

TAM ネットワークの受容野構造における機能比較

A Functional Comparison of Receptive Field Structure in TAM Network

林 勲
Isao Hayashi
関西大学
Kansai University

福島 邦彦
Kunihiko Fukushima
関西大学
Kansai University

Abstract TAM network is a biologically-motivated neural network with receptive fields of Gabor function. However, it's not good performance for rotating visual images since the receptive field's layer is a mono-type. In this paper, we formulate four types of multilayer structures with extensive receptive fields, and discuss the performance using some examples of character recognition.

1. はじめに

人間の視細胞には、単純型細胞、複雑型細胞、超複雑型細胞が存在し、初期視覚の受容野構造のモデルとしてガボール関数が提案されている [1]。網膜から視覚高次への Hubel-Wiesel の階層仮説に準拠する多くのモデル [2] が提案されており、その一つの TAM (Topographic Attentive Mapping) ネットワーク [3] はガボール関数を用いた受容野構造を入力部にもつ。TAM ネットワークは 6 階層構造からなる。網膜層から特徴マップ層の下位層では、ガボール関数の受容野構造により、対象画像の方位選択成分を抽出して、第一視覚野における特徴マップを構成し、上位層は第一次視覚野以降の視覚前野を模倣する。しかし、受容野構造が単層構造であるため、十分な方位選択性を抽出しているとは言えず、対象画像の回転特性に弱いという短所がある。

本論文では、新たな多層性の受容野構造を提案し、対象画像の十分な方位選択性を確保して、回転特性を効果的にするモデルを提案する。具体的には、視野占有型受容野構造、反復受容野構造、方位ベクトル型受容野構造、反復方位ベクトル型受容野構造の合計 4 種類を提案し、特徴マップ層で方位選択性に対応した特徴マップを構成して、基盤層以降で TAM ネットワークの学習機能により対象画像の回転性を認識する。また、ここでは簡単なアルファベットの文字認識の例を用いて、本提案モデルの有用性を検討する。

2. TAM ネットワーク

TAM ネットワークの構造を図 1 に示す。TAM ネットワークは、網膜層、外側膝状体 (LGN) 層、特徴マップ層、基盤層、カテゴリー層、出力層の 6 階層構造からなる。網膜層は二層の受容野構造をもち、第一層では、ガボールフィルタリングにより原画像の輝度情報から 16 方位選択成分を抽出する。第二層目の受容野構造では、より広範囲な方位選択成分を抽出して、広範囲な方位選択成分から対象画像の回転特性を得て、特徴マップ層での 16 方位の特徴マップを変更する。得られた特徴マップから、上位層での学習機能により対象画像を認識する。

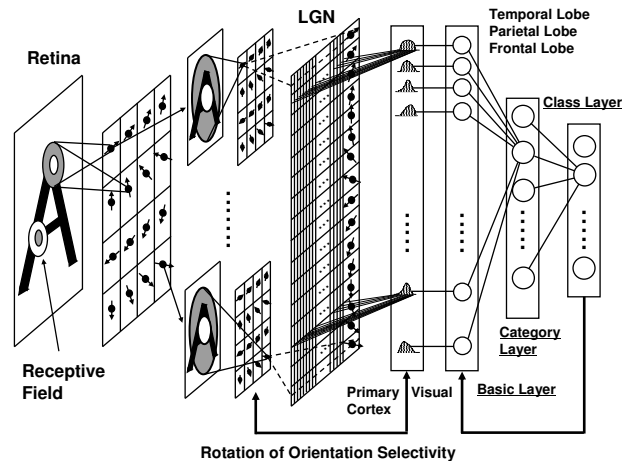


Fig. 1: TAM Network with Double Receptive Fields

3. 受容野構造の提案

方位 ϕ の 2 次元ガボール関数を次に示す。

$$G_{\phi}(x, y) = Ke^{-\frac{1}{2}\left(\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2}\right)} \times \sin(2\pi f_x x \cos \theta + 2\pi f_y y \sin \theta + \phi)$$

ただし、 K は振幅、 σ_x と σ_y は標準偏差、 f_x と f_y は周波数である。

いま、原画像 $O(p, q)$ のサイズを $R_H \times R_V$ とすると、第一層目の受容野構造における方位 ϕ の方位選択成分 $C_{\phi}^1(x, y)$ は、次の畳み込み計算により計算される。

$$C_{\phi}^1(x, y) = \sum_{q=1}^{R_V} \sum_{p=1}^{R_H} G_{\phi}(x-p, y-q) \times O(p, q)$$

第二層目でも同様に、方位 ϕ の方位選択成分 $C_{\phi}^2(x, y)$ は畳み込み計算により求める。ここでは、次の 4 種類の受容野構造を提案する。

1) 視野占有型受容野構造 (CVA):

$$C_{\phi}^2(s, t) = G_{\phi}^{max}(s - x, t - y) \times C_{\phi}^1(x, y)$$

2) 反復受容野構造 (DL):

$$C_{\phi}^2(s, t) = \sum_{x=1}^{R_V} \sum_{y=1}^{R_H} G_{\phi}(s - x, t - y) \times C_{\phi}^1(x, y)$$

3) 方位ベクトル型受容野構造 (OV):

$$C_{\phi}^2(s, t) = \frac{G_{\phi}^{max}(x - p, y - q) \times O(p, q)}{\sum_{\phi} G_{\phi}^{max}(x - p, y - q) \times O(p, q)} \times \sum_{q=1}^{R_V} \sum_{p=1}^{R_H} G_{\phi}(s - p, t - q) \times O(p, q)$$

4) 反復方位ベクトル型受容野構造 (DOV):

$$C_{\phi}^2(s, t) = \frac{G_{\phi}^{max}(x - p, y - q) \times C_{\phi}^1(p, q)}{\sum_{\phi} G_{\phi}^{max}(x - p, y - q) \times C_{\phi}^1(p, q)} \times \sum_{q=1}^{R_V} \sum_{p=1}^{R_H} G_{\phi}(s - p, t - q) \times O(p, q)$$

4 手法の概念図を図 2 に示す．視野占有型受容野構造では，第二層目に視野全体型受容野を設定し原画像の全方位選択性を獲得する．反復受容野構造では，第二層目を第一層目の受容野の反復形とし原画像の中領域の方位選択性を獲得する．方位ベクトル型受容野構造では，方位性ベクトルにより確実に方位性を検知する．反復方位ベクトル型受容野構造は第一層と第三層目の手法の合併型であり，原画像全体の方位選択性を方位性ベクトルで修正する．なお，学習後，評価用データの方位選択成分画像に対して，その特徴マップは学習用の特徴マップとの差分量が最小になる方位に入れ替えられる．

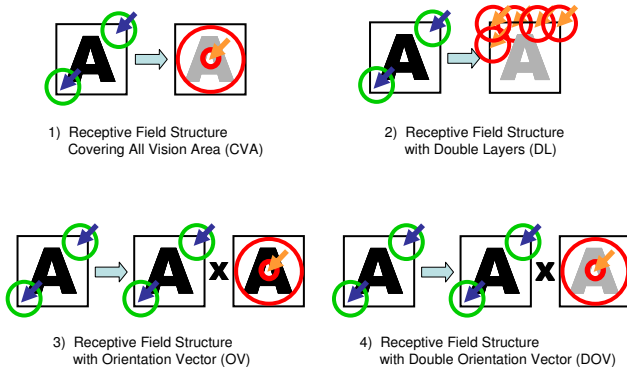


Fig. 2: Four Methods with Double Receptive Fields

4. 文字認識の例

受容野構造の有用性を確認するため，アルファベットの文字認識を行った．電子タブレットに画像データを用い，大きさを $15 \text{ pixels} \times 15 \text{ pixels}$ とした．アルファベットの学習画像に対し， 22.5° ずつの回転を与えた 32

種類の文字を評価画像として回転特性を検分した．評価用画像の認識率を表 1 に示す．認識率はデータ順序を変えた 20 回の平均値を表している．反復受容野構造 (DL) 以外の 3 手法は認識率が改善されており，特に，反復方位ベクトル型受容野構造 (DOV) が最も優れていることがわかる．図 3 に反復方位ベクトル型受容野構造の方位選択成分の例を示す．しかし，回転角度によっては十分に認識できない文字もあり，他のアルファベット文字の認識も含め，まだ検討の余地がある．

Table 1: Recognition Rate of Four Methods

	Char. 'A' (%)	Char. 'B' (%)	Average (%)
Conv. TAM	63.99	59.95	62.12
CVA	75.88	62.78	69.52
DL	49.44	59.63	54.23
OV	72.99	66.21	69.72
DOV	76.92	67.42	72.12
Average	68.80	64.01	66.40

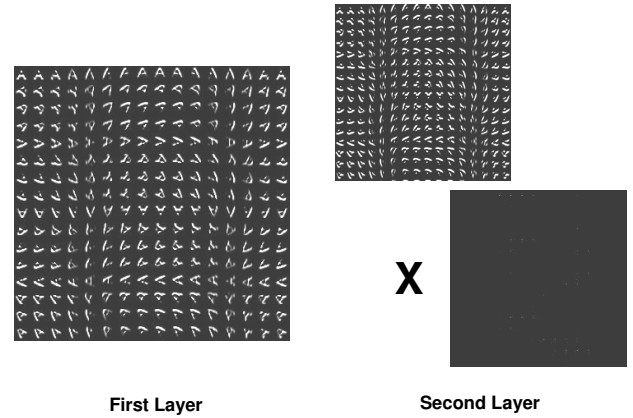


Fig. 3: Orientation Images of DOV

5. おわりに

ここでは，二層型受容野構造を導入した新たな TAM ネットワークについて議論した．なお，本研究は，文部科学省の科学研究費 (18500181) の助成を得た．

参考文献

- [1] 福島: 神経回路と情報処理, 朝倉書店 (1989)
- [2] S.Grossberg: How does the cerebral cortex work? Learning, attention, and grouping by the laminar circuits of visual cortex, *Spatial Vision*, Vol.12, No.2, pp.163-185 (1999)
- [3] 林, Williamson: TAM Network のブルーニング手法の提案, システム制御情報学会論文誌, Vol.17, No.2, pp.81-88 (2004)