



暂态混沌神经网络在人眼运动 机制模拟中的应用¹⁾

谭 营 邓 超

(中国科学技术大学电子工程与信息科学系 合肥 230027)

关键词 混沌神经网络, 计算机视觉, 人眼运动机制模拟.

APPLICATION OF TRANSIENTLY CHAOTIC NEURAL NETWORK TO SIMULATION OF HUMAN EYE MOVEMENT MECHANISM

TAN Ying DENG Chao

(Dept. of Electronic Eng. and Info. Sci., University of Science and Technology of China,
Hefei 230027)

Key words Chaos neural network, computer vision, human eye movement mechanism simulation

1 引言

人眼运动机制主要包括定位和识别两项子任务^[1]. 定位任务是利用感兴趣模式的先验知识, 在一个宽视野中定位模式的位置, 而识别任务则是在假定该模式的位置已固定的条件下确定该模式的类别. 因此, 这两项子任务是协作性任务.

为了克服 Hopfield 型动态神经网络(HNN)^[2]在信息处理和优化计算中所存在的容易陷入局部极值问题, 许多学者已提出了多种解决方法, 如随机性的和确定性的模拟退火方法. 最近通过对神经网络的复杂非线性动力学特性的深入研究表明利用这种复杂非线性动力学特性也可以达到某种退火之目的^[3], 它具有很强的退火特性和很快的搜索速度, 可以有效地求解各种复杂优化问题^[4].

通过将人眼运动机制模拟问题变换为一复杂的优化求解问题, 本文采用一种暂态混沌神经网络来求解这一复杂的优化求解问题从而达到对人眼运动机制的有效模拟.

1) 国家自然科学基金和中国博士后科学基金资助项目.

收稿日期 1997-01-20 收到修改稿日期 1997-12-01

2 暂态混沌神经网络

一种具有暂态混沌和时变增益的神经网络的动力学方程可表示为

$$Y(t+1) = kY(t) + \alpha(WX(t) + I) - Z(t)(X(t) - I_0), \quad (1)$$

$$Z(t+1) = (1 - \beta)Z(t), \quad \epsilon(t+1) = (1 - \gamma)\epsilon(t). \quad (2)$$

上式中 X, Y 分别为神经网络的输出和内部状态向量且 $X = \text{sigmoid}(1 + \exp(-Y(1 + \epsilon)))$; W 代表网络的连接权矩阵; I 为输入偏置且 $WX + I = -\partial E / \partial X$, E 是能量函数; I_0 为一正常数; α 为比例参数; $k \in [0, 1]$ 为神经薄膜的衰减因子; Z, ϵ 分别为自反馈连接权和输出函数的时变增益参量且 $0 \leq \beta, \gamma \leq 1$.

由(1)和(2)式描述的网络具有以下特点^[3]:

1) 网络具有暂态混沌特性. 图1是当各参数分别取为 $k = 0.95, I_0 = 0.84, z(0) = 0.77, \epsilon(0) = 280, \beta = 0.0001, \gamma = 0.008$ 时, 单个神经元模型的动态特性. 由其 Lyapunov 指数的正负以及其状态游动路线可以看出, 状态演化初始阶段存在一种倒分叉混沌特性, 随着 $z(t)$ 和 $\epsilon(t)$ 的减小, 这种分叉过程会很快结束, 网络表现出了暂态的混沌特性.

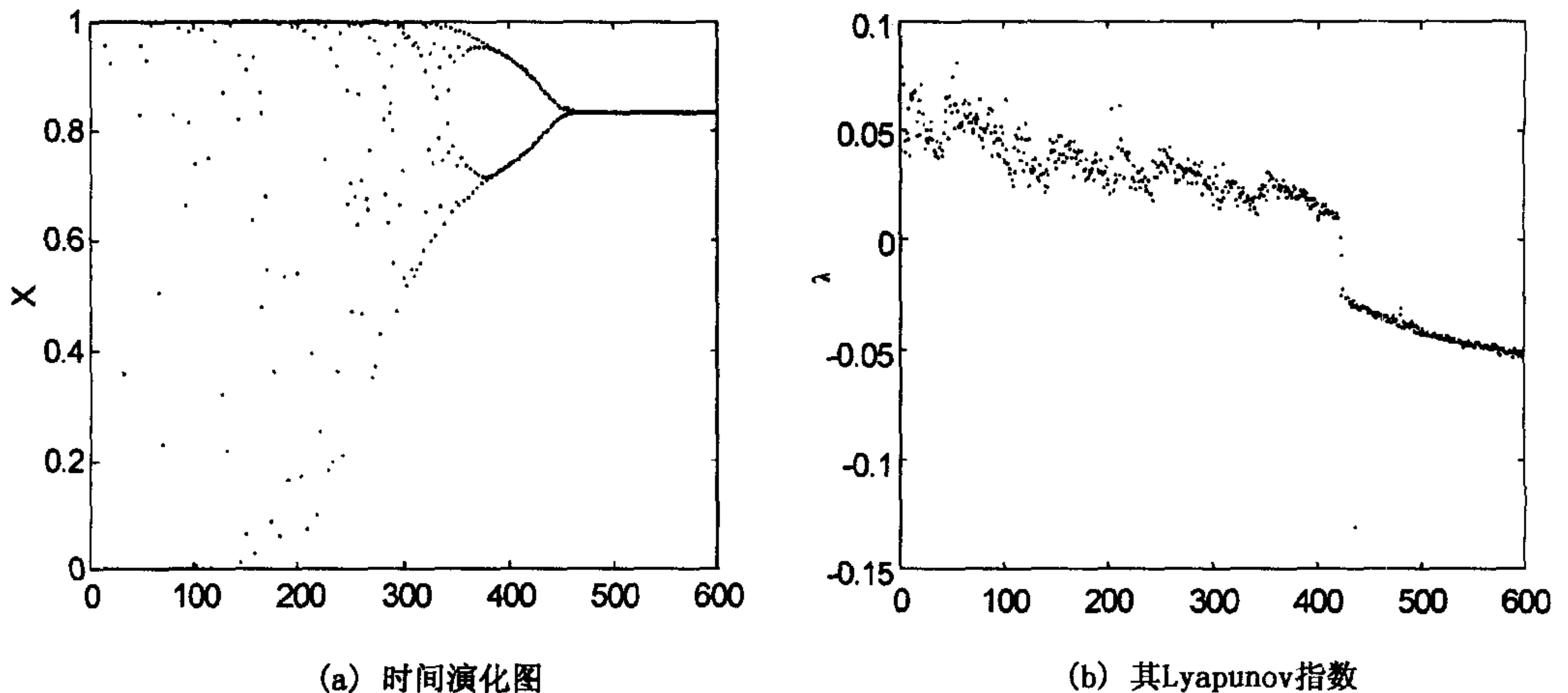


图1 单个神经元模型

2) 网络能演化到一稳定状态. 随着倒分叉暂态过程的结束, 网络的状态演化轨道将从某奇异吸引子到达一个稳定平衡点. 此时的网络逐渐趋近于 HNN, 故网络存在稳定状态.

3) 其搜索区域为一分形结构. 网络在动态演化过程中所搜索的区域相对于整个状态空间是很小的, 具有某种分形结构. 其大小相对于整个状态空间的 Lebesgue 测度等于零.

4) 具有混沌退火机制. 由于网络的倒分叉混沌过程具有丰富的动力学特性(不仅存在非周期和振荡特性而且包含有大量的奇异吸引子), 它可用于避开网络能量函数的诸多局部极值点, 有效地搜索全局最优解. 网络状态的这种搜索全局最优解的特性与随机模拟退火(SSA)具有许多的相似性, 故称其为混沌退火(chaotic annealing)特性. 混沌退火(属于确定性模拟退火(DSA)范畴)不仅在搜索效率上而且在计算速度上都比 SSA 有明显的

优势.

总之,只要适当地选取网络参数值,它就可以用于求解十分复杂的非线性优化问题.其典型的求解过程是先经过一段倒分叉暂态混沌过程,然后收敛到一个稳定的平衡点(它对应着目标函数的最小点或其一个很小的邻域).下面将其用于对人眼运动机制的模拟.

3 人眼运动机制的模拟

用于模拟人眼运动机制的神经网络结构如图2所示.它是由三个子块组成的三层网络.分别称其为模拟层(S块),隐层(H块)和输出层(O块).其中,S块和O块中的神经元是层内全连接的,S块和H块之间是层间全连接的,H块和O块之间也是层间全连接的.

网络的能量函数定义如下:

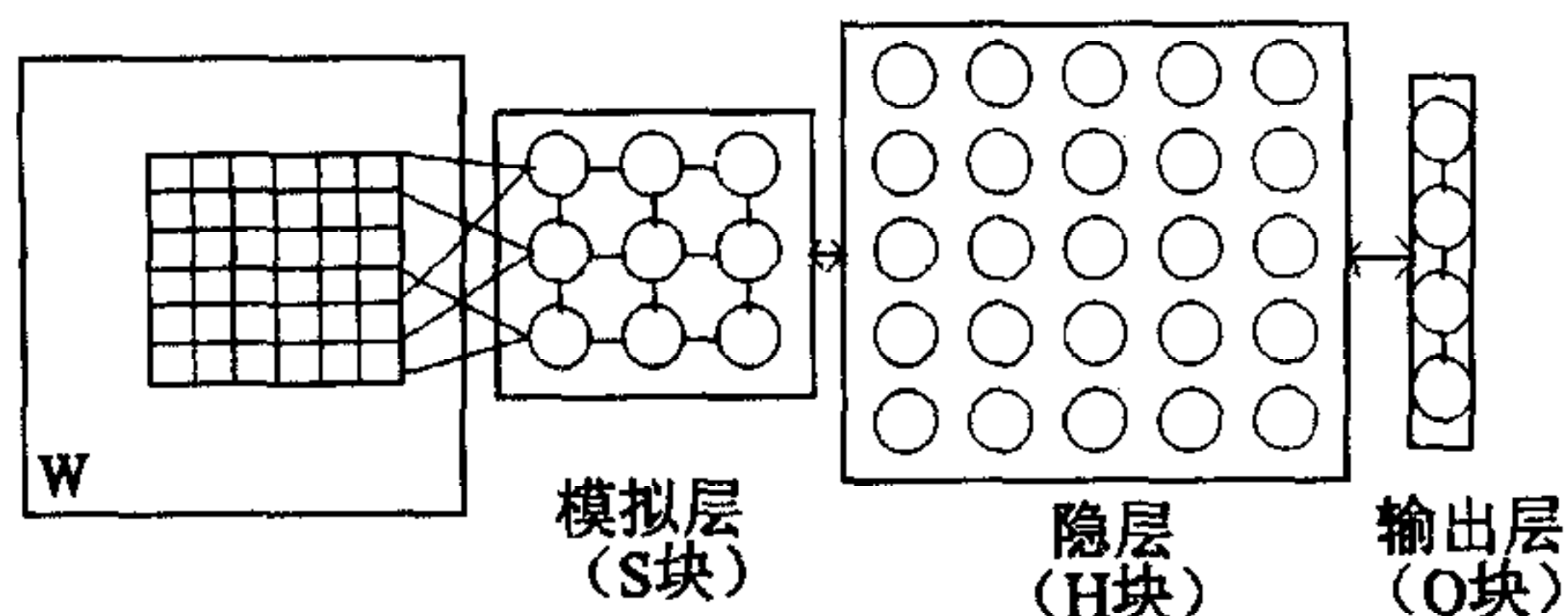


图2 模拟人眼运动机制的三层耦合神经网络

$$E = C_0 \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \left(\sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} Z_{i+x,j+y}^{l,m} V_{x,y}^S - V_{i,j}^H \right)^2 + C_1 \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} \left(\sum_{n=0}^{N-1} T_{i,j}^n V_n^O - V_{i,j}^H \right)^2 + C_2 \left(\sum_{x=0}^{X-1} \sum_{y=0}^{Y-1} V_{x,y}^S - V_{i,j}^H \right)^2 + C_3 \left(\sum_{n=0}^{N-1} V_n^O - 1 \right)^2; \tag{3}$$

$$Z_{i,j}^{l,m} = A_{l+i-(I+X)/2,m+j-(J+Y)/2} \cdot W_{l+i-(I+X)/2,m+j-(J+Y)/2}; \tag{4}$$

$$A_{l+i-(I+X)/2,m+j-(J+Y)/2} = \begin{cases} 1, 0 \leq i < I, 0 \leq j < J, \\ 0, \text{其它}. \end{cases} \tag{5}$$

上式中 (l,m) 为注视位置; W 为一个输入模式的象素图象; A 为一窗函数.为表述简捷,用 k 代表S,H和O中的某一块. $u_{i,j}^k$ 为 k 块的第 (i,j) 个神经元的激励. $V_{i,j}^k$ 为 k 块的第 (i,j) 个神经元的输出,若采用Sigmoid激活函数,则有 $V_{i,j}^k = \text{sigmoid}(u_{i,j}^k)$. T^n 为第 n 个模板.它由H块与O块之间的加权值表示. $C_i(i=0,1,2,3)$ 为正常数,用于均衡各项优化目标和约束.相对于 (L,m) 在第 $\left(\frac{x}{2}, \frac{y}{2}\right)$ 个神经元,S块实际上是一个输入模式位置的矩阵表示.这样,O块中最活跃的神经元代表了输入模式的类别,同时该输入模式的位置由S块表示.

定位子任务和识别子任务可以分别通过最小化能量函数的第一项和第二项完成.能量函数的第三项和第四项分别代表在S块和O块的胜者优先(WTA)约束.

注视位置的修正规则可表述为

$$l(n+1) = l(n) + x_m - X/2, \quad m(n+1) = m(n) + y_m - Y/2, \tag{6}$$

其中 n 表示迭代次数, (x_m, y_m) 代表 S 块中最活跃神经元的位置.

人眼运动的定位和识别协作任务的完成就是通过(6)式进行注视位置的不断迭代修正,直到注视位置不再变化为止.此时 O 块中输出最大(最活跃)的神经元就代表了输入模式的类别,同时该输入模式的位置由 S 块中的激活神经元来表示.

4 仿真结果

例1. 该模拟实验说明这里的网络能够完成人眼运动机制中的协作任务. 实验中我们选取(3)式能量函数中的各正常数和每块(层)的神经元数量分别为: $C_0=0.5, C_1=0.4, C_2=2.0, C_3=1.0, X=Y=9$ (S 块有91个神经元), $I=J=8$ (H 块有64个神经元), $N=6$ (O 块有6个神经元). 实验中所采用的六个模板 $T^i (i=1\sim6)$ 如图3所示(都为 8×8 二值象素图象).

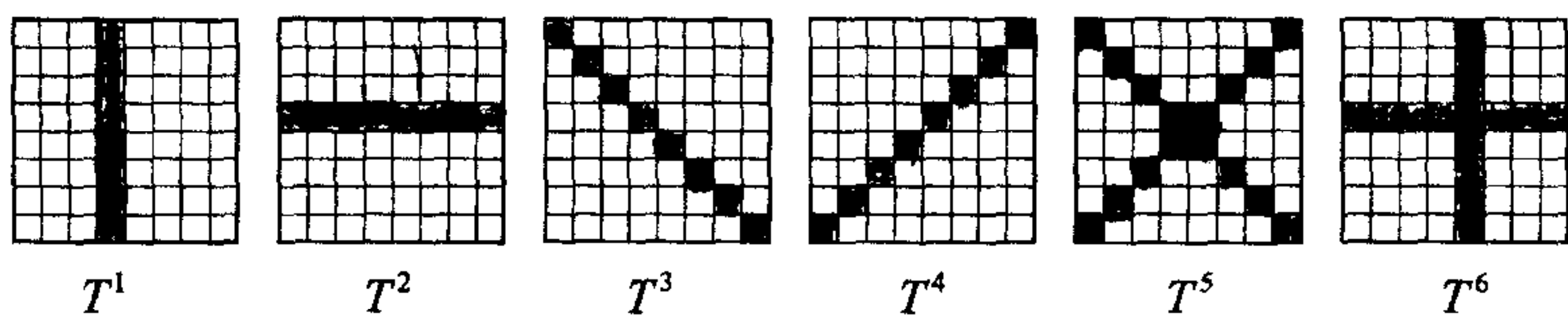


图3 例中所采用的六个模板

网络耦合参数分别选为: $k=0.9, I_0=0.5, z(0)=0.977, \epsilon(0)=340, \beta=0.0053, \gamma=0.0032$. 图2的三层网络从输入图象开始演化,经过一个倒分叉暂态混沌过程,网络最后达到的稳定状态如图4(a)–(f)所示. 图中每个方块的一边代表对应神经元的输出电平. S 块中的有向线段表示一个运动向量. 可以看出目标模式的位置和类别分别由 S 块和 O 块中的激活神经元表示. 从图中也可以看出 H 块中的激活模式与目标模式类模板十分相似. 大量实验表明,采用本文的暂态混沌动态神经网络模拟方案在求解精度和速度诸方面都比现有的 HNN 方法更为有效.

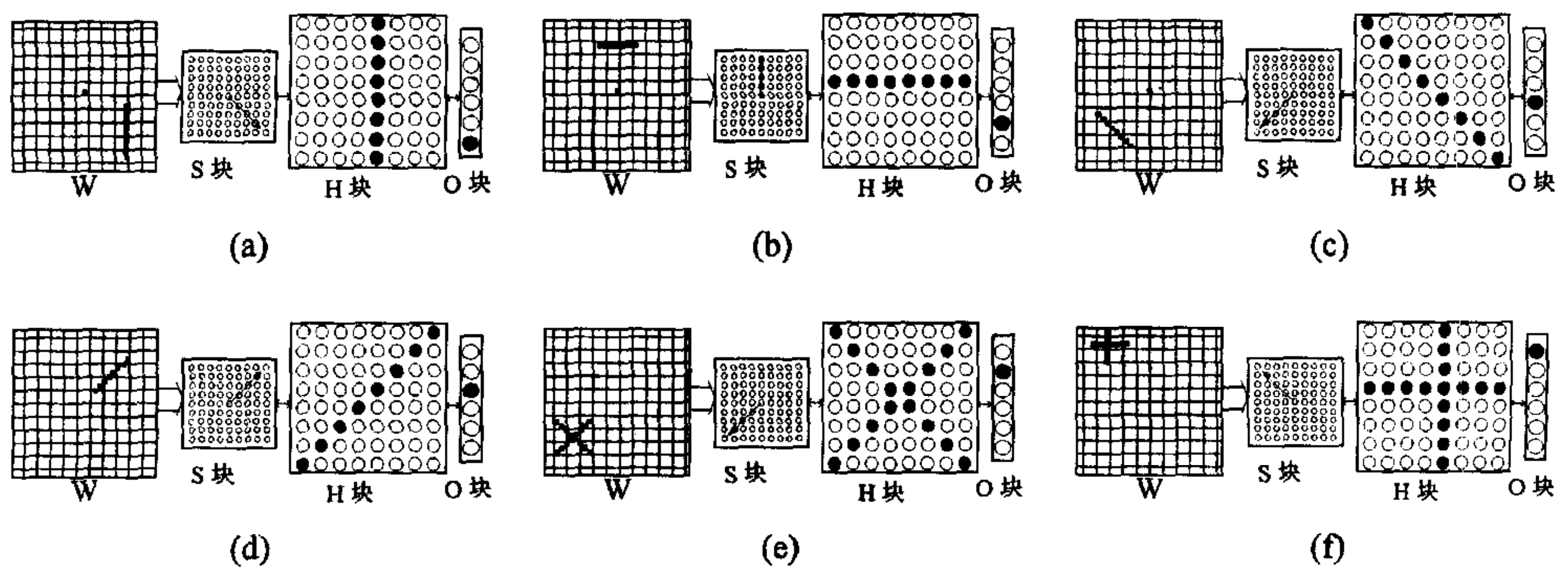


图4 完成例中输入模式的定位和识别协作任务后网络的稳定状态

参 考 文 献

1 Ballard D H. Behavioral constraints on animate vision. *Image and Vision Computing*, 1989, 7(1): 340—356

- 2 Hopfield J J. Neurons with graded response have collective computational properties like those of two-state neurons. In: Proc. Nat. Acad. Sci. U. S. , 1984, 81: 3088—3092
- 3 谭营, 邓超, 王宝云等. 一种具有暂态混沌和时变增益的神经网络及其在优化计算中的应用. 电子学报, 1998, 25(7): 123—127
- 4 Tan Ying, Deng Chao, He Z Y. A chaotic annealing neural network and its application to direction estimation of spatial signal sources. Neural Networks for Signal Processing VII, Florida, USA; IEEE Press, 1997. 233—244
- 5 Tan Ying, Liu Z K. Signal detection in digital communications in terms of a chaotic neural network. In: Proc. 9th IEEE Signal Processing Workshop on Statistical and Array Processing, USA, 1998. 260—263

谭 营 男, 1964年9月生, 博士后, 副教授. 研究领域: 神经网络, 智能计算科学, 信号与图象的处理和识别电子工程等.

邓 超 女, 博士生. 研究领域: 神经网络学习算法, 智能知识系统, KDD 等.

论文清稿规格及要求

1. **直接制版** 论文须为可供直接照相制版的清稿(此稿不退)。为此, 清稿必须用激光打印机在 A4 白纸上单面打印, 清稿的打印区尺寸统一规定: 长为210毫米, 宽为145毫米。打印区外不能印任何字迹或图纸。
2. **题目** 题目一律居中用4号黑体字, 其余均用5号字。
3. **作者姓名和单位** 文题下空两行印作者姓名, 其下(不空行)印作者单位名。
4. **摘要** 作者单位名下空两行印摘要。格式: 先印‘摘要’二字, 后空二字开始印摘要的内容。建议勿超过10行。
5. **正文** 由于论文数量较多, 建议每篇论文篇幅勿超过8页。
6. **节和节标题** 正文可以分为若干节, 每节有一个节标题。节可编号可不编号。节标题从打印区边线缩进15毫米(不要居中)。
7. **参考文献** 参考文献列于正文之后, 内容依次为: 编号、作者、题目、出版单位(或期刊, 或会议), 年份。
8. **图和图题** 若图不能直接在文中排版, 可以贴在正文中, 也可附在全稿最后(参考文献和附录之后)。图号打印在图下方, 与文呼应。图题可有可无。
9. **页码** 在清稿每页打印区以外的右上角用铅笔写上页码。