マルチエージェント物体認識システムにおける協調に関する考察

A Consideration on Cooperation at the Object Recognition System Based on Multi-agent Architecture

柳井 啓司
Keiji Yanai

電気通信大学 情報工学科
Department of Computer Science, The University of Electro-Communications

Abstract: We propose a multi-agent architecture of object recognition system to realize flexible recognition for real world image. In this architecture, each agent recognizes only one kind of object, and the system constructed as an assembly of agents achieves recognition for various objects by cooperating among agents. Each agent has shape knowledge about objects and relational knowledge of object location, sizes and so on between two objects. By integrating these two knowledge, they recognize objects and resolve conflict among them. This paper describes an outline of this architecture and result of a recognition experiment of the implemented system.

1　はじめに

我々は，より人間に近い画像認識システムの実現のために，様々な実画像に対して，画像中に含まれる物体の「人間の付けた一般名称」を認識できる物体認識システムの研究を行なっている．こうした認識を実現する場合においては「実世界には一つの一般名称で表される様々な形状の物体が存在する．」という事実があるために，従来の研究で行なわれてきたような，認識システムに予め正確な形状モデルを知識として与えておくことが難しく，形状に関する知識だけでは正確な認識を実現することは難しい．そこで，実世界で通常考えられる物体同士の位置や相対的な大きさなどの関係の知識を与えておいて，それらも利用することによって，正確に与えることができない形状知識を補ってやる必要がある．

我々は，こうした認識システムの実現を目指すために，実世界に存在する物体の多様性を「実世界には物体を表す様々な一般名称が存在する．」という問題と，「実世界には一つの一般名称で表される様々な形状の物体が存在する．」という問題の 2 つに分離して対処するシステムの構成を提案する．我々のシステム構成法では，物体の形状に関する知識を用いて「一つの一般名称で表される様々な形状の物体」を認識する認識プログラムを多数用意し，それらに物体間の関係の知識を利用する協調機構を付加して，認識プログラム同士が相互作用を行なえるようにする．そして，そうした認識プログラムの集合体としてシステムを構築することによって，「様々な一般名称」に対する認識の実現を行なう．本発表では，こうしたシステム構成法の具体的な内容と，実際にインプリメントしたシステムを用いて行なった実験の結果について述べる．

2　システムの構想と概要

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予め想定していた．例えば，Nagao[1], SIGMA[2] は航空写真，Ohta[3], The Schema System[4] 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よって，実画像特有の困難な問題を予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予め限定された対象の性質を反映させるべきではない．

我々のシステム構成法では，画像を直接扱う部分と，記号的な処理を行な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できる．

具体的には，システムを単一種類の物体のみを認識するエージェントの集合体として構築する (図 1)．そして，各エージェントは通信によって認識結果を互いに交換して，矛盾がある場合には交渉を行ない解決し，常にシステム全体の認識結果が整合性のとれたものとなる

〒182 東京都調布市調布ヶ丘 1-5-1　Tel.0424-83-2161 ex.4145
E-mail: yanai@cs.uec.ac.jp

ようにする．各エージェントは，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なうための処理を行なう通信モジュールによって構成される (図 2)．

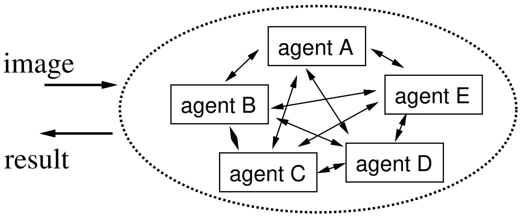

図 1: システムの構成．

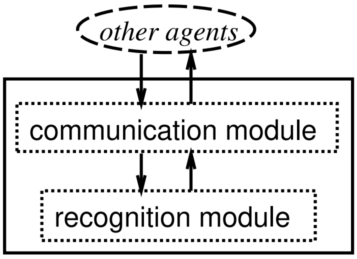

図 2: エージェントの構成．

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，知識及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．

通信モジュールは，認識モジュールの認識結果がシステム全体で整合性が保たれるように他のエージェントに結果の提示と競合解消のための交渉を行なう．つまり，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った場合には，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ので，それを互いに協力して解消するのが通信モジュールの役割である．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とって，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，常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なう．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．また，通信モジュールは関係知識を持っており，それを利用して，自分の結果を評価したり，認識プロセスに存在可能性の高い領域や物体の予想される大きさなどを教えたりする．

2.1 動作の概要

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，すべてメッセージ駆動によって行なわれる．図 3 に基づいて説明すると，初めにすべての認識モジュールに対して入力画像が送信される．そして，各認識モジュールは認識を開始し，1 つ物体を認識する度にその認識結果を物体候補として通信モジュールに送り (図 3(a))，通信モジュールは他の全エージェントに対してその物体候補の存在範囲の情報を送信する (b)．それを受けとった他のエージェントの通信モジュールは，それが既に認識されている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．もし重複があれば，そのエージェントは異義メッセージを返信する (c)．さらに，それを受けとったエージェントは，異義を受けた物体に関する詳細情報を送り返し (d)，詳細情報を得たエージェントは，互いに競合する物体候補同士を後述する方法で比較し，どちらがか取り消しか両立か，の判定をする．取り消しが発生した場合は，取り消しメッセージを全エージェントに対して送信する (e)．

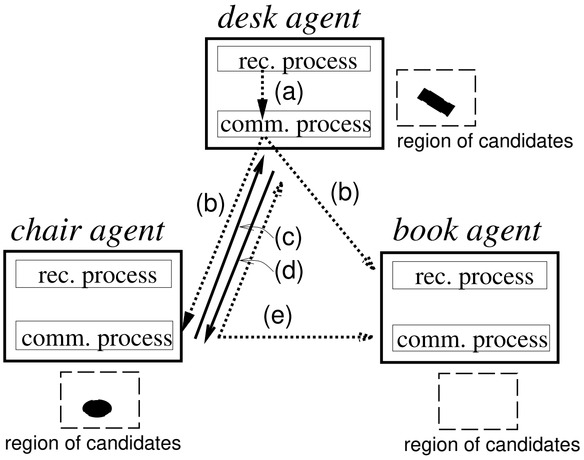

図 3: メッセージの流れ．(a) 物体候補の情報．(b) 物体候補の存在範囲の情報．(c) 異義メッセージ． (d) 異義メッセージを受けた物体候補の詳細情報．(e) 取り消しメッセージ．

一度，取り消された結果も，関係評価値が上がると再びブロードキャストされ，復活が行なわれる．そして，もう一度比較が行なわれるために，常に互いに整合性のとれた認識結果のみが残ることになる．

こうして，常にシステム全体で認識結果の整合性がチェックされ，やがて，すべてのエージェントの認識が終了し，メッセージ待ち状態になると，認識が終了する．

3　形状知識と関係知識の利用

3.1　関係知識の利用

エージェント同士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同士で比較して，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なう．しかし，各エージェントがそれぞれ異なる知識，アルゴリズムを用いて認識を行なっているために統一した基準を用いて比較を行なうことは難しい．そこで，各認識モジュールが自ら生成した物体候補の形状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なった評価にも基づいて比較を行なう．

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．

関係知識は「物体名」＋「関係」＋「物体名」という形で表現される．例えば，「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1にいくつか例を示す．なお，これらの関係知識は，実験システムでは，テキストファイル中に記述されており容易に変更追加が可能である．

実際にインプリメントした実験システムで現在使用可能な関係を表2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．

また，例えば，「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時，"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，両立することとみなすと，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いる時，つまり，オクルージョンが発生している時に，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．

他にも物体間の関係知識は，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予め用意しておく．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．

表 1: 関係知識の例．

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

表 2: 実験システムで使用できる関係．

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

3.2　競合の解決方法

本システムにおいては，競合が起こった時は，当事者であるエージェント間で交渉を行なう．その時，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なうことに決めている．判定を行なうエージェントは，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，その情報を元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，比較を行なう．この時，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で比較を行なう．

以下に，実際にインプリメントしたシステムで採用している，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．

3.3　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1,2,3,4,5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3のようなおおまか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，あまり厳密に行な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．

表 3: 形状評価値のおおよその基準．

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	最低限の構成要素と付加的な要素が見付かっている
2	最低限の構成要素だけが見付かっている
1	最低限の構成要素らしいものが見付かっている

3.4 関係の評価値

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なわれる．例えば，"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，それを元にして，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，画像上で机の上方の領域に含まれば，関係が満たされることとする．

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なわれる．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/(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ない，その値を合計したものが関係評価値となる．

3.5 結果の比較

競合が起こった時には，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なう．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り消しとなる．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り消しとなる．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り消す (図 4)．

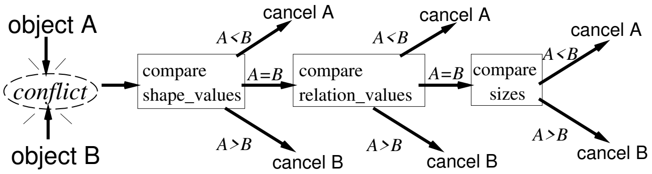

図 4: 競合する物体候補の比較．

4 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，エージェント毎に異なり，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，例えば，椅子なら座面，机なら机上面などに注目して認識を行なう [5]．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 5: 室内サンプル画像．

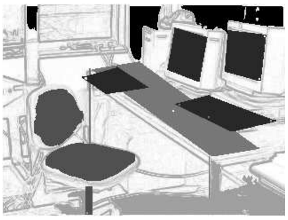

図 6: 認識結果. 椅子, 机, WS, 床, 壁がほぼ認識できている．

図 5の室内サンプル画像に対する実験では，「机」「椅子」がそれぞれ２つ (図 7, 図 8)，「WS」が２組のディスプレイとキーボード，「床」「壁」「空」「木」が１つづ

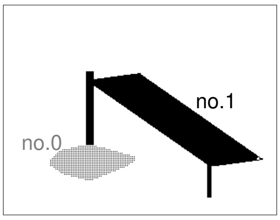

図 7: 机の候補．

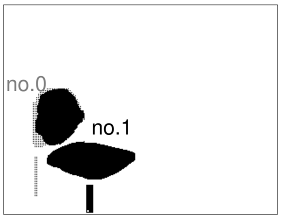

図 8: 椅子の候補．

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，取り消しが6回，復活が2回行なわれ，最終的に図6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，背もたれが見付かっているので形状の評価値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，取り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なわれており，例えば，「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，机の上方に関してWSの探索を行なっている．そのために，平行四辺形の認識という単純処理しか行なっていないのにも拘らず，ディスプレイとキーボードが認識できている．

次に，図9の屋外サンプル画像に対する実験では，「空」「自動車」「道路」「木」「壁」の候補が1つづつ生成された．これらの候補生成によって，競合が1回，取り消しが1回が行なわれ，最終的に図10の認識結果が得られた．

5 おわりに

本発表では，形状知識を扱う認識モジュールと，関係知識を扱う通信モジュールとを別々に構築し，それらを組み合わせて1つのエージェントとして，エージェントの集合体として認識システムを構築する方法を述べた．そうすることにより，室内画像と屋外画像とを同時に認識できる，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実現でき，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．

今後は，今回は非常に簡単な手法で実現していた同一一般名称で呼ばれる物体を認識する認識モジュールの実現法を研究していく必要がある．

図 9: 屋外サンプル画像．

図 10: 認識結果．道路, 木, 空, 自動車がほぼ認識できている．

参考文献

[1] M. Nagao and T. Matsuyama. *A Structure of Complex Aerial Photographs.* Plenum Press, New York, 1980.

[2] T. Matsuyama and V. S. Hwang. *SIGMA: A knowledge-based aerial image understanding system.* Plenum Press, New York, 1990.

[3] Y. Ohta. *Knowledge-Based Interpretation of Outdoor Natural Color Scenes.* Pitman Advanced Publishing Program, Boston, 1985.

[4] B. Draper, R. Collins, J. Brolio, A. Hanson, and E. Riseman. The schema system. *International Journal of Computer Vision*, Vol. 3, No. 2, pp. 209–250, 1989.

[5] L. Stark. Functionality in object recognition. *Computer Vision and Image Understanding*, Vol. 62, No. 2, pp. 145–146, 1995.