

一个具有拟神经组织的边缘检测模型¹⁾

潘士先 夏宇闻 雷志平

(北京航空航天大学电子工程系, 北京 100083)

摘要 基于视皮层简单细胞的 Gabor 函数模型, 仿照视皮层的朝向柱结构, 并辅以模式增强和处理网络, 构造了一个具有拟神经组织的边缘检测模型. 模拟结果显示, 其检测能力与平滑-空间导数型边缘检测算法相当. 分析表明, 这个基于早期边缘检测器概念的模型与平滑-空间导数型算法在检测机理上是一致的. 模型为用光学-电子方法和人工神经网络实现边缘检测提供了一个原型.

关键词 视皮层简单细胞, Gabor 函数, 边缘检测, 人工神经网络.

AN EDGE DETECTION MODEL WITH NEURON-LIKE ORGANIZATION

PAN Shixian XIA Yuwen LEI Zhiping

(Dept. of Electronic Engineering, Beijing University of Aero. & Astro., Beijing 100083)

Abstract Based on the Gabor function model of simple cortical cells, by imitating the structure of orientation columns in visual cortex, and aided by pattern enhancement and processing nets, a model for edge detection with neuron-like organization is constructed. Simulation results show that its detection capability is comparable to that of the advanced edge detection algorithms of the smoothing-spatial differentiation type in vogue. It is shown that this model as intuitively based on the early concept of edge detectors is essentially identical in the underlying mechanism to the latter. The proposed model provides a prototype for edge detection using electronic-optical techniques and artificial neural nets.

Key words Simple cortical cells, Gabor functions, edge detection, artificial neuron nets.

1 引言

视皮层简单细胞——感受野由若干个兴奋和抑制交错排列的条状子域组成的线性神

1) 国家自然科学基金资助项目(69475003)

收稿日期 1997-08-25 收修改稿日期 1999-03-31

神经元——的基本性质是对于空间频率和朝向的选择性, 并被认为具有特征检测的功能; 特别是, 其中感受野基本上由两个显著子域构成的简单细胞, 下称二极(bipartite)细胞, 具有边缘检测的功能^[1]. 然而 Marr 等人^[2]对此表示怀疑, 认为这样的特征检测器可对“太多的刺激产生响应”: 缺乏分辨率. 在后来边缘检测理论的发展中, 这种由简单细胞而来的特征检测器的想法未受重视. 然而, 近来图像的二维 Gabor 函数表示^[3]和小波分析^[4]提示了特征检测器的思想与目前在边缘检测中占主流地位的平滑-空间导数型算法, 如 Marr 等人^[2]的 $\nabla^2 G$, Canny^[5]的 $\nabla \cdot G$ 和 Mallat 等人^[6]的小波算法, 在机理上是相通的. 基于二极细胞的 Gabor 函数模型^[7]并仿照视皮层的朝向柱结构, 我们构造了一个具有拟神经组织的边缘检测模型(NMED). 模拟试验表明, 其检测能力可与先进的平滑-空间导数法相比较. 此外还论证了此模型在检测机理上与平滑-空间导数型算法的一致性.

2 模型二极细胞

Daugman^[7]建立了二极细胞的二维的 Gabor 函数模型. 二极细胞的感受野可用下列 Gabor 函数表示:

$$\psi(\mathbf{x}) = \exp\{-\pi(x^2a^2 + y^2b^2)\} \sin 2\pi u_0 x, \quad (1)$$

$$\hat{\psi}(f) = \frac{1}{2j} (\exp\{-\pi[(\frac{u-u_0}{a})^2 + (\frac{v}{b})^2]\} - \exp\{-\pi[(\frac{u+u_0}{a})^2 + (\frac{v}{b})^2]\}), \quad (2)$$

其中 $\mathbf{x} = (x, y)$, $f = (u, v)$. $\psi(\mathbf{x})$ 是高斯包络 $g(\mathbf{x}) = \exp\{-\pi(x^2a^2 + y^2b^2)\}$ 的正弦调制栅格, 其朝向与 x 轴相重合, 空间频率为 $f_0 = u_0$; $g(\mathbf{x})$ 的主轴与 x, y 轴相重合, 沿 x, y 轴的方差各为 $\sigma_x = 1/(a\sqrt{2\pi})$, $\sigma_y = 1/(b\sqrt{2\pi})$. $\hat{\psi}(f)$ 为 $\psi(\mathbf{x})$ 的傅立叶变换, 是空间频率域上以 $\pm f_0$ 为中心的带通函数, 具有椭圆状通带. 若以半峰值水平界定其通带, 它在 u, v 方向的带宽各为 $\Delta u = 2a\sqrt{\ln 2/\pi}$, $\Delta v = 2b\sqrt{\ln 2/\pi}$. 也可用“带宽” $\Delta f = \Delta u$ 和“朝向带宽”

$$\Delta\theta_{1/2} = \tan^{-1} \frac{\Delta v/2}{f_0} = \frac{b\sqrt{\ln 2}}{f_0\sqrt{\pi}} \quad (3)$$

标志此椭圆通带^[7]. 令模型二极细胞(MBC) i 的净输入为 net_i , 输出为 O_i . 其特性可表示为

$$O_i = \text{net}_i \quad \text{net}_i > T; \quad O_i = 0, \text{net}_i \leq T. \quad (4)$$

其中 T 为阈值, 以下设 $T=0$, i (的感受野中心) 位于 \mathbf{x} 处, 则

$$\text{net}_i = \int \psi(\mathbf{r} - \mathbf{x}) I(\mathbf{r}) d\mathbf{r}, \quad (5)$$

其中 $I(\mathbf{x})$ 为图像函数. 由式(1, 4, 5)表达的 MBC 的朝向(正弦栅格的方向) $\theta=0$. 可以通过旋转产生任意朝向的 MBC. 令 $\mathbf{x}' = (x', y')$, 则

$$x' = x \cos \theta - y \sin \theta, y' = x \sin \theta + y \cos \theta, \quad (6)$$

$$\psi_\theta(\mathbf{x}) = \psi(\mathbf{x}'). \quad (7)$$

其中 $\psi_\theta(\mathbf{x})$ 具有朝向 θ .

以上就是 MBC 原型. 它最显著的特性是对台阶边缘的朝向选择性. 设有理想台阶边缘穿越其感受野中心, 对感受野有效尺度而言边缘可认为是直线. 则其输出在 $\text{net}_i > 0$ 的范围内随其朝向与边缘朝向之差 $\delta\theta$ 的增大而减小, $\delta\theta=0$ 时为最大. 图1中曲线1 示一 MBC ($\Delta\theta_{1/2}=15^\circ, a/b=2$) 的归一化朝向响应, 此即表示 MBC 作为一种边缘检测器基本

特性的“朝向调谐曲线”。

3 结构和功能

检测模型为如下四层结构：

第一层. 模型二级细胞层. 设边缘朝向量化为 $2n$ 个等角间隔: $\theta_k = k\Delta\theta, k = 0, 1, \dots, 2n-1; \Delta\theta = \pi/n$. 记输入图像为 $\{I(i, j): 0 \leq i, j \leq N-1\}$. 第一层由 $2n$ 个子层组成, 其第 k 个子层由 $N \times N$ 个朝向为 θ_k 的 MBC 组成——每一元对应于一个输入节点¹⁾. 记同一位置 $p = (i, j)$ 上 $2n$ 个子层元为 $S_{1p}(k), k = 0, 1, \dots, 2n-1$; 第一层可表示为 $\{S_{1p}(k): p = (i, j), 0 \leq i, j \leq N-1; k = 0, 1, \dots, 2n-1\}$. 对于一般矩形采样数字图像, 模型的网络结构对 $n=4$ 最易实现和描述且不失一般性, 故以下讨论中均设 $n=4$. 令 $S_{1p}(k)$ 的输出和净输入各为 $O_{1p,k}, \text{net}_{1p,k}$, 有

$$\text{net}_{1p,k} = \sum_l w_{kl} I_l; \quad O_{1p,k} = \text{net}_{1p,k}, \quad \text{net}_{1p,k} > 0; \quad O_{1p,k} = 0, \quad \text{net}_{1p,k} \leq 0. \quad (8)$$

式中 w_{kl} 表示它对图像像素 I_l 的灵敏度, l 遍及其感受野 $\{w_{kl}\}$, 而 $\{w_{kl}\}$ 是 $\psi_{k\Delta\theta}(\mathbf{x})$ 的离散近似. 选择 ψ 的参数使得 n 个 $\psi_{k\Delta\theta}$ 的通带在空间频率域上构成适当的覆盖; $n=4$ 时可取 $\Delta\theta_{1/2} = 22.5^\circ, a/b = 1$, 此时通带为圆形. 注意 $\psi_{(k+n)\Delta\theta}(\mathbf{x}) = -\psi_{k\Delta\theta}(\mathbf{x}), k = 0, 1, \dots, n-1$. 这样, 每一子层构成类似于视皮层等朝向柱的等朝向域, 其所有诸元敏感于同一朝向.

第二层. 朝向响应锐化层. 结构同第一层, 记为 $\{S_{2p}(k): p = (i, j), 0 \leq i, j \leq N-1; k = 0, 1, \dots, 2n-1\}$. 从第一层到第二层的连接限于它们的对应组 $S_{1p}(k), S_{2p}(k)$ 之间. 记 $S_{2p}(k)$ 的净输入为 $\text{net}_{2p,k}$, 输出为 $O_{2p,k}$, 有

$$\text{net}_{2p,k} = \sum_l w_{kl}^1 O_{1p,l}; \quad O_{2p,k} = \text{net}_{2p,k}, \quad \text{net}_{2p,k} > 0; \quad O_{2p,k} = 0, \quad \text{net}_{2p,k} \leq 0. \quad (9)$$

连接权 $\{w_{kl}^1: l = 0, 1, \dots, 2n-1\}$ 形成如下以 k 为中心的中心开周边闭感受野

$$w_{kl}^1 = -G''(\theta_k - \theta_l), \quad G''(x) = -\frac{1}{\sigma^3 \sqrt{2\pi}} \left(1 - \frac{x^2}{\sigma^2}\right) e^{-x^2/2\sigma^2}, \quad (10)$$

其中 G'' 是高斯函数 G 的二阶导数, $-G''$ 是 $\nabla^2 G$ 滤波核^[2]的一维形式. 适当选择 G 的方差 σ 时, 这种连接模式形成一种锐化第一层朝向响应的非递归侧抑制. 为了说明这一点, 先考虑 $-G''$ 与图1中调谐函数(曲线1)的卷积. 这相当于用 G 将后者平滑后再取二阶导数的负值. 由此得图1中的曲线2 ($\sigma = 22.5^\circ$, 对峰值归一化, 正、负值部分分别用实、虚线画出). (注意其中央锐峰). 事实上, 依照文献[5]关于 $\nabla^2 G$ 是屋顶边缘最佳检测器的论述, 因调谐曲线呈近似屋顶状, 用 $-G''$ 可有效锐化其峰值. 将曲线2 的负值部分截去, 便得到中央锐峰. 可见调谐函数经 $-G''$ 滤波和截去负部的非线性操作被锐化为一尖峰. 当有边缘过 p 点时, $\{S_{1p}(k)\}$ 的 $2n$ 个元的响应(构成的)曲线显然与调谐曲线一致, 不过离散化为 $2n$ 个点. 因此, 用一个相应地离散化的 $-G''$ 滤波核作侧抑制权并通过 $S_{2p}(k)$ 的阈值作用, 就可达到锐化朝向响应的目的. 注意, 虽然 σ 越小锐化越强, 但其值不可太小, 否则 $-G''$ 不能用 $2n$ 个点有效地离散化.

1) 略去图像的边缘效应, 考虑边缘效应时可将图像镜像延拓.

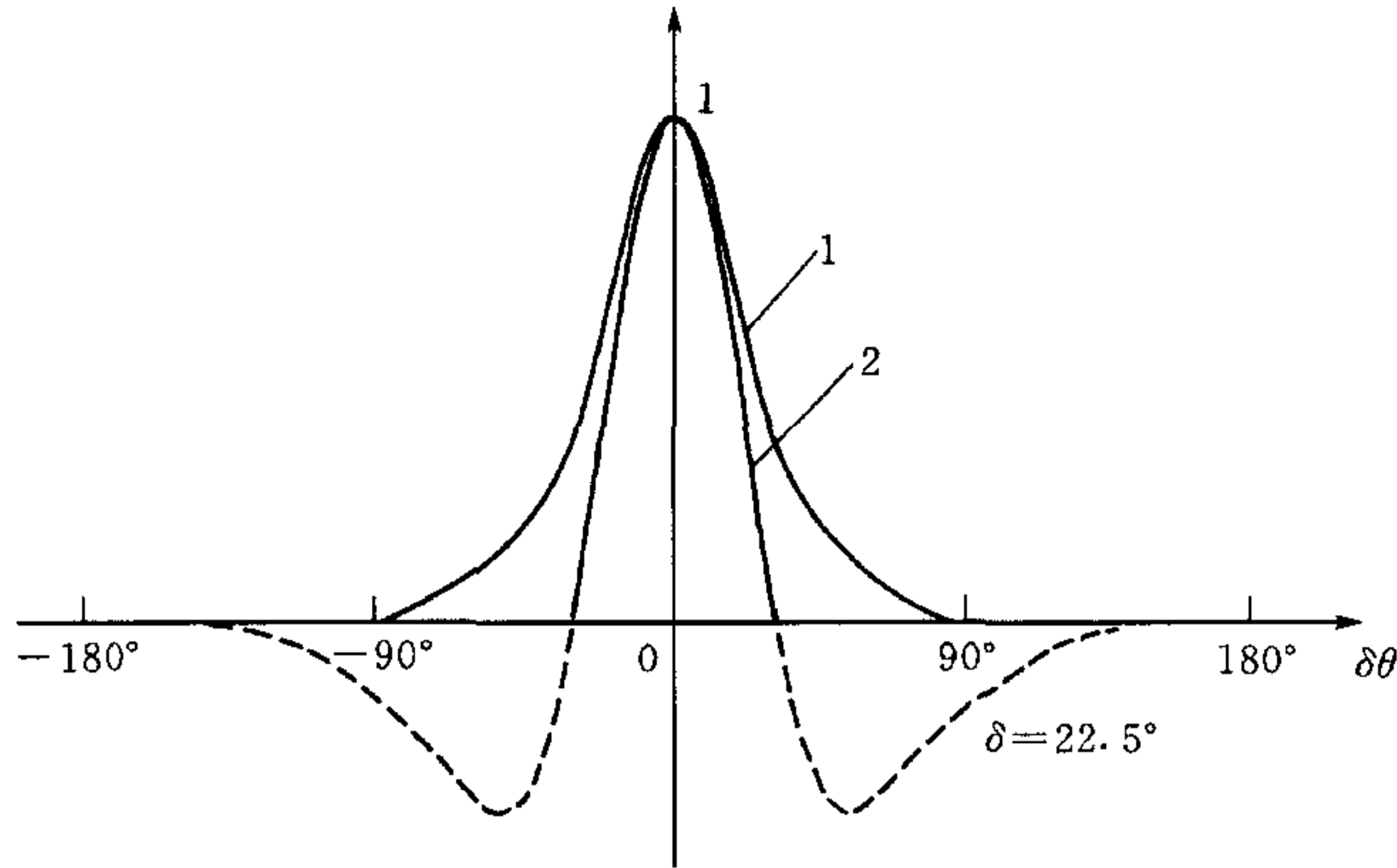


图1 曲线1为 MBC 的调谐曲线,曲线2是它与滤波核 -G'' 的卷积

第三层. 边缘提取层. 其作用是通过朝向间的侧抑制以及体现边缘连续性约束的连接以获得朝向无模糊且在空间分布上可信的边缘. 众所周知, 边缘检测是一个病态问题, 只有在适当的约束下方可求得稳定的解. 连续性约束被认为是边缘的一个一般物理约束, 它可表述为: 除了交点、顶点、端点等处外, 边缘一般是连续的. 使用这一约束有助于减少错检和漏检. 第三层神经元和结构与第二层同, 记为 $\{S_{3p}(k): p=(i, j), 0 \leq i, j \leq N-1; k=0, 1, \dots, 2n-1\}$. 记任意元 $S_{3p}(k)$ 的输出和净输入为 $O_{3p,k}, net_{3p,k}$, 有

$$O_{3p,k} = \alpha net_{3p,k}, \quad net_{3p,k} > T_1; \quad O_{3p,k} = 0, \quad net_{3p,k} \leq T_1, \tag{11}$$

其中 T_1 为阈值, α 为增益; 而

$$net_{3p,k} = - \sum_{l \neq k} O_{2p,l} + \alpha O_{2p,k} + \sum_{q \in \Omega_k} b_q O_{2q,k}. \tag{12}$$

Ω_k 是 p 在垂直于 θ_k 方向上的一个邻域, 其大小约等于 MBC 感受野的有效尺度, a, b_q 为权. 式(12)中第一, 二项表示 $S_{3p}(k)$ 除受 $S_{2p}(k)$ 的输出激励外还受 $\{S_{2p}(l)\}$ 中所有其余 $2n-1$ 个元输出的抑制. 第三项是体现上述连续性约束的邻域支持. 为了有效体现连续性约束, 应适当选择权 a, b_q , 使得只有在得到足够大的邻域支持时 $S_{3p}(k)$ 方能充分激活.

第二, 三层所形成的强朝向选择性以及适当的阈值 T_1 可产生如下效应: 在绝大多数节点上, 第三层 $2n$ 个元或者均有零输出——对应于图像的均匀区域, 或者仅一个有非零输出——对应与边缘且其值有赖于边缘的对比度.

第四层. 边缘细化层. 如上所述, 从第三层的输出已可得到边缘图(除了在顶点, 交点等特殊灰度结构处可产生错检、漏检或朝向模糊). 但这些边缘一

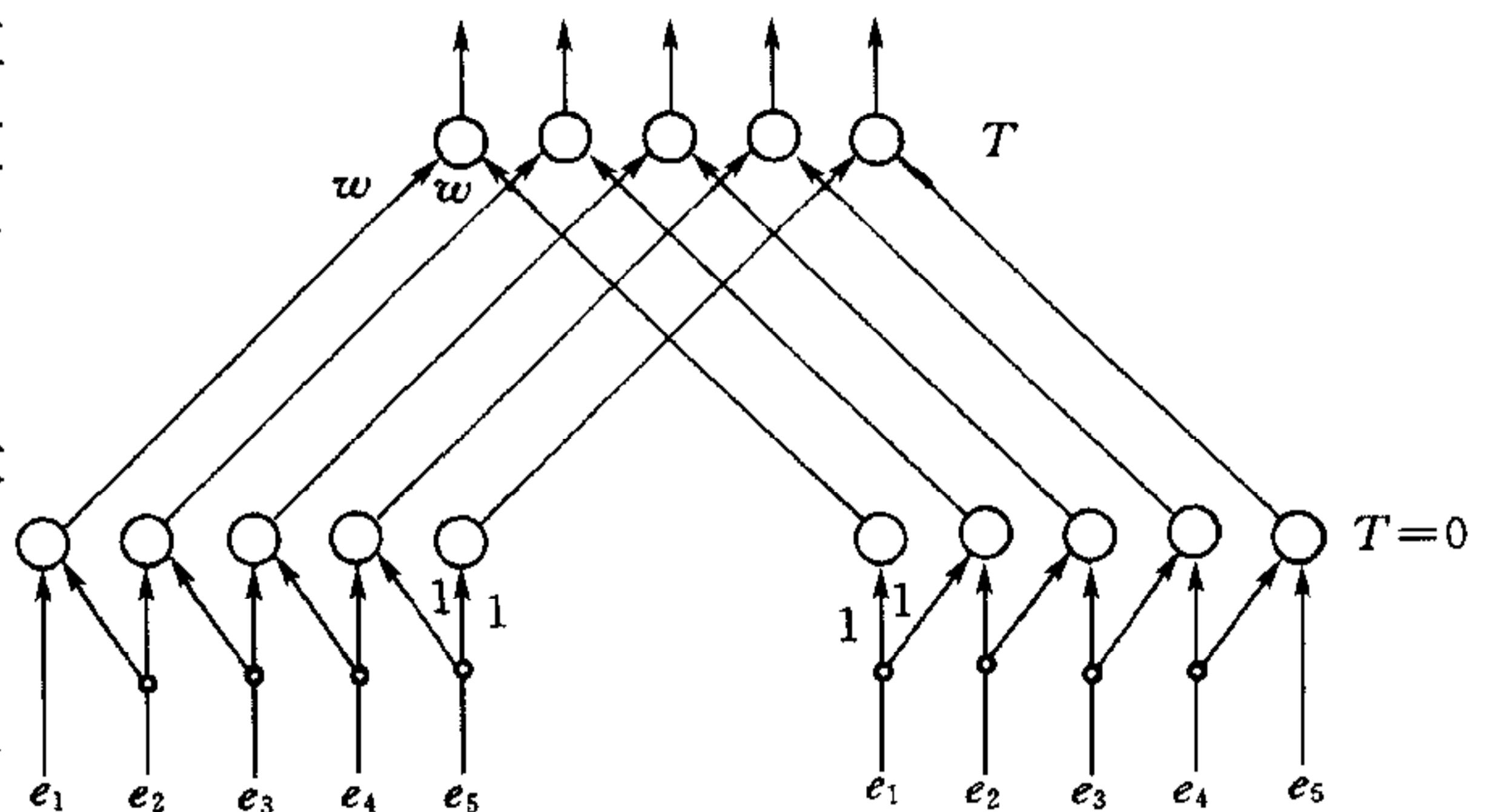


图2 两层结构的单边侧抑制网

般不够“细”(见下述模拟试验), 即同一子层内沿其朝向方向的若干个相邻元可同时有非零且在朝向上唯一的响应. 粗边缘意味着边缘位置模糊, 应予细化. 细化就是从上述相邻的非零输出中选出最大者并将其余的抑制掉. 这里采用了一种简单的单边侧抑制网络, 取得了较为理想的效果.

单边侧抑制网是两层结构,如图2所示.第一层分为两组:左(侧抑制)网和右(侧抑制)网,所有元均有阈值特性且阈值为零;左右网接受相同的输入.第二层为综合层,输入为左、右网输出之和乘以权 w ,所有元为阈值为 T 的阈值元件.显然,这种单边侧抑制通过相邻输入的比较,只要 $2w > T$,总能检出输入中的局域最大值.

第三层每一朝向子层 $\{S_{3p}(k): p=(i,j), 0 \leq i, j \leq N-1\}$ 的输出沿 θ_k 方向使用单边侧抑制网就构成细化层.它们可表示为神经元集合 $\{S_{4p}^l(k), S_{4p}^r(k), S_{4p}(k): p=(i,j), 0 \leq i, j \leq N-1; k=0, 1, \dots, 2n-1\}$, $S_{4p}^l(k), S_{4p}^r(k), S_{4p}(k)$ 各为左,右和综合层的元,且有

$$\begin{aligned} \text{net}_{4p,k}^l &= O_{3p,k} - O_{3s,k}; & O_{4p,k}^l &= 1, \text{net}_{4p,k}^l > 0; & O_{4p,k}^l &= 0, \text{net}_{4p,k}^l \leq 0. \\ \text{net}_{4p,k}^r &= O_{3p,k} - O_{3t,k}; & O_{4p,k}^r &= 1, \text{net}_{4p,k}^r > 0; & O_{4p,k}^r &= 0, \text{net}_{4p,k}^r \leq 0. \\ \text{net}_{4p,k} &= w(O_{4p,k}^l + O_{4p,k}^r); & O_{4p,k} &= 1, \text{net}_{4p,k} > T; & O_{4p,k} &= 0, \text{net}_{4p,k} \leq T. \end{aligned} \quad (13)$$

式中神经元净输入和输出的符号如前, s, t 各为 p 在 θ_k 方向的左、右邻.由上述可知,在绝大多数节点上细化层 $2n$ 个元或者均有零输出,或者仅有一个为1 其余为零.后一种情形标志了边缘.令 $e(p)$ 为边缘图, $\text{ort}[e(p)]$ 为边缘朝向,取 $e(p)=1, \text{ort}[e(p)]=\theta_k$, 若有 $O_{4p,k}=1, O_{4p,l}=0, l \neq k$; 否则 $e(p)=0$.

4 机理

现在来说明 NMED 在检测机理上与平滑-空间导数型算法的一致性.令 $\varphi(\mathbf{x}) = \varphi(-\mathbf{x})$. $\varphi(\mathbf{x})$ 满足 $\int_{-\infty}^{\infty} \varphi(\mathbf{x}) d\mathbf{x} = 0$, 故是二维小波函数.令

$$\varphi_{s,\theta}(\mathbf{x}) = \frac{1}{s^2} \varphi(x', y'), \quad (14)$$

$$x' = \frac{1}{s}(x \cos \theta - y \sin \theta), y' = \frac{1}{s}(x \sin \theta + y \cos \theta), \quad (15)$$

s 为尺度参数.图像 I 对于 $\varphi_{s,\theta}$ 的小波变换为

$$w_{s,\theta} I(\mathbf{x}) = (I * \varphi_{s,\theta})(\mathbf{x}). \quad (16)$$

注意式(5)就是一定 s 和 θ 上的这样一个变换.令 $S_{s,\theta}(\mathbf{x})$ 为 $\varphi_{s,\theta}(\mathbf{x})$ 对于 x' 的积分

$$\varphi_{s,\theta}(\mathbf{x}) = \frac{\partial}{\partial x'} S_{s,\theta}(\mathbf{x}), \quad (17)$$

$S_{s,\theta}(\mathbf{x})$ 是一平滑函数:对称于 x', y' 且沿 x', y' 平滑地衰减并趋于零.由此有

$$w_{s,\theta} I(\mathbf{x}) = (I * \frac{\partial}{\partial x'} S_{s,\theta})(\mathbf{x}) = \frac{\partial}{\partial x'} (I \cdot S_{s,\theta})(\mathbf{x}). \quad (18)$$

这意味着 $w_{s,\theta} I$ 系 I 经 $S_{s,\theta}$ 平滑后沿 θ 的方向导数.依照平滑-空间导数型算法的概念,尺度 s 上的边缘定义为 $\max_{\theta} \left| \frac{\partial}{\partial x'} (I * S_{s,\theta}) \right|$ 的局域最大点,由式(18)也即 $\max_{\theta} |w_{s,\theta} I|$ 的局域最大点.将 θ 在 $(0, \pi)$ 的范围内量化为 n 个等角间隔且如前述适当选择 φ 的参数时,令 $\varphi_{s,k} = \varphi_{s,k\Delta\theta}, w_{s,k} I = w_{s,k\Delta\theta} I, k=0, 1, \dots, n-1$. 此时检出 $\max_k |w_{s,k} I|$ 的局域最大点便得尺度 s 上的边缘. NMED 以一种神经计算方式实现了上述算法.第一层完成对于 $\varphi_{s,\theta}$ 的小波变换.然而因为 MBC 不能有负输出, $w_{s,\theta} I (k=0, 1, \dots, n-1)$ 的负部被截去,不得不用 $w_{s,k+n} = I * \varphi_{k+n}$ (注意 $\varphi_{s,k+n} = -\varphi_{s,k}$) 的正部去代替它们,从而获得“整流”了的 $w_{s,k} I$. 其余

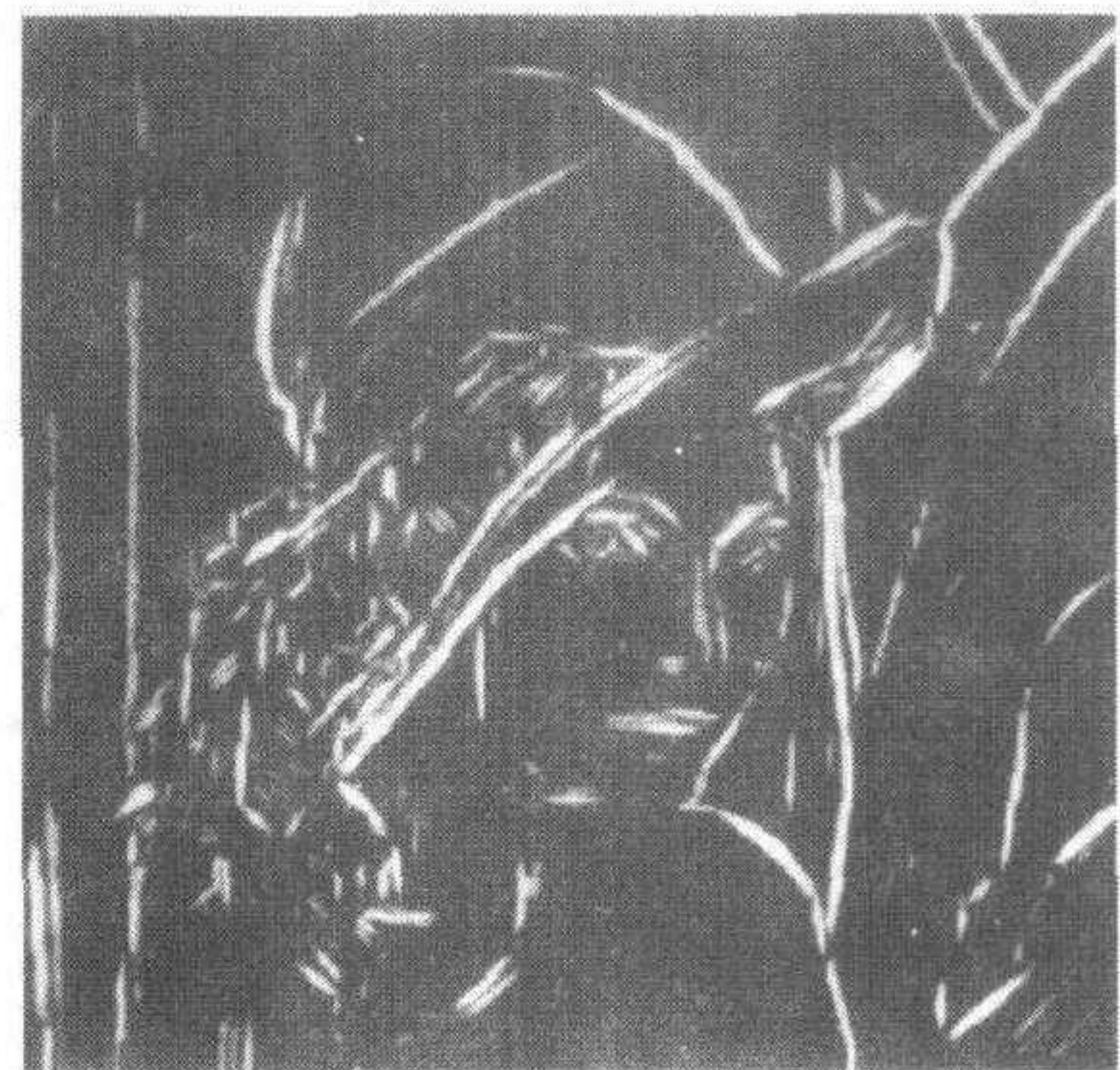
诸层实现上述局域最大值的提取,其中还融入了连续性约束.

5 模拟试验

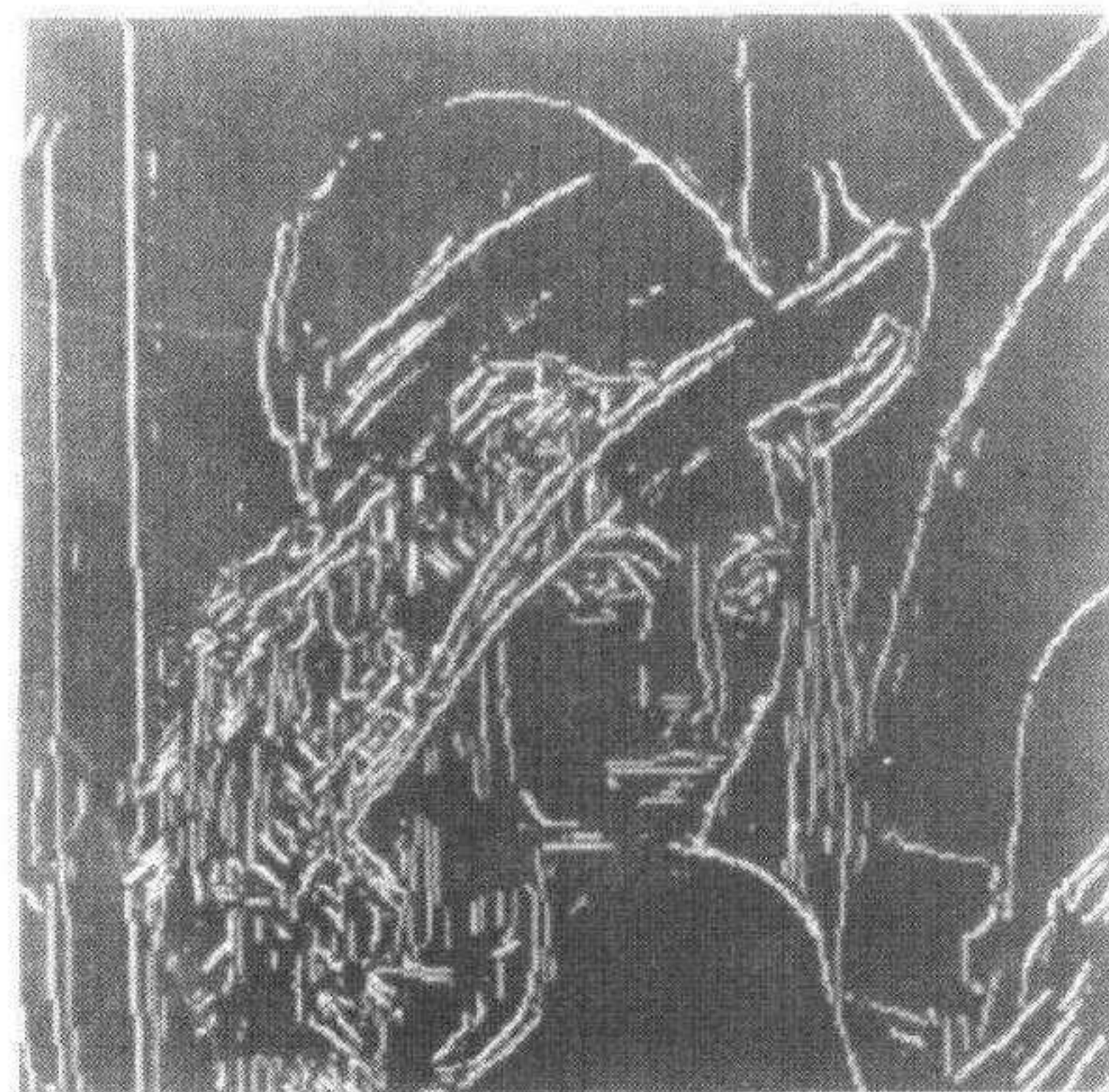
在 PC 机上作模拟试验. 所用图像为 $256 \times 256 \times 8\text{bits}$, $n=4$, $\Delta\theta=45^\circ$. 在高分辨率上, 感受野的有效半径只覆盖数个像素, ψ 不能有效离散化¹⁾, 故模拟中采用了一种简化感受野: 半径为 r 的圆形半开半闭感受野, w_{kl} 在开、闭区各为 $+1$ 和 -1 . $2r$ 略大于六个像素. 为了避免像素落在开闭分界线上, 感受野的中心取在图像采样网格的中央. $\sigma=22.5^\circ$, Ω_k 对于所有的 k 均包含六个相邻点, $a=1/2$, $b_q=1/12$, α 约为 1.0×10^{-3} , $T_1=30-40$ (有赖于图像对比度). 这里给出两个典型结果, 图3(a)是“Lena”图像; 图3(b)是从第三层输出的边缘图 ($e(p)=\max(O_{3p,0}, \dots, O_{3p,2n-1})$). 可以看到主要边缘均被检出, 但太粗. 图3(c)是细化层输出的边缘图. 这个结果可与先进的平滑-空间导数型算法相比较^[6]连续性甚至更好. 图4(a)、(b)是一计算机合成图像及所得边缘图. 从这个简单的图像可以清楚看到检测效果: 绝大部分边缘均被正确检出, 但也有些漏检 (有赖于 T_1 值的选择).



(a)Lena 图像



(b)第三层输出的边缘图



(c)细化层输出的边缘图

图3

1)除非将图像插值加密,那时可作亚像素边缘检测.

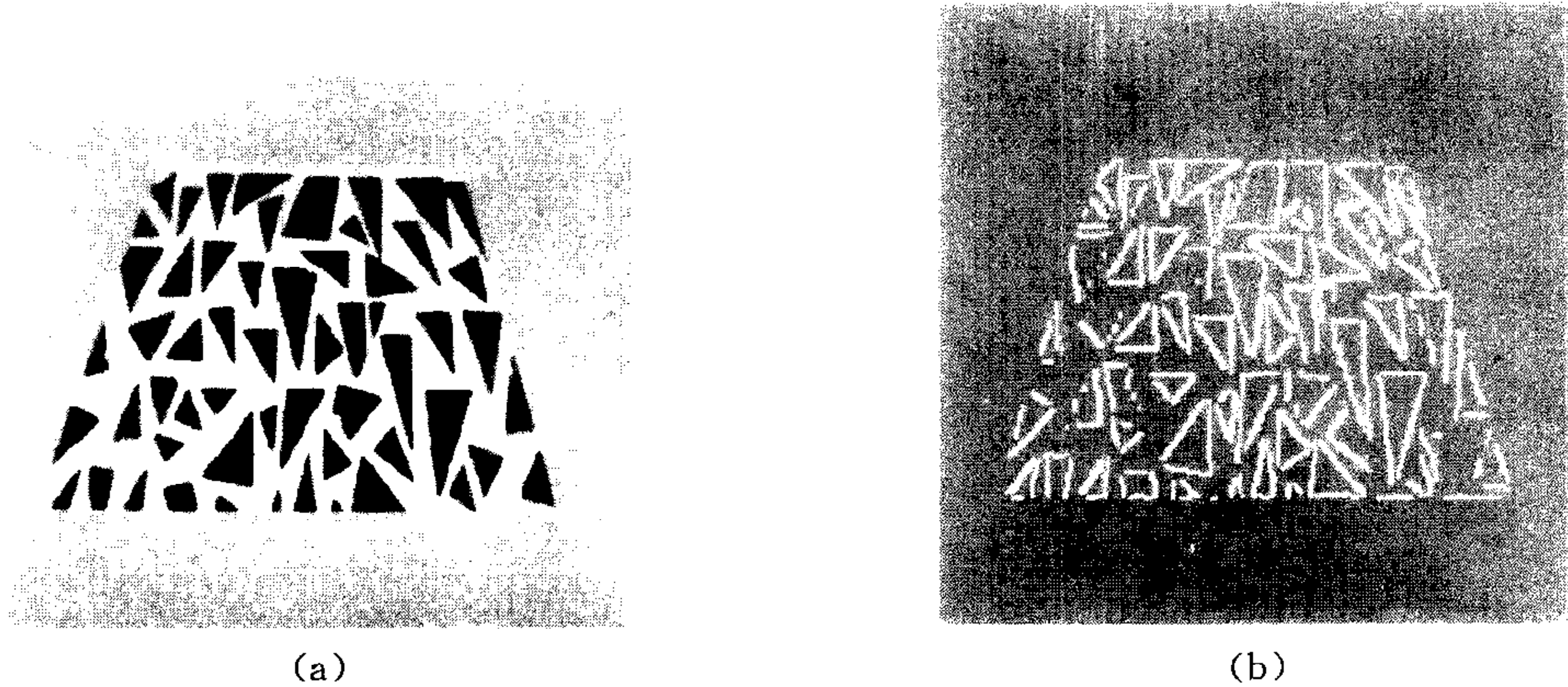


图4

以上介绍了单尺度的 NMED. 多尺度模型原则上可依照二分(dyadic)小波方式构造. NMED 可和图像分割、体视等网络^[8]相结合构成早期视觉人工网络处理系统.

参 考 文 献

- 1 Pollon D A *et al.* Visual cortical neurons as localized spatial frequency filters. *IEEE Trans. Syst. Man. Cybern.*, 1983, **13**(5): 907~916.
- 2 Marr D *et al.* Theory of edge detection. In: Proc. R. Soc. Lond. B 207, 1980: 187~217.
- 3 Porat M *et al.* The generalized Gabor scheme of image representation in biological and machine vision. *IEEE Trans. Pattern Anal. Mach. Intell.* 1988, **10**(4): 452~468.
- 4 Mallat S G. Multifrequency channel decomposition of images and wavelet models. *IEEE Trans. Acoustics, Speech, and Signal Processing*, 1989, **37**(12): 2091~2109.
- 5 Canny J. A computational approach to edge detection. *IEEE Trans. Pattern Anal. Machine Intell.*, 1986, **8**(6): 678~698.
- 6 Mallat S *et al.* Characterization of signals from multiscale edges. *IEEE Trans. Pattern Anal. Mach. Intell.*, 1992, **14**(7): 710~732.
- 7 Daugman J G. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. *J. Opt. Soc. Am. A/2*, 1985, **7**(2): 1160~1169.
- 8 李文彪等. 一个神经网络体视协同算法, 自动化学报, 1998, **24**(3): 323~330

潘士先 见自动化学报1998年24卷3期.

夏宇闻 1946年8月生于上海, 1968年毕业于北京清华大学自动控制系计算技术及装置专业, 1981年在北京航空航天大学获电子与通信专业工学硕士学位, 1986至1988年在澳大利亚新南威尔士大学通信系从事研究工作, 多年来在北航从事复杂数字系统的教学和研究工作.