

画像オブジェクトに基づく3次元構造の獲得

守田 了*・田中 稔*

3-D Structure Reasoning Based on Image Objects

Satoru MORITA* and Minoru TANAKA*

Abstract

We have been creating a spatial reasoning system using the sequential image with noise. One of the problem of the image understanding is that image processing and image understanding are often studied independently. In this paper, we introduce an image object to compute image understanding and processing in the same frame. An image object is a computer, which autonomously observe, move and reason in cooperation with other image object. They are defined as the hierarchical model, so an object changes into objects of the various hierarchical levels to observe their image in the adaptive resolution.

An example shows that our object modelling is suitable for spatial reasoning.

1. はじめに

自立移動ロボットや知的ロボティクスを目的として、カメラ視点を移動しつつ撮影した時空間画像からシーンの奥行き情報をえる運動立体視の研究がおこなわれている。厳密な復元が可能な反面、視点と対象の厳密な位置が既知であることが前提になっているため、使用できる場面は制限される。本研究では、視点と対象の厳密な位置がわからない時系列画像から大まかな空間配置を知るために、空間推論マシンを作成する。

画像から輪郭線を抽出するアプローチ、輪郭線から空間を推論するアプローチが研究されている。個々の技術は研究されているにもかかわらず、個々の技術を組み合わせた場合に、雑音、反射や影の影響で輪郭と物体の輪郭が一致しないため、実画像には適用しにくい。これは、画像の抽出と画像の理解を別の枠組で計算していることに起因する。R. Brooksは従来の知能ロボットのアーキテクチャとは異なるサブサンプシオンアーキテクチャを提案している⁽¹⁾。サブサンプシ

ンアーキテクチャは、一般化された処理ユニットの直列的実行によって知覚、制御処理を実現するのではなく、各ユニットが独立に行動を実行しようとする並列的実行によって知覚、制御処理を実現する。他方、ネットワークや計算機のアーキテクチャにおいて、分散型アーキテクチャが注目され、並列オブジェクト指向に関する研究が行われている⁽²⁾。そこで、画像の抽出と推論を同じ枠組で行う画像オブジェクトを導入し、複数のオブジェクトが並列に画像を抽出し、推論するアーキテクチャを提案する。画像オブジェクトは個々のオブジェクトが自立的に移動し推論するエージェントである。画像中に存在する複数のオブジェクトが協調しながら推論する。特に、異なる尺度で画像を同時に観測するために、階層的にオブジェクトを定義し、異なる階層の複数のオブジェクトが各オブジェクトに応じた推論結果を排出する。不確実な推論をインクリメンタリに行ない、後の時系列でその不確実な推論を修正しやすい構造に設計する。

推論は、2次元画像から決定される定性的変化のみを用いて行う。Jainらは、内部輪郭線を含む線画のトポロジーをアスペクトにとり、輪郭線の連続の仕方の違いで視点空間を分割し、CADデータからアスペクト

*知能情報システム学科

ラフ⁽¹⁰⁾が生成できることを示している⁽⁸⁾。本研究では、しゃへい輪郭の変化に着目しているが、アスペクトグラフを生成することが目的ではなく、物体の位置関係を推論することが目的である。そのため、異なる物体及び同一物体間に起こるしゃへい輪郭とその位置関係を明らかにし、推論規則として用いる。隣接オブジェクト間の関係からしゃへい輪郭の変化を抽出し推論に利用する。統合されたオブジェクトは局所的推論結果から矛盾を抽出し、推論の正しさを確信する。

2章では画像オブジェクトモデルのアーキテクチャについて述べる。3章ではオブジェクトの4つの機能（観測、自立的移動、推論手法、群化の機構）について述べる。4章では実時系列画像に対して、推論マシンを適応し、手法の有効性を示す。

2. 画像オブジェクトモデル

画像抽出を前処理として行い、その後画像理解を行うという計算の枠組みがほとんどであった。しかし、一定の解像度で観測すると、ある画素では細かく観測しすぎ、ある画素ではおおまかに観測してしまうことがおきる。このことは、画像中のすべての画素で同一の解像度で観測を行っていることに起因する。また、観測結果に無関係に推論すると、ある画素では推論を細かくしすぎ、ある画素ではおおまかに推論しすぎることがおきる。このことは、画像中のすべての画素で同一の基準で推論を行っていることに起因する。

本稿では特に次の点に着目し、推論マシンを作成する。

- 観測および推論結果に応じた画像の抽出の実現
- 観測および推論結果に応じた画像の推論の実現

予測の確信度が高い画素については、おおまかに観測する。逆に、予測の確信度が低い画素については、より細かく観測する。特に誤った推論結果に対処できるように、大局的な観測と局所的な観測を容易に変換可能にする。例えば、画像中の画素のある部分では輪郭であるという前提のもと観測し、ある部分では物体であるという前提のもと観測をする。このように、違った尺度に基づき並列に観測して、高速に推論する。

そのために本稿では、画像オブジェクトモデルを導入する。画像オブジェクトは画像抽出と空間推論を同じ枠組みで行うエージェントである。本稿で提案するオブジェクトモデルは次の性質を持つ。

- 観測画像のどの画素にも少なくとも一つの画像オブジェクトが存在する。
- 画像オブジェクトは隣接する画像オブジェクトの

情報しかわからない。

- 画像オブジェクトは自身を計測する。
- 画像オブジェクトは自身が存在する画素のみを観測する。（観測）
- 画像オブジェクトに応じた推論結果を排出する。（推論）
- 画像オブジェクトは自身が存在する画素と隣接の画像オブジェクトと協調して変形、移動する。
- 画像オブジェクトは自律的に移動する。（自立移動）
- 画像オブジェクトは様々なオブジェクトに変換可能である。
- 画像オブジェクトを変換するために推論の確信度を計算する。（群化の機構）

画像オブジェクトは以上の4つの機能（観測、移動、推論、群化）を持っている。

2.1 階層的オブジェクト

最も基本的なオブジェクトは線分である。情景画像から抽出されたエッジを曲率の不連続点をもとに分解し、得られた線分を曲率要素を基に分類する。曲率要素の分類にはprocess-grammarを用いる⁽⁷⁾。すなわち、曲率の正負および極小極大により要素をM+, M-, m+, m-に分類する。画像中からこのように分類される線分集合 $l_1, l_2, l_3, \dots, l_i, \dots, l_n$ を抽出する。但し、すべての線分集合の l_n 間には曲率の不連続点は存在しない。それぞれの線分は次の属性がある

1. 開曲線：片側及び両側が他の線分と連結していない場合
2. 閉曲線：両側が他の線分と連結している場合
3. 各線分 $\theta(1)$ は以下の計算からさらに細かい輪郭要素に分解される。

$$\phi(\sigma, l) = \theta(l) * G(\sigma, l)$$

$$\text{但し } G(\sigma, x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{x^2}{2\sigma^2}}$$

である。同一輪郭線に対して、 σ を小さくした場合連結点は保存される⁽⁵⁾⁽⁴⁾。この性質がオブジェクトが再帰的に分解できるために重要な性質である。

隣接する線分間はリンク点 $lx_1, lx_2, lx_3, \dots, lx_i, \dots, lx_n$ で関係づけられる。リンク点は少なくとも1点曲率の不連続点を含む点であり、近傍の曲率要素と微分幾何学的連続性により分類される。各リンク点の種類とリンクする線分とリンク点の2次元画像中の位置を記述するFig.1(a)は線分間を関係づけるリンク点のタイプの分類である。(1)(2)(3)はそれぞれ4本、3本、2本の線分間の接合点のタイプである。接合点のタイプ

は、接合点におけるすべての線の組合せのうち、接合点に微分幾何学的に不連続な点が現れる線の数である。Fig.1(b)は2本の線のリンク点にProcess-grammarをあてはめた場合である。Fig.1(a)はprocess-grammarを割り当てることによりさらに分類される。Fig.2(1)(2)(3)はそれぞれ2本、3本、4本の線分のリンク点のタイプの変化を表している。これらの変化は形状がなめらかに変形する場合に現れる。2本から3本、3本から4本へ変化するリンク点の変化とあわせて、線分間の変形規則（オペレータ）として利用する。

2.2 階層的特徴

さらにそれぞれ線分集合は群化の指標の計算により、次に示す階層的特徴（クラス）にまとめられる。物体性の3つのレベルに応じて、オブジェクトは3つのクラス（線分、輪郭、物体）に分けられる。それぞれ他のオブジェクトとの関係集合を保持する。

- 線分集合 [l1,l2,l3,...,li,...,ln]
リンク点 [lx1,lx2,lx3,...,lxi,...,lxn]
- 輪郭集合 [c1,c2,c3,...,ci,...,cn]
関係集合 [cx1,cx2,cx3,...,cxi,...,cxn]
- 物体集合 [o1,o2,o3,...,oi,...,on]
関係集合 [ox1,ox2,ox3,...,oxi,...,oxn]

上位特徴（クラス）は下位特徴から構成されている。各輪郭、物体はそれぞれ線分集合、輪郭と線分集合を属性として持つ。各輪郭の関係cx1,cx2,cx3,...,cnには輪郭が接触する部分にできる線分集合l1,...,lnとそのリンク点の集合lx1,...,lxnを記述する。また物体の関係は物体を形づくる輪郭集合c1,...,cnとその関係集合cx1,...,cxnが記述される。本研究では、各成分を上位

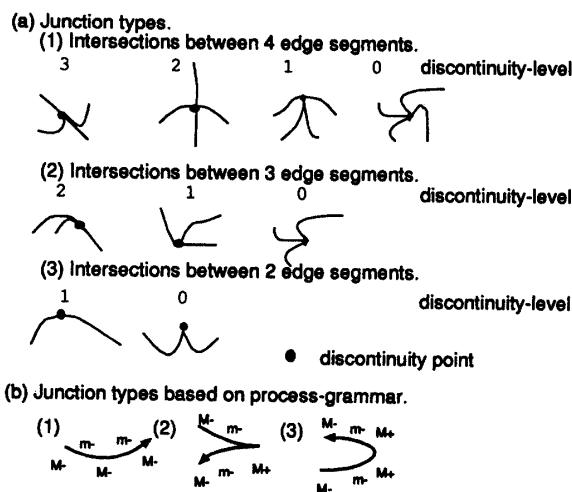


Fig.1 Junction types between edge segments.

の特徴（クラス）にまとめるために、同一物体における動きの連続性および幾何学的位置関係の拘束を用いる。

3. オブジェクトの機構

各オブジェクトは一つのエージェントであり、自身で画像中を移動し、推論する。そのため、オブジェクトは観測、自律的移動、推論、群化の4つの機能を持っている。

3.1 オブジェクトの観測

オブジェクトは自身で画像中を移動し、空間を推論するために、外界を観測する機能を保持している。画像中に複数のオブジェクトが協調して存在する。観測画像からすべての対象を抽出するために、観測画像中のすべての画素を、少なくとも一つのオブジェクトが観察していることが必要である。すなわち画像中のオブジェクトは次の条件を満たす。

$$\bigcup_{i=1}^n I_{(i)}(u^{(i)}, v^{(i)}) = \Sigma$$

$o_i^t (i = 1, \dots, n)$ は t 時刻に存在する観測画像中の n 個のオブジェクトであり、 $(u^{(i)}, v^{(i)}) (j = 1, \dots, n(i))$ は o_i が観測する画像である。 $n(i)$ は o_i が観測する画素数であり、 Σ は観測画像の画素全体である。

個々のオブジェクト o_i は次に示すデータを観測している。

- 一時刻前の $o_i^{(t-1)}$ 。
- 存在する画像中の位置 $I_{(i)}(u^{(i)}, v^{(i)}) (j = 1, 2, \dots, n)$

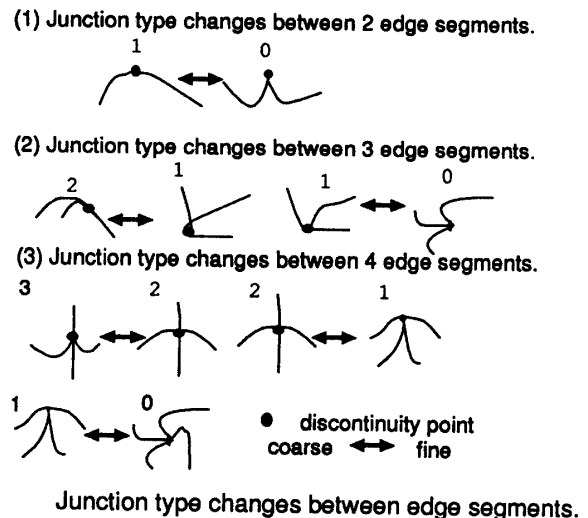


Fig.2 Deformation processes of junction types.

(i)

- 形状解析のためかけるフィルターの係数 $\sigma_shape_{i_1}^t$
- 画像にかけるフィルターの係数 $\sigma_image_{i_1}^t$.
- 隣接するオブジェクト $o^t_neigh(i, j)$ ($j=1, 2, \dots, n$).
- 隣接するオブジェクトとの2次元画像中の位置関係 $Relate_2D$.
- 隣接するオブジェクトとの3次元空間中の視線に依存する位置関係 $Relate_3D$.
- 隣接するオブジェクトとの3次元空間中の視線によらない位置関係 $Relate_3D$.
- 下位オブジェクト $o^t_child(i, j)$ ($j=1, \dots, n$).
- オブジェクトの位置関係の時間変化 $vector1(i), vector2(i)$
- 形状 $shape(i)$

但し $vector1(i) = place(o_i^t) - place(o_i^{t-1}), vector2(i) = place1(i)^t - place1(i)^{t-1}$ とする. ($placeu(i),$

$$placev(i)) = \left(\frac{1}{n} \sum_{j=1}^n u^{(j)}, \frac{1}{n} \sum_{k=1}^n v^{(k)}\right) \text{ とする. } shape(i) \text{ は,}$$

$M+M-m+m-n$ などの曲率要素を保持する.

3.2 オブジェクトの自律的移動

オブジェクトは自律的に対応点を探す. 各オブジェクトは変形オペレータを選択し, 移動, 変形する. 特に, 親クラスの推論結果に矛盾せずに, かつ動きの連続性を保持した移動変形を実現する. 特に, オブジェクトは, 親クラスの推論結果に矛盾がなく, また動きの連続性を保持しつつ移動する.

- 線分 [li]
オペレータ [lo1, lo2, lo3, ..., loi, ..., lon]
- 輪郭 [ci]
オペレータ [co1, co2, co3, ..., coi, ..., con]
- 物体 [oi]
オペレータ [oo1, oo2, oo3, ..., ooi, ..., oon]

オブジェクトは自身が存在する画素および隣接するオブジェクトが存在する画素を観測し, 自律的に移動する. オブジェクト o_1 は, $I_{(j)}^{-1}(u^{(j)}, v^{(j)})$ ($j=1, 2, \dots, n$), $I_{k(l)}^{-1}(u^{(l)}, v^{(l)})$ ($l=1, 2, \dots, n$), ($j=1, 2, \dots, n_1$) にフィルタリング係数 $\sigma_image(i)^{t-1}$ をかけた画像を観測する. 観測画像をクラスタリングし, 係数 $\sigma_shape(i)^{t-1}$ でフィルタリングをし, 形状解析を行い得られた輪郭を $\{ob_j(j=1, \dots, m)\} = OB$ に分解する. $\{o^t_child(i, j)(j=1, \dots, n)\} = OC$ と OB を比較する.

オブジェクトが移動するオペレータの集合 OP $\{op1 \dots opn\}$ を求める. n 個のオペレータをオブジェクト o_i にかける. 観測画像の一部に移動させる. $OJ = opn(\dots(op1(o_i)))$ 但しオブジェクトが消失することを許す変形では, 隠れオブジェクトとして, 画像中の隠れた位置とオブジェクトを保存する. この隠れオブジェクトを除くと, $OJ \subset OB$ である.

推論結果に基づくオペレータの制限関数 Con_Induct を次のように定義する.

$$Con_Induct(Relate_3D(j) \mid OP, Relate_3D(i)^{t-1}) = \begin{cases} 1 & Induct^t = Induct^{t-1} \text{ の場合} \\ k & \text{それ以外} \end{cases}$$

すなわち一時刻前の推論結果を導く変形オペレータは 1 それ以外を k と計算する. 例えば $Relate_3D^{t-1} = Connect$ なら $Connect$ の推論結果を導くオペレータが 1 それ以外が k となる

動きに基づくオペレータの制限関数 Con_move を以下のように定義する.

$$Con_Move(vector(j)^t \mid OP, vector(i)^{t-1}) = \prod_{i=1}^n vector1^t(i) * vector1^{t-1}(i)$$

または,

$$\prod_{i=1}^n vector2^t(i) * vector2^{t-1}(i)$$

*は内積である. この関数値が 1 にちかいほどなめらかな動きをする.

形状に基づくオペレータの制限関数 Con_shape を以下のように定義する.

$$Con_shape(shape(j)^t \mid OP, shape(i)^{t-1}) = \begin{cases} 1 & shape(j)^t \subset shape(ob_j) \{j=1, \dots, m\} \text{ の場合} \\ k & \text{それ以外} \end{cases}$$

但し $Relate_3D = Behind$ の場合は対応がない場合にも隠れを考慮し, $Con_shape = 1, Con_Move = 1$ として計算する.

Con_move と Con_induct と Con_shape を用いてオブジェクトの操作オペレータを求める.

$$Con(OP, o_i^{t-1}) = \sum_{i=1}^n (Con_Induct(OJ \mid OP, o_i^{t-1}) + Con_Move(OJ \mid OP, o_i^{t-1}) + Con_shape(OJ \mid OP, o_i^{t-1}))$$

が最小でかつ $OJ \subset OB$ を満たす OP の集合を選ぶ. 但し適応するオペレータが w 個以上の場合 (実験では $w=3$) は計算を打ち切る前の時刻のオブジェクトを保存する.

3.3 オブジェクトの推論

オブジェクトは隣接オブジェクト間の関係の変化を

用いて推論する。一時刻前の二オブジェクト間の関係の変化からわかる局所的観測に基づく推論と複数のオブジェクト間の推論の通信によってわかる大局所的観測に基づく推論がある。位置関係には、視点に依存する位置関係と視点に依存しない位置関係がある。例えば、前、後ろは前者であり、隣接は後者にあたる。本手法では、局所的観測に基づく推論から大局所観測に基づく推論を導き、視点に依存する位置関係から視点に依存しない位置関係を導く。そのため、オブジェクト間のレベルに応じて、オブジェクト間の関係 $Relate_2D, Relate_3D, \bar{Relate_3D}$ を定義する。

$Relate_2D$ は各オブジェクト o_i と隣接オブジェクト間の関係である。特に、2次元画像 (u,v) 上の幾何学的変化のみを記述する。

$Relate_3D$ はオブジェクト o_i と隣接するオブジェクト間の関係である。2オブジェクト間の3次元空間 (x,y,z) 中の幾何学的変化のみを記述する。 $Relate_3D$ は視線方向が変わるごとに変化する。2オブジェクト間の空間的位置関係を $Relate_2D$ から推論する。2オブジェクト間の画像における位置関係から推論される2オブジェクト間の空間的位置関係 $Induct(Relate_2D^t | Relate_2D^{t-1})$ を定性的に分類する。推論結果を $Induct^t = Connect(A,B), Disconnect(A,B), Front(A,B), Up(A,B), Neighbour(A,B)$ のように記述す

る。Fig. 3 はリンク点の変化に基づく推論結果である。リンク点の変化は、線分オブジェクト間の関係の変化に基づく推論に対応する。(A) はリンク点が発生する場合であり、(B) はリンクタイプが変わるときの推論結果に対応する。(a)(b) はそれぞれ1点のリンク点の変化に基づく推論、2点のリンク点の変化に基づく推論を表している。Fig. 4 は輪郭に基づく接合線の変化に基づく推論結果である。(A) は輪郭間の接合関係の変化に基づく推論結果であり、(B) は輪郭オブジェクトの生成による関係の変化に基づく推論結果である。これらの変化は $Relate_3D$ のオブジェクトが両方輪郭の場合の輪郭間の変化に基づく位置関係に対応する。

視線の影響を受ける位置関係を、視線に不変な位置関係に変換する。 $\bar{Relate_3D}$ は視線の影響を受けない4オブジェクト間の位置関係を記述する。もし $Connect(A,B), Connect(A,C), Connect(A,D)$ なら同物体間の位置関係を $pyramid_connect(A:B,C,D)$ と記述する。もし $Disconnect(A,B), Disconnect(A,C), Disconnect(A,D)$ なら異なる物体間の位置関係を $pyramid_disconnect(A:B,C,D)$ と記述する。BCS間の位置関係を同時刻における位置関係から、 $pyramid_disconnect(A:B,C,D)$ か $pyramid_disconnect(A:D,C,B)$ かを制約する。

(a) Edge segment	
(A) Generating of linking points	Reasoning
	A:front,B:back
	A:front,B:back
	A:front,B:back
	A:back,B:front
(B) Changes of linking types	Reasoning
	A,B:connect
	A,B:connect
(b) Contour	
(A) Generating of linking points	Reasoning
	A:front,B:back
	A:back,B:front
(B) Changes of linking types	Reasoning
	A,B:connect
	A,B connect
	A,B connect

Fig.3 Deformation processes of linking types and reasoning.

(A) Changes of linking types.	Reasoning
	A,B:connect
	A,B:disconnect
	A,B:connect
	A,B:disconnect
(B) Generating of a contour object	Reasoning
	A:front,B:back B:appear
	A,B:connect B:appear
	A,B:disconnect B:appear
	A:front,B:back B:hide
	A,B:connect B:hide
	A,B:disconnect B:hide

contour objects A ● contour objects B ●

Fig.4 Relation changes of contour objects and reasoning.

3.4 オブジェクトの群化（統合と分解）

推論結果と動きに基づき同一物体でない可能性が高い場合はオブジェクトを分解し、他のオブジェクトと同一物体性が高い場合はオブジェクトを統合する。まず、同一物体性の確信度関数を定義し、次に確信度関数を用いたオブジェクトの分解および統合を実現する。

3.4.1 同一物体性の指標

複数のオブジェクトが同一物体であることを推論するために確信度関数 $Uni_induct, Uni_move1, Uni_move2$ を定義する。

オブジェクト集合 $OB \subset \{ob1, ob2, \dots, obn\}$ が同一物体である確信度の計算を推論に基づいて行う。確信度関数 $Uni_move(OB)$ を次のように定義する。

$$Uni_induct(OB) = \prod_{i=1}^n Uni_induct(ob_i, OB)$$

但し $Uni_induct(ob_i, OB)$ はオブジェクト ob_i がオブジェクト集合 $OB \subset \{ob1, ob2, \dots, obn\}$ と同一物体であることを推論結果に基づいて計算する確信度関数である。

もし $Relate_3D(ob_j, ob_i) = CONNECT$ となる

$$Uni_induct(ob_j, OB) = \begin{cases} 1 & ob_i \in OB \text{ が存在する時} \\ k & ob_i \in OB \text{ が存在しない時} \end{cases}$$

オブジェクト集合 $OB \subset \{ob1, ob2, \dots, obn\}$ が同一物体であるという確信度の計算を以前の動きに基づいて行う。動きの時間微分による確信度関数、 $Uni_move1(OB)$ と時間2時微分による $Uni_move2(OB)$ を以下のように定義する。

$$Uni_move2(OB) = \prod_{i=1}^n Uni_move2(ob_i, OB)$$

但し $Uni_move2(ob_i, OB)$ はオブジェクト ob_j がオブジェクト集合 $OB \subset \{ob1, ob2, \dots, obn\}$ と同一物体であることを動きの時間微分に基づいて計算する確信度関数である。

もし $Relate_2Dmove2(ob_j, ob_i) = vector2(ob_j) * vector2(ob_i) = 1$ となる。

$$Uni_move2(ob_j, OB) = \begin{cases} 1 & ob_i \in OB \text{ が存在する時} \\ k & ob_i \in OB \text{ が存在しない時} \end{cases}$$

$$Uni_move1(OB) = \prod_{i=1}^n Uni_move1(ob_j, OB)$$

但し $Uni_move1(ob_i, OB)$ はオブジェクト ob_i がオブジェクト集合 $OB \subset \{ob1, ob2, \dots, obn\}$ と同一物体であることを動きの一次微分に基づいて計算をする確信度関数である。

もし $Relate_2Dmove1(ob_j, ob_i) = vector1(ob_j) * vector1(ob_i) = 1$ かつ $vector2(ob_j) = 0, vector2(ob_i) = 0$ となる

$$Uni_move1(ob_j, OB) = \begin{cases} 1 & ob_i \in OB \text{ が存在する時} \\ k & ob_i \in OB \text{ が存在しない時} \end{cases}$$

但し $k = 0.8$ とおいて計算を行っている。

3.5 統合と分解

同一物体性を計算し、同一物体性が高ければよりおおまかに観測するためにオブジェクトを統合し、同一物体性が低ければ、より詳細に観測するためにオブジェクトを分解する。オブジェクトの統合、分解の指標を $Uni_move1, Uni_move2, Uni_induct$ を用いて計算する。

オブジェクト o_i の分解の指標の計算は、下位オブジェクト集合を用いて計算する。 $CHILD(o_i) = o_i^{child}(i, j)$ ($j = 1 \dots n$) $Uni_induct(CHILD(o_i)) < K$ かつ $Uni_move1(CHILD(o_i)) < K$ または $Uni_move2(CHILD(o_i)) < K$ を満たせば o_i は同一物体であると確信する。この条件を満たさない場合 o_i を $CHILD(o_i)$ に分解して再計算を行う。

オブジェクト o_i の群化の指標の計算は、隣接オブジェクト集合を用いて計算する。 $o_i^{neigh}(i, j)$ ($j = 1, \dots, n$) は $Uni_Induct(o_i^{neigh}(i, j), OB) = 1$ かつ $Uni_move1(o_i^{neigh}(i, j), OB) = 1$ または $Uni_move2(o_i^{neigh}(i, j), OB) = 1$ をみたと $o_i^{neigh}(i, j)$ の集合を OB と同一物体であると仮定し、再計算を行う。

線分オブジェクトでは、下位オブジェクトとして σ_shape を $\frac{\sigma_shape}{2}$ とすることで、再帰的に分割が可能になる。また、線分オブジェクトでは、上位オブジェクトとして σ_shape を $2 * \sigma_shape$ に変えることにより、上位オブジェクトへ統合される。さらに上の条件を満たすことにより、上位クラスである輪郭オブジェクト、物体オブジェクトに統合される。逆に上位オブジェクトである輪郭オブジェクト、物体オブジェクトは線分オブジェクトに分解される。但し、3本以上の交差点においては、異なるクラスの統合を行っても、線分オブジェクトの同一クラス間の統合は行わない。これは、交差点が非常にクリティカルな点であり、推論に必要

な状態の変化がおきる可能性が高い点である。群化と分解の指標の計算を繰り返すことによって、しだいに同一物体性の高い物体は統合され、同一物体性の低い物体は分解され、物体ごとの追跡が可能になる。

4. 実験

オブジェクトの推論アルゴリズムを示す。

- 観測画像から輪郭線を抽出し、観測する。(観測)
- 前の時系列の推論結果を用いて、観測画像中の対応点に移動する変形オペレーターを選択する。(自律的移動)
- 隣接オブジェクト間の関係の変化を観測し、空間的位置関係を推論する。(推論)
- 同一物体性の確信度を計算し、確信度が低ければ、オブジェクトを分解し、確信度が高ければオブジェクトを統合する。(群化)

時空間画像の観測につれて、不確実な推論は打ち消され、確実性が高い絶対的位置関係が得られる。Fig. 5は画像オブジェクトに基づく空間推論システムの構成を示している。複数の異なるクラスのオブジェクト(物体オブジェクトa1-a5, 線分オブジェクトb1,b2, 輪郭オブジェクトc1-c2)が画像を観測し、それぞれのオブジェクトに応じた推論結果を排出する。物体オブジェクトa1-a5は2つの輪郭オブジェクトa1-a3, a4-a5を下位クラスのオブジェクトとして持つ。また、輪郭オブジェクトc1-c2は2つの線分オブジェクトc1,c2を下位クラスのオブジェクトとして持つ。推論システムによって、異なるオブジェクトが排出した推論結果を視点に依存しない位置関係に統合する。

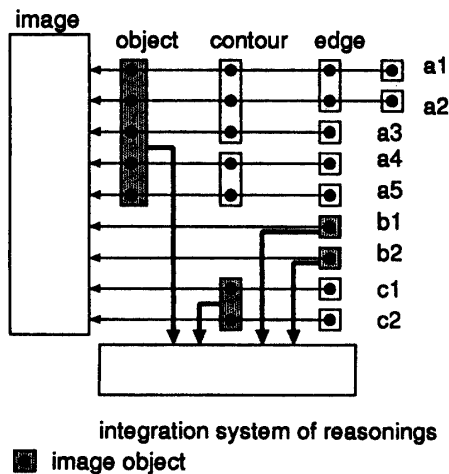
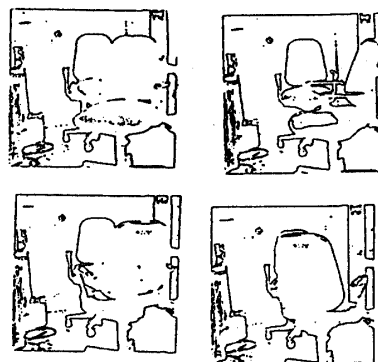


Fig.5 Architecture of spatial reasoning system based on image objects.

実験には、いすの前で他のいすを回転させた情景を撮像した時系列画像を用いた。Fig.6(A)は実験に用いた時空間画像の原画像のうちの4枚であり、Fig.6(B)は4枚の画像から観測のために抽出された輪郭線である。抽出された輪郭線は雑音反対の影響によりリンク点の欠落や連結が生じている。Fig. 7は $\sigma=1.2$ における分解された線分オブジェクトla1,la2,la3,la4,la5である。Fig 8は時刻t1からt9の画像から画像オブジェクトが統合分解される過程である。時刻t1では、オブジェクトは線分と背景であり、それぞれの線分オブジェクトは上位クラスを持たない。時刻t2で線分オブジェクトは上位クラスの輪郭オブジェクトGに統合され、さらに時刻t3でオブジェクトA,Cに統合される。オブジェクトAがいったん輪郭オブジェクトと推論されると、Aを構成する線分が隠れても、オブジェクトの前に物体が現れて隠れたことを認識して、後ろに対応する線分があることを識別する。Fig. 9は後方にあるいすの背もたれの部分 (Object A) の近傍の記述の変更と排出された推論結果である。それぞれラベルはFig. 8に対応して



(A)Original images.



(B)Contour images derived from original images.

Fig.6 Experimental Images.

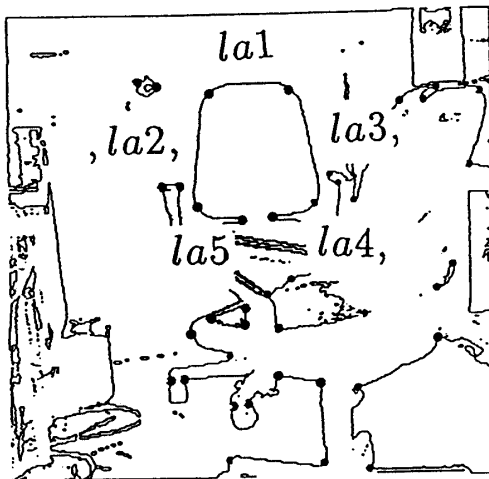


Fig.7 Edge segment objects on $\sigma=1.2$.

time	edge segment sets	added reasoning result
t1	(la1,la2)	lb1(front)
t2	(la1,la2,la3,la4)	le1(front)
t3	(la1,la2,la3,la4,la5)	la1,la2,la3,la4,la5(ObjectA)
t4	(la6,la3,la4,la5)	la1,la2[la6] ObjectJ(:c) ObjectB(front:d)
t5	(la6,la5,(la3,la4))	lb2(front) ObjectJ(hide) la3,la4(hide)
t6	(la6,(la3,la4,la5))	lb3(front) ObjectH(:c) la3(hide)
t7	(la1,la2,(la3,la4,la5))	lb4(front) la6-la1,la2 la2(hide) la3(appear)
t8	(la1,la3,la4,(la2,la5))	

((hidden line sets) subclass[parent class]
• linking point image object(reasoning)
:d disconnect :c connect

Fig.9 The reasoning results, which an object A outputs.

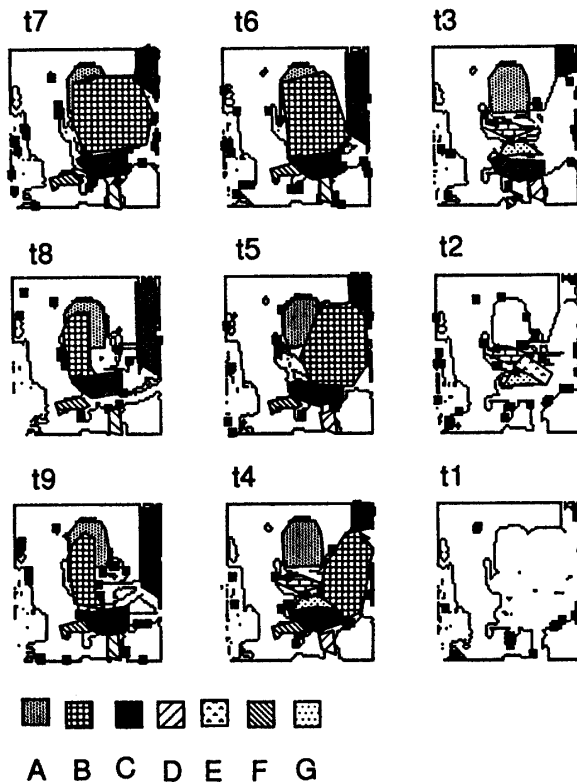
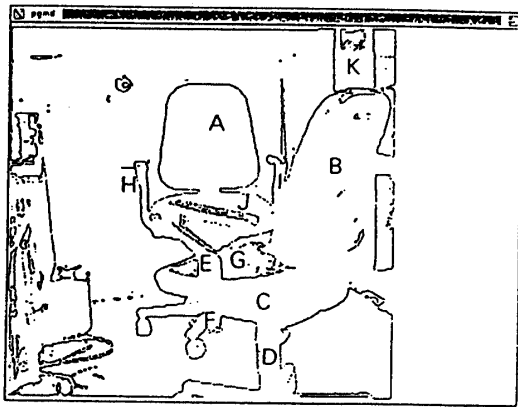


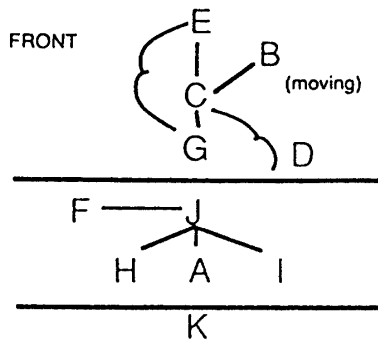
Fig.8 Composition and decomposition processes of objects for the sequential images : Fig.6.

いる。図で示された推論結果は視点に依存するrelate-3 Dである。時刻t1からt2ではlb1, 時刻t2からt3ではle1がAの前にあることを推論する。時刻t4では、 σ -shapeを $2 * \sigma$ -shapeに変えて観測し、線分オブジェクトla1とla2が統合されla6となり、時刻t7で確信度関数が小さく

なったため、 σ -shapeを $\frac{\sigma-shape}{2}$ に変えて観測をし、la6はla1,la2に分解される。時刻t3でla1からla5が上位クラスの輪郭オブジェクトに統合されているため、t5でla4,la5の隠れを推論する。また、時刻t1では、線分オブジェクトlb1が前にあることを推論したが、時刻t4,t5ではlb1をオブジェクトBと一部の輪郭であることを識別し、オブジェクトBが前にあることを推論する。Fig.10は数回のいすの回転により推論された視点に不変な位置関係である。図 (b) 中のアルファベットのリンクは接続関係を表している。Fig.10(b) はFig. 9で排出された推論結果を視点に不変な位置関係に変換した。時系列の初期には、雑音のため、統合と分解を繰り返すが、いったん物体を認識すると推論能力がアップする。また、時系列の初期においてもオブジェクトが線分であるか物体であるかにかかわらず、推論結果を出力するので、おおまかな推論が可能になっている。物体が識別されると対応点が少なくなるため、おおまかな位置関係を高速に把握することが可能になる。対応点のとりにくい雑音の多い輪郭が画像を記号要素に変換し、記号要素の変化から3次元構造を推論した。雑音による輪郭の欠落およびリンクの欠落のため、実画像では多義性が増す。多義性の高い部分は可能性をすべて記述し、次の時刻に3次元構造を補正する。本手法は曲線による記述を根底においているため、物体の曲面によって生じる遮蔽輪郭から、曲面の凹凸構造の把握が可能である。



(a) A labeled image



(b) The final reasoning results constructed by reasoning results, which is outputted by many objects.

Fig.10 The final reasoning results.

5. おわりに

本研究では、雑音の多い時系列画像からおおまかな空間配置を推論するために、画像オブジェクトを導入した。画像オブジェクトは画像処理と空間の推論を内在したエージェントである。オブジェクトが階層的に定義され、異なる階層の複数のオブジェクトが異なる尺度で画像を同時に観測し、各オブジェクトに応じた推論結果を排出する。時系列の初期において、オブジェクトが線分であるか物体であるかにかかわらず、オブジェクトに応じた推論結果を出力するため、おおまかな推論が可能になる。いったん物体が識別されると対応点が少なくなるため、おおまかな位置関係を高速に把握することが可能になる。オブジェクトが複数の物体から構成されていると判断されると、自律的に異なるオブジェクトに分解され、複数のオブジェクトが同

一物体と判断されると統合されるため対応計算の効率化がはかられている。本手法は、雑音、反射により、輪郭の一部が欠落したり、連結された画像から、インタリメンタリにおおまかな空間配置を推論する場合に有効である。

本研究の一部は実吉奨学会研究助成金（研助第0529号）の援助によった。

参考文献

- 1) A. Yonezawa and Tokoro, "Object-Oriented Concurrent Programming"; The MIT Press (1987)
- 2) R.A.Brooks, "A Robust Layered Control System For a Mobile Robot"; IEEE Journal of Robotics and Automation RA-2, 1, pp.14-23, (1986)
- 3) D. Marr, "Vision"; W.H.Freeman, Sanfransisco (1982)
- 4) J. Babaud, A.P. Witkin, M.Baudin, and R. O. Duda, "Uniqueness of the Gaussian kernel for scale space filtering"; IEEE Trans. Pattern. Anal.&Mach. Intell.-8, 1, pp.26-33, (1986)
- 5) A.P.Witkin, "Scale-space filtering"; Proceeding of International Joint Conference Artificial Intelligence, Karlsruhe, West Germany, pp.1019-1022, (1983)
- 6) R.Reiter, "A Logic for Default Reasoning"; Artificial Intelligence, 13, pp.81-132, (1980)
- 7) Leyton M, "A Process-grammar for shape"; Artificial Intelligence, 34, pp.213-247, (1988)
- 8) Rieger, J, "On the Classification of Vision of Piecewise Smooth Objects"; Image and Vision Computing,, pp.91-97, (1987)
- 9) J. Ponce and D.J.Kriegman, "Computing exact aspect graphs of curved objects: parametric patches"; Technical Report UIUC-DCSR-90-1579, University of inois at Urbana-Champaign, (1990)
- 10) Koenderink, J.J and van Doorn, A.J, "The Internal Representation of Solid Shape with Respect to Vision"; Biological Cybernetics, 32, pp211-216, (1984)

(平成5年10月15日受理)