

Распознавание образов. Обзорная лекция.

Содержание

Распознавание образов. Обзорная лекция.....	1
Введение.....	2
Типичная схема работы системы распознавания.....	2
Типы задач фазы функционирования.....	2
Задача распознавания.....	3
Математическая постановка	3
Способы определения классов объектов.....	4
Обучение с учителем, обучающая выборка.....	4
Примеры систем, решающих задачу распознавания.....	5
Классификация.....	6
Математическая постановка задачи классификации.....	6
Способы определения критериев различия.....	6
Обучение без учителя.....	6
Классификация методов распознавания.....	7
Процедура предъявления обучающих выборок.....	7
Фиксированная выборка.....	7
Последовательная выборка.....	7
Вид правил классификации.....	8
Параллельная классификация.....	8
Последовательная классификация.....	9
Способ описания объектов.....	9
Евклидово пространство.....	9
Списки признаков.....	11
Структурное описание.....	11
Распознавание скорописных текстов.....	13
Распознавание текста.....	13
Машинопечатные и скорописные тексты.....	13
Анализ методов распознавания в применении к скорописным текстам.....	15
Фреймовые модели.....	15
Список литературы.....	16

Введение

Распознавание — это способность живых организмов обнаруживать в потоке информации, поступающей от органов чувств, определённые объекты, закономерности, явления. Оно может осуществляться на основе зрительной, слуховой, тактильной информации. Так, человек без труда может узнать другого знакомого ему человека, взглянув на него или услышав его голос. Некоторые животные активно используют обоняние для узнавания других особей и поиска пищи.

Возможность распознавания опирается на схожесть однотипных объектов. Несмотря на то, что все предметы и ситуации уникальны в строгом смысле, между некоторыми из них всегда можно найти сходства по тому или иному признаку. Отсюда возникает понятие классификации — разбиения всего множества объектов на непересекающиеся подмножества - классы, элементы которых имеют некоторые схожие свойства, отличающие их от элементов других классов. И, таким образом, задачей распознавания является отнесение рассматриваемых объектов или явлений по их описанию к нужным классам. Т.е. понятие распознавания можно расширить, если говорить об обнаружении объектов в потоке не только чувственной, но и любой другой информации. Например, можно говорить о распознавании болезни по её симптомам у больного или о распознавании социальных явлении по статистической информации.

Развитие и распространение компьютерной обработки информации привели к возникновению в середине XX века потребностей в технологиях, позволяющих машинам осуществлять распознавание в обрабатываемой ими информации. Разработка методов машинного распознавания позволяет расширить круг выполняемых компьютерами задач и сделать машинную переработку информации более интеллектуальной. Примерами сфер применения распознавания могут служить системы распознавание текста, машинное зрение, распознавание речи, отпечатков пальцев и прочее. Несмотря на то, что некоторые из этих задач решаются человеком на подсознательном уровне с большой скоростью, до настоящего времени ещё не создано компьютерных программ, решающих их в столь же общем виде. Существующие системы предназначены для работы лишь в специальных случаях со строго ограниченной областью применения.

Типичная схема работы системы распознавания

1. Фаза настройки (обучения);
2. Фаза функционирования:
 - a) Этап предобработки
 - b) Этап принятия решения

Типы задач фазы функционирования

1. Распознавание (обучение с учителем) — отнесение предъявляемых объектов к определённым классам с помощью применения известных правил классификации. Это наиболее типичная задача систем распознавания. Перед тем, как система сможет выполнять данную функцию, предполагается её обучение на множестве примеров — *обучающей выборке* объектов распознавания.
2. Классификация (обучение без учителя) — разбиение множества объектов на

непересекающиеся классы по их формализованным описаниям. Данная задача решается в тех случаях, когда от системы не требуется отнесения входных образов к каким-либо определённым классам, а требуется лишь способность различать их каким-либо способом по определённым признакам. Можно говорить об оперировании безымянными классами объектов.

3. Создание формализованного описания объектов распознавания, пригодного для использования алгоритмами собственно распознавания. Как правило, исходные данные о наблюдаемых объектах представлены в форме, непригодной непосредственно для распознавания. Это могут быть растровые изображения, звуковые файлы, статистические данные (числовые наборы), видео-записи и пр — т.н. «код 0». Алгоритмы распознавания как правило требуют более высокоуровневого представления. Это приводит к необходимости произвести одно или более преобразований исходных данных, переходя от кода 0 к коду 1, 2 и т.д. В качестве примера можно рассмотреть процедуру сегментации изображения, т.е. выделения на нём одноцветных областей.

Задача распознавания

Математическая постановка

Дано множество M объектов ω . Объекты задаются значениями некоторых признаков x_i , $i=1, \dots, N$, наборы которых одинаковы для всех объектов. Совокупность признаков объекта ω определяет некоторым образом его описание $I(\omega) = (x_1(\omega), x_2(\omega), \dots, x_N(\omega))$. Признаки могут выражаться в терминах да/нет, да/нет/неизвестно, числовыми значениями, значениями из набора возможных вариантов и т.д.

На всём множестве M существует разбиение на подмножества (классы объектов):

$$M = \bigcup_{i=0}^m \Omega_i$$

Разбиение на классы может быть задано полностью или определяться некоторой априорной информацией I_0 о классах Ω_i - например, характеристическим описанием входящих в них объектов.

Задача распознавания состоит в том, чтобы для каждого данного объекта ω по его описанию $I(\omega)$ и априорной (обучающей) информации I_0 вычислить значения предикатов

$$P_i = (\omega \in \Omega_i), \quad i=1, \dots, m$$

Для описания невозможности распознавания объектов предикаты P_i заменяются величинами $\alpha_i \in \{0(\omega \notin \Omega_i), 1(\omega \in \Omega_i), \Delta(\text{неизвестно})\}$.

Таким образом, для рассматриваемого объекта ω необходимо вычислить его информационный вектор

$$\alpha(\omega) = (\alpha_1(\omega), \dots, \alpha_m(\omega))$$

Процедура, строящая информационный вектор $\alpha(\omega)$ в данном случае выражает алгоритм принятия решения об отнесении объекта ω к тому или иному классу и называется «решающей функцией».

Способы определения классов объектов

Разбиение рассматриваемого множества объектов на классы Ω_i может быть задано следующими способами:

1. *Перечисление.* Каждый класс задаётся путём прямого указания его членов. Такой подход используется в том случае, если доступна полная априорная информация о всех возможных объектах распознавания. Предъявляемые системе образы сравниваются с заданными описаниями представителей классов и относятся к тому классу, которому принадлежат наиболее сходные с ними образцы. Такой подход называют методом *сравнения с эталоном*. Он, к примеру, применим при распознавании машинопечатных символов определённого шрифта. Его недостатком является слабая устойчивость к шумам и искажениям в распознаваемых образах.

2. *Задание общих свойств.* Класс задаётся указанием некоторых признаков, присущих всем его членам. Распознаваемый объект в таком случае не сравнивается напрямую с группой эталонных объектов. В его первичном описании выделяются значения определённого набора признаков, которые затем сравниваются с заданными признаками классов. Такой подход называется *сопоставлением по признакам*. Он экономичнее метода сравнения с эталоном в вопросе количества памяти, необходимой для хранения описаний классов. Кроме того, он допускает некоторую вариативность распознаваемых образов. Однако, главной сложностью является определение полного набора признаков, точно отличающих членов одного класса от членов всех остальных.

3. *Кластеризация.* В случае, когда объекты описываются векторами признаков или измерений, класс можно рассматривать как кластер. Распознавание осуществляется на основе расчёта расстояния описания объекта до каждого из имеющихся кластеров. Если кластеры достаточно разнесены в пространстве, при распознавании хорошо работает метод оценки расстояний от рассматриваемого объекта до каждого из кластеров. Сложность распознавания возрастает, если кластеры перекрываются. Обычно это является следствием недостаточности исходной информации и может быть разрешено увеличением количества измерений объектов. Для задания исходных кластеров целесообразно использовать процедуру обучения.

Обучение с учителем, обучающая выборка

Альтернативой ручному вводу в систему описаний классов объектов и параметров решающей функции является обучение системы. Оно представляет собой процедуру самонастройки системы распознавания на основе воспринимаемой информации и может происходить как при подготовке системы, так и в процессе её работы по мере «накопления опыта». Вид обучения, в котором системе представляется набор образцов распознаваемых объектов с указанием их принадлежности классам, называется *обучением с учителем*. Набор образцовых объектов называется *обучающей выборкой*. Необученная система производит распознавание предлагаемых объектов и сравнивает свои результаты с правильными ответами, поступающими от учителя. По результатам сравнения система корректирует параметры решающей функции. В случае использования метода перечисления для задания

классов система просто сохраняет в памяти всю доступную об объектах информацию. Если классы задаются описанием общих свойств, то система определяет различия в значениях признаков объектов разных классов. При использовании кластерного подхода по анализу векторов признаков объектов обучающей выборки формируются исходные кластеры в пространстве признаков.

Итак, обучающая выборка в задаче распознавания является априорной информацией о множестве распознаваемых объектов и представляет описание всех классов объектов:

$$T = I_0(\Omega_1, \dots, \Omega_m)$$

Её составляют описания предложенных учителем объектов с указанием их принадлежности классам, т.е. можно определить её как совокупность описаний объектов

$$(I(\omega_1), \dots, I(\omega_{r_1}), I(\omega_{r_1+1}), \dots, I(\omega_{r_2}), \dots, I(\omega_{r_m})) ,$$

где объекты $\omega_1, \dots, \omega_{r_1}$ принадлежат классу Ω_1 , объекты $\omega_{r_1+1}, \dots, \omega_{r_2}$ - классу Ω_2 и т.д.

Таким образом, обучающая выборка представляет собой таблицу, строки которой помечены названиями объектов $\omega_i, i=1, \dots, r_m$, а столбцы — названиями признаков $x_j, j=1, \dots, N$. Элементами таблицы являются значения признаков объектов $x_j(\omega_i)$. Строки таблицы сгруппированы по классам $\Omega_j, j=1, \dots, m$.

	x_1	x_2	...	x_N	Классы
ω_1	$x_1(\omega_1)$	$x_2(\omega_1)$...	$x_N(\omega_1)$	Ω_1
ω_2	$x_1(\omega_2)$	$x_2(\omega_2)$...	$x_N(\omega_2)$	
...	
ω_{r_1}	$x_1(\omega_{r_1})$	$x_2(\omega_{r_1})$...	$x_N(\omega_{r_1})$	
...
$\omega_{r_{m-1}+1}$	$x_1(\omega_{r_{m-1}+1})$	$x_2(\omega_{r_{m-1}+1})$...	$x_N(\omega_{r_{m-1}+1})$	Ω_m
...	
ω_{r_m}	$x_1(\omega_{r_m})$	$x_2(\omega_{r_m})$...	$x_N(\omega_{r_m})$	

Примеры систем, решающих задачу распознавания

- OCR;
- распознавание отпечатков пальцев;
- машинное зрение робототехнических систем;
- распознавание речи;
- биомедицинские приложения: анализ и автоматическая обработка рентгенограмм, электрокардиограмм, исследования хромосом, диагностика заболеваний, интерпретация опросников;
- контроль и мониторинг технических систем.

Классификация

Математическая постановка задачи классификации

Дано множество M объектов ω , имеющих свои описания $I(\omega) = (x_1(\omega), x_2(\omega), \dots, x_N(\omega))$. Задан критерий $K(I(\omega))$, позволяющий отличать объекты друг от друга в соответствии с определённым условием. Например, такой критерий может выдавать целое число, определяющее положение объекта в некоторой ранговой системе.

Требуется в соответствии с заданным критерием K построить разбиение множества M на классы $\Omega_i : M = \bigcup_{i=0}^m \Omega_i$. Действие критерия в этом случае определяется следующим выражением:

$$\forall i, j (i, j = 1, \dots, m) \quad i \neq j \Leftrightarrow K(I(\omega \in \Omega_i)) \neq K(I(\omega \in \Omega_j))$$

Следует отметить, что число классов m изначально неизвестно и определяется по результатам выполнения классификации.

Способы определения критериев различия

В задачах классификации для задания правил различия объектов могут быть использованы те же методы, что и при распознавании.

1. *Прямым перечислением* можно задать примеры заведомо различающихся объектов. Решение об отнесении поступающих объектов к тому или иному классу принимается на основе меры сходства с эталонами. Решающая функция строится на основе анализа эталонных объектов и априорной информации об их принадлежности разным классам.
2. *Описание общих свойств* позволяет сформировать правило классификации на основе анализа характеристик входного объекта.
3. *Кластеризация* позволяет различать входные объекты на основе анализа характерных свойств разнесённых в пространстве кластеров.

Обучение без учителя

Для обучения системы классификации, как и системы распознавания, можно воспользоваться обучающей выборкой. На этот раз от системы не требуется запоминания конкретных классов объектов и последующего отнесения к ним поступающих образов. Достаточно лишь определения, относятся ли представленные объекты к одному или к различным классам. Поэтому при обучении системе должна быть представлена выборка объектов с указанием, какие из них различны, а какие — сходны. Другими словами, системе нужно показать примеры отличия образов. Не важно, какие значения будет вырабатывать решающая функция, важно, чтобы эти значения были разными для разнотипных объектов и

одинаковыми для однотипных. Такое обучение, в котором отсутствуют подсказки правильных ответов учителем, называется *обучением без учителя*. Во время обучения без учителя система строит кластеры объектов обучающей выборки, а на этапе функционирования производит определение близости входных образов к построенным кластерам.

Назначение классифицирующих систем:

- кластеризация множества объектов;
- понижение размерности описаний для упрощения дальнейшего анализа.

Классификация методов распознавания

Процедура предъявления обучающих выборок

- *фиксированная выборка* - построение правил классификации по единственному обучающему множеству, предъявляемому до начала классификации;
- *последовательная выборка* - коррекция правила классификации с каждой новой предъявляемой выборкой на основе оценки критерия качества распознавания (самообучение, автоподстройка).

Фиксированная выборка

На основе обучающей выборки, представленной системе до начала выполнения ей функции распознавания, строится правило классификации. В дальнейшем это правило не меняется; все подлежащие распознаванию объекты оцениваются на основе полученных при обучении данных. Ошибки, совершаемые системой при распознавании, не учитываются и носят постоянный характер. Системы, построенные на этом принципе, требуют тщательного обучения и используются в основном там, где множество распознаваемых объектов хорошо известно, вариативность объектов предсказуема или хорошо отрабатывается правилами классификации. В целом, такие системы более простые и надежные.

Последовательная выборка

При таком подходе первоначальная выборка формирует лишь предварительную версию правил классификации. При последующей работе система оценивает качество распознавания и подстраивает правила классификации или вырабатывает новые с целью повышения качества распознавания. Процесс прекращается тогда, когда правила классификации начинают удовлетворять некоторому критерию оптимальности. Этот подход называется *адаптивным распознаванием* или *машинным обучением*.

Свойства процедуры последовательного обучения:

- сходимости — процесс изменения правил классификации должен приводить через конечное число корректировок к окончательному варианту;

- оптимальность — решающее правило, к которому сходится процесс обучения, должно быть оптимальным в смысле минимизации функции стоимости возможных ошибок;

- вычислительная сложность — финальное правило классификации должно минимизировать число шагов, необходимых для распознавания образа, время их выполнения.

Системы с последовательной процедурой обучения проявляют большую гибкость при распознавании объектов с высокой степенью варьированности. Кроме того, «накопление опыта» позволяет им достичь высокого качества и скорости распознавания. Однако, алгоритм работы таких систем более сложен и требует обоснования с точки зрения выполнения указанных выше свойств.

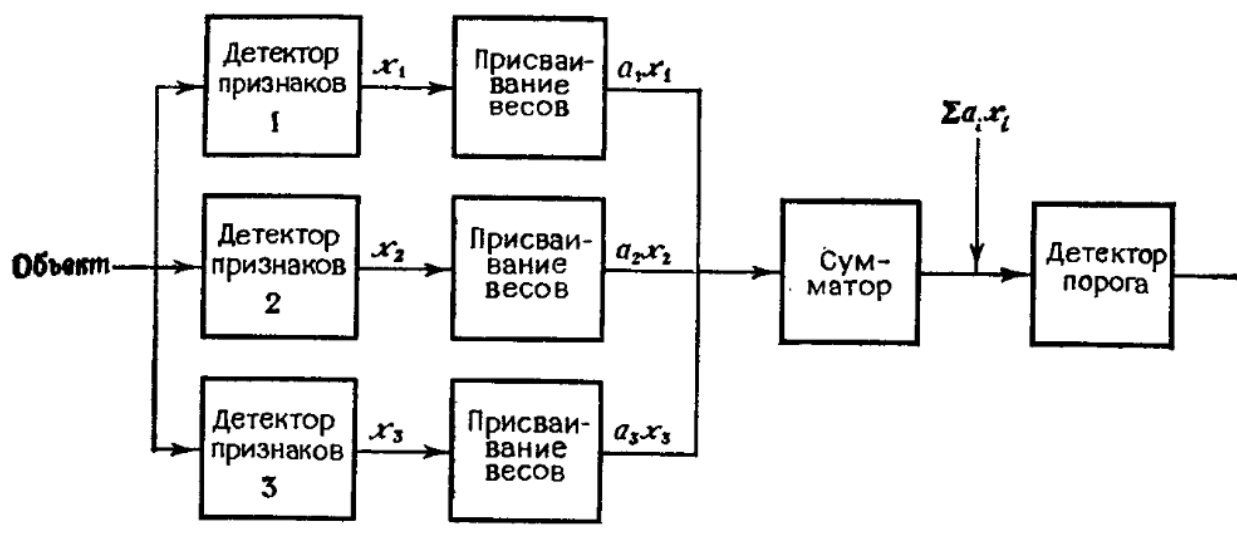
Вид правил классификации

- *параллельные* - проведение ряда тестов над всей совокупностью выявленных данных об объекте и принятие решения на основе их результатов;

- *последовательные* - проведение последовательности тестов над подмножествами выявленных данных; выбор очередного теста определяется результатами предыдущих тестов.

Параллельная классификация

Систему с параллельным способом обработки входного образа можно проиллюстрировать следующим образом:



Для выполнения распознавания система производит ряд тестов над всеми компонентами $x_i, i=1, \dots, N$ описания входного объекта одновременно. Решающая функция в этом случае представляется функцией не более, чем N переменных $g(x_1, \dots, x_N)$.

Система может быть организована в виде множества параллельных функций $F = \{f_j(x_1, \dots, x_N)\}, j=1, \dots, m$, каждая из которых производит оценку принадлежности

объекта к соответствующему ей классу. В таком случае решающая функция принимает решение на основе максимального полученного значения - $g = \max_j (f_j)$.

Параллельная процедура является достаточно надёжной и требует для распознавания постоянного времени, равного времени выполнения самой продолжительной из процедур распознавания. Однако, она не обладает гибкостью, свойственной последовательным процедурам.

Последовательная классификация

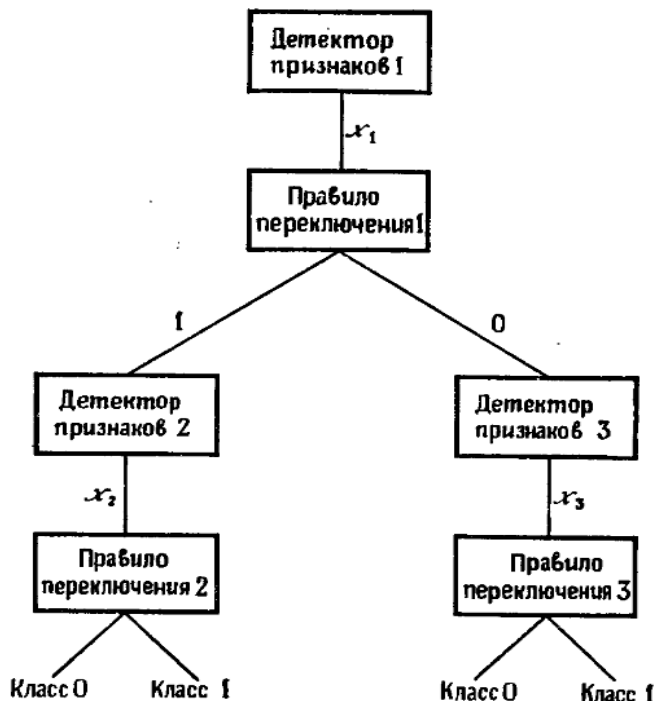
Последовательный способ классификации, как и параллельный, предусматривает ряд тестов над признаками распознаваемого объекта. Однако, выполняются они не одновременно, а последовательно, причём порядок их выполнения может зависеть (и чаще всего зависит) от получаемых результатов.

Последовательная классификация позволяет сократить число необходимых для распознавания тестов. Время выполнения в сравнении с параллельной процедурой зависит от выполняющего её вычислительного устройства. Если

устройство позволяет производить параллельные вычисления, то последовательная процедура потребует значительно большего времени. Если же устройство последовательное, то в худшем случае на выполнение последовательной процедуры потребуется ровно столько же времени, сколько и на выполнение параллельной.

В целом последовательная процедура позволяет реализовать более сложные правила классификации.

Возможно комбинирование этих двух принципов с целью достижения компромисса между простотой параллельных процедур и возможностями последовательных.



Способ описания объектов

- *евклидово пространство* - объекты представляются точками в евклидовом пространстве их вычисленных параметров, представление в виде набора измерений;
- *списки признаков* - выявление качественных характеристик объекта и построение характеризующего вектора;
- *структурное описание* - выявление структурных элементов объекта и определение их взаимосвязи.

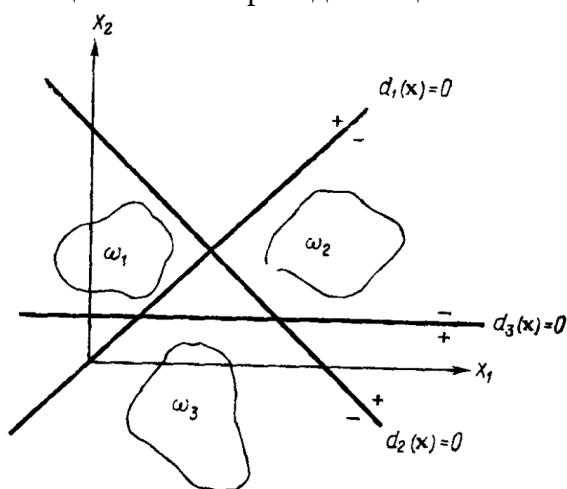
Евклидово пространство

Представление распознаваемых объектов в виде точек Евклидова пространства строится следующим образом. Над первичным представлением распознаваемого образа

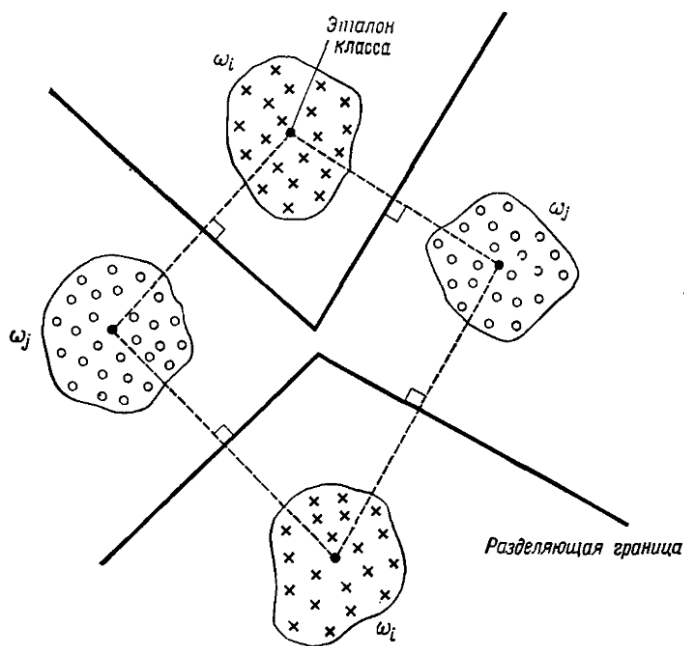
производится серия вычислений, определяющих необходимые для классификации характеристики. Далее в многомерном Евклидовом пространстве (параметрическом пространстве, или пространстве характеристик), каждое измерение которого соответствует одной из вычисляемых характеристик, строится точка, соответствующая совокупности полученных измерений. По совокупностям точек, Евклидово расстояние между которыми мало, выделяют в область пространства, соответствующую данному классу изображений.

Как правило, методы, работающие с Евклидовым пространством, используют параллельную процедуру подачи входного изображения. Распознавание основывается на проведении ряда математических вычислений над полным множеством точек изображения. Можно выделить следующие основные используемые виды правил классификации:

- *Решающие (дискриминантные) функции.* Построение правил классификации заключается в поиске функций, описывающих разделение параметрического пространства на соответствующие классам объектов области. При этом реализуется способ задания классов в виде указания общих свойств принадлежащих им объектов.



- *Функции расстояния.* Определение принадлежности объектов различным классам производится на основе анализа расстояний между точками объектов одного класса и между точками объектов различных классов. Классы представляются в виде кластеров в параметрическом пространстве. Построение правил классификации заключается в поиске функций, обеспечивающих оптимальное построение кластеров объектов. Данный подход активно использует методы кластерного анализа.



- *Функции правдоподобия.* Используется математическая статистика, Байесовские оценки, теория игр. Процесс распознавания представляется как игра распознающего устройства с реальным миром, в которой при распознавании каждого символа машина пытается угадать символ, задуманный природой. Построение правила классификации сводится к поиску функций выигрышей и потерь игроков. Принятие решения о классе

распознаваемого объекта строится на значении функции правдоподобия предполагаемого решения при заданных значениях параметров объекта. Данный метод применим при наличии достаточного количества априорной информации о распознаваемом множестве.

К данному классу методов можно также отнести методы, использующие в своей основе нейронные сети. Они активно используют процедуру самообучения (последовательная обучающая выборка) и относятся к параллельным методам. Нейронные сети способны решать как задачи распознавания, так и автоматической классификации, допуская обучение как с учителем, так и без него.

Списки признаков

Методы данной категории опираются на возможность классификации распознаваемых объектов на основе наличия в них некоторых характерных признаков. Существуют два основных подхода. Первый основывается на предположении, что простые измерения, проводимые над изображением, есть результат действия совокупности небольшого числа порождающих признаков. Задачей разработки метода распознавания является определение пространства признаков, к которому требуется свести пространство прямых измерений. При этом пространство признаков имеет меньшую размерность. При таком подходе используется аппарат факторного анализа.

Второй подход определяет признаки как подмножества множеств простых измерений. При распознавании бинарных изображений такими признаками могут служить наличие черных точек в определённых областях изображения (например, диагональ или горизонтальная черта в середине изображения), число черных точек на характеристической линии, проводимой через изображение и т.д. Распознаваемые объекты представляются как различные совокупности наблюдаемых признаков.

Методы данного класса являются в большой степени эвристическими. Их эффективность определяется правильностью выбора набора рассматриваемых признаков.

Пример из работ Кадакина. При распознавании рукописных шифров источников картотеки для каждой строки шифра осуществляется определение значений нескольких признаков. В роли этих признаков выступают количество чёрных точек на средней и нижней линии строки с учётом её возможного наклона, а также число верхних и нижних выступающих элементов букв. Далее по полученным данным производится оценка числа букв в словах. Число букв является сложным признаком более высокого уровня. Он используется далее для сопоставления полученных данных со сведениями о имеющихся шифрах и идентификация наблюдаемого шифра по результатам сопоставления.

The image shows the handwritten word "Разум" in a cursive script. Below the word, there are several small circles and lines indicating feature extraction points, likely corresponding to the "middle and bottom lines" mentioned in the text.This image is similar to the one above, showing the word "Разум" with feature markers. The markers here appear to be more focused on the lower part of the word, possibly representing the "bottom line" feature mentioned in the text.

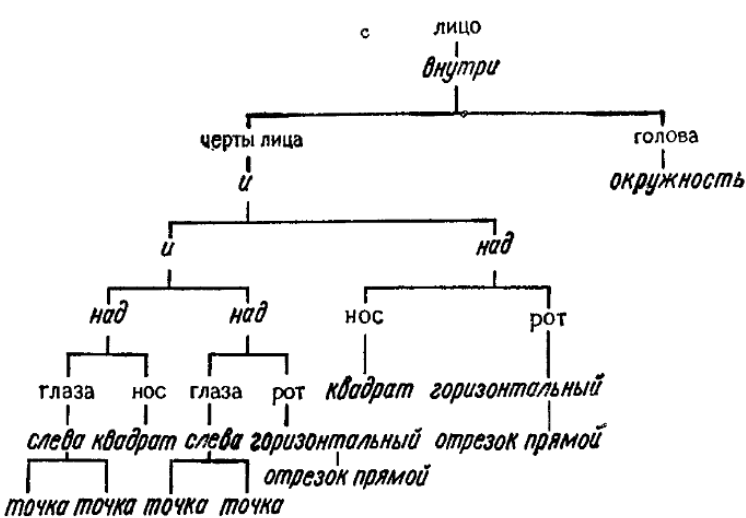
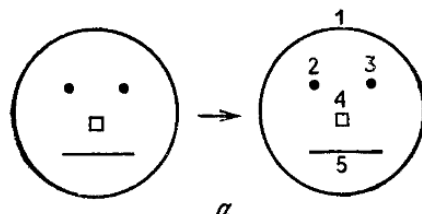
Структурное описание

Структурное описание основывается на представлении объектов в виде совокупности «непроизводных элементов» и отношений между ними. Под непроизводными элементами понимаются фрагменты распознаваемых образов, которые, с одной стороны, формируют эти образы, с другой — просты в смысле собственной структуры, т.е. не содержат других непроизводных элементов, сколь-нибудь значимых для описания образа.

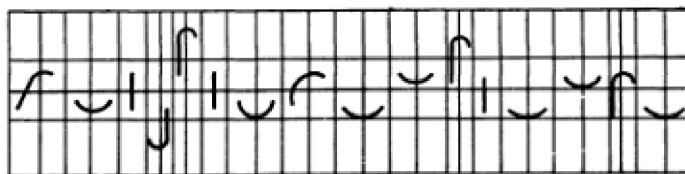
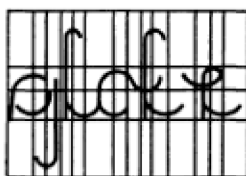
Как правило системы, использующие структурное описание объектов, реализуют последовательную процедуру распознавания. В процессе решения информационной задачи они обрабатывают входной образ, обходя его структуру элемент за элементом. Процесс

распознавания делится на два потока: выделение в образе структурных элементов определенного вида и согласование получаемой структурной информации с имеющимися в системе моделями для классов изображений.

Одним из видов структурных методов являются методы грамматической классификации образов. Они используют представление образов в виде предложений специального языка. На этапе инициализации алгоритма требуется определить виды возможных структурных элементов изображений. Построение правила классификации сводится к выводу грамматики, описывающей язык классифицируемых образов. Распознавание заключается в определении выводимости рассматриваемого предложения с помощью найденной грамматики. Примерами грамматического описания изображений служат: язык описания изображений PDL (Picture Definition Language, Язык Описания Изображений), плекс-грамматики, веб-грамматики, синтаксическое описание рукописных символов. Грамматики Эванса используют в качестве терминальных символов простые элементы изображений, такие как круг, прямоугольник, линия. В качестве терминалов, описывающих взаимоотношения элементов, используются специальные операторы, такие как "над", "внутри" и т.д.



В работе Идена приводится модель английской скорописи. В результате анализа английских скорописных текстов был выявлен ограниченный набор типов линий, с помощью комбинации которых можно изобразить любую скорописную букву. Подход нашел своё применение в системах распознавания Parascript FormXtra.



В качестве структурных элементов изображений могут использоваться траектории движения пишущего или рисующего инструмента. В таком случае эти элементы могут характеризоваться последовательностью направлений шагов перемещения пера. Направления могут кодироваться кодами из конечного набора наперёд заданных направлений,

распределённых на диапазоне $0^\circ - 360^\circ$. Другим способом является описание элементов с помощью кривых Безье.

Распознавание скорописных текстов

Распознавание текста является достаточно широко востребованной задачей. В настоящее время существуют системы, осуществляющие распознавание машинопечатных текстов (например, ABBYY FineReader). Распознавание рукописного текста является более сложной задачей и в достаточной мере на сегодняшний день не решена.

В применении к распознаванию изображений можно говорить о двух принципиально разных подходах:

- онлайн-распознавание — процесс распознавания ведётся параллельно с процессом синтеза изображения;
- офлайн-распознавание — распознавание производится на уже сформированном изображении.

Онлайн-режим подразумевает совмещение процесса формирования входных изображений и их ввода в систему. Таким образом обеспечивается возможность "наблюдения" системой за процессом начертания вводимых символов, что предоставляет ей возможность получать помимо графической информации ещё и данные о структуре входных изображений. Онлайн-задача является более простой в решении, и сегодня реализовано достаточное количество приложений, осуществляющих онлайн-распознавание рукописного текста. Офлайн-распознавание рукописи реализовано лишь отчасти, в виде специальных приложений, работающих с формами, заполняемыми от руки.

Распознавание текста

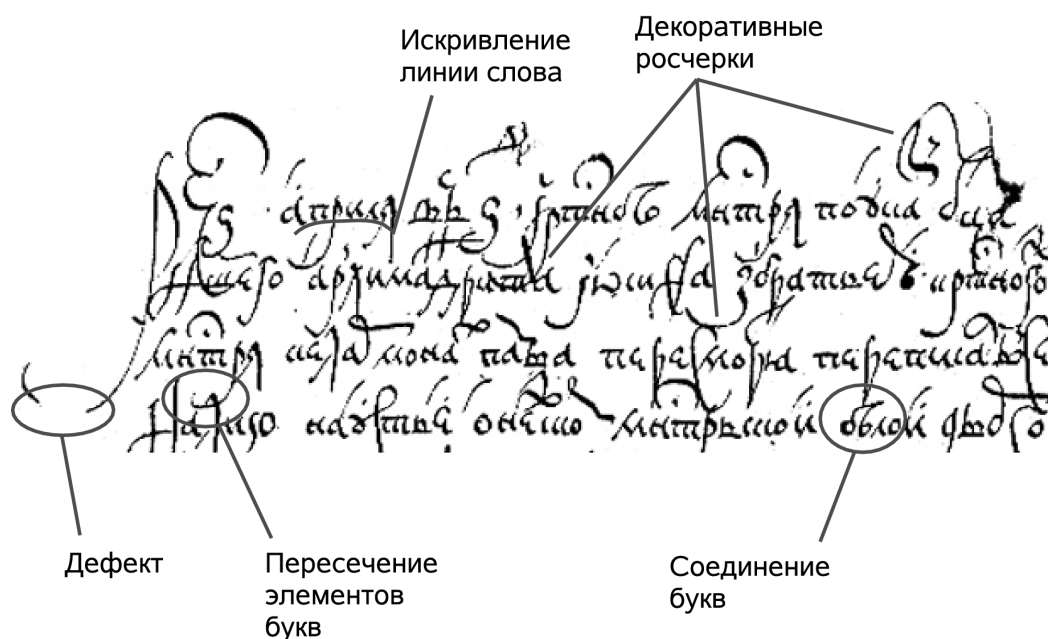
Задачу офлайн-распознавания текстовой информации можно определить следующим образом. На вход подаётся изображение текста. Требуется определить его кодированное электронное представление, т.е. произвести его перевод из растрового графического представления в текстовое. Последнее означает, что для каждого изображения буквы текста нужно определить соответствующее кодирующее число. А это означает, что каждый фрагмент исходного изображения, соответствующий букве, должен быть отнесён к одному из классов букв, набор которых задаётся известным алфавитом. Изображения букв разных классов отличаются по начертанию. Выходное текстовое представление рукописи должно содержать весь набор букв, представленный в исходном изображении, в соответствующем порядке их следования.

Таким образом, решаемая задача является задачей множественного распознавания изображений букв текста.

Машинопечатные и скорописные тексты

Сложность решения задачи распознавания текста находится в сильной зависимости от особенностей его графического представления. В текстах, полученных типографским способом, буквы располагаются в ровных строках, имеют одинаковое начертание в пределах всего документа и, как правило, имеют достаточно четко различимые промежутки между собой. При решении задачи распознавания эта информация позволяет использовать

допущения, снижающие сложность разрабатываемых алгоритмов и методов. В рукописных текстах, и особенно в скорописных, таких ограничений нет. Здесь присутствует множество особенностей, связанных с процессом синтеза текста. Назначением скорописи является быстрое нанесение текста на бумагу, при этом без внимания оказывается эстетическая сторона оформления документа. Поэтому в скорописных текстах велика вариативность начертания букв и они часто связываются соединительными линиями, строки не располагаются на ровных горизонталях, могут встречаться помарки и кляксы. Поскольку буквы часто имеют соединительные линии и могут иметь случайные пересечения, задача распознавания расширяется необходимостью выделения отдельных букв в изображении перед их классификацией. Кроме того, процесс получения электронных изображений рукописей обычно включает в себя несколько этапов копирования (фотографирования и сканирования), каждый из которых вносит свою долю помех и дефектов в выходное изображение.



Виды вариативности в рукописном начертании символов:

- элемент выведен не точно, с отклонениями (особенности мышечной деятельности писца);
- отсутствует предполагаемое пересечение элементов;
- присутствует лишнее пересечение (например, с другой буквой);
- элемент дополнен декоративным сегментом;
- добавлен декоративный элемент;
- добавлен соединительный штрих;
- дефект начертания (пробельный участок);
- изменены относительные размеры и положение элементов.



Различие начертания



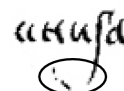
Отсутствует пересечение



Лишнее пересечение



Декоративный элемент



Дефекты

Анализ методов распознавания в применении к скорописным текстам

Методы параллельного распознавания (в частности, методы, использующие Евклидово пространство описаний и списки признаков) опираются на фиксированность размера принимаемой на вход выборки. Условием их применения для распознавания является предварительное выделение символов из исходного изображения. В контексте скорописного текста это обстоятельство является недостатком в связи со связанностью букв между собой. Определить границу между двумя буквами параллельным методом можно только уже распознав сами буквы.

Структурные методы позволяют системам распознавания перемещаться по изображению, постепенно и последовательно собирая необходимую информацию. Необходимость в выделении отдельных символов для их распознавания отпадает, так как в процессе анализа полученных на данный момент сведений об изображении распознаватель может выбирать область дальнейшего исследования и самостоятельно определять границы наблюдаемого символа. В настоящее время структурные методы распознавания изображений применяются в наиболее развитых системах распознавания скорописного текста.

Фреймовые модели

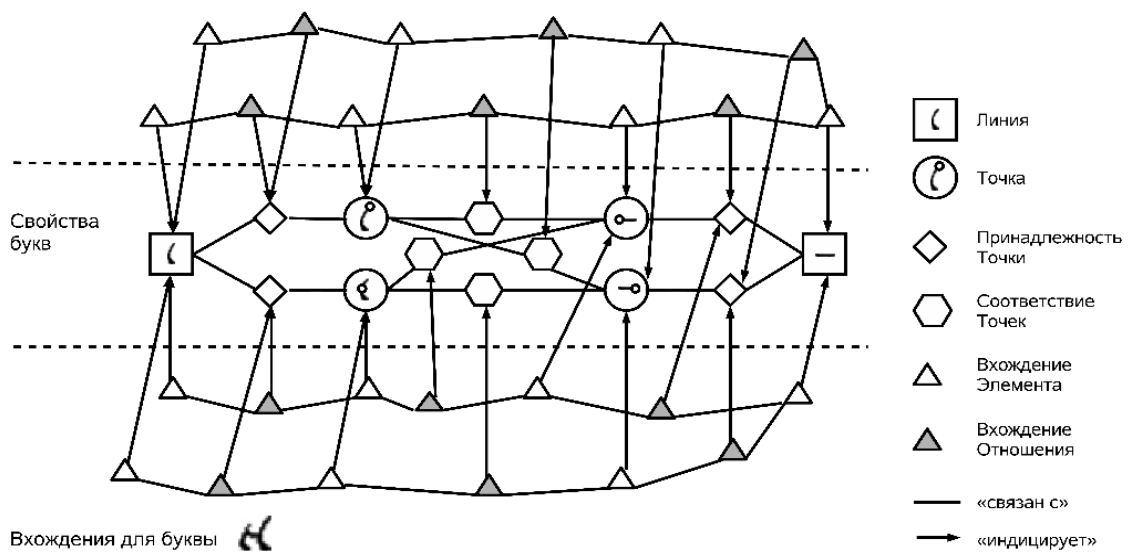
Одним из способов анализа и согласования структурной информации об изображении является построение фреймовых моделей, предложенных в работе Минского. Информация о структурных элементах изображения и их взаимосвязи представляется в виде фреймов -- сетевых структур, узлы которых описывают объекты изображения, а именованные дуги -- их взаимоотношения. В процессе распознавания в изображении выделяются структурные элементы и определяются связи между ними. Процесс сегментации изображения управляется процессом проверки выбранного в текущий момент фрейма -- гипотезы. Изображение признаётся как описываемое каким-либо фреймом, если удаётся связать его терминальные узлы с наблюдаемыми элементами изображения. Преимуществом такой модели является возможность описания изображений на разных концептуальных уровнях, от наиболее абстрактных понятий к конкретным фрагментам изображения. Это позволяет проводить распознавания на разных уровнях детализации, способствуя решению проблемы вариативности начертания рукописных символов. Кроме того, существует возможность разделения несколькими высокоуровневыми объектами общих низкоуровневых элементов. Другой полезной особенностью является возможность задания для некоторых терминалов фрейма значений по умолчанию, которые могут быть изменены при обнаружении противоречия, или признаются действительными в противном случае.

Пример модели описания структуры букв в виде фреймовой сети

Структуру изображения буквы можно определить с помощью следующих типов производных элементов:

- линии различных видов, соответствующие росчеркам пишущего инструмента;
- точки пересечения линий;
- пространственные отношения (выше-ниже, слева-справа, больше-меньше);
- отношения принадлежности точек линиям;
- отношения соответствия (совпадения) точек на изображении;
- маркеры вхождения элементов и отношений в фрейм конкретной буквы.

Вхождения для буквы **т**



Список литературы

1. Хант Э. Искусственный интеллект. М.: Мир, 1978.
2. Фу К. Структурные методы в распознавании образов. М.: Мир, 1977
3. Ту Дж, Гонсалес Р. Принципы распознавания образов. М.: Мир, 1978.
4. Уоссермен Ф. Нейрокомпьютерная техника: теория и практика. Пер. с англ. Ю.А.Зуев, В.А. Точенов, 1992.
5. Мински М. Фреймы для представления знаний. Пер. с англ. 1979.
6. Журавлёв Ю.И. Распознавание. Классификация. Прогноз. Математические методы и их применение. Вып.2. М.: Наука, 1989.