



微小レンズアレイによるマッチトフィルタリングと ニューラルネットワークを用いた文字認識

山口 茂*・龜丸 俊一**

茨城大学 *大学院工学研究科精密工学専攻, **工学部システム工学科
〒316 日立市中成沢町 4-12-1

(1993年1月6日受付, 1993年4月16日受理)

Character Recognition by Matched Spatial Filtering Using a Microlens Array and Neural Networks

Shigeru YAMAGUCHI* and Shun-ichi KAMEMARU**

* Graduate Student, Department of Precision Engineering, Faculty of Engineering,

** Department of Systems Engineering, Faculty of Engineering,
Ibaraki University, 4-12-1, Naka-narusawa, Hitachi 316

(Received January 6, 1993; Accepted April 16, 1993)

A hybrid image recognition system by a microlens array (MLA) and neural networks is described in this paper for realization of a biological visual perception system. The system is composed of an optical processing system and a digital processing system. VanderLugt-type matched spatial filters (MSFs) by the MLA with 3.3 mm diameters, which are arrayed 4 by 3, are used for optical processing. The feature extraction technique is used in matched filtering. The neural networks based on back propagation algorithm is used for digital processing. Five characters such as D, K, O, X and a blank space are used for the targets to be recognized. In the results of the experiment, all input patterns are recognized by the system, which shows feasibility of the proposed system.

1. 緒 言

われわれ人間をはじめとした多くの生物が、視覚を通して目の前にある物体を認識するプロセスが非常にすぐれていることは周知の事実である。たとえば、目の前の物体あるいはパターンの一部が欠けていたり変形していた場合であっても、それらがこれまで見慣れた物であれば、われわれは極めて短時間でその対象を認識することが可能である。

そこでこれら生物のもつ、視覚を通しての認識処理をベースとしたシステムを、既存の光学系や電子回路・装置等を用いて実現することにより、われわれが有するに等しい認識能力をもつシステムの構築が期待できる。このシステムを実現するため、視覚を通じた認識に関する研究が心理学、生理学、工学などさまざまな領域で行わ

れてきた^{1,2)}。

しかし生物が視覚を通して行っている認識処理は、多数のニューロンを介した並列処理を基本とするものと考えられている。したがって、この認識システムを実際の機器・装置によって実現することは、現在のデジタル処理、すなわち逐次処理のみでは処理速度および取り扱う情報の量などの点から、困難であると考えられる。

一方、このデジタル処理のみに依存したシステム構築の限界に対し、現在着目されているのが光コンピューティングなどに代表される³⁾光情報処理を応用したシステムの構築である。光情報処理は二次元の画像情報を扱うことが可能であり、しかもその処理速度が高速であるという特徴を有している。

この光情報処理の中に、ある画像の中から所望のパターンの有無と位置を同時に検出可能とする処理として、

マッチトフィルタリングがある⁴⁾。このマッチトフィルタリングはディジタル数字やアルファベット文字の同時認識^{5,6)}、あるいは指紋の照合、珪藻の識別⁷⁾などに応用され、その報告がなされている。

また、光学システムの微小化、集積化そしてアライメントフリー化など、よりその実用性を考慮した分野にマイクロオプティクスがある。その中の重要なキーデバイスの一つとして、マイクロレンズアレイ (MLA) が注目されている⁸⁾。

そこでわれわれはマッチトフィルタリングを MLA により微小並列化し、これによって生体系での視覚による特徴抽出機構を光学的に実現可能であることを提案してきた^{9,10)}。しかしこれまで MLA によるマッチトフィルタリングに関する報告⁹⁻¹¹⁾においては、相関信号ピークを出力させた状態における認識結果が示されているのみであった。これはシステムの微小化にともなって生じるノイズの影響により、MLA によって得られた相関信号をさらに処理する手法が確立されず、提案されたシステムの認識結果とするまでには至っていなかったためである。

そこでわれわれはこれを改良するため、ピーク信号をデジタル処理する手法を提案した。さらにこの手法によって提案するシステムのハイブリッド化を可能にし、得られた結果をすでに報告している¹²⁾。

一方われわれはこれまで単レンズによるマッチトフィルタを用いたハイブリッド認識システムにおいて、光学処理系で出力された相関信号の分布から被認識物体を選択する方法を提案している¹³⁾。しかしこの手法は学習した被認識物体に対する理想的な相関信号の分布と、出力されたピーク信号の分布を単純に比較して認識とするものであった。そのため、外的の要因（光軸のわずかなズレ、入力未知物体上のノイズなど）により所望の相関信号が一部出力されない場合や、出力信号と同時に、強いレベルのノイズが生じた場合などは認識が不可能であった。

そこでわれわれは、相関信号の強度分布（これは、システムに入力された物体の特徴の分布と考えることができる）から被認識物体を識別する部分に、神経回路を模倣したニューラルネットワークを用いることが、提案する認識システムを構築する上で有効であると考えた。そして単レンズによるマッチトフィルタリングとニューラルネットワークを組み合わせたハイブリッド認識システムを用いることにより、入力文字のかけ、かすれ等に対する認識実験を行い良好な結果を得た¹⁴⁾。

本報告では、これまで構築されてきた MLA によるハイブリッドシステムに、生物が行っている視覚認識処理と同様の、入力像の分割、並列特徴抽出、特徴の統合という三段階の処理を適用する。その際にデジタル処理部にはニューラルネットワークを導入する。そして、空白を含む 5 文字に対し認識実験を行い、提案するシステムの構成の有効性と、ニューラルネットワーク導入による光処理部とデジタル処理部との整合性について検討を行ったので報告する。

2. 生物の視覚認識システム

生物の視覚認識システムのモデルは数多く提案されているが、いまだ解明されていない点も多い。そこでわれわれは、Fig. 1 に示すブロック図に基づく視覚認識システムの実現⁹⁾を方針としている。このブロック図は、入力物体の分割、部分パターンの特徴抽出、それらの特徴の統合、比較、認識の三工程から構成されている。

この処理の大部分はパラレルな処理であり、現在のシーケンシャル処理を基本とするコンピュータのみでは、その処理速度、一度に取り扱い可能な情報量等を考えるならば、理想とするシステムの実現は容易とは言えない。

そこでわれわれは、このシステムを光情報処理と情報の取扱いが簡便であるデジタル処理を組み合わせたハイブリッド処理により実現する。すなわち、入力物体の分割、並列特徴抽出は MLA によるマッチトフィルタリングにより光学的に実行し、これによって得られた部分的な特徴の統合、比較、認識はニューラルネットワークのアルゴリズムを有するデジタル処理より行う。この一連の処理による認識を、生物による実際の認識プロセスと比較し、その対比を Fig. 1 において B の部分に示す。

以上のコンセプトに基づき、本実験で用いる実際の認

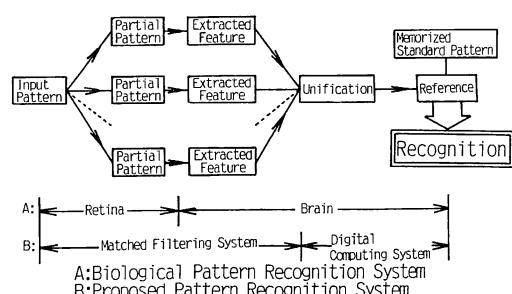


Fig. 1 Schematic diagram of a concept of a biological recognition system.

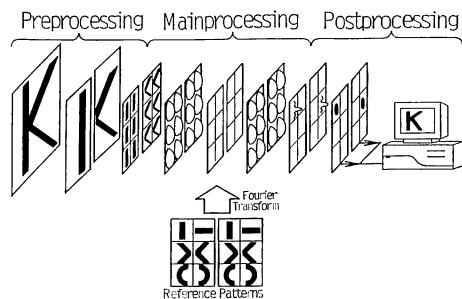


Fig. 2 Schematic diagram of biological visual perception.

識システムの概念図を Fig. 2 に示す。このシステムにより、文字「K」を認識する場合について説明する。

本システムはその処理によって分類した場合、入力物体の分割を行う preprocessing、各部分の特徴を抽出する mainprocessing、特徴の統合、比較、認識を行う postprocessing の三つの部分から構成される。

まず、前処理である preprocessing では、入力像を幾つかの部分要素に分割する。さらにこの分割に加えて、次段のマッチトフィルタへの入力のため、各々の部分要素について複製を行う。すなわち、「K」がシステムへ入力された場合、入力物体が左側の「|」と右側の「<」の2種類の部分要素に分割され、そして各々が縦3個×横2個に複製される。

続いて mainprocessing では、各複製されたパターンからその特徴を並列に抽出する。具体的には MLA によりマッチトフィルタを並列化し、複製されたパターンとフィルタに記録した参照パターンとの相関を実行させる。その結果、出力面には分割パターンと参照パターンの相関度に応じた強度を持つ相関信号が空間的に分布する。「K」を入力した場合、フィルタに記録された参照パターンに対応し、左側の「|」、右側の「<」、それぞれに自己相関信号が出力され、「K」固有の相関強度分布が得られる。

最後の postprocessing では、各入力物体に固有である相関信号の相関強度分布から、これまで学習した情報を基に入力物体の認識を行う。本処理ではこのアルゴリズムとして、学習が可能であり並列処理性を有するニューラルネットワークを使用した。この処理を行うため、mainprocessing の光学出力面のデータに対し二値化処理を施し、各微小レンズについて出力される相関信号の検出を行う。この相関信号分布をニューラルネットワークの入力層とし、学習を行ったネットワークにより入力物体を認識する。今回実験で使用したニューラルネット

ワークには、提案するシステム実現にあたっての導入段階として、数理工学的解析が比較的進んでおり工学的にも数多く応用されている、三層のバックプロパゲーション学習則を用いたモデルを使用した。

3. 実 験

今回は提案するシステムの有効性を示すため、Fig. 2 のシステムを光学素子とマイクロコンピュータ等によつて構築し、「D, K, O, X」の4文字と、1個の空白を含む計5物体に相当する入力に対し認識実験を行った。

本実験の preprocessing では、入力物体を左右に2分割、また各分割像に対する複製数を縦3×横2=6個とした。この処理は二重回折系と MLA を組み合わせた複製光学系¹⁵⁾とマスキングの組み合わせにより実現可能である。

MLA によるマッチトフィルタに記録する参照物体を Fig. 3 に示す。この参照パターンの決定手法について説明する。まず被認識物体「D, K, O, X」を左右に分割すると、これらの4文字を構成する基本パターンとして「|」、「<」、「>」、「(」、「)」の5パターンが得られる。この5種類の基本パターンにさらに多数の文字の基本となっている「-」を加え、それらを縦3個×横2個に配列する。これを一つの分割部分に対する参照物体とし、左右にこの参照パターンを配置する。これに相当する物体を Konica high resolution plate 上に透過物体として記録し、参照物体 (Fig. 3) とした。各パターンの寸法は乾板上で、高さ約 1 mm、間隔は MLA のレンズ間隔と等しくなっている。

Mainprocessing で用いるマッチトフィルタ作製光学系を Fig. 4 に示す。

ここで P₁ 面は参照物体面、P₂ 面はマッチトフィルタ面であり、それぞれに Fig. 3 の参照物体およびホログラム乾板 (Agfa 社製 8 E 75 HD) を配置する。MLA は

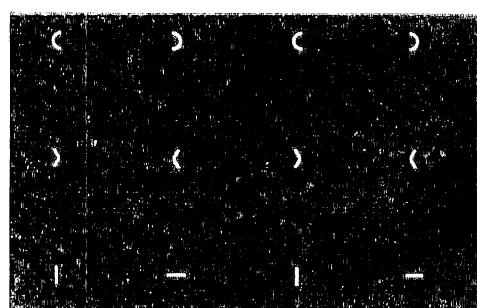


Fig. 3 Reference objects for synthesizing the MSF with an MLA.

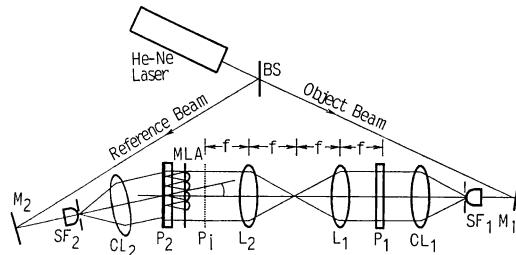


Fig. 4 A setup for synthesizing the MSF with the MLA.

焦点距離 3.65 mm、開口 3.3 mm の微小単レンズが、ピッチ 6.0 mm で縦 3 個 × 横 4 個に配置されている。また認識の際に問題となるノイズを相関信号から分離するため、各パターンは各微小レンズの光軸より約 0.7 mm オフセットして入力されている¹²⁾。

光源には 5 mW の He-Ne レーザーを用いた。このとき参照光の入射角度は物体光の光軸に対して $\theta = \text{約 } 5^\circ$ の角度であり、ホログラム乾板を裏面から照射する。この光学系によってフィルタを露光した後、この乾板に現像、漂白処理を施すことにより反射型マッチトフィルタを得る。

次に認識処理を行うためのシステムについて説明する。本実験では、現在のところ入力物体の preprocessing である分割・複製光学系の作製段階には至っていない。そこで入力物体はすでに preprocessing を施されていると仮定し、本実験での認識システムを mainprocessing 部分と postprocessing 部分により構成した。そのため、mainprocessing への入力物体には、「D, K, O, X」の文字に対しそれらの分割画像を作製し、これを用いた。すなわち入力が「K」である場合には、本システムの mainprocessing への入力物体は **Fig. 5** となる。この入力物体は、参照物体と同様の処理により乾板上に作製される。

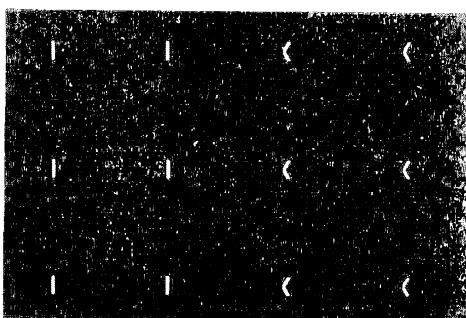


Fig. 5 An input pattern by division and duplication of the letter 'K' for mainprocessing.

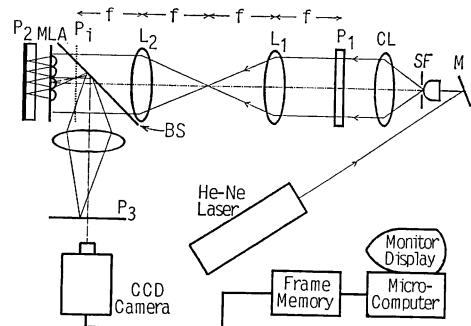


Fig. 6 A setup for the hybrid recognition system with the MLA and neural networks.

Fig. 6 に今回構築した認識システムを示し説明する。光学処理を実行する mainprocessing は He-Ne レーザーより P_3 面まで、デジタル処理を行う postprocessing は CCD カメラからコンピュータまでとなっていいる。

特徴抽出を行う mainprocessing と、そこでの処理について説明する。 P_1 面は入力物体面であり、分割、複製された像 (**Fig. 5**) を入力物体として挿入する。 P_2 面には先に作製したマッチトフィルタを配置し、入力物体と参照物体の相関信号を P_3 面で観測する。本システムでは特徴抽出を行う光学素子として反射型マッチトフィルタを採用した。そのため、 P_2 上面上のフィルタからの回折波はフィルタ上で反射し、MLA 中の同一レンズにより再度フーリエ変換される。その結果、 P_1 面に入力物体と参照物体の相関信号が出力される。この各レンズに対して出力された相関信号の分布をビームスプリッタ BS により検出し、 P_3 面で相関信号の強度およびその分布を観測する。

次に特徴の統合を行う postprocessing と、そこでの処理について説明する。 P_3 面での相関信号は CCD カメラにより 256 階調でフレームメモリへ取り込まれる。この信号に対し二値化処理を施し、ニューラルネットワークへの入力層とする。今回の実験では、バックプロパゲーション学習則による三層のニューラルネットワークを使用した。MLA に対するニューラルネットワークの関係を **Fig. 7** に示す。ここでネットワークの構成は、入力層を MLA の微小レンズの個数に等しい $4 \times 3 = 12$ セル、中間層を 12 セル、出力層を認識対象となる「D, K, O, X, 空白」のカテゴリの 5 セルとした。学習方法には「K」～「空白」の各学習パターンを 1 個入力するごとに結合係数を修正する逐次修正法を使用した¹⁶⁾。またネットワークの誤差の自乗和が十分減少するまで学習を

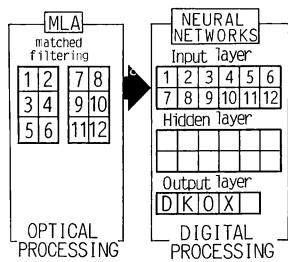


Fig. 7 The relationship between the MLA and an input layer of neural networks.

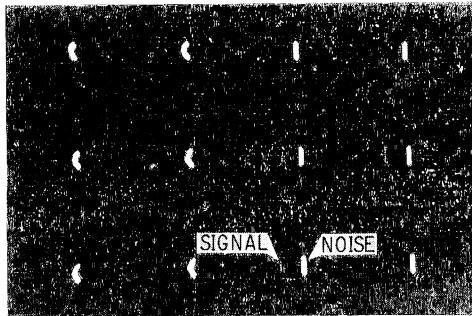


Fig. 8 Observed peak signals when a pattern 'K' is input.

繰り返し行った。

入力物体として Fig. 5 に示す文字「K」が入力された場合の mainprocessing における光学出力面 P_3 を Fig. 8 に示す。この出力面には、図示するように分割、複製された像に対して出力される相関信号とノイズが同時に出力されている。このノイズは、 P_1 面の入力物体が L_1 、 L_2 の二重回折系により全体が天地逆となった後、MLA により各パターンが天地逆となって生じているものである。

一方、所望の参照パターンと入力パターンの相関信号は、各ノイズの左側に出力されている。この相関信号とノイズの間隔はマッチトフィルタ記録時のオフセット値に依存している。

入力文字が「K」の場合、Fig. 7 に示す MLA の配列 No. 3 の位置に「<」の自己相関信号が、No. 11 の位置に「|」の自己相関信号が現れている。その他の位置に生じているのはすべて相互相関信号である。

この光学出力面を CCD カメラにより 256 階調のデータとしてフレームメモリへ取り込み、そのデータに対し適当なしきい値に基づいて二値化処理を施し、入力物体を認識した結果を Fig. 9 に示す。この画面は、システムによって最終的に認識結果を得るときの CRT の画面を記録したものである。ここで画面左側には二値化処理

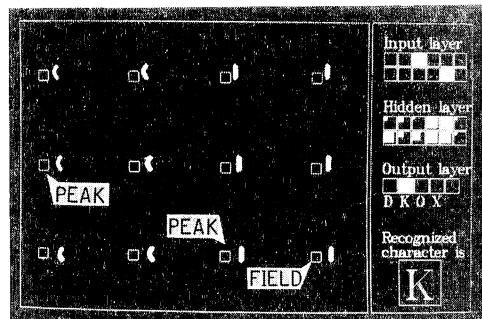


Fig. 9 Result of recognition when a pattern 'K' is input.

理された相関信号が、右側には認識結果が表示されている。二値化処理を施した画面には、各レンズの位置に相当する部分に四角の枠が設定されているが、これは相関信号を検出するための領域で、処理に先立ってあらかじめ設定されている。これをフィールドと呼ぶ。この場合、画面上で図示した No. 3 と No. 11 の位置のみにピーク信号が output されている。

次にニューラルネットワークでの処理について説明する。Fig. 9 における二値化の画面と Fig. 7 の関係に基づいてニューラルネットワークの入力層が形成される。Fig. 9 の画面右側は、ニューラルネットワークによる認識処理を表示している。上段が入力層、中段が中間層、下段が出力層である。

入力されたピーク信号の分布により、出力層では「K」のカテゴリに信号が集中している。本実験における処理として、出力が最大値の 70% 以上であれば入力物体に対する認識結果とするようプログラムを作成した。この実験結果の場合、入力物体が「K」であると表示され、その認識結果が正しいことがわかる。

他の入力として「D」が入力された場合のディジタル処理結果を Fig. 10 に示す。この場合には「D」を構成す

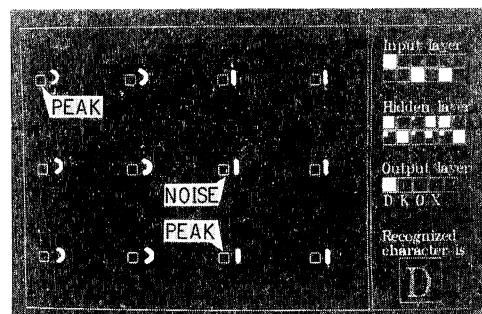


Fig. 10 Result of recognition with noise when input pattern is 'D'.

る No. 1 と No. 11 の位置の自己相関信号以外に、No. 9 の位置に「|」に対する「<」の相互相関信号がノイズとして現れている。ニューラルネットワークを有しないシステムでは、光学処理で得られた相関ピークの信号分布を、理想的に得られる相関分布と比較するため、処理する信号が完全に一致する分布でなければ正しい認識結果が得られない。しかし本実験においてニューラルネットワークをディジタル処理部に用いることにより、Fig. 10 に示すノイズをともなった出力信号からも、入力された物体「D」を正しく認識することが可能となった。

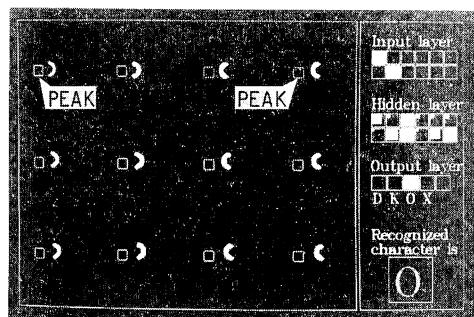


Fig. 11 Result of recognition with noise when input pattern is 'O.'

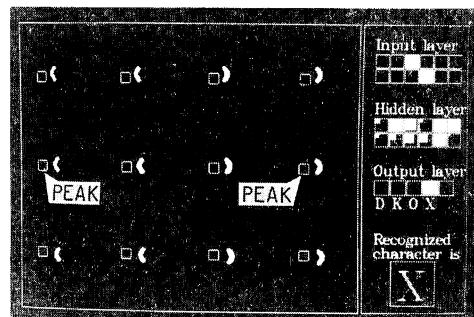


Fig. 12 Result of recognition with noise when input pattern is 'X.'

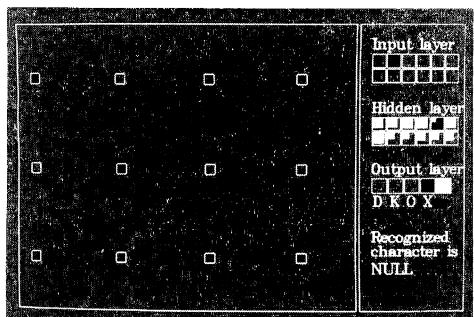


Fig. 13 Result of recognition when input pattern is a blank space.

同様に「O」、「X」、「空白」についてそれぞれ認識実験を行った結果、各入力に対し所望の自己相関信号分布が得られ、正しい認識結果が得られた。Fig. 11~13 にそれらの結果を示す。

4. 結 言

本研究では、生物の視覚認識システムが行っている入力物体の分割、複製、並列特徴抽出、特徴の統合、認識の一連の処理を、光学素子等によって実現すること目的とした。このため、 4×3 個の MLA によるマッチトフィルタリングとニューラルネットワークを用いたハイブリッドシステムを提案し、そのシステムによって認識実験を行った。今回その基礎実験として「D, K, O, X」そして「空白」の 5 文字に相当する入力を認識対象として実験を行った。その結果、対象としたすべての入力に対して正しい認識結果を得ることができた。これにより提案するシステムの有効性が確認できた。

またディジタル処理部にニューラルネットワークを導入したことにより、従来のシステムでは対応できないノイズを含む光学信号が得られた場合にも認識処理が可能になった。この結果、MLA による光学処理とニューラルネットワークの整合性が良く、これらを組み合わせてそれぞれの処理の特徴を生かしたシステムの構築が期待できることが明らかとなった。

今回提案したシステムによる認識実験では、分割方向が左右 2 分割、六つの特徴を抽出するシステムであり、認識対象も 5 文字であった。そこで今後認識対象を一般化した場合の最適な分割方法、およびその数、抽出する参照パターンの決定などが課題である。またマッチトフィルタに用いる参照パターンに関しては、物体の基本構成要素となるさまざまな角度を持つ線分要素を用いることが有効であると考えられ、次の段階の研究ではこれらを検討してゆきたい。

本研究を進めるに当たりレンズを提供していただいた及川正尋博士（日本板硝子（株））に深く感謝します。

文 献

- 1) 外山敬介、甘利俊一、金子隆芳、村松正実、和田英一：「生体における情報処理」、南雲仁一編（岩波書店、東京、1982）。
- 2) 福島邦彦：“ネオコグニトロン”，Computrol, No. 24 (1988) 61-69.
- 3) 北山研一：“光コンピューティングシステム”，光学, 20 (1991) 657-663.
- 4) A. VanderLugt: "Signal detection by complex spatial filtering," IEEE Trans. Inform. Theory, IT-10 (1964) 139-145.
- 5) S. Kamemaru, J. Yano and H. Itoh: "Matched

- spatial filtering by feature extracted reference patterns using cross-correlated signals," Proc. SPIE, **1564** (1991) 143-154.
- 6) 矢野潤一, 龍丸俊一: "特徴抽出された参照パターンによるマッチトフィルタとそれを用いた文字認識", 光学, **21** (1992) 319-326.
- 7) S.P. Almeida and J.K. -T. Eu: "Water pollution monitoring using matched spatial filters," Appl. Opt., **15** (1976) 510-515.
- 8) 伊賀健一: "微小光学の展望", 微小光学の物理的基礎, 応用物理学学会日本光学会編 (朝倉書店, 東京, 1991) pp. 1-16.
- 9) M. Agu, A. Akiba and S. Kamemaru: "A parallel-processing optical-digital recognition system as a model of biological visual perception," Opt. Commun., **66** (1988) 69-73.
- 10) M. Agu, A. Akiba, T. Mochizuki and S. Kamemaru: "Multimatched filtering using a microlens array for an optical-neural pattern recognition system," Appl. Opt., **29** (1990) 4087-4091.
- 11) P. Asthana, A. Akiba, T. Yamaki, K. Nishizawa and M. Oikawa: "Graded-index lens array matched filtering," Opt. Lett., **13** (1988) 84-86.
- 12) S. Kamemaru, S. Yamaguchi and J. Yano: "Fabrication of a biological visual perception system using a microlens array in a hybrid pattern recognition system," Jpn. J. Appl. Phys., **31** (1992) 1682-1688.
- 13) S. Kamemaru, T. Nagaoka, M. Kakuta and I. Shimizu: "Multiple matched spatial filtering with corresponding symbols by the hybrid system," Jpn. J. Appl. Phys., **29** (1990) L1557-L1560.
- 14) 馬場 勉, 山口 茂, 矢野潤一, 龍丸俊一: "マッチトフィルタとニューラルネットワークを用いたパターン認識システム(I)", 第39回応用物理学関係連合講演会講演予稿集第3分冊, 28a-SC-6 (1992) p. 795.
- 15) K. Hamanaka, H. Nemoto, M. Oikawa, E. Okuda and T. Kishimoto: "Multiple imaging and multiple fourier transformation using planar microlens arrays," Appl. Opt., **29** (1990) 4064-4070.
- 16) 中野 騒, ニューランネットグループ, 桐谷 滋: "バックプロパゲーション", 入門と実習ニューロコンピュータ, 飯沼一元編 (技術評論社, 東京, 1989) pp. 28-84.