

一种神经网络体视协同算法¹⁾

李文彪 宁 静 潘士先

(北京航空航天大学电子工程系 北京 100083)

摘 要 描述一种多分辨率体视协同算法. 该算法用二分正交小波作多分辨率分解, 采取边匹配边重建, 二者互相修正协同进行的策略, 并以神经网络主要是 Hopfield 网络为其实现形式. 试验结果表明, 该算法是合理和有效的.

关键词 体视, 特征匹配, 表面重建, 多分辨率分析, 神经网络.

1 引言

Marr 等在 70 年代末发表的体视计算理论^[1,2]迄今仍是最重要和最有影响的体视理论框架. 10 多年来, 这一理论所包含的诸算法模块有了许多改进和发展, 在边缘检测、(特征)匹配和(表面)重建方面都出现了许多先进和精巧的算法. 特别值得指出的是, Poggio 等的初级视觉的正则化理论将计算视觉提高到一个新水平^[3,4]; 人工神经网络的发展为正则化理论的实现提供了新途径; 最近兴起的小波分析为多分辨率分析提供了严密的数学基础^[5]. 然而, 迄今仍未见到完善的多分辨率体视协同算法. 在上述进展的基础上, 本文构造了一个以人工神经网络实现的多分辨率体视协同算法. 在多分辨率特征(边缘)驱动下, 该算法以逐级融合的方式自动完成由粗到细, 直至全分辨率的匹配和重建.

2 算法框架

Marr 等提出了利用 $\nabla^2 G$ 滤波器的多尺度边缘检测和匹配的概念^[1,2], 但它缺乏图像多分辨率分解的严格数学基础. 最近 Mallat 提出的基于二分正交小波的信号分解理论^[5], 将图像多分辨率分析置于严格的数学基础之上, 提供了严密且易操作的多分辨率分析方法, 并将这一理论应用于多尺度边缘检测以及由边缘恢复图像等问题^[6]. 因此, 我们采用 Mallat 的小波多分辨率表达, 即用文[6]中算法提供多尺度边缘. 设用 n 个尺度级, 由细到粗分别为 $0, 1, \dots, l, \dots, n-1$ 级. 匹配和重建在这几个分辨率级上由粗到细逐级进行. 第 l 级重建 f^l 是由该级匹配所得视差样本对第 $l+1$ 级重建 f^{l+1} 更新得到. 在二分小波分解中, 图像尺寸依 2^2 的倍数随分辨率级的降低而递减—— f^l 张在一个比 f^{l+1} 稠密四倍的网格上. 若是没有第 l 级匹配提供的视差样本, f^l 应由 f^{l+1} 经重建滤波器^[5,6]插值乘以 2(分辨率提高一级时以像素数计算的视差值增加一倍)得到. 当有第 l 级视差样本时, 由此得到的视差表面便是 f^l 的初始估计.

1) 航空科学基金资助课题.

收稿日期 1996-04-29

图 1 表示第 l 级匹配和重建的算法框图. 第 $l+1$ 级左、右视差表面估计 f_L^{l+1} 和 f_R^{l+1} 经扩展算子 EXPAND 即前述重建滤波器插值后馈给第 l 级匹配网络以引导其进行匹配,同时也使第 l 级左、右重建网络初始化. 匹配网络是一个自动寻求满足匹配约束的全局一致匹配解的联接主义网络. 它输出的视差样本 d^l 馈给左、右重建网络,两者均由一个视差更新网络和一个间断检测网络构成. 前者执行分片平滑重建,后者从前者给出的结果 f_L^l 和 f_R^l 检测可能存在的沿水平和垂直方向的深度间断 h_L^l 和 v_L^l (h_R^l 和 v_R^l),并将结果馈给前者以修正 f_L^l 和 f_R^l . 经若干次迭代后得到收敛的 f_L^l 和 f_R^l 以两种方式修正匹配——匹配网络的视差一致性兴奋输入和遮挡抑制输入. 如果一个匹配给出的视差与 f_L^l 和 f_R^l 充分接近,它将受到视差一致性兴奋输入的支持. f_L^l 和 f_R^l 提供了遮挡识别线索,一个匹配中只要有一个边缘被判定有遮挡可能时,它将受遮挡抑制输入的抑制. 匹配和重建经若干(通常为 2-3)次迭代后给出收敛解.

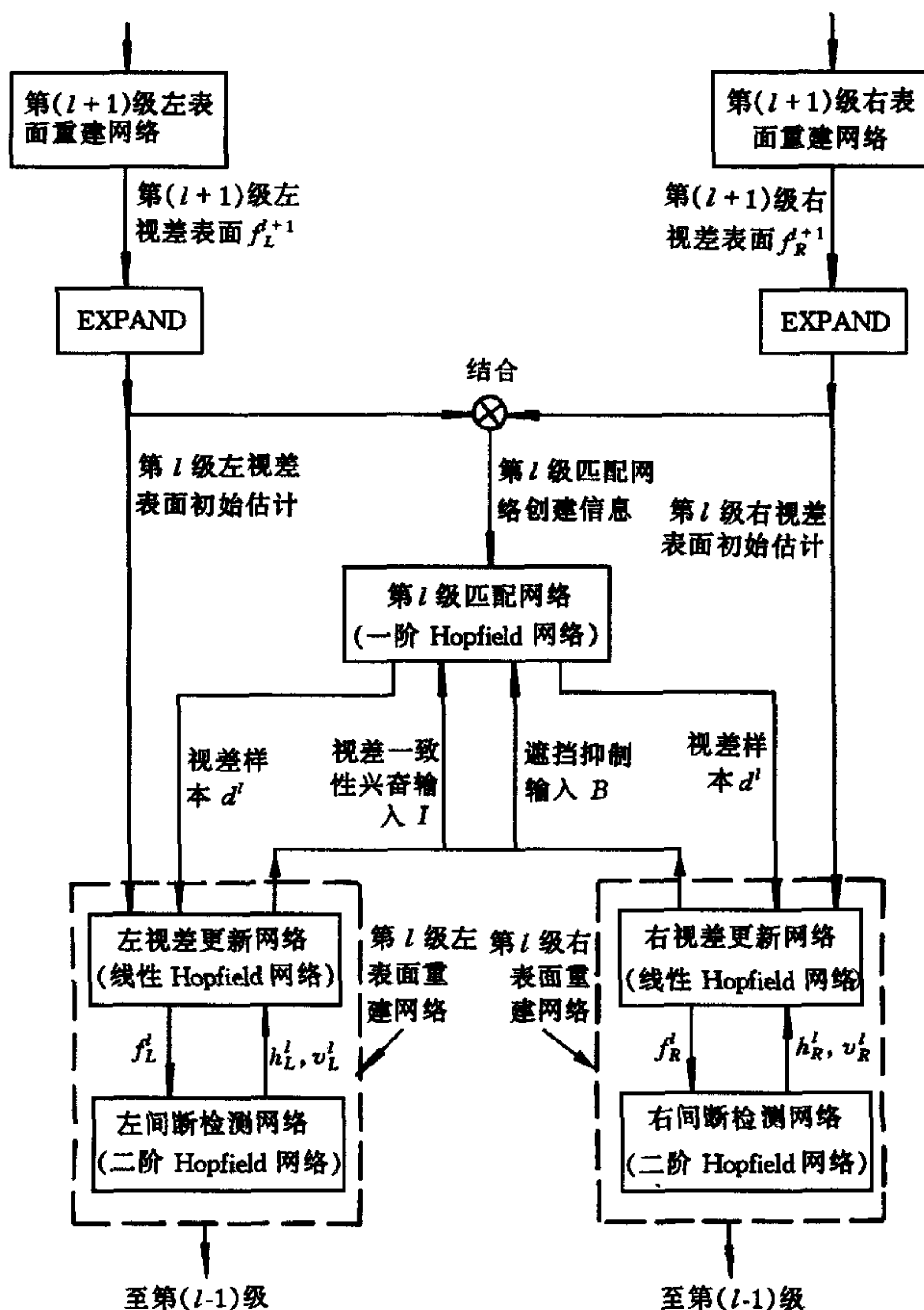


图 1 算法框图

3 算法模块

算法包含边缘检测、匹配网络和重建网络模块. 边缘检测算法见文[6],它产生尺寸依 2^2 递减的多尺度边缘地图序列. 一般说来,边缘地图包含许多噪声(伪边缘),故宜在匹配前用适

当的修剪算法(参见文[7])剔除噪声.

匹配网络类似于 Stewart 和 Dyer^[8]的联接主义匹配网络. 在每一级匹配网络中, 每个神经元代表一个可能匹配, 神经元之间的联接直接体现匹配诸约束. 下面分别说明其神经元设置、内部联接和外部输入.

3.1 神经元设置

第 l 级匹配网络中神经元的设置基于该级的边缘地图以及第 $l+1$ 级的重建结果. 如上所述, 第 l 级重建的初始估计为

$$f_L^l = 2\text{EXPAND}(f_L^{l+1}), \quad f_R^l = 2\text{EXPAND}(f_R^{l+1}). \quad (1)$$

对第 l 级左图边缘 P_L , 右图上可能匹配搜索窗为 $[-W + f_L^l(P_L), W + f_L^l(P_L)]$, 其中 $2W$ 为搜索窗宽. 若窗内有边缘 P_R 与 P_L 朝向相似(例如朝向差 $\leq 60^\circ$), 则记 P_R 为 P_L 的一个待选可能匹配. 对右图的每个边缘可类似地搜索其待选可能匹配. 若 P_L 和 P_R 互为待选可能匹配, 记 (P_L, P_R) 为一可能匹配, 并在第 l 级匹配网络中设置一个神经元与之对应.

3.2 内部联接

我们曾在文[9]中总结了以下的匹配规则: 相容性(朝向相似的边缘方可匹配); 唯一性(两匹配边缘不能同时再有其他对应点); 平滑性(相邻边缘段上的视差变化应较为平缓); 连续性(连续边缘段如果确系三维边缘的投影, 其上视差应是连续的). 这里仍采用这四个规则. 相容性已体现在神经元的设置中, 现在要将其余的三个规则体现在神经元的联接中. 令神经元 i 表示可能匹配 (l_i, r_i) , l_i 和 r_i 分别表示位置为 (x_{L_i}, y_{L_i}) 和 (x_{R_i}, y_{R_i}) ($y_{L_i} = y_{R_i}$) 的左、右边缘点; 记作 $i = (l_i, r_i)$. 它给出的视差为

$$d_i = x_{R_i} - x_{L_i}. \quad (2)$$

1) 唯一性联接. 对于神经元 i 和 j , 如果 $l_i = l_j$ 或 $r_i = r_j$, 则其间有对称的抑制性联接, 联接权 ϵ 为一适当常数. 上述联接体现了单边唯一性. 这意味着若 l_i 和 r_i 中只有一个有竞争匹配时, 唯一性抑制比两者皆有竞争时弱, 因此满足单边唯一性的匹配有可能得到认可.

2) 视差平滑性联接. 依照 Julesz^[10]计算视差梯度, 先定义两匹配 i 和 j 的中眼距离

$$D(i, j) = \max(|x_{c_i} - x_{c_j}|, |y_{c_i} - y_{c_j}|), \quad (3)$$

式中 $x_{c_i} = (x_{L_i} + x_{R_i})/2$, $y_{c_i} = y_{L_i} = y_{R_i}$, $x_{c_j} = (x_{L_j} + x_{R_j})/2$, $y_{c_j} = y_{L_j} = y_{R_j}$. 令 i 和 j 上的视差之差为

$$\Delta d(i, j) = |d_i - d_j|. \quad (4)$$

再定义视差梯度为

$$g_{ij} = \frac{\Delta d(i, j)}{D(i, j)}. \quad (5)$$

用下述方式实现视差平滑性约束: 若 $D(i, j) < W$ (距离阈值, 取为可能匹配搜索窗宽之半) 且 l_i 和 l_j 或 r_i 和 r_j 不在同一局部边缘段上, 则当 $g_{ij} < T$ (视差梯度阈值, $T = 1 - 2$) 时称 i 和 j 满足视差平滑性约束. 那时神经元 i 和 j 之间有兴奋性互联, 联接权 W_{ij} 用 Prazdny^[11]的高斯相似性函数确定

$$W_{ij} = \frac{\alpha}{D(i, j)} \exp\left(-\frac{g_{ij}^2}{2C^2}\right), \quad g_{ij} < T. \quad (6)$$

式中 C 为正常数, 一般可取 $C \in [0.55, 0.85]$; α 为适当正常数. 注意, 我们将平滑性约束的作用限制在阈值 W 之内, 是因为距离过大时两匹配不在此约束复盖范围之内的可能性较大, 同

时这也有利于网络联接的局域化.

3) 视差连续性联接. 当 l_i 和 l_j, r_i 和 r_j 分别位于同一局部边缘段时, 由视差连续性来评价匹配 i 和 j 的合理性. 若 $\Delta d(i, j) \leq \theta(1-2)$, 则认为它们满足连续性, 相应的神经元之间有对称的兴奋性联接, 联接权为

$$W_{ij} = \frac{\beta}{D(i, j)}, \Delta d(i, j) \leq \theta, \quad (7)$$

β 为适当正常数.

在实现上述两种联接时需先构造局部边缘段. 这里用文[9]中的方法构造局部边缘段.

以上平滑性和连续性联接的设计是谨慎的, 即只对满足这些约束的匹配依它们满足这些约束的程度提供支持, 但不因不满足这些约束而对可能匹配加以抑制, 因为当两匹配不满足这些约束时无法判定是它们不正确, 还是这些约束的假设条件未能复盖它们.

3.3 外部输入

1) 视差一致性兴奋输入. 对匹配 $i = (l_i, r_i)$, 若有 $|d_i - f_L(l_i)| < \eta$ 和 $|d_i - f_R(r_i)| < \eta$, 其中 η 为一适当阈值, 则匹配应受重建结果的支持, 给神经元 i 以兴奋性输入

$$I_i = A(|d_i - f_L(l_i)| + |d_i - f_R(r_i)|) / \eta, \quad (8)$$

式中 A 为适当常数.

2) 遮挡抑制输入. 依照 Olsen^[12], 对匹配 i , 如有 $(d_L)_x(l_i) < -0.7$ 或 $(d_R)_x(r_i) > 0.7$, 则此匹配中至少有一个边缘有被遮挡的可能. 以上 $(d_L)_x$ 和 $(d_R)_x$ 分别为 f_L 和 f_R 沿外极线(x)方向的偏导数. 故此时神经元 i 受抑制性输入

$$I_i = -B, \quad (9)$$

B 为适当常数. 当 i 抑制性输入时不应再有上述兴奋性输入.

给定一个适当的状态更新规则, 网络状态便自行更新且最终收敛于一个稳定平衡状态, 给出诸约束相容且与重建一致的匹配解. 将匹配网络当作一个模拟 Hopfield 网络^[13]来实现, 其中神经元具有典型的 S 形特性(如 $U = g(v) = (1 + e^{-2\lambda v})^{-1}$). 内部联接的对称性保证给出一个稳定的收敛解.

重建网络采用了 Koch 等^[14]用来实现带间断检测的表面重建的一种神经网络. 这种网络由一个实现分片平滑重建的线性 Hopfield 网络(这里称视差更新网络)和一个实现间断检测的二阶 Hopfield 网络组成. 下面分别加以说明.

1) 视差更新网络. 左、右重建网络是一样的, 这里的符号不必区分左右. 记 f 为重建视差表面, d 为匹配所得稀疏采样网格 Ω_D 上的视差样本. h 和 v 标志 f 的间断点, $h_{ij} = 1$ 表示在 (i, j) 和 $(i, j+1)$ 间存在一个间断点, $v_{ij} = 1$ 表示在 (i, j) 和 $(i+1, j)$ 间存在一个间断点; 否则 $h_{ij} = v_{ij} = 0$. 网络通过实现能量函数

$$E_d(f) = C_D \sum_{(i,j) \in \Omega_D} (f_{ij} - d_{ij})^2 + \sum_{i,j} (f_{i,j+1} - f_{ij})^2 (1 - h_{ij}) + \sum_{i,j} (f_{i+1,j} - f_{ij})^2 (1 - v_{ij}) \quad (10)$$

在给定 h 和 v 下的最小化给出分片平滑的重建 f . 式中 C_D 为权系数. 由式(10)得

$$\frac{df_{i,j}}{dt} = -\frac{1}{2} \frac{\partial E_d(f)}{\partial f_{i,j}} = -(\delta_D(i, j)C_D + T_{i,j})f_{i,j} + \bar{h}_{i,j}f_{i,j+1} + \bar{h}_{i,j-1}f_{i,j-1} +$$

$$\bar{v}_{i,j}f_{i+1,j} + \bar{v}_{i-1,j}f_{i-1,j} + \delta_D(i,j)C_D d_{i,j}. \quad (11)$$

式中 $\delta_D(i,j)=1, (i,j) \in \Omega_D$; 否则 $\delta_D(i,j)=0$; $\bar{h}=1-h, \bar{v}=1-v$; $T_{ij}=\bar{h}_{ij}+\bar{h}_{i,j-1}+\bar{v}_{ij}+\bar{v}_{i-1,j}$. 可见每个神经元仅与其四个最近邻以权 $\bar{h}_{ij}, \bar{v}_{ij}, \bar{h}_{i,j-1}$ 和 $\bar{v}_{i-1,j}$ 相联, 具有负自反馈 ($\delta_D(i,j)C_D+T_{ij}$), 且在样点处有外部输入 $C_D d_{i,j}$.

2) 间断检测网络. 对于给定的 f 实现能量函数

$$E_f(h,v) = E_f + E_{L,h} + E_{L,v} + E_G \quad (12)$$

的最小化, 式中

$$E_f = \sum_{i,j} [(f_{i,j+1} - f_{i,j})^2(1 - h_{i,j}) + (f_{i+1,j} - f_{i,j})^2(1 - v_{i,j})],$$

$$E_G = C_G \left\{ \sum_{i,j} \left[\int_0^{h_{i,j}} g^{-1}(t) dt + \int_0^{v_{i,j}} g^{-1}(t) dt \right] \right\},$$

$$E_{L,h} = C_v \sum_{i,j} h_{i,j}(1 - h_{i,j}) + C_p \sum_{i,j} h_{i,j}h_{i,j+1} + C_c \sum_{i,j} h_{i,j} + C_L \sum_{i,j} h_{i,j} [(1 - h_{i+1,j} - v_{i,j} - v_{i,j+1})^2 + (1 - h_{i-1,j} - v_{i-1,j} - v_{i-1,j+1})^2],$$

$$E_{L,v} = C_v \sum_{i,j} v_{i,j}(1 - v_{i,j}) + C_p \sum_{i,j} v_{i,j}v_{i+1,j} + C_c \sum_{i,j} v_{i,j} + C_L \sum_{i,j} v_{i,j} [(1 - v_{i,j+1} - h_{i,j} - h_{i+1,j})^2 + (1 - v_{i,j-1} - h_{i,j-1} - h_{i+1,j-1})^2],$$

这里 $g(t) = (1 + e^{-2t})^{-1}$ 为网络神经元特性, $h_{ij} = g(m_{ij}), v_{ij} = g(n_{ij})$. 这一最小化给出在几何结构上具有适度可信性的间断^[14]. 由式(12)有

$$\begin{aligned} \frac{dm_{ij}}{dt} = - \frac{\partial E_f(h,v)}{\partial h_{ij}} = & (f_{i,j+1} - f_{ij})^2 - C_v(1 - 2h_{ij}) - C_p(h_{i,j-1} + h_{i,j+1}) - C_c - \\ & C_L[(1 - h_{i+1,j} - v_{ij} - v_{i,j+1})^2 + (1 - h_{i-1,j} - v_{i-1,j} - v_{i-1,j+1})^2] + \\ & 2C_L h_{i-1,j}(1 - h_{ij} - v_{i-1,j} - v_{i-1,j+1}) + 2C_L h_{i+1,j}(1 - h_{ij} - v_{ij} - v_{i,j+1}) + \\ & 2C_L v_{ij}(1 - v_{i,j+1} - h_{ij} - h_{i+1,j}) + 2C_L v_{i,j+1}(1 - v_{ij} - h_{ij} - h_{i+1,j}) + \\ & 2C_L v_{i-1,j}(1 - v_{i-1,j+1} - h_{i-1,j} - h_{ij}) + 2C_L v_{i-1,j+1}(1 - v_{i-1,j} - h_{i-1,j} - h_{ij}) - \\ & C_G m_{ij}, \end{aligned} \quad (13)$$

dn_{ij}/dt 与式(13)类似, 从略. 由式(13)可见, 每个神经元与其相邻的八个元有一阶和二阶对称联接. 局域联接使网络收敛较快.

第 l 级视差更新网络依式(1)初始化, 以该级的视差样本启动平滑重建过程. 间断检测网络在由此得到的结果下检测间断, 其结果返回视差更新网络重作分片平滑重建. 如此经 2-3 次迭代后完成一轮重建. 在最低分辨级上, 可用一个常数视差表面来初始化.

4 模拟结果

这里用随机点图对来检验上述算法的有效性. 所用两个随机点图对的尺寸为 256×256 , 所包含的“景物”分别为“楔子”和“两层平台”, 如图 2 和 3 所示. 在 256 个灰度级中, 背景灰度为 0, 随机点灰度为 200, 随机点密度为 20%. 采用如此低的密度是为了使匹配网络的规模不致太大, 但这增加了重建的难度, 特别是“景物”轮廓的重建. 为了缓解这一困难, 赋予景物区

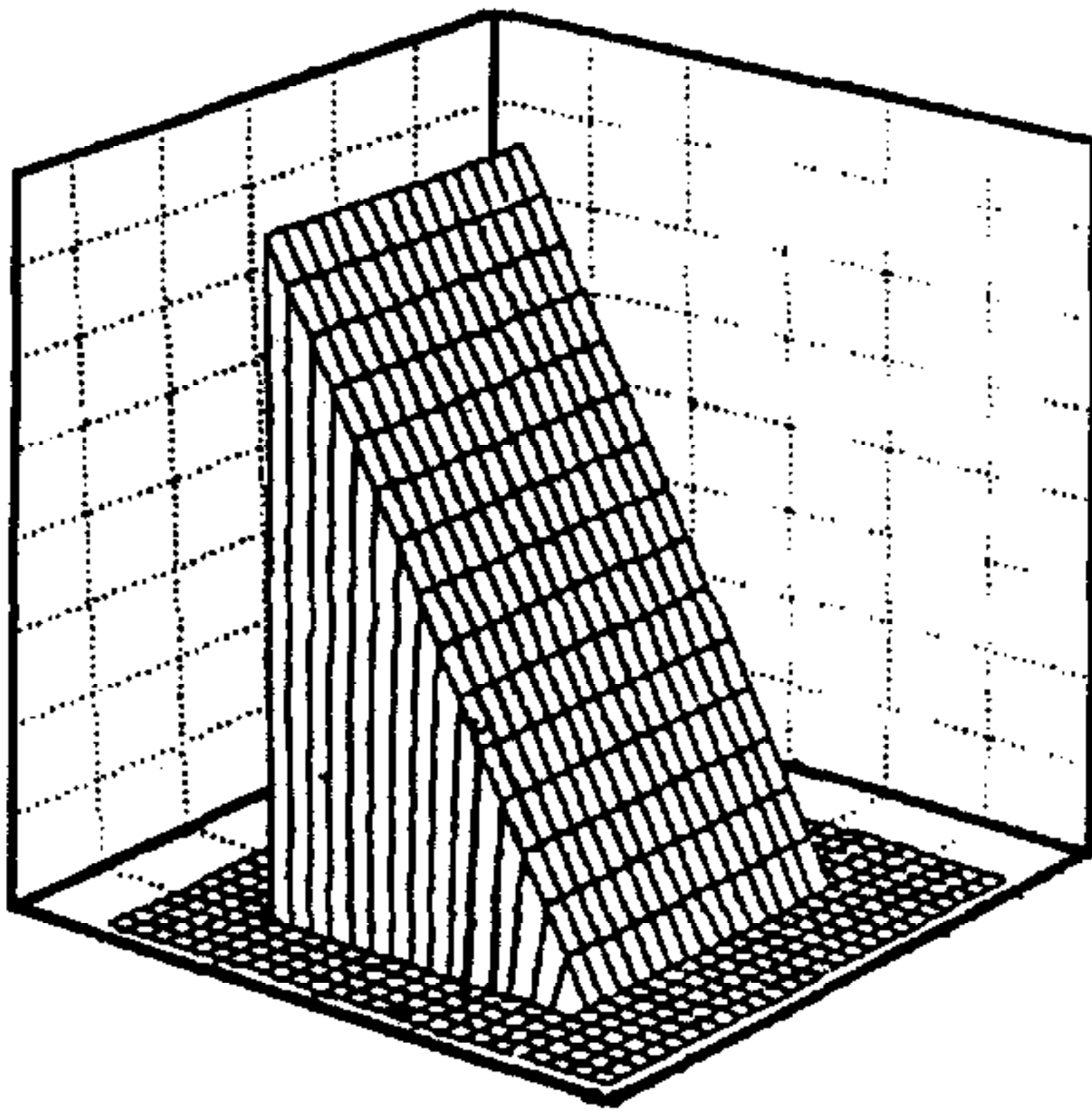


图 2 “楔子”

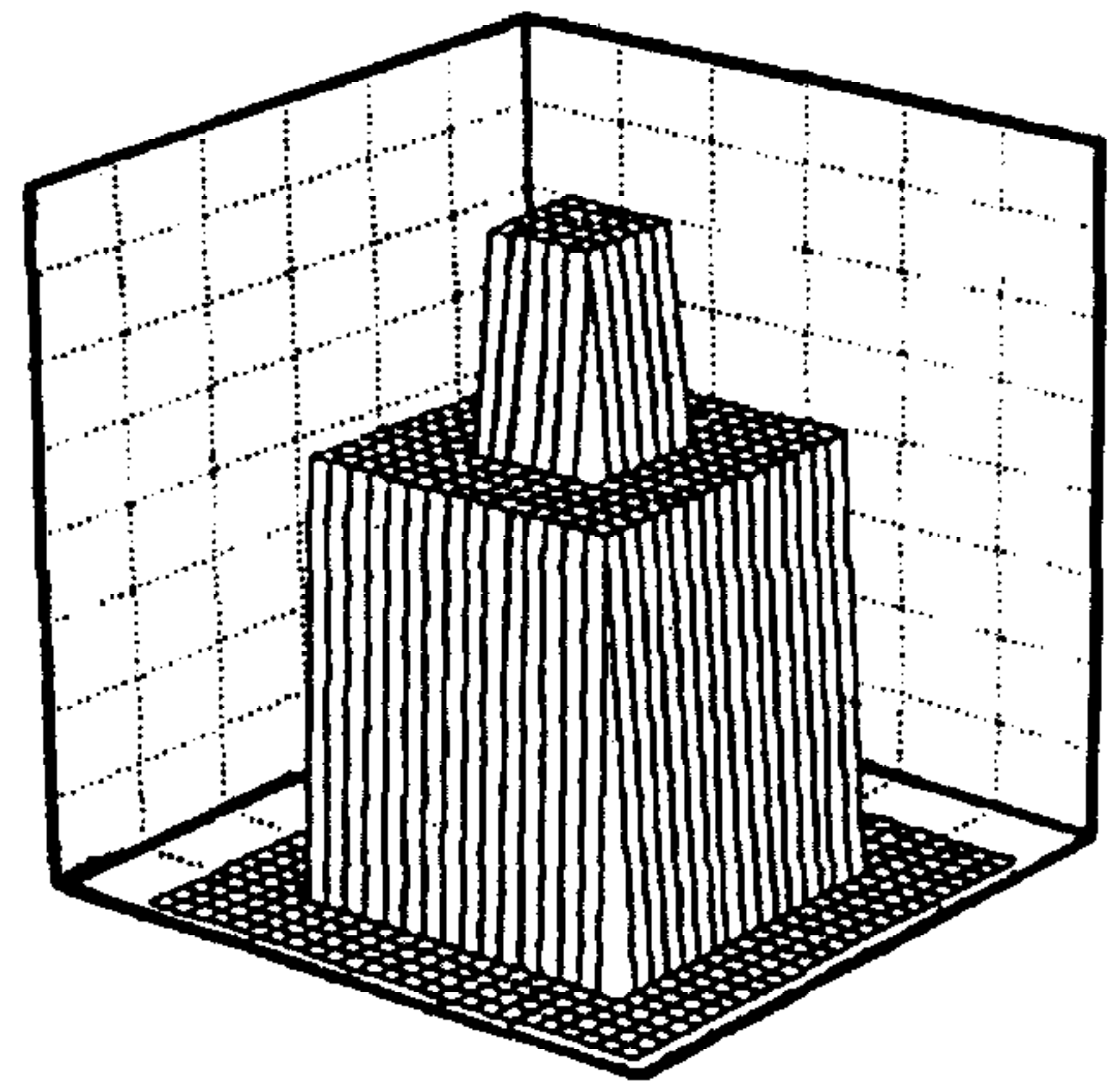
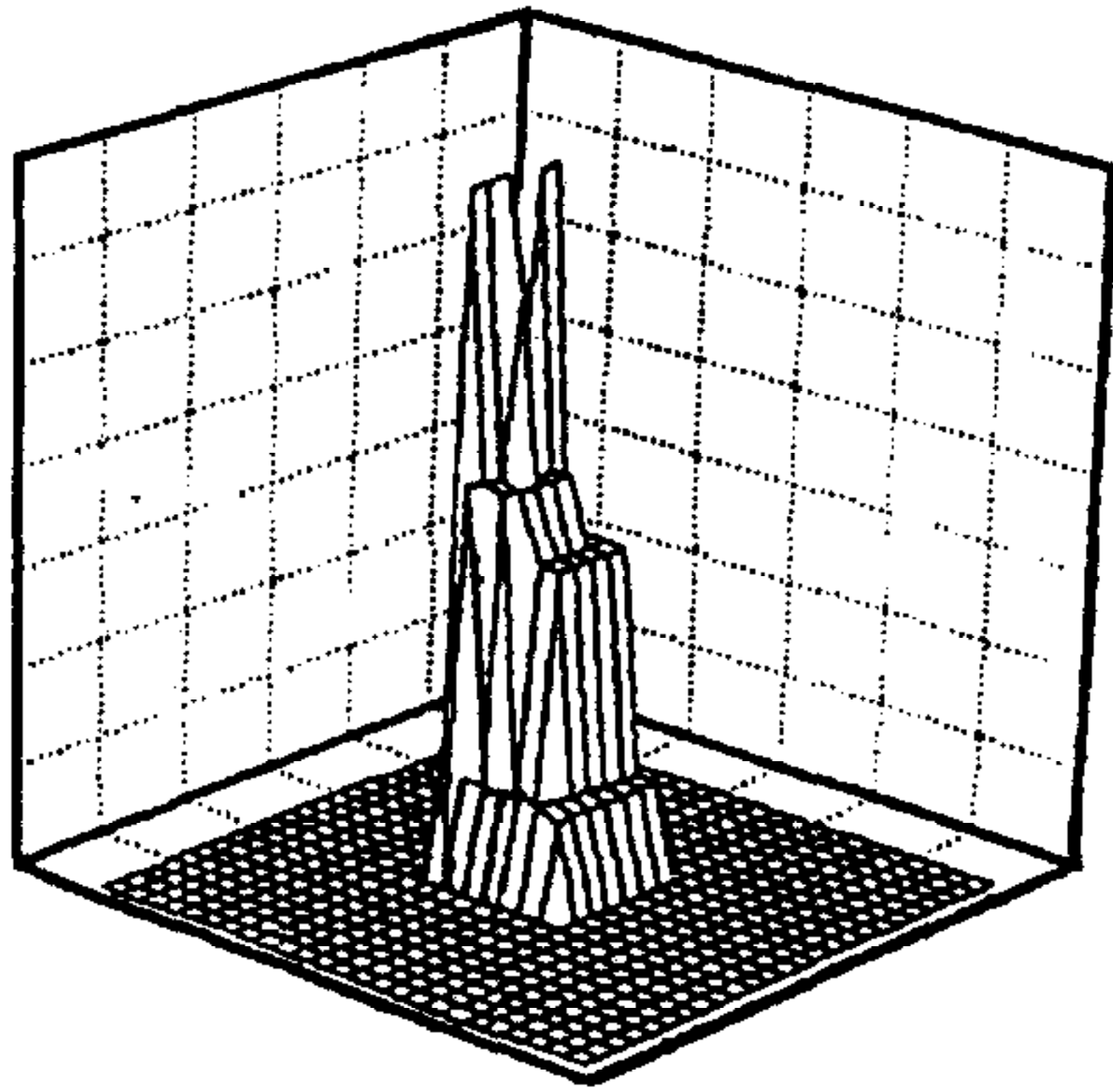
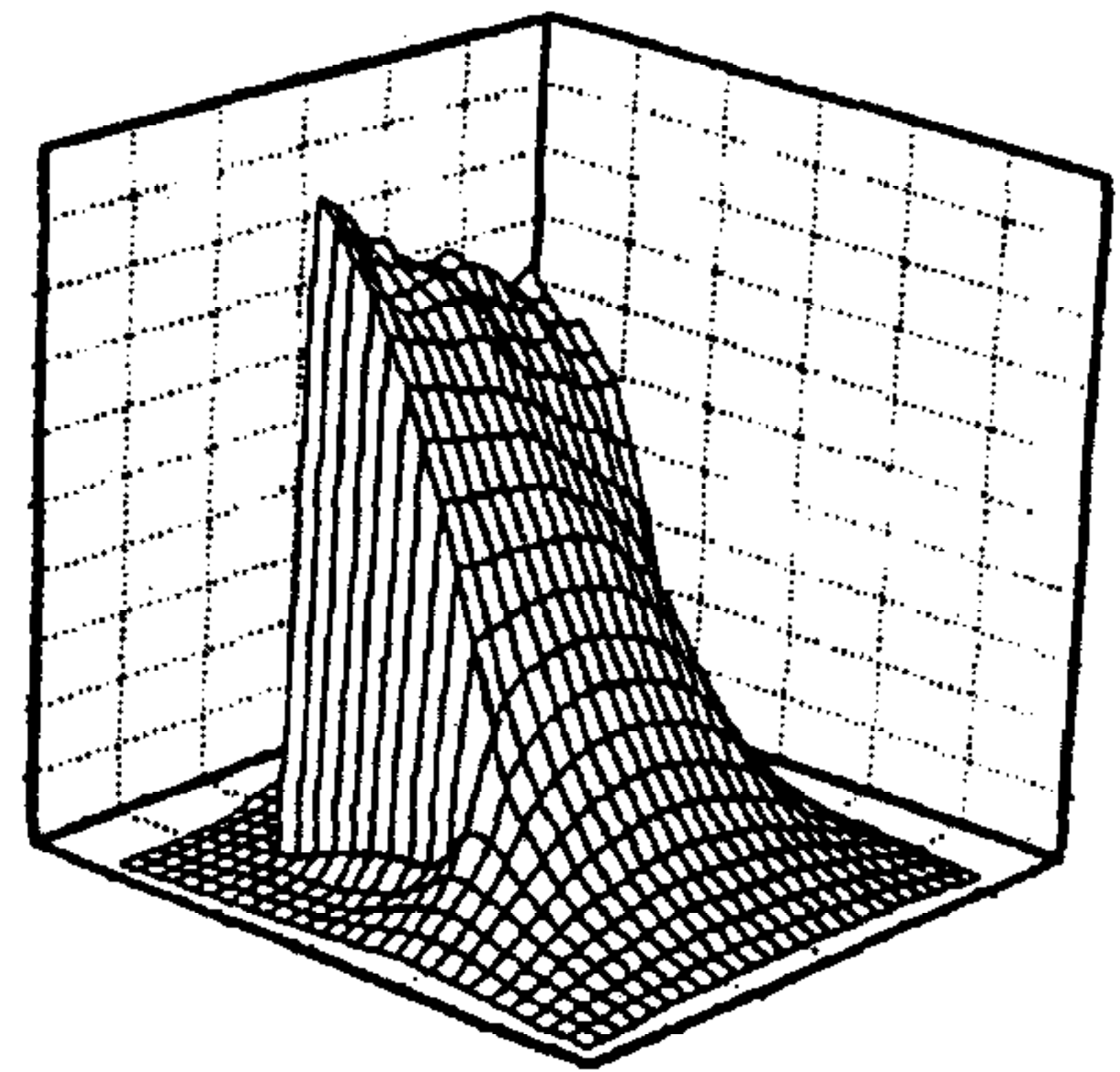


图 3 “两层平台”

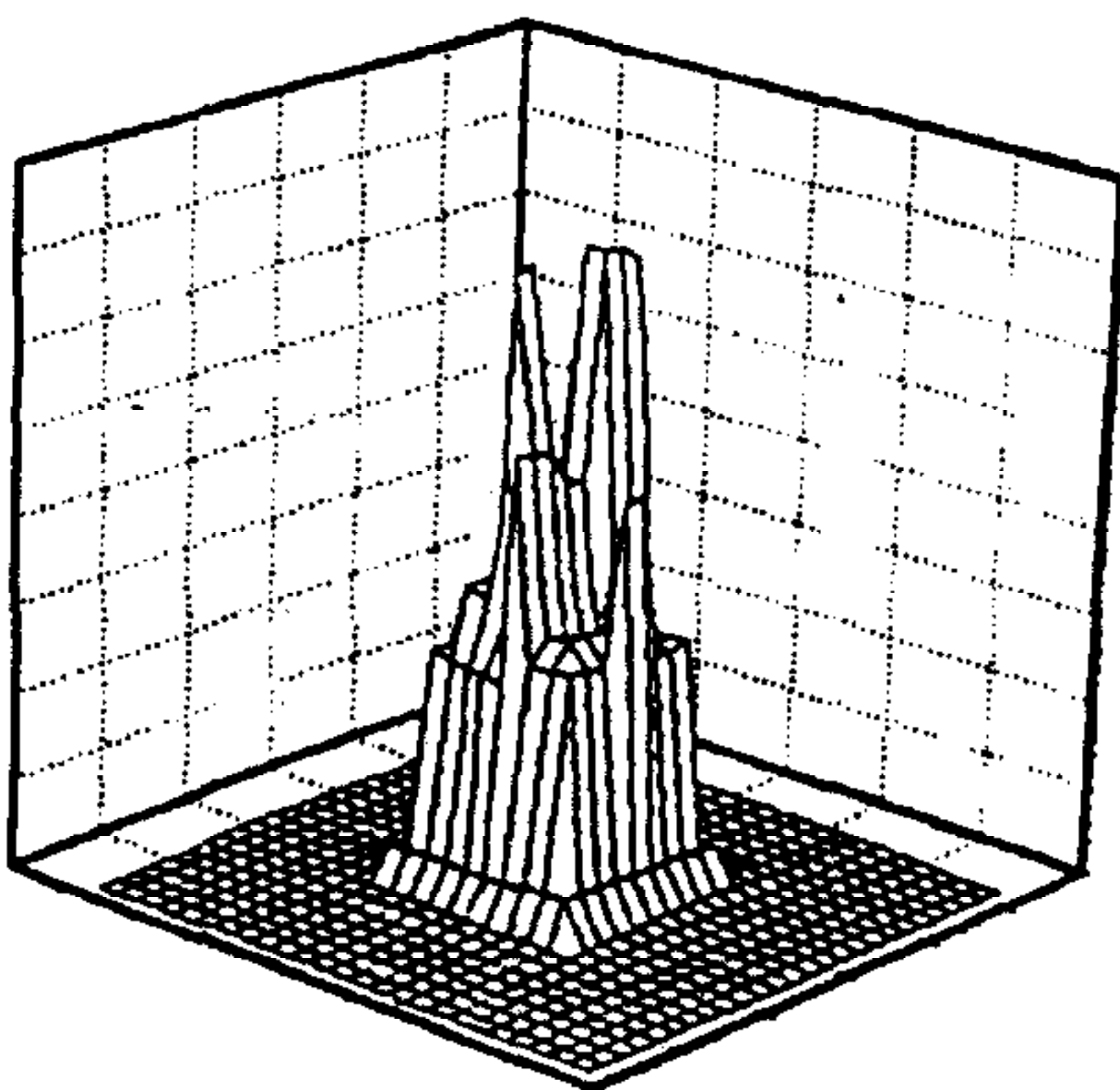


(a) f_L^1

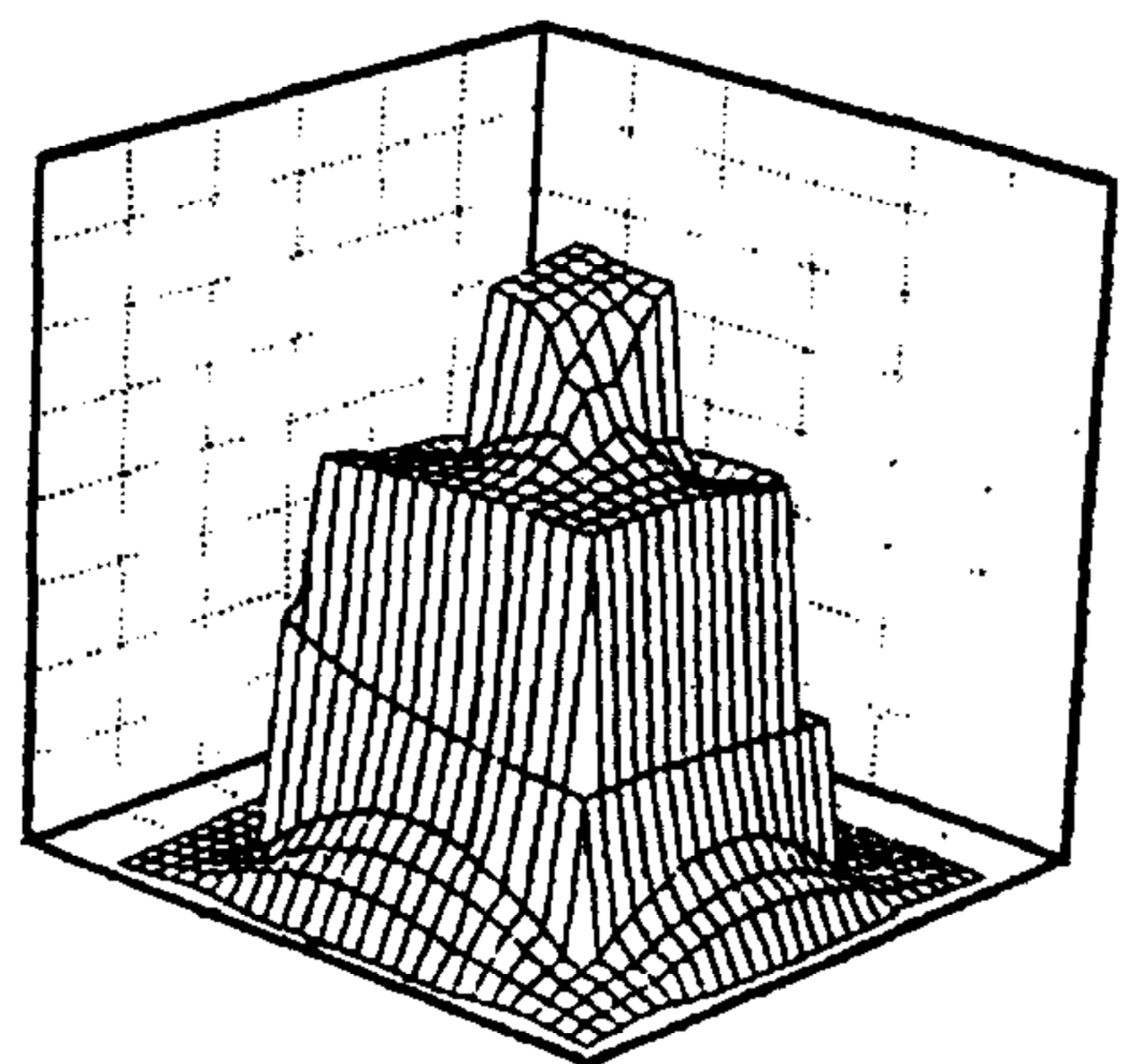


(b) f_L^0

图 4 “楔子”两级重建结果



(a) f_L^1



(b) f_L^0

图 5 “两层平台”两级重建结果

以有别于背景的灰度,从而使“景物”有了完整的边界. 只用两个分辨率级,匹配网络的参数为 $\epsilon=0.1$, $\alpha=0.1$, $\beta=0.3$, $C=0.65$, $\eta=3.0$, $A=1.0$, $B=0.5$, W 对第 0 和 1 级各为 20 和 10. 重建网络的参数为 $C_D=0.5$, $C_v=0.05$, $C_p=0.005$, $C_c=0.01$, $C_L=0.005$, $C_G=0.05$, $\lambda=16$. 重建结果示于图 4 和 5(为节省篇幅,只给出左视差表面).

基本试验表明,上述算法能够得到相当好的视差表面估计. 进一步的工作主要是通过对随机点及适当的自然图像的试验改进和细化算法以及优化其参数选择.

参 考 文 献

- 1 Marr D, Poggio T. A computational theory of human stereo vision. In: Proc. R. Soc., London;1979. B 204:301—328
- 2 Marr D. Vision. New York: Freeman, 1982
- 3 Poggio T *et al.* Computational vision and regularization theory. *Nature*, 1985, 317:314—319
- 4 Berteto M *et al.* Ill-posed problems in early vision. In: Proc. IEEE, 1988,76:869—889.
- 5 Mallat S G. A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Trans. PAMI*, 1989, 11:674—693
- 6 Mallat S, Zhong S. Characterization of signals from multiscale edges. *IEEE Trans. PAMI*, 1992, 14:711—732
- 7 李文彪,潘士先. 弱正则化边缘检测. 自动化学报,1996,22(5):545—553
- 8 Stewart C V, Dyer C R. A connectionist model for stereo vision. In: Proc. IEEE 1st Conf. on Neural Networks. San Diego, CA:1987, 215—223
- 9 李文彪,潘士先. 一个两级迭代体视匹配算法. 自动化学报,1992,18(4):464—469
- 10 Julesz B. Foundations of Cyclopean Perception. Chicago; IL. Univ. Press, 1977
- 11 Prazdny K. Dectection of binocular disparities. *Biol. Cybern.*, 1985, 52:93—99
- 12 Olsen S I. Stereo correspondence by surface reconstruction. *IEEE Trans. PAMI*, 1990, 12:309—315
- 13 Hopfield J. Neurons with graded response have collective computational properties like those of two-state neurons. In: Proc. Natl. Acad. Sci. USA, Biophysics, 1984, 81:3088—3092
- 14 Koch C *et al.* Analog neuronal networks in early vision. In: Proc. Natl. Acad. Sci. USA, Biophysics, 1986, 83:4263—4267

A COMBINATORIAL STEREO VISION ALGORITHM IMPLEMENTED AS NEURAL NETS

LI WENBIAO NING JING PAN SHIXIAN

(*Beijing Univ. of Aeronautics & Astronautics, Beijing 100083*)

Abstract Described here is a combinatorial multiresolution stereo vision algorithm. In the algorithm, dyadic orthogonal wavelet is used for multiresolution decomposition, and, a matching while reconstruction strategy is taken, in which matching and surface reconstruction correct each other and proceed cooperatively. The algorithm is implemented in the form of neural nets, mainly, Hopfield net. Results of primary tests show its rationality and effectiveness.

Key words Stereo vision, feature matching, surface reconstruction, multiresolution analysis, neural network.

李文彪,潘士先 简介见本刊第 22 卷第 5 期.