

doi 10.18699/vjgb-25-64

## Определение числа соприкасающихся зерен пшеницы на изображениях на основе эллиптической аппроксимации

Д.Р. Авзалов<sup>1,2</sup>, Е.Г. Комышев<sup>1,3</sup>, Д.А. Афонников <sup>1,2,3</sup> <sup>1</sup> Федеральное исследовательское учреждение Институт цитологии и генетики Сибирского отделения Российской академии наук, Новосибирск, Россия<sup>2</sup> Новосибирский национальный исследовательский государственный университет, Новосибирск, Россия<sup>3</sup> Курчатowski геномный центр ИЦиГ СО РАН, Новосибирск, Россия ada@bionet.nsc.ru

**Аннотация.** Количество зерен растения напрямую характеризует его урожайность, а размер и форма тесно связаны с массой семян. Для оценки количества зерен, их формы и размеров в настоящее время, как правило, используют анализ цифровых изображений. Зерна на таких изображениях могут быть полностью разделены, соприкасаться или быть плотно упакованными. В случае разделенных зерен высокую точность выделения и подсчета зерен на изображении позволяют получить самые простые алгоритмы бинаризации/сегментации, например алгоритм водораздела. Но в случае соприкасающихся зерен простые алгоритмы машинного зрения могут приводить к неточностям в определении контуров отдельных зерен. В этой связи актуальными являются методы точного определения контуров индивидуальных зерен в случае их соприкосновения. Один из подходов основан на поиске пикселей области соприкосновения зерен, в частности с помощью поиска угловых точек на границе контура зерен. Однако зерна могут иметь сколы, впадины и выпуклости, что приводит к идентификации угловых точек, которые не соответствуют области контакта зерен. Это влечет за собой ошибки и для их устранения требует дополнительной обработки данных, фильтрации ложных угловых точек. В настоящей работе мы предлагаем алгоритм идентификации зерен пшеницы на изображении, который позволяет идентифицировать касающиеся зерна и определять их границы на изображении. Он базируется на модификации алгоритма поиска угловых точек и использует метод отнесения пикселей границы контура к одному зерну на основе аппроксимации контуров зерен эллипсами. Мы показали на тестовых изображениях, что предложенный алгоритм позволяет более точно идентифицировать зерна на изображении по сравнению с алгоритмом без такой аппроксимации и алгоритмом водораздела. Однако временные затраты для такого алгоритма существенны и быстро растут с увеличением количества зерен и контуров, включающих несколько зерен.

**Ключевые слова:** пшеница; зерна; подсчет; цифровые изображения; сегментация; алгоритм; угловые точки

**Для цитирования:** Авзалов Д.Р., Комышев Е.Г., Афонников Д.А. Определение числа соприкасающихся зерен пшеницы на изображениях на основе эллиптической аппроксимации. *Вавиловский журнал генетики и селекции*. 2025;29(4): 608-614. doi 10.18699/vjgb-25-64

**Финансирование.** Работа выполнена за счет финансирования Курчатовского геномного центра Федерального исследовательского центра ИЦиГ СО РАН, соглашение с Министерством образования и науки РФ № 075-15-2019-1662.

## Counting touching wheat grains in images based on elliptical approximation

D.R. Avzalov<sup>1,2</sup>, E.G. Komyshev<sup>1,3</sup>, D.A. Afonnikov <sup>1,2,3</sup> <sup>1</sup> Institute of Cytology and Genetics of the Siberian Branch of the Russian Academy of Sciences, Novosibirsk, Russia<sup>2</sup> Novosibirsk State University, Novosibirsk, Russia<sup>3</sup> Kurchatov Genomic Center of ICG SB RAS, Novosibirsk, Russia ada@bionet.nsc.ru

**Abstract.** The number of grains of a cereal plant characterizes its yield, while grain size and shape are closely related to its weight. To estimate the number of grains, their shape and size, digital image analysis is now generally used. The grains in such images may be completely separated, touching or densely packed. In the first case, the simplest binarization/segmentation algorithms, such as the watershed algorithm, can achieve high accuracy in segmentation and counting grains in an image. However, in the case of touching grains, simple machine vision algorithms may lead to inaccuracies in determining the contours of individual grains. Therefore, methods for accurately determining the contours of individual grains when they are in contact are relevant. One approach is based on the search for pixels of the grain contact area, in particular, by identification of concave points on the grain contour boundary. However, some grains may have chips, depressions and bulges, which leads to the identification of the corner points that do not correspond to the grain contact region. Additional data processing is required to avoid these errors. In this paper, we propose an algorithm for the identification of wheat grains in an image and determine their boundaries in the case when they are touching. The algorithm is based on using a

modification of the concave point search algorithm and utilizes a method of assigning contour boundary pixels to a single grain based on approximation of grain contours by ellipses. We have shown that the proposed algorithm can identify grains in the image more accurately compared to the algorithm without such approximation and the watershed algorithm. However, the time cost for such an algorithm is significant and grows rapidly with increasing number of grains and contours including multiple grains.

**Key words:** wheat; grains; counting; digital images; segmentation; algorithm; concave points

**For citation:** Avzalov D.R., Komyshev E.G., Afonnikov D.A. Counting touching wheat grains in images based on elliptical approximation. *Vavilovskii Zhurnal Genetiki i Selekcii = Vavilov J Genet Breed.* 2025;29(4):608-614. doi 10.18699/vjgb-25-64

## Введение

Одно из важнейших направлений селекционно-генетических исследований злаков – поиск и изучение генов, контролирующих показатели урожайности. Количество зерен растения напрямую характеризует его урожайность, а размер и форма тесно связаны с массой семян (Zhang X. et al., 2014; Brinton, Uauy, 2019). Оценка числа зерен, их формы и размеров в настоящее время, как правило, выполняется при помощи высокопроизводительного фенотипирования (Афонников и др., 2016; Li et al., 2020; Kolhar, Jagtap, 2023), которое основано на анализе цифровых изображений (Tanabata et al., 2012; Whan et al., 2014; Komyshev et al., 2017). Эти методы отличаются высокой производительностью, низкой стоимостью оборудования и простотой протоколов получения изображений и их обработки. Дополнительное преимущество методов заключается в том, что можно точно определить не только число зерен, но и их характеристики (размер, форму, цвет) (Tanabata et al., 2012; Cervantes et al., 2016; Комышев и др., 2020), что при ручном подсчете сделать затруднительно или невозможно. Кроме того, полученные изображения могут храниться в электронном виде без изменений, тогда как характеристики зерен в зависимости от длительности хранения могут изменяться (Afonnikov et al., 2022).

Типичные изображения для анализа – зерна на светлом фоне, полученные с помощью цифровой камеры, настольного сканера или смартфона (Herridge et al., 2011; Tanabata et al., 2012; Whan et al., 2014; Komyshev et al., 2017). Зерна на таких изображениях могут быть полностью разделены, соприкасаться или быть плотно упакованными. В случае разделенных зерен для выявления зерен на изображении можно использовать самые простые алгоритмы бинаризации/сегментации, например алгоритм водораздела (Roerdink, Meijster, 2000). При этом можно оценить размер и форму каждого зерна (Mebatsion et al., 2013). Однако такой протокол требует от пользователя заметных временных затрат на аккуратное размещение зерен, что затрудняет анализ большого количества образцов. В случае плотной упаковки определить контуры индивидуальных зерен сложно из-за их возможного наложения, и объективную оценку можно ожидать только для количества зерен на изображении, но не для размера и формы. Во втором варианте простые алгоритмы машинного зрения могут приводить к неточностям в определении контуров отдельных зерен, но характеристики их формы и размера могут быть оценены; при этом не требуется аккуратного разделения зерен, что снижает временные затраты при анализе. В этой связи разработка методов определения контуров индивидуальных зерен в случае их соприкосновения на изображениях является актуальной.

Для решения этой задачи применяют подходы на основе машинного зрения (Wang, Paliwal, 2006; Qin et al., 2013) и глубокого машинного обучения (Yang et al., 2021). Методы глубокого обучения в настоящее время активно развиваются, но требуют для обучения больших выборок изображений, размеченных вручную, что является трудозатратным. Алгоритмы машинного зрения менее требовательны к размеру обучающих выборок и их аннотации и тоже активно разрабатываются (Liang et al., 2022; Lin et al., 2023). Для выделения зерен они используют алгоритмы бинаризации, а после выделения контуров зерен применяют анализ формы сложных контуров, которые включают два и более зерен. Один из подходов для анализа сложных контуров основан на поиске пикселей области соприкосновения зерен. Для этого используется алгоритм идентификации угловых точек на границе контура зерен (Gao et al., 2017; Liu et al., 2017; Tan et al., 2019; Liang et al., 2022; Zhang J. et al., 2022). Данный алгоритм позволяет быстро идентифицировать точки резкого изгиба границы контура. Пиксели, в которых изгиб наибольший, считаются возможными местами касания зерен. Зерна на изображении могут иметь выпуклости, впадины и выпуклости, что приводит к идентификации угловых точек, которые не соответствуют области контакта зерен. Для устранения ошибок требуется дополнительная обработка данных, фильтрация ложных угловых точек.

В настоящей работе мы предлагаем алгоритм идентификации зерен пшеницы на изображении, который позволяет идентифицировать касающиеся зерна и определять их границы на изображении. Он является модификацией алгоритма поиска угловых точек и относит пиксели границы контура к одному зерну на основе аппроксимации контуров зерен эллипсами.

## Материалы и методы

**Идентификация контуров зерен на изображении.** Общая схема метода анализа изображений представлена на рис. 1. Она включает следующие этапы:

- предварительная обработка изображений (уменьшение размера и Гауссов фильтр);
- бинаризация (выделение областей контуров зерен и фона);
- поиск угловых точек на контуре;
- постобработка пикселей границы контура на основе аппроксимации эллипсами.

Каждый из этапов включает в себя использование различных алгоритмов компьютерного зрения, которые будут описаны далее.

Исходные изображения имеют высокое разрешение, 3968×2976 пикселей, что отрицательно сказывается на



Рис. 1. Основные этапы обработки изображений для подсчета зерен пшеницы.

времени работы алгоритмов классического компьютерного зрения. Эмпирически было установлено, что уменьшение разрешения изображения в два раза (до размеров 1984×1488 пикселей) в несколько раз сокращает время работы алгоритмов без существенной потери точности подсчета количества зерен. Для этого использовался метод, основанный на билинейной интерполяции (Гонсалес, Вудс, 2005). В исходном изображении выбирались сканирующие неперекрывающиеся окна размером 2×2 пикселя, которые в преобразованном изображении заменялись на один пиксель. Значения интенсивности красной, зеленой и синей компонент этого пикселя вычислялись на основе соответствующих компонент пикселей сканирующего окна.

Пусть  $(x_i, y_j)$ ,  $i, j = 1, 2$ , – четыре соседние точки, расположенные «квадратом», в каждой из которых значение пикселя равно  $f(x_i, y_j)$ . Эти четыре значения заменяются одним по следующему алгоритму:

1. Линейная интерполяция в направлении  $X$ :

$$f(x, y_i) = f(x_1, y_i) + f(x_2, y_i). \quad (1)$$

2. Линейная интерполяция полученных значений в направлении  $Y$ :

$$f(x, y) = f(x, y_1) + f(x, y_2). \quad (2)$$

В итоге разрешение изображения уменьшается в два раза.

К полученному изображению с пониженной размерностью применяется Гауссов фильтр для устранения шумов (Gedraite, Hadad, 2011). Отфильтрованное изображение переводится в цветовое пространство HSV, в которой различия между зернами и фоном более явно выражены (Фисенко В.Т., Фисенко Т.Ю., 2008; Домасев, Гнатюк, 2009).

Следующий этап предварительной обработки заключается в дополнительном сглаживании изображений с помощью алгоритма сдвига среднего (Comaniciu, Meer, 1999). Результирующее изображение преобразуется в мо-

нохромное и бинаризуется с помощью метода Оцу (Otsu, 1979).

Примеры выполнения этих шагов для масштабированного изображения после Гауссовой фильтрации показаны на рис. 2.

**Анализ контуров и идентификация угловых пикселей.** После получения бинарной маски на изображении идентифицируются все контуры, соответствующие областям зерен. Контур – это кривая, соединяющая все пиксели границы маски области зерен. Далее каждый контур анализируется независимо от других.

Производится обход контура и для каждого пикселя  $p$  границы вычисляется значение функции углового отклика ( $CRF$  – corner response function) (Tan et al., 2019):

$$CRF(p) = \frac{n_p}{A}, \quad A = \pi R^2, \quad (3)$$

где  $n_p$  – число пикселей, которые принадлежат контуру внутри окружности с радиусом  $R$  и центром в точке  $p$ ;  $A$  – полное число пикселей в этой окружности (рис. 3, а).

Методом перебора из множества  $[3, 4, \dots, 10]$  было подобрано значение  $R = 7$ , при котором обеспечивается наиболее высокая точность (при  $R > 10$  точность падала). Величина  $CRF$  близка к  $\sim 0.5$  на «ровной» части границы контура зерна. Если внутри окружности большая доля пикселей зерен, что характерно для угловых пикселей, то  $CRF$  достигает пика (см. рис. 3, б). Это позволяет идентифицировать угловые точки на изображении (см. рис. 3, в).

Мы использовали критерий  $CRF > 0.6$  для отнесения пикселя границы к угловым, как это сделано в работе (Zhang J. et al., 2022). Зная количество угловых точек  $N_{corners}$  и число сегментов границы контура между угловыми точками  $R_{closed}$ , можно оценить количество зерен как  $N_{grains} = N_{corners}/2 - R_{closed} + 1$  (Liu et al., 2017). Недостаток алгоритма заключается в предположении об идеальном

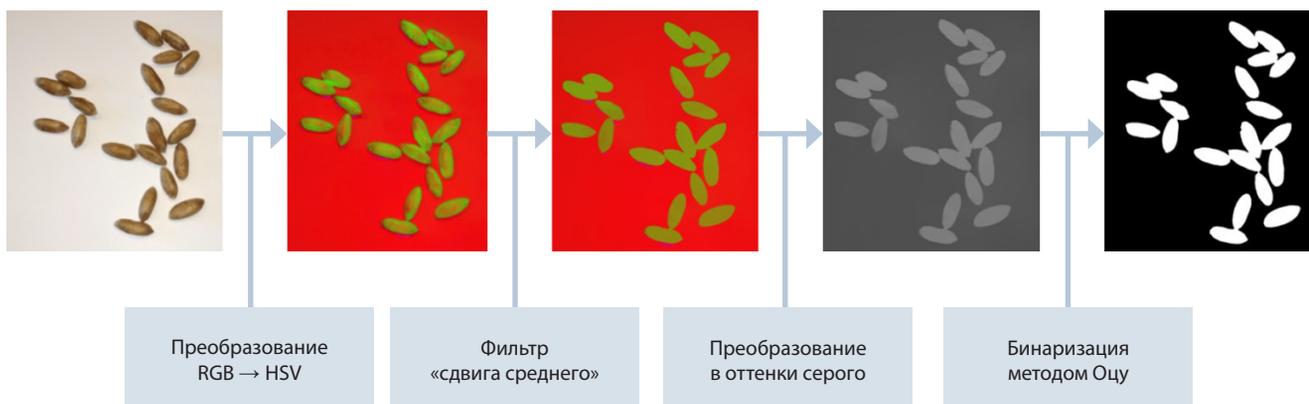
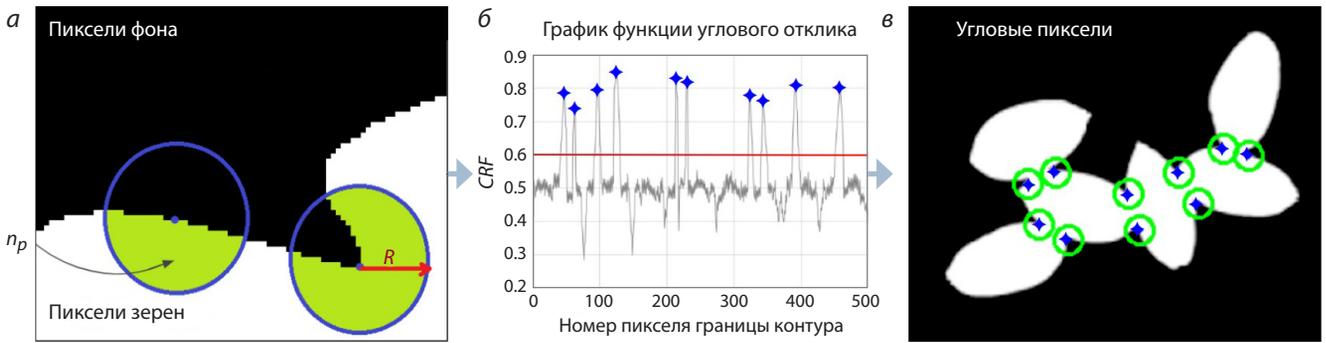
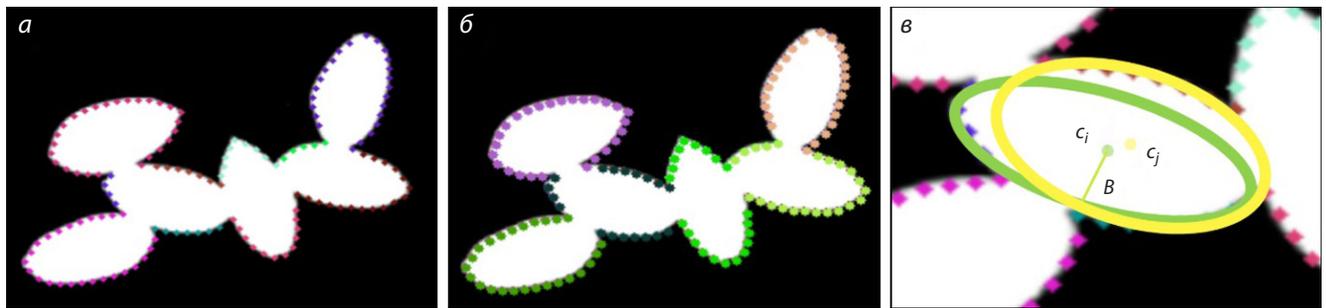


Рис. 2. Результаты предварительной обработки и бинаризации изображений зерен.



**Рис. 3.** Использование функции углового отклика для идентификации угловых пикселей контура.

*a* – визуализация расчета *CRF*; пиксели фона отображены черным цветом, пиксели зерен – белым цветом за пределами окружности расчета *CRF*, зеленым – в пределах окружности; *b* – график функции *CRF* (ось *Y*) в зависимости от номера пикселя контура (ось *X*); пики, соответствующие угловым точкам, показаны синими крестиками; *в* – демонстрация определения угловых пикселей (синие крестики, обведенные зелеными окружностями) на изображении маски зерен.



**Рис. 4.** Алгоритм аппроксимации эллипсами для идентификации сегментов границы контура, принадлежащих одному зерну.

*a* – точки, принадлежащие одному сегменту границы между угловыми пикселями, окрашены одним цветом; *b* – точки границы контура, принадлежащие одному зерну, окрашены одинаковым цветом; *в* – альтернативное положение двух эллипсов для сегментов одного зерна, их центры  $c_i$  и  $c_j$  показаны желтой и зеленой точками, показана ось эллипса  $B$ .

эллиптическом контуре зерна. На реальных же примерах из-за сколов и неровностей на зернах детектируются ложные угловые точки.

**Аппроксимация контуров зерен эллипсами.** Поскольку зерна имеют форму, близкую к эллиптической, их можно аппроксимировать эллипсами, за счет чего разбить общий контур группы зерен на отдельные зерна. Знание местоположения угловых точек (даже если среди них есть ложные) позволяет разбить весь контур на множество сегментов, каждый из которых принадлежит только одному зерну. Задачей является разбиение множества сегментов на подмножества таким образом, чтобы каждое подмножество представляло сегменты границы одного зерна (рис. 4, *a*, *б*).

Алгоритм основывается на поиске оптимального разбиения сегментов контура по принадлежности к эллипсам, в котором ошибка разбиения будет минимальной. Ошибка разбиения – суммарная ошибка аппроксимации пикселями каждого из эллипсов. Дополнительно к значению ошибки добавляли штраф, учитывающий тот факт, что для нескольких сегментов контура можно выбрать несколько эллипсов, которые будут давать близкие значения аппроксимации эллипсов; он позволял выбрать из них тот, который имеет положение центра, наиболее близкое к среднему.

Рассмотрим набор пикселей границы контура  $p_{ik}$ ,  $i = 1, \dots, m_k$ . Уравнения эллипса – квадратичная форма  $a_{11}x^2 + 2a_{12}xy + a_{22}y^2 + 2b_1x + 2b_2y + 1 = 0$ . Подставим каждую из точек в уравнение и получим переопределенную систему:

$$A = \begin{pmatrix} x_1^2 & 2x_1y_1 & y_1^2 & 2x_1 & 2y_1 \\ \dots & \dots & \dots & \dots & \dots \\ x_{m_k}^2 & 2x_{m_k}y_{m_k} & y_{m_k}^2 & 2x_{m_k} & 2y_{m_k} \end{pmatrix}, b = (-1, \dots, -1)^T, \quad (4)$$

$$\|A \cdot \alpha - b\|^2 \rightarrow \min_{\alpha}, \quad \alpha = (a_{11}, a_{12}, a_{22}, b_1, b_2). \quad (5)$$

Оптимальное решение такой системы ищется методом наименьших квадратов и разложения SVD (Brinton, Уауу, 2019). Решение  $\alpha^*$  можно найти как

$$\alpha^* = A^+b = UD^{-1}V^Tb. \quad (6)$$

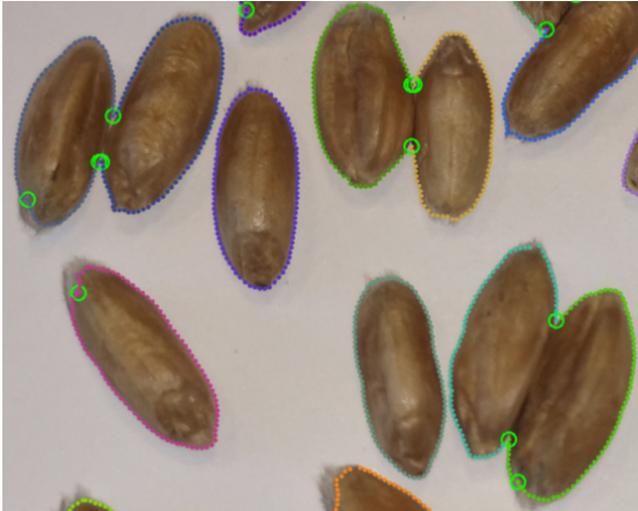
В этом случае суммарную ошибку разбиения находим по следующей формуле:

$$E_{\text{partition}} = \sum_k E_k + P, \quad E_k = \|A \cdot \alpha_k^* - b\|, \quad (7)$$

где

$$P = \sum_{i=1}^N \sum_{j=i+1}^N \frac{2}{B \cdot \|\text{center}_i - \text{center}_j\|} \quad (8)$$

– это штраф за «неправильное» разбиение, а  $B$  – минимальная ось построенных эллипсов (рис. 4, *в*).



**Рис. 5.** Результат разбиения контуров нескольких групп зерен на основе алгоритма аппроксимации эллипсами.

Идентифицированные угловые пиксели показаны зелеными окружностями, пиксели контура, принадлежащие одному зерну, – одним цветом.

В результате нахождения разбиения с наименьшей ошибкой количество зерен равняется количеству подмножеств в разбиении. Такой подход решает проблему с лишними угловыми точками, что сказалось на точности работы метода. Пример работы алгоритма при идентификации зерен приведен на рис. 5. Как можно видеть, для двух нижних зерен справа (бирюзовый и зеленый контуры) разбиение было построено правильно (цветами обозначены сегменты, отнесенные к отдельным подмножествам), хотя присутствует лишний угловой пиксель. То же самое относится и к другим контурам.

Описанные алгоритмы были реализованы на языке Python v.3.9 с использованием библиотек OpenCV v. 4.6.0 (Howse, 2013) и Numpy v. 1.21 (<https://numpy.org/>).

**Оценка точности идентификации зерен на изображении.** Для каждого изображения было найдено количество зерен с помощью описанного алгоритма и произведена оценка точности алгоритма с использованием метрики

$$CR = 100 \cdot \left[ 1 - \frac{|N - N^*|}{N} \right], \quad (9)$$

где  $N^*$  – число зерен, определенное с помощью предложенного алгоритма;  $N$  – истинное количество зерен. Оценивали точность для алгоритма угловых точек без дополнительного анализа контуров на основе использования эллипсов ( $CR_{cp}$ ) и с анализом контуров на основе использования эллипсов ( $CR_{cpe}$ ).

Дополнительно мы оценили среднюю точность для всех изображений идентификации зерен при использовании методов эрозии ( $CR_e$ ) и водораздела ( $CR_w$ ) (Zhang J. et al., 2022). Также для каждого изображения было посчитано полное время работы  $T$  с фильтрацией контуров на основе эллипсов. Вычисления производились на ноутбуке с процессором Intel i5 4 \* 2.9 GHz и 6 GB RAM под управлением операционной системы Windows 10.



**Рис. 6.** Пример изображения соприкасающихся зерен пшеницы.

## Результаты и обсуждение

Для анализа был взят набор из девяти изображений зерен пшеницы на белом листе бумаги. Съемку производили на смартфон HUAWEI P20, камера Sony IMX380, размер изображения 11.8 Мп (3968×2976 пикс.). Исходные изображения были сделаны с использованием вспышки. При съемке был установлен режим «авто» (по умолчанию), режим широкого динамического диапазона (high dynamic range, HDR) был отключен. Зерна на изображении были размещены неплотно, некоторые из них соприкасались (рис. 6).

На шести изображениях было по 20 зерен; выборка также включала изображения 31, 46 и 51 зерна.

Результаты подсчета количества зерен с помощью двух алгоритмов: используя только угловые точки (cp) и используя алгоритм угловых точек с построением эллипсов (cpe), представлены в таблице. Даны оценки числа зерен, идентифицированных алгоритмом, а также меры точности и времени вычислений.

Среднее значение точности ( $CR$ ) для алгоритма cp составило 0.90, для алгоритма cpe – 0.96, для алгоритма водораздела – 0.77, для алгоритма эрозии – 0.93. Видно, что алгоритм коррекции числа зерен на основе эллипсов дает наиболее точные результаты. Можно также сделать вывод, что точность алгоритма зависит от числа зерен, которые формируют сложные контуры: чем больше зерен в контуре, тем ниже точность. Точность зависит и от числа зерен, что объясняется, прежде всего, наличием большей вероятности контактирующих зерен.

Точность метода с учетом эллипсов сравнима с точностью некоторых ранее опубликованных методов, тоже учитывающих угловые точки. Так, в работе (Tan et al., 2019) для подсчета зерен риса использовали алгоритм детекции угловых точек в сочетании с нейронной сетью с обратным распространением ошибки для последующей коррекции результатов сегментации. В среднем точность сегментации зерен для различных сортов риса составила 94 %. W. Wang, J. Paliwal (2006) предложили алгоритм

Оценка точности подсчета зерен алгоритмами поиска угловых точек без и с фильтрацией контуров на основе метода эллипсов для девяти тестовых изображений

Номер изображения	$N$ , шт.	$M$ , шт.	$N_{\text{ср}}$ , шт.	$CR_{\text{ср}}$ , %	$N_{\text{срe}}$ , шт.	$CR_{\text{срe}}$ , %	$T$ , с
1	20	2	21	95.0	20	100.0	13.57
2	20	3	22	90.0	20	100.0	20.01
3	20	3	20	100.0	20	100.0	19.54
4	20	4	21	95.0	20	100.0	31.33
5	20	5	21	95.0	18	90.0	80.09
6	20	5	24	80.0	21	95.0	82.87
7	31	7	33	93.5	32	96.8	270.11
8	46	4	53	84.8	45	97.8	43.50
9	53	9	66	75.5	45	85.0	1121.30

Примечание.  $N$  – число зерен;  $M$  – максимальное число зерен в контуре;  $N_{\text{ср}}$  – число зерен, определенное алгоритмом ср;  $CR_{\text{ср}}$  – точность алгоритма ср;  $N_{\text{срe}}$  – число зерен, определенное алгоритмом срe;  $CR_{\text{срe}}$  – точность алгоритма срe;  $T$  – время работы алгоритма срe.

водораздела для изображений, полученных из исходных методом сегментации, затем преобразованным в зависимости от расстояний между пикселями фона и зерен. С его помощью проведен подсчет зерен на изображениях для шести типов растений: озимой, твердой белозерной и твердой янтарной пшеницы, ячменя, овса и ржи. Доля правильно идентифицированных зерен варьировала от 88.6 до 94.4 % для пшеницы и от 55.4 до 79.0 % для остальных трех видов. T. Liu с коллегами (2017) применили алгоритм, основанный на детекции характерных точек на изображении (feature points) и определении корреляционной зависимости между их количеством и количеством зерен на изображении. Авторы провели тестирование алгоритма на зернах пшеницы и риса и показали, что ошибка их метода составляет от 0 до 4.7 % (в среднем 0.1 % для пшеницы и 1.5 % для риса), тогда как использование обычного алгоритма водораздела приводило к ошибке от 14 до 40 % у пшеницы и от 20 до 50 % у риса. N. Liang с коллегами (2022) предложили комплексный подход, который идентифицирует число не касающихся зерен методом К-средних, послыйный алгоритм водораздела использован для идентификации и подсчета редко касающихся зерен, для плотно лежащих зерен применялся алгоритм разделяющих линий. Точность метода составила 99.65 %.

Таким образом, наш алгоритм в целом дает точность, сравнимую с существующими алгоритмами (особенно в случае, когда контуров с большим числом зерен мало). Недостатком является быстрый рост времени обработки изображения с увеличением как числа сложных контуров, так и числа анализируемых зерен. Для числа зерен более 20 и контуров со множеством зерен это время становится малопримемлемым для анализа. При увеличении количества соприкасающихся зерен в одном контуре растет число возможных комбинаций разбиений на подмножества. Более того, возрастает число пикселей контура, координаты которых используются для составления систем линейных уравнений (4). Уменьшения времени счета (в случае описанной реализации метода) можно достичь распараллеливанием алгоритма по данным (контурам на изображении). Дальнейшая оптимизация возможна за

счет предварительной грубой оценки количества зерен, например, по площади контура. Это позволит уменьшить количество разбиений.

Отметим также, что алгоритм основан на аппроксимации формы зерен эллипсами, что может быть неприменимо к зернам более сложной формы, например, бобов. Однако в общем случае данный алгоритм можно модифицировать для контуров зерен, отличных от овальной формы, что в перспективе может быть реализовано для зерен других видов растений.

## Выводы

Предложен алгоритм для идентификации и подсчета зерен на цифровых изображениях в случае, если они соприкасаются. Алгоритм основан на бинаризации изображения для выделения контуров, содержащих зерна, и последующей обработке этих контуров. Обработка заключается в поиске угловых точек на контуре и дальнейшем их отборе за счет отнесения пикселей границы контура к одному зерну на основе аппроксимации контуров зерен эллипсами. В результате такой обработки удастся исключить из анализа ложные точки соприкосновения зерен в контуре. Анализ тестовых изображений показал, что в случае, когда число касающихся зерен невелико, алгоритм позволяет получить высокую точность подсчета зерен (до 100 %, и систематически лучше, чем без аппроксимации эллипсами). Однако при наличии контуров, включающих большое число контактирующих зерен, время расчета алгоритма увеличивается во много раз (до нескольких минут), что делает его практическое применение затруднительным.

## Список литературы / References

- Фонников Д.А., Генаев М.А., Дорошков А.В., Комышев Е.Г., Пшеничникова Т.А. Методы высокопроизводительного фенотипирования растений для массовых селекционно-генетических экспериментов. *Генетика*. 2016;52(7):788-803. doi 10.7868/S001667581607002X
- [Afonnikov D.A., Genayev M.A., Doroshkov A.V., Komyshev E.G., Pshenichnikova T.A. Methods of high-throughput plant phenotyping for large-scale breeding and genetic experiments. *Russ J Genet*. 2016;52(7):688-701. doi 10.1134/S1022795416070024]

- Гонсалес Р., Вудс Р. Цифровая обработка изображений. М.: Техносфера, 2005  
[Gonzalez R.C., Woods R.E. Digital Image Processing. CRC Press, Boca Raton, FL, 2004]
- Домасев М.В., Гнатюк С.П. Цвет, управление цветом, цветовые расчеты и измерения. СПб.: Питер, 2009  
[Domasev M.V., Gnatyuk S.P. Color, Color Management, Color Calculations and Measurements. St. Petersburg: Piter Publ., 2009 (in Russian)]
- Комышев Е.Г., Генаев М.А., Афонников Д.А. Анализ цветовых и текстурных характеристик зерен злаков на цифровых изображениях. *Вавиловский журнал генетики и селекции*. 2020;24(4): 340-347. doi 10.18699/VJ20.626  
[Komyshev E.G., Genaev M.A., Afonnikov D.A. Analysis of color and texture characteristics of cereals on digital images. *Vavilovskii Zhurnal Genetiki i Seleksii = Vavilov J Genet Breed*. 2020;24(4): 340-347. DOI 10.18699/VJ20.626]
- Фисенко В.Т., Фисенко Т.Ю. Компьютерная обработка и распознавание изображений. СПб.: СПбГУ ИТМО, 2008  
[Fisenko V.T., Fisenko T.Yu. Computer Processing and Image Recognition. St. Petersburg, 2008 (in Russian)]
- Afonnikov D.A., Komyshev E.G., Efimov V.M., Genaev M.A., Koval V.S., Gierke P.U., Börner A. Relationship between the characteristics of bread wheat grains, storage time and germination. *Plants*. 2022;11(1):35. doi 10.3390/plants11010035
- Brinton J., Uauy C. A reductionist approach to dissecting grain weight and yield in wheat. *J Integr Plant Biol*. 2019;61(3):337-358. doi 10.1111/jipb.12741
- Cervantes E., Martín J.J., Saadaoui E. Updated methods for seed shape analysis. *Scientifica*. 2016;2016(1):5691825. doi 10.1155/2016/5691825
- Comaniciu D., Meer P. Mean shift analysis and applications. In: Proceedings of the Seventh IEEE International Conference on Computer Vision. Vol. 2. Kerkyra, Greece, 1999;1197-1203. doi 10.1109/ICCV.1999.790416
- Gao L., Zhao C., Liu M. Segmentation of touching seeds based on shape feature and multiple concave point detection. In: 2017 IEEE International Conference on Imaging Systems and Techniques (IST), Beijing, 2017;1-5. doi 10.1109/IST.2017.8261448
- Gedraite E.S., Hadad M. Investigation on the effect of a Gaussian Blur in image filtering and segmentation. In: Proceedings ELMAR-2011, Zadar, Croatia, 2011;393-396
- Herridge R.P., Day R.C., Baldwin S., Macknight R.C. Rapid analysis of seed size in *Arabidopsis* for mutant and QTL discovery. *Plant Methods*. 2011;7(1):3. doi 10.1186/1746-4811-7-3
- Howse J. OpenCV Computer Vision with Python. Birmingham: Packt Publishing, 2013
- Kolhar S., Jagtap J. Plant trait estimation and classification studies in plant phenotyping using machine vision. A review. *Inf Process Agric*. 2023;10(1):114-135. doi 10.1016/j.inpa.2021.02.006
- Komyshev E.G., Genaev M.A., Afonnikov D.A. Evaluation of the SeedCounter, a mobile application for grain phenotyping. *Front Plant Sci*. 2017;7:1990. doi 10.3389/fpls.2016.01990
- Li Z., Guo R., Li M., Chen Y., Li G. A review of computer vision technologies for plant phenotyping. *Comput Electron Agric*. 2020;176: 105672. doi 10.1016/j.compag.2020.105672
- Liang N., Sun S., Yu J., Taha M.F., He Y., Qiu Z. Novel segmentation method and measurement system for various grains with complex touching. *Comput Electron Agric*. 2022;202:107351. doi 10.1016/j.compag.2022.107351
- Lin W., Ma D., Su Q., Liu S., Liao H., Yao H., Xu P. Image segmentation method for physically touching soybean seeds. *Software Impacts*. 2023;18:100591. doi 10.1016/j.simpa.2023.100591
- Liu T., Chen W., Wang Y., Wu W., Sun C., Ding J., Guo W. Rice and wheat grain counting method and software development based on Android system. *Comput Electron Agric*. 2017;141:302-309. doi 10.1016/j.compag.2017.08.011
- Mebatsion H.K., Paliwal J., Jayas D.S. Automatic classification of non-touching cereal grains in digital images using limited morphological and color features. *Comput Electron Agric*. 2013;90:99-105. doi 10.1016/j.compag.2012.09.007
- Otsu N. A threshold selection method from gray-level histograms. In: IEEE Transactions on Systems, Man, and Cybernetics. 1979;9(1): 62-66. doi 10.1109/TSMC.1979.4310076
- Qin Y., Wang W., Liu W., Yuan N. Extended-maxima transform watershed segmentation algorithm for touching corn kernels. *Adv Mech Eng*. 2013;5:268046. doi 10.1155/2013/268046
- Roerdink J.B.T.M., Meijster A. The watershed transform: definitions, algorithms and parallelization strategies. *Fundam Inform*. 2000; 41(2):187-228
- Tan S., Ma X., Mai Z., Qi L., Wang Y. Segmentation and counting algorithm for touching hybrid rice grains. *Comput Electron Agric*. 2019;162:493-504. doi 10.1016/j.compag.2019.04.030
- Tanabata T., Shibaya T., Hori K., Ebana K., Yano M. SmartGrain: high-throughput phenotyping software for measuring seed shape through image analysis. *Plant Physiol*. 2012;160(4):1871-1880. doi 10.1104/pp.112.205120
- Wang W., Paliwal J. Separation and identification of touching kernels and dockage components in digital images. *Can Biosyst Eng*. 2006;48:7
- Whan A.P., Smith A.B., Cavanagh C.R., Ral J.P.F., Shaw L.M., Howitt C.A., Bischof L. GrainScan: a low cost, fast method for grain size and colour measurements. *Plant Methods*. 2014;10:23. doi 10.1186/1746-4811-10-23
- Yang S., Zheng L., He P., Wu T., Sun S., Wang M. High-throughput soybean seeds phenotyping with convolutional neural networks and transfer learning. *Plant Methods*. 2021;17(1):50. doi 10.1186/s13007-021-00749-y
- Zhang J., Liu S., Wu W., Zhong X., Liu T. Research on a rapid identification method for counting universal grain crops. *PLoS One*. 2022;17(9):e0273785. doi 10.1371/journal.pone.0273785
- Zhang X., Deng Z., Wang Y., Li J., Tian J. Unconditional and conditional QTL analysis of kernel weight related traits in wheat (*Triticum aestivum* L.) in multiple genetic backgrounds. *Genetica*. 2014; 142(4):371-379. doi 10.1007/s10709-014-9781-6

**Конфликт интересов.** Авторы заявляют об отсутствии конфликта интересов.

Поступила в редакцию 28.12.2024. После доработки 22.01.2025. Принята к публикации 06.02.2025.