

10

54-60

基于子波分解的多通道神经网络 纹理分割方法*

张军 戚飞虎

(上海交通大学计算机科学与工程系, 上海, 200030)

TP391.4J
TN919.8

摘要 描述了一种体现多通道滤波技术的神经网络纹理分割方法, 决策神经网络(DBN)可提高纹理分类的精度, 同时纹理的子波变换降低了图像数据间的相关性, 提高了网络的学习效率, 实验表明本文提出的方法分类误差较低, 获得了令人满意的纹理分割效果。

关键词 纹理分割 | 子波变换 | 多通道滤波 | 神经网络

图像

引言

图像分割在图像分析和计算机视觉应用中是一项十分困难而又非常重要的任务, 在很多情况下, 灰度和颜色在小区域上的差别并不足以提供进行图像分割的依据, 而且在对自然图像的分析中, 有时能够利用的唯一信息只有纹理特征, 因此基于纹理分割的图像分割方法在图像的分析 and 处理中显示出越来越重要的地位。

纹理分割就是在给定的图像上划分出具有均匀纹理的区域, Sklansky^[1]认为如果图像中某一区域的局部统计或其他局部特征保持不变、缓慢变化或周期性变化, 那么这一区域就体现出均匀的纹理特征, 因此纹理不仅取决于某一给定像素的灰度值, 而且取决于围绕这一像素点领域的灰度值分布, 近年来, 多通道滤波方法在纹理分析的应用中得到众多研究人员的关注^[2], 这不但是因为它的实验方法简单, 而且其原理在神经生理学实验中得到了证实^[3], 另一方面, 人工神经网络技术所具有并行计算能力和高度的自适应性在纹理分析中得到了广泛的应用^[4-5]。

1 多通道滤波

纹理图像的多通道滤波过程由特征提取和分类两部分组成, 特征提取中, 纹理图像首先经一组滤波器的滤波, 然后经非线性变换和空域平滑得到一组对应滤波器的特征图像, 最后由分类器进行分类, 完成对纹理的分类和分割, 滤波器的每个通道都能够得到输入纹理的某种局部特征, 例如空间频率、方向、边缘等等, 实际上每个滤波器在空域中表示一个模板, 对输入图像的滤波就是和相应模板的卷积, 用得较多的滤波器组是 Gabor 滤波器, 因为复值 Gabor 滤波器能够提供空域和频域上的最佳局部化。

* 国防预研基金(编号 96J2.4.2)及国家自然科学基金(编号 69572026)资助项目
稿件收到日期 1997-06-17, 修改稿收到日期 1997-08-04

2 神经网络实现

神经网络的典型结构如图 1 所示,我们注意到神经网络的每个单元执行的操作非常近似于多通道滤波过程中的一个通道.神经元计算输入和权向量之间的点积,然后再进行非线性变换.相似地,多通道滤波过程中,一个输入像素的特征值首先由模板系数和输入图像中相应像素的邻域进行点积,再进行非线性变换得到.考虑到这种相似性,我们可以用神经元来实现通道的功能.这样做,一方面使得特征提取和分类在单一的神经网络结构中实现,能够提取到更加适合于分类的特征;另一方面可以应用许多适用于神经网络的技术来提高分类器的性能.神经网络的输入直接来自于输入图像中的 $M \times M$ 区域(M 称为输入模板的大小, M 必须为奇数),输入节点实际上就是这 $M \times M$ 个节点.模板大小 M 的选择取决于纹理分类精度和纹理边界精度之间的权衡考虑.我们采用图 2 所示的图像输入模板,只使用模板中有数字的像素作为 $M \times M$ 窗口中的输入节点,其他像素不予考虑.本文使用纹理图像的

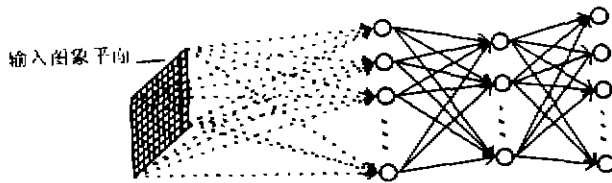


图 1 神经网络的典型结构

Fig. 1 The typical structure of neural networks

3			2			3
	3		2		3	
		3	2	3		
1	1	1	0	1	1	1
		3	2	3		
	3		2		3	
3			2			3

图 2 纹理图像输入模板

Fig. 2 Input template of texture image

子波变换作为神经网络的输入,这样做有两个优点:(1)子波变换可以对输入纹理数据解相关,从而起到与 M. Unser^[6]所使用线性变换相同的作用;(2)子波变换提供了纹理分析中所必需的方向敏感信息.子波变换利用一簇子波函数以及相应的尺度函数把原始信号分解成不同的子带.为了实现子波变换,利用一组正交镜像滤波器 H (低通滤波器)和 G (高通滤波器)对原始信号进行滤波. H 和 G 的冲激响应分别为 $h(n)$ 和 $g(n)$,它们的关系为

$$g(k) = (-1)^k h(1-k). \quad (1)$$

具体的分解方法见图 3,图 3 中 $A_{2^{j-1}}^d f$ 、 $A_{2^j}^d f$ 分别表示图像在分辨率为 2^{j-1} 和 2^j 时的近似, $D_{2^j}^v f$ 、 $D_{2^j}^h f$ 、 $D_{2^j}^d f$ 分别表示图像在垂直、水平、对角线方向上,分辨率为 2^j 时的细节部分.图 2 中,图像输入模板上标数字像素的意义为:0 表示纹理图像的灰度值 $A_{2^j}^d f$,1 表示纹理图像在垂直方向的细节 $D_{2^j}^v f$,2 表示纹理图像在水平方向的细节 $D_{2^j}^h f$,3 表示纹理图像在对角线方向的细节 $D_{2^j}^d f$.

3 神经网络结构

本文所使用的神经网络是 DBNN (Decision-Based Neural Networks),它是由 Koho-

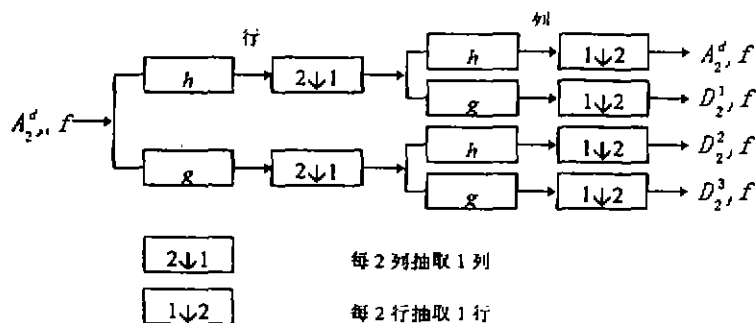


图 3 图像的子波分解方法结构图

Fig. 3 The structure of wavelet decomposition of images

nen^[7]等人提出的. DBNN 又分隐节点 DBNN 和子簇 DBNN 两种形式, 由于隐节点 DBNN 所具有的抗干扰能力和较好的稳定性, 我们采用它作为纹理的分类模型. 它的网络结构如图 4 所示, 整个网络由 L 个子网组成, L 是样本的分类数目. 从结构上网络可分为 4 层: 输入层、隐节点层、累加层和输出层. 每个子网的输入都来自于输入层的原始数据, 隐节点层节点的局部判别函数为 $\varphi_1(\vec{X}, \vec{w}_{1k_1})$, 整数 1 表示第 1 个子网, 整数 $k_1 \in 1, \dots, K_1$, K_1 表示第 1 个子网所拥有的隐节点数目. \vec{X} 表示输入向量 \vec{w}_{1k_1} 为第 1 个子网的第 K_1 个隐节点对应的输入权向量.

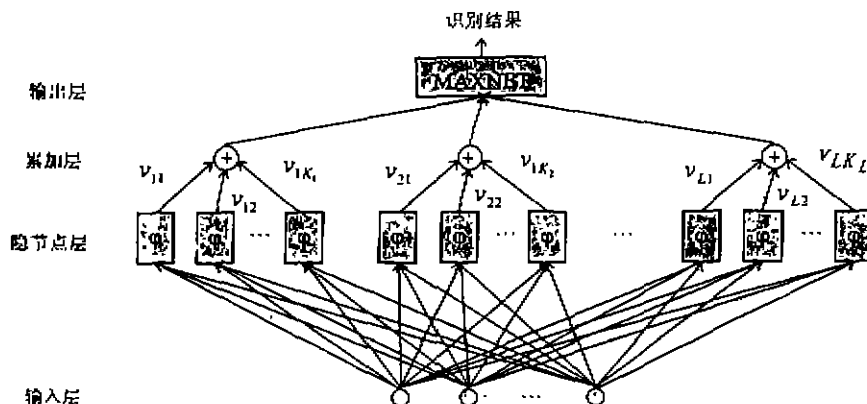


图 4 隐节点 DBNN 网络结构

Fig. 4 The structure of hidden subnodes DBNN

4 学习算法

在 DBNN 中, 样本分类取决于各子网累加层的差别式 $\varphi_1(\vec{X}, W_1)$, 其表达式为

$$\varphi_1(\vec{X}, W_1) = \sum_{k_1=1}^{K_1} v_{1k_1} \varphi_1(\vec{X}, \vec{w}_{1k_1}). \quad (2)$$

式(2)中的 W_1 为第 1 个子网考虑输入层和累加层权值的综合权矩阵, $\varphi_1(\vec{X}, \vec{W}_{1k_1})$ 采用高斯 RBF(Radial Basis Function)的形式,即

$$\varphi_1(\vec{X}, \vec{W}_{1k_1}) = \frac{e^{-\|\vec{X} - \vec{W}_{1k_1}\|^2}}{2\sigma_{1k_1}^2} \quad (3)$$

式(3)中的 σ_{1k_1} 是第 1 子网 K_1 隐节点的标准偏差.

权值应该在判别函数的正梯度方向(强化学习)或它的反梯度方向(弱化学习)上进行调整,即

$$\Delta \vec{W}_{1k_1} = \pm \eta \nabla \varphi_1(\vec{X}, W_1) \quad (4)$$

式(4)中的 η 为正学习常数. 函数 φ 对 \vec{W}_{1k_1} 的梯度向量表示为

$$\begin{aligned} \nabla \varphi_1(\vec{X}, W_1)_{K_1} &= \frac{\partial \varphi_1(\vec{X}, W_1)}{\partial \vec{W}_{1k_1}} \\ &= \left[\frac{\partial \varphi_1(\vec{X}, W_1)}{\partial \vec{W}_{1k_1, 1}}, \frac{\partial \varphi_1(\vec{X}, W_1)}{\partial \vec{W}_{1k_1, 2}}, \dots, \frac{\partial \varphi_1(\vec{X}, W_1)}{\partial \vec{W}_{1k_1, N}} \right] \end{aligned} \quad (5)$$

式(5)中 N 是权向量的维数,也是输入样本的维数.

由式(2)和(3),第 1 子网输入权向量 \vec{W}_{1k_1} 的第 i 个元素的修正量为

$$\Delta W_{1k_1, i} = \pm \eta \frac{V_{1k_1}(\vec{X}_i - W_{1k_1, i})}{\sigma_{1k_1}^2} e^{-\frac{\|\vec{X} - \vec{W}_{1k_1}\|^2}{2\sigma_{1k_1}^2}}, \quad 1 \leq i \leq N \quad (6)$$

相应第 1 个子网的第 i 个累加权值的修正量为

$$\Delta V_{1k_1} = \pm \mu \cdot e^{-\frac{\|\vec{X} - \vec{W}_{1k_1}\|^2}{2\sigma_{1k_1}^2}} \quad (7)$$

式(7)中的 μ 是累加层权值的学习常数.

5 实验结果

为了使实验结果具有可比性,在所有的实验中都使用图 5 所示的 8 种纹理图像,并且,每种纹理都对灰度均值和方差做了规一化,每个纹理图像的大小为 128×128 . 随机地从输入样本中抽选 10%,作为训练的样本. 8 种纹理图像中的任意 $M \times M$ 大小区域都能作为输

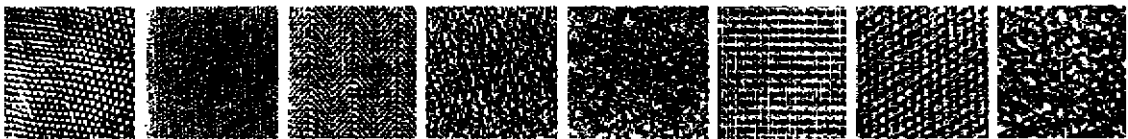


图 5 实验采用 8 种纹理图像
Fig. 5 Eight texture images used for experiment

入样本.在实际应用中,必须利用中值滤波器或高斯滤波器对 DBNN 分类的纹理区域进行平滑滤波,以去除由于局部纹理灰度变化引起的斑点噪声.平滑滤波器模板大小的选择也必须考虑到它对总体分类误差和纹理边界定位精度的影响.本文实验为了使最终的结果具有可比性并避免问题的复杂性,我们得出的误差率都直接来源于神经网络的输出,没有经过任何的后续处理或空域的平滑滤波.

DBNN 的网络参数有输入权向量 \vec{w}_{1k} 和累加权向量 \vec{v}_1 ,以及输入权学习常数 η 和累加权学习常数 μ . \vec{w}_{1k} 和 \vec{v}_1 的每个分量都可以随机初始化为一个大于 0 小于 1 的数, η 和 μ 一般凭经验取值.实验中,我们取 η 为 0.25, μ 为 0.05,径向基函数的标准偏差 σ 取为 0.5.图像输入模板大小分别为 5×5 、 7×7 、 9×9 .

图 6 列出使用 A. K. Jain^[9]方法和本文方法对 3 种不同大小模板下的分类误差率情况.可以看出,用我们的方法得到的误差率更小,这一方面是由于用纹理的子波变换代替了灰度值作为训练样本,使得数据间的相关性更小,另一方面由于我们用高斯径向基函数 DBNN 代替了 A. K. Jain 使用的 BP 网络.BP 神经网络的机理是不断调节联结权值,使得网络的实际输出和目标值之间的误差最小,它有一个逼近的程度问题.而 DBNN 训练的目的是获得一组使得分类正确的联结权值,因此 DBNN 提供了一种更有效的神经网络分类方法.从实验中可以看出,无论是灰度值还是子波变换作为输入,所用的输入模板越大,分类的误差率就越小.在相同的模板下,使用纹理的子波变换作为输入比直接使用其灰度值要来得小,但由于要做一次子波变换,所消耗的时间要多一些.另外模板也不是越大越好:一方面模板的增大会增加网络的规模,从而增加计算量;另一方面在检测纹理边界时,其定位精度降低,会产生较粗糙的边界.所以,在纹理分割时应权衡考虑分类误差和定位精度的影响,根据实际情况选择合适的模板尺寸.图 7 是纹理分割的一个例子,采用的是 7×7 的模板.图 7(a)是与图 5 中的 4 种纹理组合所产生的待分割图像,图 7(b)是用子波 DBNN 方法分割后形成的图像,图 7(c)是图 7(b)经过平衡滤波后的分割结果.

分类方法	模板大小		
	5×5	7×7	9×9
A. K. Jain 的方法	0.348	0.336	0.322
本文采用的方法	0.316	0.302	0.284

图 6 纹理分类误差结果比较

Fig. 6 Comparison of the error of texture classification

6 结语

本文中采用的“细粒度”纹理都可以在较小模板的情况下用神经网络进行分类.多通道滤波技术之所以优于其他小窗口纹理特征提取方法的原因,是在于它既能分割细粒度纹理又能分割粗粒度纹理.“多通道”通常意味着“多分辨率”,细和粗的纹理只是简单地在不同的

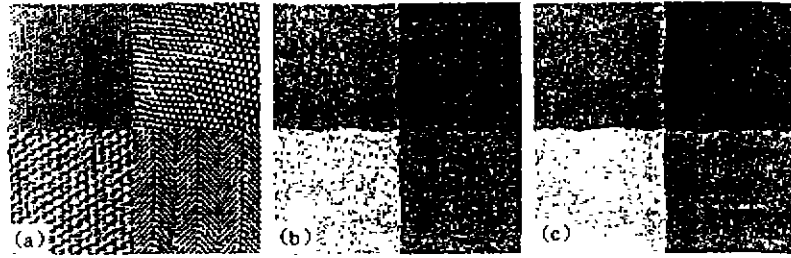


图 7 纹理分割结果 (a) 待分割的纹理图像;(b) 不经后续平滑处理的 DBNN 分割结果,
(c) 经平滑滤波后的纹理分割结果

Fig. 7 The result of texture segmentation (a) input texture image for segmentation (b) the DBNN
segmentation result without post processing (c) the final segmentation result after smoothing

尺度下进行处理.从原理上说,依靠网络的自学习功能可以选择一个最佳的尺度.但实际上,更多的是根据一些网络结构上的先验知识对参数空间加以限制,如图 2 对网络输入的选择就是这样.多尺度模板可以靠增加 $M \times M$ 窗口尺寸再对模板进行亚抽样加实现.当然更通用的方法是用一组参数表示模板,然后在计入分类误差和纹理边界定位精度影响的代价函数曲面的负梯度方向上调整这些参数,使得系统分割的代价函数最小,但在实际运用中这是非常复杂的.

REFERENCES

- 1 Sklansky J. *IEEE Trans. Syst. Man Cybern.*, 1978, **8**(3): 237~247
- 2 Farrokhnia J F. *Pattern Recognition*, 1991, **24**(12): 1167~1186
- 3 Valois D, Albrecht D G, Thorell L G. *Vision Res.*, 1982, **22**(3): 545~559
- 4 Lu H X. *Pattern Recognition*, 1995, **28**(12): 1807~1817
- 5 A K Jain Mao. *Pattern Recognition*, 1992, **25**(2): 173~188
- 6 Unser M, Eden M. *IEEE Trans. PAMI*, 1989, **11**(11): 717~728
- 7 Kohonen T. *Series in Information Science*, 1984, **8**(4): 234~241
- 8 Jain A K, Karu K. *IEEE Trans. PAMI*, 1996, **18**(2): 195~205

A WAVELET TRANSFORMATION-BASED MULTICHANNEL NEURAL NETWORK METHOD FOR TEXTURE SEGMENTATION*

ZHANG Jun QI Fei-Hu

(*Department of Computer Science and Engineering, Shanghai Jiaotong University, Shanghai 200030, China*)

Abstract A neural network texture segmentation method, in which multichannel filtering is embodied, was proposed. Multichannel filtering technology is a very effective method for texture segmentation. Instead of using a general filter bank, the texture feature extraction and classification tasks were performed in this paper by the same unified neural network. Decision-based neural network was adopted to improve the accuracy of classification. Wavelet transformation of texture was used to decrease the correlation of texture data and increase the efficiency of networks learning. Experiments show that the proposed method achieves lower error rates than other methods and a satisfactory result is obtained.

Key words texture segmentation, wavelet transformation, multichannel filtering, neural networks.

* The project supported by the Preliminary Research Foundation of National Defence and the National Natural Science Foundation of China
Received 1997-06-17, revised 1997-08-04