

Тема 4.

Методы распознавания образов

Основные принципы

- Ту Дж., Гонсалес Р. Принципы распознавания образов / Пер. с англ. Гуревич И.Б. - М.: «Мир», 1978

Одной из базовых функций нервной системы живых организмов, которая обеспечивает не только возможность ориентироваться во внешней среде, но и выживание — это способность различать и классифицировать предметы в этой среде.

Вместе с этим биологические механизмы решения задачи распознавания до конца не изучены. Решением этой задачи занимались и продолжают заниматься специалисты самых разных областей:

- Психологи
- Лингвисты
- Физиологи
- Математики
- Инженеры

... и результаты, к которым они приходят, зачастую имеют мало общего. Единого подхода к решению разнообразных задач, связанных с распознаванием, пока не найдено.

Работы по распознаванию образов можно разбить на три группы:

- 1) **Исследования живых распознающих систем**, и прежде всего человека (психология, физиология, биология)
- 2) **Создание специализированных распознающих устройств**, таких, как считывающие автоматы, устройства для автоматической постановки диагноза и т. д. (техника, вычислительные машины, информатика)
- 3) Анализ абстрактных моделей, отражающих, по мнению авторов, наиболее существенные свойства систем распознавания.

*вторую и третью группы иногда объединяют в одну

Существует значительное концептуальное различие между:

- созданием машины, которая будет обнаруживать световые лучи с определенной длиной волны, звуковые сигналы определенной частоты, определенные величины электрического напряжения или веса предметов, и...
- созданием машины, которая будет обнаруживать объект, исходя из некоторой совокупности его признаков.

Значительная часть работ посвящена описанию «классической» задачи распознавания образов.

Ее постановка предполагает следующее:

- **Объекты распознавания заданы набором признаков. Известно некоторое число эталонных объектов, и их описания составляют исходную (обучающую) информацию.**
- **На основе этой информации синтезируется алгоритм, определяющий для вновь поступивших объектов, к какому (или к каким) из конечного числа классов они принадлежат.**

Первые попытки автоматизации процесса распознавания образов относятся к **началу 50-х годов**, когда цифровые вычислительные машины стали общедоступным средством обработки информации. Часть этих исследований была направлена на разработку систем принятия решений для аппаратов автоматического считывания печатных буквенно-цифровых знаков.

В **конце 50-х годов** Розенблатт предложил **перцептронный алгоритм**, который представлял собой одну из первых моделей запоминания и организации информации, реализуемых мозгом.

Исследования в области **синтеза систем распознавания** набирали темп на протяжении **60-х годов** по мере расширения использования вычислительных машин и роста потребности в **скорости и эффективности взаимодействия человека и машины**.

Для использования результатов теории машинных языков при решении некоторых типов задач распознавания зрительных образов, был предложен **синтаксический** подход как дополнение к аналитическим методам.

Одна из основных проблем, повлекших повышенное внимание и развитие теории систем автоматического распознавания образов — **информационный взрыв**.

Эта проблема особо обращала на себя внимание в **70-е годы**. По предположениям психологов Гарвардского университета, к 2000 году возможности человеческого мозга воспринимать возросший поток информации могут оказаться исчерпанными.

Согласно статистике (США), в это время ежегодно насчитывалось более 60 000 научных журналов на 50 языках, содержащих более 2,5 миллиона статей. Для сравнения в 1830г. выходило около 300 технических и научных журналов. Ежегодно во всем мире выпускалось 80 000 книг.

Что касается банковской деятельности, то ежегодно через банки проходит 20 млрд. чеков, каждый из которых обрабатывается по 4-5 раз. Ведущие банки выполняют в день около 25 млн. операций.

Почта США за одну секунду обрабатывает 27 000 единиц отправок, что 84 млрд. в год, а к 80 году ожидалось около 116 млрд.

Информационные системы играют важную роль в процессах принятия решений. Окончательная оценка таких решений будет всегда производиться человеком, однако появившиеся концепции и методы, предусматривающие использование вычислительных машин, обеспечивают выдвижение и оценку такого количества вариантов, какое ни один руководитель не мог осмыслить.

Возможность использовать, кроме того, методы анализа риска позволяет руководителю оценивать потенциальную опасность, связанную с каждой из множества изучаемых стратегий.

Информационные системы обеспечат руководителю возможность более ясно представлять результаты принимаемых им решений.

Способность распознавать считается основным свойством человеческих существ и других живых организмов.

Образ представляет собой описание объекта. В каждое мгновение бодрствования человек совершает акты распознавания:

- распознает окружающие объекты и принимает решение о перемещении и других действиях,
- заметить в толпе знакомого человека, понять, что он говорит,
- узнать голос знакомого человека
- прочесть рукописный текст
- отличить улыбку от злобной гримасы

В соответствии с характером распознаваемых образов акты распознавания можно разделить на два основных типа:

- распознавание **конкретных объектов** (зрительные слуховые образы - «сенсорное» распознавание) и
- распознавание **абстрактных объектов** (опознать старый довод, обдумать ситуацию, найти решение задачи - «понятийное» распознавание)

В сущности **распознавание** человеком **образов** можно свести к вопросу оценки относительных шансов на то, что **исходные данные соответствуют** тому или иному из **известных множеств** статистических совокупностей, определяющихся **прошлым опытом** человека и представляющих **ориентиры на априорную информацию** для распознавания.

Таким образом, **задачу распознавания** образов можно рассматривать как **задачу установления различий** между исходными данными, причем не посредством отождествления с отдельными образами, но с **их совокупностями**.

Последнее осуществляется при помощи **поиска признаков** (инвариантных свойств) на множестве объектов, **образующих определенную совокупность**.

Образ — это описание любого элемента как представителя соответствующего класса образов.

Искусственные и естественные образы

Искусственные образы создаются человеком, как правило, для коммуникации между собой (передачи информации) с учетом некоторых взаимных соглашений относительно этих условных образов. Типичными примерами являются формулы, музыкальные ноты, дорожные знаки, карты, чертежи и другие условные системы.

Естественные образы — образы, встречающиеся в природе и которым человек присваивает наименования (озеро, река, дерево, гора, облако).

Свойства классов условных образов:

- Конечность набора элементарных символов
- Первичная абстрактная структура
- Конечная совокупность правил формирования
- Значение контекста
- Избыточность
- Введение новых символов и новых правил

- **Конечность набора элементарных символов:**

Крайне трудно привести пример условной системы символов, в которой количество элементарных символов не является конечным. Однако это не всегда характерно для задач распознавания естественных образов.

- **Первичная абстрактная структура:**

Среди изобилия различных форм имеется некое общее абстрактное представление. Для многих классов естественных образов первичная абстрактная структура неизвестна

Основные задачи при разработке СРО

- Представление исходных данных, полученных в результате измерений для подлежащего распознаванию объекта.

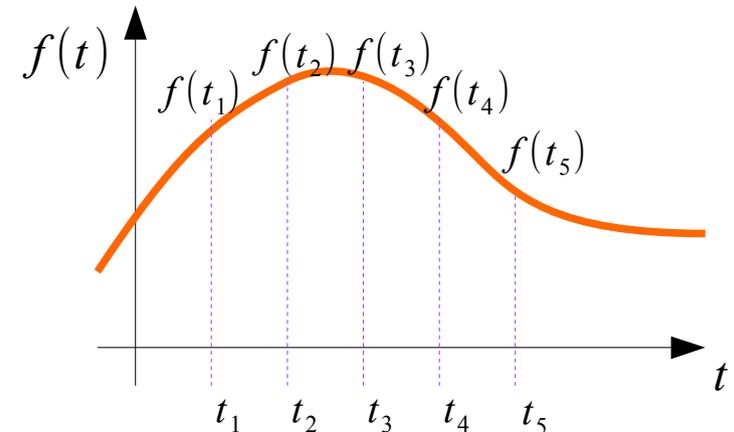
Каждая измеренная величина является некоторой характеристикой образа или объекта. В случае распознавания символов в датчике может быть реализована измерительная сетчатка. Если сетчатка состоит из n элементов, то результаты измерений представляются в виде вектора — вектора образа:

$$\mathbf{x} = (x_1, x_2, \dots, x_n)^T$$

- Образами служат непрерывные функции (например, звуковые сигналы).

Если измерение значений функции производится в дискретных точках t_1, t_2, \dots, t_n , то вектор образа формируется в виде

$$x_1 = f(t_1), \quad x_2 = f(t_2), \quad \dots, \quad x_n = f(t_n)$$



Векторы образов содержат всю поддающуюся измерению информацию о самих образах. Процесс измерения обычно понимается как процесс кодирования характеристик образов в компонентах вектора x .

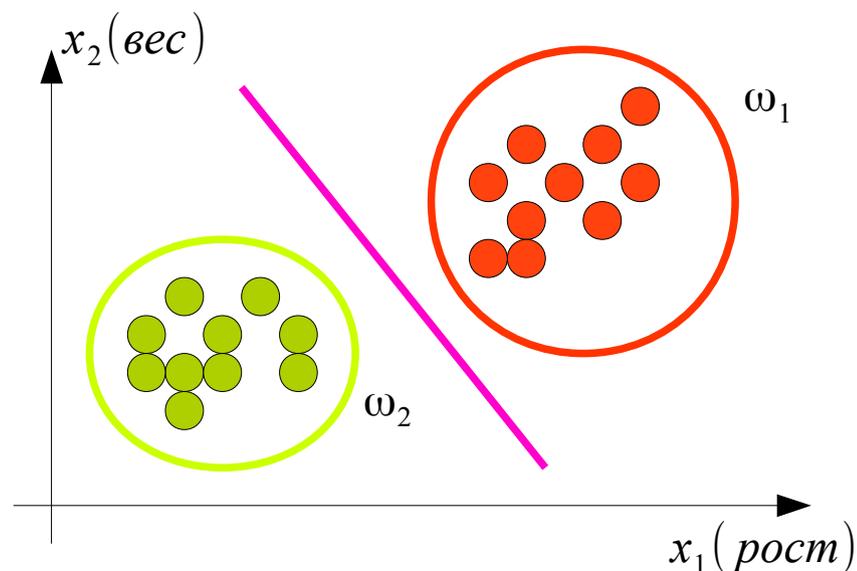
Когда измерения приводят к информации в форме действительных чисел, полезно рассматривать векторы образов в качестве точек n -мерного евклидова пространства. Множество образов, принадлежащих одному классу, соответствует совокупности точек, рассеянных в некоторой области пространства измерений.

Например, имеются два класса образов, каждый из которых характеризуется двумя измерениями: ростом и весом.

ω_1 - класс группы футболистов-профессионалов

ω_2 - класс группы жокеев

Векторы образов имеют вид: $x = (x_1, x_2)^T$



Вторая задача распознавания образов связана с выделением характерных признаков или свойств из полученных исходных данных и снижением размерности векторов образов. Эту задачу часто определяют как задачу предварительной обработки и выбора признаков.

Признаки класса образов представляют собой характерные свойства, общие для всех образов данного класса. Признаки, характеризующие различия между отдельными классами, можно интерпретировать как межклассовые признаки.

Внутриклассовые признаки, общие для всех рассматриваемых классов, не несут полезной информации с точки зрения распознавания и могут не приниматься во внимание.

Выбор признаков считается **одной из важных задач** в теории распознавания образов и при построении распознающих систем.

Для большинства практических задач процедура выделения различающих признаков является чрезвычайно трудной, однако ее можно использовать для упрощения автоматического распознавания.

Третья задача, связанная с построением СРО, состоит в отыскании оптимальных решающих процедур, необходимых при идентификации и классификации.

Задача сводится к построению границ областей решений, разделяющих классы точек в пространстве измерений.

Решающие функции можно получить различными способами. В случае, когда имеются полные априорные сведения о распознаваемых образах, решающие функции могут быть точно определены на основе этой информации.

Если относительно образов имеются лишь качественные сведения, то о виде решающих правил могут быть выдвинуты лишь разумные допущения, при этом границы областей решений могут существенно отклониться от истинных.

В этом случае применяется поэтапная корректировка решающих функций. В частности может использоваться дополнительная информация в контексте рассматриваемых образов, измеренная с помощью условных вероятностей, лингвистических статистик и других количественных характеристик.

Объекты, подлежащие распознаванию и классификации (с помощью СРО), должны обладать набором измеримых характеристик.

Цель работы СРО заключается в том, чтобы на основе собранной информации определить класс объектов с характеристиками, аналогичными измеренным у распознаваемых объектов.

Корректность распознавания зависит от объема различающей информации, содержащейся в измеряемых характеристиках, и эффективности использования этой информации.

Если было бы возможно измерить все возможные характеристики и время на обработку собранной информации было бы не ограничено, то достичь адекватного уровня распознавания было бы возможным, используя самые примитивные методы.

Основные концепции и методологии

В основе идеи синтеза систем распознавания лежат способы, с помощью которых описываются и разделяются классы образов.

- ★ Принцип принадлежности перечню: класс характеризуется перечнем входящих в него членов.
- ★ Принцип общности свойств: класс характеризуется некоторыми общими свойствами, присущими всем его членам.
- ★ Принцип кластеризации: при рассмотрении классов, обнаруживается тенденция к образованию образами кластеров в пространстве их признаков.

Принцип перечисления членов класса

В случае задания класса перечислением входящих в него классов механизм распознавания основывается на процедуре сравнения с эталоном.

Множество образов, принадлежащих одному классу, хранятся системой распознавания. При предъявлении этой системе новых образов, она последовательно сравнивает их с хранящимися в ее памяти эталонами. В результате распознавания новый образ будет относиться к тому классу, с эталонным образом которого совпал предъявленный к распознаванию образ.

Пример принципа перечисления членов класса — литеры различных шрифтов при условии отсутствия их зашумления и искажения.

Идея принципа перечисления элементов класса **достаточно проста**, позволяет строить **недорогие системы распознавания**, успешно справляющимися со своими задачами, однако эта задача становится нетривиальной в **случае большого числа элементов с большим количеством признаков**.

Принцип общности свойств

Задание класса с помощью свойств, общих для всех входящих в его состав членов, предусматривает реализацию процесса автоматического распознавания путем **выделения подобных признаков и работы с ними.**

Основное допущение в этом методе заключается в том, что образы, принадлежащие одному и тому же классу, обладают рядом общих свойств или признаков, отражающих подобие таких образов.

Когда системе предъявляется новый образ, то выделяется набор описывающих его признаков, которые в закодированном виде сравниваются с признаками, хранящимися в памяти распознающей системы. В результате сравнения система зачислит предъявленный образ в класс, характеризующийся набором признаков, подобных признакам предъявленного образа.

Основной задачей метода является **выделение ряда общих свойств** по конечной выборке образов, принадлежащих определенному классу.

Для запоминания признаков класса требуется значительно меньше ресурсов, чем для хранения всех объектов, входящих в класс.

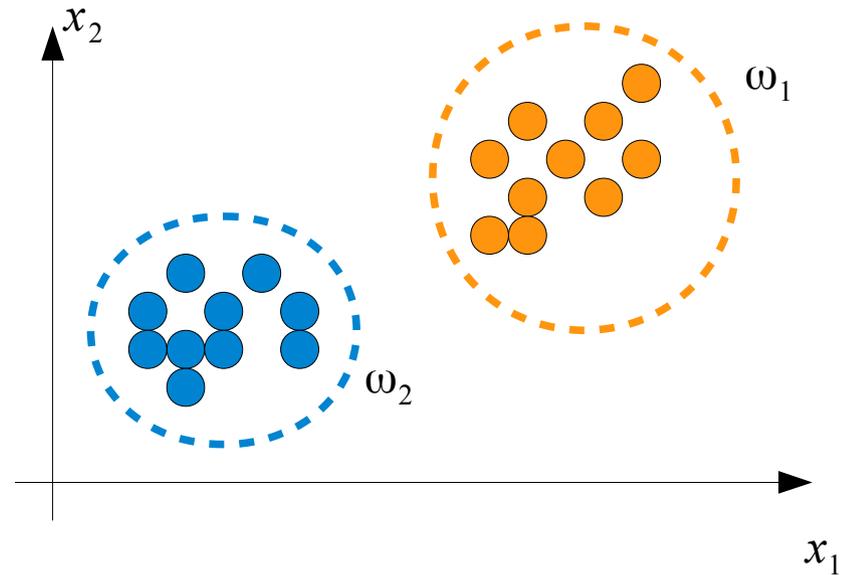
Принцип сопоставления признаков допускает вариацию характеристик отдельных образов в отличие от процедуры сопоставления с эталоном, не допускающей существенной вариации признаков.

Если все признаки, характеризующие класс, можно найти по имеющейся выборке образов, то процесс распознавания сводится к сопоставлению по признакам.

Очень трудоемкой является задача отыскания полного набора различающих признаков класса. В некоторых случаях решение этой задачи не представляется возможным.

Принцип кластеризации

Под кластером понимается группа образов (в общем случае — объектов), образующих в пространстве признаков в некотором смысле компактную область.



Методы разбиения объектов на классы основаны на принципе геометрической близости, что в терминах характеристик образов означает близость или схожесть их признаков.

Построение СРО, основанных на принципе кластеризации, определяется взаимным расположением кластеров в пространстве признаков.

Если кластеры, соответствующие различным классам, располагаются достаточно далеко друг от друга, то можно воспользоваться достаточно простыми методами классификации образов, например, основанными на принципе минимального расстояния.

Для реализации рассмотренных выше принципов построения СРО выделяются 3 методологических подхода:

- ★ Эвристический
- ★ Математический
- ★ Лингвистический

Нередко процесс разработки СРО включает одновременно несколько подходов.

Эвристический подход

В основе подхода — интуиция и опыт человека.

В рамках построения СРО эвристический подход подразумевает **перечисление членов** класса и **принцип общности свойств**.

СРО, построенные с использованием эвристических методов, содержат ряд специфических процедур, выполняющих конкретные функции применительно к конкретной задаче распознавания.

Например, в задаче распознавания символов, эвристический метод может быть направлен на выделение, подсчет и определение последовательности определенных штрихов.

Недостаток эвристического подхода — ориентированность на конкретную задачу. Качество решения зависит от опыта и квалификации разработчиков.

Математический подход

В основе подхода лежат правила классификации, которые формулируются и выводятся в рамках определенного математического формализма с помощью принципов общности свойств и кластеризации.

Математические методы построения СРО можно разделить на:

- ★ **Детерминистские:** Базируются на математических методах, не использующих в явном виде статистических свойств изучаемых объектов. Например, итерационные алгоритмы.
- ★ **Статистические:** основываются на математических правилах классификации, которые формулируются и выводятся в терминах математической статистики. Например, построение статистического классификатора в общем случае предполагает использование байесовского классификационного правила. Требуются плотности распределения для всех совокупностей образов и вероятности появления образа для каждого из классов.

Лингвистический подход

Образ можно описать с помощью иерархической структуры подобразов, аналогичной синтаксической структуре языка.

Это обстоятельство позволяет применять теорию формальных языков. Предполагается, что грамматика образов содержит конечные множества элементов, называемых переменными, производными элементами и правилами подстановки. Характер правил подстановки определяет тип грамматики.

Среди наиболее изученных грамматик можно выделить:

- ★ Регулярные
- ★ Бесконтекстные
- ★ Грамматики составляющих

Ключевыми элементами данного подхода являются:

- ★ Выбор непроеизводных элементов образа
- ★ Объединение этих элементов и связывающих их отношений в грамматики образов
- ★ Реализация процессов анализа и распознавания

Лингвистический подход особенно полезен при работе с образами, которые либо не могут быть описаны числовыми измерениями, либо настолько сложны, что их локальные признаки идентифицировать не удастся и приходится обращаться к глобальным свойствам объектов.

Примеры автоматических СРО

- ★ Распознавание символов
- ★ Классификация спутниковых данных
- ★ Биомедицинские приложения
- ★ Распознавание отпечатков пальцев
- ★ Технический надзор за состоянием ядерного реактора

Простая модель распознавания образов

Простая схема распознавания содержит два основных блока:

★ Датчик

★ Классификатор

Датчик представляет собой устройство, преобразующее физические характеристики объекта, подлежащего распознаванию, в набор признаков, характеризующих объект:

$$\mathbf{x} = (x_1, x_2, \dots, x_n)^T$$

Классификатор представляет собой устройство, относящее каждый поступающий на его вход допустимый набор значений к одному из конечного числа классов, вычислив множество значений решающих функций.

Считается, что система распознавания допускает ошибку в том случае, если она относит к классу ω_j объект, на самом деле принадлежащий отличному от ω_j классу.

Считается, что система распознавания R_1 лучше системы распознавания R_2 , если вероятность совершить ошибку для системы R_1 меньше, чем для системы R_2 .

Датчик выдает информацию в виде вектора $\mathbf{x} = (x_1, x_2, \dots, x_n)^T$

Предполагается, что вектор измерений \mathbf{x} принадлежит одному из M классов образов

$$\omega_1, \omega_2, \dots, \omega_n$$

Принимается допущение о том, что априорные вероятности появления объектов каждого класса одинаковы, т. е. вектор \mathbf{x} может с равной вероятностью относиться как к одному, так и любому другому классу.

Пусть $p(\mathbf{x}|\omega_i) = p_i(\mathbf{x})$ есть плотность распределения для вектора \mathbf{x} при условии, что он принадлежит классу ω_i .

В таком случае вероятность того, что на самом деле вектор \mathbf{x} принадлежит классу ω_j , определяется выражением:

$$p_j = \frac{p(\mathbf{x} | \omega_j)}{\sum_{k=1}^M p(\mathbf{x} | \omega_k)}$$

Вероятность того, что вектор \mathbf{x} не принадлежит классу ω_j , определяется выражением:

$$1 - p_j = 1 - \frac{p(\mathbf{x} | \omega_j)}{\sum_{k=1}^M p(\mathbf{x} | \omega_k)}$$

задающим вероятность ошибки.

Решающая функция представляет собой функцию $d(x)$, относящую x точно к одному из M заданных классов.

Оптимальной считается решающая функция $d^o(x)$, которая дает наименьшую вероятность ошибки при всех допустимых значениях x .

Значение j , при котором величина $1 - p_j$ будет наименьшей, совпадает с тем значением j , которому соответствует наибольшее значение вероятности $p(x | \omega_j)$.

Оптимальная решающая функция $d^o(x)$ относит набор x к классу ω_i в том и только том случае, если выполняются неравенства

$$p(x | \omega_i) > p(x | \omega_j) \quad \forall i \neq j$$

или

$$\frac{p(x | \omega_i)}{p(x | \omega_j)} > 1 \quad \forall i \neq j \quad \omega_j$$

Решающая функция представляет собой функцию $d(x)$, относящую x точно к одному из M заданных классов.

Оптимальной считается решающая функция $d^o(x)$, которая дает наименьшую вероятность ошибки при всех допустимых значениях x .

Значение j , при котором величина $1 - p_j$ будет наименьшей, совпадает с тем значением j , которому соответствует наибольшее значение вероятности $p(x | \omega_j)$.

Оптимальная решающая функция $d^o(x)$ относит набор x к классу ω_i в том и только том случае, если выполняются неравенства

$$p(x | \omega_i) > p(x | \omega_j) \quad \forall i \neq j$$

или

$$\frac{p(x | \omega_i)}{p(x | \omega_j)} > 1 \quad \forall i \neq j \quad \omega_j$$