

# 关于一般存储器神经网络的学习收敛性研究<sup>\*</sup>

彭宏京<sup>1</sup> 陈松灿<sup>2</sup>

(南京工业大学计算机系 南京210009)<sup>1</sup> (南京航空航天大学计算机系 南京210016)<sup>2</sup>

**摘要** 本文在一类称为一般存储器神经网络(General Memory Neural Network (GMNN))的统一框架下来研究学习收敛性。该一般模型类的结构由三部分组成:输入空间量化、存储器地址产生器、查表式某种组合输出。当产生的地址是固定有限的个数以及网络输出是线性求和时,可以证明 GMNN 能在最小平方误差意义下收敛。CMAC (Cerebellar Model Articulation Controller)、SLLUP(Single-Layer Lookup Perceptrons)是该类模型的典型代表。本文的意义在于为构造新的基于局部学习的神经网络模型提供理论指导,最后给出了这种构造的两个例子——SDM (Sparse Distributed Memory)和 SLLUP 的两个推广模型。

**关键词** 存储器网络,CMAC,n-tuple,SLLUP,SDM,学习收敛性

## Research on Learning Convergence of General Memory Neural Network

PENG Hong-Jing<sup>1</sup> CHEN Song-Can<sup>2</sup>

(Dept. of Computer Science, Nanjing University of Technology, Nanjing 210009)<sup>1</sup>

(Dept. of Computer Science, Nanjing University of Aeronautics and Astronautics, Nanjing 210016)<sup>2</sup>

**Abstract** In this paper, we will concentrate on learning convergence of a class of neural network architectures named general memory neural network (GMNN) that consists of: input space quantization, memory address generator, combination output by memory lookup operations. If the number of generated addresses is fixed the output of network is given by summation operator, the learning convergence of GMNN to the least square error can be proved. Both CMAC(Cerebellar Model Articulation Controller)and SLLUP(Single-Layer Lookup Perceptrons)are examples of GMNN. The main purpose of this paper is that it can provide a theoretical instruction about how to construct a new neural network model with local learning. Finally two constructive examples, generalized SDM (Sparse Distributed Memory)and generalized SLLUP models, are given.

**Keywords** Memory network,CMAC,n-tuple,SLLUP,SDM, Learning convergence

## 1. 引言

目前,在人工神经网络领域,像多层感知器(MLPs)和径向基函数网络(RBFN)等那些具有普遍逼近特性的网络结构能充分地逼近任意连续光滑的映射而备受人们的关注<sup>[1,2]</sup>,这些结果为它们的广泛应用提供了理论基础。但理论分析并未解决实际应用中高昂的计算代价问题,人们不得不寻求优化的结构、优化的算法,尽管它们形式化的分析可能很困难。

然而,还有一类应用广泛的网络,如 CMAC<sup>[3]</sup>、n-tuple<sup>[4,5]</sup>,结构简单、计算量小、便于实时和硬件实现,但对它们的非线性映射能力却缺乏足够的认识。这一类网络包含典型的三个部分:输入空间量化、局部感受野、查表式线性求和(求平均等)输出。物理上由一层地址选择器和一层存储器单元(其内容是可调的权重)构成。对于任意一个输入,通过地址选择器,查找相应的权重单元,通过线性求和操作得输出响应。如果在网络整个运行过程中,地址选择器产生的地址只有有限的少数几个,那么存储单元层中需要调整的权重也就是有限的几个,因此,计算量少,速度快。同时,该类网络从拓扑结构看类似于单层线性网络,但后者无任何非线性神经元,而前者的固有的非线性性质隐式地含于地址的选择中(这个特点正好与 RBFN 显式地在隐层使用局部非线性单元来实现

非线性映射形成对照),地址选择器保证了该类网络能实现一大类非线性映射。

Kolcz<sup>[6-9]</sup>等人注意到上述网络类的隐式非线性性质,并从 CMAC 映射中所表现出来的隐含的输入空间量化结合常数权重选择(计算输出响应)原理,延伸到一大类称为一般存储器神经网络框架结构中,差别仅仅是特定的量化形式不同而已。希望在此统一的模型框架下获得对这类网络的一致认识,并致力于把该类网络的隐式非线性显式化(像 RBFN 的基函数那样),文[8,10,11]给出了 CMAC、n-tuple、SLLUP 的那种代表它们如何实现非线性映射的基函数(平均意义下)。换句话说,通过地址选择操作实现非线性映射相当于直接使用该基函数达到的效果。

本文试图在这一统一的框架下来研究该模型类的学习收敛性。主要的出发点源自于以下几点:

1、CMAC 的学习收敛性的理论分析比较充分<sup>[12,13]</sup>。

2、n-tuple 获得了许多成功的应用,理论分析却非常困难,因为难于形式化<sup>[10]</sup>。Tattersall 在文[11]中将 n-tuple 与感知器结合,得到了 SLLUP 作为低通内插器的解释,证实了内插核的存在。因为内插核的大小强烈地依赖输入编码的选择,所以这种解释也只能是渐进意义的。对 n-tuple 以及它的变种,包括使用不同的学习算法下的学习能力还是认识不够。

<sup>\*</sup>国家自然科学基金资助项目(编号:69973021)。彭宏京 博士,研究领域:模式识别、机器学习、数据挖掘。陈松灿 博士,博士生导师,研究领域:模式识别、神经网络、机器学习、SVM。

3、CMAC 的许多结论可以在一般存储器神经网络框架中作纯形式化的推导,从而有望获得包括 SLLUP 在内的一类模型的学习能力的分析。

## 2. GMNN 结构及其学习收敛性

假设我们的学习目标是逼近光滑、连续的函数  $f: \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}$

$$t_s = f(x_s) + \epsilon$$

学习样本对  $\{(x_s, t_s) : s = 1, \dots, N_s\}$ ,  $\epsilon$  是零均值的随机误差分量。

### 2.1 结构

GMNN 的框架结构包括:

1. K 个存储器结点集,每个结点有  $|A_k|$  个地址,  $k = 1, \dots, K$ ;

2. 一组地址产生器,对每个输入  $x$ ,由它形成一个地址向量  $A(x) = [A_1(x), A_2(x), \dots, A_K(x)]$ ,  $A_k(x)$  是第  $k$  个存储器结点地址集;

3. 一个输出函数  $F$ :

$$F: W(x) \rightarrow \mathbb{R}$$

这里  $W(x) = [W_1(x), W_2(x), \dots, W_K(x)]$ ,  $W_k(x)$  是  $A_k(x)$  存储地址单元的内容;

因此,整个网络实现的输入输出映射为:

$$g: \Omega \ni x \rightarrow A(x) \rightarrow W(x) \rightarrow g(x) \in \mathbb{R}$$

4. 通过学习,网络不断调整存储器中的权重,使得某种误差指标  $\pi(f, g)$  最小。

如图1,地址产生部分划分为指针产生器  $I(x)$  和地址解码部分,  $S_k(x) = [S_k^1(x), S_k^2(x), \dots, S_k^{|A_k|}(x)]$ ,  $k = 1, \dots, K$ ; 是选择函数,对于  $S_k^i(x)$ ,若输入属于第  $i$  个量化单元,则返回1,否则返回0;指针产生器表征输入点属于的量化单元,  $I(x) = [I_1(x), I_2(x), \dots, I_K(x)]$ ,  $I_k: x \rightarrow \{1, \dots, |A_k|\}$ ;地址解码器根据指针值获得物理地址,从而,  $N_k(x)$  个存储单元被选中(激活)

参与学习,因此每次参与调整的存储权重总数是  $\sum_{k=1}^K N_k(x) \equiv N_s(x)$  个,相对于  $\sum_{k=1}^K |A_k| \equiv N$  是少量的。GMNN 的地址产生器操作实际执行了输入空间的矢量量化(VQ),因为有限的地址空间中,  $\Omega$  内的某个邻域内的点在每个存储结点被赋予相同的地址,从而,所有  $K$  个结点的地址集相同。

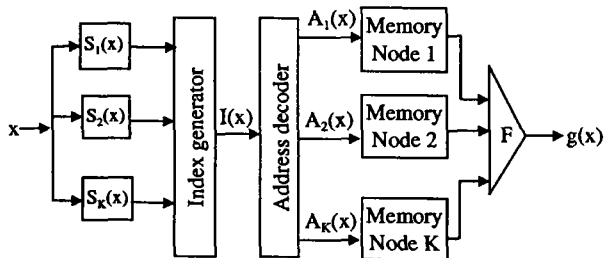


图1 GMNN 的框架结构

本文只考虑网络输出是线性求和的 GMNN。很显然,一旦某种量化方式确定下来,即实现了从输入空间  $\Omega$  到存储地址空间的映射,那么 GMNN 学习收敛性问题就转化为线性优化问题,而它的非线性性质就隐含在上述这种映射中。

为便于分析,将 GMNN 的输入输出关系重新描述为(注意:这里暂不考虑具体的量化方式,我们只在变换后的空间中

讨论。)

第  $s$  个样本的网络输出为

$$y_s = g(x_s) = c_s^T w = [c_{s,1}, c_{s,2}, \dots, c_{s,N}] \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_N \end{bmatrix} = \sum_{j=1}^N c_{s,j} w_j \quad (1)$$

$N = \sum_{k=1}^K |A_k|$ , 即输入空间经  $K$  重量化后,所需存储地址的总和,  $W$  是相应的存储内容(权重),  $c_s$  是存储位置指示向量,

有  $\sum_{k=1}^K N_k(x_s) \equiv N_s(x_s)$  (以下再简记为  $N_s(s)$ ) 个1,其余为0;  $c_{s,j} = 1$  意味着某个地址被选中(激活),比如某个  $A_k(x_s)$  中的某个地址。  $N_s$  个训练样本,网络输出写成向量形式

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_{N_s} \end{bmatrix} = \begin{bmatrix} c_1^T \\ \vdots \\ c_{N_s}^T \end{bmatrix} w \equiv Cw \quad (2)$$

### 2.2 学习收敛性

#### 2.2.1 最小平方回归

目的是寻找  $w$  使

$$E(w) = \frac{1}{2} \|T - y\|^2 = \frac{1}{2} \|T - Cw\|^2 \quad (3)$$

最小,  $T = [t_1, t_2, \dots, t_{N_s}]^T$  是目标输出向量。

$$E(w) = \frac{1}{2} (T^T T - 2T^T Cw + w^T C^T Cw) \quad (4)$$

$$\nabla E = -C^T T + C^T Cw \quad (5)$$

$E$  是连续、可微、有界的,所以对于满足  $C^T T = C^T Cw$  的  $w$  一定能达到其最小点。如果  $C^T C$  是正定的,则有唯一解:

$$w = (C^T C)^{-1} C^T T \quad (6)$$

此时,在几何上相当于找一个  $(N+1)$  维过原点的超平面按平方误差意义最好地拟合  $N_s$  个点  $(c_s^T, t_s)$ ,  $s = 1, \dots, N_s$ 。下节将证明 GMNN 通过误差迭代学习能逼近式(6)这个解。

2.2.2 最小平方误差收敛 CMAC 作为一般存储器神经网络的特例,采取分层分块(超矩形)的量化方式,它的学习行为已经有很多结果。Wong and Sideris<sup>[12]</sup> 的学习收敛性证明是在假定输入空间量化足以保证  $C^T C$  是正定的情况下给出的。相反,Chun-Shin Lin and Ching-Tsan Chiang<sup>[13]</sup> 放松了这一要求,只要矩阵  $C$  的任何元素都不会大于或等于2,其意义是对任一输入,  $K$  个结点地址映射,决不会有两个映射到同一地址的情况出现,显然,GMNN 的定义严格保证了这一条件满足,并且文[13]的推导方法完全可以移植过来。

设  $N_s$  个训练样本依次反复地送入 GMNN,第  $s$  个样本在第  $i$  次迭代的权值修正,由(5)式

$$w_s^{(i)} = w_s^{(i-1)} + \Delta w_s^{(i)} = w_s^{(i-1)} + \frac{\alpha}{N_s(s-1)} c_{s-1} (t_{s-1} - c_{s-1}^T w_s^{(i-1)}) \quad (7)$$

利用(7)式可得  $w_s$  的相继两次迭代的变化量为

$$\begin{aligned} \Delta w_s^{(i)} &= w_s^{(i+1)} - w_s^{(i)} = w_s^{(i-1)} + \Delta w_s^{(i+1)} - (w_s^{(i-1)} + \Delta w_s^{(i)}) = \\ &= \Delta w_s^{(i)} + \frac{\alpha}{N_s(s-1)} c_{s-1} (t_{s-1} - c_{s-1}^T w_s^{(i+1)}) - \\ &= \frac{\alpha}{N_s(s-1)} c_{s-1} (t_{s-1} - c_{s-1}^T w_s^{(i)}) = \Delta w_s^{(i)} - \\ &= \frac{\alpha}{N_s(s-1)} c_{s-1} c_{s-1}^T (w_s^{(i+1)} - w_s^{(i)}) = (I - \frac{\alpha}{N_s(s-1)} \\ &= c_{s-1} c_{s-1}^T) \Delta w_s^{(i)} \quad (8) \end{aligned}$$

定义  $\Delta w_s^{(i)} \equiv \Delta w_s^{(i-1)}$  和  $c_0 \equiv c_{N_s}$ ,再令

$$E_s \equiv I - \frac{\alpha}{N_s(s)} c_s c_s^T \quad (9)$$

由式(8)、(9)有

$$Dw_i^{(0)} = (E_{i-1}E_{i-2}\cdots E_1E_{N_i}\cdots E_i)Dw_i^{(i-1)} = (E_{i-1}E_{i-2}\cdots E_1E_{N_i}\cdots E_i)'Dw_i^{(0)} \quad (10)$$

$$\text{令 } G_i \equiv E_{i-1}E_{i-2}\cdots E_1E_{N_i}\cdots E_i \quad (11)$$

$$\text{这样,式(10)就变为 } Dw_i^{(0)} = G_i'Dw_i^{(0)} \quad (12)$$

下面的推导同文[13]的一致,这里简要列出结论,证明详见文[13]。根据  $N_a(s)$  是否固定常数分两种情况:

情形1:当对于任意样本  $s, N_a(s) \equiv N_a$  为常数时,GMNN 有如同文[13]的结论。

**定理1** 矩阵  $E_i$  是对称的,其元素  $e_{ii}$  为

$$e_{ii} = \begin{cases} 1 & \text{如果 } c_i \text{ 的第 } i \text{ 个元素是 } 0 \\ 1 - \frac{\alpha}{N_a} & \text{如果 } c_i \text{ 的第 } i \text{ 个元素是 } 1 \end{cases}$$

$$e_{ij} = \begin{cases} 0 & \text{如果 } c_i \text{ 的第 } i \text{ 或第 } j \text{ 个元素是 } 0 \\ -\frac{\alpha}{N_a} & \text{如果 } c_i \text{ 的第 } i \text{ 和第 } j \text{ 个元素都是 } 1 \end{cases}$$

**定理2** 如果学习速率  $0 < \alpha < 2$ ,对任意矩阵  $B, \|B \cdot E_i$  的第  $i$  行  $\| \leq \|B$  的第  $i$  行  $\|$

证明:设  $Q = BE_i, E_i$  第  $p$  行所有的非零元素记为  $e_{pp}, e_{pq}, \dots, e_{pr}$ 。定义  $S = \{p, q, \dots, r\}$  来表示这些元素的位置。

$$\|Q \text{ 的第 } i \text{ 行}\|^2 = \sum_{j=1}^N q_{ij}^2 = (b_{ip} \cdot e_{pp} + b_{iq} \cdot e_{qp} + \dots + b_{ir} \cdot e_{rp})^2 + (b_{ip} \cdot e_{pq} + b_{iq} \cdot e_{qq} + \dots + b_{ir} \cdot e_{rq})^2 + \dots + (b_{ip} \cdot e_{pr} + b_{iq} \cdot e_{qr} + \dots + b_{ir} \cdot e_{rr})^2 + \sum_{j \in S} b_{ij}^2$$

根据定理1,再对上式通过展开、重新合并,最后得

$$\|Q \text{ 的第 } i \text{ 行}\|^2 = \sum_{j=1}^N b_{ij}^2 + \frac{\alpha(\alpha-2)}{N_a} \cdot (\sum_{i \in S} b_{ii})^2 \quad (13)$$

如果  $0 < \alpha < 2$ ,由(13)定理2得证。同时,有  $\sum_{j=1}^N q_{ij}^2 \leq \sum_{j=1}^N b_{ij}^2$ , 等号只有当  $\sum_{i \in S} b_{ii} = 0$  时才成立。

**定理3**  $\lim_{k \rightarrow \infty} \|G_k^i\|$  的任意行  $\| = \lim_{k \rightarrow \infty} \|G_k^i \cdot G_i$  的同样行  $\|$ ,且是有界的。

**定理4** 若  $0 < \alpha < 2$ ,则当  $k$  趋于无穷,  $G_k^i \cdot E_j = G_j^i, j = 1, 2, \dots, N_a; G_k^i$  收敛到零或常数矩阵。

**定理5**  $\lim_{k \rightarrow \infty} G_k^i \cdot c_j = 0, j = 1, \dots, N_a$ 。

由上述定理,可得

**定理6** 如果  $0 < \alpha < 2$ ,GMNN 的迭代学习收敛到一个极限环。

**定理7** 当迭代次数趋于无穷,学习速率  $\alpha$  趋于0, GMNN 学习到最小平方误差,即(6)式。

情形2:当至少有两个样本  $s, s', N_a(s) \neq N_a(s')$  时,定理6 仍然成立,用文[13]的方法推不出定理7的结论。但并不意味着最后收敛的误差界就一定更差,根据式(5)、(6),只要量化能获得恰当的  $C$ ,就可以任意精度学习。那么如何找到这样的  $C$  呢?或者说它与量化方式有怎样的关系?下面的分析针对情形1,对情形2可做同样的分析。

### 2.3 精确学习的进一步分析

虽然文[12,13]获得了 CMAC 的学习收敛性的结果,并且上节我们把这个结果延伸到包含 CMAC 在内的一大类网络 GMNN 中,但正如我们上面看到的那样,我们实际所作的只是在假定输入空间量化已经完成,即从原始样本数据空间到变换空间(地址空间)这种蕴涵了非线性的映射也就被确定下

来后,在变换空间中,寻求  $w$  使得平方误差最小,这是一个线性回归问题(见2.2.1节),而 GMNN(包括 CMAC)的结构保证了误差修正迭代学习可以逼近这个解。可见,量化确定后,最后的误差的界也就确定。那么,对于一个给定的学习问题,学习精度完全由量化方式、量化程度来决定。换句话说,如果把由量化决定的地址指示向量(即矩阵  $C$ )看成是隐层输出,即网络的内部表示的话,为了达到学习目标,要寻求一种最优的内部表示,使学习速度最快、网络规模(对于 GMNN 来说,所需存储空间)最小。给定一个学习精度,输入空间到底该离散化到什么程度,实际应用中,一般是凭经验确定。这里给出一般 GMNN 的精确学习同输入量化、而由此决定的中间内部表示之间的关系的一般描述。

由定理7,GMNN 收敛后,  $w = (C^T C)^{-1} C^T T, (C^T C)^{-1} C^T$  是广义逆  $C^+$ ,代入(4)

$$\text{得 } E = \frac{1}{2} T^T (I - CC^+) T \quad (14)$$

$$\text{最小化 } E \text{ 就相当于 } \arg \min_{\text{量化}} E = \arg \max_{\text{量化}} T^T CC^+ T \quad (15)$$

$$\text{根据奇异值分解理论 } C = UH V^T \quad C^+ = V H^+ U^T \quad (16)$$

$$H H^+ = \begin{bmatrix} I_{r \times r} & 0 \\ 0 & 0 \end{bmatrix} \quad (17)$$

$U, V$  是正交矩阵,  $H$  是对角矩阵。

由(16)、(17),则(15)可进一步表示为

$$\arg \min_{\text{量化}} E = \arg \max_{\text{量化}} T^T U \begin{bmatrix} I_{r \times r} & 0 \\ 0 & 0 \end{bmatrix} U^T T = \arg \max_{\text{量化}} \sum_{j=1}^r (T^T u_j)^2 \quad (18)$$

从式(18)知,  $T^T u_j$  代表目标向量到隐层输出空间的正交基的投影,式(18)代表了此投影向量的范数,即如果量化能获得此范数,则就实现了精确学习。

## 3. 两个推广模型

### 3.1 基于稀疏 RAM 的逼近型 n-tuple 网络和逼近型 SDM 网络模型

图2是一个单输出的 SLLUP 模型<sup>[11]</sup>。它的工作原理:将输入  $x$  进行二值化成0,1串或0,1网(retina),然后,对其做  $n$ -tuple 采样,形成若干个  $n$  位长地址(假设  $m$  个,每个是  $n$  位长,则每个 RAM 地址个数为  $2^n$  个,共需要  $m \times 2^n$  个存储单元),相应那些地址处的存储内容(即权重)求和得输出响应。图中为直观起见,在地址单元阵列后,加了一个地址指示器向量,'1'表示相应的 RAM 地址被选中,否则记'0',它们到输出单元的连接弧的强度即权重  $w$ 。从 SLLUP 的结构可知,它的量化过程是由  $n$ -tuple 采样来完成。

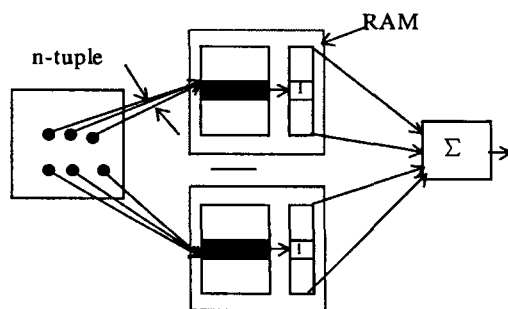


图2 SLLUP 拓扑结构

SLLUP 是将  $n$ -tuple 分类器与感知器结合的新模型,在文[11]中,把 MLPs、RBFN、SLLUP 置于低通内插滤波器的统一认识下,并在输入编码为 Barchart 时,给出了其核函数,

证实了 SLLUP 作为低通内插滤波器解释的合理性。现在,显然可以把 SLLUP 视为 GMNN 的一个成员,因此而获得其学习能力的进一步认识。

SLLUP 通常也被称为逼近型 n-tuple 网络(它是基于 RAM 的)。现在,我们把 RAM 换成稀疏 RAM,得到的模型称之为基于稀疏 RAM 的逼近型 n-tuple 网络(以下简称 SN-TUPLE)。显然,SLLUP 可视为 SN-TUPLE 的特例。

所谓的稀疏 RAM 是来自文[14]的 SDM 的概念。常规的 RAM 是存储位置的阵列,其主要特征是在读写数据时,对每一个输入地址,仅有一个存储位置被选择。SDM 模型中的地址存储阵列(稀疏 RAM)与常规的 RAM 的区别主要在于它的稀疏性和分布性,依据某种近邻准则,激活存储结点中的若干地址单元;对于 n 位长的地址字,RAM 的所有可能的地址个数为  $M=2^n$ ,而对于 SDM,可取  $M \ll 2^n$ 。反映在图2中,就是每个 n-tuple 相应的地址指示器的‘1’的个数可多于一个。对于任意一个输入  $x_i$ ,先转化为 0、1 二值向量,再经 n-tuple 采样形成 m 个长为 n 位的地址,从而,地址指示向量  $c^T = (c_{i,1}, \dots, c_{i,M}, \dots, c_{i,(l-1)M+1}, \dots, c_{i,(l-1)M+M}, \dots)$ 。对于 SLLUP,  $M = 2^n$ ,  $N = m \times M$ ,每个 n-tuple 相应的指示器只有一个‘1’,其余为‘0’,那么,  $c_i$  的‘1’的个数(激活的地址个数)为  $N_i(s) = m$ ;至于 SN-TUPLE 模型的  $c_i$  中的‘1’的个数则依据某种近邻规则来定。

规则1:以 n-tuple 采样形成的每个 n 位长地址向量为中心,半径为 d 的海明球内的地址集中的地址被激活,相应的地址指示位置‘1’。

规则2:n-tuple 采样形成的每个 n 位长地址向量的 D 个最近邻被激活,相应的地址指示位置‘1’。

采用上述两种规则的 SLLUP 的推广模型分别简记为 SN-TUPLE1 和 SN-TUPLE2。前者的  $c_i$  中‘1’的个数难于确定,那么(9)式中的  $N_i(s)$  不是固定值,因此只能有定理6的结果;后者则有  $N_i(s) = m \times D$ (固定常数),定理6、7均成立。当把整个样本作为一个 n-tuple 时( $m=1$ ),上述两种情形可看成逼近型 SDM 模型(简记为 ASDM1, ASDM2)——相当于修改原 SDM 的外积法学习规则而得,这个改进的意义在于

突破了原来 SDM 只用于联想记忆的应用范围。

第2节及以上的分析表明推广模型 SN-TUPLE 较之 SLLUP 的优越之处:比 SLLUP 多了可调参数 M、d 或 D,可以避免有些测试样本的输出因为空置而输出 0;通过选择适当的可调参数,可以构造出规模较小的网络以达到学习目标,表现出更大的灵活性;SN-TUPLE2 可以直接处理实数或多级输入而不必二值化,即直接在原始样本数据上执行 n-tuple 采样,此时地址矩阵也可以是实数编码。因此,很适合如灰度人脸图像那种大维数样本的直接自动识别而无需降维等预处理,具体实现另文再述。

### 3.2 实验

考虑如下的非线性映射:

$$f(x_1, x_2) = \sin(\pi x_1) \cos(\pi x_2) \quad (19)$$

$x_1 \in [-1, 1], x_2 \in [-1, 1]$ 。由于输入只是二维,因此对于 SN-TUPLE 不宜在原始输入上直接执行 n-tuple 采样,下面的实验针对 SN-TUPLE1;对于逼近型 SDM,用 ASDM2 更合适。

我们按  $\Delta x_1 = \Delta x_2 = 0.1$  等间隔取点,由(19)式得到 400 组输入输出样本对来训练 SN-TUPLE1 和 ASDM2。性能指标取均方根误差 RMS。

对于 SN-TUPLE1,把  $x_i$  归一化到 [0, 1] 上、离散化 L 个等级,Barchart 编码为二进制串,各个  $x_i$  级联,合并成 0、1 串或网(retina),再进行 n-tuple 划分,划分为 m 个这样的 n 元子模式;然后到各自对应的稀疏 RAM 的地址矩阵按规则1激活地址单元,每个地址矩阵均是均匀随机预置的 0、1 阵,有 M 个长为 n 的行(或列)。

对于 ASDM2,我们不进行输入编码转换,将实数向量输入归一化到  $[-1, 1]^2$  上;均匀随机预置地址矩阵也不是 0、1 阵,是  $[-1, 1]$  区间上取值的实数矩阵,有 M 个与输入向量等长的行(或列)。按规则2激活地址单元,距离度量使用城市距离或欧氏距离。

为了显示网络的泛化能力,对未训练过的数据( $\Delta x_1 = \Delta x_2 = 0.12$ 取点),如图3(b)、(c)分别是 SN-TUPLE1 和 ASDM2 输出的估计结果,图3(a)是由(19)算出的期望结果;实验结果汇总在表1。

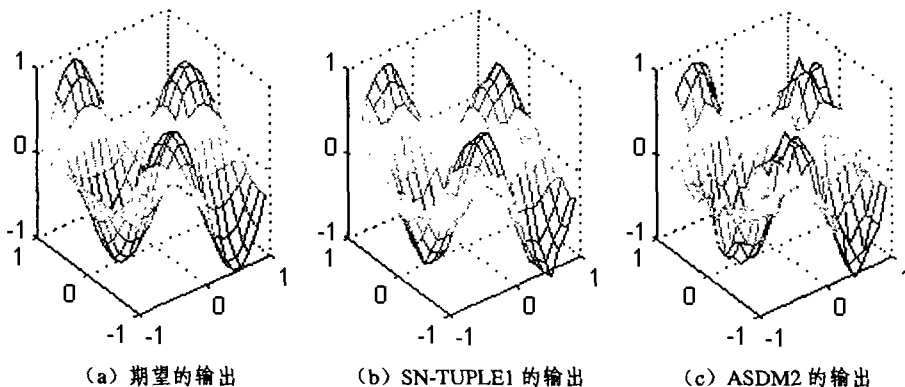


图3 期望输出与网络的实际输出对照

从实验结果可以看出,正如我们分析的,最后的收敛误差跟量化方式、程度有着密切的关系。无论选择什么样的参数,通过地址选择操作,都要形成 0、1 矩阵 C,即隐层的中间内部表示,它的规模大小、性质直接反映了所作的量化,同时,也决定了该网络最后的学习能力。SLLUP 的地址空间大小为  $m \times 2^n$ ,因此 C 是非常稀疏的大矩阵,而 SN-TUPLE1 的地址空间

大小为  $m \times M, M \ll 2^n$ ,这时的 C 规模比前者大大减小了,不仅如此,从表1知,通过选择适当的参数 d,学习精度,尤其是泛化能力均比 SLLUP 更优。ASDM2 可以看作 SN-TUPLE2 的取  $m=1$  时的情形。

总之,如何量化可以获得最小规模的 C,而学习能力最

(下转第132页)

抗 MAP 图像去噪方法的性能。

表1 抗低通滤波试验结果

水印算法	未处理前输出相关值	谱能量衰减比值	局部平均后输出相关值	谱能量衰减比值	中值滤波后输出相关值
Cox	1.0	0.016	0.9590	0.0048	0.9963
Kim	0.9995	0.242	0.2147	0.337	0.2498
Wang	0.9999	0.0465	0.4424	0.0519	0.5748
Xia	1.0	0.3247	0.1339	0.5556	0.1668

表2 维纳滤波试验结果

水印算法	均方误差( $\times 10^5$ )	Wiener 滤波后输出相关值
Cox	0.4747	0.9989
Kim	1.5329	0.5062
Wang	1.5293	0.8458
Xia	1.451	0.8965

**结论** 本文从水印攻击的角度,分析了去噪攻击的原理,得出了如下结论:如果差值图像的谱能量在低通滤波器低频带通区域越集中,差值图像谱能量衰减比值越小,那么差值图像被低通滤波滤除的可能性越小,算法抗 ML 估计去噪攻击的性能也越好;如果差值图像和原始图像的局部方差分布越相似,那么算法抗 MAP 估计去噪攻击的鲁棒性越好。相反则水印算法抵御去噪攻击的鲁棒性越差。通过实验我们证明了

结论的正确性,实际工作中,可以利用上述结论有效地增强水印算法抗去噪攻击的能力,从而设计出性能更好的水印算法。

参考文献

- Xia X-G, Boncelet C G. Wavelet transform based watermark for digital images. *Optics Express*, 1998, 3(12): 497
- Wang H-J, Su P-C, Kuo C-C J. Wavelet-based blind watermark retrieval technique. In: Proc. of the SPIE Phonics East Symposium on Voice, Video and Data Communications, Boston, MA, USA, 1998, 3528: 440~451
- Kim J R, Moon Y S. A robust wavelet-based digital watermark using level-adaptive thresholding. In: Proc. of the 6<sup>th</sup> IEEE Intl. Conf. on Image Processing, ICIP'99, page 202, Kobe, Japan, Oct. 1999
- Kim Y-S, Kwon O-H, Park R-H. Wavelet based watermarking method for digital images using the human visual system. *Electronic Letters*, 1999, 35(6): 466~467
- Cox I, Kilian J. Secure spread spectrum watermarking for multimedia. *IEEE Trans. Image Processing*, 1997, 6: 1673~1687
- Voloshynovskiy S, Pereira S. Attack modeling: towards a second generation watermarking benchmark. *Signal Processing of the ACM*, 2001, 81: 1177~1214
- Voloshynovskiy S, Deguillaume F. Content Adaptive watermarking based on a stochastic multiresolution image modeling. *EUSIPCO 2000, Tampere, Finland, Sep. 2000*
- Voloshynovskiy S, Herrigel A. A stochastic approach to content adaptive digital image watermarking. In: *Third Intl. Workshop on Information Hiding, Dresden, Germany, Sep. 29~Oct. 1, 1999*
- 刘振华, 尹萍. 信息隐藏技术及其应用. 科学出版社

(上接第121页)

强,这就是我们在使用此类网络(GMNN)最关心的,前面的分析以及实验结果说明确实存在某种最优的量化。

表1 实验结果

网络结构		RMS (训练集)	RMS (测试集)
SN-	L=512 m=10 n=31 M=500 d=8	0.0186	0.0848
	L=512 m=50 n=15 M=128 d=2	0.0287	0.1052
TUPLE1	L=512 m=50 n=15 M=128 d=4	0.0158	0.0585
	L=128 m=10 n=15 M=128 d=6	0.0835	0.1338
SLLUP	L=128 m=10 n=15	0.0783	0.1169
	L=512 m=10 n=31	0.0113	0.1124
ASDM2	M=500 D=5	0.0393	0.0874
	M=500 D=8	0.0463	0.0834
	M=900 D=8	0.0426	0.0817

**结论** 我们避开 CMAC、SLLUP 的具体的量化方式,它们按各自的方式实现从原始数据到地址空间的映射,其后的学习就是一个线性优化问题了,而非线性性质就隐含在上述这种映射之中。这一特点可以延伸到一大类称为一般存储器神经网络(GMNN)框架中,并且证明了它通过误差迭代学习能收敛到平方误差最小。由于 GMNN 的性能主要取决于如何量化以及量化程度,而因此决定了中间内部表示(即地址指示向量)。分析表明,确实存在(可以找到)一种量化方式,对于任意一个给定的学习问题,可以最小的网络规模达到目标,并给出了一个一般性的描述。当然,尚待研究的问题很多,对那类基于存储器的网络究竟作何评价、与其它普通的网络的关系,以及量化方式蕴含的非线性性质如何通过抽象出来的中间内部表示来揭示(已经有人作了部分工作),还有过拟合(过学习)的问题等等,这些问题都会影响到这类网络的广泛应用。

参考文献

- Funahashi K. On the approximate realization of continuous mappings by neural networks. *Neural Networks*, 1989, 2: 183~192
- White H. Learning in artificial neural networks. *Neural Comput.* 1989, 1(4): 425~464
- Albus J S. A theory of cerebellar function. *Mathematical Biosciences*, 1971, 10: 25~61
- Bledsoe W, Browning I. Pattern recognition and reading by machine. *IRE Joint Computer Conference*, 1959. 225~232
- Aleksander I, Thomas W V, Bowden P A. WISARD-a radical step forward in image recognition. *Sensor Review*, July, 1984. 120~124
- Kolcz A, Allinson N M. N-tuple Regression Network. *Neural Networks*, 1996, 9(5): 855~869
- Kolcz A, Allinson N M. General memory neural network-extending the properties of basis networks to RAM-based architectures. In: Proc. 1995 IEEE Int. Conf. on Neural Networks (ICNN'95), 1995. 1638~1643
- Kolcz A, Allinson N M. Basis function models of the CMAC network. *Neural Networks*, 1999, 12: 107~126
- Kolcz A, Allinson N M. The general memory neural network and its relationship with basis function architectures. *Neurocomputing*, 1999, 29: 57~84
- Rohwer R, Morciniec M. The Theoretical and Experimental Status of the n-tuple Classifier. *Neural Networks*, 1998, 11(1): 1~14
- Tattersall G D, Foster S, Johnston R D. Single-layer lookup perceptrons. *IEE PROCEEDINGS-F*, 1991, 138(1): 46~54
- Wong Y-F, Sideris A. Learning convergence in the Cerebellar model articulation controller. *IEEE Trans. On Neural Networks*, 1992, 3(1): 115~121
- Lin C-S, Chiang C-T. Learning convergence of CMAC technique. *IEEE Trans. On Neural Networks*, 1997, 8(6): 1281~1292
- Kanerva P. *Sparse Distributed Memory*. MIT Press, Cambridge, Massachusetts, 1988