

# Сравнение формы бинарных растровых изображений на основе скелетизации\*

*Кушнир О. А.*

kushnir-olesya@rambler.ru

Тула, Тульский государственный университет

Данная работа посвящена проблеме сравнения формы бинарных растровых изображений на основе скелетных графов. Проводится анализ существующих подходов к сравнению скелетных графов, заключающихся в применении к этим графам различных методов классификации на основе векторов признаков, мер, метрик. Также ставится задача нахождения метрики, заданной в пространстве скелетных графов, которая позволила бы эффективно сравнивать формы произвольных объектов в реальном времени путем применения универсального классификатора, построенного на методе опорных векторов.

*Ключевые слова:* форма бинарного растрового изображения, скелетный граф, метрика, потенциальная функция.

## Binary bitmaps shape comparison based on skeletization\*

*O. A. Kushnir*

Tula State University

The problem to solve is to compare shapes of the binary bitmaps using the skeletal graphs. We analyzed existing approaches to the skeletal graphs comparison. This approaches are to apply to this graphs various methods of classification basing on the features, measures, metrics. Also the considered problem is to find a skeletal graphs space metrics which lets to compare the shapes of the arbitrary objects effectively in the real time by using the general classifier based on the support vectors method.

*Keywords:* binary bitmap shape, skeletal graph, metric, potential function.

## Введение

Задача распознавания формы объектов, представляемых в виде бинарных изображений, возникает во многих приложениях, например, при биометрической идентификации личности, распознавании позы и жестов человека, в системах оптического распознавания символов и пр. Наиболее удачной моделью формы является описание в виде скелетного графа, представляющего собой срединные оси фигуры. Существуют различные методы построения этого графа (они описаны в разделе 1 данной работы), и на основе результатов скелетизации фигуры различными методами можно выделять различные признаки для решения задачи классификации формы фигур.

В настоящее время не существует универсального метода для сравнения форм бинарных растровых изображений. Почти все имеющиеся меры сходства строятся на основе априорной информации о классифицируемых объектах, исходя из специфики конкретной прикладной задачи. Следовательно, для каждого нового приложения вопрос определения меры сходства решается заново. В качестве универсальной метрики в пространстве графов рассматривается только редакционное расстояние (edit distance) между двумя скелетными

---

Научный руководитель О. С. Середин

графами. Но при построении данной метрики для графов с большим количеством вершин возникает проблема большой вычислительной сложности, решаемая только приближенными методами (см. раздел 2 данной работы).

Следовательно, существует открытая задача нахождения метрики в пространстве скелетных графов. Причем необходим эффективный алгоритм ее вычисления, позволяющий сравнивать объекты в реальном времени. Последнее требование предъявляется исходя из практических приложений распознавания образов.

Наиболее эффективным на сегодняшний день методом классификации данных является метод опорных векторов (Support Vector Machines – SVM). Он является линейным классификатором, но для применения его на линейно неразделимых выборках используется переход от используемых в SVM скалярных произведений к произвольным потенциальным функциям. Потенциальные функции могут быть построены, опираясь на метрики специального вида (так называемые евклидовы метрики [8]).

В третьей части данной работы дан краткий обзор теоретических положений метода опорных векторов на основе потенциальных функций и приведен пример потенциальной функции, построенной на метрике редакционного расстояния между графами.

В качестве цели для дальнейших исследований ставится задача нахождения приемлемой метрики и на ее основе – потенциальной функции, которая позволила бы эффективно решать задачу сравнения форм произвольных объектов в реальном времени путем применения универсального классификатора для форм бинарных растровых изображений, оставаясь в рамках теории линейных методов анализа эмпирических данных на попарных методах сравнения объектов.

## 1. Скелетизация бинарных растровых изображений для задач описания формы

Если аппроксимировать черные точки бинарного растрового изображения непрерывной замкнутой областью, то *скелетом* замкнутой области называется геометрическое место точек области, являющихся центрами максимальных по включению вписанных окружностей [5] (см. рис. 1).

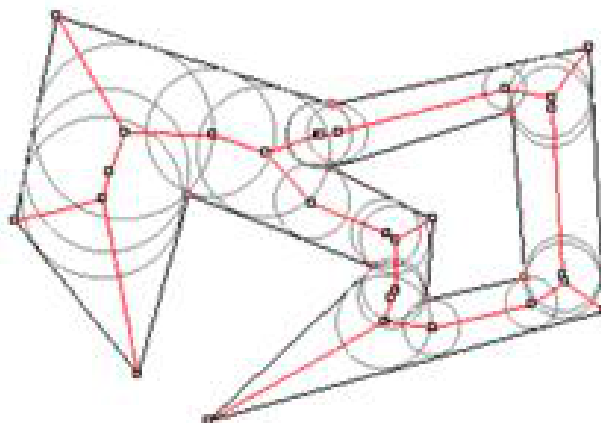


Рис. 1. Скелет замкнутой области

Скелет представляет собой плоский граф, структура которого очень хорошо отражает особенности формы объекта на изображении. Вследствие этого скелетизация широко

используется для решения задач анализа и распознавания формы объектов в системах компьютерного зрения.

Для того чтобы распространить понятие скелета на дискретные растровые изображения, применяются два основных принципиально различных подхода, которые условно можно назвать дискретным и непрерывным.

Дискретный подход к скелетизации использует определение скелета, основанное на метафоре «пожара в прерии». Предполагается, что по границе области одновременно вспыхивает огонь, который распространяется внутри нее по всем направлениям с постоянной скоростью. Те точки области, в которых сходятся два или более огненных фронта, являются по определению точками скелета. Точка схода фронтов равноудалена от ближайших точек возгорания на границе. «Пожар в прерии» служит основой для определения и построения скелета в терминах растровых бинарных изображений. Подмножество черных точек растра рассматривается как дискретный образ некоторой замкнутой области, граничные точки растрового пятна – как образ границы области, а непрерывное распространение огня моделируется дискретным процессом последовательного «сжигания» соседних черных точек растра.

Процесс построения скелета дискретного образа реализуется в различных алгоритмах «уточнения» [4, 5], в алгоритме «дистанционных карт» [5], в «волновом» алгоритме [9]. Последовательное «сжигание» соседних черных точек растра на примере метода топологического уточнения приведен на рис. 2.

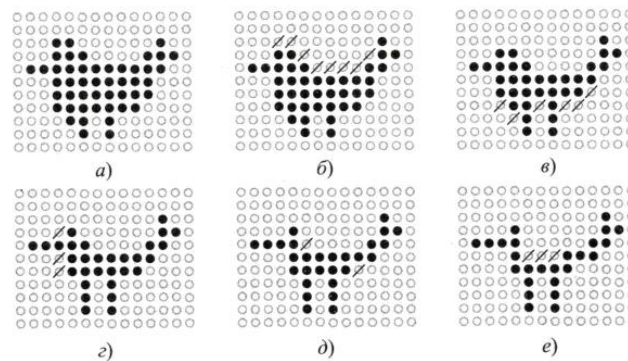


Рис. 2. Метод топологического уточнения

В результате построения скелета для дискретного бинарного изображения получается новое дискретное бинарное изображение.

Следует отметить, что дискретные скелеты обладают следующими недостатками. Во-первых, расстояние между точками растра при их построении измеряется не в евклидовой метрике, а в зависимости от используемого понятия 4- или 8-соседства точек растра, в метрике  $l_1$  («манхэттенской») или  $l_\infty$  («шахматной»). Это значит, что в качестве вписанной в область окружности выступает квадрат. Неевклидовы метрики приводят к неустойчивости получаемых скелетов по отношению к простым преобразованиям исходного изображения.

Во-вторых, дискретный скелет вычисляется неоднозначно и зависит от последовательности анализа граничных точек образа. Успешное построение дискретных аналогов скелета возможно, как правило, если дискретные фигуры имеют простую структуру.

Непрерывный подход к построению скелета растрового образа развит и хорошо исследован в работах школы Л.М. Местецкого [5, 6, 7, 11, 13] и состоит в аппроксимации

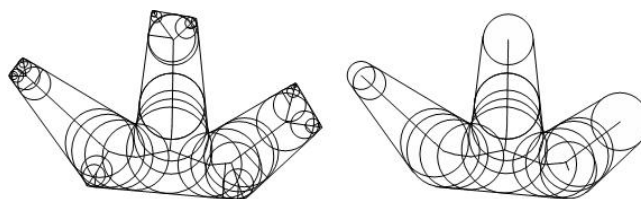
границы дискретной фигуры многоугольником минимального периметра и построении скелета области, ограниченной этим многоугольником, в соответствии с формальным его определением, что скелет – это множество тех точек фигуры, для которых существует не менее двух равноудаленных ближайших точек границы области. Итогом применения данного подхода к растровому изображению является математическое описание скелета аппроксимирующей данное изображение «непрерывной» замкнутой области.

Скелет чрезвычайно чувствителен к локальным свойствам границы образа. С каждой точкой локального максимума кривизны границы связана отдельная ветвь скелета. Две области, имеющие несущественные для глаз различия границы, например, за счет шумов, имеют принципиально различные в смысле топологической структуры скелеты.

Поэтому после построения скелета фигуры проводится выделение в скелете части, которая называется базовым скелетом, и удаление элементов, появление которых обусловлено шумовыми эффектами. Метод получения этой информативной части скелета называется регуляризацией скелета с контролируемой точностью.

Объединение скелета фигуры и максимальных вписанных в фигуру окружностей, центры которых и составляют скелет, называется в рамках непрерывного подхода к скелетизации циркулярным представлением фигуры. Объект может быть описан в виде совокупности примитивов определенного вида. В качестве таких примитивов используются жирные линии – объединения однопараметрического семейства кругов переменного радиуса на непрерывных кривых. Жирная линия представляет собой след от перемещения окружности переменного радиуса вдоль непрерывной кривой. Для произвольного растрового бинарного образа вначале строится скелет, затем каждую отдельную ветвь скелета аппроксимируют жирными линиями. Циркулярное представление является развитием понятия скелета и дает возможность не только анализировать форму объекта, но и осуществлять ее преобразования.

На рис. 3 показан многоугольник минимального периметра, скелет (слева), базовый скелет и итоговое циркулярное представление фигуры (справа).



**Рис. 3.** Многоугольник минимального периметра, скелет (слева), базовый скелет и итоговое циркулярное представление фигуры (справа)

Достоинствами непрерывного подхода к скелетизации являются математическая корректность, адекватность человеческому представлению о форме и ее преобразованиях, широкие возможности для преобразования и сравнения форм, высокая вычислительная эффективность.

Недостатком же является существенное усложнение алгоритмов с точки зрения их математического содержания и программной реализации по сравнению с алгоритмами дискретной скелетизации.

Анализ существующих подходов к скелетизации позволяет сделать предположение о том, что непрерывный подход является наиболее приоритетным для построения матема-

тической модели решения задачи сравнения формы бинарных растровых изображений в силу перечисленных выше достоинств.

## 2. Анализ существующих методов сравнения и классификации формы бинарных растровых изображений

В зависимости от подхода к построению скелета, при разработке методов сравнения фигур скелет рассматривается либо как взвешенный граф, состоящий из вершин и дуг (дискретный подход к построению скелета), либо как циркуляр (непрерывный подход к построению скелета).

Методы, рассматривающие скелет как взвешенный граф, можно условно разделить на использующие для сравнения скелетов топологические признаки графов и использующие редакционное расстояние между двумя графами.

Топологические методы применяются, например, для решения задачи распознавания печатных символов [1]. Каждый контур скелета описывается в виде набора особых точек и так называемого цепного кода, состоящего из точки привязки, числа кодов и массива направлений из очередной точки на следующую точку. Особые точки – это концевые точки и точки ветвления (триоды), т.е. точки, соседи которых образуют не менее трех связных областей. Для каждой особой точки вычисляются следующие топологические признаки:

- нормированные координаты особой точки (вершины графа);
- длина ребра до следующей вершины в процентах от длины всего графа;
- нормированное направление из данной точки на следующую особую точку;
- нормированное направление входа в точку, выхода из точки (для триодов эти признаки различаются, для точек индекса «1» совпадают с точностью до знака);
- кривизна дуги, точнее «левая» и «правая» кривизна дуги, соединяющей особую точку со следующей вершиной (кривизна слева и справа). Кривизна вычисляется как отношение максимального расстояния от точек дуги (находящихся соответственно слева/справа от прямой) до прямой, соединяющей вершины, к длине отрезка, соединяющей те же вершины.

На рис. 4 условно показаны некоторые из топологических признаков.

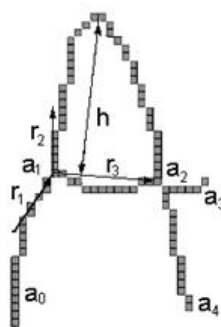


Рис. 4. Примеры скелетных признаков

Обучение метода состоит в построении деревьев распознавания для каждого из определенных заранее (вручную или автоматически) топологических кодов. Распознавание символов ведется на основе обхода построенных деревьев. Метрики для попарного сравнения символов данным методом не предоставляет. Поэтому его нельзя использовать для построения потенциальной функции классификатора по методу опорных векторов (см. следующий раздел).

В методах сравнения формы объектов, использующих редакционное расстояние, в качестве метрики рассматривается соответственно редакционное расстояние (edit distance) между двумя скелетными графами [15].

Понятие редакционного расстояния изначально использовалось для сравнения строк и определяется как минимальное количество операций вставки одного символа, удаления одного символа и замены одного символа на другой, необходимых для превращения одной строки в другую. Чтобы измерить различие между строками, вводится понятие неотрицательных весов для каждой из операций. Следовательно, необходимо найти такую последовательность замен в строке, которая минимизировала бы суммарный вес всех операций [3].

Граф можно представить как множество вершин, дополненное отношением смежности между вершинами, представленное ребрами, соединяющими вершины. Редакционное расстояние для графов определяется аналогично строкам как минимальное количество операций вставки одной вершины, удаления одной вершины и замены одной вершины на другую, вставки одного ребра, удаления одного ребра и замены одного ребра на другое, необходимых для превращения одного графа в другой. В случае взвешенного редакционного расстояния между графами необходимо найти такую последовательность замен в графе, которая минимизировала бы суммарный вес всех операций. В качестве весов вершин для скелетных графов может употребляться их численно выраженная позиция на изображении.

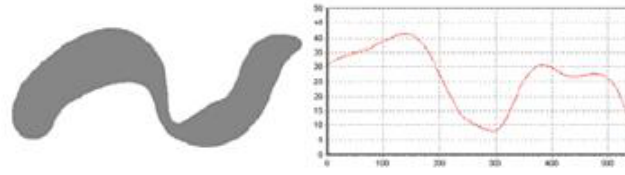
Задача поиска оптимальной последовательности замен решается методом динамического программирования, но имеет экспоненциальную вычислительную сложность в зависимости от количества вершин в графе. На практике с приемлемой скоростью можно вычислить редакционное расстояние между графами, которые имеют не более 10 вершин [15]. Соответственно, такая метрика не подходит для сравнения изображений сложной формы, в скелетах которых количество вершин превышает 10. Чтобы обойти данное ограничение, для сравнения скелетных графов используются алгоритмы приближенных расчетов редакционного расстояния [14].

В работах, посвященных классификации формы объектов на основе непрерывного скелета – циркулярной модели фигуры, используются меры сходства объектов на основе сравнения составляющих их жирных линий и на основе вычисления расхождения граничных функций ширины двух силуэтов.

*Мера сходства объектов на основе сравнения составляющих их жирных линий*

Пусть имеется произвольная составная жирная B-сплайновая кривая  $C$ , заданная в виде вектор-функции  $C(t) = [x(t), y(t), r(t)]$ , где  $P(t) = [x(t), y(t)]$  задает ось жирной линии, а  $r(t)$  – ее ширину,  $t \in [0, T]$ . Обозначим через  $l = f(t)$  функцию длины осевой линии, зависящую от параметра  $t$ , тогда  $t = f^{-1}(l)$ . Рассмотрим функцию  $r(l) = r(f^{-1}(l))$  – зависимости ширины жирной кривой от длины осевой линии. Пример жирной линии и соответствующий ей график функции  $r(l)$  представлен на рис. 5, где по оси абсцисс отложена длина осевой линии, а по оси ординат – ширина жирной кривой.

Рассмотрим теперь две произвольные жирные линии  $C_1(t)$  и  $C_2(t)$ . Построим в общей системе координат графики функций  $r_1(l)$  и  $r_2(l)$ . Обозначим через  $L_1$  длину осевой линии жирной кривой  $C_1(t)$ , а через  $L_2$  – длину осевой линии жирной кривой  $C_2(t)$ . Пусть  $L = \max(L_1, L_2)$ , и доопределим функции  $r_i(l)$  на  $[0, L]$  как  $r_i(l) = 0$  при  $l > L_i$ , если  $L_i < L$ . Пусть  $D_1$  – множество точек плоскости, ограниченной графиком функции  $r_1(l)$  и осью абсцисс, а  $D_2$  – ограниченное графиком  $r_2(l)$  и осью абсцисс. Обозначим через  $s(D_1 \cup D_2)$  – площадь объединения множеств  $D_1$  и  $D_2$ , а через  $s(D_1 \cap D_2)$  – площадь их пересечения.



**Рис. 5.** Жирная В-сплайновая кривая и график зависимости ширины жирной кривой от длины осевой линии

Под мерой сходства двух жирных линий  $C_1(t)$  и  $C_2(t)$  будем понимать величину

$$\Delta = \frac{s(D_1 \cap D_2)}{s(D_1 \cup D_2)} = \frac{\int_0^L \min(r_1(l), r_2(l)) dl}{\int_0^L \max(r_1(l), r_2(l)) dl}$$

Поскольку площадь пересечения двух множеств не превышает площади их объединения, имеем  $s(D_1 \cup D_2) \geq s(D_1 \cap D_2)$  и величина  $\Delta \leq 1$ . Очевидно, что чем больше значение  $\Delta$ , тем две жирные линии более похожи. Следует также заметить, что у абсолютно идентичных жирных линий величина сходства будет равна единице. Тем самым, нахождение меры близости двух жирных линий сводится к нахождению отношения двух определенных интегралов [6].

Далее для сравнения образов в конкретных приложениях проводится обработка составляющих их жирных линий. Например, для задачи попарного сравнения образов человеческих ладоней предлагается следующий алгоритм:

1. Для каждого образа строится его циркулярное разложение,
2. Из каждого разложения выбираются и идентифицируются пять самых длинных жирных линий,
3. Производится попарное сравнение соответствующих жирных линий. Результат сравнения  $i$ -ой пары есть значение  $\Delta_i \in [0, 1]$ ,  $i = 1, \dots, 5$ , которое характеризует меру сходства двух жирных линий.
4. Результатом сравнения двух образов ладоней будет значение  $\omega = \frac{\sum_{i=1}^5 \Delta_i}{5}$ . Полученное значение  $\omega \in [0, 1]$  показывает величину относительной меры сходства двух образов ладоней [13].

*Мера сходства объектов на основе вычисления расхождения граничных функций ширины двух силуэтов*

Возьмем точку на скелете и начнем обход скелета по часовой стрелке. Зависимость ширины скелета от длины пройденного пути будем называть *граничной функцией ширины*. Обход закончится, когда вернемся в точку, откуда начинали движение. Отмасштабируем фигуру таким образом, чтобы длина полного пути обхода равнялась 1, т.е. будем считать, что граничная функция определена на отрезке  $[0, 1]$ . В зависимости от того, с какой точки мы начинали движение, получим разные функции, но они будут отличаться друг от друга лишь циклическим сдвигом в области аргумента. Кроме ширины построим аналогичную функцию для степени вершины скелета. Для узловых вершин скелета степень вершины равна количеству исходящих из нее ребер, для всех остальных точек скелета (это точки, которые находятся строго внутри ребра) определим ее равным 2.

Пусть имеется два силуэта,  $r_1 : [0, 1] \rightarrow R^2$  и  $r_2 : [0, 1] \rightarrow R^2$  – граничные функции ширины для них, а  $\text{deg}_1 : [0, 1] \rightarrow \{1, 2, 3\}$  и  $\text{deg}_2 : [0, 1] \rightarrow \{1, 2, 3\}$  граничные функции

для степени вершины скелетов соответственно. Объединим две такие функции в двухкомпонентный вектор  $R_i = (r_i, \text{deg}_i) : i \in \{1, 2\}$ . Обозначим через  $T_\delta$  оператор циклического сдвига аргументов на  $\delta$  для функций, определенных на отрезке  $[0, 1]$ , т.е.  $\forall w : [0, 1] \rightarrow [0, 1]$ :

$$T_\delta \circ f(t) = \begin{cases} f(t + \delta), & \text{при } t + \delta \leq 1 \\ f(t + \delta - 1), & \text{при } t + \delta > 1 \end{cases}$$

Здесь значком  $\circ$  обозначена операция суперпозиции. Задача выравнивания заключается в построении непрерывного монотонного отображения  $w : [0, 1] \rightarrow [0, 1]$  и нахождения сдвига  $\delta \geq 0$ , которые сопоставляют граничные функции ширины  $R_1(t) \leftrightarrow T_\delta \circ R_2(w(t))$  и при этом минимизируют расхождение, заданное в виде функционала:

$$d(R_1, R_2) = \min_{\delta \in [0, 1]} \min_{\substack{w \in C[0, 1] \\ w(0)=0 \\ w(1)=1 \\ w\text{-монот.}}} \int_0^1 \rho(R_1(t), R_2(T_\delta \circ w(t))) \sqrt{1 + w'(t)^2} dt$$

Будем требовать, чтобы узловые вершины скелетов по возможности совпадали (топология скелетных графов может отличаться), для этого в функционал добавлен член  $k_1 \cdot |\text{deg}_1 - \text{deg}_2|$ . Также потребуем, чтобы граничные функции ширины как можно лучше совпадали после подгонки. В итоге получаем следующую функцию штрафа:

$$\rho = (|r_1 - r_2| + k_1 \cdot |\text{deg}_1 - \text{deg}_2| + k_2)$$

Здесь  $k_2$  константа, которая отвечает за гладкость кривой  $w(t)$ .

Отметим некоторые свойства выравнивания  $R_1(t) \leftrightarrow T_\delta \circ R_2(w(t))$ :

1. Очевидно, что выполняется условие рефлексивности:  $d(R, R) = 0$ .
2. Легко проверить, что такое выравнивание симметрично относительно  $R_1, R_2$ , т.е.  $d(R_1, R_2) = d(R_2, R_1)$ .
3. Однако неравенство треугольника не выполняется:  $d(R_1, R_2) + d(R_2, R_3)$  не всегда больше  $d(R_1, R_3)$ .

Следовательно,  $d(R_1, R_2)$  не является метрикой. Тем не менее, значение функционала  $d(R_1, R_2)$  показывает, насколько похожи силуэты.

Дискретизируем задачу относительно параметра  $\delta$ . Выберем  $N$  точек  $\frac{i}{N}, i = 0..N$  на отрезке  $[0, 1]$  и будем перебирать значение  $\delta$  среди них. Тогда задача сводится к  $N$  подзадачам, каждая из которых решается методом динамического программирования за время  $O(N^2)$ . Следовательно, задача минимизации решается за время  $O(N^3)$  [11].

### 3. Задача сравнения форм объектов при помощи потенциальной функции скелетного графа

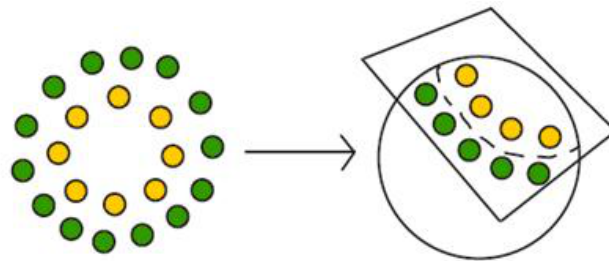
Наиболее эффективным на сегодняшний день методом классификации данных является метод опорных векторов (Support Vector Machines – SVM). В основу его положен алгоритм построения оптимальной разделяющей гиперплоскости, предложенный в 1963 году В.Н. Вапником [2]. Метод опорных векторов является линейным классификатором, а следовательно, применение его на линейно неразделимых выборках ведет к появлению большого количества неверно классифицированных объектов. Для решения этой проблемы в 1992 году Бернхард Босер, Изабель Гийон и В.Н. Вапник предложили способ создания нелинейного классификатора, в основе которого лежит переход от используемых



в SVM скалярных произведений к произвольным ядрам линейного преобразования, так называемый *kernel trick* (предложенный впервые М.А. Айзерманом, Э.М. Броверманом и Л.В. Розоноером для метода потенциальных функций).

Это переход от исходного пространства признаков описаний объектов  $X$  к новому, так называемому спрямляющему, пространству  $H$  с помощью некоторого преобразования  $\psi : X \rightarrow H$ . Если пространство  $H$  имеет достаточно высокую размерность, то можно предположить, что в нём выборка окажется линейно разделимой.

Пример перехода к расширенному пространству показан на рис. 6. Точки двух классов (внешняя и внутренняя окружность) на плоскости линейно не разделимы. Если же мы перенесем эти точки на сферу (трехмерное пространство), то тогда они разделяются плоскостью, которая срезает часть сферы вместе с точками на внутренней окружности. Таким образом, «выгнув» пространство вместе с точками с помощью отображения  $\psi : X \rightarrow H$ , можно легко найти разделяющую гиперплоскость.



**Рис. 6.** Пример перехода к расширенному пространству для построения разделяющей гиперплоскости

Результирующий алгоритм похож на алгоритм линейной классификации, с той лишь разницей, что каждое скалярное произведение в методе опорных векторов заменяется нелинейной потенциальной функцией (скалярным произведением в пространстве с большей размерностью). В этом пространстве уже может существовать оптимальная разделяющая гиперплоскость. Так как размерность получаемого пространства может быть больше размерности исходного, то преобразование, сопоставляющее скалярные произведения, будет нелинейным, а значит функция, соответствующая в исходном пространстве оптимальной разделяющей гиперплоскости, будет также нелинейной.

Отсюда вытекает естественное требование: пространство  $H$  должно быть наделено скалярным произведением, в частности, подойдёт любое евклидово, а в общем случае и гильбертово, пространство.

Определение потенциальной функции (ядра): функция  $K : X \times X \rightarrow R$  называется *ядром* (*kernel function*), если она представима в виде  $K : (x, x') = (\psi(x), \psi(x'))$  при некотором отображении  $\psi : X \rightarrow H$ , где  $H$  – пространство со скалярным произведением.

Алгоритм классификации по методу опорных векторов зависит только от скалярных произведений объектов, но не от самих признаков описаний. Это означает, что скалярное произведение  $(x, x')$  можно формально заменить потенциальной функцией  $K(x, x')$ . Поскольку потенциальная функция в общем случае нелинейна, такая замена приводит к существенному расширению множества реализуемых алгоритмов. Более того, можно вообще не строить спрямляющее пространство  $H$  в явном виде, и вместо подбора отображения  $\psi$  заниматься непосредственно подбором потенциальной функции.

Можно и вовсе отказаться от признаковых описаний объектов. Для такого подхода был придуман термин беспризнаковое распознавание (featureless recognition). Объекты можно изначально задать информацией об их попарном взаимоотношении, например, отношении сходства. Если эта информация допускает представление в виде двуместной функции  $K(x, x')$ , удовлетворяющей аксиомам скалярного произведения, то задача может решаться методом SVM. Поэтому необходимо найти отношение сходства между двумя объектами, являющееся метрикой.

Функция  $K(x, x')$  является ядром тогда и только тогда, когда она симметрична,  $K(x, x') = K(x', x)$ , и неотрицательно определена:

$$\int_X \int_X K(x, x')g(x)g(x')dx dx' \geq 0 \text{ для любой функции } g : X \rightarrow R.$$

Существует эквивалентное определение неотрицательной определённости: функция  $K(x, x')$  неотрицательно определена, если для любой конечной выборки  $X_p = (x_1, \dots, x_p)$  из  $X$  матрица  $K = \|K(x_i, x_j)\|$  размера  $p \times p$  неотрицательно определена:  $z^T K z > 0$  для любого  $z \in R^p$ .

Проверка неотрицательной определённости функции в практических ситуациях может оказаться делом нетривиальным. Часто ограничиваются перебором конечного числа функций, про которые известно, что они являются ядрами. Среди них выбирается лучшая, как правило, по критерию скользящего контроля. Очевидно, что это не оптимальное решение. На сегодняшний день проблема выбора ядра, оптимального для данной конкретной задачи, остаётся открытой.

Существуют правила порождения, позволяющие строить ядра в практических задачах:

1. Произвольное скалярное произведение  $K(x, x') = \langle x, x' \rangle$  является ядром.
2. Константа  $K(x, x') = 1$  является ядром.
3. Произведение ядер  
 $K(x, x') = K_1(x, x')K_2(x, x')$  является ядром.
4. Для любой функции  $\psi : X \rightarrow R$  произведение  $K(x, x') = \psi(x)\psi(x')$  является ядром.
5. Линейная комбинация ядер с неотрицательными коэффициентами  $K(x, x') = \alpha_1 K_1(x, x') + \alpha_2 K_2(x, x')$  является ядром.
6. Композиция произвольной функции  $\varphi : X \rightarrow X$  и произвольного ядра  $K_0$  является ядром:  $K(x, x') = K_0(\varphi(x), \varphi(x'))$ .
7. Если  $s : X \times X \rightarrow R$  - произвольная симметричная интегрируемая функция, то  $K(x, x') = \int_X s(x, z)s(x', z)dz$  является ядром.
8. Функция вида  $K(x, x') = k(x - x')$  является ядром тогда и только тогда, когда Фурье-образ  $F[k](\omega) = (2\pi)^{\frac{n}{2}} \int_X e^{-i(\omega, x)} k(x)dx$  неотрицателен.
9. Предел локально-равномерно сходящейся последовательности ядер является ядром.
10. Композиция произвольного ядра  $K_0$  и произвольной функции  $f : R \rightarrow R$ , представимой в виде сходящегося степенного ряда с неотрицательными коэффициентами  $K(x, x') = f(K_0(x, x'))$ , является ядром. В частности, функции  $f(z) = e^z$  и  $f(z) = \frac{1}{1-z}$  от ядра являются ядрами.

Наиболее распространённые ядра:

— полиномиальное (однородное)

$$K(x, x') = (x \cdot x')^d$$

— полиномиальное (неоднородное)

$$K(x, x') = (x \cdot x' + 1)^d$$

— радиальная базисная функция

$$K(x, x') = \exp(-\gamma \|x - x'\|^2),$$

для  $\gamma > 0$

— радиальная базисная функция Гаусса

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{2\sigma^2}\right)$$

— сигмоид

$$K(x, x') = \tanh(kx \cdot x' + c),$$

для почти всех  $k > 0$  и  $c < 0$  [10, 16].

На основании материала предыдущего раздела можно утверждать, что для сравнения форм бинарных растровых изображений задача поиска метрик, на основе которых может быть построена потенциальная функция, остается очень актуальной: на данный момент метрика может быть получена только методом редакционного расстояния, но этот метод учитывает преимущественно количество вершин и ребер в скелетном графе, опуская другие его не менее важные особенности, как например, ширина или кривизна ветвей. Следовательно, часть признаков объектов не будет использована для построения классификатора, что ведет к понижению надежности распознавания.

Рассмотрим в качестве примера одну из потенциальных функций, которая может быть построена на основе редакционного расстояния между графами.

Пусть  $X$  – признаковое пространство скелетных графов и  $X^t \subset X$  – обучающая выборка образов. Для фиксированного образа  $x_0 \in X^t$  потенциальная функция  $k_{x_0} : X \times X \rightarrow R$  может быть определена как

$$k(x, x') = k_{x_0}(x, x') = \frac{1}{2}(d(x, x_0)^2 + d(x_0, x')^2 - d(x, x')^2),$$

где  $d(\cdot, \cdot)$  – редакционное расстояние между двумя образами.

Эта потенциальная функция может быть интерпретирована как мера квадратного расстояния от образа  $x$  до  $x_0$  и от образа  $x_0$  до  $x'$  относительно к квадратному расстоянию от образа  $x$  до  $x'$  напрямую.

Данная потенциальная функция действительна, потому что существует пространство со скалярным произведением  $H$ , где каждый граф  $x \in X$  представлен единственным вектором  $\Phi(x) \in H$  и скалярное произведение векторов эквивалентно потенциальной функции. Образ  $x_0$  называется нулевым графом, поскольку он обладает свойствами начала координат признакового пространства скелетных графов.

Эксперименты, проведенные авторами данной потенциальной функции, показывают, что применение даже такой простой функции в методе опорных векторов дает более корректные результаты классификации изображений по сравнению с методом ближайших соседей [15]. Так что поиск метрик в пространстве скелетных графов и построение потенциальных функций на их основе представляется очень перспективным направлением для классификации изображений на основе скелетного представления их формы.

## Выводы

Для сравнения и классификации форм бинарных растровых изображений задача поиска метрик в пространстве скелетных графов остается очень актуальной: на данный момент метрика может быть получена только методом редакционного расстояния, но этот

метод учитывает преимущественно количество вершин и ребер в скелетном графе, опуская другие его не менее важные особенности, как например, ширина или кривизна ветвей. Следовательно, часть признаков объектов не будет использована для построения классификатора, что ведет к понижению надежности распознавания.

Наиболее перспективной представляется метрика для непрерывных скелетов фигуры, которые описываются скелетным графом и вписанными в фигуру максимальными окружностями с центрами на скелете, поскольку такое математическое описание более адекватно человеческому представлению о форме и ее преобразованиях. Построенная на основе такой метрики функция ядра должна повысить возможности использующего ее классификатора.

Поскольку такой метрики в настоящее время нет, то актуальна задача ее нахождения и построения на ее основе потенциальной функции, применяемой в методе опорных векторов для классификации объектов.

## Литература

- [1] *Афонасенко А. В., Елизаров А. И.* Обзор методов распознавания структурированных символов. // Доклады ТУСУРа, 2008, № 2, часть 1.
- [2] *Вапник В. Н., Червонекис А. Я.* Об одном классе алгоритмов обучения распознаванию образов. // Автоматика и телемеханика, 1964, № 4.
- [3] *Гасфилд Д.* Строки, деревья и последовательности в алгоритмах: Информатика и вычислительная биология. // СПб: Невский диалект; БХВ-Петербург, 2003.
- [4] *Гонсалес Р., Вудс Р.* Цифровая обработка изображений. // М: Техносфера, 2005.
- [5] *Местецкий Л. М.* Непрерывная морфология бинарных изображений: фигуры, скелеты, циркуляры. // М: ФИЗМАЛИТ, 2009.
- [6] *Местецкий Л. М.* Компьютерная графика на основе жирных линий. // Труды межд. конф. «Графикон-2000», Москва, 2000.
- [7] *Местецкий Л. М., Рейер И. А.* Непрерывная гранично-скелетная модель дискретного изображения с контролируемой точностью аппроксимации. // Доклады Всероссийской конференции «Математические методы распознавания образов», (ММРО-11). Москва, 2003, с. 367-371.
- [8] *Моттль В. В.* Метрические пространства, допускающие введение линейных операций и скалярного произведения. // Доклады Академии наук, 2003, том 388, № 3, с.1-4.
- [9] Применение волнового алгоритма для нахождения скелета растрового изображения. // [Электронный ресурс]. Режим доступа: <http://osgai.narod.ru>. Загл. с экрана.
- [10] *Середин О. С., Моттль В. В.* Методы беспризнакового распознавания образов // Тула: Изд-во ТулГУ, 2004.
- [11] *Цискаридзе А. К.* Восстановление пространственных циркулярных моделей по силуэтным изображениям // Диссертация на соискание ученой степени кандидата физико-математических наук. МФТИ, 2010.
- [12] *Dupe F. X., Brun L.* Shape classification using a flexible graph kernel // Proceedings of Academic Press 2009, September 2009.
- [13] *Mestetskiy L., Semenov A.* Palm Shape Comparison Based on Fat Curves // Proceedings of 7th International conference on Pattern recognition and image analysis: new information technologies, St. Petersburg, 2004, pp. 788–791.
- [14] *Neuhaus M., Bunke H.* An error-tolerant approximate matching algorithm for attributed planar graphs and its application to fingerprint classification, in: Proceedings of the 10th International Workshop on Structural and Syntactic Pattern Recognition. // Lecture Notes in Computer Science, vol. 3138, Springer, Berlin, 2004, pp. 180–189.
- [15] *Neuhaus M., Bunke H.* Edit-distance based kernel for structural pattern classification. // Pattern Recognition 39 (2006), pp. 1852–1863.
- [16] *Vapnik V.* Statistical Learning Theory. // NY.: J. Wiley, 1998. 768 p.