

一种基于不变矩和 BP 网络的目标识别方法

孙红辉, 王红霞, 田 涛

(第二炮兵工程学院, 陕西 西安 710025)

摘 要: 为了有效地提高旋转、尺度变化目标的识别率, 首先提取目标图像的不变矩, 以此作为目标识别的特征向量, 然后利用将附加动量项与自适应学习速率相结合的改进 BP 算法实现对目标的分类识别。字符图像仿真实验表明, 这种针对旋转、尺度变化目标的识别方法是有效的, 可行的。

关键词: 不变矩; BP 神经网络; 目标识别; 改进 BP 算法

中图分类号: TP391

文献标识码: A

文章编号: 1000-7180(2011)03-0063-03

The Recognition Method of Objects Based on Moment Invariant and BP Neural Network

SUN Hong-hui, WANG Hong-xia, TIAN Tao

(The Second Artillery Engineering Institute, Xi'an 710025, China)

Abstract: In order to improve the recognition rate of rotation and scale changed objects, firstly the moment invariant of object images was extracted, which was taken as the characteristic vector, then improved BP algorithm that connects additional momentum with self-adaptive learning rate was used to classify and recognize the objects. The emulation experimental results of character images show the object recognition method of rotation and scale changed images is effective and feasible.

Key words: moment invariant; BP neural network; object recognition; improved BP algorithm

1 引言

目标图像的识别技术在工农业生产和国防军事上都有着重要的应用价值。由于图像中包含大量的信息, 数据量大, 导致目标识别速度不尽人意, 而对于同一个目标, 从不同的方位角度和高度观测时, 目标必然发生平移、旋转和尺度缩放等变化, 也为其识别带来困难。如何使识别系统实现对目标图像的准确和快速识别, 一直是国内外学者着力解决的问题。随着人工智能理论和计算机技术的发展, 神经网络因其具有实时性和鲁棒性等优点, 使得它在模式识别方面的应用研究日益活跃起来^[1-2]。

2 目标图像不变矩特征提取

应用神经网络进行目标识别, 特征向量的提取尤为关键, 特征提取过程是去除冗余信息的过程, 具

有提高识别精度、减少运算量和提高运算速度的作用, 良好的特征应具有可分性、稳定性和独立性等特点。矩是一种非常重要的表示目标总体形状的特征量, 二维图形的几个关键特征均直接与矩有关, 如目标的大小、形心和旋转情况等等。由于不变矩概念清晰、识别率稳定, 对具有平移、旋转和缩放变化的目标有良好的不变性及抗干扰性, 能有效地反映目标的本质特征^[3], 所以文中选取不变矩作为神经网络的输入向量。一幅目标图像可以用二维函数 $f(x, y)$ 表示, 其二维的 $(p+q)$ 阶矩定义如下^[4]:

$$m_{pq} = \sum_{x=1}^M \sum_{y=1}^N x^p y^q f(x, y) \quad (1)$$

从而中心矩表示为

$$m_{pq} = \sum_{x=1}^M \sum_{y=1}^N (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (2)$$

式中, $\bar{x} = \frac{m_{10}}{m_{00}}, \bar{y} = \frac{m_{01}}{m_{00}}$. 归一化的中心矩定义为

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^r} \tag{3}$$

式中, $r = (p+q)/2 + 1, p+q = 2, 3, \dots$. Hu M. K 利用规格化中心矩导出了 7 个不高于三阶的函数式, 称之为 Hu 不变矩, 它们对于图像平移、旋转和尺度缩放都具有不变性.

$$\phi_1 = \eta_{20} + \eta_{02} \tag{4}$$

$$\phi_2 = (\eta_{20} - \eta_{02})^2 + 4\mu_{11}^2 \tag{5}$$

$$\phi_3 = (\eta_{30} - 3\eta_{12})^2 + (\eta_{03} - 3\eta_{21})^2 \tag{6}$$

$$\phi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \tag{7}$$

$$\phi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \tag{8}$$

$$\phi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{21} + \eta_{03})(\eta_{12} + \eta_{30}) \tag{9}$$

$$\phi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \tag{10}$$

实验中发现, 7 个不变矩的数值很小且变化范围很大, 不利于作为识别目标的特征向量, 为了便于比较, 对结果采用取对数的方法进行数据压缩, 并取其绝对值, 实际采用的不变矩为^[3]

$$\phi_k = |\log_{10} |\phi_k|| \quad k = 1, 2, 3, 4, 5, 6, 7 \tag{11}$$

实验中, 用字符代表需要识别的目标, 表 1 中给出了两个灰度字符图像 A、B(如图 1 所示)旋转和尺度缩放时不变矩提取的部分数据.

表 1 图像的不变矩数据

旋转	A 字符的 7 个不变矩						
10°	2.838	6.415	8.864	11.161	22.356	14.369	23.460
20°	2.838	6.416	8.863	11.165	21.484	14.373	23.360
30°	2.837	6.415	8.862	11.152	21.189	14.360	23.172
40°	2.837	6.415	8.864	11.142	21.186	14.351	22.267
尺度	B 字符的 7 个不变矩						
0.8	2.833	6.439	10.534	12.593	24.187	16.274	24.981
1.0	2.834	6.452	10.569	12.788	24.552	16.158	24.591
1.2	2.833	6.445	10.549	12.704	24.368	16.160	24.600
1.4	2.834	6.451	10.553	12.750	24.505	16.098	24.502

由表 1 中数据可以看出, 当两个字符发生旋转和尺度变化时, 其不变矩的波动范围较小, 特征向量具有良好的稳定性和可分性, 这为后续目标图像的准确识别奠定了基础. 数据发生波动变化主要是由于图像分辨率有限, 仅在整数坐标处才有精确值,

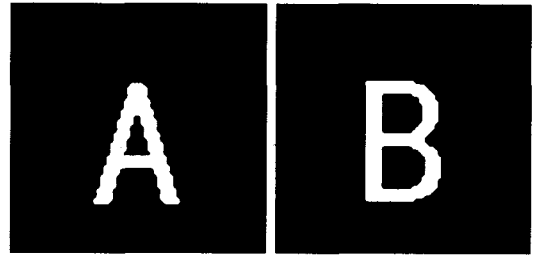


图 1 字符 A 和 B 图像

因此计算所得的不变矩只能保证其平移不变性, 并不能随原图像的几何变换恒定不变, 而是有一定误差, 图像的分辨率越高, 这种误差就越小, 但计算量也会相应地变大.

3 BP 神经网络算法

1986 年, Kumelhart 等人提出误差反向传播神经网络, 简称 BP 网络. 目前, BP 网络已成为各个领域广泛使用的网络模型^[5]. BP 算法其实质上是把一组样本输入和输出问题转化为一个非线性优化问题, 已经证明, 具有非线性 Sigmoid 函数的三层神经网络可以以任意精度逼近任何连续函数. 传统 BP 网络无论在理论上还是实用上都已经相当成熟, 但它还存在以下一些缺陷: 训练学习的速度非常慢; BP 学习算法存在陷入局部极小点的问题.

为了加速 BP 网络的收敛速度和尽量避免陷入局部极小点等问题, 对 BP 算法主要进行了以下的改进算法方案^[2,6].

(1) 附加动量法. 附加动量法使网络在修正其权值时, 不仅考虑误差在梯度上的作用, 而且考虑在误差曲面上变化趋势的影响. 该方法是在反向传播过程中, 在每一个权值的变化上加上一项正比于前次权重变化量的值, 以产生新的权值变化. 其调节公式为

$$w(k+1) = w(k) + \alpha[(1-\eta)D(k) + \eta D(k-1)] \tag{12}$$

式中, $w(k)$ 表示权值矩阵; $D(k) = -\partial E/\partial w(k)$ 为第 k 次的负梯度; α 为学习率, $\eta(0 < \eta < 1)$ 为动量系数. 附加动量法能降低网络对于误差曲面局部细节的敏感性, 有效地抑制了网络陷于局部极小, 可提高目标识别的精度.

(2) 自适应学习速率法. 学习率控制着权值空间中权值对应每步沿梯度方向变化的大小, 学习率较大时, 学习的速度比较快, 但有时可能产生振荡, 即误差总不能小于某个特别小的值, 而当学习率取得较小时, 学习速度较慢, 但一般比较平稳. 为了解

决这一问题,人们想到在训练过程中自动调整学习速率,通常调节学习速率的准则是:检查权值的修正值,看其是否真正降低了误差函数,如果确实如此,则说明所选的学习速率值小了,可对其再增大一个量,若不是这样,即产生了过调,那么应减小学习速率的值,这样,自适应调整学习速率能有效地缩短训练时间.

文中采用将附加动量项与自适应学习率相结合的改进 BP 算法^[6],它在增加动量项的同时,也对学习速率进行调节,这种算法可以综合附加动量法与自适应学习速率法的优点,能够有效地提高传统 BP 算法的训练速度和识别精度.

4 基于不变矩和 BP 网络的目标图像识别

4.1 网络训练样本

用 A、B、C、D 四个字符代表需要识别的四个目标,对四个字符在(−90°~ 90°)范围内间隔 10° 旋转,得到 76 个旋转图像,再对每个旋转目标分别进行 0.8、1.0、1.2、1.4 倍尺度缩放,这样一共得到 304 个样本图像,分别提取其不变矩,得到神经网络的训练样本.

4.2 网络参数设置

实验中采用单隐层的三层神经网络模型,网络设置中,输入层 7 个神经元节点,对应 7 个不变矩参数;输出层 4 个神经元节点,分别对应 4 个目标. 隐含层单元数的选取与识别要求、输入输出节点的多少都有直接的关系. 一般来说,隐含层的神经元越多,结果越精确,但是隐含层的神经元过多会大大增加训练时间,并会造成网络的容错能力下降;神经网络隐含层的神经元个数过少,会使迭代次数增多,权值来回调整,不易稳定,导致训练时间过长^[7]. 因此,权衡以上因素,文中隐含层神经元的个数选取为 25 个.

网络中间传递函数为 tansig 和 logsig 函数,训练的目标误差为 0.002,动量常数为 0.8,初始自适应学习率为 0.01,学习率增加的比率为 1.05,减少的比率为 0.7.

4.3 网络训练及目标识别结果

将训练样本送入设置好的 BP 网络中进行训练学习,直到平方和误差下降到目标误差以下时迭代停止,训练过程如图 2、图 3 所示,改进 BP 算法迭代了 1 733 次达到了目标误差的要求.

为了测试网络的鲁棒性和识别精度,将 5 个测试样本输入网络中,样本如图 4 所示:分别是对 A

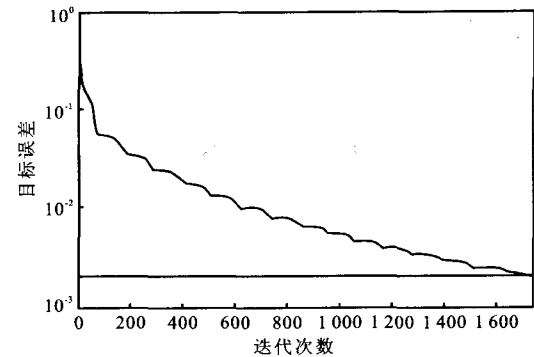


图 2 BP 神经网络训练过程

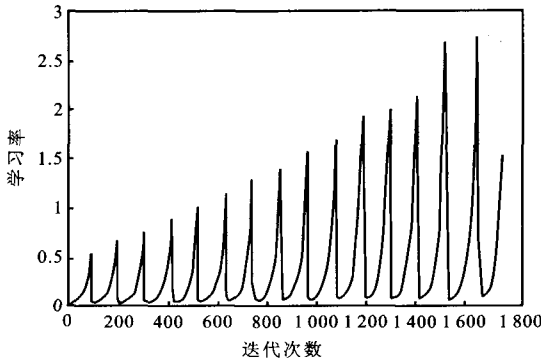


图 3 学习率自适应调整过程

字符旋转 25°、缩放 1.1 倍;B 字符旋转 125°、缩放 1.3 倍;C 字符旋转 195°、缩放 0.7 倍;D 字符旋转 305°,缩放 0.9 倍;E 字符旋转 0°,缩放 1 倍,其识别输出结果如表 2 所示.

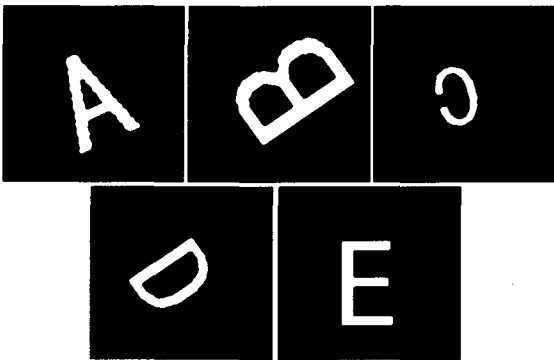


图 4 测试字符样本图像
表 2 网络识别输出结果

样本	旋转	尺度	网络识别结果			
A	25°	1.1	0.940	0.026	0.018	0.057
B	125°	1.3	0.029	0.957	0.013	0.026
C	195°	0.7	0.016	0.011	0.968	0.024
D	305°	0.9	0.038	0.029	0.061	0.867
E	0°	1	0.002	0.440	0.412	0.000

(下转第 69 页)

出,即使是物体在图像中没有完整出现,或者阴影只显示了一部份,本文算法都能达到很好的阴影检测效果.这是因为本文算法并不需要像基于模型的方法那样,对物体的形状、大小和位置、光照的方向等信息有事先的了解,这提高了算法的鲁棒性.

4 结束语

本文提出了一种基于组合特征和 HSI 颜色空间的阴影检测算法.首先通过扩展的不变矩变换和 Gabor 小波变换分别获得待识别区域的全局特征和局部特征来构成组合特征向量;然后通过 HSI 阴影颜色模型,对阴影区域进行第二次分析;最后通过连通区域标记和形态学操作,并再通过误检测区域分析,确定最终的阴影检测结果.实验结果表明,本算法具有较高的阴影检测效果.

在本文算法中,扩展的不变矩变换和 Gabor 小波变换的运算量相对较大.再就是,对全局特征和局部特征所起的作用需要根据实际情况区别对待而给予不同的权重.下一步的研究将改进该算法,以实现阴影更快更准确的检测和识别.

参考文献:

- [1] Song K, Tai J. Image-based traffic monitoring with shadow suppression [C]// Proceedings of the IEEE.

(上接第 65 页)

由表 2 中粗体部分可以看出,BP 网络能准确地对 4 个变化的字符目标进行分类识别,并能排除非目标 E 字符图像的干扰.大量的测试样本实验表明,BP 网络对 4 个字符目标在旋转($0^{\circ}\sim 360^{\circ}$)、尺度缩放(0.7~1.5)之间的识别率可达 95.5%.

5 结束语

相对其他矩特征提取算法,Hu 不变矩算法较为简单,便于实现,但其矩值易受目标背景噪声的干扰而变的不稳定,因此,实际应用中目标图像的除噪预处理是有必要的.采用附加动量项和学习速率自适应调整相结合的改进 BP 算法,能有效地提高网络的训练速度和识别精度.实验结果表明,这种将不变矩特征与神经网络算法相结合的目标识别方法对平移、旋转、尺度缩放变化目标图像的识别是有效的、可行的.

New York: IEEE, 2007, 95(2): 413-426.

- [2] Leone A, Distant C, Ancona N, et al. Texture analysis for shadow removing in video-surveillance systems [C]// Proceedings of IEEE International Conference on Systems, Man and Cybernetics. Hague, Netherlands: IEEE, 2004: 6325-6330.
- [3] 赵永锋,王毅刚.采用手指的三维人机交互方法[J].微电子学与计算机,2009,26(6):21-24.
- [4] 张哲源,史颖欢,高阳.基于色彩和形态特征的人阴影检测方法[J].微电子学与计算机,2009,26(10):86-88.
- [5] Hu M K. Visual pattern recognition by moment invariants[J]. IEEE Trans on Information Theory, 1962(8): 179-187.
- [6] Li Zongmin, Hou Kunpeng, Liu Yujie, et al. The shape recognition based on structure moment invariants[C]// The 2005 International Conference on Intelligent Computing. Hefei: ICIC, 2005.
- [7] Manjunath B S, Ma W Y. Texture features for browsing and retrieval of image data [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1996, 18(8): 837-842.

作者简介:

李子龙 男,(1979-),博士生,讲师.研究方向为图像处理、模式识别与目标跟踪.

刘伟铭 男,(1963-),教授,博士生导师.研究方向为智能交通、图像处理.

参考文献:

- [1] 孙宝琛,时银水,朱岩.基于模糊神经网络的目标识别[J].电光与控制,2005,12(3):50-54.
- [2] 张蕾,普杰信.一种基于改进 BP 网络的物体识别方法[J].微电子学与计算机,2008,25(4):153-155.
- [3] 任毅.用于激光雷达图像识别的不变矩特征向量研究[D].哈尔滨:哈尔滨工业大学,2007.
- [4] Hu M K. Visual pattern recognition by moment invariant[J]. IRE Trans Information Theory, 1962, 18(2): 179-187.
- [5] 焦李成.神经网络的应用与实现[M].西安:西安电子科技大学出版社.1993.
- [6] 周开利,康耀红.神经网络模型及其 MATLAB 仿真程序设计[M].北京:清华大学出版社,2005.
- [7] 汤茂斌,谢渝平.基于神经网络算法字符识别方法研究[J].微电子学与计算机,2009,26(8):19-93.

作者简介:

孙红辉 男,(1975-),博士,讲师.研究方向为图像处理与模式识别.