

INFOMAX による視覚野細胞 GABOR 型受容野の導出

岡島健治

NEC 基礎研究所探索研究部

入力像に対して視覚野の細胞がどう応答するかについては「Gabor 型」と呼ばれる受容野モデルによって実験結果が良く再現できることが知られている。本稿ではこの Gabor 型受容野が持つ特徴抽出能力を情報理論的な観点から解析する。そして Gabor 型受容野を規定している「Gabor 関数」がある意味で最適な特徴抽出関数になっていることを示す。この結果は、この結果は、生物の視覚野細胞（の受容野）が、情報理論的に見ても最適に構成（自己組織化）されている事を強く示唆するものであると言える。

A Derivation of a Gabor-type Receptive Field of a Cortical Neuron by Infomax

Kenji Okajima

Fundamental Research Laboratories, NEC Corporation

The receptive field of a simple cell in the primary visual cortex is well approximated by the Gabor function. This Gabor-type receptive field is analyzed by an information theoretic approach. It is shown that, under general assumptions, the Gabor-type receptive field extracts the maximum (mutual) information from input local images, and in this sense, is optimally designed.

1. はじめに

入力像に対して視覚野（第一次視覚野、V1野）の細胞がどう応答するかについては「Gabor 型」と呼ばれる受容野モデルによって実験結果が良く再現できることが知られている[1]。本報告書では、「生物の視覚系はなぜこのような関数を用いて入力像の処理を行っているのか？」という問題について情報理論的な観点から考察する。

ここで関連する研究について簡単に述べておく。

Atick[4]は網膜神経節細胞（や外側膝状体細胞）の同心円型の受容野が、“minimum-entropy coding”と呼ばれる一種の効率的符号化の観点から説明できることを示している。Linsker[5]は相互情報量最大化 (Infomax) という指針に基づいて同様の結論が導かれることを示している。Bialek[6]も情報理論に基づいた受容野の研究を行っており、シュレーディンガー方程式に類似の方程式を導いている。しかし V1 野細胞の Gabor 型受容野に関しては、それを情報理論的な考察から導いた例は、筆者の知る限りこれまではない。

相互情報量の最大化と Hebb 型の神経結合学習則とは互いに密接に関連している事が知られており[5,8]、これまでに Hebb 則に基づく受容野の自己組織化の問題が多数議論されている[8-10]。また、Kohonen[9]は自己組織化特徴マップの一種である Adaptive-Subspace Self-Organizing feature Map (ASSOM)によって Gabor 関数様の受容野が自己組織化されることをシミュレーションによって示している。さまざまな神経結合学習則の能力を情報理論的な観点から解析する研究は、今後重要だろう。

2. GABOR 型受容野

サルなどの高等動物では、眼から入力した視覚情報は、途中で視床の一部である外側膝状体 (LGN) と呼ばれる中継核を経て後頭部にある第一次視覚野 (V1 野) と呼ばれる部位に入力し、そこで大脳皮質における最初の処理を受ける。

入力像に対して V1 野の (単純型) 細胞がどう応答するかについては次のような、空間周波数にチューンされた Gabor (ガボル) 関数型の受容野モデルによって実験結果が良く近似できることが知られている [1]。

$$F(\mathbf{k}, \mathbf{X}) = \sum_x f(\mathbf{x}) \phi(\mathbf{x} - \mathbf{X}; \mathbf{k}); \quad \phi(\mathbf{x}; \mathbf{k}) = G_{\sigma}(\mathbf{x}) \exp(i\mathbf{k} \cdot \mathbf{x}) \quad (2.1)$$

ここで $f(\mathbf{x})$ は網膜上に映った像 (の光強度分布) を表す。細胞の出力は (2.1) 式の $F(\mathbf{k}, \mathbf{X})$ (の実部或いは虚部) によって良く近似できる。この式で $f(\mathbf{x})$ にかかる重みの部分 $\phi(\mathbf{x} - \mathbf{X}; \mathbf{k})$ が Gabor 関数と呼ばれる関数である。すなわち Gabor 関数とは平面波をガウシアン of 包絡線

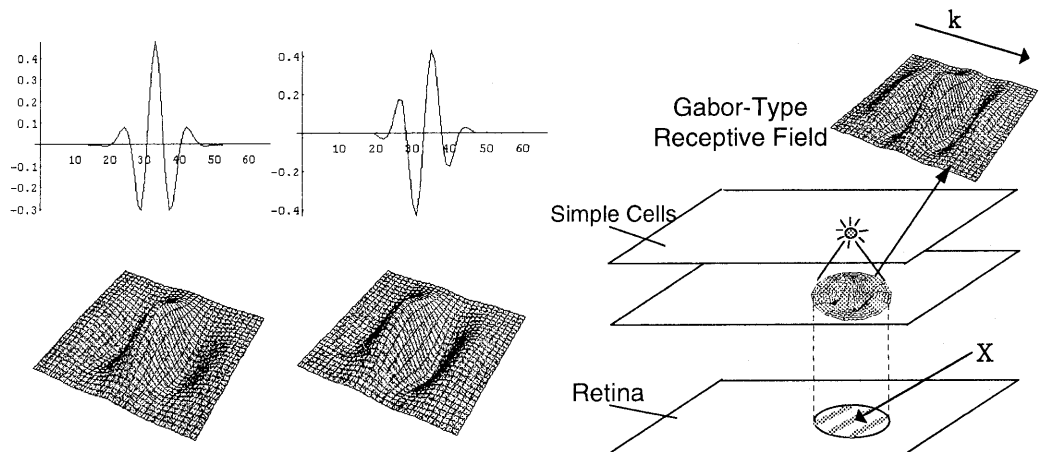


図 2.1 1次元の場合 (上段)、2次元の場合 (下段) の対称型 (cos 型、左列) および反対称型 (sin 型、右列) の Gabor 関数と Gabor 型受容野の概念図 (右側)

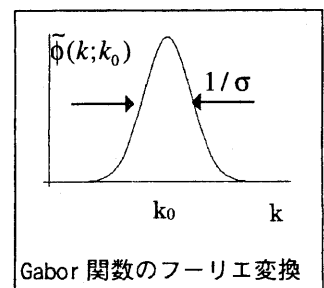
$G_{\sigma}(x)$ で局在化させたものである [図 2.1]。

例えば、図 2.1 右側のような受容野を持つ細胞を発火させるためには、受容野の山の部分と丁度一致するような傾き (方位) と間隔 (空間周波数) を持った縞模様パターンを見せてやれば良い。細胞の受容野中心位置 \mathbf{X} (及び最適空間周波数ベクトル \mathbf{k}) は細胞ごとに異なっている。V1 野の細胞全体で視野全体がカバーされている。

○不確定性の関係と Gabor 関数

ここで Gabor 関数が持つ良く知られた性質を一つ説明しておく。

Gabor 関数、 $\phi(x; k_0) = G_{\sigma}(x) \exp(ik_0 \cdot x)$ のフーリエ変換を $\tilde{\phi}(k; k_0)$ で表すと、 $\tilde{\phi}(k; k_0)$ は k_0 を中心とする偏差 $\sigma_k = 1/\sigma$ のガウシアンになる [右図]。すなわち Gabor 関数は周波数領域においても局在した関数になっている [従ってフィルターとしては、これはバンド・パス型の特性を持っている]。この時、実空間領域、周波数領域それぞれにおける Gabor 関数の局在幅を $\Delta x, \Delta k$ とする



すると ($\Delta x^2 \equiv \sum_x x^2 |\phi(x; k_0)|^2$ 、 $\Delta k^2 \equiv \sum_k (k - k_0)^2 |\tilde{\phi}(k; k_0)|^2$) 両者の間には $\Delta k \Delta x = 1/2$ の不確定性の関係が成り立っていることが分かる。ところで、偏差の積のこの値はとり得る値のうちで最小の値である事が知られており[11]、従って Gabor 関数は不確定性の積、 $\Delta x \Delta k$ を最小にするという性質を持つ事が分かる。また、この性質と関連して、二乗和 $\Delta k^2 + \gamma^2 \Delta x^2$ (γ はある定数) を最小にする関数もやはり Gabor 関数となることが証明できる[11]。この後者の性質については後で用いる [この時、Gabor 関数の包絡線であるガウシアン幅 σ は γ から $\sigma^2 = 1/\gamma$ で与えられる]。

さて、ここでの問題は「生物の視覚系は何故このような関数を用いて入力像の分析をしているのか?」という点である。以下、本報告書ではこの問題について考察を行う。

3. 受容野関数の最適性に関する評価基準

本節では、本報告書で考察する問題を数式を用いて定式化する。

○相互情報量

受容野関数を $\phi(x)$ とし、これによって (画像) 信号 $f(x)$ から特徴 a を測定することを考えよう。

$$a = (\phi, f) \equiv \sum_x \phi(x) f(x) \quad (3.1)$$

ここでは Linsker[5]に従い、このような特徴 a を測定する事によって得られる情報量を最大化するように最適受容野関数 ϕ を決定する。特徴 a を測定する事によって、信号 f に関して我々が得る情報量は、通常は、次式の相互情報量 MI で与えられる[12]。

$$MI = H[a] - H[a|f] \quad (3.2)$$

ここで、 $H[a]$ は a のエントロピーであり、周知の通り、これは a の生起確率 $P(a)$ を用いて次式で表される。

$$H[a] = -\sum_a P(a) \log[P(a)] \quad (3.3)$$

(3.2)式、第2項の条件付きエントロピーは条件付き確率 $P(a|f)$ を用いて次式により計算される。

$$H[a|f] = -\sum_f P(f) \sum_a P(a|f) \log[P(a|f)] \quad (3.4)$$

○並進対称性

以下、本報告書では画像を記述する統計に対して並進対称性を仮定する。すなわち、ある画像パターン $f(x)$ が網膜上に映ったとすると、それは常に同じ位置に映る訳ではなく、別の位置にずれたパターン $f(x - x_0)$ も同じ確率で生じ得るものと仮定する。従ってこの場合、入力画像のパターンは形式的には $f_{\alpha, x_0}(x) = f_{\alpha}(x - x_0)$ と表され、2つのパラメータ、すなわちパターンの種類を表すパラメータ (α) とパターンの位置を指定するパラメータ (x_0) で指定されることになる。

○二つの評価基準

最適な受容野関数を決定するための評価基準として、以下の2つの相互情報量について検討を行った。

$$(A) MI_1 = H[a] - H[a\{\alpha, x_0\}] \quad (= H[a] - H[a|f]) \quad (3.5)$$

$$(B) MI_2 = H[a] - H[a\{\alpha\}] \quad (3.6)$$

MI_1 は、特徴 a を測定する事によって、今どのようなパターン f_{α, x_0} が映っているのかに関して我々が得る事ができる情報量（相互情報量）である [これは(3.2)式で表される通常の相互情報量と同じである]。(A)の基準は、これを最大化する受容野関数を最適とする基準である。

他方、基準(B)では、特徴 a を測定する事によってどの種類のパターン α が提示されているか [その位置は問わない] に関して得られる情報量を評価基準とし、これを最大化する受容野関数を最適とする。

結論を先に述べておくと、

[1] 一次元の場合には、いずれの評価基準からも、比較的広い条件の下で Gabor 関数が最適受容野関数として導出されるが、

[2] 二次元の場合には（画像の統計が等方的な場合）(B)の評価基準からは、やはり Gabor 関数が最適受容野として導出されるが、(A)の基準からは Gabor 関数は導かれないことが分かった。従って、生物の視覚系では (B) の評価基準に従って視覚野細胞の受容野が構成（自己組織化）されている可能性が強いと言える。

本報告書では (A) の評価基準を中心に考察を行う。また、実際にはこれらの評価基準に対して、さらに受容野を局在化させるためのペナルティ項を付け加えて考察を行うが、これについては次節において詳しく説明する。

4. 評価基準(A)の解析

本節では(3.5)式で表される評価基準 (A) についての解析を行う。当面のゴールは(4.11)式の評価関数を導くことである。(4.11)式から最適受容野関数として Gabor 関数が導出されることを示すのは比較的容易である。

4. 1 局所像に対する Infomax と「局在化ポテンシャル」

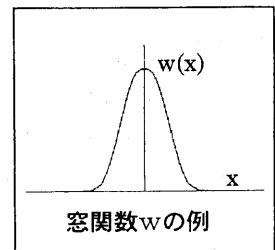
経験的には良い特徴量は局所的である。そこで以下では局所的な受容野関数に焦点を絞って考察を進めることにする。以下ではまず、局所像に対して Infomax を行うことによって、(4.10)式の評価関数、及びそこに現れる「局在化ポテンシャル」の項が導かれることを示す。

像 f を局所窓 $w(x)$ からのぞき、局所像 $g(x) = w(x)f(x)$ を得るとしよう [以下の議論ではこの窓関数 $w(x)$ の詳細は結論には影響しない。 $w(x)$ が中心 ($x=0$ とする) で滑らかな最大値をとり、周辺にいくに従って 0 に減衰していけば良い]。ベクトル形式で表すと

$$g = Wf, \quad W_{x,x'} = w(x)\delta_{x,x'} \quad (4.1)$$

である。ここで $\delta_{x,x'}$ は Kronecker のデルタである。

以下では、この局所像 g から受容野関数 ϕ によって特徴を抽出することを考え [$a = (\phi, g)$] その場合について、相互情報量の最大化 (Infomax) を行う。まず、簡単化のため当面の間、画像 f は (従って局所像 g も) ガウシアン¹⁾の統計で記述される信号であるものと仮定する (平均値は $\langle f(x) \rangle = 0$ とする)。また、画像 g に重畳するノイズ n は、画像とは独立で、無相関のガウシアン・ノイズであるとする (やはり平均は $\langle n(x) \rangle = 0$ とし、分散



は $\langle n(x)n(x') \rangle = \sigma_n^2 \delta_{x,x'}$ とする)。以上のような仮定の下では (3.5) 式の相互情報量は、

$$MI_1 = 1/2 \log \left[\frac{(\varphi, C' \varphi)}{\sigma_n \|\varphi\|^2} + 1 \right] \quad (4.2)$$

と計算される事が分かる [12]。ただし、ここで $C' \equiv \langle g g' \rangle$ は局所像 g の共分散行列である [g' はベクトル g の転置ベクトルを、 $\langle \rangle$ は平均値を求める操作を表す]。ここで (4.2) 式を眺めると、(4.2)式は受容野関数 φ を定数倍しても不変である事が分かる。従って (4.2) 式を最大化する事は規格化条件、 $\|\varphi\|^2 = 1$ の下で次式、

$$(\varphi, C' \varphi) \quad (4.3)$$

を最大化する事と等価になる。

ここでは、解析的な取り扱いを容易にするために、評価関数(4.3)式を原画像 f の共分散行列 $C = \langle f f' \rangle$ を用いて書き直しておく。局所像 g の共分散行列 C' は、原画像 f の共分散行列 C を用いて以下のように表される。

$$C' = \langle g g' \rangle = W \langle f f' \rangle W = W C W, \quad C = \langle f f' \rangle \quad (4.5)$$

従って(4.3)式は、

$$(\varphi, C' \varphi) = (\varphi, W C W \varphi) = (W \varphi, C W \varphi) \quad (4.6)$$

と変形される。ここで新たに関数 $\phi = W \varphi$ (4.7) を導入すると (4.3) 式は結局、

$$(\varphi, C' \varphi) = (\phi, C \phi) \quad (4.8)$$

と変形されることになる [測定される特徴 a は $a = (\varphi, g) = (\phi, f)$ と表されることに注意]。また、拘束条件 $\|\varphi\|^2 = 1$ は、 ϕ を用いて表すと、

$$\|W^{-1} \phi\|^2 = \sum_x u(x) |\phi(x)|^2 = 1 \quad (4.9)$$

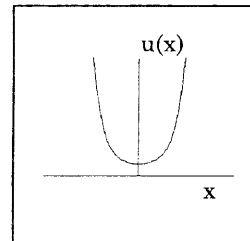
と書き表される。ただし、ここで $u(x)$ は、 $u(x) \equiv 1/w(x)^2$ で定義される関数である [以後これを「局在化ポテンシャル」と呼ぶことにする]。

従って、ラグランジュの未定乗数法を用いると、最適受容野関数を決定する評価関数として最終的には次式の評価関数が得られる。

$$(\phi, C \phi) - \mu \sum_x u(x) |\phi(x)|^2 \quad (4.10)$$

ここで μ はラグランジュの未定乗数である。

以下では、(4.10)式を最大化するように最適受容野関数 ϕ を決めていくが、この時注意すべき点は、拘束条件 $\sum_x u(x) |\phi(x)|^2 = 1$ によって $\sum_x |\phi(x)|^2$ の値は有限区間に限定されており、その結果、実際上は $\sum_x |\phi(x)|^2$ の値は一定と見なして(4.10)式の最大化を行っても差し支えないという点である。



4. 2 最適特徴抽出関数としての Gabor 関数の導出

さて、ここで画像 f を記述する統計に対して並進対称性を仮定する。共分散行列 $C = \langle ff' \rangle$ が並進対称性を持つ場合（すなわち行列要素 $C_{x,x'}$ が $x-x'$ だけの関数で表される場合）に成り立つ恒等式

$$(\phi, C\phi) = \sum_k p(k) |\tilde{\phi}(k)|^2 \quad \begin{array}{l} p(k): \text{power spectrum of } f \\ \tilde{\phi}(k): \text{Fourier transform of } \phi(x) \end{array}$$

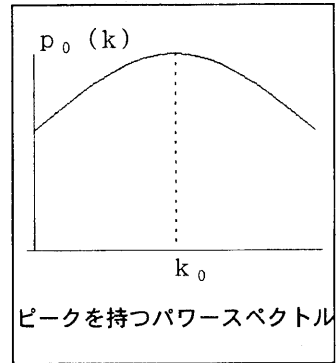
を用いると、最大化すべき評価関数 (4.10) 式はさらに

$$\sum_k p(k) |\tilde{\phi}(k)|^2 - \mu \sum_x u(x) |\phi(x)|^2 \quad (4.11)$$

と変形される。ただし、ここで $p(k)$ は画像のパワースペクトル、 $\tilde{\phi}(k)$ は求めるべき受容野関数の Fourier 変換である。

ここで(4.11)式を見ると、まず、第1項は画像のパワースペクトル $p(k)$ が最大となる周波数に成分が集中しているような関数に対して有利に働くことが分かる。すなわち第1項は受容野関数を周波数領域で局在化させる方向に働く。それに対して第2項は、受容野関数を実空間で $[u(x)$ の最小値付近に] 局在化させる効果を持つ。結果として周波数領域、実空間領域双方で局在している関数が最適受容野関数となることになる。

ここでパワースペクトル p と「局在化ポテンシャル」 u をその最大値、あるいは最小値のまわりで2次までの中に展開すると ($p(k) = p_0 - \alpha(k - k_0)^2$, $u(x) = u_{\min} + \beta x^2$)、評価関数(4.11)式は



$$-\left[\sum_k \alpha(k - k_0)^2 |\tilde{\phi}(k)|^2 + \mu\beta \sum_x x^2 |\phi(x)|^2 \right] + (p_0 - \mu u_{\min}) \sum |\phi(x)|^2. \quad (4.12)$$

と表される。

$\sum_x |\phi(x)|^2 = \text{一定}$ の下では、この(4.12)式の第2項は定数となる。第1項に対しては不等式、

$$-\left[\sum_k \alpha(k - k_0)^2 |\tilde{\phi}(k)|^2 + \mu\beta \sum_x x^2 |\phi(x)|^2 \right] \leq -\sqrt{\mu\beta\alpha} \quad (4.13)$$

が成り立つ [11]。ここで等号は $\phi(x) = c \exp(-\gamma x^2 / 2 + ik_0 x)$; $\gamma = \sqrt{\mu\beta/\alpha}$ の時、すなわち ϕ が Gabor 関数である時に成立する [周波数空間、実空間における局在幅の二乗和を最小化する関数は Gabor 関数である事を用いる]。従って、パワースペクトルが最大値をとる k_0 を中心周波数とする Gabor 関数が (4.13) 式の評価関数を正確に最大化する最適受容野関数となることが分かった。

○数値解

これまでの議論で得られた結論を確かめるために (4.2) 式の最大化問題を数値的に解いてみた。これは次の固有方程式の最大固有値に属する固有関数を求める問題と等価である。

$$C' \phi = \lambda \phi \quad (4.4)$$

パワースペクトルの持つ対称性 $p(k)=p(-k)$ から、この固有方程式の最大固有値は少なくとも二重に縮退していることが予想される。図 4.1 に数値計算の結果を示す。予想通り、最大固有値はほぼ二重に縮退していた。また、最大固有値に属する 2 本の固有ベクトル [を ϕ として $\phi = W\varphi$ で

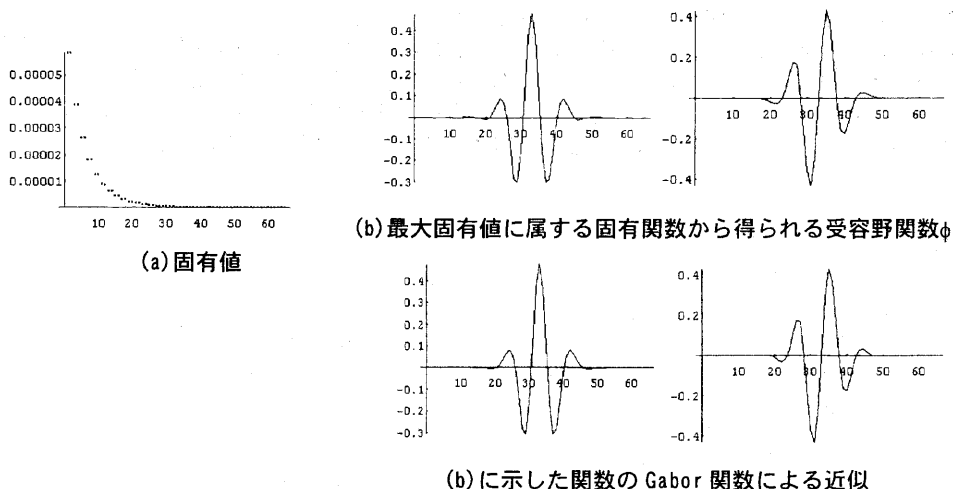


図 4.1 最適受容野関数の数値解

与えられる受容野関数]として図に示すように、対称型(cos型)および反対称型(sin型)の Gabor 関数(様の関数)が確かに得られることが分かった。[図 4.1 では信号 f としては DoG フィルターを通過させ直流成分を除いた無相関ノイズを、窓関数 w としては Gaussian 窓を用いた。しかし結果はこれらの条件の詳細には依存しない。条件を変えても同様の結果が得られる]。

○画像の統計がガウシアンでない場合

ここまでの議論では、信号 f がガウシアン統計で記述されることを仮定していた。画像の統計がガウシアンでない場合には、エントロピーの計算は一般には容易ではないが、その場合でも S/N 比が小さい場合には、前節までに述べた議論がやはり成り立ち Gabor 関数が最適受容野関数として得られることが分かる。

○Gabor 関数のセット

ここまでの考察から、パワースペクトルがピークを持つ時、そのピーク周波数 k_0 を中心周波数とする Gabor 関数が(比較的広い条件下で)最適な受容野関数となることが分かった。

ところで、こうして得られた Gabor 関数は実空間においても局在しているので、実際にはさまざまな位置 X を中心位置とする Gabor 関数を用意して、さまざまな位置で特徴を測定する必要がある。すると、この結果、全体として像 f から k_0 付近の周波数成分が抽出されて、残った周波数成分のパワースペクトルは元のパワースペクトル p から k_0 付近の成分が「掘られた」形状を持つことになる。このパワースペクトルを持つ信号を新たに入力信号と考えて同様の操作を行うこと

によって、新たなパワースペクトルのピーク周波数 k_0 を中心周波数とする新しい Gabor 関数のセットが得られることになる。同様な操作を続けていくことによって、さまざまな中心周波数を持つ Gabor 関数のセットが得られることになる。

5. おわりに

相互情報量の最大化 (Infomax) という基準から、比較的広い条件下で Gabor 関数が最適受容野関数として導出されることを示した。主要な仮定は、①局所的な受容野関数に焦点を絞り、その範囲で最適受容野関数を決めること、②信号の統計に対して並進対称性が成り立つことの2点である。さらに、計算の都合上 S/N 比が小さい場合を仮定した。

神経生理学の問題としては2次元の場合が重要である。しかし、画像を記述する統計が等方的な場合、本稿で考察した3節、(A)の評価基準からは異方的な2次元の Gabor 関数は導かれないことが分かる。その場合でも3節、(B)の相互情報量の基準からは、(S/N 比が小さい極限で)やはり Gabor 関数が最適受容野関数として導出されることを示すことができる。

文献

- [1]Daugman, J.G. (1985) J. Opt Soc. Am. A2, 1160-1169
- [2]Okajima, K (1995) Extnded Abstracts of the 5th International Symposium on BMED, Okinawa
- [3]Okajima, K (1996) The 2nd International Brainware Workshop, Tokyo
- [4]Atick, J (1992) in "Princeton lectures on biophysics", W. Bialek Ed., World Scientific, Singapore
- [5]Linsker, R (1993) Advances in Neural Information Processing Systems 5, 953-960:Linsker, R (1988) Computer 21, 105-117
- [6]Bialek, W. (1992) in "Princeton lectures on biophysics", W. Bialek Ed., World Scientific, Singapore
- [7]Daugmann, JG (1989) IEEE BME36, 107-114
- [8]Tanaka, S (1990) NEC R&D, 98, 1-14
- [9]Kohonen, T (1994) Self-Organizing Feature Map, Springer-Verlag, New York
- [10]M. Miyashita and S. Tanaka: NeuroReport 3(1992)69
- [11]ランダウ、リフシッツ、量子力学1 (日本語版) p54、東京図書
- [12]有本卓(1986)、カルマン・フィルター、産業図書:有本卓(1985)、情報理論、共立出版