Neural-Network Algorithms for Data Processing
in an Adaptive Optical System for Wave-Front Correction

Andrei B. Utkin

Institute for Laser Physics, Scientific Center S.I. Vavilov State Optics Institute,
Birzhevaya 12, St.Petersburg 199034, Russia, Tel.: (812)-218-9982, fax: (812)-218-1093

Abstract

The possibility of application of a neural network to data processing in an adaptive optical system is investigated. In case of a linear input/output transformation, the simple two-layer feedforward neuron structure is shown to be a good choice for heteroassociative reconstruction of the driving signals from the current data on the phase distribution. Several learning algorithms for the neural net are presented and the opportunity to construct the training set with the help of experiment is discussed.

1. Introduction

Correction of the wave-front distortions using adaptive optics technique, see Fig.1., typically requires a wave-front sensor (the Hartmann sensor, shearing interferometer, etc.) and an active optical element (for example, deformable mirror). In addition, the data-processing system is needed to convert the array of data on the current beam quality $\vec{g} = \{g_1, g_2, \ldots, g_N\}$ (in case of the Hartmann sensor it is the vector of two transverse components of the phase gradient, $\partial \varphi / \partial x$ and $\partial \varphi / \partial y$, detected in $N/2$ sample points of the beam cross-section) into the vector of driving signals $\vec{u} = \{u_1, u_2, \ldots, u_Q\}$ applied to $Q$ control points (actuators) of the active element.

Precise wave-front control requires several sample points and actuators, that results in complex systems and, consequently, in intricate algorithms for the wave-front correction (i.e. derivation of $\vec{u}$ from $\vec{g}$). The requirements of hardware simplicity and reliability from one side and high efficiency from another side are hardly compatible for traditional systems. One of the novel approaches to the problem of the active element control is the introduction of the data-processing system based on an artificial neural network. This exploratory study is to call attention to the relatively simple class of neural networks which may be successfully adapted to the above optical trains for the wave-front correction.

The two components required for definition of a neural network are processing nodes, or neurons, and links. The former are represented in algorithmic descriptions by their response function $f$. Response function for the $i$ -th neuron represents the dependency between signals on its input $s_i$ and output $a_i$ (the latter is also called the state of the neuron)

$$a_i = f \left( s_i \right).$$  \hspace{1cm} (1)
This function is supposed to be the same for all neurons of the network. Typical response functions of so-called binary, bipolar, and continuous types are shown in Fig.2.

Links between neurons are defined by neural-network structure. They are assumed to be linear, i.e. independent from the level of transmitted signal, so the link which transmits signal from the \( j \) -th to \( i \) -th neuron can be characterized by one scalar value – the interconnection weight \( T_{ij} \). Since the system is implemented in the adaptive optical train for direct heteroassociative derivation of the driving signal \( \vec{u} \) from the input data \( \vec{g} \), a natural choice for the network architecture is multi-layer feedforward structure (Fig.3a) consisting of several layers of neurons. The simplest feedforward network implementation is two-layer structure which consists of the input and output layers only (Fig.3b). According to the neural-network approach, where is no need to know the underlying algorithm of input/output mapping. Desired \( \vec{g} \rightarrow \vec{u} \) transformation must be learned from examples which form the training set \( ( \vec{G}_1, \vec{U}_1 ), ( \vec{G}_2, \vec{U}_2 ), ..., ( \vec{G}_M, \vec{U}_M ) \). Each of \( M \) elements of the set represents a pair consisting of input signal from the wave-front sensor \( \vec{G}_M \) and corresponding driving signal \( \vec{U}_M \) for the best phase correction. One should adjust the interconnection weights \( T_{ij} \) so what the network derives output training signal \( \vec{U}_M \) for every input pattern \( \vec{G}_M \). This process can be considered as writing the training set into the internal network memory, so the matrix of the interconnection weights is also called memory matrix.

Since ones of the simplest neurons are those of bipolar type, the less intricate neural network for heteroassociation is two layers of bipolar neurons. Each neuron may be only in one of two possible states \( ( a_i = -1 \) or \( a_i = +1 \) ) and can encode only one bit of input/output information, so the digital data representation is required. Unfortunately, there are two unsurmountable drawbacks which make this simple network unusable for us. First, the storage capacity of such a system \( SC \) is not greater than \( \min(N_i,N_o) \) pairs of input/output patterns (here \( N_i \) and \( N_o \) are numbers of neurons in the input and output layers respectively) while for the exact reproduction of arbitrary \( \vec{g} \rightarrow \vec{u} \) mapping it is necessary to store all possible input signals (together with associated driving signals) whose number \( M \) may be defined at once from combinatorics as \( 2^{N_i} \). To increase the storage capacity we are forced to introduce the additional neurons into the scheme, which are not involved in the input/output operations. These neurons form so-called hidden layers, each of which extends the storage capacity at the expense of increasing neural-network complexity (Fig.3a). In our case, however, this cannot be treated as a good remedy because of the second drawback of the system that is combinatoric explosion in the number of possible input patterns \( \vec{G}_M \). The number of neurons in the input layer \( N_i \) must be big enough even for rather rough wave front description. For example, if we use the Hartmann sensor with matrix of only \( 3 \times 3 \) primary detectors and only 7-bit code for each of \( N = 2 \times 9 = 18 \) data entries (1 for sign and 6 for magnitude), we need \( N = 7 \times 18 = 126 \) neurons in the input layer for binary representation of \( \vec{g} \). So our network should be able to operate over the field of \( M = 2^{N_i} = 2^{126} \approx 8.5 \times 10^{37} \) possible input images. This cause the introduction of at least \( M \) hidden neurons. One can hardly imagine the real dimensions of such a system and the realization of learning algorithm with the training set of such a length. The problem is
very close to one of gray-level discrete associative memory design\textsuperscript{1,2}. The trouble is that we have too many «gray» levels and, which is more important, no similarity between binary images of kindred phase distributions.

It is possible to circumvent the above impediments using neurons with continuous response function (Fig.2c.). The state of a neuron of such type provides us the information both for sign and magnitude of the input signal and we can restrict the network structure to one neuron per one component of the input/output signal, \( N_i = N \) and \( N_o = Q \). Then the \( j \)-th component of the signal from the wave-front sensor is input to the \( j \)-th neuron and its output signal becomes \( a_j = f(g_j) \). Transported by the link \( j \rightarrow i \), it causes the partial signal \( T_{ij}a_j \) in the input of the \( i \)-th neuron of the output layer. The total signal in the input of this neuron is

\[
s_i = \sum_{j=1}^{N} T_{ij}a_j = \sum_{j=1}^{N} T_{ij} f(g_j) \quad (2)
\]

and one obtains for the output pattern

\[
u_i = f(s_i) = f \left( \sum_{j=1}^{N} T_{ij} f(g_j) \right), \quad i = 1, 2, \ldots, Q. \quad (3)
\]

Pioneer investigations of a neural-net data-processing system for an on-line adaptive optical train with atmospheric wavefront sensing are reported by Lloyd-Hart \textit{et al.} in their recent paper\textsuperscript{3}. In place of traditional wave-front sensors, they use direct input of a pair of in-focus and out-of-focus pixel images of the wave front to be corrected. This approach result in substantially non-linear input-to-output signal mapping and usage of a complex multi-layer feedforward structure of neurons with a sigmoid threshold function. Notably, if traditional wave-front sensors are used, most applications lead to the linear \( \vec{g} \rightarrow \vec{u} \) transformations. For these important particular cases we can restrict ourselves to the neurons with linear response functions, say, \( f(s) = s \). Being very efficient, the linear-neuron nets are very simple in point of architecture and learning algorithms. Below we discuss them in details.

2. Neuron Nets for Linear Transformations

**Basic relations**

For the case \( f(s) = s \), one gets from (3) \( u_i = \sum_{j=1}^{N} T_{ij}g_j, i = 1, 2, \ldots, Q \), or, in matrix notation, \( \vec{u} = T \vec{g} \). Being intended for reproduction of linear dependences, this neural-network algorithm for derivation of the output signal exhibits the linear degeneracy: if we have defined the network responses to the training input patterns \( \vec{G}_1 \) and \( \vec{G}_2 \) as \( \vec{U}_1 \) and \( \vec{U}_2 \), then the response of the system to an arbitrary linear combination \( \alpha_1 \vec{G}_1 + \alpha_2 \vec{G}_2 \) is also defined as \( \alpha_1 \vec{U}_1 + \alpha_2 \vec{U}_2 \). Thus, we can restrict the training set to the series of linearly independent input patterns \( \vec{G}_m \) corresponding to some \( M \) basic wave-front aberrations (for example, to the phase distributions proportionate to Zernike polynomials). After the choice of the training set \( (\vec{G}_m, \vec{U}_m) \), the learning rule leads at ones to

\[
T \vec{G}_m = \vec{U}_m, \quad m = 1, 2, \ldots, M \quad (4)
\]

This linear variant of neural-network structure is also called heteroassociative memory\textsuperscript{4} and the most common notation for (4) is

\[
TG = U \quad (5)
\]
where $G = \{ \vec{G}_1, \vec{G}_2, \ldots, \vec{G}_M \}$ is the matrix of size $N \times M$ with the $\vec{G}_m$ as its columns and $U = \{ \vec{U}_1, \vec{U}_2, \ldots, \vec{U}_M \}$ is the matrix with $Q$ rows and $M$ columns which are the vectors of the driving signals $\vec{U}_m$ from the training set.

Since the dimensionality of the matrix $T$ is $Q \times N$, the conditions (4) are equivalent to the $MQ$ scalar equations with $NQ$ unknown interconnection weights. Thus, the definition of $T_{ij}$ is trivial for $M = N$ only.

Overdetermined conditions

Then the length of the training set is greater than the number of neurons in the input layer, $M > N$, (5) represents an overdetermined system of equations, wherein are more equations than unknowns, and the exact solution does not exist. One can use the best fit to $U$ in the least-square sense that is given by the equation

$$T = U \, G^T \, (G \, G^T)^{-1}$$

(6)

where the superscript denotes the operation of transposition.

We should realize that, owing to the limited storage capacity, the system in question fails to process correctly more than $N$ types of linearly independent aberrations. The solution (6) is not more than the best possible.

Underdetermined conditions

Suppose we know from the underlying physics that the incident primary beam of radiation may be actually affected by $M$ types of linearly independent aberrations corresponding to $M$ definite transverse phase distributions $\Phi_1(x,y), \Phi_2(x,y), \ldots, \Phi_M(x,y)$, i.e. every phase perturbation $\varphi(x,y)$ of the beam can be represented with enough accuracy as a linear combination

$$\varphi(x,y) = \sum_{m=1}^{M} \alpha_m \, \Phi_m(x,y).$$

(7)

In point of mathematics, for exact correction of any phase perturbation of type (7) one needs the wave-front sensor with $N = M$ sampling points and, correspondingly, $N_i = M$ neurons in the input layer. Thus, we obtain $M$ input signals $\vec{G}_1, \vec{G}_2, \ldots, \vec{G}_M$ of dimensionality $N = M$ for the training set. They can be made linearly independent by the appropriate choice of the sampling points. Then the equation (5), containing the square matrix $G$ of full rank, have the unique exact solution

$$T = U \, G^{-1}$$

(8)

for the memory matrix $T$ of dimensionality $Q \times M$.

In practice, we have to detect noise signals and more than $M$ sample points are needed to provide reliable information on the wave-front distortion, which is the unknown mixture of $M$ standard aberrations. Hereupon we will consider more actual case $N > M$ that corresponds to the underdetermined matrix equation (5) with more unknowns than equations, where $Q$ is the nonsquare matrix with linearly independent columns. For $N > M$ Eq.(5) has a family of exact solutions

$$T = U \, G^+ + Y + Y \, G \, G^+$$

(9)

where $Y$ is an arbitrary matrix with the same dimensions as $T$ and

$$G^+ = (G^T \, G)^{-1} \, G^T$$

(10)

denotes the pseudoinverse matrix for the matrix $G$ with linearly independent columns. The solution
\[ T = U G^+ = U (G^T G)^{-1} G^T, \]  

(11)
corresponding to \( Y = 0 \), stands out as one having the minimum norm. Below we discuss some of its properties.

3. Minimum Norm Solution

Choice of the training set, precision, and stability

The minimum norm solution (11) is said to be ill-conditioned if the elements of the matrix of interconnection weights \( T_{ij} \) are very sensitive to small changes in the components of the original equation (5). The analytical form of the solution enables us to determine at once that the only possible reason of instability is the inversion of the matrix

\[ V = G^T G. \]  

(12)

In fact, since the columns of \( G \), i.e. \( \vec{G}_1, \vec{G}_2, \ldots, \vec{G}_M \), are real (physical) signals, we may not worry about the possibility of their linear dependency which for \( M < N \) requires exact relationship between \( \vec{G}_m \) and hence equals zero. Matrix \( G \) has the rank \( r(G) = M \) and so the square matrix \( V \) of size \( M \times M \) is of full rank. Nevertheless it may happen that \( \vec{G}_m \) are nearly linearly dependent and, consequently, \( V \) is nearly singular. This can be characterized quantitatively by the invariant measure of inherent instability of the inversion which is the condition number \( CN(V) = \|V\| \cdot \|V^{-1}\| \) where the matrix norm definition corresponds to the Euclidean vector's norm \( \|\vec{g}\| = (\vec{g}, \vec{g})^{1/2} \)

\[ \|V\| = \max_{\vec{g} \neq 0} \frac{\|GV\|}{\|\vec{g}\|}. \]  

(13)

Since \( V \) is a symmetrical matrix, \( V^T = (G^T G)^T = G^T (G^T)^T = G^T G = V \), \( CN(V) \) can be expressed as [7]

\[ CN(V) = \frac{|\lambda_{\max}(V)|}{|\lambda_{\min}(V)|}. \]  

(14)

Here \( \lambda_{\max}(V) \) and \( \lambda_{\min}(V) \) stand for the maximum and minimum eigenvalues of the matrix \( V \). It has been universally accepted that the eigenvalues may be allowed about 3 orders of magnitude variation without significantly amplifying the instability [7,8]. For bigger variation one has to change the vectors \( \vec{G}_m \) in the training set (by changing the basic aberrations or the sampling locations, or both).

Using the definition (12), one can obtain for the \( kn \)-th element of the matrix \( V \)

\[ V_{kn} = \sum_{j=1}^{N} (G^T)_{kj} G_{jn} = \sum_{j=1}^{N} G_{jk} G_{jn} = (\vec{G}_k, \vec{G}_n), \]  

(15)

where the inner product \( (\vec{G}_k, \vec{G}_n) \) can be treated as the measure of correlation between \( \vec{G}_k \) and \( \vec{G}_n \). If using the appropriate orthogonalization and scale procedures we can choose the basic set of aberrations \( \Phi_m \) in (7) so that generated \( \vec{G}_m \) are orthogonal to each other and have the same norm \( \gamma \),

\[ (\vec{G}_k, \vec{G}_n) = \delta_{kn} \gamma^2 = \begin{cases} \gamma^2 & \text{for } k = n, \\ 0 & \text{for } k \neq n. \end{cases} \]  

(16)
then the matrix \( V \) becomes proportionate to the identity matrix \( I \), \( V = \gamma^2 I \). In this case we obtain the best conditions for stability, \( \lambda_{\max}(V) = \lambda_{\min}(V) = \gamma^2 \), \( CN(V) = 1 \), and the minimum norm solution yields \( T = \gamma^{-2} U G^T \), or
\[ T_{ij} = \gamma^{-2} \sum_{m=1}^{M} U_{im} (G^T)_{mj} = \gamma^{-2} \sum_{m=1}^{M} U_{im} G_{jm} \]  

where the index \( i = 1, 2, \ldots, Q \) enumerates actuators (or components of the driving signal) while the index \( j = 1, 2, \ldots, N \) enumerates the sample points (or components of the array of data on the phase distribution), \( U_{im} \) is the \( i \)-th component of the \( m \)-th output signal \( \hat{U}_m \) of the training set, and \( G_{jm} \) is the \( j \)-th component of the corresponding input pattern \( \hat{G}_m \).  

**Decision capabilities and noise reduction**  

When the above two-layer network with the minimum-norm-solution memory matrix \( (11) \) is used for data processing, the input signal \( \vec{g} \) from the wave-front sensor can be represented as a sum

\[ \vec{g} = \vec{g}_V + \vec{g}_{ER} \]  

of the valid information on the wave-front distortions \( \vec{g}_V \) and an erroneous signal \( \vec{g}_{ER} \) resulted from noise (dynamic error) and system dysfunction (static error).  

Any arbitrary vector of the erroneous signal can be uniquely decomposed into two components

\[ \vec{g}_{ER} = \vec{g}_{ER\|} + \vec{g}_{ER\perp} \]  

of which one, \( \vec{g}_{ER\|} \), belongs to the linear \( M \)-dimensional subspace \( \mathcal{I} \) spanned by the vectors of the training set \( \hat{G}_1, \hat{G}_2, \ldots, \hat{G}_M \) and another, \( \vec{g}_{ER\perp} \), is orthogonal to \( \mathcal{I} \) (i.e. \( \vec{g}_{ER\|} \in \mathcal{I} \) and \( \vec{g}_{ER\perp} \perp \mathcal{I} \)).  

The first component of the erroneous signal \( \vec{g}_{ER\|} \) corresponds to the verisimilar part of the error which is a linear combination of \( \hat{G}_m \) and thus cannot be separated from the valid signal \( \vec{g}_V \). We will show, however, that our network system is «smart» enough to throw off the last term in Eq. (19) as an obvious error.  

One can check by direct calculations that the operator of orthogonal projection \( P_\mathcal{I} \) of a vector \( \vec{g} \) on the subspace \( \mathcal{I} \) spanned by the vectors \( \hat{G}_m \), \( P_\mathcal{I} \vec{g} = \vec{g}_\| \), can be expressed via the matrix \( G \) as

\[ P_\mathcal{I} = G (G^TG)^{-1}G^T = GG^+ \]  

Using the above consideration, one obtains \( P_\mathcal{I} \vec{g}_{ER\|} = \vec{g}_{ER\|} \), \( P_\mathcal{I} \vec{g}_{ER\perp} = 0 \). The output driving signal

\[ \vec{u} = T \vec{g} = T \vec{g}_V + T \vec{g}_{ER} = T \vec{g}_V + T \vec{g}_{ER\|} + T \vec{g}_{ER\perp} = \vec{u}_V + \vec{u}_{ER\|} + \vec{u}_{ER\perp} \]  

consists of three parts: the valid driving signal \( \vec{u}_V \) and two erroneous signals \( \vec{u}_{ER\|} \) and \( \vec{u}_{ER\perp} \) resulted from the errors of two above mentioned kinds. Let us calculate the contribution of the obvious input error. Using Eq. (11) we get

\[ \vec{u}_{ER\perp} = T \vec{g}_{ER\perp} = U G^+ \vec{g}_{ER\perp} \]  

Since \( G^+ G = (G^TG)^{-1}G^TG = (G^TG)^{-1}(G^TG) = I \) is the identity matrix of size \( M \times M \),

\[ \vec{u}_{ER\perp} = U (G^+ G)^+ \vec{g}_{ER\perp} = (UG^+)(GG^+) \vec{g}_{ER\perp} = (UG^+) P_\mathcal{I} \vec{g}_{ER\perp} = 0 \]  

Thus, only the \( M \)-dimensional part of the erroneous input signal, corresponding to the orthogonal projection on the subspace \( \mathcal{I} \), brings about non-zero output error. Other \( N-M \) components of the noise and static error are thrown off, so the less the ratio \( M/N \), the better the noise performance and the static error correction (or at least there is a reasonable assurance that this will occur).
An extensive investigation of the heteroassociative system's noise performance of nearly the same type has been made by Cassasent and Tefler\textsuperscript{4}. Their results indicate that computer simulation is the best way of getting the noise characteristics of the system because: 1) the type of output vector encoding has a significant effect on performance and 2) it is difficult to obtain analytical estimates for the input/output noise variances without statistical approach to the elements of the heteroassociative memory matrix $T$ and without rather bold suggestions about their distribution (see, for example, Appendix D of Ref. [4]).

4. Scheme of experimental definition of the training signals

For traditional adaptive optical schemes the description of desired $\mathbf{g} \rightarrow \mathbf{u}$ mapping brings about a close inspection of the performance of the wave-front sensor and the adaptive optical element. As it was mentioned in the Introduction, for the neural-network implementation the problem of input/output mapping reduces to the choice of the training set. This enables us to obtain the most reliable training set with the help of experiment. When all preliminary theoretical results have been obtained and tested and the appropriate basic aberrations $\Phi_m(x,y)$ have chosen, it is possible to measure the corresponding input vectors $\mathbf{G}_{jm}$ by registration of the set of standard beams with the transverse phase distributions $\varphi(x,y) = \Phi_m(x,y)$. Moreover, applying «manually» different driving signals $\mathbf{u}$ to the actuators and registering by another wave-front sensor the quality of the improved standard beams, we can find the best combinations $\mathbf{u}^* = \mathbf{U}_m$ (see Fig.4.). After that one has the training set which automatically include the proper description of some dysfunctions of the above two elements of the scheme (for example, irregularities in characteristics of different primary wave-front detectors and different actuators of the active element).

5. Summary

Some possible techniques for neural-network data processing in adaptive optical systems have been discussed. It has been shown that the neurons with continuous response function result in much less complicated system than binary ones, although they are more difficult to realize.

For linear input-output transformations, the basic learning rules for two-layer feedforward structure have been considered. We have investigated the best solution for the memory matrix having the minimum norm. It is possible to formulate some criterion of the numerical stability of this solution, so one has the opportunity to estimate the reliability of the system at the design stage. The orthonormal (in terms of the condition (16)) set of the input training vectors turn out to be the best choice. The minimum-norm-solution system has the unique capability to tell the input signals which are obviously erroneous and throw them off, that is very important when the entire system operates in autonomous conditions (for example, in outer space) and one has no opportunities to correct unexpected dysfunctions.

Neural-network algorithms enable us to define both input and output training signals experimentally and to obtain the description of the input/output mapping that is the best match for a particular hardware. This advantage can be realized in traditional computer-controlled adaptive optical systems, where it is possible to emulate a neural network.

Fig.4. Conceptual sketch of the experiment for measurement of the training signals $\mathbf{G}_{jm}$ and $\mathbf{U}_{jm}$.

1 - wave-front sensor; 2 - active optical element with actuators; 3 - data-processing system (not used); 4 - beam splitter; 5 - initial beam with the plane wave front; 6 - phase modulator for generation of standard phase distribution $\Phi_m(x,y)$; 7 - wave-front quality detector; 8 - autonomous source of the driving signals.
References


