

工件表面缺陷图像检测中的自适应聚类

周友行^{a,b}, 马逐曦^{a,b}, 石弦韦^{a,b}, 刘汉江^{a,b}

(湘潭大学 a.复杂轨迹加工工艺及装备教育部工程研究中心;
b.机械工程学院, 湖南 湘潭 411105)

摘要: **目的** 针对工件表面形态复杂、干涉交叠缺陷难以实现自动分离、分类图像检测的情况, 提出一种工件表面交叠缺陷自适应图像聚类方法。 **方法** 首先提取工件表面缺陷二值图像, 采用混合概率主成分分析器估计缺陷位置各像素点局部切空间信息, 并改进局部切空间之间的相似性矩阵; 然后通过改进局部密度峰值自适应方法, 基于相似性矩阵确定聚类中心点和数目; 最后通过谱多流形聚类, 将各分析器所包含的像素点分配至不同缺陷流形结构中, 实现多个交叠缺陷分离、检测。 **结果** 首先通过比较计算与实际测量的长度、宽度来验证该方法对相互交叠结构缺陷良好、准确的分离效果, 平均相对误差分别为 0.957% 和 0.650%。其次为了体现该方法对于分离工件表面相互交叠缺陷的有效性及其优越性, 使用 k-means 聚类、谱聚类与该方法进行对比实验, 证明了该方法良好的聚类效果。最后对所设计方法的稳定性进行测试, 统计检测结果的平均 ME 值均在 6% 以下, 正确聚类数目率高达 99%~100%。 **结论** 该方法能够较为准确地自动识别工件表面图像中存在相互干涉的不同缺陷, 并进行分离。

关键词: 工件表面质量; 缺陷检测; 自适应; 流形聚类

中图分类号: TP161+.14 **文献标识码:** A **文章编号:** 1001-3660(2019)09-0327-09

DOI: 10.16490/j.cnki.issn.1001-3660.2019.09.040

Adaptive Clustering Method of Image Detection for Work-piece Surface Defect

ZHOU You-hang^{a,b}, MA Zhu-xi^{a,b}, SHI Xian-wei^{a,b}, LIU Han-jiang^{a,b}

(a.Engineering Research Center for Complex Track Processing Technology and Equipment under Ministry of Education,
b.School of Mechanical Engineering, Xiangtan University, Xiangtan 411105, China)

ABSTRACT: The work aims to propose an adaptive image clustering method for overlapping defects on work-piece surface, so as to solve the problem that the complex and mutually interfered defects on work-piece surface are difficult to be separated automatically and classified and identified by images. Firstly, the binary image of the workpiece surface defects was extracted. The principal component analyzer of mixed probabilities was used to estimate the local tangent space of each pixel on the defects and improve the similarity matrix between the local tangent space of each defect location. Then, the clustering center point and quantity were determined through the improved density peaks adaptive method based on the similarity matrix. Finally, the pixels included in each analyzer were assigned to the different defect manifolds through SMMC to realize the separation and

收稿日期: 2018-12-21; 修订日期: 2019-03-21

Received: 2018-12-21; Revised: 2019-03-21

基金项目: 国家自然科学基金资助项目 (51775468, 51375419)

Fund: Supported by the National Natural Science Foundation of China (51775468, 51375419)

作者简介: 周友行 (1971—), 男, 博士, 教授, 博士生导师, 主要研究方向为数字化设计与制造、制造过程监测与控制、机器人学。
邮箱: zhouyouhang@xtu.edu.cn

Biography: ZHOU You-hang (1971—), Male, Doctor, Professor, Ph. D. supervisor, Research focus: digital design and manufacturing, manufacturing process monitoring and control, and robotics. E-mail: zhouyouhang@xtu.edu.cn

detection of workpiece defects. Firstly, the good and accurate separation effect of this method on the interfered structure defects was verified by comparing the calculated and actual length and width. The average relative errors were 0.957% and 0.650%. Secondly, in order to reflect the effectiveness and superiority, this method was compared with k-means clustering and spectral clustering, which proved a good clustering effect. Finally, the stability of the method was tested. The average ME value was below 6% for the statistical test results, and the correct cluster number rate was as high as 99%~100%. The experimental results show that this method can automatically separate different defects that interfere with each other in the surface image of the workpiece more accurately.

KEY WORDS: workpiece surface quality; defect detection; adaptive detection; manifold clustering

采用机器视觉技术结合 AI 算法进行工件表面质量的缺陷分类和无损检测,能够准确识别工件表面缺陷形态,及时整改、减少生产中加工工件的返工率和废品率,提高生产质量和生产效率^[1-4]。

如何从工件表面图像中准确辨识缺陷区域,即通常意义上的“缺陷提取”,是表面缺陷图像检测的重要步骤。目前缺陷提取方式为:基于灰度不连续变化的边缘检测和基于相似区域的聚类^[5-6]。如 SHI Tian 等人^[7]通过改进 Sobel 算子,提取铁轨表面缺陷 6 个方向的缺陷特征,使检测边缘信息更全面和检测精度更准确;郭皓然等人^[8]构建 Canny 最优边缘检测器,结合类间方差法和形态学,完成高亮金属表面缺陷自适应缺陷识别;贺振东等^[9]根据缺陷边缘特征和面积滤波分割出缺陷图像,提出了一种基于反向 P-M 扩散的钢轨表面缺陷检测方法。依据灰度不连续变化进行边缘检测可以很好地提取出缺陷的形状特征,但是工件表面的走刀纹理过于明显时,则不能很好地区分相似的缺陷与纹理。

基于相似区域的聚类方法可以解决无法区分缺陷和纹理的问题,被大量应用于缺陷检测的图像分割中。如刘涵等人^[10]采用基于排序点的聚类算法(OPTICS)对石油钢管焊缝区域内任意形状大小的缺陷和噪声干扰点进行分割,实现了对焊缝缺陷检测与识别;钱彬等人^[11]提出一种结合多特征流形学习和矩阵分解的路面裂缝检测算法,解决无法从复杂背景噪声中准确提取裂缝信息的问题;Jaina George 等人^[12]采用平滑滤波器进行图像预处理后,使用模糊 C 均值聚类对玻璃缺陷进行检测。但目前的聚类研究都必须先标定缺陷样本参数(如缺陷类别、缺陷形态、缺陷个数等信息),再进行监督学习训练。由于工件表面缺陷种类、形状、成因的复杂性和多种工序叠加进行等原因,导致代表不同类型缺陷数据间存在线性及非线性、良分离及相互交叠结构,在一定程度上难以实现工件表面缺陷分类检测和加工过程监控。

针对上述工件表面缺陷的复杂情况,近年来提出了混合流形聚类方法^[13],以解决这类数据分离问题。R. Souvenir^[14]首次提出将流形聚类方法,使用 ISOMAP 方法计算数据点的距离和使用 EM 方法求解测度距离进行聚类。CHEN 等人^[15]提出了一种用于分

割相似子空间的谱曲率聚类方法。Wang Y 等人^[16]提出了一种新的流形聚类方法 mumCluster,来解决数据流形交叠和聚类数量、固有尺寸未知的问题。Amir Babaeian 等人^[17]通过对 ISOMAP 最短路径法施加曲率约束,来解决不同曲面相交点聚类的问题。混合流形聚类可直接将整体数据中的不同流形结构从流形交叠部分拆分出来,并构造出更忠实于流形结构的近邻图,以实现交叠缺陷的图像分割。

因此,针对目前图像分割问题中无法将相互交叠的不同缺陷分离,影响后续各种缺陷识别、分类的问题,本文改进基于谱多流形的混合流形聚类方法,解决工件表面不同交叠缺陷相互分离的问题,同时利用局部密度峰值实现工件表面缺陷图像自适应聚类分割,能够较为准确地自动分离、识别工件表面相互干涉的不同缺陷。

1 基于改进谱多流形聚类的交叠缺陷分割

谱多流形聚类方法(Spectral Multiple-Manifolds Clustering, SMMC)^[18]利用缺陷区域像素点所包含的局部切空间信息定义各缺陷像素集合局部几何结构。按同一缺陷像素点的特征具有较大的相似性,而不同缺陷像素点的特征具有较大的相异性,辅助构造缺陷区域像素点的相似性矩阵。利用拉普拉斯变化将缺陷区域像素点的相似性矩阵映射到不同分割特征的低维空间中,最终聚类得到不同流形空间的相交缺陷。

1.1 缺陷区域像素点局部切空间求解

局部切空间求解通常根据样本点的局部采样协方差矩阵进行计算。但是,当两个来自不同流形的数据点之间的欧式距离较小时,会导致两点的局部协方差矩阵相似,无法适用于混合流形聚类的计算。

根据文献[19-20]所述,全局非线性流形在局部能被多个局部线性流形近似表示,相似于同一个分析器的数据点具有相似的局部切空间,并且可以被局部分析器的主子空间近似。因此,训练混合概率主成分分析器(Mixtures of probabilistic principal component analyzers, MPPCA)将处于不同缺陷流形的点分离,

并估计像素点的局部切空间^[21], 可解决上述问题。

将工件表面图像缺陷区域像素点定义为 $X\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$ (N 为像素点数量), 混合概率主成分分析器的数量为 M (聚类数量), 则各分析器可定义为^[18]:

$$\theta_m = \{\mu_m, V_m, \sigma_m^2\} \quad (1)$$

其中, $m=1, 2, \dots, M$, μ_m 表示第 m 个分析器的均值向量, V_m 表示第 m 个分析器的权重矩阵, σ_m^2 表示第 m 个分析器的噪声方差。

对于第 m 个分析器, 每个像素点 i 坐标向量 \mathbf{x}_i 对应一个降维后潜在向量 \mathbf{y}_i , 则有:

$$\mathbf{x}_i = V_m \mathbf{y}_i + \mu_m + \varepsilon_m \quad (2)$$

式中: 潜在向量 \mathbf{y}_i 服从正态分布 $\mathbf{y}_i \sim N(0, \mathbf{I})$, 数据集的各向同性噪声 ε_m 同样服从正态分布 $\varepsilon_m \sim N(0, \sigma_m^2 \mathbf{I})$ 。

模型参数 μ_m 、 V_m 、 σ_m^2 使用 K 均值聚类方法求得迭代初值后, 采用 E-M 方法求解最优表述缺陷像素点的 M 个线性分析器和每个点对应的概率分布。

E-step: 使用已知模型参数 μ_m 、 V_m 、 σ_m^2 计算式 (3), 式中 R_{im} 表示用于生成像素点 \mathbf{x}_i 的第 m 个分析器的后验概率; π_m 为混合比例, 且满足 $\sum_{m=1}^M \pi_m = 1$ 和 $\pi_m \geq 0$ 。

$$\begin{cases} R_{im} = \frac{\pi_m p(\mathbf{x}_i | m)}{\sum_{m=1}^M \pi_m p(\mathbf{x}_i | m)} \\ \pi_m^{\text{new}} = \frac{1}{N} \sum_{i=1}^N R_{im} \\ \mu_m^{\text{new}} = \frac{\sum_{i=1}^N R_{im} \mathbf{x}_i}{\sum_{i=1}^N R_{im}} \end{cases} \quad (3)$$

M-step: 以式 (4) 更新模型参数 V_m 、 σ_m^2 :

$$\begin{cases} T_m = \sigma_m^2 \mathbf{I} + V_m^T V_m \\ S_m = \frac{1}{\pi_m^{\text{new}} N} \sum_{i=1}^N R_{im} (\mathbf{x}_i - \mu_m^{\text{new}})(\mathbf{x}_i - \mu_m^{\text{new}})^T \\ V_m^{\text{new}} = S_m V_m (\sigma_m^2 \mathbf{I} + T_m^{-1} V_m^T S_m V_m)^{-1} \\ (\sigma_m^2)^{\text{new}} = \text{tr} \left[S_m - S_m V_m T_m^{-1} (V_m^{\text{new}})^T \right] \end{cases} \quad (4)$$

迭代收敛后, 像素点 i 以式 (5) 关系分配至第 j 个主成分分析器, 其局部切空间为式 (6)。

$$p(\mathbf{x}_i | j) = \max_m p(\mathbf{x}_i | m) \quad (5)$$

$$\mathcal{O}_i = \text{span}(V_j) \quad (6)$$

与主成分分析器相比, 混合概率主成分分析器不仅克服噪声影响提取主成分, 而且通过全局迭代最优解避免了 PCA 对邻域点选取的依赖, 有良好的抗噪性和稳定性^[22]。解决工件表面缺陷图像数据集局部采

样协方差矩阵求解数据集局部切空间时, 难以区分相近图像缺陷流形边界点的问题。

1.2 相似性权值改进及相似性矩阵构造

由于工件表面缺陷种类多形状不一, 存在不同缺陷区域相互干涉的情况, 若采用单一相似性权值构造相似性矩阵对缺陷区域像素点数据进行谱聚类, 不同缺陷像素点可能会被误判。如划痕缺陷上各像素点往往具有较强的流形关联性, 而烧伤等缺陷上各像素点则更注重欧氏距离关联。同时, 根据谱聚类原理相似性权值为零或相对小时, 才可以区分来自不同流形结构的像素点。然而, 在缺陷交叠处像素点之间的欧氏距离小, 使得其相似性权值较大。因此采用 SMMC 进行缺陷提取时, 需要综合考虑像素点流形关联性和欧氏距离关联性的差异。利用 1.1 节所求得像素点所内含的局部几何结构信息来辅助构造新的相似性矩阵 W 。当两个像素点的欧式距离相对较小且具有相似的局部切空间时, 判断两个像素点是来自同一个流形聚类。

由式 (6) 所得, 像素点 i 的局部切空间为 \mathcal{O}_i , 则图像像素点 i 与像素点 j 之间的相似性矩阵见式 (7), 其中两点之间局部切空间的流形相似性见式 (8), 两点之间的欧氏距离关联见式 (9)。

$$W = w_{ij} = a_{ij} b_{ij} = \frac{1}{\ln(1 + d_{ij})} (\cos \theta)^\theta \quad (7)$$

$$a_{ij} = (\cos \theta)^\theta \quad (8)$$

$$b_{ij} = \frac{1}{\ln(1 + d_{ij})} \quad (9)$$

式中: θ 表示局部切空间夹角; O 表示可调节参数, $O \in N^+$, 当 O 值增大时, 局部切空间较欧氏距离影响增大; d_{ij} 表示两点欧式距离, 当 d_{ij} 值增大时, 欧氏距离关联性影响呈对数降低。

式 (7) 满足以下性质: 若两像素点分别来自两个相近流形的边界点时, 其相似性权值较小; 若两像素点局部切空间夹角较大但欧氏距离极小时, 其相似性权值较大。

1.3 基于改进相似性矩阵的谱聚类分割

利用式 (7) 获得像素点间相似性矩阵 W 后, 给定聚类数目 k , 通过谱聚类方法^[26]构造 N-cut 全局评价函数, 得到分割最优解, 能够较为简单地将各主成分分析器分配至各个流形, 得到最终的聚类结果。首先, 以式 $d_{ii} = \sum_i w_{ij}$ 构建相似性矩阵 W 的对角矩阵 D ,

定义 $L = D - W$ 为 Laplacian 矩阵。其次, 根据拉普拉斯矩阵的性质, 可将问题转化为最小化 L 矩阵的特征值问题, 即计算 L 矩阵的前 k 个最小特征值及对应的特征向量 $\mathbf{u}_1, \dots, \mathbf{u}_k$, 且 L 矩阵特征值所对应特征向量的各行向量之间存在较明显的欧氏距离关联。因此谱

聚类计算时,求得最小 l 个特征值对应的特征向量集合 $U=[u_1, \dots, u_l]$ (其中 l 为大于 k 的正整数), 并以其行向量作为对应像素点在该 R^l 空间下的坐标。最后, 使用 k -means 算法对 $U=[u_1, \dots, u_l]$ 的行向量进行聚类, 得到分割后的各缺陷。

谱聚类把缺陷区域像素点转换到谱空间中对其进行分析, 将 R^l 空间下的坐标嵌入低维向量, 相较于直接对相似性矩阵 W 进行聚类分割, 谱空间能够更好地表达不同流形空间的缺陷特征, 得到更好的聚类结果。

2 基于局部密度峰值的自适应聚类

对工件表面图像缺陷像素点数据进行谱多流形聚类分析时, 需要人为确定缺陷类型数目 (聚类数目), 无法实现工件表面图像缺陷的自适应聚类问题。

针对欧氏空间中数据点的聚类问题, Alex Rodriguez 等人^[24]提出了一种基于局部密度峰值的聚类方法, 与传统 K 均值、谱聚类等聚类方法相比, 此方法能够综合考虑数据点局部密度 (ρ_i) 和欧氏距离 (δ_i) 两种因素, 自动确定聚类中心及数目, 解决聚类中心及数目无监督学习问题。使用谱多流形聚类过程中, 各主成分分析器的切空间坐标存在欧氏空间关联, 因此借鉴基于局部密度峰值聚类方法, 解决工件表面各种图像缺陷聚类中心无法自适应选取问题。

根据各数据点的局部密度特性, 基于密度峰值自动确定欧氏空间聚类中心算法的具体步骤如下。

1) 计算数据点两者之间欧式距离 d_{ij} , 取各点间欧氏距离的均值作为截断距离 d_c 。

2) 使用高斯函数描述数据点 i 的局部密度 ρ_i , 其公式为:

$$\rho_i = \sum_j \exp \left(- \left(\frac{d_{ij}}{d_c} \right)^2 \right) \quad (10)$$

其中 d_{ij} 表示数据点 i 与数据点 j 之间的欧氏距离, d_c 表示所设定的截断距离。与数据点 i 的距离小于 d_c 的数据点越多, ρ_i 值越大。

3) 对于数据点 i 与其他具有更高局部密度的数据点之间的最小距离 δ_i , 以式 (11) 进行计算:

$$\delta_i = \min_{j: \rho_j > \rho_i} (d_{ij}) \quad (11)$$

4) 根据聚类中心的性质, 当数据点 i 所对应的 ρ_i 与 δ_i 均较大时, 数据点 i 作为聚类中心的概率较高。 ρ_i 与 δ_i 之间数值差异较大, 对其进行归一化处理, 综合考虑 ρ_i 与 δ_i 的值定义为:

$$\gamma_i = \rho_i \delta_i \quad (12)$$

对每个数据点来说, 数值较大的 γ_i 表明此点有较大密度的同时, 距离其他更高密度点较远。当 γ_i 大于给定阈值时, 数据点 i 作为聚类中心。

基于局部密度峰值聚类方法得到聚类中心点及其数目, 结合上述谱多流形聚类方法, 将属于不同流形空间的各种缺陷从同一张表面图像中分离出来, 实现工件表面各种缺陷的自适应检测, 算法过程如图 1 所示。

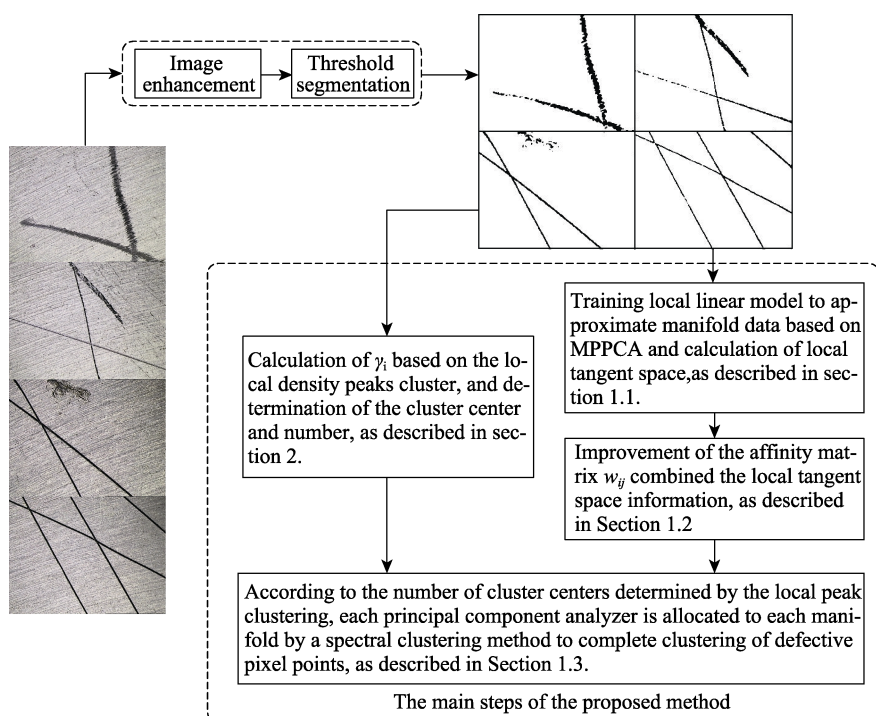


图 1 本文聚类算法流程图

Fig.1 Flow chart of proposed clustering method

3 实验与分析

3.1 实验过程与结果

实验使用天准 VMC 影像仪 VMC222 作为图像采集平台, 平台集成数字相机 (SENTECH 彩色工业相机、NAVITAR 全自动连续变焦镜头)、照明系统 (6 环 8 区 LED 白色冷光源)、运动平台以及计算机控制系统 (影像仪 Vispec 测量软件), 如图 2 所示。

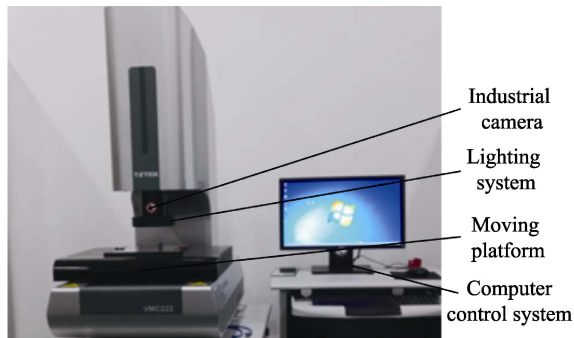


图 2 图像采集平台
Fig.2 Image acquisition platform

当成像设备像素、安装位置和焦距确定后, 即可明确工件表面图像尺寸与分辨率之间的关系。本文中采集的图像格式为 RGB 真彩色, 图像允许畸变,

9.6 mm×7.2 mm 的工件表面图像分辨率为 256 像素×192 像素。以 Intel(R) Xeon(R) CPU E5-1620 0、8 GB 的计算机作为硬件平台, 利用 MatlabR2014b 作为软件平台, 对本文所设计的方法进行实验分析。

考虑工件表面图像颜色信息较为单一, 见图 3 所示。将采集的图像转化为灰度图像后, 进行图像增强和阈值分割技术处理, 提取到工件表面缺陷二值图像, 如图 4 所示 (为便于后续说明, 对图中缺陷进行编号)。

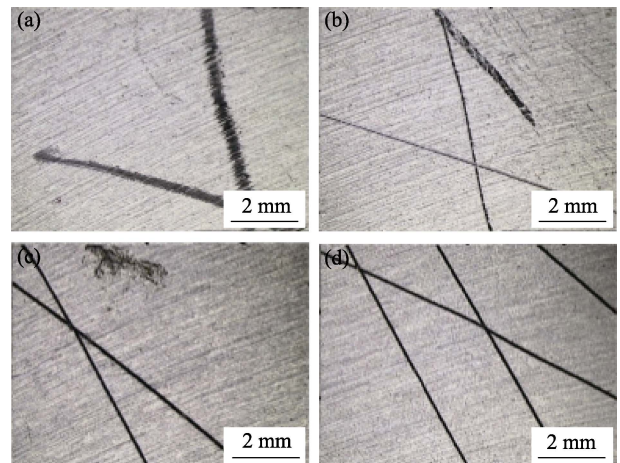


图 3 工件表面图像
Fig.3 Surface image of workpiece

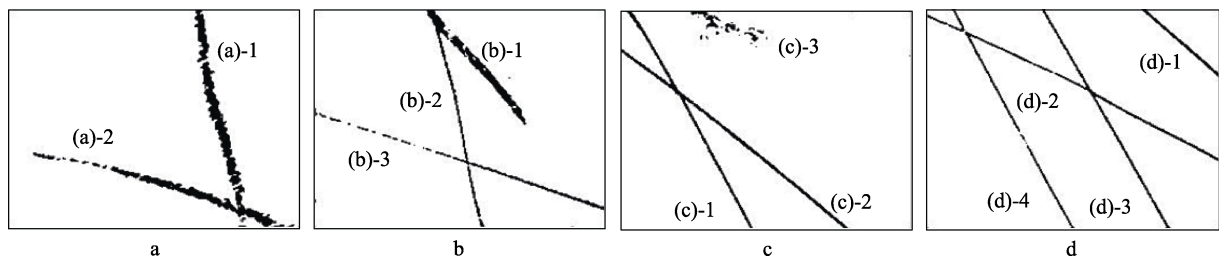


图 4 表面缺陷二值图像
Fig.4 Binary image of surface defect

对于图 4 所示工件表面缺陷二值图像, 首先采用 MPPCA 方法和改进的相似性权值构造各像素点的相似性矩阵。然后基于局部密度峰值进行自适应聚类数目选择, 设定 γ_i 阈值为 ρ 与 δ 各自最大值 1/4 的乘积, 将全部缺陷像素点的 $\{\rho, \delta\}$ 信息在平面中表示出来, 如图 5 所示。

可以看出, 图 5a 中两个像素点 a①、a②同时具有较大的 ρ 与 δ , 因此可知图 4a 中的聚类中心数目 $k=2$ 。同理可得图 4b—d 中的聚类中心数目分别为 $k=3$ 及 $k=4$ 。

最后, 将求得的聚类中心点作为聚类迭代初值, 将聚类中心数目代入谱聚类算法, 对所求得的 $U=[u_1, \dots, u_k]$ 的行向量进行聚类, 即得到如图 6 所示的缺陷检测结果。对比图 3, 从图 6 中可看出, 工件表面图像上各缺陷均被较准确地分离出来。

3.2 实验结果分析

首先验证本文算法对相互交叠结构缺陷良好、准确的分离效果。对图 3 中各缺陷宽度、长度, 根据影像仪自带软件系统 Vispec 测量软件进行人工测量, 测量各缺陷总长度, 并选取缺陷上 3 点测量宽度并求取平均值, 记为实际测量值 (Actual measurement)。同时根据采集图像实际尺寸范围及其对应的像素个数, 可以求出单个像素对应的实际尺寸为 $0.0375 \text{ mm} \times 0.0375 \text{ mm}$, 通过统计图 6 中各缺陷所占像素点个数, 计算出每个缺陷的宽度、长度, 记为计算测量值 (Calculation measurement), 结果见表 1。

通过比较计算与实际测量值, 其长度平均相对误差为 0.957%, 宽度平均相对误差为 0.650%。从计算结果可见, 个别缺陷误差较大主要原因为提取二值图

像过程中, 阈值分割时, 阈值选择对灰度变化不明显

缺陷边界造成的影响, 从而导致测量、计算时的误差。

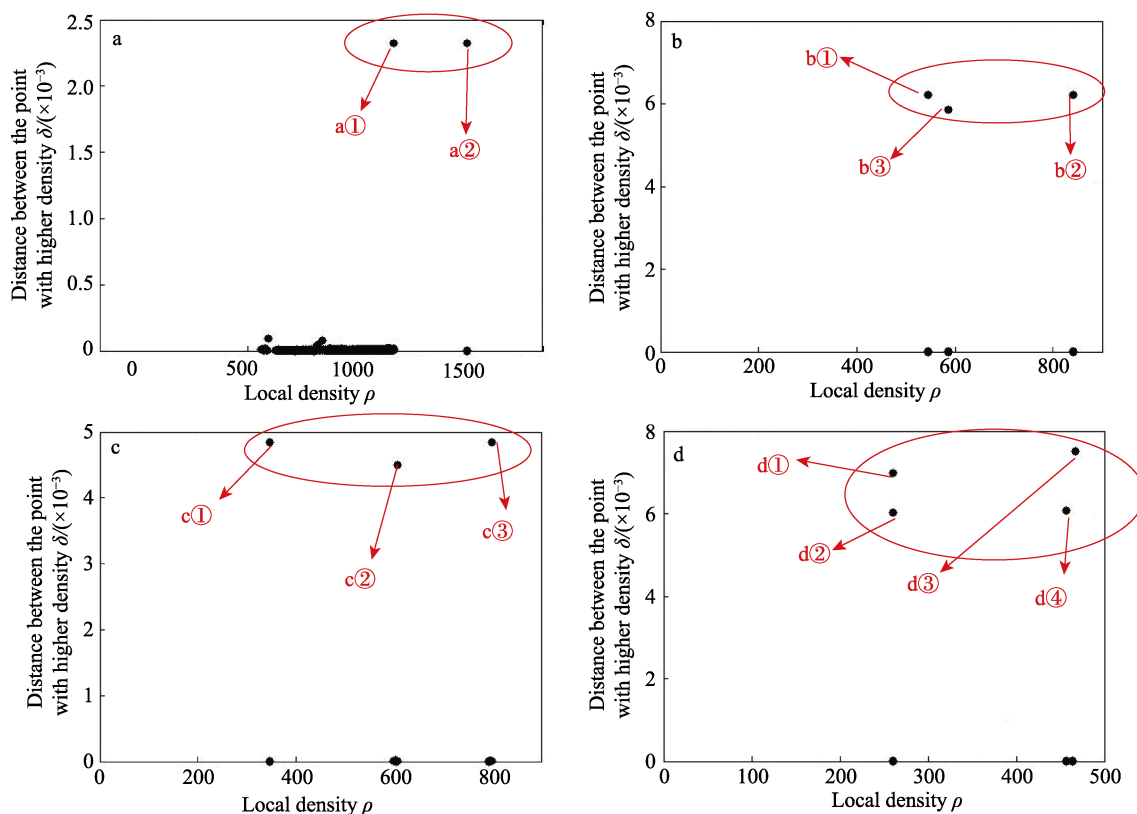


图5 聚类中心选择图

Fig.5 Selection chart of clustering center

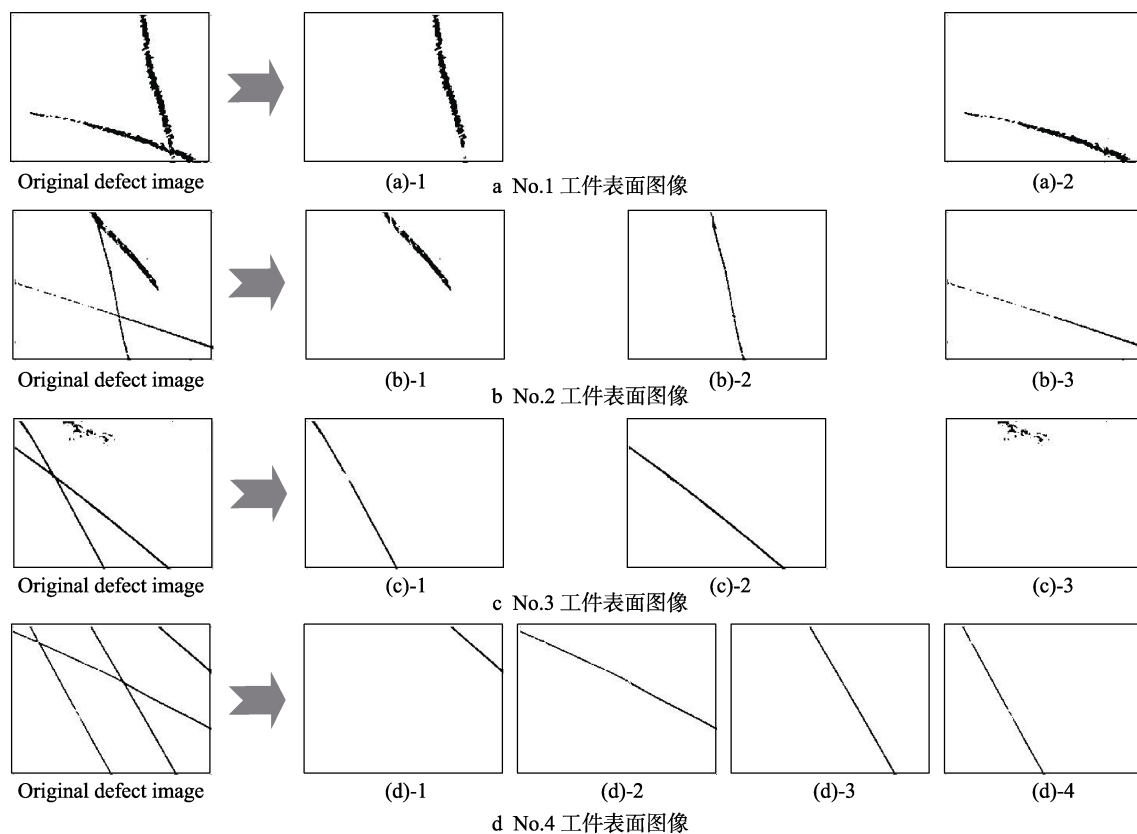


图6 工件表面缺陷检测分割结果

Fig.6 Segmentations of defects on workpiece surface: (a) no.1 workpiece surface image, (b) no.2 workpiece surface image, (c) no.3 workpiece surface image, (d) no.4 workpiece surface image

表 1 缺陷实际测量与计算测量结果
Tab.1 Actual and calculation measurement results of defects

No.		Value/mm												Average relative error
		(a)-1	(a)-2	(b)-1	(b)-2	(b)-3	(c)-1	(c)-2	(c)-3	(d)-1	(d)-2	(d)-3	(d)-4	
Width	Actual measurement	0.43	0.37	0.29	0.08	0.09	0.1	0.11	0.5	0.08	0.08	0.08	0.08	0.957%
	Calculation measurement	0.4193	0.3861	0.2730	0.0750	0.0839	0.0839	0.1186	0.5045	0.0912	0.0874	0.0836	0.0886	
Length	Actual measurement	7.32	8.6	4.93	6.61	10.09	8.18	9.65	2.36	3.28	10.65	8.28	8.15	0.650%
	Calculation measurement	7.3257	8.6026	4.9217	6.6137	10.0798	8.1735	9.6401	2.3539	3.2949	10.6496	8.2842	8.1555	

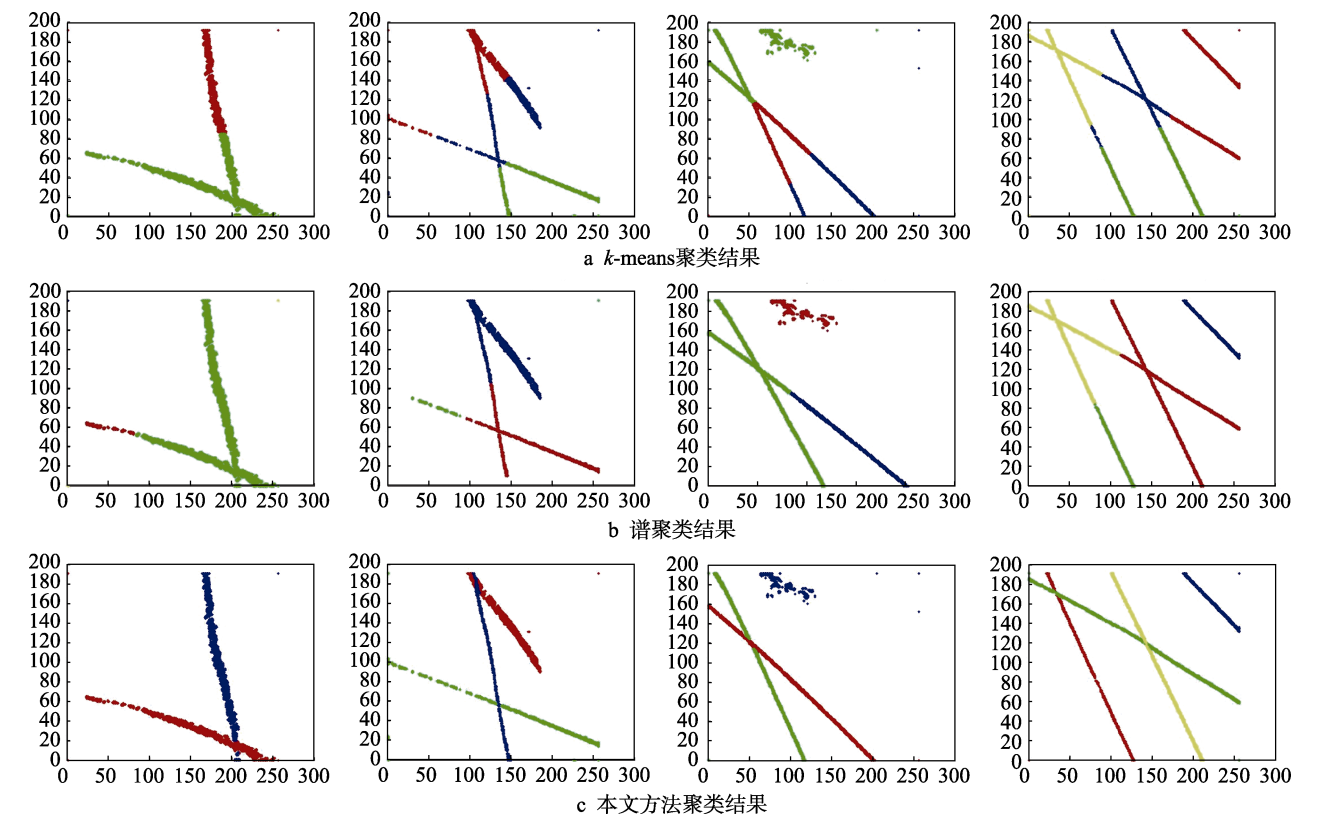


图 7 结果不同聚类算法对工件表面缺陷图像的分割结果
Fig.7 Segmentation results of defects on workpiece surface by different clustering algorithms:
(a) results of *k*-means clustering, (b) results of spectral clustering, (c) results of clustering based on the proposed method

为了体现本文算法对于分离工件表面相互交叠缺陷的有效性及优越性，使用 *k*-means 聚类、谱聚类与本文聚类算法进行对比实验，如图 7 所示。*k*-means 聚类随机划分多个类，采用平方误差准则不断重复迭代直至其最小。谱聚类采用隶属度确定每个数据点属于某个类相似度，通过构建相似度矩阵实现划分。对比图 7a—c 可以看出，*k*-means 聚类只能根据图像像素点之间的欧氏距离进行分类，当数据呈流形或相互交叠时，完全无法准确聚类；谱聚类对于流形像素点的效果好于 *k*-means 聚类，在相对小的范围内可以识别不同类的流行数据，但对于相互交叠的情况，依然无法进行正确聚类。反观图 7c 可以清楚地看到，不同流形结构且相互交叠的缺陷数据点被准确分类。综

上所述，通过对比实验验证，本文算法具有良好的聚类效果。

为了进一步对本文所设计方法的稳定性进行测试，通过在 400 幅含有缺陷的图像中进行 10 次随机抽样（一次抽取 30 张含缺陷图像）测试本文方法分割效果，并统计检测结果的平均 ME 值^[25]与正确聚类数目率。

$$ME = 1 - \frac{|B_o \cap B_T| + |F_o \cap F_T|}{|B_o| + |F_o|} \tag{13}$$

式中： B_o 与 F_o 是标准二值掩码（ground truth）图像中背景纹理与缺陷，由人工观察制作而成； B_T 与 F_T 为提取后二值图像的背景和缺陷； \cap 为取交操

作; || 为统计各部分像素点个数。ME 值主要用于衡量提取结果与原图像的差值, 该值越小表示提取效果越好。检测统计结果如表 2 所示。可知本文方法对相互交叠结构缺陷自适应确定聚类数目的正确率高达 99%~100%, 第 4 次和第 9 次出现错误的原因是缺陷中含有大面积锈斑, 且没有明显的密度集中分布情况。同时在聚类的分割效果上, ME 值均在 6% 以下。

表 2 缺陷检测 ME 值和正确聚类数目结果
Tab.2 Results of ME and correct clusters number of defects

No.	Value of ME	Number of correct clusters
1	2.79%	100.00%
2	5.67%	100.00%
3	2.12%	100.00%
4	5.93%	99.95%
5	3.12%	100.00%
6	3.27%	100.00%
7	3.97%	100.00%
8	2.71%	100.00%
9	2.17%	99.98%
10	3.37%	100.00%

4 结论

1) 改进谱多流形聚类方法, 通过 MPPCA 估计缺陷区域像素点所含局部切空间, 并辅助构造包含流形空间和欧氏距离信息的相似性矩阵, 进而发现利用谱聚类完成不同流形空间相交缺陷分割, 可解决工件表面缺陷区域像素点数据集局部采样协方差矩阵求解数据集局部切空间时, 难以区分相近图像缺陷流形边界点的问题, 能够自动区分各类工件表面缺陷形状, 提高存在相互干涉的各类缺陷分离识别效果。

2) 采用局部密度峰值确定自适应选择聚类中心点及其数目, 改进谱多流形聚类方法, 可有效实现工件表面各种图像缺陷的自适应分类和识别。

参考文献:

- [3] DUTTA Samik, PAL S K, SEN R. On-machine tool prediction of flank wear from machined surface images using texture analyses and support vector regression[J]. Precision engineering, 2016, 43: 34-42.
- [4] PRABUWONO A S, BESARI A R A, ZAMRI R, et al. Surface defects classification using artificial neural networks in vision based polishing robot[C]// International conference on intelligent robotics and applications. Aachen: Intelligent robotics and applications, 2011: 599-608.
- [5] 徐科, 王磊, 王璟瑜. 基于 Tetrolet 变换的热轧钢板表面缺陷识别方法[J]. 机械工程学报, 2016, 52(4): 13-19.
- XU Ke, WANG Lei, WANG Jing-yu. Surface defect recognition of hot-rolled steel plates based on tetrolet transform[J]. Journal of mechanical engineering, 2016, 52(4): 13-19.
- [6] CUI D, XIA K. Dimension reduction and defect recognition of strip surface defects based on intelligent information processing[J]. Arabian journal for science & engineering, 2017(1-2): 1-8.
- [7] 段瑞玲, 李庆祥, 李玉和. 图像边缘检测方法研究综述[J]. 光学技术, 2005, 31(3): 95-99.
- DUAN Rui-ling, LI Qing-xiang, LI Yu-he. Summary of image edge detection[J]. Optical technique, 2005, 31(3): 95-99.
- [8] 何俊, 葛红, 王玉峰. 图像分割算法研究综述[J]. 计算机工程与科学, 2009, 31(12): 58-61.
- HE Jun, GE Hong, WANG Yu-feng. Survey on the methods of image segmentation research[J]. Computer engineering & science, 2009, 31(12): 58-61.
- [9] SHI T, KONG J, WANG X, et al. Improved sobel algorithm for defect detection of rail surfaces with enhanced efficiency and accuracy[J]. Journal of Central South University, 2016, 23(11): 2867-2875.
- [10] 郭皓然, 邵伟, 周阿维, 等. 全局阈值自适应的高亮金属表面缺陷识别新方法[J]. 仪器仪表学报, 2017(11): 2797-2804.
- GUO Hao-ran, SHAO Wei, ZHOU A-wei, et al. Novel defect recognition method based on adaptive global threshold for highlight metal surface[J]. Chinese journal of scientific instrument, 2017(11): 2797-2804.
- [11] 贺振东, 王耀南, 刘洁, 等. 基于背景差分的高铁钢轨表面缺陷图像分割[J]. 仪器仪表学报, 2016, 37(3): 640-649.
- HE Zhen-dong, WANG Yao-nan, LIU Jie, et al. Background differencing-based high speed rail surface defect image segmentation[J]. Chinese journal of scientific instrument, 2016, 37(3): 640-649.
- [12] 刘涵, 郭润元. 基于 X 射线图像和卷积神经网络的石油钢管焊缝缺陷检测与识别[J]. 仪器仪表学报, 2018, 39(4): 247-256.
- LIU Han, GUO Run-yuan. Pavement crack detection based on multi-feature manifold learning and matrix factorization[J]. Chinese journal of scientific instrument, 2018, 39(4): 247-256.
- [13] 钱彬, 唐振民, 沈肖波, 等. 基于多特征流形学习和矩阵分解的路面裂缝检测[J]. 仪器仪表学报, 2016, 37(7): 1639-1646.
- QIAN Bin, TANG Zhen-min, SHEN Xiao-bo, et al. Pavement crack detection based on multi-feature manifold learning and matrix factorization[J]. Chinese journal of scientific instrument, 2016, 37(7): 1639-1646.
- [14] GEORGE J, JANARDHANA S, JAYA J, et al. Automatic defect detection in spectacles and glass bottles based on fuzzy c means clustering[C]// International conference on current trends in engineering and technology. India: Coimbatore, 2013: 8-12.
- [15] 王勇. 基于流形学习的分类与聚类方法及其应用研究

- [D]. 长沙: 国防科学技术大学, 2011.
WANG Yong. Manifold learning based classification and clustering approaches with their applications[D]. Changsha: National University of Defense Technology, 2011.
- [16] SOUVENIR R, PLESS R. Manifold clustering[J]. Proceedings, 2005, 1(1): 648-653.
- [17] CHEN G, LERMAN G. Spectral curvature clustering (SCC) [J]. International journal of computer vision, 2009, 81(3): 317-330.
- [18] WANG Y, JIANG Y, WU Y, et al. Multi-manifold clustering[C]// Pacific rim international conference on trends in artificial intelligence. Korea: Daegu, 2010: 280-291.
- [19] BABAEIAN A, BAYESTEHTASHK A, BANDARABADI M. Multiple manifold clustering using curvature constrained path[J]. Plos one, 2015, 10(9): 0137986.
- [20] WANG Y, JIANG Y, WU Y, et al. Spectral clustering on multiple manifolds[J]. IEEE transactions on neural networks, 2011, 22(7): 1149-1161.
- [21] ROWEIS S T, SAUL L K. Nonlinear dimensionality reduction by locally linear embedding[J]. Science, 2000, 290(5500): 2323-2326.
- [22] SAUL L K, ROWEIS S T. Think globally, fit locally: Unsupervised learning of low dimensional manifolds[J]. J machine learning research, 2003, 4(2): 119-155.
- [23] TIPPING M E, BISHOP C M. Mixtures of probabilistic principal component analyzers[J]. Neural computation, 2014, 11(2): 443-482.
- [24] 王帅, 孙华燕, 郭惠超, 等. 激光点云的混合流形谱聚类自适应分割方法[J]. 光学学报, 2017(10): 125-133.
WANG Shuai, SUN Hua-yan, GUO Hui-chao, et al. Mixed manifold spectral clustering adaptive segmentation method for laser point cloud[J]. Acta optica sinica, 2017(10): 125-133.
- [25] SHI J, MALIK J. Normalized cuts and image segmentation[J]. IEEE transactions on pattern analysis & machine intelligence, 2000, 22(8): 888-905.
- [26] RODRIGUEZ A, LAIO A. Clustering by fast search and find of density peaks[J]. Science, 2014, 344(6191): 1492.
- [27] SEZGIN M, SANKUR B. Survey over image thresholding techniques and quantitative performance evaluation[J]. Journal of electronic imaging, 2004.
- [28] NG A Y, JORDAN M I, WEISS Y. On spectral clustering: Analysis and an algorithm[J]. Proc nips, 2001, 14: 849-856.
- (上接第 320 页)
- [8] 刘海江, 景施博. 基于案例推理的航天大型薄壁件加工过程质量追溯[J]. 制造业自动化, 2018, 40(4): 3-7.
LIU Hai-jiang, JING Shi-bo. Quality traceability of machining process of aerospace large thin-wall parts based on CBR[J]. Manufacturing automation, 2018, 40(4): 3-7.
- [9] NUMTHONG C, BUTDEE S. The knowledge based system for forging process design based on case-based reasoning and finite element method[J]. International journal of applied science and technology, 2013, 5(2): 45-54.
- [10] XU T, WADE N S, DAVIDSON E M, et al. Case-based reasoning for coordinated voltage control on distribution networks[J]. Electric power systems research, 2011, 81(12): 2088-2098.
- [11] 陈维克, 范微微, 李忠群, 等. 基于混合推理的大型机床零件切削数据库系统[J]. 中国机械工程, 2015, 26(7): 923-928.
CHEN Wei-ke, FAN Wei-wei, LI Zhong-qun, et al. Cutting database system of large machine tool parts based on hybrid reasoning method[J]. China mechanical engineering, 2015, 26(7): 923-928.
- [12] 周帆, 江志刚, 张华. 基于改进实例推理的废旧零部件再制造工艺设计[J]. 湖北工业大学学报, 2014(5): 30-34.
ZHOU Fan, JIANG Zhi-gang, ZHANG Hua. The remanufacturing process design method for used machine tools based on improved case-based reasoning[J]. Journal of Hubei University of Technology, 2014(5): 30-34.
- [13] 相克俊, 刘战强, 艾兴, 等. 基于混合推理的高速切削数据库系统的建立[J]. 计算机集成制造系统, 2006, 12(3): 420-427.
XIANG Ke-jun, LIU Zhan-qiang, AI Xing, et al. Development of high-speed cutting database system based on hybrid reasoning[J]. Computer integrated manufacturing systems, 2006, 12(3): 420-427.
- [14] CAI R, MORGAN M N. Development of intelligent grinding database[J]. Key engineering materials, 2007, 329(7): 21-26.
- [15] 甘露华. 小孔相贯线超声去毛刺工艺试验研究和数值分析[D]. 太原: 太原理工大学, 2015.
GAN Lu-hua. Experimental research and numerical analysis on the intersection line of pinhole ultrasonic deburring[D]. Taiyuan University of Technology, 2015.
- [16] QIAO Z, QU D, WANG H, et al. Experimental investigation of the influence of chip interference on burr height in machining micro V-grooves on electroplated copper roll die[J]. Proceedings of the institution of mechanical engineers part B: Journal of engineering manufacture, 2016, 232(8): 583.
- [17] 韩敏, 沈力华. 基于 FCM 与神经网络的案例推理方法[J]. 控制与决策, 2012, 27(9): 1421-1424.
HAN Min, SHEN Li-hua. Case-based reasoning based on FCM and neural network[J]. Control and decision, 2012, 27(9): 1421-1424.
- [18] 张传辉, 田建艳, 高伟, 等. 基于模糊聚类分析的风电功率预测研究[J]. 太原理工大学学报, 2018(1): 133-139.
ZHANG Chuan-hui, TIAN Jian-yan, GAO Wei, et al. Research on wind power forecasting based on fuzzy clustering analysis[J]. Journal of Taiyuan University of Technology, 2018(1): 133-139.