

基于随机分布估计的点云密度提取

叶爱芬, 龚声蓉, 王朝晖, 刘纯平

(苏州大学计算机科学与技术学院, 苏州 215006)

摘 要:针对目前密度提取方法提取的密度信息不能表现点云局部分布信息和分布随机性的缺陷, 提出结合随机分布估计的密度提取方法。该方法采用分块计数法得到每个小分块的密度, 结合点云总体的密集度得到一个能够反映点云局部积聚特征的参数, 为判别点云分布的随机性、均匀性等提供较好的特征依据。

关键词: 离散点云; 点云密度; 点云随机分布估计

Point Cloud Density Extraction Based on Stochastic Distribution Estimation

YE Ai-fen, GONG Sheng-rong, WANG Zhao-hui, LIU Chun-ping

(School of Computer Science & Technology, Soochow University, Suzhou 215006)

【Abstract】 Density extraction method has difficulty in representing local distribution and its stochastic feature from the extracted density information. This paper proposes a solution to solve this problem, combing density method with stochastic distribution estimation. The method computes the density of each single small plot, and combines it with the overall density of the point cloud. A parameter is obtained, which can reflect the local aggregation feature. Tests show that this parameter can satisfactorily provide reliable data on estimating stochastic distribution and homogeneity of the point cloud.

【Key words】 scattered point cloud; point cloud density; stochastic distribution estimation of point cloud

1 概述

随着光学点阵扫描仪和其他形状获取设备的广泛应用, 常常需要对各种扫描点云^[1]做处理, 如点云精简、点云分割、点云聚类、点云识别。

对于提取能够反映点云分布随机性的密度参数方面, 国内外并未进行深入系统的研究, 但其实际应用需求却很多。如在企业的自动化生产流程控制中, 常用激光扫描^[2]确定产品品质是否正常; 机械电子企业常需判别单晶硅片表面是否有划痕或缺陷, 对此可以通过检测其表面光学扫描点云^[3]中是否包含非随机均匀分布的特征结构来判别其质量是否正常。由于在生产过程中, 原本光洁的硅片表面可能因为生产线故障而产生划痕和缺陷, 或在存储过程中其表面积聚了灰尘, 因此扫描产生的二维扫描点云中常包含分布均匀的灰尘点云和非随机分布的划痕点云、缺陷点云等多种类型的点云, 如图 1~图 4 所示。在这 4 种扫描点云中, 只有图 1 可以被视为质量正常的产品, 其他 3 种都应被视为质量异常而需要发出警报。

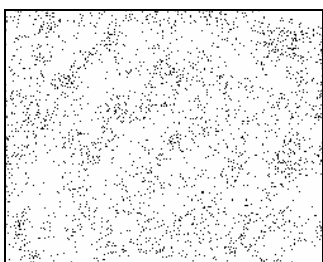


图 1 均匀分布的灰尘点云

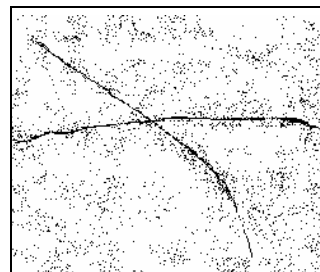


图 2 包含划痕和灰尘的点云

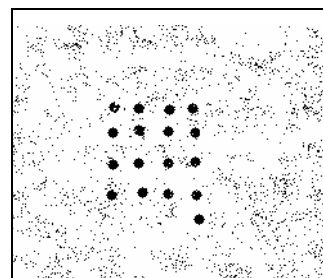


图 3 包含缺陷和灰尘的点云

基金项目: 国家自然科学基金资助项目(60673092); 教育部科研基金资助重点项目(205059); 江苏省高校自然科学基金资助项目(07KJ D520186)

作者简介: 叶爱芬(1982 -), 女, 硕士研究生, 主研方向: 模式识别; 龚声蓉, 教授、博士; 王朝晖, 副教授、硕士; 刘纯平, 副教授、博士

收稿日期: 2008-07-10 **E-mail:** ye209@cs.uregina.ca

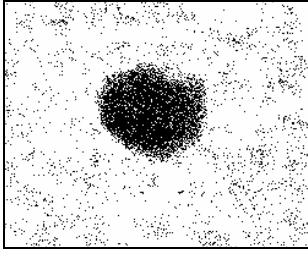


图4 局部大量沉积的灰尘点云

目前对包含非随机分布的点云的研究主要集中于针对其中特征结构的高效提取方法,对点云本身分布是否随机很少关注。然而,在一些实际应用中,往往对点云所包含的结构并不感兴趣,相反,需要对该点云的分布是否随机进行判别,此时提取出能表现这些点云分布和密度信息的特征参数是精确判别这些点云的关键。本文提出的结合随机性判别的点云的密度就是一个可以有效区分点云分布的随机性和密集性的参数。

本文在归纳基密度的提取方法(基于分块的点云密度提取和基于距离的点云密度表示方法)的基础上,提出了能够满足随机性判别、充分体现局部密度特征的改进密度提取方法,并将其应用于单晶硅片表面扫描点云的判别中。

2 现有的点云密度特征提取方法

目前,点云密度特征提取方法分为2类:基于距离的点云密度表示方法和基于分块的点云密度提取方法。

2.1 分块点云密度提取方法

分块点云密度提取方法是将点云空间均匀地分成若干块,然后分别计算出落在各块内的点数目,以这些点的数目为基础做取舍组合,得到点云的密度特征。根据空间划分的不同和最后各块数据整合方式的不一样,分块点云密度提取有多种不同的方法,如可以依据各点为中心将该点周围一定范围内的领域划分为一个计数块,也可以将空间平均地分成具有一定边长的小方块来计数。对计数块数据的整合取舍也可以有多种方式,如取各分块内点个数的平均值或各分块点数的最大值、最小值。

2.1.1 以点为中心的邻域分块法

以点为中心的邻域分块法^[4]按照如下方法定义点云的密度:

定义 1 点云: $S = \{p_i\}_{i=1}^N$ 为已知平面上一组离散的带误差的无序采样数据点。

定义 2 点的密度:对空间中任意点 p 和距离 $Radius$,以 p 点为中心、半径为 $Radius$ 的区域内的点的个数称为点 p 基于距离 $Radius$ 的密度,记作 $density(p, Radius)$ 。

定义 3 包含 N 个点的空间点云 S 的平均密度为

$$\overline{density}_S = \frac{1}{N} \sum_{i=1}^N density(p_i, radius) \quad (1)$$

其中, $density(p_i, radius)$ 是第 i 个点的点密度。

对于某些应用,可能需要得到点云的最大、最小密度。

定义 4 对包含有 N 个点的空间点云,其最大密度为

$$density_{\max} = \max(density(p_i, radius), i = 1, 2, \dots, N) \quad (2)$$

最小密度为

$$density_{\min} = \min(density(p_i, radius), i = 1, 2, \dots, N) \quad (3)$$

2.1.2 点云空间平均分块法

点云空间平均分块法是目前应用得最多也是最简单的点云密度提取方法。分块的思想比较直观,假设二维点云 S 的

大小为 $N \times M$,拟将它分为大小为 $n \times m$ 的小块,则总共可以分为 $\left\lceil \frac{N}{n} \right\rceil \times \left\lceil \frac{M}{m} \right\rceil$ 块,然后分别计算落在各块内点的个数。假设用 $density_p$ 表示落在第 p 块内点的个数,则有如下定义:

定义 5 若将大小为 $N \times M$ 的二维空间点云 S 划分为 $n \times m$ 的小分块,则点云的平均密度为

$$\overline{density}_S = \frac{1}{N} \sum_{p=1}^{\left\lceil \frac{N}{n} \right\rceil \times \left\lceil \frac{M}{m} \right\rceil} density_p \quad (4)$$

定义 6 将大小为 $N \times M$ 二维空间点云 S 划分为 $n \times m$ 的小分块,则点云的最大密度为

$$density_{\max} = \max(density_p, p = 1, 2, \dots, \left\lceil \frac{N}{n} \right\rceil \times \left\lceil \frac{M}{m} \right\rceil) \quad (5)$$

最小密度为

$$density_{\min} = \min(density_p, p = 1, 2, \dots, \left\lceil \frac{N}{n} \right\rceil \times \left\lceil \frac{M}{m} \right\rceil) \quad (6)$$

可以发现,该方法对于点云的分布信息的表现能力非常局限。若用点云的平均密度 $\overline{density}_S$ 表示一个点云的分布,则 $\overline{density}_S$ 是全局密度的平均值,完全无法表现局部密集的点云 $\overline{density}_S$ 的分布具有局部密集性。反之,如果用 $density_{\max}$ 或 $density_{\min}$ 表示,则对局部太敏感,以偏概全的表现结果往往是与实际分布相背离的,因此,这2个参数只能作为得到点云全局分布信息后的辅助参数。

2.2 基于距离的点云密度表示方法

基于距离的点云密度表示是以点云内各点间的距离为度量标准的一种密度表示方法。由于各点的距离有最小距离、最大距离、平均距离等多个可取参数,而对各点的距离也可以不同的方式进行整合取舍,如取其平均值、最大值、最小值,因此基于距离的点云密度表示方法也有多种。下文只对平均距离密度表示法给予简要的介绍。

平均距离密度表示法对点云 S 内各点的距离取平均值得到平均距离,并以此表现点云分布的紧密程度。此处点云 S 内各点的距离可以根据需求自行选取,一般取点 p 与点云 S 中距离点 p 最近的点之间的距离为点 p 的距离。

对于由 N 个点构成的点云 S ,以 $distance(p, q)$ 表示 p 点和 q 点之间的距离,则有如下定义:

定义 7 p 为点云 S 中的任一点,定义点 p 的 r -邻域为 $B_r(p) = \{q \mid distance(p, q) < r, q \in S\}$ (7)

定义 8 假设以 d_p 表示点云 S 中 p 点与其他点的最小距离,则

$$d_p = \min(distance(p, q), q = 1, 2, \dots, N \text{ 且 } q \neq p) \quad (8)$$

定义 9 点云 S 中点间的平均距离为

$$\bar{d} = \frac{1}{N} \sum_{p=1}^N d_p \quad (9)$$

定义 10 点云 S 中点间的最大距离为

$$d_{\max} = \max(d_p, p = 1, 2, \dots, N) \quad (10)$$

最小距离为

$$d_{\min} = \min(d_p, p = 1, 2, \dots, N) \quad (11)$$

点云中各点间的距离越小,则点的分布越集中,因此,点云的密度就比较大;相反,点间距离越大,说明点云的密度越小。然而该方法在提取某些稀疏点云密度时存在很大的局限性。例如对于一个成对出现的稀疏点云,假设各点对之间的距离都很大。由于成对出现,在各自很小的领域内都有一个

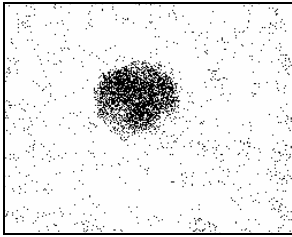
与之距离很近的点,因此各点的 d_p 都很小,整体 \bar{d} 也就很小,由此出现了点云的 \bar{d} 很小但却是稀疏点云的情形,从而导致对点云密度性质的误判。

3 改进的点云密度提取算法

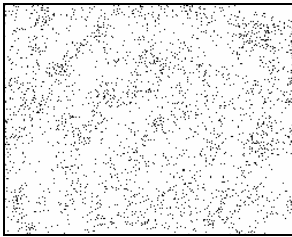
通过以上分析可以发现,现有的方法都存在密度均匀化现象,即局部的分布信息被均匀化,以至最终得到的密度信息既不能确切地表现总体的分布状态,也不能表现局部的分布密集度情况^[5]。对此本文提出距离密度结合距离方差和改进整体分块密度2个解决方案。

3.1 距离密度结合距离方差方案

图5(a)和图2所示的点云通过距离密度表示和分块密度表示法提取得到的密度没有太大区别,但是2个点云的分布有明显的差异,如何表现这2个点云的分布和密度差异便是需要解决的问题。



(a)局部密集但总体密度不大的点云



(b)分布均匀但总体密度与图5(a)相当点云

图5 2个整体平均密度相近的点云

从统计学的角度看,方差是表现数据分布均匀性的有力工具,而点云中各点的距离也是一个可以统计的数据集,因此,可以引进方差衡量点云中各点距离的均匀性

$$Var(d) = \frac{1}{N} \sum_{p=1}^N (d_p - \bar{d})^2 \quad (12)$$

方差 $Var(d)$ 越小,说明点云内各点分布得越均匀。于是通过 \bar{d} 和 $Var(d)$ 这2个参数就能分析得到点云的分布及密集度状况,具体如表1所示。

表1 \bar{d} , $Var(d)$ 与点云分布的关系

\bar{d} 取值	$Var(d)$ 取值	点云的分布状态
大	小	均匀分布的稀疏点云
大	大	分布不太均匀的稀疏点云
小	小	分布较均匀的密集点云
小	大	分布不均匀的密集点云

3.2 改进的整体分块密度特征提取方法

对于类似于交叉线形分布的点云,如果用平均最小距离+距离方差法表示该点云的密度,则它的平均最小距离和方差都不大,但它并不是密集均匀分布的点云。通过观察就可以发现,其局部(构成点云中线状的点云)是由非常密集的点组成的,而其他区域(线以外的区域)几乎没有点。因此,本文引进另一个密度计算方法——改进的分块密集度提取方法。

设 S 是具有 N_T 个点,分布范围为 $N \times M$ 的点云,并被均

匀分割成大小为 $n \times m$ 的块,用 B_i 表示有点落在其中块的小分块个数,用 B_t 表示总的分块数,则 $B_t = \left\lceil \frac{N}{n} \right\rceil \times \left\lceil \frac{M}{m} \right\rceil$ 。

紧密度 ρ 表示空间中点云 S 的密集情况,这里用 $\frac{B_d}{B_t}$ 表示,即有点落在其间的分块个数占总体的比例。

块内密度:表示每个分块内部点的密集程度,本文仅用块内的点个数 V_{ij} 表示,其中, i, j 块表示分块索引,即位于点云中第 i 行 j 列的小分块。

平均块内点密度:表示整个点云内落在大小为 $n \times m$ 区域内的平均点的个数。用 V_a 表示,则

$$V_a = \frac{1}{B_t} \sum_{i=1}^{\left\lceil \frac{N}{n} \right\rceil} \sum_{j=1}^{\left\lceil \frac{M}{m} \right\rceil} V_{ij} \quad (13)$$

块内点密度方差:用以衡量各块内点云分布密度的差异性。用 Var_S 表示,则

$$Var_S = \frac{1}{B_t} \sum_{i=1}^{\left\lceil \frac{N}{n} \right\rceil} \sum_{j=1}^{\left\lceil \frac{M}{m} \right\rceil} (V_{ij} - V_a)^2 \quad (14)$$

实验表明,应对点云局部的密集分布信息给予特别重视,因此,在描述其分布密度信息的过程中必须给予它较大的权重;而稀疏部分是噪声点的可能性比较大,因此,需要适当降低其对整体密度的影响。本文采用各分块点数的加权平均数衡量总体密度。另外,点云总体的紧密度参数对密度也有反面影响。如果点云中包含的点比较多、整体紧密度小,则说明点云内大部分点分布在较少的几个分块内,因此,局部密集度非常大。本文给出如下点云密度计算的方法:

$$density_S = \frac{1}{N_T} \frac{B_d}{B_t} \sum_{i=1}^{\left\lceil \frac{N}{n} \right\rceil} \sum_{j=1}^{\left\lceil \frac{M}{m} \right\rceil} V_{ij}^2 \quad (15)$$

实验显示,具有相似分布的点云具有相似的 $density_S$ 属性,而不同分布的点云其 $density_S$ 属性区别也较大。并且分布越不均匀、越紧密, $density_S$ 越大。

值得注意的是, $density_S$ 受分块大小的影响较大,本文采用的分块方法是先确定分块大小,从而自动确定分块的个数,实际应用中可以考虑实际情况加以选用。

对于分为 $B_t = \left\lceil \frac{N}{n} \right\rceil \times \left\lceil \frac{M}{m} \right\rceil$ 块的点云 S ,本文结合 $V_a, Var_S, density_S$,采用模糊分类策略判别点云的性质类型。其中, $V_a, Var_S, density_S$ 取值大小的判定阈值分别为 $\varphi_v, \varphi_{var}, \varphi_{density}$,即取值大于该阈值时将其标记为大,否则为小,如表2所示。

表2 $V_a, Var_S, density_S$ 的取值大小与点云类型的判定关系

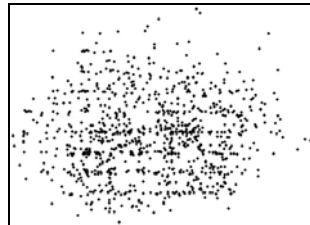
V_a	Var_S	$density_S$	点云类型	P
大	大	大	密集、非均匀、非随机	大
大	大	小	密集、随机、非均匀	大
大	小	大	密集、较均匀、非随机	大
大	小	小	密集、均匀、随机	小
小	大	大	局部密集、非随机	大
小	大	小	非均匀点云、不含非随机结构	小
小	小	大	局部密集,包含非随机结构	大
小	小	小	稀疏随机均匀分布点云	小

在生产线上自动判别单晶硅片表面扫描点云是否异常的过程中,均匀分布的灰尘点云和光洁的表面都应判为质量正常;而带有严重局部沉积的灰尘、划痕以及缺陷特征点云系统应发出警告,自动判为质量异常。因此,以 P_S 表示某一产品

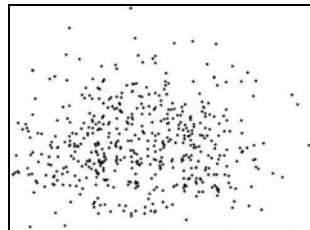
为质量异常的可能性，初始化为 0，即默认为质量正常。 P_s 越大，产品质量异常的可能性越大。根据 $V_a, Var_s, density_s$ 的取值，通过模糊逻辑给出 P_s ，当 $density_s$ 很大时，产品异常的可能性非常大，因此，为 P_s 赋较大值； V_a, Var_s 取值大而 $density_s$ 取值较小的点云一般是具有严重局部沉积的灰尘点云，密集随机但非均匀分布，此时，也为 P_s 赋较大的值。

3.3 实验结果分析

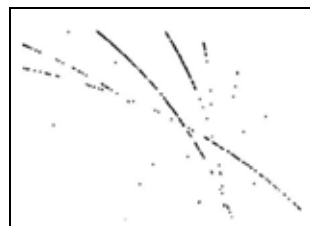
通过证明该方法具有很强的区分点云分布特性的能力，表 3 分别列出图 6 所示点云的实验结果。



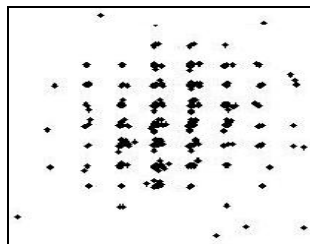
(a) 无定型随机分布的点云



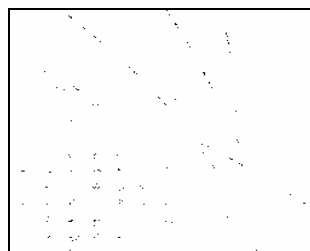
(b) 另一无定型随机分布的点云



(c) 交叉线形分布的点云



(d) 不规则非随机分布的点云



(e) 随机稀疏点云

图 6 随机均匀分布点云和非随机分布点云

表 3 图 6 中各点云的数据

点云	V_a	Var_s	$density_s$	点云类型	P
图 6(a)	0.332 9	0.332 9	3.259 6	密集随机均匀	小
图 6(b)	0.237 3	0.371 2	2.790 2	密集随机均匀	小
图 6(c)	0.061 2	0.199 5	8.125 0	局部包含非随机(线形)结构	大
图 6(d)	0.187 1	0.677 4	7.294 7	非随机非均匀(缺陷)	大
图 6(e)	0.031 5	0.057 9	4.870 5	稀疏随机	小

由上述结果可以发现，如果仅以 V_a 和 Var_s 表示点云的分布密度信息，则图 6(c) 被判别为均匀分布的稀疏点云，这明显有违实际情况。而引入 $density_s$ 后，可以非常准确地判定分布密度相似的点云和分布密度不同的点云，本方法的优越性显而易见。

表 3 的 $V_a, Var_s, density_s$ 证明该方法在用于区分随机均匀分布的点云和非随机分布点云时表现出很好的性能。对于局部积聚比较明显的点云， $density_s$ 取值较大；对于均匀分布的点云， $density_s$ 的取值较小，应用中可以用 $V_a, Var_s, density_s$ 共同判定点云的密度和分布状态。

实验取 $\varphi_v = 0.1, \varphi_{var} = 0.5, \varphi_{density} = 5.0$ 。由表 3 可知，图 6(a)、图 6(b) 的 V_a 较大， Var_s 和 $density_s$ 都较小，由表 2 可知是密集均匀分布点云，这与事实相符；图 6(c) 的 V_a 和 Var_s 取值小，但 $density_s$ 很大，因此，是局部密集包含非随机结构的点云，这也与实际相符。类似地，图 6(d)、图 6(e) 可以分别判断为密集非随机非均匀分布点云和系数的随机均匀分布点云。 P_s 的值只在图 6(c)、图 6(d) 中取大，因此，系统只对这 2 件产品做警报，将其判为质量异常，这与实际情况相符。图 6(c) 为划痕点云，图 6(d) 为非随机的缺陷点云。

4 结束语

基于距离和基于空间分块的密度提取方法在表现点云密度的能力方面都有其独到之处，但对于非随机均匀分布的点云和局部密集的点云无能为力。本文在归纳密度提取方法的同时，针对这些局限性提出改进的密度特征提取方法，实验表明其在表现点云分布的密集度、随机性等方面具有良好的性能，特别是对于局部非常密集而其他区域非常稀疏的点云，判别效果非常显著。因此，可以作为一个很好的点云分布密度参数。

参考文献

- [1] Theodoridis S, Koutroumbas K. Pattern Recognition[M]. 3rd ed. [S. l.]: China Machine Press, 2006.
- [2] Sotoodeh S. Outlier Detection in Laser Scanner Point Clouds[J]. [2008-08-10]. http://www.isprs.org/commission5/proceedings06/paper/SOTO_653.pdf.
- [3] Klein J A, Zachmann G. Point Cloud Surfaces Using Geometric Proximity Graphs[J]. Computers & Graphics, 2004, 28(6): 839-850.
- [4] 马 帅, 王腾蛟, 唐世渭, 等. 一种基于参考点和密度的快速聚类算法[J]. 软件学报, 2003, 14(6): 1089-1095.
- [5] Kobler A, Pfeifer N, Ogrinc P, et al. Using Redundancy in Aerial Lidar Point Cloud to Generate DTM in Steep Forested Relief[C]// Proc. of Workshop on 3D Remote Sensing in Forestry. Vienna, Austria: [s. n.], 2006.