

Курс лекций «Введение в ИИ.». Часть II.
Распознавание образов
Лекция 11. Введение в распознавание образов

О.Г. Чанышев

Содержание

1 Введение: понятие образа	1
2 Проблема обучения распознаванию образов (ОРО)	2
3 Геометрический и структурный подходы	3
3.1 Геометрический подход	3
3.2 Структурный (лингвистический) подход	4
4 Обучение и самообучение. Адаптация и обучение	5

1 Введение: понятие образа

Термины "идентификация", "классификация", кластеризация" широко распространены в различных областях науки и техники. Центральное место среди этих понятий занимает понятие (или задача) распознавания образов. Эту задачу человек решает постоянно, даже во сне, когда он видит сны. Воспринимая внешний мир или его отображения в нашем мозге, мы всегда производим классификацию воспринимаемых ощущений, т. е. разбиваем их на группы похожих, но не тождественных явлений. Например, несмотря на существенное различие, к одной группе относятся все буквы А, написанные различными почерками, или все звуки, соответствующие одной и той же ноте, взятой в любой октаве и на любом инструменте, а оператор, управляющий техническим объектом, на целое множество состояний объекта реагирует одной и той же реакцией.

Распознаванием образов постепенно овладевают системы, наделенные элементами искусственного интеллекта, например, роботы. Методы распознавания образов играют ведущую роль в системах распознавания речи. В системах понимания текста формальные синтаксические методы уступают место методам распознавания образов.

Как мы увидели ранее, самым узким местом экспертных систем является процесс извлечения знаний у эксперта. Для того, чтобы они вышли на качественно новый уровень, научившись самостоятельно приобретать, структурировать и иерархизировать знания, их конструкция обязательно должны быть дополнена "блоком распознавания образов".

В 1985 г. Алан Фог [Fog:89] писал, что если сопоставить процесс отыскания функции, разбивающей множество признаков на требуемые классы, с процессами распознавания в мозгу

Рис. 1: Образы, сгруппированные по сходству

человека, то каждый этап ее создания можно соотнести с процессом, протекающим в соответствующем отделе нервной системы. Сканер будет соответствовать периферии нервной системы, а по мере прохождения сигнала более высоким отделам, операции приобретают все более высокую степень обобщения.

Если оставить в стороне техническую часть, т.е., устройство и функционирование сканеров (датчиков), то останется задача систематизации результатов их работы - множества признаков - или задача таксономии (гр. taxis (расположение в порядке) + nomos (закон)). Содержательную постановку задачи таксономии дал еще Демокрит [Zagor:99] в "Письме ученому соседу"[MatDrGr:57]: "Если тебе, дорогой друг, нужно разобраться в сложном нагромождении фактов или вещей, ты сначала разложи их на небольшое число куч по схожести. Картина прояснится и ты поймешь природу этих вещей". Сейчас уместно попрактиковаться в распознавании визуальных образов, отдавая себе отчет в собственных умственных действиях (см. рис.1). Здесь представлены 12 задач, в которых следует отобрать признаки, при помощи которых можно отличить левую триаду картинок от правой. Решение данных задач требует моделирования логического мышления в полном объеме.

Образы обладают характерным свойством, проявляющимся в том, что ознакомление с конечным числом явлений из одного и того же множества дает возможность узнавать сколь угодно большое число его представителей.

Образы объективны в том смысле, что разные люди, обучающиеся на различном материале наблюдений, большей частью одинаково и независимо друг от друга классифицируют одни и те же объекты, что и составляет основу для взаимопонимания.

В литературе, посвященной проблеме обучения распознавания образов (ОРО), часто вместо понятия образа вводится понятие класса.

2 Проблема обучения распознаванию образов (ОРО)

В целом проблема распознавания образов состоит из двух частей: обучения и распознавания. Обучение осуществляется путем показа отдельных объектов с указанием их принадлежности тому или другому образу (классу).

В результате обучения распознающая система должна приобрести способность реагировать одинаковыми реакциями на все объекты одного образа и различными — на все объекты различных образов. Важно, что в процессе обучения указываются только сами объекты и их принадлежность образу. За обучением следует процесс распознавания новых объектов обученной системой.

Автоматизация этих процедур и составляет проблему обучения распознаванию образов. В том случае, когда человек сам разгадывает или придумывает, а затем навязывает машине правило классификации, проблема распознавания решается частично, так как основную и главную часть проблемы (обучение) человек берет на себя.

Проблема обучения распознаванию образов интересна как с прикладной, так и с принципиальной точки зрения. С прикладной точки зрения решение этой проблемы важно прежде всего потому, что оно открывает возможность автоматизировать многие процессы, которые до сих пор связывали лишь с деятельностью живого мозга. Принципиальное значение проблемы тесно связано с вопросом, который все чаще возникает в связи с развитием идей кибернетики: что

может и что принципиально не может делать машина? В частности, может ли машина развить в себе способность перенять у человека умение производить определенные действия в зависимости от ситуаций, возникающих в окружающей среде? Пока стало ясно только то, что если человек может сначала сам осознать свое умение, а потом его описать, т. е. указать, почему он производит действия в ответ на каждое состояние внешней среды или как (по какому правилу) он объединяет отдельные объекты в образы, то такое умение без принципиальных трудностей может быть передано машине. Если же человек обладает умением, но не может объяснить его, то остается только один путь передачи умения машине — обучение на примерах.

Круг задач, которые могут решаться с помощью распознающих систем, чрезвычайно широк. Сюда относятся не только задачи распознавания зрительных и слуховых образов, но и задачи распознавания сложных процессов и явлений, возникающих, например, при выборе целесообразных действий руководителем предприятия или выборе оптимального управления технологическими, экономическими, транспортными или военными операциями. В каждой из таких задач анализируются некоторые явления, процессы, состояния внешнего мира, всюду далее называемые объектами наблюдения. Прежде чем начать анализ какого-либо объекта, нужно получить о нем определенную, каким-либо способом упорядоченную информацию. Такая информация представляет собой характеристику объектов, их отображение на множестве воспринимающих органов распознающей системы. Каждое отображение какого-либо объекта на воспринимающие органы распознающей системы, независимо от его положения относительно этих органов, принято называть изображением объекта (или отражением), а множества таких изображений, объединенные какими-либо общими свойствами, представляют собой образы.

При решении задач управления методами распознавания образов вместо терминов "изображение" или "отражение" применяют термин "состояние". Состояние — это множество отображений измеряемых текущих (или мгновенных) характеристик наблюдаемого объекта в следящей системе. Совокупность состояний определяет ситуацию. Понятие "ситуация" является аналогом понятия "образ". Но эта аналогия не полная, так как не всякий образ можно назвать ситуацией, хотя всякую ситуацию можно назвать образом. Если в качестве объекта наблюдения рассматривается некоторый объект управления, то понятие "ситуация" объединяет такие состояния этого объекта, в которых следует применять одни и те же управляющие воздействия. Если объектом наблюдения является военная игра, то ситуация объединяет все состояния игры, которые требуют, например, мощного танкового удара при поддержке авиации.

Выбор исходного описания объектов является одной из центральных задач проблемы ОРО. При удачном выборе исходного описания (пространства признаков) задача распознавания может оказаться тривиальной и, наоборот, неудачно выбранное исходное описание может привести либо к очень сложной дальнейшей переработке информации, либо вообще к отсутствию решения. Например, если решается задача распознавания объектов, отличающихся по цвету, а в качестве исходного описания выбраны сигналы, получаемые от датчиков веса, то задача распознавания в принципе не может быть решена.

3 Геометрический и структурный подходы

Каждый раз, когда сталкиваются с незнакомыми задачами, появляется естественное желание представить их в виде некоторой легко понимаемой модели, которая позволяла бы осмыслить задачу в таких терминах, которые легко воспроизводятся нашим воображением. А так как мы существуем в пространстве и во времени, наиболее понятной для нас является пространственно-временная интерпретация задач.

Рис. 2: Границы между образами

3.1 Геометрический подход

Любое изображение, которое возникает в результате наблюдения какого-либо объекта, можно представить в виде вектора, а значит и в виде точки некоторого пространства признаков. Если утверждается, что при показе изображений возможно однозначно отнести их к одному из двух (или нескольких) образов, то тем самым утверждается, что в некотором пространстве существует две (или несколько) области, не имеющие общих точек, и что изображения — точки из этих областей. Каждой такой области можно приписать наименование, т. е. дать название, соответствующее образу.

Проинтерпретируем теперь в терминах геометрической картины процесс обучения распознаванию образов, ограничившись пока случаем распознавания только двух образов. Заранее считается известным лишь только то, что требуется разделить две области в некотором пространстве и что показывааются точки только из этих областей. Сами эти области заранее не определены, т. е. нет каких-либо сведений о расположении их границ или правил определения принадлежности точки к той или иной области. В ходе обучения предъявляются точки, случайно выбранные из этих областей, и сообщается информация о том, к какой области принадлежат предъявляемые точки. Цель обучения состоит либо в построении поверхности, которая разделяла бы не только показанные в процессе обучения точки, но и все остальные точки, принадлежащие этим областям, либо в построении поверхностей, ограничивающих эти области так, чтобы в каждой из них находились только точки одного образа. Иначе говоря, цель обучения состоит в построении таких функций от векторов-изображений, которые была бы, например, положительны на всех точках одного и отрицательны на всех точках другого образа.

Понятно, что это - математическая задача об аппроксимации функций. Разумеется, решение таких задач требует введения определенных ограничений на классе рассматриваемых функций, а выбор этих ограничений зависит от характера информации, которую может добавить учитель в процессе обучения. Одной из таких подсказок является гипотеза о компактности образов, которая может быть сформулирована так: "реализация одного и того же образа отображается в пространстве признаков в геометрически близкие точки".

Интуитивно ясно, что аппроксимация разделяющей функции будет задачей тем более легкой, чем более компактны и чем более разнесены в пространстве области, подлежащие разделению. Так, например, в случае, показанном на Рис. 2а, разделение заведомо более просто, чем в случае, показанном на Рис. 2б. Действительно, в случае, изображенном на Рис. 2а, области могут быть разделены плоскостью, и даже при больших погрешностях в определении разделяющей функции она все же будет продолжать разделять области. В случае же на Рис. 2б, разделение осуществляется замысловатой поверхностью и даже незначительные отклонения в ее форме приводят к ошибкам разделения. Именно это интуитивное представление о сравнительно легко делимых областях привело к гипотезе компактности.

3.2 Структурный (лингвистический) подход

Наряду с геометрической интерпретацией проблемы обучения распознаванию образов существует и иной подход, который назван структурным, или лингвистическим. Поясним лингвистический подход на примере распознавания зрительных изображений. Сначала выделяется

набор исходных понятий — типичных фрагментов, встречающихся на изображениях, и характеристик взаимного расположения фрагментов — "слева", "снизу", "внутри" и т. д. Эти исходные понятия образуют словарь, позволяющий строить различные логические высказывания, иногда называемые предположениями. Задача состоит в том, чтобы из большого количества высказываний, которые могли бы быть построены с использованием этих понятий, отобрать наиболее существенные для данного конкретного случая. Далее, просматривая конечное и по возможности небольшое число объектов из каждого образа, нужно построить описание этих образов. Построенные описания должны быть столь полными, чтобы решить вопрос о том, к какому образу принадлежит данный объект. При реализации лингвистического подхода возникают две задачи: задача построения исходного словаря, т. е. набор типичных фрагментов, и задача построения правил описания из элементов заданного словаря.

В рамках лингвистической интерпретации проводится аналогия между структурой изображений и синтаксисом языка. Стремление к этой аналогии было вызвано возможностью использовать аппарат математической лингвистики, т. е. методы по своей природе являются синтаксическими. Использование аппарата математической лингвистики для описания структуры изображений можно применять только после того, как произведена сегментация изображений на составные части, т. е. выработаны слова для описания типичных фрагментов и методы их поиска. После предварительной работы, обеспечивающей выделение слов, возникают собственно лингвистические задачи, состоящие из задач автоматического грамматического разбора описаний для распознавания изображений. При этом проявляется самостоятельная область исследований, которая требует не только знания основ математической лингвистики, но и овладения приемами, которые разработаны специально для лингвистической обработки изображений.

4 Обучение и самообучение. Адаптация и обучение

Все картинки, представленные на Рис. 1, характеризуют задачу обучения. В каждой из этих задач задается несколько примеров (обучающая последовательность) правильно решенных задач. Если бы удалось подметить некое всеобщее свойство, не зависящее ни от природы образов, ни от их изображений, а определяющее лишь их способность к делимости, то наряду с обычной задачей обучения распознаванию, с использованием информации о принадлежности каждого объекта из обучающей последовательности тому или иному образу можно было бы поставить иную классификационную задачу — так называемую задачу обучения без учителя. Задачу такого рода на описательном уровне можно сформулировать следующим образом: системе одновременно или последовательно предъявляются объекты без каких-либо указаний об их принадлежности к образам. Входное устройство системы отображает множество объектов на множество изображений и, используя некоторое заложенное в нее заранее свойство делимости образов, производит самостоятельную классификацию этих объектов. После такого процесса самообучения система должна приобрести способность к распознаванию не только уже знакомых объектов (объектов из обучающей последовательности), но и тех, которые ранее не предъявлялись. Процессом самообучения некоторой системы называется такой процесс, в результате которого эта система без подсказки учителя приобретает способность к выработке одинаковых реакций на изображения объектов одного и того же образа и различных реакций на изображения различных образов. Роль учителя при этом состоит лишь в подсказке системе некоторого объективного свойства, одинакового для всех образов и определяющего способность к разделению множества объектов на образы.

Оказывается, таким объективным свойством является свойство компактности образов. Вза-

имное расположение точек в выбранном пространстве уже содержит информацию о том, как следует разделить множество точек. Эта информация и определяет то свойство делимости образов, которое оказывается достаточным для самообучения системы распознаванию образов.

Большинство известных алгоритмов самообучения способны выделять только абстрактные образы, т. е. компактные множества в заданных пространствах. Различие между ними состоит, по-видимому, в формализации понятия компактности. Однако это не снижает, а иногда и повышает ценность алгоритмов самообучения, так как часто сами образы заранее никем не определены, а задача состоит в том, чтобы определить, какие подмножества изображений в заданном пространстве представляют собой образы. Хорошим примером такой постановки задачи являются социологические исследования, когда по набору вопросов выделяются группы людей. В таком понимании задачи алгоритмы самообучения генерируют заранее неизвестную информацию о существовании в заданном пространстве образов, о которых ранее никто не имел никакого представления. К этой группе задач относится и задача об автоматической кластеризации текстов.

Кроме того, результат самообучения характеризует пригодность выбранного пространства для конкретной задачи обучения распознаванию. Если абстрактные образы, выделяемые в процессе самообучения, совпадают с реальными, то пространство выбрано удачно. Чем сильнее абстрактные образы отличаются от реальных, тем "неудобнее" выбранное пространство для конкретной задачи.

Обучением обычно называют процесс выработки в некоторой системе той или иной реакции на группы внешних идентичных сигналов путем многократного воздействия на систему внешней корректировки. Такую внешнюю корректировку в обучении принято называть "поощрениями" и "наказаниями". Механизм генерации этой корректировки практически полностью определяет алгоритм обучения. Самообучение отличается от обучения тем, что здесь дополнительная информация о верности реакции системе не сообщается.

Обучение — это процесс, в результате которого система постепенно приобретает способность отвечать нужными реакциями на определенные совокупности внешних воздействий, а адаптация — это подстройка параметров и структуры системы с целью достижения требуемого качества управления в условиях непрерывных изменений внешних условий.

Следующая лекция

[Лекция 12. Элементы теории измерений. Таксономия](#)