

用冠层叶色偏态分布模式 RGB 模型预测大豆产量

张 佩^{1,2}, 陈郑盟³, 马顺登¹, 尹 帝¹, 江海东^{1*}

(1. 南京农业大学农业部作物生理生态与生产管理重点实验室、江苏省现代作物生产协同创新中心、国家信息农业工程技术中心, 南京 210095; 2. 江苏省气象局, 南京 210008; 3. 福建省烟草公司龙岩市公司, 龙岩 364000)

摘 要: 为了探索加色混色 (Red-Green-Bule, RGB) 模型偏态分布模式在大豆产量预测上的可行性, 并验证其在不同肥料运筹、不同品种上的通用性, 该研究选用曲茎和徐豆 18 两个大豆品种, 设计了不同种植密度和氮肥水平的大田裂区试验, 以无人机搭载数码相机, 在花期及以后的 2 个重要生殖生长期采集大豆冠层数据。研究证实了大豆冠层数码图像的光学三原色 RGB 模型色阶遵循偏态分布, 并利用偏态分析得到 5 类共 20 个叶色偏态参数。花期、荚期和鼓粒期的冠层叶色偏态参数普遍存在明显差异, 其中表征叶色深浅的均值、中位数、众数从花期至鼓粒期呈现先降后升的变化趋势, 而表征叶色偏向性的偏度和表征叶色集中度的峰度则普遍呈现相反的变化趋势。基于偏态参数构建的大豆产量预测模型的预测准确度平均达 91.30% (建模组), 对氮肥运筹验证组的预测准确度平均为 87.33%, 对不同品种验证组的预估准确度虽然低于建模组和氮肥运筹验证组, 但也接近 80%。这说明 RGB 模型偏态参数可准确地描述不同生育期大豆冠层叶色状况, 基于偏态参数构建的产量预测模型有了更多的冠层颜色信息输入, 对选用同品种但采用不同氮肥运筹措施和选用不同品种下的产量预测准确率均较高, 可广泛用于不同生产条件的大豆产量预测, 具有较好的适用性与较高的推广价值。

关键词: UVA; 氮; 图像分析; 大豆; RGB 图像; 偏态参数; 产量预估

doi: 10.11975/j.issn.1002-6819.2021.09.014

中图分类号: S565.1

文献标志码: A

文章编号: 1002-6819(2021)-09-0120-07

张 佩, 陈郑盟, 马顺登, 等. 用冠层叶色偏态分布模式 RGB 模型预测大豆产量[J]. 农业工程学报, 2021, 37(9): 120-126.

doi: 10.11975/j.issn.1002-6819.2021.09.014 <http://www.tcsae.org>

Zhang Pei, Chen Zhengmeng, Ma Shundeng, et al. Prediction of soybean yield by using RGB model with skew distribution pattern of canopy leaf color[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(9): 120-126. (in Chinese with English abstract) doi: 10.11975/j.issn.1002-6819.2021.09.014 <http://www.tcsae.org>

0 引 言

随着数码图像技术的成熟和高分辨率相机设备的普及, 通过分析彩色数码图像来描述植物表型性状和生长状况的工作越来越多^[1-4]。数字彩色图像包含丰富的植物形态、结构和颜色信息^[5-6], 已经被用来研究作物叶绿素含量^[7]、营养状况^[8]、逆境胁迫程度^[9]。由于数码图像信息与植物生长已经被证实有着较好的相关性, 因此, 逐渐有学者开始利用作物图像来开展产量估测^[10-11]。

加色混色模型 (Red-Green-Bule, RGB) 模型是数码图像最常用的颜色分析方法^[12], 其叶色分布模式有两种: 正态分布或者偏态分布^[13-15]。在统计学中, 常用的偏态分布特征参数有均值、中位数、众数、偏度和峰度等, 正态分布是一种特殊的数据分布模式, 其特点是均值、中位数、众数三者合一, 偏度为 0^[16]。在以往的多数研究中, 研究人员把植株冠层叶色分布当作正态分布或者近

似正态分布处理^[17-19], 这导致获得的统计量只有各通道均值一项, 信息量较少, 只能以各通道均值及其组合参数近似描述植物叶色深浅情况。同时, 也忽略了其他参数在叶色定量描述中的意义, 无法真实、准确地描述叶色分布情况, 极大地限制了 RGB 模型的应用。仅有少量研究者开展了叶色色阶按偏态分布进行处理的尝试, 发现茶叶新叶和老叶的偏度明显不同^[13], 玉米叶片中的水分状况跟叶片 RGB 模型中灰度通道的偏度相关^[14], 烟草叶片 SPAD 值 (Soil and Plant Analyzer Development) 与红色 (Red, R) 通道均值和偏度相关^[15]。这些研究基于偏态分布得到了均值、中位数、众数、偏度和峰度 5 种共 20 个常用的 RGB 模型偏态参数。这些参数不仅极大地拓展植物叶片颜色信息, 而且能从颜色深浅、分布的偏向性及集中度三个方面系统地描述叶色。但是这些偏态分布的研究都是器官层面的, 在作物生产上的应用需要上升到群体层面。因此, 本研究拟进一步验证作物群体图像是否符合偏态分布模式, 并探索 RGB 模型偏态分布模式在大豆产量预估上的应用前景。不同的大豆栽培措施能够产生不同的群体特征并影响最终的产量^[20-22]。为了尝试利用 RGB 模型偏态参数无损快速地预测大豆产量, 本研究以大豆新品系曲茎为材料, 通过不同的种植密度和氮肥用量处理形成不同的群体, 利用无人机分别于花期、荚期和鼓粒期采集冠层图像, 分析群体图像色

收稿日期: 2021-03-15 修订日期: 2021-05-01

基金项目: 国家重点研发计划课题 (2018YFD1000900); 江苏省“333 工程”

高层次人才培养科研项目和江苏省气象局科技项目 (KM201905)

作者简介: 张佩, 高级工程师, 研究方向为应用气象及设施农业。

Email: 78073954@qq.com

*通信作者: 江海东, 副教授, 研究方向为作物栽培生理生态。

Email: hdjiang@njau.edu.cn

阶偏态分布的特征, 然后分别提取 RGB 图像 R 、绿 (Green, G)、蓝 (Blue, B) 3 个颜色通道及灰度 (Gray, Y) 图像的正态参数、偏态参数, 运用多元回归法构建基于无人机 RGB 图像偏态参数的大豆产量预测模型, 并基于不同栽培措施和不同品种条件下大豆的产量进行验证, 以期为无损快速的智能化大豆产量预测提供一种可靠的思路与方法。

1 材料与方法

1.1 供试材料

大豆新品系曲茎由南京农业大学国家大豆改良中心提供, 用于建模和同品种不同栽培措施间验证; 徐豆 18 由江苏徐淮地区徐州农业科学研究所提供, 用于不同品种间验证。试验在安徽省滁州市琅琊区三官镇 ($32^{\circ}32'N$, $118^{\circ}30'E$) 实施, 试验地面积 $12\text{ m}\times 50\text{ m}$, 壤土, 全氮 1.01 g/kg 、速效磷 16.33 mg/kg 、速效钾 101.3 mg/kg 和有机质 16.8 g/kg 。试验地气象数据来自安徽省气象信息中心。

1.2 试验设计

2020 年 6 月 12 日播种, 10 月 1 日收获。大豆生育期间 (6 月 12 日—9 月 30 日) 累计降水量为 843.6 mm , 完全满足生长的水分需求, 无灌溉。氮肥基施和花期追肥, 以尿素施用。氯化钾 150 kg/hm^2 和过磷酸钙 300 kg/hm^2 基施。行距 0.5 m , 每穴播 $5\sim 6$ 粒, 留苗 3 株。小区面积 $3\text{ m}\times 4\text{ m}$ 。

建模试验。试验品种为曲茎, 采用裂区试验, 主区为种植密度: $15\times 10^4\text{ 株/hm}^2$ (穴距 0.4 m)、 $22.5\times 10^4\text{ 株/hm}^2$ (穴距 0.27 m)、 $30\times 10^4\text{ 株/hm}^2$ (穴距 0.20 m) 和 $37.5\times 10^4\text{ 株/hm}^2$ (穴距 0.16 m)。副区为氮肥用量: 75 、 150 、 225 kg/hm^2 (以 N 计, 下同)。基追比为 $5:5$ 。每个试验处理设 3 个重复小区 (下同)。

验证试验 1 (氮肥运筹)。试验品种为曲茎, 采用裂区试验, 主区为氮肥用量: 75 、 150 、 225 kg/hm^2 。副区为基追比: $0:10$ 、 $5:5$ 和 $10:0$ 。密度为 $22.5\times 10^4\text{ 株/hm}^2$ (穴距 0.27 m)。

验证试验 2 (不同品种)。试验品种为徐豆 18, 采用裂区试验, 主区为密度: $24\times 10^4\text{ 株/hm}^2$ (穴距 0.25 m) 和 $36\times 10^4\text{ 株/hm}^2$ (穴距 0.17 m)。副区为氮肥基追比: $0:10$ 、 $5:5$ 、 $10:0$ 。

每个小区取 1 m^2 大豆植株用于最终产量测定。

1.3 大豆冠层无人机数码图像采集

无人机数码图像采集于 2020 年大豆花期 (7 月 29 日)、荚期 (8 月 15 日) 及鼓粒期 (9 月 5 日) 3 个关键生育期 09:00 左右进行。7 月 29 日、8 月 15 日和 9 月 5 日 09:00 天气晴好, 气温分别为 29.2 、 34.8 和 $31.5\text{ }^{\circ}\text{C}$, 风速分别为 1.8 、 2 和 1.0 m/s 。无人机 (Hudson zino, 中国) 飞行高度设置为 30 m , 静止拍摄, 图像采集镜头 (Sony, 日本) 分辨率为 3840×2160 像素, 光圈进光量 $F=0.8$, 曝光时间 0.05 s , ISO 感光度为 100 , 源图像以 *.JPG 格式保存。

1.4 冠层图像叶色色阶分布的偏度和峰度参数计算方法

图像色阶分布的偏度和峰度参数均由高阶方程计算

得到^[16]。在概率论中, 矩是用来描述统计样本的某些特征数值, 即求平均值, 用 E 表示。中心矩 (Central moment) 定义为, 对于正整数 k , 如果 $E(X)$ 存在, 且 $|E((X-E(X))^k)| < \infty$, 则称 $E((X-E(X))^k)$ 为整体样本统计量 X 的 k 阶中心矩。其计算公式^[16]如下:

$$\mu_k = E((X-E(X))^k) = \int_{-\infty}^{\infty} (x-\mu)^k f(x) dx \quad (1)$$

其 3 阶中心矩 μ_3 用来定义 X 的偏度 $S(x)$:

$$S(x) = \mu_3 = \int_{-\infty}^{\infty} (x-E(x))^3 f(x) dx \quad (2)$$

其 4 阶中心矩 μ_4 用来定义 X 的峰度 $K(x)$,

$$K(x) = \mu_4 = \int_{-\infty}^{\infty} (x-E(x))^4 f(x) dx \quad (3)$$

式中 x 为单个样本值, μ 为中心矩, $f(x)$ 为概率密度函数。

1.5 冠层无人机数码图像 RGB 模型颜色特征参数的提取与分析

1.5.1 图像处理及其 RGB 模型颜色特征参数提取

采用 Photoshop 软件根据试验区域对样品原始图像进行切割, 并将切割后的图像保存为 .JPG 图像格式。参照 Chen 等^[15]的方法, 采用 MATLAB 2016R 软件对叶片图像色阶分布进行偏态分析, 共得到叶色偏态参数 (20 个), 即 R 通道色阶的均值 (R_{Mean})、中位数 (R_{Median})、众数 (R_{Mode})、偏度 (R_{Skewness}) 及峰度 (R_{Kurtosis}), G 通道色阶的均值 (G_{Mean})、中位数 (G_{Median})、众数 (G_{Mode})、偏度 (G_{Skewness}) 及峰度 (G_{Kurtosis}), B 通道色阶的均值 (B_{Mean})、中位数 (B_{Median})、众数 (B_{Mode})、偏度 (B_{Skewness}) 及峰度 (B_{Kurtosis}), 及 Y 图像色阶的均值 (Y_{Mean})、中位数 (Y_{Median})、众数 (Y_{Mode})、偏度 (Y_{Skewness}) 及峰度 (Y_{Kurtosis}), 其中 R_{Mean} , G_{Mean} , B_{Mean} , Y_{Mean} 构成叶色正态参数 (4 个)。

1.5.2 不同通道色阶分布的正态检验

采用 MATLAB 运用 lillietest 及 jbtest 函数分别对大豆冠层彩色图像 R 、 G 、 B 三个通道及 Y 图像色阶分布情况进行 Lilliefors 和 Jarque-Bera 正态性检验。

1.5.3 色阶累积直方图制作

采用 MATLAB 运用 imhist 函数获取图像色阶累积直方图。

1.6 数据分析及模型构建

1.6.1 方差分析

选取建模试验的 36 个小区样本, 以 3 个发育期 (花期、荚期和鼓粒期) 为分类因子, 采用 SPSS 软件对 20 个叶色偏态参数进行方差分析, 采用最小显著性差异 (Least-Significant Difference, LSD) 方法及杜肯 (Duncan's) 方法^[23]进行多重比较 (显著水平 $\alpha=0.05$), 以分析花期、荚期和鼓粒期大豆冠层图像的叶色偏态特征参数的差异性。

1.6.2 回归模型构建

选取曲茎 36 个小区样本建模, 采用 SPSS 软件, 以大豆产量作为因变量, 分别以 3 个发育期的叶色正态参数 (12 个)、叶色偏态参数 (60 个) 作为自变量, 采用基于最小二乘法的逐步回归方式^[24]建立线性模型 $F1$ 、 $F2$ (表 1)。当回归方程显著性检验 $P \leq 0.05$ 时, 自变量移

入回归方程；当回归方程显著性检验 $P \geq 0.1$ 时，自变量移出回归方程。同时，对备选回归模型进行决定系数优选、回归模型及回归系数显著性检验、回归模型自变量共线性诊断^[23]，最终确定最佳回归模型表达式。

表 1 大豆产量预测模型结构

模型 Models	自变量或输入层因子 Independent variable or input layer factors	因子数量 Number of factors	因变量或输出层因子 Dependent variable or output layer factor
F1	叶色正态参数	12	产量
F2	叶色偏态参数	60	产量

1.6.3 模型预测准确度比较

为了验证预测模型的准确度，运用模型对建模试验组样本和 2 个验证样本组（氮肥运筹试验样本和品种试验样本）产量进行预测，并对其预测准确度进行计算，如式（4）^[25]所示：

$$PA = \left(1 - \frac{|y_p - y|}{y}\right) \times 100\% \quad (4)$$

式中 PA 为模型预测准确度，%； y_p 为模型计算得到的产量预测值， kg/hm^2 ； y 为产量实测值， kg/hm^2 。

2 结果与分析

2.1 大豆群体冠层无人机图像叶色色阶分布的偏态分析及正态性检验

随机选取一个建模小区分析大豆群体冠层无人机图像的色阶累积分布情况，结果如图 1 所示。

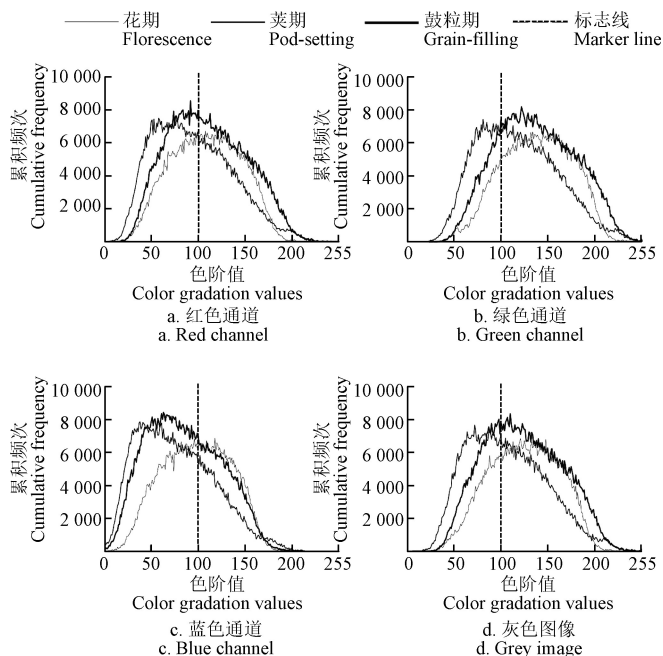


图 1 不同生育期大豆冠层叶色色阶偏态分布直方图

Fig.1 Histogram of color gradation skewness distribution for soybean canopy at different growth stages

图 1 所示，不同生育期大豆群体冠层图像红（R）、绿（G）、蓝（B）三个通道及灰度图像色阶累积直方图均呈现偏态分布；参照色阶值为 100 的标志线可看出，

不同生育期、不同通道的分布直方图还表现出不同的偏离程度，花期后随着生育期的推移，各通道色阶分布先向横坐标轴 0 端移动后向 255 端移动。进一步对各通道色阶分布数组进行 Lilliefors 和 Jarque-Bera 正态性检验，结果表明，两种检查的正态假设统计量 H 值均为 1，即否定正态分布假设；而统计量分布概率检测显著性 P 值均为 0.001，小于 0.05，说明不同生育期大豆群体冠层图像不同通道的叶色色阶分布均不符合正态分布，结合图 1 可以看出，叶色分布为偏态分布。

2.2 不同生育期大豆群体冠层无人机图像色阶分布特征差异分析

不同生育期大豆群体冠层具有不同的外观表型，反映到叶色参数上，则呈现出在不同生育期间的差异性，且不同参数的变化特征不同（表 2）。从叶色深浅来看，R、G、B 三个通道及 Y 图像的均值、中位数及众数在花期-荚期-鼓粒期呈现先降后升的变化特征，其中，三个生育期冠层图像 R、G 通道及 Y 度图像色阶众数和 B 通道色阶平均值、众数间具有显著差异，荚期冠层图像 R、G 通道及 Y 图像的平均值和中位数显著低于花期和鼓粒期。

表 2 不同生育期大豆冠层图像的叶色偏态特征参数差异

通道/图像 Channel/Image	参数 Parameters	花期 Florescence	荚期 Pod-setting	鼓粒期 Grain-filling
红色通道 Red channel	均值 R_{Mean}	108.58a	95.82b	111.12a
	中位数 R_{Median}	108.75a	93.04b	108.00a
	众数 R_{Mode}	107.42a	81.11c	90.81b
	偏度 R_{Skewness}	-0.01c	0.36a	0.27b
	峰度 R_{Kurtosis}	2.30b	2.60a	2.34b
绿色通道 Green channel	均值 G_{Mean}	136.21a	118.42b	136.21a
	中位数 G_{Median}	136.92a	116.11b	134.00a
	众数 G_{Mode}	139.28a	104.81c	122.11b
	偏度 G_{Skewness}	-0.10c	0.27a	0.14b
	峰度 G_{Kurtosis}	2.31b	2.47a	2.26c
蓝色通道 Blue channel	均值 B_{Mean}	89.12a	79.95c	85.86b
	中位数 B_{Median}	99.56a	73.81c	82.81b
	众数 B_{Mode}	101.58a	57.83c	85.06b
	偏度 B_{Skewness}	-0.04c	0.41a	0.29b
	峰度 B_{Kurtosis}	2.28b	2.64a	2.37b
灰色图像 Grey image	均值 Y_{Mean}	123.72a	107.03b	123.18a
	中位数 Y_{Median}	124.15a	104.42b	120.56a
	众数 Y_{Mode}	123.97a	90.94c	104.72b
	偏度 Y_{Skewness}	-0.06c	0.33a	0.21b
	峰度 Y_{Kurtosis}	2.29b	2.55a	2.29b

注：不同小写字母表示同一参数在不同生育期间差异显著（ $P < 0.05$ ）。
Note: Different small letters represent significantly difference of same parameter among growing stages ($P < 0.05$).

从色阶分布偏向性来看（表 2），R、G、B 三个通道及 Y 图像的偏度与均值、中位数、众数呈现相反的变化特征，二者在花期-荚期-鼓粒期一致呈现先升后降的变化特征，且 R、G、B 通道和 Y 图像色阶偏度在三个生育期间均具有显著差异。

从色阶分布集中度来看（表 2），与偏度类似，各通道峰度在花期-荚期-鼓粒期也呈现先升后降的变化特征，且荚期各通道色阶峰度和其他 2 个生育期间均具有显著差异。

由此可以看出，花期、荚期和鼓粒期大豆群体冠层图像具有不同的颜色特征，基于偏态分析得到的 20 个偏态参

数可全面系统地反映这些差异特征，这为将 RGB 模型的偏态参数用于预测大豆产量提供了生理解释及数学基础。

2.3 大豆产量预测模型的构建

分别以叶色正态参数和叶色偏态参数为自变量，构建叶色参数与大豆产量的多元回归模型，并综合考虑方程的拟合优度及自变量共线性检测结果，确定最优模型。结果发现，采用花期、荚期和鼓粒期三个时期 12 个冠层 RGB 图像正态参数无法建立大豆产量的预测模型，这可能是由于正态参数反映的大豆冠层叶色信息有限，只反映了颜色的深浅，无法全面地表达大豆冠层叶色信息特征。采用花期、荚期和鼓粒期三个时期共 60 个冠层图像偏态参数构建的叶色偏态参数-大豆产量的预测模型 F_2 ，如式（5）所示：

$$F_2=483.893PR_{Kurtosis}-459.390PY_{Kurtosis}+136.567 \quad (5)$$

式中 F_2 为大豆产量， kg/hm^2 ； $PR_{Kurtosis}$ 为荚期冠层 RGB 图像 R 通道色阶的峰度； $PY_{Kurtosis}$ 为 Y 图像色阶的峰度。从模型的拟合效果来看，模型 F_2 的决定系数 R^2 为 0.235， $P=0.012$ 且 $P<0.05$ ，通过了回归模型显著性检验，可用于产量预测。采用式(5)计算建模组大豆产量预测结果，如表 3 所示，对建模组 36 个处理小区的预测准确度普遍超 85%，平均为 91.30%，其中 23 个小区的预测准确度超 90%，13 个小区样本的预测准确度不小于 95%。

表 3 建模集基于回归模型的大豆产量预测准确度

Table 3 Accuracy of regression model for soybean yield prediction based on establishment dataset

小区号 Plot No.	实际产量 Actual yield/ (kg·hm ⁻²)	预测产量 Prediction yield/ (kg·hm ⁻²)	准确度 Accuracy/%
1	3 086	3 388	90.23
2	3 287	3 670	88.35
3	2 912	3 386	83.74
4	3 311	3 153	95.25
5	3 454	3 239	93.76
6	3 199	3 350	95.28
7	2 510	3 166	73.85
8	4 009	3 421	85.34
9	3 739	3 336	89.24
10	3 174	3 311	95.69
11	2 906	3 298	86.51
12	3 165	3 249	97.35
13	3 022	3 369	88.51
14	2 914	3 074	94.51
15	3 158	3 179	99.34
16	3 938	3 541	89.93
17	3 214	3 374	95.00
18	3 405	3 338	98.04
19	3 413	3 483	97.96
20	3 793	3 560	93.85
21	3 477	3 418	98.29
22	2 726	3 254	80.61
23	3 331	3 211	96.39
24	3 545	3 263	92.04
25	3 056	3 250	93.67
26	2 988	3 240	91.58
27	3 163	3 296	95.81
28	3 978	4 097	97.00
29	4 230	3 595	84.97
30	4 031	3 595	89.18
31	3 800	3 556	93.58
32	3 251	3 165	97.37
33	4 094	3 045	74.36
34	3 765	3 476	92.31
35	3 348	3 676	90.19
36	3 036	3 408	87.74
平均值 Average			91.30

2.4 大豆产量预测模型验证

应用模型 F_2 分别对 2 个验证组（氮肥运筹试验样本和不同品种试验样本）的产量进行预测，比较预测产量与实测值，结果见表 4。

表 4 验证集基于回归模型的大豆产量预测准确度

Table 4 Accuracy of regression model for soybean yield prediction based on validation dataset

小区号 Plot No.	氮肥运筹组 Fertilizer-N group			品种组 Variety group		
	实际值 Actual value/ (kg·hm ⁻²)	预测值 Prediction value/ (kg·hm ⁻²)	准确度 Accuracy/ %	实际值 Actual value/ (kg·hm ⁻²)	预测值 Prediction value/ (kg·hm ⁻²)	准确度 Accuracy/ %
1	3 251	3 573	90.09	3 345	3 072	91.85
2	3 056	3 847	74.12	3 792	3 098	81.70
3	3 094	3 554	85.13	2 987	3 405	86.02
4	3 024	3 517	83.71	3 077	3 056	99.33
5	2 910	3 768	70.50	3 551	2 994	84.32
6	3 127	3 631	83.90	3 176	3 177	99.99
7	3 401	3 413	99.67	2 954	2 938	99.46
8	2 945	3 496	81.30	3 417	2 983	87.29
9	3 147	3 309	94.86	2 927	3 497	80.52
10	3 485	3 346	96.01	5 233	3 153	60.26
11	2 758	3 448	74.95	4 255	3 284	77.19
12	3 035	3 213	94.13	1 940	3 441	22.68
13	3 272	3 263	99.71	4 435	2 989	67.39
14	3 815	3 344	87.66	4 112	3 118	75.81
15	3 872	3 135	80.97	4 803	3 143	65.43
16	3 457	3 363	97.27	3 388	3 076	90.80
17	3 415	3 207	93.91	2 949	3 188	91.89
18	3 848	3 234	84.03	2 529	3 679	54.52
平均值 Average			87.33			78.69

由表 4 可知，本文模型对氮肥运筹试验 18 个处理小区的预测准确度普遍超 85%，平均为 87.33%，其中 8 个小区的预测准确度超 90%，4 个小区样本的预测准确度超 95%；对不同品种试验 18 个处理小区的预测准确度平均为 78.69%，其中 6 个小区的预测准确度超 90%，3 个小区样本的预测准确度超 95%。结合建模组和验证组总体看来（表 3 和表 4），采用叶片偏态参数的模型对大豆产量的预测准确度平均可达 85.50%。

3 讨 论

基于 RGB 模型进行植物生长性状评估及产量估计已在水稻、小麦、玉米等作物中广泛应用^[14,26]。传统的图像色阶分析方法是基于正态分布的假设前提下对数据进行分析，它是一种方便的近似取值的方法，所能获取的颜色信息较少，主要以 R 、 G 、 B 等三个颜色通道的色阶均值为主，虽然有不少研究人员提出了这三个基础参数的多种组合参数，但仍不能全面、真实地反映叶色的分布情况^[27]，也无法描述叶色变化的生理学意义^[13-14]，限制 RGB 模型的更广泛应用。在本试验条件下，花期、荚期和鼓粒期等三个生育期只有 12 个冠层 RGB 图像正态参数，无法运用多元逐步回归法拟合大豆产量的线性变化，这可能是由于正态参数只是反映了冠层颜色的深浅，不能反映其他生长特征。

单张叶片图像 RGB 模型各通道色阶遵循偏态分布^[13-15]。而本研究应用大豆群体冠层图像,验证了大豆在花期、荚期和鼓粒期三个不同生育期的冠层数码图像的 RGB 模型色阶同样遵循偏态分布。本研究中三个生育期各通道色阶峰度均小于 3 (见表 2),说明了色阶分布曲线较扁平^[16],即冠层叶色色阶分布分散,集中度低,反映田间大豆个体植株长势存在一定不均一性。不同生育时期各通道色阶的偏度在-0.5~0.5 之间(见表 2),说明各通道色阶近似正态分布^[16],但有一定的偏向性。可见基于偏态分析得到的叶色参数可系统地从叶色深浅程度、分布偏向性和集中度 3 个方面定量地描述冠层颜色信息。进一步分析不同生育期大豆冠层图像颜色的偏态参数特征发现,大豆冠层图像颜色的 5 类偏态参数在花期、荚期和鼓粒期间普遍具有显著差异,且反映叶色深浅信息的色阶均值、中位数、众数从花期至鼓粒期的变化趋势与反映叶色偏向性的偏度和反映叶色集中度的峰度相反。这说明基于数码图像 RGB 模型的色阶偏态参数在各生育期具有不同的特征,它们可较精准地反映各生育期大豆冠层颜色特征,为将 RGB 模型色阶的偏态参数用于预估大豆产量提供了生理解释及数学基础。

对 3 个生育时期的 60 个偏态参数进行逐步回归分析,可构建基于大豆荚期冠层 RGB 图像 R 通道色阶的峰度 ($PR_{Kurtosis}$) 和 Y 图像色阶的峰度 ($PY_{Kurtosis}$) 的多元逐步回归方程,且方程通过了显著性水平检验。这可能是由于作物群体是由许多个体有机组成的,个体与群体之间、个体与个体之间彼此制约、相互影响^[28-29],从而形成有着不同叶色均匀程度的冠层结构,反映到参数值上,表现为不同的峰度。同时,模型在建模组和验证组均具有较高的预估精度,其中对建模组(曲茎)预测准确度平均达 91.30%,各处理小区的预测准确度普遍超 85%;对不同栽培措施(氮肥运筹)处理的预测准确度也普遍超 85%,平均为 87.33%;对不同品种的预测准确度略低,平均为 78.69%。

同时也发现,基于偏态参数构建的大豆产量预测模型虽然通过了显著性检验,但 R^2 较低,这说明模型精度还有更多地提升空间。除了 RGB 模型参数外,基于数码图像可提取的颜色信息还包括 HSV 颜色模型参数(色调、饱和度和明度)^[30]、Lab 颜色模型参数(亮度、a 通道值及 b 通道值)^[31];而除了颜色信息外,还可提取叶面纹理参数(如对比度、相关度、熵、同质性和对角二阶矩)^[32]。这些多维参数能够更加全面准确地反映大豆冠层图像信息,可以为产量预测模型的构建提供更多的信息输入因子,可能进一步提高模型精度,为利用图像信息进行大豆产量模型构建提供更多的可能。

综上所述,基于偏态分布模式的 RGB 模型能提供更详细的大豆冠层图像信息,系统地从深浅程度和分布偏向性、集中度 3 个方面定量地描述冠层颜色信息,从而建立具有较高准确性的产量预测模型,可广泛用于不同生产条件的大豆产量预测。同时无人机^[33]和数码相机的使用提高了图像获取的效率并降低了图像获取的成本,更有利于本方法的推广应用。

4 结 论

本文探索了偏态分布模式加色混色 (Red-Green-Bule, RGB) 模型在大豆产量预测上的应用。结果如下:

1) 大豆冠层数码图像的光学三原色 RGB 模型色阶遵循偏态分布。通过偏态分析得到的偏态参数可很好地表征不同生育期冠层颜色信息特征,且不同类型叶色信息参数随着生育期推移呈现不同的变化规律。

2) 基于 RGB 模型偏态参数可构建大豆产量预测模型,模型通过了显著性检验。且模型对不同品种和不同栽培措施下大豆产量预测效果均较好,预测准确度平均达 85.50%。基于大豆冠层图像信息进行实时、快速和无损地预测大豆产量,可弥补传统实地取样测产耗时耗力、破坏性和主观误差大等不足。

模型精度还有更多的提升空间。未来将继续探索可描述冠层信息的其他颜色参数和纹理参数,为产量预测模型的构建提供更多更全面的冠层信息,最终为无损快速的智能化的大豆产量预测提供更多的思路与方法。

致谢:感谢加拿大卡尔顿大学的安子越使用 MATLAB 软件对数据作处理。

[参 考 文 献]

- [1] 李红,李松,李民赞,等. 农业信息成像感知与深度学习研究进展[J]. 农业机械学报, 2020, 51(5): 1-17.
Li Hong, Li Song, Li Minzan, et al. Research progress of image sensing and deep learning in agriculture[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(5): 1-17. (in Chinese with English abstract)
- [2] 张慧春,周宏平,郑加强,等. 植物表型平台与图像分析技术研究进展与展望[J]. 农业机械学报, 2020, 51(3): 1-17.
Zhang Huichun, Zhou Hongping, Zheng Jiaqing, et al. Research progress and prospect in plant phenotyping platform and image analysis technology[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(3): 1-17. (in Chinese with English abstract)
- [3] 袁培森,李润隆,任守纲,等. 表示学习技术研究进展及其在植物表型中应用分析[J]. 农业机械学报, 2020, 51(6): 1-14.
Yuan Peisen, Li Runlong, Ren Shougang, et al. State-of-the-art review for representation learning and its application in plant phenotypes[J]. Transactions of the Chinese Society for Agricultural Machinery, 2020, 51(6): 1-14. (in Chinese with English abstract)
- [4] Chen D J, Neumann K, Friedel S, et al. Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis[J]. Plant Cell, 2014, 26(46): 36-55.
- [5] 黄芬,高帅,姚霞,等. 基于机器学习和多颜色空间的冬小麦叶片氮含量估算方法研究[J]. 南京农业大学学报, 2019, 43(2): 364-371.
Huang Fen, Gao Shuai, Yao Xia, et al. Estimation of winter wheat leaf nitrogen concentration using machine learning

- algorithm and multi-color space[J]. Journal of Nanjing Agricultural University, 2019, 43(2): 364-371. (in Chinese with English abstract)
- [6] Vasseur F, Bresson J, Wang G, et al. Image-based methods for phenotyping growth dynamics and fitness components in *Arabidopsis thaliana*[J]. Plant Methods. 2018, 14: 63.
- [7] Dey A K, Sharma M, Meshram M R. An analysis of leaf chlorophyll measurement method using chlorophyll meter and image processing technique[J]. Procedia Computer Science, 2016, 85: 286-292.
- [8] Barbedo J G A. Detection of nutrition deficiencies in plants using proximal images and machine learning: A review[J]. Computers and Electronics in Agriculture, 2019, 162: 482-492.
- [9] Humplik J F, Lazár D, Husíková A, et al. Automated phenotyping of plant shoots using imaging methods for analysis of plant stress responses-a review[J]. Plant Methods. 2015a, 11, 29.
- [10] Wang L, Cen H Y, Zhu J P, et al. Grain yield prediction of rice using multi-temporal UAV-based RGB and multispectral images and model transfer – a case study of small farmlands in the South of China[J]. Agricultural and Forest Meteorology, 2020, 291: 1-15.
- [11] Li B, Xu X M, Zhang L, et al. Above-ground biomass estimation and yield prediction in potato by using UAV-based RGB and hyperspectral imaging[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 162: 161-172.
- [12] Barker J, Zhang N Q, Sharon J, et al. Development and evaluation of a field-based high-throughput phenotyping platform[J]. Computers and Electronics in Agriculture, 2016, 122: 74-85.
- [13] 吴雪梅, 张富贵, 吕敬堂. 基于图像颜色信息的茶叶嫩叶识别方法研究[J]. 茶叶科学, 2013, 33(6): 584-589.
Wu Xuemei, Zhang Fugui, Lv Jingtang. Research on recognition of tea tender leaf based on image color information[J]. Journal of Tea Science, 2013, 33, 584-589. (in Chinese with English abstract)
- [14] Han W T, Sun Y, Xu T F, et al. Detecting maize leaf water status by using digital RGB images[J]. International Journal of Agricultural and Biological Engineering, 2014, 7: 45-53.
- [15] Chen Z M, Wang F Z, Zhang P, et al. Skewed distribution of leaf color RGB model and application of skewed parameters in leaf color description model[J]. Plant Methods, 2020, 16: 23.
- [16] 贾俊平. 统计学[M]. 北京: 中国人民大学出版社. 2018.
- [17] Rigon J P G, Capuani S, Fernandes D M, et al. A novel method for the estimation of soybean chlorophyll content using a smartphone and image analysis[J]. Photosynthetica, 2016, 54 (4): 559-566.
- [18] Matloobi M, Tahmasebi S, Dadpour M. Predicting cut Rose stages of development and leaf color variations by means of image analysis technique[J]. Journal of Ornamental Plants, 2017, 7(1): 25-36.
- [19] Bai G, Jenkins S, Yuan W, et al. Field-based scoring of soybean iron deficiency chlorosis using RGB imaging and statistical learning[J]. Frontiers in Plant Science, 2018, 9: 1002.
- [20] 邱野, 王瑄. 耕作模式对坡耕地土壤水分和大豆产量的影响[J]. 农业工程学报, 2018, 34(22): 128-137.
Qiu Ye, Wang Xuan. Effects of tillage patterns on soil moisture and soybean yield in sloping fields[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2018, 34(22): 128-137. (in Chinese with English abstract)
- [21] 朱元刚, 高凤菊, 曹鹏鹏, 等. 种植密度对玉米-大豆间作群体产量和经济产值的影响[J]. 应用生态学报, 2015, 26(6): 1751-1758.
Zhu Yuangang, Gao Fengju, Cao Pengpeng, et al. Effect of plant density on population yield and economic output value in maize-soybean intercropping[J]. Chinese Journal of Applied Ecology, 2015, 26(6): 1751-1758. (in Chinese with English abstract)
- [22] 符小文, 张永杰, 杜孝敬, 等. 麦-豆轮作体系周年施氮量对夏大豆氮素利用效率和产量的影响[J]. 植物营养与肥料学报, 2020, 26(3): 453-460.
Fu Xiaowen, Zhang Yongjie, Du Xiaojing, et al. Effect of annual nitrogen application rate on nitrogen use efficiency and yield of summer soybean in winter wheat-summer soybean rotation system[J]. Journal of Plant Nutrition and Fertilizers, 2020, 26(3): 453-460. (in Chinese with English abstract)
- [23] 盖钧益. 试验统计方法[M]. 北京: 中国农业出版社, 2000.
- [24] 尼加提·卡斯木, 师庆东, 王敬哲, 等. 基于高光谱特征和偏最小二乘法的春小麦叶绿素含量估算[J]. 农业工程学报, 2017, 33(22): 208-216.
Nijat Kasim, Shi Qingdong, Wang Jingzhe, et al. Estimation of spring wheat chlorophyll content based on hyperspectral features and PLSR model[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2017, 33(22): 208-216. (in Chinese with English abstract)
- [25] 张佩, 陈郑盟, 刘春伟, 等. 冬小麦产量结构要素预报方法[J]. 农业工程学报, 2020, 36(8): 78-87.
Zhang Pei, Chen Zhengmeng, Liu Chunwei, et al. Method for the prediction of wheat yield components[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2020, 36(8): 78-87. (in Chinese with English abstract)
- [26] Ali M M, Al-Ani A, Eamus D, et al. New image processing based technique to determine chlorophyll in plants[J]. American-Eurasian Journal of Agricultural and Environmental Sciences, 2012, 12(10): 1323-1328.
- [27] Li L, Zhang Q, Huang D F. A review of imaging techniques for plant phenotyping[J]. Sensors. 2014, 14(11): 20078.
- [28] 李毅念, 杜世伟, 姚敏, 等. 基于小麦群体图像的田间麦穗计数及产量预测方法[J]. 农业工程学报, 2018, 34(21): 185-194.
Li Yinian, Du Shiwei, Yao Min, et al. Wheat and maize yield model estimation based on MODIS and meteorological data in shaanxi province[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2018, 34(21): 185-194. (in Chinese with English abstract)
- [29] 樊高琼, 郑亨, 陈溢, 等. 耕作方式、播深及覆土对机播

- 套作小麦群体质量和产量的影响[J]. 农业工程学报, 2011, 27(14): 20-25.
- Fan Gaoqiong, Zheng Ting, Chen Yi, et al. Effects of tillage managements, sowing depth and soil-covering on population quality and yield of mechanical sowing wheat under relay intercropping condition[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2011, 27(14): 20-25. (in Chinese with English abstract)
- [30] 张凯兵, 章爱群, 李春生. 基于 HSV 空间颜色直方图的油菜叶片缺素诊断[J]. 农业工程学报, 2016, 32(19): 179-187.
- Zhang Kaibing, Zhang Aiqun, Li Chunsheng. Nutrient deficiency diagnosis method for rape leaves using color histogram on HSV space[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2016, 32(19): 179-187. (in Chinese with English abstract)
- [31] 曹英丽, 林明童, 郭忠辉, 等. 基于 Lab 颜色空间的非监督 GMM 水稻无人机图像分割[J]. 农业机械学报, 2021, 52(1): 162-169.
- Cao Yingli, Lin Mingtong, Guo Zhonghui, et al. Unsupervised GMM for rice segmentation with UAV images based on Lab color space[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(1): 162-169
- [32] 陶华伟, 赵力, 奚吉, 等. 基于颜色及纹理特征的果蔬种类识别方法[J]. 农业工程学报, 2014, 30(16): 305-311.
- Tao Huawei, Zhao Li, Xi Ji, et al. Fruits and vegetables recognition based on color and texture features[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2014, 30(16): 305-311. (in Chinese with English abstract)
- [33] 杨雪峰, 咎梅, 木尼热·买买提. 基于无人机和卫星遥感的胡杨林地上生物量估算[J]. 农业工程学报, 2021, 37(1): 77-83.
- Yang Xuefeng, Zan Mei, Munire·Maimaiti. Estimation of above ground biomass of Populus euphratica forest using UAV and satellite remote sensing[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2021, 37(1): 77-83. (in Chinese with English abstract)

Prediction of soybean yield by using RGB model with skew distribution pattern of canopy leaf color

Zhang Pei^{1,2}, Chen Zhengmeng³, Ma Shundeng¹, Yin Di¹, Jiang Haidong^{1*}

(1. Key Laboratory of Crop Physiology and Ecology in Southern China, Ministry of Agriculture, Jiangsu Collaborative Innovation Center for Modern Crop Production, National Engineering and Technology Center for Information Agriculture, Nanjing Agricultural University, Nanjing 210095, China; 2. Jiangsu Meteorological Bureau, Nanjing 210008, China; 3. Longyan Company of Fujian Provincial Tobacco Corporation, Longyan 364000, China)

Abstract: With the increasing maturity of digital imaging technology and the increasing popularity of high resolution camera equipment, the advantages of high resolution and low cost have prompted the use of digital imaging technology to conduct more qualitative and quantitative descriptions of phenotypic traits for plant appearance. The RGB model is the most commonly used color representation for digital images. In order to explore the feasibility of using color gradation distribution parameters of the RGB model in soybean yield prediction, and to verify the universality of the method in different fertilizer operations and varieties, two soybean varieties, Qujing and Xudou 18, were selected to design field fissure experiments with different densities and nitrogen fertilizer levels in this study. Digital cameras were carried by Unmanned Aerial Vehicle (UAV) to collect soybean canopy digital images during three important reproductive growth stages. The results showed that the cumulative distribution of canopy color gradation of soybean at the florescence, pod-setting and grain-filling stages, all conformed to the skewed distribution, and a total of 20 Color Gradation Skewness-Distribution (CGSD) parameters were obtained by skew analysis. These parameters simultaneously described the shade and distribution of the canopy leaf color. The 20 CGSD parameters were significantly different among the florescence, pod-setting and grain-filling stages. And the variation trend of color depth parameters (mean, median, and mode) was opposite to that of the distribution parameters (skewness and kurtosis). The prediction model of soybean yield by using prediction model multiple stepwise regression method was constructed based on CGSD parameters with *P* value of 0.012. The model had high estimation accuracy in both the modeling group and the verification groups. The prediction accuracy of the model in modeling group reached 91.30% on average; the average prediction accuracy of 18 plots in the nitrogen operation research validation group was 87.33%. Although the prediction accuracy of the validation group for different varieties was lower than that of the modeling group and the validation group for nitrogen fertilizer operation research, it was also close to 80%. In conclusion, the RGB color model based on skewness distribution provided detailed soybean canopy image information, and the canopy color information quantitatively described systematically from the degree of depth, distribution bias and uniformity. And thus the yield prediction model based on CGSD parameters had high prediction accuracy, which can be widely used to predict yield of soybean grown in different production conditions. At the same time, the use of UAV and digital cameras improves the efficiency of image acquisition, while reduces the cost of image acquisition, which is more conducive to the popularization and application of this method.

Keywords: UAV; nitrogen; image analysis; soybean; RGB models; skewness-distribution parameters; yield prediction