

·信号与信息处理·

## 基于不变矩和改进BP神经网络的目标识别

文华荣<sup>1</sup>, 李 秩<sup>2</sup>, 冯 一<sup>3</sup>, 吴晓迪<sup>3</sup>

(1.西藏军区指挥自动化站, 西藏 拉萨 850000; 2.73677部队, 江苏 南京 210000; 3.电子工程学院, 安徽 合肥 230037)

**摘要:**基于Hu不变矩的尺度不变性,以图像的不变矩特征作为输入,建立基于批训练的改进型误差反向传播(BP)神经网络。运用基于Bayesian正则化的Levenberg-Marquardt算法优化误差函数计算精度,改进网络,实现参数最优化组合。通过MATLAB环境,建立了基于不变矩的改进BP神经网络目标识别模型。实验表明,该方法实现了对目标的准确识别和对干扰图像的正确判断。

**关键词:**不变矩;BP神经网络;目标识别;Levenberg-Marquardt算法

中图分类号:TP391

文献标识码:A

文章编号:1673-1255(2013)-04-0049-06

## Object Recognition Based on Invariant Moment and Improved Back Propagation Neural Network

WEN Hua-rong<sup>1</sup>, LI Zhi<sup>2</sup>, FENG Yi<sup>3</sup>, WU Xiao-di<sup>3</sup>

(1. Tibet Military Command Automation Station, Lhasa 850000, China; 2. 73677 Army Unit, Nanjing 210000, China; 3. Electronic Engineering Institute, Hefei 230037, China)

**Abstract:** Based on the geometrical invariability of Hu invariant moment, and taking the characteristics of image invariant moment as an input, an improved error back propagation (BP) neural network for batch training is established. According to Bayesian normalization Levenberg-Marquardt algorithm, the calculation accuracy of error functions is optimized and the network is improved so as to realize the optimized parameter combination. An improved BP neural network object recognition model based on invariant moment is built in MATLAB environment. Experimental results show that accurate target recognition and correct interferential image estimation are implemented with the method.

**Key words:** invariant moment; back propagation (BP) neural network; target recognition; Levenberg-Marquardt algorithm

对于传统的目标识别方式,由于受环境、视角、时间以及成像手段等因素的影响,导致在目标高速运动过程中获取的实时图像与基准图像之间可能产生几何失真,再加上数据运算量大,难以满足实际应用需要。

文献[1-5]对修正的Hu不变矩进行了研究,发现其不论在连续状态还是在离散状态下都对平移、缩放、旋转具有尺度不变性,可以应用于图像识别领域;神经网络技术具有并行信息处理等能力,实时性

和鲁棒性较强,可以有效地提高系统响应速度和目标识别率,在模式识别方面的应用研究日益活跃起来。基于这两点,文中提出了一种结合不变矩和改进BP神经网络的目标识别方法,提取目标图像的七个不变矩作为训练样本,建立基于批训练的改进型BP神经网络模型,进而实现目标图像的识别和分类。实验结果表明,该不变矩特征库训练的神经网络分类器目标识别精度较高,具有较好的分类效果。

## 1 目标图像的不变矩

最早的不变量概念是在19世纪初 Cayley、Salmon、Clebsch 等人提出来的代数不变量。Hu 于1961年首先提出了基于直角坐标系的规则矩的概念。之后,他运用规则矩的非线性组合,得到一套具有尺度、平移和旋转不变性的矩不变量。

一幅数字图像  $f(x, y)$  的二维  $(p+q)$  阶矩和中心矩分别定义为

$$m_{pq} = \sum_x \sum_y x^p y^q f(x, y) \quad (1)$$

$$\mu_{pq} = \sum_x \sum_y (x - \bar{x})^p (y - \bar{y})^q f(x, y) \quad (2)$$

其中,  $f(x, y)$  是图像在  $(x, y)$  处灰度值,  $\bar{x} = m_{10}/m_{00}$ ,  $\bar{y} = m_{01}/m_{00}$  是图像重心坐标。

定义归一化  $(p+q)$  阶中心矩为

$$\eta_{pq} = \mu_{pq} / \mu_{00}^\gamma \quad (3)$$

其中,  $p, q = 0, 1, 2, \dots$ ;  $\gamma = (p+q)/2 + 1$ 。

对平移、缩放、镜像和旋转都不敏感的七个二维不变矩的集合可以由下列公式推导出来,它们为

$$\begin{aligned} \Phi_1 &= \eta_{20} + \eta_{02} \\ \Phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \Phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ \Phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \Phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ & \quad (3\eta_{21} - \eta_{03}) \times (\eta_{21} + \eta_{03}) [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (4)$$

$$\begin{aligned} \Phi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + \\ & \quad 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \Phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - \\ & \quad 3(\eta_{21} + \eta_{03})^2] + \\ & \quad (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03}) \times [3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{aligned} \quad (5)$$

由于不变矩的变化范围很大,为了便于比较,可利用取对数的方法进行数据压缩;同时考虑到不变矩有可能出现负值的情况,因此,在取对数之前先取绝对值。实际采用的不变矩为

$$\Phi_k = \lg|\Phi_k| \quad (k=1, 2, \dots, 7) \quad (6)$$

为验证不变矩对于图像平移、旋转和缩放的尺度不变性,文中选取图1所示的尺度变换后的目标图像与干扰图像,进行了特征值计算,结果如表1所示。

分析表1可得出两点结论:第一,不变矩对缩放、

旋转等具有不变性,但只是一种近似的不变性,理论表明矩精确不变性仅仅在模拟图像上才可能体现,所以图数字化虽然有利于计算机处理,提高处理效率,但也增加了数字误差,这也会影响图像识别率;第二,不同目标的组合矩不变量具有较大的差别,可对不同目标进行分类识别,即具有可分性。因此,引入神经网络理论,将不变矩作为目标图像的特征量,输入神经网络分类器,具有可行性和必要性。

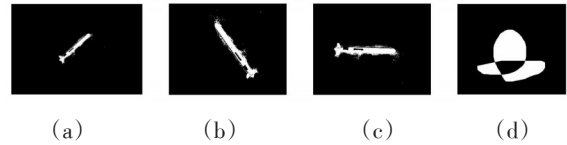


图1 目标与干扰图像示例

表1 目标与干扰图像特征提取结果

Num	a	b	c	d
$\Phi_1$	5.64	5.65	5.63	7.32
$\Phi_2$	11.37	11.37	11.33	16.77
$\Phi_3$	22.46	23.27	23.06	38.25
$\Phi_4$	23.13	22.89	22.66	42.08
$\Phi_5$	46.39	46.03	45.52	82.31
$\Phi_6$	29.23	28.64	28.34	50.56
$\Phi_7$	46.18	47.13	48.54	63.85

## 2 BP神经网络模型的改进

### 2.1 传统BP神经网络

BP网络(back propagation network)是当前工程应用最广泛的一种人工神经网络,包括三层前馈网络,即输入层、隐含层和输出层,相邻层之间的各神经网络实现完全连接,连接强度构成网络的权值矩阵。它的最大特点是仅仅借助样本数据,无需建立系统的数学模型,就可对系统实现输入神经元的模式向量空间到输出的高度非线性映射<sup>[6]</sup>。其单元结构如图2所示。

图2中  $p_i (i=1, 2, \dots, n)$  表示输入层中的输入向量,  $W_{i,j}$  表示输入层与隐藏层或隐藏层与输出层连接的权值,  $b$  为单个神经元的阈值,  $f$  是激活函数,  $a$  为单个神经元的输出。用公式可表示为

$$a = f(WP + b) \quad (7)$$

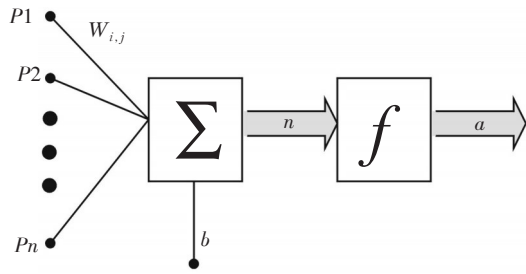


图2 单个神经元结构

根据误差的反向传播方式,利用梯度下降的方法,不断地修改层与层之间的连接权值,寻找权值的最佳值来减小目标输出和实际输出之间的误差。其误差的计算方法为

$$E = \sum_I (\sum_S (d_{s,i} - a_{s,i})^2) \quad (8)$$

其中,  $d_{s,i}$  是样本  $s$  在神经元  $i$  的期望输出,  $a_{s,i}$  是样本  $s$  在神经元  $i$  的实际输出,  $I$  为输出层中神经元节点的序列,  $S$  是样本个数。

## 2.2 基于批训练的BP神经网络

在传统BP算法中,每输入一个样本,都要回传误差并调整权值,这种对每个样本轮训的权值调整方法又被称为单样本训练<sup>[7]</sup>。由于单样本训练只针对每个样本产生的误差进行调整,使整个训练次数增加,导致收敛速度过慢。而基于批训练的BP神经网络是在所有样本输入之后计算网络的总误差,然后根据总误差计算各层的误差信号并调整权值

$$E = \sum_{k=1}^p E_k \quad (9)$$

式中,  $p$  为样本总数。

## 2.3 基于 Bayesian 正则化的 Levenberg-Marquardt 算法

Bayesian 正则化的 Levenberg-Marquardt 算法(简称 L-M 算法)是避免神经网络出现过拟合并提高精度的有效方法<sup>[8-10]</sup>。L-M 算法使用 Quasi-Newton 算法来修正网络误差。首先引入 Hessian 矩阵

$$H = J^T J \quad (10)$$

式中,  $J$  为雅克比行列式矩阵,该矩阵包含所有网络误差及其相应的权重、阈值的初阶倒数。

于是得到

$$\Delta S = H^{-1} J^T E \quad (11)$$

式中,  $\Delta S$  为网络权重、阈值的修正值。牛顿算法中, Hessian 矩阵一般不正定,因此,在修正迭代计算中,牛顿方向可能是指向局部极大点或者某个鞍点。

L-M 算法通过在 Hessian 矩阵上加一个小正数,使 Hessian 矩阵成为正定矩阵

$$\Delta S = (J^T J + \mu I)^{-1} J^T E \quad (12)$$

式中,  $\mu$  为小正数,  $I$  为单位阵。显然,当  $\mu=0$  时,式(13)就是 Quasi-Newton 算法;当  $\mu$  趋于无穷大时,  $\Delta S$  是以小步长急剧下降的梯度下降算法。当  $\Delta S$  达到最小或者满足某一标准时,网络训练完成。

采用上述改进后,牛顿算法会更快更准确地接近误差最小值。Bayesian 正则化的引入是假设网络的权重和偏差都具有某种确定性分布,而正则化参数与这些未知变量及其对应的分布具有一定关联。因而可以使用统计学方法来估计这些参数。

构造 BP 网络误差方程

$$F(\omega) = \beta E_{sse} + \alpha E_{ssw} \quad (13)$$

式中,  $E_{sse}$  为误差平方和,  $E_{ssw}$  为网络所有权重、阈值的平方和,  $\alpha$  和  $\beta$  为目标函数参数。当  $\beta > \alpha$  时,则训练算法使网络误差较小,当  $\beta < \alpha$  时,则训练强调权值的减小,而网络误差较大。最终使得神经网络的输出更加平滑,防止出现极小点。

确定网络输入后,根据 Bayesian 公式可以得到网络权值的概率密度函数为

$$P(\omega | D, \alpha, \beta, M) = \frac{P(D | \omega, \beta, M) P(\omega | \alpha, N)}{P(D | \alpha, \beta, M)} \quad (14)$$

式中,  $\omega$  为网络权值,  $D$  为数据集,  $M$  为网络模型,  $N$  为神经网络参数总数。

假设标准化因子  $P(D | \alpha, \beta, M)$  与权值向量无关,且数据中的噪声和权向量均服从高斯分布,由式(14)最终可得到网络的最优目标函数参数为

$$\alpha^* = f_1(r, \omega_{\min}) = Nr / (2E_{sse} \omega_{\min}) \quad (15)$$

$$\beta^* = f_2(r, \omega_{\min}) = r / (2E_{ssw} \omega_{\min}) \quad (16)$$

式中,  $r$  为有效网络参数数目,  $\omega_{\min}$  为  $\omega$  的最小值。由此可得具有  $E_{sse}$  和  $E_{ssw}$  最优组合的  $F(\omega)$ 。

在 BP 神经网络模型运算过程中先由 L-M 分别计算网络权重、阈值,再由 Bayesian 对已计算出来的权重、阈值及误差最小值进行优化组合。两者结合使得改进后的 BP 网络可以更好地逼近输出端数据,同时大幅度地提高了稳定性。

### 3 基于图像矩特征的神经网络模型的建立

神经网络模型的确定是通过构建样本集和确定网络结构中各个参数来实现的。参数主要包括神经网络的层数、各层的结点数、激活函数、各层间权值和学习率等等。由于文中选用学习率自适应调整的梯度下降动量法对网络进行训练,因此,学习率无需确定。

#### 3.1 构建样本集

样本集是建立神经网络的基础,样本选择的优劣对神经网络的识别效率有重要的影响。为了更好地实现样本的识别,文中依据下列准则选择样本。

首先,训练样本要尽量多。只有足够多训练样本才能提取目标更多本质特征,才能使网络得到充分的训练,才能得到最佳的权值进行具有高识别率的图像识别。可以从不同角度、不同距离、不同姿态等方面采集目标,一方面丰富了样本,另一方面保证了使采集后的图像尽可能多包含目标信息。其次,进行训练时,对不同类型的干扰图像以交叉方式输入。在前向反馈神经网络的训练过程中,前面的权值的改变影响后边的权值调整,后面样本权值会对前面样本权值进行修正。当同一类样本过于集中时,后边的不同目标样本会对前面一类目标样本建立的输入输出映射产生重大的修正,致使网络训练仅仅对最后一类集中样本敏感,而对前面的训练样本不具有敏感作用。第三,不同形状的干扰图像的样本数目相等。这样能保证各种姿态的图像能得到充分均等训练,避免不均匀的权值记忆。

在实验中,共采集了经旋转、缩放、平移等处理的目标、干扰图像共480张,基本可以满足识别精度的要求。

#### 3.2 确定神经网络的输入节点

神经网络输入节点的选择对识别效率有很大的影响。在实际中神经网络的输入节点与输入到神经网络的数据源的维数相等,此时神经网络的输入神经源相当于计算机的缓冲存储器,它存储数据源发出的数据。由于文中基于图像特征进行识别,神经网络的输入节点数应等于提取出的图像特征的维数,所提取的图像特征是七个特征不变矩,因而输入节点数目就是七个。

#### 3.3 神经网络输出节点的确定

常用确定输出节点数的方法有两种。一种就是待识别图像目标种类比较少时,直接把待识别图像目标的种类数作为神经网络的输出节点数。输出节点用长度为输出结点数,元素值仅有0和1的向量来表示。第二种方法借鉴了计算机技术的二进制方法。若待识别图像种类比较多并且种类数为 $n$ 的整数次方个,输出节点个数就为 $\log_2 n$ ,其表示方式用二进制编码法来表示。由于文中待识别图像分为目标和干扰图像两种,因此只有一个神经网络的输出节点,当图像为目标时输出1,为干扰时输出0。

#### 3.4 神经网络层数的确定

在实际中网络层数一般低于四层,四层以上的神经网络用的比较少。神经网络的选择和神经网络的输入输出和激活函数有关系。根据Cybenko的证明,当各节点都采用S型激活函数时,一个隐含层就足够实现任意的分类识别问题,两个隐含层足以表示输入图像的任意输出。由于文中是用神经网络实现图像识别分类的,所以根据这一结论,一个隐含层,即三层即神经网络就足够实现图像识别。

#### 3.5 神经网络激活函数的确定

神经网络各层神经元的输入输出映射是通过激活函数来实现的。文中采用的是S型函数: $f(x)=1/[\exp(-x)]$ ,此函数单调递增,当 $x$ 趋于正负无穷时,分别达到最大值和最小值1和0,即该函数输出区域为 $[0,1]$ ,与期望输出数值范围一致。

#### 3.6 神经网络隐含层神经元数的确定

目前还没有成熟的理论,指导如何确定隐含层的神经元数量,但一些研究者通过研究给出了一些经验公式如下。

$$(1) \sum_{i=0}^n C_n^i > k, \text{ 其中, } k \text{ 为样本数, } n_i \text{ 为隐层神经元数, } n \text{ 为输入单元数。如果 } i > n_i, C_n^i = 0。$$

$$(2) n_1 = \sqrt{n+m} + a, \text{ 其中, } m \text{ 为输出神经元数, } n \text{ 为输入单元数, } a \text{ 为}[1, 10]\text{之间的常数。}$$

$$(3) n_1 = \log_2 n, \text{ 其中, } n \text{ 为输入单元数。}$$

文中在借鉴上述公式的基础上,经过多次实验调试,最后采用以下方法确定隐层神经元数:使隐层神经元数可变,通过学习将那些不起作用的隐单元



去除,直到不可收缩为止,最后确定隐层神经元数目为4。并成功建立了基于不变矩和改进BP神经网络的目标识别模型,其算法流程图如图3所示。

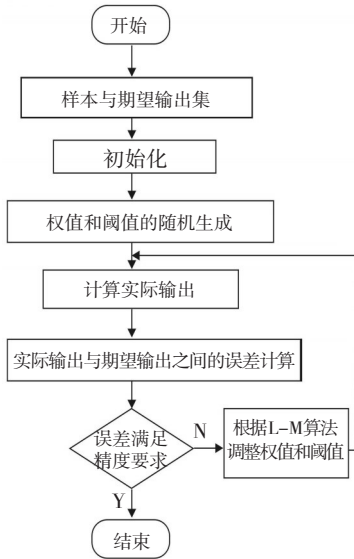


图3 文中算法流程图

### 4 实验结果与分析

#### 4.1 神经网络训练结果

在MATLAB软件环境下,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。图4、图5所示为一次训练过程,在该次训练中,改进的BP算法迭代了5318次,达到了目标误差0.001的要求。通过多次训练实验结果统计分析,因训练样本量较大,完成训练的平均时间大约为90~110 s。

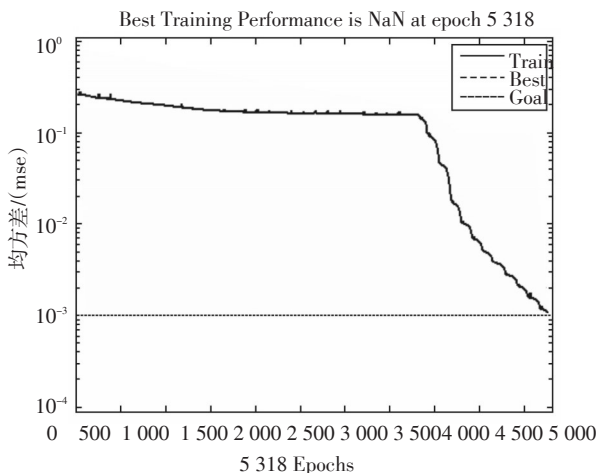


图4 误差函数曲线

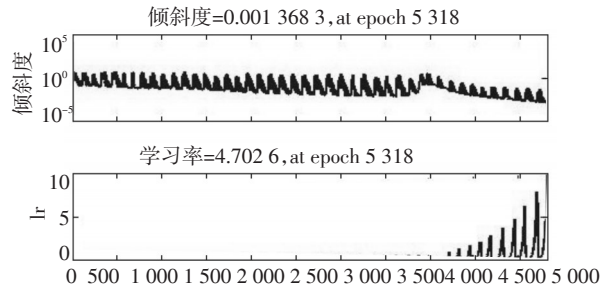


图5 训练倾斜度和学习率自适应曲线

#### 4.2 目标识别实验结果

为验证训练结果,从样本集中随机抽取目标和干扰图像各60幅进行仿真。结果表明,实验建立的神经网络对目标识别的输出平均值为0.975 1,对干扰图像的输出平均值为0.026 3。达到了预期的识别效果。图6为部分实验图像,图6a~图6d是目标图像,图6e~图6h是干扰图像,表2给出了它们对应的识别结果。

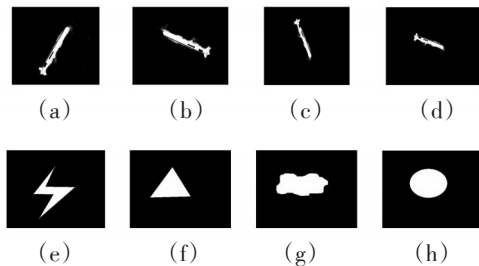


图6 部分测试图像

表2 神经网络识别输出结果

序号	期望输出	实际输出	平均值
A		0.970 1	
B		0.977 1	
C	1	0.974 8	0.973 6
D		0.977 1	
E		1.190 0e-004	
F		1.481 4e-004	
G	0	0.041 2	0.027 1
H		0.041 6	

通过表2可以看出,仿真结果与期望输出结果拟合度较高,达到了准确识别的目的。其中,对于干扰图像闪电图6e、三角图6f的识别精度明显高于云彩图

6g、圆盘图6h,表明干扰外形特点与目标轮廓差异越大,识别精度越高,这也进一步说明文中采用Hu不变矩作为特征量的科学性。

## 5 结论与展望

围绕不变矩的尺度不变性进行目标辨识研究,结合BP神经网络可并行信息处理的特点,运用批训练理论优化训练速度,运用基于Bayesian正则化的L-M算法优化误差函数计算精度,在MATLAB软件环境下成功建立基于不变矩的改进BP神经网络目标识别模型。实验结果表明,该模型对缩放、旋转、平移的目标和干扰图像能够正确分类,识别精度较高。

同时应该看到,文中选取的目标图像形状比较简单,背景单纯,没有考虑真实环境中存在的噪声干扰、运动模糊等不利因素,而且L-M算法改进后的神经网络模型较难兼顾计算精度与运算速度,迭代次数偏多。如何克服这些不足,是下一步研究的重点。

## 参考文献

- [1] HU M K. Visual pattern recognition by moment invariants[J]. IEEE Transactions on Information Theory, 1962, 8(2): 179-187.
- [2] WONG W H, SIU W C. Improved digital filter structure for fast moment computation[J]. IEEE Proceedings in Vision, Image and Signal Processing, 1999, 146:73-79.
- [3] DUDANI S A, BREEDING K J, MCGHEE R B. Aircraft identification by moment invariants[J]. IEEE Transactions on Computers, 1983, 26(1):39-45.
- [4] LIAO S X, PAWLAK M. On the accuracy of Zernike moments for image analysis[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1998, 20 (12): 1358-1364.
- [5] ZION B, SHKLYA A, KARPLUS I. In-vivo fish sorting by computer vision[J]. Aquaculture Engineering, 2000, 22: 165-179.
- [6] 肖潇,杨国光,白剑.基于最优参数的全景图像三次样条差值复原[J]. 红外与激光工程, 2007, 36(5):725-728.
- [7] 董长虹. Matlab神经网络与应用[M].北京:国防工业出版社, 2005:278-280.
- [8] 蒋天一,胡德金,许开州.改进型BP神经网络对球面磨削最高温度的模拟与预测[J]. 上海交通大学学报, 2011, 45(6):901-906.
- [9] 闫滨. 基于L-M算法的BP神经网络在大坝安全监控预报中的应用[J]. 沈阳农业大学学报, 2009, 40(4): 506-509.
- [10] 项文强. 基本L-M算法的BP网络在变压器故障诊断中的应用[J]. 电力系统保护与控制, 2011, 39(8):100-103.

### 欢迎利用期刊网站浏览本刊已发表文章

为了满足读者对《光电技术应用》期刊文章的快速、方便、阅读需求,《光电技术应用》期刊网站(网址为:<http://www.gdjsyy.com>),为读者提供了《光电技术应用》期刊2009~2012年已发表文章的在线浏览。读者可在过刊目录下,查阅2009~2012年期刊各期目录,点击文章题目或摘要,阅读文章全文。欢迎广大读者登陆期刊网站,及时了解《光电技术应用》期刊已发表文章的最新信息。

《光电技术应用》编辑部