

图像测量中平行线的自适应拟合^①

蒋爱德

(河南牧业经济学院 信息工程学院, 郑州 450011)

摘要: 对于图像测量中平行线边缘点的拟合问题, 本文通过聚类分析法中的 *DBSCN* 算法去除边缘点图像中的噪声数据, 并结合最小二乘法对平行线边缘点进行分类拟合, 解决了传统算法中根据图像特点人为划分区域进行分别拟合的问题, 达到了机器对图像进行自动处理的目的。

关键词: 图像测量; 最小二乘法; 聚类分析

Parallel Lines Adaptive Fitting in Image Measurement

JIANG Ai-De

(Department of Information and Engineering, Henan University of Animal Husbandry and Economy, Zhengzhou 450011, China)

Abstract: For the parallel edge points fitting in image measurement, this paper combines with the least square method to classify parallel edge points fitting through clustering analysis algorithm to eliminate the noise in the edge points image. The traditional algorithm based on image characteristics dividing area artificially to fitting respectively is solved. The result achieves the purpose of automatic image processing by machine.

Keywords: image measurement; least squares; clustering analysis

0 引言

随着计算机技术、数字图像处理技术以及广电技术的快速发展, 图像测量技术^[1]在工农业、遥感、医学、气象、军事、通信、智能机器人等领域得到了广泛应用, 成为现代科学技术和现代工农业生产快速发展的重要技术支撑和高新技术之一。图像测量^[2]就是测量被测对象时, 把图像当做检测盒传递信息的手段和载体加以利用的测量方法。其目的是从图像中提取有用的特征信号, 然后对图像中目标或区域的特征进行量测和估计。测量时, 由光学测量系统、图像输入设备和计算机, 在线非接触的获取被测对象的原始图像, 应用图像处理技术对原始图像进行处理, 获取图像的边缘点数据并进行数据处理, 从而获得对象的几何参数。

由于受诸多条件的影响, 边缘检测得到的边缘点并不一定完全符合原对象图像的边缘, 并且检测到的边缘不完整, 为了便于对图像进行测量, 就需要对这些边缘点进行拟合。由于图像测量过程中很多图像中

的目标是直线, 甚至是两条或多条直线, 并且得到的图像特征边缘点包含有噪声, 这就为这类图像中边缘点的拟合带来很多问题。对于这类问题, 传统的方法是人工确定区域, 将近似在一条直线上的像素点划为一个区域, 然后在该区域内对像素进行拟合。这种人工划分区域的方法违背了图像测量的目的。图像测量技术的研究主要源于两个应用目的, 其一便于人们分析而对图像信息进行改进; 其二是为了使机器自动理解而对图像数据进行存储、传输及显示。本文分析了最小二乘法数据拟合算法, 并利用聚类分析中的 *DBSCN* (*Density-Based Spatial Clustering of Applications with Noise*) 算法对图像中直线的噪声数据进行去除, 并自动对平行线边缘点进行分组拟合。

1 最小二乘法数据拟合^{[3][4]}

对离散数据点序列 $(x_j, y_j) (j=1, 2, \dots, m)$, 选择 $\{\varphi_{i(x)}\} (i=0, 1, \dots, n)$ 作函数的基函数, 于是得到下面的拟合函数:

① 收稿时间:2013-10-07;收到修改稿时间:2013-11-28

$$P_n(x) = a_0\varphi_0(x) + a_1\varphi_1(x) + \dots + a_n\varphi_n(x) \quad (1.1)$$

在这组数据点上, 首先构造一个偏差的平方和函数

$$S(a_0, a_1, \dots, a_n) = \|\delta\|_2^2 = \sum_{j=1}^m [p_n(x_j) - y_j]^2 \quad (1.2)$$

此处偏差向量 $\delta = (y_1 - p_n(x_1), \dots, y_m - p_n(x_m))$, 显然 S 是待定系数 (a_0, a_1, \dots, a_n) 的 $n+1$ 元函数; 然后要求在 (1.2) 式所有可能的系数中求一组系数 $(a_0^*, a_1^*, \dots, a_n^*)$ 使得 $S(a_0^*, a_1^*, \dots, a_n^*)$ 取得最小值, 即

$$S(a_0^*, a_1^*, \dots, a_n^*) = \min \sum_{j=1}^m [p_n(x_j) - y_j]^2 \quad (1.3)$$

这种选取系数并求和函数的方法称为最小二乘法. 由这组系数构造的拟合函数

$P_n^*(x) = a_0^*\varphi_0(x) + a_1^*\varphi_1(x) + \dots + a_n^*\varphi_n(x)$ 就是最小二乘意义下的误差最小的拟合函数, 其示意图见图1所示.

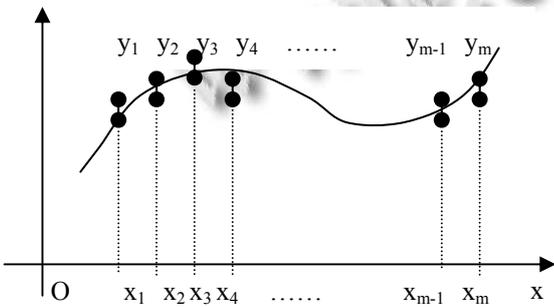


图 1 最小二乘法拟合示意图

在实际问题中, 常在每个点 x_j 处给偏差的平方赋予一个权值 $\rho(x_j)$, 结合(1.1)式, 则 S 可表示为:

$$S(a_0, a_1, \dots, a_n) = \sum_{j=1}^m \rho(x_j) [y_j - \sum_{i=0}^n a_i \varphi_i(x_j)]^2 \quad (1.4)$$

取(1.1)中的 $\{\varphi_0(x), \varphi_1(x), \dots, \varphi_n(x)\}$ 为 $\{1, x, \dots, x^n\}$, 则拟合函数 $P_n(x)$ 为拟合多项式, 这里取 $n=1$ 和 2 , 此时 $P_1(x)$ 为拟合一次多项式(线形拟合), $P_2(x)$ 为拟合二次式. 对于(1.4)式, 为简化起见, 这里令权函数 $\rho(x) \equiv 1$, 由内积定义, 可得以下内积:

$$\begin{aligned} (\varphi_0, \varphi_0) &= m, \quad (\varphi_0, \varphi_1) = (\varphi_1, \varphi_0) = \sum_{j=1}^m x_j, \\ (\varphi_1, \varphi_1) &= \sum_{j=1}^m x_j^2, \quad (y, \varphi_1) = \sum_{j=1}^m x_j y_j \end{aligned} \quad (1.5)$$

令 $P_1(x) = a_0 + a_1 x$, 由(1.11)式此时的法方程为:

$$\begin{pmatrix} m & \sum_{j=1}^m x_j \\ \sum_{j=1}^m x_j & \sum_{j=1}^m x_j^2 \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} = \begin{pmatrix} \sum_{j=1}^m y_j \\ \sum_{j=1}^m x_j y_j \end{pmatrix} \quad (1.6)$$

用消去法或克莱姆法解之, 得到:

$$a_0 = \frac{D_1}{D}, \quad a_1 = \frac{D_2}{D} \quad (1.7)$$

$$D = \begin{vmatrix} m & \sum_{j=1}^m x_j \\ \sum_{j=1}^m x_j & \sum_{j=1}^m x_j^2 \end{vmatrix} = m \sum_{j=1}^m x_j^2 - (\sum_{j=1}^m x_j)^2 \quad (1.8)$$

$$D_1 = \begin{vmatrix} \sum_{j=1}^m y_j & \sum_{j=1}^m x_j \\ \sum_{j=1}^m x_j y_j & \sum_{j=1}^m x_j^2 \end{vmatrix} = \sum_{j=1}^m y_j \sum_{j=1}^m x_j^2 - \sum_{j=1}^m x_j \sum_{j=1}^m x_j y_j \quad (1.9)$$

$$D_2 = \begin{vmatrix} m & \sum_{j=1}^m y_j \\ \sum_{j=1}^m x_j & \sum_{j=1}^m x_j y_j \end{vmatrix} = m \sum_{j=1}^m x_j y_j - \sum_{j=1}^m x_j \sum_{j=1}^m y_j \quad (1.10)$$

2 基于 DBSCN 算法^{[5][6][7]}的图像边缘点分类拟合

DBSCN 是一个基于密度的聚类算法. 该算法通过不断生长足够高密度区域来进行聚类, 它能从含有噪声的空间数据库中发现任意形状的聚类. DBSCN 算法将一个聚类定义为一组“密度连接”的点集. 其算法思想如下:

(1) 一个给定对象 ε 的半径内的近邻就称为对象的 ε -近邻.

(2) 若一个对象的 ε -近邻至少包含一定数目 ($MinPts$) 的对象, 该对象称为核对象.

(3) 给定一组对象集 D , 若对象 p 为另一个对象 q 的 ε -近邻且 q 为核对象, 那么就称 p 是从可以“直接密度可达”.

(4) 对于一个 ε 而言, 一个对象 p 是从对象 q 可“密度可达”; 一组对象集 D 有 $MinPts$ 个对象; 若有一系列对象 p_1, p_2, \dots, p_n , 其中 $p_1 = q$, 且 $p_n = p$, 从而使得(对于 ε 和 $MinPts$ 来讲) p_{i+1} 是从 p_i 可“直接密度可达”. 其中有 $p_i \in D, 1 \leq i \leq n$.

(5) 对于 ε 和 $MinPts$ 来讲, 若存在一个对象 $o (o \in D)$, 使得 o 从可“密度可达”对象 p 和对象 q . 对象 p 是“密度连接”对象 q .

基于密度聚类就是一组“密度连接”的对象, 以实现最大化的“密度可达”. 不包含在任何聚类中的对象为噪声数据.

图像测量中平行线边缘点一般分为双平行线和多平行线边缘点, 以双平行线为例, 如下图所示.

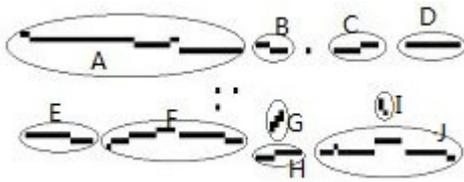


图 2 平行线边缘点的聚合

对于得到的测量图像边缘点，每一个边缘点就是一个像素。对于图 2 中的平行线边缘点的拟合根据 DBSCAN 算法，每一个边缘像素点在给定 ϵ 半径的邻域内有不少于 N_{min} 边缘点的边缘点保存下来作为核对象，由于许多噪声数据点没有核对象，因此可以很容易将噪声数据与边缘数据区分开来。以核对象为基础点，在给定的 ϵ 半径内搜寻“密度可达”的边缘点并划分为一类。图 2 中的图像可以分为 A、B、C、D、E、F、G、H、I 和 J 10 个小聚类。对于将要拟合成的直线 $P_i(x) = m_i + k_i x$ ，根据公式(1.7~1.10)，首先每个聚类内部进行拟合，得到的不同的斜率 k_1, k_2, \dots, k_{10} 和截距 m_1, m_2, \dots, m_{10} ，然后分别给定阈值 T_k 和 T_m ，根据 $|k_i| \leq T_k (i = 1, 2, \dots, n)$ ，可以得到 G 和 I 聚类的边缘点属于噪声点。根据 $|m_i| \leq T_m (i = 1, 2, \dots, n)$ ，A、B、C 和 D 重新划分一个大聚类，E、F、G、H 和 J 重新划分一个大聚类，并对大聚类内的边缘点最终拟合。在实际应用中，计算 A、B、C、D 小聚类斜率和截距的平均值作为上面的直线的斜率和截距，计算 E、F、G、H、J 小聚类斜率和截距的平均值作为下面的直线的斜率和截距，然后生成直线。

根据以上算法思想，对具有平行线特征的图像测量过程步骤如下：

- (1) 对原始图像预处理；
- (2) 对预处理过后的图像进行边缘检测，获得待检测图像的特征边缘点；
- (3) 对获得的边缘点图像进行聚类分析划分小的聚类；
- (4) 对小聚类内的边缘点利用最小二乘法公式获得斜率和截距；
- (5) 根据斜率和截距重新聚类，最终拟合成直线。

3 应用实例

根据以上算法思想，通过以 VC++ 为平台编程进行实验验证，对于给定的一组工件装配时的上下模间隙图像进行间隙测量。

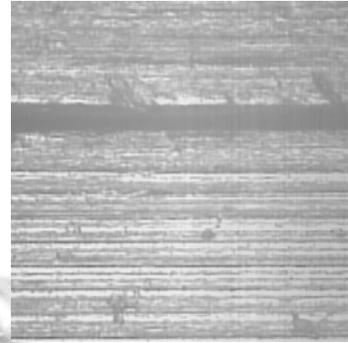


图 3 工件上下模间隙原图

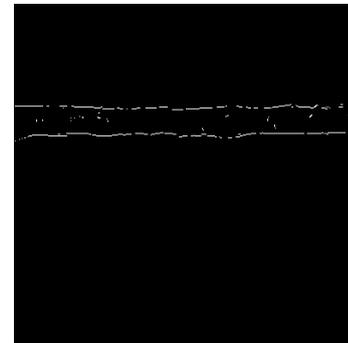


图 4 边缘检测后得到的图像



图 5 传统方法拟合后的图像

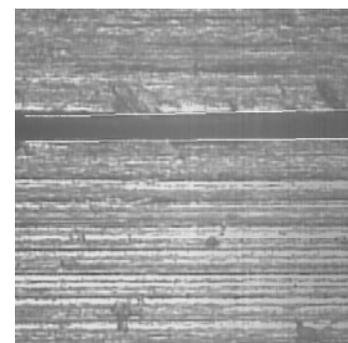


图 6 图 5 与图 3 相加后的图像



图7 本文方法拟合后的图像

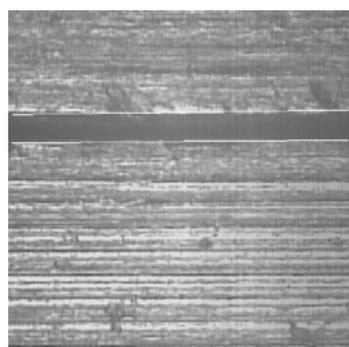


图8 图6与图3相加后的图像

实验结果显示,人工划分区域的传统方法拟合后的图像与边缘检测的图像中的噪声有关系。如果噪声多,则拟合后的图像与原图像相加显示有偏差(图6所示)。本文方法获得的拟合后的图像与原图像相加显示无偏差(图8所示),并且在边缘检测后的图像扫描过程中边扫描边聚类,增加的时间可以忽略不计,达到了图像测量的目的。

4 结束语

一般在图像测量中,图像的特征边缘比较明显,

所以预处理后边缘检测的图像中噪声有多有少,利用传统方法拟合直线,不同的图像需要划分不同的区域,鲁棒性不高,而且拟合后的直线与原图像边缘有偏差。

由于本文算法能对具有平行线特征的图像进行自适应图像测量,尤其适应于批量处理的图片,只要图像得到的边缘点不相交,利用此方法都能进行自适应拟合。所以利用最小二乘法与聚类分析中的DBSCAN算法结合进行数字图像测量中的平行线拟合,实验结果符合图像测量的要求,达到了对图像进行自动处理的目的。

参考文献

- 1 许琳,曹茂永,冯秀霞.图像测量技术及其在无损检测中的应用.电子测量技术,2008,(1):137-140.
- 2 池福俊,陈敏.图像测量系统及关键技术研究.制造技术与装备市场,2011,(4):43-44.
- 3 党兴菊,吴文良.最小二乘法拟合直线公式的初等推导.重庆科技学院学报(自然科学版),2010,(4):185-187.
- 4 程玉民.移动最小二乘法研究进展与述评.计算机辅助工程,2009,(2):5-11.
- 5 李乐,陈鸿昶,李鹏.一种改进的基于密度的聚类算法.电子技术应用,2009,(3):101-104.
- 6 王桂芝,王广亮.改进的快速DBSCAN算法.计算机应用,2009,(9):2505-2508.
- 7 易宝林,伍仪强,丰大洋,张小莉.基于DBSCAN的批量更新聚类算法.计算机工程,2009,(2):63-67.