

Контурная глубина плоской фигуры: применение для определения симметрии и классификации форм

Д.С. Белобородов¹

dmitribeloborodov@yandex.ru

¹МГУ им. М.В. Ломоносова, Москва, Россия

В статье предлагается новый подход к выделению признаков формы из плоских фигур и бинарных изображений — контурной глубины силуэта. Эти признаки вычисляются при помощи непрерывного скелета плоской фигуры и позволяют описать произвольный силуэт без отверстий в виде вектора конечной длины. Рассматривается применение контурной глубины для выявления симметрии объектов, а также для классификации листьев растений из набора данных Flavia.

Ключевые слова: распознавание формы, выделение признаков, непрерывный скелет, поиск симметрии.

Contour depth of a flat figure: application for symmetry detection and shape classification

D.S. Beloborodov¹

dmitribeloborodov@yandex.ru

¹Lomonosov MSU, Moscow, Russia

In this paper I propose a novel approach to feature extraction from flat figures and binary images — contour depth of a silhouette. These features are constructed from continuous skeleton of a flat figure, they allow to represent any silhouette without holes as a finite-sized vector. These features are used for symmetry detection, as well as for plant leaves classification in Flavia dataset.

Keywords: shape recognition, feature extraction, continuous skeleton, symmetry detection.

1. Введение

Задача распознавания формы является одной из важных задач в области компьютерного зрения. Как правило, эта задача предполагает определение класса различных форм или силуэтов, а также извлечение другой полезной информации. Роль исходных данных играют силуэты объектов, чаще всего представленные бинарными изображениями.

Вначале фигуру необходимо представить в виде, допускающем некоторый способ сравнения с другими фигурами, например, выделить из нее числовые признаки. Для задач анализа формы к этим признакам обычно предъявляются требования инвариантности относительно поворотов и сдвигов, а иногда и относительно деформаций. Наиболее распространенный подход заключается в том, чтобы вычислить некоторые статистики с помощью самой фигуры или ее скелета. Такие признаки инвариантны к поворотам и сдвигам, однако они не учитывают важную информацию о структуре фигуры. Существующие подходы можно разделить на три группы: основанные на данных контура, основанные на данных скелета, а также использующие одновременно контур и скелет.

Подходы, использующие контур, часто применяют алгоритмы выравнивания к последовательностям контура, также существуют алгоритмы на основе преобразования (изгиба) контуров, либо подсчета каких-либо статистик контура. К данным работам относятся [5, 6, 9, 10, 12]. Во втором подходе чаще всего применяются алгоритмы выравнивания скелетов фигур: [3, 4, 7, 11, 15]. В работах [2, 16] фигуры сравниваются

по морфологическим спектрам (распределение ширины в фигуре). В [13, 14] для подсчета статистических признаков фигуры используются и скелет, и граница. В работе [8] производится выравнивание последовательностей, построенных при обходе скелета и учитывающих также информацию из контура. Предложенный в данной работе метод можно также отнести к последней группе.

Предлагается новый дескриптор плоской фигуры с гладкой границей, «функция контурной глубины», имеющая смысл распределения площади фигуры по ее границе. Этот дескриптор обобщается на случай многоугольной фигуры, граница которой уже не является гладкой. Предлагается способ выделения из силуэта вектора признаков конечной длины, «вектора контурной глубины», при помощи данного дескриптора. С помощью этого вектора можно, например, устанавливать симметрию объектов, а также сравнивать их напрямую в некоторых случаях. В следующих разделах будет подробно описан алгоритм вычисления вектора контурной глубины.

2. Функция контурной глубины

Рассмотрим плоскую фигуру с гладкой границей, не имеющую отверстий. Тогда для каждой точки границы мы можем рассмотреть множество точек внутри фигуры, для которых данная точка границы будет ближайшей. Введем отображение $S(Y)$: любому подмножеству Y границы фигуры оно ставит в соответствие множество точек внутри фигуры, для которых расстояние до множества Y меньше, чем до осталь-

ных точек границы (локус). Можно выбрать на границе фигуры произвольную точку отсчета и параметризовать границу фигуры некоторой переменной $x \in [0, L)$, где L — длина границы фигуры. Пусть $\partial[x_1, x_2]$ — участок границы, параметризованный значениями $x \in [x_1, x_2]$. Тогда мы можем рассмотреть функцию $G(x) = |S(\partial[0, x])|$, $x \in [0, L)$ — она имеет смысл площади фигуры, заметаемой при проходе вдоль границы от начальной точки до x . Если у этой функции существует производная $G'(x) = \tilde{F}(x)$ (для фигур с гладкой границей она будет существовать), будем называть ее функцией контурной глубины. Эта функция имеет смысл распределения площади фигуры относительно ее границы. Поскольку $x = 0$ и $x = L$ параметризуют одну точку на контуре, а $\tilde{F}(x)$ задает прирост площади в данной точке, можно доопределить функцию в L : $\tilde{F}(0) = \tilde{F}(L)$. Из тех же соображений можно считать функцию $\tilde{F}(x)$ периодической с периодом L (что имеет смысл при обходе контура более одного раза, а также обходе в обратном направлении). Стоит заметить, что некоторые точки внутри фигуры являются ближайшими для нескольких точек границы, и эти точки можно не учитывать при подсчете $G(x)$: множество всех таких точек является множеством центров максимальных вписанных кругов (скелетом фигуры) и имеет нулевую площадь. Поэтому $G(L) = \int_0^L \tilde{F}(x) dx = S$.

При смене точки начала отсчета на границе сама функция $\tilde{F}(x)$ принципиально меняться не будет, произойдет циклический сдвиг: $\tilde{F}(y) = \tilde{F}(x - l)$, где x, y — параметризации границы фигуры с точками начала отсчета x_0, y_0 , а l — длина участка границы, заключенного между ними (при обходе от x_0 к y_0). Таким образом, данная функция обладает некоторой инвариантностью к поворотам фигуры: при смене точки начала обхода она сдвигается циклически. Также очевидна инвариантность относительно перемещения фигуры на плоскости.

Функция $F(x) = \tilde{F}(x\sqrt{S})/\sqrt{S}$, $x \in [0, L/\sqrt{S}]$ должна быть инвариантна относительно масштаба фигуры, это функция контурной глубины аналогичной фигуры с единичной площадью. Функцию $F(x)$ будем называть нормированной контурной глубиной.

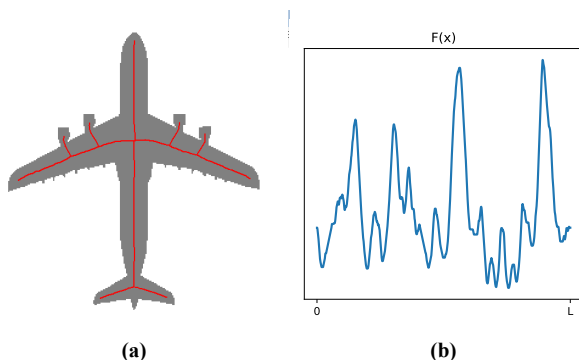


Рис. 1. Плоская фигура и ее скелет (a) и функция контурной глубины (b)

Пример функции контурной глубины можно видеть на рисунке 1.

3. Контурная глубина многоугольных фигур

Для многоугольной фигуры определить функцию контурной глубины будет сложнее: вершины при вогнутых углах обладают множествами ближайших точек ненулевой площади, из-за чего функция $G(x)$ будет разрывной. В это же время вершины при выпуклых углах будут всегда вносить нулевой вклад в площадь. Таким образом, любые углы в контуре будут очень сильно влиять на значения функции.

Очевидным решением будет как-либо сгладить границы фигуры, и в таком случае будет гораздо удобнее рассматривать $F(x)$ как распределение площади по границе, и саму эту функцию подсчитывать с помощью непрерывного скелета фигуры и площадей ячеек Вороного.

Для описания алгоритма построения необходимо использовать уже введенное понятие непрерывного скелета: непрерывным скелетом плоской фигуры называется множество центров максимальных кругов, вписанных в данную фигуру. Для многоугольной фигуры ее скелет может быть представлен в виде плоского геометрического графа, узлы которого являются точками на плоскости, а ребра — отрезками прямых или парабол. При этом каждая точка на графе соответствует некоторому вписанному в эту фигуру максимальному кругу. Чтобы построить непрерывный скелет бинарного изображения, необходимо сначала приблизить его границы некоторой многоугольной фигурой, для которой впоследствии и строится скелет.

Существуют эффективные алгоритмы построения непрерывного скелета для бинарных изображений, они подробно описаны в [1].

Понятие непрерывного скелета тесно связано с понятием диаграммы Вороного: граф скелета является подмножеством границ ячеек Вороного. Диаграммой Вороного для некоторого набора элементов (сайтов) называется разбиение плоскости на области (ячейки Вороного), каждая из которых состоит из точек, ближайших к конкретному сайту. В терминах многоугольных фигур сайтами являются вершины и стороны фигуры. При этом каждая ячейка Вороного сопоставлена некоторому сайту, а каждый участок границы между ячейками Вороного представляет собой линию, в каждой точке которой расстояния до некоторых двух сайтов равны. Если сайты — это две точки или два отрезка прямых, разделяющая кривая будет отрезком прямой, если же эти два сайта являются точкой и отрезком, то кривая будет параболической. Несложно заметить, что непрерывный скелет фигуры будет подмножеством границ ячеек Вороного (все, лежащие внутри фигуры, за исключением границ между ребрами, соединяющимися под углом более 180°). Зная скелет, можно также восстановить границы ячеек Вороного.

Таким образом, при построении диаграммы Вороного внутренняя часть фигуры разбивается на ячейки,

соответствующие точкам и сегментам границы. Границы этих ячеек будут состоять из отрезков прямых и парабол. При этом, поскольку сайты границы неизбежно соприкасаются со своими соседями, сам сайт также обязательно будет частью границы ячейки. Зная непрерывный скелет фигуры, можно вычислить площади всех ячеек и установить их принадлежность к сайтам границы. Другими словами, мы знаем некоторый закон распределения площади фигуры по ее границе.

Вспомним, что на границе фигуры есть два типа вершин: выпуклые (внутренний угол при вершине не более 180°) и вогнутые (внутренний угол при вершине больше 180°). Ячейка Вороного для выпуклой вершины внутри фигуры всегда имеет нулевую площадь, ячейка для вогнутой вершины имеет ненулевую площадь и содержит саму эту вершину как часть своей границы (но не содержит никаких других точек из границы фигуры). В этом случае получается, что ячейка с ненулевой площадью отвечает участку границы нулевой длины.

Пусть для некоторой многоугольной фигуры известна диаграмма Вороного. Тогда можно составить список последовательных сегментов границы и площадей ячеек Вороного, соответствующих им. Такой же список можно составить для вершин границы и площадей их ячеек (для выпуклых вершин эти площади нулевые). Интуитивно это можно представить как разрезание фигуры на «дольки» и разворачивание в линию. Параметр $x \in [0, L]$ параметризует линию границы начиная с выбранной точки, где L — длина границы фигуры.

Будем считать, что площадь ячеек вершин не сконцентрирована в одной точке границы, а разделяется соседними точками границы некоторым образом. Пусть x_i — координата данной вершины на параметризованной линии границы, а S_i — площадь ячейки. Положим, что площадь ячейки распределена по точкам границы следующим образом: основная часть сконцентрирована в x_i , а при удалении от нее доля площади, приходящаяся на каждую точку, убывает линейно. Это можно формализовать в виде симметричной треугольной положительной функции с центром в x_i , определенной на $[0, L]$, интеграл от которой равен S_i :

$$f_i(x) = \begin{cases} k(x - x_i) + \sqrt{kS_i}, & x \in \left[x_i - \sqrt{\frac{S_i}{k}}, x_i \right] \\ k(x_i - x) + \sqrt{kS_i}, & x \in \left[x_i, x_i + \sqrt{\frac{S_i}{k}} \right] \\ 0, & \text{else} \end{cases}$$

где $k > 0$ — параметр размытости площади вокруг точки (наклон сторон треугольника), одинаковый для всех сайтов фигуры.

Теперь построим аналогичную функцию для сегментов границы. Вместо треугольника используем трапецию с тем же наклоном сторон k : это значит, что площадь ячейки равномерно распределена по ребру в центре, а по краям размывается. Пусть x_j^b, x_j^e — координаты

начала и конца сегмента, $x_j^c = \frac{x_j^b + x_j^e}{2}$ — координата центра ребра. Также S_j — площадь ребра, l_j — его длина. Если $l_j \geq \frac{S_j}{kl_j}$ (ребро достаточно длинное, чтобы построить трапецию), зададим распределение площади следующей функцией:

$$g_j(x) = \begin{cases} k(x - x_j^b) + \frac{S_j}{2l_j}, & x \in \left[x_j^b - \frac{S_j}{2kl_j}, x_j^b + \frac{S_j}{2kl_j} \right] \\ \frac{S_j}{l_j}, & x \in \left[x_j^b + \frac{S_j}{2kl_j}, x_j^e - \frac{S_j}{2kl_j} \right] \\ k(x_j^e - x) + \frac{S_j}{2l_j}, & x \in \left[x_j^e - \frac{S_j}{2kl_j}, x_j^e + \frac{S_j}{2kl_j} \right] \\ 0, & \text{else} \end{cases}$$

где $k > 0$ — тот же параметр размытости.

Заметим, что площадь этой трапеции равна S_j , центры наклонных сторон имеют координаты x_j^b, x_j^e , а центр трапеции находится в x_j^c .

Если же условие $l_j \geq \frac{S_j}{kl_j}$ не выполняется, вместо трапеции используется треугольник:

$$g_j(x) = \begin{cases} k(x - x_j^c) + \sqrt{kS_j}, & x \in \left[x_j^c - \sqrt{\frac{S_j}{k}}, x_j^c \right] \\ k(x_j^c - x) + \sqrt{kS_j}, & x \in \left[x_j^c, x_j^c + \sqrt{\frac{S_j}{k}} \right] \\ 0, & \text{else} \end{cases}$$

Такой треугольник центрирован относительно x_j^c и имеет площадь S_j .

Эти формулы выписаны для случая, когда сайт границы находится достаточно далеко от начальной точки. Для сайтов, находящихся рядом с точкой старта, эти функции должны циклически продолжаться между концом и началом отрезка $[0, L]$ (поскольку граница является циклическим объектом, и любая точка может быть начальной).

Теперь функция контурной глубины может быть выражена через сумму всех полученных функций:

$$\tilde{F}(x) = \sum_i f_i(x) + \sum_j g_j(x),$$

суммирование ведется по всем вершинам и сегментам границы соответственно.

Ее можно нормировать: $F(x) = \tilde{F}(x\sqrt{S})/\sqrt{S}$, в терминах введенных выше функций это означает, что значения $x_i, x_j^b, x_j^e, x_j^c, l_i$ необходимо масштабировать на значение $1/\sqrt{S}$, а S_i — на значение $1/S$.

4. Вектор контурной глубины

Полученная нормированная функция контурной глубины является информативным дескриптором формы, однако саму эту функцию сложно использовать в задачах распознавания форм. Вместо этого предлагается использовать вектор фиксированной длины n . Для этого достаточно просто взять n точек t_n , расположенных равномерно на отрезке $[0, L/\sqrt{S}]$ и посчитать в них функцию $F(x)$. При этом $\sum_n F(t_n) \frac{L}{n} \approx 1$. Параметр n регулирует соотношение между информативностью контурной глубины и длиной вектора (сложностью вычислений). Также, поскольку длины контуров

нормированных фигур могут быть разными, правильнее будет зафиксировать шаг вычисления d_i , на который отстоят значения t_i и t_{i+1} — в этом случае для разных фигур длины векторов будут отличаться.

Полученные признаки инвариантны относительно сдвигов и масштабирования изображения, а также очевидно меняются при повороте: можно считать, что при этом просто происходит циклический сдвиг вектора. Они также устойчивы к упругим деформациям объектов. Однако, увеличение длины контура за счет появления дополнительных элементов может привести к заметному изменению признаков.

Параметр k отвечает за степень размытости площади по границе: чем больше k , тем больше наклон у сторон треугольников и трапеций, тем более детальной получается функция F . Наоборот, маленькие значения k позволяют сгладить эту функцию. Параметр n отвечает за длину вектора, в который преобразуется фигура, и также влияет на детализацию. Эти параметры связаны: при маленьких значениях k параметр n также можно уменьшить.

Теперь, если для контуров двух фигур фиксированы начальные точки, для них можно вычислить вектора контурной глубины a, b , которые можно сравнивать напрямую. Поскольку эти вектора отражают распределение площади по границе, для их сравнения можно использовать функции близости для множеств, например аналог коэффициента Жаккара: $J(a, b) = \frac{\sum_i \min(a_i, b_i)}{\sum_i \max(a_i, b_i)}$. Это значение имеет смысл отношения пересечения к объединению фигур (поскольку вычисленные признаки представляют площади).

Другим важным вопросом является поиск правильных начальных точек, с которых можно начинать сравнение. Помимо очевидного способа перебрать все n точек в каждом контуре или их подмножество, в некоторых случаях можно использовать специфичную структуру объектов, например, их симметрию.

5. Определение симметрии

Вектор контурной глубины можно использовать для простого поиска симметрии: для этого достаточно перебрать все циклические сдвиги вектора и найти сдвиг, максимизирующий коэффициент Жаккара для первой половины вектора и второй половины в обратном порядке, это займет $O(n^2)$. Точка сдвига определяет одну из точек на контуре, через которую проходит ось симметрии. В качестве второй точки оси симметрии берется точка, отстоящая от первой на $n/2$. По индексам компонент вектора можно восстановить и саму ось симметрии.

При нарушении симметрии фигуры (наличии дефектов контура) такой наивный алгоритм приведет к неточному результату. Для повышения точности можно воспользоваться описанной далее процедурой. Сначала необходимо циклически сдвинуть вектор на значение, найденное наивным алгоритмом. Затем производится симметрическое выравнивание вектора: выравнивание самого вектора контурной глубины с этим

же вектором, записанным в обратном порядке, с помощью Алгоритма Нидлмана-Вунша выравнивания последовательностей. (Этот алгоритм изначально использовался для выравнивания символьных последовательностей с табличными функциями стоимости, однако он легко обобщается и на вещественные последовательности. В данной работе используются функции стоимости $f(a, b) = \min\{a, b\} / \max\{a, b, \epsilon\}$, $g(a) = -0.1$, где ϵ имеет маленькое значение для численной устойчивости). Тогда вторая точка оси симметрии будет задаваться элементом вектора, который был выравнен с самим собой (либо с соседним элементом). Этот процесс изображен на рисунке 2.

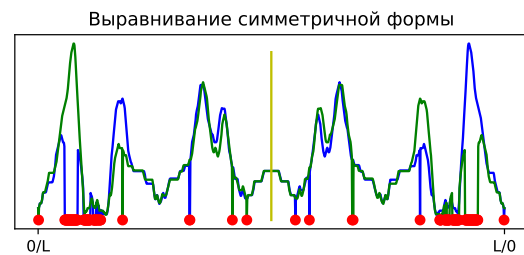


Рис. 2. Симметрическое выравнивание вектора. Красные точки — пропуски выравнивания, желтая линия — искомая точка оси симметрии.

Поскольку наивный алгоритм мог неверно определить первую точку оси симметрии, этот процесс можно повторить, используя в качестве первой точки, найденную с помощью выравнивания. Результат выравнивания можно видеть на примере сложного контура на рисунке 3: в качестве контура выступает изображение самолета, один из двигателей которого соединился с корпусом из-за ошибки бинаризации. Это сильно влияет на симметрию (контур делится осью симметрии на неравные части), однако с помощью выравнивания можно точно восстановить ось симметрии. Сложность такого подхода составляет $O(n^2)$, где n — число компонент вектора контурной глубины.



Рис. 3. Оси симметрии наивного метода (a) и с использованием выравнивания (b).

6. Классификация

Векторы контурной глубины были также применены в задаче классификации листьев растений в наборе данных Flavia [17], который содержит 1907 изображений листьев растений 32 разных видов.

Предварительно изображения были бинаризованы по пороговому значению. Для подсчета векторов контурной глубины использовалось значение $k = 0.1$ и шаг пересчета компонент вектора $d_t = 0.015$. Каждый вектор циклически сдвигался таким образом, чтобы наименьшая компонента (отвечающая одному из заостренных участков листа) оказалась в начале вектора. Также из каждого вектора вычиталось его минимальное значение — это позволяет перенести акцент с формы в целом на более тонкие детали. Затем для каждого двух векторов a, b считался аналог меры сходства Жаккара: $J_s = J(a, b) = \frac{\sum_i \min(a_i, b_i)}{\sum_i \max(a_i, b_i)}$. При этом пары векторов выравнивались по центру, а кратчайший вектор дополнялся нулями с двух сторон. Поскольку данные вектора оценивают только форму листа в целом, дополнительно была использована информация о цвете листьев: для этого были использованы гистограммы цвета (26 ячеек на каждый канал), для которых также использовалась мера сходства $J_c = J(u, v)$, где u, v — распределения цветов по ячейкам гистограммы. Для классификации был использован метод ближайших соседей с 10 соседями и мерой сходства $0.85J_s + 0.15J_c$. Полученная точность (ассигасу) составила 0.866.

7. Заключение

Предложенные в данной статье признаки контурной глубины описывают распределение площади фигуры по ее границе. Эти признаки легко интерпретируемы, а гиперпараметры позволяют подбирать как детализацию, так и размер вектора. С помощью этих векторов можно сравнивать формы фигур, а также точно находить ось симметрии даже при нарушении контура.

8. Благодарности

Данная работа выполнена при поддержке гранта РФФИ 17-01-00917.

9. Литература

- [1] Местецкий, Л. М., 2009. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры. Москва, Физматлит.
- [2] Местецкий, Л. М., 2014. Медиальная ширина фигуры — дескриптор формы изображений. Машинное обучение и анализ данных, 1(9), 1291-1318.
- [3] Aslan, C., Erdem, A., Erdem, E., Tari, S., 2008. Disconnected skeleton: shape at its absolute scale. IEEE Trans. Pattern Analysis and Machine Intelligence 30, 2188–2203.
- [4] Bai, X., Latecki, L., 2008. Path similarity skeleton graph matching. IEEE Trans. Pattern Analysis and Machine Intelligence 30, 1282–1292.
- [5] Belongie, S., Malik, J., Puzicha, J., 2002. Shape matching and object recognition using shape contexts. IEEE Trans. Pattern Analysis and Machine Intelligence 24, 509–522.
- [6] Felzenszwalb, P.F., Schwartz, J., 2007. Hierarchical matching of deformable shapes, in: CVPR.
- [7] Klein, P., Tirthapura, S., Sharvit, D., Kimia, B., 2000. A tree-edit-distance algorithm for comparing simple, closed shapes. Proceedings of the eleventh annual ACM-SIAM symposium on Discrete algorithms. Society for Industrial and Applied Mathematics, pp. 696–704.
- [8] Kushnir, O., Seredin, O., 2015. Shape matching based on skeletonization and alignment of primitive chains. International Conference on Analysis of Images, Social Networks and Texts, 123-136, Springer, Cham.
- [9] Ling, H., Jacobs, D.W., 2007. Shape classification using the innerdistance. IEEE Trans. Pattern Analysis and Machine Intelligence 29, 286–299.
- [10] Neuhaus, M., Bunke, H., 2006. Edit distance-based kernel functions for structural pattern classification. Pattern Recognition 39.10, pp. 1852–1863.
- [11] Sebastian, T., Klein, P., Kimia, B., 2004. Recognition of shapes by editing their shock graphs. IEEE Trans. Pattern Analysis and Machine Intelligence 26, 550–571.
- [12] Sederberg, T.W., Greenwood, E., 1992. A Physically Based Approach to 2-D Shape Blending, Comput. Graphics, no. 26 (2), pp. 25–34.
- [13] Shen, W., Wang, X., Yao, C., Bai, X., 2014. Shape Recognition by Combining Contour and Skeleton into a Mid-Level Representation. Pattern Recognition, Springer Berlin Heidelberg, pp. 391–400.
- [14] Shen, W., Jiang, Y., Gao, W., Zeng, D., Wang, X., 2016. Shape recognition by bag of skeleton-associated contour parts. Pattern Recognition Letters, 83, 321–329.
- [15] Siddiqi, K., Shokoufandeh, A., Dickinson, S., Zucker, S., 1999. Shock graphs and shape matching. Int'l J. Computer Vision 35, 13–32.
- [16] Vizilter, Yu.V., Sidyakin, S.V., Rubis, A.Yu., Gorbatshevich, V.S., 2012. Morphological shape comparison based on skeleton representations. Pattern Recognition and Image Analysis 22.3, pp. 412–418.
- [17] Wu, S.G., Bao, F.S., Xu, E.Y., Wang, Y.-X., Chang, Y.-F., Xiang, Q.-L., 2007. A leaf recognition algorithm for plant classification using probabilistic neural network. Signal Processing and Information Technology, 2007 IEEE International Symposium, 11–16.

Об авторах

Белобородов Дмитрий Сергеевич, студент магистратуры кафедры Математических методов прогнозирования факультета Вычислительной математики и кибернетики Московского государственного университета им. М.В. Ломоносова. Его e-mail dmitribeloborodov@yandex.ru.