2006 International Joint Conference on Neural Networks
Sheraton Vancouver Wall Centre Hotel, Vancouver, BC, Canada
July 16-21, 2006

Locally Asymptotically Stable Fixed-point Assignment Problems in Neural Networks and Application to Associative Memory

Hiroshi Inaba, *Member, IEEE,* Takafumi Ooki and Keylan Alimhan

***Abstract*— This paper proposes and studies in the frame work of systems and control theory the problem of how to construct a dynamical system defined over a Hilbert space such a way that any given vectors are assigned to locally asymptotically stable fixed-points of the system. Some basic properties of such systems are investigated, and further the results are applied to neural networks to implement associative memory or pattern recognition for two-dimensional images in such a way that a certain structural deformation of images is acceptable. Finally some numerical examples for associative memory are presented to illustrate the performance.**

I. Introduction

In the dynamical systems theory, the problem of analyzing the stability of fixed-points of a system is one of most important issues. On the other hand, for instance, in the neural network theory, the problem of assigning given vectors to fixed-points of a system together with their stability analysis becomes a major issue. In fact, in implementing *associative memory* or *pattern recognition* using a dynamical neural network, each vector representing information to be stored need be assigned as a locally asymptotically stable fixed-point of the network. In this way, the content of information can be recalled by only giving an incomplete content or a portion of the memorized information as an initial state of the network because the state converges to the fixed-point storing the true information.

There have been studied two types of neural networks, i.e., the discrete state space type with discrete time and the continuous state space type with discrete or continuous time. For the discrete type, a variety of methods for assigning given vectors to locally asymptotically stable fixed-points have been studied [1]-[6]. Further, for the continuous type, quite a number of investigations have appeared in the literature, see, e.g., [7]-[9] and the references there. However, the problem has not been well understood in the system theoretical setting and seems to be worthwhile to be reformulated and further investigated in a unified manner from the viewpoint of systems theory.

This paper deals with systems of the continuous state type with continuous time. Among many important research subjects in dynamical systems, this paper particularly focuses on the fixed-point assignment problem by reformulating and studying it from the viewpoint of the mathematical system theory, and further applies the results for implementing associative memory using a neural network.

In Section II, the fixed-point assignment problem is formulated in the framework of Hilbert space and the stability of the assigned fixed-points is discussed, particularly emphasizing the locally asymptotical stability. Then, in Section III, further properties of the fixed-point assigned system are investigated, and in Section IV the result is applied for implementing associative memory using a neural network, including some numerical examples to demonstrate the performance. Finally, some concluding remarks are given in Section V.

II. Construction of Dynamical Systems with Given Fixed-points

Let $\mathcal{H}$ denote a real separable Hilbert space and consider a dynamical system over $\mathcal{H}$ described as

$$\dot{x}(t) = f(x(t)), \quad x(0) = x_0 \in \mathcal{H} \tag{1}$$

where $f : \mathcal{H} \to \mathcal{H}$ is a sufficiently smooth function. Further, let $\mathcal{P} = \{v_1, \cdots, v_r\}$ be a set of linearly independent vectors in $\mathcal{H}$ and consider the problem of assigning all the vectors in $\mathcal{P}$ as locally asymptotically stable fixed-points of a system of the system (1).

To begin with, first denote by $\mathcal{H}_{\mathcal{P}}$ the subspace spanned by the vectors in $\mathcal{P}$, and considering $\mathcal{P} = \{v_1, \cdots, v_r\}$ as a basis of $\mathcal{H}_{\mathcal{P}}$ and identifying its dual space $\mathcal{H}_{\mathcal{P}}^*$ as itself $\mathcal{H}_{\mathcal{P}}$, introduce its dual basis $\mathcal{P}^* = \{v_1^*, \cdots, v_r^*\} \subset \mathcal{H}_{\mathcal{P}}$, which is defined as the following duality conditions:

$$\langle v_i, v_j^* \rangle = \delta_{ij} := \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \tag{2}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product on $\mathcal{H}$. Then arbitrary vector $x \in \mathcal{H}$ can be uniquely expressed in the form

$$x = y + w = \sum_{i=1}^{r} \xi_i v_i + w, \quad y \in \mathcal{H}_{\mathcal{P}}, w \in \mathcal{H}_{\mathcal{P}}^{\perp} \tag{3}$$

where $\mathcal{H}_{\mathcal{P}}^{\perp}$ denotes the orthogonal complement of and $\xi_i := \langle x, v_i^* \rangle$. Further the dual vector $x^* \in \mathcal{H}$ of x expressed in (3) is defined as

Hiroshi Inaba is with the Department of Information Sciences, Tokyo Denki University , Hatoyama-machi, Hiki-gun, Saitama 350-0394 Japan (e-mail: inaba@cck.dendai.ac.jp).

Takafimi Ooki is a graduate student in the Information Sciences Department, Tokyo Denki University , Hatoyama-machi, Hiki-gun, Saitama 350-0394 Japan (e-mail: cfa27470@syd.odn.ne.jp)

Keylan Alimhan is with the Department of Information Sciences, Tokyo Denki University , Hatoyama-machi, Hiki-gun, Saitama 350-0394 Japan. (e-mail: Keylan@j.dendai.ac.jp)

$$x^* = y^* + w^* = \sum_{i=1}^{r} \xi_i v_i^* + w^*, \quad y^* \in \mathcal{H}_{\mathcal{P}}, w^* \in \mathcal{H}_{\mathcal{P}}^{\perp} \quad (4)$$

where the vector $w^* \in \mathcal{H}_{\mathcal{P}}^{\perp}$ is determined via the relation that

when w is represented as $w = \sum_{\forall k} \gamma_k u_k$ for a basis $\{u_k\}$ of $\mathcal{H}_{\mathcal{P}}^{\perp}$

then

$$w^* = \sum_{\forall k} \gamma_k u_k^* . \quad (5)$$

where $\{u_k^*\}$ is the dual basis. The norm often used in this study is given as

$$\|x\| := \langle x, x^* \rangle^{\frac{1}{2}}, \quad x \in \mathcal{H}. \quad (6)$$

It should be noted that this norm is different from the one $\|x\|_{\mathcal{H}} := \langle x, x \rangle^{\frac{1}{2}}$ induced by the inner product on $\mathcal{H}$, but these norms are equivalent. Further, it should be also noted that the following relations hold: For arbitrary vector $x \in \mathcal{H}$ expressed as in (3) one has

$$\begin{aligned} \|x\|^2 &= \langle x, x^* \rangle = \langle y, y^* \rangle + \langle w, w^* \rangle \\ &= \|y\|^2 + \|w\|^2 = \sum_{i=1}^{r} \xi_i^2 + \|w\|^2 \end{aligned} \quad (7)$$

and moreover for arbitrary vectors $x, z \in \mathcal{H}$

$$\langle x, z^* \rangle = \langle x^*, z \rangle . \quad (8)$$

Now, based on the works [7]-[9] introduce a dynamical system defined over $\mathcal{H}$ as follows:

$$\begin{aligned} \dot{x}(t) &= a \sum_{i=1}^{r} \langle x(t), v_i^* \rangle v_i - \sum_{i=1}^{r} \sum_{j=1}^{r} b_{ij} \langle x(t), v_j^* \rangle^2 \\ &\times \langle x(t), v_i^* \rangle v_i - a \langle x(t), x(t)^* \rangle x(t), \quad x(0) = x_0 \in \mathcal{H} \end{aligned} \quad (9)$$

where $a \in \mathbb{R}$ is a constant and $b_{ij} = b_{ji} \in \mathbb{R}$ are constants with the property $b_{ii} = 0$. Then for arbitrary initial condition $x_0 \in \mathcal{H}$ there exists a unique solution $x(t)$ in $\mathcal{H}$ [10]. Further it is easily seen that the origin $x = 0$ is a fixed-point and every vector v_i in $\mathcal{P} = \{v_1, \cdots, v_r\}$ and the negative "$-v_i$" are also fixed-points of (9). In general, if $p \in \mathcal{H}$ is a fixed-point, then so is the negative "$-p$". Therefore, when some statement on a fixed-point is obtained it will be interpreted hereafter that the same is applied for the negative point.

Furthermore, it can be shown that under certain assumptions on the system (9) each v_i is locally asymptotically stable. The proof requires several steps. First, introduce a functional $V : \mathcal{H} \to \mathbb{R}$, called a *generalized energy functional*, of the form

$$\begin{aligned} V(x) &:= -\frac{1}{2} a \sum_{i=1}^{r} \langle x, v_i^* \rangle^2 + \frac{1}{4} \sum_{i=1}^{r} \sum_{j=1}^{r} b_{ij} \langle x, v_i^* \rangle^2 \langle x, v_j^* \rangle^2 \\ &+ \frac{1}{4} a \langle x, x^* \rangle^2 \end{aligned} \quad (10)$$

and compute the Fréchet derivative of $V(x)$, denoted $V_x(x)$, which is obtained as

$$\begin{aligned} V_x(x) &= -a \sum_{i=1}^{r} \langle x, v_i^* \rangle v_i^* + \sum_{i=1}^{r} \sum_{j=1}^{r} b_{ij} \langle x, v_j^* \rangle^2 \langle x, v_i^* \rangle v_i^* \\ &+ a \langle x, x^* \rangle x^* . \end{aligned} \quad (11)$$

Notice that $V_x(x)$ is a continuous linear functional on $\mathcal{H}$, depending on $x \in \mathcal{H}$ [10].

Now, for notational simplicity, introduce the following notations:

$$\begin{aligned} &\xi := [\xi_1 \quad \cdots \quad \xi_r]^T \text{ where } \xi_i := \langle x, v_i \rangle \text{ for } i = 1, \cdots, r \\ &B := \left[b_{ij} \right]_{i,j=1}^{r} \in \mathbb{R}^{r \times r}. \end{aligned} \quad (12)$$

Then the system (9) and the functional (10) together with the Fréchet derivative (11) can be expressed as

$$\begin{aligned} \dot{x} &= a \sum_{i=1}^{r} \xi_i v_i - [\xi_1 v_1 \quad \cdots \quad \xi_r v_r] \\ &\times B \left[\xi_1^2 \quad \cdots \quad \xi_r^2 \right]^T - a x \|x\|^2 \end{aligned} \quad (13)$$

$$\begin{cases} V(x) = -\frac{1}{2} a \sum_{i=1}^{r} \xi_i^2 + \frac{1}{4} \left[\xi_1^2 \quad \cdots \quad \xi_r^2 \right] \\ \qquad \times B \left[\xi_1^2 \quad \cdots \quad \xi_r^2 \right]^T + \frac{1}{4} a \|x\|^4 \\ V_x(x) = -a \sum_{i=1}^{r} \xi_i v_i^* + [\xi_1^2 \quad \cdots \quad \xi_r^2] \\ \qquad \times B [\xi_1 v_1^* \quad \cdots \quad \xi_r v_r^*]^T + a \|x\|^2 x^* . \end{cases} \quad (14)$$

Now, the following lemma is proved.

LEMMA I. Let $\mathcal{P} = \{v_1, \cdots, v_r\}$ be a set of linearly independent vectors in $\mathcal{H}$ with the dual basis $\mathcal{P}^* = \{v_1^*, \cdots, v_r^*\} \subset \mathcal{H}_{\mathcal{P}}$ as characterized by (2), and consider the functional $V(x)$ given by (10).

Then if $a > 0$, $b_{ij} = b_{ji} > 0$ and $b_{ii} = 0$ for $i, j = 1, \cdots, r$, $V(x)$ attains a local minimum at each point v_k, and all the minimum values are equal and are given as

$$V(v_k) = -\frac{1}{4} a < V(0) = 0, \quad k = 1, \cdots, r . \quad (15)$$

PROOF. First, it follows from (11) and (2) that

$$V_x(v_k) = -a\sum_{i=1}^{r}\langle v_k, v_i^*\rangle v_i^* + \sum_{i=1}^{r}\sum_{j=1}^{r} b_{ij}\langle v_k, v_j^*\rangle^2 \langle v_k, v_i^*\rangle v_i^* + a\langle v_k, v_k^*\rangle v_k^* = 0 \tag{16}$$

and hence each v_k and its negative $-v_k$ are extreme points of $V(x)$. Further (15) follows easily from (10). To show that each extreme point v_k attains a local minimum, first observe from the identity (3) and (14) that, for an arbitrary $x \in \mathcal{H}$ expressed as

$$x = \sum_{i=1}^{r} \xi_i v_i + w \text{ with } \xi_i := \langle x, v_i^*\rangle \text{ and } w \in \mathcal{H}_{\mathcal{P}}^{\perp},$$

one has

$$\begin{aligned} V(x) &= -\frac{1}{2}a\sum_{i=1}^{r}\xi_i^2 + \frac{1}{4}\begin{bmatrix}\xi_1^2 & \cdots & \xi_r^2\end{bmatrix} \\ &\quad \times B\begin{bmatrix}\xi_1^2 & \cdots & \xi_r^2\end{bmatrix}^T + \frac{1}{4}a\left(\sum_{i=1}^{r}\xi_i^2 + \|w\|^2\right)^2 \\ &\geq -\frac{1}{2}a\sum_{i=1}^{r}\xi_i^2 + \frac{1}{4}\begin{bmatrix}\xi_1^2 & \cdots & \xi_r^2\end{bmatrix} \\ &\quad \times B\begin{bmatrix}\xi_1^2 & \cdots & \xi_r^2\end{bmatrix}^T + \frac{1}{4}a\left(\sum_{i=1}^{r}\xi_i^2\right)^2 =: \bar{V}(\xi) \end{aligned} \tag{17}$$

where the assumption $a > 0$ has been used. Therefore $V(x)$ is bounded from the below and its all local minimums occur on the subspace $\mathcal{H}_{\mathcal{P}}$.

To proceed further, first compute the first and second derivatives of the lower bound $\bar{V}(\xi)$ defined in (17) to obtain

$$\begin{cases} \dfrac{\partial \bar{V}}{\partial \xi_k}(\xi) = \displaystyle\sum_{i=1}^{r}(a + b_{ki})\xi_i^2\xi_k - a\xi_k, \quad \text{for } k = 1, \cdots, r \\ \dfrac{\partial^2 \bar{V}}{\partial \xi_k \partial \xi_l}(\xi) = \begin{cases} \displaystyle\sum_{i=1}^{r}(a + b_{ki})\xi_i^2 + 2a\xi_k^2 - a, & k = l \\ 2(a + b_{kl})\xi_k\xi_l, & k \neq l \quad \text{for } l = 1, \cdots, r. \end{cases} \end{cases} \tag{18}$$

Then it follows from (18) that, since each extreme point v_i corresponds to the vector $\xi = [\xi_1 \ \cdots \ \xi_r]$ with $\xi_i = \pm 1$ and all the other elements $\xi_j = 0$, the matrix $\frac{\partial^2 \bar{V}}{\partial \xi \partial \xi}$ formed by the second derivatives evaluated at this extreme point v_i is obtained as

$$\left.\frac{\partial^2 \bar{V}}{\partial \xi \partial \xi}\right|_{\substack{\xi_i = \pm 1 \\ \xi_j = 0, j \neq i}} = \mathrm{diag}\left\{b_{1i}, \cdots, b_{(i-1)i}, 2a, b_{(i+1)i}, \cdots, b_{ri}\right\} > 0,$$

for $i = 1, \cdots, r$ where the assumptions $a > 0$ and $b_{ij} = b_{ji} > 0$ have been used to get the positivity. Therefore, each extreme point v_k attains a local minimum of $\bar{V}(\xi_1, \cdots \xi_r)$. □

Now our main theorem is stated and proved as follows. The proof is omitted and will be given elsewhere.

THEOREM 1. Let $\mathcal{P} = \{v_1, \cdots, v_r\}$ be a set of linearly independent vectors in $\mathcal{H}$ with $\mathcal{P}^* = \{v_1^*, \cdots, v_r^*\} \subset \mathcal{H}_{\mathcal{P}}$ of the dual vectors as characterized by (2), and consider the functional $V(x)$ given by (10).

Then, if $a > 0$, $b_{ij} = b_{ji} > 0$ and $b_{ii} = 0$ for $i, j = 1, \cdots, r$, then each fixed-point v_k of the system (9) is locally asymptotically stable.

PROOF. First notice from (9) and (11) that $\dot{x}(t) = V_x^*(x(t))$. Using this fact, one easily obtains

$$\begin{aligned} \frac{dV}{dt}(x(t)) &= \langle V_x(x(t)), \dot{x}(t)\rangle = -\langle V_x(x(t)), V_x^*(x(t))\rangle \\ &= -\|V_x(x(t))\|^2 \leq 0 \end{aligned} \tag{19}$$

which implies that $V(x(t))$ always decreases along the trajectory $x(t)$ of the system (9). Now recall that each v_k in $\mathcal{P} = \{v_1, \cdots, v_r\}$ is a fixed-point of the system (9) and further Lemma 1 ensures that the point attains a local minimum of $V(x)$. Therefore (19) implies that every trajectory $x(t)$ starting from a sufficiently small vicinity of any fixed-point v_k converges asymptotically to the point v_k. This completes the proof. □

III. BASIC PROPERTIES OF THE CONSTRUCTED SYSTEM

In the previous section, a method for constructing a dynamical system for which a set of prescribed vectors is assigned as its fixed-points was studied, and further it was shown under certain assumptions that each assigned fixed-point is locally asymptotically stable. In this section, more properties of the constructed system are investigated.

First, we note that using the expression (3) it is possible to describe the system (9) as the following two coupled subsystems:

$$\dot{\xi}_i(t) = \left\{a - \sum_{j=1}^{r}(a + b_{ij})\xi_j^2(t) - \|w\|^2\right\}\xi_i(t), \ i = 1, \cdots, r \tag{20a}$$

$$\dot{w}(t) = -a\left(\sum_{j=1}^{r}\xi_j^2(t) + \|w(t)\|^2\right)w(t) \tag{20b}$$

where the solution $x(t)$ is given by

$$x(t) = \sum_{i=1}^{r}\xi_i(t)v_i + w(t).$$

The subsystem (20b) implies that the solution $w(t)$ belonging to the orthogonal subspace $\mathcal{H}_{\mathcal{P}}^{\perp}$ always satisfies

$$\lim_{t \to \infty} w(t) = 0. \tag{21}$$

The following theorem states more about fixed-points. The proof is omitted and will be given elsewhere. Hereafter, for notational simplicity, the argument "t" in variables may be omitted when no confusion seems to be possible.

THEOREM 2. Consider the system (9) with $a>0$, $b_{ij}=b_{ji}>0$ and $b_{ii}=0$ for $i,j=1,\cdots,r$. Then:

(i) Every fixed-point of the system (9) satisfies $w=0$, and therefore every fixed-point lies on the subspace $\mathcal{H}_{\mathcal{P}}$ and is expressible in the form

$$p=\sum_{i\in Q_1}\xi_i v_i \tag{22}$$

(ii) The number N_{fixed} of all possible fixed-points of the system (9) is given as

$$N_{fixed}=\sum_{m=0}^{r}2^m\ {}_rC_m. \tag{23}$$

(iii) Further among these fixed-points the r fixed-points $v_1,\cdots,v_r$ as well as their negatives are locally asymptotically stable, and the fixed-point $x=0$ (i.e., all $\xi_i=0$ and $w=0$) is unstable. □

REMARK. First notice from THEOREM 2 (i) that each fixed-point v_i corresponds to the vector $\xi:=[\xi_1\ \cdots\ \xi_r]^T$ with $\xi_i=1$ and $\xi_j=0$ for all $j\neq i$. Therefore, THEOREM 2 (iii) states that the fixed-points corresponding to $\xi_i=\pm1$ and $\xi_j=0$ for all $j\neq i$ are locally asymptotically stable and the fixed-point corresponding to $\xi_j=0$ for all $j=1,\cdots,r$ is unstable. But it says no statement on the stability for other fixed-points. The authors' conjecture for these fixed-points is as follows:

(i) For a fixed-point corresponding to the vector with only m components $\xi_{i_1},\cdots,\xi_{i_m}\neq 0$,

$$0<\xi_{i_1}{}^2=\cdots=\xi_{i_m}{}^2<1. \tag{24}$$

(ii) Such a fixed-point is not asymptotically stable but a saddle point.

This conjecture will be discussed elsewhere in the near future. □

IV. APPLICATION TO ASSOCIATIVE MEMORY

The basic idea of implementing *associative memory* or pattern recognition using a neural network is as follows:

(i) First, the desired information to be stored is represented by some vectors in an appropriately chosen abstract space, and a neural network defined over the abstract space is constructed in such a way that these vectors are assigned to locally asymptotically stable fixed-points.

(ii) Then, any desired information can be recalled by giving only an incomplete content or a portion of the memorized information as an initial state of the network so that the state of the network converges to a fixed-point, which represents the desired information.

4.1 Associative Memory for 2-Dimensional Images

In our example, a certain number of two-dimensional images (pictures, fingerprints, letters, etc.) are taken as a set $\mathcal{P}=\{v_1,\cdots,v_r\}$ of prototype vectors to be memorized in a neural network. More precisely, introduce a compact domain $\mathcal{D}\subset\mathbb{R}^2$ together with the Hilbert space $\mathcal{H}=\mathfrak{L}_2(\mathcal{D})$ of square integrable functions where these functions represent all possible two-dimensional images, and choose a set $\mathcal{P}=\{v_1,\cdots,v_r\}$ of prototype vectors in $\mathfrak{L}_2(\mathcal{D})$. Then, construct a neural network of the form (9) with $a>0$, $b_{ij}=b_{ji}>0$ and $b_{ii}=0$ for $i,j=1,\cdots,r$.

In practical situations, an initial image $x(0)\in\mathfrak{L}_2(\mathcal{D})$ is a corrupted version of some prototype vector v_k due to a noise and/or a deformation. In this study, we assume that the noise is additive and the deformation is described by *a deformation operator* $T(\theta):\mathfrak{L}_2(\mathcal{D})\to\mathfrak{L}_2(\mathcal{D})$ depending on a parameter vector $\theta\in\mathbb{R}^m$. Thus, we assume that

$$x(0)=T(\theta)v_k+\eta \tag{25}$$

for some prototype vector v_k where $\eta\in\mathfrak{L}_2(\mathcal{D})$ is an additive noise. It is assumed that $\theta\in\mathbb{R}^m$ is unknown but the functional structure $T(\bullet)$ is known and mapping $\theta\mapsto T(\theta)$ is sufficiently smooth. Further it is assumed that the inverse $S(\theta)=T^{-1}(\theta)$ exists for all possible $\theta\in\mathbb{R}^m$ and there exists a unique nominal parameter vector $\bar{\theta}\in\mathbb{R}^m$ such that

$$T(\bar{\theta})x=x\ \text{ for all }x\in\mathcal{H}.$$

The deformation operator $T(\theta)$ we consider is described via an Affine transformation of the two-dimensional domain $\mathcal{D}\subset\mathbb{R}^2$ as follows: Let $x(\mu)\in\mathfrak{L}_2(\mathcal{D})$ be an arbitrary image, and denote $\mu=[\mu_1\ \mu_2]^T\in\mathcal{D}\subset\mathbb{R}^2$ and $\rho=[\rho_1\ \rho_2]^T\in\mathbb{R}^2$. Then our operators $T(\theta)$ and $S(\theta)$ are given by

$$\begin{aligned}T(\theta)x(\mu)&:=[\det\Omega]^{-\frac{1}{2}}x(\Omega\mu+\rho)\\ S(\theta)x(\mu)&=[\det\Omega]^{\frac{1}{2}}x[\Omega^{-1}(\mu-\rho)]\end{aligned} \tag{26a}$$

where

$$\begin{cases}\Omega\mu+\rho:=\begin{bmatrix}\omega_{11}&\omega_{12}\\ \omega_{21}&\omega_{22}\end{bmatrix}\mu+\begin{bmatrix}\rho_1\\ \rho_2\end{bmatrix}\\ \theta:=[\omega_{11}\ \omega_{12}\ \omega_{21}\ \omega_{22}\ \rho_1\ \rho_2]^T.\end{cases} \tag{26b}$$

4.2 The Stability Analysis for Deformation

In this case, to improve the performance of associative memory, the deformation structure is taken into account. To make such an improvement, we introduce *a modified generalized energy functional* as follows:

$$\begin{aligned} V(x,\theta) := & -\frac{1}{2}a\sum_{i=1}^{r}\left\langle x,\hat{v}_i^*\right\rangle^2 \\ & +\frac{1}{4}\sum_{i=1}^{r}\sum_{j=1}^{r}b_{ij}\left\langle x,\hat{v}_i^*\right\rangle^2\left\langle x,\hat{v}_j^*\right\rangle^2 \\ & +\frac{1}{4}a\left\langle x,x^*\right\rangle^2+\frac{1}{2}c\left\langle \theta-\bar{\theta},\theta-\bar{\theta}\right\rangle \end{aligned} \quad (27)$$

where $c > 0$ is a constant and

$$\hat{v}_i := S(\theta)v_i \text{ and } \hat{v}_i^* := S(\theta)v_i^*, \quad i=1,\cdots,r.$$

Then it can be shown that the duality for $\hat{\mathcal{P}} = \{\hat{v}_1,\cdots,\hat{v}_r\}$ is preserved, i.e.,

$$\left\langle \hat{v}_i,\hat{v}_j^*\right\rangle = \left\langle v_i,v_j^*\right\rangle = \delta_{ij}.$$

Corresponding to (27), a modification of the system (9) is introduced as follows:

$$\begin{aligned} \dot{x}(t) = & \, a\sum_{i=1}^{r}\left\langle x(t),\hat{v}_i^*\right\rangle\hat{v}_i \\ & -\sum_{i=1}^{r}\sum_{j=1}^{r}b_{ij}\left\langle x(t),\hat{v}_j^*\right\rangle^2\left\langle x(t),\hat{v}_i^*\right\rangle\hat{v}_i \\ & -a\left\langle x(t),x(t)^*\right\rangle x(t), \quad x(0)=x_0\in\mathcal{H}. \end{aligned} \quad (28)$$

Further, a dynamics for the parameter vector θ is introduced as follows:

$$\dot{\theta}(t) = -\gamma V_\theta^*(x(t),\theta(t)), \quad \theta(0)=\theta_0\in\mathbb{R}^m \quad (29)$$

where $\gamma > 0$ is some constant. It should be noticed that if $\theta\in\mathbb{R}^m$ coincides with the nominal value $\bar{\theta}$ then (27) and (28) become the original system (9) and the original generalized energy functional (10), respectively.

THEOREM 3. Consider *the combined system* of (28) and (29) with $a>0$, $b_{ij}=b_{ji}>0$ and $b_{ii}=0$ for $i,j=1,\cdots,r$. Then each $(v_k,\bar{\theta})$ is a locally asymptotically stable fixed-point of this combined system.

PROOF. First, compute the Fréchet derivatives of $V(x,\theta)$ with respect to x and θ, respectively, to obtain

$$\begin{aligned} V_x(x,\theta) = & -a\sum_{i=1}^{r}\left\langle x,\hat{v}_i^*\right\rangle\hat{v}_i^* \\ & +\sum_{i=1}^{r}\sum_{j=1}^{r}b_{ij}\left\langle x,\hat{v}_j^*\right\rangle^2\left\langle x,\hat{v}_i^*\right\rangle\hat{v}_i^*+a\left\langle x,x^*\right\rangle x^* \end{aligned} \quad (30)$$

$$\begin{aligned} V_\theta(x,\theta) := & -a\sum_{i=1}^{r}\left\langle x,\hat{v}_i^*\right\rangle\left\langle x,\frac{\partial}{\partial\theta}\hat{v}_i^*\right\rangle \\ & +\sum_{i=1}^{r}\sum_{j=1}^{r}b_{ij}\left\langle x,\hat{v}_i^*\right\rangle^2\left\langle x,\hat{v}_j^*\right\rangle\left\langle x,\frac{\partial}{\partial\theta}\hat{v}_i^*\right\rangle \\ & +a\left\langle x,x^*\right\rangle\sum_{i=1}^{r}\left\langle x,\frac{\partial}{\partial\theta}x^*\right\rangle+c(\theta-\bar{\theta})^*. \end{aligned} \quad (31)$$

Then notice that the right hand side of (31) evaluated at $(v_k,\bar{\theta})$ vanishes and from (34) that $V_\theta^*(v_k,\bar{\theta})=0$. Therefore, each $(v_k,\bar{\theta})$ is a fixed-point of the combined system. Further one obtains

$$\begin{aligned} \frac{dV}{dt}(x(t),\theta(t)) & = \left\langle V_x(x,\theta),\dot{x}\right\rangle+\left\langle V_\theta(x,\theta),\dot{\theta}\right\rangle \\ & = -\left\langle V_x(x,\theta),V_x^*(x,\theta)\right\rangle-\gamma\left\langle V_\theta(x,\theta),V_\theta^*(x,\theta)\right\rangle \\ & = -\|V_x(x,\theta)\|^2-\gamma\|V_\theta(x,\theta)\|^2\le 0 \end{aligned} \quad (32)$$

and hence, if the initial state $(x(0),\theta(0))$ for (28) and (29) is sufficiently close to $(v_k,\bar{\theta})$ for some k, then

$$\lim_{t\to\infty}(x(t),\theta(t)) = (v_k,\bar{\theta}), \quad (33)$$

showing that $(v_k,\bar{\theta})$ is locally asymptotically stable. □

V. NUMERICAL EXAMPLES

Our simple example used to demonstrate the results in the previous sections is described as follows: First, we choose $r=10$ two-dimensional images as the set $\mathcal{P}=\{v_1,\cdots,v_{10}\}$ of prototype vectors in $\mathfrak{L}_2(\mathcal{D})$, including letters, flowers, bird and some simple pictures. Then, each 2-dimensional image v_k is divided into $128\times128 = 16{,}384$ pixels, and is approximately represented by an order 128×128 matrix with the entries equal to the values of $v_k(\mu)$ evaluated at the pixels.

The parameters of the combined system of (28) and (29) are chosen as $a=1, b_{ij}=1, c=2, \gamma=3$. The initial state $x(0)=T(\theta)v_9+\eta$ given by (26) is constructed from a flower image v_9 by setting $\theta_1=\omega_{11}=0.7, \omega_{22}=1$ with all the other parameters equal to zero and adding a white noise with zero mean and variance 1. Therefore, the structural deformation in this case contains only an enlargement in the horizontal direction by $1/0.7\cong1.429$ and no other deformation is introduced. The initial value $\theta(0)$ is set as

$$\theta(0)=\bar{\theta}=\begin{bmatrix}1 & 0 & 1 & 0 & 0 & 0\end{bmatrix}^T.$$

Fig. 1 shows the result using (20) or equivalently (9) with this initial state $x(0)$, i.e., the one without taking into account the deformation structure (25). First, notice that $\xi_8(t)\to1$ implies $x(t)\to v_8$ where v_8 is another flower similar to v_9, and that the other $\xi_j(t)$'s converge to zero and in particular the noise term $\|w(t)\|^2$ sharply reduces to zero. It is also seen that the initial values satisfy $\xi_9(0) < \xi_8(0)$ which leads to $\xi_8(t) \rightarrow 1$, converging to an incorrect prototype image v_8.

On the other hand, Figs. 2-3 show the results using (28) and (29), i.e., the case taking into account the deformation structure (25). In this case, $\xi_9(t) \rightarrow 1$ which implies the convergence to the correct prototype image v_9, as expected.

VI. Concluding Remarks

This paper studied from the viewpoint of the systems and control theory the problem of constructing a dynamical system having given vectors as its locally asymptotically stable fixed-points, and also discussed basic properties of the resultant system. Further, the result obtained was applied to neural networks to implement associative memory for 2-dimensional images, and a new method for handling structural deformation of images was proposed. Finally, some numerical examples were presented to demonstrate the performance.

Acknowledgement

This work was supported in part by the Japanese Ministry of Education, Science, Sports and Culture under the 21st Century Center of Excellence (COE) Program.

References

[1] L. Personnaz, I. Guyou and G. Dreyfun, Collective computational properties of neural networks: New learning mechanism, *Physical Reviews A*, vol. 34, pp. 4217-4228, 1986.

[2] T. Kohonen, *Self-Organization and Associative Memory*, Springer-Verlag, 1988 (Second Edition).

[3] Yves Kamp and Martin Hasler, *Recursive Neural Networks for Associative Memory*, Wiley, 1990.

[4] A. Michel and J. A. Farrell, Associative memories via artificial neural networks, *IEEE Control Magazine*, pp. 6-17, April 1990.

[5] K. Ishii and H. Inaba, Associative memories using dynamical neural networks with limit cycles, *Proceedings of the ISCIE International Symposium on Stochastic Systems Theory and Its Applications*, pp. 175-180, November 1993, Osaka, Japan.

[6] Y. Shoji and H. Inaba, A module neural network and its basic behaviors, *Proceedings of IEEE International Conference on Neural Networks*, vol. 2, pp. 1305-1310, June 1997, Houston, Texas, USA.

[7] H. Haken, *Synergetic Computers and Cognition*, Springer-Verlag, Berlin, 2004.

[8] A. Fuchs and H. Haken, Pattern recognition and associative memory as dynamical processes in a synergetic system, I, Translational invariance, selective attention, and decomposition of scenes, *Biological Cybernetics*, vol. 60, pp. 17-22, 1988.

[9] A. Fuchs and H. Haken, Pattern recognition and associative memory as dynamical processes in a synergetic system, II, decomposition of complex scenes, simultaneous with respect to translation, rotation, and scaling with A, *Biological Cybernetics*, vol. 60, pp. 107-109, 1988.

[10] G. E. Ladas and V. Lakshmikantham, *Differential Equations in Abstract Spaces*, Academic Press, New York, 1972.

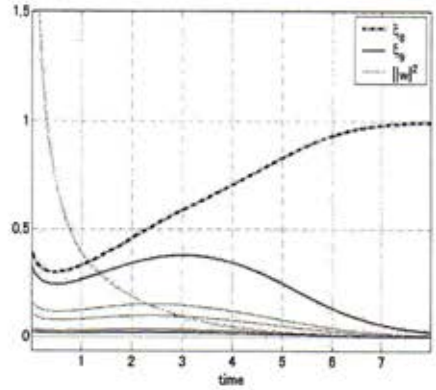

Fig. 1. The results using (20) or (9), converging to an incorrect prototype vector v_8

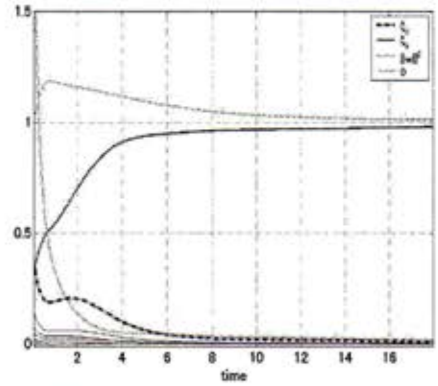

Fig. 2. The results using (28) and (29), converging to the correct prototype vector v_9

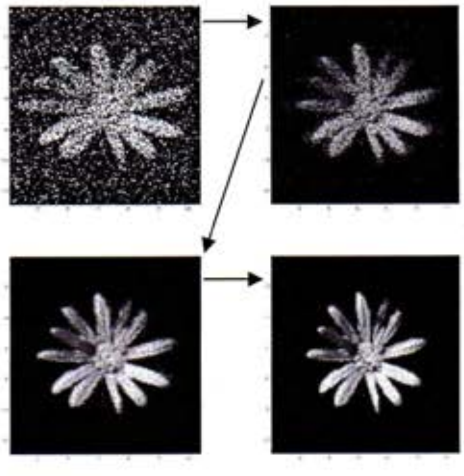

Fig.3. The sequence of images converging to the correct image v_9