

# 具有旋转不变性的纹理分割方法<sup>1)</sup>

汪涛 吴吟 邢小良 庄新华

(浙江大学计算机科学与工程系, 杭州 310027)

## 摘 要

本文提出了一种具有旋转不变性的纹理分割算法。纹理图象被模拟为多层次过程综合作用形成的, 它们是区域过程、纹理过程和退化过程。整个方法包括具有旋转不变性的纹理特征抽取、纹理特征的光滑、基于多层次神经网络的初始分割以及利用多值 Boltzmann 机的进一步分割。对自然纹理图象的分割结果证明了算法的有效性。

**关键词:** 纹理分割, 旋转不变性, 层次神经网络, 多值 Boltzmann 机。

## 一、前 言

边缘检测和区域检测技术是图象分割的两种主要技术<sup>[1]</sup>, 它主要依据于图象灰度变化, 因此无法用于纹理图象分割。由 Gibbs 分布表征的马尔可夫随机场模型在图象统计处理技术中得到广泛运用<sup>[2]</sup>, 通常包括图象模型参数估计和图象搜索两个步骤。

本文结合马尔可夫随机场模型、旋转不变性纹理特征抽取和光滑技术以及神经网络, 发展了具有旋转不变性的纹理分割算法。用区域过程、纹理过程和退化过程描述和模拟纹理图象。笔者选择局部纹理特征<sup>[3]</sup>描述每个象素, 并且引入了边缘保护的光滑技术<sup>[4]</sup>。为了避免对信号结构的假设, 采用了神经网络进行学习和分割。对真实纹理图象的实验结果表明了算法的有效性。

## 二、层次纹理图象模型

假设图象定义在  $L = \{(i, j) | 1 \leq i, j \leq N\}$ , 观察图象是层次模型在  $L$  上的一个实现 (realization)。

### 1. 区域过程

区域过程控制具有相同纹理特征的象素聚合成区域, 是定义在图象平面  $L$  上的一个马尔可夫随机场, 表达为  $X = \{X(i, j) | (i, j) \in L\}$ , 且  $X(i, j) \in \{0, 1, \dots, N_r - 1\}$ ,  $N_r$  是区域类型数目。那么它的一个实现  $x = \{x(i, j) | (i, j) \in L\}$  表示图象的一个分割, 其中  $x(i, j) = k$  表示象素  $(i, j)$  属于区域类型  $k$ 。

本文于1990年7月10日收到。

1) 国家自然科学基金和国防科技基金资助课题。





反映了纹理的方向性特征,分别定义为

$$\begin{aligned} \text{CMDM}(i, j) &= \sum_{l=1}^8 dc(l)/8, \\ \text{CVDM}(i, j) &= \sum_{l=1}^8 [dc(l) - \text{CMDM}(i, j)]^2/8, \end{aligned} \tag{3.1}$$

其中  $DC = [dc(1), \dots, dc(8)]$  是循环方向测度,  $DC = \sum_{N(i,j)} |G \cdot D|$ ,  $N(i, j)$  表示以象素  $(i, j)$  为中心的窗口, 绝对值符号代表对每个分量求绝对值, 求和号有类似作用.  $G$  是  $8 \times 4$  矩阵,  $D$  是 4 维向量  $[d(1), \dots, d(4)]$ ,

$$G = \begin{bmatrix} 1 & 0 & 0 & 0 \\ g_1 & g_2 & g_3 & 0 \\ a & b & a & 0 \\ g_3 & g_2 & g_1 & 0 \\ 0 & 0 & 1 & 0 \\ g_3 & 0 & g_1 & g_2 \\ -a & 0 & a & b \\ -g_1 & 0 & g_3 & g_2 \end{bmatrix}, \begin{aligned} d(1) &= z(i, j-1) - z(i, j), \\ d(2) &= z(i-1, j-1) - z(i, j), \\ d(3) &= z(i-1, j) - z(i, j), \\ d(4) &= z(i-1, j+1) - z(i, j), \end{aligned}$$

其中  $g_1 = 0.422$ ,  $g_2 = 0.265$ ,  $g_3 = 0.140$ ,  $a = 0.217$ ,  $b = 0.4$ .  $\alpha(i, j)$  是循环 AR 模型参数, 即

$$z(i, j) = \alpha(i, j) \cdot \sum_{(k,l) \in N_0} g(k, l) z(i+k, j+l) + e(i, j), \tag{3.2}$$

其中  $N_0$  和  $g(k, l)$  的定义和取值参见文献[3]. 方差  $\rho^2(i, j)$  表示剩余噪声  $e(i, j)$  的方差. 本文采用最小均方算法估计 AR 模型参数  $\alpha(i, j)$ .

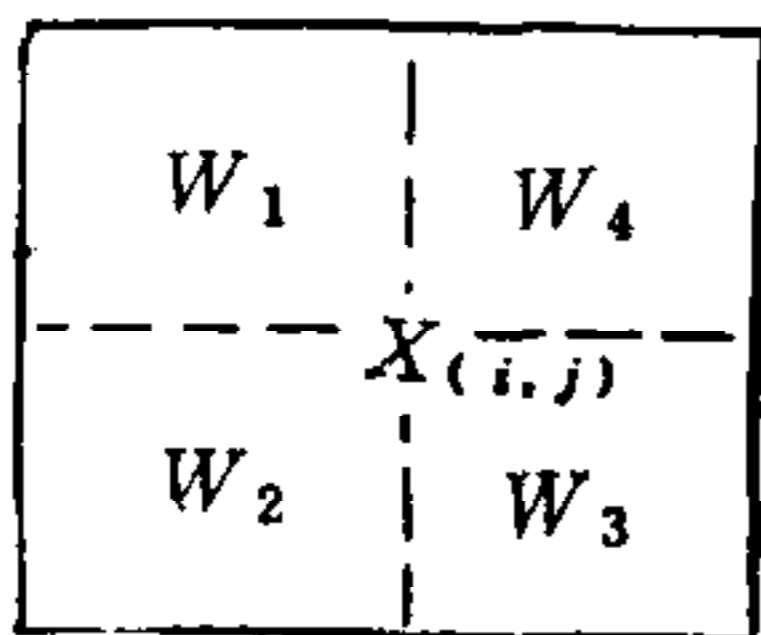


图 1

由于噪声等因素的影响,即使是同一种类纹理,纹理特征  $s(i, j)$  也存在一定变化. 因此, 本文采用了一种边缘保护的光滑技术<sup>[4]</sup>. 对于每个象素  $(i, j)$  定义四个窗口如下. 令  $m(W_1), \dots, m(W_4)$ ,  $v(W_1), \dots, v(W_4)$  分别表示四个窗口中纹理特征  $s(i, j)$  的均值和方差, 即

$$\begin{aligned} m(W_k) &= \sum_{(i,j) \in W_k} s(i, j)/W^2, \\ v(W_k) &= \sum_{(i,j) \in W_k} [s(i, j) - m(W_k)]^2/W^2, \quad k = 1, 2, 3, 4. \end{aligned} \tag{3.3}$$

那么象素  $(i, j)$  的纹理特征  $S(i, j)$  被修正为

$$s(i, j) = m(W_k), \text{ 若 } v(W_k) = \min_{1 \leq l \leq 4} v(W_l), \tag{3.4}$$

其中  $W$  是光滑窗口大小(实验中取  $W = 7$ ). 根据实际需要. 四个窗口大小可以不相同.

### 四、纹理图象分割过程

在纹理图象初始分割中,引入了层次神经网络,包括输入单元层, 隐含单元层和输

出单元层. 输入层神经元数目等于  $s(i, j)$  分量个数, 输出层神经元数目等于可能出现的类别数. 隐含单元数目是由学习算法<sup>1)</sup>动态确定的. 在纹理特征  $s(i, j)$  作为输入值前, 首先必须进行归一化, 即  $s(i, j) = s(i, j) / \max_{(i, j)} s(i, j)$ .

初始分割过程只是根据  $s(i, j)$  确定类别的过程, 类别定义为输出层中状态值最大的神经元定义的类别, 这样就获得了初始分割图象  $v = \{v(i, j) | (i, j) \in L\}$ , 其中  $v(i, j)$  是  $s(i, j)$  所确定的输出类别.

由层次图象模型可知, 初始分割  $v$  是区域过程的一个实现. 因此合理地分割  $v$  应使  $p(X = v)$  达到最大值. 根据 Boltzmann 机的学习算法<sup>[5]</sup>, 要估计一组参数  $\gamma = \{\alpha_k, k = 0, 1, \dots, N_r - 1; \beta_k, k = 1, 2, 3, 4\}$ , 使

$$\max_{\gamma} Q = \ln p(X = v). \quad (4.1)$$

根据本文采用的邻域系统  $\eta$  和象素小集团集合  $C$ , 能量函数  $E(x)$  可表达为

$$\begin{aligned} E(x) = & - \sum_{k=0}^{N_r-1} \sum_i \sum_j \alpha_k \phi[x(i, j), k] - \beta_1 \sum_i \sum_j \phi[x(i, j), x(i, j+1)] \\ & - \beta_2 \sum_i \sum_j \phi[x(i, j), x(i+1, j)] - \beta_3 \sum_i \sum_j \phi[x(i, j), x(i+1, j+1)] \\ & - \beta_4 \sum_i \sum_j \phi[x(i, j), x(i+1, j-1)], \end{aligned} \quad (4.2)$$

其中当  $x_1 = x_2$  时,  $\phi(x_1, x_2) = 1$ ; 当  $x_1 \neq x_2$ ,  $\phi(x_1, x_2) = 0$ . 函数  $Q$  对  $\gamma$  的偏导数为

$$\partial Q / \partial \gamma = - \partial E(v) / \partial \gamma + \sum_x p(X = x) \cdot \partial E(x) / \partial \gamma. \quad (4.3)$$

根据梯度方法, 要使  $\max_{\gamma} Q$ , 迭代估计过程为

$$\gamma^{(n+1)} = \gamma^{(n)} + \partial Q / \partial \gamma, \quad (4.4)$$

其中  $\gamma^{(n)}$  是第  $n$  步的迭代值. 但是, 直接计算  $p(X = x)$  是不可能的, 因为它包含了  $N_r^{N \times N}$  种组合. 下面将以  $\beta_1$  为例说明其计算方法.

$$\begin{aligned} \partial Q / \partial \beta_1 = & \sum_i \sum_j \phi[v(i, j), v(i, j+1)] - \sum_x p(X = x) \\ & \cdot \sum_i \sum_j \phi[x(i, j), x(i, j+1)], \end{aligned} \quad (4.5)$$

右边第一项是常量, 与初始分割有关. 右边第二项表示  $\sum_i \sum_j \phi[x(i, j), x(i, j+1)]$  的期望值, 因为  $p(X = x)$  表示出现  $x$  的概率. 因此可以近似计算  $\partial Q / \partial \beta_1$ : 给定迭代值  $\gamma^{(n)}$ , 利用搜索算法对图象进行搜索, 并且统计每次搜索后  $\sum_i \sum_j \phi[x(i, j), x(i, j+1)]$  的值, 这里一次搜索是指所有象素访问一次; 将统计值对搜索次数求平均值, 近似计算期望值. 当搜索次数趋向于  $+\infty$  时, 两者相同. 实验中, 取搜索次数等于 20 次.

1) 庄新华、吴吟、汪涛, 自动生成层次神经网络的反推 (Back-propagation) 学习算法, 第一届中国人工智能联合学术会议, 1990 年, 458--463.



显然,参数估计和搜索过程构成一个统一整体:参数估计依赖于搜索算法(即分割过程),同时搜索算法也依赖于当前参数估计值.也就是说,纹理图象分割和参数估计同时完成.下面给出多值 Boltzmann 机的搜索算法.

根据 Bayes 规则和 Gibbs 分布式(2.1),象素  $(i, j)$  状态值  $x(i, j) = k$  的概率  $p_k$  定义为

$$p_k = p(x(i, j) = k | x(k, l) \neq x(i, j), (k, l) \in L) \\ = \frac{p(x(i, j) = k)}{\sum_{x(i, j)} p(x)} = \frac{\exp[-E(x(i, j) = k)/T]}{\sum_{x(i, j)} \exp[-E(x(i, j))/T]}, \quad (4.6)$$

其中  $E(x(i, j) = k)$  是局部能量函数,只与象素  $(i, j)$  及其邻域有关,故可以并行处理.搜索算法如下:

- 1) 分别计算概率  $p_k, k = 0, 1, \dots, N_r - 1$ ;
- 2) 产生均匀分布在  $[0, 1]$  间的随机数  $\xi$ ;
- 3) 置状态  $x(i, j) = k^*$ , 如果满足

$$\sum_{l=0}^{k^*-1} p_l < \xi \leq \sum_{l=0}^{k^*} p_l. \quad (4.7)$$

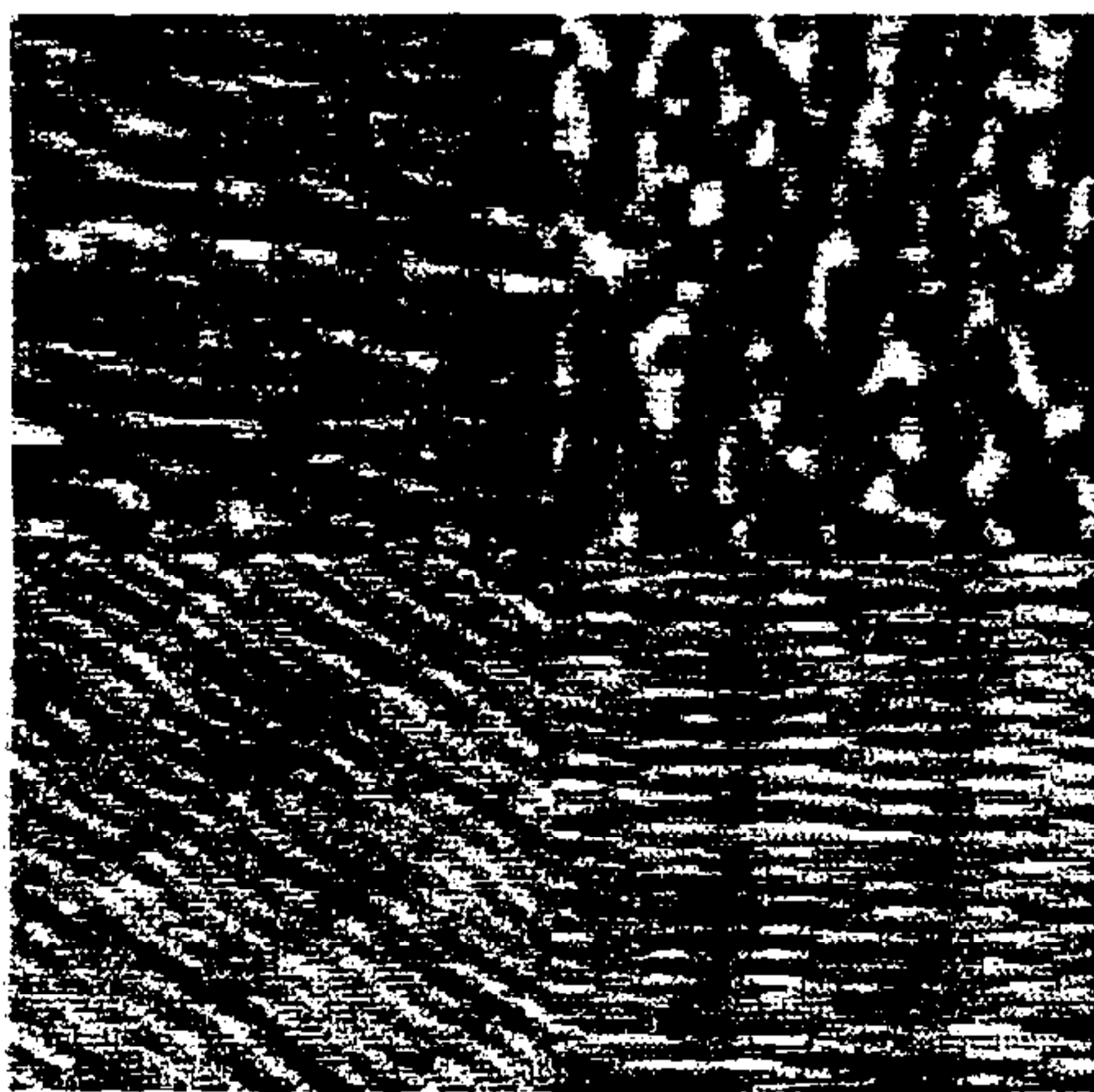
注意当  $N_r = 2$  时,多值 Boltzmann 机转化为双值 Boltzmann 机.温度  $T$  在搜索过程中逐渐降低,

$$T = c_0 / \ln(m + 1), \quad (4.8)$$

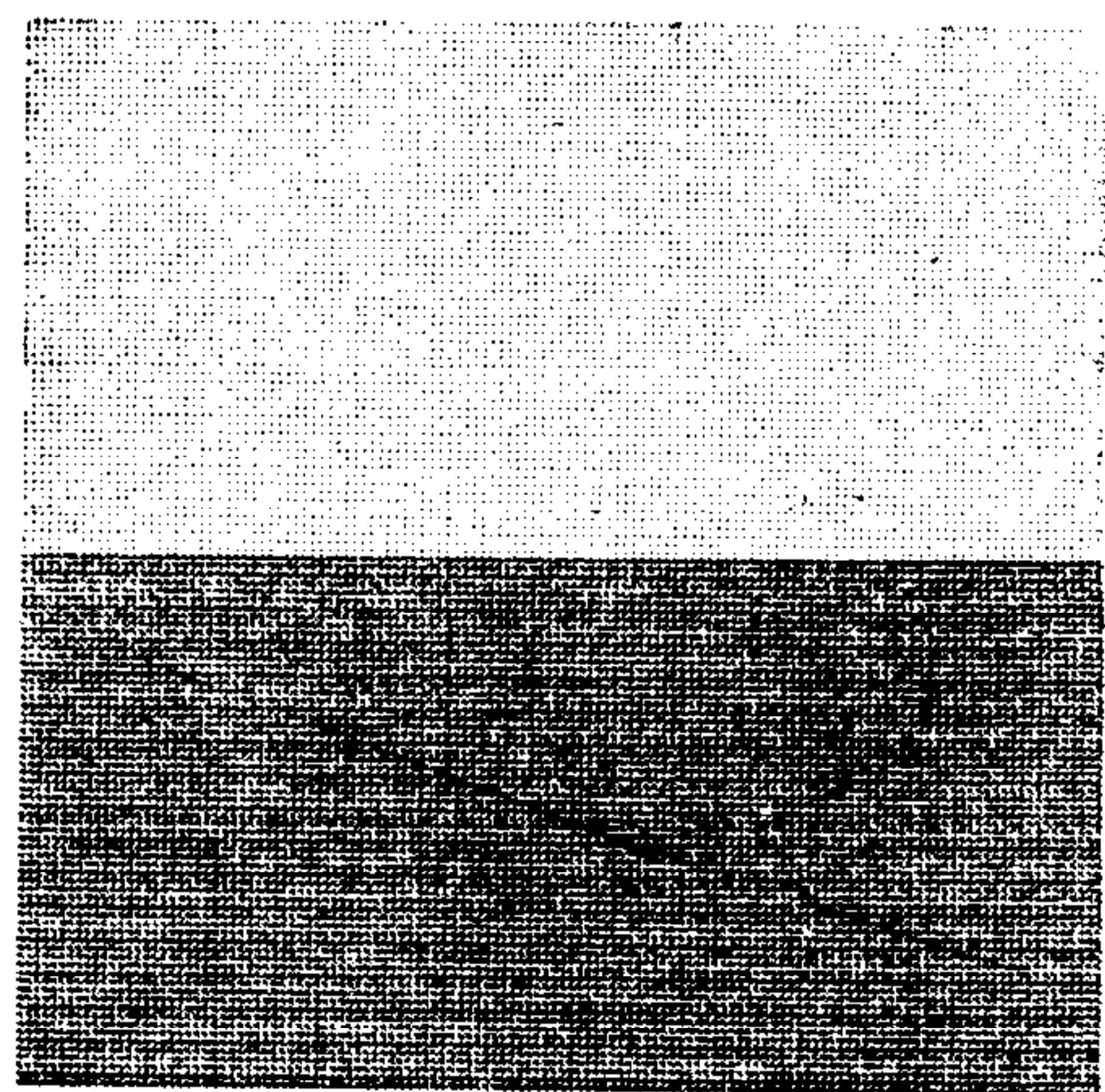
其中  $c_0$  为常数(实验中  $c_0 = 5.0$ ),  $m$  为搜索次数.

## 五、实验结果

图 2 是由两种自然纹理构成的待分割图象和真正的分割图象.



(a) 待分割图象



(b) 真实分割图象

图 2



层次神经网络的训练样本集包括五类纹理，其中的两类构成了待分割图象。对每一类自然纹理，取了 10 幅图象。对每幅图象求光滑后特征向量的均值作为一个训练样本，其类别已知。因此，层次神经网络的训练样本集包括 50 个学习样本。图 3 表示其中的 5 幅图象。在这种情况下，层次神经网络的输出层有 5 个神经元，每个神经元对应一类。例如，对第二类学习样本，对应的期望输出为 01000。当层次神经网络学习完成

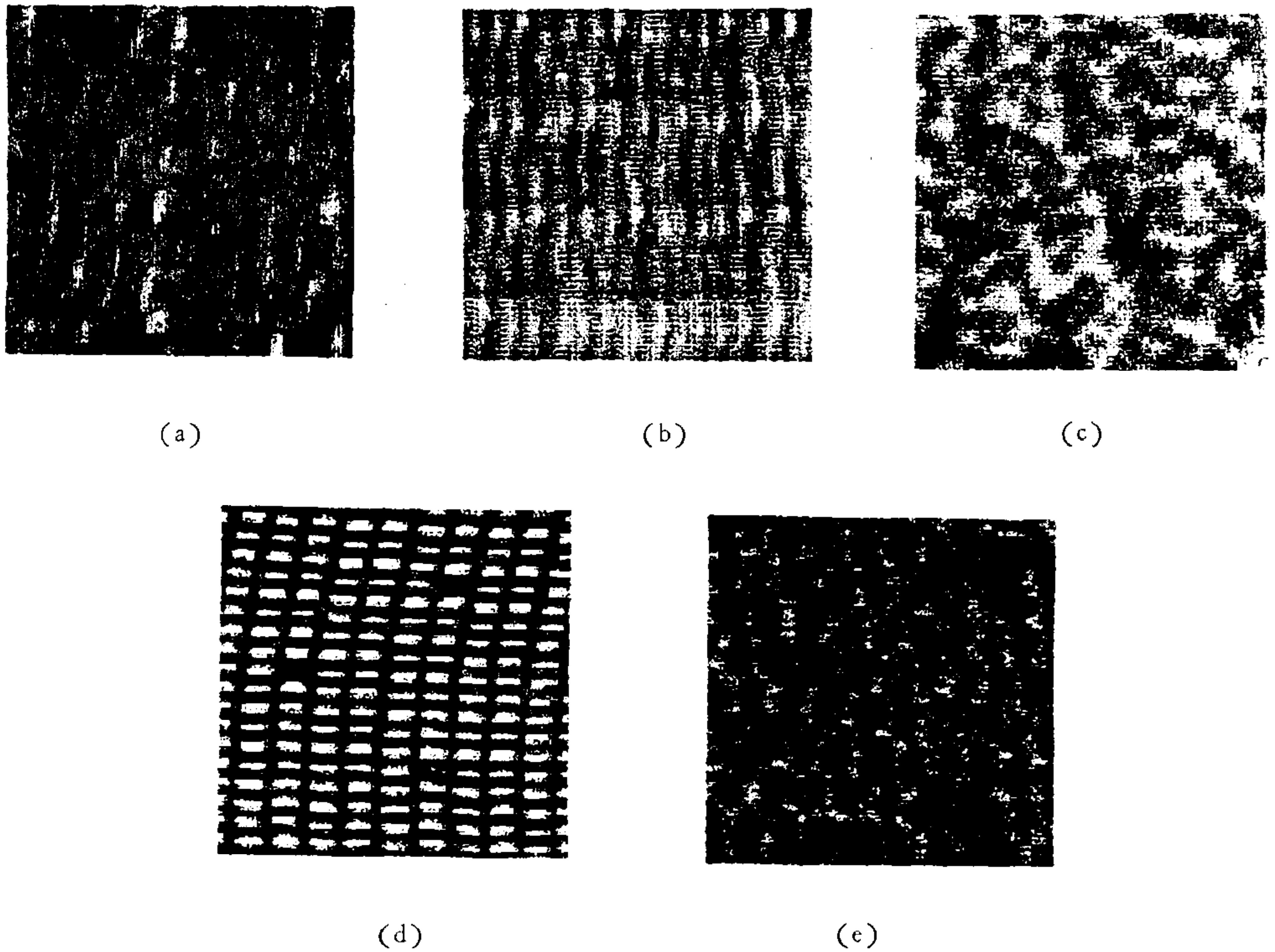


图 3

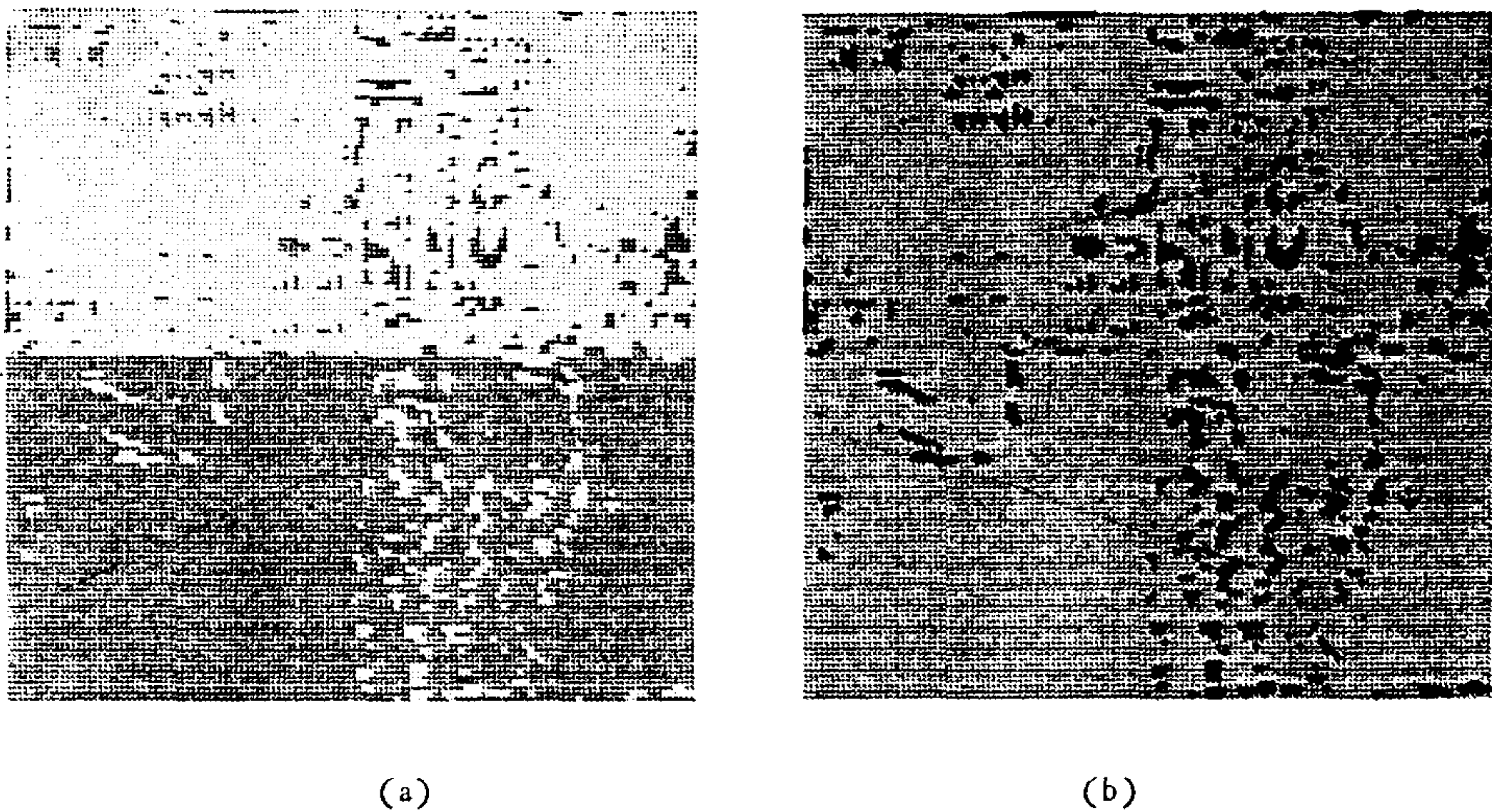


图 4



后,对待分割纹理图象(图 2)进行纹理特征抽取和光滑,然后进行初始分割,与图 2(b)的相对误差是 7.9%。初始分割结果如图 4(a)所示,其中颜色最深的标号表示错分类为其它三种纹理之一的情况。图 4(b)是误分类象素的分布图。

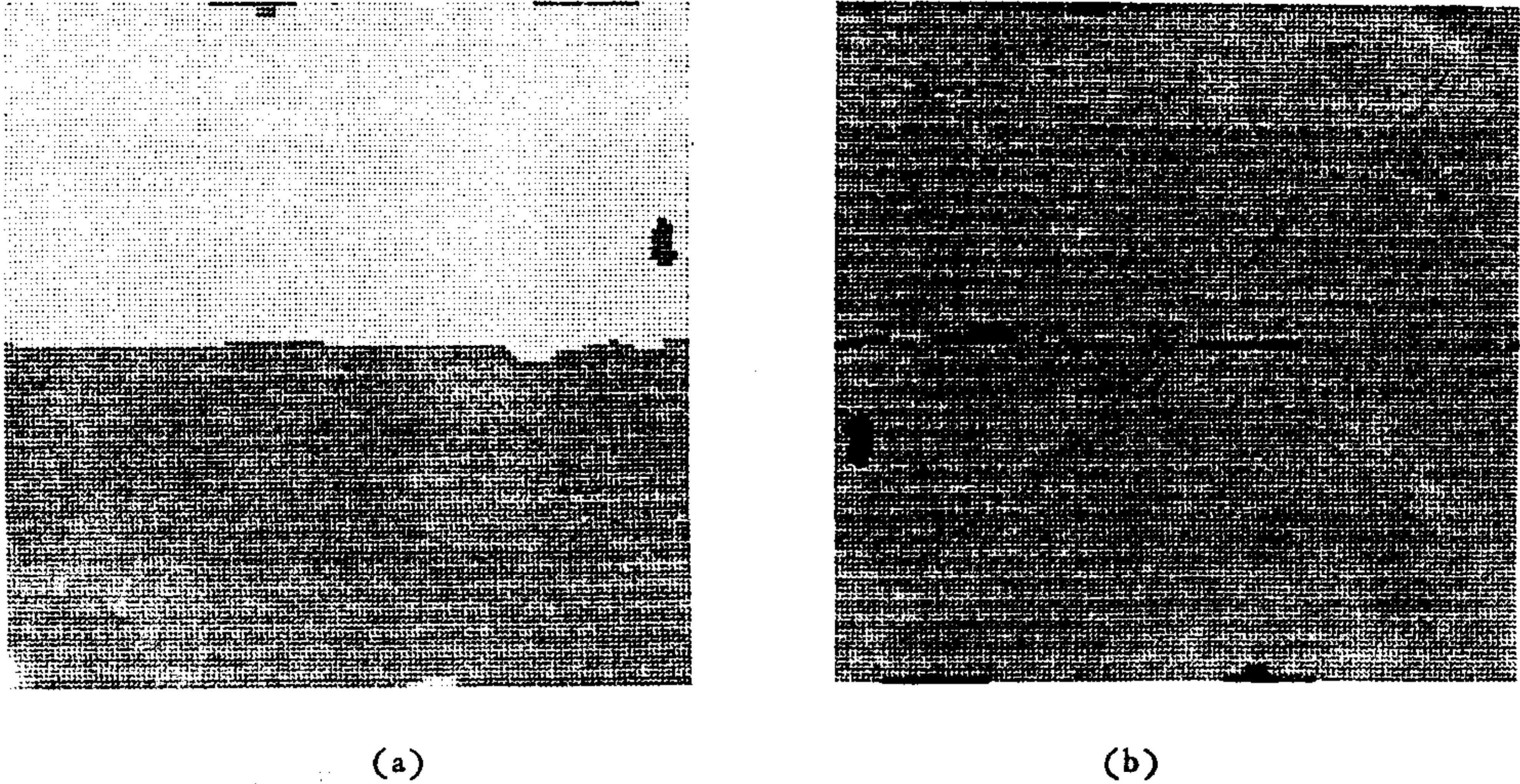


图 5

将初始分割结果作为多值 Boltzmann 机进一步分割的学习样本,使其最有可能出现。图 5(a)表示最终的搜索结果,相对误差是 0.9%。图 5(b)是误分类象素的分布图。

### 参 考 文 献

- [1] Rosenfeld, A. and Kak, A. C., *Digital Picture Processing (2nd)*, New York: Academic, 1982.
- [2] Derin, H. and Elliott, H., "Model and Segmentation of Noisy and Textured Images using Gibbs Random Fields, *IEEE Trans. on Pattern Analy. Machine Intell.*, **9**(1987), 39—55.
- [3] 于晓晗、袁保宗,一种新的旋转不变纹理特征,电子学报, **3**(1990),49—54.
- [4] Hsiao, J. and Sawchuk, A. A., Superrised Textured Image Segmentation Using Feature Smoothing and Probabilistic Relaxation Techniques, *IEEE Trans. on Pattern Analy. Machine Intell.*, **11** (1989), 1279—1292.
- [5] Ackley, D.H., Hinton, G.E. and Sejnowski, T. J., A Learning Algorithm for Boltzmann Machines, *Cognitive Science*, **9** (1985), 147—169.



## A ROTATION-INVARIANT APPROACH FOR THE SEGMENTATION OF TEXTURED IMAGES

WANG TAO    WU YIN    XING XIAOLIANG    ZHUANG XINHUA

(Department of Computer Science & Engineering Zhejiang University, Hangzhou 310027)

### ABSTRACT

This paper presents a rotation-invariant approach for segmenting textured images. The textured images are modeled as three hierarchical processes, namely, the region process, the texture process, and the degradation process. The whole procedure includes extraction of rotation-invariant features, smoothing of these features, initial segmentation based on three-layered neural networks, and extensive segmentation using multi-valued Boltzmann machines. Experiments on natural textures have shown its efficiency.

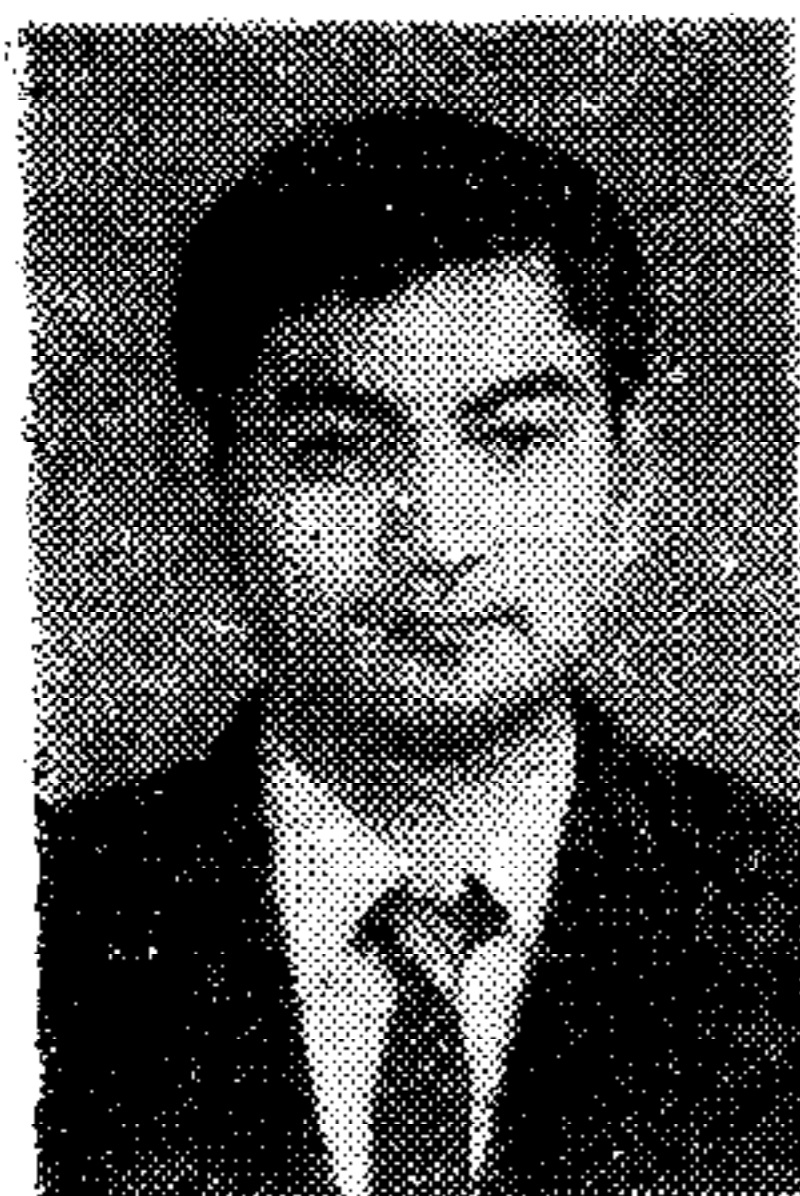
**Key words :** Texture segmentation; rotation-invariancy; layered neural network; multi-valued Boltzmann machine.



**汪 涛** 1967年11月出生于杭州。1989年获浙江大学计算机系学士学位,1991年起攻读计算机系博士学位。研究兴趣为人工神经网络的理论和应用、计算机视觉、模式识别以及图象处理等领域。



**吴 吟** 1988年毕业于浙江大学计算机系,1991年获得计算机系硕士学位。对人工神经网络及其应用有浓厚的兴趣。目前在浙江大学 CAD&CG 国家重点实验室从事图形学方面的研究。



**邢小良** 1981年毕业于浙江大学计算机系,1983年赴法国留学。1987年获得法国博士学位,现任浙江大学计算机系副教授,从事的研究工作包括数字图象处理、计算机视觉和人工神经网络等方面。