

文章编号:1001-9081(2008)12-3147-03

一种基于竞争层神经网络的新型图像轮廓线分组算法

田树华, 宋立明, 王 庚

(唐山学院东校区 信息工程二系, 河北 唐山 063020)

(tsjx_tsh@163.com)

摘 要:提出了一种基于竞争层神经网络(CLNN)的新型图像轮廓线分组算法。该算法借助于 CLNN 对生物大脑视觉皮层侧向连接和 WTA 连接模式的功能性模拟,采用与生物神经元类似的视觉信号处理机制对图像的相应区域进行分组感知。实验证明,所提出的算法具有分组能力强、抗噪声能力突出的优点。

关键词:竞争层神经网络;轮廓线分组;图像处理

中图分类号: TP301 **文献标志码:** A

Novel image contour grouping algorithm based on competitive layer neural network

TIAN Shu-hua, SONG Li-ming, WANG Geng

(Department two of Information Engineering, Tangshan College, Tangshan Hebei 063020, China)

Abstract: An image contour grouping algorithm based on Competitive Layer Neural Network (CLNN) was proposed. The presented algorithm, which is inspired from the lateral interactions and WTA interaction within the biological visual cortex, performs the grouping perception on the image by using the visual signal processing mechanisms similar to the biological neurons. The results of experiments show that the proposed algorithm exhibits stronger grouping capability and promising anti-noise capability.

Key words: Competitive Layer Neural Network (CLNN); contour grouping; image processing

0 引言

竞争层神经网络(Competitive Layer Neural Network, CLNN)源自于对生物大脑视觉皮层的功能性模拟^[1]。它是一种典型的反馈式神经网络,其神经元按层次排列,层内神经元之间相互激励,而层间神经元之间相互抑制。这样的结构简洁而富有规律,且已被证明与生物大脑视觉皮层神经元的结构具有某种相似性。而更为重要的是,这种激励和抑制相互耦合的网络赋予了神经元丰富的动力学特性^[2-3],在这些特性的支撑下,CLNN 体现出强大的信息——特别是图像信息处理功能,因此在诸多工程领域得到了广泛的应用^[4-6]。

目前,关于图像轮廓线的提取已有多种高效的方法,但尚没有一种非常有效的方法能够对提取的轮廓线进行正确的分组。本文提出了一种基于竞争层神经网络的新型图像轮廓线分组算法。该算法采用与生物神经元类似的视觉信号处理机制对图像的相应区域进行分组感知,通过采用基于 Gestalt 原理^[7]的神经元侧向连接建立机制和基于 WTA(Winner Takes All)的横向竞争机制提取出图像的轮廓线并对其进行分组。另外,在 CLNN 现有算法的基础上提出了一种基于退火的并行随机优化算法,该算法具有更短的收敛时间和更高的收敛精度。采用真实图片对所提出的算法进行了验证,并将 CLNN 方法同著名的 Potts spin^[8]方法进行了对比。实验结果证明:CLNN 方法具有强于传统方法的轮廓线分组性能和抗噪声能力。

1 竞争层神经网络

竞争层神经网络的结构如图 1 所示。网络总共包含 L 层,

每一层由 R 个神经元组成。任意一个神经元的状态由 x_{ij} 表示,其中 $1 \leq i < R, 1 \leq j < L$ 。第 l 层的任意两个神经元 x_{il} 和 x_{sl} 之间存在侧向连接 f_{rs}^l ,表示两个神经元在同一层的耦合程度;而第 r 列的任意两个神经元 x_{rl} 和 x_{ml} 之间存在抑制性的 WTA 连接 I_r^m ,代表层与层的相互竞争强度。此外,网络具有一个单独的输入层,共包含 R 个输入。每一个输入 h_r 被并行地输送到对应列 r 的 L 个神经元中,连接强度为 J_r 。

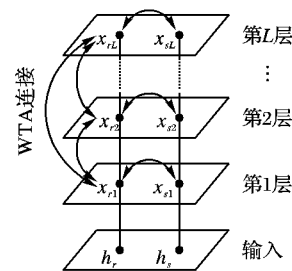


图 1 CLNN 结构

基于以上描述,CLNN 的学习算法可以被严格定义如下:

$$\frac{dx_{rl}}{dt} = -x_{rl} + \sigma(J_r h_r - \sum_m I_r^m x_{ml} + \sum_s f_{rs}^l x_{sl} + x_{rl}) \quad (1)$$

其中 $\sigma(x) = \max(0, x)$ 是非饱和线性阈值函数。容易证明,虽然 $\sigma(\cdot)$ 函数是非线性无界函数,然而在(1)式所定义的动力系统演变下,每一个神经元的状态变量却是有界的。实际上,将式(1)改写如下:

$$\frac{dx_{rl}}{dt} = -x_{rl} + \sigma(E_{rl} + x_{rl}) \quad (2)$$

其中 $E_{rl} = J_r h_r - \sum_m I_r^m x_{ml} + \sum_s f_{rs}^l x_{sl}$ 代表神经元 x_{rl} 的外部净

收稿日期:2008-06-25;修回日期:2008-08-07。

作者简介:田树华(1973-),男,河北唐山人,讲师,硕士,主要研究方向:数据库开发、人工智能软件;宋立明(1979-),男,河北唐山人,助理讲师,硕士,主要研究方向:网络信息系统;王庚,男,助理讲师,硕士,主要研究方向:网络信息系统。

输入。可以构造能量函数如下:

$$E = - \sum_i J_i h_i x_{ri} + \frac{1}{2} \sum_r \sum_m I_{rm}^l x_{ri} x_{rm} - \frac{1}{2} \sum_i \sum_{rs} f_{rs}^l x_{ri} x_{si} \quad (3)$$

显然 $\partial E / \partial x_{ri} = -E_{ri}$, 因此有:

$$\frac{dE}{dt} = \sum_i \frac{\partial E}{\partial x_{ri}} \frac{dx_{ri}}{dt} = - \sum_i E_{ri} (-x_{ri} + \sigma(E_{ri} + x_{ri})) \quad (4)$$

另一方面,通过 $\sigma(\cdot)$ 的非负特性,容易证明在式(1)的动力系统作用下,神经元 x_{ri} 保持非负,即 $x_{ri} \geq 0$ 。此时式(4)的计算可以分为两种情况。若 $E_{ri} + x_{ri} \geq 0$, 可知 $\sigma(E_{ri} + x_{ri}) = E_{ri} + x_{ri}$, 从而 $E_{ri}(-x_{ri} + \sigma(E_{ri} + x_{ri})) = E_{ri}^2 \geq 0$ 。若 $E_{ri} + x_{ri} < 0$, 结合 $x_{ri} \geq 0$ 可知 $E_{ri} < 0$, 另外有 $\sigma(E_{ri} + x_{ri}) = 0$, 从而 $E_{ri}(-x_{ri} + \sigma(E_{ri} + x_{ri})) = -E_{ri}x_{ri} \geq 0$ 。基于以上分析可知 $dE/dt \leq 0$ 。此外,容易证明当 x_{ri} 趋于正无穷时,能量函数趋于无穷。综合以上分析可知:式(1)所定义的 CLNN 动力方程使得神经元的状态在有界区域内变化。这一结论对于我们将 CLNN 应用于图像处理具有十分重要的意义。因为只有神经元的状态有界,将图像特征绑定为神经元状态才会具有实际操作含义。否则,即便在理论上可行,无限增长的神经元活跃度也将极大地影响计算的实时效率,降低图像处理能力。

2 基于 CLNN 的轮廓线分组

本文提出的基于竞争层神经网络的图像轮廓线分组算法分三个步骤,即图像轮廓线的特征提取,基于轮廓线特征的竞争层神经网络权值生成,以及基于式(1)的神经网络动力系统迭代。

2.1 图像轮廓线特征提取

将输入图像划分为 R 个不相交的矩形子图。对于矩形子图 r , 采用 3×3 的 Sobel 算子对其进行预处理,从而得出其灰度变化的梯度信息。基于此,在每一个子图内选取具有最大 Sobel 平方响应的坐标点,过该点绘出代表该子图的特征单位向量 \mathbf{n}_r 。该向量将作为 CLNN 的第 r 个输入特征,特征 r 的网络输入强度定义为 $h_r = \text{sobel}_x^2 + \text{sobel}_y^2$, 其中 sobel_x 和 sobel_y 分别是子图 r 在特征点沿水平和垂直方向的 sobel 响应值。图 2 描绘了采用该方法对 Lena 图像进行特征提取的效果。图(a)为原始输入图像,图(b)为采用 10×10 子图窗口进行特征提取的效果。图中特征向量 \mathbf{n}_r 的粗细正比于网络输入值 h_r 。

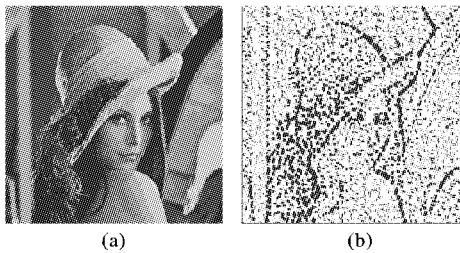


图 2 特征提取示意图

2.2 网络连接权值

CLNN 轮廓线分组模型共包含 L 层神经元,其中 L 是图像轮廓线的分组个数,需预先设定。第 1 层为背景层,用以分离图像背景,第 2 至 L 层为分组层。每一组代表了图像轮廓线的一组分类。正如上一节所介绍的,同层神经元之间存在强度各异的侧向连接。实际上,这样的连接强度将代表不同图像

特征的神经元归为一组的“相容度”,这正是“轮廓线分组”的关键所在。因此,同一层神经元之间的连接权值 f_{rs}^l 就成为了关系到网络性能的重要参数,从而如何对其进行确定以及调节也就是本文研究工作的重点。

设两个图像特征由 $\mathbf{m}_r = (\mathbf{p}_1, \mathbf{n}_1)$ 和 $\mathbf{m}_s = (\mathbf{p}_2, \mathbf{n}_2)$ 表示,其中 $\mathbf{p}_1 = (p_1^x, p_1^y)$ 和 $\mathbf{p}_2 = (p_2^x, p_2^y)$ 分别代表两个特征在二维图像上的坐标, $\mathbf{n}_1 = (n_1^x, n_1^y)$ 和 $\mathbf{n}_2 = (n_2^x, n_2^y)$ 为 2.1 节中所得到的特征向量。再令 $\mathbf{d} = \mathbf{p}_1 - \mathbf{p}_2, d = \|\mathbf{p}\|, \hat{\mathbf{d}} = \frac{\mathbf{d}}{d}$ 。

则代表着两个图像特征的同层神经元之间的连接权值可计算如下:

$$f_{rs}^l = \delta(k_1 k_2 k_3) \exp\left(-\frac{d^2}{R^2} - C^2 S\right) - I \exp\left(-\frac{2d^2}{R^2}\right) \quad (5)$$

其中 $k_1 = n_1^x \hat{d}^y - n_1^y \hat{d}^x, k_2 = n_2^x \hat{d}^y - n_2^y \hat{d}^x, k_3 = \mathbf{n}_1 \cdot \mathbf{n}_2, C = |\mathbf{n}_1 \cdot \hat{\mathbf{d}}| - |\mathbf{n}_2 \cdot \hat{\mathbf{d}}|, \delta(\cdot)$ 为 Kronecker Delta 函数。可调参数 R 控制同一层中具有较远距离的神经元之间的相容度, S 控制两个特征之间基于共圆特性的相容度, I 控制同层近距离神经元之间的抑制程度。经实验测试, $R = 0.2, S = 200, I = 0.1$ 是一个较为有效的参数组合。

图 3 描绘了单位正方形内一条线段(位于图中央部分的粗线段)与周围位置(排列为方格分布)的神经元之间的侧向连接强度。若通过式(5)计算,任意一条线段所代表的神经元与中央线段所代表神经元的侧向连接强度大于 0,即相互激励,则画出该线段。因此,图 3 中所描绘的线段集合也就是与中央线段相互激励的线段集合。可以看出,这样的激励模式和相关的 Gestalt 原理,即相似性、相近性、包含性、连续性和对称性是非常吻合的。表明式(5)的设计是较为合理的,并且为轮廓线分组提供了坚实的基础。

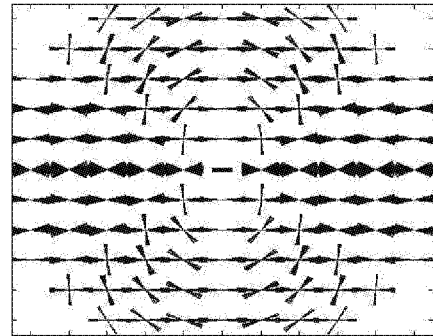


图 3 CLNN 侧向连接权值计算结果

对于第 2 至 L 层,采用上述法则计算神经元之间的侧向连接权值。而对于第 1 层,即背景层而言,其任意两个神经元之间的连接权值 $f_{rs}^1 = 0, f_r^1 = m$ 。对于 WTA 连接而言,其设定则相对简单。将所有的 WTA 连接设为相同值,即 $I_{rs}^m = J_r = J_s$ 。

2.3 CLNN 的动力系统迭代

CLNN 的原始学习算法基于式(1)所定义的确定性动力系统,其计算过程为串行方式,存在收敛慢、计算复杂度高的缺陷,这些缺陷将导致图像轮廓线分组算法性能降低,从而限制其实际应用。本文从所提出的基于 CLNN 的图像轮廓线分组实际要求出发,基于式(1),并结合随机优化的相关理论设计了一种新的基于模拟退火的 CLNN 并行随机学习算法,该算法具有收敛快、计算简单的特点。实验中发现,采用新算法所需的平均时间大约为采用原始算法所需时间的 $1/6$,且收敛精度更高。算法具体过程如下。

1) 随机初始化所有神经元:

$$x_{rl}(t=0) \in \left[\frac{h_r}{L} - \varepsilon, \frac{h_r}{L} + \varepsilon \right]; 0 < \varepsilon \ll 1$$

2) 初始化退火温度 T 。

3) 将神经元随机编号。

4) For $i = 1$ to RL

① 提取编号为 i 的神经元 x_{rl} 。

$$\varphi = \frac{J_r h_r - \sum_{m \neq l} I_r^{lm} x_{rm} + \sum_{s \neq r} f_{rs}^l x_{sl}}{I_r^{ll} - f_{rr}^l + T}.$$

③ 更新神经元状态 $x_{rl} = \max(0, \varphi)$ 。

4) 降温: $T = (1 - \eta)T, 0 < \eta \ll 1$ 。

5) 若收敛条件尚未满足,跳至 2); 否则结束算法。

在算法中引入降温的目的在于适当降低算法的收敛速度,以免分组结果过早地陷入局部最优。初始温度的取值对于算法的结果影响不大,而关于 η 的取值将在实验环节进行分析。

另外,可以证明:我们所提出的算法与式(1)所定义的动力系统在数学上是等价的。事实上,在上一小节所引入的 $I_r^{lm} = J_r = J$ 设置下,式(3)定义的能量函数可以改写为如下形式:

$$E^* = \frac{J}{2} \sum_r \left(\sum_l x_{rl} - h_r \right)^2 - \frac{1}{2} \sum_r \sum_{rs} f_{rs}^l x_{rl} x_{sl} \quad (6)$$

经推导可以证明式(6)和式(3)所定义的能量函数仅相差一个常数项,这意味着 E 和 E^* 沿着 CLNN 动力系统轨线的渐进行为是拓扑等价的。因此只需证明以上所提出的算法能够以概率 1 收敛到式(6)的局部最小点即可,这可通过随机优化中的 Heat Bath Update 方法的相关理论加以证明(详见文献[9])。基于式(6)易知当网络位于平衡态的时候有:

$$\frac{dx_{rl}}{dt} = J(h_r - \sum_{\alpha} x_{r\alpha}) + \sum_s f_{rs}^l x_{sl} \quad (7)$$

将式(7)中的求和项展开、拆分并移项整理,可得:

$$x_{rl} = \frac{J(h_r - \sum_{\alpha \neq l} x_{r\alpha}) + \sum_{s \neq r} f_{rs}^l x_{sl}}{J - f_{rr}^l} \quad (8)$$

在 $I_r^{lm} = J_r = J$ 的条件下,式(8)便等价于上述算法中的 CLNN 动力系统迭代公式。

3 仿真实验

通过仿真实验分别从定性和定量的角度对所提出的方法进行对比验证。如图4(a)所示,采用一个真实的图像数据

(命名为 Fruits),分别使用 CLNN 方法和 Potts spin 方法对其进行轮廓线提取以及分组。图4(b)描绘了经预处理后神经元的大致分布图。所采用的矩阵子图窗口大小为 5×5 。图4(c)、(d)展示了采用 CLNN 和 Potts spin 方法的结果。CLNN 中计算侧向抑制连接权值的参数设置为 $R = 0.2, S = 200, I = 0.1$ 。图中使用不同的符号代表分组结果,例如正方形代表香蕉。可以看出,Potts spin 方法基本上提取出了图像的大致轮廓,但是其运算结果存在以下主要问题:1) 分组结果不好,例如香蕉的左半部和右半部轮廓线分别被划分到了不同的组中,这显然是错误的;2) 噪声过多,出现了一些本不应该被提取出的轮廓线,且这些多余的轮廓线在运算中对正常轮廓线的分组计算产生了干扰。例如前面提到的香蕉轮廓线的错误,就是因为香蕉中部出现了噪声轮廓线的缘故。CLNN 方法的轮廓线提取能力和 Potts spin 方法相当,但是其分组质量更好。由于本文采用了基于 Gestalt 原理的侧向连接权值设定规则,使得网络具备了类似于生物视觉神经元的信号处理机制,对于不同的物体 CLNN 方法可以正确地分辨轮廓线的归属。另外,在 CLNN 方法的实验结果中不存在噪声数据,这得益于 CLNN 网络中所采用的 WTA 连接机制,每一列神经元中最多只有一个能被最终激活,这样就使得出现噪声数据的可能性大大降低。此外,背景层的引入也使得一部分噪声数据被归入背景层,这进一步过滤了噪声数据的影响。

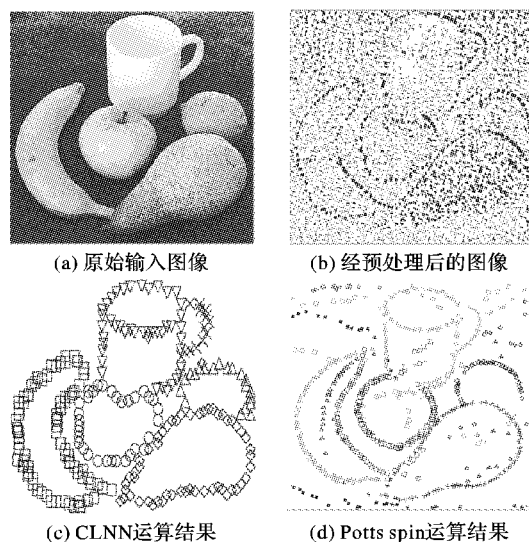


图4 基于真实图像数据 Fruits 的对比实验结果

表 1 退火速度对分组性能的影响

输入图像		性能指标	η				
			0	0.000 1	0.001	0.01	0.1
Fruits	原始图像	E 的均值	503.6	396.4	150.7	1 081.0	2 240.5
		E 的标准差	144.2	212.8	89.9	280.2	287.4
	Speckle 噪声图像	E 的均值	621.5	602.4	271.3	789.4	1 350.8
		E 的标准差	200.4	155.2	112.4	338.7	397.9
	高斯噪声图像	E 的均值	760.0	667.4	282.6	1 220.0	2 400.8
		E 的标准差	251.7	180.9	120.1	306.5	402.3
lena	原始图像	E 的均值	4 045.7	1 700.4	503.6	1 922.8	6 050.3
		E 的标准差	902.5	753.4	199.3	435.7	2 111.0
	Speckle 噪声图像	E 的均值	5 023.6	1 932.5	556.1	2 256.8	8 087.4
		E 的标准差	1 134.7	254.7	98.2	430.9	2 097.6
	高斯噪声图像	E 的均值	4 303.7	1 880.6	702.5	2 091.4	9 932.1
		E 的标准差	740.2	409.8	154.9	508.4	3 088.3

其中:

$$T_{\min} = \min\{T[R[j].L] \mid R[j] \in M\} \quad (11)$$

2) 同时需要对标记连通等价表 T 进行操作:

$$\text{Assignment}(T, T[R[j].L], R[i].L) \quad (12)$$

3) 扫描完成后,对连通标记连接表应用函数 $\text{flatten}()$ 进行扁平化操作。这一操作使得每个连通区域的全部标记都等价于同一标记。

4) 最后,根据游程标记信息及 T 数组信息将游程重新转换为图像,并让像素值 m 等于 $T[m]$ 。

flatten 函数表示为:

```
void flatten(vector T)
{ for(int i = 0; i < size(T); i++)
  { T[i] = T[T[i]]; } }
```

2.5 优化算法在实验中的应用和分析

为了验证改进后的标记算法在运行速度和所需存储空间上的改善,在 DM642 硬件平台上分别运行了优化算法, Suzuki 算法和 Kesheng Wu 算法来进行二值钢珠图像对比实验,其中 DSP 的主频是 600 MHz,结果如表 1 所示。

表 1 3 种算法运行结果比较

算法	对象数量/像素	所需存储空间/kB	运行时间/s
Suzuki 算法	413280	807	0.2164
KeshengWu 算法	413280	807	0.0535
本文优化算法	2460 游程	15	0.0027

表 1 中运行时间仅考虑标记算法的时间,不考虑数据传输时间。由于基于像素的算法中临时标记等于像素值,为满

足标记总数需要,在实验中像素所占空间为 2 B;而游程所占空间是 6 B。从实验结果看,优化算法所需存储空间仅为另两种算法的 2% 左右,并且运行速度分别提高了近 80 倍和近 20 倍。

3 结语

本文针对液滴实时图像反馈控制系统的环境,提出了基于游程编码的连通区域标记优化算法来实现边缘提取,在运行速度和对内存容量的需求上都有较大的改善,可满足空间液滴蒸发实验对图像实时处理的需要。

参考文献:

- [1] 于强, 宁乔. 蒸发液滴空间实验研究的图像反馈控制系统 [J]. 空间科学学报, 2008 28(1): 64-68.
- [2] SUZUKI K, HORIBA I, SUGIE N. Linear-time connected-component labeling based on sequential local operations [J]. Computer Vision and Image Understanding, 2003, 89(1): 1-23.
- [3] WU KE-SHENG, OTTO E, SHOSHANI A. Optimizing connected component labeling algorithms [C]// Proceedings of SPIE 5747. Bellingham, WA: Society of Photo-Optical Instrumentation Engineers, 2005: 1965-1976.
- [4] FIORIO C, GUSTEDT J. Two linear time union-find strategies for image processing [J]. Theoretical Computer Science, 1996, 154(2): 165-181.
- [5] SHIMA Y, MURAKAMI T, KOGA M, et al. A high-speed algorithm for propagation-type labeling based on block sorting of runs in binary images [C]// Proceedings of 10th International Conference on Pattern Recognition. Washington: IEEE, 1990: 655-658.

(上接第 3149 页)

在本问题中引入了基于退火的并行随机优化方法对网络进行训练,其中退火速度 η (确切地说是 η 的数量级) 是影响网络性能的重要参数。注意到在式(3)中引入了能量函数 E , 用以讨论 CLNN 动力系统的有界性。而网络的平衡点恰好是能量函数的全局最小点,因此可以使用 E 作为评价 CLNN 分组性能的指标。分别使用 Fruits 和 Lena 图作为输入图像。对于每幅图像,分别加入均值为 0、方差为 0.04 的随机 speckle 噪声和均值为 0、方差为 0.01 的高斯噪声,连同原始图像共组成 6 幅输入图像。在 $\eta = 0, 0.0001, 0.001, 0.1$ 四种条件下分别进行 50 次试验,求得网络收敛时能量的平均值和标准差。实验结果如表 1 所示。可以看出,对于全部六幅图像,在 $\eta = 0.001$ 时 CLNN 均取得了最好的性能(即最小的能量值),因此认为 η 的数量级取为 10^{-3} 是一个较为理想的设置。

4 结语

本文提出了一种基于竞争层神经网络的新型图像轮廓线分组算法。该算法借鉴了生物神经元的视觉信号处理机制,采用 Gestalt 原理对网络权值进行计算。实验证明,所提出的算法具有分组能力强、抗噪声能力突出的优点。在下一步的工作中,将继续对 CLNN 进行理论分析,在更深的层面探讨网络参数对网络性能的影响。另外,进一步拓宽 CLNN 的应用范围也是非常具有意义的。

参考文献:

- [1] RITTER H. A spatial approach to feature linking [C]// Proceedings of International Conference on Neural Networks. Paris: [s. n.],

- 1990.
- [2] ZHANG YI, TAN K K. Multistability of discrete-time recurrent neural networks with unsaturating piecewise linear activation functions [J]. IEEE Transactions on Neural Networks, 2004, 15(2): 329-336.
- [3] ZHANG YI, TAN K K, LEE T H. Multistability analysis for recurrent neural networks with unsaturating piecewise linear transfer functions [J]. Neural Computation, 2003, 15(3): 639-662.
- [4] WENG S, WERSING H, STEIL J J, et al. Learning lateral interactions for feature binding and sensory segmentation from prototypic basis interactions [J]. IEEE Transactions on Neural Networks, 2006, 17(4): 843-862.
- [5] ONTRUP J, WERSING H, RITTER H. A computational feature binding model of human texture perception [J]. Cognitive Processing, 2004, 5(1): 31-44.
- [6] WERSING H, KIRSTEIN S, SCHNEIDERS B, et al. Online learning for bootstrapping of object recognition and localization in a biologically motivated architecture [C]// ICVS 2008: Proceedings of International Conference on Computer Vision Systems. Santorini, Greece: [s. n.], 2008: 383-392.
- [7] MCCAFFERTY J D. Human and Machine Vision [M]. [S. l.]: Ellis Horwood, 1990.
- [8] WU F Y. The Potts model [J]. Review of Modern Physics, 1983, 54: 235.
- [9] KOSEKI O, MATSUBARA F. Cluster heat bath method on a Quasi-One-Dimensional Ising Model [J]. Journal of the Physical Society of Japan, 1997, 66(2): 322-325.