

逆向工程中数据精简技术的研究

刘德平^{1,2}, 陈建军¹

(1. 西安电子科技大学机电工程学院, 陕西 西安 710071;

2. 郑州大学机电一体化研究所, 河南 郑州 450002)

摘要: 由于激光扫描技术的进步,可以方便地以较高精度和速度获取零件模型表面信息,对于产生的大量扫描数据,需要对其进行精简处理.为此提出了自适应最小距离精简方法,首先通过中值滤波和小波包滤波实现数据点的噪声去除,然后通过曲率分析,实现数据点云区域分割,对于分割后的区域选用不同的距离准则,如果数据点间的距离小于设定的准则,则该点删除从而实现数据精简.自适应最小距离法不仅保留了造型所需数据点的精度,同时数据精简效率为36%,最后通过实例验证了该方法的有效性.

关键词: 逆向工程;曲率分析;最小距离方法;数据精简

中图分类号: TP16 **文献标识码:** A **文章编号:** 1001-2400(2008)02-0334-06

Point data reduction technique in reverse engineering

LIU De-ping^{1,2}, CHEN Jian-jun¹

(1. School of Mechano-electronic Engineering, Xidian Univ., Xi'an 710071, China;

2. Mechatronics Institute of Zhengzhou Univ., Zhengzhou 450001, China)

Abstract: Advances in laser scanning technologies have facilitated sampling part surface data with speed and accuracy. It is necessary to manipulate these large amounts of point data. The adaptive minimum distance (AMD) method is proposed in this paper which is a kind of data reduction algorithm that balances efficiency and precision. Firstly the noise data is eliminated by median filtering and wavelet packet filtering, and then the curvature of the point data are analyzed and different zones are separated where a suitable minimum distance is selected. The points where the distance is larger than the given distance are neglected. This novel algorithm has merits of both precision and efficiency with the ratio of reduction being 36%. The method is applied to two sample models, and experimental results illustrate the feasibility of the new algorithm.

Key Words: reverse engineering; curvature analyzing; minimum distance method; data reduction

在逆向工程中,非接触方法可快速获取具有复杂曲面形状零件的大量几何数据,但是所获得的测量数据点云一般密度很大,需要很大的存储空间同时也降低了计算效率.在利用数据点云进行曲面重构、三角网格构造或者用于评价被测曲面的误差时,一般都不需要过密的数据点,特别是在被测曲面的曲率较小处.在重构曲面时,过密的点云不但计算量大,而且可能影响其光顺性,同时存储、处理和显示都将消耗大量的时间和计算机资源,因此,在逆向工程中测量数据的精简是重要的研究内容之一.

Sun^[1]等人采用包围盒法来简化测量点云,K. H. Lee^[2]等人通过点云法矢量的估算,运用八叉树数据结构,根据点云法矢量规则进行空间网格的细分,实现空间点云的精简.Eck^[3]利用简化和重构(STL)模型法来简化测量点云.笔者在分析前述数据精简方法的基础上,对于扫描线存储的数据点云,提出了自适应最小距离精简方法.首先对测量点的噪声数据采用中值滤波和小波包滤波处理,滤除干扰数据点得到实物模型的真实数据点云,然后利用MATLAB语言完成数据点的曲率计算并绘制数据点对应的曲率图形,根据曲率图形实现数据点云区域分段,对于不同的区域根据其特性设定不同的最小距离来实现数据精简.该方法不但保留了造型的精

收稿日期:2007-09-20

基金项目:国家“863”计划资助(2006AA04Z402);河南省教育厅自然科学基金资助(200510459060)

作者简介:刘德平(1966-),男,西安电子科技大学博士研究生,E-mail: liudeping66@163.com.

度,同时也提高了数据精简的效率,故可以处理曲率变化大、具有复杂附加特征的散乱测量数据.

1 数据精简准则

一般“点云”数据的排列形式大致可分为 3 类:① 阵列数据,数据具有行×列的特点;② 按光刀扫描线组织的数据,数据点基本上位于同一等截面线上,可认为是部分散乱数据;③ 完全散乱的、无组织数据.对于完全散乱数据点的压缩方法,大都集中在如何减少三角网格的数目问题上,对于第 1 类和第 2 类数据噪声点的处理,常按扫描线逐行实现数据的精简.

1.1 均匀和非均匀网格法

Martin^[4]等提出了一种基于“中值滤波”原理的网格法,现已广泛应用于图像处理.首先在垂直于扫描方向的平面上建立一系列均匀或非均匀的小方格,每一个扫描得到的点都被分配给某一个方格,计算出这一点到方格的距离 d ,按距离大小排列所有分配到同一方格的数据点,取距离位于中间值的数据点代表所有分配于这个方格中的数据点,其余点则被删除.过程如图 1 所示,C 点被选中,其余点则被删除.

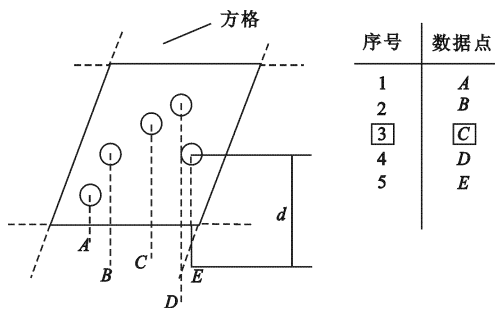


图 1 均匀方格法

1.2 三角网格法

在可视化或者分析一个扫描物体时,三角形网格模型被广泛使用,对于这些三角形模型,数据减少方法是基于减少三角形数量,然后对剩余的节点重新进行三角化.Chen^[5]等提出了一种减少三角网格数目从而删除部分数据点的方法.首先将扫描获取的数据点云直接网格化,生成 STL 文件,然后比较数据点所在三角面片的邻近三角面片法向量,根据一种向量加权算法,在平面或近似平面的较平坦区域用大的三角面片取代小的三角面片,删除多余点,从而实现数据精简.

1.3 按照数据点比例精简

对扫描线数据点按给定比例进行精简,其主要思想是根据零件的精度要求以及数据点的多少来确定数据的精简比例.一旦确定了精简比例,就可以以每条扫描线为单位,按照数据点的序列以给定的比例进行精简^[6].

假如一组序列 $X = \{X_0, X_1, \dots, X_n\}$,若精简比例为 3:1,那么 $Y = \{X_0, X_3, \dots, X_{3,i}\}$ 序列为保留的数据序列,即每隔两个点保留一个点.这种方法原理简单,易于实现.只是没有考虑零件的形状特征,也容易造成零件的严重失真.所以,适合于精度要求比较低的场合.

1.4 按平均距离精简

基本原理仍然以每条扫描线为单位,依次计算点与点之间的距离 d_i ,平均距离为

$$d = \frac{1}{n} \sum_{i=0}^n d_i \quad (1)$$

若 $d_i \leq d$,则点 x_i 将被删除;相反,则保留.这种方法原理也很简单,但是无法保证零件的细节特征,不适合数据密集以及精度要求很高的场合.

1.5 按最小距离精简

按最小距离精简原理和按平均距离精简原理近似:只是前者是先设定一个最小距离 d_{\min} ,然后沿扫描线方向顺序比较相邻两点之间的距离 d ,若 $d < d_{\min}$,则两点中的一个点将被删除,依次判断所有的扫描点.其中,最小距离的确定应根据具体数据精简的精度要求来选取.当采用最小距离法进行数据精简时,为了保证产品精度,只能采用较小的最小距离,数据的精简效率会大大减弱;为了保证效率,选用较大的最小距离,则不能更好地保持产品细节特征,这是该方法的不足之处.

2 自适应最小距离法

在分析前述数据点精简的基础上,笔者提出了自适应最小距离法,首先对测量点的噪声数据采用中值滤

波和小波包滤波处理,再按照曲率的变化情况把数据点分成不同的区域,在此基础上分别对不同区域的数据点采取不同的最小距离进行数据精简,也就是自适应最小距离法.如图 2 所示为采用自适应最小距离法进行数据精简的流程.

2.1 噪音点去除

在数据获得过程中,由于受到人为或随机因素的影响,会引入不合理的噪音点.为了降低或消除其对后续数据处理的影响,有必要对所获数据点进行噪音点去除.针对扫描线点云数据提出先加权中值滤波,再小波包分析滤波的新滤波方法.加权中值滤波是把数据点的 z 坐标看成图像处理中的灰度,从而借鉴图像处理的成熟理论进行数据优化,可以把分散的脉冲噪声完全除去.小波包分析是将信号的小波包分解、计算最佳小波包基、分解系数的阈值量化处理及重构原来信号实现消噪,以去除彼此靠近的脉冲噪声.

2.1.1 加权中值滤波

中值滤波是把数字图像或数字序列中一点的值用该点邻域中各点值的中值替代,其二维中值滤波的数学表达式为

$$y_{ij} = \text{Med}\{x_{ij}\} = \text{Med}\{x_{(i+r),(j+s)}, (r,s) \in A, (i,j) \in I^2\} \quad (2)$$

其中 A 为滤波窗(常用的有方形、圆形、十字形等); $\{x_{ij}, (i,j) \in I^2\}$ 表示数字图像各点的灰度值.

$$Y = \text{Med}[W_1 \diamond X_1, W_2 \diamond X_2, \dots, W_N \diamond X_N] \quad (3)$$

其中 Med 表示中值操作, X_1, X_2, \dots, X_N 为按大小顺序排好的序列, W_i 为其对应权值(权值为非负整数), \diamond 表示复制,即如下所示

$$K \diamond X = \overbrace{X, \dots, X}^{K \text{次}} \quad (4)$$

范围为 N 的加权中值滤波输出 Y 可用式(3)表示,加权中值滤波是对滤波窗口内的像素排序,然后对每一像素 X_i 相应其权值 W_i 进行复制,从新的序列中选择中值作为输出.

在简单的中值滤波中,由于没有权重,滤波对于给定窗口内的所有 X 值,具有相同的影响.若选用加权中值滤波的方法,使窗口的中心点具有较大的权值,从而使之对输出结果具有较大的影响.

利用复杂度低的矩形窗口,即以矩形窗 9 个点的加权中值来代替其中心点.如图 3 所示,移动矩形窗,遍历 3 条扫描线的所有点,直到 3 条扫描线的其中一条达到末尾,比如窗 A ,则其中心点 $P_{i+1,j+1}$ 的坐标 Z 值以其 9 个点的中值代替,其中考虑到中心点的邻近点对其影响较大,所以对矩形窗对角线上 4 个点的权值取为 1,对中心点取 3,另外 4 个点则取为 2.然后将矩形窗下移一条扫描线,从头开始对其进行同样的步骤,遍历所有扫描线,进行加权中值滤波.各点权重的最佳值可由遗传算法通过优化来确定,以达到改善滤波器精度的效果.

由于中值滤波是用滤波器窗的中值代替其中心点,故充分分散的噪声脉冲将完全被加权中值滤波器去掉.而彼此靠近的脉冲可能保留下来,所以有些情况通过中值滤波不能得到良好的效果,可在加权中值滤波之后,采用小波包分析方法再进行滤波以消除这种噪声.

2.1.2 小波包分析

小波包分析是从小波分析延伸出来的一种对信号进行更加细致的分析和重构方法.它将信号分解为近似与细节两部分,近似和细节同时又可以继续分解,依次类推.不但对低频部分分解,还对高频细节进行分解.应用小波包进行消噪处理主要包括以下几步:

- (1) 小波包分解 选择一个小波并确定所需分解的层次,然后对信号 X (扫描线)进行小波包分解.
- (2) 确定最优小波包基 对于一个给定的熵标准,计算最佳树.
- (3) 小波包分解系数的阈值量化 对每一个小波包分解系数,选择一个恰当的阈值进行阈值处理.

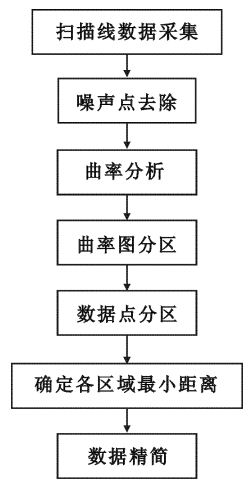


图 2 自适应最小距离法流程图

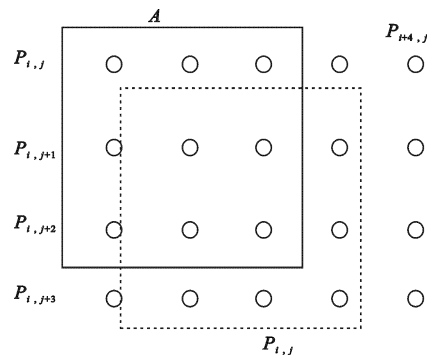


图 3 矩形窗口的作用

(4) 小波包重构 根据最低层的小波包分解系数和量化处理系数,进行小波包重构。

在 WindowXP 系统下,通过 Matlab6.5 Add-in 实现 VC++ 与 Matlab 的混合编程,既保持了 Matlab 的优良算法,提高了执行速度,又能很好地运用 VC++ 编制界面的优越性。

2.2 数据点分区

运用 MATLAB 进行数据点曲率分析来实现数据点分区,可把数据点分为平坦区、突变区和过渡区。具体实现步骤:

(1) 曲率计算^[7] 对于二维曲线 $y = f(x)$, 曲率可以被定义为

$$k = \left(\frac{d^2y}{dx^2} \right) / \left[1 + \left(\frac{dy}{dx} \right)^2 \right]^{3/2} = \frac{f''}{[1 + (f')^2]^{3/2}} \quad (5)$$

同一条扫描线上的数据点可以认为在同一平面上,取任意连续 3 点可确定一个圆如图 4 所示,其圆心坐标可通过这 3 个点的坐标求得

$$x_0 = \frac{a-b+c}{d}, \quad y_0 = \frac{e-f+g}{-d} \quad (6)$$

其中 $a = (x_1 + x_2)(x_2 - x_1)(y_3 - y_2)$,
 $b = (x_2 + x_3)(x_3 - x_2)(y_2 - y_1)$,
 $c = (y_1 - y_3)(y_2 - y_1)(y_3 - y_2)$,
 $d = 2[(x_2 - x_1)(y_3 - y_2) - (x_3 - x_2)(y_2 - y_1)]$,
 $e = (y_1 + y_2)(y_2 - y_1)(x_3 - x_2)$,
 $f = (y_2 + y_3)(y_3 - y_2)(x_2 - x_1)$,
 $g = (x_1 - x_3)(x_2 - x_1)(x_3 - x_2)$ 。

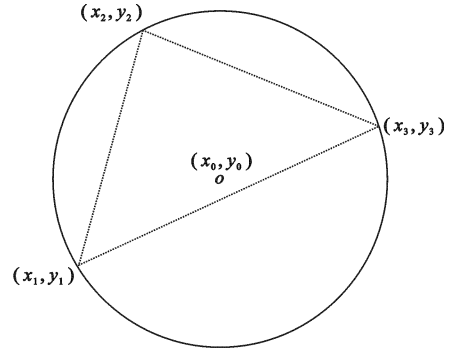


图 4 曲率计算原理图

则 X_2 点的曲率值可定义为 $k = 1/r = 1 / \left((x_0 - x_2)^2 + (y_0 - y_2)^2 \right)^{1/2}$ 。 (7)

(2) 曲率分析 在 MATLAB 中,根据上一步曲率计算公式,通过编程来实现数据点曲率计算,运行后即可显示曲率图。

(3) 由曲率分析图实现对应数据点分区 通过曲率分析图可以看出曲率的变化情况,同时结合精度要求把曲率分析图分成不同的区域:即曲率变化大的区域(突变区);曲率变化较大的区域(过渡区);曲率变化平缓的区域(平坦区)。曲率图分区的同时,则对应数据点也就被分成相对应的区域,从而实现数据点的分割。当然,考虑到区域之间的过渡问题,可以采取包容原理:即突变区可以包容过渡区和平坦区;过渡区可以包容平坦区。在分区时,可稍延长包容区的两端,从而实现区域之间的平滑过渡。

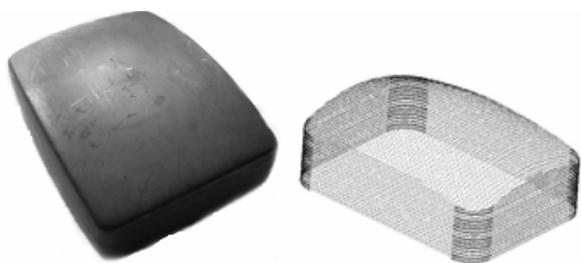
2.3 自适应最小距离数据点精简

首先,根据精度要求和曲率变化对每一区域的数据点选取合适的最小距离。其基本原则为:在突变区和过渡区采用较小的最小距离,能够更好地保留曲面的细节特征;而在平坦区采用较大的最小距离,能在保持精度的前提下提高数据精简的速度。用最小距离法分别对不同区域的数据点进行精简。这样,也就实现了在曲面曲率较小的区域保留较少的点,而在曲率较大或尖锐棱边处保留较多的点,不但继承了最小距离法速度快而且有效的优点,同时又很好地保留了原始数据的细节特征。

3 应用实例

采用智泰公司的 LSH300 激光扫描测量机测量某显像管玻壳模具,该测量机基于光学三角形原理,其测量精度为 $\pm 0.05 \text{ mm}$,图 5 所示为显像管玻壳实物模型和扫描点云。

图 6 是从图 5 特征线扫描点云中提取出的一条特征线扫描点(46 个点)和去除噪音点后的扫描点(41 个点)。用 MATLAB 软件编程实现扫描点云的曲率分析,如图 7 所示,然后把扫描点分为 3 段:第 1~14 个点和第 27~41 个点曲率变化非常大(属突变区),选取最小距离 d_1 进行数据精简;第 14~27 个点曲率变化较大(属过渡区),选取最小距离 d_2 进行数据精简,其中 $d_1 < d_2$ 。



(a) 玻璃模具实物模型 (b) 测量点云数据

图 5 测量点云和实物模型

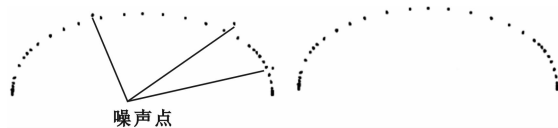


图 6 一条扫描线点云和去除噪声后的扫描线点云

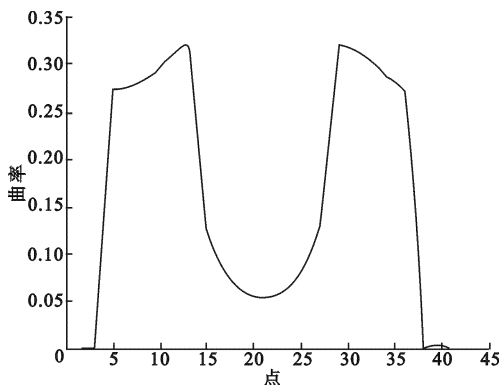


图 7 曲率分析图

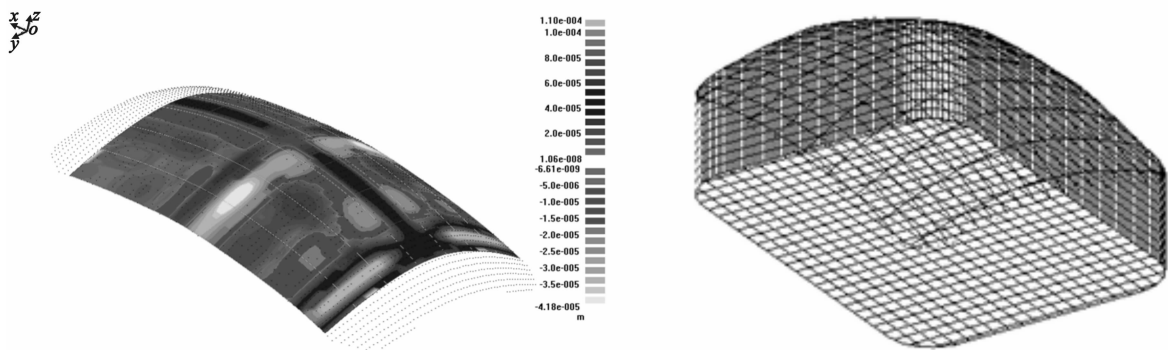


图 8 最小距离法精简后和自适应最小距离法精简后的扫描点

图 8 是按最小距离法进行数据精简后的扫描点(32 个点)和按自适应最小距离 $d(d_1 < d < d_2)$ 进行数据精简后的扫描点(26 个点). 从数据精简处理结果可以看出自适应最小距离法的处理过程, 与最小距离法进行了比较(见表 1). 数据点处理后重构的曲面模型与数据点云误差如图 9(a)所示, 重构的曲面模型如图 9(b).

表 1 结果比较

精简方法	最小距离/mm	突变区点数		过渡区点数		突变区点数		总点数		精简效率/%
		前	后	前	后	前	后	前	后	
最小距离法	0.23	14	10	13	13	14	9	41	32	21.95
自适应最小距离法	0.15 0.83 0.15		10		6		10		26	36.59



(a) 重构曲面与扫描点云误差对比图

(b) 曲面模型

图 9 重构误差和曲面模型

4 结 论

(1) 实践应用表明, 采用笔者提出的自适应最小距离法可以对逆向工程中非接触式获取的密集测量数

据进行直接、有效的精简. 由于针对不同的曲率分段区域, 曲率大的区域采用较小的最小距离, 而曲率小的区域采用较大的最小距离, 则该方法既具有较高的精简效率, 也保留了数据点的精度, 适用于三维扫描线数据点云, 精简效果良好.

(2) 从表 1 和图 9 可以看出, 当采用最小距离法进行数据精简时, 为保证精度, 只能采用较小的最小距离, 数据的精简效果会大大减弱; 为了保证效率, 选用较大的最小距离, 则不能更好地保持产品细节特征, 其精简效果不如笔者提出的自适应最小距离法.

(3) 笔者提出的算法还存在一些不足, 例如对数据点密度变化过大的点云精简效果有待进一步提高, 如何根据曲面的特性, 运用优化方法自动确定每个区域的最小距离, 这将是未来研究的目标.

参考文献:

- [1] Sun W, Bradley C, Zhang Y F, et al. Cloud Data Modeling Employing a Unified Non-redundant Triangular Mesh[J]. Computer-Aided Design, 2001, 33(2):183-193.
- [2] Lee K H, Woo H, Suk T. Point Data Reduction Using 3D Grids [J]. The International Journal of Advanced Manufacturing Technology, 2001, 18(3): 201-210.
- [3] Eck M, DeRose T, Duchamp T, et al. Multiresolution Analysis of Arbitrary Meshes[C]//The Proceedings of SIGGRAPH Computer Graphics. New York: ACM Press, 1995: 82-90.
- [4] Lee K H, Woo H, Suk T. Data Reduction Methods for Reverse Engineering[J]. The International Journal of Advanced Manufacturing Technology, 2001, 17(3): 735-743.
- [5] Chen Y H, Neg C T, Wang Y Z. Data Reduction in Integrated Reverse Engineering and Rapid Prototyping [J]. The International Journal of Computer Integrated Manufacturing, 1999, 12(2): 97-103.
- [6] 张丽艳, 周儒荣, 蔡伟斌, 等. 海量测量数据简化技术研究[J]. 计算机辅助设计与图形学学报, 2001, 10(9): 20-40.
Zhang Liyan, Zhou Rurong, Cai Weibin, et al. Research on Cloud Data Simplification[J]. Journal of Computer Aided Design & Computer Graphics, 2001, 10(9): 20-40.
- [7] Budak I, Hodolic J, Sokovic M. Development of a Programme System for Data-point Pre-processing in Reverse Engineering[J]. Journal of Materials Processing Technology, 2005, 162-163(5): 730-735.

(编辑: 齐淑娟)

简 讯

- ❖ 美国莱斯大学计算与应用数学系印卧涛博士于 2007 年 12 月 26 日~27 日来我校进行学术交流活动.
- ❖ 美国福特汽车公司戴扬博士 2007 年 12 月 30 日来我校讲学. 戴扬博士多年来一直从事汽车变速箱振动噪声的 CAE 分析, 在振动哭声分析以及结构优化方面有着丰富多彩实践经验.

摘自《西电科大报》2008. 1. 10