

**Дмитренко Е.Н., Тимонина К.А.**

Технологический университет им. А.А. Леонова, Королев, Россия

### **Особенности распознавания объектов с использованием *Matlab***

Распознавание изображений – в технике и технологиях – научно-техническое направление, связанное с разработкой методов и построением систем для установления принадлежности некоторого изображения (предмета, явления, ситуации, сигнала) к одному из заранее выделенных классов изображений. Распознавание изображений является междисциплинарной областью в искусственном интеллекте, объединяющей усилия инженеров и математиков, программистов и системотехников в деле разработки теории и практики создания систем распознавания. В статье рассматриваются методы распознавания объектов и их идентификации: глубокое обучение и машинное обучение, их преимущества и недостатки, возможности применения с использованием программного продукта MathLab.

**Ключевые слова:** машинное обучение, глубокое обучение, идентификация объектов, нейронные сети, корреляционный анализ.

**Dmitrienko E.N., Timonina K.A.**

A.A. Leonov Technological University, Korolev, Russia

### **Features of object recognition using *Matlab***

Image recognition in engineering and technology is a scientific and technical direction associated with the development of methods and the construction of systems to establish the belonging of a certain image (object, phenomenon, situation, signal) to one of the pre-selected classes of images.. Image recognition is an interdisciplinary field in artificial intelligence that unites the efforts of engineers and mathematicians, programmers and system engineers in developing the theory and practice of creating recognition systems. The article discusses methods of object recognition and identification: deep learning and machine learning, their advantages and disadvantages, as well as the possibilities of application using various software products MathLab).

**Keywords:** machine learning, deep learning, object identification, neural networks, correlation analysis.

#### **Рецензия:**

Статья посвящена адаптации алгоритма корреляционной обработки изображения, представленных цифровыми отсчётами, для выполнения в режиме реального времени. Новизна предложенного решения состоит в возможности уточнения значений результирующей корреляционной функции на основе обработки поступающих порций данных. Важной особенностью алгоритма являются сравнительно невысокие вычислительные затраты при сохранении возможности применения наложения при сегментации данных, а также нормализации корреляционной функции. Работоспособность решения в реальном масштабе времени показана в ходе моделирования в Simulink Matlab.

Статья представляет интерес для специалистов, аспирантов и магистрантов в области информатики и нейронных сетей. Рекомендуется к опубликованию.

## 1. Введение

Распознавание объектов — это метод компьютерного зрения для идентификации объектов на изображениях или видео. Распознавание объектов является основным результатом алгоритмов глубокого и машинного обучения [1].

Многие отрасли техники, имеющие отношение к получению, обработке, хранению и передаче информации, в значительной степени ориентируются в настоящее время на развитие систем, в которых информация имеет характер изображений. Изображение, которое можно рассматривать как двумерный сигнал, является значительно более емким носителем информации, чем обычный одномерный (временной) сигнал. Вместе с тем, решение научных и инженерных задач при работе с визуальными данными требует особых усилий, опирающихся на знание специфических методов, поскольку традиционная идеология одномерных сигналов и систем мало пригодна в этих случаях [2]. Это проявляется при создании новых типов информационных систем, решающих такие проблемы, которые до сих пор в науке и технике не решались, и которые решаются сейчас благодаря использованию информации визуального характера [3].

Теория распознавания образов базируется на гипотезе компактности: образам соответствуют компактные множества («сгустки» точек), объединенные в классы в пространстве признаков [4].

В целом, автоматизация обучения и распознавания неизвестных объектов составляет общую проблему распознавания образов, решение которой включает:

- определение полного перечня признаков, характеризующих образ (детерминированных, вероятностных, логических, структурных);
- первичная классификация образов и составление априорного алфавита, то есть выбор принципа классификации;
- разработка априорного словаря признаков, в который вносятся лишь те признаки, по которым может быть получена априорная информация;
- описание априорного алфавита классов на языке априорного словаря признаков;
- разбиение пространства признаков на области по классам алфавита с помощью решающих (разделяющих) функций;

- выбор и применение алгоритмов распознавания;
- определение рабочего алфавита классов и рабочего словаря признаков;
- разработка алгоритмов управления системой распознавания;
- выбор показателей эффективности работы системы распознавания (вероятность правильного решения, время распознавания, величина ресурсов и т.д.)
- обработка изображений с целью их распознавания является одной из центральных задач при создании систем искусственного интеллекта (СИИ) [5].

## 2. Материалы и методы

Комплексный характер задачи обработки изображений иллюстрирует рис. 1, на котором указаны основные процедуры от начального этапа восприятия поля зрения посредством датчиков, например, телекамеры, до конечного, которым является распознавание.

Обработка изображения представляет собой трудоемкий процесс, включающий процедуры фильтрации, сегментации, масштабирования, распознавания.



Рисунок 1- Этапы обработки изображения

Для того, чтобы существенно упростить исследования, разработку новых, оптимизацию существующих алгоритмов и тестирование полученных результатов, целесообразно использовать программные пакеты, предоставляющие готовые функции обработки изображений и алгоритмы распознавания. Пакет Image Processing Toolbox (IPT) предлагает широкий спектр средств для цифровой обработки и анализа изображений. Он представляет собой набор функций, которые расширяют возможности числовых вычислений в среде MatLab [6]. IPT поддерживает различные операции обработки изображений и включает:

- пространственные преобразования изображений;
- морфологические операции;
- скользящую и блочную обработку;
- линейную фильтрацию различными фильтрами;
- анализ и улучшение изображений;

- восстановление изображений;
- удаление размытостей.

### *Особенности работы с изображениями в Matlab*

Изображение можно определить как двумерную функцию  $f(x, y)$ , где  $x$  и  $y$  – это пространственные (плоскостные) координаты, а амплитуда  $f$  для каждой пары координат  $(x, y)$  называется интенсивностью или яркостью изображения в точке с этими координатами. Словосочетание “уровень серого” часто используется для обозначения яркости монохромного изображения. Цветные изображения формируются комбинацией нескольких монохромных изображений. Изображение может иметь непрерывные  $x$  – и  $y$  – координаты, а также непрерывную амплитуду  $f$ . Преобразование такого изображения в цифровую форму требует представления координат и значений амплитуды некоторыми дискретными отсчетами. Представление координат конечным множеством отсчетов называется дискретизацией, а представление амплитуды значениями из конечного набора называется квантованием. Таким образом, если координаты  $x$  и  $y$ , а также величины амплитуды  $f$  выбираются из фиксированных конечных наборов элементов, то изображение называется цифровым изображением.

Результатом дискретизации и квантования является матрица чисел.

Цифровое изображение в MATLAB имеет естественное представление в виде матрицы:

$$f = \begin{bmatrix} f(1,1) & f(1,2) & \dots & f(1,N) \\ f(2,1) & f(2,2) & \dots & f(2,N) \\ \dots & \dots & \dots & \dots \\ f(M,1) & f(M,2) & \dots & f(M,N) \end{bmatrix},$$

где  $f(p, q)$  обозначает точку, расположенную в строке  $p$  и в столбце  $q$ .

Полученные размеченные матрицы представляют собой выделенный объект на изображении. Чтобы его каким-либо образом классифицировать, необходимо определить параметры этого объекта, которые могут характеризовать его как геометрический объект конкретного класса, к ним относятся:

- Центр масс изображения;
- Deskрипторы областей;
- Deskрипторы границ;
- Фурье-deskрипторы;
- Морфометрические признаки;

*Один из вариантов алгоритма распознавания можно представить следующим образом:*

1. Локализовать изображения всех объектов на изображении.
2. Вычислить признаки объектов.
3. Сравнить полученные величины с эталонными значениями и вынести решение об отнесении фигуры к тому или иному классу геометрических фигур.

Существует несколько подходов к распознаванию изображений на основе вычисления корреляции.

#### *Классический корреляционный метод*

Под классическим алгоритмом понимается вычисление интеграла типа свертки с последующим поиском максимума этой функции:

1. Выделение габаритов фигуры, вычисление центра масс.
2. Вычисление максимально удаленной точки от центра масс.
3. Вырезание квадратного участка со стороной  $L = 2 * R_{max} + 1$ .
4. Нормализация.
5. Вычисление спектра стандартной проекции.
6. Нахождение спектров эталонов из базы  $T_i(x, y)$ .
7. Вычисление функции взаимной корреляции

$$R_i(x, y) = F^1 \langle G(f_x, f_y) \cdot T_i^*(f_x, f_y) \rangle$$

8. Определение изображенной фигуры, путем сопоставления с эталоном

$$R_i(0,0) = \max R(0,0).$$

Корреляционная обработка образов в пространстве Фурье обеспечивает определенные преимущества, особенно в аналоговых оптических системах. В цифровых системах применяются алгоритмы быстрого преобразования Фурье (БПФ), обеспечивающие сокращение времени вычислений в 10 – 20 раз.

Определив центр масс фигуры, необходимо вычислить корреляционную функцию лишь для одного положения. Также можно найти расстояние до максимально удаленной точки фигуры, что позволит вырезать участок изображения таким образом, чтобы в дальнейшем при расчете центральных моментов различных порядков не использовался весь массив данных изображения. Вид вырезаемой области – квадрат, у которого центр масс совпадает с центром самого квадрата [7].

В таком случае, при повороте фигуры центр масс остается на месте, и ни одна точка фигуры не выходит за границу вырезанной области.

Вырезав нужный участок изображения указанным способом, нужно выполнить нормализацию, то есть найти центральные моменты различных порядков, после этого необходимо рассчитать угол, на который должна повернуться фигура. Рассчитав угол поворота, и повернув изображение, для получения стандартной проекции исходной фигуры, остается провести масштабирование.

Получив стандартную проекцию, можно определить, какая фигура была на исходном изображении. Для этого ищется эталон из базы “знаний” системы (т.е. фигуры, которые она может опознать), наиболее похожий на полученную стандартную проекцию. Мерой сходства в данном методе является корреляционная функция [8].

#### *Разностный корреляционный метод*

Разностный корреляционный метод основан на поэлементном вычислении разностей интенсивности изображения. Для данного метода в целях сокращения объема вычислений необходимо провести предварительную обработку такую же, как и для предыдущего метода. Предварительная обработка заключается в нормализации.

Метод выглядит следующим образом:

1. Выделение габаритов фигуры, вычисление центра масс.
2. Вычисление максимально удаленной точки от центра масс.
3. Вырезание квадратного участка со стороной  $L = 2 * R_{max} + 1$ .
4. Нормализация.
5. Вычисление меры сходства.
6. Определение изображенной фигуры путем сопоставления с эталоном, мера сходства с которым принимает максимальное значение.

Далее, получив стандартную проекцию, можно определить, какая фигура была на исходном изображении. Для этого необходимо найти эталон  $(t(x, y))$  из базы эталонов, наиболее похожий на полученную стандартную проекцию  $(g(x, y))$ .

Общее выражение для разностных алгоритмов имеет вид

$$R(m, n) = \sum_i \sum_j |g(i, j) - t(i - m, j - n)|^p$$

где  $p = 0, 1, 2, \dots$

#### *Контурный метод*

Под контуром обычно понимается местоположение локального изменения или резкого перепада яркости на изображении. Подобные перепады возникают на границах объектов.

Процедуры построения контурных описаний изображений можно разделить на глобальные и локальные. Для глобальных процедур характерно разбиение изображения на однородные области, на основе которых и строятся контуры, например, как границы этих областей либо как их срединные оси. Локальные процедуры основываются либо на определении цепочек максимумов на градиентном поле изображения, либо на непосредственной аппроксимации яркостных переходов в рамках функциональных или стохастических моделей изображений. Построение градиентного поля полутонового изображения осуществляется с помощью операторов, выполняющих дискретное дифференцирование. Классическими операторами, служащими для этой цели, являются операторы Робертса, Превитт, Собела и др [9].

#### *Распознавание объектов и обнаружение объектов*

Обнаружение объектов и распознавание объектов являются схожими методами для идентификации объектов, но они различаются по своему исполнению. Обнаружение объектов — это процесс обнаружения экземпляров объектов на изображениях. В случае глубокого обучения методы обнаружения объектов — это подмножество методов распознавания объектов, когда объект не только идентифицируется, но и располагается на изображении. Это позволяет идентифицировать несколько объектов и размещать их на одном изображении [10].

Возможно использование различных подходов для распознавания объектов. В последнее время методы машинного и глубокого обучения стали популярными подходами к проблемам распознавания объектов. Обе технологии учатся распознавать объекты на изображениях, но они различаются по своему исполнению.

#### *Методы распознавания объектов*

##### *Распознавание объектов с использованием глубокого обучения*

Методы глубокого обучения стали популярным методом распознавания объектов. Модели глубокого обучения, такие как сверточные нейронные сети (или CNN), используются для автоматического изучения присущих объекту свойств, чтобы идентифицировать этот объект. Существует два подхода к распознаванию объектов с использованием глубокого обучения:

##### *Распознавание объектов с помощью машинного обучения*

Распространенными примерами методов машинного обучения являются:

- извлечение функций HOG с помощью модели машинного обучения SVM;
- алгоритм Виолы-Джонса, который можно использовать для распознавания различных объектов, включая лица и части тела [11].

### 3. Результаты исследований

С помощью всего нескольких строк кода MATLAB возможно создавать модели машинного и глубокого обучения для распознавания объектов, не будучи экспертом.

Процедура распознавания состоит из двух шагов:

1. Выбор изображения для распознавания (рис. 2).
2. Выбор метода распознавания (рис. 3).

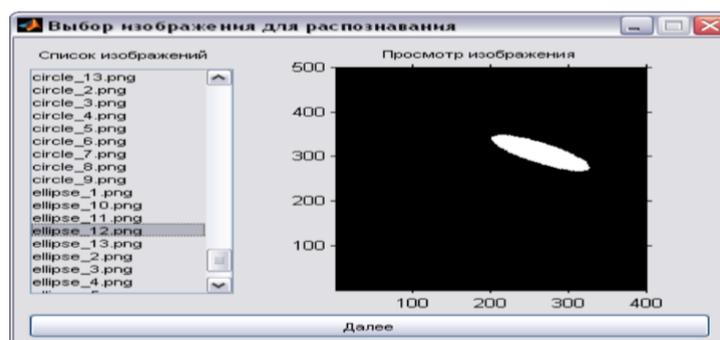


Рисунок 2 – Первый шаг – выбор распознаваемого изображения



Рисунок 3 – Второй шаг – выбор метода распознавания

В каждом конкретном методе распознавания реализован выбор эталона для анализа значения корреляционного признака «вручную» и кнопка запуска выбранного метода распознавания, при нажатии на которую происходит вычисление коэффициента корреляции, его минимизация/максимизация и выбор наиболее подходящего эталона.

### 4. Обсуждение

MATLAB автоматизирует развертывание моделей в корпоративных системах, кластерах, облаках и встроенных устройствах [12].

В зависимости от задачи может быть достаточно других, более простых подходов к распознаванию объектов:

- сопоставление с шаблоном, при котором используется небольшое изображение или шаблон для поиска совпадающих областей на большом изображении.

– сегментация изображений, при котором используются простые свойства объекта, такие как размер, цвет или форма, что может обеспечить надежное решение, которое не требует сотен или тысяч обучающих изображений [13].

## 5. Заключение

Использование MATLAB для распознавания объектов позволяет добиться успеха за меньшее время, поскольку позволяет:

- использовать свой опыт в предметной области и изучать науку о данных с MATLAB;
- использовать приложения для разметки данных и построения моделей;
- MATLAB позволяет создавать модели машинного и глубокого обучения с минимальным кодом;
- MATLAB может объединить распознавание объектов в единый процесс [14-15].

## Литература

1. Барский А.Б. Введение в нейронные сети [Электронный ресурс] / А.Б. Барский. — Электрон. текстовые данные. — М. : Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. — 358 с. — 2227-8397. — Режим доступа: <http://www.iprbookshop.ru/52144.html>
2. Волков, В.Ю. Адаптивные и инвариантные алгоритмы обнаружения объектов на изображениях и их моделирование в Matlab [Электронный ресурс] : учеб. пособие — Электрон. дан. — Санкт-Петербург : Лань, 2014. — 12 192 с. — Режим доступа: <https://e.lanbook.com/book/68475>. — Загл. с экрана.
3. Вьюгин, В.В. Математические основы машинного обучения и прогнозирования [Электронный ресурс] : учеб. пособие — Электрон. дан. — Москва : МЦНМО, 2013. — 304 с. — Режим доступа: <https://e.lanbook.com/book/56397>
4. Емельянов В.В. и др. Теория и практика эволюционного моделирования. Серия «Проблемы искусственного интеллекта». Научное издание. – М.: Физматлит, 2013. – 432 с.
5. Местецкий, Л.М. Математические методы распознавания образов [Электронный ресурс] : учеб. пособие — Электрон. дан. — Москва : , 2016. — 156 с. — Режим доступа: <https://e.lanbook.com/book/100634>. — Загл. с экрана
6. Неделько В.М. Основы статистических методов машинного обучения [Электронный ресурс] : учебное пособие / В.М. Неделько. — Электрон. текстовые данные. — Новосибирск: Новосибирский государственный технический университет, 2010. — 72 с. — 978-5-7782-1385-2. — Режим доступа: <http://www.iprbookshop.ru/45418.html>
7. Р. Гонсалес. Цифровая обработка изображений [Электронный ресурс] / Гонсалес Рафаэл, Вудс Ричард. – Электрон. текстовые данные. – М.: Техносфера, 2012. – 1104 с. – 978-5-94836-331-8.– Режим доступа: <http://www.iprbookshop.ru/26905.html>.
8. Родзин С.И. и др. Интеллектуальные системы. Генетические алгоритмы: Базовая концепция, когнитивные возможности и проблемные вопросы теории / Коллективная монография. – М.: Физматлит, 2017. – 295 с.
9. Родзин С.И. и др. Интеллектуальные системы. О некоторых эвристических алгоритмах, инспирированных природными системами / Коллективная монография. – М.: Физматлит, 2009. – 248 с.
10. Родзин С.И. и др. Системы искусственного интеллекта. Лабораторный практикум: Учебное пособие. – Таганрог: Изд-во ТТИ ЮФУ, 2017. – 128 с.
11. Сотник С.Л. Проектирование систем искусственного интеллекта [Электронный ресурс] / С.Л. Сотник. — 2-е изд. — Электрон. текстовые данные. — М. : Интернет-Университет Информационных Технологий (ИНТУИТ), 2016. — 228 с. — 2227-8397. — Режим доступа: <http://www.iprbookshop.ru/73716.html>
12. Тарков М.С. Нейрокомпьютерные системы [Электронный ресурс] / М.С. Тарков. — Электрон. текстовые данные. — М. : Интернет-Университет 10 Информационных Технологий

(ИНТУИТ), 2016. — 170 с. — 5-9556-0063-9. — Режим доступа: <http://www.iprbookshop.ru/52200.html>

13. Флах, П. Машинное обучение. Наука и искусство построения алгоритмов, которые извлекают знания из данных [Электронный ресурс] — Электрон. дан. — Москва : ДМК Пресс, 2015. — 400 с. — Режим доступа: <https://e.lanbook.com/book/69955>. — Загл. с экрана.
14. Шапиро, Л. Компьютерное зрение [Электронный ресурс] / Л. Шапиро, Д. Стокман. — Электрон. дан. — Москва : Издательство "Лаборатория знаний", 2015. — 763 с. — Режим доступа: <https://e.lanbook.com/book/84096>. — Загл. с экрана.
15. Bengio, Y. (2012). Practical recommendations for gradient-based training of deep architectures. arXiv:1206.5533