

363-368
⑤

动态连接权的目标搜索神经网络模型*

余波 张立明

(复旦大学电子工程系, 上海, 200433)

R318.18
TP18

摘要 提出一种基于初级视觉皮层的图像目标搜索模型, 该模型只采用方位选择性细胞和皮层内有限范围水平连接等 V1 基本单元, 以链码表示的目标轮廓作为知识, 允许该知识以时间脉冲形式控制 V1 区内神经细胞的动态活动, 和网络中的动态连接权相配合, 使与知识轮廓形状相符合的轮廓内细胞逐步进入并维持在兴奋状态, 最终实现对视野中特定目标轮廓的提取。

关键词 知识, 初级视觉皮层, 轮廓, 动态连接权。

神经网络模型 目标搜索

引言

近年来, 人们对视觉皮层的研究有了很大进展, 但对知识是如何存贮, 如何被用来控制视觉信息处理过程的, 还没有一个清楚的认识, 这主要是因为较高级视觉皮层的高度复杂性。我们提出一种基于视觉生物学知识基础之上的、用于目标识别定位的模型, 尝试对知识的表达和利用。该模型的目的是在视野中搜索出一个已经记忆的目标, 它只采用了初级视觉皮层的单元, 而存贮在较高级皮层的知识则表达为目标图像的轮廓。

在视觉系统的早期阶段, 单个神经元只对处于它们感受野中的刺激做出反应, 比如特定方向的边缘, 而感受野的范围一般比较小, 视觉系统必须把这些局部的单元信息汇集成具有全局意义的特征, 才能感知视野中的目标^[1]。李兆平^[2]对如何将边缘段集成为轮廓的神经机制进行了研究, 即轮廓集成, 她的模型只采用初级视觉皮层的简单单元就可以实现对平滑封闭轮廓的提取, 但没有解决知识表达和利用的问题, 它只提取视野中所有的轮廓, 还不能根据知识有选择地提取轮廓。我们提出的目标搜索模型 OSM (Object Searching Model) 也只采用了 V1 单元, 即方位选择性细胞和皮层内有限范围水平连接, 它以目标轮廓作为知识, 并且允许该知识控制 V1 区神经元的活动, 和网络中的动态连接权相配合, 最终实现对视野中特定目标的提取。

1 作为知识的链码

在 OSM 中的知识是由目标的轮廓链码编码来表示的, 设目标的链码为 $C = (c_0, c_1, \dots, c_1, \dots, c_{L-1}, c_L)$, 其中 $c_k \in \{\theta_k \mid \theta_k = k\pi/K \text{ for } k = 1, 2, \dots, K\}$, L 是轮廓链码的长度, θ_k 表示 V1 区内神经的第 k 个可能的最优朝向。这里 $c_0 = c_L$, 以形成一个封闭的轮廓链码。

由于链码 C 用于控制神经元的动态活动时, 是首尾交替送入神经网络的, 所以要调整为 $C' = (c'_0, c'_1, \dots, c'_1, \dots, c'_{L-1}) = (c_0, c_{L-1}, c_1, c_{L-2}, c_2, \dots)$ 。调整的规则是: 对于 mod

* 国家自然科学基金(编号: 39870194) 和高等学校博士学科专项科研基金资助项目
稿件收到日期 1999-03-09, 修改稿收到日期 1999-06-02

$(t, 2) = 0$, 则 $c'_t = c_{t,2}$; 对于 $\text{mod}(t, 2) \neq 0$, 则 $c'_t = c_{t-1,1,2}$. 在 C' 中包含了目标的轮廓形状信息, 它将被存贮在较高级视觉中枢, 用以控制匹配过程.

2 目标搜索模型(OSM)

初级皮层的一些神经元主要对它们感受野内的边缘段做出反应, 具有相互覆盖的感受野, 不同最优朝向的神经元往往是聚集在一起的^[1], 另外, 神经元的活动还受它们的感受野外的刺激活动的影响, 这种影响被归结于皮层内的水平联接^[1-4].

在我们的功能模型中, 视觉区由三维结构的神经网络组成, 如图 1 所示, 左图中每个位置 i 上的棘状体都是由 K 个右图所示的神经元构成的超柱, 其中每条小线段代表位于 i 的一个神经元, 线段的朝向为该神经元的最优朝向 (orientation), 棘状体的二维分布情况和视野中的二维输图象是相对应的. 每个在位置 i 并且具有最优朝向 $\theta \in \{\theta_k | \theta_k = k\pi/K \text{ for } k=1, 2, \dots, K\}$ 的神经元, 用 $(i\theta)$ 来表示, 神经元具有膜电位 $x_{i\theta}$ 和模拟发放率输出 $g(x_{i\theta}) \geq 0$. 视野中的图像经过 V1 内的边缘检测机构后^[5], 其边缘状况作为视觉输入 $I_{i\theta}$ 送至视区相应的空间位置 i 上. 神经元 $(i\theta)$ 的邻近神经元 $(j\theta)$ 可通过发送单突触兴奋信号 $w_{i\theta, j\theta} g(x_{i\theta})$ 而使之兴奋, 这里 $w_{i\theta, j\theta}$ 表示水平连接的突触强度. 另外, 每个神经元还接收来自较高级视觉中枢的输入.

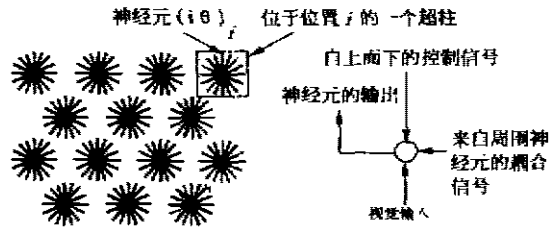


图 1 视觉区、超柱和神经元示意图
Fig. 1 Visual area, hypercolumn and neuron model

2.1 神经元动态方程

在 OSM 中, 神经元的动态特性描述为

$$\dot{x}_{i\theta} = -\alpha \cdot x_{i\theta} + \beta \cdot [I_{i\theta} \cdot J_{i\theta} \sigma(x_{i\theta}, 0) - T] + \gamma \cdot S_{i\theta} \quad (1)$$

式中 $I_{i\theta}$ 为视觉输入, 如果在处于 $(i\theta)$ 的神经元的感受野内, 存在与该神经元的最佳朝向 θ 相同朝向的边缘, 则 $I_{i\theta} = \hat{I} > 0$, \hat{I} 为常数, 否则, $I_{i\theta} = 0$; $J_{i\theta}$ 是自上而下的控制信号, 知识就是通过它来控制网络中的神经元的, $J_{i\theta} = \hat{J} \delta(c'_t, \theta) \geq 0$, 其中 \hat{J} 为常数, 函数 $\delta(c'_t, \theta)$ 满足当 $c'_t = \theta$ 时, $\delta = 1$, 当 $c'_t \neq \theta$ 时, $\delta = 0$; 函数 $\sigma(x_{i\theta}, 0)$ 则在 $x_{i\theta} > 0$ 时, $\sigma = 1$, 当 $x_{i\theta} \leq 0$ 时, $\sigma = 0$, 它保证了神经元只有处于静止状态时, 才受到高层信号的控制; T 是阈值, 用于控制神经元对知识信号变化的敏感程度; $1/\alpha$ 为神经元的膜时间常数, 参数 β, γ 分别控制神经元的活动相对于视觉输入和相邻神经元耦合的灵敏度.

$S_{i\theta}$ 表示由周围神经元来的耦合, 表示为

$$S_{i\theta} = \sum_{j \neq i, \theta} w_{i\theta, j\theta} g(x_{j\theta}). \quad (2)$$

其中 $w_{i\theta, j\theta} \in [0, \hat{W}]$ 表示神经元 $(i\theta)$ 和它周围相邻神经元的权连接, 我们定义 $w_{i\theta, j\theta} \geq \hat{W}$ 为权处于激活状态; $g(x_{i\theta})$ 是神经元的输出, 它是一个 Sigmoid 函数: $g(x_{i\theta}) = 1/[1 + e^{-\varphi(x_{i\theta} - \theta)}]$; 常数 φ 决定了曲线在 x 轴上的位置, 各参数满足 $T - \hat{I} - \hat{J} < \hat{W} < T - \hat{J}$ 和 $\hat{I} < \hat{J} < \hat{W}$.

2.2 动态联接权

2.2.1 神经元间连接方式

在OSM中,每个神经元并不是和所有相邻的神经元都有连接,如果可以找到一条平滑连续轮廓通过神经元 $(i\theta)$ 和 $(j\theta')$,则它们之间存在权连接, $W_{i\theta,j\theta'}$ 可以动态变化,否则 $W_{i\theta,j\theta'}$ 将恒为0,即没有连接,如图2所示.图中间的神经元(粗实线)只与周围最优方向相近的神经元(细实线)有权连接,而与其它神经元(虚线)无连接.于是,在不同视觉方位柱之间,只有最优朝向相近的神经元连接,这是符合生物实验的^[6,7].OSM中的权连接对于各个不同最优朝向的神经元是完全对称的.这样构造 $W_{i\theta,j\theta'}$ 的目的是使神经元能够根据周围神经元的上下文状况做出响应,并且在同一平滑轮廓内的神经元能相互增强彼此的活动,而由噪声所引起的孤立的边缘段则会因得不到增强信号而不能保持兴奋.

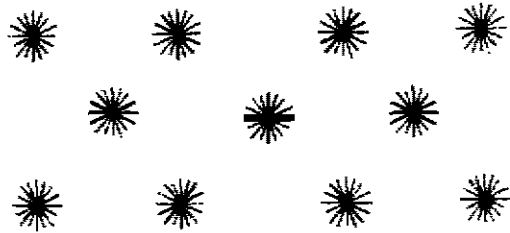


图2 神经元连接模式

Fig. 2 The connection pattern of neurons

我们设计连接权的动态方程为

2.2.2 连接权的动态方程

我们设计连接权的动态方程为

$$\begin{cases} \dot{w}_{i\theta,j\theta'} = -\epsilon[\mu(1 + \tanh(z/\rho)) + \eta + w_{i\theta,j\theta'}], \\ \dot{z} = 3z - z^3 - 3 + w_{i\theta,j\theta'} - g(t). \end{cases} \quad (4)$$

这是一个松弛振荡子方程^[8],这里 $g(t) = g(x_{i\theta}) + g(x_{j\theta'})$, z 是辅助变量, $\epsilon, \mu, \rho, \eta$ 均为常量.方程(4)中存在由参数 $g(t)$ 控制的霍夫分叉, $g(t)$ 用来调整 z 零梯度曲线(三次曲线)的位置,当它和 w 零梯度曲线(S曲线)只有一个交点时,为极限环状态;如果曲线有3个交点,则其中一个稳定点,如图3所示,其中LB和LK分别是S曲线的左分支及其左拐点,RB和RK分别是S曲线的右分支及其右拐点.

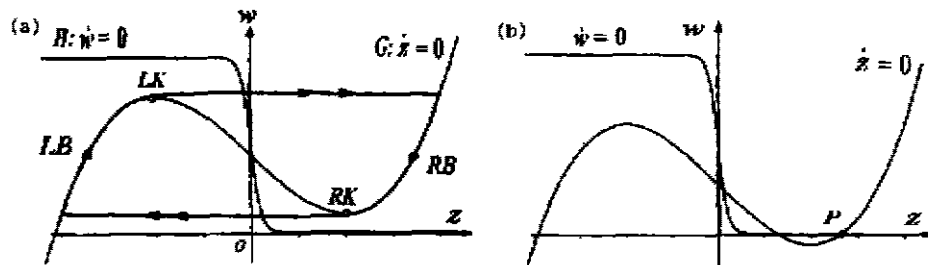


图3 相空间零梯度曲线图,(a)极限环状态,(b)稳定点状态

Fig. 3 Two states of the weight equation, (a)limit cycle state, (b)stable point state

连接权可以根据它所连接的两个神经元的状态,动态调整自身的值.当2个神经元都处于静止状态时, P_1 是稳定点,对应权值为0,如图4(a)所示.当有一个神经元进入兴奋状态后,三次曲线上抬,稳定点 P_1 消失,进入极限状态,状态沿LB分支向上演变,权值开始增大,如图4(b).如果紧接着另一个神经元也开始兴奋,则如图4(c),三次曲线继续上抬和S曲线相交,出现稳定点 P_2 ,对应于权值为 \dot{w} ,于是连接权从第一个神经元兴奋开始,经过一定时间延迟进入了激活状态.如果两个神经元都保持兴奋,则权值将锁定在激活状态;否则,

一个神经元的兴奋还不足以使两曲线再相交,状态将越过LK向RB演变,使权值减小,权很快退出激活状态.与此同时,已经兴奋的神经元由于得不到另一神经元的兴奋耦合,也会向静止状态演变,三次曲线持续下落,如图4(d)所示,在状态越过RK前,和S曲线相交,又出现图4(a)所示的稳定点 P_1 ,使权值锁定在0.

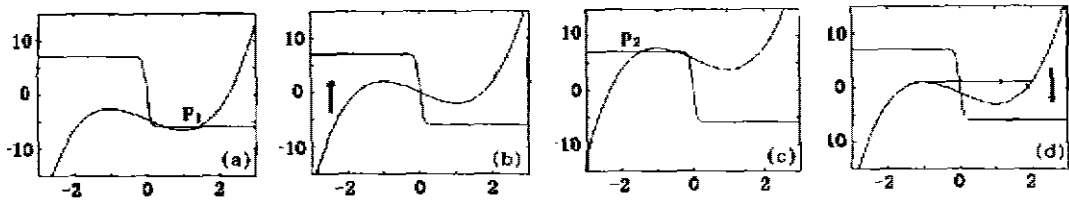


图4 权值在相空间的演变图(纵坐标为 w ,横坐标为 z)
Fig. 4 Evolution of the weight value in phase space (z, w)

3 目标搜索过程

目标搜索实际上是轮廓增长的过程,最终在视区中有一组相邻的兴奋神经元组成一条封闭的链,这些神经元的最优方向连在一起,就形成了目标的轮廓.由于视区中神经元的分布和视野是拓扑对应的,因而这轮廓的位置就是待搜索目标在视野中的位置.

假设链码 C' 中的每个 c'_i 在送入神经网络后保持的时间都为 τ ,在该期间神经元有足够的时间对其做出适当的反应,称之为神经元反应周期.在匹配过程中,作为知识的 c'_i 将会在 $t \wedge \tau$ 时刻通过 J 影响视区中的神经元,如果这时一个 $\theta=c'_i$ 的神经元($i\theta$)恰好也有视觉输入 \hat{I} ,并且可以通过激活的权从在它的一个相邻神经元(jc'_{i-1})得到兴奋耦合,则该神经元将会进入兴奋状态.这里 θ, c'_i 的值均为一个表示朝向的角度,为便于分析,也把 $\theta=c'_i$ 的神经元($i\theta$)表示为 (ic'_i) ,象 (ic'_i) 、 (jc'_{i-1}) 这样的神经元连接在一起,就会形成一条候选边缘链CEC(Candidate Edge Chain).由于现实图像的随机性,在某一时刻,在视区内的不同位置有可能已增长出多条CEC,然而在经过时间 $L \wedge \tau$ 后,只有形成封闭链的CEC才是要匹配的目标轮廓OC(Object Chain).如果CEC在某个 τ 内没有增长,则组成它的神经元将会逐渐地返回静止状态.

如果某CEC只是和目标轮廓是部分匹配的话,则最终不会形成封闭的轮廓.因为在CEC开始与OC不符的那一刻,处于CEC头部的神经元将会在 τ 时间内退化到静止状态,随后CEC内的神经元会一个接一个地退化到静止状态,CEC将最终在视区内消失.由于每个CEC都有2个端点,所以CEC的增长是在它们两端交替进行的,否则会出现一端增长,而一端退化的情况,这也是第一节中对链码进行调整的原因.

连接权经过一个周期的延迟才能激活,保证了在一个周期内CEC只可能有一端得到增长.连接权及时退出激活状态,以及式(1)中只有静止神经元才受高层影响的条件,保证了CEC的某一端一旦开始退化就不会再在这端增长,将引起连锁退化反应,使该CEC中的其它神经元将一个个进入静止状态,最终使该CEC在视区中消失.

4 计算机仿真结果

我们取参数 $\hat{I}=3, \hat{I}=2, T=9.5, \hat{W}=4.3, \alpha=0.1, \beta=1, \gamma=2, \varphi=0.03, \varepsilon=0.015, \mu=$

6, $\rho=0.1, \gamma=1.1$. 先用二值图像测试模型. 结果如图 5 所示.

在图 5(a)为匹配的目标图像;(b)为对(a)进行边缘提取后得到的轮廓,从该轮廓就可以得到作为知识的链码;(c)为待搜索的图像,我们的目的就是要在其中找出目标图像(a);(d)表示视区内的神经元接收视觉输入的情况,图中有方向的小线段表示一个接收到视觉输入的神经元,当图像(c)出现在视野中时,按照我们前面的定义,如果一个神经元 RF 内有与它的最优朝向相同的边缘出现时,该神经元有视觉输入,从(d)可以看出,有视觉输入的

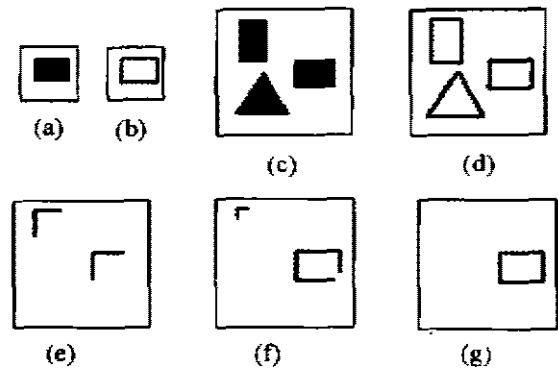


图 5 仿真结果
Fig. 5 Simulation results

神经元的分布是和图像(c)中的边缘分布情况相一致的.(e)、(f)、(g)表示了视区中的神经元的输出 $g(x, y)$ 随时间演变的情况,图中的小线段表示处于兴奋状态的神经元.(e)表示在 24τ 时刻的情况,可以看到视区内有两条 CEC 在同时增长,而到达时刻 67τ 时,如(f)所示,由于左上角的 CEC 不和目标轮廓相符,它的神经元已经开始退化到静止状态了,而右边的那条 CEC,也就是 OC,将增长成一个封闭的轮廓.到达 85τ 时刻时,(g)中左上角的 CEC 中的神经元已经完全返回静止状态,整个视区中只有 OC 中的神经元保持在兴奋状态,从而完成了图像的匹配工作.

图 6 为采用灰度图像的目标识别情况,被搜索的目标是图 6(a)所示的停放在街道中的一辆汽车.(b)是视野中待搜索图像,(c)是接收视觉输入的神经元分布,(d)为目标定位结果,兴奋的神经元指示出了目标的位置.

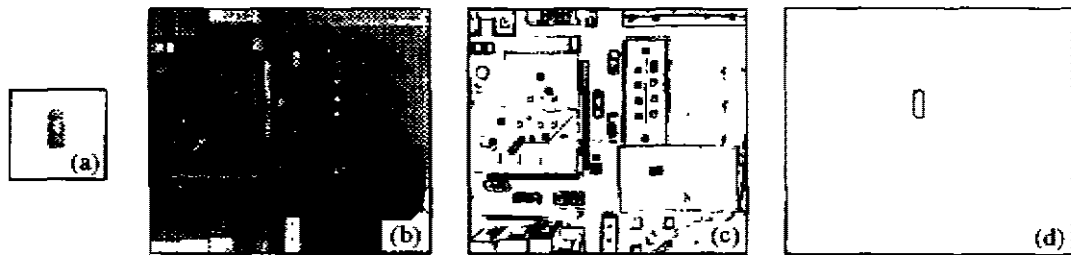


图 6 用于灰度图像目标的识别
Fig. 6 Object recognition of gray level image

5 讨论

OSM 模型的最终效果是经过一定时间后,在视区内浮现出目标的轮廓,匹配速度与视区大小无关,只取决于时间周期 τ 和轮廓提取精度.可以看到,在该模型中,知识表示为目标轮廓链码,这不论对知识存贮和图像匹配都是高效率的,目标轮廓的信息被转变为时间脉冲串送入视区内的每个神经元,实现了神经网络中的并行图像匹配.

OSM 模型的匹配精度取决于轮廓提取的精度,这是由图 1 所示的视区边缘检测尺度决定的,由于本模型采用一个固定的尺度,所以还不能进行多分辨率的图像匹配,本模型现已

经用于处理灰阶图像,并实现了平移不变性,下一步研究工作是实现旋转不变性。

由于到目前为止,人们对知识在脑内是如何存贮和被利用的,还没有一个较清楚的认识,本文假设知识是作为目标的轮廓来存贮的,希望能在这方面的做进一步探索。

致谢 作者感谢复旦大学生理和生物物理学系顾凡及教授对本文的指正。

REFERENCES

- 1 Hubel D H, Wiesel T N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex, *J. Physiol.*, 1962, **160**:106-154
- 2 Li Zhaoping. A neural model of contour integration in the primary visual cortex, *Neural Computation*, 1998, **10**(4): 903-940
- 3 Knierim J J, Van Essen D C. Neuronal responses to static texture patterns in area V1 of the alert macaque monkeys *J. Neurophysiol.*, 1992, **67**:961-980
- 4 Kapadia M K, Ito M, Gilbert C D, et al. Improvement in visual sensitivity by changes in local context: parallel studies in human observers and in V1 of alert monkeys, *Neuron*, 1995, **14**(4):843-56
- 5 Marr D Vision. *A Computational Investigation into the Human Representation and Processing of Visual Information*. San Francisco:Freeman, 1982
- 6 Gilbert C D. Horizontal integration and cortical dynamics, *Neuron*, 1992, **9**(1):1-13
- 7 Weliky M, Kandler K, Fitzpatrick D, et al. Patterns of excitation and inhibition evoked by horizontal connections in visual cortex share a common relationship to orientation columns, *Neurons*, 1995, **15**: 541-552
- 8 David Terman, Wang De-Liang, Wang Glo-Bal. Competition and local cooperation in a network of neural oscillators, *Physica D*, 1995, **81**:148

NEURAL NETWORK MODEL WITH DYNAMIC WEIGHTS FOR OBJECT SEARCHING*

YU Bo ZHANG Li-Ming

(Department of Electrical Engineering, Fudan University, Shanghai 200433, China)

Abstract A new model based on the structure of V1 area in the biological visual system was presented for image searching. The model only uses the known V1 elements, such as orientation selective cells and horizontal intra-cortical connections. The knowledge, which will control the dynamic of the neurons, is represented by the chain code of the object's contour. The chain code is presented to the neural network in timing pulse form. By cooperating with the dynamic weights, only the neurons of the contour according with the object contour will reach the active state and stay there. Finally the contour that conforms to the knowledge contour emerges in the visual area, and the object is located.

Key words knowledge, primary cortex, contour, image searching.

* The project supported by the National Natural Science Foundation and by the Ph. D Program Foundation of State Education Committee of China
Received 1999-03-09, revised 1999-06-02