたシステム構成法として、それぞれに異なった単一種類の物体を認識するプログラムをマルチエージェントによって統合する方法を提案する。我々の方法では、単に各認識プログラムの結果をそれぞれの認識プログラムが算出した評価値に基づいて統合するだけでなく、物体どうしの位置や相対的な大きさなどの関係の知識も利用して統合する。そして、本方法によって実験システムを並列計算機上に実装し、室内画像でも屋内画像でも認識対象に依存することのないシステムが構築できることを確認した。

## A Multi-agent Architecture of Object Recognition System for Various Image

KEIJI YANAI† and KOICHIRO DEGUCHI†

An image understanding system for real world images must have ability to recognize various kinds of objects. We propose a multi-agent architecture to integrate object recognition programs for respective kind of target objects. In our method, recognized results by several different agents are integrated based not only on evaluations made by each programs themselves but also on relations of object locations, sizes and so on. This integration is carried out autonomously between the agents concerned, and the most reliable result is selected after the arbitration between them. We implemented an experimental system on a parallel computer, and achieved recognition for both indoor and outdoor images. The system performs recognition for various kind of objects in either of target images.

### 1. はじめに

実世界に対応した画像理解システムを実現するためには、多種多様な画像に対して認識ができることが必要である。我々はそうしたシステムの実現のための研究として、単一の濃淡画像中に含まれる一般名称を持つ物体の名前と画像中におけるその物体の占める領域を認識するシステムの構築を行っている。ここで的一般名称とは、人間が見たとき通常真っ先に思い浮かんだり、幼児が最初に覚える名称、つまり基本認識レベルでの名称のことであり、たとえば「机」「椅子」などのことである。つまり、本研究の目的は、正確な形状モデルが与えられている状況で認識を行う従来のモデルベースによる物体認識ではなく、存在する物体の種類やその正確な形状などがあらかじめ分からないという実世界シーンの画像に対して認識を行うという、

より人間に近い認識システムの実現である。

本論文では、こうした認識を実現するためのシステムの構築法として、単一種類の物体のみを認識する認識プログラム（認識モジュール）を複数用意し、それぞれに通信機構（通信モジュール）を付加することによってエージェントを構築し、その集合体として認識システムを構築する方法を提案する。

本論文では、まず、提案するシステムの構成の概要について述べる。次に、システムの動作について簡単に説明し、エージェント間での競合解消および関係知識を利用した認識について述べる。そして、システムの停止性と終了判定などの検討すべき問題について議論し、最後に、並列計算機上に実装した実験システムにより、その有効性を示す。

### 2. 対象画像に依存しない認識システムの構成の基本構想

従来の実画像を対象とした画像理解システムの研究では、ほとんどの場合、認識対象とする画像をあらかじめ

† 東京大学大学院工学系研究科  
Graduate School of Engineering, University of Tokyo

じめ想定していた。たとえば、Nagao<sup>1)</sup>、SIGMA<sup>2)</sup>は航空写真、Ohta<sup>3)</sup>、The Schema System<sup>4)</sup>は風景画像をそれぞれ対象としていた。対象画像を限定しておくことによって、実画像特有の困難な問題をあらかじめ想定してシステムを構成することが可能であった。しかし、認識対象を限定することがない、より人間に近いシステムを構築するためには、システムの構成法自体にあらかじめ限定された対象の性質を反映させるべきではない。

そこで、このようなシステムの構成法として、画像を直接扱う部分と、記号的な処理を行う部分に分けてシステムを構成することによって、対象画像に依存しないシステムを実現することを提案する。基本的な考え方としては、単一種類の物体を認識する認識プログラムを複数個用意し、それぞれに協調機構を付加することによって、プログラムどうしが相互作用を行えるようにし、その集合体としてシステムを構築するというのである。これを実現するために、マルチエージェントによってシステムを構築する。

従来のマルチエージェントによる画像理解システムでは、複数アルゴリズムの統合を目的としていたり<sup>5)</sup>、空間構造の柔軟な利用を目的としていた<sup>6)</sup>が、エージェント間の相互作用を個々に考える必要があったり、対象の階層性を利用しているために、システムの構築が複雑であったりした。そのため、認識対象を限定していることが多く、我々の目標とするような一般的な認識システムを構築する場合には、現実的ではなかった。しかし、本システム構成法では、単一種類の物体を認識する認識プログラムをシステムの他の部分とは独立に構築し、それらを共通の協調機構を付加して統合するだけなので、従来より容易に一般的な認識システムが構築可能である。

## 2.1 システムの構成

システムは、エージェントの集合体として構築される(図1)、さらに、各エージェントは、認識モジュールと、通信モジュールによって構成される(図2)。

認識モジュールは画像中に含まれるある単一種類の対象物体を認識するプログラムであれば、その内部に関しては自由であり、それぞれの対象に適した認識手法、知識およびその表現方法を使用する。認識モジュールは、物体を認識すると、その認識結果に対する確信度を自己評価値として付けて、認識結果と評価値を物体候補の情報として通信モジュールに送信する。

各エージェントは、それ自体が独立した認識システムであり、自分の対象とする物体をできるだけ多く認識できるように動作する。しかし、画像には「1つの

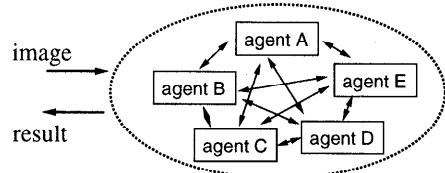


図1 システムの構成。単一種類の物体を認識する認識プログラムの集合体として構成される。

Fig. 1 Structure of the system. It is constructed as an assembly of recognition programs each of which treats with one kind of object.

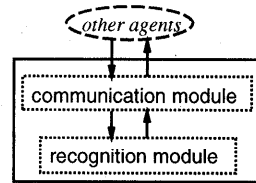


図2 エージェントの構成。画像を直接扱う認識モジュールと、エージェント間で協調を行うための処理を行う通信モジュールによって構成される。

Fig. 2 Structure of an agent. It consists of a recognition module, which recognize objects in the image, and a communication module, which makes cooperation among agents.

画像の領域は1つの物体に対応する」という原則があるために、複数の認識エージェントが同じ領域を異なる物体として認識してしまうと、結果の競合が発生し、システム内で矛盾した結果が存在することになる。そこで、競合を解消するためにエージェント間で交渉を行う必要がある。この交渉を実際に行うのが通信モジュールである。通信モジュールは、認識モジュールが生成した物体候補の情報を受け取って、システム内で矛盾した結果が存在しないように、互いに情報を交換し合い、つねにシステム内のすべてのエージェントの認識結果の整合性が保たれるように処理を行う。この通信モジュールは、システムの構築を容易にするために、各エージェントで原則として同一のものを使用する。

本システム構成では、エージェントの内部を通信モジュールと認識モジュールに分けているために、通信と認識を並列に動作させることが可能であり、また、1つのエージェント内に複数の認識モジュールを用意し、状況に応じて使い分けることも容易に可能になる。たとえば、物体が認識しやすい状況では画像の2次元的な解釈のみによる簡単な認識を行い、複雑な状況では3次元的なモデルを用いた認識を行って、さらに、完全に未知の形状に対しては人工物であればその機能を推測するような高次の認識を行うようにすることも

考えられる<sup>7)</sup>。つまり、各エージェントをさらにマルチエージェントによって構築することも可能である。

このようなマルチエージェント型のシステム構成を採用した場合、全体を制御する機構がないために、認識モジュールを新たに構築するのみで、認識可能な物体の数を増やすことができる。ただし、非同期通信型のマルチエージェントシステムであるため、1エージェントがリアルタイムにシステム全体の状態を知ることができない。そのため、後述のような取消しと復活のループの発生防止や、認識の終了の判定に工夫が必要である。

## 2.2 競合解消のために関係知識を利用

エージェントどうしの認識結果の競合の解消は、通常は競合した結果を出したエージェントどうしで比較して、最も可能性の高い結果のみを残し、それ以外の認識結果は取り消すことによって行う。しかし、各エージェントがそれぞれ異なるアルゴリズムを用いて認識を行っているために統一した基準を用いて比較を行うことは難しい。そこで、各エージェントの認識モジュールが持っている対象物体の形状に関する知識に加えて、通常考えられる物体どうしの位置や相対的な大きさなどの関係の知識を通信モジュールに与える。そして、各認識モジュールが認識結果に対して独自の基準で付けた自己評価だけでなく、他の物体との関係の知識を用いて行った評価にも基づいて比較を行う。

また、たとえば、「机」と「本」が競合し、両者の認識結果の形状が明らかに異なるとき、「book on desk」という関係知識が存在していれば、両者は矛盾がなく、両立することと見なす。すると、前にある小さい物体が後ろの大きな物体の一部を隠してオクルージョンが発生しているときに、そのオクルージョンを無視して、本来の形状のまま認識するような結果が出力されても、矛盾が起きないようにすることができる。このように、関係知識を用いることによって、一見矛盾している結果でも、実は整合性がとれていることが判定できる。

他にも物体間関係知識は、認識モジュールに物体が存在する可能性の高い領域を示したり、一度取り消された認識結果の復活などに利用される。なお、これらの処理に利用される関係知識はすべてのエージェントで共通のものであり、システム全体で共通である関係知識データベースにあらかじめ用意しておく。そして、認識動作を開始する前に、関係のあるエージェントの通信モジュールに分配しておく。

本システム構成法で用いられている関係知識は、2物体間の通常考えられる相対的な関係を記したものである。具体的には、2物体間の位置関係、大きさの関

係、形状の複雑度の関係、濃淡の関係など、様々な関係がある。人間は、関係知識をいわゆる「常識」として扱っていて、辞書などには通常出ていない知識であり、正確には定義できない曖昧な知識である。しかし、認識には有用な知識である。実験システムにおいては、これらの関係はすべて定性的なものとして表現されている。「常識」を表現するのに、数値によって絶対的に定義することは難しい。

我々は、以上のような構想に基づいて実験システムを構築した。次章以降は、構築した実験システムの詳細について述べる。

## 3. 認識の流れ

本システムにおける物体認識の基本的な流れは、次のとおりである。あるエージェントが物体を1つ認識すると、それを他のすべてのエージェントに対してブロードキャストする。他のエージェントから異義が出なければ、承認されたと見なす。もし、異義が出れば、そのエージェントと交渉して、結果の両立、取消しの決定を行い、システム全体で整合性を保つ。これは、エージェント間の協調作用によって、つねにシステム全体で認識結果の整合性を保つという考えに基づいている。

こうしたエージェントおよびそれを構成するモジュールの動作は、すべてメッセージ通信によって行う。各モジュールは、初めにメッセージ待ちに入り、メッセージを受け取るとそれについての処理を行う。そして、再びメッセージ待ちに入る。基本的な動作は、このサイクルの繰返しである(図3)。

実際のメッセージの流れを図4の例に基づいて述べる。

- (1) 認識対象画像がすべてのエージェントの認識モジュールに送られ、認識モジュールの認識が始まる。通信モジュールは、メッセージ待ちに入る。

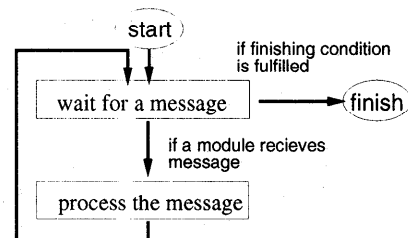


図3 メッセージ駆動によるモジュールの動作。各モジュールは、終了条件が満たされるまで、メッセージ受信とその処理のループを繰り返す。

Fig. 3 Processing by message-driven mechanism. Each module repeats a loop of receiving a message and processing it, until finishing condition is fulfilled.

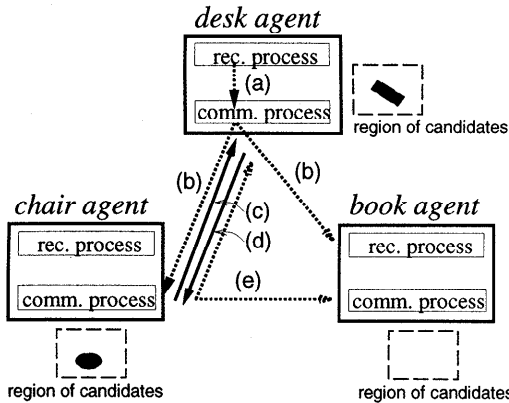


図4 メッセージの流れ。(a) 物体候補の情報, (b) 物体候補の存在範囲の情報, (c) 異義メッセージ, (d) 異義メッセージを受けた物体候補の詳細情報, (e) 取消しメッセージ

Fig. 4 Flow of messages. (a) Information of an object candidate, (b) Information of existence bound of an object candidate, (c) Objection message, (d) Detail information of candidate received objection message, (e) Cancellation message.

- (2) 認識モジュールが物体候補を1つ生成すると、その情報を通信モジュールに送る(図4(a)).
- (3) 通信モジュールは候補の情報を記憶し、その存在範囲(実験システムでは矩形で表現)の情報を他エージェントへブロードキャストする(図4(b)).
- (4) 候補の存在範囲情報を受け取ったエージェントの通信モジュールは、自分の物体候補の領域と重複していないかチェックする。
- (5) もし重複があれば、物体候補をブロードキャストしたエージェントに、異義メッセージを送る(図4(c)). 異義メッセージには、自分より相手の面積が小さいときは自分の重複する結果の詳細情報(物体候補の領域の画素と後述する形状の評価値など)も含める。そうでないときは、相手の詳細情報要求を含める。
- (6) 異義メッセージに詳細情報要求が含まれているときは、重複する認識結果の詳細情報を送る(図4(d)).
- (7) 詳細情報を受け取ったエージェントは、両立か、どちらか取消しの判定する。取消しの場合、全エージェントに対して、取消しメッセージを送る(図4(e)). 両立の場合、相手のエージェントのみに結果を送る。
- (8) 再び、通信モジュールはメッセージ待ちに入る。
- (9) すべてのエージェントのすべてのモジュールが待ち状態になり、通信路上にメッセージが存在しなければ、認識が終了する。

表1 形状評価値のおおよその基準

Table 1 Rough standard of shape evaluation value.

評価値	おおよその基準
5	すべての形状が得られている
4	ほぼ全体が得られている
3	最低限の構成要素に加えて、付加的な要素も見ついている
2	最低限の構成要素だけが見ついている
1	最低限の構成要素らしいのみが見ついている

#### 4. 競合の解決

本システムにおいては、競合が起こったときは、当事者であるエージェント間で交渉を行う。そのとき、実際の判定は画像での面積の小さい方の物体候補のエージェントで行うことに決めている。判定を行うエージェントは、相手の物体候補が実際に占める画素の情報を詳細情報として受け取り、その情報を基に自分の物体候補との重なりを調べる。もし、ある一定以上重なっているものがあれば、比較を行う。このとき、ある一定以上重なっているものが複数あれば、複数に対して、それぞれ1対1で比較を行う。なお、異義は送られてきた存在範囲情報と、自分の物体候補の占める領域とに重なりがあるとき発生するので、実際に画素どうしを比較した場合には、重なっていないこともある。その場合は、競合は起こっていないこととする。

以下に、比較で必要な形状の評価値、関係知識とその評価の仕方、関係評価値について触れ、その後比較の方法について説明する。

##### 4.1 形状の評価値

形状の評価値は、それぞれの認識モジュールが自分の認識結果に対して、その確信度を評価した値で、実験システムでは、形状の評価値を1~5の5段階にしてある。評価の付け方はそれぞれの認識モジュールによって異なるが、おおよそ表1のようなおおよそな基準に基づいて値を決定することとしている。

このように形状の評価値を5段階としてあるのは、形状評価値が、それぞれ異なる認識手法、知識を持った認識モジュールの結果の自己評価値であり、異なる対象を異なる基準で評価していることになるために、それらを同列に比較することは、あまり厳密に行っても意味がないからである。後は、関係知識を活用して決定するという方針を実験システムではとった。

たとえば、「椅子」についてこの基準を当てはめると表2のようになる。基本的に「椅子」のような構造がはっきりしている人工物の評価値は、その物体が機能を発揮するのに最も必要な要素を最低限の構成要素と見なし、それ以外にどれだけの構成要素が見つかるかで、評価値を付ける。

表2 「椅子」の形状評価値の例

Table 2 Example of shape evaluation value for "chair".

評価値	検出されて構成要素
5	座面+背もたれ+安定した足
4	座面+安定した足 (2本以上 or 1本+足元の安定した構造)
3	座面+足 or 座面+背もたれ
2	正方形か円に近い形状を持った座面
1	形が正方形でも円でもない座面

表3 関係知識の例

Table 3 Examples of relational knowledge.

desk near chair
desk larger_than chair
desk higher_than chair
ws on desk
book on desk

表4 実験システムで使用できる関係

Table 4 Relation installed on the experimental system.

no.	relation name	内容	type
1	near	近くにある	A
2	on	上に載っている	A
3	larger_than	面積が広い	B
4	higher_than	位置がより高い	B
5	more_complex_than	複雑度が高い	B
6	more_straight_than	境界線が直線的	B
7	lighter_than	平均輝度値が高い	B
8	high_texture	輝度値の分散が大きい	B

## 4.2 関係知識とその評価

関係知識は「物体名」+「関係」+「物体名」という形で表現される。たとえば、「本は机の上にある」という関係は、「book on desk」というように表現される。表3にいくつか例を示す。

実験システムで現在使用可能な関係を表4に示す。なお、2つのタイプの関係があるが、Aは位置関係、Bはその他の関係を表している。

## 4.3 関係評価値

関係知識の評価は、各エージェントが自分の物体候補と他の物体候補の間に、関係が成立しているかどうかチェックすることによって行われる。たとえば、「book on desk」という関係知識を「机」エージェントが持っていた場合、初めに、「机」エージェントは、過去に「本」エージェントから物体候補のブロードキャストがあったかどうか、自分が保持している他のエージェントの物体候補の情報の中から調べる。ブロードキャストされた情報には、本の占める存在範囲の情報が含まれているので、それを基にして、自分の物体の認識結果との位置関係を調べる。この場合、関係「on」なので、本が、画像上で机の上方か内部の領域に含まれば、関係が満たされることとする。

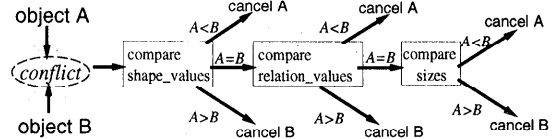


図5 競合する物体候補の比較

Fig. 5 Comparison of conflicting candidates.

評価はエージェントが持っている関係知識を1つずつチェックすることによって行われる。まず、関係知識を1つ取り出し、それに記されている相手の認識結果がないかどうか調べる。もし、あった場合には、関係を調べる。関係も成り立っていた場合には、評価値にある値が足されることになる。実験システムでは、関係を2種類に分けて、加算する値を区別している。Aタイプの関係では関係の相手が複数あった場合に、1つでも成り立てば1が加算される。2つ以上成り立っていても1しか加算されない。Bタイプの関係では、関係が成り立った場合、 $1/(\text{関係の相手の候補の個数})$ が加算され、したがって、関係がすべての相手について成り立った場合には1が加算されることとなる。これは、前者が必ずしも成り立つわけではない関係、後者がほとんどの場合成り立っているべき関係ということで、区別している。

このような関係知識の評価をすべての関係知識について行い、その値を合計したものが関係評価値となる。

## 4.4 結果の比較

競合が起こったときには、最終的には、どちらかを取消す必要がある。ただし、前述したように、両者の間に両立しうる関係が存在する場合もある。

比較は、まず初めに5段階の形状評価値によって評価を行う。ここで、どちらか一方の形状評価値が小さければ小さい方は取消しとなる。もし、形状評価値が同じであった場合は、関係評価値を比較し、評価値が小さい方が取消しとなる。さらに形状評価値も等しかった場合は、単純に面積の小さい方を取消す(図5)。

## 4.5 結果の取消し、復活

競合を解決するための比較が終了すると、その結果をすべてのエージェントに知らせるために、取消しメッセージがブロードキャストされる。取消しメッセージには取り消す物体候補の識別番号が含まれていて、それを受け取ったエージェントは、それに基づいて、他エージェントが生成した物体候補情報を修正する。それと同時に、取消しメッセージを受け取った各エージェントは、取消しメッセージによって取り消された物体候補によって、以前に自分の生成した物体候補が取り消されていなかったかどうか調べる。もし、取り消され

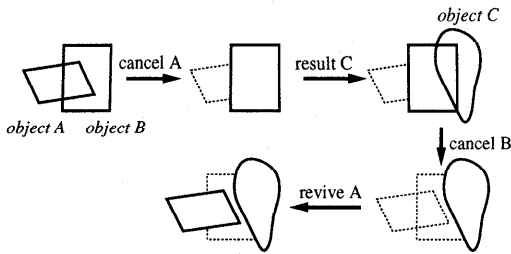


図6 物体候補 A を取り消した物体候補 B が取り消された場合、A は復活する。

Fig. 6 If candidate B that canceled candidate A was canceled, the candidate A revives.

ていたら、取消しの理由がなくなったわけであるから、復活することができる (図6)。

復活は新たな物体候補が見つかったときと同様に、物体候補の存在範囲情報をメッセージとしてブロードキャストすることによって行われる。その後のエージェント間の協調は、新たな物体候補情報をブロードキャストしたときと同様である。

また、復活は、一度取り消された物体候補と、関係のある物体候補の情報を受信したときにも行われる。そのため、各エージェントは他のエージェントの生成した物体候補の情報を受け取ると、過去に取り消された自分の候補と関係が成り立つかチェックしている。ただし、この復活は、他のエージェントの新たな候補生成によって、関係評価値が増えた結果、一度取消しに決まった比較の結果を覆すことができるとエージェントが判断できる場合のみ行われる。この判断は取消しメッセージに含まれている取消し理由の情報を基にエージェントが行う。復活の仕方は、前述の場合と同じである。

こうして、一度取り消された結果も、復活が行われることによって、もう一度比較が行われ、つねに両立可能なすべての認識結果が残ることになる。

## 5. 関係知識の利用によるトップダウン的認識

エージェントの認識結果はブロードキャストされて、他のすべてのエージェントに伝えられる。このとき、自分の物体と関係のある物体候補の情報を受け取ると、その両方の物体の間に成り立っている関係から、自分の物体の領域を探すための何らかの手がかりが得られることがある。たとえば、図7の例では「机」エージェントが机候補を生成し、そのことを知った「本」エージェントは「book on desk」という関係知識を用いて、机の上にある四角い形の物体は本だと推測している。このように、関係知識と他物体候補の存在情報を合わせることによって、自分の物体の存在する可

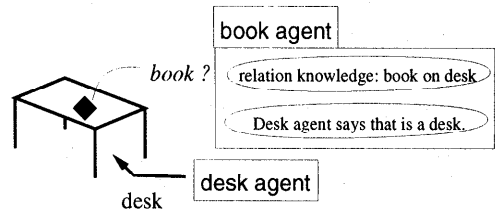


図7 トップダウン認識の例。机の候補と「本は机の上にある」という関係知識から、机の上の本を探す。

Fig. 7 Top-down recognition using relational knowledge. Book agent searches candidates of book using information of a candidate of desk and a knowledge of "book on desk".

能性が高い領域や、大きさなどの、認識に役立つ手がかりを知ることができ、他のエージェントの認識結果からのトップダウン的な認識が実現できる。

これを実際に実現するために、本システム構成法では、すべてのエージェントの通信モジュールは、他のエージェントの物体候補の情報を受け取ると、自分の関係知識と照合し、もし何らかの手がかりが得られる場合には、通信モジュールから認識モジュールへその手がかりの情報を送信する。たとえば、図7の例では、「本」エージェント内の通信モジュールが、机の位置と"on"という関係から予想され得る本の存在領域を認識モジュールに送信する。そうすると、認識モジュールは予想存在領域に対して、再認識を行う。初期認識のときに認識できなかった物体に対して再認識を行うわけであるから、通常は、閾値などの認識パラメータを変更して、初期認識で認識できなかった物体を認識しようと試みる。そして、認識ができた場合は、結果を物体候補として通信モジュールに送り返す。

## 6. システムの停止性の問題と終了の判定

### 6.1 システムの停止性の問題

本システムでは、各エージェントの通信モジュール間の通信は基本的に認識モジュールからの物体候補出力に基づいて行われるので、キャンセルと復活の無限ループが発生せず、かつ、認識モジュールの物体候補出力の数が有限であれば、システムは必ず停止する。ここでは、後者の問題は個々の認識モジュールの設計の問題なので、前者の方の問題のみを議論する。

キャンセルと復活のループとは、たとえば、A, B, Cの3つの物体候補が互いに競合している場合に、AがBを取り消し、BがCを取り消し、さらにCがAを取り消す。すると、取り消された原因となった物体候補がそれぞれ取り消されたことになるので、A, B, Cがともに復活する。すると最初の状態に戻って

しまい、再び互いに取り消し合い、復活するという無限ループの発生である。これを回避するには、ループを構成するような比較の結果が出ないようにすればよい。つまり、A, B, C に対して、それぞれのエージェントが  $(A < B) \& (B < C) \& (A < C)$  のように、比較の結果の順序が一意に決まるような評価を出せばよい。また、そうになっていないと、比較の順序によって、結果が変わってしまうという問題も発生してしまう。

本システムでの競合解消のための比較は、つねに同じ方法によって数値である評価値の比較として行われているので、このように比較の順序が必ず一意に決まるようになっており、原則としてループは発生しない。ただし例外的に、関係の評価値が変化した場合に、この原則が守られないことがある。たとえば、 $(A < B)$ ,  $(B < C)$  という比較が順番に行われた後に A の関係評価値が変化し、A が復活し比較が行われると、 $(A > C)$  となる場合がある。しかし、このような関係評価値の変化が連続的に発生して、無限ループが形成されることは、非常に稀なケースであり、本システムで行った実験中には特に問題は起きていない。

## 6.2 終了の判定法

すべてのエージェントのすべてのモジュールが待ち状態になり、かつ、通信路上にメッセージが存在しないという条件が満たされたときに、認識が終了する。終了判定は、システム全体で送信メッセージの数と受信メッセージの数が同一であるということを利用して行っている。具体的には、各エージェントは、待ち状態に入ったとき、返答を必要とするメッセージを送信した回数と、それに対する返答の受信回数の差が 0 になると、そのエージェントのメッセージの総受信数と総送信数をメッセージとしてブロードキャストする。そして、すべてのエージェントがそのメッセージを送信すると、各エージェントは全エージェントの総受信数と総送信数が合計して等しくなるか調べ、等しければ終了、そうでなければ、メッセージが通信路上に存在するか、待ち状態でないモジュールがあるので、待ちを続ける。こうすることによって、全エージェントが待ち状態になったときに、同時に終了できる。

## 7. 実験システムの実装と実験

画像中の「机」「椅子」「ワークステーション（以下 WS と略す）」「壁」「床」「道路」「自動車」「空」「木」を認識する 9 種類のエージェントを構築し、分散メモリ型並列計算機 AP1000+ 上で、1 プロセッサを 1 エージェントに割り当てて実装した。用意した関係知識の数は 17 である、

各エージェントの認識モジュールの実現法は、エージェントごとに異なり、人工物なら、物体の機能など認識対象の本質を表しているような部分、たとえば、椅子なら座面、机なら机上面などに注目して認識を行う<sup>7)</sup>。床や道路などの形のないものは領域分割による方法などを用いている。基本的には、その物体の一般的な見え方を認識できるように実装した。

図 8 の室内画像に対する実験では、「机」「椅子」がそれぞれ 2 つ (図 10, 11)、「WS」が 2 組のディスプレイとキーボード、「床」「壁」「空」「木」が 1 つずつ、それぞれのエージェントによって候補として生成された。これらの 10 個の候補生成によって、競合が 7 回、取消しが 6 回、復活が 2 回行われ、最終的に図 9 のように、5 種類 6 個の物体候補が残り、認識結果が得られた。処理時間は 76 秒である。

次に、競合について見てみると、机候補 no.0 と椅子候補 no.1 の間で競合が発生している。椅子 no.1 は必要条件の座面に加えて、足、背もたれが見つかったので形状の評価値 5、一方机 no.0 は机上面のみなので形状評価値 3 となって、形状評価値の比較によって机 no.0 が取り消される。また、壁と空も競合が起こっており、壁、空の形状評価値はどちらも 4 であるが、関係評価値は "wall higher than floor" によって壁が 1、空は 0 なので、空の候補が取り消される。他には、椅子候補 no.0 (図 11 中の背もたれの領



図 8 実験画像 1 (室内画像)  
Fig. 8 Sample image 1  
(indoor image).

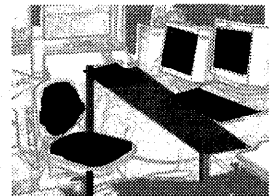


図 9 認識結果。椅子、机、WS、床、壁がほぼ認識できている。  
Fig. 9 Recognition result. A desk, a chair, a workstation, floor and wall are recognized.

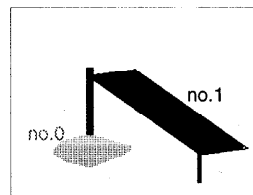


図 10 机の候補  
Fig. 10 Candidates of desk in Fig. 8.

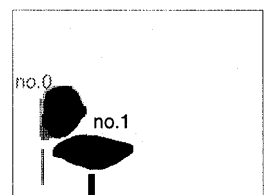


図 11 椅子の候補  
Fig. 11 Candidates of chair in Fig. 8.

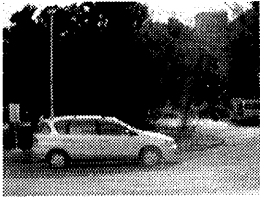


図12 実験画像2 (屋外画像)  
Fig. 12 Sample image 2  
(outdoor image).



図13 認識結果. 道路, 木, 空,  
自動車がほぼ認識できて  
いる.  
Fig.13 Recognition result.  
Road, tree, sky and  
a car are recog-  
nized.

域とそのすぐ下の縦線)が椅子 no.1 と重なっているが,これはエージェント内での競合で,形状評価値と面積によって比較される.ここでは形状評価値によって,椅子 no.0 が取り消される.他にも WS のキーボードと机が重なっている部分があるが,これは "ws on desk" という両立する関係があるので,取消しは起こらず両立する.このようにして,オクルージョンのある物体でも認識可能となっている.

また,関係知識によるトップダウン的な認識も行われており,たとえば,「WS」エージェントは机が認識されたことを知ってから, "ws on desk" という関係情報を用いて,机の上方面して WS の探索を行っている.そのために,平行四辺形の認識という単純処理しか行っていないのにもかかわらず,ディスプレイとキーボードが認識できている.

次に,図12の屋外画像に対する実験では,「空」「自動車」「道路」「木」「壁」の候補が1つずつ生成された.これらの候補生成によって,競合が1回,取消しが1回行われ,最終的に図13の認識結果が得られた.

## 8. ま と め

本論文では,認識対象とする画像の種類に依存しないシステム構成を,マルチエージェントを用いることによって実現する方法について述べた.本システム構成法では,エージェントを認識部分と通信部分に分けて構築し,さらにエージェント間の協調に関係知識を用いることで,単一種類の物体の認識を行うエージェントの集合体として,室内画像と屋外画像とを同時に認識できる,多様な画像に対応した物体認識システムが構築可能であることを示した.

今後の課題としては,物体候補に対する評価方法や,さらに高度な協調メカニズムの検討,また,今回は非常に簡単な手法で実現していた単一種類の物体を認識する認識モジュールの実現法の研究などがあげられる.

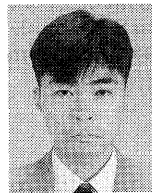
謝辞 本研究の一部は,日本学術振興会未来開拓学術研究推進事業,「分散協調視覚による動的3次元状況理解」によったものである.

## 参 考 文 献

- 1) Nagao, M. and Matsuyama, T.: *A Structure of Complex Aerial Photographs*, Plenum Press, New York (1980).
- 2) Matsuyama, T. and Hwang, V.S.: *SIGMA: A knowledge-based aerial image understanding system*, Plenum Press, New York (1990).
- 3) Ohta, Y.: *Knowledge-Based Interpretation of Outdoor Natural Color Scenes*, Pitman Advanced Publishing Program, Boston (1985).
- 4) Draper, B., Collins, R., Brolio, J., Hanson, A. and Riseman, E.: *The Schema System*, *International Journal of Computer Vision*, Vol.3, No.2, pp.209-250 (1989).
- 5) Ohta, Y., Watanabe, M. and Sumi, Y.: *Approaches to Parallel Computer Vision*, *IEICE Trans.*, Vol.E74, No.2, pp.417-426 (1991).
- 6) 松山隆司, 大田友一(編): 小特集「並列・分散協調画像理解」, *人工知能学会誌*, Vol.9, No.5, pp.629-643 (1994).
- 7) Stark, L.: *Functionality in Object Recognition*, *Computer Vision and Image Understanding*, Vol.62, No.2, pp.145-146 (1995).

(平成9年5月12日受付)

(平成9年9月10日採録)



柳井 啓司 (正会員)

1995年東京大学工学部計数工学科卒業. 1997年, 東京大学大学院情報工学専攻修士課程修了. 1997年10月より電気通信大学情報工学科助手. 画像理解システム, マルチエージェント, 並列処理等に興味がある. 人工知能学会会員.



出口光一郎 (正会員)

1976年東京大学大学院修士課程修了(計数工学). 同年より東京大学工学部助手, 講師を経て, 1984年, 山形大学工学部情報工学科助教授, 1988年, 東京大学工学部計数工学科助教授, 現在に至る. この間, 1991~1992年, 米国ワシントン大学客員準教授. コンピュータビジョン, 画像計測, 並列コンピュータの研究に従事. 計測自動制御学会, 電子情報通信学会, 形の科学会, IEEE等各会員.