

重み付き最小共通スーパーグラフを用いたシルエット画像認識

宮崎 智[†] 大町 真一郎[†](正会員) 阿曾 弘 具[†][†] 東北大学大学院工学研究科

〈あらまし〉 画像中の物体を認識することはコンピュータビジョンの分野において主要な問題である。本論文では穴のないシルエット画像を対象とし、画像からグラフを抽出してグラフの類似性により認識する手法を提案する。シルエット画像をグラフで表現して認識する手法はいくつか提案されているが、従来手法の多くは学習データの中からカテゴリごとにプロトタイプを選び、未知画像とプロトタイプとの類似度を計算することで認識を行なっている。提案手法はプロトタイプを学習データの中から選ぶのではなく、カテゴリごとにすべての学習データを用いて共通的な構造をもつグラフ（代表グラフ）を求め、未知画像をこの代表グラフとの類似性により認識する。実際のシルエット画像を用いた認識実験を行い、提案手法の有効性を示す。

キーワード：中心軸変換、シルエット画像、重み付き最小共通スーパーグラフ

〈Summary〉 It is desired to recognize objects in images correctly. Recognition using structural feature has been studied. A single structure extracted by the medial axis transform from a digital silhouette image is not always an essential feature as a prototype that represents silhouettes in a category, because of noise and distortion. In this paper, we propose a method for recognizing silhouette images by obtaining an essential structure from the images of a category. The essential structure is defined as a weighted minimum common supergraph of graphs which are extracted from silhouette images. To show the validity of the proposed method, experiments are carried out for categorizing silhouette images.

Key words: medial axis transform, silhouette image, weighted minimum common supergraph

1. はじめに

画像中の物体を認識することはコンピュータビジョンの分野において主要な問題であるとともに、安全・安心・快適な社会を実現するキーワードである。画像中の物体をシルエット化したものは物体の基本レベルでのカテゴリ分類に有用であることが知られており¹⁾、シルエット画像を認識する研究が行われている。物体のシルエット画像をカテゴリ分類することは物体認識の第一段階と考えられる。シルエット画像を正しくカテゴリ分類することができれば、さらに詳細で的確な認識を行うことが可能になる。

物体のシルエットの形状を用いた認識手法として、Be-

longie らは shape-context マッチング法²⁾を提案している。これは輪郭上のある点の特徴を他の輪郭上の点との位置関係により得られるヒストグラムで定め、輪郭どうしのマッチングを行うものである。これらの手法は若干の幾何学的変形やオクルージョンに対しては頑健だが、非剛体などの物体で生じる著しいパーツの曲げや伸びなどの構造的な変形に対応することは困難である。

一方、シルエット画像の形状を記述する手法の1つに、中心軸を用いる方法がある³⁾。中心軸とは図形の輪郭線までの距離が等しい内部点の集合である。中心軸変換³⁾、⁴⁾により物体のシルエットから中心軸が得られると、抽出された中心軸を用いて、物体の構造の特徴をグラフにより記述することができる。そのグラフを用いることにより、回転やオクルージョン、スケール変化などの形状の幾何学的な変形に加え、非剛体の物体のシルエットに見られるパーツの折れ曲がりなどの変形にも頑健な形状特徴の記述が可能である。その一方で輪郭線上の細か

“Silhouette Image Recognition with Weighted Minimum Common Supergraph” by Tomo MIYAZAKI, Shinichiro OMACHI, (Member), and Hirotomo ASO (Graduate School of Engineering, Tohoku University).

な凹凸に対応するノイズに弱い等の難点もあるが、中心軸を用いてシルエット画像を認識する手法が研究されている^{5),6),8)~11)}。Siddiqi らは中心軸の各点にショックラベル (shock-label) と呼ばれるラベルを割り当て、それを用いてシルエット画像の構造をラベル付き木で表わし、グラフの局所的な違いによりシルエット画像を分類した⁵⁾。シルエット画像の構造を表わすグラフが得られると、グラフの比較によりシルエット画像の認識を行うことができる。グラフを比較する手法の 1 つが、編集距離を用いて類似度を計算する手法である^{6),8)}。編集距離は削除、挿入、置換といったグラフを編集する操作にコストを割り当てた場合の一連の編集操作に必要なコストの総和であり、2 つのグラフを一致させる編集系列のコストを動的計画法により得てその大きさにより認識を行なうことができる。各編集操作に割り当てられるコストとしては、輪郭線の長さ⁹⁾、パーツの輪郭線と中心軸各点の弾性マッチングコスト¹⁰⁾、中心軸に対応するシルエット画像の面積¹¹⁾等が提案されている。

これらの従来手法ではすべて、カテゴリを代表するプロトタイプを学習データから選び、未知のシルエット画像が与えられたときは各カテゴリのプロトタイプとの類似性を測り認識を行なう。しかし、同一カテゴリでも様々に変形した画像が含まれる場合、プロトタイプが必ずしも未知画像と似ている形状であるとは限らないため、高精度な認識は困難であると考えられる。カテゴリを適切に表わすグラフをいかに得るかは重要な問題であり、カテゴリの構造をよく表わすグラフを用いることによってシルエット画像をより適切に認識することができる¹²⁾。

本論文では穴のないシルエット画像を対象とし、プロトタイプを学習データの中から選ぶのではなく、カテゴリに含まれる全ての学習データを用いて共通的な構造をもつグラフ (代表グラフ) を求め、未知入力とこの代表グラフとの類似性を測ることでシルエット画像を認識する手法を提案する。代表グラフを求めるために、Bunke らによって提案されている、複数のグラフの重み付き最小共通スーパーグラフを求める手法¹³⁾を用いる。文献¹³⁾では小規模な人工グラフを対象としていたが、実際のシルエット画像から得られるグラフは大規模であり、直接的な適用は困難であった。そこで、シルエット画像の中でグラフの各エッジに対応する部分の面積 (「パーツ面積」と呼ぶ) を求め、このパーツ面積をエッジの重要度とみなし、グラフの不要なノードやエッジを削除することで、より認識に適したグラフを構築する。さらに、複数のグラフが似ている度合を表わすエントロピーとこのパーツ面積を組み合わせることで、シルエット画像を認識するアルゴリズムを与える。そして、実際のシルエット画像

を用いた認識実験を行い、提案手法の有効性を示す。

2. 重み付き最小共通スーパーグラフ¹³⁾

本節では複数のグラフを代表するグラフを得る手法¹³⁾について述べる。重み付きグラフ g を

$$g = (\mathbf{v}, \mathbf{e}, \lambda, \varepsilon) \quad (1)$$

で表わす。 \mathbf{v} はノードの集合、 \mathbf{e} はエッジの集合である。 λ, ε は重み関数で、 $\lambda(v_i)$ はノード v_i の重み、 $\varepsilon(e_j)$ はエッジ e_j の重みである。グラフの集合 $\mathbf{G} = \{g_1, \dots, g_n\}$ が与えられたとき、これらのグラフの共通スーパーグラフとは、 \mathbf{G} に含まれるすべてのグラフを部分グラフとしてもつグラフのことである。また、そのうちノード数が最小のものを最小共通スーパーグラフ (minimum common supergraph; MCS) と呼ぶ。さらに、最小共通スーパーグラフの各ノードとエッジの重みとして、 \mathbf{G} のグラフの対応するノードの数、対応するエッジの数を設定したものを重み付き最小共通スーパーグラフ (weighted MCS; WMCS) と呼ぶ。

グラフ G とその部分グラフ g が与えられたとき、その差分グラフ $G - g$ とは、 G から g のノードとそれに隣接するエッジを取り除くことによって得られるグラフである。 $G - g$ と g を連結していたエッジの集合を $emb(G, g)$ と記す。WMCS の構築アルゴリズムを記述するために、グラフ g_1 と g_2 の $MCS(g_1, g_2)$ を次のように定める。

$$MCS(g_1, g_2) = mcs(g_1, g_2)$$

$$\cup_{emb(mcs(g_1, g_2), g_1)}(g_1 - mcs(g_1, g_2))$$

$$\cup_{emb(mcs(g_1, g_2), g_2)}(g_2 - mcs(g_1, g_2)). \quad (2)$$

ここで、 $mcs(g_1, g_2)$ は g_1 と g_2 に共通に含まれる部分グラフのうちノード数が最大のグラフであり、backtracking アルゴリズム¹⁴⁾により求められる。また、 \cup_{emb} は emb により 2 つのグラフを連結する操作である。 $MCS(g_1, g_2)$ は、 g_1 と g_2 を部分グラフとして含む最小のグラフであり、すなわち、これらの 2 つのグラフの最小共通スーパーグラフである。 $MCS(g_1, g_2)$ のノード v とエッジ e の重み $\lambda(v)$ 、 $\varepsilon(e)$ は、以下の式により定める。

$$\lambda(v) = \lambda(v_1) + \lambda(v_2) \quad (3)$$

$$\varepsilon(e) = \varepsilon(e_1) + \varepsilon(e_2) \quad (4)$$

ただし、 v_1 と v_2 は v に対応する g_1 および g_2 のノードであり、 e_1 と e_2 は e に対応する g_1 および g_2 のエッジである。対応するノードやエッジが存在しない場合 (\emptyset と表す) は、 $\lambda(\emptyset) = 0, \varepsilon(\emptyset) = 0$ とする。

\mathbf{G} の重み付き最小共通スーパーグラフ $WMCS(\mathbf{G})$ は、まず \mathbf{G} に含まれるすべてのグラフの全ノードと全エッジの重みを 1 とおき、式 (2) を逐次的に適用することに

よって構築する。 $WMCS(\mathbf{G})$ を構築する具体的なアルゴリズムを以下に示す。

procedure $WMCS(\mathbf{G})$

input: $\mathbf{G} = \{g_1, \dots, g_n\}$

output: $WMCS$ of \mathbf{G}

begin

$W = g_1$;

 for $i = 2$ to n

 begin for

$W = MCS(W, g_i)$;

 calculate λ and ε of W ;

 end for

$WMCS(\mathbf{G}) = W$;

end procedure.

$WMCS(\mathbf{G})$ のノード v とエッジ e の重み $\lambda(v)$, $\varepsilon(e)$ は、それぞれ対応するノードとエッジの \mathbf{G} における出現度数に相当する。従って、 $WMCS(\mathbf{G})$ から重みが高いノードとエッジのみを抽出したグラフは \mathbf{G} に属するグラフの細かな違いを無視した大局的共通構造とみなせる。閾値係数 t ($0 \leq t < 1$) を定め、閾値 $t \times n$ 以上の重みのノードとエッジのみを抽出したものを $RWMCS(\mathbf{G})$ と表わす。 \mathbf{G} があるカテゴリに属するグラフ集合のとき、 $RWMCS(\mathbf{G})$ は \mathbf{G} という学習データに極端に依存しない汎化性をもったカテゴリの特徴を表すとみなせる。

\mathbf{G} に含まれるグラフがどの程度似ているかを表わす尺度の1つに、エントロピーがある。 $WMCS(\mathbf{G})$ の重みの最大値は n であることから、 $WMCS(\mathbf{G})$ の各ノード、エッジの \mathbf{G} における出現確率はそれぞれ、 $p(v_i) = \frac{\lambda(v_i)}{n}$, $q(e_j) = \frac{\varepsilon(e_j)}{n}$ により与えられる。 $WMCS(\mathbf{G})$ のエントロピーを、

$$EN(WMCS(\mathbf{G})) = - \sum_{i=1}^p \frac{\lambda(v_i)}{n} \log \frac{\lambda(v_i)}{n} - \sum_{j=1}^q \frac{\varepsilon(e_j)}{n} \log \frac{\varepsilon(e_j)}{n} \quad (5)$$

と定義する。ただし、 p と q はそれぞれ $WMCS(\mathbf{G})$ のノード数およびエッジ数である。 \mathbf{G} が似た構造のグラフを多く含む場合、 $\lambda(v_i)$, $\varepsilon(e_j)$ が n に一致、あるいは近い値となり、エントロピーは小さくなる。

3. 提案手法

提案手法では、シルエット画像をその画像の構造を表わすグラフに変換し、グラフを用いて認識を行なう。認識対象となるカテゴリごとに複数のシルエット画像を用意しておき、これらの画像からそのカテゴリの画像の特徴を表わす代表グラフを1つ定める。未知のシルエット画

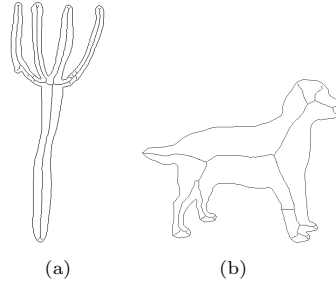


図1 シルエットから抽出した中心軸の例。

Fig. 1 Example of medial axes obtained from silhouette images.

像が与えられたとき、その画像から得られるグラフと各カテゴリの代表グラフとの類似性を測ることににより、認識を行なう。

3.1 シルエット画像のグラフ表現

まずシルエット画像に対して中心軸変換を施し、中心軸を抽出する。シルエット画像の中心軸（スケルトン、骨格ともいわれる）は、シルエット画像の輪郭に少なくとも2点以上で内接する円の中心点の集合として定義される。中心軸の各点は8近傍で連結しており、各点における内接円の半径の情報を保持させる。輪郭線から中心軸を求める古典的な手法は草焼きのイメージの細線化³⁾であるが、Kimiaらは輪郭線上から流体が湧き出る波面の発展方程式を用いて、この発展方程式の特異点が中心軸に対応することを示した¹⁵⁾。Siddiqiらはこの特異点を簡便に求める手法⁴⁾を提案した。本論文ではSiddiqiらの手法を用いてシルエットの中心軸を抽出する。図1にシルエット画像から中心軸を抽出した例を示す。

次に、中心軸上の各点をノードまたはエッジに割り当てることによりシルエット画像の構造を表わすグラフを作成する。点 (i, j) の8近傍に中心軸の点が2点あるならエッジ点、そうでなければノード点とする。これにより、中心軸の端点や分岐点がノード点となり、それらをグラフのノードに対応させる。また、隣り合うエッジ点の集合をグラフのエッジに対応させる。1つのエッジは2つのノードに接続する。なお、中心軸変換によって部分的に中心軸が集中し、複数のノード点が隣接することがあるが、その場合は隣り合った複数のノード点の集合をグラフにおける1つのノードとみなす。

図1に示すように、実際のシルエット画像から中心軸変換を行なうと、画像の輪郭の微妙な凹凸によって非常に多くの中心軸が抽出される。これらは画像の構造を反映しているとは言えず、 $WMCS$ の構築や認識の際に悪影響を及ぼす。そこで、得られたそれぞれのエッジの重要度を定め、ノイズ的なエッジを除去することを考える。

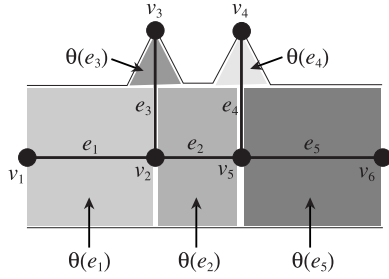


図 2 エッジに対応する領域.

Fig. 2 Regions corresponding to edges.

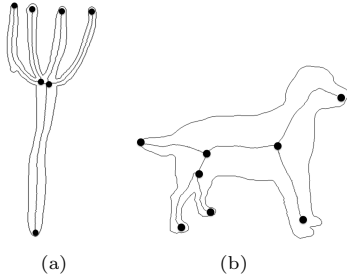


図 3 シルエットから得られたグラフの例.

Fig. 3 Example of graphs obtained from silhouette images.

重要度は、文献¹⁾の方法に倣い、エッジに対応する元のシルエット画像の領域の面積を総面積で割って正規化したもの(「パーツ面積」と呼ぶ)を用いた。詳細は省略するが、この方法ではノードとエッジの接続関係をもとに元の画像を分割し、領域を割り当てる。エッジに対応する領域を求めた例を図 2 に示す。 $e_1 \sim e_5$ がエッジであり、それぞれのエッジに対応する領域は $\theta(e_1) \sim \theta(e_5)$ で示される長方形または三角形の領域となる。ノイズに対応すると思われるエッジ (e_3 と e_4) に対しては小さな領域が割り当てられる。そして、各領域の面積をシルエット画像の総面積で割って正規化し、パーツ面積を求める。エッジ e_i のパーツ面積を $\theta(e_i)$ で表わす。

パーツ面積をエッジの重要度とみなし、 $\theta(e_i)$ の値が小さいものをノイズエッジとみなし除去する。例として、図 3 に、図 1 のシルエット画像から得られたグラフからノイズを除去したグラフを示す。各グラフ (a), (b) のノード数は、ノイズ除去前はそれぞれ 33, 32, ノイズ除去後はそれぞれ 7, 8 となった。ノイズを除去したグラフはシルエット画像の大局的な構造を表わしている。

なお、このパーツ面積 $\theta(e)$ はノイズ除去のみならず、認識にも用いる。そこで、 $\theta(e)$ もグラフの特徴を表わすパラメータとして保持するものとする。式 (1) の定義を拡張し、以降はグラフ g を

$$g = (\mathbf{v}, \mathbf{e}, \lambda, \varepsilon, \theta) \tag{6}$$

と表わす。

3.2 認識

あらかじめカテゴリごとにシルエット画像を n 枚用意して学習データとし、カテゴリを代表するグラフを構築する。各シルエット画像から 3.1 で述べた方法で重み付きグラフを構築し、そのグラフの集合を $\mathbf{G} = \{g_1, \dots, g_n\}$ とおく。 \mathbf{G} に含まれるすべてのグラフから 2. で述べた方法で最小共通スーパーグラフを求める。グラフを式 (6) のように拡張したので、式 (3) と式 (4) による重みの統合に加え、 θ を求める必要がある。ここでは、 $WMCS(\mathbf{G})$ のエッジ e に対応する、グラフ g_1, \dots, g_n のエッジをそれぞれ e_1, \dots, e_n として、

$$\theta(e) = \text{median}(\theta(e_1), \dots, \theta(e_n)) \tag{7}$$

と定める。このようにして定めた重み付き最小共通スーパーグラフ $WMCS(\mathbf{G})$ あるいは $RWMCS(\mathbf{G})$ を \mathbf{G} の代表グラフと呼ぶ。

未知のシルエット画像の認識のために、式 (5) で定義されるグラフのエントロピーを用いた認識アルゴリズムを与える。既に構築されている $WMCS(\mathbf{G})$ にもう 1 つグラフ g を加えて新たな $WMCS$ を構築しようとする場合、新たに加えるグラフが \mathbf{G} に含まれるグラフと似ていればエントロピーの増加は小さく、似ていなければエントロピーの増加が大きい。この性質を用い、提案手法ではエントロピーの増分 $\Delta EN(WMCS(\mathbf{G}), g)$ を用いてグラフを認識する。

$$\begin{aligned} \Delta EN(WMCS(\mathbf{G}), g) = \\ EN(WMCS(WMCS(\mathbf{G}), g)) \\ - EN(WMCS(\mathbf{G})) \end{aligned} \tag{8}$$

$WMCS(\mathbf{G})$ とテストデータのシルエット画像から得られるグラフ g に対して、対応するエッジの妥当性をパーツ面積 θ の差分を用いて定義し、上述のエントロピーの増分と合わせて認識を行なう。具体的には、ペナルティ pn を次式により定義する。

$$pn(WMCS(\mathbf{G}), g) = \sum |\theta(e_{WMCS(\mathbf{G})}) - \theta(e_g)| \tag{9}$$

ここで、 e_g は $WMCS(\mathbf{G})$ のエッジ $e_{WMCS(\mathbf{G})}$ に対応する g のエッジを表わす。 $e_{WMCS(\mathbf{G})}$ に対応するエッジがない場合は θ の値をそのまま pn に加える ($\theta(0) = 0$)。カテゴリ \mathbf{G} と g との距離 (非類似度) $d(\mathbf{G}, g)$ を式 (10) で定義する。 α はある定数で、2 つの評価値を統合するためのパラメータである。 $d(\mathbf{G}, g)$ は g が $WMCS(\mathbf{G})$ の部分グラフであれば 0 に近い小さい値をとる。 $d(\mathbf{G}, g)$ が最も小さいカテゴリを認識結果とする。

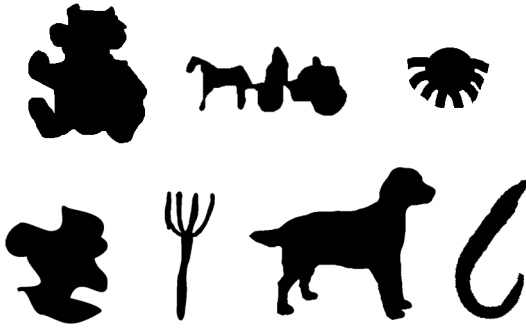


図4 実験に用いたシルエット画像の例. 上段左から teddy, carriage, octopus. 下段左から Misk, fork, dog, sea_snake.

Fig. 4 Silhouette images used for experiments: teddy, carriage, octopus, Misk, fork, dog, and sea_snake.

$$d(\mathbf{G}, g) = \Delta EN(WMCS(\mathbf{G}), g) + \alpha \cdot pn(WMCS(\mathbf{G}), g) \quad (10)$$

4. 実験

提案手法の有効性を確認するために実験を行なった. データとして, MPEG7 CE-Shape-1⁷⁾に含まれる teddy, Misk, carriage, octopus, fork, dog, sea_snake の7カテゴリを用いた. 各カテゴリには20枚の画像が含まれ, 1から20までの通し番号が付けられている. 以下の実験ではこの番号で画像を参照する. 各カテゴリのシルエットの例を図4に示す. MPEG7 CE-Shape-1 データセットには穴のあいている画像も含まれているが, 提案手法は穴のないシルエット画像を対象としているため, 穴のあいている画像については輪郭線から内側を塗りつぶす加工を施した.

まず, 各カテゴリから WMCS, RWMCS を抽出する実験を行った. 各カテゴリごとに10枚のシルエット画像(画像1,2,3,4,5,6,7,8,9,10)を学習データとし, これらの画像からグラフを抽出し, WMCS または RWMCS を構築することでカテゴリの代表グラフを構築した(ただし $t = 0.5$ とした). 図5および図6に, WMCS および RWMCS を用いて構築された各カテゴリの代表グラフを示す. 図中の数字はノードの重み λ を表す. WMCS と RWMCS を比較すると, WMCS は重みの小さいノードのトポロジが学習されているのに対して, RWMCS では重みの大きいノードのトポロジのみを保持している. これは WMCS は学習データすべてのトポロジが反映しているが, RWMCS では学習データ内に共通する平均的なトポロジ有していることを示している.

次に, WMCS, RWMCS の有効性を評価するため,

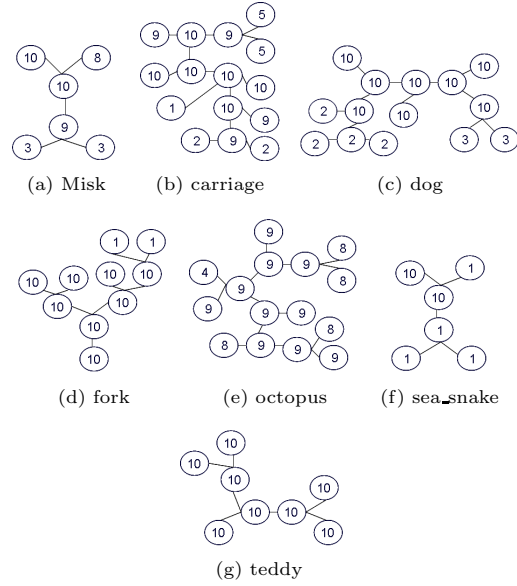


図5 WMCS を用いて構築された各カテゴリの代表グラフの例. 数字は λ .

Fig. 5 Examples of representative graph for each category constructed by WMCS. The number is λ .

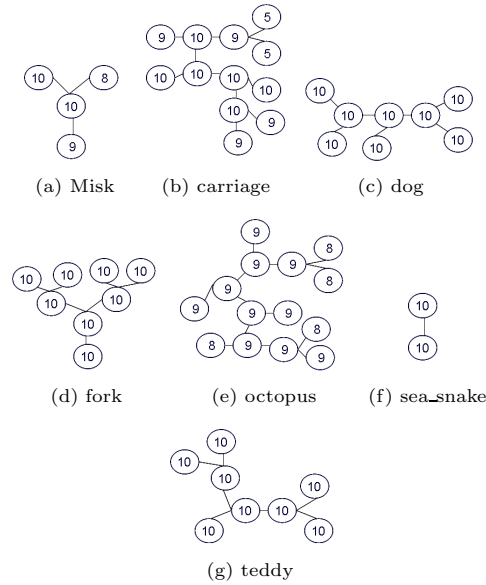


図6 RWMCS を用いて構築された各カテゴリの代表グラフの例. 数字は λ .

Fig. 6 Examples of representative graph for each category constructed by RWMCS. The number is λ .

ランダムに10枚の学習データを選択して WMCS, RWMCS を構築し, 学習データとは異なるデータを認識

表 1 正しいカテゴリに認識された数 ($\alpha = 3.0$)。ランダムに各カテゴリ 10 枚の学習データを用いて代表グラフを構築した認識実験を 20 回行ったときの平均値と最大値, 最小値を示す。

Table 1 Number of correctly recognized images ($\alpha = 3.0$).

		Misk	carriage	dog	fork	octopus	sea_snake	teddy	認識率
WMCS	平均	7.1	8.1	5.4	7.2	5.1	9.5	10	74.6(%)
	最良の結果	10	8	9	9	6	10	10	88.6(%)
	最悪の結果	6	6	2	5	2	9	10	57.1(%)
RWMCS	平均	6.1	8.6	6.5	8.8	7.2	9.5	9.8	80.5(%)
	最良の結果	7	9	9	9	8	10	10	88.6(%)
	最悪の結果	1	10	7	8	4	9	10	70.0(%)
手法 ⁸⁾	平均	5.8	5.3	6.6	7.7	6.8	9	9	71.6(%)
	最良の結果	8	5	7	9	7	9	9	77.1(%)
	最悪の結果	3	5	6	6	6	9	9	62.9(%)

表 2 提案手法と手法⁸⁾の比較。

Table 2 Comparison of the proposed method and method⁸⁾.

		Misk	carriage	dog	fork	octopus	sea_snake	teddy	認識率
学習データ A	提案手法	9	10	7	10	8	10	10	91.4(%)
	従来手法	4	6	4	7	3	9	9	60.0(%)
学習データ B	提案手法	6	6	2	5	2	9	10	57.1(%)
	従来手法	3	5	6	6	6	9	9	62.9(%)

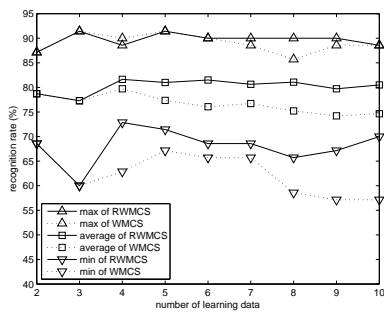


図 7 学習データ数と認識率の関係。

Fig. 7 Relations between number of learning data and recognition rate.

する実験 20 回行った。比較のため、従来手法としてカテゴリのプロトタイプとテストデータのグラフ編集距離により認識する手法⁸⁾を用いた。この手法は、プロトタイプとテストデータの 2 つのグラフを一致させるために必要な最小の編集コストを距離とし、認識を行なうものである。認識率の平均値と最大値, 最小値を表 1 に示す。表中の数字は 10 枚のテストデータ中、正しいカテゴリに認

識された画像の数である。また、代表グラフ作成に用いた学習データ数と認識率を図 7 に示す。各学習データ数 (2~10 枚) においてランダムに学習データを選択して代表グラフを作成し、認識する試行を 20 回行いその平均値を示している。すべての試行の中で、学習データ 5 枚 (画像 8,12,3,10,15) を用いて WMCS を作成したとき、最高の認識率 91.4% が得られ、最低の認識率 57.1% が学習データ数 10 枚 (画像 7,16,20,19,18,2,6,10,17,5) を用いた WMCS により得られた。

提案手法との条件をあわせるために、従来手法においても提案手法と同じ複数のプロトタイプを用い、テストデータを最小の距離を与えるプロトタイプを含むカテゴリに認識する。RWMCS, WMCS, 従来手法⁸⁾を用いた認識実験の平均認識率を図 8 に示す。図 8 より、提案手法である RWMCS や WMCS の認識率は、従来手法よりもはるかに高いことが分かる。また、提案手法が最良の結果となった学習データ A (画像 8,12,3,10,15)、このときのテストデータ (画像 2,4,5,6,7,11,13,16,17,18)、提案手法が最悪の認識結果を得た学習データ B (画像 7,16,20,19,18,2,6,10,17,5)、このときのテストデータ (画

像 1,3,4,8,11,12,13,14,15) をそれぞれ用いた場合の認識結果を表 2 に示す。表 2 から、提案手法は従来手法と比較して、データセットのよっては認識率が大幅に向上することと、認識率が下がる場合もその差はわずかであることが分かる。

実験結果から提案手法はプロトタイプを用いた従来手法よりも多くの画像が正しく認識されており、提案手法の有効性が確認された。表 1 の平均認識数から、特に Misk や carriage などでの複雑な形状をもつ画像において、提案手法は従来手法に対して約 13%~33%の認識数が向上した。提案手法により得られるグラフは複数のプロトタイプを単純に用いて認識する場合よりも良い認識精度を得ることができる。

提案手法により学習データ (画像 1,2,3,4,5,6,7,8,9,10) から得られた代表グラフを用いて、テストデータ (画像 11,12,13,14,15,16,17,18,19,20) を認識実験を行った場合、dog カテゴリが最低の認識数 6 となった。dog カテゴリのテストデータの多くは fork や teddy カテゴリに誤認識していた。これは dog や fork, teddy カテゴリのテストデータはほぼ同じトポロジとなっていることが原因と考えられる。しかし、 $\alpha = 0.4$ とすると、10 画像中 9 画像が正しく認識された。 α が大きいとき、同一カテゴリでペナルティ項が小さい部分でも距離が大きくなってしまい異なるカテゴリとみなされてしまうことがあるが、小さい場合は距離も小さくなり正しく認識されると考えられる。これは、式 (9) によって定義されるペナルティ項の有効性を示すものである。 α を変化させたときの認識結果を図 9 および図 10 に示す (学習データ数は 10 に固定)。 α の値が大きすぎると認識率が一定ないし下がる傾向にあり、ペナルティ項が効かないことを示している。今回は α の値を固定して実験を行なったが、カテゴリに応じて適切な α の値を決めることができれば認識性能がさらに向上することが期待される。

また、表 1 において Misk カテゴリでは、WMCS を用いた方が RWMCS を用いるよりも精度が良く、一方、carriage では RWMCS の方が WMCS を上回っている。これは構築した RWMCS の閾値係数 t により定まる閾値で除去されたノードとエッジが、Misk では必要なノードとエッジであったが、carriage では不必要なノードとエッジであったためと思われる。したがって、 t の値について、カテゴリに応じて適切な値を決めることで認識性能が向上することが期待される。

本論文ではシルエット画像の中心軸から得られるグラフを用いた手法を提案しているため、比較手法としてグラフを用いた代表的な手法⁸⁾を用いた。シルエット画像認識においては、中心軸を用いる以外にも多くの手法が

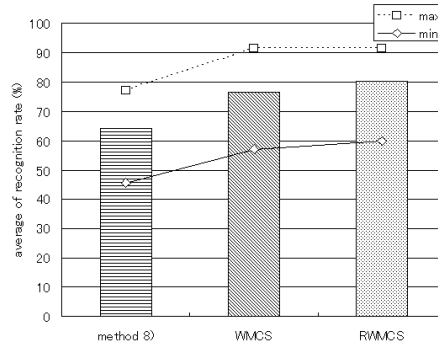


図 8 手法⁸⁾と WMCS, RWMCS を用いた平均認識率。
Fig. 8 Average recognition rate using method⁸⁾, WMCS and RWMCS.

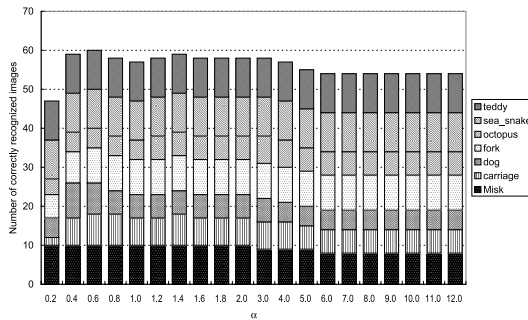


図 9 WMCS を用いた場合の様々な α における認識結果
Fig. 9 Results using WMCS with various α .

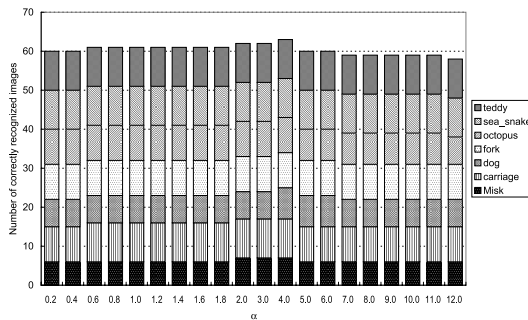


図 10 RWMCS を用いた場合の様々な α における認識結果
Fig. 10 Results using RWMCS with various α .

提案されており、これらの手法と提案手法を比較することは今後の課題とする。

5. むすび

本論文では学習データのシルエット画像からカテゴリを代表するグラフを構築し、認識する手法を提案した。提案手法は、学習用シルエット画像から共通的なグラフを

抽出し、これを認識に用いるものであり、学習データからプロトタイプを選んで認識する従来手法とは大きく異なるものである。7 カテゴリーのシルエット画像を用いて実験を行ない、各カテゴリーの画像の構造を表すグラフが構築できること、および、提案手法により従来手法よりも良い精度で認識ができることを確認した。

今後の課題として、RWMCS の閾値係数 t によって必要なノードやエッジが除去されてしまう場合があることから、適切な t を定める方法を検討することが挙げられる。さらに、パラメータ α が認識性能を変化させることから、カテゴリに適した α の定め方を検討することも課題である。

参考文献

- 1) E. Rosch, C.B. Mervis, W.D. Gray, D.M. Johnson, and P. Boyes-Bream: "Basic objects in natural categories," *Cognitive Psychology*, vol.8, no.3, pp.382-439, July 1976.
- 2) S. Belongie, J. Malik and J. Puzicha: "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.24, no.4, pp.509-522, 2002.
- 3) H. Blum: "Biological shape and visual science," *Journal of Theoretical Biology*, vol.38, pp.205-287, 1973.
- 4) K. Siddiqi, S. Bouix, A. Tannenbaum, and S.W. Zucker: "Hamilton-Jacobi skeleton," *Int. J. Comput. Vis.*, vol.48, no.3, pp.215-231, 2002.
- 5) K. Siddiqi, A. Shokoufandeh, S. J. Dickenson, and S. W. Zucker: "Shock graphs and shape matching," *Int. J. Comput. Vis.*, vol.35, no.1, pp.13-32, 1999.
- 6) S. Tirthapura, D. Sharvit, P. Klein, B. B. Kimia: "Indexing based on edit-distance matching of shape graphs," *SPIE Conference on Multimedia Storage and Archiving System*, pp.25-36, 1998.
- 7) L. J. Latecki, R. Lakaemper, and U. Eckhardt: "Shape Descriptors for Non-rigid Shapes with a Single Closed Contour," *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR 2000)*, pp.424-429, 2000.
- 8) P. N. Klein, T. B. Sebastian, and B. B. Kimia: "Shape matching using edit-distance: an implementation," *Proc. Twelfth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA-01)*, pp.781-790, Jan. 2001.
- 9) A. Torsello and E. R. Hancock: "A skeletal measure of 2D shape similarity," *Compt. Vis. Image Understand.*, vol.95, pp.1-29, 2004.
- 10) T. B. Sebastian, P. N. Klein and B. B. Kimia: "Recognition of shapes by editing their shock graphs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.26, no.5, pp.550-571, May 2004.
- 11) 影山勝彦, 大町真一郎, 阿曾弘具: "中心軸変換を用いたシルエット図形の構造化," *信学論 (D)*, vol.J90-D, no.6, pp.1568-1578, Jun. 2007.
- 12) A. Torsello and E.R. Hancock: "Learning shape-classes using a mixture of tree-unions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.28, no.6, pp.954-967, June 2006.
- 13) H. Bunke, P. Foggia, C. Guidobaldi, and M. Vento: "Graph clustering using the weighted minimum common supergraph," *Lecture Notes in Computer Science*, vol.2726, pp.235-246, Aug. 2003.
- 14) J. J. McGregor: "Backtrack search algorithms and the

maximal common subgraph problem," *Software Practice and Experience*, vol.12, pp.23-34, 1982.

- 15) B. Kimia, A. Tannebaum, and S. Zucker: "Shapes, shocks, and deformations I: the components of two-dimensional shape and the reaction-diffusion space," *International Journal of Computer Vision*, vol.15, no.3, pp.189-224, July 1995.

(2009 年 4 月 8 日受付)

(2009 年 7 月 14 日再受付)

宮 崎 智



2006 年山形大学工学部情報科学科卒。現在、東北大学大学院博士後期課程在学中。画像認識の研究に従事。

大 町 真一郎 (正会員)



1988 年東北大学工学部情報工学科卒。1993 年同大学院工学研究科情報工学専攻博士後期課程了。同年同大情報処理教育センター助手。1996 年同大工学部助手。1999 年同大学院工学研究科助教授。現在准教授。博士(工学)。その間、2000 年~2001 年米国ブラウン大学客員准教授。パターン認識、コンピュータビジョン、並列処理、文字認識システムの開発などの研究に従事。2007 年画像の認識・理解シンポジウム MIRU 長尾賞

および IAPR/ICDAR Best Paper Award を受賞。IEEE、電子情報通信学会、情報処理学会、人工知能学会などの会員。

阿 曾 弘 具



1968 年東北大学工学部電気工学科卒。1974 年同大学院工学研究科電気及び通信工学専攻博士後期課程了。1973 年同大工学部助手。1979 年名古屋大学工学部講師。1982 年同大助教授。1986 年東北大学工学部助教授を経て、1991 年同大教授。2009 年日本大学工学部教授。工博。その間、学習オートマトン、セル構造オートマトン、並行処理理論、ストリクアルゴリズム設計論、文字認識、音声認識、ニューラルネット

ワークなどの研究に従事。1991 年度電子情報通信学会業績賞受賞。IEEE、ACM、EATCS、電子情報通信学会、情報処理学会、人工知能学会、LA 各会員。