

Р. А. Завьялов, Н. С. Сенюшкин

АВТОМАТИЧЕСКОЕ РАСПОЗНАВАНИЕ ОБРАЗОВ БОРТОВЫМИ ОПТИЧЕСКИМИ СИСТЕМАМИ МНОГОЦЕЛЕВЫХ КОМПЛЕКСОВ МОНИТОРИНГА

В статье описана методика работы и алгоритм распознавания объектов расположенных на контролируемом пространстве, подлежащему мониторингу. Дана постановка задачи, показан метод потенциальных функций, а так же его модификация под поставленную задачу. Приведен практический пример его применения. *Мониторинг; распознавание объектов; метод потенциальных функций*

1. ПОСТАНОВКА ЗАДАЧИ

В общем случае распознавание образов – это отнесение исходных данных к определенному классу с помощью выделения существенных признаков, характеризующих эти данные, из общей массы несущественных данных. Под распознаванием оптическими системами объектов подразумевается имитация машиной свойственного человеку умения классифицировать воспринимаемые ощущения.

Человек, воспринимая явления внешнего мира, классифицирует их, т. е. разбивает на классы «похожих», но не тождественных явлений, наделяя каждый класс специальным наименованием. Если иметь ввиду зрительные ощущения, то примерами такого рода классов могут служить классы: «портрет», «пейзаж», «рукописная буква а», «окружность», «диск», «печатный текст» и т. д.

Предположим, что некоторая классификация зрительных восприятий произведена; например, просматривая изображения, человек может выделить среди них «портреты» и «пейзажи». Это значит, что такое выделение этот человек может производить не только по отношению к ранее виденным объектам (в данном примере – изображениям), но и по отношению к объектам, ранее не виденным. Именно благодаря этому человек умеет читать текст, написанный разными почерками или напечатанный не встречавшимся ранее шрифтом, а так же опре-

делять что пред ним именно автомобиль вне зависимости от его марки.

Выделяют четыре основные группы методов распознавания: прямое сравнение нормализованного изображения с имеющимся образом; статистический метод – строится распределение для каждого класса и классифицируется по правилу Бейеса, обучение производится с помощью тренировочной коллекции; нейронная сеть – настройка ведется формированием структуры сети и подбором коэффициентов, чем больше будет отработано образов, тем выше результат распознавания; структурные и синтаксические методы: строится правило в зависимости от вхождения / не вхождения конкретных элементов [1].

2. МЕТОД ПОТЕНЦИАЛЬНЫХ ФУНКЦИЙ

Задача распознавания образов, помимо прочих, имеет и вероятностную постановку, которая связана с восстановлением некоторой функции, характеризующей вероятность принадлежности объекта тому или иному образу. В связи с этим, на взгляд авторов, наиболее перспективным для применения в автономных комплексах мониторинга является метод потенциальных функций. Его преимуществами является минимальная потребность в вычислительных ресурсах и оперативной памяти (в отличие от нейронных сетей), высокая скорость и точность работы (в отличие от методов прямого сравнения образов).

Рассмотрим случай, когда имитируется процесс обучения разделению объектов на классы, когда показываются примеры объектов и сообщается точная информация о том, к какому классу они принадлежат. В геометрической ин-

Контактная информация: 8(347)273-79-54

Работа выполнена при поддержке ФЦП «Научные и научно-педагогические кадры инновационной России»

терпретации – в некотором пространстве X каждому объекту соответствует точка; классам объектов в этом пространстве соответствуют непесекающиеся области; задача сводится к построению по показываемым точкам и по сообщаемой о них информации такой поверхности, которая разделяет эти области, т. е. функции, принимающей положительные значения на точках из одной области и отрицательные – на точках из второй области [2, 3].

Введем в рассмотрение функцию двух переменных $K(x, y)$, где x и y – точки из пространства X . Если зафиксировать точку y , положив ее $y = x^*$, то функция $K(x, y^*)$ станет функцией точки пространства X и будет зависеть от того, как выбрана точка x^* . Примером подобной функции в физике является потенциал, определенный для любой точки пространства, но зависящий от того, где расположен источник потенциала. Имея ввиду эту аналогию, назовем функцию $K(x, y)$ потенциальной функцией [4].

Чтобы упростить наглядную интерпретацию задачи (рис. 1), будем представлять себе, что в пространстве X как-либо введено расстояние между точками и что в качестве потенциальной функции $K(x, y)$ выбрана некоторая функция, удовлетворяющая следующим условиям:

- функция $K(x, y)$ всюду положительна,
- она убывает при удалении точки x от $y = x^*$, т. е. в частности, при фиксированном x^* достигается максимума при $x = x^*$.

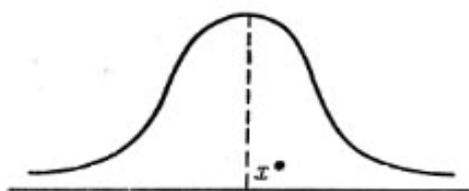


Рис. 1. График потенциальной функции

Удобно представить себе, что K есть функция расстояния $\rho(x, y)$ между точками x и y , т. е. $K = K[\rho(x, y)]$. Например, можно положить $K = e^{-\alpha\rho}$ или $K = \frac{1}{1 + \alpha\rho^2}$, где $\alpha > 0$ – постоянная и т. д.

Рассмотрим теперь следующую процедуру, которая сама по себе дает решение задачи лишь в простейших случаях и не пригодна для случаев сколь-либо сложных, но удобна для иллюстраций излагаемых далее идей.

Пусть надо научиться относить точки к одному из двух классов, которые условно назовем

A и B . Предположим, что задана точка $x = x^1$ и сообщено, что она принадлежит к классу A . Примем точку $x = x^1$ за «источник потенциала», положив $x^* = x^1$, т. е. построим «холм» с вершиной в этой точке и запомним, что этот холм относится к точке из A . При предъявлении следующих точек x^s из A или из B , каждый раз строятся подобные же «холмы» с вершинами в показанных точках и запоминается, к какому классу, A или B , этот холм принадлежит.

Сложим отдельно потенциалы, которые были построены над точками, принадлежащими классу A , и над точками, принадлежащими классу B , т. е. построим функции

$$K_A(x) = \sum_{x^s \in A} K(x, x^s);$$

$$K_B(x) = \sum_{x^s \in B} K(x, x^s).$$

Используя нашу геометрическую иллюстрацию, можно сказать, что холмы, построенные над показанными точками из A , «складываются» и образуют «гору» над областью, где расположены эти точки. Аналогично выстраивается «гора» над областью, где расположены точки B .

Таким образом, в результате процедуры оказались построенными две функции, $K_A(x)$ и $K_B(x)$, которые можно назвать потенциалами образов A и B .

Теперь, после окончания процесса обучения, начинается «экзамен», т. е. предъявляются новые точки и требуется дать ответ на вопрос «к какому классу они относятся?». В методе потенциальных функций предлагается относить показанную точку $x = \bar{x}$ к A , если

$$K_A(\bar{x}) > K_B(\bar{x})$$

и к B при обратном знаке неравенства.

Естественность такой процедуры оправдывается следующими интуитивными соображениями. Если области A и B компактны, и если потенциальные функции достаточно быстро убывают с расстоянием, то можно ожидать, что значения $K_A(x)$ и $K_B(x)$ будут большими для точек «своего» образа, отмеченного индексом, и меньшими для другого образа.

Вводя функцию

$$\Phi(x) = K_A(x) - K_B(x),$$

замечаем, что она положительна над точками из A и отрицательна над точками из B , т. е. разделяет знаком множества A и B .

Если надо разделить объекты не на два, а на большее число классов, то можно совершенно

аналогично построить потенциалы для всех образцов порознь и при появлении в процессе экзамена новой точки относить ее к тому образцу, чей потенциал в этой точке наибольший.

3. АДАПТАЦИЯ МАТЕМАТИЧЕСКОЙ МОДЕЛИ К ПОСТАВЛЕННОЙ ЗАДАЧЕ

Для большей наглядности рассмотрим черно-белое изображение. Если изображение цветное то его надо сделать монохромным, тем самым повысив качество распознавания объекта, так в этом случае контрастность границ максимальная.

Пусть изображение состоит всего из двух пикселей. Тогда множество всех объектов, которое можно будет изобразить (универсальное множество), состоит из четырех объектов: (0,0), (0,1), (1,0), (1,1), где 1 — черный пиксель, 0 — белый (рис. 2).

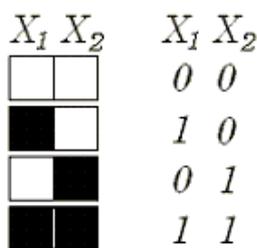


Рис. 2

Все объекты универсального множества можно разместить в вершинах единичного квадрата, таким образом, множеству фигур, изображенных на двухпиксельном поле, может быть сопоставлено множество точек в двумерном пространстве. Ребру этого квадрата будет соответствовать переход от одного изображения к другому. Для перехода от (1,1) к (0,0) нужно будет пройти два ребра, для перехода от (0,1) к (0,0) — одно (рис. 3). Отметим, что число ребер в нашем переходе — это количество несовпадающих пикселей двух изображений. Отсюда можно сделать вывод, что расстояние от одного рисунка до другого равно числу несовпадающих пикселей в них. Это расстояние называется расстоянием по Хэммингу [5].

Теперь представим себе, что у нас рисунок состоит из трех пикселей. Коды изображений тогда будут состоять из трех значений, универсальное множество — из восьми элементов, которые разместим в вершинах единичного куба. Но принципиально ничего не изменится, и расстояние по Хэммингу вычисляется так же.

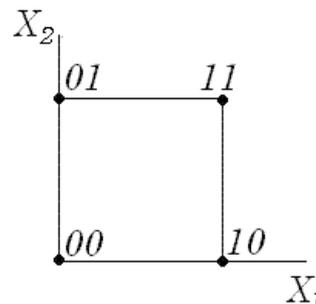


Рис. 3

В программе, демонстрирующей рассматриваемый метод, используется рисунок $200 \times 200 = 40\,000$ пикселей. Нетрудно заметить, что в этом случае код любого изображения состоит из 40 000 значений, универсальное множество — из $2^{40\,000}$ элементов, которые будем размещать в вершинах единичного 40 000-мерного куба. Представить себе такой 40 000-мерный куб нелегко, но смысл от этого не меняется абсолютно. Основная идея заключается в том, что в этом многомерном кубе изображения, соответствующие какому-то определенному образу, лежат недалеко друг от друга. Эта идея получила название «Гипотеза о компактности образов» [6].

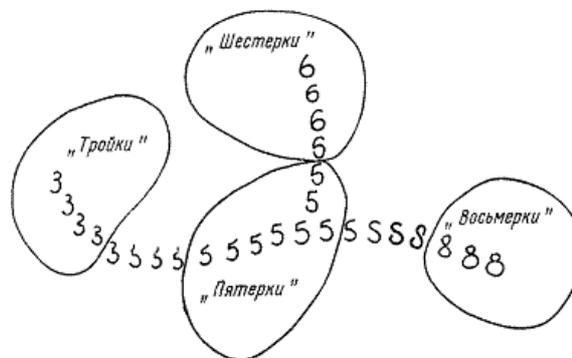


Рис. 4

Переформулируем нашу задачу: нужно универсальное множество разбить на «куски», компактные множества, каждому из которых соответствует образ.

4. РЕАЛИЗАЦИЯ МЕТОДА ПОТЕНЦИАЛЬНЫХ ФУНКЦИЙ

В программу в процессе обучения загружаются изображения (точки многомерного куба) и указания, к какому образу каждое изображение относится. При распознавании программа просто проверяет, в какую из известных компактных областей попало входное изображение.

Скорее всего, все указанные машине изображения лягут более-менее компактно, поэтому универсальное множество будет можно разделить. Собственно, разделять универсальное множество не будем, а будем пользоваться некоторой характеристикой, которая показывает удаленность одного рисунка (точки в вершине многомерного куба) до группы таких же изображений. В качестве меры удаленности рисунка от группы рисунков используется потенциал.

Известно, что электрический заряд создает вокруг себя поле, одной из характеристик которого является потенциал. В любой точке он может быть вычислен по формуле:

$$P = \alpha \frac{q}{R^2},$$

где α – некоторый постоянный коэффициент, q – величина заряда, R – расстояние от данной точки до заряда. Если электрическое поле образовано двумя или более зарядами, то потенциал в данной точке равен сумме потенциалов каждого заряда (принцип суперпозиции). Аналогия очевидна – каждый рисунок, на котором программа обучалась, создает в пространстве универсального множества потенциал. После обучения программе дают распознать какой-либо рисунок (точку в вершине многомерного куба), программа вычисляет потенциал, создаваемый в этой точке всеми объектами образа «а», образа «б»..., на которых программу учили, и распознаваемый рисунок относится к образу, который создал наибольший потенциал (рис. 4).

Для повышения надежности распознавания можно предусмотреть сильно различающиеся между собой эталоны (для того, чтобы как можно сильнее разбросать эталонные точки в пределах компактного множества).

Любое изображение представляет собой матрицу пикселей, ячейки которой могут принимать значения, соответствующие цвету пикселя. Повышению надежности распознавания сильно способствует введение так называемой приведенной матрицы. Образ, которому обучается программа, назовем эталонным, и будем формировать его на основе матрицы фиксированных размеров пикселей, назовем эту матрицу приведенной.

5. ВЫВОДЫ О ПРИМЕНИМОСТИ МЕТОДА ПОТЕНЦИАЛЬНЫХ ФУНКЦИЙ

В настоящее время уже во многих областях техники используются различные автоматы и устройства, более или менее удачно решаю-

щие задачу распознавания. К таким устройствам можно отнести и систему сортировки почтовых конвертов по индексу, и голосовой набор номера сотового телефона, и распознавание рукописного ввода на КПК с экранами сенсорного типа, и различные системы анализа спутниковых снимков, и многое другое. Особую пикантность алгоритму придает тот факт, что его математическое обоснование было предложено советскими математиками в начале 60-х гг., а еще лет через 15 была доказана его весьма глубокая аналогия с нейросетевыми методами.

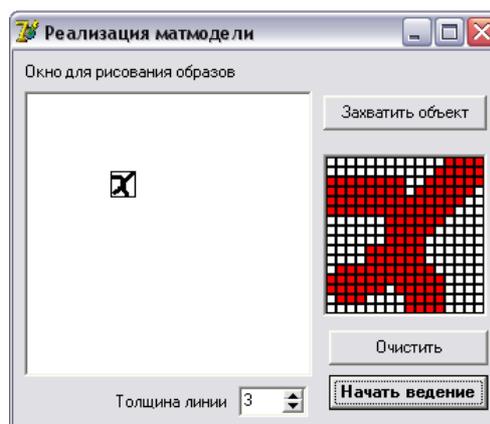


Рис. 5. Ведение объекта

В данной статье рассмотрен метод потенциальных функций применительно к захвату и ведению объекта, однако такую же его реализацию можно применить к проблеме обнаружения номеров движущихся в потоке автомобилей.

На основе сформированного метода создана и успешно испытана программа для ПЭВМ (рис. 5), которая позволяет построить объект произвольной формы, осуществить его захват и сопровождение в автоматическом режиме при его произвольных перемещениях. Успешное испытание данной программы (срывы сопровождения отсутствуют полностью) показывает правильность выбранного метода и его реализации и подтверждает его надежность.

В настоящее время коллектив ведет работу по разработке программы для ПЭВМ на основе этого метода по распознаванию и удержанию образа на подстилающей поверхности, позволяющей производить автоматическое оптическое сопровождение объекта системой, находящейся на подвижном аппарате.

СПИСОК ЛИТЕРАТУРЫ

1. **Ту Дж., Гонсалес Р.** Принципы распознавания образов = Pattern recognition principles. М.: Мир, 1978. 411 с.
2. **Козлов В. Н.** Элементы математической теории зрительного восприятия. М.: МГУ, 2001, 122 с.
3. **Местецкий Л. М.** Математические методы распознавания образов. М.: МГУ, 2004.
4. **Айзерман М. А.** Метод потенциальных функций в теории обучения машин. М.: Наука, 1970. 384 с.
5. **Мазуров В. Д.** Математические методы распознавания образов. Екатеринбург, 2010. 101 с.
6. **Потапов А. А.** Новейшие методы обработки изображений. М.: ФИЗМАТЛИТ, 2008. 496 с.

ОБ АВТОРАХ

Сенюшкин Николай Сергеевич, доц. каф. авиац. теплотехники и теплоэнергетики, ст. науч. сотр. НИЛ САПР-Д каф. авиац. двигателей, зам. декана ФАД. Дипл. инженер по авиац. двигателям и энергетическ. установкам (УГАТУ, 2005). Канд. техн. наук по тепл., электроракетн. двигателям и энергоустановкам ЛА (УГАТУ, 2009). Иссл. в обл. моделирования и проектирования камер сгорания авиац. ГТД, проектирования БПЛА.

Завьялов Роман Алексеевич, асп. каф. авиац. двигателей. Дипл. инженер по авиац. двигателям и энергетическ. уст. (УГАТУ, 2010). Иссл. в обл. автоматизации испытаний и проектирования беспилотных летательных аппаратов.