

基于点云的谷粒高通量表型信息自动提取技术

黄霞^{1,2} 郑顺义^{1,2} 桂力^{1,2} 赵丽科¹ 马浩¹

(1. 武汉大学遥感信息工程学院, 武汉 430079; 2. 地球空间信息技术协同创新中心, 武汉 430079)

摘要: 在进行水稻的数字化考种、表型与基因关联分析和数字农业仿真模拟时,需要大量的谷粒表型信息作数据支撑。本文提出了一种基于三维点云的谷粒高通量表型信息自动提取方法,能同时自动获取谷粒的三维模型和40个表型参数,实现谷粒形状的定量和定性描述。首先,通过对谷粒点云数据进行聚类分析,完成谷粒点云的分类;其次,实现谷粒的三维重建,对谷粒离散点云进行柱面构网,获取谷粒点云的三维模型数据;最后,根据不同表型参数的特点,实现了谷粒的三维表面积和体积、长、宽、高、3个主成分剖面的周长和面积等11个基本参数与长宽比、长高比和体积比等11个衍生参数以及18个形状因子的自动提取。利用 Handyscan 700 型手持式激光扫描仪获取的谷粒高精度点云数据进行实验,成功实现了谷粒表型参数的自动提取,测量结果可达毫米级。基于主成分方法分析了各表型参数的权重。以游标卡尺测量值和 Geomagic Studio 测量值作为真值,长、宽、高的平均相对误差为 1.14%、1.15% 和 1.62%, 体积和表面积的平均相对误差为零,3个主成分剖面面积的平均相对误差为 1.82%、2.12% 和 2.43%。本文方法与人工测量方法及软件测量方法相比,精度相当,且具有批量、自动、人工干预少(仅数据采集阶段需要人工操作)以及效率高的特点。

关键词: 谷粒; 表型信息; 点云; 自动化; 批处理

中图分类号: TP391.1; S818 **文献标识码:** A **文章编号:** 1000-1298(2018)04-0257-08

Automatic Extraction of High-throughput Phenotypic Information of Grain Based on Point Cloud

HUANG Xia^{1,2} ZHENG Shunyi^{1,2} GUI Li^{1,2} ZHAO Like¹ MA Hao¹

(1. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China

2. Collaborative Innovation Center of Geospatial Technology, Wuhan 430079, China)

Abstract: Large amount of grain phenotypic information is needed in researches such as digital grain traits investigation, phenotype and gene association analysis and digital agriculture simulation. A method for automatic extraction of grain high-throughput phenotypic information based on point cloud was proposed, aiming to automatically obtain three-dimensional (3D) grain model and 40 phenotypic parameters. Firstly, the classification of grain point cloud was completed through cluster analysis. Secondly, 3D grain model was reconstructed with cylindrical mesh method. Finally, according to the characteristics of different phenotypic parameters, 11 primary parameters, 11 derived parameters and 18 shape factors were automatically extracted. Experiment using data obtained by hand-held laser scanner (Handyscan 700) showed that the measurement result could reach millimeter level. The weight of each phenotypic parameter was analyzed based on principal component analysis method. With parameters measured by vernier caliper and Geomagic Studio as the true value, the average relative error of length, width, height, surface area and volume, the cross-sectional area of three principal component sections was 1.14%, 1.15%, 1.62%, 0, 1.82%, 2.12% and 2.43%, respectively. Compared with the manual measurement method and the software measurement method, the results of the proposed method was competitively accurate, which had advantages of batch processing, automation, less manual intervention (only in data acquisition) and high efficiency.

Key words: grain; phenotypic information; point cloud; automation; batch processing

收稿日期: 2017-08-16 修回日期: 2017-11-16

基金项目: 国家自然科学基金项目(41671452、41701532)、中央高校基本科研业务费专项资金项目(2042016kf0012)和中国博士后科学基金项目(2017M612510)

作者简介: 黄霞(1989—),女,博士生,主要从事三维测量研究, E-mail: yuxiuhx@whu.edu.cn

0 引言

表型组学^[1]是近年来植物学研究的热点,它是一门在基因组水平上系统研究某一生物或细胞在各种不同环境条件下所有表型的学科。表型指生物体个别或少数性状以至全部性状的表现,基因型是指生物的全部遗传物质(基因)组成^[2]。表型是基因的直接反映,基因是表型的内因^[3],通过研究作物的表型数据,可以建立基因与表型之间的关联分析(Genome wide association study, GWAS)^[4],进而通过基因筛选达到品种改良的目的^[5]。水稻谷粒的形状和长、宽、高、表面积以及体积等表型信息,对反映水稻的生长发育、生理生化以及遗传等生命过程现象至关重要^[6],也是建立谷粒离散元模型的基础^[7]。在水稻表型组学研究中,需要庞大的表型信息作为基础数据,因此,高通量自动获取谷粒表型信息十分重要。

传统的谷粒表型信息主要是人工测量^[8-9]。由于谷粒的尺寸较小,手工测量操作困难,极易出错,且耗时长^[10],测量的表型信息仅局限于谷粒的长和宽,不能测量更多的信息,不适合进行大规模谷粒表型信息的采集。

计算机视觉技术的飞速发展,使谷粒信息的自动测量成为了可能^[11]。常见的基于图像的谷粒信息测量的开源软件有 SmartGrain^[12]、ImageJ^[13]、CellProfiler^[14]和 P-TRAP^[15]以及 WinSeedle 等商业软件^[16]。这些软件都能实现谷粒表型信息的测量,但是需要对每一粒谷粒进行一一量取,并不能实现谷粒表型信息的批量测量。因此,有学者基于手机图像进行谷粒信息的提取^[17],实现了谷粒表型信息的自动和批量输出。但是,在图像上直接量取几何信息是不准确的,因为二维俯视图是一个中心投影,离投影中心越远,图像变形越大。

近年来 3D 技术日新月异,小目标的三维测量也获得了较高的精度^[18-19],3D 技术在种子测量方面也有了成功的应用^[20]。OGAWA 等^[21]利用 3D 技术获取了稻谷的多个剖面影像,实现了谷粒的三维建模,但此方法仅适合单个谷粒的三维建模,不适合批量建模。青克乐其其格等^[22]利用数字显微摄影测量技术采集了稻谷的显微图像,基于多角度匹配技术实现了谷粒的三维重建,但此方法需要对单粒谷粒进行多角度拍照,在数据获取时耗时较长,且实验设备昂贵。刘彩玲等^[23]利用全自动高精度立体扫描仪 CF-30 获取了 500 粒水稻种子的三维点云数据,精度可达毫米级,但其长度、宽度及厚度等参数是用 Geomagic Studio 软件逐一量取的,没能实

现谷粒表型信息的自动化提取和批量处理。

当前谷粒表型信息参数自动计算多采用图像技术,不能实现谷粒三维表面积和体积等参数的提取^[22]。为增加表型参数的种类和数量,改进谷粒表型参数提取的算法,本文提出一种基于轴对齐包围盒(Axis-aligned bounding box, AABB)算法的长宽高测量方法,基于三角网格的表面积算法,带符号投影的体积算法,基于凸包的剖面周长和面积的计算方法。

综上,为了实现谷粒表型信息批量和自动提取,增加谷粒表型信息的种类和数量,本文根据数字化考种、水稻种子建模和谷粒表型库建设等农业应用实际需求,整合三维扫描技术、信息处理技术以及计算机视觉技术等,提出一种基于点云的谷粒高通量表型信息自动提取方法。该方法旨在批量、快速和自动地获取谷粒的三维模型和 40 个表型参数,以期为科学研究和生产应用提供一种谷粒高通量表型信息自动提取方法。

1 材料与方法

1.1 数据获取

将 100 粒干粳稻随意铺撒在桌面上(避开标志点),采用 Handyscan 700 型手持式激光扫描仪进行扫描,获得谷粒的三维点云数据。

Handyscan 700 型扫描仪的测量原理是三角测量法,是一种非接触测量法,利用激光二极管发出激光,通过透镜聚焦投射到目标上,反射或漫射后的激光按照一定的三角关系成像,从而获取目标的空间位置信息^[24]。Handyscan 700 型扫描仪共有 7 束交叉激光线,精度为 0.030 mm。在扫描谷粒的同时,其附带的 VXelements 软件可实时显示谷粒的三维数据采集情况。

图 1 为 100 粒谷粒和单个谷粒的扫描数据,从形态上可以看出组成谷粒的点云均匀分布,谷粒的三维形态逼真。

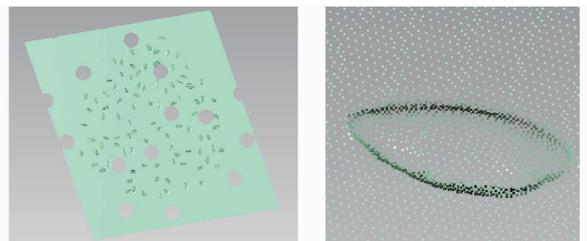


图 1 谷粒点云数据

Fig. 1 Cloud point data of grain

1.2 方法

本文的谷粒表型信息提取方法主要分为 3 部分:谷粒数据预处理,即点云分类,将组成单个谷粒

的点云单独分为一类；三维重建，获取谷粒的三维模型；表型信息提取，实现谷粒 40 个表型参数的自动提取。

1.2.1 分类

对谷粒的原始点云数据进行平面与谷粒的分离。平面上的点采用最小二乘法进行平面拟合，平面和每个点到平面的距离分别为

$$Ax + By + Cz + D = 0 \quad (1)$$

$$d(i) = \frac{|Ax_i + By_i + Cz_i + D|}{\sqrt{A^2 + B^2 + C^2}} \quad (2)$$

式中 $A、B、C$ ——平面方程的系数

D ——常数

$x_i、y_i、z_i$ —— i 点的坐标值

如果 $d(i)$ 大于给定的阈值 T (T 取 2 倍点云密度)，则 i 点为谷粒的点，如图 2a 中红色的点；如果 $d(i)$ 小于给定的阈值 T ，则 i 点为平面的点，如图 2a 中灰色的点。

实现谷粒点的提取后，如图 2b 所示，采用基于欧氏距离的最短聚类分析方法对谷粒进行分类。基本原理是：假设总共有 n 个样本，第 1 步，将每个样

本单独聚成一类，共有 n 类；第 2 步，根据确定样本的距离（本文采用欧氏距离）公式，把距离较近的 2 个样本聚成一类，其他样本为另一类，共有 $n - 1$ 类；第 3 步，按照第 2 步将样本聚成 $n - 2$ 类，依次循环，直到类间距离达到给定的阈值，如图 2c 所示。

基于欧氏距离的最短聚类结果一般会出现过分类的现象，如图 2d 所示，有单独的孤点自成一类或者几个飞点聚成一类。则需要对这类点进行处理。计算每类样本点的个数 N_i 和平均每类样本的个数 N_0 ，如果 N_i 小于阈值 T_{n0} ($T_{n0} = 0.5N_0$)，则判定此类点为飞点，直接删除，结果如图 2e 所示。

在实际铺撒谷粒时，往往会出现有谷粒重叠或者粘连在一起的情况，且由于其重叠的随意性，导致其在几何上没有规律可循。对于这样的谷粒，本文在计算谷粒表型信息时，直接剔除。判定方法为： N_i 大于 T_{n1} ($1.5N_0$)，则判定该目标不止一个谷粒，为相连的多个谷粒，直接删除该类点。如图 2e 所示，紫色的点为 2 粒粘连在一起的谷粒，图 2f 为删除粘连谷粒点云后的结果，每一类点云就是一粒谷粒，点云的分类个数就是谷粒的个数。

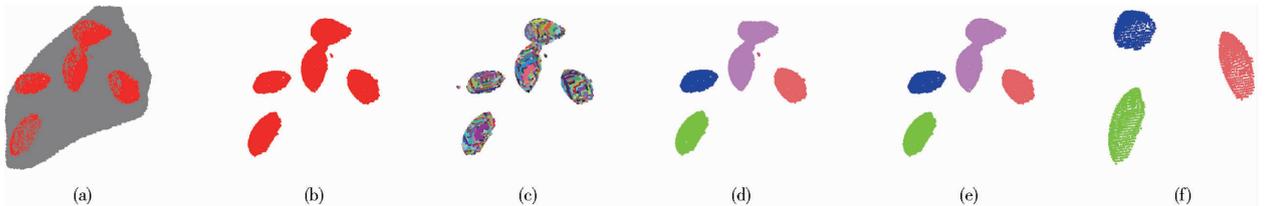


图 2 谷粒点云分类

Fig. 2 Classifications of grain cloud point

1.2.2 三维重建

要获取谷粒的表面积、体积等信息，需要进行三维重建。根据谷粒的三维特征可知，谷粒是一个较为简单的凸几何体，则采用柱面拟合的方式构建谷粒的曲面。

采用最近邻域法得到点的邻域 ($k = 15$)，由点的邻域集计算拟合平面。定义邻域 k 范围内点云的协方差矩阵

$$M = \sum_{i=1}^{n_1} (p_i - p_{i_0})(p_i - p_{i_0})^T \quad (3)$$

式中 n_1 ——某邻域内点的个数

p_i ——某邻域内 i 点的坐标

p_{i_0} ——某邻域内点云的质心坐标

式(3)中矩阵 M 为质量分布矩阵，此矩阵为对称矩阵，亦为半正定矩阵，其对应的最小特征值的特征向量 n_j 为拟合平面的法向量。进一步构建邻域点集的法矢方差矩阵

$$M' = \sum_{j=1}^{n_2} n_j n_j^T \quad (4)$$

式中 n_2 ——法向量的个数

求得该矩阵的 3 个特征值 $\lambda_1、\lambda_2、\lambda_3$ ($\lambda_1 \leq \lambda_2 \leq \lambda_3$)，归一化后为 $t_k = \lambda_k / (\lambda_1 + \lambda_2 + \lambda_3)$ ($k = 1, 2, 3$)。则 2 个较小的特征值对应 2 个主曲率，相应的特征向量为曲率对应的方向。

其较小的特征值对应的曲率方向作为柱面的母线方向，将邻域点集投影到法矢方差矩阵的另 2 个特征值对应的方向上，再将投影得到的平面点集拟合为一条圆锥曲线。平面点的坐标记为 (x_1, x_2) ，一般的圆锥曲线方程

$$G_i(x) = a_1 x_1^2 + a_2 x_1 x_2 + a_3 x_2^2 + a_4 x_1 + a_5 x_2 + a_6 \quad (5)$$

式中 $a_1 \sim a_5$ ——方程系数 a_6 ——常数

记 $x = [x_1^2 \ x_1 x_2 \ x_2^2 \ x_1 \ x_2 \ 1]^T$ ， $a = [a_1 \ a_2 \ a_3 \ a_4 \ a_5 \ a_6]^T$ ，则拟合的目标为

$$G_i(x) = \min_{j \in S_i} (a^T x_j)^2 = \min_{j \in S_i} a^T \left(\sum_{j \in S_i} x_j x_j^T \right) a \quad (6)$$

式中 S_i —— i 点局部邻域点集

记 $X = \sum_{j \in S_i} x_j x_j^T$, 该矩阵是对称、半正定矩阵, 则

X 的最小特征值对应的特征向量为拟合目标的解。

1.2.3 参数计算

1.2.3.1 基本参数

(1) AABB 包围盒法提取长宽高参数

扫描仪自带的参考坐标系和信息提取的全局坐标系往往是不一致的。为简化谷粒表型信息计算, 构建 AABB 包围盒^[25], 实现表型信息的批量提取, 本文先对谷粒进行坐标转换, 使在全局坐标系下任意方向的谷粒有统一的三轴(X 轴、 Y 轴和 Z 轴) 方向。校正后的谷粒在全局坐标系统的姿态为: 谷粒的几何中心为原点, 主对称面(谷粒长轴与短轴构成的平面)平行于 XOY 平面, 长轴为 X 轴, 短轴为 Y 轴, Z 轴遵循右手坐标法则。如图 3 所示, 左图为坐标转换前的谷粒, 右图为经坐标旋转后的谷粒。

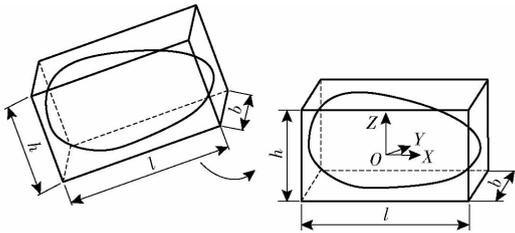


图 3 坐标旋转

Fig. 3 Coordinate rotation

分析算法进行坐标旋转。根据谷粒的几何特性, 本文采用谷粒的长轴、短轴及高作为谷粒几何分布的第一、二、三主成分量。坐标旋转公式为

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \mathbf{R} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \quad (7)$$

式中 (x_0, y_0, z_0) ——原始坐标

(x, y, z) ——全局坐标

\mathbf{R} ——旋转矩阵

按照式(3)构建所有谷粒三维点的协方差矩阵, 求得该矩阵的 3 个特征向量构成的矩阵即为所求的旋转矩阵

$$\mathbf{R} = [\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{r}_3] \quad (8)$$

式中 $\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3$ ——谷粒点云协方差矩阵的特征值对应的特征向量

根据旋转矩阵 \mathbf{R} , 可实现任意坐标系下的谷粒在全局坐标系中有统一的测量姿态。此时, 求出谷粒三轴的最大最小坐标即可得到 AABB 包围盒的 8 个角点, 进而构建 AABB 包围盒。

构建谷粒包围盒后, 可知一个谷粒的 AABB 包围盒就是一个简单的六面体, 每一边都平行于一个

坐标平面, 谷粒在这个矩形边界框内, 则测出边界框的三边长度即为谷粒的长、宽、高。如图 3 中的 l, b, h 即为谷粒的长、宽、高。

(2) 三角面求和计算表面积

谷粒的三维表面积可看作三角网格模型的总表面积 S , 即所有构网的三角形面积之和。

$$S = \sum_{i=1}^k s_i \quad (9)$$

式中 k ——三角形面的总数

s_i ——第 i 个三角形面的面积

(3) 带符号投影的凸五面体法求体积

谷粒的三维模型是三角网格围成的封闭空间, 其体积 V 为所有带符号三角形面的投影体积之和。

$$V = \sum_{i=1}^k (-1)^N V_i \quad (10)$$

式中 N ——正负号的判定符号, 取值为 0 或 1

V_i ——第 i 个三角形面的投影体积

N 是体积的符号判定, 0 和 1 分别为该三角形的法向量背向和指向投影平面。设定一个与网格模型中所有三角形面均不相交的投影平面, 投影体积为三角形面与投影平面围成的凸五面体的体积。如图 4 所示, A, B, C 为三角形的 3 个顶点, A_0, B_0, C_0 为 A, B, C 在投影平面上的投影顶点, 将此凸五面体剖分为 3 个四面体, 凸五面体 $ABC A_0 B_0 C_0$ 体积计算公式为

$$V_i = V(A_0 A C B) + V(A_0 B C C_0) + V(A_0 B C_0 B_0) \quad (11)$$

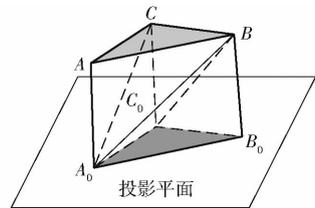


图 4 凸五面体

Fig. 4 Convex pentahedron

对于任意的四面体, 若 4 个顶点的三维坐标分别为 $(x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3)$ 和 (x_4, y_4, z_4) , 则其体积计算公式为

$$V((x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3), (x_4, y_4, z_4)) = \frac{1}{6} \begin{vmatrix} x_2 - x_1 & x_3 - x_1 & x_4 - x_1 \\ y_2 - y_1 & y_3 - y_1 & y_4 - y_1 \\ z_2 - z_1 & z_3 - z_1 & z_4 - z_1 \end{vmatrix} \quad (12)$$

因此, 确定此三角网格模型体积的关键问题在于确定各三角形面投影体积的符号。如图 4 所示, $\triangle ABC$ 的顶点按 A, B, C 依次存储, 按照右手坐标系法则, 法向量方向背向投影面, 则该三角形面的投影体积为正, 若 $\triangle ABC$ 的顶点按 A, C, B 依次存储, 则

该三角形面的投影体积为负。

(4) 凸包法求剖面周长与剖面面积

本文基于凸包法,计算谷粒的剖面周长(C)和剖面面积(A)。本文计算谷粒的 3 个主要剖面,即谷粒几何形状第一、二、三主成分分量构成的横剖面、纵剖面以及水平剖面。以水平剖面为例,先截取质心 δ (δ 为 2 倍点云密度)邻域的点集,通过正投影直接把三维点云转换为二维数据,如图 5a 所示,绿色的点为组成谷粒的平面投影点,对其用凸包法进行排序,算法如下:

(1) 以 X 轴为主序, Y 轴为副序对点从左到右进行升序排列。

(2) 使用逐点插入法寻找包围盒:以升序方式从左到右逐步插入上凸包点,再以逆序方式从右到左逐步插入下凸包点,凸点的判断条件为凸点总是位于上一条边的右侧。

(3) 连接所有的凸包点(图 5a 中红色的点),蓝色曲线为所求的周长曲线,剖面的周长计算公式为

$$C = \sum_{i=1}^m \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} \quad (13)$$

式中 m ——凸包点的数目

剖面面积为所有凸包点与中心点构成的三角形面积总和(图 5b),计算公式为

$$A = \sum_{i=1}^m a_i \quad (14)$$

式中 a_i ——第 i 个剖面三角形的面积

本文的基本参数有 11 个,见表 1 中参数 1 ~ 11。其中,包围盒体积 $V_{_AABB} = lbh$ 。

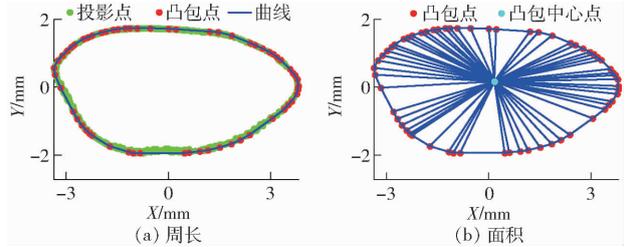


图 5 剖面周长和面积

Fig. 5 Perimeter and area of a profile

1.2.3.2 衍生参数和形状因子

本文衍生参数 11 个(表 1 中参数 12 ~ 22)。

传统谷粒形状描述因子为水平剖面的 5 个形状因子和 1 个紧凑度指数^[26],本文计算 3 个主成分剖面(水平剖面 xy 、纵剖面 xz 和横剖面 yz)对应的形状因子和紧凑度指数共计 18 个参数,见表 1 中参数 23 ~ 40。形状因子($f_1 \sim f_5$)和紧凑度指数(c)计算公式为

$$f_1 = \frac{4\pi A}{C^2} \quad (15)$$

$$f_2 = \frac{l_p}{A} \quad (16)$$

$$f_3 = \frac{A}{l_p^3} \quad (17)$$

$$f_4 = \frac{4A}{\pi l_p^2} \quad (18)$$

$$f_5 = \frac{4\pi A}{\pi w} \quad (19)$$

$$c = \frac{C^2}{4\pi A} \quad (20)$$

式中 l_p ——剖面的长 w ——剖面的宽

表 1 40 个表型参数

Tab. 1 40 phenotypic parameters

序号	符号	名称	序号	符号	名称	序号	符号	名称
1	S	表面积	15	$V_{_AABB}$	包围盒体积	29	c_{xz}	纵剖面紧凑度指数
2	V	体积	16	S/V	面积体积比	30	f_{1xz}	纵剖面形状因子 1
3	l	长	17	S/l	面积长度比	31	f_{2xz}	纵剖面形状因子 2
4	b	宽	18	S/b	面积宽度比	32	f_{3xz}	纵剖面形状因子 3
5	h	高	19	S/h	面积高度比	33	f_{4xz}	纵剖面形状因子 4
6	C_{xy}	水平剖面周长	20	V/l	体积长度比	34	f_{5xz}	纵剖面形状因子 5
7	A_{xy}	水平剖面面积	21	V/b	体积宽度比	35	c_{yz}	横剖面紧凑度指数
8	C_{xz}	纵剖面周长	22	V/h	体积高度比	36	f_{1yz}	横剖面形状因子 1
9	A_{xz}	纵剖面面积	23	c_{xy}	水平剖面紧凑度指数	37	f_{2yz}	横剖面形状因子 2
10	C_{yz}	横剖面周长	24	f_{1xy}	水平剖面形状因子 1	38	f_{3yz}	横剖面形状因子 3
11	A_{yz}	横剖面面积	25	f_{2xy}	水平剖面形状因子 2	39	f_{4yz}	横剖面形状因子 4
12	l/b	长宽比	26	f_{3xy}	水平剖面形状因子 3	40	f_{5yz}	横剖面形状因子 5
13	l/h	长高比	27	f_{4xy}	水平剖面形状因子 4			
14	b/h	宽高比	28	f_{5xy}	水平剖面形状因子 5			

2 实验与分析

在实际考种中,对谷粒表型信息的测量通常选

用 15 ~ 30 粒种子^[27],本文选用 100 粒水稻谷粒进行谷粒高通量表型信息自动提取实验。经分类后,除去 4 粒粘连在一起的谷粒,本文实现了对 96 粒谷

粒的三维建模和40个表型参数的提取。

2.1 实验结果与谷粒形态分析

图6为谷粒的三维模型。图6a为96粒谷粒的三维模型,图6b为单粒谷粒的三维模型,从图中可以看出,本文所构建的谷粒三维模型仿真度较高。



图6 谷粒三维模型

Fig.6 3D models of grain

表2~4分别是谷粒表型信息的基本参数、衍生参数及形状因子的最大值、最小值和均值。如表2所示,谷粒的表面积、长、宽、水平剖面面积和周长的

值差异较小,体积、高、横纵剖面面积和周长的值差异较大。

表2 基本参数

Tab.2 Primary parameters

	S/mm^2	V/mm^3	l/mm	b/mm	h/mm	C_{xy}/mm	A_{xy}/mm^2	C_{xz}/mm	A_{xz}/mm^2	C_{yz}/mm	A_{yz}/mm^2
最大值	68.11	44.14	7.77	4.20	3.17	18.85	23.83	17.67	17.77	11.41	9.60
最小值	46.10	18.26	5.89	3.19	1.34	15.29	16.45	14.23	8.73	7.77	3.72
均值	59.46	33.43	7.04	3.72	2.59	20.70	17.53	14.71	16.34	7.44	10.17

表3 衍生参数

Tab.3 Derived parameters

	l/b	l/h	b/h	V_{-AABB}	S/V	S/l	S/b	S/h	V/l	V/b	V/h
最大值	2.25	5.51	2.87	89.56	2.69	9.60	18.54	40.26	6.24	11.33	15.55
最小值	1.64	2.00	1.13	36.23	1.54	6.59	12.94	16.58	2.54	5.22	8.95
均值	1.90	2.84	1.49	67.79	1.82	8.45	16.00	23.61	4.75	8.98	12.97

表4 3个主成分剖面的形状因子

Tab.4 Shape factors of three principal component profiles

		f_1	f_2	f_3	f_4	f_5	c
水平剖面(xy)	最大值	0.91	0.40	0.08	0.61	1.10	1.27
	最小值	0.79	0.30	0.05	0.46	0.93	1.10
	均值	0.85	0.34	0.06	0.53	1.01	1.18
纵剖面(xz)	最大值	0.82	0.84	0.07	0.49	1.23	2.29
	最小值	0.44	0.40	0.02	0.21	0.91	1.22
	均值	0.69	0.49	0.04	0.38	1.04	1.47
横剖面(yz)	最大值	0.96	0.96	0.23	0.94	1.18	1.59
	最小值	0.63	0.41	0.07	0.35	0.85	1.04
	均值	0.89	0.52	0.15	0.69	0.99	1.13

研究发现,体积、高、横纵剖面面积和周长的值较小均为瘪粒,而这些参数直接与谷粒的高度信息有关,因此,谷粒的高度信息可用于区分谷粒实粒和瘪粒,在一定程度上能直接反映谷粒的灌浆程度。如表3和表4所示,谷粒的衍生参数和形状因子的差异均不大,反映了谷粒表型信息的细节差异,例如,相同长宽的谷粒的衍生参数和形状因子却不一定相同。

2.2 基于主成分分析的参数权重分析

本文提取的40个参数很好地定量描述了谷粒的几何形态,可用于逆向工程谷粒建模^[23,28]。本文利用主成分分析法对谷粒表型参数的权重进行分析,建立谷粒形状模型^[29]。

利用SPSS软件对40个参数进行主成分分析,获得6个主成分,累积贡献率为93.356%,满足主成分分析特征向量累积达到85%的要求。由SPSS软件可得到各主成分特征值、主成分贡献率和主成

分载荷矩阵。以主成分的方差贡献率为权重,建立谷粒形状综合得分模型

$$\begin{aligned}
 Y = & 0.0965X_1 + 0.0636X_2 + 0.0991X_3 + \\
 & 0.1186X_4 + 0.0104X_5 + 0.1463X_6 + \\
 & 0.1298X_7 + 0.0850X_8 + 0.0975X_9 + \\
 & 0.1283X_{10} + 0.131X_{11} + 0.0402X_{12} + \\
 & 0.0118X_{13} + 0.0239X_{14} + 0.0644X_{15} + \\
 & 0.0235X_{16} + 0.0593X_{17} + 0.0241X_{18} + \\
 & 0.0445X_{19} + 0.0704X_{20} + 0.0601X_{21} + \\
 & 0.0946X_{22} - 0.0603X_{23} + 0.0601X_{24} - \\
 & 0.1341X_{25} - 0.0232X_{26} + 0.0459X_{27} + \\
 & 0.007X_{28} + 0.0645X_{29} - 0.0684X_{30} - \\
 & 0.0105X_{31} - 0.0787X_{32} - 0.0659X_{33} - \\
 & 0.0064X_{34} + 0.1123X_{35} - 0.1122X_{36} - \\
 & 0.0472X_{37} - 0.1012X_{38} - 0.0784X_{39} + \\
 & 0.0012X_{40}
 \end{aligned}
 \tag{21}$$

式中 Y ——谷粒形状综合得分

$X_1 \sim X_{40}$ ——表 1 中对应的 40 个表型参数

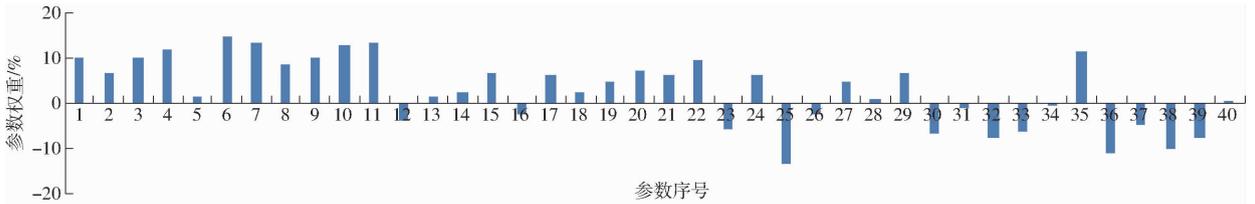


图 7 谷粒表型参数权重

Fig. 7 Weight of grain phenotypic parameter

重图。从图 7 可知,绝对值大于平均权重 ($1/40 \times 100\% = 2.5\%$) 的参数有表面积、体积与长等 30 个参数,小于平均权重的有长高比、宽高比和水平剖面形状因子等 10 个参数。

2.3 精度与效率分析

为了验证算法的测量精度,本文利用精度为 0.01 mm 的游标卡尺,测量 96 粒谷粒(除去粘连的 4 粒谷粒)的长、宽和高作为直线参数精度评价的参考值(真值),考虑到人为测量误差,实测值取 3 次人工测量的均值;表面积、体积和剖面面积的参考值由 Geomagic Studio 逐一量取,由于 Geomagic Studio 没有直接计算周长的功能,本文的周长不做精度评价分析。

图 8 是谷粒表型测量值与参考值的误差分析。从图 8 可以看出,谷粒的体积和表面积参数与

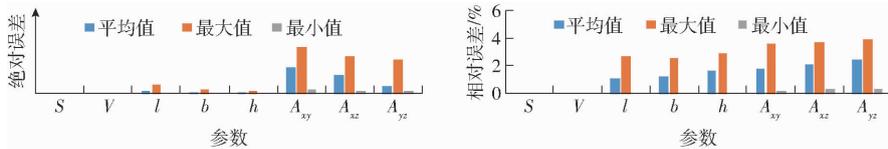


图 8 谷粒表型测量值与参考值的误差分析

Fig. 8 Error analysis of measured and reference phenotypic values of grain

表 5 是本文方法与其他方法的精度对比,实验对象都为粳稻,但品种略有不同。方法 1^[30] 是 2011 年 DUAN 基于图像对谷粒的长和宽进行测量,方法 2^[17] 是 2016 年 MA 等基于手机图像进行的谷粒的长宽测量。这两种方法和本文的真值均是用游标卡尺的测量值进行检测的。

表 5 本文方法与其他方法的精度对比

Tab. 5 Accuracy comparison between proposed method and two state-of-the-art methods

精度参数	本文方法	方法 1	方法 2
长平均绝对误差/mm	0.08	0.11	0.10
宽平均绝对误差/mm	0.04	0.04	0.04
长平均相对误差/%	1.14	1.47	1.30
宽平均相对误差/%	1.15	1.31	1.23

从表 5 可以看出,本文方法的精度均不低于另

由该模型可知谷粒参数权重分布,如图 7 所示。

图 7 是本文提取的表 1 中 40 个表型参数的权

Geomagic Studio 测量值相同。长、宽和高的最小绝对误差均为 0 mm,最小相对误差均为零,最大绝对误差依次为 0.20、0.10、0.07 mm,最大相对误差为 2.63%、2.59% 和 2.89%。长、宽和高的平均绝对误差为 0.08、0.04、0.04 mm,平均相对误差为 1.14%、1.15% 和 1.62%。3 个剖面面积与 Geomagic Studio 测量值的平均绝对误差分别为 0.62、0.44、0.19 mm²,最大绝对误差为 1.1、0.91、0.81 mm²,最小绝对误差为 0.11、0.08、0.06 mm²。3 个剖面面积与 Geomagic Studio 测量值的平均相对误差分别是 1.82%、2.12% 和 2.43%,最大相对误差为 3.64%、3.71% 和 3.92%,最小相对误差为 0.14%、0.25% 和 0.29%。一般表型测量的精度要求是 5%^[17,30],可以看出,在精度上,本文提出的方法可以满足应用需求。

外 2 种方法。且本文可以测量出由这 2 种方法不能测量的谷粒表面积、体积、剖面周长和剖面面积等其他表型参数。

本次实验的数据采集从铺撒到三维数据采集完成,共计约 160 s。数据分析和处理是在华硕台式计算机(2.50 GHz,8.0 GB RAM)上完成,算法实现平台为 Microsoft Visual Studio 2010,共耗时 98.497 s。包括数据采集和信息提取,平均 2.69 s 完成 1 粒谷粒的测量,耗时大大少于人工测量,可用于大规模数据采集和分析。

3 结论

(1) 批量实现谷粒的三维建模和提取上百粒谷粒各 40 个表型参数,表型信息提取不需要人工干预,不需要人工设置参数。

(2) 以手工实测值和软件测量值为真值,本文

方法所提取的长、宽、高和3个主成分剖面面积的平均相对误差在2.43%以内。与相关文献方法相比,精度略高且表型信息种类更丰富。

(3)所提取的40个参数可建立谷粒形状综合得分模型。包括数据采集和信息提取,平均2.69s完成单粒谷粒40个表型参数的测量。

参 考 文 献

- 1 HOULE D, GOVINDARAJU D R, OMHOLT S. Phenomics: the next challenge[J]. *Nature Reviews Genetics*, 2010, 11(12): 855 - 866.
- 2 GJUUSLAND A B, VIK J O, BEARD D A, et al. Bridging the genotype-phenotype gap: what does it take[J]. *Journal of Physiology*, 2013, 591(8): 2055 - 2066.
- 3 BARH D, KHAN M, DAVIES E. *PlantOmics: the omics of plant science*[M]. India: Springer, 2015.
- 4 PANTALIÃO G F, NARCISO M, GUIMARÃES C, et al. Genome wide association study (GWAS) for grain yield in rice cultivated under water deficit[J]. *Genetica*, 2016, 144(6): 651 - 664.
- 5 GLASZMANN J C, KILIAN B, UPADHYAYA H D, et al. Accessing genetic diversity for crop improvement[J]. *Current Opinion in Plant Biology*, 2010, 13(2): 167 - 173.
- 6 UPADHYAYA H D, REDDY K N, SINGH S, et al. Phenotypic diversity in *Cajanus* species and identification of promising sources for agronomic traits and seed protein content[J]. *Genetic Resources and Crop Evolution*, 2013, 60(2): 639 - 659.
- 7 赵湛,李耀明,陈义,等. 水稻籽粒碰撞力学特性研究[J/OL]. *农业机械学报*, 2013, 44(6): 88 - 92. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx?file_no=20130617&flag=1. DOI:10.6041/j.issn.1000-1298.2013.06.017.
ZHAO Zhan, LI Yaoming, CHEN Yi, et al. Impact mechanical characteristics analysis of rice grain[J/OL]. *Transactions of the Chinese Society for Agricultural Machinery*, 2013, 44(6): 88 - 92. (in Chinese)
- 8 FURBANK R T, TESTER M. Phenomics-technologies to relieve the phenotyping bottleneck[J]. *Trends in Plant Science*, 2011, 16(12): 635 - 643.
- 9 HITALMANI S, HUANG N, COURTOIS B, et al. Identification of QTL for growth and grain yield-related traits in rice across nine locations of Asia[J]. *Theoretical and Applied Genetics*, 2003, 107(4): 679 - 690.
- 10 AFONNIKOV D A, GENAEV M A, DOROSHKOV A V, et al. Methods of high-throughput plant phenotyping for large-scale breeding and genetic experiments[J]. *Russian Journal of Genetics*, 2016, 52(7): 688 - 701.
- 11 BHAGWAT S G, SAINIS J K, SHOUCHE S P, et al. Grain morphometric characterisation of genetically related wheat selections using image analysis[J]. *Cereal Research Communications*, 2003, 31(1-2): 205 - 212.
- 12 TANABATA T, YANO M. SmartGrain: high-throughput phenotyping software for measuring seed shape through image analysis[J]. *Plant Physiology*, 2012, 160(4): 1871 - 1880.
- 13 IGATHINATHANE C, PORDESIMO L O, COLUMBUS E P, et al. Shape identification and particles size distribution from basic shape parameters using ImageJ[J]. *Computers & Electronics in Agriculture*, 2008, 63(2): 168 - 182.
- 14 LAMPRECHT M R, SABATINI D M, CARPENTER A E. CellProfiler: free, versatile software for automated biological image analysis[J]. *Biotechniques*, 2007, 42(1): 71 - 75.
- 15 ALTAM F, ADAM H, ANJOS A D, et al. P-TRAP: a panicle trait phenotyping tool[J]. *BMC Plant Biology*, 2013, 13(1): 122 - 135.
- 16 GROVES F E, BOURLAND F M. Estimating seed surface area of cotton seed[J]. *Journal of Cotton Science*, 2010, 14(2): 74 - 81.
- 17 MA Z, MAO Y, GONG L, et al. Smartphone-based visual measurement and portable instrumentation for crop seed phenotyping[J]. *IFAC-PapersOnLine*, 2016, 49(16): 259 - 264.
- 18 ZHENG S, ZHOU L, LI C, et al. Automatic 3D model reconstruction of the diamond object[J]. *Measurement*, 2012, 45(5): 844 - 855.
- 19 CHAKRABORTY J, SARKAR D, SINGH A, et al. Measuring the three-dimensional morphology of crystals using regular reflection of light[J]. *Crystal Growth & Design*, 2012, 12(12): 6042 - 6049.
- 20 MOORE C R, GRONWALL D S, MILLER N D, et al. Mapping quantitative trait loci affecting *Arabidopsis thaliana* seed morphology features extracted computationally from images[J]. *G3 Genes Genomes Genetics*, 2013, 3(1): 109 - 118.
- 21 OGAWA Y, KUENSTING H, SUGIYAMA J, et al. Structure of a rice grain represented by a new three-dimensional visualisation technique[J]. *Journal of Cereal Science*, 2002, 36(1): 1 - 7.
- 22 青克乐其其格,郭新宇,肖伯祥,等. 基于多角度图像的水稻稻谷几何建模方法研究[J]. *中国传媒大学学报:自然科学版*, 2011, 18(3): 72 - 75.
QINGKELE Qiqige, GUO Xinyu, XIAO Boxiang, et al. Rice grain modeling based on multi-perspective images[J]. *Journal of Communication University of China: Science and Technology*, 2011, 18(3): 72 - 75. (in Chinese)
- 23 刘彩玲,王亚丽,宋建农,等. 基于三维激光扫描的水稻种子离散元建模及试验[J]. *农业工程学报*, 2016, 32(15): 294 - 300.
LIU Cailing, WANG Yali, SONG Jiannong, et al. Experiment and discrete element model of rice based on 3D laser scanning[J]. *Transactions of the CSAE*, 2016, 32(15): 294 - 300. (in Chinese)

- fingerprinting[J]. *Pervasive & Mobile Computing*, 2012, 8(2):292-316.
- 7 BRUNATO M, BATTITI R. Statistical learning theory for location fingerprinting in wireless LANs[J]. *Computer Networks*, 2005, 47(6):825-845.
- 8 石柯, 陈洪生, 张仁同. 一种基于支持向量回归的 802.11 无线室内定位方法[J]. *软件学报*, 2014, 25(11):2636-2651. SHI Ke, CHEN Hongsheng, ZHANG Rentong. Indoor location method based on support vector regression in 802.11 wireless environments[J]. *Journal of Software*, 2014, 25(11):2636-2651. (in Chinese)
- 9 张勇, 黄杰, 徐科宇. 基于 PCA-LSSVR 算法的 WLAN 室内定位方法[J]. *仪器仪表学报*, 2015, 36(2):408-414. ZHANG Yong, HUANG Jie, XU Keyu. Indoor positioning algorithm for WLAN based on principal component analysis and least square support vector regression[J]. *Chinese Journal of Scientific Instrument*, 2015, 36(2):408-414. (in Chinese)
- 10 徐玉滨, 邓志安, 马琳. 基于核直接判别分析和支持向量回归的 WLAN 室内定位算法[J]. *电子与信息学报*, 2011, 33(4):896-901. XU Yubin, DENG Zhian, MA Lin. WLAN indoor positioning algorithm based on KDDA and SVR[J]. *Journal of Electronics & Information Technology*, 2011, 33(4):896-901. (in Chinese)
- 11 李华亮, 钱志鸿, 田洪亮. 基于核函数特征提取的室内定位算法研究[J]. *通信学报*, 2017, 38(1):158-167. LI Hualiang, QIAN Zhihong, TIAN Hongliang. Research on indoor localization algorithm based on kernel principal component analysis[J]. *Journal on Communications*, 2017, 38(1):158-167. (in Chinese)
- 12 DWIYASA F, LIM M H, ONG Y S, et al. Extreme learning machine for indoor location fingerprinting[J]. *Multidimensional Systems and Signal Processing*, 2017, 28(3):867-883.
- 13 ZOU H, LU X, JIANG H, et al. A fast and precise indoor localization algorithm based on an online sequential extreme learning machine[J]. *Sensors*, 2015, 15(1):1804-1824.
- 14 ENGEL Y, MANNOR S, MEIR R. The kernel recursive least-squares algorithm[J]. *IEEE Transactions on Signal Processing*, 2004, 52(8):2275-2285.
- 15 VAERENBERGH S V, VIA J, SANTAMARIA I. Nonlinear system identification using a new sliding-window kernel RLS algorithm[J]. *Journal of Communications*, 2007, 2(3):1-8.
- 16 CHEN B, ZHAO S, ZHU P, et al. Quantized kernel least mean square algorithm[J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2012, 23(1):22-32.
- 17 VAERENBERGH S V, SANTAMARIA I, LIU W, et al. Fixed-budget kernel recursive least-squares [C] // 2010 IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP), 2010: 1882-1885.
- 18 HASSIBI B, STORK D G, WOLFF G J. Optimal brain surgeon and general network pruning[C] // IEEE International Conference on Neural Networks, 1993: 293-299.
- 19 KRUIF B J D, VRIES T J A D. Pruning error minimization in least squares support vector machines[J]. *IEEE Transactions on Neural Networks*, 2003, 14(3):696-702.
- 20 ROSIPAL R, TREJO L J. Kernel partial least squares regression in reproducing kernel hilbert space[J]. *Journal of Machine Learning Research*, 2001, 2(2):97-123.
- 21 李军, 李大超. 基于优化核极限学习机的风电功率时间序列预测 [J]. *物理学报*, 2016, 65(13):130501. LI Jun, LI Dachao. Wind power time series prediction using optimized kernel extreme learning machine method[J]. *Acta Physica Sinica*, 2016, 65(13):130501. (in Chinese)
- 22 LAOUDIAS C. Device self-calibration in location systems using signal strength histograms [J]. *Journal of Location Based Services*, 2013, 7(3):165-181.

(上接第 264 页)

- 24 成思源, 刘俊, 张湘伟. 基于手持式激光扫描的反求设计实验[J]. *实验室研究与探索*, 2011, 30(8):153-155. CHENG Siyuan, LIU Jun, ZHANG Xiangwei. Experimental teaching of the reverse design based on hand-held laser scanning[J]. *Research and Exploration in Laboratory*, 2011, 30(8):153-155. (in Chinese)
- 25 GINO V D B. Efficient collision detection of complex deformable models using AABB trees[J]. *Journal of Graphics Tools*, 1997, 2(4):1-13.
- 26 SHOUCHE S P, RASTOGI R, BHAGWAT S G, et al. Shape analysis of grains of Indian wheat varieties. [J]. *Computers & Electronics in Agriculture*, 2001, 33(1):55-76.
- 27 IGATHINATHANE C, PORDESIMO L O, BATCHELOR W D. Major orthogonal dimensions measurement of food grains by machine vision using ImageJ[J]. *Food Research International*, 2009, 42(1):76-84.
- 28 陈晨, 吴崇友, 江涛. 基于逆向工程的水稻精准模型构建及试验验证[J]. *农机化研究*, 2017, 39(10):46-52. CHEN Chen, WU Chongyou, JIANG Tao. Precise modeling method of rice based on reverse engineering and test verification[J]. *Journal of Agricultural Mechanization Research*, 2017, 39(10):46-52. (in Chinese)
- 29 杨春周, 滕克难, 程月波. 作战效能评估指标权重的确定[J]. *计算机仿真*, 2008, 25(10):5-7. YANG Chunzhou, TENG Ke'nan, CHENG Yuebo. Determination of the weight values of indexes for the combat capability[J]. *Computer Simulation*, 2008, 25(10):5-7. (in Chinese)
- 30 DUAN L, YANG W, HUANG C, et al. A novel machine-vision-based facility for the automatic evaluation of yield-related traits in rice[J]. *Plant Methods*, 2011, 7(1):44-57.