РОССИЙСКАЯ АКАДЕМИЯ НАУК
СИБИРСКОЕ ОТДЕЛЕНИЕ
АВТОМЕТРИЯ
№ 1 1996

УДК 007.001.362 : 681.3.019

Н. Г. Федотов, А. А. Кадыров
(Пенза)

НОВЫЕ МЕТОДЫ ФОРМИРОВАНИЯ ПРИЗНАКОВ РАСПОЗНАВАНИЯ ОБРАЗОВ С ПОЗИЦИЙ СТОХАСТИЧЕСКОЙ ГЕОМЕТРИИ

Проанализировано применение методов стохастической геометрии в распознавании образов. Рассуждения проводились на базе введенного авторами Trace-преобразования исходных изображений в изображения на листе Мебиуса. Предложен новый подход к формированию признаков распознавания, не зависящих от движений изображений, а также от их линейных деформаций. Отличительной чертой группы рассматриваемых признаков является представление каждого из них в виде последовательной композиции трех функционалов.

В распознавании образов традиционно выделяют формирование признаков и решающую процедуру. В кибернетической литературе исторически сложилось так, что подавляющее большинство работ по распознаванию образов посвящено решающим правилам и практически нет работ по формированию признаков. По общепринятому мнению это объясняется тем, что процесс формирования признаков является эмпирическим и зависит от интуиции проектировщика распознающей системы.

Подход с позиций стохастической геометрии, развитый в [1], позволяет восполнить этот пробел и, наряду с конструктивной теорией признаков, дать практические методы генерации большого числа новых признаков распознавания изображений. Столь мощное смещение акцента с решающих правил на новые признаки распознавания сближает данный подход с нейрокомпьютингом.

В работе [1] предложено в качестве признаков распознавания изображений использовать вероятности геометрических событий, под которыми понимают результат взаимодействия геометрических объектов: пересечения, покрытия и т. п. Роль геометрических объектов выполняют, с одной стороны, сложные траектории сканирования со случайными параметрами (отрезки, линии, кривые, фигуры и т. п.), с другой — фрагменты распознаваемого изображения. Рассматривается структура подобных распознающих систем, примеры конкретных технических реализаций. Рассмотрены также возможные расширения базисного метода распознавания, основанного на стохастической геометрии. Одно из расширений связано с усложнением наблюдений случайного события — пересечения линии развертки с изображением, т. е. с применением более сложных признаков распознавания.

В статье представлены начала новой теории формирования признаков распознавания, не зависящих от движений изображений, а также от аффинных преобразований. Отличительной чертой группы рассматриваемых признаков является представление каждого из них в виде последовательной композиции трех функционалов.

Рассмотрим входную сетчатку распознающего устройства, под которой будем понимать сканируемую часть плоскости изображения. В этой части плоскости располагается некоторое изображение, тогда как оставшаяся часть плоскости — фоновая. Таким образом, изображение финитно. Рассмотрим случайную прямую l, которая может пересекать изображение. Предположим, что пересечение прямой l и изображения позволяет нам вычислить некоторое число g, характеризующее их взаимное расположение. Производя серию случайных бросаний прямой l на плоскость, получаем выборку для случайной величины g. Далее, можно определить какую-нибудь эмпирическую характеристику n случайной величины g. Описанную процедуру следует реализовать в радиоэлектронной системе, осуществляющей распознавание изображений [1].

Математическая сторона рассмотренной процедуры интенсивно исследовалась в стохастической геометрии. Было выяснено, что при некоторых условиях характеристика n может иметь явный геометрический смысл. Для нас важно, что, легко реализуясь в устройствах, эта идея может служить исходной точкой для получения новых признаков распознавания образов как в теоретическом анализе, так и в практической сфере.

В [1] приводятся формулы, на основе которых строятся критерии распознавания. Рассматриваются только бинарные изображения (черные фигуры на белом фоне).

1. Рассмотрим изображение в виде кусочно-дифференцируемой кривой, которая может быть границей фигуры. Пусть g — число пересечений этой кривой со случайной прямой l. Тогда математическое ожидание Mg пропорционально длине кривой.

2. Рассмотрим изображение в виде выпуклой фигуры. Это может быть выпуклая оболочка некоторой другой фигуры. Пусть g — длина пересечения выпуклой фигуры со случайной прямой l. Тогда средние величины Mg^0, Mg, Mg^2 пропорциональны соответственно периметру, площади и собственному потенциалу однородного слоя.

Trace-преобразование. Приведенные выше формулы и их многочисленные аналоги имеют для распознавания образов следующие недостатки: 1) число этих формул ограничено, поскольку ясно выраженных геометрических характеристик не так много, а признаков требуются тысячи и более; 2) формулы применимы только для бинарных изображений. К достоинствам следует отнести возможности параллельных вычислений (одновременно обрабатывается несколько прямых сразу) и стохастической реализации, последнее позволяет оборвать процесс при достижении нужной точности, кроме того, вычисленные признаки не зависят от движений объектов. Известно, что обычно признаки сильно зависят от поворота и сдвига объекта, в то время как во многих задачах распознавания поворот и сдвиг объектов совершенно неинформативны.

В данной работе мы предлагаем обобщение приведенного выше подхода с целью преодоления его недостатков и с сохранением достоинств, причем это обобщение в некотором смысле полное.

Обозначим буквой F финитное изображение. Если дана прямая l, то число g, характеризующее взаимное расположение прямой l и изображения, будем вычислять согласно некоторому правилу T: $g = T(l, F)$; отображение T будем называть функционалом. Для нас желаемым свойством является независимость вычислений от движения объекта, поэтому единственное требование, которое мы накладываем на T, формулируется следующим образом. Пусть изображение претерпело сдвиг и поворот, при этом возникло новое изображение F'. При этом же сдвиге и повороте прямая l перейдет в прямую l', оставаясь, таким образом, «вмороженной» в изображение. Требуется, чтобы $T(l, F) = T(l', F')$. Это равенство должно быть верным для всех прямых и всех допустимых изображений. Такое свойство назовем полной инвариантностью функционала T. Следует отметить, что понятие полной инвариантности весьма сильно расширяет возможности распознавания образов, ибо это не обязательно число пересечений, длина секущей и т. д. Например, если изображение цветное, переменной яркости, то таких функционалов можно найти довольно много. Итак, круг функционалов и обрабатываемых изображений значительно расширен.

Аналогично, как и в стохастической геометрии, определена случайная величина $g = T(l, F)$, распределение которой не зависит от сдвигов и поворотов изображения. Поэтому числовые характеристики этой случайной величины опять могут служить признаками изображений, которые определяются специальными техническими устройствами и системами. Недостаток нового семейства признаков — первоначальное отсутствие ясного геометрического смысла, и заранее не известна их различающая сила. Однако для распознавания образов это не так важно, ибо решающей все-таки является экспериментальная проверка.

Отметим еще одно свойство вполне инвариантного функционала T (Trace): он не обязательно определяется лишь сечением прямой изображения. Для его вычисления может быть привлечена также и другая информация, например, свойства окрестности этого сечения.

Чтобы понять, что предложенное обобщение в некотором смысле исчерпывает все его возможности, изложим теорию Trace-преобразований (или Tr-преобразований). Прямая l, если введены полярные координаты на плоскости, характеризуется расстоянием p от начала координат до нее и углом φ (с точностью до 2π) ее направляющего вектора:

$$l = \{(x, y)\colon x\cos\varphi + y\sin\varphi = p\}, \quad l = l(\varphi, p),$$

где x, y — декартовы координаты на плоскости. Если позволить параметру p принимать также и отрицательные значения, то

$$l(\varphi, p) = l(\varphi + \pi, -p).$$

Таким образом, множество всех направленных прямых, пересекающих круг радиусом R с центром в начале координат («сетчатку»), однозначно параметризуется множеством

$$M = \{(\varphi, p)\colon 0 \le \varphi \le \pi,\ -R \le p \le R\}$$

при условии, что параметры $(0, p)$ и $(\pi, -p)$ задают одну прямую. Видно, что множество прямых на сетчатке есть в топологическом смысле не что иное, как лист Мебиуса [2]. Множество чисел $T(l(\varphi, p), F)$, зависящее от точки на листе Мебиуса M, есть некоторое преобразование изображения, которое назовем Tr-преобразованием. Если, например, при численном анализе Tr-преобразование представлено матрицей, то будем называть ее Tr-матрицей. Если направить ось 0φ горизонтально, а ось $0p$ вертикально, то в точке φ_j, p_i будет расположен элемент матрицы с номером (i, j), т. е. значение $T(l(\varphi_j, p_i), F)$. Здесь φ_j, p_i — некоторые значения равномерных дискретных сеток на указанных осях. Матрица будет 2π-периодична в направлении горизонтальной оси, причем через каждый интервал длины π столбцы ее переворачиваются.

Будем считать дополнительно, что если прямая l не пересекает изображения, то $T(l, F)$ есть заданное число (например, 0), или другой фиксированный элемент, если функционал T нечисловой. В этом случае первоначальному изображению F соответствует $\mathrm{Tr}(F)$ — новое изображение (можно трактовать $T(l(\varphi, p), F)$ как изображение, характеристики которого в точке (φ, p)) — его Tr-образ.

Заметим, что известное преобразование Радона [3] может рассматриваться как пример Tr-преобразования.

Кратко остановимся на том, как меняется изображение $\mathrm{Tr}(F)$ при сдвигах и вращениях исходного изображения F. Если первоначальное изображение поворачивается, то его Tr-образ сдвигается по горизонтальной оси 0φ. Если же происходит сдвиг исходного изображения на некоторый вектор, то его Tr-образ претерпевает следующие преобразования. Лучше их изложить в терминах Tr-матриц. Столбцы остаются неизменными, на своих местах, но могут сдвигаться вверх или вниз. Вектор сдвига определяет числа a и b такие, что

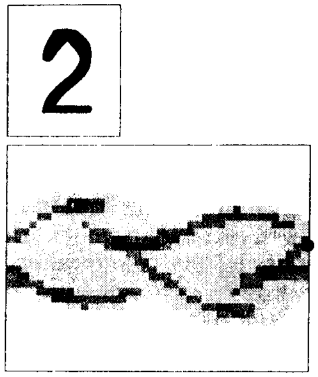

Рис. 1

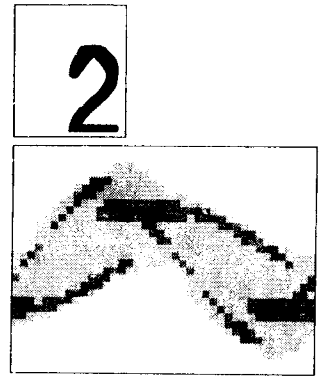

Рис. 2

столбец с координатой φ_i сдвигается в вертикальном направлении на $a \cdot \cos(\varphi_i - b)$. Следует отметить, что вполне строгим это описание будет лишь в том случае, если T-матрицу считать непрерывной, т. е. i и j — непрерывные параметры. На рис. 1 представлены изображение цифры «2» и результат его Tr-преобразования — Tr-трансформанта, на рис. 2, 3 — изображение цифры «2» при сдвигах и соответствующие Tr-трансформанты, на рис. 4 — изображение цифры «2», претерпевшее масштабное преобразование, и соответствующая ему расширенная Tr-трансформанта, на рис. 5 — повернутое изображение этой цифры и ее Tr-трансформанта, сдвинутая по горизонтальной оси.

Обычная евклидова мера $d\varphi dp$ листа Мебиуса инвариантна к указанным преобразованиям, поэтому плотность распределения всякой функции, заданной на листе Мебиуса, в данном случае функции изображения $\mathrm{Tr}(F)$, не зависит от указанных преобразований, т. е. если изображение F сдвинуто и повернуто до состояния F', то распределения значений функций изображений $\mathrm{Tr}(F)$ и $\mathrm{Tr}(F')$ одинаковы. Именно поэтому их значения могут трактоваться как случайные функции, не зависящие от движений исходного изображения. Этим доказано, что при данном выше обобщении признаков, действительно, сохраняется инвариантность.

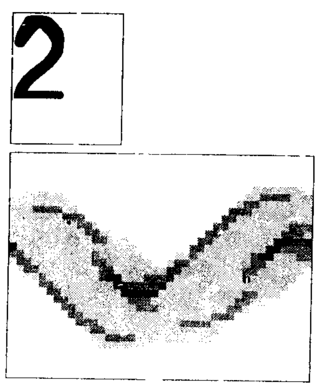

Рис. 3

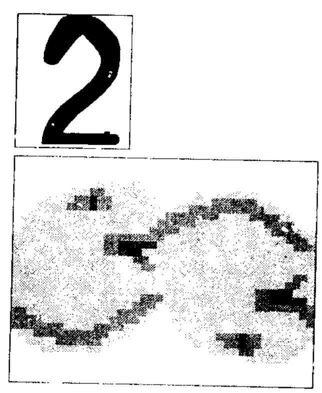

Рис. 4

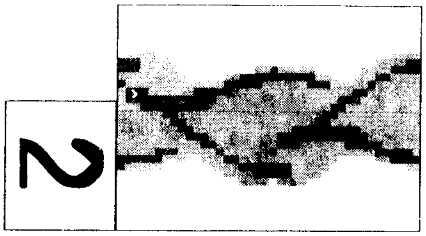

Рис. 5

Триплексные признаки. Рассмотрим формирование триплексных признаков, представляющих последовательную композицию трех функционалов:

$$\Pi(F) = \Phi \circ P \circ T(F \circ l(\varphi, p, t)).$$

Каждый функционал (Φ, P и T) действует на функции одной переменной (φ, p и t) соответственно. Для каждого из трех функционалов легко можно придумать десятки разных конкретизаций, удовлетворяющих требуемым условиям. Следовательно, сразу получаем тысячи новых признаков, инвариантных к движениям. Для распознавания 2π объектов требуется порядка n признаков, следовательно, мы получаем возможность распознавать очень большое число изображений, например идеограмм.

Функционал T, соответствующий Tr-преобразованию, подробно рассмотрен выше. В дискретном варианте вычислений результат этого преобразования, или Tr-трансформанта $T(F \circ l(\varphi, p, t))$, представляет собой матрицу, элементами которой являются, например, значения яркости изображения F на пересечениях со сканирующей линией $l(\varphi, p)$. Параметры сканирующей линии p определяют позицию этого элемента в матрице. Последующее вычисление признака заключается в последовательной обработке столбцов матрицы с помощью функционала P, а затем в преобразовании полученной периодической функции с помощью функционала Φ в число-признак $\Pi(F)$.

Рассмотренные триплексные признаки распознавания могут быть вычислены в высшей степени параллельном процессе. Подобно признакам, формируемым нейронными сетями, данные признаки не имеют наперед заданного смысла, их отбор осуществляется в ходе машинного эксперимента, принимая во внимание исключительно лишь их полезность для классификации.

СПИСОК ЛИТЕРАТУРЫ

1. **Федотов Н. Г.** Методы стохастической геометрии в распознавании образов. М.: Радио и связь, 1990.
2. **Масси У., Столлингс Дж.** Алгебраическая топология. Введение. М.: Мир, 1977.
3. **Хелгасон С.** Преобразование Радона. М.: Мир, 1983.

Поступила в редакцию 30 октября 1995 г.