第 6 章　自组织特征映射

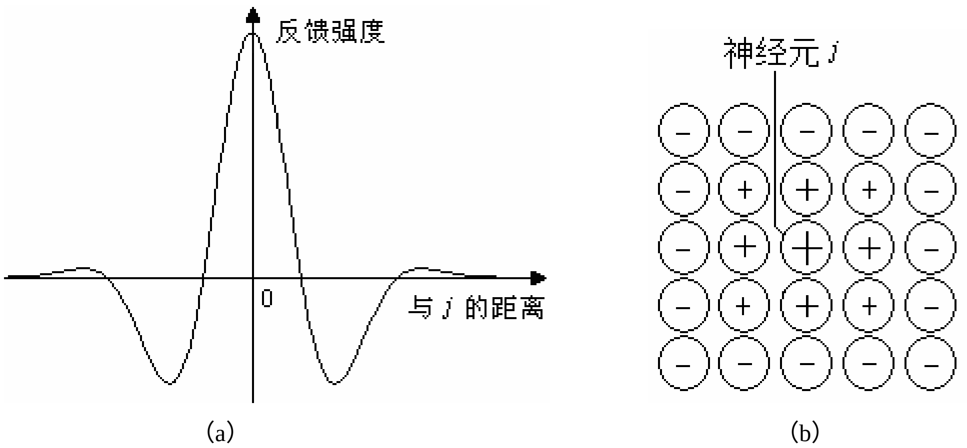
本章介绍 Kohonen 的自组织特征映射（Self-Organizing Feature Mapping，简称 SOFM）[Koho1984]。自组织特征映射是一种竞争学习网络，可以通过神经元之间的竞争实现大脑神经系统中的“近兴奋远抑制”功能，并具有把高维输入映射到低维的能力（拓扑保形特性）。我们先介绍生物系统中的竞争现象，然后介绍 SOFM 的网络结构和学习算法，最后通过仿真例子演示 SOFM 的拓扑保形特性。

6.1　生物系统中的竞争

在第 4 章介绍 RBF 网的生理学基础时，我们曾提到，某些视觉神经细胞在视网膜上有特定的感受野，并具有近兴奋远抑制(on-center off-surround)功能，因此我们用径向基函数建模这样的近兴奋远抑制神经元。在本章，我们从神经元之间互相竞争的角度再来看这一现象。

生物神经网络的研究发现，大脑皮层中，神经元是呈 2 维空间排列的，而且邻近神经元之间通过侧反馈的方式紧密互联。因此每个神经元既有外部区域的输入信号，也有来自同一区域其它神经元的反馈输入信号。而邻近神经元之间侧反馈信号的强度体现为这些神经元之间的连接强度，因此而这些连接权值的分布也体现出明显的“近兴奋远抑制”现象。更具体的说，以某个激活的神经元为圆心，邻近其它神经元根据与该神经元的距离，与之的连接权值呈三个区域的分布：对较邻近的神经元呈强的兴奋性侧反馈；对远邻的神经元呈抑制性侧反馈；对更远的神经元又呈弱的兴奋性侧反馈。通常情况下，可以不考虑第三区的弱侧反馈。

这里所说的邻近神经元，在大脑皮层中是指以某兴奋神经元为圆心．半径约为 50－500μm 左右的其它神经元，而远邻神经元是指半径为 200μm－2mm 左右的神经元。



(a)　(b)

图 6.1　生物神经元中的“近兴奋远抑制”

另外，神经元之间连接强度从增强到抑制的过渡是平滑的，因此相邻神经元连接强

度的分布呈"墨西哥帽"式分布，如图 6.1（a）所示。图 6.2（b）所示为"墨西哥帽"式分布在二维神经元空间的演示。图中 j 为兴奋神经元，神经元标"＋"号表示该神经元接受兴奋性侧反馈信号，"－"号表示接受抑制性侧反馈信号。

大脑神经元正是通过上述侧反馈及其它协同作用，从而识别外界信号。如果在人工神经网络中也以 2 维空间组织神经元，且邻近神经元之间也通过侧向交互作用产生竞争过程，则也能自适应地形成了针对特殊信息的组织结构。于是，神经网格的这种自组织过程和行为，就可以用于检测各种不同信息模式。

6.2　SOFM 结构

如果在学习过程中逐步缩小神经元之间的作用邻域，并用 hebb 学习规则增强中心神经元的激活程度，则去掉各神经元之间的侧向连接也能得到"近兴奋远抑制"的效果。这就是 Kohonen 的自组织特征映射（SOFM）的思路。

图 6.2 为二维自组织特征映射的连接图，可见 SOFM 网络可以将任意的多维输入信号变换到二维离散网格上。图中的两维离散网格即为网络的输出层，SOFM 的所有计算节点都在这一层。离散网格中的每个神经元都与输入层中的所有输入相连，这些神经元排成行和列的形式。图 6.2 中输入为 3 维，输出为 2 维，离散网格中共有 5 行 4 列共 20 个神经元。

应该指出，一维自组织特征映射是二维情况的一个特例。

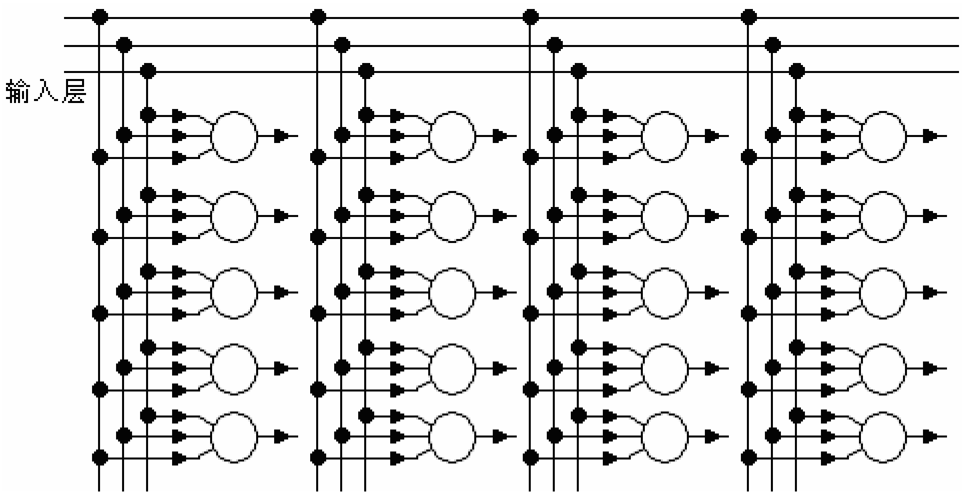


图 6.2　二维 SOFM 网络

自组织特征映射的每个输入模式均对应于二维网格上的一个局部化区域，而且随着输入模式的不同，该区域的位置和性质也各不相同。因此，必须有充分数量的输入模式，才能保证网格中所有的神经元都受到训练，并确保自组织过程的正确收敛。

SOFM 的一个重要特点是具有拓扑保形特性，即最终形成的以输出权矢量所描述的特征映射能反映输入的模式的分布。

6.3　SOFM 的学习算法

算法应首先对权值进行初始。类似于 BP 网，SOFM 的初始权值常取小的随机数。权值初始化后，SOFM 还应完成两个基本过程：竞争过程和合作过程。竞争过程就是最优匹配神经元的选择过程，合作过程则是网络中权系数的自组织过程。选择最优匹配神经元实质是选择输入模式对应的中心神经元，权系数的自组织过程则是以“墨西哥帽”的形态来使输入模式得以存放。这两部分是密切相关的，它们共同作用才能完成自组织特征映射的学习过程。

每执行一次学习，SOFM 网络中就会对外部输入模式执行一次自组织适应过程，其结果是强化现行模式的映射形态，弱化以往模式的映射形态。下面分别对自组织特征映射 SOFM 的学习算法两个过程进行介绍。

(1)　竞争过程

假定从输入空间随机选取的输入模式为 $\boldsymbol{x}=\left[x_1,x_2,\cdots,x_n\right]^T$，自组织特征映射的输出层神经元 j 的对应的权系数向量为 $\boldsymbol{w}_j=\left[w_{j1},w_{j2},\cdots,w_{jn}\right]^T$，$j=1,2,\cdots,l$，其中 l 为输出神经元数目。

竞争过程的主要目的是寻找与输入模式匹配的最佳权矢量，从而确定与输入模式最优匹配的神经元，即竞争获胜的神经元。输入模式 $\boldsymbol{x}$ 和权系数 $\boldsymbol{w}_j$ 的匹配程度可用两者的内积表示，即 $\boldsymbol{x}^T\boldsymbol{w}_j$。内积最大处正是与输入模式均对应的二维网格上的一个局部化区域的中心。由于内积 $\boldsymbol{x}^T\boldsymbol{w}_j$ 最大等价于向量 $\boldsymbol{x}$ 和 $\boldsymbol{w}_j$ 之间的距离 $\left\|\boldsymbol{x}-\boldsymbol{w}_j\right\|$ 最小，因此可根据该最小矩离准则确定最优匹配的神经元 i，即寻找满足下式的神经元：

$$i(\boldsymbol{x})=\arg\min_j\left\|\boldsymbol{x}-\boldsymbol{w}_j\right\|,\quad j=1,2,\cdots,l \tag{6.1}$$

该式子即匹配规则。

上面式子说明：局部化区域的中心就是神经元 $i(\boldsymbol{x})$，该神经元的权系数向量 $\boldsymbol{w}_i$ 与输入模式 $\boldsymbol{x}$ 产生最优匹配。

(2)　合作过程

根据匹配规则求出最优匹配的神经元 i 后，接着就应考虑神经元 i 的邻域神经元的权系数的自组织过程。这是因为虽然 $\boldsymbol{w}_i$ 是求出的对输入模式 $\boldsymbol{x}$ 的最优匹配；但 $\boldsymbol{w}_i$ 仍然不足以充分表示 $\boldsymbol{x}$，故还应对权向量 $\boldsymbol{w}_i$ 进行自组织学习。

如前所述，SOFM 的基本思想是去掉各神经元之间的侧向连接，并在学习过程中逐

步缩小最优匹配的神经元 i 所作用的领域，然后用 hebb 学习规则增强中心神经元的激活程度，以达到"近兴奋远抑制"的效果。因此必须在自组织过程中引入一个领域函数 $\Lambda_{j,i}$，其中 i 为最优匹配神经元，j 为其邻近神经元。该邻域函数应该是学习时间 n 的函数，典型的领域函数可选用 Gaussian 函数：

$$\Lambda_{j,i}(n)=\exp\left(-\frac{d_{j,i}^2}{2\sigma(n)^2}\right),\quad n=0,1,2,\cdots \tag{6.2}$$

上式中 $d_{j,i}^2=\left\|r_j-r_i\right\|^2$，其中 r_i 和 r_j 表示神经元 i 和 j 在二维网格中的离散位置，参数 $\sigma(n)$ 则是 Gaussian 函数的有效宽度。$\sigma(n)$ 应随着学习时间的增加逐步衰减，比如说可采用指数衰减形式：

$$\sigma(n)=\sigma_0\exp\left(-\frac{n}{\tau_1}\right),\quad n=0,1,2,\cdots \tag{6.3}$$

其中 σ_0 为 $\sigma(n)$ 的初始值，τ_1 为指数衰减时间常数。

图 6.3 显示了在 Gaussian 邻域函数作用下四个不同学习时刻 $t_1<t_2<t_3<t_4$ 最佳匹配神经元 i 的邻域变化情况。一般情况下，开始时 $\Lambda_{j,i}(n)$ 较大，随着学习的进行，$\Lambda_{j,i}(n)$ 缩小到只包含最佳匹配神经元 i 邻近的 1－2 个神经元。

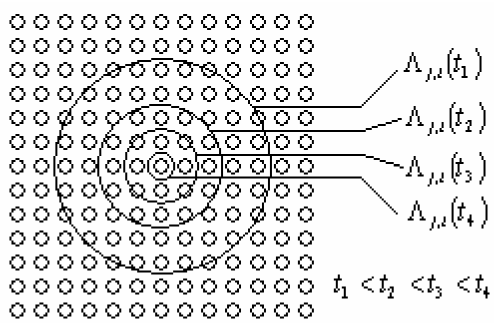


图 6.3　最佳匹配神经元在不同学习时刻的领域

邻域函数确定后，最后应考虑对邻域内神经元 j 的权值 $\boldsymbol{w}_j$ 进行修正。由于直接使用 Hebb 学习规则将导致权值朝着一个方向发展（一直增加或减小），最终导致系统不稳定，因此可考虑增加一个权值的"遗忘"项 $g(y_j)\boldsymbol{w}_j$，其中 y_j 为神经元 j 的输出。于是神经元 j 权系数的调整公式如下：

$$\boldsymbol{\Delta w}_j=\eta(n)y_j\boldsymbol{x}-g(y_j)\boldsymbol{w}_j \tag{6.4}$$

其中$\eta(n)$为时变的学习率，$\boldsymbol{x}$为输入模式。$\eta(n)$可取为

$$\eta(n)=\eta_0\exp\left(-\frac{n}{\tau_2}\right)，\ n=0,1,2,\cdots \tag{6.5}$$

其中η_0为$\eta(n)$的初始值，τ_2为指数衰减时间常数。

为简单起见，如果令$g(y_j)$和y_j分别为

$$g(y_j)=\eta(n)y_j \tag{6.6}$$

$$y_j=\Lambda_{j,i}(n) \tag{6.7}$$

于是神经元j权系数的调整公式简化为

$$\boldsymbol{w}_j(n+1)=\boldsymbol{w}_j(n)+\eta(n)\Lambda_{j,i}(n)(\boldsymbol{x}-\boldsymbol{w}_j(n)) \tag{6.8}$$

综上所述，SOFM 的学习算法如下：

（1）权值初始化：可令各权矢量$\boldsymbol{w}_j$，$j=1,2,\cdots,l$，初始值取小的各不相同的随机数，或干脆随机选择样本输入作为初始权值。令叠代次数$n=1$。

（2）在样本集中随机选择一个模式$\boldsymbol{x}$作为 SOFM 的输入；

（3）在时刻n，根据下式的最小欧氏距离准则选择$\boldsymbol{x}$的最佳匹配神经元i（竞争过程）：$i(\boldsymbol{x})=\arg\min\limits_j\|\boldsymbol{x}-\boldsymbol{w}_j\|$，$j=1,2,\cdots,l$；

（4）根据下式确定邻域函数（合作过程）：$\Lambda_{j,i}(n)=\exp\left(-\dfrac{d_{j,i}^2}{2\sigma(n)^2}\right)$，其中$d_{j,i}^2=\|r_j-r_i\|^2$，$\sigma(n)=\sigma_0\exp\left(-\dfrac{n}{\tau_1}\right)$；

（5）令学习率为$\eta(n)=\eta_0\exp\left(-\dfrac{n}{\tau_2}\right)$，然后按下式对邻域内的所有神经元进行权值修正：$\boldsymbol{w}_j(n+1)=\boldsymbol{w}_j(n)+\eta(n)\Lambda_{j,i}(n)(\boldsymbol{x}-\boldsymbol{w}_j(n))$；

（6）如果已形成稳定的特征映射，则结束学习；否则令$n=n+1$，转（2）。

在 SOFM 的学习算法中，学习率$\eta(n)$、邻域函数$\Lambda_{j,i}(n)$的选择对学习效果影响很大，Kohonen 讨论了 SOFM 的权值自适应过程的两个阶段[Koho1984，Koho1997]，并给出了这些参数的参考值：

1. 自组织阶段（排序阶段）。这是自适应过程的第一阶段，发生在 SOFM 的学习算法的开始约 1000 此叠代中。在该阶段，权矢量从无序走向有序。对学习率$\eta(n)$，其初始值应在 0.1 附近，此后逐渐减小，但仍应大于 0.01，故建议$\eta_0 = 0.1$，$\tau_2 = 1000$；对邻域函数$\Lambda_{j,i}(n)$，开始时应包含以最佳匹配神经元为中心的网络中的几乎所有的神经元，然后逐步减小，直至只剩最佳匹配神经元，或者其周围一两个神经元为止，故建议σ_0取输出神经元网格的半径，$\tau_1 = \dfrac{1000}{\log \sigma_0}$。
2. 收敛阶段。这是自适应过程的第二阶段，需要的时间较长，主要完成对特征映射的精调。收敛阶段的$\eta(n)$应保持较小值，大致为 0.01 左右，对$\Lambda_{j,i}(n)$，应从只包含以最佳匹配神经元为中心的多个神经元，缩小到只剩周围一个或零个神经元。

6.4 仿真例子

这里，我们首先演示了二维平面上均匀分布的数据映射到二维网格上的情况。SOFM 的输出神经元排列成 10 行 10 列，共 100 个神经元；输入训练模式则为平面区域$[-1,+1]\times[-1,+1]$内均匀分布的样本。图 6.4 为用到的输入训练模式集。

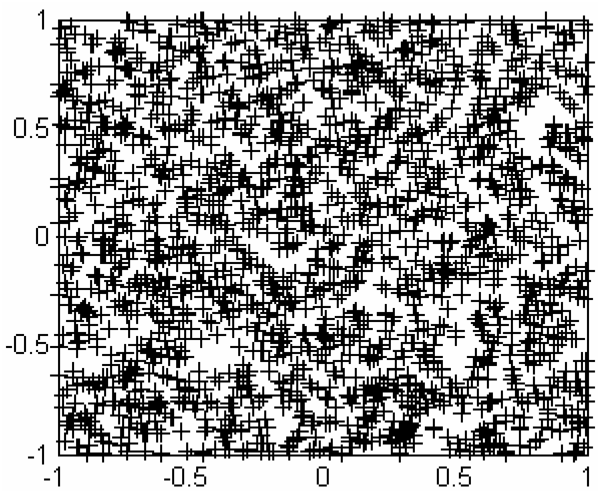


图 6.4 训练二维 SOFM 的输入模式

该例子中 SOFM 的训练参数设置如下：学习率$\eta(n)$的初始值为$\eta_0 = 0.1$，衰减常数$\tau_2 = 1000$，意味着$\eta(n)$经过 1000 次训练后衰减到 0.0368；邻域函数$\Lambda_{j,i}(n)$的初值$\sigma_0 = 3$，衰减常数$\tau_1 = 1000/(\log\sigma_0)$，意味着$\Lambda_{j,i}(n)$经过 1000 次训练后衰减到 1；所有神经元初始权值取[-0.1,+0.1]内的随机数；共训练 2000 次后结束。

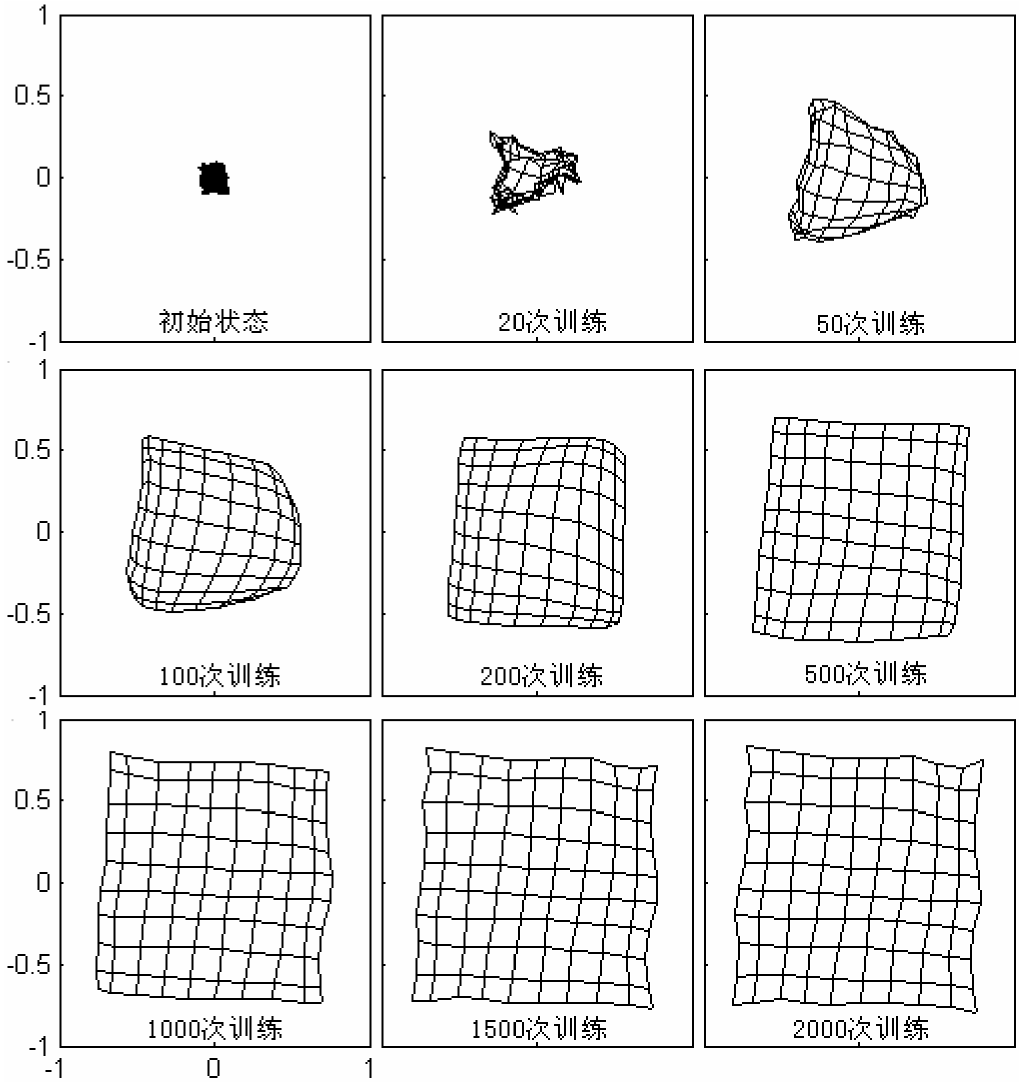


图 6.5　训练二维 SOFM 时映射图的演变

图 6.5 演示了在映射输入分布时，SOFM 学习算法在不同学习阶段的映射图，该映射图通过连接邻近神经元的二维权矢量得到。由图可见，训练开始时，网格中所有神经元的权值分布在[-0.1,+0.1]×[-0.1,+0.1]的较小区域内，随着学习的进行，权矢量很快从无序走向有序（排序阶段），在大约进行了 1000 次训练后，排序阶段结束，算法进入收敛阶段，此时算法主要完成精调，因此映射图已经没有明显的变化。与输入模式的分布

相比较，可见图 6.5 中最终的映射图已经完全实现了对输入模式的逼近。

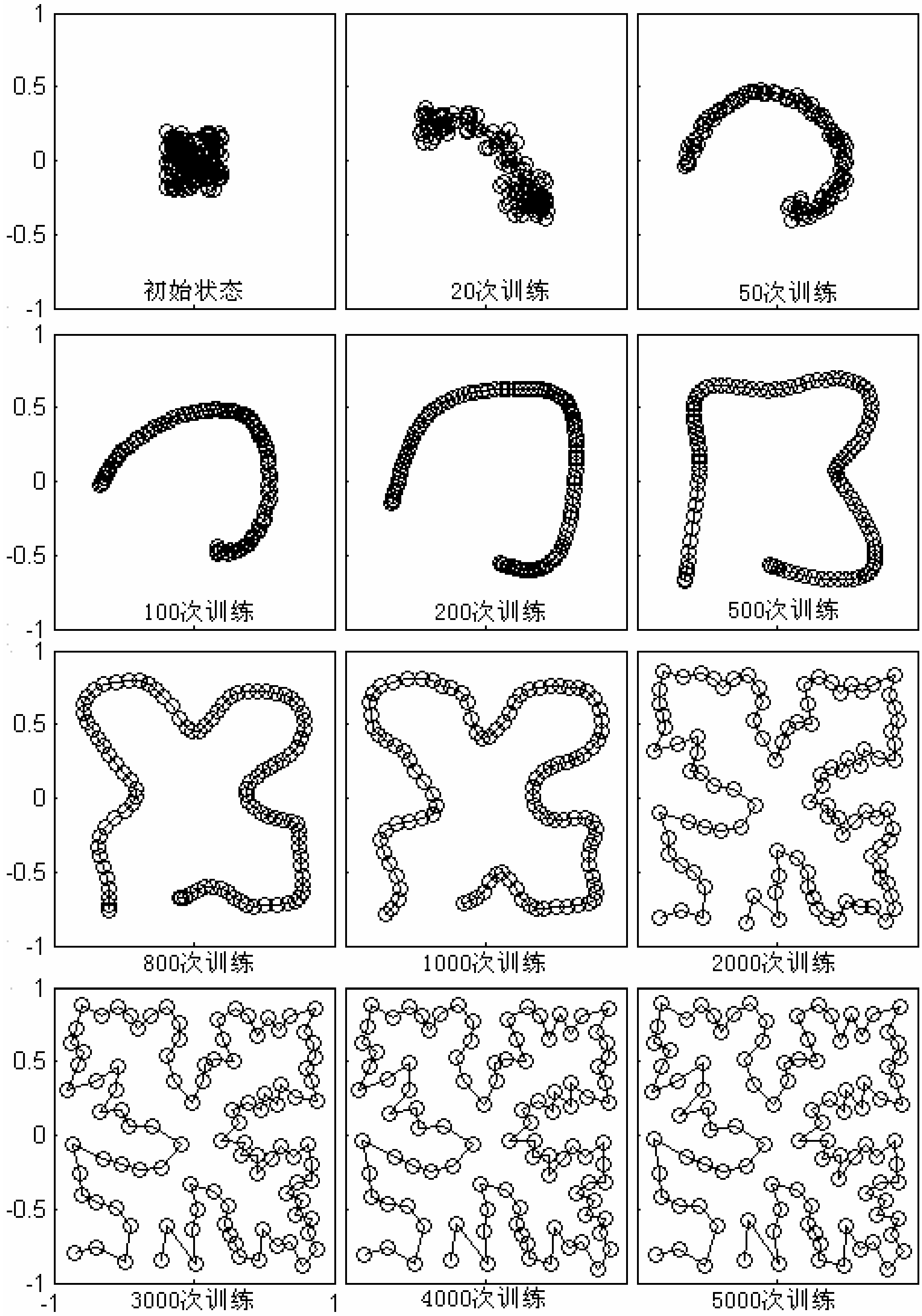


图 6.6　训练一维 SOFM 时映射图的演变

为了研究输入模式维数大于特征映射维数时的映射，我们接着仿真了二维平面上均匀分布的数据映射到一维网格的情况，如图 6.6 所示。此时我们也使用 100 个输出神经

元，但只在一维空间（即直线上）平均排列和编号。该例子中训练模式的分布，以及其它训练参数设置同二维的情况完全一样，差别是学习率衰减常数 $\tau_2 = 3000$，邻域函数 $\Lambda_{j,i}(n)$ 的初值 $\sigma_0 = 18$，所有神经元初始权值取 $[-0.2, +0.2]$ 内的随机数，共训练 5000 次后结束。

与二维 SOFM 相比，在逼近均匀分布二维数据时，一维 SOFM 的实现方式是使各输出权矢量尽可能均匀分布在 $[-1, +1] \times [-1, +1]$ 区域内。由此也可以看到，所谓的 SOFM 能把多维空间映射到一维，实际上是映射到同一多维空间内的一条曲线上（而不是直线）；同样，所谓的 SOFM 能把多维空间映射到二维，实际上也是映射到同一多维空间内的一个二维曲面上（而不是平面）。