ガボール関数を用いたTAMネットワークの受容野入力構造

Receptive Field Type Input Layer of TAM Network Using Gabor Function

林　勲
Isao Hayashi
関西大学
Kansai University

前田 裕正
Hiromasa Maeda
情報科学センター
Information Science Research Center Corp.

Abstract　TAM (Topographic Attentive Mapping) network is a biologically-motivated neural network. Gabor function is a receptive field model and can detect orientations of objects using a Gaussian envelope and sine/cosine function. In this paper, we formulate a receptive field type input layer for TAM network using Gabor filtering to receptive field of human visual cortex.

1. はじめに

網膜の視覚情報は受容野に対応した各細胞で処理され，対象の知覚はより上位の視覚連合野で認識される [1]．視覚細胞には，単純型細胞，複雑型細胞，超複雑型細胞等がある．特に，超複雑型細胞はオン反応とオフ反応の受容野をもち，スリット状の光刺激に対して方位選択性を示し，対象の縁を検出する．最近，多くの受容野モデルが提案されている [2–7]．特に有用なモデルとしてガボール関数 [6,7] がある．ガボール関数はガウス関数と正弦・余弦関数からなる関数であり，任意の周波数成分を抽出するフィルタリング機能をもつ．一方，多くの視覚系ニューラルネットワークが提案されている [8–11]．特に，TAM(Topographic Attentive Mapping) ネットワーク [11,12] は有用である．入力層，基盤層，カテゴリー層，出力層の 4 層からなり，神経節細胞から外側膝状体，第一次視覚野を模擬している．

本論文では，TAM ネットワークの入力層にガボール関数を用いた受容野構造を定式化する．具体的には，受容野の輝度情報からガボール関数のフィルタリング機能によって知覚対象の縁の 16 方位選択成分を抽出し，濃度化して TAM ネットワークの属性マップを構成する．この画像の入力値と中間層のカテゴリー層，上位層のクラス層における共振学習により画像対象を認識する．ここでは，定式化の有用性を確認するため，文字認識の例を用いて，ガボール関数のパラメータ値に対する方位選択成分画像の感度，TAM ネットワークの認識率と頑健性について議論する．

2. ガボール関数による受容野入力構造

2 次元正弦関数型ガボール関数 $G(x,y)$ は次のように定義される．

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

ただし，K は振幅，(μ_x,μ_y) はガボール関数の中心，σ_x と σ_y は標準偏差，f_x と f_y は周波数を表す．$\phi = \pi/2$ の場合には，正弦関数型ガボール関数は余弦関数型ガボール関数となる．図 1 にガボール関数の例を示す．ただし，$(\mu_x,\mu_y) = (0.0, 0.0)$，$\sigma_x = 2.0$，$\sigma_y = 2.0$，$f_x = 1.0\ Hz$，$f_y = 1.0\ Hz$，$\theta = 0°$，$\phi = 0.0$ としている．

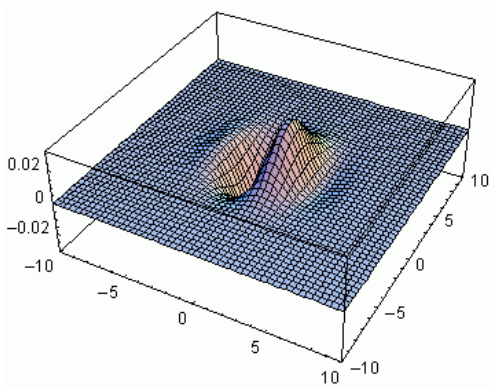

Fig. 1: Gabor Function

ガボール関数を用いた受容野の構成と方位選択成分の検出方法を定式化する．方位選択成分の検出過程を図 2 に示す．いま，原画像の横幅と縦幅をそれぞれ，$R_H\ pixels$ と $R_V\ pixels$ とする．原画像 $I(p,q),\ 0 \le p \le R_H,\ 0 \le q \le R_V$ における受容野の i 番目の方位選択成分を $C_i(x,y),\ \ i = 1,2,\cdots 16$ とし，次の畳み込み計算により計算する．

$$C_i(x,y) = \sum_{q=1}^{R_V}\sum_{p=1}^{R_H} G_i(x-p, y-q) \times I(p,q)$$

画像に配置された受容野を移動し，$R_H \times R_V$ 個の方位選択成分 $C_i(x,y)$ を計算する．TAM ネットワークの i 番目の属性マップは，方位選択成分として得られた輝度を各方位ごとに正規化して構成される．

$$f_{ih} = \frac{\sum_{\{x,y|O_i(x,y)=h\}} C_i(x,y)}{\sum_{y=1}^{R_V}\sum_{x=1}^{R_H} C_i(x,y)}, \quad h = 1,2,\cdots,256$$

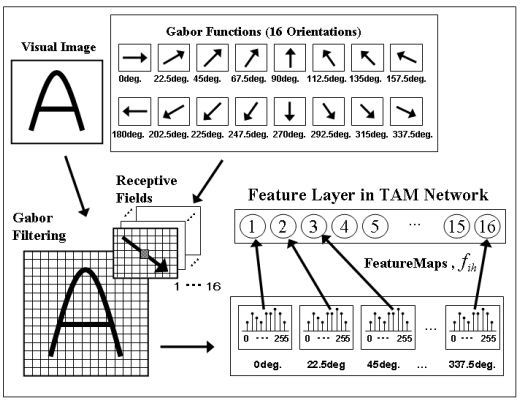

Fig. 2: Gabor Filtering Process

3. TAM ネットワーク

TAM ネットワークの構成を図 3 に示す．カテゴリー層では，属性マップ f_{ih} と学習荷重 w_{jih} との活性値 x_{ji} を計算し，出力層への出力 y_j を得る．

$$y_j = \prod_{i=1}^{M} x_{ji} = \prod_{i=1}^{M} \frac{\sum_{h=1}^{L} f_{ih} w_{jih}}{1 + \rho^2 b_{ji}}$$

ただし，ρ はビジランスパラメータであり，b_{ji} は抑制荷重である．

TAM ネットワークの出力値 K は次式より計算される．

$$K = \{k | \max_k z_k\} = \{k | \max_k \sum_{j=1}^{N} y_j p_{jk}\}$$

ただし，p_{jk} はカテゴリー層ノードに対する学習荷重である．

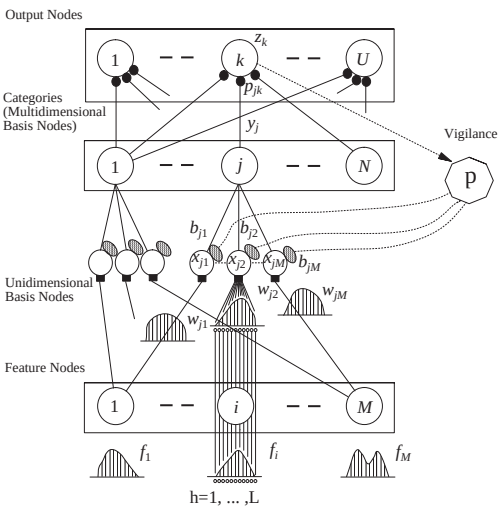

Fig. 3: TAM Network

いま，K^* を正しい教師値としよう．$K \neq K^*$ の場合には，ビジランスパラメータ ρ を初期値 $\rho = 0$ から上昇させ，$z_{K^*}/z_K \geq OC$ の条件が満足されるか最大値になるまで上昇さる．最大値になった場合には，カテゴリー層ノードを 1 個分増加させる．

一方，$z_{K^*}/z_K \geq OC$ を満足する場合には，学習モードに入り，学習荷重 w_{jih}，p_{jk}，b_{ji} を次式で更新する．

$$\begin{aligned}
\triangle w_{jih} &= \frac{\alpha y_j^* (1 - \lambda^{1/M})(f_{ih} - w_{jih})}{(\alpha - 1)\lambda^{1/M} + n_j}, \lambda \in (0, 1) \\
\triangle p_{jk} &= \frac{\alpha y_j^* (z_k^* - p_{jk})}{\alpha + n_j} \\
\triangle b_{ji} &= b_j^{(rate)} y_j^* (x_{ji} - b_{ji}) \\
\triangle n_j &= \alpha y_j^* (1 - n_j) \\
y_j^* &= \frac{\prod_{i=1}^{M} x_{ji} \times \sum_{k=1}^{U} z_k^* p_{jk}}{\sum_{j'=1}^{N} \prod_{i=1}^{M} x_{j'i} \times \sum_{k=1}^{U} z_k^* p_{j'k}} \\
z_k^* &= 1 \text{ if } k = K^*; \quad z_k^* = 0 \text{ otherwise}
\end{aligned}$$

ただし，$b_j^{(rate)}$ は学習定数であり，α と λ は係数，y_j^* はカテゴリー層へのフィードバック信号である．

パラメータの調整はネットワークへデータが逐次的に投入されるごとに行い，学習データを 1 エポックとして数エポック回数分の学習を行う．

4. 文字認識による例題

ガボール型受容野をもつ TAM ネットワークの有効性を確認するため，アルファベット 'A' と 'B' を用いた文字認識を行った．入力画像として電子タブレットによって作成された画像データを用いた．画像の大きさを $15\ pixels \times 15\ pixels$ とする．ここでは，'A' と 'B' の文字画像から属性マップを構成し，TAM ネットワークの学習用データを作成して，学習後の評価用データに対する認識率を計算した．次の 3 つの点から TAM ネットワークの有用性を評価した．

- フィルタリング後の方位選択成分を解析する．
- TAM ネットワークの認識率を評価する．
- 評価用データに対する頑健性を評価する．

学習用画像を図 4 と図 5 に示し，評価用画像を図 6 に示す．評価用画像は学習用画像に対して，位置をずらした文字画像，回転させた文字画像，ノイズを加えた文字画像，ピクセルを欠如させた文字画像の 4 種類を用意した．TAM ネットワークの学習パラメータは次の通りである．

L	=	255	$\rho\ init$	=	0.0
OC	=	0.8	$\rho\ step$	=	0.1
α	=	0.0000001	$\rho\ max$	=	100.0
λ	=	0.33	$b_j^{(rate)}$	=	0.01
η	=	0.8	θ	=	0.03
ξ	=	0.5	ϕ	=	0.
μ_x	=	0.0	μ_y	=	0.0

図 4 の左端の 'A' の画像に対して，ガボール型受容野を用いて方位選択成分を抽出した結果を図 7 に示す．左上画像が方位 $0°$，その右側は $22.5°$ の方位選択成分で

あり，同様に，$45.0°$，$67.5°$，$\cdots$ ，$337.5°$ の方位選択成分を表す．これらの画像から文字の方位選択成分が良好に抽出されていることがわかる．

Fig. 4: Training Image of 'A'

Fig. 5: Training Image of 'B'

ガボール関数のパラメータ変化による方位選択成分の感度を議論するため，パラメータを次のように変化させ，方位選択成分を解析した．

$$\begin{aligned} \sigma_x &= 0.5, 0.75, \cdots, 4.0 \\ \sigma_y &= 0.5, 0.75, \cdots, 4.0 \\ f_x &= 0.01, 0.015, \cdots, 0.3 \\ f_y &= 0.01, 0.015, \cdots, 0.3 \\ K &= 0.1, 0.2, \cdots, 1.0. \end{aligned}$$

アルファベット 'A' に対して，標準偏差の σ_x を変化させた場合の $135.0°$ の方位選択成分を図 8 に示す．また，そのガボール関数を図 9 に示す．標準偏差 σ_x の値は，左から 0.5，1.0，1.5，2.0，2.5，3.0，3.5，4.0 である．ただし，他のパラメータ値を $\sigma_y = 1.92$，$f_x = f_y = 0.127$，$K = 1.0$ とした．標準偏差 σ_x の値が大きくなるほど，方位選択成分の画像は不明瞭となっていることがわかる．これより，標準偏差 σ_x は方位選択成分の画像の明瞭性を表現するパラメータといえる．同様に，標準偏差 σ_y を変化させた場合の方位選択成分も解析した．σ_y は標準偏差 σ_x と同じく方位選択成分の画像の明瞭性を表現するパラメータとなっているが，ガボール関数の長軸を規定するパラメータであるので，σ_x ほど感度は高くない．その他，周波数と振幅についても解析した．周波数 f_x と f_y は方位選択成分に対してそれほど感度は高くない結果が得られた．振幅 K も同様であった．

次に，TAM ネットワークの認識率を評価した．学習用画像と評価用画像に対する TAM ネットワークの認識率を図 10 に示す．学習用画像の認識率は全般的に高いが，評価用画像の認識率は標準偏差 σ_x の値が大きくなるほど低下している．特に，$\sigma_x = 0.75$ のとき認識率が 90.0% で最大となった．同様に，標準偏差 σ_y に対しては，$\sigma_y = 0.75$ のとき認識率が 82.0% で最大となった．画像の方位選択成分の解析結果でも明らかなように，標準偏差に対する認識率の低下は方位選択成分の明瞭性に影響を与える．標準偏差の値が小さいほど明瞭性があるので，標準偏差の値を比較的小さな範囲で調整した場合に高い認識率を得るという結果は納得できる．一方，周波数 f_x と f_y を変化させた場合の認識率は，学習用画像に対してほぼ 100% となった．しかし，評価用画像に対

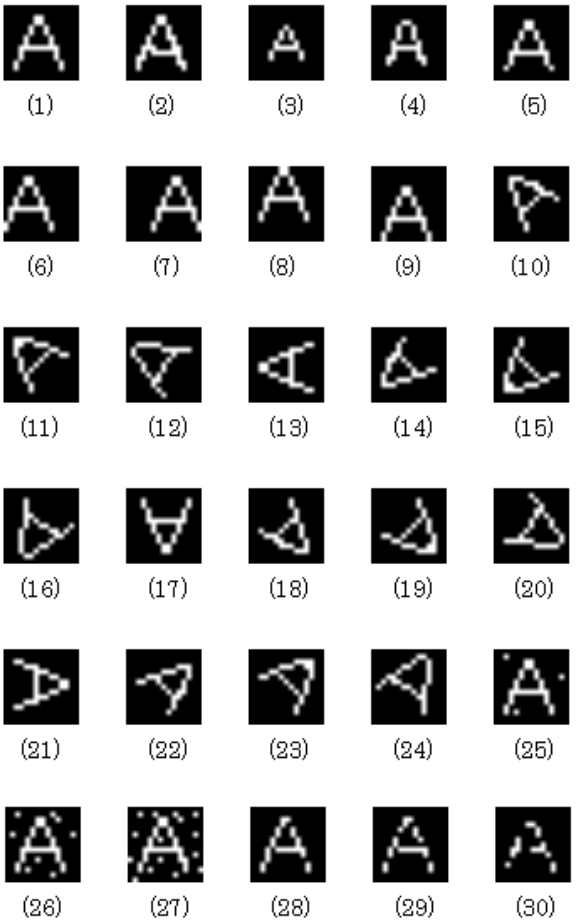

Fig. 6: Checking Image

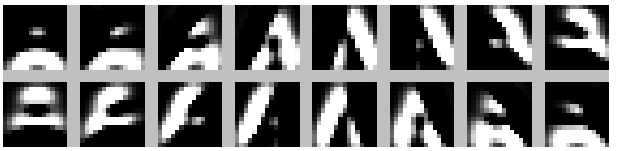

Fig. 7: Image after Gabor Filtering

Fig. 8: Detecting of Contourss of Input Images

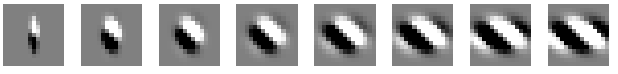

Fig. 9: Gabor Functions for Detecting Contours

しては，周波数 f_x と f_y の変化に対して認識率が鋭敏に反応し，高い認識率を得るために周波数を調整することが難しい．なお，振幅 K を変化させた場合の認識率は，学習用画像と評価用画像に対して，あまり変化がなかった．

最後に，TAM ネットワークの頑健性について検討した．4 種類のアルファベット 'A' の評価用画像に対して，標準偏差 σ_x を変化させた場合の認識率を図 11 に示す．位置ずれ文字と回転文字に対する認識率はほぼ 70.0%以上と高く，頑健性に優れている．しかし，ノイズを加えた文字に対する認識率は低く頑健性が低い．他の標準偏差 σ_y，周波数 f_x と f_y，振幅 K を変化させた場合の頑健性についても同様に，ノイズ追加文字以外の 3 つの文字に対する認識率は高く頑健性が高い結果を得た．一方，アルファベット 'B' の評価用画像に対しては，標準偏差，周波数，振幅の変化に対して，位置ずれ文字とピクセル欠如文字に対する頑健性が高い結果を得た．

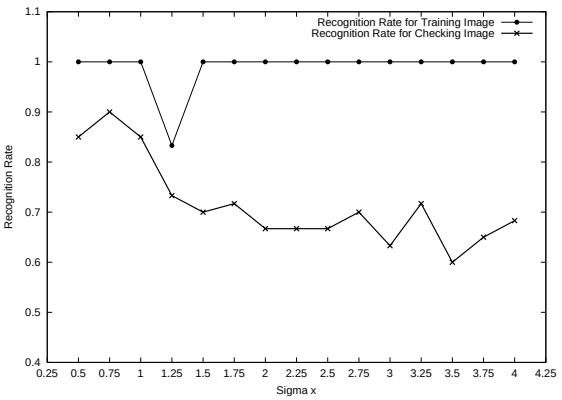

Fig. 10: Recognition Rate

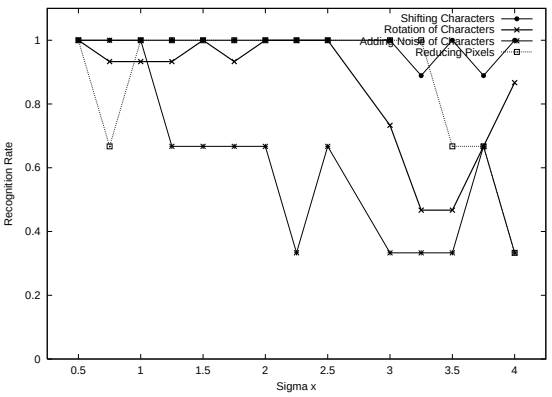

Fig. 11: Recognition Rate of Four Types of Image 'A'

5. おわりに

ここでは，ガボール関数を用いて TAM ネットワークの受容野入力構造を定式化した．本研究の一部は，文部科学省科学研究補助金 (基盤研究 C) 課題番号 14580433 の一環として行われた．

参考文献

[1] 松田 隆夫: 知覚心理学の基礎, 培風館 (2000)

[2] 岡島 健治, 今岡 仁: 情報量最大化と生体視覚細胞の受容野, 電子情報通信学会論文誌 *A*, No.J83-A, Vol.6, pp.620-628 (2000)

[3] A.D.Pollen and S.F.Ronner: Visual cortical neurons as localized spatial frequency filters, *IEEE Transactions of System, Man and Cybernetics*, Vol.SMC13, pp.907-916 (1983)

[4] W.Urushihara and T.Nagano: A model for the detection of second-order motion, *Technical report of IEICE*, Vol.NC98-191, pp.293-298 (1999)

[5] D.C.Lee: Adaptive processing for feature extraction: Application of two-dimensional Gabor function, *Remote Sensing*, Vol.17, No.4, pp.319-334 (2001)

[6] S.Marčelja: Mathematical description of the responses of simple cortical cells, *Optical Society of America*, Vol.70, No.11, pp.1297-1300 (1980)

[7] J.Daugman: Uncertainty relation for resolution in space, spatial frequency, and orientaion optimized by two-dimensional visual cortical filters, *Optical Society of America*, Vol.2, No.7, pp.1160-1169 (1985)

[8] 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)

[9] H.Neumann and W.Sepp: Recurrent V1-V2 interaction in early visual boundary processing, *Biological Cybernetics*, Vol.81, pp.425-444 (1999)

[10] K.Fukushima: Recognition of partly occluded patterns: a nural network model, *Biological Cybernetics*, Vol.84, No.4, pp.251-259 (2001)

[11] J.R.Williamson: Self-organization of topographic mixture networks using attentional feedback, *Neural Computation*, Vol.13, pp.563-593 (2001)

[12] 林 勲，J.R.Williamson: TAM Network のプ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)

[連絡先]

Isao Hayashi, Kansai University
2-1-1, Ryozenji-cho, Takatsuki, Osaka 569-1095
tel 072-690-2448
e.mail ihaya@res.kutc.kansai-u.ac.jp