

РАСПОЗНАВАНИЕ ПОЛА ДРОЗОФИЛЫ ПО ФОТОИЗОБРАЖЕНИЯМ

Л.М. Местецкий¹, Д.С. Гуру², Х.С. Нагендрасвами², Чаннабасава Чола²

¹ Московский государственный университет, Москва

² Университет Майсура, Карнатака, Индия

mestlm@mail.ru, dsg@compsci.unimysore.ac.in,

hsnswamy@compsci.uni-mysore.ac.in, channuchola@gmail.com

Аннотация. Рассматривается задача распознавания пола дрозофилы по фотоснимкам, полученным с помощью микроскопа. Необходимость в автоматизации решения этой задачи возникает в генетических исследованиях при массовой обработке большого числа изображений. Сложность алгоритмической классификации определяется большим разнообразием условий съёмки, освещенности, ракурса, а также произвольным положением дрозофилы в кадре. Предлагается метод решения задачи на основе анализа формы брюшка дрозофилы с использованием непрерывной морфологической модели. Модель состоит из многоугольной фигуры, аппроксимирующей тело дрозофилы, и скелета этой многоугольной фигуры. Скелет имеет вид геометрического графа, образованного множеством точек – центров вписанных в фигуру окружностей.

Предлагаемый метод включает следующие этапы:

1. Сегментация дрозофилы на изображении. Исходное цветное изображение преобразуется в бинарную картинку, на которой черным цветом выделено тело мушки.
2. Построение непрерывной морфологической модели полученного бинарного изображения. Строится многоугольная фигура, аппроксимирующая с высокой точностью тело дрозофилы по бинарному изображению. Для полученной многоугольной фигуры строится скелет методами вычислительной геометрии.
3. Определение частей тела дрозофилы и осевой линии брюшка. Эта задача решается на основе анализа скелетного графа и множества инцидентных вписанных окружностей.
4. Определение границы брюшка. В качестве границы принимается огибающая семейства окружностей с центрами на осевой линии брюшка.
5. Оценка заостренности хвостового конца брюшка дрозофилы. Заостренность описывается оценкой кривизны линии границы брюшка.

Сравнение показателя заостренности с порогом составляет суть решающего правила для классификации пола дрозофилы.

Предложенный метод реализован и прошёл экспериментальную проверку в вычислительных экспериментах на базе из 100 изображений. Доля правильного определения пола дрозофил составила 98 %.

Ключевые слова: дрозофила, форма тела, морфологическая модель, срединная ось, распознавание пола.

Задача определения пола дрозофилы по фотоснимкам, полученным с помощью микроскопа, возникает в различных генетических исследованиях [1, 2]. Опытный исследователь легко решает эту задачу классификации, руководствуясь несколькими наблюдаемыми признаками. Необходимость в разработке алгоритма для автоматизации решения задачи распознавания возникает при массовом анализе большого числа изображений.

Дрозофила имеет ярко выраженный половой диморфизм, самки и самцы морфологически отличаются друг от друга по ряду признаков [1]. Самки несколько крупнее самцов, однако, размеры мушки во многом зависят от условий питания, особенно на личиночной стадии. Брюшко у самки немного округлое с заостренным концом, у самца оно цилиндрическое с притупленным концом (рис.1). Эти признаки хорошо различимы на изображениях при визуальном анализе. Другие морфологические признаки связаны с пигментацией и количеством хитиновых

щетиннок груди (тергитов) и хитиновых пластинок с брюшной стороны (стернитов). Имеют место также некоторые различия в пигментации самцов и самок.

При разработке алгоритма классификации пола дрозофил по их фотографиям следует учитывать, что перечисленные признаки не всегда хорошо наблюдаются. Так, истинные размеры и особенности пигментации дрозофил по фотографии определяются неточно, они зависят от положения мушек в кадре, дальности съёмки и освещенности. Точный подсчет тергитов и стернитов по фотографии может быть затруднителен из-за недостаточного разрешения и окклюзий. Наиболее надёжным признаком, хорошо наблюдаемым в большинстве случаев, является форма брюшка.

Предлагаемый алгоритм основан на анализе формы брюшка дрозофил по имеющимся фотоизображениям. Идея алгоритма состоит в том, чтобы отследить форму брюшка и определить, является ли его концевая часть заострённой или округлой. Сложность решения этой задачи определяется несколькими

ми факторами. Во-первых, брюшко на изображении не имеет чётко очерченной границы, что хорошо видно при увеличении (рис.2). Во-вторых, брюшко покрыто хитиновыми щетинками, имеющими вид волосков. Оба этих фактора существенно искажают изображение края брюшка. В-третьих, положение,

ориентация и ракурс съёмки варьируются в широких пределах. Поэтому даже простое определение края брюшка и его концевой точки может вызвать затруднение. В конечном счёте, эти факторы существенно усложняют точную сегментацию и классификацию формы брюшка дрозофилы на изображении.



Рис. 1. Примеры фотографий дрозофил под микроскопом. Верхний ряд – самки, нижний ряд – самцы.

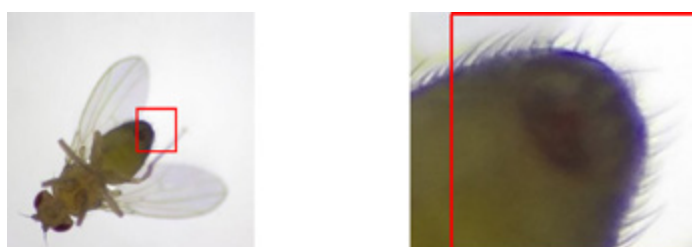


Рис. 2. Увеличенное изображение концевой части брюшка дрозофилы.

Предлагаемое решение основано на построении непрерывной морфологической модели изображения [3]. Непрерывная морфологическая модель позволяет построить кривую, описывающую с хорошей точностью форму концевой части брюшка дрозофилы и на основе геометрических свойств этой кривой оценить, является ли край округлым или заострённым.

Алгоритм включает в себя следующие шаги.

1. Сегментация дрозофилы на изображении

Для сегментации выбирается синяя компонента RGB-изображения. Полученная гистограмма яркости имеет выраженный локальный минимум, который выбирается в качестве порога для бинаризации. Полученное в результате бинарное изображение выделяет тело мушки, отсекая фон и крылья. Резуль-

татом сегментации является бинарное изображение, на котором черным цветом на белом фоне изображен силуэт тела мушки без крыльев (рис. 3).

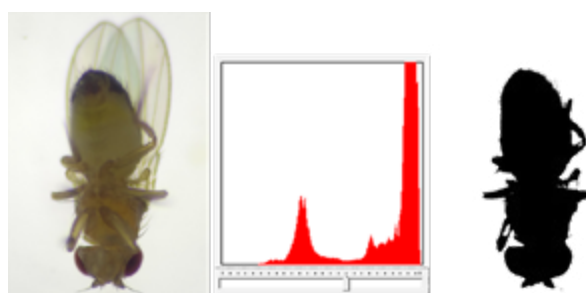


Рис. 3. Сегментация тела дрозофилы: исходное фото, гистограмма яркости синего цвета, бинарное изображение.

2. Построение непрерывной морфологической модели бинарного изображения

Непрерывная морфологическая модель бинарного изображения состоит из многоугольной фигуры, аппроксимирующей бинарное изображение, множества вписанных в фигуру окружностей и множества срединных осей, называемого скелетом этой фигуры (рис. 4). Скелет представляет собой множество точек фигуры, являющихся центрами вписанных в фигуру пустых окружностей. Окружность называется пустой, если все её внутренние точки являются внутренними точками фигуры. Пустая окружность называется вписанной, если она не содержится целиком внутри другой пустой окружности. Построение аппроксимирующей многоугольной фигуры и её скелета осуществляется с помощью методов, описанных в [3].

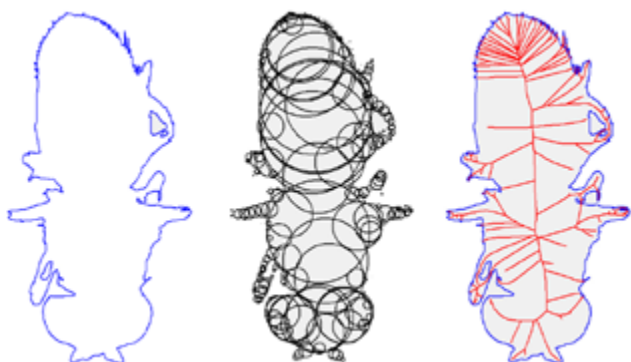


Рис. 4. Непрерывная морфологическая модель силуэта дрозофилы: аппроксимация границы многоугольной фигурой, вписанные пустые окружности, скелет.

3. Определение частей тела дрозофилы и осевой линии брюшка

Части тела дрозофилы – голову, грудь и брюшко можно выделить на основе анализа размеров вписанных окружностей. Скелет можно рассматривать как геометрический граф, каждая точка которого имеет вес, равный радиусу вписанного круга с центром в этой точке. Этот вес называется радиальной функцией скелета. Очевидно, что радиальная функция имеет три локальных максимума в силуэтах головы, груди и брюшка дрозофилы соответственно. Вписанные окружности с центрами в точках локального максимума позволяют локализовать эти части тела (рис. 5).

Голове соответствует меньший из трёх локальных максимумов. Это даёт возможность определить, где голова, а где брюшко, а также установить направление вектора из центра груди в центр брюшка. С помощью этого вектора легко выделяется ветвь скелета,

являющаяся осевой линией брюшка. На рис. 5 вектор показан жёлтым, а ось – красным цветом.

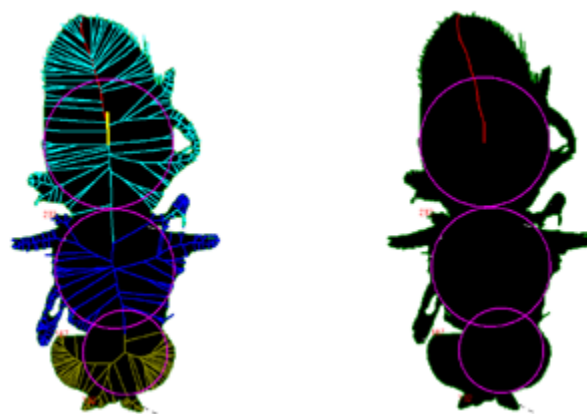


Рис. 5. Сегментация тела дрозофилы, выделение головы, груди и брюшка.

4. Определение границы брюшка

Линия границы брюшка находится на основе сглаживания края изображения, однако сделать это непосредственно путем морфологических операций дилатации и эрозии не представляется возможным. В предложенном алгоритме эта задача решается путём выделения всех вписанных окружностей, имеющих центры на оси брюшка, и вычисления огибающей для этого семейства окружностей (рис.6).



Рис. 6. Определение граничной линии брюшка на основе стрижки скелета и построения огибающей семейства вписанных окружностей.

Представленные на рисунке изображения иллюстрируют последовательные шаги алгоритма: бинаризация, непрерывная морфологическая модель (верхний ряд), выделение вписанных окружностей на оси брюшка и силуэт, образованный огибающей этого семейства окружностей (нижний ряд).

5. Оценка заостренности хвостового конца брюшка дрозофилы

На основе полученной линии границы строятся диаграммы ширины брюшка относительно оси (рис. 7). По оси абсцисс откладывается расстояние от концевой точки брюшка до точки на оси брюшка,

а по оси ординат – радиус вписанной окружности с центром в этой точке. Полученная кривая изображает границу брюшка в фиксированной системе координат. Заострённость края можно оценить на

основе угла касательной к этой кривой в начале координат. Оценка угла осуществляется численно путем аппроксимации полученной кривой полиномом 3 порядка.

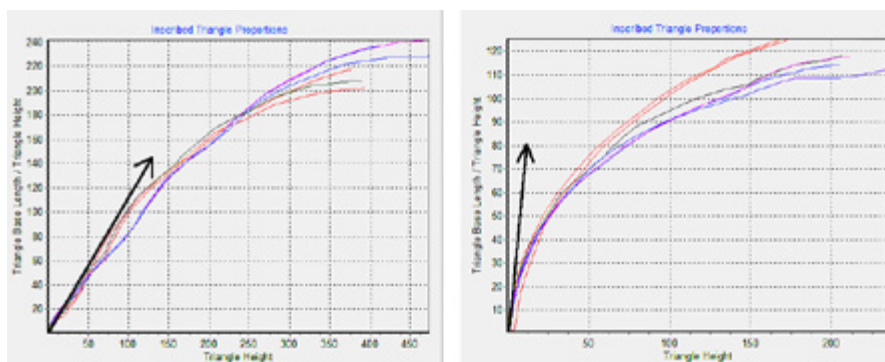


Рис. 7. Определение заостренности граничной линии брюшка.

Левая диаграмма – самки, правая – самцы.

В результате выполнения перечисленных действий для исходного изображения дрозофилы получается угловой коэффициент касательной к оси абсцисс. На рис.7 на левой диаграмме показаны типичные линии огибающей брюшка у самок, на правой – у самцов. Из диаграмм видно, что угловой коэффициент касательной у самок существенно меньше, чем у самцов. Для построения решающего правила нужно выбрать пороговое значение углового коэффициента. Выбор порога был выполнен на основе экспериментов с базой из 100 образцов фотоизображений 50 самок и 50 самцов. Полученное

значение порога углового коэффициента 1.7.

Построенный алгоритм показал высокий уровень качества классификации. Количество правильных ответов в экспериментах получилось 98 %. При этом скорость работы алгоритма составляет 0.4 сек. на одно изображение, что позволяет обрабатывать большие массивы изображений на обычном настольном компьютере.

Работа выполнена при поддержке РФФИ и Департамента науки и технологий правительства Индии гранты 16-57-45054 ИНД_а и 17-01-00917.

Литература

1. <https://www.wikihow.com/Distinguish-Between-Male-and-Female-Fruit-Flies>
2. Козак М.Ф. Дрозофила – модельный объект генетики: учебно-методическое пособие. Астрахань: Издательский дом «Астраханский университет», 2007.
3. Местецкий Л.М. Непрерывная морфология бинарных изображений. Москва, Физматлит, 2009.

GENDER IDENTIFICATION OF FRUIT FLIES BY MICROSCOPIC IMAGES

L.M. Mestetskiy¹, D.S. Guru², H.S. Nagendraswamy², Channabasava Chola²

¹ Moscow State University, Moscow

² University of Mysore, Karnataka, India

mestlm@mail.ru, dsg@compsci.unimysore.ac.in,

hsnswamy@compsci.uni-mysore.ac.in, channuchola@gmail.com

Abstract. The problem of *Drosophila* gender recognition from photographs taken with a microscope is considered. The need for automation to solve this problem arises in genetic research during mass processing of a large number of images. The complexity of the algorithmic classification is determined by the wide variety of shooting conditions, lighting, angle, as well as the arbitrary position of the *Drosophila* in the frame. A method for solving the problem is proposed based on the analysis of the shape of the *Drosophila* abdomen using a continuous morphological model. The model consists of a polygonal figure approximating the body of a *Drosophila* and a skeleton of this polygonal figure. The skeleton has the form of a geometric graph formed by the set of centers of the circles inscribed in the figure.

The proposed method includes the following steps:

1. *Drosophila* segmentation in the image. The original color image is converted to a binary image, in which the body of the fly is highlighted in black.
2. Construction of a continuous morphological model of the obtained binary image. A polygonal figure is constructed that approximates with high accuracy the body of a *Drosophila* in a binary image. For the obtained polygonal figure, a skeleton is constructed by methods of computational geometry.
3. Determination of *Drosophila* body parts and the abdominal centerline. This problem is solved on the basis of the analysis of the skeletal graph and the set of incident inscribed circles.
4. Determination of the border of the abdomen. The boundary is the envelope of a family of circles with centers on the centerline of the abdomen.
5. Assessment of the pointedness of the caudal end of the *Drosophila* abdomen. The sharpness is described by the assessment of the curvature of the abdominal border line.

Comparison of the pointedness indicator with the threshold is the essence of the decisive rule for the classification of *Drosophila* sex.

The proposed method is implemented and passed an experimental test in computational experiments based on 100 images. The proportion of the correct determination of the sex of *Drosophila* was 98 %.

Keywords: *Drosophila*, body shape, morphological model, median axis, gender recognition

References

1. <https://www.wikihow.com/Distinguish-Between-Male-and-Female-Fruit-Flies>
2. Kozak M.F. *Drosophila* – a model object of genetics: a teaching aid. Astrakhan: Publishing House Astrakhan University, 2007. (In Russian)
3. Mestetsky L.M. Continuous morphology of binary images. Moscow, Fizmatlit, 2009. (In Russian)