

## 関係モデルを用いた物体認識の研究

勝間 大輔<sup>†</sup> 清水 大輔<sup>†</sup> 渡邊 陸<sup>†</sup>

<sup>†</sup> 鹿児島大学 大学院理工学研究科 〒890-0065 鹿児島市郡元 1-21-40

E-mail: † {sc097012, sc101022, mutty}@ics.kagoshima-u.ac.jp

あらまし 室内に存在する物体の種別・位置姿勢を2次元画像認識により求める際、複数物体の重なり・遮蔽や影などの影響によるコントラスト低下が生じた部分の特徴抽出が失敗し、正しく認識が行われないという問題がある。本論文では、認識対象個々の形状・属性情報を記述した物体モデルに加え、物体相互の位置関係、および人間動作との干渉関係を記述した『関係モデル』を適用することにより、上記の問題に対処する手法を提案する。撮影したシーンからまずエッジを抽出・ラスタベクトル変換・補完した後、連結性解析により閉面領域抽出を行う。次に物体モデルと照合することによりまず閉面単位のスコアを計算した後、隣接する閉面単位スコアを統合した統合スコアを求め、認識対象物体の候補を抽出する。ここで関係モデルを適用することにより、認識対象候補完の位置関係の整合性を検証し、矛盾の無い閉面領域の組み合わせを認識対象物体領域として得る。最後に物体同士の位置関係や人間動作との干渉関係を用いることにより、未認識領域に対する解釈当て嵌めを行う。現在実装しているのは認識対象物体の候補抽出部分までであり、室内環境シーンを対象とした実験結果について述べる。

キーワード 物体認識, 物体モデル, 関係モデル, モデルベースビジョン

## Object Recognition using Relation Models

Daisuke KATSUMA<sup>†</sup> Daisuke SHIMIZU<sup>†</sup> and Mutsumi WATANABE<sup>†</sup>

<sup>†</sup> Kagoshima University 1-21-40 Kohrimoto, Kagoshima-shi, Kagoshima, 890-0065 Japan

E-mail: † {sc097012, sc101022, mutty}@ics.kagoshima-u.ac.jp

**Abstract** 2D Object recognition system tends to fail when detected features are vastly lost caused by overlapping of multiple objects, occlusion and decline in contrast. This paper proposes a method to manage the *incomplete feature detection* problem by applying a “relation model” in addition to the conventional object model. First, detection of edge segments, raster-to-vector transformation and supplementation of edge segments are executed and closed region-based descriptions are created. Next, candidates of recognized objects are extracted by calculating the scores of each closed region based on matching with the object model, and unified scores are calculated by unifying the closed region scores in the neighborhood. Then, the consistency among recognized object candidates are examined by applying the relation model, and combinations of closed region are selected as recognized object regions. Finally recognition of unknown regions is executed by applying the relations among object positions and the relations between objects and human motion. Experimental results in indoor scenes by the implemented system are shown.

**Keyword** Object Recognition, Object Model, Relation Model, Model-based Vision

### 1. はじめに

物体認識は、画像の中に存在する物体の種別・位置・姿勢を求める問題[1]であり、認識対象である物体モデルを画像から抽出した特徴に基づく記述と照合することにより行うのが一般的である。しかし、ビデオカメラなど通常の撮影装置は、3次元空間を2次元画像へと投射する仕組みであるため、2次元画像認識に適用した場合は奥行き情報の欠如、複数物体の重なり・遮蔽や影などの影響によるコントラスト低下によ

り特徴抽出が失敗した場合に、不完全な記述結果しか得られず、物体認識が正しく行われないという基本的な課題がある。

遮蔽などを含む複雑シーンで生じる上記の問題に対処するため、『物体が他の物体の上に乗っている』という支持制約を適用することにより整合性の取れた認識結果を得る手法[2]が提案されている。これは、3次元情報を直接復元することなしに、モデル当て嵌めにより定性的な3次元構造を推定し制約を適用する点で、

2次元画像認識において有効な手法である。ただし支持制約は物体同士の関係制約の一部であり、他にもコンポーネントを構成する物体同士の近接関係など、様々な制約が存在する。また画像特徴が大幅に欠落し初期段階のモデル当て嵌めに失敗すると、支持制約の適用による整合性の確保が困難になる課題がある。

一方、人間が料理において器具を使用している動作を調理器具の認識に用いる研究[3]が行われている。屋内物体の大多数は人間が使用することを考えれば、人間動作の利用は自然であり有効であると考えられる。しかしながらこの研究例では調理シーンがズーム撮影されたテレビの料理番組映像を対象としており、また調理器具に特化した動作認識を組み込んでいることから、未だ一般的な物体認識へ適用できる段階には至っていない。

本論文ではこれらの提案を更に発展させ、認識対象個々の形状・属性情報を記述した物体モデルに加え、物体同士の支持関係のみならず、例えばPC筐体の周辺にはモニタやキーボード、マウスが存在するといった物体相互の近接関係、およびモニタの電源をオンにしたりマウスを操作するといった人間動作と物体の干渉関係を統一的に記述した『関係モデル』を適用することにより、上記の問題に対処する手法を提案する。

## 2. 提案手法

### 2.1. 全体の流れ

図1に、提案手法全体の流れを示す。

撮影したシーンからまずエッジを抽出・ラスタベクトル変換・細線化・特徴点(端点, 分岐点, 隣接接点, 角点)抽出・補完した後, 連結性解析により閉面領域抽出を行う。次に物体モデルと照合することによりまず閉面単位のスコアを計算した後, 隣接する閉面単位スコアを統合した統合スコアを求め, 認識対象物体の候補を抽出する。ここで関係モデルを適用することにより, 認識対象候補の位置関係の整合性を検証し, 矛盾の無い閉面領域の組み合わせを認識対象物体領域として得る。最後に物体同士の位置関係や人間動作との干渉関係を用いることにより, 未認識領域に対する解釈当て嵌めを行う。

以下, 個々の構成要素について説明を行う。

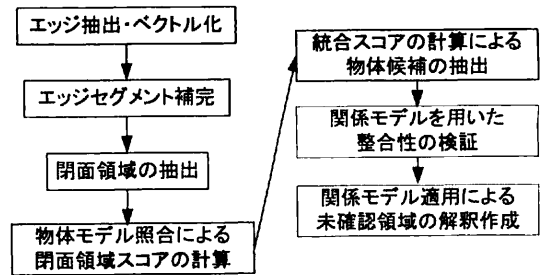


図1 処理の流れ

### 2.2. エッジ抽出

Canny Edge Detector[4]を用いることによりエッジ抽出を行う。

Canny Edge Detectorの特徴として

- ガウシアンフィルタの利用により, ノイズに対してロバスト
- 極大点追跡処理により, エッジの幅はほぼ1
- 二つの閾値を使っているため, 途切れにくい

一方, 問題点として

- 3つ以上の領域が重なる部分は一方のエッジに断線が生じる可能性がある

という点ももつ。この原因は, 等方的なガウシアンフィルタリングによりエッジ方向上でない領域の影響を受けてしまうためであり, この断線を補完する必要がある。

### 2.3. ラスタベクトル変換

入力される画像はラスタデータであり, エッジ情報もまたラスタデータである。ラスタデータは単なる点の集合であるため, 断線補完を行うためには直線情報などのベクトルデータが必要である。Canny Edge Detectorで抽出されるエッジはラスタ型データであるものの辺の部分においては連結性が高く, 隣接するエッジを線とみなすことで図2のように擬似的にベクトルデータに変換することが可能である。そこで, エッジ情報を端点・角点および分岐点を節点とし, 節点同士の結合状態を枝とするグラフによって表現することによりエッジ情報をベクトル化し, 断線補完に利用する。

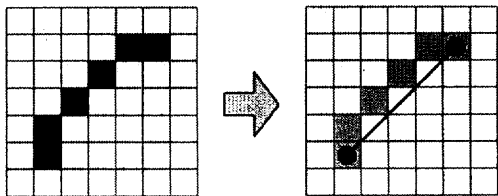


図 2 直線変換の例

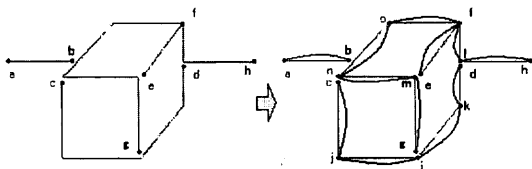


図 3 エッジのベクトル化

### 2.3.1. 細線化

次に細線化を行う。Canny Edge Detector で抽出されたエッジはほぼ太さ1ピクセルの線である。しかし、斜線については図4のように2種類出力される。エッジ追跡には8近傍調べるため、図4の(a)では経路が複数現れてしまい、追跡処理効率が大幅に低下する。そのため、経路を単一化するために(a)を(b)に変換する細線化を行う。

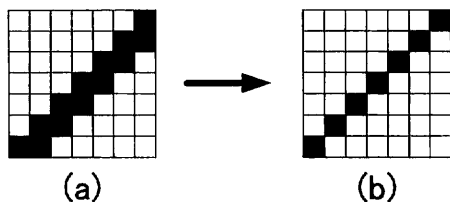


図 4 斜線のエッジパターン

細線化には図5に示す2種類のフィルタを用いた。ここで、黒はエッジ有り、白はエッジ無し、灰はエッジ有・無どちらかを示す。



図 5 細線化フィルタ

### 2.3.2. 端点・分岐点の抽出

エッジ周囲8方向に対し、エッジが一つの点(端点)および3つ以上の点(分岐点)があるポイントを節点として登録する(図6)。

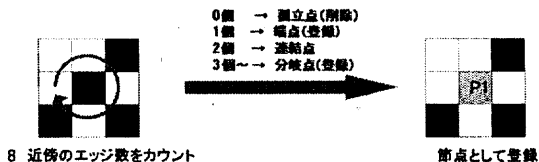


図 6 節点の登録

### 2.3.3. 連結点の抽出

節点に対しエッジをラベリングしながら追跡し、到達した節点に枝として登録する(図7)。

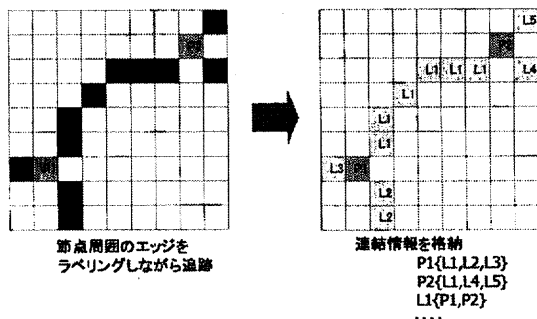


図 7 枝の登録

### 2.3.4. 隣接接点の扱い

エッジによっては図8のように隣接した節点が抽出される場合がある。この場合は隣接した節点の一つの節点として扱う。

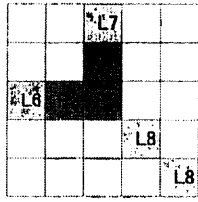


図 8 隣接接点の扱い

### 2.3.5. 環状エッジの対応

環状のエッジに対しては節点が存在しないため、グラフ化が出来ない。そこで図 17 のように仮の節点を設置し、節点と同じ処理を行うことで対応する。

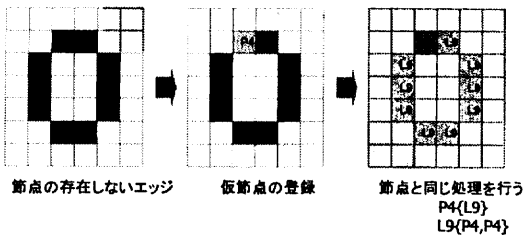


図 9 環状エッジの対応

### 2.3.6. 角点抽出

構築されたグラフはエッジの端点および分岐点を連結しただけのものであるため、コーナーを含むエッジの場合、枝は輪郭に沿ったものではなくなる、そこで、沿ったものにするべく節点(角点)を追加する。

まず、節点を結ぶ直線と中継点の距離を計算し、最も離れた中継点を記憶する(図 10 左図)。次に最長距離が閾値以上の場合、新しく節点を生成し、新たに枝を再構築する(図 10 右図)。

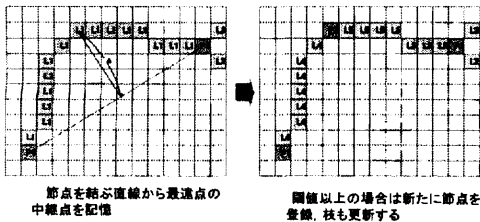


図 10 角点の抽出

## 2.4. エッジセグメント補完

断線は主に3つ以上の領域が重なる部分で生じる。しかし、一般的な物体において3つの領域が重なる部分は頂点部分であり、その頂点パターンに対応した補完を行うことで本来のエッジを復元することができる。まず、断線パターンを述べ、その後対応した補完方法を述べる。

### 2.4.1. 断線パターン

頂点パターンは図 11 に示すように3種類に分けられる[5]。領域の明度差によって、図 12 のようにそれぞれ3つのエッジの内、1つの線が断線する可能性がある。

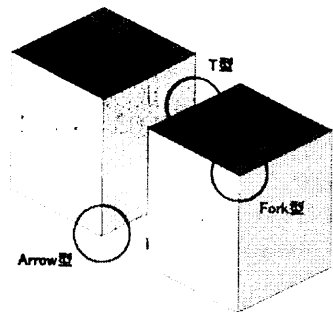


図 11 頂点パターン(T型 Arrow型 Fork型)

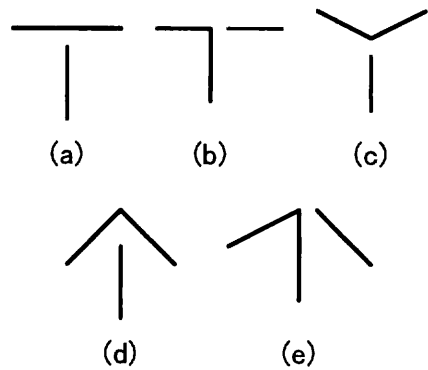


図 12 各頂点パターンによる断線パターン

### 2.4.2. 端線延長

断線パターンのうち、図 12 の(a)(c)および(d)は断線されたエッジの直線延長上にエッジが存在する。断線されたエッジは端線として抽出することができるので図 13 によりその端線の直線延長上にエッジを延長ことにより補完できる。

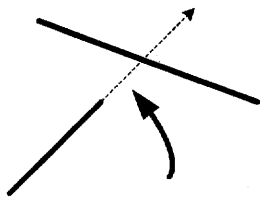


図 13 T型頂点の断線補完例

### 2.4.3. 端点—角点結合

図 12 の(b)(e)においては図 14 に示すように端線延長上にエッジが存在しない場合がある。しかし、(b)(e)の断線パターンでは、断線されていないエッジには角点が存在するため、図 15 に示すように端点—角点間を結合することにより補完できる。

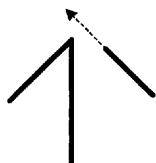


図 14 端線延長による断線補完の失敗例

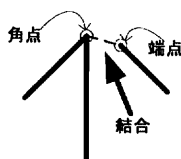


図 15 端点—角点結合例

### 2.5. 閉面領域の抽出

エッジで囲まれた領域を図 16 に示すように塗りつぶすことによりラベリングを行い、閉面領域を抽出する

F1	F1	L2	F2	F2	F2	F2	F2	F2
F1	F1	F1	L2	F2	F2	F2	F2	F2
F1	F1	F1	F1	P1	F2	F2	F2	F2
F1	F1	F1	L1	F3	L3	F2	F2	F2
F1	F1	F1	L1	F3	F3	L3	F2	F2
F1	L1	L1	F3	F3	F3	F3	L3	F2
L1	F3	F3	F3	F3	F3	F3	F3	L3

図 16 塗りつぶしによる面の抽出

### 2.6. 属性抽出

閉面領域単位での属性抽出を行う。抽出する属性は主に色・形状に関する情報であるが、辞書モデルとのスコア演算を行うため、それぞれの情報を数値化する必要がある。今回抽出する属性は黒・白・四角・棒状・横向き・上向きの6つの属性を抽出する。それぞれの属性は0~1の値になるように調整する。

#### 2.6.1. 色属性

色属性においては領域内のピクセルの平均明度を用いる。平均明度 $\bar{I}$ を用いて、白属性、および黒属性を式(1)(2)に示すように設定する。

$$A_{黒} = \frac{255 - \bar{I}}{255} \quad (1)$$

$$A_{白} = \frac{\bar{I}}{255} \quad (2)$$

#### 2.6.2. 形状属性

形状属性においては領域の頂点から代表点を4つ抽出し、その代表点の位置関係から属性値を算出する。代表点の抽出は以下の方法で行う

1. 領域の中心点から最遠点を P1 とする
2. P1 からの最遠点を P2 とする
3. P1-P2 を結ぶ直線からの最遠点を P3 とする
4. P1-P2-P3 を結ぶ三角形の辺からの最遠点を P4 とする

四角属性とは平行四辺形に近い形状を表した属性である。そこで、以下の式(3)に示すように設定する。

代表点で構成される辺の長さを $l_{1-4}$  ( $l_1$ の対線を $l_3$ ,

$l_2$ の対線を $l_4$ ) とすると

$$A_{四角} = \frac{(\min(l_1, l_3) / \max(l_1, l_3)) + (\min(l_2, l_4) / \max(l_2, l_4))}{2} \quad (3)$$

棒状属性とは細長い形状を現した属性である。そこで、以下の式(4)に示すように設定する。

$$A_{棒状} = 1 - \frac{\min(\max(l_1, l_3), \max(l_2, l_4))}{\max(\max(l_1, l_3), \max(l_2, l_4))} \quad (4)$$

#### 2.6.3. 方向属性

方向属性とは、閉面の垂線が3次元空間上でどの方向に向いているかを表す。形状属性抽出で使われた代表点を用いて算出する

上向き属性とは、閉面の垂線が3次的に上方向に

向いている状態を表す。

代表点において左側にある点で構成された線を  $S_L$ 、右側にある点で構成された線を  $S_R$  とし、直線式  $x=ay+b$  で表したときの  $a$  の値をそれぞれ  $a_L$ 、 $a_R$  とした場合、上向き属性の式(5)は以下のように設定する。

$$A_{\text{上向}} = |a_L| * 0.5 + |a_R| * 0.5 \quad (5)$$

(ただし、1を超える値は1とする)

また、横向き属性は以下の式(6)に設定する。

$$A_{\text{横向}} = 1 - A_{\text{上向}} \quad (6)$$

## 2.7. 認識に用いるモデル

### 2.7.1. 物体モデル

物体モデルとは、主に形状や色、および閉面領域の隣接関係などを記述したモデルである。図17にディスプレイの物体モデルを記述した例を示す。

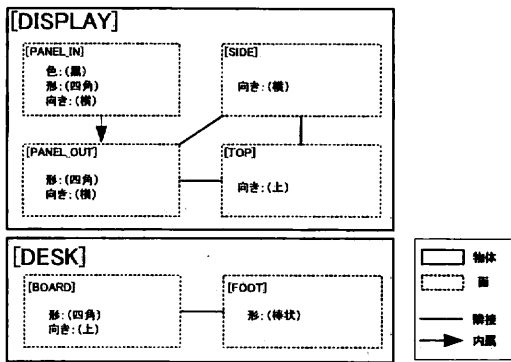


図17 物体モデルの具体例

### 2.7.2. 関係モデル

関係モデルとは、物体同士の関係を記述したモデルである。図18にそれぞれの関係を記述した例を示す。

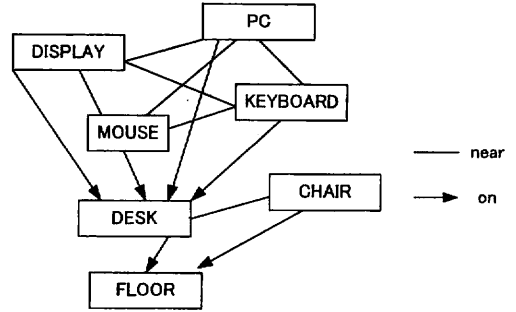


図18 物体の位置関係を記述した関係モデル

関係モデルにおいて、他の物体を利用するモデルであるため、認識済みの物体が必要となる。

## 2.8. モデル照合

モデル照合によるスコアの計算方法について述べる。

### 2.8.1. 閉面単位スコア

閉面単位スコアとは、1つの閉面に対する物体モデルの面単位での照合スコアである。この照合で重要なのは色や形状のみで認識可能な物体を抽出することである。ここで物体モデルにおいて多くの属性を持っている物体が、より正しく認識される可能性が高いと考えられる。よって閉面単位スコアの演算においては属性の数に比例するよう加算型のスコア計算式(7)とした

$$S_f = \sum_{i=1}^n A_{fi} M_{fi} \quad (7)$$

ここで  $n$  は属性の数、 $S_f$  は面  $f$  に対する閉領域スコア、 $A_{fi}$  は抽出した面  $f$  のそれぞれの属性に対応す

る属性値、 $M_{fi}$  は物体モデルに記述されたそれぞれの属性値である。加算型であるため、モデルに属性が記述されている数が多く、一致も多ければスコアは高くなる。物体モデルに属性が多く記述されていない物体に対しては閉領域スコアでの認識は控え、隣接情報を用いた統合スコアや関係モデルを用いた照合により認識を行う。

## 3. 実験

これまでに述べた方式に基づく認識システムを汎用パソコン(PC)上のソフトウェアとして実装し、実験を行った。

実験に使用したPCのCPUはPentium4 3.2GHz、OSはWindowsXP Professional、プログラミング言語は

Microsoft 社製 Visual C++.NET 2005 である。

### 3.1. 実験内容

今回は研究室内を撮影した画像を用いた。図 19 にその 1 例を示す。



図 19 原画像例

### 3.2. 実験結果

#### 3.2.1. エッジ抽出

図 20 に Canny Edge Detector を使用したエッジ抽出結果を示す。また、使用したパラメータは

$\sigma = 2.0$ ,  $T_{high} = 0.75$ ,  $T_{low} = 0.3$  である。

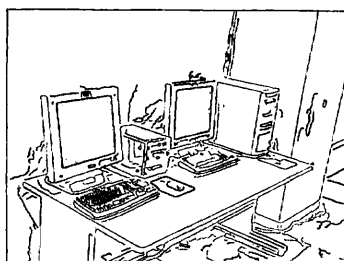


図 20 Canny Edge Detector の結果 ( $\sigma = 2.0$ ,  
 $T_{high} = 0.75$ ,  $T_{low} = 0.3$ )

#### 3.2.2. ベクトル化

ベクトル化した結果を図 21 に示す。

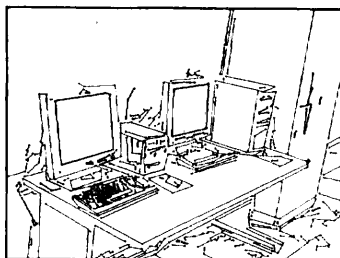


図 21 エッジ画像をベクトル化した結果

#### 3.2.3. エッジセグメント補完

エッジセグメント補完を行った結果を図 22 に示す。背景との明度がほとんど一致しているディスプレイの上端部、PC 筐体の上端部を除き、ほぼ忠実にセグメントが得られていることが確認できる。

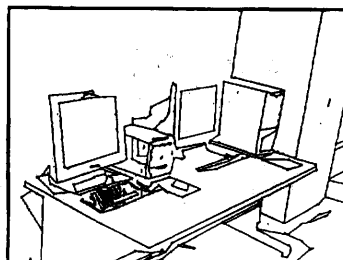


図 22 エッジセグメント補完の結果

#### 3.2.4. 閉面領域の抽出

閉面領域の抽出を行った結果を図 23 に示す。また、比較対象としてエッジセグメント補完を行わなかった場合の領域抽出の結果を図 24 に示す。エッジ補完処理が有効であることが確認できる。

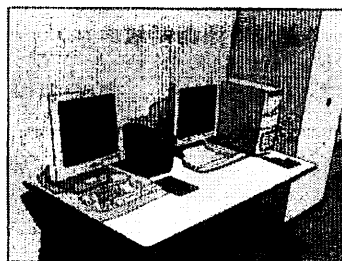


図 23 エッジセグメント補完を行った場合の  
閉領域抽出の結果

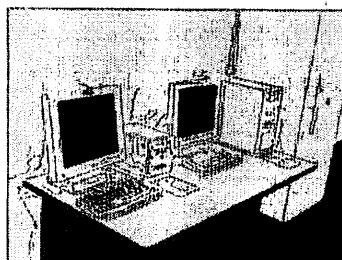


図 24 エッジセグメント補完を行わなかった  
場合の閉領域抽出の結果

### 3.2.5. 属性抽出・閉面領域スコアの計算

属性抽出を行った結果を表1に示す、また、閉面領域スコアの計算結果を表2に示す。対象とする面は図25のF1, F2, F3の3つの面とした

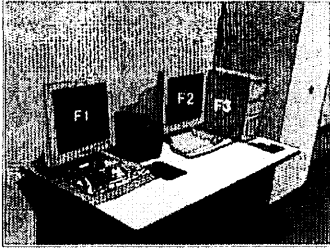


図 25 対象の面

表 1 属性抽出結果

面	黒	白	四角	棒状	上向	横向
F1	0.86	0.13	0.92	0.03	0.07	0.93
F2	0.85	0.14	0.96	0.07	0.04	0.96
F3	0.55	0.44	0.94	0.07	0.19	0.81

表 2 閉面領域スコア

F1	物体名[面]	閉面領域スコア	正答差
1位	DISPLAY[PANEL_IN]	2.71	0.00
2位	DISPLAY[PANEL_OUT]	1.98	0.73
3位	PC[SIDE]	1.85	0.86
正解面	DISPLAY[PANEL_IN]	2.71	-

F2	物体名[面]	閉面領域スコア	正答差
1位	DISPLAY[PANEL_IN]	2.77	0.00
2位	DISPLAY[PANEL_OUT]	2.06	0.71
3位	PC[SIDE]	1.92	0.85
正解面	DISPLAY[PANEL_IN]	2.77	-

F3	物体名[面]	閉面領域スコア	正答差
1位	DISPLAY[PANEL_IN]	2.30	-0.55
2位	DISPLAY[PANEL_OUT]	2.19	-0.44
3位	PC[SIDE]	1.75	0.00
正解面	PC[SIDE]	1.75	-

### 4. 考察

図23, 24の比較に示すように、エッジセグメント補完処理が閉面領域検出において有効であることは確認できたと考える。一方、図22に示すように、セグメント補完に失敗している部分も若干存在した。失

敗している箇所はエッジ抽出の特性による断線ではなく、背景との明度一致や影などによって部分的にエッジが抽出されないことに因る。この対処は今後の課題である。

また、閉面領域スコアは、ほぼ妥当に算出された。

(7)式に示すようにスコアの計算式が加算型だったため、領域別では物体モデルにおいて属性数の多いDISPLAY[PANEL\_IN]やDISPLAY[PANEL\_OUT]の面がほとんどの領域で上位を占める結果となった。しかし、閉面領域スコアは対応する領域が何の物体の領域かを判断するためだけではなく、認識しやすい物体から認識し、認識結果を必要とする関係モデルと適用させるためのものでもあるため、関係モデルとの整合性をとることにより問題は解決するものと考えられる。

### 5. まとめ

認識対象個々の形状・属性情報を記述した物体モデルに加え、物体相互の位置関係、および人間動作との干渉関係を記述した関係モデルを適用することにより、複数物体の重なり・遮蔽や影などの影響によるコントラスト低下が生じた部分の特徴抽出が失敗し不完全な記述結果しか得られない可能性が大きい複雑シーンにおいても安定な認識を実現する手法を提案した。

今後は、関係モデルの適用による整合性チェックの部分を実装し、有効性評価実験を行っていく。

謝辞 本研究の一部は、日本学術振興会・科学研究費補助金(「分散視覚認識エージェントの統合に基づく親和的情報空間の展開」, 課題番号:15500115)の援助による。

### 文献

- [1] 出口光一郎, “画像認識論講義,” 昭晃堂, Jun. 2002.
- [2] 柳井啓司, 出口光一郎, “物体間の支持関係を利用した室内画像の認識,” 信学論 D-II, Vol. J-48-D-II, pp.1741-1752, Aug. 2001.
- [3] 中島隆行, 白井良明, Michael Hild, “人間の動作の解析に基づく物体の認識,” 信学研資, PRMU2003-208, pp.59-64, Jan. 2004.
- [4] J. Canny, “A Computational Approach to Edge Detection”, IEEE Transaction on Pattern Analysis and Machine Intelligence. Vol. PAMI-8, No.6, pp.679-697. 1986.
- [5] 杉原厚吉, “不可能物体の数理,” 森北出版, 1993.