

# 手書きシルエット画像のトップダウン認識

高田英司 山村毅  
愛知県立大学大学院情報科学研究科

あらまし: 手書きのシルエット画像を認識する手法を提案する。手書きのような書く人によって形が大きく変化する対象から、認識するうえで必要な特徴を得ることは容易ではない。そこで、本研究では、知識に合うようにトップダウンに対象の構造を理解していく。まずはじめに、シルエットを意味があると思われるパーツに分割し、各パーツの幾何学的特徴量を計算する。そして、個々のパーツの形状特徴とパーツ間の位置関係などをグラフで記述し、モデルとの間でグラフマッチングを行う。このとき、モデルの構造に合わせてパーツの対応をとっていくことで、構造に多少の違いがあってもマッチングをとれるようにする。何枚かの手書きのシルエット画像を対象に実験を行い、本手法の有効性を示す。

## Top-Down Recognition of Hand-Drawn Silhouettes

Eiji Takada Tsuyoshi Yamamura  
Graduate School of Information Science and Technology, Aichi Prefectural University

**Abstract:** We propose a method for recognizing hand-drawn silhouettes. In such figures, it is not easy to obtain the features which are required for object recognition, because the object shapes are variable depending on the drawers, even if they drew the same object. One way to solve this problem is to interpret the object in a top-down manner so that it fits into a model. In this manner, the silhouette is first divided into meaningful parts, and geometric properties are calculated. Then those properties and relations among the parts are described using a graph, and the matching is performed between the graph and a model considering subtle differences and errors. Experiments shows effectiveness of our method.

### 1 はじめに

さまざまな形を認識することは、パターン認識における基本的問題のひとつである。工業製品のような形がある程度決まった人工物のようなものなら、従来のテンプレートマッチングの手法で認識できるが、姿勢の変化によって形状が変化する関節物体や、書く人によって形が変化する手書きの図形のような対象では、そのような手法ではうまくいかない。例えば、図1の2つのシルエット

を見比べた時、形状そのものには、はっきりとした違いがある。しかし、我々人間は2つのシルエットが同じ対象(キリン)を表したものだとして理解できる。これは我々が、対象に関する知識を持っているので、対象の姿勢が変わったり、部分的な形状が多少変わっても、知識を柔軟に対象に当てはめて認識できるからだと考えられる。

対象のシルエット、つまり外輪郭情報のみを用いて、対象を認識する手法が提案されている。Gdalyahu

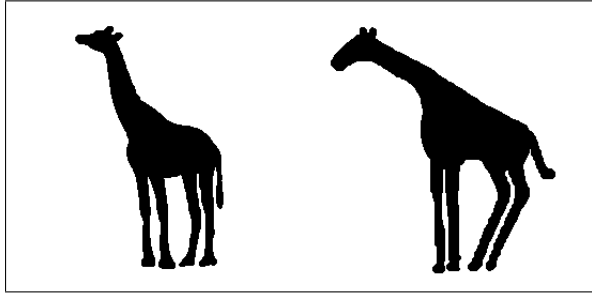


図 1: 形状の認識

ら [1] は画像データベースから、目的の形状を検索するための、シルエット間の形状類似度を定義している。これは、輪郭線を直線近似し、その線分群を挿入、削除、代替といった操作によって、一方の輪郭線を他方に置き換えるコストを非類似度とするものである。しかし、構造そのものが大きく違った図形は区別できるが、構造が似ている図形は類似度あまり差がないという問題がある。Tyngら [2] はシルエットの中心軸を求めて、シルエットを分割し、その接続関係をツリー構造で表現する。そして、ツリー・マッチングでは、図形の伸縮、オクルージョンに対応するため、カットとマージという操作を定義して、柔軟なマッチングを行っている。しかし、対象に関する知識を用いていないので、図形の部分的な伸縮などの単純な変化には対応できるが、関節物体の姿勢の変化のような、複雑な変化には対応できない。阿部 [5] らは、入力図形を意味のある要素に分解し、それらから部品を構成、部品間の関係をグラフの形で記述して、対象物のクラスを表すモデル記述との間でグラフマッチングを行い、類似度を求めている。しかし、モデルの構造を手作業で細かく記述する必要があるため、モデルの構築に手間がかかるほか、マッチングを行いながら部品の合成を行うわけではないので、必ずしもモデルに合った部品を構築できるとは限らない。

本研究では、手書きのシルエット画像の認識を行う。まず、対象とするシルエットを意味があると思われるパーツに分割し、各パーツの特徴抽出をする。そして、実際のシルエット画像を分割して構築したモデルとマッチングを行う。ここで、マッチングを行いながら、モデルの構造に合うように対象のパーツを再分割することによって、対象の解釈を柔軟に変化させて、認識を行う。

## 2 提案手法

### 2.1 処理の流れ

本研究では対象とするシルエットがいくつかのパーツから構成されているものとし、そのパーツの関係を構造的に記述することによって、シルエット全体を表現する。図 2 に処理の流れを示す。まずはじめに、入力画像から輪郭線を抽出し、直線近似する。次に、直線近似された輪郭線のある評価関数を用いて、パーツに分割する。分割された各パーツの特徴量を抽出し、対象シルエットの構造を記述する。そして、それとあらかじめ用意された複数のモデルとマッチングを行う。ここで、対象のパーツのすべてがうまく抽出されていない場合にも対応するため、モデルと対象の間でパーツの構造に違いがあってもパーツの対応がとれるような柔軟なマッチングを行う。具体的には、隣接するパーツ同士を併合するマージ操作と余分なパーツを削除するカット操作によって、モデルと対象のパーツの対応を 1 対 1 にとることができるようにする。そして、モデルの構造に合うように対象のパーツを分割し直して、マッチングコストを計算し、コストが小さい(類似度が高い)モデルを認識結果とする。

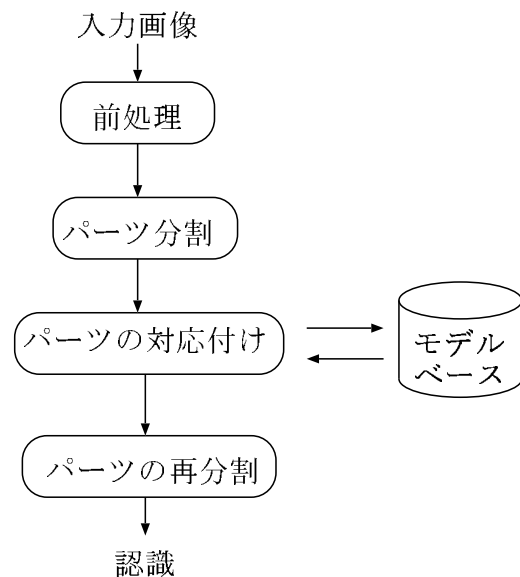


図 2: 処理のながれ

## 2.2 モデルの記述

モデルはどんなシルエットでも簡単に構築できるように、実際の画像からシルエットを抽出したもので、図3のように手作業でパーツに分割してある。そして、3で説明するように、特徴量を各パーツについて計算し、パーツの関係をグラフで記述する。

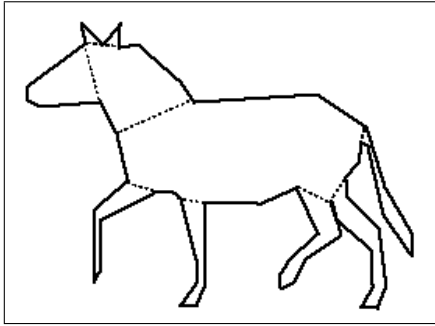


図 3: モデル画像

## 3 シルエットの分割と特徴抽出

### 3.1 パーツ分割

シルエットの構造を記述するために、意味のあるパーツに分割する。分割の手順としては、シルエットから得られる輪郭線を直線近似し、凹頂点を検出する。すべての凹頂点のペアについて(1)のような評価関数を計算する。

$$E = K(\cos \theta_i + \cos \theta_j) - d_{ij} \quad (1)$$

ここで、 $d_{ij}$ は凹頂点*i*と*j*との間の距離、 $\theta_i, \theta_j$ はそれぞれ凹頂点を結ぶ直線と凹頂점에隣接する線分のなす角度、 $K$ は定数である(図4)。分割線

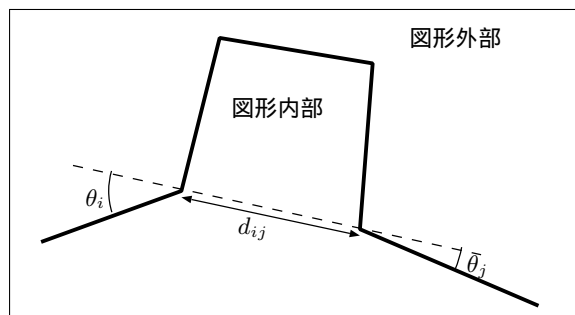


図 4: 分割関数

の長さが短く、かつ分割線と頂点に隣接する線分のなす角度が小さいほど、評価関数の値は大きくなる。

評価関数があるしきい値を越えたペアに分割線を生成し、二つのパーツに分割していく。

### 3.2 パーツの特徴量

各パーツの特徴量として、書く人による形状の差異を吸収するため、大まかな形状特徴を用いる。具体的には、以下の6つの特徴量を計算する。(図3.2参照)

- 面積  $S$   
領域の大きさを示し、画素数を用いる。
- 周囲長  $L$   
領域の輪郭線の長さ
- 円形度  $\frac{4\pi S}{L^2}$   
領域の円らしさ。真円のとき1となる。
- 重心  $(x, y)$   
領域の物体内の位置を示す。画像内の座標系における、その領域の重心座標。
- 慣性主軸  $\theta$   
領域の2次モーメントを求め、長軸とx軸とのなす角度で表す。
- 細長さ  $\frac{l}{w}$   
領域の外接長方形の長軸と短軸の長さの比。

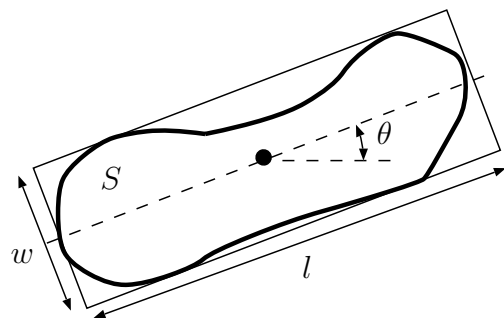


図 5: 特徴量

### 3.3 グラフ表現

パーツ間の大小関係や位置関係をうまく記述するためにグラフ構造を用いる。グラフのノードには個々のパーツの特徴量を記述し、アークにはパーツ間の関係を記述する。ここで、グラフのノードおよびアークには、画像の拡大・縮小、回転に対して不変な特徴量を用いる。ノード  $i$  の特徴量は  $\vec{N}_i = \{ \text{円形度, 細長さ} \}$ 、ノード  $i$  とノード  $j$  のアークの特徴量は  $\vec{A}_{ij} = \{ \text{面積比, 慣性主軸のなす角度, 重心間の距離, 重心間の角度} \}$  とする。

## 4 パーツの対応付け

モデルと対象との間でグラフのノード（パーツ）の対応をとる。1対1だけでなく、多対1，多対多の対応をとることによって、モデルの構造に合わせた柔軟なマッチングを行えるようにする。

### 4.1 初期マッチング候補

モデルと対象のグラフのノード数を  $N, M$  とすると、初期のマッチング候補は  $N \times M$  通り存在する。ここで、対象とする図形の特徴として、胴体部分を中心として、いくつかのパーツが周りに接していることがあげられる。これらの特徴から、各パーツについて、(接続しているパーツの数)  $\times$  (全体に占める面積の割合) を計算し、値が最大のパーツを胴体部分と推定する。そして、その胴体部分同士をグラフマッチングの初期マッチング候補とする。

### 4.2 マッチングの探索

モデルと対象のグラフのノードの対応を1対1だけでなく、多対1，多対多の対応をとれるようにするため、ノードのマージ（併合）操作を行う。ノードのマージが行われたとき、パーツの特徴量を再計算する。そして、次のようなコストをマージのペナルティとして与える。

$$cost_M = \alpha \times (\cos \theta'_i + \cos \theta'_j) \quad (2)$$

ここで、 $\theta'_i, \theta'_j$  は図6に示す角度、 $\alpha$  は定数である。初期マッチングから開始して、これらのノードのマージも考慮に入れて、マッチングを全探索によって求める。対応のとれないノードに関しては、余分なパーツと見なして、カット操作を行い、次の

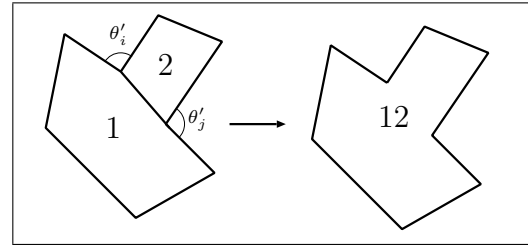


図 6: パーツのマージ（併合）

ようなコストをペナルティとして与える。

$$cost_C = \beta \times \frac{s_i}{S} \quad (3)$$

ここで、 $s_i, S$  はそれぞれ、パーツ  $i$  の面積とシルエット全体の面積、 $\beta$  は定数である。

### 4.3 マッチングコスト

ノードの対応がとれたら、ノードとアークについて、以下のようなマッチングコストを計算する。ノード  $i$  とノード  $j$  の特徴量ベクトルをそれぞれ、 $\vec{x}_i = (x_1^i, \dots, x_n^i), \vec{y}_j = (y_1^j, \dots, y_n^j)$  とするとき、それらのマッチングコストは、

$$cost(\vec{x}_i, \vec{y}_j) = \sum_{k=1}^n g_k \frac{p_k(x_k^i, y_k^j)}{s_k} \quad (4)$$

と重み付き線形和として定義する。ここで重み係数  $g_k$  は、人間が各特徴量を主観的に重視する程度を表現する定数である。また、 $s_k$  は、各重み係数の値を同程度の大きさにするための補正項であり、 $k$  番目の特徴量の標準偏差とする。

アークについても同様に、特徴量ベクトルの距離を計算し、マッチングコストに加える。

### 4.4 パーツの再分割

モデルの複数のパーツと対象のパーツが対応しているとき、モデルのパーツの分割情報を用いて、対応する対象のパーツを分割する。モデルのパーツの輪郭線と対象のパーツの輪郭線の対応点が求まれば、モデルの分割点と対応する対象の輪郭点を分割点として、対象のパーツを分割できる。2つの輪郭線の対応点を求めるために、Gdalyahu ら [1] のアルゴリズムを用いる。

## 4.5 Gdalyahu らのアルゴリズム

直線近似された2つの図形の輪郭線を次のように線分群で表す。

$$A = \{a_0, a_1, a_2, \dots, a_{N-1}\}$$

$$B = \{b_0, b_1, b_2, \dots, b_{M-1}\}$$

ここで、 $a_i, b_j$  は図形 A 及び B の適当な位置から時計回りに数えた  $i$  および  $j$  番目の線分を表し、パラメータとして、長さ、絶対角度の2つの要素を持つとする。そして、エディットコストを最小にする線分の対応を求める。

### 4.5.1 線分間の類似度

対応付けの良否を判定するために、2本の線分間の類似度を導入する。いま、2本の線分を  $a_i, b_j$ 、線分  $a_i$  の長さを  $l$ 、角度を  $\theta$ 、線分  $b_j$  の長さを  $l'$ 、角度を  $\theta'$  とする。また、対応付けの開始点を  $\{a_0, b_0\}$  それぞれの長さ、角度を  $(l_0, \theta_0), (l'_0, \theta'_0)$  とする。このとき、長さに対する類似度を次の式で計算する。

$$S_l = \cos 2\delta = \frac{4cc_0 + (c^2 - 1)(c_0^2 - 1)}{(c^2 + 1)(c_0^2 + 1)} \quad (5)$$

ここで、 $c = \frac{l}{l'}$ 、 $c_0 = \frac{l_0}{l'_0}$  であり、それぞれ、 $a_i$  と  $b_j$  の長さの比、 $a_0$  と  $b_0$  の長さの比である。長さの比が近いほど  $S_l$  は大きな値をとる。

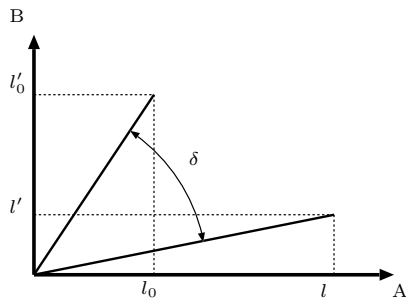


図 7: 長さに関する類似度

また、角度に関する類似度を次の式で計算する。

$$S_\theta = \cos[(\theta \rightarrow \theta') - (\theta_0 \rightarrow \theta'_0)] \quad (6)$$

ここで、 $(\theta \rightarrow \theta'), (\theta_0 \rightarrow \theta'_0)$  は、それぞれ、 $\theta$  から  $\theta'$  に向かう角度、 $\theta_0$  から  $\theta'_0$  に向かう角度である。

最終的な線分間の類似度  $S$  は次のように計算する。 $w$  は重み係数である。

$$S = wS_l + S_\theta \quad (7)$$

### 4.5.2 動的計画法によるエディットコストの最小化

1対1の他に多対1、1対多、多対多などの対応付けを許して、2つの線分群の線分を過不足なく対応づける。対応関係にある線分の各組について類似度を計算し、それらの合計をとったエディットコストを最小にする対応付けを動的計画法を用いて求める。

2つの輪郭線の線分群を  $A = \{a_0, a_1, \dots, a_n\}$ 、 $B = \{b_0, b_1, \dots, b_m\}$  とする。 $a_k$  と  $b_l$  までのエディットコスト  $R[k, l]$  を式 (6) のように求める。 $\overline{\alpha k}$  は連続する線分  $\{a_{\alpha+1}, a_{\alpha+2}, \dots, a_k\}$  の始点と終点を直線で結んだものを表す。 $K = 4$  の場合、図 8 に示す6つの斜線部分の中からエディットコストが最小となる対応を求める。ここで、式 (9) に示す  $K$  は1つの直線に  $K - 1$  本の直線までを対応させることができるという範囲を表す。

$$R[k, l] = \min_{\alpha, \beta \in \Omega} \{R[\alpha, \beta] - S(\overline{\alpha k}, \overline{\beta l})\} \quad (8)$$

$$\Omega = \{\alpha, \beta \mid 0 < \alpha < k, 0 < \beta < l, (k - \alpha) + (l - \beta) \leq K\} \quad (9)$$

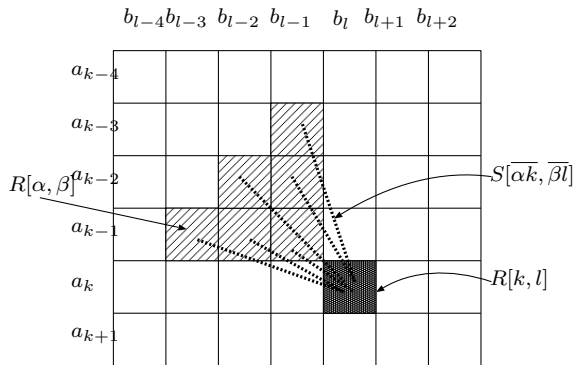
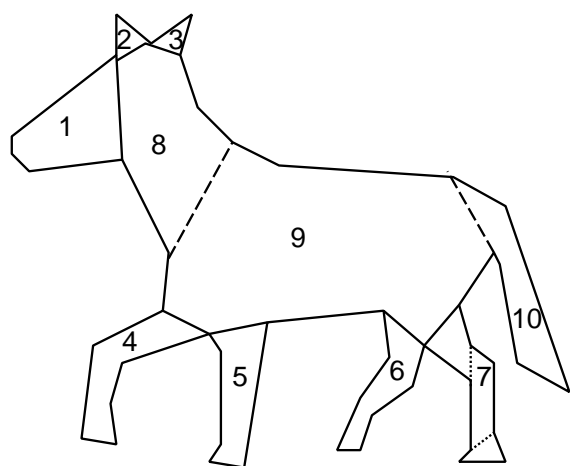


図 8:  $K = 4$  の場合の  $R[k, l]$  の算出

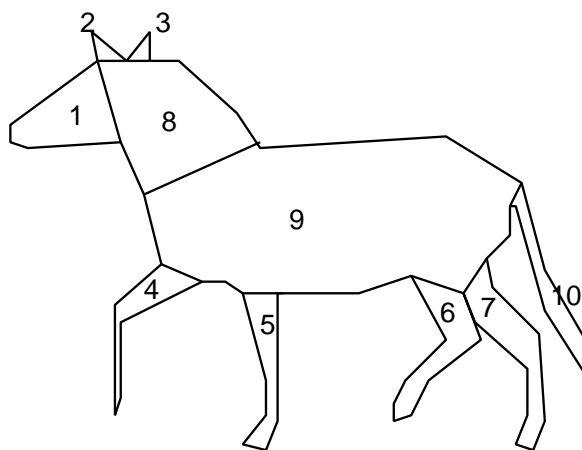
## 5 実験結果

実際に手書きで書いた動物の画像を入力画像として、モデルとマッチングを行い、パーツの対応付けと、マッチングコストを求める実験を行った。

- ..... マージされた分割線
- カットされた分割線
- モデルと一致した分割線
- 再分割でつくられた分割線

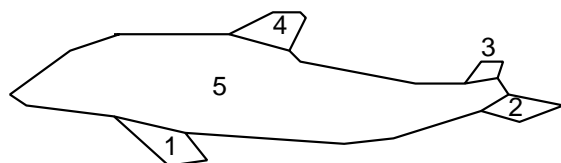
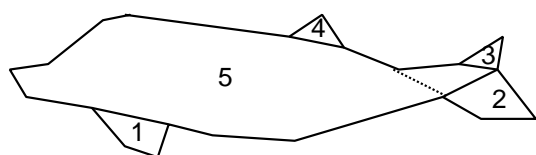


入力画像

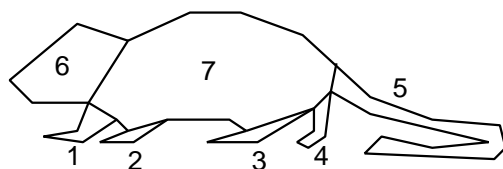
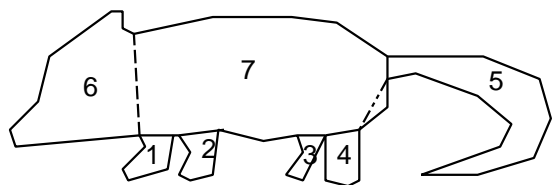


モデル

(a) マッチングコスト : 65



(b) マッチングコスト : 25



(c) マッチングコスト : 54

図 9: モデルと対象図形のパーツの対応とマッチングコスト

実験の結果を図9に示す。左の画像が入力画像で、右の画像がモデルである。最終的に対応のとれたパーツを同じ番号で図中に示す。また、パーツの分割線を4種類の線で図形内に示す。ここで、マージされた分割線とカットされた分割線とは、はじめのパーツ分割で生成されたが、モデルとのマッチングでマージまたはカットされたパーツの分割線である。モデルと一致した分割線とは、はじめのパーツ分割で生成され、モデルとのマッチングでモデルと対応のとれたパーツの分割線である。再分割でつくられた分割線とは、はじめのパーツ分割では生成されなかったが、パーツの再分割の過程でつくられたパーツの分割線である。

(a)では、はじめのパーツ分割で、対象図形の後ろ足が過分割されている。また、モデルのように、しっぽや首のパーツが分割されていない。しかし、モデルとのマッチングで、後ろ足の分割されたパーツがマージされている。そして、パーツの再分割の過程で、1つのパーツだった首、胴体、しっぽ(パーツ番号8,9,10)のパーツが、モデルのパーツの輪郭線の対応をとることによって、モデルに近い形で、3つのパーツに分割されている。

(b)でも(a)同様に、パーツ分割で過分割が生じたが、モデルとのマッチングでマージされている。マッチングコストは、パーツの数が少ないため、小さな値になっている。

(c)では、パーツ分割で後ろ足と、しっぽの間に小さなパーツができたが、モデルとのマッチングで、カットされた。これは、パーツの面積が小さいため、マージコストよりカットコストの方が小さくなったからである。また、頭(パーツ番号6)のパーツが再分割の過程で生成されている。

3つの実験結果を見ると、モデルと対象図形のパーツの対応がとれていることがわかる。

## 6 おわりに

本研究では、手書きのシルエット画像を認識する手法を提案した。対象の姿勢の変化や部分的な形状の変化に対応するため、シルエットを意味があると思われるパーツに分割して、パーツの大まかな形状特徴や、パーツ間の位置関係などの特徴量を計算し、シルエット全体を構造的に記述した。

モデルは、どんな対象でも簡単に構築できるように、実際の画像から、シルエットを抽出して、パーツと思われる部分に分割した。モデルと対象のマッチングでは、パーツのマージ(併合)操作を行うことにより、構造に多少の違いがあってもパーツの対応がとれるようにし、対応がとれたら、モデルのパーツの分割情報に合わせて、対象のパーツを分割する。こうすることで、モデル(知識)に合わせて、対象を解釈するという認識ができる。

今後の課題としては、はじめのパーツ分割で、モデルと対応するパーツがうまく抽出されないと、パーツの再分割によって生成されるパーツが、モデルの構造に合っていないことがあるので、はじめのパーツの分割の精度を上げる必要がある。また、今回、モデルにはパーツの分割情報しか与えていないが、例えば、関節物体の関節の駆動領域をある程度制限するような、関節モデルをモデルに導入したり、その図形を認識する上で重要となるようなパーツには、重要度を記述するといったことが考えられる。また、パーツの対応付けの探索は、全探索で行ったが、予測コストを導入してA\*アルゴリズムなどで、効率的に探索を行えば、探索時間の短縮が期待できる。

## 参考文献

- [1] Yoram Gdalyahu and Daphna Weinshall, "Flexible Syntactic Matching of Curves and Its Application to Automatic Hierarchical Classification of Silhouettes," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, pp. 1185-1190, 2000
- [2] Tyng-Luh Liu and Davi Geiger, "Aproximate Tree Matching and Shape Similarity," Proc. Seventh IEEE International Conference on Computer Vision, pp. 456-462, 1999
- [3] Longin Jan Latecki and Rolf Lakämper, "Shape Similarity Measure Based on Correspondence of Visual Parts," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, pp. 1185-1190, 2000

- [4] Serge Belongie, Jitendra Malik, and Jan Puzicha, "Shape Matching and Object Recognition Using Shape Contexts," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 24, NO. 24, pp. 509-522, 2002
- [5] 久嶋毅, 阿部圭一, "2次元図形の構造的記述による認識," 信学技報 PRMU96-18, pp. 45-50, 1996
- [6] 山村毅, 滝沢和明, 大西昇, 杉江昇, "主観的輪郭に示唆を得た、重なり合った物体の内部輪郭の生成," 電子情報通信学会論文誌 D-II, Vol.J78-D-II, No.5, pp. 880-883, 1995
- [7] 藤富知行, 大西昇, 杉江昇, "可変形状物体の姿勢推定についての検討," 信学技報 PRU93-115, pp. 17-24, 1994
- [8] 遊佐 博幸、兵庫 明、関根 慶太郎, "手の大きさ、方向の変化に対応可能な手形状認識方法の一提案," 信学技報 HC94-93, pp. 47-54, 1995
- [9] 亀田 能成、美濃 導彦、池田 克夫, "シルエット画像からの関節物体の姿勢推定法," 電子情報通信学会論文誌 D-II, Vol.J79-D-II, No. 1, pp. 26-35, 1996