

Министерство образования Республики Беларусь
БЕЛОРУССКИЙ НАЦИОНАЛЬНЫЙ ТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ

Кафедра «Системы автоматизированного проектирования»

СТРУКТУРНЫЙ ПОДХОД
ПРИ РАСПОЗНАВАНИИ ИЗОБРАЖЕНИЙ

Методические указания
к лабораторной работе для студентов специальности
1-40 01 02 «Информационные системы и технологии»

Минск
БНТУ
2011

УДК 004.932.2 (076.5)(0/75.8)

ББК 32.97я7

С 87

Составитель *И.Л. Ковалева*

Рецензенты:

Н.Н. Гурский, В.Т. Придухо

Методические указания касаются одного из разделов теории обработки и распознавания изображений – структурного подхода. В теоретической части вводятся основные понятия структурного подхода, определяются его этапы.

Изучение теоретического материала, изложенного в указаниях, и выполнение лабораторной работы дает студентам возможность понять правила решения двух основных задач структурного подхода: задачи построения исходного словаря, т. е. наборов характерных элементов и описаний их взаимного расположения, и задачи построения правил описания нового объекта из элементов заданного словаря.

Лабораторная работа направлена на освоение студентами изложенного материала в ходе самостоятельной разработки приложения для экспериментального исследования заданного типа изображений, меры близости и способа распознавания.

Цель работы: изучить основы структурного подхода при распознавании изображений.

1. Применение структурного подхода при решении задач распознавания образов

Математические методы решения задач распознавания образов разделяются на две группы, одну из которых можно трактовать с позиций теории решений (дискриминантный подход), а другую – в рамках структурного (или синтаксического) подхода. В первом подходе объекты описываются значениями признаков, характеризующих их. При структурном подходе объекты описываются не множеством числовых значений признаков, а структурой. Структурные методы обладают малой чувствительностью к сдвигу, растяжению, сжатию, повороту, изменению наклона и другим искажениям.

Структурный подход применяется к задачам распознавания образов, в которых важна информация, описывающая структуру каждого объекта. От процедуры распознавания требуется, чтобы она давала возможность не только отнести объект к определенному классу (классифицировать его), но и описать те его стороны, которые исключают отнесение к другому классу. Типичным примером таких задач служит распознавание изображений или, говоря шире, анализ сцен. Рассматриваемые в этом классе задач объекты обычно сложны и число требуемых признаков часто велико. Это делает привлекательней идею описания сложного объекта в виде иерархической структуры более простых объектов.

Иерархия предполагает описание сложных объектов с помощью более простых подобъектов. Те, в свою очередь, могут быть описаны с помощью подобъектов следующего уровня и т. д. Этот подход основан на аналогии между структурой объектов и синтаксисом языков. Он приемлем тогда, когда простейшие подобъекты вычленять и распознавать легче, чем изображение (объект) в целом.

На рис. 1 представлено изображение и описание его иерархической структуры.

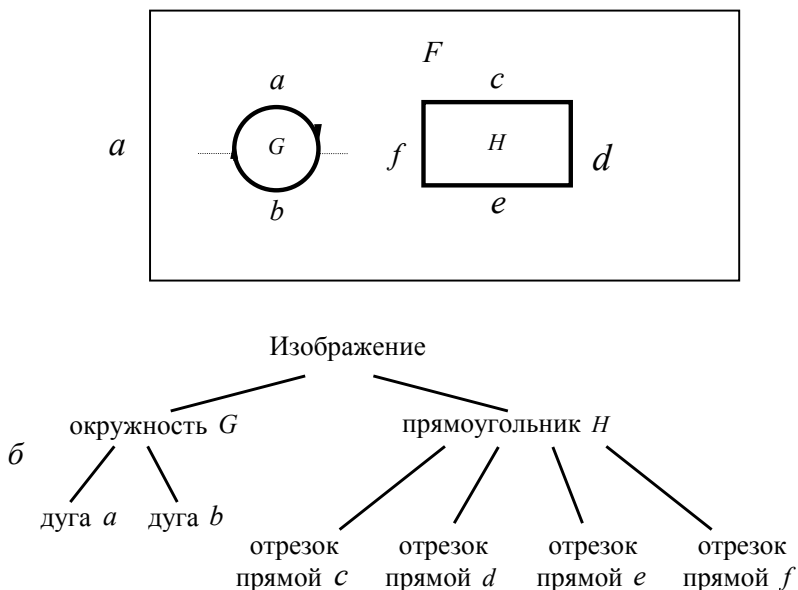


Рис. 1. Изображение (а) и его иерархическое структурное описание (б)

Правила композиции простейших (непроизводных) элементов при описании объекта в целом называют *грамматикой языка описания объектов*. Таким образом, каждый объект представляется совокупностью непроизводных элементов, «соединенных» между собой теми или иными способами, или, другими словами, «предложением» некоторого «языка». Путем синтаксического анализа (грамматического разбора) «предложения» устанавливается его синтаксическая «правильность» или, что эквивалентно, определяется, может ли некоторая фиксированная грамматика, описывающая класс, породить имеющееся описание объекта. Грамматический разбор производится так называемым «синтаксическим анализатором», который

представляет полное синтаксическое описание объекта в виде дерева грамматического разбора, если объект является синтаксически правильным, т. е. принадлежит классу, описываемому данной грамматикой. В противном случае, объект либо отклоняется, либо подвергается анализу с помощью других грамматик, описывающих иные классы объектов. Известны бесконтекстные, автоматные и другие типы грамматик. Однако задача восстановления (определения) грамматик по некоторому множеству высказываний (предложений-описаний объектов), порождающих данный язык, является трудно формализуемой. В литературе приводится описание эвристических правил автоматического восстановления грамматик для конструирования и применения лингвистических алгоритмов распознавания образов.

Преимущество структурного подхода проявляется в том случае, если удастся большое количество сложных объектов представлять с помощью небольшого множества непроеизводных элементов и грамматических правил (например, распознавание устных слов по последовательности фонем).

На рис. 2 представлен пример описания объекта (а) при помощи операции композиции «составление цепочки» из непроеизводных элементов (б) $a + a + a + b + b + c + c + c + d + d$.

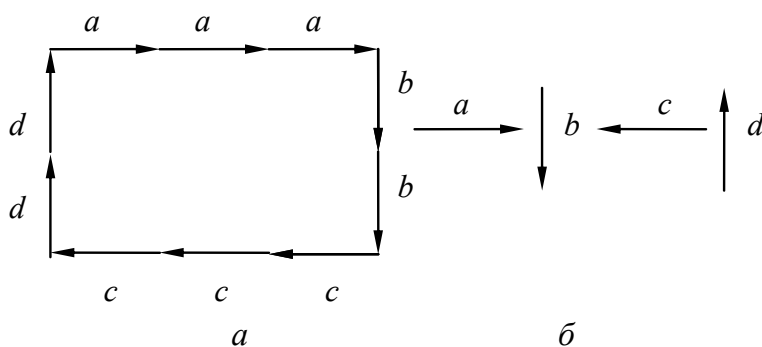


Рис. 2. Прямоугольник (а) и его непроеизводные элементы (б)

На рис. 3 приведен более сложный пример структурного описания цифры 8 (восемь).

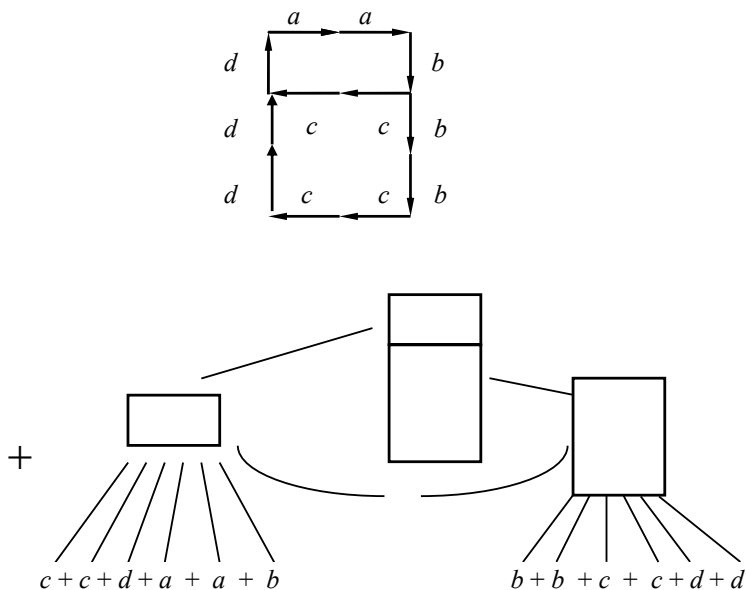
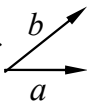


Рис. 3. Изображение цифры 8 и его структурное описание

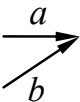
Грамматика языка описания объектов формируется на этапе обучения на основе обучающей выборки. Теоретической базой данного подхода является теория формальных языков и лежащих в их основе порождающих грамматик.

В качестве примера приведены фрагменты языка описания изображений PDL (Picture Description Language). Определены непроеизводные элементы \xrightarrow{a} \nearrow^b \searrow^c \uparrow^d , имеющие различающиеся головную и хвостовую точки, а также четыре бинарных оператора соединения элементов в цепочки:

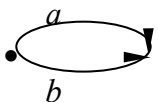
$a + b \Rightarrow \xrightarrow{a} \nearrow^b \Rightarrow$ головная точка a примыкает к хвостовой точке b ;

$a \times b \Rightarrow$  \Rightarrow хвостовая точка a примыкает к хвостовой

точке b ;

$a - b \Rightarrow$  \Rightarrow головная точка a примыкает к головной

точке b ;

$a * b \Rightarrow$  \Rightarrow головная точка a примыкает к го-

ловной точке b и (\wedge) хвостовая точка a примыкает к хвостовой точке b .

На рис. 4 приведено выражение на языке PDL, описывающее букву A .

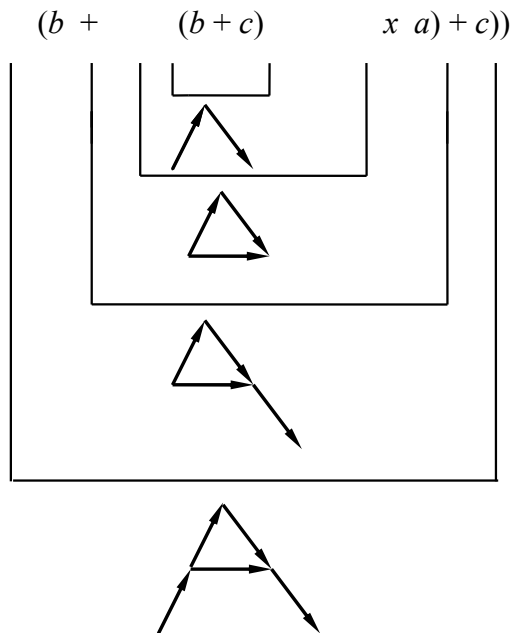


Рис. 4. Структурное описание буквы A на языке PDL

2. Основные этапы и требования к выполняемой работе

В ходе выполнения лабораторной работы необходимо разработать подсистему распознавания сложных зрительных образов (сложных изображений) на основе изложенного структурного подхода.

Разрабатываемая подсистема является подсистемой распознавания «с учителем», поэтому она должна состоять из двух частей: обучения и распознавания.

А. Обучение.

1. Пользователю должна быть предоставлена возможность классификации учебного материала, в ходе которой он последовательно вводит в машину изображения обучающей выборки и информацию о том, к какому образу эти изображения относятся. Как правило, на этом этапе используются утоньшенные изображения.

Рассмотрим методы утоньшения бинарного изображения. Любые полутоновые и цветные могут быть приведены к бинарным.

1.1. Утоньшение методом LRTB

Граничным элементом (ГЭ) называется пиксель, который не обладает четырехсвязностью по крайней мере с одним из ближайших соседей. Точка дуги (ТД) – пиксель, который обладает четырехсвязностью только с верхним и нижним элементом, или правым и левым. Дуговой концевой элемент (ДКЭ) обладает четырехсвязностью только с одним своим соседом.

Утоньшение выполняется в 4 этапа:

– ГЭ удаляются с левой стороны объекта, если они не являются ТД и их удаление не ведет к нарушению восьмисвязности;

– ГЭ удаляются с правой стороны объекта, если они не являются ТД и их удаление не ведет к нарушению восьмисвязности;

– ГЭ удаляются с верхней стороны объекта, если они не являются ТД и их удаление не ведет к нарушению восьмисвязности;

– ГЭ удаляются с нижней стороны объекта, если они не являются ТД и их удаление не ведет к нарушению восьмисвязности.

И так до тех пор, пока нельзя будет удалить не один элемент без нарушения восьмисвязности.

Пример работы алгоритма приведен на рис. 5.

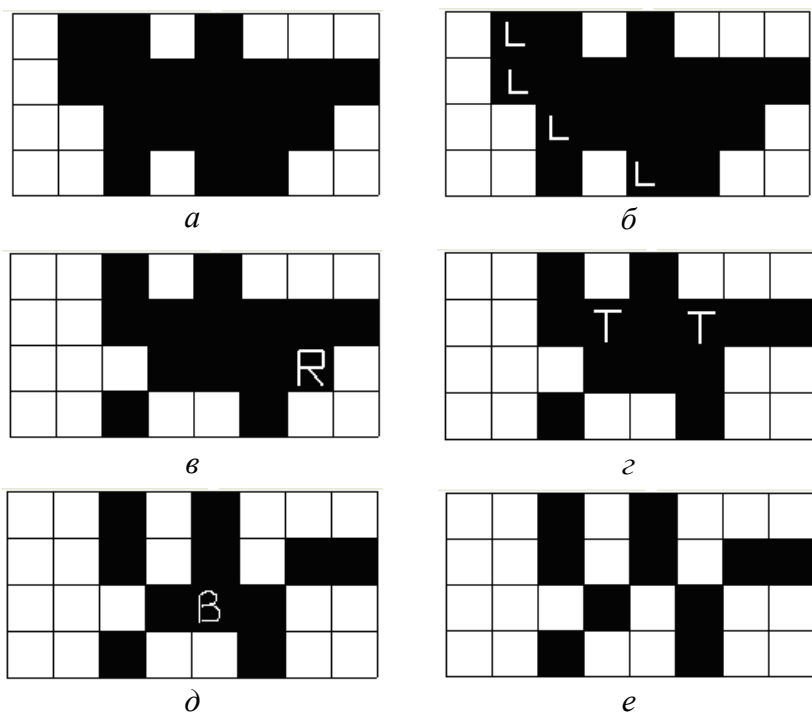


Рис. 5. Утоньшение методом LRTB

a – исходное изображение; *б* – утоньшение с левой стороны объекта; *в* – утоньшение с правой стороны объекта; *г* – утоньшение с верхней стороны объекта; *д* – утоньшение с нижней стороны объекта

1.2. Утоньшение методом Зонга–Суня

Вводим матрицу 3×3 , представленную на рис. 6.

$$\begin{bmatrix} P1 & P2 & P3 \\ P8 & P1 & P4 \\ P7 & P6 & P5 \end{bmatrix}$$

Рис. 6. Матрица нумерации пикселей для утоньшения

Накладываем матрицу на изображение, совмещая интересующий нас пиксел с $P1$.

Каждая итерация состоит из двух шагов.

Шаг 1. Пиксел $P1$ удаляется из изображения, если выполняются следующие условия:

- $2 \leq B(P1) \leq 6$;
- $A(P1) = 1$;
- $P2 \cdot P4 \cdot P6 = 0$;
- $P4 \cdot P6 \cdot P8 = 0$,

где $A(P1)$ – число конфигураций 01 в последовательности $P2, P3, P4, P5, P6, P7, P8, P9$ замыкает эту цепочку на $P2$, т. е. вокруг этого пиксела существует только один переход от 0 к 1;

$B(P1)$ – количество пикселей со значением, равным 1, определяется по формуле $B(P1) = \sum_{i=2}^9 P_i$.

Шаг 2. Выполняется аналогично, только

- $P2 \cdot P4 \cdot P8 = 0$;
- $P2 \cdot P6 \cdot P8 = 0$.

Таким образом, в процессе выполнения шага 1 происходит удаление точек на юго-восточной границе и северо-западных угловых точек, а в процессе выполнения шага 2 – удаление точек на северо-западной границе и юго-восточных угловых точек.

Однако этими условиями алгоритм не охватывает некоторые случаи.

Рассмотрим ситуацию, приведенную на рис. 7.

$$\begin{bmatrix} \blacksquare & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ 1 & \blacksquare & \blacksquare & \blacksquare & \blacksquare \\ 1 & 0 & 0 & \blacksquare & \blacksquare \\ 1 & p & 0 & \blacksquare & \blacksquare \\ 0 & 1 & 1 & 1 & \blacksquare \end{bmatrix}$$

Рис. 7. Случай некорректной работы алгоритма

Пиксел P , если он был единицей, не удаляется на основании тех условий, которые были описаны выше. Поэтому проводится еще одна итерация, которая устраняет подобные недочеты.

На этой итерации ищутся два единичных пиксела по вертикали или горизонтали, которые окружены нулями.

Итак, точка P удаляется, если выполняется одно из условий:

1) $!P9 \cdot P4 \cdot P6 = 1$;

2) $!P5 \cdot P8 \cdot P2 = 1$;

3) $!P3 \cdot P6 \cdot P8 = 1$;

4) $!P7 \cdot P2 \cdot P4 = 1$,

где $!P9$ – отрицание $P9$.

2. На каждом изображении необходимо выявить заранее определенное пользователем число характерных простейших элементов, а затем для каждого из найденных характерных элементов вычислить код его направлений. Характерные элементы и коды направлений запоминаются. Этот этап может выполняться либо пользователем в диалоговом режиме, либо автоматически. Для реализации этапа в автоматическом режиме можно использовать нижеперечисленные алгоритмы.

2.1. Определение характерных элементов

Наиболее информативными элементами с точки зрения описания изображений являются участки с сильным изменением формы – характерные элементы.

Для поиска характерных элементов необходимо осуществить сканирование изображения с помощью окошка, размер которого меньше, чем размер исследуемого изображения («рассматривая» изображение через такое окошко, программа «вырезает» из него отдельные элементы). Затем среди «вырезанных» элементов с помощью функции информативности отбираются характерные. Для этого предварительно на поле рецепторов окошка можно нарисовать стандартный фрагмент, представляющий собой пятно черноты, которое убывает от центра окошка к его периферии. Тогда в качестве функции информативности можно принять евклидово расстояние в пространстве рецепторов окошка между точкой, соответствующей фрагменту, нарисованному в окошке, и точкой, соответствующей фрагменту, «видимому» на изображении через окошко.

Будем считать, что когда центр окошка расположен так, что функция информативности имеет локальный экстремум, то в этот момент окошко «вырезает» на изображении характерный элемент. Выделенные таким способом характерные элементы являются характерными простыми фигурами, для которых можно подобрать простые названия. Нахождение экстремумов характерных элементов может быть выполнено с помощью одного алгоритма оптимизации, например, Хука–Дживса.

2.2. Определение кодов направлений

После выделения всех характерных элементов на каждом изображении обучающей выборки определяется их взаимное расположение на соответствующем изображении. Для этого каждому элементу присваивается код направлений, указывающий, в каких направлениях от данного элемента присутствуют, а в каких отсутствуют другие характерные элементы. Код направлений может иметь 8 разрядов, каждый из которых характеризует один из секторов «мишени» (рис. 8).

Центр «мишени» помещается в центр характерного элемента. В разряды кода направлений ставится 1, если в направлении сектора, соответствующего номеру разряда, есть другие характерные элементы, 0 – если их нет. Иногда в разряды этого кода ставится количество характерных элементов, существующих в данном направлении. Полученный таким образом код характеризует положение данного элемента на исследуемом изображении независимо от размеров изображения и его расположения. Сформированные коды направлений для всех характерных элементов с геометрической точки зрения можно интерпретировать как точки в 8-мерном пространстве.

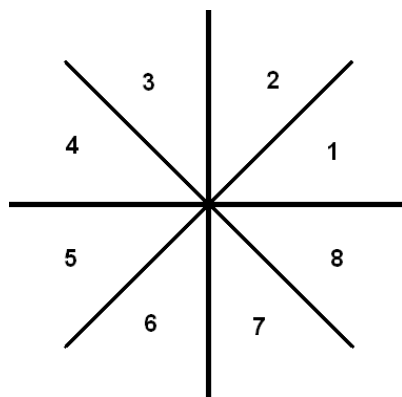


Рис. 8. «Мишень» для определения кода направлений

2.3. Классификация характерных элементов и кодов направлений

Множество отобранных характерных элементов и множество кодов направлений классифицируется на заданное число классов. Каждому классу элементов и направлений присваивается условный номер. Этот этап также может выполняться либо пользователем в диалоговом режиме, либо автоматически. Для реализации этапа в автоматическом режиме можно использовать один из алгоритмов объективной классификации.

Совокупность классов элементов и направлений может в дальнейшем использоваться при распознавании новых изображений. Однако для оптимизации процесса распознавания можно использовать и эталонные фрагменты каждого класса. Они содержат в себе нечто общее, характерное для всего данного класса.

Рассмотрим процесс формирования эталонного фрагмента для класса, состоящего из трех фигур, представленных на рис. 9.

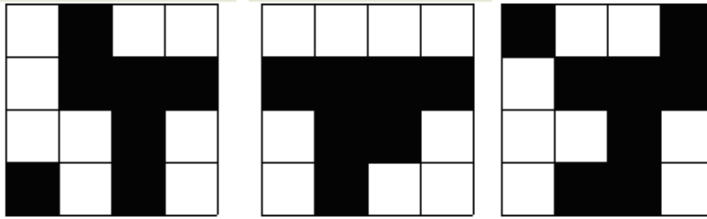


Рис. 9. Набор изображения для формирования эталонного фрагмента

Для каждой из фигур запишем ее код и сложим поразрядно коды всех трех фигур, затем определим среднее значение значимых разрядов, (т. е. отличных от нуля). Результат приведен на рис. 10.

$$\begin{array}{r}
 0100011100101010 \\
 + 0000111101100100 \\
 1001011100100110 \\
 \hline
 11011133301301220
 \end{array}$$

$$\sum = 22 \quad 16-4=12 \text{ значимых} \\
 \text{разрядов}$$

$$\begin{array}{r}
 22 \overline{) 12} \\
 \underline{11} \\
 18
 \end{array}$$

Рис. 10. Определение среднего значения значимых разрядов

После этого сравним каждый разряд суммарного кода со средним значением значимых разрядов и сформируем код эталонного фрагмента, поставив единицы в тех разрядах, где разряд суммарного кода больше среднего значения значимых разрядов, а ноль – в остальных разрядах. Полученный эталонный фрагмент и его код приведены на рис. 11.

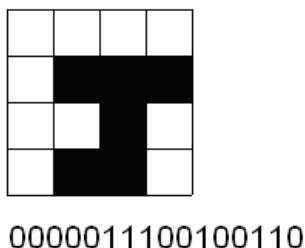


Рис. 11. Эталонный фрагмент для набора изображений с рис. 9

3. По результатам выполнения предыдущих этапов необходимо сформировать характеристические таблицы для всех изображений обучающей выборки. На предыдущих этапах было получено некоторое количество m различных эталонных фрагментов и некоторое количество n эталонов направлений. Тогда любое изображение можно представить с помощью его описания на основании характеристической таблицы, состоящей из m строк и n столбцов. В клетке таблицы ставится 1, если в изображении имеется хотя бы один соответствующий данной строке фрагмент на соответствующем данному столбцу месте. Ноль в клетке таблицы означает, что соответствующего фрагмента на соответствующем месте изображение не содержит.

В. Распознавание

1. Для распознавания нового изображения необходимо обеспечить возможность его загрузки и поиска на нем заданного числа характерных элементов, а затем – определения кодов направлений для каждого элемента. Поиск характерных элементов необходимо выполнить аналогично поиску характерных элементов на обучающей выборке, размер окошка не изменять.

2. Выполнить классификацию найденных характерных элементов и кодов направлений. Они должны быть отнесены к одному из классов, полученных на втором этапе обучения. Для этого можно использовать различные меры близости и алгоритмы. Мерой близости может служить средний потенциал, создаваемый в точке x множеством точек класса A :

$$\Phi(x, A) = \frac{1}{N_A} \sum_{i=1}^{N_A} \Phi(x, x_i),$$

полагая при этом, что потенциал характеризует среднее расстояние от точки x , соответствующей распознаваемому характерному элементу или коду, до всех точек x_i одного из сформированных классов, например, класса A . Если же для классификации используются эталонные фрагменты и эталоны направлений, то в качестве меры близости можно использовать потенциал между двумя точками x_1 и x_2 :

$$\Phi(x_1, x_2) = \frac{1}{1 + \sqrt{(x_1^1 - x_2^1)^2 + (x_1^2 - x_2^2)^2 + \dots + (x_1^n - x_2^n)^2}},$$

полагая при этом, что потенциал характеризует расстояние от точки x_1 , соответствующей распознаваемому характерному элементу или коду, до точки x_2 , соответствующей одному из эталонных фрагментов или направлений.

3. Составить характеристическую таблицу распознаваемого изображения.

4. Выполнить сравнение характеристической таблицы распознаваемого изображения с характеристическими таблицами изображений обучающей выборки. Отнесение распознаваемого изображения к определенному классу (образу) обучающей выборки можно производить двумя способами.

Первый способ состоит в определении сходства между таблицами. В этом случае характеристические таблицы изображений трактуются как точки в некотором пространстве с размерностью, равной числу клеток в каждой из таблиц ($m \cdot n$). В этом пространстве каждому классу изображений обучающей выборки будет соответствовать группа точек, изображающих объекты этого класса. Новый объект (расознаваемое изображение)

относится к тому классу (образу), к которому он ближе, т. е. которому отвечает наибольшее значение меры близости.

Второй способ более экономный и заключается в логическом сравнении таблиц. В этом случае составляется табличное описание класса (образа) изображений, состоящее из двух таблиц, совпадающих по форме с характеристическими таблицами отдельных изображений. Пусть эти таблицы условно называются таблицей обязательных элементов (кодов) *O* и таблицей исключаемых элементов (кодов) *I*. В ячейке таблицы *O* ставится 1, если характеристические таблицы всех изображений данного класса (образа) в соответствующей ячейке имеют единицы. В ячейке таблицы *I* ставится ноль, если ни одна из характеристических таблиц в соответствующей ячейке не содержит единицы. Таким образом, таблица *O* указывает, что должно присутствовать, а таблица *I* – что должно отсутствовать в изображениях данного класса (образа).

Распознаваемое изображение должно быть отнесено к тому классу (образу), который обеспечивает наилучшее по сравнению с другими классами (образами) выполнение следующего критерия принадлежности: характеристическая таблица распознаваемого изображения должна содержать единицы везде, где есть единицы в таблице *O* данного класса (образа), и нули везде, где есть нули в таблице *O* этого класса (образа). Это эквивалентно требованию: распознаваемое изображение должно содержать «все обязательное» и не содержать «ничего исключаемого» для данного класса (образа).

Литература

1. Абламейко, С.В. Обработка изображений: технология, методы, применение / С.В. Абламейко, Д.М. Лагуновский. – Минск, 1999. – 368 с.
2. Аркадьев, А.Г. Обучение машины классификации объектов / А.Г. Аркадьев, Э.С. Браверман. – М.: Наука, 1981. – 464 с.
3. Дуда, Р. Распознавание образов и анализ сцен / Р. Дуда, П. Харт. – М.: Мир, 1976. – 213 с.
4. Павлидис, Т. Алгоритмы машинной графики и обработки изображений / Т. Павлидис. – М.: Мир, 1986. – 400 с.
5. Фу, К. Структурные методы в распознавании образов / К. Фу. – М.: Мир, 1977. – 320 с.

Оглавление

1. Применение структурного подхода при решении задач распознавания образов.	3
2. Основные этапы и требования к выполняемой работе.	8
Литература.	18

Учебное издание

СТРУКТУРНЫЙ ПОДХОД
ПРИ РАСПОЗНАВАНИИ ИЗОБРАЖЕНИЙ

Методические указания
к лабораторной работе для студентов специальности
1-40 01 02 «Информационные системы и технологии»

Составитель
КОВАЛЕВА Ирина Львовна

Редактор Т.В. Кипель
Компьютерная верстка Н.А. Школьниковой

Подписано в печать 02.11.2011.

Формат 60×84¹/₁₆. Бумага офсетная.

Отпечатано на ризографе. Гарнитура Таймс.

Усл. печ. л. 1,16. Уч.-изд. л. 0,91. Тираж 100. Заказ 996.

Издатель и полиграфическое исполнение:

Белорусский национальный технический университет.

ЛИ № 02330/0494349 от 16.03.2009.

Проспект Независимости, 65. 220013, Минск.