

基于最小二乘法的几何元素拟合算法研究

Research and Implementation of Geometric Element Fitting Algorithm Based on Least-Squares

北京航空航天大学机械工程及自动化学院 杜福洲 王小强 段桂江

[摘要] 误差评定是零件检测环节中的重要检验项目之一,几何元素拟合是基于坐标测量法评定误差的核心内容。在对常见几何元素分析研究的基础之上,构建了其参数化数学模型,基于最小二乘法提出其拟合算法的整体框架,对具体的算法进行了研究并给出了优化算法初始值的计算过程,基于 MATLAB 平台对拟合算法进行了实现,并将其拟合结果与 NIST 的结果进行比较分析,结果验证了该算法的可行性。

关键词: 误差评定 最小二乘法 几何元素 拟合算法

[ABSTRACT] Error evaluation is one of the important tests in the process of part inspection. Geometric fitting is the core content in the error evaluation based on the coordinate measuring. Based on the analysis of the common geometric elements, the parameterized mathematic model of geometric elements is constructed and the frame of the fitting algorithms is proposed. Meanwhile the detail of fitting algorithms and the initial value of the optimization algorithms are illustrated. Then algorithms are carried out on the MATLAB platform and the fitting results are compared with the results of NIST. The comparison shows that the fitting algorithms are feasible.

Keywords: Error evaluation Least squares method Geometric element Fitting algorithm

在生产制造过程中,经过加工的零件,不仅会产生尺寸误差,而且会产生形位误差,这些误差的存在对零件的性能及装配质量会产生重要的影响,因此必须将其控制在一定的范围内。坐标测量法由于其通用性强、精度高等特点,目前被广泛应用于零件的误差检测中。坐标测量法中涉及几何元素的实际形状和几何元素理想形状的获取,实际形状一般用实测点云数据代替,而理想形状往往是不可直接获得的,一般用点云数据基于一定规则拟合而成的元素代替,因此元素拟合是误差评定的核心内容。

几何元素的拟合方法有最小区域法、最小二乘法和与具体元素相关的其他方法,比如圆的拟合有最小外接圆法、最大内切圆法等。其中,最小二乘法由于其理论

相对比较成熟,工程应用简单易行,因此在一些精度要求不是很高的情况下,其具有一定的可行性,有关研究表明,按最小二乘法评定的形位误差大于实际误差的 1.14 倍^[1]。

本文在对误差评定时常见的几何元素:圆、圆柱面和球面分析研究基础上,构建其参数模型,基于最小二乘法对其拟合算法进行了研究,并在 MATLAB 平台下对算法进行了实现,最后以 NIST(National Institute of Standards and Technology)提供的标准测试数据集对算法进行了验证和比较,其拟合结果与 NIST 提供的拟合结果的误差在 1×10^{-4} 内,证明了算法的可行性。

1 几何元素的参数化建模

几何元素的拟合过程就是求解其特征参数的过程,对于某一元素可以采用不同的参数模型来描述。因此,在元素拟合时应该构建一个合适的参数模型,它是元素拟合算法的前提和基础。

对于常见的几何元素:圆、圆柱面和球面,构建其参数模型如下:

(1) 圆。

圆可以用圆心的空间位置以及圆的半径来描述,同时还应该给出圆所在的平面,其参数模型可以表示为:

$$\text{circle}(x, y, z, a, b, c, r)$$

其中,向量 $\vec{p}(x, y, z)$ 表示圆心的空间位置,向量 $\vec{r}(a, b, c)$ 表示圆所在平面法向量的方向余弦, r 表示圆的半径。

(2) 圆柱面。

圆柱面可以用轴线的位姿以及圆柱面的半径来描述,其中轴线的位姿又可以用轴线上一点及轴线的方向向量来描述,因此,其参数模型可以表示为:

$$\text{cylinder}(x, y, z, a, b, c, r)$$

其中,向量 $\vec{p}(x, y, z)$ 表示圆柱面轴线上一点,向量 $\vec{r}(a, b, c)$ 表示圆柱面轴线方向向量的方向余弦, r 表示圆柱面的半径。

(3) 球面。

球面可以用球心的空间位置以及球的半径来描述,其参数模型可以表示为:

$$\text{sphere}(x, y, z, a, b, c, r)$$

其中,向量 $\vec{p}(x,y,z)$ 表示球心的空间位置, r 表示球面的半径。

2 几何元素的拟合算法

2.1 算法整体框架

最小二乘法的原理是基于偏移误差的平方和最小,因此拟合算法的核心是基于元素的参数模型构建其偏移误差。偏移误差一般又可以分为代数偏移误差和几何偏移误差,代数偏移误差是指点到其代数方程的偏差;几何偏移误差是指点到其几何形状的法向距离偏差。以二维直线为例,两种偏移误差如图 1 所示。

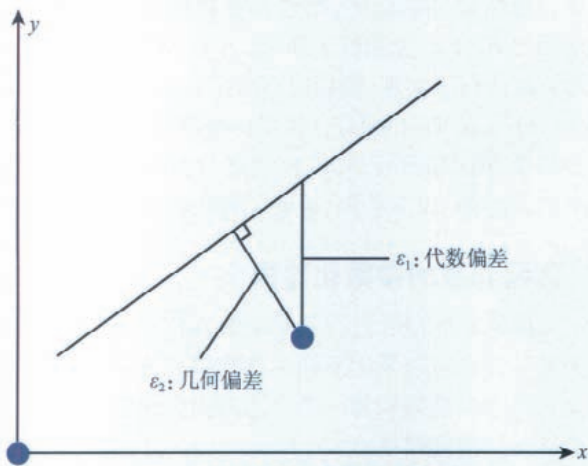


图 1 代数偏差与几何偏差示意图

Fig.1 Diagram of algebraic bias and geometrical bias

由于几何元素的代数方程表示并不唯一,因此其代数偏移误差很难确定,而几何偏移误差只与元素的特征参数有关,具有唯一性,适合优化算法的求解。因此本文选择基于几何偏移误差的最小二乘法进行元素的拟合,其算法的整体框架如图 2 所示。

2.2 拟合算法目标函数

对于常见几何元素的拟合,可以基于优化理论以及 MATLAB 的相关函数^[2]进行实现,而构建拟合算法的优化目标函数是算法的核心内容。对于圆、圆柱面和球面,其优化目标函数^[3]构建如下。

(1) 圆。

空间三维圆的拟合,其几何偏移误差不仅包括径向误差,还包含法向误差。径向误差是指点到过圆心且垂直于圆所在平面的直线之间的距离与拟合圆半径之差;法向误差是指点到拟合圆所在平面的距离。

设 $n(n \geq 3)$ 个点的坐标测量值为 $\vec{p}_i(x_i, y_i, z_i)$ ($i=1, 2, \dots, n$),则径向误差 $d_{i,1}$ 为 $d_{i,1} = |(\vec{p}_i - \vec{p}) \times \vec{r}| - r$; 法向误差 $d_{i,2}$ 为: $d_{i,2} = \vec{p}_i \cdot \vec{r}$ 。因此,各点到拟合圆的几何偏

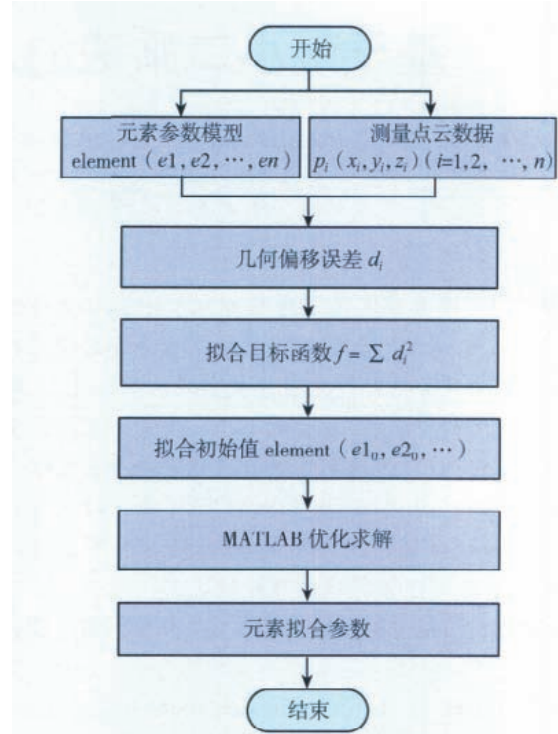


图 2 几何元素拟合算法整体框架

Fig.2 Architecture of geometrical elements fitting algorithm

移误差为:

$$d_i = \sqrt{d_{i,1}^2 + d_{i,2}^2} \quad (1)$$

基于最小二乘法,圆拟合的优化目标函数为:

$$f(x, y, z, a, b, c, r) = \min \left[\sum_{i=1}^n d_i^2 \right] \quad (2)$$

(2) 圆柱面。

设 $n(n \geq 5)$ 个点的坐标测量值为 $\vec{p}_i(x_i, y_i, z_i)$ ($i=1, 2, \dots, n$),则各点到拟合圆柱面的几何偏移误差为:

$$d_i = |(\vec{p}_i - \vec{p}) \times \vec{r}| - r \quad (3)$$

基于最小二乘法,圆柱面拟合的优化目标函数为:

$$f(x, y, z, a, b, c, r) = \min \left[\sum_{i=1}^n d_i^2 \right] \quad (4)$$

(3) 球面。

设测得 $n(n \geq 4)$ 个点的坐标测量值为 $\vec{p}_i(x_i, y_i, z_i)$ ($i=1, 2, \dots, n$),则各点到该拟合球面的几何偏移误差为:

$$d_i = |\vec{p}_i - \vec{p}| - r \quad (5)$$

基于最小二乘法,球面拟合的优化目标函数为:

$$f(x, y, z, r) = \min \left[\sum_{i=1}^n d_i^2 \right] \quad (6)$$

2.3 优化算法初始值

在调用 MATLAB 的优化函数^[4]进行元素拟合时,优化初始值的选择对于问题的求解速度和精度会产生很大的影响,如果选择不当,可能会造成算法的时间开销过大,或者得到局部最优解。因此,初始值的选择应该尽可能地接近问题的最优解,本文基于元素的特征属性构建其约束方程,通过解此方程获取优化函数的初始值。

(1) 圆。

圆拟合的初始值计算分为两个步骤,一是圆所在平面法向量 $\vec{R}_0(a_0, b_0, c_0)$ 的计算,它是一个平面拟合问题,基于 Lagrange 乘子以及约束函数寻找满足优化问题的 K-T 点,再从 K-T 点中确定问题的最优解,整个过程主要是矩阵特征值的相关计算;二是圆心 $\vec{R}_0(a_0, b_0, c_0)$ 及半径 r_0 的确定,其计算过程如下:

设测量数据 $\vec{P}_i(x_i, y_i, z_i)$ 所在的坐标系为 $o-xyz$,对 \vec{P}_i 进行旋转变换,设变换后的坐标系为 $o-x'y'z'$,使得 z' 轴与向量 \vec{R}_0 重合,其旋转变换矩阵 R 为:

$$R = \begin{bmatrix} 1 - \frac{a_0^2}{1+c_0} & \frac{-a_0b_0}{1+c} & -a_0 \\ \frac{-a_0b_0}{1+c_0} & 1 - \frac{b_0^2}{1+c_0} & -b_0 \\ a_0 & b_0 & c_0 \end{bmatrix},$$

设变换后的数据为 $\vec{P}'_i(x'_i, y'_i, z'_i)$,由各点到圆心的距离为半径值可得其约束方程为:

$$(x'_i - x'_0)^2 + (y'_i - y'_0)^2 = r_0^2, \quad (7)$$

将其化为矩阵形式为:

$$\begin{bmatrix} x'_1 & y'_1 & 1 \\ x'_2 & y'_2 & 1 \\ \vdots & \vdots & 1 \\ x'_n & y'_n & 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} -(x_1 + y_1) \\ -(x_2 + y_2) \\ \vdots \\ -(x_n + y_n) \end{bmatrix}, \quad (8)$$

其中, $A = -2x'_0$, $B = -2y'_0$, $C = (x'_0)^2 + (y'_0)^2 - (r'_0)^2$ 。解此线性方程组就可得到坐标系 $o-x'y'z'$ 下的圆心 \vec{P}'_0 和半径 r'_0 ,将其变换到坐标系 $o-xyz$ 就得到圆拟合时的圆心初始值 $\vec{p}_0(x_0, y_0, z_0)$ 和半径初始值 r_0 。

(2) 圆柱面。

圆柱面拟合初始值的计算首先是轴线方向向量 $\vec{r}_0(a_0, b_0, c_0)$ 的计算,设测量数据为 $\vec{p}_i(x_i, y_i, z_i)$,其约束方程为 $|(\vec{p}_i - \vec{p}) \times \vec{r}| = r$,坐标表示形式为:

$$A^2 + B^2 + C^2 = r^2, \quad (9)$$

其中,

$$\begin{aligned} A &= c(y_i - y_0) - b(z_i - z_0); \\ B &= a(z_i - z_0) - c(x_i - x_0); \\ C &= b(x_i - x_0) - a(y_i - y_0). \end{aligned}$$

将方程(9)化简为一般曲面方程为:

$$Ax^2 + By^2 + Cz^2 + Dxy + Eyz + Fyz + Gx + Hy + Iz + J = 0,$$

其中, $A = (b^2 + c^2)$, $B = (a^2 + c^2)$, $C = (a^2 + b^2)$, $D = -2ab$, $E = -2ac$, $F = -2bc$ 。G、H、I 和 J 在计算轴线方向向量时涉及到,因此这里不再给出具体的表示。为了便于处理,给方程两边同时除以 A,则有:

$$x^2 + \frac{B}{A}y^2 + \frac{C}{A}z^2 + \frac{D}{A}xy + \frac{E}{A}yz + \frac{F}{A}yz + \frac{G}{A}x + \frac{H}{A}y + \frac{I}{A}z + \frac{J}{A} = 0,$$

即:

$$\frac{B}{A}y^2 + \frac{C}{A}z^2 + \frac{D}{A}xy + \frac{E}{A}yz + \frac{F}{A}yz + \frac{G}{A}x + \frac{H}{A}y + \frac{I}{A}z + \frac{J}{A} = -x^2.$$

将上述方程写为矩阵形式为:

$$\begin{bmatrix} y_1^2 & z_1^2 & x_1y_1 & x_1z_1 & y_1z_1 & x_1 & y_1 & z_1 & 1 \\ y_2^2 & z_2^2 & x_2y_2 & x_2z_2 & y_2z_2 & x_2 & y_2 & z_2 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ y_n^2 & z_n^2 & x_ny_n & x_nz_n & y_nz_n & x_n & y_n & z_n & 1 \end{bmatrix} \begin{bmatrix} B/A \\ C/A \\ D/A \\ E/A \\ F/A \\ G/A \\ H/A \\ I/A \\ J/A \end{bmatrix} = \begin{bmatrix} -x_1^2 \\ -x_2^2 \\ \vdots \\ -x_n^2 \end{bmatrix}. \quad (10)$$

通过求解线性方程组(10),就可得到 9 个参数值,假设其解为: $B'=B/A$, $C'=C/A$, $D'=D/A$, $E'=E/A$, $F'=F/A$, $G'=G/A$, $H'=H/A$, $I'=I/A$, $J'=J/A$ 。

如果 $|D'|$, $|E'|$, $|F'|$ 近似为 0,此时如果 $|B'|$ 近似为 0,则有 $(a, b, c) = (0, 0, 1)$; 此时如果 $|C'|$ 近似为 0,则有 $(a, b, c) = (0, 1, 0)$ 。否则,令:

$$k = \frac{2}{1 + B' + C'},$$

则可以求出圆柱面一般曲面方程的前 6 个系数项,分别为: $A=k$, $B=kB'$, $C=kC'$, $D=kD'$, $E=E'$, $F=F'$ 。

此时,如果 A, B 都近似等于 1,则有:

$$c' = \sqrt{1 - C}, a' = E' - 2c', b' = F' - 2c',$$

如果 A 近似为 1, B 不近似等于 1,则有:

$$b' = \sqrt{1 - B}, a' = D' - 2b', c' = F' - 2b',$$

如果 A, B 都不近似等于 1,则有:

$$a' = \sqrt{1 - A}, b' = D' - 2a', c' = F' - 2a'.$$

对向量 (a', b', c') 进行单位化处理,得到圆柱面拟合的初始轴线方向向量 $\vec{r}_0(a_0, b_0, c_0)$ 。

得到 \vec{r}_0 后,对 $\vec{p}_i(x_i, y_i, z_i)$ 进行旋转坐标变换,使得变

换后的 Z 轴与向量 r_0 重合, 并将变换后的数据在 OXY 平面上投影, 对投影的数据进行二维圆的拟合得到其圆心和初始半径 r_0 。然后将圆心的坐标值进行旋转反变换到原坐标系下, 得到轴线一点的初始值 $\vec{p}_0(x_0, y_0, z_0)$, 具体的变换计算过程与圆拟合初始值的计算类似。

(3) 球面。

设测量数据为 $\vec{p}_i(x_i, y_i, z_i)$, 则球面的约束方程为:

$$(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2 = r_0^2, \quad (11)$$

转化为矩阵形式为:

$$\begin{bmatrix} x'_1 & y'_1 & z'_1 & 1 \\ x'_2 & y'_2 & z'_2 & 1 \\ \vdots & \vdots & \vdots & 1 \\ x'_n & y'_n & z'_n & 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = \begin{bmatrix} -(x_1 + y_1 + z_1) \\ -(x_2 + y_2 + z_2) \\ \vdots \\ -(x_n + y_n + z_n) \end{bmatrix}, \quad (12)$$

其中, $A=-2x_0, B=-2y_0, C=-2z_0, D=x_0^2+y_0^2+z_0^2-r_0^2$ 。求解线性方程组(12)可以得到球面拟合时的球心初始值

表1 圆的拟合结果比较 1×10^{-4}

误差	序号				
	1	6	8	10	18
Δx	0.0187	0.1455	0.0004	0.0061	0.0001
Δy	0.0090	0.2615	0.0111	0.0139	0.0001
Δz	0.0601	0.1817	0.0383	0.0408	0.0059
Δa	0.0383	0.1632	0.0006	0.0002	0.0087
Δb	0.0007	0.3443	0.0007	0.0018	0.0274
Δc	0.0001	0.2549	0.0005	0.0019	0.0001
Δr	0.0001	0.0019	0.0001	0.0001	0.0001

表2 圆柱面的拟合结果比较 1×10^{-6}

误差	序号				
	1	4	5	13	18
Δd	0.7149	0.0088	0.0001	0.0008	0.8046
Δa	0.0123	0.0001	0.0003	0.0012	0.0759
Δb	0.0027	0.0001	0.0001	0.0067	0.2090
Δc	0.0134	0.0001	0.0002	0.0001	0.0653
Δr	0.0031	0.0078	0.0001	0.0050	0.0033

表3 球面的拟合结果比较 1×10^{-7}

误差	序号				
	1	2	4	5	13
Δd	0.1082	0.0001	0.0036	0.0104	0.3148
Δa	0.0546	0.0029	0.0058	0.0099	0.0001
Δc	0.0936	0.0020	0.0047	0.0164	0.0067
Δr	0.5579	0.0026	0.0097	0.0202	0.5302

$\vec{p}_0(x_0, y_0, z_0)$ 和球半径初始值 r_0 。

3 实例计算与比较

基于上述几何元素拟合算法的目标函数以及初始值的计算, 在 MATLAB 平台下对其进行了具体的实现, 以 NIST 提供的几何元素拟合测试数据集^[5]作为输入对算法进行了验证, 将其拟合结果与 NIST 的拟合结果进行了比较。从 NIST 的数据集中选取 5 组具有代表性的数据, 其各项参数的拟合误差如表 1~ 表 3 所示, 其中序号表示测试数据集中测试数据序号。

对于圆柱面的参数拟合误差, 由于轴线上的一点不是唯一确定的, 因此轴线的位置误差以拟合点到 NIST 给出的轴线距离 Δd 来表示。从上述比较结果可以看出, 算法的参数拟合误差都小于 1×10^{-4} , 对于同一元素的拟合, 不同测试数据其拟合算法的误差会有一些的区别, 这主要是由于采点策略造成的, 包括采样点的数目及布局, 采样点在元素轮廓上分布越均匀, 采点数目越多, 其拟合效果越接近实际情况。因此, 在实际应用中, 应尽可能地避免对元素的局部轮廓进行采样, 特别是对小片段的圆弧或曲面采样, 此时的拟合结果以及误差评定结果很难保证其稳定性和唯一性。

4 结论

本文通过对常见几何元素的分析研究, 构建了其参数化的数学模型, 基于最小二乘法提出了其拟合算法的整体框架, 并对其拟合算法及优化求解时的初始值计算进行了详细论述, 最后以 NIST 提供的数据集作为测试数据, 对算法的拟合结果与 NIST 的拟合结果进行了比较分析, 其拟合误差在 1×10^{-4} 内, 表明算法的可行性, 可以将其作为零件形位误差评定的方法。

参考文献

[1] 史立新, 朱思洪. 基于 MATLAB 的平面度误差最小区域法评定. 控制与检测, 2005, 9:58-59.
 [2] 褚洪生, 杜增吉, 阎金华, 等. MATLAB7.2 优化设计实例指导教程. 北京: 机械工业出版社, 2007.
 [3] Craig M. Shakarji. Least-Squares Fitting Algorithms of the NIST Algorithm. Testing System. Journal of Research of the National Institute of Standards and Technology, 1998, 103(6):633-641.
 [4] The MathWorks. Optimization Toolbox User's Guide[EB/OL]. http://www.mathworks.com/access/helpdesk_r13/help/pdf_doc/optim/optim_tb.pdf. 2010.8.
 [5] New River Kinematics. NIST Test Data files, NIST paper, XIT files and more details[EB/OL]. <http://www.kinematics.com/products/spatialanalyzer/spatialanalyzer-documentation.html>. 2010.8.

(责编 小城)