視覚特徴による日常的対象のカテゴリー形成

○畑本恵子, 石原茂和, 松原行宏, 長町三生 (広島大学工学部)

Category Formation of Everyday Things Based on Visual Properties

Keiko HATAMOTO, Shigekazu ISHIHARA, Yukihiro MATSUBARA, and Mitsuo NAGAMACHI (Faculty of Engineering, Hiroshima University)

1. はじめに

我々人間の視覚情報の認識の特性として、非常に高速である、形状の変形やノイズに強いということはマシンビジョンでの対象認識の難しさとの対比から広く知られている。見ているものをすぐに理解することができるのは、視覚対象の大まかな特徴を把握して記憶とマッチするからであると考えられる。また、対象は見る角度が変化したり、他のものが重なっていたりする場合にも認識できるのは、我々が対象物の形状を抽象化したモデルとして記憶の中に表象しているからである。

また、筆者らは主観的輪郭 (物理的に輪郭がない位置に知覚される輪郭) が知覚される場合にも記憶中に保持されている対象の形状がトップダウンの情報として利用されることを示した¹⁾。

対象モデルの記述方法の1つとして、Biedermanはgeonと呼ぶ単純な3次元形状を持つ要素の組合せによる記述を提案している²⁾。

本研究では、日常よく見られる対象について、入力された形状から対象の名前を高速に検索する、また人力が部分的であっても対応できる記憶モデルを構築することを目的とする。対象の視覚的特徴をgeonによって記述し、類似した対象のカテゴリーを自発的に形成し検索するメカニズムを考える。

2. 対象モデルの記述

視覚対象は単純な体積的要素 (geon) とその結合関係で記述される。各geonは断面の対称/非対称、曲線/直線、軸の直/曲といった、視点の変化に左右されにくい、我々にとって判別の容易な属性に基づいており、24種類が提案されている²⁾。Geon間の接続関係は、"TOP-OF", "SIDE-CONNECTED" など6種類が提案されている (全部は明らかにしていない²⁾)。

本研究では、視覚対象をgeonとその接続関係で記述するため、新たにgeon間の関係を定義する。すなわち、各物体の要素の中で最も大きいもの、機能的な中心をなすものを中心として、それとの相対的關係として、"上にある"、"下にある"、"前にある"、"横にある"、"角にある"、"中心geonより大きい"、の6種類を用いる。Geonをノード、関係をリンクとす

ると、日常的な対象はほとんど中心的geonを根にした2階層の木構造で記述可能である (図1)。

このように各物体をモデルとして抽象的に記述することによって、個々の物体間、たとえばボールペンと万年筆のような形状の類似性をモデルの中を含めることができる。したがって類似したものをカテゴリーにしておき、新たに見た対象に対してマッチングを行えば、形状の小さな変化に対応して認識することができる。

3. ニューラルネットを使ったカテゴリー形成

視覚システムには、入力された形状が、分類基準を外部から与えられることなく、人力の集合によって自己組織的に分類されるメカニズムが必要である。本研究ではこれを具体化するものとしてART 1.5 (Adaptive Resonance Theory 1.5)^{3,4)}の学習方法を改良したネットワーク⁵⁾を用いる。

3.1 ART 1.5による分類の学習

ネットワークは図2に示す構造をもっており、 F_1 にベクトル $(x_1, \dots, x_m)$ が入力される度に、 F_2 の中でカテゴリー・ノード v_j が1つ選択される。 x_i と v_j との間の結合係数 z_{ij} は次式によって修正される。

図1. 対象のgeonによる記述と対応するgeon-tree

図2. ART 1.5 の構造

$$\Delta z_{ji} = (x_i - z_{ji}) / k_j$$

ここで k_j はカテゴリ j が選択された回数であり、分類における入力順序による影響を同じにするよう改良したもの⁵⁾である。

3.2 geon-treeで記述された対象のコーディング

前章で述べた geon-tree を次のように長さ204ビットのベクトルとして表現した。

rootのgeon名(24bits) leaf1の関係(6bits) leaf1のgeon名(24bits) leaf6の関係 leaf6のgeon名

3.3 カテゴリーの形成

家の中に日常的にあるものの種類と色、その形状を上記述べた方法で記述して、ネットワークに入力し、カテゴリーを形成した。結果の一部を図3に示す。構成する geon と関係に対応したカテゴリーが形成された。

また、このネットワークに新しい対象を入力すると、それが動的に妥当なカテゴリーに分類されることを示した。

Category 1 rootが2である：鉛筆 ボールペン シャープペンシル 灰皿 針 糸 湯呑 盆
Category 2 leaf1が横にある：万年筆 黒板 ソファ ベッド
.....

図3. 形成されたカテゴリーの例

4. 視覚表象データベースの検索

前章で構築した視覚表象データベースを用いて、カテゴリーのプロトタイプに対応する視覚表象および、部分的な視覚表象から可能性のある対象の名前を検索することができる。現在は拡張prologであるMacCupで記述したインターフェースを用いて検索する。

4.1 プロトタイプの視覚表象

図4に検索結果の例を示す。(a)はカテゴリーのプロトタイプとして学習された表象の例である。インスタンスの間で共通する属性が抽出されており、共通度の高い geon や接続関係に大きい値が割り当てられている。(b)は部分的な入力から対象名を検索した例である。

_q(グループ1のプロトタイプは何ですか?).
_?-start.
leaf1が10である?1
leaf1が下にある?2
rootが2である?4
no.
Elapsed Total time = 6.45 sec (Constraints Handling = 0.00 sec)

(a) プロトタイプの属性

_q(rootが2であるものは何ですか).
_?-start.
鉛筆
ボールペン
シャープペンシル
灰皿
針
糸
湯呑
盆
鍋
no.
Elapsed Total time = 0.95 sec (Constraints Handling = 0.00 sec)

(b) 部分的な入力からの検索

図4. データベースの検索結果

5. おわりに

日常的な対象は geon-tree のような抽象化したモデルによって表象することが十分可能であることが示された。また抽象化することによって、表象の中に小さな形状の変化を含めることができる。このように視覚特徴によって類似したものをカテゴリーとしておくことによって、探索範囲を小さくして検索を高速化することができ、また形の小さな変化に対応することができる。

【参考文献】

- 1) 畑本, 石原, 長町, 松原: 視覚対象の認識における記憶表象の役割, 日本人間工学会誌, 第27巻, 特別号, 224/225 (1991).
- 2) Biederman, I.: Higher-level vision, In: Osherson, D. N. et al. (Eds.), Visual cognition and action, Vol. 2, MIT Press, 41/72 (1990)
- 3) Carpenter, G. A., Grossberg, S.: The ART of adaptive pattern recognition by a self-organizing neural network, IEEE Computer, 21, 3, 77/88 (1988).
- 4) Levine, D. S. & Penz, P. A.: ART 1.5 - A simplified adaptive resonance network for classifying low-dimensional analog data, IJCNN-90-WASH DC, Vol. 2, 639/642 (1990).
- 5) 上田中: ニューラルネットワークと記号処理を用いた人間の記憶のモデルに関する研究, 平成3年度広島大学工学部卒業論文 (1992).