Chaotic Modeling and Simulation (CMSIM) 1: 133-146 , 2013

Nonlinear Analysis of the Hopfield Network Dynamical States Using Matrix Decomposition Theory

Alexander M. Krot and Ryhor A. Prakapovich

United Institute of Informatics Problems of National Academy of Sciences of Belarus, Laboratory of Self-Organization System Modeling, Minsk, Belarus
(E-mail: alxkrot@newman.bas-net.by, rprakapovich@robotics.by)

Abstract: Nonlinear analysis of dynamical states based on matrix decomposition theory for the Hopfield's neural network is developed in this paper. A formula for determining the values of approximating vector function increments for any number of network neurons is derived. The procedure of noisy monochrome image restoration using Hopfield's network is simulated here.
Keywords: Hopfield's neural network, Nonlinear dynamical systems, State space, Attractors, Matrix decomposition theory.

1 Introduction

Hopfield's neural networks have been widely used as a simple and intuitive understanding of the associative memory model. This follows from the fact that artificial neural networks (ANNs) have some similar characteristics of the human associative memory, namely: 1) information retrieval is carried out not by means of a memory address supply but through a data measure determining the similarity with the standard pattern; 2) data distribution of the stored patterns are located throughout the memory space; 3) data access to memory space are represented by a dynamical process. Due to the stable states of Hopfield's AAN correspond to local minimum of the Hopfield's energy function, they have well been used to solve various optimization problems [2].

Despite its advantages, the Hopfield's ANNs have several drawbacks. These include a small memory capacity for the stored standard patterns and higher sensitivity to the correlation between the input patterns. For example, in [1] it has been experimentally proved that the memory tends to 0,15N where N is a number of neurons in the network. However, in the paper [3] it has been shown that the number of stored patterns can not exceed $N/4\log N$, besides the memory capacity is decreased sharply in case of correlation between the stored reference patterns.

It should be noted that these formulas are not very effective for practical application. For example, it is impossible to determine the input pattern affecting on the behavior of all trained network. In this connection an investigation of dynamical states of Hopfield's ANN is very important problem.

Received: 19 March 2012 / Accepted: 10 September 2012

ISSN 2241-0503

Generally speaking, methods of nonlinear dynamics including calculation of minimal attractor embedding dimension, the Lyapunov characteristic exponents etc. are basic tools for characterizing behavior of complex systems [4], [5], [6]. To quantify more exactly a dynamics of complex system quantitatively more exactly these methods have to take into account higher order nonlinearities. In the papers [7-14] higher order nonlinearities have been described by means of matrix series in state space of complex system. Entirely, decomposition methods of nonlinear operators describing the behavior of system in state space (phase space) are very important for analysis, identification and modeling of nonlinear dynamical systems, especially complex nonlinear dynamical systems [5], [6].

In this context, the purpose of this paper is to study the behavior of the Hopfield network based on the developed in [7-14] nonlinear analysis methods for attractors of complex dynamical systems. This paper investigates the stability of the convergence of retrieval binary vectors processes using the matrix series expansion theory [7-14].

2 Analysis of the Hopfield' ANN Dynamics on the Basis of the Matrix Decomposition Theory

Let us consider the Hopfield ANN as a nonlinear dynamical system consisting of three neurons u_1, u_2 and u_3 (Figure 1). It is known [1], the dynamics of the Hopfield ANN functioning is given by the following rule:

$$u_i(t+1) = \sum_{l=1,\, i\neq l}^{N} w_{i,l} \cdot F\left(u_l(t)\right) - \mathrm{T}_i,\ i = \overline{1,N} \tag{1a}$$

where $w_{i,l}$ are elements of synaptic weights matrix $\mathrm{W}_{3\times 3}$, N is the input vector length (in particular, $N = 3$), F is an activation function, T_i is a bias value of the i-th neuron (as a rule, $\mathrm{T}_i = 0$). As activation function $F(u_l)$ we choose the hyperbolic tangent.

Let us describe the dynamics of states of each neuron u_i :

$$\dot{u}_i = \sum_{l=1,\, i\neq l}^{N} w_{i,l} F(u_l) - u_i - \mathrm{T}_i \tag{1b}$$

Let us investigate the dynamics of the state changing for all output neurons in the Hopfield's ANN for N=3 in accordance with Figure 1:

$$\begin{cases} \dot{u}_1 = f_1(u_1,u_2,u_3) = w_{12}F(u_2) + w_{13}F(u_3) - u_1; \\ \dot{u}_2 = f_2(u_1,u_2,u_3) = w_{21}F(u_1) + w_{23}F(u_3) - u_2; \\ \dot{u}_3 = f_3(u_1,u_2,u_3) = w_{31}F(u_1) + w_{32}F(u_2) - u_3. \end{cases} \tag{2}$$

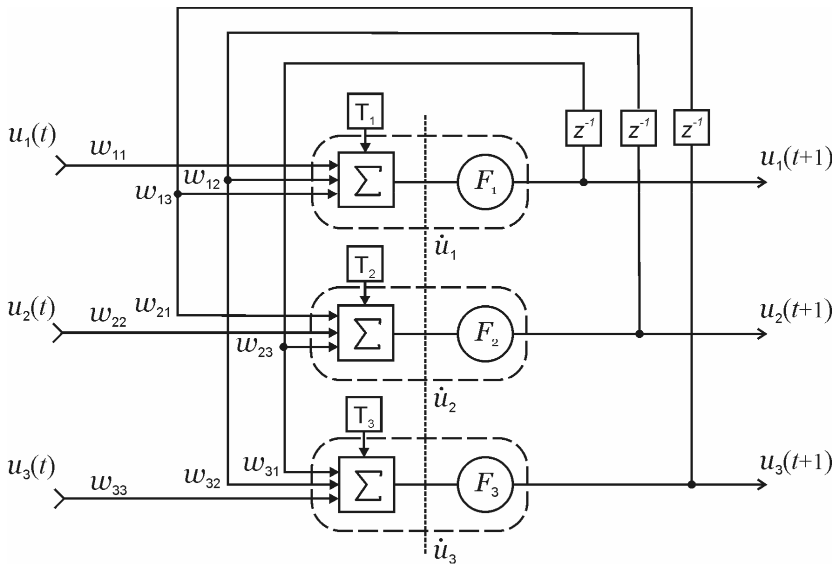

Fig. 1. The Hopfield's neural network architecture

In order to use theory of matrix decomposition [7-14] we represent the system of equations (2) by the following vector functions:

$$\dot{\vec{u}} = \begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \\ \dot{u}_3 \end{bmatrix} \tag{3a}$$

$$\vec{f}(\vec{u}) = \begin{bmatrix} w_{12}F(u_2) + w_{13}F(u_3) - u_1 \\ w_{21}F(u_1) + w_{23}F(u_3) - u_2 \\ w_{31}F(u_1) + w_{32}F(u_2) - u_3 \end{bmatrix} \tag{3b}$$

According to the nonlinear analysis based on matrix decomposition [7-14] we study the solution of equation (1*b*) near a specific standard state $\{u_i^*\}$, where $u_i^* = u_i^*(t)$, permanently disturbed by value $v_i = v_i(t)$ of external perturbations or internal fluctuations. In result, instead of u_i^* a new solutions becomes

$$u_i = u_i^* + v_i \tag{4}$$

Taking into account (4) we can find the increment of the vector function in the state space of the Hopfield's ANN (3b) in the form

$$\Delta\vec{f}(\vec{v},\vec{u}^*) = \vec{f}(\vec{u}^* + \vec{v}) - \vec{f}(\vec{u}^*) =$$

$$= \begin{bmatrix} w_{12}\left(F(u_2^* + v_2) - F(u_2^*)\right) + w_{13}\left(F(u_3^* + v_3) - F(u_3^*)\right) - v_1 \\ w_{21}\left(F(u_1^* + v_1) - F(u_1^*)\right) + w_{23}\left(F(u_3^* + v_3) - F(u_3^*)\right) - v_2 \\ w_{31}\left(F(u_1^* + v_1) - F(u_1^*)\right) + w_{32}\left(F(u_2^* + v_2) - F(u_2^*)\right) - v_3 \end{bmatrix} \quad (5)$$

According to the matrix decomposition theory [7-14] let us represent the increment of the vector function in the space of states by the matrix series expansion:

$$\Delta\vec{f}(\vec{u}^*,\vec{v}) = \vec{f}(\vec{u}^* + \vec{v}) - \vec{f}(\vec{u}^*) =$$

$$= L_{N\times N}^{(1)}\vec{v} + \frac{1}{2!}L_{N\times N^2}^{(2)}(\vec{v}\otimes\vec{v}) + \frac{1}{3!}L_{N\times N^3}^{(3)}(\vec{v}\otimes\vec{v}\otimes\vec{v}) = \sum_{k=1}^{\infty}\frac{1}{k!}L_{N\times N^k}^{(k)}\cdot\vec{v}^{\otimes k}, \quad (6a)$$

$$L_{N\times N^k}^{(k)} = \left(\frac{\partial}{\partial\vec{v}^{\mathrm{T}}}\otimes\left(\frac{\partial}{\partial\vec{v}^{\mathrm{T}}}\otimes\ldots\otimes\left(\frac{\partial}{\partial\vec{v}^{\mathrm{T}}}\otimes\vec{f}\right)\ldots\right)\right)\Bigg|_{\vec{u}^*}, \quad (6b)$$

where $L_{N\times N^k}^{(k)}$ are matrix kernels of homogeneous nonlinear operators of the system into the state space; $\vec{v}^{\otimes k} = (\overbrace{\vec{v}\otimes\vec{v}\otimes\ldots\otimes\vec{v}}^{k})$ is k-th Kronecker degree of the vector $\vec{v}$ [7-14].

In particular, the kernels of elements of the first order in accord with (1b) and (4) can be expressed by the following formula:

$$L_{ij}^{(1)} = \frac{\partial f_i}{\partial v_j}\Bigg|_{u_j=u_j^*} = \frac{\partial}{\partial v_j}\left(\sum_{l=1,l\neq i}^{3} w_{il}F(u_l) - u_i\right) =$$

$$\sum_{l=1,l\neq i}^{3} w_{il}\frac{\partial F(u_l)}{\partial v_j}\Bigg|_{u_j=u_j^*} - \frac{\partial u_i}{\partial v_j} = w_{il}\frac{\partial F(u_j^*)}{\partial v_j}(1-\delta_{ij}) - \delta_{ij}, \quad (7a)$$

where δ_{ij} is the Kronecker's delta-symbol. Taking into account (7a) let us write the kernel of the first order in the matrix form:

$$L_{3\times 3}^{(1)} = \begin{bmatrix} -1 & w_{12}F'(u_2^*) & w_{13}F'(u_3^*) \\ w_{21}F'(u_1^*) & -1 & w_{23}F'(u_3^*) \\ w_{31}F'(u_1^*) & w_{32}F'(u_2^*) & -1 \end{bmatrix}. \quad (7b)$$

Similarly, we find elements of the second order kernel by means of the formula:

$$L_{ijk}^{(2)} = \frac{\partial^2 f_i}{\partial v_j\partial v_k}\Bigg|_{\substack{u_j=u_j^* \\ u_k=u_k^*}} = \frac{\partial^2}{\partial v_j\partial v_k}\left(\sum_{l=1,l\neq i}^{3} w_{il}F(u_l) - u_i\right)\Bigg|_{\substack{u_j=u_j^* \\ u_k=u_k^*}} =$$

$$= \frac{\partial}{\partial v_k}\left(\sum_{l=1,l\neq i}^{3} w_{il}\frac{\partial F(u_l^*)}{\partial v_j} - \frac{\partial u_i}{\partial v_j}\right)\Bigg|_{\substack{u_j=u_j^* \\ u_k=u_k^*}} =$$

$$= \sum_{l=1,l\neq i}^{3} w_{il}\frac{\partial^2 F(u_l^*)}{\partial v_j\partial v_k} = w_{ij}\frac{\partial^2 F(u_j^*)}{\partial^2 v_j}(1-\delta_{ij})\cdot\delta_{ik}. \quad (8a)$$

So, the corresponding matrix for the second order kernel takes the following form:

$$L^{(2)}_{3\times 9}=\left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & w_{12}F''(u_2^*) & 0 & 0 & 0 & w_{13}F''(u_3^*) \\ w_{21}F''(u_1^*) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{23}F''(u_3^*) \\ w_{31}F''(u_1^*) & 0 & 0 & 0 & w_{32}F''(u_2^*) & 0 & 0 & 0 & 0 \end{array}\right] \tag{8b}$$

By analogy with (7*b*), (8*b*) we can obtain the matrix form of kernel of the third order:

$$L^{(3)}_{3\times 27}=$$

$$=\left[\begin{array}{ccc|ccc|ccc|ccc|ccc|ccc|ccc|ccc|ccc} 0 & 0 & 0 & 0&0&0 & 0&0&0 & 0&0&0 & 0 & w_{12}F'''(u_2^*) & 0 & 0&0&0 & 0&0&0 & 0&0&0 & 0 & 0 & w_{13}F'''(u_3^*) \\ w_{21}F'''(u_1^*) & 0 & 0 & 0&0&0 & 0&0&0 & 0&0&0 & 0 & 0 & 0 & 0&0&0 & 0&0&0 & 0&0&0 & 0 & 0 & w_{23}F'''(u_3^*) \\ w_{31}F'''(u_1^*) & 0 & 0 & 0&0&0 & 0&0&0 & 0&0&0 & 0 & w_{32}F'''(u_2^*) & 0 & 0&0&0 & 0&0&0 & 0&0&0 & 0 & 0 & 0 \end{array}\right] \tag{9}$$

Restricting number of terms in the matrix series (6*a*) up to the 3-rd order inclusively, we approximate the increment of the vector function (5) into state space of the Hopfield's ANN:

$$\Delta\vec{f}(\vec{v},\vec{u}^*)\approx L^{(1)}_{3\times 3}(\vec{u}^*)\vec{v}+\frac{1}{2!}L^{(2)}_{3\times 9}(\vec{u}^*)\cdot(\vec{v}\otimes\vec{v})+\frac{1}{3!}L^{(3)}_{3\times 27}(\vec{u}^*)\cdot(\vec{v}\otimes\vec{v}\otimes\vec{v})\,. \tag{10}$$

To estimate the accuracy of the approximation, let us we find the following three terms of the matrix series in analytical form:

$$L^{(1)}_{3\times 3}(\vec{u}^*)\vec{v}=$$

$$=\begin{bmatrix} -1 & w_{12}F'(u_2^*) & w_{13}F'(u_3^*) \\ w_{21}F'(u_1^*) & -1 & w_{23}F'(u_3^*) \\ w_{31}F'(u_1^*) & w_{32}F'(u_2^*) & -1 \end{bmatrix}\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}=$$

$$=\begin{bmatrix} w_{12}F'(u_2^*)\cdot v_2+w_{13}F'(u_3^*)\cdot v_3-v_1 \\ w_{21}F'(u_1^*)\cdot v_1+w_{23}F'(u_3^*)\cdot v_3-v_2 \\ w_{31}F'(u_1^*)\cdot v_1+w_{32}F'(u_2^*)\cdot v_2-v_3 \end{bmatrix}. \tag{11a}$$

$$L^{(2)}_{3\times 9}(\vec{u}^*)\cdot(\vec{v}\otimes\vec{v})=$$

$$=\left[\begin{array}{ccc|ccc|ccc} 0 & 0 & 0 & 0 & w_{12}F''(u_2^*) & 0 & 0 & 0 & w_{13}F''(u_3^*) \\ w_{21}F''(u_1^*) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{23}F''(u_3^*) \\ w_{31}F''(u_1^*) & 0 & 0 & 0 & w_{32}F''(u_2^*) & 0 & 0 & 0 & 0 \end{array}\right]\begin{bmatrix} v_1^2 \\ v_1v_2 \\ v_1v_3 \\ v_2v_1 \\ v_2^2 \\ v_2v_3 \\ v_3v_1 \\ v_3v_2 \\ v_3^2 \end{bmatrix}=$$

$$=\begin{bmatrix} w_{12}F''(u_2^*)\cdot v_2^2+w_{13}F''(u_3^*)\cdot v_3^2 \\ w_{21}F''(u_1^*)\cdot v_1^2+w_{23}F''(u_3^*)\cdot v_3^2 \\ w_{31}F''(u_1^*)\cdot v_1^2+w_{32}F''(u_2^*)\cdot v_2^2 \end{bmatrix} \tag{11b}$$

$$L^{(3)}_{3\times 27}(\vec{u}^*)\cdot(\vec{v}\otimes\vec{v}\otimes\vec{v})=$$

$$
= \left[\begin{array}{ccc|ccc|ccc|ccc|ccc|ccc|ccc|ccc|ccc}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{12}F'''(u_2^*) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{13}F'''(u_3^*) \\
w_{21}F'''(u_1^*) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{23}F'''(u_3^*) \\
w_{31}F'''(u_1^*) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & w_{32}F'''(u_2^*) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}\right]
\begin{bmatrix} v_1^3 \\ v_1v_1v_2 \\ v_1v_1v_3 \\ v_1v_2v_1 \\ v_1v_2^2 \\ v_1v_2v_3 \\ v_1v_3v_1 \\ v_1v_3v_2 \\ v_1v_3^2 \\ v_2v_1^2 \\ v_2v_1v_2 \\ v_2v_1v_3 \\ v_2v_2v_1 \\ v_2^3 \\ v_2v_2v_3 \\ v_2v_3v_1 \\ v_2v_3v_2 \\ v_2v_3^2 \\ v_2v_1^2 \\ v_3v_1v_2 \\ v_3v_1v_3 \\ v_3v_2v_1 \\ v_3v_2^2 \\ v_3v_2v_3 \\ v_3v_3v_1 \\ v_3v_3v_2 \\ v_3^3 \end{bmatrix} =
$$

$$
= \begin{bmatrix} w_{12}F'''(u_2^*)\cdot v_2^3 + w_{13}F'''(u_3^*)\cdot v_3^3 \\ w_{21}F'''(u_1^*)\cdot v_1^3 + w_{23}F'''(u_3^*)\cdot v_3^3 \\ w_{31}F'''(u_1^*)\cdot v_1^3 + w_{32}F'''(u_2^*)\cdot v_2^3 \end{bmatrix} \qquad (11c)
$$

Substituting (11a)–(11c) in (10) we find an approximating function $\vec{g}_M(\vec{v},\vec{u}^*)$ for $\Delta\vec{f}(\vec{v},\vec{u}^*)$ as a vector sum of three terms of this matrix series ($M = 3$):

$$
\Delta\vec{f}(\vec{v},\vec{u}^*) \approx \vec{g}_3(\vec{v},\vec{u}^*) =
$$

$$
= \begin{bmatrix} w_{12}F'(u_2^*)v_2 + w_{13}F'(u_3^*)v_3 - v_1 \\ w_{21}F'(u_1^*)v_1 + w_{23}F'(u_3^*)v_3 - v_2 \\ w_{31}F'(u_1^*)v_1 + w_{32}F'(u_2^*)v_2 - v_3 \end{bmatrix} + \frac{1}{2!}\begin{bmatrix} w_{12}F''(u_2^*)v_2^2 + w_{13}F''(u_3^*)v_3^2 \\ w_{21}F''(u_1^*)v_1^2 + w_{23}F''(u_3^*)v_3^2 \\ w_{31}F''(u_1^*)v_1^2 + w_{32}F''(u_2^*)v_2^2 \end{bmatrix} +
$$

$$
+ \frac{1}{3!}\begin{bmatrix} w_{12}F'''(u_2^*)v_2^3 + w_{13}F'''(u_3^*)v_3^3 \\ w_{21}F'''(u_1^*)v_1^3 + w_{23}F'''(u_3^*)v_3^3 \\ w_{31}F'''(u_1^*)v_1^3 + w_{32}F'''(u_2^*)v_2^3 \end{bmatrix} =
$$

$$
= \begin{bmatrix} -v_1 + w_{12}\left(F'(u_2^*)v_2 + \frac{1}{2}F''(u_2^*)v_2^2 + \frac{1}{6}F'''(u_2^*)v_2^3\right) + w_{13}\left(F'(u_3^*)v_3 + \frac{1}{2}F''(u_3^*)v_3^2 + \frac{1}{6}F'''(u_3^*)v_3^3\right) \\ -v_2 + w_{21}\left(F'(u_1^*)v_1 + \frac{1}{2}F''(u_1^*)v_1^2 + \frac{1}{6}F'''(u_1^*)v_1^3\right) + w_{23}\left(F'(u_3^*)v_3 + \frac{1}{2}F''(u_3^*)v_3^2 + \frac{1}{6}F'''(u_3^*)v_3^3\right) \\ -v_3 + w_{31}\left(F'(u_1^*)v_1 + \frac{1}{2}F''(u_1^*)v_1^2 + \frac{1}{6}F'''(u_1^*)v_1^3\right) + w_{32}\left(F'(u_2^*)v_2 + \frac{1}{2}F''(u_2^*)v_2^2 + \frac{1}{6}F'''(u_2^*)v_2^3\right) \end{bmatrix} \qquad (12)
$$

To determine the vector function of an approximation error $\vec{\varepsilon}_M(\vec{v},\vec{u}^*)$, i.e. the so-called residual vector, we find the difference between the right-hand sides of equations (5) and (12) for $M=3$:

$$\vec{\varepsilon}_M(\vec{v},\vec{u}^*)=\Delta\vec{f}(\vec{v},\vec{u}^*)-\vec{g}_M(\vec{v},\vec{u}^*)=$$

$$=\begin{bmatrix} w_{12}\left(F(u_2^*+v_2)-F(u_2^*)-F'(u_2^*)v_2-\dfrac{F''(u_2^*)v_2^2}{2}-\dfrac{F'''(u_2^*)v_2^3}{6}\right) \\ w_{21}\left(F(u_1^*+v_1)-F(u_1^*)-F'(u_1^*)v_1-\dfrac{F''(u_1^*)v_1^2}{2}-\dfrac{F'''(u_1^*)v_1^3}{6}\right) \\ w_{31}\left(F(u_1^*+v_1)-F(u_1^*)-F'(u_1^*)v_1-\dfrac{F''(u_1^*)v_1^2}{2}-\dfrac{F'''(u_1^*)v_1^3}{6}\right)\end{bmatrix}+$$

$$+\begin{bmatrix} w_{13}\left(F(u_3^*+v_3)-F(u_3^*)-F'(u_3^*)v_3-\dfrac{F''(u_3^*)v_3^2}{2}-\dfrac{F'''(u_3^*)v_3^3}{6}\right) \\ w_{23}\left(F(u_3^*+v_3)-F(u_3^*)-F'(u_3^*)v_3-\dfrac{F''(u_3^*)v_3^2}{2}-\dfrac{F'''(u_3^*)v_3^3}{6}\right) \\ w_{32}\left(F(u_2^*+v_2)-F(u_2^*)-F'(u_2^*)v_2-\dfrac{F''(u_2^*)v_2^2}{2}-\dfrac{F'''(u_2^*)v_2^3}{6}\right)\end{bmatrix} \tag{13}$$

Further we estimate the approximation error δ_M of vector function (5) in the state space of Hopfield's ANN based on a length of vector discrepancy $\vec{\varepsilon}_M(\vec{v},\vec{u}^*)$:

$$\delta_M=\frac{\left\|\vec{\varepsilon}_M(\vec{v}_N,\vec{u}_N^*)\right\|}{\sqrt{N}}\cdot 100\%\,. \tag{14}$$

Then in the case of $M=3$, the residual vector δ_3 is equal to

$$\delta_3=\frac{1}{\sqrt{3}}\sqrt{\vec{\varepsilon}_1^{\,2}(\vec{u}_2^*,\vec{u}_3^*,\vec{v}_2,\vec{v}_3)+\vec{\varepsilon}_2^{\,2}(\vec{u}_1^*,\vec{u}_3^*,\vec{v}_1,\vec{v}_3)+\vec{\varepsilon}_3^{\,2}(\vec{u}_1^*,\vec{u}_2^*,\vec{v}_1,\vec{v}_2)}\cdot 100\%\,.$$

Before calculating (13) and (14) should be noted that

$$F(u_l)=\tanh(u_l)=\frac{e^{u_l}-e^{-u_l}}{e^{u_l}+e^{-u_l}};$$

$$F'(u_l)=\frac{1}{\cosh^2(u_l)}=\frac{4}{\left(e^{u_l}+e^{-u_l}\right)^2};$$

$$F''(u_l)=-\frac{2}{\cosh^2(u_l)}\tanh(u_l)=-8\frac{\left(e^{u_l}-e^{-u_l}\right)}{\left(e^{u_l}+e^{-u_l}\right)^3};$$

$$F'''(u_l)=-\frac{2}{\cosh^4(u_l)}+\frac{4}{\cosh^2(u_l)}\tanh^2(u_l)=-\frac{32+16\left(e^{u_l}-e^{-u_l}\right)^2}{\left(e^{u_l}+e^{-u_l}\right)^4},$$

where the values u_l are chosen equal to -1, 0 or +1. Then it follows directly that

$$F(x)=\begin{cases}-0.7616 \text{ if } u_l=-1;\\ 0.0 \text{ if } u_l=0;\\ 0.7616 \text{ if } u_l=1.\end{cases}$$

$$F'(x)=\begin{cases}0.4200 \text{ if } u_l=-1;\\ 1.00 \text{ if } u_l=0;\\ 0.4200 \text{ if } u_l=1.\end{cases}$$

$$F''(x)=\begin{cases}0.6397 \text{ if } u_l=-1;\\ 0.00 \text{ if } u_l=0;\\ \text{-}0.6397 \text{ if } u_l=1.\end{cases}$$

$$F'''(x)=\begin{cases}0.0618 \text{ if } u_l=-1;\\ \text{-}2.00 \text{ if } u_l=0;\\ \text{-}0.7673 \text{ if } u_l=1.\end{cases}$$

One can see from (12) that in general form the elements of a vector approximating functions $\vec{g}_M(\vec{v},\vec{u}^*)$ can be described as follows:

$$g_M^i(v_j,u_j^*)=-v_i+\sum_{j=1,\ j\neq i}^{N} w_{i,j}\sum_{k=1}^{M}\frac{1}{k!}F^{(k)}(u_j^*)v_j^k\,, \qquad (15)$$

where $F^{(k)}(u_j^*)$ denotes the k-th derivative of the activation function $F(u_j)$ calculated at the point $\vec{u}_j^*$, N is a number of neurons in the input layer and M is a number of kernels of the matrix series (6a).

Thus, owing to (12) and (15) it has become possible to calculate the increment of the vector function $\Delta\vec{f}(\vec{v},\vec{u}^*)$ in the state space of the Hopfield's ANN under condition of input vector of arbitrary length N with accuracy to M-th kernel.

To obtain the residual vector $\vec{\varepsilon}_M(\vec{v},\vec{u}^*)$ and the error value δ_M let us simulate the Hopfield's ANN.

3 A Computational Experiment to Determine the Approximation Error for Binary Patterns

As an example, let us consider the process of restoration of binary vector $\mathbf{a}=[1\quad 1\quad 1]^{\mathrm{T}}$ using the Hopfield's ANN (Figure 1). To this end, we use the Hebb's learning rule [1] to calculate the weight matrix $\mathrm{W}_{3\times 3}$. In general, the formation of the weight matrix is carried out by means of the Hebb's learning rule [1], [2]:

$$\mathrm{W}_{N\times N}=\frac{1}{N}\sum_{i=1}^{p}\left(\mathbf{a}_i'\otimes\mathbf{a}_i'-\mathrm{E}_{N\times N}\right), \qquad (16)$$

where $\mathrm{W}_{N\times N}$ is a weight matrix with size of $N\times N$, besides N is a length of the input vector, p is a number of trained pairs of vectors, $\mathrm{E}_{N\times N}$ is a diagonal identity matrix with size of $N\times N$. The operation $\mathbf{a}'_i=(2\mathbf{a}_i-1)$ is performed to convert the binary vectors $\mathbf{a}_i$ in the bipolar form, i.e. to find $\mathbf{a}'=[1\quad 1\quad 1]^{\mathrm{T}}$. For the considered case N=3 and p =1, we obtain that

$$\mathrm{W}_{3\times 3}=\frac{1}{3}\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\begin{bmatrix}1 & 1 & 1\end{bmatrix}-\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}\right)=\frac{1}{3}\begin{bmatrix}0 & 1 & 1\\1 & 0 & 1\\1 & 1 & 0\end{bmatrix}, \tag{17}$$

i.e. weight matrix element values are calculated as follows: $w_{12}=w_{13}=1/3$, $w_{21}=w_{23}=1/3$, $w_{31}=w_{32}=1/3$, a $w_{11}=w_{22}=w_{33}=0$.
Considering the obtained element values w_{ij} of the weights matrix $\mathrm{W}_{3\times 3}$ accord with (17), we rewrite the system (2) as follows:

$$\begin{cases}\dot{u}_1=\frac{1}{3}\big(F(u_2)+F(u_3)\big)-u_1;\\ \dot{u}_2=\frac{1}{3}\big(F(u_1)+F(u_3)\big)-u_2;\\ \dot{u}_3=\frac{1}{3}\big(F(u_1)+F(u_2)\big)-u_3.\end{cases} \tag{18}$$

According to the above mentioned statements of the matrix decomposition theory (3*a*) – (12) with respect to a Hopfield's ANN, an external disturbance vector is interpreted as a disturbance vector $\vec{v}$ distorting the standard vector $\vec{u}^*$. In other words, according to formula (4) let us assume that the vector $\vec{v}$ is the disturbance from behind the input vector $\vec{u}$ differs from the reference vector $\vec{u}^*$, i.e. $\vec{u}-\vec{u}^*=\vec{v}$.
Thus, the elements of the vector $\vec{v}$ belong to the set «–1», «0» and «+1» that defines the following: if $v_i=0$ then $u_i=u_i^*$, i.e. the test vector elements completely coincide with elements of the standard vector; if $v_i=\pm 1$ then $u_i\neq u_i^*$. Consequently, the vector magnitude $\vec{v}$ can be estimated on the basis of the Hamming' distance $d(\vec{u},\vec{u}^*)$ between vectors $\vec{u}$ and $\vec{u}^*$, i.e. by the number of positions in which these vectors are different. In other words, the Hamming distance $d(\vec{u},\vec{u}^*)$ is the norm of vector $\vec{v}$:

$$d(\vec{u},\vec{u}^*)=\|\vec{v}\|=\sum_{i=1}^{N}|v_i|, \tag{19}$$

where $d(\vec{u},\vec{u}^*)$ is the Hamming distance and N is a length of the vector $\vec{v}$.

A computational simulation of the Hopfield' ANN permits to determine the numerical values of the vector $\vec{\varepsilon}_M$ components of the approximation error. The data of computational simulation are presented in the Table 1 which displays the approximation error δ_M of vector function in the state space of the Hopfield's ANN on the basis of the theory of matrix decomposition and computer modeling.

The fifth column of this Table 1 shows the kernels number *M* used for function approximation. As can be seen from the Table 1, the maximum error occurs in the case of linear approximation ($M = 1$) and matches to 9.3% (in the first example), and the minimum error occurs under taking into account the nonlinear terms of higher order in equation (6a), besides it is equal to 1.05% (see the second example). The first 5 kernels have been used in the simulation only. However, the values δ_M of the approximation error for the first pair of vectors $\vec{u}^*$ and $\vec{v}$ lead to an assumption about periodic behavior.

At the same time, a series of experiments estimating the residual vector $\vec{\varepsilon}_M$ and the approximation error δ_M have been carried with an activation function as the sigmoidal function $F(u_l) = 1/(1+e^{-l})$. As a result, the values $\vec{\varepsilon}_M$ and δ_M are found slightly higher but the behavior error is remained the same.

Table 1. Calculation of the approximation error in computational experiments with Hopfield's ANN

No	$\vec{u}$	$\vec{u}^*$	$\vec{v}$	M	Theoretical estimation $\Delta\vec{f}$	Computational estimation $\Delta\vec{f}$	$\vec{\varepsilon}_M$	δ_M, %
1	2	3	4	5	6	7	8	9
1				1	$\begin{bmatrix}-0.1400\\1.0\\-0.1400\end{bmatrix}$		$\begin{bmatrix}-0.1139\\0.0\\-0.1139\end{bmatrix}$	9,30
2				2	$\begin{bmatrix}-0.2466\\1.0\\-0.2466\end{bmatrix}$		$\begin{bmatrix}-0.0073\\0.0\\-0.0073\end{bmatrix}$	0,60
3	$\begin{bmatrix}1\\0\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\\1\end{bmatrix}$	$\begin{bmatrix}0\\-1\\0\end{bmatrix}$	3	$\begin{bmatrix}-0.2040\\1.0\\-0.2040\end{bmatrix}$	$\begin{bmatrix}-0.2539\\1.0\\-0.2539\end{bmatrix}$	$\begin{bmatrix}-0.0499\\0.0\\-0.0499\end{bmatrix}$	4,07
4				4	$\begin{bmatrix}-0.1947\\1.0\\-0.1947\end{bmatrix}$		$\begin{bmatrix}-0.0591\\0.0\\-0.0591\end{bmatrix}$	4,83
5				5	$\begin{bmatrix}-0.1793\\1.0\\-0.1793\end{bmatrix}$		$\begin{bmatrix}-0.0746\\0.0\\-0.0746\end{bmatrix}$	6,09

Table 1 (continuation)

1	2	3	4	5	6	7	8	9
6	$\begin{bmatrix}1\\1\\1\end{bmatrix}$	$\begin{bmatrix}1\\1\\0\end{bmatrix}$	$\begin{bmatrix}0\\0\\1\end{bmatrix}$	1	$\begin{bmatrix}-0.3333\\-0.3333\\-1.0\end{bmatrix}$	$\begin{bmatrix}-0.2539\\-0.2539\\-1.0\end{bmatrix}$	$\begin{bmatrix}-0.0795\\-0.0795\\0.0\end{bmatrix}$	6,49
7				2	$\begin{bmatrix}-0.3333\\-0.3333\\-1.0\end{bmatrix}$		$\begin{bmatrix}-0.0795\\-0.0795\\0.0\end{bmatrix}$	6,49
8				3	$\begin{bmatrix}-0.2222\\-0.2222\\-1.0\end{bmatrix}$		$\begin{bmatrix}-0.0316\\-0.0316\\0.0\end{bmatrix}$	2,58
9				4	$\begin{bmatrix}-0.2222\\-0.2222\\-1.0\end{bmatrix}$		$\begin{bmatrix}-0.0316\\-0.0316\\0.0\end{bmatrix}$	2,58
10				5	$\begin{bmatrix}-0.2667\\-0.2667\\-1.0\end{bmatrix}$		$\begin{bmatrix}0.0128\\0.0128\\0.0\end{bmatrix}$	1,05

4 Computer Simulation of Stages of the Hopfield's ANN functioning

One of the main applications of ANN is the classification and pattern recognition. The task of classification is the reference of the input vector to one of the known classes. A stable functioning of the classifier depends on a measure of similarity of the input vector with the standard one storing in the memory of the classifier. This stability also depends on a level of noise imposed on the input vector when the latter can be still recognized correctly.

The process of patterns retrieving based on the Hopfield's ANN is to suppress the distortions presenting in the input vectors. Due to the known difficulties of the mathematical analysis of complex dynamical behavior of recurrent ANN, the question of the maximal possible level determining has been not enough iterpretive in the scientific literature. Therefore, one of purposes of this paper is to develop a method of nonlinear analysis based on matrix decomposition allowing predicting the behavior of the Hopfield's ANN under recognizing the input vectors.

However, the above illustrated example for recording and recovery (with the help of Hopfield's ANN) of binary vector (consisting of 3 elements only) does not allow to fully estimate the benefits of the proposed method. Therefore, let us consider a typical problem of Hopfield's ANN concerning associative restoration of noisy patterns.

For example, Figure 2 shows two noisy images (2*b*) and (2*c*) represented by vectors $\vec{u}_1$, $\vec{u}_2$, and one distorted image (2*d*) represented by a vector $\vec{u}_3$, which have been obtained by applying to the standard image (2*a*) (represented by a vector $\vec{u}^*$) disturbances in the form of noises (2*e*)–(2*g*) encoded by vectors $\vec{v}_1$, $\vec{v}_2$ and $\vec{v}_3$ respectively. Two-dimensional vector with size of 32×32 represents different types of monochrome images of the letter "A".

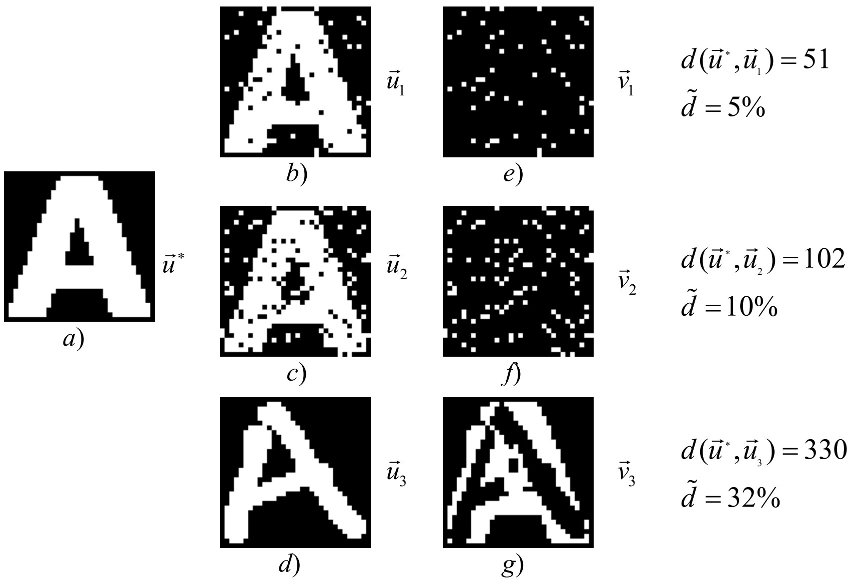

Fig. 2. The process of applying distortion and noise on the standard image

For each pair of vectors, the values of the Hamming distance $d(\vec{u},\vec{u}_i^*)$ [15] are shown. However, as it follows from this example, the analysis of vectors consisting of a large number of elements (32×32=1024) is not always convenient to use the specified value. Therefore, we introduce a new value, so-called a relative Hamming distance $\tilde{d}$, besides its value does not depend on the length of vectors $\vec{u}$ and $\vec{u}^*$:

$$\tilde{d}(\vec{u},\vec{u}^*)=\frac{d(\vec{u},\vec{u}^*)}{N}\cdot 100\%. \tag{20}$$

In this regard, the vector $\vec{v}$ is to be characterized by a variable $\tilde{d}(\vec{u},\vec{u}^*)$ which is calculated by the ratio of the Hamming distance $d(\vec{u},\vec{u}_i^*)$ to the value of elements of this vector $\vec{v}$

Even in the case of presence 32% of distortions the trained Hopfield' ANN is able to qualitatively recover the input image to the standard values. This is achieved due to the fact that one image is recorded by the ANN only. However, the practical implementation of the Hopfield' ANN shows that an increase in the number of images recorded in the network leads to decreases of the ability of the Hopfield' ANN to restore them, i.e. an image can be restored only in the case of a slight distortion.

Due to (15) it is possible to numerically analyze the influence of the initial distortions in the process of pattern restoring. For example, Figure 3 shows the gradient of the increment of the function describing the dynamics of the Hopfield's ANN consisting of 1024 neurons in which each image represents a two-dimensional monochrome image of the letter "A". The standard vector $\vec{u}^*$ and input vector $\vec{u}_3$ for these images are illustrated in Figure 2*a* and 2*d*, respectively. As one can see from these data, a new vector is characterized by the small perturbations therefore it has properly been restored.

It should be noted that in result of the restoration of pattern, the image parts are presented by black and dark gray colors in Figure 3, i.e. by the values -0.1470 and -0.1464, have been assigned to the object, whereas these shown by light gray and white, i.e. by values -0.0007 and 0.0, have been interpreted as background. Thus, knowing the threshold value, it is possible to determine in advance how the input vector can influence the process of recovery through the ANN.

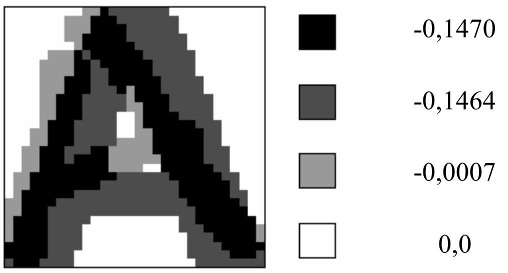

Fig. 3. The gradient values of the function increments

On the other hand, knowing the value of the standard elements of recognizable images, we can estimate based on formula (15) what percentage of the distortions for already trained network can be overcome. In other words, the Hopfield' ANN should consistently apply to the images in which the percentage of distortion increases with each time. Then, analyzing the values g_M^i we can determine the threshold value when the system would no longer be able to adequately restore the image.

4 CONCLUSIONS

In this paper we propose a new approach for the numerical determination of the perturbations of the Hopfield' ANN on the stages of the restoration of previously unknown pattern with usage of the matrix decomposition theory [7-14]. The approximating function g_M^i has been derived besides its accuracy of restoration depends on the number M of the matrix kernels of homogeneous non-linear operators of the complex dynamical system as the Hopfield's ANN.

References

1. J. Hopfield. Neural networks and physical systems with emergent collective computational abilities, *Proceedings of National Academy of Sciences*, vol. 79, no. 8, 2554–2558, 1982
2. I.I. Melamed. Neural networks and combinatorial optimization. *Automation and Remote control*, vol. 55, no. 11, 1553-1584, 1994.
3. R.J. McEliece, E.C. Posner, E.R. Rodemich, S.S. Venkatesh. The capacity of the Hopfield associative memory, *IEEE Transactions on Information Theory*, vol. 33, no 4, 461-482, 1987.
4. E.N. Lorenz. Deterministic nonperiodic flow, *Journal of Atmospheric Sciences*, vol. 20, 130-141, 1963.
5. G. Nicolis, I. Prigogine. Self-organization in nonequilibrium systems: From Dissipative Structures to Order through Fluctuation, John Willey & Sons, New York etc., 1977.
6. P. Berge, Y. Pomeau and C. Vidal. L'ordre dans le chaos: vers une approach deterministe de la turbulence, Hermann, Paris, 1988.
7. A.M. Krot. The decomposition of vector functions in vector-matrix series into state-space of nonlinear dynamic system. EUSIPCO-2000 : *Proc. X European Signal Processing Conference*, Tampere, Finland, 3: 2453–2456, 2000.
8. A.M. Krot. Chaotic dynamic methods based on decomposition of vector functions in vector-matrix series into state-space. *Proc. 10th Mediterranean Electrotechnical Conference*, Lemesos, Cyprus, 2: 643–646, 2000.
9. A.M. Krot. Matrix decompositions of vector functions and shift operators on the trajectories of a nonlinear dynamical system, *Nonlinear Phenomena in Complex Systems*, vol. 4, no. 2, 106–115, 2001.
10. A.M. Krot. Application of expansion into matrix to analysis of attractors of complex nonlinear dynamical systems. *Proc. 14th IEEE International Conference on Digital Signal Processing*, Santorini, Greece, 959–962, 2002.
11. A.M. Krot, H.B. Minervina. Minimal attractor embedding dimension for discrete dynamic system using state-space method: theoretical ground, *Proc. of the 6th IEEE Int. Conf. on Electronics, Circuits and Systems*, Pafos, Cyprus, 2: 941-944, 1999.
12. A.M. Krot, H.B. Minervina. Minimal attractor embedding estimation based on matrix decomposition for analysis of dynamical systems, *Nonlinear Phenomena in Complex Systems*, vol. 5, no. 2., 161–172, 2002.
13. A.M. Krot. The development of matrix decomposition theory for nonlinear analysis of chaotic attractors of complex systems and signals. *Proc. 16th IEEE International Conference on Digital Signal Processing*, Thira, Santorini, Greece, 2009.
14. A.M. Krot. Bifurcation analysis of attractors of complex systems based on matrix decomposition theory. *Proc. of IEEE Intern. Conference on Industrial Engineering and Management*, Zhengzhou, China, 2011.
15. R.E. Blahut. Fast algorithms for digital signal processing. Addison-Wesley Publ. Co.: Reading etc., 1985.