

基于 RANSAC-LSSVM 回归的惯性约束聚变 光学元件损伤在线检测技术

冯 博¹ 陈凤东¹ 刘炳国¹ 彭志涛² 刘 楠^{1,2} 刘国栋¹ 孙和义¹

(¹哈尔滨工业大学电气学院, 黑龙江 哈尔滨 150001
²中国工程物理研究院激光聚变研究中心, 四川 绵阳 621900)

摘要 针对惯性约束聚变(ICF)光学元件损伤问题,提出了一种基于随机抽样一致性(RANSAC)及最小二乘支持向量机(LSSVM)回归的高精度检测方法。建立了损伤区域总灰度与实际尺寸的回归模型,通过该回归模型对待检测损伤区域的尺寸进行预测,得到损伤区域的高精度尺寸。为剔除回归模型建立过程中离群样本点的影响,采用RANSAC方法对训练样本进行优化选择。针对抽样组中样本数对检测精度及检测效率的影响进行了相关实验,确定了抽样组中样本数的合适区间。RANSAC-LSSVM方法可通过改变误差评价函数得到不同评价体系下的最优回归模型。实验证明,在传统像素级检测方法的基础上,该方法将损伤尺寸检测的平均相对误差降低了近90%。

关键词 图像处理; 光学元件损伤; 在线检测; 回归; 最小二乘支持向量机; 随机抽样一致性

中图分类号 TN247 **文献标识码** A **doi**: 10.3788/CJL201340.0808002

RANSAC-SVM Based Online Inspection Technique for Optics Damage in Inertial Confinement Fusion System

Feng Bo¹ Chen Fengdong¹ Liu Bingguo¹ Peng Zhitao² Liu Nan^{1,2}

Liu Guodong¹ Sun Heyi¹

¹Department of Automatic Measurement and Control, Harbin Institute of Technology,
Harbin, Heilongjiang 150001, China

²Laser Fusion Research Center, China Academy of Engineering Physics, Mianyang, Sichuan 621900, China

Abstract A high precision method, based on random sampling consensus (RANSAC) and least squares support vector machine (LSSVM) regression, is advanced to improve the inspection for inertial confinement fusion (ICF) optics damage. In present study, the measurement precision of damaged areas is improved significantly, via analyzing the regression model of grey value and size. RANSAC is adopted for optimizing the training samples to avoid the effect of outliers. In addition, special experiment is proceeded to assess the effect of sample size on the measurement precision and the efficiency, and the appropriate interval of sample size is obtained as well. RANSAC-LSSVM can get optimal regression models which is adopted in different evaluative systems, by using different error evaluation functions. The results show that RANSAC-LSSVM reduces the mean relative error by 90% approximately, in measuring of damaged areas, compared with traditional pixel-level detection.

Key words imaging processing; optics damage; online inspection; regression; least squares support vector machine; random sampling consensus

OCIS codes 100.2000; 150.1835; 330.3350; 350.1820

收稿日期: 2013-03-12; 收到修改稿日期: 2013-04-08

基金项目: 国家自然科学基金(51275120,61275096)、航空科学基金(20120177004)、中央高校基本科研业务费专项资金(HIT.NSRIF.2013012)

作者简介: 冯 博(1983—),男,博士研究生,主要从事精密视觉测量技术、传感器及信号处理等方面的研究。

E-mail: yy_llfeng@163.com

导师简介: 刘国栋(1974—),男,教授,博士生导师,主要从事光电信息及光电检测技术,传感器及信号处理等方面的研究。E-mail: lgd @ hit. edu. cn(通信联系人)

1 引言

惯性约束聚变(ICF)系统内集成了数量庞大的大口径光学元件。由于光束自聚焦等原因,光学元件的表面及内部会产生激光损伤^[1]。为保证系统安全运行、降低运行成本,需要在损伤形成的初始阶段就对其进行探测和跟踪,在损伤尺寸达到一定阈值前及时维修^[2]。传统离线检测方法^[3-4]由于需要将待检测光学元件逐一拆卸,其效率无法满足 ICF 系统的检测要求,因此目前普遍采用基于机器视觉的在线检测方法^[5]。

在 ICF 系统中,某些光学元件的损伤区域需要在尺寸超过 $50\ \mu\text{m}$ (目前一般采用等价圆直径,即与损伤区域等面积圆的直径作为损伤尺寸的评价单位)时被检测出来, $170\ \mu\text{m}$ 以上损伤区域需要被定量检测^[5],仅依靠硬件实现上述检测分辨率十分困难。另外,由于跨像素成像、色差等问题,损伤区域在 CCD 上的像素尺寸往往要比其理论尺寸大,这些误差对于损伤检测的精度影响很大。

最小二乘支持向量机(LSSVM)^[6-7]是支持向量机(SVM)^[8]的一个改进版本,因其准确率高、泛化性强等优点,目前被广泛应用于各个领域的分类、回归问题当中。随机抽样一致性(RANSAC)是最有效的离群点剔除算法之一^[9-10],对于高错误率样本仍然有很好的效果,被广泛应用于机器视觉等领域当中。以“神光-III原型装置”为平台,基于 RANSAC-LSSVM 回归原理提出了一种光学元件损伤在线检测方法,其核心思想是将一定数量损伤区域的精确尺寸与其在线检测图像中的灰度参数作为样本建立回归模型,检测时通过该模型对待检测损伤区域的尺寸进行预测,得到其亚像素级精确损伤尺寸。

2 算法原理

2.1 原理及数据获取

在“神光-III原型装置”光学元件损伤在线检测系统中,由于成像分辨率限制,仅通过传统像素级图像处理方法难以满足 $50\ \mu\text{m}$ 以上损伤定性检测、 $170\ \mu\text{m}$ 以上损伤精确定量检测的要求。另外,由于跨像素成像等原因,损伤区域在 CCD 上所成像一般大于理论尺寸,该误差对于小尺寸损伤区域的检测精度影响很大。某些尺寸小于成像系统像素当量的损伤区域在检测图像中甚至大于 1 pixel。图 1(a)与(b)分别为“神光-III原型装置”光学元件损伤在线检测图像中等价圆直径为 $252.9\ \mu\text{m}$ 及 $140.5\ \mu\text{m}$ 的损伤区域,检测系统的像素当量为 $170\ \mu\text{m}/\text{pixel}$ 。

传统像素级光学元件损伤图像处理对于损伤的自动识别及区域确定^[11-13]效果很好,但是通过图 1 可以发现,像素级检测算法不能得到在线检测图像中损伤的高精度尺寸。

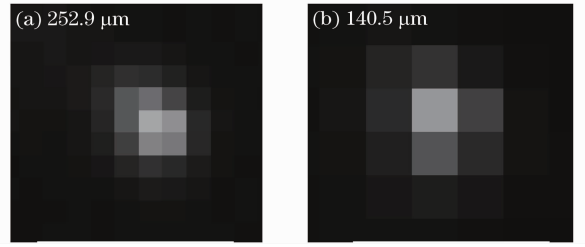


图 1 光学元件在线检测图像中的损伤区域

Fig. 1 Defects in optics online inspection image

为了获得高对比度的损伤图像,在“神光-III原型装置”中,光学元件在线检测系统采用内全反射照明^[14]。内全反射照明下,每一个损伤区域相当于一个朗伯型光源,辐射出的光通量 Φ 与其面积 S 近似成正比,体现在损伤图像上即为其总灰度

$$G = \sum_{i=1}^n g_i, \quad (1)$$

式中 n 为该损伤区域的像素个数, g_i 为像素点 i 的灰度值。于是可以定义损伤计算函数为

$$L = \Psi(G), \quad (2)$$

式中 $\Psi(G)$ 为损伤计算函数, L 为损伤区域尺寸(等价圆直径),损伤区域的总灰度 G 可以通过传统像素级图像处理方法获得,将 G 代入(2)式即可得到损伤区域的准确尺寸。

由于中间影响环节过多,损伤计算函数 $\Psi(G)$ 的精确代数表达式难以得到,本文采用回归的方法得到其近似模型。即通过一定数量损伤区域的离线、在线参数,直接标定出损伤区域的尺寸 L 与其检测图像中的总灰度 G 的关系为

$$L = f(G), \quad (3)$$

式中 $f(G)$ 为通过回归方法得到的 $\Psi(G)$ 的近似模型,建模训练样本的损伤尺寸 L 可以通过高精度离线检测系统获得^[15]。实际检测时,通过像素级图像处理算法得到损伤区域在检测图像中的总灰度 G ,将 G 代入到损伤检测模型(3)式中即可得到损伤区域的亚像素级预测尺寸 L 。实际检测时的照明、CCD 曝光时间等条件需调整到与建模样本获取时相同,不同光学元件分别建立损伤检测模型。

2.2 LSSVM 损伤检测模型

LSSVM 通过线性 Karush-Kuhn-Tucker (KKT)条件来代替支持向量机中二次规划问题的求解,大大提高了计算效率,目前被广泛应用于各个

领域的分类及回归问题当中。从检测效率考虑,采用基于 LSSVM 的回归方法建立损伤检测模型。

通过离线与在线系统得到样本集 $D = \{(x_1, y_1), \dots, (x_k, y_k), \dots, (x_N, y_N)\}$, $x_k, y_k \in \mathbb{R}$, 其中 x_k 为在线图像中各损伤区域的总灰度, y_k 为损伤区域在高精度离线系统下获得的高精度尺寸。于是回归问题被表述为约束优化问题

$$\min_{\omega, b, e} [J(\omega, e)] = \frac{1}{2} \|\omega\|^2 + \frac{\gamma}{2} \sum_{k=1}^N e_k^2, \quad (4)$$

subject to

$$y_k = \langle \omega, \varphi(x_k) \rangle + b + e_k, \quad k = 1, \dots, N, \quad (5)$$

式中 $\varphi(x_k)$ 为非线性映射函数, 将线性不可分问题转化为高维空间中的线性问题。 ω 为权值向量, γ 为误差控制变量, e_k^2 为误差平方, b 为偏置项。为得到目标函数的极小值, 定义拉格朗日函数为

$$L = \frac{1}{2} \|\omega\|^2 + \frac{\gamma}{2} \sum_{i=1}^N e_k^2 - \sum_{i=1}^N \alpha_k \{ \langle \omega, \varphi(x_k) \rangle + b + e_k - y_k \}, \quad (6)$$

式中 α_k 为拉格朗日乘子。根据 KKT 条件, 最优化问题可以被表示为

$$\begin{bmatrix} 0 & \mathbf{1}^T \\ \mathbf{1} & \langle \varphi(x_k), \varphi(x_l) \rangle + \gamma^{-1} I_N \end{bmatrix} \begin{bmatrix} b \\ \alpha \end{bmatrix} = \begin{bmatrix} 0 \\ \mathbf{y} \end{bmatrix}, \quad (7)$$

$$k, l = 1, \dots, N,$$

式中 $\alpha = [\alpha_1, \dots, \alpha_N]^T$, $\mathbf{y} = [y_1, \dots, y_N]^T$, $\mathbf{1} = [1, \dots, 1]^T$ 。于是得到最优回归模型为

$$f(x) = \sum_{k=1}^N \alpha_k K(x, x_k) + b. \quad (8)$$

根据样本特点, (8) 式中核函数选择径向基 (RBF) 函数

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma}\right), \quad (9)$$

式中 σ 为高斯核函数的宽度参数, 其最优值 σ_{opt} 与 (6) 式中误差控制变量 γ 的最优值 γ_{opt} 可通过交叉法获得。

2.3 损伤检测模型的 RANSAC 优化

由于光噪声、光学元件表面灰尘等原因, 建模样本中会存在一些离群点, 这些离群点对于回归模型的准确性会产生很大影响。RANSAC 是解决样本中大量离群点的一种稳健的方法。不同于一般迭代剔除离群点的方法, RANSAC 随机抽取小样本建立一系列的模型, 从中选出最优模型对样本进行离群点剔除。采用 RANSAC 对 LSSVM 的训练样本进行预处理, 提高了回归、预测的精度。

在样本集 D 中随机抽样一系列的小样本 $\{W_1(x_{1,1}, y_{1,1}, \dots, x_{1,I}, y_{1,I}), \dots, W_m(x_{m,1}, y_{m,1}, \dots, x_{m,I}, y_{m,I}), \dots, W_M(x_{M,1}, y_{M,1}, \dots, x_{M,I}, y_{M,I})\}$, $W_m \in D$ 。其中 M 为需要抽样的次数, I 为每组抽样所需要的样本数。抽样次数 M 需要保证在一定的置信概率 P 下至少有一组抽样中不含有离群点, 于是

$$M = \lg(1 - P) / \lg(1 - \epsilon^I), \quad (10)$$

式中非离群点概率 ϵ 的定义为

$$\epsilon = f/F, \quad (11)$$

式中 f 表示非离群点样本数, F 表示总样本数。样本的 ϵ 一般无法得到确切值, 但是可以通过对原始数据的分析得到其近似值。不同于一般 RANSAC 方法, 用于建立 LSSVM 模型的最小样本数 I 只与样本质量有关, 合适的 I 值需要通过实验获得。

综上所述, RANSAC-LSSVM 算法的建模流程如下:

1) 通过传统在线检测方法得到全部损伤区域的总灰度后, 随机抽样 N 个损伤区域, 在离线检测系统中得到其准确尺寸, 建立样本集 $D = \{(x_1, y_1), \dots, (x_k, y_k), \dots, (x_N, y_N)\}$, $x_k, y_k \in \mathbb{R}$ 。

2) 估算离群点概率 ϵ , 通过 (10) 式计算需要进行抽样的次数 M 。

3) 随机抽样 M 组小样本 $\{W_1(x_{1,1}, y_{1,1}, \dots, x_{1,I}, y_{1,I}), \dots, W_m(x_{m,1}, y_{m,1}, \dots, x_{m,I}, y_{m,I}), \dots, W_M(x_{M,1}, y_{M,1}, \dots, x_{M,I}, y_{M,I})\}$, $W_m \in D$ 。

4) 选取合适的误差评价函数 $Z(x)$ 。

5) 通过交叉验证法建立一系列 LSSVM 回归模型 $\{f(W_1), \dots, f(W_m), \dots, f(W_M)\}$ 。

6) 通过整体样本 D , 逐一验证回归模型并得到误差评价结果 $\{Z[f(W_1)], \dots, Z[f(W_m)], \dots, Z[f(W_M)]\}$, 其中 $Z_{\min} = \min\{Z[f(W_1)], \dots, Z[f(W_m)], \dots, Z[f(W_M)]\}$ 所对应的样本集 W_{\min} 即为一组无离群点样本集, 由该样本集建立的回归模型 $f(W_{\min})$ 即为最优回归模型。

误差评价函数 $Z(x)$ 根据检测要求确定, 常用的误差评价函数有均方误差 (MSE)、平均绝对误差 (MAE)、平均相对误差 (MRE)、平方相关系数 (R^2) 等。对于本文实验, 由于损伤区域尺寸跨度较大, 因此检测精度的评价采用平均相对误差^[16]

$$R_{\text{MRE}} = \frac{1}{N} \sum_{i=1}^N \left(\frac{|\hat{y}_i - y_i|}{y_i} \right), \quad (12)$$

式中 \hat{y}_i 为由回归模型预测的损伤尺寸, y_i 为损伤的实际尺寸。

3 实验结果与分析

针对 RANSAC-LSSVM 算法中抽样组样本数对检测精度及检测效率的影响进行了相关实验。如图 2 所示,总样本数量为 200,非离群点概率为 85%,设定置信概率为 99%。抽样组样本数从总样本的 5%到 30%,抽样间隔为 5%。实验结果表明,针对本组样本,随着抽样组中样本数的增加,MRE 大致呈现降低的趋势。在 5%~10%阶段 MRE 降低较大,说明 5%抽样不能体现出整体样本性质,所建立的回归模型存在较大误差。在 10%~25%抽样率阶段,MRE 稳定下降且变化稳定,说明该抽样率下的样本已经可以表征整体样本的性质,25%抽样率的情况下所建立的模型最优。30%抽样率相对于 25%抽样率 MRE 已经没有降低,说明 25%抽样率已经可以实现无离群点样本选取。

抽样次数决定了算法的检测效率。在 5%至 20%抽样率阶段,抽样次数增加较慢;在 20%至 30%抽样率阶段抽样次数增加较快。因此对于本组样本,抽样组样本数可以综合考虑检测精度及检测效率后在总样本的 10%~25%区间内选取。

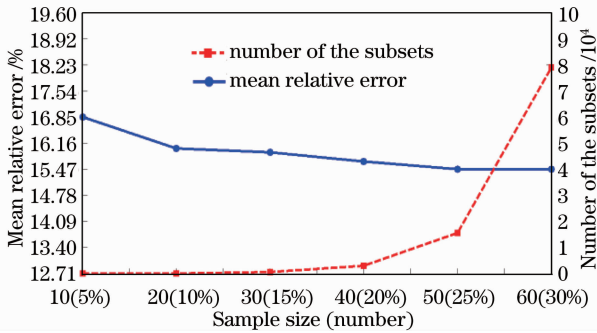


图 2 抽样组样本数对检测精度及检测效率的影响

Fig. 2 Effect of sample size on detection precision and the number of subsets

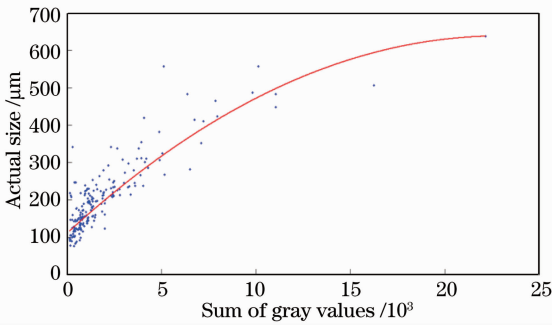


图 3 200 个训练样本点分布及 RANSAC-LSSVM 回归曲线

Fig. 3 200 training samples and the regression curve of RANSAC-LSSVM

图 3 为 200 个样本点分布以及建模样本取总样本 20%时,RANSAC-LSSVM 方法得到的回归曲线。可以看出,回归模型在全尺度范围可以对损伤区域的尺寸进行准确预测。

为了验证文中算法的检测精度,分别通过 200 个训练样本与 50 个测试样本对 RANSAC-LSSVM 方法与传统像素级方法中性能较优的“局部信号强度算法(LASS)^[13]进行了实验对比,实验结果如表 1 所示。可以看出,相对 LASS 方法,该算法将损伤尺寸检测的 MRE 降低了近 90%。

表 1 RANSAC-LSSVM 方法与传统像素级方法实验结果对比

Table 1 Performance comparison of RANSAC-LSSVM and traditional method

Samples	Evaluation function	LASS	RANSAC-LSSVM
Training samples	MRE	105.72%	15.44%
	MSE	38937	2198
Test samples	MRE	106.03%	17.99%
	MSE	46465	1963

采用的误差评价体系为平均相对误差,如果需要其他误差评价体系下获得最优回归模型,可以改变算法中的误差评价函数。表 2 为分别采用 MRE、MSE、 R^2 误差评价函数得到的结果。实验证明,通过应用不同误差评价函数,RANSAC-LSSVM 方法可以得到不同误差评价体系下的最优回归模型。该方法推广到其他回归、预测的应用领域当中同样可以得到很好的结果。

表 2 不同误差评价函数下 RANSAC-LSSVM 算法实验结果对比

Table 2 Test results of RANSAC-LSSVM with different error evaluation functions

Methods (evaluation function)	MRE	MSE	R^2
RANSAC-LSSVM (MRE)	15.54%	2198	0.79
RANSAC-LSSVM (MSE)	17.46%	1916	0.80
RANSAC-LSSVM (R^2)	16.62%	2001	0.82

4 结 论

提出了一种基于 RANSAC-LSSVM 回归的光学元件损伤在线检测方法,通过回归、预测的方式实现了损伤的高精度检测。对整体样本随机抽样得到一系列的小样本,分别建立 LSSVM 回归模型。根据误差评价函数通过整体样本从一系列回归模型中得到最优的模型,从而剔除离群点的影响,提高了检测精度。通过实验验证了抽样组中样本数与检测精

度及检测效率的关系,确定了抽样组中样本数的合适区间。实验证明,通过选取不同的误差评价函数,RANSAC-LSSVM 方法可以得到不同误差评价体系下的最优结果。综上所述,本文提出的方法可以实现光学元件损伤在线检测图像中损伤区域的高精度亚像素检测,在传统像素级方法的基础上显著提高了检测精度。

参 考 文 献

- 1 Li Dawei, Zhao Yuan'an, He Hongbo, *et al.*. Exponential fitting of laser damage threshold and analysis of testing errors [J]. *Chinese J Lasers*, 2008, 35(2): 273-275.
李大伟, 赵元安, 贺洪波, 等. 光学元件激光损伤阈值的指数拟合法以及测试误差分析[J]. *中国激光*, 2008, 35(2): 273-275.
- 2 Li Xibin, Wang Haijun, Yuan Xiaodong, *et al.*. Effects of CO₂ laser irradiation on surface morphology and stress distribution of fused silica[J]. *Chinese J Lasers*, 2011, 38(9): 0903002.
李熙斌, 王海军, 袁晓东, 等. CO₂ 激光辐照对熔石英表面形貌与应力分布的影响[J]. *中国激光*, 2011, 38(9): 0903002.
- 3 Wang Xue, Xie Zhijiang, Sun Hongyan, *et al.*. Study on automatic flaw inspection system for large caliber precision optical surface[J]. *Chinese J Scientific Instrument*, 2006, 27(10): 1262-1265.
王雪, 谢志江, 孙红岩, 等. 大口径精密光学元件表面疵病检测系统研究[J]. *仪器仪表学报*, 2006, 27(10): 1262-1265.
- 4 Deng Yan, Xu Qiao, Chai Liqun, *et al.*. Total internal reflection microscopy: a subsurface defects identification technique in optically transparent components [J]. *High Power Laser and Particle Beams*, 2009, 21(6): 835-840.
邓燕, 许乔, 柴立群, 等. 光学元件亚表面缺陷的全内反射显微检测[J]. *强激光与粒子束*, 2009, 21(6): 835-840.
- 5 彭志涛. 强激光复杂光机组件光学元件激光损伤在线检测技术研究[D], 绵阳: 中国工程物理研究院, 2011. 1-8, 16-19.
- 6 Yu Zhongdang, Wang Longshan. Salt pepper noise switching filter based on LS-SVR convolution mask[J]. *Acta Optica Sinica*, 2008, 29(1): 163-168.
于忠党, 王龙山. 基于回归型最小二乘支持向量机卷积模板的椒盐噪声开关滤波器[J]. *光学学报*, 2008, 29(1): 163-168.
- 7 Guo Zhiming, Chen Liping, Huang Wenqian, *et al.*. Application of genetic algorithm-least squares support vector regression with near infrared spectroscopy for prediction of nicotine content in tobacco[J]. *Laser & Optoelectronics Progress*, 2012, 49(2): 021201.

- 郭志明, 陈立平, 黄文倩, 等. 基于 FLS-SVM 背景预测的红外弱小目标检测[J]. *激光与光电子学进展*, 2012, 49(2): 021201.
- 8 Zhao Shusen, Chen Sijia, Shen Jingling. Identification of terahertz absorption spectra of illicit drugs using support vector machines[J]. *Chinese J Lasers*, 2009, 36(3): 752-757.
赵树森, 陈思嘉, 沈京玲. 用支持向量机识别毒品的太赫兹吸收光谱[J]. *中国激光*, 2009, 36(3): 752-757.
- 9 Liu Guixi, Liu Dongmei, Liu Fengpeng, *et al.*. A robust image registration algorithm based on feature points matching[J]. *Acta Optica Sinica*, 2008, 28(3): 454-461.
刘贵喜, 刘冬梅, 刘凤鹏, 等. 一种稳健的特征点配准算法[J]. *光学学报*, 2008, 28(3): 454-461.
- 10 Zhang Ning, Chang Lei, Xu Xiping. Research on the technology of three-dimensional reconstruction based on machine vision[J]. *Laser & Optoelectronics Progress*, 2012, 49(5): 051001.
张宁, 常雷, 徐熙平. 基于机器视觉的三维重建技术研究[J]. *激光与光电子学进展*, 2012, 49(5): 051001.
- 11 Zhang Ji, Li Dahai. Algorithm of optics damage inspection from its dark-field image[J]. *Chinese J Lasers*, 2006, 33(8): 1109-1111.
张际, 李大海. 光学元件损伤暗场成像检测的算法[J]. *中国激光*, 2006, 33(8): 1109-1111.
- 12 Xie Yaping, Sun Zhihong, Cheng Ze, *et al.*. Image processing in online inspection of damage in optics[J]. *High Power Laser and Particle Beams*, 2006, 18(7): 1085-1089.
解亚平, 孙志红, 成泽, 等. 光学元件损伤在线检测中的图像处理[J]. *强激光与粒子束*, 2006, 18(7): 1085-1089.
- 13 Feng Bo, Chen Fengdong, Liu Bingguo. Segmentation of small defects in final optics damage online inspection images[C]. *IEEE International Conference on Image Analysis and Signal Processing*, 2012. 110-113.
- 14 Peng Zhitao, Wei Xiaofeng, Yuan Haoyu, *et al.*. Signal noise ratio of total internal reflection edge illumination for optics damage inspection[J]. *Infrared and Laser Engineering*, 2011, 40(6): 1111-1114.
彭志涛, 魏晓峰, 元浩宇, 等. 全内反射照明光学元件损伤检测信噪比分析[J]. *红外与激光工程*, 2011, 40(6): 1111-1114.
- 15 Xiao Bing. Research on the Key Issues of Automatic Defects Measurement on Large Fine Optics [D]. Hangzhou: Zhejiang University, 2010.
肖冰. 大口径光学元件表面疵病自动化检测系统关键问题讨论与研究[D]. 杭州: 浙江大学, 2010.
- 16 Shen Qinmei, Zhou Weidong, Li Kexue. Quantative elemental analysis using laser induced breakdown spectroscopy and neuro-genetic approach[J]. *Chinese J Lasers*, 2011, 38(3): 0315001.
沈沁梅, 周卫东, 李科学. 基于遗传神经网络的激光诱导击穿光谱元素定量分析技术[J]. *中国激光*, 2011, 38(3): 0315001.

栏目编辑: 何卓铭