

НАУЧНО • ТЕХНИЧЕСКАЯ ИНФОРМАЦИЯ

Серия 2. ИНФОРМАЦИОННЫЕ ПРОЦЕССЫ И СИСТЕМЫ
ЕЖЕМЕСЯЧНЫЙ НАУЧНО-ТЕХНИЧЕСКИЙ СБОРНИК

Издается с 1961 г.

№ 8

Москва 2020

ИНФОРМАЦИОННЫЙ АНАЛИЗ

УДК 81'367

Н.И. Сидняев, Ю.И. Бутенко, Е.Е. Болотова

Теории формальных грамматик в методах распознавания неизвестных объектов

Рассматриваются вопросы формирования контекстных грамматик, описывающих как структурную информацию образа, так и взаимодействие образов в сложной сцене. Предлагается использование многоуровневой грамматики, включающей задачу синтаксического анализа последовательности изображений, а также задачу синтаксического анализа объектов различного назначения, когда природа источника не ясна. Показано, что формирование грамматики, описывающей как структурную информацию образа, так и взаимодействие образов, связано с необходимостью разработки алгоритма восстановления грамматики по заданному множеству динамических изображений, представляющих собой обучающую выборку. Излагаются некоторые основные положения, присущие структурным методам описания и распознавания сцены.

Ключевые слова: распознавание, грамматики, синтаксический метод, оценка, решение, образ, обучение

DOI: 10.36535/0548-0027-2020-08-1

ВВЕДЕНИЕ

Компьютерное зрение – одно из наиболее востребованных направлений в области интеллектуальных систем, включающих сложные алгоритмы распознавания и комплексные механизмы глубокого машинного обучения. Компьютерное зрение основано на воспроизведении механизмов сложной системы человеческого зрения, позволяющих получать значимую информацию из изображений, видео и других визуальных средств, а также принимать необходимые меры или давать рекомендации на основе полученной информации. До недавнего времени компьютерное зрение работало только в рамках ограниченных возможностей, однако благодаря достижениям в области искусственного интеллекта и инновациям глубокого обучения эта область в последние годы получила активное развитие и во многом превзошла человека в выполнении некоторых задач, связанных с обнаружением и распознаванием объектов [1–3]. Одним из движущих факторов развития компьютерного зрения является объем данных, который используется для обучения и улучшения подобных систем [4–7].

Проблемы, возникшие при решении задач распознавания изображений, привели к разработке различных алгоритмов, одним из которых является структурный, получивший название лингвистического или синтаксического метода [4, 5]. Его особенность заключается в том, что априорными описаниями классов являются структурные описания – формальные конструкции, при которых последовательно проводится принцип учета иерархичности структуры объекта и учета отношений, существующих между отдельными элементами этой иерархии в пределах одних и тех же уровней и между ними. Структурно-

ориентированное распознавание применяется для самых разнообразных задач: распознавание символов, речи и отпечатков пальцев, анализ сцены, химических и биологических структур [6].

Во многих случаях апостериорная информация о распознаваемых объектах хранится в записях соответствующих сигналов – электрокардиограмм, энцефалограмм, инфракрасных и акустических сигналов и т.д. Синтаксический метод распознавания позволяет обнаруживать упорядоченность структур в сигналах, порождаемых распознаваемыми объектами, и использовать для описания объектов способы комбинаторной регулярности структур, заключающиеся в том, что, оперируя весьма ограниченным количеством атомарных (непроизводных) элементов и правил комбинирования, можно получать значительное разнообразие описаний объектов с помощью неограниченного (например, рекуррентного) применения к исходным элементам правил комбинирования и результатов соответствующих рекуррентных процедур. Использование комбинаторной регулярности в качестве принципа описания структуры объектов обеспечивает экономное расходование средств [7].

На современном этапе развития интеллектуальных систем стало возможным применить теорию формальных грамматик для распознавания изображений путем описания объекта в виде иерархической структуры [8]. Теория формальных языков и грамматик была разработана в середине 1950-х гг. и применялась для описания алгоритмов обработки естественных языков, однако в последующем стало очевидно, что подобный механизм может быть перенесен и на распознавание изображений, обладающих сложной многоуровневой структурой.

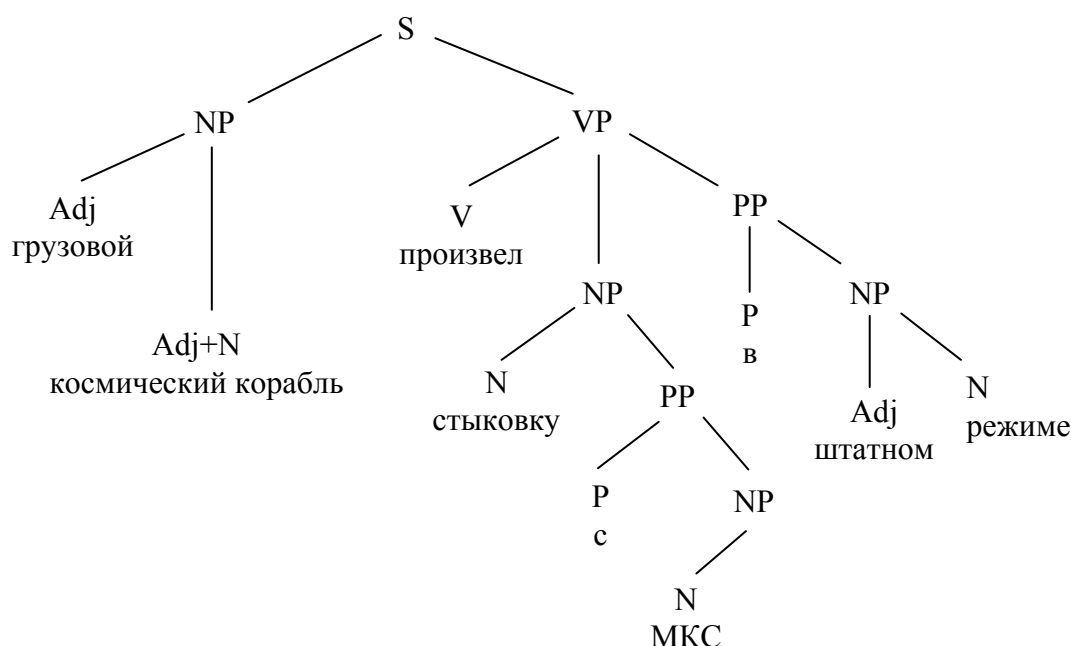


Рис. 1. Синтаксическая структура предложения

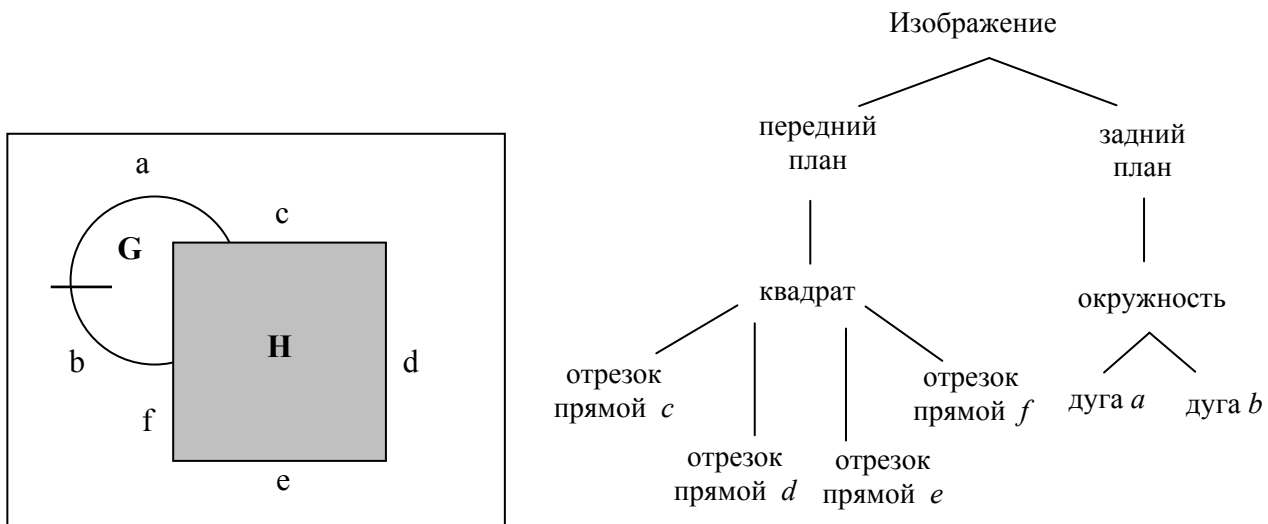


Рис. 2. Синтаксический анализ изображения

В процессе распознавания изображения выделяются такие элементы объекта, как подобразы, которые в свою очередь подразделяются на более простые подобразы, находящиеся в некоторых отношениях с другими выделенными подобразами. Это позволяет представить информацию об объекте в виде иерархической структуры, чаще всего, с помощью графа, в узлах которого находятся подобразы, а дуги обозначают отношения между выделенными подобразами распознаваемого объекта. Таким образом, структурный метод распознавания предельно схож с методом анализа синтаксической структуры предложений естественного языка [9]. Соответствие между синтагматикой и линейной упорядоченностью слов в предложении естественного языка можно представить в виде дерева составляющих. Глубинная структура пропозиционального компонента любого простого предложения представляет собой конструкцию, состоящую из глагольной (VP) и именной (NP) группы, которые находятся в некоторых семантико-синтаксических отношениях. Как показано на рис. 1. дерево составляющих позволяет представить синтаксическую структуру предложения в виде иерархии, составленной из связанных друг с другом непересекающихся единиц.

Схожим образом может быть описано распознаваемое системой изображение. На рис. 2 проиллюстрирована аналогия между синтаксическим разбором предложения на естественном языке и структурным разбором изображения, содержащего геометрические фигуры.

В рамках представленного подхода решение задачи распознавания сводится к использованию лингвистических конструкций, содержащих некоторый словарь признаков распознаваемого объекта и грамматику, по правилам которой описываются элементы этого объекта. После того как каждый подобраз внутри объекта идентифицирован, выполняется его распознавание путем синтаксического анализа предложения, описывающего этот объект в виде древовидной структуры, с целью определить, является ли

такое предложение синтаксически правильным по отношению к указанной грамматике. Шаблоны предложений определяются как выстраиваемые последовательности из выделенных подобразов различными способами так же, как и фразы предложения выстраиваются путем конкатенации слов, а слова – путем конкатенации морфем. Синтаксический подход к распознаванию образов дает возможность описать множество подобразов с помощью небольших наборов простых грамматических правил формальной грамматики [6, 9].

ОСНОВНЫЕ ПОНЯТИЯ ФОРМАЛЬНОЙ ГРАММАТИКИ

Поскольку реализация структурных методов распознавания образов основана на аппарате формальной грамматики, рассмотрим некоторые ее основные понятия, необходимые для понимания сущности структурного подхода.

Теория формальных языков в качестве самостоятельной дисциплины, как правило, берет свое начало из работ знаменитого американского лингвиста Ноама Хомского, когда в 1950-х гг. он предпринял попытки дать точную и уникальную характеристику структуры естественных языков. Его целью было определение синтаксиса языков с применением простых и точных математических правил. Позже было обнаружено, что синтаксис языков программирования может быть описан с использованием одной из грамматических моделей Хомского, называемой свободно-контекстной грамматикой [10]. Вскоре после появления современных ЭВМ стало очевидно, что все формы информации – будь то числа, имена, изображения или звуковые волны – могут быть представлены в виде записанных в строку по определенным правилам элементов. Теория формальных грамматик заняла центральное место в прикладных исследованиях в области искусственного интеллекта. Более того, с помощью аппарата формальных грамматик возможно не только описать структуру объек-

та, но и создать синтаксический анализатор для определения того, является ли предложение, описывающее объект, синтаксически верным, т.е. принадлежит ли оно конкретному языку или является грамматически неправильным. Для описания таких распознающих моделей в теории формальных языков используются отдельные формализмы, известные как теория автоматов [11].

В теории формальных языков грамматика G – это набор правил для порождения строк на формальном языке. Правила описывают, как из алфавита языка формировать строки, которые действительны в соответствии с синтаксисом языка. Алфавитом является конечный непустой набор символов. Предполагается, что символы неделимы. Слово (строка) в алфавите Σ представляет собой конечную последовательность элементов. Порождающую грамматику G определяют как производственную систему, состоящую из упорядоченной четверки $G = \langle T, N, I, P \rangle$, где T – набор терминальных символов алфавита Σ , т.е. исходные элементы словаря, N – конечный набор нетерминальных (вспомогательных) символов, I – начальный символ грамматики и P – конечный набор правил.

Построение грамматики начинается с начального символа I . Применяя правила сначала к I , а затем рекурсивно к результату предыдущего преобразования, можно сгенерировать корректное "предложение" формального языка, определенного грамматикой. Грамматика строится с помощью специальных правил подстановки выражений вида « $x \rightarrow y$ », что означает «заменить x на y » или «подставить x вместо y », где x и y – цепочки, содержащие любые терминальные или нетерминальные символы. Алгоритм преобразования останавливается, когда выражение больше не содержит нетерминальные символы алфавита. Таким образом, в соответствии с правилами грамматики объект представляется предложением в этом языке (рис. 3). Для дальнейшего описания процесса порождения необходимо введение таких понятий, как непосредственная выводимость, выводимость и язык, порождаемый грамматикой [10, 11].

Если имеются цепочки x и y , которые можно представить в виде $x = z_1 a z_2$ и $y = z_1 b z_2$, где $a \rightarrow b$ –

одно из правил грамматики G , то цепочка y непосредственно выводима из цепочки x в грамматике G . Другими словами, цепочка x может быть переработана в цепочку y за один шаг путем применения одной подстановки: x получается из y подстановкой b на место некоторого вхождения цепочки a . Это обозначается как $x \mid = y$ или $x \vdash y$.

Если имеется последовательность цепочек x_0, x_1, \dots, x_n , в которой каждая следующая цепочка непосредственно выводима из предыдущей, то цепочка x_n выводима из цепочки x_0 . Последовательность цепочек x_0, x_1, \dots, x_n называется выводом x_n из x_0 в грамматике G . Это означает, что цепочка x_0 перерабатывается в цепочку x_n необязательно за один шаг, а последовательным применением нескольких подстановок. Очевидно, что непосредственная выводимость есть частный случай выводимости. Вывод, начинающийся начальным символом и заканчивающийся цепочкой, состоящей только из терминальных символов, называется полным выводом. Естественно, не всякий вывод, начинающийся начальным символом, является полным; возможны выводы, начинающиеся начальным символом, которые невозможно продолжить до полного вывода, – тупиковые выводы. Невозможность продолжать вывод, хотя он и не является полным, т.е. не кончается цепочкой терминальных символов, определяется тем, что в соответствующей грамматике отсутствует правило, левая часть которого содержалась бы в последней цепочке данного вывода. Считаем, что от «хорошей» грамматики вовсе не требуется, чтобы любой вывод в ней заканчивался правильной терминальной цепочкой: достаточно, чтобы любой полный вывод давал правильную цепочку. В результате построения полных выводов из начального символа в грамматике получается цепочка, состоящая только из терминальных символов. Такую цепочку называют языком, порождаемым грамматикой [9–11].

Важно учитывать, что правильно построенная грамматика не должна содержать бесполезных и недостижимых символов.

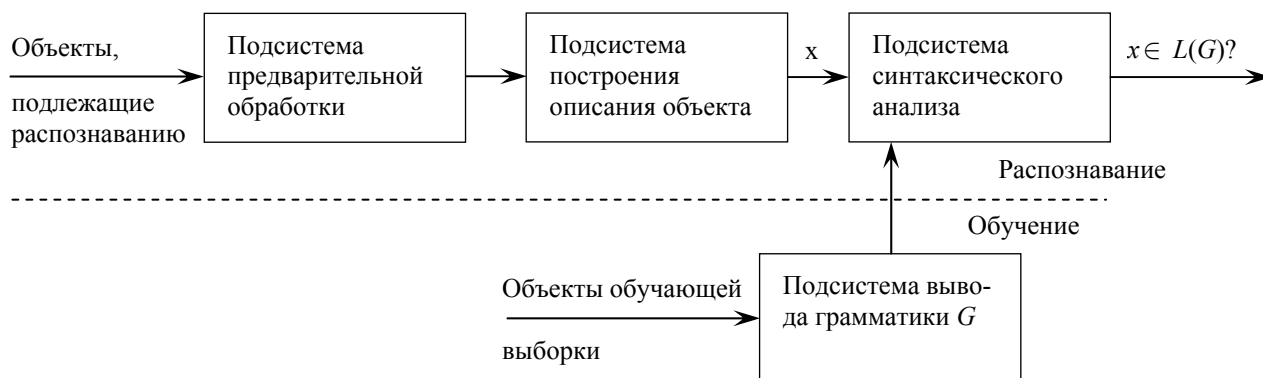


Рис. 3. Грамматика представления объекта

Определение. Символ X называется полезным в грамматике $G = \{T, N, I, P\}$, если существует вывод вида $A \Rightarrow * \alpha X \beta \Rightarrow * w$, где $w \in N^*$.

Отметим, что X может быть как переменной, так и терминалом, а выводимая цепочка $\alpha X \beta$ – первой или последней в выводе.

Определение. Если символ X не является полезным, то он называется бесполезным. Очевидно, что исключение бесполезных символов не изменяет порождаемого грамматикой языка, поэтому все такие символы можно удалить.

Определение. Символ X называется порождающим в грамматике $G = \{T, N, I, P\}$, если существует вывод вида $X \Rightarrow * w$, где $w \in N^*$.

Заметим, что каждый терминал является порождающим символом, поскольку выводится сам из себя за 0 шагов.

Определение. Символ X называется достижимым в грамматике $G = \{T, N, I, P\}$, если существует вывод вида $A \Rightarrow * \alpha X \beta$ для некоторых α, β .

Полезный символ, как следует из определения, является и порождающим, и достижимым. Если удалить из грамматики сначала непорождающие, а затем недостижимые символы, то останутся только полезные.

Пример. Рассмотрим грамматику $G = \langle \{A, B, C, a, b\}, \{a, b\}, I, P \rangle$ с правилами $A \rightarrow BC; A \rightarrow a; B \rightarrow b$.

Все символы, кроме C , являются порождающими, поскольку терминалы a и b порождают самих себя, A порождает a и B порождает b . Удаление B приводит к удалению правила $A \rightarrow BC$, т.е. к грамматике $G = \langle \{A, B, a, b\}, \{a, b\}, P, A \rangle$ с правилами $A \rightarrow a; B \rightarrow b$.

По иерархии Н. Хомского [12], формальные грамматики делятся на 4 типа, где каждый последующий тип является подмножеством предыдущего:

грамматика нулевого типа (рекурсивно перечислимая, неограниченная). Такая грамматика включает в себя формальные грамматики всех типов и допускает наличие любых правил вывода. Правила можно записать в виде: $a \rightarrow b$, где на строки a и b не устанавливаются какие-либо ограничения. Разбор предложения такого языка можно выполнять с помощью машины Тьюринга;

грамматика 1-го типа (контекстно-зависимая грамматика). Для контекстно-зависимых грамматик левая сторона правила может указывать контекст, в котором применяется определенное правило, поэтому один и тот же нетерминальный элемент может быть преобразован по-разному в зависимости от контекста, в котором он появляется. Контекстно-зависимые грамматики имеют правила вида: $xYz \rightarrow xyz$, где подстановка $Y \rightarrow y$ допустима только в контексте x, z . Разбор предложения такого языка можно выполнять с помощью линейно ограниченной машины;

грамматика 2-го типа (контекстно-свободная грамматика). Подобная грамматика генерирует контекстно-свободные языки, все правила которых принимают форму $A \rightarrow \beta$, где $A \in T, \beta \in (T \cup N)^*$. Подстановки правил внутри грамматики не зависят от контекста. Разбор предложения данного языка можно выполнять с помощью автомата с магазинной (стековой) памятью;

грамматика 3-го типа (регулярная). Правила записываются в виде $A \rightarrow a, A \rightarrow a, B$, где $A, B \in N, a \in T^*$. Разбор предложения языка третьего типа можно выполнять с использованием конечного автомата.

РЕАЛИЗАЦИЯ ПРОЦЕССА РАСПОЗНАВАНИЯ ОБЪЕКТОВ НА ОСНОВЕ СТРУКТУРНЫХ МЕТОДОВ

Для распознавания неизвестного объекта на основе структурных методов необходимо прежде всего найти его непроецируемые элементы и отношения между ними, а затем с помощью синтаксического анализа (грамматического разбора) установить, согласуется ли описание образа с грамматикой, которая, по предположению, могла его породить [13, 14].

Для формирования соответствующей грамматики можно воспользоваться либо априорными сведениями о распознаваемых объектах, либо результатами изучения конечного выборочного множества «типичных» в некотором смысле объектов. В первом случае говорят о задании грамматики на основе эвристических соображений, во втором – о выводе грамматики.

Допустим, что имеем дело с двумя классами объектов (Ω_1 и Ω_2), и объекты, входящие в эти классы, можно описать с помощью признаков, принадлежащих некоторому конечному множеству. Эти признаки можно считать основными символами (терминальными) и обозначать через T . Каждый объект может рассматриваться как цепочка или предложение, поскольку он составлен из элементов основного словаря.

Пусть также существует грамматика G такая, что порождаемый ею язык состоит из предложений (объектов), принадлежащих исключительно одному из классов, например классу Ω_1 . Эта грамматика может быть использована для классификации объектов, так как предъявленный неизвестный объект можно отнести к классу Ω_1 , если он является предложением языка $L(G)$. В противном случае объект приписывается классу Ω_2 . В данном случае объект попадает в класс Ω_2 исключительно потому, что не входит в класс Ω_1 : если оказывается, что объект не является грамматически правильным предложением в смысле грамматики G , то предполагается, что он должен принадлежать классу Ω_2 . На самом же деле объект может не относиться и к классу Ω_2 . Он может также представлять собой зашумленную или искаженную цепочку, которую в ряде случаев предпочтительно просто изъять из рассмотрения. Для этого необходимо располагать грамматиками G_1 и G_2 , порождающими языки $L(G_1)$ и $L(G_2)$ соответственно. Объект зачисляется в тот класс, в языке которого он оказывается грамматически правильным предложением. Если последнее не выполняется, то очевидно, что данный объект не принадлежит ни одному из двух заданных классов и, следовательно, требуется еще одна грамматика G_3 и т. д. В случае m классов рассматриваются m грамматик и связанных с ними языков $L(G_i), i = 1, 2, \dots, m$. Распознаваемый объект относится к классу Ω_2 в том и только в том случае, если он является грамматически правильным предложением языка $L(G_i)$. Если объект оказывается грамматически правильным предложением более чем одного языка, то решение относительно его принадлежности принимается точно таким же образом, как это делается при использовании иных методов распознавания, когда оказывается, что результаты реализации процедуры распознавания не позволяют определенно отнести объект к одному из заданных классов.

Для практического использования структурного метода требуется разрешить следующие основные проблемы:

- 1) построение адекватного описания распознаваемых объектов;
- 2) выбор грамматики;
- 3) реализация процесса распознавания посредством процедур синтаксического анализа;
- 4) использование процедур обучения для вывода грамматик;
- 5) применение в рамках структурного подхода процедур из других методов распознавания (например, статистических для учета помех и искажений случайного характера, кластер-анализа и т. д.).

Рассмотрим основные приемы, используемые при практическом применении структурных методов распознавания.

Методы математической морфологии

В процессе распознавания изображений на основе структурных методов важным является этап предварительной обработки анализируемого изображения. На этом этапе предъявленный для распознавания объект подвергается кодированию и фильтрации, восстановлению и улучшению качества изображения. Подобные методы фильтрации рассматриваются в рамках математической морфологии, обеспечивающей эффективный подход к анализу цифровых изображений [15, 16]. В рамках морфологической фильтрации используются не аналитические свойства сигналов, а их геометрические характеристики. Объект кодируется или аппроксимируется таким образом, чтобы дальше с ним было удобно работать. Например, черно-белое изображение можно закодировать с помощью сетки (или матрицы) нулей и единиц. Чтобы повысить эффективность обработки на последующих этапах, часто прибегают к какой-нибудь разновидности «сжатия данных». Затем с помощью методов фильтрации и восстановления устраняются искажения с целью улучшения качества изображения. Предполагается, что по окончании предварительной обработки воспроизводятся образы достаточно хорошего качества.

Выбор производных элементов

Первым шагом в построении структурного описания объекта распознавания является определение набора производных элементов, с помощью которых можно было бы описывать рассматриваемый объект. Этот выбор зависит от характера объекта, вида исходных данных, области приложений и способа практической реализации системы распознавания. В настоящее время не существует некоего универсального решения задачи выбора производных элементов. Обычно при выборе производных элементов используются следующими двумя основными критериями:

- 1) производные элементы должны служить основными элементами образа, чтобы обеспечивалось получение компактного и адекватного описания исходных данных с помощью вполне определенных структурных отношений (например, отношения соединения);

- 2) производные элементы должны легко поддаваться выделению или распознаванию с помощью известных методов неструктурного характера, так как предполагается, что они представляют собой простые и компактные образы, «внутренняя» структурная информация которых не принимается во внимание.

Так, для речи естественно считать фонемы «хорошим» набором производных элементов, на котором задается отношение соединения. Подобным же образом штрихи (черточки) выбираются в качестве производных элементов для представления рукописных символов.

Однако для изображения вообще нет какого-либо универсального элемента, подобного фонеме или штриху для конкретных объектов – речи или рукописных символов. Иногда для адекватного описания необходимо, чтобы производные элементы содержали информацию, существенную при распознавании объектов или явлений. Если, в частности, при решении задачи распознавания объекта существен его размер (или форма, или местоположение), то производные элементы должны нести информацию о размере (или форме, или местоположении), чтобы объекты, принадлежащие различным классам, оказывались различными независимо от того, какой метод используется для анализа соответствующих им описаний. Это условие, однако, часто приводит к необходимости прибегать к семантической информации при описании производных элементов.

Пример. Допустим, необходимо уметь отличать прямоугольники различных размеров от других фигур.

Выбирается следующий набор производных элементов: $a' - 0^\circ$ – отрезок горизонтальной линии; $b' - 90^\circ$ – отрезок вертикальной линии; $c' - 180^\circ$ – отрезок горизонтальной линии; $d' - 270^\circ$ – отрезок вертикальной линии.

Множество всех возможных прямоугольников (различных размеров) задается с помощью одного предложения – цепочки $a'b'c'd'$ (рис. 4). Если же требуется различение прямоугольников различных размеров, то такое описание оказывается неадекватным.

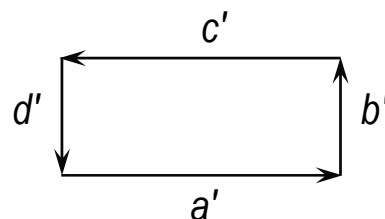


Рис. 4. Пример предложения для описания прямоугольника

В этом случае в качестве производных элементов следует использовать отрезки единичной длины. Множество прямоугольников различных размеров можно описывать с помощью языка:

$$L = \{a^n b^m c^n d^m, n, m = 1, 2, \dots\}.$$

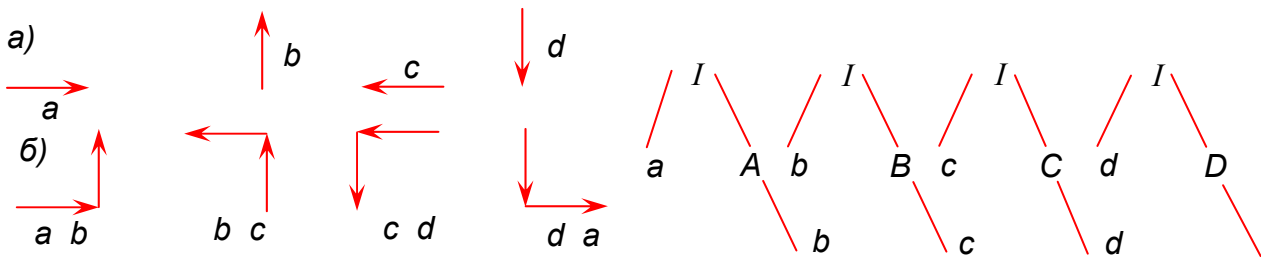


Рис. 5. Пример вывода с использованием дерева

В некоторых случаях второй критерий выбора производных элементов может вступать в противоречие с первым критерием из-за того, что производные элементы, выбранные в соответствии с этим первым критерием, могут плохо поддаваться распознаванию с помощью известных в настоящее время методов. Однако второй критерий допускает выбор достаточно сложных производных элементов, если только они могут быть распознаны. Усложнение производных элементов позволяет упрощать структурные описания, т. е. позволяет пользоваться менее сложными грамматиками. Отыскание компромиссного решения имеет существенное значение при построении реальных систем распознавания, основанных на использовании структурных методов.

Использование грамматик и языков для структурного описания объектов

Граматику можно использовать для порождения предложений, представляющих собой некий объект, и для грамматического разбора предложений, чтобы установить, является ли структура этих предложений допустимой с точки зрения соответствующей грамматики.

Пример. Рассмотрим грамматику

$G = \{T, N, I, P\}$, где $T = \{a, b, c, d\}$; $N = \{I, A, B, C, D\}$;

$P: I \rightarrow aA \quad A \rightarrow b \quad I \rightarrow cC \quad C \rightarrow d$

$I \rightarrow bB \quad B \rightarrow c \quad I \rightarrow dD \quad D \rightarrow a$.

В данном случае можно вывести четыре предложения (\Rightarrow – символ вывода):

$I \Rightarrow aA \Rightarrow ab; \quad I \Rightarrow cC \Rightarrow cd;$

$I \Rightarrow bA \Rightarrow bc; \quad I \Rightarrow dD \Rightarrow Da.$

Если производными элементами считать ориентированные отрезки, изображенные на рис. 5а, то с помощью этой грамматики можно получить четыре описания отрезков, образующих прямой угол (рис. 5б). Здесь использован прием, позволяющий описывать двумерные объекты с помощью цепочечных грамматик. Этот прием заключается в соединении структур только в особых точках. Один из способов реализации этого требования заключается в том, что в каждой структуре выделяются только две точки. В нашем случае выделенные точки интерпретируются как «головой» и «хвостовой» концы стрелки.

Деревья вывода

В информатике под деревом понимают широко используемый абстрактный тип данных, аналогичный иерархической структуре дерева, с корневым значением и поддеревьями – дочерними элементами с родительским узлом, представленными в виде набора связанных узлов [17, 18]. Каждый узел дерева имеет определенную степень, характеризующую число поддеревьев узла. Узел с нулевой степенью называется листом. Листья представляют собой узлы, из которых не выходит ни одной ветви.

Процесс вывода предложения, описывающего прямоугольник в виде дерева разбора, изображен на рис. 5. Начальный символ I выполняет роль корня дерева, концевые узлы, прочитываемые слева направо, образуют выведенное предложение.

В сущности, любая иерархически упорядоченная схема позволяет представить соответствующий объект в виде дерева. Так, на рис. 6 упорядочение состоит в группировке областей распознаваемого объекта, причем область b находится в области a , в свою очередь находящейся в области r . На рис. 6б изображена древовидная структура, естественно вытекающая из этой схемы упорядочения. Каждое дерево вывода грамматики задает одно предложение. Большинство грамматик позволяет получать значительное, а при введении вероятностных правил вывода – и бесконечное, число предложений.

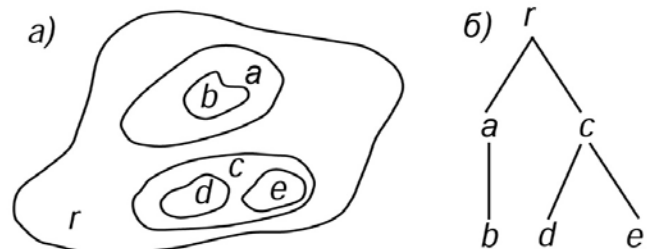


Рис. 6. Упорядоченная схема объекта:

а) – группировка областей распознаваемого объекта;
б) – дерево вывода

В простейшем случае в таких деревьях от каждого узла отходят две ветви: одна соответствует наличию у объекта определенного признака, другая – его отсутствию. Подобные деревья не несут какой-либо информации о процессе вывода решений в указанном выше смысле – они дают иерархическое представление о том, какие признаки (в случае структурного распознавания – какие производные элементы и отношения) характеризуют изучаемый объект. Однако можно сформировать грамматику, описывающую процесс вывода в соответствии со структурой дерева решений (обычно в рамках этого метода строится и используется только дерево).

Процедуры распознавания, связанные с использованием деревьев решений, относят к группе структурных методов, хотя в строгом смысле деревья решений являются инструментом иерархического метода разделения [18], который используется в случаях, когда дерево решений «содержит» много признаков. В каждом узле дерева изучается один признак и в зависимости от результатов этого изучения определяется очередная ветвь дерева. Результат классификации выявляется на нижнем ярусе дерева. Подобная схема распознавания очень удобна для учета априорных сведений об объектах, однако для нее отсутствуют оптимальные процедуры обучения.

РЕАЛИЗАЦИЯ ПРОЦЕССА РАСПОЗНАВАНИЯ ОБЪЕКТОВ С ПОМОЩЬЮ ГРАММАТИЧЕСКОГО РАЗБОРА

Значение формальных грамматик для распознавания объектов заключается, в частности, в том, что они позволяют видеть синтаксическую правильность или неправильность применительно к определенной грамматике представления изучаемого образа с помощью заданных производных элементов и отношений. Подобные элементы и отношения позволяют получить процедура [19], использующая два основных вида грамматического разбора: разбор сверху–вниз и разбор снизу–вверх. Разбор сверху–вниз состоит в последовательных попытках путем использования правил соответствующей грамматики получить заданное терминальное предложение исходя из начального символа. При разборе снизу–вверх необходимо восстанавливать дерево вывода, начиная с элементов основного словаря и применяя инвертированные правила подстановки. Эта процедура начинается с конкретного предложения и заканчивается получением начального символа. Так, производный элемент a , входящий в предложение ab , может быть получен как в результате перехода $D \rightarrow a$, так и в результате перехода $I \rightarrow aA$; элемент D возникает в результате подстановки $I \rightarrow dD$, так что этот вывод не приводит к получению предложения ab . Так как подстановки $I \rightarrow aA$ и $A \rightarrow b$ действительно порождают предложение ab , то это предложение считается грамматически правильным; предложения же ac и ad отклоняются. Описанные схемы грамматического разбора, в принципе, неэффективны, так как предусматривают проведение полного перебора при применении правил подстановки. Часто можно избежать воспроизведения всей последовательности применения правил подстановки, так как можно проверять допустимость промежу-

точных результатов, выясняя тем самым, обеспечивает ли данная последовательность правил успешный грамматический разбор.

Повышение эффективности грамматического разбора можно обеспечить с помощью синтаксических правил, устанавливающих допустимые или запрещающие какие-то специфические отношения между объектами.

Пример. Рассмотрим грамматику $G = \{T, N, I, P\}$, где $T = \{a_1, a_2\}$, $N = \{I, O_1, O_2\}$, $P: I \rightarrow$ выше $(a_1, O_2), O_2 \rightarrow$ выше $(O_1, a_1), O_1 \rightarrow$ слева (a_2, a_1) ,

способную порождать квадраты (непроизвольные элементы, представляющие собой горизонтальный и вертикальный отрезки прямой фиксированной длины); отношения выше (x, y) и слева (x, y) читаются соответственно « x расположен над y » и « x расположен слева от y », при этом часть x обязательно находится непосредственно над отрезком y и часть отрезка x во втором случае находится непосредственно слева от отрезка y .

Последовательное применение правил грамматики приводит к получению структур, построение которых начинается с горизонтального отрезка, затем вводится еще один горизонтальный отрезок, расположенный над первым, после чего между ними помещаются два вертикальных отрезка (см. рис. 2). Разнообразие конструкций обеспечивается наложением ограничений на позиционные отношения выше и слева.

На этом же примере покажем порядок выполнения синтаксически ориентированного грамматического разбора. Пусть требуется установить, принадлежит ли некоторая конструкция к классу объектов, порождаемых данной грамматикой. При разборе снизу–вверх в первую очередь делается попытка обнаружить объект O_1 с помощью вхождения в объект производного элемента a_2 слева от производного элемента a_2 . В положительном случае разбор продолжается, в отрицательном – исследуемый объект отклоняется. Поскольку процедура разбора снизу–вверх начинается с терминального предложения, на первом шаге рассматриваются только правила подстановки, приводящие исключительно к производным элементам. На следующем шаге разбора необходимо получить объект O_2 , состоящий из объекта O_1 , расположенного над производным элементом a_1 . При положительном результате предпринимается попытка вывести начальный символ I , отыскивая производный элемент a_1 , расположенный над объектом O_2 . При положительном результате образ считается допустимым, при отрицательном – отклоняется.

Вывод грамматик

Рассмотрим на конкретном примере способ получения грамматики, пригодной для описания объектов в рамках структурного подхода к распознаванию, с помощью процедуры обучения. В сущности, искомая грамматика формируется на основании изучения конечной выборки предложений некоторого языка, причем иногда при этом учитывается и конечная выборка предложений, не относящихся к данному языку. Элементы первой выборки должны обладать

структурной полнотой в том смысле, что каждое правило искомой грамматики используется при порождении по крайней мере одного предложения выборки.

На рис. 6 представлена схема вывода цепочечных грамматик. Задача заключается в том, что множество выборочных цепочек $\{x_i\}$ подвергается обработке с помощью адаптивного обучающего алгоритма (*алгоритмом вывода*). На выходе блока в результате процесса обучения воспроизводится грамматика G , согласованная с заданными цепочками. Процедура, позволяющая решить указанную задачу в общем случае, до сих пор неизвестна; существует, однако, множество частных алгоритмов. Типичным подходом можно считать следующий: сначала строятся правила подстановки, порождающие именно те цепочки, которые были заданы; затем посредством сращивания элементов основного словаря – более простая рекурсивная грамматика, обеспечивающая порождение бесконечного числа цепочек. Процедура состоит из трех этапов:

- 1) формирование правил подстановки;
- 2) преобразование правил подстановки в грамматику;
- 3) упрощение грамматики.

На примере двумерной грамматики продемонстрируем её вывод, поскольку, как указывалось выше, именно этот тип грамматик представляет в настоящее время наибольший интерес с прикладной точки зрения. Предполагается, что заданы конечный набор отношений, конечный список элементов основного словаря, т. е. непроединных элементов, и некоторое множество описаний объектов $\{S_1, \dots, S_n\}$. Каждое описание содержит список входящих в него непроединных элементов, каждый из которых может быть снабжен какой-либо дополнительной информацией. Задача состоит в определении «хорошего» набора правил грамматики, записанных через заданные отношения и согласующихся с заданным набором объектов. Таким образом, следует сформировать такую грамматику, чтобы при введении в грамматический анализатор любого из заданных объектов и разборе его в соответствии с полученной грамматикой на выходе анализатора воспроизводилось одно или несколько структурных описаний объекта.

Первый шаг процесса вывода выполняется отдельно и независимо для каждого «входного» объекта. В результате этого должна выявиться структура, задаваемая выбранной системой отношений – в данном случае предикатов, характеризующих как свойства непроединных элементов, так и отношения между ними. На этом же шаге для каждого элемента объектов проверяется справедливость их предикатов. При положительном исходе такой проверки формируется новый элемент и информация о соответствующих импликантах и истинности предиката фиксируется в виде некоторого правила грамматики. Затем выполняется следующий цикл, в котором предикаты проверяются на элементах, построенных из новых элементов и элементов входных объектов; процедура повторяется до тех пор, пока построение новых элементов оказывается невозможным. В каждом очередном цикле, естественно, проверяются лишь те описания, которые содержат хотя бы один элемент, сформированный в предыдущем цикле [20].

По окончании первого шага проверяется, какие из элементов включают в качестве своих импликант все непроединные элементы соответствующего входного объекта. Все не прошедшие проверку описания и описания, входящие в качестве импликант в описания, прошедшие проверку, из дальнейшего рассмотрения исключаются. Соответственно исключаются и все правила грамматики, не использовавшиеся при построении оставленных описаний. Полученные таким образом описания являются некоей разновидностью деревьев структурных описаний, которые сначала подвергаются «чистке» (в частности, отождествляются деревья, оказывающиеся эквивалентными в результате замены конкретных непроединных элементов их типами).

Для каждого входного объекта таким образом строится одна или несколько грамматик, соответствующих различным структурным описаниям, получаемым для входного объекта. Высшему уровню описания – описанию объекта в целом соответствуют лишь те элементы вспомогательного словаря, которые встречаются более чем в одной из построенных грамматик.

Второй шаг процесса вывода состоит в следующем. Строится объединение грамматик, полученных на первом шаге для каждого из объектов. К каждой из этих грамматик применяется специальная процедура преобразования и согласно некоторому критерию из полученных модифицированных грамматик выбирается наилучшая. Предпочтение, естественно, отдается коротким грамматикам, поскольку они действительно являются единой грамматикой для заданной обучающей выборки, а не просто набором частных случаев. Однако целесообразно, чтобы правила подстановки были как можно конкретнее в том, что касается возможных типов непроединных элементов и отношений. Очевидно, что окончательный критерий выбора грамматики должен базироваться на какой-то комбинации двух этих эвристических приемов.

Преобразование грамматики

Для эффективного распознавания неизвестных объектов с помощью формальной грамматики, её необходимо сократить, т. е. избавиться от избыточных правил. Для этого применяется процедура преобразования грамматики, которая включает несколько основных этапов:

1) устраняются все избыточные вхождения правил, если какое-то правило входит в преобразуемую грамматику G несколько раз;

2) отыскивается такая пара элементов вспомогательного словаря, что отождествление этих элементов (замена одного другим во всей грамматике) приведет к появлению избыточных вхождений правил. При наличии различных вариантов выбора таких пар следует выбирать пару, приводящую к наибольшему числу избыточных вхождений. После отождествления правил следует вернуться к первому этапу;

3) отыскивается такая пара N, n , включающая элемент вспомогательного словаря N и непроединный элемент n , что включение в грамматику G правила $N \rightarrow n$ (если оно в нее не входит) и выборочная замена N на n в грамматике Γ приводят после устранения множественности вхождений правил к сокращению

их числа. При наличии различных вариантов выбора таких пар следует выбирать пару, обеспечивающую наибольшее сокращение числа правил в грамматике. Затем следует вернуться к первому этапу;

4) вводится порог по числу отождествляемых на этапах 1–3 правил: замена производится, если число вхождений или отождествляемых правил не меньше значения порога. Это делается для того, чтобы не снижать существенно разделяющую силу грамматики;

5) вводится порог по числу правил, оказывающихся идентичными во всем (исключение составляет тип непроизводного элемента): если число таких правил больше значения порога или равно ему, то соответствующие правила заменяются одним, содержащим непроизводный элемент произвольного типа. Затем следует вернуться к первому этапу.

Если дальнейшие упрощения невозможны, то процедура прекращается. В результате выводится одна грамматика. После того как такая процедура выполняется для каждой из сформированных на первом этапе грамматик, необходимо выбрать «наилучшую» грамматику. В качестве критерия оптимальности грамматики может быть использовано, например, отношение разделяющей силы грамматики к квадрату числа ее правил, а разделяющая сила грамматики определяется как сумма разделяющих сил каждого из ее правил, которые определяются как сумма входящих в него непроизводных элементов и термов, представляющих собой отношение или свойство.

В качестве примера рассмотрим ситуацию, когда предъявляется только один входной объект, представленный на рис. 7. Здесь непроизводные элементы принадлежат к следующим типам: окружность, точка, квадрат, отрезок прямой. Предположим, заданы следующие предикаты: слева (x, y) (выполняется, если элемент x находится слева от элемента y); выше (x, y) (выполняется, если элемент x находится выше элемента y); внутри (x, y) (выполняется, если элемент x находится внутри элемента y). В данном случае входной объект представляется перечнем, включающим пять непроизводных элементов: точка (два), окружность, квадрат и отрезок прямой. Вначале выделяются все случаи, когда элементы входного списка удовлетворяют одному из заданных предикатов. Каждый выполненный предикат образует новый элемент. Далее проверяется, какие новые элементы можно построить из исходных вновь образованных, затем эта процедура повторяется с очередными построенными элементами и так до тех пор, пока возможности построения новых элементов не окажутся исчерпанными. После этого удаляются все элементы, не являющиеся компонентами некоторого элемента, содержащего все непроизводные элементы входного объекта. Практически это выглядит следующим образом. Непроизводные элементы снабжены номерами 1–5. Пусть, например, обнаружен элемент выше (4, 5) – соответствующий предикат «выше (4, 5)» выполняется. Выполненный предикат образует новый элемент, компонентами которого служат непроизводные элементы 4 и 5.

Каждое из полученных таким образом структурных описаний непосредственно переводится в грамматику (при этом символы G_1, G_2, G_3 , используются по мере необходимости для обозначения элементов

вспомогательного словаря, а элементы основного словаря заменяются непроизводным элементом определенного типа).

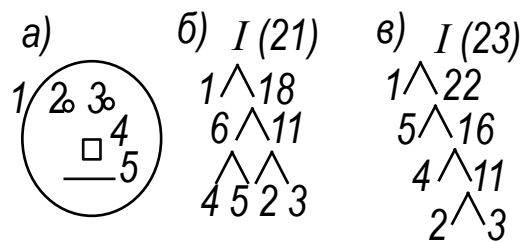


Рис. 7. Представление предикатов с использованием грамматик.

Итак, две построенные грамматики имеют следующий вид:

грамматика 1) $I \rightarrow (x, y): G_1(x)$, окружность (y): внутри (x, y): $G_1 \rightarrow (x, y): G_2(x)$, отрезок прямой (y): выше (x, y): $G_2 \rightarrow (x, y): G_3(x)$, квадрат (y): выше (x, y): $G_3 \rightarrow (x, y)$: точка (x), точка (y): слева (x, y);

грамматика 2) $I \rightarrow (x, y): G_4(x)$, окружность (x): внутри (x, y): $G_4 \rightarrow (x, y): G_5(x), G_6(y)$: выше (x, y): $G_5 \rightarrow (x, y)$: точка (x), точка (y): слева (x, y): $G_6 \rightarrow (x, y)$: квадрат (x), отрезок прямой (y): выше (x, y).

Структурный подход к распознаванию объектов не располагает еще стройной и строгой математической теорией. Основные области приложения структурного распознавания в настоящее время – обработка и анализ сигналов различной природы, двумерных и трехмерных изображений. Практика построения систем распознавания свидетельствует о том, что структурные методы целесообразно использовать совместно с другими методами распознавания объектов и явлений. Степень применимости структурных методов будет определяться развитием их математической теории и наличием ориентированных на него нейронных сетей.

Можно отметить новые нормальные формы контекстно-свободных грамматик, в которых все правила имеют одну из двух форм:

$$U \rightarrow u;$$

$$U \rightarrow TN,$$

где U, N, T – нетерминальные символы, u – терминальный символ. Эта форма называется нормальной формой Хомского.

ЗАКЛЮЧЕНИЕ

Для распознавания многоуровневых и плохо структурируемых объектов необходимо выбрать наилучшую грамматику. Распознавание неизвестного объекта на основе структурных методов требует прежде всего нахождения его непроизводных элементов и отношений между ними, а затем с помощью синтаксического анализа (грамматического разбора) установления, согласуется ли описание образа с

грамматикой, которая, по предположению, могла его породить. Для формирования соответствующей грамматики можно воспользоваться априорными сведениями о распознаваемых объектах или результатами изучения конечного выборочного множества понятных в некотором смысле объектов. Каждый объект рассматривается как цепочка или предложение, поскольку он составлен из элементов основного словаря. Представленные грамматики могут быть использованы для классификации объектов. Проявленный неизвестный объект можно отнести к определенному классу, если он является предложением языка. В противном случае объект приписывается другому классу. Как правило, объект зачисляется в тот класс, в языке которого он оказывается грамматически правильным предложением. Если это условие не выполняется, то очевидно, что данный объект не принадлежит ни одному из заданных классов и, следовательно, требуется еще одна грамматика и т. д. Расознаваемый объект относится к интересующему классу в том и только в том случае, если он является грамматически правильным предложением языка. Если объект оказывается грамматически правильным предложением более чем одного языка, решение относительно его принадлежности принимается точно таким же образом, как это делается при иных методах распознавания, когда оказывается, что результаты реализации процедуры распознавания не позволяют определенно отнести объект к одному из заданных классов.

В качестве критерия оптимальности грамматики может быть использовано отношение разделяющей силы. Слова могут объединяться в более сложные конструкции – предложения. Здесь язык рассматривается как множество предложений. Предложения строятся из слов и более простых предложений по правилам синтаксиса. Синтаксис языка представляет собой описание правильных предложений. Алфавит, лексика и синтаксис полностью определяют набор допустимых конструкций языка и внутренние взаимоотношения между конструкциями. Набор правил синтаксиса образует грамматику языка. Правила синтаксиса могут описывать либо процедуру получения правильных предложений, либо процедуру распознавания «правильности» предложений (т. е. их принадлежности данному языку). В первом случае грамматика будет порождающей, во втором – распознающей. Критерием выбора метода распознавания объекта может служить простота определения меры близости, сложности списания границ классов и образов, разрешающая способность и т. д. При этом весьма важно знать индивидуальные физические особенности распознаваемых объектов, информативность выбранных признаков, количество и качество априорной и текущей информации, возможность введения адапционных (весовых) коэффициентов и др.

СПИСОК ЛИТЕРАТУРЫ

1. Иванько А.Ф., Иванько М.А., Горчакова Я.В. Методы распознавания образов и задачи логического выделения объектов // Научное обозрение. Технические науки. – 2019. – № 3. – С. 36-40.
2. Ситников В.В., Люминарский В.В., Коробейников А.В. Обзор методов распознавания объектов, используемых в системах машинного зрения // Вестник ИжГТУ им. М.Т. Калашникова. – 2018. – Т. 21, № 4. – С. 222-229.
3. Костылев Д.А., Федотов О.В. Машинное зрение в робототехнических системах // Наука, техника и образование. – 2016. – № 7. – С. 55.
4. Васендина И.С. и др. Разработка методики распознавания образов на изображениях на основе структурного подхода // Вестник Брянского государственного технического университета. – 2017. – № 1(54). – С.171-177.
5. Прокофьева Е.В., Прокофьева О.Ю. Кластерный, корреляционный и структурно-лингвистический методы в распознавании образов // Современные исследования в сфере естественных, технических и физико-математических наук. – 2018. – С. 679-685.
6. Хаустов П.А. Алгоритмы распознавания рукописных символов на основе построения структурных моделей // Компьютерная оптика. – 2017. – Т. 41, № 1.
7. Сидняев Н.И., Бутенко Ю.И., Болотова Е.Е. Экспертная система производственного типа для создания базы знаний и конструкций летательных аппаратов // Авиакосмическое приборостроение. – 2019. – №6. – С. 38-52.
8. Фаворская М.Н. К вопросу об использовании формальных грамматик при распознавании объектов в сложных сценах // Решетневские чтения: материалы XIII междунар. науч. конф. Ч. 2. – Красноярск, 2009. – С. 540-541.
9. Singh M., Khan H., Gupta R. Digital Image Processing a Formal Grammar Approach // International Journal for Research in Applied Sciences and Biotechnology (IJRASB). – 2018. – Vol. 5, №2. – С. 3-5.
10. Ходжабекян М.С., Авраменко А.А., Зейтунян В.М. Порождающая грамматика Н. Хомского // Развитие науки и техники: механизм выбора и реализации приоритетов. – 2017. – С. 92.
11. Глушков В.М. Абстрактная теория автоматов // Успехи математических наук. – 1961. – Т. 16, № 5(101). – С. 3-62.
12. Рубцов А.А. Обобщение иерархии Хомского // Дискретные модели в теории управляющих систем. – 2018. – С. 234-237.
13. Гороховатский В.А. Структурный анализ и интеллектуальная обработка данных в компьютерном зрении: монография. – Харьков: Компания СМИТ, 2014. – 316 с.
14. Тупиков В.А., Павлова В.А., Крюков С.Н., Созинова М.В., Шульженко П.К. Лингвистические методы в задачах распознавания изображений // Известия ТулГУ. Технические науки. – 2015. – №11-2. – С. 28-37.
15. Визильтер Ю.В. Обобщенная проективная морфология // Компьютерная оптика. – 2008. – Т. 32. – № 4.
16. Шашев Д.В., Шидловский С.В. Морфологическая обработка бинарных изображений с ис-

пользованием перестраиваемых вычислительных сред // Автометрия. – 2015. – Т. 51, № 3. – С. 19-26.

17. Сидняев Н.И., Храпов П.В. Нейросети и нейроматематика: учебное пособие / под ред. Н.И. Сидняева. – М.: Изд-во МГТУ им. Н.Э. Баумана, – 2016. – 83 с.
18. Финогеев А.Г., Четвергова М.В. Разработка и исследование методики распознавания изображений для систем расширенной реальности // Известия Волгоградского государственного технического университета. – 2012. – Т. 15, № 15. – С. 130-136.
19. Фаворская М.Н., Горошкин А.Н. Модель распознавания изображений рукописного текста // Сибирский журнал науки и технологий. – 2008. – № 2(19). – С. 52-57.
20. Соловьев С.Ю. Эквивалентные преобразования контекстно-свободных грамматик // Информационные процессы. – 2010. – Т. 10, № 3. – С. 292-302.

Материал поступил в редакцию 24.05.20.

Сведения об авторах

СИДНЯЕВ Николай Иванович – доктор технических наук, профессор, заведующий кафедрой «Высшая математика» Московского государственного технического университета имени Н.Э. Баумана (национальный исследовательский университет)
e-mail: sidnyaev@yandex.ru

БУТЕНКО Юлия Ивановна – кандидат технических наук, доцент, доцент кафедры «Романо-германские языки» Московского государственного технического университета имени Н.Э. Баумана (национальный исследовательский университет)
e-mail: iuliiabutenko2015@yandex.ru

БОЛОТОВА Елизавета Евгеньевна – студент Московского государственного технического университета имени Н.Э. Баумана (национальный исследовательский университет)
e-mail: bolotovae@mail.ru