

観察角度に依存しない物体認識のニューラルネットモデル

小林 遼太[†] 山根 健[†] 森田 昌彦[†]

[†]筑波大学大学院システム情報工学研究科 〒305-8573 つくば市天王台 1-1-1

1 はじめに

画像を用いた物体認識が困難である主な要因の一つに、観察角度によって物体の見え (アピアランス) が大きく変化してしまうことが挙げられる。このため同一クラス内の変動が大きくなり、識別境界面が特徴空間の中で複雑に絡み合う非線形なものとなる。

従って、単純なテンプレートマッチングのような従来手法では、1つの物体の異なる視点に対応したプロトタイプが多数必要となり、物体が増えるにつれて記憶すべきデータ量や認識に要する時間が飛躍的に増大してしまう。これに対して、人間や高等動物は、知っている物体であれば、ほぼどんな視点からでも正しく認識できる上に、記憶している物体の数がいくら増えても認識に要する時間はほとんど変わらない。

本研究では、非単調ニューラルネット [1] に形成した線状のアトラクタを用いることによって、人間と同様な性質をもつ物体認識のモデルを提案する。また、脳における特徴抽出や物体表現に関する知見を参考にして、画像情報の符号化を行い、実際にこのモデルにいくつかの 3 次元物体を識別させる実験を行った。

2 非単調ニューラルネットを用いた物体認識モデル

2.1 画像情報の記述

従来のパターン認識では、特定のクラスに特有な特徴をできるだけ少数抽出し、特徴空間を低次元化することが求められることが多い。しかし、3次元物体認識の場合、クラス内に共通する特徴が少ないため、次元の圧縮によってしばしばモデルの認識性能が低下してしまう。これに対し本モデルでは、観察角度の変化による見えの連続的な変化に伴って、非単調ニューラルネットの状態が連続的に変化することが重要である。そこで、特徴空間を必要以上に低次元化するのではなく、対象物体の情報が広く分散して含まれるような特徴量を多数用意し、見えの変化に伴って特徴パターンが徐々に変化するようする。

本研究では、以下に述べるように、形状に関する特

徴パターンと色に関する特徴パターンをそれぞれ別々に生成した上で、両者を統合することとした。

形状特徴に関しては、まず入力画像に 8 方向の Gabor フィルタを掛けエッジ情報を取得する。ある程度位置ずれを許容するために、Gabor 特徴画像にぼかし操作を行う。一般に観察角度の変化によるアピアランスの変化は画像の中心部より周辺部の方が大きいので、画像の位置によって異なる大きさの受容野を用い、周辺部では中心部より 3 倍ほど大きくぼかすこととした。こうして得られた画素値とそれらの組み合わせの値を、実験的に得られた閾値によって ± 1 に 2 値化することで、3136 次元の形状特徴パターンを得る。

色特徴の抽出では始めに入力画像を HSV 系に変換する。そして形状特徴抽出過程と同様のぼかし操作を行い、画像の中心部領域から色相の値を取得する。画素毎の色相値を 11 段階に量子化し、 ± 1 を等確率にとる 10 個の素子によって各段階を表現することで色特徴パターンを得る。このとき、色相の値は色相環に沿った角度によって与えられるため、各色相を表現するパターンも環状となるように符号化を行っている。

抽出された二つの特徴パターンを選択的の不感化 [2] によって統合する。ここでは色特徴を修飾情報としたため、各画像の形状特徴パターンは、色特徴パターンに対応する部分空間に射影されることになる。

2.2 学習と認識

上記のような符号化を行うと、ある物体を徐々に回転させたとき、それに対応する特徴パターンは特徴空間内で連続的な軌道を描く。このような軌道は、学習によって非単調ニューラルネットの軌道アトラクタ [1] にすることができる。

そこで、ある物体をさまざまな角度から見た画像がある場合に、それぞれの形状特徴パターンを補間したパターン系列をいくつか作成し、その軌道を同一クラス内の色特徴パターンに対応した部分空間において学習させる。このとき、各軌道の終点はその物体のプロトタイプ (ここでは正面画像) の特徴パターンとする。これをすべての物体について何度か繰り返すと、同じ物体の特徴パターンを結んだ軌道アトラクタ群が形成される (図 1)。

学習後の回路網にある軌道アトラクタの近くに位置する特徴パターンを初期状態として与えると、回路網

A neural network model of view-invariant object recognition

[†] Ryota KOBAYASHI(ryo@bcl.esys.tsukuba.ac.jp)

[†] Ken YAMANE

[†] Masahiko MORITA(mor@bcl.esys.tsukuba.ac.jp)

Graduate School of Systems and Information Engineering,
Tsukuba University ([†])

1-1-1 Ten-nodai, Tsukuba-shi, Ibaraki 305-8573, Japan

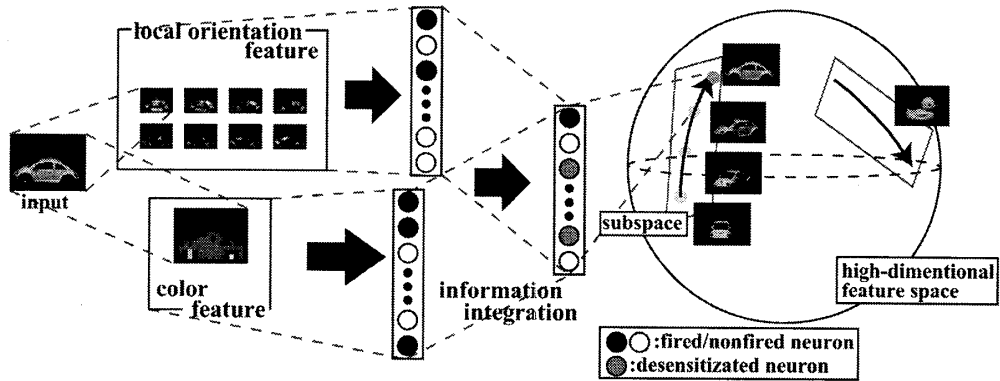


図 1: 特徴抽出過程とモデルの状態空間

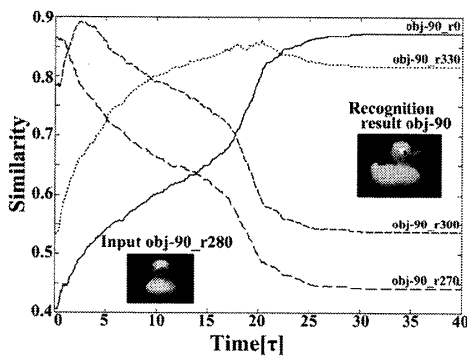


図 2: 認識過程

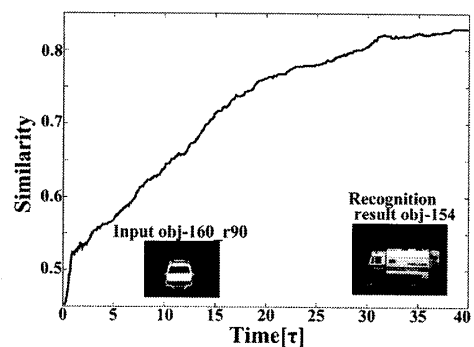


図 3: 未学習物体の類推

の状態はその軌道に引き込まれた後、軌道に沿って遷移し、正面画像の特徴パターンに達することが期待される。これにより、出力特徴パターンとプロトタイプの特徴パターンを比較することで、最初に与えた画像がどの物体の画像であるかを識別できる。

3 実験

実験には、Amsterdam Library of Object Images[3]の Image-Database よりランダムに取得した 50 個の物体に関する画像群を用いた。それぞれの物体について、観察方向が水平方向 -90 度から 90 度まで 30 度ごとの画像 7 枚を用いて学習系列を生成し、十分な回数だけ学習を行った。

学習後のモデルに未学習の見えを含む -90 度から 90 度までの 5 度ごとの画像を入力したところ、すべての物体を正しく識別することができた。認識過程の一例を図 2 に示す。このグラフの縦軸は、回路網の状態パターンと学習に用いた画像の特徴パターンとの類似度、横軸はネットワークのダイナミクスの時定数 τ を単位とする時間である。また、未知の物体を入力したときの認識の過程を示したのが図 3 である。学習したプロトタイプの特徴パターンの一つが想起されてい

ることがわかる。

4 おわりに

物体の見えの連続的な変化を、特徴空間内に軌道として表現し、その軌道への引き込みによって物体認識を行うモデルを提案した。このモデルは単純な構造でありながら、見え方の大きく異なる物体を一つの内部表現に対応させることができる。また、認識に要する時間は、観察方向に依存する（プロトタイプと異なるほど長い）が物体の数にはほとんど依存しない。このように、このモデルは従来のモデルよりも人間に近い性質をもつと考えられる。

参考文献

- [1] 森田昌彦：非単調ダイナミクスを用いた時系列パターンの連想記憶，電子情報通信学会論文誌，Vol.J78-D-II，No.4，pp.679-688 (1995)。
- [2] 森田昌彦，松沢浩平，諸上茂光：非単調神経素子の選択的不感化を用いた文脈依存的連想モデル，電子情報通信学会論文誌，Vol.J85-D-II，No.10，pp.1602-1612 (2002)。
- [3] J.M. Geusebroek, G.J. Burghouts, and A.W.M. Smeulders: The Amsterdam Library of Object Images, International Journal of Computer Vision, Vol.61, No.1, pp.103-112 (2005)。