В качестве модуля общения (связи) в интеллектуальной системе выступает система распознавания образов, преобразующая воздействия внешнего мира во внутренние (машинные) представления о нем. Такими внешними воздействиями могут быть, например, свет, звук, механическое давление и т. п. Соответствующие им сигналы формируются при помощи разнообразных устройств. В простейшем случае, такие сигналы формируются посредством стандартных внешних устройств компьютера: посредством клавиатуры, «мыши», модема, сканера. В других случаях таким внешним устройством может быть видеокамера, микрофон, а также датчики самых разнообразных физических величин. Часто выделение необходимой информации из поступивших сигналов не представляет трудностей. Однако, существуют ситуации, когда «понимание» поступивших сигналов, их правильная интерпретации представляет собой серьезную проблему. В частности, такая ситуация возникает при необходимости правильно интерпретировать изображения, речевые сообщения.

Далее основное внимание будет уделено проблеме правильной интерпретации компьютером внешних воздействий и выделении полезной информации из соответствующих им сигналов подобно тому, как это происходит в системах зрения, слуха, осязания живых организмов.

ОСНОВНЫЕ ПОНЯТИЯ

Образ

Образ представляет собой описание (отражение) некоторого объекта, представляемое в виде сигналов.

Сигналы поступают от сенсоров (датчиков), реагирующих на различные воздействия (свет, звук, давление и т.п.), порождаемые наблюдаемым объектом. Соответственно можно говорить о зрительных, слуховых, тактильных и т.п. сигналах, составляющих образ. Образ может быть представлен сигналами, порожденными воздействиями различной природы.

Например, кошка, в нашем сознании предстает в виде образа, составленного из порождаемой ею совокупности зрительных, слуховых, тактильных сигналов, если мы соответственно ее видим, слышим и можем дотрагиваться до нее.

Распознавание образов

Конкретная кошка чем-то похожа на всех других кошек. Конкретные яблоко, автомобиль, стол похожи чем-то на все прочие яблоки, автомобили и столы соответственно.

Именно из-за наличия у конкретных объектов некоторых общих признаков мы объединяем их в классы, обозначаемые тем или иным термином (именем): яблоко, автомобиль, стол и т.п.

Если наблюдаемый образ обладает совокупностью признаков, присущих некоторому классу, мы называем этот образ именем этого класса.

Распознавание образов есть отнесение конкретного образа к некоторому классу. Иными словами распознавание образов есть не что иное как их классификация.

Примеры:

- буква А, написанная конкретным шрифтом или почерком, распознается и относится к классу букв А, куда входят буквы А, написанные любым другим шрифтом или почерком;

- звук А, произнесенный конкретным человеком, распознается как звук А и относится к классу звуков А, имеющих различный тембр звучания, длительность, громкость и т.п.; - конкретная книга распознается как книга и относится к классу книг, в который могут входить книги различного формата, толщины, содержания;

- - отпечаток пальца или голос конкретного человека узнается и относится к конкретному человеку, как к классу, объединяющему множество возможных отпечатков пальцев или звуков голоса, принадлежащих данному человеку.

В широком смысле распознавание образов включает не только отнесение их к некоторому классу, но и выделение самих классов и тех общих свойств и признаков, которые лежат в основе выделения классов.

Признаки распознавания

Из сказанного выше следует, что распознавание основано на выделении общих свойств у образов, относящихся к данному классу.

Каждое свойство общее для всех образов данного класса называют признаком распознавания.

Например, такими свойствами (признаками) являются

для автомобиля наличие: двигателя, четырех колес, рулевого управления, посадочных мест для водителя и (возможно) пассажиров и т.д.;

для яблока наличие специфической округлой формы, окраски, запаха и т.д.

для воды наличие знакомых нам зрительных, осязательных, вкусовых ощущений, типичных для данного вещества, а также физических и химических свойств.

Признаки, используемых в технических системах распознавания, описываются и представляются на некотором языке, как правило, на языке математики.

Описывая признаки распознавания на том или ином языке, мы тем самым пытаемся формализовать осуществляемую нами процедуру распознавания. Учитывая ограниченные возможности формализации, процедура распознавания образов, реализованная техническими средствами, неизбежно будет в чем-то отличаться от реализуемой нами.

В некоторых случаях, выделение и формализованное описание признаков достаточных для надежного распознавания объектов заданного класса, оказывается трудноразрешимой, а порою и неразрешимой задачей. Задумайтесь, как нам удается узнавать знакомых людей по голосу, по старым фотографиям, запечатливших их в те годы, когда мы их не знали? Как мы узнаем на ощупь различные предметы?. Мы умеем это делать (как и многое другое), но не всегда можем объяснить и описать в строгих и точных терминах как мы это делаем, какие признаки при этом мы используем.

Определение признаков распознавания представляет собой процедуру абстрагирования – выделение наиболее существенного, стабильного, повторяющегося в образах данного класса. При этом исключается из дальнейшего рассмотрения все второстепенное, малосущественное и переменчивое.

ФОРМИРОВАНИЕ ОБРАЗА

В ТЕХНИЧЕСКИХ СИСТЕМАХ РАСПОЗНАВАНИЯ

Выбор физических эффектов и стратегии распознавания

Первой задачей, с которой сталкивается разработчик системы распознавания, является выбор физических эффектов, на которых будет основано распознавание. Такими эффектами могут быть воздействия объекта на систему, предполагающие регистрацию излучаемого (отражаемого) объектом света, звука, электромагнитных волн и т.д. Либо взаимодействия объекта и системы, предполагающие, например, облучение объекта звуковыми и электромагнитными волнами и регистрацию их отражений, измерения электро и термопроводности объекта (его фрагментов), его звуко и газопроницаемости, веса и т.п.

При этом помимо выбора используемых физических эффектов одновременно производится выбор соответственно между пассивной и активной стратегией распознавания. Стратегия распознавания в технических системах, как и у живых организмов, является активной (лизнуть, потрогать, толкнуть и оценить реакцию), если предполагает взаимодействие системы и объекта, и является пассивной (смотреть, слушать, нюхать), если предполагает одностороннее воздействие объекта на систему распознавания.

Проблема выбора совокупности физических эффектов, обеспечивающей наиболее рациональное построение системы распознавания, является трудно формализуемой задачей. Для ее решения не удается предложить сколь либо универсальный метод или систему методов. Здесь на принятие решения влияет ранее накопленный опыт проектировщика, широта его кругозора (в частности, глубина и широта его знаний физических эффектов, на которых может базироваться распознавание), его изобретательность, склонность к новаторству либо консерватизм и склонность следовать сложившиеся традиции.

Вектор образа

Далее будем полагать, что данные, поступающие от датчиков первичной информации, могут быть представлены вектором конечной размерности. Этот вектор далее будем называть вектором измерений или вектором образа.

Такой подход к формализованному представлению образа достаточно универсален.

Если первичная информация поступает от сенсорной сетчатки (например, светочувствительной), состоящей из n элементов, то результаты измерений можно представить вектором

x = (x1, …, xi ,…., xn),

где xi - интенсивность сигнала от i – го элемента сетчатки. При этом вектор x – вектор образа.

В случае, если первичная информация поступает от датчика непрерывно изменяющегося воздействия (например, от микрофона, от магнитной головки магнитофона и т.п.), то вектор образа получают в результате записи значений сигнала xi в дискретные моменты на заданном интервале времени.

Если первичная информация получается в результате измерений различных физических величин (электропроводности, упругости, магнитных свойств объекта, его веса и т. д.), вектор образа получают в результате записи в определенном порядке результатов xi проведенных измерений.

ВЫДЕЛЕНИЕ ПРИЗНАКОВ РАСПОЗНАВАНИЯ

Как уже отмечалось, выделение признаков распознавания представляет собой процедуру абстрагирования – выделение наиболее существенного, стабильного, повторяющегося в образах данного класса.

Эвристическое выделение признаков.

Во многих случаях выделение признаков является творческой задачей, решаемой проектировщиком системы распознавания без привлечения каких либо типовых методов. В таких случаях говорят о эвристическом формировании признаков распознавания.

Рассмотрим пример. Пусть среди множества изображений необходимо распознавать круг, квадрат и кольцо. Исходные данные поступают от светочувствительной сетчатки. Выделим характерные признаки, выделяющие перечисленные фигуры из множества всех других плоских фигур. При этом потребуем неизменности (инвариантности) результатов распознавания от местоположения изображения в кадре, а также от его размеров, ориентации и освещенности.

В качестве первого признака распознавания x1 можно использовать отношение периметра внешнего контура фигуры к расстоянию от центра фигуры до максимально удаленной от него точки фигуры.

Это отношение для круга равно 2pi = 6,28; для квадрата 8\*2-1/2 = 5,66; для кольца 2pi = 6,28.

Вычисление данного признака позволяет отличить круг и кольцо от квадрата, но не позволяет отличит круг от кольца. Поэтому необходимо использование дополнительного признака.

В качестве второго признака x2 можно использовать отношение квадрата периметра внешнего контура фигуры к ее площади. Это отношение для круга равно 4pi = 12,56; для квадрата 16; для кольца > 4p.

Вычисление второго признака позволяет отличать друг от друга все три фигуры. Однако, использование совместно со вторым еще и первого признака делает распознавание более надежным в условиях наличия помех, обусловленных в том числе и ограниченной разрешающей способностью сенсорной сетчатки.

В геометрической интерпретации каждому из рассматриваемых образов соответствует точка на плоскости (x1, x2). Классу образов «квадрат » - точка (8\*2-1/2; 16). Классу «круг» - точка (2pi; 4 pi). Классу «кольцо» - отрезок [2 pi; (2 pi, бесконечность)].

На практике используют приближенное определение перечисленных классов, как окрестностей указанных точек и отрезка. Такое определение класса позволяет игнорировать малые отклонения координат x1, x2 от их идеальных значений, вызванные неидеальностью формы, а также влиянием ошибок округления и помех на результат вычисления признаков реально наблюдаемой фигуры.

Очевидно, каждой точке плоскости x1, x2 соответствует своя фигура.

Для надежного распознавания фигур, более разнообразных в сравнении с рассмотренными в данном примере, указанных выше признаков может оказаться недостаточно. В таких случаях приходится привлекать дополнительные признаки.

Типовые процедуры выделения признаков.

Наряду с эвристическим подходом к формированию признаков распознавания проектировщик может использовать ряд типовых методов выявления признаков, освещенных в соответствующей литературе. Рассмотрим один из них.

Формирование признаков для распознавания полутоновых изображений

на основе Фурье - преобразований

Пусть имеется черно-белое полутоновое изображение, содержащее распознаваемый образ. При этом F1(ф) – функция изменения яркости изображения при обходе замкнутого контура в виде окружности с радиусом r1 и центром, совпадающим с центром распознаваемого образа, ф - угол в полярной системе координат, определяющий положение точки на окружности с радиусом r1. Функция F1(ф) – периодическая функция с периодом 2pi. Аппроксимируем F1(ф) отрезком ряда Фурье:

F1\*(ф)=C11sin(ф+фо11)+ C12sin(2ф+фо12)+… Ci1sin(iф+фоi1)+…+C1N sin(2ф+фо1N)=

=sum[Ci1sin(iф+фоi1)], i=1,…,N.

Использование нескольких контуров с радиусами r1,…,rj,…,rM позволяет по значениям соответствующих им коэффициентов Cij различать соответствующие образы.

Коэффициенты Cij являются, таким образом, признаками распознавания, позволяющими при достаточно больших значениях N и M надежно отличать один образ от другого.

При выборе контуров обхода изображения в виде окружностей распознавание изображения не зависит от его ориентации (поворотов) относительно начала координат. Это объясняется следующим. При повороте изображения форма графика функции Fj(ф) не изменяется. Он лишь сместится вдоль оси ф, т.е. изменится лишь фаза фоij функции Fj(ф), а значения коэффициентов Cij сохраняются неизменными.

При помещении начала координат в центре распознаваемой фигуры, значения признаков становятся инвариантными к смещению рамки экрана относительно изображения, т.к. при этом система координат остается “привязанной” к одной и той же точке изображения и, следовательно, неизменными остаются графики функций Fj(ф) и значения коэффициентов Cij.

Инвариантность значений признаков к размеру изображения может быть достигнута путем выбора значений радиусов rj контуров как определеннной и неизменной доли характерного размера изображения, например, расстояния от центра тяжести фигуры до ее максимально удаленной части.

Для обеспечения независимости процедуры распознавания от общей яркости картинки вместо функции яркости картинки Fj(ф) целесообразно использовать нормализованную функцию

F нj(ф) = Fj(ф)/Fj max,

где Fj max – наибольшее значение функции Fj(ф) при обходе j-го контура.

Увеличение числа используемых признаков позволяет расширить набор классов объектов, поддающихся распознаванию. Однако, на определенном этапе увеличение числа признаков может приводить к увеличению вероятности ошибочного распознавания.

Это ставит задачу выделения из множества возможных признаков наиболее важных. Решение этой задачи осуществляется на основе оценки для каждого признака степени влияния их изменения на решения, принимаемые системой распознавания. После чего отсеиваются те признаки, изменения которых наименее существенно сказывается на результатах.

Рассмотренные выше подходы к формированию признаков распознавания ориентированы на узнавание изображений, геометрических подобных одному из заданных эталонов. При этом признаки носят количественный характер и выражаются при помощи чисел.

Однако, существуют задачи распознавания, в которых опознание объектов основано не на их геометрическом подобии эталону, а на логическом анализе структуры изображения. Типичным примером, таких задач, является распознавание букв и цифр. Мы узнаем буквы и цифры, несмотря на их существенно различное начертание в силу различия почерков, используемых стилей и т.д. Мы узнаем их не потому, что они подобны некоторому эталону, а потому что их строение, соответствует некоторым правилам. Например, буква А по нашим представлениям обязательно должна имеет две ножки, расположенные ниже тела буквы, имеющего вид замкнутого контура. Буква В должна иметь два замкнутых контура, расположенных один под другим, причем левая грань каждого контура, должны быть образованы вертикальными прямыми линиями.

Соответственно признаки распознавания в таком случае выражаются уже не числами, а гораздо более сложными математическими структурами, которые качественно, а не количественно характеризуют распознаваемый образ.

На применении таких признаков базируются лингвистические (грамматические) методы распознавания образов.

Формирование лингвистических признаков распознавания

В рамках лингвистического анализа образа формирование признаков основано, во-первых, на выделении в изображении типовых элементов, в частности, замкнутых контуров, концевых точек, точек ветвления линий, и, во-вторых, на описании их взаимного расположения и ориентации.

Рассмотрим пример формирования признаков для лингвистического описания и распознавания буквы А.

На первом этапе производится переход от исходного полутонового изображения к контурному. При этом вначале осуществляется сегментация изображения, т.е. отнесение каждого элемента либо к образу либо к фону. С этой целью используют следующие методы.

1. Разделение по порогу яркости или степени зачерненности изображения. Если зачернение выше порогового уровня, компоненту относят к образу, если ниже, то к фону.

2. Обнаружение края. Компоненты относятся к фону либо к образу в зависимости от того, на какую сторону от границы перепада зачерненности они находятся.

3. Разделение изображения на области с одинаковыми значениями зачерненности.

Из полученного сегментированного изображения получают контурное. При этом широкие сплошные линии заменяют линиями толщиной в один пиксел, проходящими через их середину. Исключают случайные изолированные точки, сливают разорванные линии, спрямляют отдельные участки и т.д.

На втором этапе каждой точке изображения в соответствии с геометрическими свойствами ее окрестности приписывается определенный код. Код учитывает как направлена линия, проходящая через эту точку. При этом коды L, N, R, E означают соответственно направления: вверх и влево (leFt), вверх = (на север(north)), вверх и вправо (right), влево = (на восток(east)). Инверсии указанных кодов ~L, ~N, ~R, ~E указывают соответственно противоположные направления. Данный этап называют маркировкой изображения.

На третьем этапе осуществляется анализ макркированного изображения. При этом выделяют все особые точки, к которым относят: концы линий, точки излома и точки ветвления линий, т.е такие точки, с которыми соседствует точка с кодом, отличающимся от кода данной точки. После этого все особые точки заполняют прямолинейными отрезками. В результате получается граф, вершинами которого являются особые точки, а его дугами – отрезки, соединяющие вершины.

Вершины графа нумеруются в некотором порядке. Каждому такому номеру соответствует символ, обозначающий тип особой точки (концевая точка, угол, ветвление), и набор символов, характеризующих направление линий, исходящих из этой точки.

Полученный граф запоминается в виде соответствующей ему матрицы инцендентностей.

Каждый элемент полученной матрицы может рассматриваться как признак распознавания.

Важной положительной особенностью рассмотренной методики формирования признаков распознавания является то, что на каждом шаге указанного процесса все точки могут обрабатываться параллельно и независимо одна от другой.

Главным результатом, достигаемым в результате обработки данных на этапе выделения признаков, является снижение объема информации, используемой процедурой принятия решений в системе распознавания образов.