

КОДИРОВАНИЕ КОНТУРОВ, ПРЕДСТАВЛЕННЫХ КРИВЫМИ БЕЗЬЕ, В ЗАДАЧАХ НЕЙРОСЕТЕВОЙ КЛАССИФИКАЦИИ

Abstract: The article is devoted to describe method and algorithms of coding the contours represented with Bezier curves for neural network pattern recognition. Because a contour description can consist arbitrary number and sequence of the curves, neural networks utilization requires extra coding. For the coding there are used histograms based on parametric notion of curves in the contour. There are proposed limitations on Bezier curves that allow significantly reduce the length of obtained codes. Efficiency of developed method is shown with example of classification of plane figures.

Key words: pattern recognition, coding of the contours, neural networks.

Анотація: Стаття присвячена опису методу та алгоритмів кодування контурів, які представлені кривими Безьє, для розпізнавання візуальних образів за допомогою нейронних мереж. Оскільки опис контура може мати довільну кількість і послідовність кривих Безьє, то для використання нейронних мереж необхідне додаткове кодування. Додаткове кодування виконується за допомогою гістограм з використанням параметричного запису кривих у контурі. Запропоновані обмеження на криві Безьє, які дозволяють суттєво зменшити довжину отриманого коду. Працездатність та ефективність розробленого методу показана на прикладі класифікації плоских фігур.

Ключові слова: розпізнавання образів, кодування контурів, нейронні мережі.

Аннотация: Статья посвящена описанию метода и алгоритмов кодирования контуров, представленных кривыми Безье, для нейросетевого распознавания зрительных образов. Поскольку описание контура может содержать произвольное количество и последовательность кривых Безье, то для использования нейронных сетей требуется дополнительное кодирование. Дополнительное кодирование осуществляется при помощи гистограмм на основе параметрического представления кривых контура. Предложены ограничения на кривые Безье, позволяющие существенно сократить длину получаемого кода. Работоспособность и эффективность разработанного метода показана на примере классификации плоских фигур.

Ключевые слова: распознавание образов, кодирование контуров, нейронные сети.

1. Введение

Распознавание зрительных образов широко используется в автоматических системах диагностики, безопасности, управления и многих других. Одной из центральных проблем распознавания зрительных образов является идентификация объекта на изображении. Проблема идентификации может ставиться как задача точной идентификации объекта, например, в дактилоскопии, или в виде задачи классификации, то есть отнесение объекта к одному из известных системе классов.

Обычно в алгоритмах идентификации используются не сами изображения, а признаки объектов. В зависимости от постановки задачи и типа изображений (например, цветное или монохромное) применяют различные наборы признаков. При поиске и идентификации объектов наиболее информативными признаками являются признаки формы. К таким признакам относятся соотношения характерных размеров и характеристики контуров как самого объекта, так и его элементов.

Для решения многих задач распознавания оказывается достаточно выделить и описать контуры объектов и их элементов. Одним из наиболее мощных средств описания контуров, как кривых произвольной формы, являются сплайны и, в частности, кривые Безье. Однако в описании контура может содержаться произвольное число кривых, что затрудняет использование таких описаний в задачах распознавания, например, с помощью нейронных сетей.

Для решения этой проблемы в работе предлагается использовать гистограммное кодирование контуров, заданных кривыми Безье. Нас интересовали возможность построения и эффективность применения алгоритма кодирования при использовании описания контуров в качестве информативных признаков на входе нейронной сети.

2. Постановка задачи

Кривые Безье позволяют точно описать произвольный контур, благодаря чему эти кривые широко применяются в векторной графике [1]. Однако прямое использование контуров, представленных кривыми Безье, для решения задачи распознавания затруднено. Проблема заключается в том, что, в зависимости от качества изображения и погрешностей преобразования, описание контура содержит разное количество кривых. Кроме того, последовательность кривых зависит от того, какая из характерных точек объекта выбрана в качестве начальной при описании контура.

Кривая Безье записывается как пара якорных (anchor) точек, задающих начало и конец кривой, и, в общем случае, произвольное количество управляющих (control) точек, задающих форму кривой. Иногда якорные и управляющие точки не различают и говорят об опорных точках.

Для построения кривой Безье N -ого порядка необходимо $N+1$ точек. В компьютерной графике наибольшее распространение получили квадратичные кривые Безье, строящиеся по трём точкам, и кубические кривые Безье, строящиеся по четырём точкам. Квадратичные кривые Безье используются, например, в шрифтах TrueType при определении контуров символов. Кубические кривые Безье задаются следующей формулой:

$$P(t) = A \cdot (1-t)^3 + 3 \cdot B \cdot t \cdot (1-t)^2 + 3 \cdot C \cdot (1-t) \cdot t^2 + D \cdot t^3,$$

где A – начало кривой; D – её конец, а B и C – первая и вторая управляющие точки. Прямая AB является касательной к кривой в точке A ; прямая CD – в точке D . Параметр t изменяется от 0 до 1. При $t = 0$ $P(t) = A$, при $t = 1$ $P(t) = D$.

Одним из важнейших свойств кривой Безье является её делимость. Если кривую разделить на две кривых в точке $t = 0,5$, каждая из полученных кривых также будет являться кривой Безье.

Другие, широко используемые свойства кривой:

- 1) инвариантность относительно аффинных преобразований;
- 2) инвариантность относительно линейных замен параметризации;
- 3) кривая Безье принадлежит выпуклой оболочке опорных точек;
- 4) кривая проходит через начальную и конечную точки;
- 5) симметричность, то есть, если рассматривать опорные точки в обратном порядке, то кривая не изменится;
- 6) степень многочлена, представляющего кривую в аналитическом виде, на единицу меньше числа опорных точек.

После того как изображение проанализировано и контуры объектов описаны, используется алгоритм распознавания, позволяющий по выделенным признакам идентифицировать или классифицировать объект. Для решения прикладных задач распознавания зрительных образов

широко применяются искусственные нейронные сети. Особенно успешно нейронные сети применяются при решении задач классификации [2].

Нейронные сети имеют фиксированное число входов. Причем каждый вход нейронной сети рассматривается как отдельный признак объекта (в идеале – независимый признак). Поэтому для использования в качестве входов нейронной сети коды объектов должны удовлетворять следующим двум условиям:

- 1) любой код объекта имеет всегда одну и ту же длину;
- 2) каждый вход сети (элемент кода объекта) всегда отражает одну и ту же характеристику этого объекта.

В данной работе ставилась задача разработки метода, позволяющего использовать описания контуров кривыми Безье для классификации объектов при помощи нейронных сетей.

3. Кодирование контуров

К алгоритму кодирования предъявлялись два требования. Первое – сохранение инвариантности получаемого кода к аффинным преобразованиям объекта. Проблема обусловлена тем, что хотя описание в виде кривых Безье и не изменяет относительные характеристики контура при масштабировании и поворотах объекта в плоскости изображения, но масштаб контура, представленного описанием, заранее не известен. Первое требование полностью удовлетворяется применением гистограммного кодирования.

Второе требование – единообразие представления контуров объектов одного класса (или конкретного объекта, при точной идентификации). Данная проблема вызвана произволом в выборе типов кривых и неточностью алгоритмов при аппроксимации контуров.

Если не накладывать определенные ограничения на кривые Безье, используемые при описании контуров, то понадобится код бесконечной длины для точного описания произвольного контура. Под кодом бесконечной длины в данном случае подразумевается, что при выборе любого фиксированного количества параметров описания контура всегда можно либо построить контур, требующий большего количества кривых Безье с постоянным числом параметров, либо описать тот же контур кривыми Безье с большим числом параметров (управляющих точек).

В задачах распознавания зрительных образов требование единообразного представления может быть выполнено, если на кривые Безье, используемые в описании контура, наложены следующие ограничения:

- 1) при кодировании любых объектов используются кривые с фиксированным числом управляющих точек;
- 2) общая длина контура, или наибольший размер объекта, ограниченного контуром, должен быть больше некоторой заданной величины, измеряемой в единицах дискретизации изображения. Например, размер объекта должен быть не менее 20 пикселей.

Расстояния между опорными точками в кривых Безье (в простейшем случае расстояние между якорными точками) больше некоторой величины δ , которая задается либо как доля от общей длины контура, либо как процент от наибольшего размера объекта. Например, расстояние между якорными точками в кривых Безье должно быть не менее 0,05 от наибольшего размера

объекта. Очевидно, что для реальных задач это расстояние не может быть меньше шага растра на изображении.

Последние два ограничения призваны обеспечить устойчивость описания контуров к шумам на изображении и неточностям самого алгоритма аппроксимации. Эти ограничения существенно зависят от решаемой задачи и качества исходных изображений.

Наиболее широкое распространение получили кривые Безье с четырьмя опорными точками. Это связано с тем, что кубические сплайны, соответствующие таким кривым, позволяют описать функцию, имеющую непрерывную вторую производную в точках сопряжения двух соседних кривых, что является приемлемым для векторной графики. Однако в связи с конечными размерами объектов, дискретностью изображений и неточностью выделения контуров, для прикладных задач распознавания достаточно, чтобы в местах сопряжения аппроксимирующих кривых можно было получить просто гладкую функцию.

Кривые Безье обладают тем свойством, что если рассматривать опорные точки в обратном порядке, то кривая не изменится. То есть, для описания контура не важно направление обхода. С точки зрения нейронной сети, направление обхода контура тоже не имеет значения, но для выполнения второго условия, накладываемого на входы нейронной сети, при описании контуров всегда должен соблюдаться единый порядок обхода.

3.1. Гистограмма сегментов контура

Обозначим начальную и конечную опорные точки кривой Безье через P_0 и P_2 , а управляющую – как P_1 . Построим векторы $\vec{b} = \overrightarrow{P_0P_2}$, $\vec{c} = \overrightarrow{P_0P_1}$ и обозначим угол между ними как α . Два параметра, отношение норм $d = \|\vec{c}\|/\|\vec{b}\|$ и угол α , полностью определяют независимые от положения и масштаба характеристики кривой с тремя опорными точками. Аналогично, кривая Безье с числом управляющих точек N определяется $2N$ параметрами.

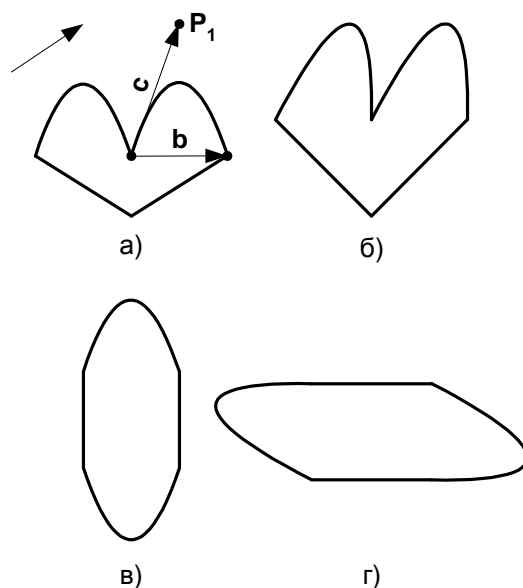


Рис. 1. Примеры контуров

С точки зрения признаков формы, принципиальное значение имеет не столько наличие определенной кривой в описании, сколько последовательность различных кривых при обходе контура. Этот факт иллюстрируется рисунком 1, где фигуры 1а), 1б) и 1в), 1г) образованы двумя парами идентичных кривых. На рис. 1а) показаны векторы \vec{b} и \vec{c} для одной из кривых Безье, а стрелка сверху указывает направление обхода контура.

Перенумеруем кривые последовательно в направлении обхода контура и введем два дополнительных параметра $g^i = \|\vec{b}^i\| / \|\vec{b}^{i+1}\|$ и $\beta^i = \angle(\vec{b}^i, \vec{b}^{i+1})$, где верхний индекс i – номер кривой в контуре. Тогда для точного описания замкнутого контура потребуется $4k$ параметров $(d^i, \alpha^i, g^i, \beta^i \quad i = \overline{1, k})$, где k – количество кривых Безье. Заметим, что данный вид кодирования по числу параметров не превышает классическую запись контура в кривых Безье и для точного восстановления размеров и положения контура достаточно задать координаты якорных точек одной из кривых.

Если зафиксировать количество кривых Безье, используемых при описании контуров, то предложенное кодирование полностью удовлетворяет всем рассмотренным выше требованиям. При этом любой контур будет содержать $4k$ независимых признаков и, соответственно, столько же входов будет иметь нейронная сеть. Однако на практике требование фиксированного количества кривых Безье в описании контура является слишком жестким и в большинстве случаев неприемлемым.

С другой стороны, в задачах классификации зрительных образов требуется не точное воспроизведение контура, а определение, к какому классу принадлежит объект. Для многих приложений контуры на рис. 1а) и 1б) могут считаться либо искаженным контуром одного и того же объекта (например, сфотографированного в разных ракурсах), либо контурами двух объектов, принадлежащих к одному классу. Контур на фигурах 1а) и 1б) отличаются тем, что выпуклые кривые имеют разное значение параметров d и α . Изменяя эти параметры в некотором небольшом диапазоне, мы будем оставаться в пределах одного класса объектов (контуров).

Следовательно, при классификации контуров параметры d, α, g и β можно рассматривать как дискретные. Углы имеют естественное ограничение $-\pi < \alpha, \beta < \pi$. При распознавании визуальных образов размеры контуров объектов ограничены максимальным размером изображения L сверху и шагом растра δ снизу, то есть, соотношения d и g линейных размеров также ограничены $d, g < L/\delta$.

Поскольку параметры представления кривых ограничены, то, задав шаг дискретизации по d и α , получим конечное число возможных типов кривых Безье в представлении произвольного контура. Пусть каждый из этих параметров принимает n значений, тогда получаем n^2 возможных типов кривой. Аналогично, если каждый из параметров g и β может принимать m значений, то получаем m^2 возможных комбинаций двух соседних кривых. При этих условиях число возможных

уникальных комбинаций двух соседних кривых будет $n^4 \cdot m^2$. Такую уникальную пару будем называть сегментом контура.

Если построить гистограмму сегментов, где элементом гистограммы служит номер сегмента, а значением элемента – количество таких сегментов в контуре, то получим код фиксированной длины, не зависящий от количества кривых Безье в описании контура.

Утверждение 1: Гистограмма сегментов однозначно задает контур объекта, кроме случаев, когда:

– контур содержит три одинаковых сегмента, не стоящих в описании непосредственно друг за другом, расстояния между ними равны, и перестановка участков контура, расположенных между этими сегментами, приводит к изменению контура (переставляемые участки не одинаковы);

– в описании контура существуют одинаковые пары не стоящих друг за другом сегментов, расстояния между сегментами пар равны, и перестановка участков контура, расположенных между этими сегментами, не меняет тип сегментов на краях переставляемых участков, но приводит к изменению контура.

Условие равенства расстояний позволяет поменять местами участки контура, ограниченные этими сегментами, но при этом делает приведенные ограничения очень слабыми. На практике не часто можно встретить объекты, которые отличались бы только переменной мест двух отличающихся друг от друга участков контура одинаковой длины и при этом относились бы к различным классам. Кроме того, эти исключения легко найти в процессе кодирования и сделать контуры различимыми путем введения дополнительного признака.

Утверждение 1 иллюстрируется рис. 2, где показаны контуры, имеющие пары одинаковых сегментов (одинаковые сегменты – прямой угол).

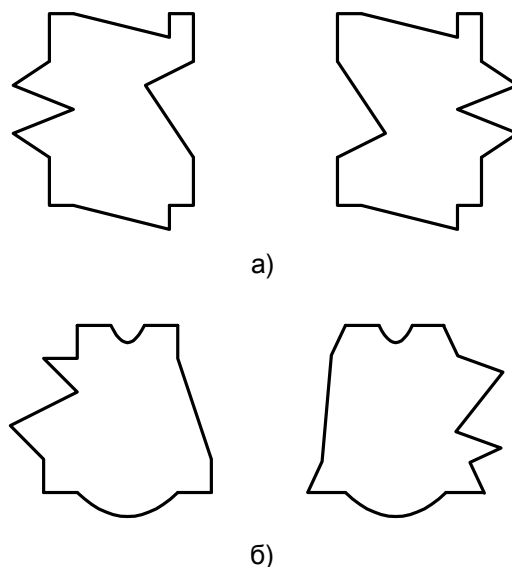


Рис. 2. Пример замены участков контура

На рис. 2а) показана ситуация, когда для обоих контуров гистограмма сегментов будет одинаковой, а на рис. 2б) – случай, когда перестановка участков контура изменяет гистограмму. В

случае 2б) меняется тип сегментов на краях переставляемых участков, а, следовательно, и гистограмма сегментов.

Проблема прямого использования гистограммы сегментов контура заключается в длине получаемого кода при малом шаге дискретизации. Так, например, уже при значениях настраиваемых параметров алгоритма гистограммирования n и m равных 10, количество элементов в гистограмме становится 10^6 , что неприемлемо для большинства приложений. Чтобы избежать этой проблемы, для классификации можно использовать упрощенные гистограммы, где параметры алгоритма подбираются исходя из конкретных условий задачи.

Если в качестве аппроксимирующих кривых выбрать отрезки прямой, то параметры d и α становятся ненужными. Выбирая достаточно короткие отрезки, можно описать произвольный контур с любой наперед заданной точностью. Конечно, такое описание становится чувствительным к масштабированию. Однако возможность линейной аппроксимации показывает, что параметры g и β более чувствительны к шагу дискретизации, чем d и α . То есть, для применений гистограммного кодирования можно выбирать $m > n$.

Существенное уменьшение длины гистограммы сегментов можно получить, накладывая дополнительное ограничение на кривые Безье, описывающие контур. Потребуем, чтобы для кривых выполнялись условия $-\pi/2 < \alpha < \pi/2$ и $d \cdot \cos(\alpha) < 1$. Геометрически это означает, что точка P_1 лежит внутри полосы, ограниченной параллельными прямыми, перпендикулярными \vec{b} и проходящими через точки P_0 и P_2 соответственно. Требование не жесткое, поскольку любая кривая Безье с тремя опорными точками, которая не удовлетворяет этому условию, может быть представлена двумя кривыми, которые данному условию удовлетворяют.

Используя введенные условия и заменяя параметры d и α одним – $A = d \cdot \sin(\alpha)$, заметим, что с точки зрения задач классификации, формы кривых с $A = const$ можно считать эквивалентными, или кривыми одного типа. Более того, во многих случаях будет достаточно ввести всего три типа кривых ($n = 3$): $\varepsilon < A$ – «выпуклые», $-\varepsilon < A < \varepsilon$ – «прямые» и $A < -\varepsilon$ – «вогнутые». Тут ε – некоторая малая величина, являющаяся настраиваемым параметром алгоритма кодирования. Значение ε выбирается при решении конкретной задачи и, в основном, зависит от точности построения описания контура и характеристик контуров классифицируемых объектов. С учетом замены двух параметров одним и считая $n = 3$, получаем длину гистограммы сегментов, равную $9 \cdot m^2$.

Дополнительное сокращение длины кода можно получить, используя гистограмму расстояний.

3.2. Гистограмма расстояний

При решении задач классификации нас не интересует точное восстановление контура и, следовательно, точное соотношение размеров соседних кривых Безье в описании. Уберем из определения сегмента контура параметр g . Тогда сегмент будет определяться тремя

параметрами: типами соседних кривых, определяемых величиной $A = d \sin(\alpha)$, и углом между соседними кривыми β .

В таком случае длина гистограммы сегментов будет $9 \cdot m$, однако теряется информация о характерных размерах объекта и некоторые контуры становятся не различимы. На рис. 3 даны примеры контуров, которые становятся не различимы в гистограмме сегментов при удалении параметра g . На рис. 3б) показаны аппроксимации окружности и эллипса с помощью четырех кривых Безье и выделены начальные точки P_0 этих кривых.

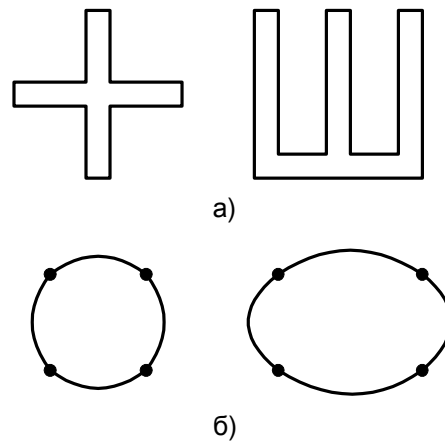


Рис. 3. Пример контуров, не разделяемых упрощенной гистограммой сегментов

Решить эту проблему позволяет гистограмма расстояний. Гистограмма расстояний $B = \{b_1, \dots, b_K\}$ $b \in N$ строится следующим образом:

- 1) находим максимальное R_{\max} расстояние между начальными опорными точками кривых Безье P_0 ;
- 2) задаем количество элементов гистограммы K и вычисляем шаг дискретизации $\Delta = R_{\max} / K$;
- 3) для всех попарных комбинаций кривых Безье находим

$$b_k = \sum_{i=1}^C a_i^k;$$

$$a_i^k = \begin{cases} 1 & \text{если } k-1 \leq \frac{r_i}{\Delta} < k, \\ 0 & \end{cases}$$

где C – количество попарных комбинаций; r_i – Евклидово расстояние между начальными точками i -й пары кривых.

Нетрудно проверить, что контуры, представленные на рис. 3, дают различные гистограммы расстояний. Кроме того, гистограмма расстояний позволяет сделать различимыми контуры, представленные на рис. 2а). К сожалению, не все контуры, попадающие в исключения

Утверждения 1, различимы при помощи гистограмм расстояний. Такая возможность зависит от симметрии расположения начальных точек кривых Безье в описании контура.

3.3. Построение гистограмм

Важнейшее значение при решении проблем распознавания визуальных образов имеет сохранение «меры близости» (конгруэнтности) похожих объектов. Однако, поскольку каждый элемент гистограммы с точки зрения алгоритма классификации рассматривается как независимый признак, при непосредственном использовании дискретизации параметров возникают ситуации, когда «похожие» объекты описываются различными признаками.

Рассмотрим пример построения гистограммы расстояний. Пусть расстояние между двумя начальными точками $r = \Delta$, где Δ – шаг дискретизации расстояний и ε – точность, с которой расстояние вычисляется (зависит от точности построения кривых Безье в описании контура). Тогда, в зависимости от знака ошибки кодирования, для одного и того же расстояния можем получить либо код $B = \{1, 0, \dots\}$, либо $B = \{0, 1, \dots\}$.

Чтобы избежать подобных ситуаций на границах областей дискретизации, можно ввести дополнительный параметр, степень перекрытия соседних областей θ , который удобно задавать в долях от шага дискретизации. Тогда определение переменной a_i^k на третьем шаге алгоритма построения гистограммы расстояний можно записать в виде

$$a_i^k = \begin{cases} 2 & \text{если } k + \theta \leq \frac{r_i}{\Delta} < k + 1 - \theta \\ 1 & \text{если } \left| \frac{r_i}{\Delta} - k \right| < \theta \quad \text{либо} \quad \left| \frac{r_i}{\Delta} - (k + 1) \right| < \theta. \\ 0 & \text{иначе} \end{cases}$$

Если известна точность, с которой составлено описание контура, то, выбирая $\theta = \varepsilon$ для случая, когда измеренное расстояние попадает на границу областей дискретизации, получим код $B = \{1, 1, \dots\}$. Внутри областей код будет $B = \{2, 0, \dots\}$ и $B = \{0, 2, \dots\}$. То есть, при кодировании с перекрытием один признак добавляет в гистограмму 2 единицы. Аналогично в случае, если один кодируемый признак определяется n параметрами, то каждый признак будет добавлять в гистограмму 2^n единиц.

Кодирование с перекрытием позволяет несколько уменьшить ошибку, вносимую дискретизацией, не изменяя длину получаемого кода. Увеличивая количество перекрытий с различными диапазонами, то есть, увеличивая число единиц, добавляемых в гистограмму на один шаг дискретизации, можно улучшить такую аппроксимацию.

Другой проблемой кодирования контуров, описанных кривыми Безье, является неоднозначность получаемого описания. В данном случае под неоднозначностью описания имеется ввиду тот факт, что участок контура, который описывается одной кривой, может быть представлен в виде нескольких последовательных одинаковых (или того же типа) кривых. Такая неоднозначность может проявляться при описании зашумленного контура или являться

следствием неточного выбора параметров алгоритма выделения кривых на контуре. При этом в гистограмму дают вклад «неинформативные» признаки, то есть такие, которые не являются определяющими для данного контура.

Проблема иллюстрируется рис. 4, где показаны три варианта последовательного расположения двух кривых одного типа (на рисунке выделены точки P_0). Причем, на рис. 4а) участок контура не может быть описан одной кривой, то есть обе кривых являются информативными, а на рис. 4б) и 4в) показаны кривые, которые разбиты на две в результате неточности описания контура, причем кривые 4б) описывают «выпуклую» кривую, а 4в) – отрезок прямой.

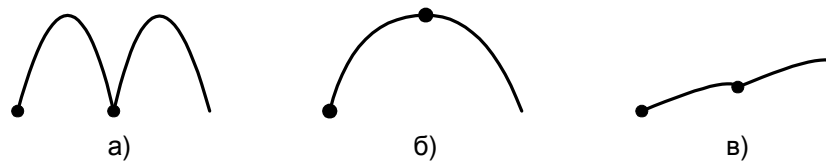


Рис. 4. Примеры двух соседних кривых в описании контура

Неточности в кодировании отрезков прямой вызваны тем, что отрезки прямых контура, непараллельные границам растрового изображения, выглядят как серия «штрихов». При описании контура кривыми Безье необходимое требование – совпадение конечной точки предыдущей и начальной точки последующей кривой. Поэтому при построении описания наклонных прямых управляющая точка P_1 не лежит внутри отрезка, образованного начальной P_0 и конечной P_2 точками кривой, описывающей «штрих».

Рассматриваемая неоднозначность частично устраняется условием минимального размера, накладываемым на алгоритм описания контуров кривыми Безье. При построении гистограмм данная проблема устраняется предварительным анализом и коррекцией кривых, составляющих контур.

Неоднозначность описания контура при построении диаграмм можно частично устранить, используя два настроечных параметра: ε – определяющий «меру схожести» соседних кривых и γ – малый угол, определяющий допустимую разность направлений. Придерживаясь обозначений $\vec{b} = \overrightarrow{P_0P_2}$, $\vec{c} = \overrightarrow{P_0P_1}$, $d^i = \|\vec{c}^i\|/\|\vec{b}^i\|$ и $\alpha^i = \angle(\vec{b}^i, \vec{c}^i)$, где верхний индекс – номер кривой в описании, можем определить правила коррекции контура:

1. Если в описании контура имеем $|d^i \cdot \sin(\alpha^i)| < \varepsilon$ и $|d^{i+1} \cdot \sin(\alpha^{i+1})| < \varepsilon$, случай двух отрезков прямых и их направления совпадают: $\angle(\vec{b}^i, \vec{b}^{i+1}) < \gamma$, то проводим замену $P_2^i = P_2^{i+1}$. При этом кривая с номером $i+1$ исключается из описания контура. То есть, два последовательных отрезка одного направления заменяются одним. При записи в параметрической форме точка P_2^i нужна, чтобы вычислить новое значение параметра g , при этом можно считать $\alpha^i = 0$.

2. Если в описании контура имеем $\varepsilon < |d^i \cdot \sin(\alpha^i)|$ и $\varepsilon < |d^{i+1} \cdot \sin(\alpha^{i+1})|$, случай двух дуг и они могут быть заменены одной дугой, то есть, выполняются условия $|\angle(\overrightarrow{P_1P_2}^i, \overrightarrow{P_0P_1}^{i+1})| < \gamma$ и $|\angle(\overrightarrow{P_0P_1}^i, \overrightarrow{P_1P_2}^{i+1})| < \pi - \gamma$, то для вычисления параметра g считаем $P_2^i = P_2^{i+1}$ и исключаем кривую $i+1$ из описания контура. Новую точку P_1^i можно найти как пересечение прямых, заданных парами точек P_0^i, P_1^i и P_1^{i+1}, P_2^{i+1} , однако для построения упрощенной гистограммы сегментов можно просто считать, что кривая осталась того же типа.

Коррекция описаний контуров существенно улучшает гистограммы, однако полностью не устраняет неточности кодирования. Таким образом, полученные гистограммы контуров содержат некоторый уровень «шума» и распознавание контуров простейшими методами, например, сравнением с шаблонами гистограмм, затруднено. Нейронные сети позволяют эффективно выделять полезную информацию при наличии шумов в коде, поэтому в качестве классификатора удобно использовать, например, многослойный персептрон [3].

При использовании гистограмм в качестве входного вектора нейронной сети необходимо учитывать, что большинство алгоритмов обучения работают устойчиво, если входные данные приведены к определенному диапазону значений. Чаще всего используется диапазон $[0;1]$. Чтобы привести элементы гистограммы к заданному диапазону, обычно используется нормировка

$$b_i = \frac{b_i - b_{\min}}{b_{\max} - b_{\min}}, \quad i = \overline{1, K},$$

где K – число элементов гистограммы, а b_{\max} и b_{\min} – максимальное и минимальное значения соответственно. Максимум и минимум для элементов гистограммы определяются условиями решаемой задачи. Если абсолютное количество появлений данного признака в коде объекта, например, количество одинаковых расстояний, не является информативным для решения задачи, то экстремумы вычисляют для каждого контура отдельно (нормировка вдоль вектора):

$$b_{\max} = \max_i(b_i), \quad b_{\min} = \min_i(b_i), \quad i = \overline{1, K}.$$

В том случае, если частота встречаемости информативна, экстремумы вычисляют для всего набора контуров N , которые используются в обучающей выборке (нормировка по выборке):

$$b_{\max} = \max_{i,j}(b_{ij}), \quad b_{\min} = \min_{i,j}(b_{ij}), \quad i = \overline{1, K}, \quad j = \overline{1, N}.$$

Следует отметить, что нормировка по полному массиву обучающей выборки имеет тот недостаток, что при независимом тестировании могут появляться контуры с элементами, значения которых превосходят максимальное для обучающей выборки.

4. Экспериментальная проверка

Экспериментальное исследование предложенного кодирования осуществлялось на модельной задаче, в которой распознавались контуры для десяти классов плоских фигур. Фигуры располагались на рисунке произвольным образом, имели различный размер и ориентацию.

Исходными данными исследования явились контуры фигур, представленные в виде списков кривых Безье. Выбор фигур, выделение и описание контуров кривыми Безье проведены в отделе №235 ИПММС НАНУ.

На рис. 5 показаны примеры представленных для распознавания видов фигур, где под каждой из них проставлен класс, к которому она относится. Для фигур 0-го и 9-го классов представлены по два примера. Рис. 5а) содержит исходные растровые изображения фигур, а рис. 5б) – контуры этих же фигур, восстановленные по описаниям контуров кривыми Безье. Контуры фигур, относящиеся к классам 3 и 7, совпадают с контурами, приведенными на рис. 3а), поэтому для решения поставленной задачи использовалась как гистограмма сегментов, так и гистограмма расстояний.

В связи с наличием неточностей в описаниях перед построением гистограмм была проведена коррекция контуров по алгоритмам, описанным в п.3.3, и к полученным гистограммам была применена нормировка вдоль вектора.

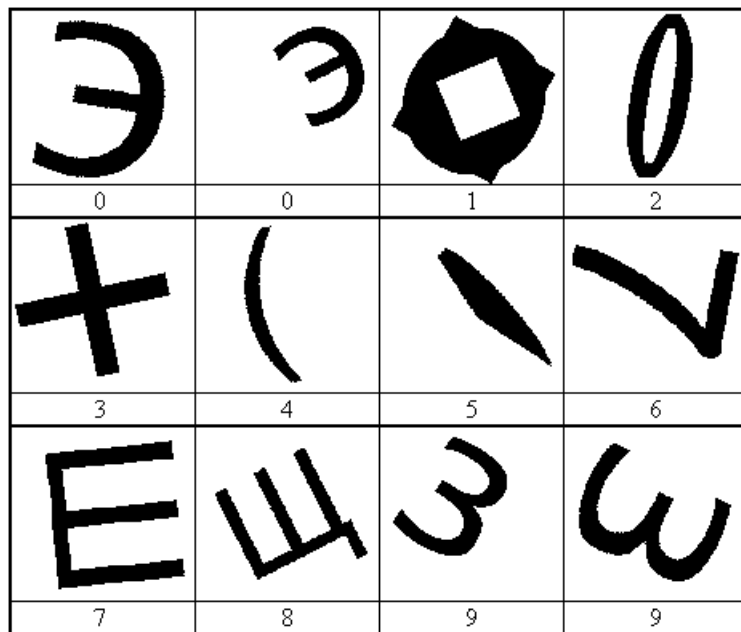
В экспериментах использовалась упрощенная гистограмма сегментов, в которой каждый сегмент описывался тремя параметрами A^i, A^{i+1} и β . Причем параметр A принимает три значения: “выпуклый”, “прямой” и “вогнутый”, а угол между кривыми принимает 8 значений, то есть шаг дискретизации был выбран $\pi/4$. Для угла β использовалось кодирование с перекрытием, где параметр перекрытия выбирался $\theta = \pi/18$. То есть на каждый сегмент контура в гистограмму добавлялись 2 единицы.

Поскольку описания в виде кривых Безье позволяют привести исходные контуры к единому масштабу, что иллюстрируется рис. 5, то гистограмма расстояний строилась с шагом дискретизации $0,1 \cdot L$, где L – максимальный размер контура по вертикали. Таким образом, полный вектор, описывающий один контур, содержит 82 элемента, где 72 – длина гистограммы сегментов и 10 – гистограммы расстояний.

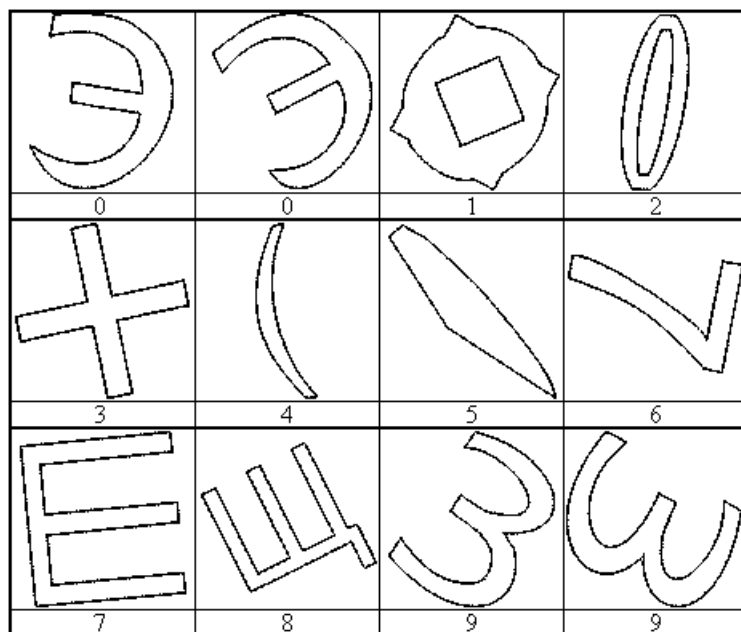
Распознавание рассматриваемых классов фигур производилось при помощи нейросетевого классификатора. В качестве классификатора применялся многослойный персептрон с правилом обучения Extended Delta-Bar-Delta [3]. В связи с достаточно высокой точностью кодирования и простым характером контуров персептрон содержит всего один слой нейронов. Таким образом, архитектура нейронной сети была 82x10, где первая цифра – число входов, а вторая – количество нейронов, равное числу распознаваемых классов.

Для экспериментальной оценки возможности и эффективности кодирования при помощи гистограмм исходный массив данных был разбит случайным образом на обучающую и тестовую выборки. В обучающую выборку было отобрано 80% исходных данных, а в тестовую – 20%. Все эксперименты выполнялись в среде САПР МНН [4].

Для статистической достоверности полученных результатов разбиение на обучающую и тестовую последовательности проводилось 5 раз, с выделением в тест различных образов. Нейронная сеть обучалась на тренинговой последовательности, а затем тестировалась.



а)



б)

Рис. 5. Примеры исходных фигур а) и их контуров, построенных по кривым Безье б)

Поскольку общее число фигур в эксперименте было относительно мало – 230 образов, методы повышения обобщающих свойств сети, такие как «периодическое сохранение лучшей», не использовались. Во всех 5 экспериментах сеть показала 100% распознавание на обучающих выборках. На тестовых выборках в одном из пяти экспериментов была получена всего одна ошибка, в остальных случаях результат классификации равнялся 100%.

На рис. 6 показаны ошибочно распознанный контур и образец класса, к которому этот контур был отнесен нейронной сетью.

Анализ результатов тестирования показывает, что существуют две причины полученной ошибки распознавания. Первая – это неточность при описании контура кривыми Безье. В данном случае неточность описания контура не может быть устранена методами, предложенными в п. 3.3. Вторая причина – ограниченный размер обучающей выборки.

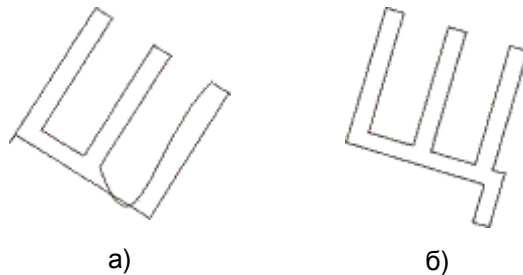


Рис. 6. Ошибочно распознанный контур а) и класс, к которому он был отнесен б)

Этот факт подтверждается тем, что, во-первых, весь набор данных может быть запомнен простейшей сетью, которая использовалась в эксперименте, и, во-вторых, дополнительные эксперименты показали, что данный контур распознается правильно при другом выборе образцов в обучающую последовательность.

5. Выводы

1. Предложено гистограммное кодирование контуров, представленных кривыми Безье. Определены ограничения, накладываемые на кривые, позволяющие существенно сократить длину получаемого кода.

2. Экспериментальная проверка разработанного метода кодирования показала эффективность метода представления кривых и алгоритмов построения гистограмм.

3. Предложенные гистограммы, разделяющие кривые Безье по типам сегментов и расстояниям между опорными точками, являются достаточными для решения поставленной задачи и могут быть эффективно использованы при распознавании зрительных образов.

Развитием данного направления будет использование гистограммного кодирования контуров в прикладных задачах, таких как оптические системы классификации движущихся объектов в системах дорожной безопасности и других.

СПИСОК ЛИТЕРАТУРЫ

1. Роджерс Д., Адамс Дж. Математические основы машинной графики: Пер. с англ. – М.: Мир, 2001. – 604 с.
2. Duda R.O., Hart P.E., Stork D.G. Pattern Classification. – 2nd ed. – New York: John Wiley & Sons, Inc., 2001. – 654 p.
3. Haykin S. Neural Networks: A Comprehensive Foundation. – 2-nd ed. – Englewood Cliffs, NJ. – Prentice-Hall, 1999.
4. Система проектирования модульных нейронных сетей САПР МНС / Резник А.М., Кусскуль М.Э., Сычев А.С., Садовая Е.Г., Калина Е.А. // Математические машины и системы. – 2002. – № 3. – С. 28 – 37.