

СРАВНЕНИЕ ФОРМЫ ИЗОБРАЖЕНИЙ НА ОСНОВЕ СПЕКТРА ШИРИНЫ ОБЪЕКТОВ

Е. Зак, Н. Ломов, Л. Местецкий

Факультет вычислительной математики и кибернетики

Московский государственный университет им. М.В. Ломоносова, Москва, Россия

e_zak@mail.ru, nikita-lomov@mail.ru, mestlm@mail.ru

Аннотация

Работа посвящена разработке и сравнительному анализу методов вычисления так называемых морфологических спектров для решения задач классификации изображений гибких объектов. В качестве базового примера гибкого объекта, у которого форма может изменяться в широких пределах, рассматривается изображение человеческой ладони. Необходимость в сравнении ладоней по форме возникает при решении задач биометрической идентификации личности. Морфологические спектры, как показано в работе, позволяют выявить информативные признаки формы таких объектов и построить метрики для определения их сходства и различия.

Приводится сравнительное исследование трех типов спектров: морфологический спектр (pattern spectrum), предложенный в работе П. Марагоса, медиальный спектр и спектр дистанционной функции формы, предложенные авторами статьи.

Ключевые слова: морфологический спектр, медиальный спектр, спектр дистанционной ширины, анализ формы.

1. ВВЕДЕНИЕ

Ширина объектов на изображениях – важная характеристика их формы. Поэтому естественным является стремление использовать информацию о ширине объектов для распознавания формы в задачах компьютерного зрения. Поскольку ширина объекта различна для разных его частей, она не может быть описана какой-нибудь скалярной величиной, например, «средней» шириной. В качестве дескриптора ширины желательно построить такое описание, которое характеризовало бы весь спектр ее значений. В этой роли может выступить описание диапазона вариации значений ширины в различных частях объекта с указанием «частоты присутствия» каждого значения в данном объекте. Такое интегральное описание ширины удобно представить в виде функции, определяющей меру, например площадь, той части объекта, в которой ширина не превосходит заданного порога. При этом аргументом функции будет величина этого порога. Таким образом, дескриптор ширины приобретает вид, аналогичный функции распределения вероятности, и его вполне уместно назвать функцией распределения ширины объекта.

Для того чтобы полностью формально определить такой дескриптор, нужно сделать следующее:

- ввести понятие ширины объекта в каждой его точке;

- построить описание области «заданной ширины», представляющей собой множество точек объекта, в которых ширина не превосходит заданного порога;

- построить меру для подмножества заданной ширины, как функцию от величины порога.

В данной статье мы рассматриваем три возможных варианта построения функции распределения ширины, различающиеся способами формализации понятия ширины объекта в точке. Все три способа основываются на понятии пустого круга.

В качестве объектов изображения, для которых нужно строить дескриптор ширины, будем использовать понятие фигуры замкнутой ограниченной области на евклидовой плоскости. Пустым кругом называется круг, целиком принадлежащий фигуре. Максимальным пустым кругом или вписанным кругом будем называть пустой круг, который не содержится ни в одном другом пустом круге. Вписанный круг всегда имеет общие точки с границей фигуры. Радиус-вектор, соединяющий центр вписанного круга с точкой, в которой он касается границы, называется спицей.

Мы рассматриваем три варианта определения ширины фигуры в точке: морфологическая, медиальная и дистанционная ширина.

Морфологической шириной фигуры в точке называется радиус наибольшего пустого круга, покрывающего эту точку.

Дистанционной шириной фигуры в точке называется радиус наибольшего пустого круга с центром в этой точке.

Медиальной шириной фигуры в точке называется максимальное значение длины спицы, проходящей через эту точку.

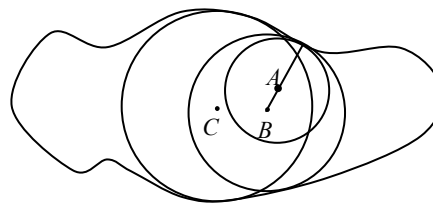


Рис. 1. Пример морфологической, медиальной и дистанционной ширины для точки A фигуры

В примере на рис. 1 для точки A морфологическая ширина равна радиусу круга C, дистанционная ширина – радиусу круга A, медиальная ширина – радиусу круга B.

Для этой же фигуры на рис. 4 серым цветом выделены области заданной ширины, соответствующие этим определениям. Области представляют собой подмножества точек фигуры, в которых значение ширины не превосходит радиуса изображенного на рисунке круга.

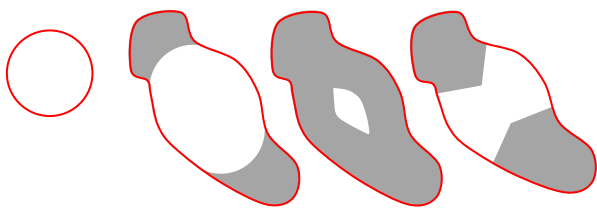


Рис. 2. Области заданной ширины (морфологической, дистанционной, медиальной), соответствующие размеру круга (слева)

2. ОПРЕДЕЛЕНИЯ И ПОДХОДЫ К ВЫЧИСЛЕНИЮ ШИРИНЫ И СПЕКТРА

Понятие морфологического спектра было введено Марагосом в работе [3]. Приведем определение спектра для дискретного случая, то есть когда мы работаем с объектом, заданным на растровой решетке.

Определение 1. Морфологическим спектром PS_X фигуры X относительно примитива (структурного элемента) B называется функция

$$PS_X(r_i, B) = - \frac{S(X \circ B(r_i)) - S(X \circ B(r_{i+1}))}{r_{i+1} - r_i}$$

где $(r_{i+1} - r_i)$ – шаг растровой решетки; $r_i \geq 0$, $X \square B(r)$ – операция морфологического открытия фигуры X с примитивом B ; $S(X)$ – площадь X .

Если в качестве примитива $B(r)$ взять диск радиуса r , то результатом операции открытия $X \square B(r)$ будет подмножество X , покрываемое множеством дисков радиуса r , которые целиком лежат в X , то есть подмножество равно объединению всех дисков, вписанных в X . Следовательно, подмножество остальных точек из X – это те точки, которые покрываются вписанными кругами меньшего радиуса. Таким образом, функция $F(r) = S(X \setminus X \square B(r))$ есть функция распределения морфологической ширины фигуры X в терминологии, введенной в предыдущем разделе. В определении 1 фигурирует разностная производная этой функции.

Другой подход к описанию ширины объекта основывается на понятии срединной оси, или скелета, которые вводятся с помощью непрерывной морфологии [7]. Скелетом фигуры называется множество точек-центров всех вписанных кругов фигуры. Скелет имеет вид плоского геометрического графа. Введем необходимые для понятия медиальной ширины определения.

Определение 2. Скелетом фигуры называется множество точек-центров всех вписанных кругов фигуры.

Определение 3. Радиальная функция определена в точках скелета и ставит в соответствие точке скелета радиус вписанного круга с центром в этой точке.

Определение 4. Спицей называется отрезок прямой, соединяющий точку скелета с ближайшей точкой границы фигуры.

Отметим некоторые важные свойства, связывающие точки фигуры с множеством спиц этой фигуры.

Утверждение 1. Через каждую точку фигуры проходит хотя бы одна спица. Следовательно, спицы покрывают всю фигуру.

Утверждение 2. Через каждую внутреннюю точку фигуры, не являющуюся точкой скелета, проходит только одна спица.

Теперь определим понятие медиальной ширины [8] фигуры в точке.

Определение 5. Медиальная ширина фигуры в точке равна длине минимальной спицы, инцидентной этой точке.

Из утверждения 1 следует, что медиальная ширина определена для всех точек фигуры. Все спицы, инцидентные одной точке скелета, имеют одинаковую длину, равную радиусу пустого круга с центром в этой точке. Поэтому для точек скелета медиальная ширина просто равна радиальной функции. Для каждой внутренней точки, не являющейся точкой скелета, существует, согласно утверждению 2, единственная инцидентная спица, поэтому медиальная ширина в такой точке определена однозначно.

Точки фигуры, имеющие несколько инцидентных спиц разной длины, могут находиться лишь на границе фигуры. Именно для таких случаев в определении 5 медиальная ширина определена через понятие минимальной спицы.

Введем обозначения:

- $\phi(g)$ – медиальная ширина в точке g ;
- $G_r = \{g \in G, \phi(g) \leq r\}$ – подмножество точек фигуры, в которых медиальная ширина не превосходит заданное значение r ;
- $F(r)$ – функция медиальной ширины фигуры, площадь множества точек G_r .

Тогда, под медиальным спектром будем понимать

$$f(r) = dF(r) / dr.$$

Еще одним способом описать фигуру с помощью дескрипторов, основанных на ширине фигуры, является вычисление дистанционной ширины с помощью карты расстояний (Distance transform map, далее по тексту DT) [1, 2]. Карта расстояний (Distance map) – это такое представление изображения фигуры, в котором каждой точке объекта ставится в соответствие минимальное расстояние до точек фона.

Определение 6. Функция построения карты расстояний DT является отображением фигуры X , при котором каждой точке фигуры x ставится в соответствие минимальное расстояние до точек фона, то есть:

$$DT(x) = \min_p (d(x, p) | x \in X, p \notin X).$$

Функция $DT(x)$ описывает дистанционную ширину в точке x .

Таким образом, функция распределения дистанционной ширины строится следующим образом:

$$F(r) = \sum_{x \in X} [DT(x) \leq r].$$

3. РЕАЛИЗАЦИЯ И ЭКСПЕРИМЕНТЫ

Для проверки применимости спектра для решения задач анализа формы объектов и классификации была собрана выборка из 160 фотографий ладоней 36 различных людей.

Выборка состоит из изображений с разрешением 640x480 px. Примеры некоторых объектов из выборки приведены на рис. 3. Как видно из примеров, положение ладони на изображении и другие ограничения не фиксированы.



Рис. 3. Элементы выборки

Все три способа вычисления функции распределения ширины были реализованы в соответствии с данными выше определениями. Сравнение этих методов выполнялось по критериям скорости вычислений и по качеству построенных на их основе алгоритмов классификации.

Скоростная характеристика алгоритма является одним из самых главных критериев, с точки зрения использования его в системах машинного зрения. Для сравнения скоростей вычисления функций распределения ширины были проведены эксперименты для 100 случайно выбранных изображений (табл. 1). Вычисление проводилось на процессоре Intel Core i5 2.67 GHz.

Эксперимент показал, что вычисление морфологической ширины существенно отстает от других видов спектров по времени расчета. Разница между временем вычисления медиальной ширины и дистанционной ширины в 4 раза связана с относительной сложностью операции построения карты глубины, которая сравнима по времени с алгоритмом обхода в глубину графа. Результаты показывают, что время построения функции распределения медиальной ширины позволяет использовать ее в анализе видео последовательностей, так как алгоритм работает достаточно быстро для видео с 24 кадрами в секунду.

Морфологическая ширина (сек.)	Медиальная ширина (сек.)	Дистанционная ширина (сек.)
2.548	0.025	0.093

Таблица 1. Усредненное по 100 вычислениям на CPU время вычисления ширины на изображениях выборки (640x480 px)

Для сравнения классификационной способности в экспериментах был использован метрический алгоритм kNN . Оценка качества классификации (Q) выполнена с помощью

скользящего контроля по схеме leave-one-out как отношение числа верно классифицированных объектов (N_{true}) к общему числу объектов (N):

$$Q = N_{true} / N.$$

Алгоритм k -ближайших соседей имеет 3 параметра:

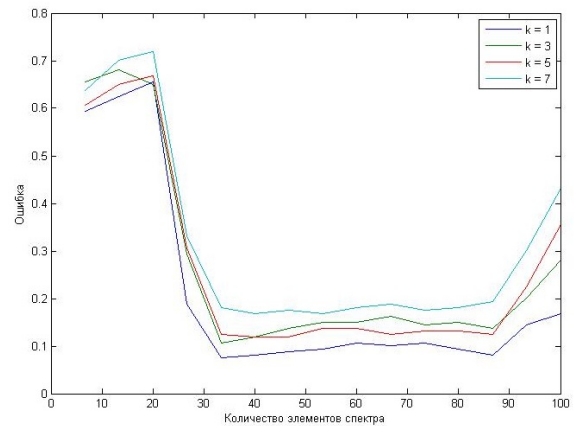
- метрика – $d(H_1, H_2)$;
- количество «соседей» – k ;
- набор элементов спектра – $I = i_{p1}, \dots, i_{pm}$.

Метрика сравнения является одним из основных параметров модели классификации. В исследовании были рассмотрены 4 метрики (табл. 2). На данной выборке лучшее качество показала так называемая транспортная метрика (EMD – метрика).

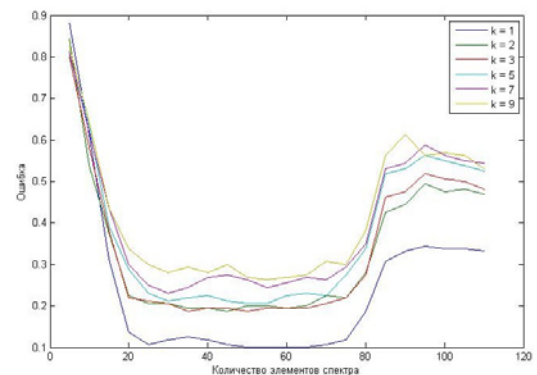
$d(H_1, H_2)$	Q
Simple	0.8313
χ^2	0.8663
Бхаттачария	0.7375
EMD	0.9012

Таблица 2. Сравнение метрик сходства морфологических спектров на выборке по качеству классификации

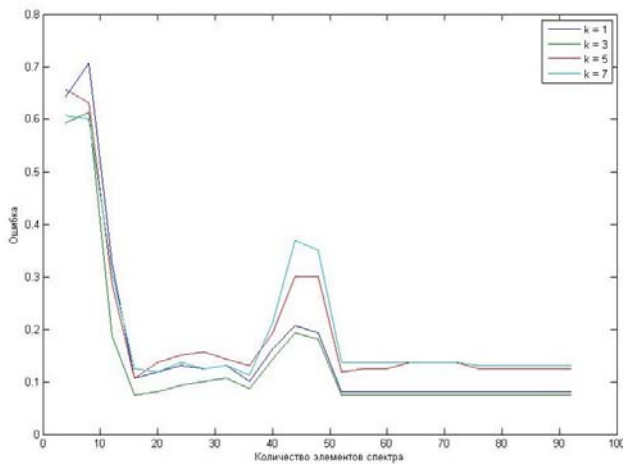
Для подбора оптимального количества соседей и количества первых элементов спектра, участвующих в классификации, была проведена серия экспериментов, в рамках которых для каждого типа спектров оценивалось качество классификации при различных значениях параметров. Результаты приведены на рис. 4.



(a)



(б)



(в)

Рис. 4. Значение ошибки классификации алгоритмом kNN в зависимости от выбора параметров: количество «соседей», количество первых элементов спектра, участвующих в рассмотрении алгоритма для: морфологической ширины (а); медиальной ширины (б); дистанционной ширины (в)

Из графиков на рис. 4 можно определить оптимальные значения параметров. Важно отметить, для всех трех методов эксперименты показали, что на имеющихся данных наиболее эффективно работает алгоритм, определяющий класс по одному ближайшему объекту обучающей выборки. Количество первых элементов спектра, используемых в классификации, для каждого из видов спектра разное, что объясняется их различиями при построении.

С учетом подобранных параметров были проведены эксперименты по вычислению качества идентификации с помощью каждого вида спектра. На рис. 5 представлены графики зависимости ошибки классификации с ростом объема выборки для моделей классификации на основе каждого из 3 спектров. Графики построены на основании проведения 100 экспериментов, в ходе каждого из которых исходная выборка случайным образом делилась на 40 блоков. После чего, в каждом из экспериментов рассматривались результаты классификации для моделей, обученных на 1, 2, 3 и т.д. блоках.

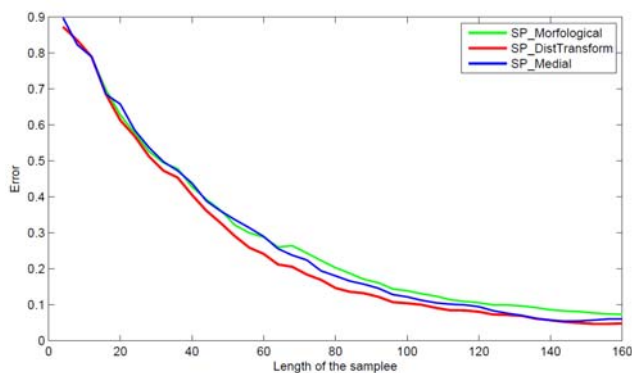


Рис. 5. Сравнение качества классификации в зависимости от длины выборки с использованием морфологического спектра (синий цвет), медиального спектра (красный цвет) и дистанционного спектра (зеленый цвет)

Итоговое значение качества классификации для моделей на основе морфологического спектра Марагоса, медиального спектра и спектра на основе карты расстояний составило **93.7 %, 94.5 %, 95.2 %**, соответственно.

4. ЗАКЛЮЧЕНИЕ

В данной работе рассмотрен один из подходов к решению задачи классификации объектов по их форме. Метод, основанный на построении спектров ширины, то есть дескрипторов ширины фигуры, не подразумевает существенных ограничений к объектам выборки и вычислительным ресурсам; при этом он обладает набором полезных свойств, например регулярность, инвариантность относительно положения объекта на изображении и т.д.

Приведены скоростные оценки спектров, показано, что вычисление морфологического спектра работает слишком медленно, и поэтому не может использоваться в прикладных системах реального времени, в отличие от медиального и дистанционного спектров. Показано также, что по критериям качества классификации все три спектра показывают качественно близкие результаты.

Эти результаты позволяют сделать следующие выводы: во-первых, спектры ширины являются хорошим инструментом, который применим как признаковое описание формы объекта. Это позволяет рассчитывать на его использование не только в задаче биометрической идентификации личности на основе формы ладони, но и в других задачах классификации изображений. Использование медиального спектра позволяет решать задачи анализа формы в режиме реального времени, так как время построения спектра составляет $\approx 1/40$ секунды, что пригодно для обработки видеопоследовательностей с высокой частотой кадров.

Спектр на основе дистанционной ширины также вычисляется за время, существенно меньшее, чем морфологический спектр, но уступает медиальному – примерно в 4 раза. При этом он обладает рядом полезных свойств: построение карты расстояний является чисто растровой операцией, эффективная реализация которой присутствует во многих современных прикладных системах обработки изображений, например в Matlab. Также экспериментально показано, что спектр на основе карты расстояний показывает относительно лучший результат классификации среди трёх рассматриваемых спектров на имеющейся выборке.

Вместе с тем, продемонстрированные результаты исследования (качество классификации 93 - 95%) не являются высокими по сравнению с существующими методами биометрической идентификации. Это говорит о том, что спектр целесообразно использовать в мультимодальных системах как один из нескольких признаков для описания объекта, отвечающий за свойства ширины объекта.

5. ССЫЛКИ

[1] Fabbri R., Costa L.Da F., Torelli J. C., Bruno O. M., 2D Euclidean Distance Transform Algorithms: A Comparative Survey // ACM Computing Surveys, Vol. 40, No. 1, Article 2, 2008.

[2] Meijster A., Roerdink J. B. T. M., Hesselink W. H., A General Algorithm for Computing Distance Transforms in Linear Time // *Mathematical Morphology and its Applications to Image and Signal Processing Computational Imaging and Vision* Volume 18, 2000.

[3] Petros Maragos. Pattern Spectrum and Multiscale shape representation // *IEEE transaction on pattern analysis and machine intelligence*, VOL II, № 7 – 1989.

[4] Ramirez-Cortes J.M., Gomez-Gil P., Sanchez-Perez G., Baez-Lopez D., A Feature Extraction Method Based on the Pattern Spectrum for Hand Shape Biometry // *Proceedings of the World Congress on Engineering and Computer Science, WCECS 2008, October 22 - 24, San Francisco, USA – 2008.*

[5] Serra J. *Image Analysis and Mathematical Morphology* // Academic Press, London, 1982.

[6] Визильтер Ю.В., Сидякин С.В. Робастное сравнение морфологических спектров двумерных фигур и изображений, семинар «Морфологический анализ данных» под руководством проф. Ю.П. Пытьева, 2012.

[7] Местецкий Л.М. Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры. – М.: Физматлит, 2009.

[8] Местецкий Л.М., Чернышев В.Г. Медиальная ширина объектов – дескриптор формы изображений // *Сб. докл. всероссийской конференции «Математические методы распознавания образов» (ММРО-16)*, 2013.

[9] Сидякин С.В. Разработка алгоритмов построения морфологических спектров для анализа цифровых изображений и видеопоследовательностей // *Диссертация на соискание ученой степени кандидата технических наук.* – М., 2013.

ОБ АВТОРАХ

Местецкий Леонид – д.т.н., профессор ВМК
МГУ им. М.В. Ломоносова.
E-mail: mestlm@mail.ru

Зак Евгений – студент ВМК МГУ им. М.В. Ломоносова.
E-mail: e_zak@mail.ru

Ломов Никита – студент ВМК МГУ им. М.В. Ломоносова.
E-mail: nikita-lomov@mail.ru