УДК 004.93.51

Т.М. Косовская

РАСПОЗНАВАНИЕ ИЗОБРАЖЕНИЯ НА ЭКРАНЕ ДИСПЛЕЯ И ОЦЕНКИ ЧИСЛА ШАГОВ АЛГОРИТМОВ В РАМКАХ ЛОГИКО-ПРЕДМЕТНОЙ РАСПОЗНАЮЩЕЙ СИСТЕМЫ

Косовская Т.М. **Распознавание изображений на экране дисплея и оценки числа шагов алгоритмов в рамках логико-предметной распознающей системы.**

Аннотация. Статья посвящена получению оценок числа шагов логико-предметных алгоритмов распознавания сложных изображений на экране дисплея. Доказана полиномиальность задачи выделения и распознавания эталонного изображения на сложной сцене. Для задачи выделения и распознавания объекта из класса, описание которого содержит только характерные признаки этого класса, доказана её принадлежность классу *NP*. Для уменьшения числа шагов работы алгоритма предложено понятие „размытого изображения“. Рассмотрена задача инвариантного (относительно изменения масштаба) распознавания изображения.

Ключевые слова: формулы исчисления предикатов, распознавание изображений, сложность алгоритмов.

Kosovskaya T.M. **Display screen image recognition and bounds of number of steps of an algorithm in the frameworks of logic-objective recognition system.**

Abstract. The paper is devoted to the proof of upper bounds of steps of logic-objectiv algorithms for recognition of a complicated image situated on a display screen. It is proved that the problem of separation and recognition of an etalon object from a complicated scene has a polynomial algorithm. The problem of separation and recognition of an object from a class the description of which contains only distinctive attributes of this class belongs to *NP*. To decrease the algorithm number of steps a notion of „fuzzy image“ is introduced. The problem of invariant (under rescaling) image recognition is regarded.

Keywords: predicate calculus formulas, image recognition, complexity of algorithm.

1. Введение. Для решения различных задач распознавания образов в [1] предложен логико-предметный подход. При формулировке этого подхода предполагается, что признаки характеризуют не весь распознаваемый объект целиком, а лишь его отдельные части. Распознаваемый объект задаётся как множество составляющих его элементов $\omega = \{\omega_1, \ldots, \omega_t\}$. Множество всех распознаваемых объектов Ω разбито на K классов $\Omega = \bigcup_{k=1}^{K} \Omega_k$. На элементах распознаваемых объектов ω задана совокупность предикатов

$\{p_1, \ldots, p_n\}$, описывающих свойства элементов и отношения между ними.

Описанием $S(\omega)$ распознаваемого объекта ω называется множество истинных формул вида $p_i(\overline{\tau_i})$ или $\neg p_i(\overline{\tau_i})$ при $\tau_i \subset \omega$. Здесь и далее посредством $\overline{\tau}$ обозначается упорядоченный набор элементов из τ.

Описанием $A_k(\overline{x})$ класса Ω_k называется формула, представленная в виде дизъюнкции элементарных конъюнкций атомарных формул с предикатами $p_1, \ldots, p_n$, такая, что $\omega \in \Omega_k \Leftrightarrow A_k(\overline{\omega})$.

Для того чтобы записать, что значения для переменных списка $\overline{x}$, удовлетворяющие формуле $A(\overline{x})$, различны, ниже будет использоваться обозначение $\exists \overline{x}_{\neq} A(\overline{x})$.

Рассматриваются следующие задачи распознавания.

Задача идентификации. *Проверить, принадлежит ли объект ω или его часть классу Ω_k.*

Задача классификации. *Найти все такие номера классов k, что $\omega \in \Omega_k$.*

Задача анализа сложного объекта. *Найти и классифицировать все части τ объекта ω, для которых $\tau \in \Omega$.*

Эти задачи в [1] сведены к доказательству соответственно логических следований

$$S(\omega) \Rightarrow \exists \overline{y}_{\neq} A_k(\overline{y}), \tag{1}$$

$$S(\omega) \Rightarrow \bigvee_{k=1}^{K} A_k(\overline{\omega}), \tag{2}$$

$$S(\omega) \Rightarrow \bigvee_{k=1}^{K} \exists \overline{y}_{\neq}\ A_k(\overline{y}), \tag{3}$$

с указанием всех частей объекта ω, удовлетворяющих описанию класса (для (1) и (3)) и всех номеров классов (для (2) и (3)).

В работе [2] доказано, что все эти задачи NP-трудны. Оценки числа шагов различных алгоритмов, решающих эти задачи, в показателе степени имеют различные параметры описаний классов. Для алгоритма, основанного на полном переборе, в оценке числа шагов решения любой из этих задач экспоненциальный сомножитель имеет вид $O(t^m)$, где t — количество элементов в распознаваемом объекте, m — максимальное количество переменных в элементарных конъюнкциях, входящих в описания классов.

2. Изображение на экране дисплея. При распознавании изображения на экране дисплея исследуемый объект представлен не как множество своих элементов, а как упорядоченный по строкам и по столбцам список пикселей. Соответственно, и оценки числа шагов решения задач распознавания должны измениться.

Экраном дисплея T назовем таблицу размером $m \times n$ пар чисел, описываемую одним предикатом p, характеризующим положение пикселя на экране, яркость этого пикселя и, быть может, цвет. То есть атомарная формула $p(i, j, x, y)$ истинна тогда и только тогда, когда пиксель с координатами (i, j) имеет яркость x и цвет y. При этом $i = 1, \ldots, n$, $j = 1, \ldots, m$, $x = 1, \ldots, X$, $y = 1, \ldots, Y$. Атомарную формулу вида $p(i, j, a, b)$ будем называть описанием пикселя с координатами (i, j), имеющего яркость a и цвет b.

Описание $S(T)$ экрана T представляет из себя упорядоченный (по строкам и столбцам) набор описаний составляющих его пикселей.

Изображением (или объектом) ω_D назовем такую часть экрана, что $(i, j) \in D$. В этом случае задача идентификации (1) состоит в выделении такой области D на экране, что ω_D является изображением из заданного класса. Задача классификации (2) — это задача определения, каким классам принадлежит выделенное изображение. Задача анализа сложного объекта (3) — это задача анализа сцены, то есть выделение всех таких областей, изображение на которых поддаётся классификации.

При создании описания отдельного изображения в качестве атомарных формул могу выступать как атомарные формулы с предикатом p, так и неравенства между аргументами различных четвёрок. Кроме этого, в каждой четвёрке могут находиться не только переменные, но и термы с использованием арифметических операций. Так, например, для того чтобы описать два соседних по горизонтали пикселя, достаточно задать формулу $p(i, j, x_1, y_1) \,\&\, p(i, j+1, x_2, y_2)$, а для задания условия, что два пикселя расположенны на одной горизонтали, достаточно задать формулу $p(i, j_1, x_1, y_1) \,\&\, p(i, j_2, x_2, y_2) \;\&\; (j_1 < j_2)$. Для того, чтобы описать, что пиксель с номером (i_1, j_1) вдвое ярче пикселя с номером (i_2, j_2), достаточно задать формулу $p(i_1, j_1, x, y_1) \,\&\, p(i_2, j_2, 2 \cdot x, y_2)$, а для описания того, что второй пиксель более яркий, чем первый, достаточно задать формулу $p(i_1, j_1, x_1, y_1) \,\&\, p(i_2, j, x_2, y_2) \,\&\, (x_1 < x_2)$.

В данной работе под шагом работы алгоритма будем понимать выполнение одной из следующих операций: присваивание переменным, входящим в описание класса, значений, взятых из описания изображения; сравнение двух постоянных атомарных формул вида $p(i,j,a,b)$ на их графическое совпадение; проверка неравенства, находящегося в описании класса, после замены входящих в него переменных на константы.

Если в каждой элементарной конъюнкции описания класса все переменные упорядочены по первым двум аргументам четверок, и каждая пара таких аргументов (кроме первой) находится в функциональной или предикатной зависимости от некоторых предыдущих, то задача классификации (2) принимает в этом случае вид $S(\omega_D) \Rightarrow A_k(\omega_D)$ и имеет линейную от длины записи описания класса верхнюю оценку числа шагов её решения.

Задача анализа (3) представляет собой многократное (при $k = 1, ..., K$) решение задачи идентификации, и число шагов её решения легко выражается через число шагов решения задачи идентификации.

При решении задачи идентификации (1) изображения на сложной сцене можно рассмотреть две различные постановки задачи идентификации изображени (ЗИИ).

ЗИИ1. В процессе присваивания переменным описания класса $A_k(\overline{x})$ значений из T переменным с соседними номерами присваиваются значения соседних элементов из T, т.е. присваивания $(i_1, j_1) := (\alpha_1, \beta_1)$ и $(i_2, j_2) := (\alpha_2, \beta_2)$ допустимы тогда и только тогда, когда $(i_2 = i_1 + 1 \Leftrightarrow \alpha_2 = \alpha_1 + 1)$ & $(j_2 = j_1 + 1 \Leftrightarrow \beta_2 = \beta_1 + 1)$.

ЗИИ2. В процессе присваивания переменным описания класса $A_k(\overline{x})$ значений из T переменным с большими номерами присваиваются значения элементов из T с большими номерами. Т.е. присваивания $(i_1, j_1) := (\alpha_1, \beta_1)$ и $(i_2, j_2) := (\alpha_2, \beta_2)$ допустимы тогда и только тогда, когда $(\alpha_1 < \alpha_2 \Leftrightarrow i_1 < i_2)$ & $(\beta_1 < \beta_2 \Leftrightarrow j_1 < j_2)$.

Пусть $A(\overline{x})$ — элементарная конъюнкция, входящая в описание заданного класса, $\overline{x}$ — список четверок вида (i, j, x, y). Длиной формулы $A(\overline{x})$ будем называть число $|A|$, равное количеству атомарных формул в формуле $A(\overline{x})$.

3. Поиск эталонного изображения на экране дисплея. Задача идентификации изображения на сцене в первой постановке соответствует нахождению эталонного изображения, заданного формулой $A(\overline{x})$. Так как в этом случае процесс идентификации сводится к нахождению значений для первой переменной списка

$\overline{x}$, то справедлива следующая теорема.

Теорема 1. *Если в элементарной конъюнкции $A(\overline{x})$ переменные упорядочены по первым двум аргументам четверок, каждая пара таких аргументов (кроме первой) находится в функциональной или предикатной зависимости от некоторых предыдущих, то для задачи идентификации в первой постановке существует алгоритм её решения, число шагов которого не превосходит*

$$(n \cdot m - i_{max} \cdot j_{max}) \cdot |A|,$$

где i_{max} и j_{max} — наибольший номер первой и, соответственно, второй координат четвёрок, являющихся переменными формулы $A_k(\overline{x})$.

Доказательство. Если значение для первой пары аргументов четвёрок задано, то при сравнении первой атомарной формулы вида $p(i, j, x, y)$ из $A(\overline{x})$ с постоянной атомарной формулой из описания изображения $S(T)$ будут определены значения третьего и четвёртого аргументов рассматриваемой четвёрки. Кроме этого, будут определены значения первой пары аргументов второй атомарной формулы этого же вида.

Аналогичным образом, после того как значения для первой пары аргументов l-ой атомарной формулы вида $p(i, j, x, y)$ из $A(\overline{x})$ уже определены, в силу функциональной или предикатной зависимости первой пары аргументов $(l + 1)$-ой атомарной формулы вида $p(i, j, x, y)$ из $A(\overline{x})$ будут определены значения первой пары её аргументов.

Всего возможностей для первичного значения первой пары аргументов четвёрок не превосходит $(n \cdot m - i_{max} \cdot j_{max})$. ■

Следствие. *ЗИИ1 принадлежит классу P .*

Пример 1.

Пусть имеется монохромный экран дисплея размером 10×10 с двумя градациями яркости (0 – белый, 1 – чёрный). Предикат, задающий описание пикселя, имеет три аргумента (четвёртый аргумент, задающий цвет, отсутствует). В качестве эталонного объекта на экране дисплея задано изображение буквы **Т** (Рис.1).

По этому изображению составлено описание всех букв **Т**, отличающихся от эталонного изображения только положением

$$A_T(i, j) = p(i, j, 1) \ \& \ p(i, j + 1, 1) \ \& \ p(i, j + 2, 1) \ \&$$

$$p(i + 1, j + 1, 1) \ \& \ p(i + 2, j + 1, 1) \ \& \ p(i + 3, j + 1, 1).$$

Для идентификации буквы **Т** предъявлено изображение, представленное на Рис. 2.

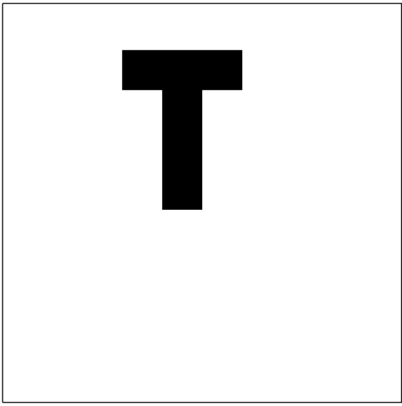

Рис. 1.

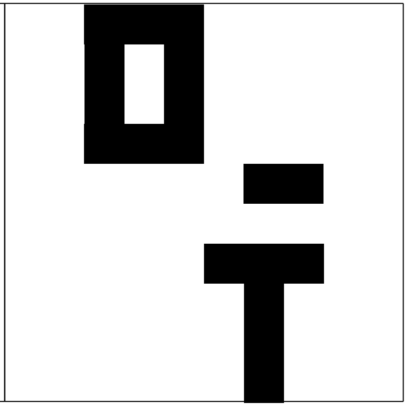

Рис. 2.

Экраном дисплея будет таблица

$$\begin{matrix}
0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0
\end{matrix}$$

Описанием этого изображения будет последовательность атомарных формул $S(T)$ вида

$$(p(1,1,0),\ p(1,2,0),\ p(1,3,1),\ p(1,4,1),\ p(1,5,1),\ p(1,6,0),...$$

$$...p(10,7,1),\ p(10,8,0),\ p(10,9,0),\ p(10,10,0)).$$

При проверке логического следования $S(T) \Rightarrow \exists ij\ A_T(i,j)$ последовательно присваиваем

$i := 1$, $j := 1$ — в $A_T(1,1)$ $p(1,1,0)$ из $S(T)$ отсутствует. (Один шаг.)

$i := 1$, $j := 2$ — в $A_T(1,2)$ $p(1,2,0)$ из $S(T)$ отсутствует. (Один шаг.)

$i := 1$, $j := 3$ — в $A_T(1,3)$ присутствуют $p(1,3,1)$, $p(1,4,1)$, $p(1,5,1)$, но $p(2,4,1)$ из $S(T)$ отсутствует. (Четыре шага.)

.....

$i := 7$, $j := 6$ — в $A_T(7,6)$ присутствуют только атомарные формулы из $S(T)$. (Шесть шагов.)

$i := 7$, $j := 7$ — в $A_T(7,7)$ присутствуют $p(7,7,1)$, $p(7,8,1)$, но $p(7,9,1)$ из $S(T)$ отсутствует. (Три шага.)

$i := 7$, $j := 8$ — в $A_T(7,8)$ присутствует $p(7,8,1)$, но $p(7,9,1)$ из $S(T)$ отсутствует. (Два шага.)

Таким образом, область $D = \{(7,6),\ (7,7),\ (7,8),\ (8,7),\ (9,7),\ (10,7)\}$ представляет изображение буквы **Т**.

4. Поиск изображения на экране дисплея, описанного характерными признаками. Задача идентификации изображения на сцене во второй постановке соответствует нахождению на сцене изображения, в котором присутствуют характерные черты изображений из заданного класса, отражённые в описании класса $A_k(\overline{x})$.

Теорема 2. *Если в элементарной конъюнкции $A(\overline{x})$ переменные упорядочены по первым двум аргументам четверок, каждая пара таких аргументов (кроме первой) находится в функциональной или предикатной зависимости от некоторых предыдущих, то для задачи идентификации во второй постановке существует алгоритм её решения, число шагов которого не превосходит*

$$\sum_{i=1}^{n-i_{max}} \sum_{j=1}^{m-j_{max}} C_{n_i}^{i_{max}} \cdot C_{m_j}^{j_{max}} \cdot |A|,$$

где i_{max} и j_{max} — наибольший номер первой и, соответственно, второй координат четвёрок, являющихся переменными формулы $A(\overline{x})$, $n_i = n - i_{max} - i + 1$, $m_j = m - j_{max} - j + 1$.

Доказательство. Для задания значения первого и второго аргумента четвёрок существует не более, чем $n - i_{max}$ и, соответственно, $m - j_{max}$ возможностей.

Если значения (i, j) для этих аргументов задано, то определение значений остальных аргументов может быть выполнено не более чем за $C_{n_i}^{i_{max}} \cdot C_{m_j}^{j_{max}}$ шагов, что равно количеству выборок упорядоченных i_{max} и j_{max} наборов из упорядоченных наборов из n_i и m_j элементов соответственно. Проверка того, что в формуле $A(\overline{x})$ после присвоения значений всем аргументам первых пар четвёрок остальные аргументы могут быть присвоены в соответствии с описанием $S(T)$, может быть осуществлена не более чем за $|A|$ шагов. ■

Теорема доказана.

Следствие. *ЗИИ2 принадлежит классу NP.*

Замечание. Несмотря на то, что задача идентификации во второй постановке принадлежит классу *NP*, число шагов её решения меньше, чем для аналогичной задачи в случае, если исследуемый объект представлен как неупорядоченное множество своих элементов. В случае неупорядоченного множества в оценках числа шагов вместо числа сочетаний $C_{n_i}^{i_{max}}$ и $C_{m_i}^{j_{max}}$ стоит соответственно число размещений $A_{n_i}^{i_{max}}$ и $A_{m_i}^{j_{max}}$ [2].

Пример 2.

Пусть имеется монохромный экран дисплея размером 10×10 с двумя градациями яркости (0 – белый, 1 – чёрный). Предикат, задающий описание пикселя, имеет три аргумента (четвёртый аргумент, задающий цвет, отсутствует). На экране могут появляться изображения букв О и Т.

Для различения этих букв в описаниях классов записаны их характерные особенности, а именно: для буквы О — наличие светлых пикселей между тёмными

$$A_O(i, j, k) = p(i, j, 1) \ \& \ p(i, j + k, 0) \ \& \ p(i, j + k + 1, 1),$$

для буквы Т — место соединения горизонтальной и вертикальной линий

$$A_T(i, j, k) = p(i, j, 1) \ \& \ p(i, j + 1, 1) \ \& \ p(i, j + k, 1) \ \&$$
$$p(i + 1, j, 0) \ \& \ p(i + 1, j + 1, 1) \ \& \ p(i + 1, j + k, 0).$$

Для идентификации буквы О на рис. 2 при проверке логического следования $S(T) \Rightarrow \exists ijk \ A_O(i, j, k)$ последовательно присваиваем

$i := 1, j := 1, k := 1$ — в $A_O(1, 1, 1)$ $p(1, 1, 0)$ из $S(T)$ отсутствует. (Один шаг.)

$i := 1, j := 2, k := 1$ — в $A_O(1,2,1)$ $p(1,2,0)$ из $S(T)$ отсутствует. (Один шаг.)

$i := 1, j := 3, k := 1$ — в $A_O(1,3,1)$ присутствует $p(1,3,1)$ из $S(T)$, но $p(1,4,0)$ отсутствует. (Два шага.)

$i := 1, j := 3, k := 2$ — в $A_O(1,3,1)$ присутствует $p(1,3,1)$ из $S(T)$, но $p(1,5,0)$ отсутствует. (Два шага.)

$i := 1, j := 3, k := 3$ — в $A_O(1,3,1)$ присутствуют $p(1,3,1)$ и $p(1,6,0)$ из $S(T)$, но $p(1,7,1)$ отсутствует. (Три шага.)

.....

$i := 1, j := 3, k := 6$ — в $A_O(1,3,1)$ присутствуют $p(1,3,1)$ и $p(1,9,0)$ из $S(T)$, но $p(1,10,1)$ отсутствует. (Три шага.)

.....

$i := 2, j := 3, k := 1$ — в $A_O(2,3,1)$ присутствуют только атомарные формулы из $S(T)$. (Три шага.) Для выделения буквы O требуется дополнительно найти связную область экрана с тёмными пикселями, содержащую пиксели (2,3) и (2,5).

.....

$i := 3, j := 3, k := 1$ — в $A_O(3,3,1)$ присутствуют только атомарные формулы из $S(T)$. (Три шага.) Для выделения буквы O требуется дополнительно найти связную область экрана с тёмными пикселями, содержащую пиксели (3,3) и (3,5). Эта область совпадёт с полученной при $i = 2, j = 3, k = 1$.

.....

И т.д., причём каждый раз, когда вслед за тёмным пикселем следует светлый ((i,j) совпадающий с $(5,8)$, или с $(7,8)$, или с $(8,7)$, или с $(9,7)$, или с $(10,7)$, параметр k будет меняться до максимально возможного значения (от 1 до соответственно 1, 1, 2, 2, 2).

5. Размытое изображение. В [3] введено понятие многоуровневого описания классов и доказаны условия, при которых выделение „часто“ встречающихся подформул „небольшой сложности“ позволяют уменьшить число шагов работы алгоритмов решения задач (1), (2) и (3). Для изображений такими подформулами описаний классов объектов могут являться характерные обобщённые характеристики изображения.

Если моделировать восприятие изображения человеком, то следует, в первую очередь, отметить, что мы воспринимаем изображение как некое „размытое изображение“ с выделенными частями, а лишь затем начинаем различать детали изображения. Для формализации такого рода восприятия введем определения.

Определение. Множеством наибольшей яркости $Light(D)$ (соответственно, яркости не меньшей l $light_l(D)$) изображения ω_D называется множество координат пикселей изображения наибольшей яркости

$$Light(D) = \{(i,j) : (i,j) \in D \ \& \forall x \ y \ i_1 \ j_1 \ x_1 \ y_1$$

$$((i_1, j_1) \in D \ \& \ p(i,j,x,y) \ \& \ p(i_1, j_1, x_1, y_1) \to x \geq x_1)\}$$

(соответственно, множество координат пикселей изображения, яркость которых не менее, чем l

$$light_l(D) = \{(i,j) : (i,j) \in D \ \& \forall x \ y$$

$$(p(i,j,x,y) \to x \geq l)\} \ ).$$

Следующее утверждение очевидно.

Утверждение.

$$Light(D) = \bigcap_{l: light_l(D) \neq \emptyset} light_l(D) \ .$$

Определение. Границей $\Gamma(D)$ области D называется множество, состоящее из координат пикселей из D, с одной стороны от которых (по горизонтали или по вертикали) пикселя изображения нет, а с другой стороны пиксель принадлежит D.

$$\Gamma(D) = \{(i,j) :$$

$$(i-1,j) \notin D \ \& \ (i,j) \in D \vee (i,j-1) \notin D \ \& \ (i,j) \in D \vee$$
$$(i+1,j) \notin D \ \& \ (i,j) \in D \vee (i,j+1) \notin D \ \& \ (i,j) \in D\}.$$

Определение. Границей $\Gamma_h(D)$ области D толщины h называется множество, состоящее из координат пикселей из D, с одной стороны от которых (по горизонтали или по вертикали) пикселей изображения нет, а с другой стороны $h-1$ пиксель принадлежит D, а также сами эти $h-1$ пиксель.

$$\Gamma_h(D) = \{(i,j) : \exists k \ (0 \leq k \leq h-1 \ \&$$

$$(((i-1,j) \notin D \ \& \ (i,j) \in D \ \& \ \ldots \ \& \ (i+h-1,j) \in D) \ )\vee$$

$$((i,j-1)\notin D \;\&\; (i,j)\in D \;\&\; \ldots \;\&\; (i,j+h-1)\in D)\;))\vee$$
$$(((i+1,j)\notin D \;\&\; (i,j)\in D \;\&\; \ldots \;\&\; (i-h+1,j)\in D)\;)\vee$$
$$((i,j+1)\notin D \;\&\; (i,j)\in D \;\&\; \ldots \;\&\; (i,j-h+1)\in D)\;))\}.$$

Будем называть „размытым изображением“, например, такие объекты:

— граница области D небольшой толщины h вместе с множеством наибольшей яркости

$$\Gamma_h(D)\bigcup Light(D);$$

— граница области D небольшой толщины h вместе с множеством большой (большей, чем заданная яркость l) яркости

$$\Gamma_h(D)\bigcup light_l(D);$$

— граница области D небольшой толщины h вместе с множеством малой (не большей, чем заданная яркость l) яркости

$$\Gamma_h(D)\bigcup(D\setminus light_l(D));$$

— граница области D небольшой толщины h вместе с множеством большой (большей, чем заданная яркость l_1) яркости и множеством малой (не большей, чем заданная яркость l_2) яркости

$$\Gamma_h(D)\bigcup light_{l_1}(D)\bigcup(D\setminus light_{l_2}(D)).$$

Размером „размытого изображения“ будем называть мощность соответствующего множества.

Определение. Описанием размытых изображений k-го класса называется такая формула $A_k^f(\overline{x}_1)$, что, во-первых, каждая элементарная конъюнкция формулы $A_k^f(\overline{x}_1)$ является подформулой описания объектов класса $A_k(\overline{x})$ и, во-вторых, всякое „размытое изображение“ объекта из этого класса удовлетворяет формуле $A_k^f(\overline{x}_1)$.

Если для каждой элементарной конъюнкции $A(\overline{x})$, входящей в описание класса $A_k(\overline{x})$, соответствующая ей подформула $A^f(\overline{x_1})$ много короче её (т.е. $|A^f|\ll|A|$), то время идентификации распознаваемого изображения на экране дисплея существенно уменьшится, если сначала идентифицировать „размытое изображение“, а лишь затем полностью идентифицировать объект.

Теорема 3. *Если в каждой из элементарных конъюнкций $A(\overline{x})$ и $A^f(\overline{x_1})$ переменные упорядочены по первым двум аргументам четверок, каждая пара таких аргументов (кроме первой) находится в функциональной или предикатной зависимости от некоторых предыдущих, то для задачи идентификации существует алгоритм её решения, число шагов которого в первой постановке не превосходит*

$$(n \cdot m - i_{max} \cdot j_{max}) \cdot |A^f| + (|A| - |A^f|),$$

во второй постановке не превосходит

$$\sum_{i=1}^{n-i_{max}} \sum_{j=1}^{m-j_{max}} C_{n_i}^{i_{max}} \cdot C_{m_j}^{j_{max}} \cdot |A^f| + (|A| - |A^f|),$$

где i_{max} *и* j_{max} *— наибольший номер первой и соответственно второй координат четвёрок, являющихся переменными формулы* $A_k(\overline{x})$, $n_i = n - i_{max} - i + 1$, $m_j = m - j_{max} - j + 1$.

Доказательство основано на оценках, полученных в теоремах 1 и 2.

Описания размытых изображений можно получать несколькими способами. Например:

— при построении описания класса изображений по эталонным одновременно строить описания соответствующих „размытых изображений“ эталонных изображений;

— если описание класса изображений уже имеется, то выделить из него подформулы, соответствующие границе изображения и множествам наибольшей (наименьшей) яркости.

6. Изменение масштаба изображения

Естественным требованием к решению задачи распознавания изображений является инвариантность алгоритма решения к такому преобразованию, как изменение масштаба.

При восприятии человеком отдельного изображения обычно изменение масштаба рассматривается относительно некоего „центра“ этого изображения. Для формализации этого понятия введём определения.

Определение. Горизонтальным и вертикальным размерами изображения ω_D называются соответственно числа $s_g(D)$ и $s_h(D)$, вычисляемые по формулам

$$s_g(D) = \max_{j:\ \exists i\ (i,j)\in D} j - \min_{j:\ \exists i\ (i,j)\in D} j \ ,$$

$$s_h(D) = \max_{i:\ \exists j\ (i,j)\in D} i - \min_{i:\ \exists j\ (i,j)\in D} i \ .$$

То есть горизонтальный и вертикальный размеры изображения задают размер наименьшего прямоугольника, в котором расположено изображение.

Определение. Горизонтальным и вертикальным диаметрами изображения называются соответственно числа $d_g(D)$ и $d_h(D)$, вычисляемые по формулам

$$d_g(D) = \max_{j_1,j_2:\ \exists i\ (i,j_1)\in D\ \&\ (i,j_2)\in D} (j_2 - j_1) \ ,$$

$$d_h(D) = \max_{i_1,i_2:\ \exists j\ (i_1,j)\in D\ \&\ (i_2,j)\in D} (i_2 - i_1) \ .$$

То есть горизонтальный и вертикальный диаметры задают минимальные размеры наибольших подстроки и подстолбца экрана, которые содержат точки изображения.

Определение. Координатами (i_0, j_0) геометрического центра изображения называются числа, вычисляемые по формулам

$$i_0 = \min_{i:\ \exists j\ (i,j)\in D} i + [\frac{s_h}{2}] \ ,$$

$$j_0 = \min_{j:\ \exists i\ (i,j)\in D} i + [\frac{s_g}{2}] \ .$$

То есть геометрическим центром изображения является центр наименьшего прямоугольника, в котором расположено изображение.

Определение. Координатами (i_v, j_v) визуального центра изображения D называются координаты точки из $Light(D)$, ближайшей к его геометрическому центру.

Пусть E_k — растяжение в k раз относительно пикселя с координатами (i_0, j_0) (по горизонтали и по вертикали) изображения ω_D. Тогда его описанием $S(E_k(\omega_D))$ будет множество атомарных формул вида

$$S(E_k(\omega_D)) = \{p(i_1, j_1, x_{ij}, y_{ij}) : (i,j) \in D \ \&$$

$$[(i_0 - i)(k-1)] < i_0 - i_1 \leq [(i_0 - i)k] \ \&$$
$$[(j_0 - j)(k-1)] < j_0 - j_1 \leq [(j_0 - j)k]))\}.$$

При изменении масштаба изображения по сравнению с эталонным изменятся также и множества $Light(D)$ и $\Gamma_h(D)$. Так, например, при растяжении E_k в k раз (по горизонтали и по вертикали) самого объекта ω_D можно считать, что это растяжение произошло относительно визуального центра и

$$Light(E_k(D)) = \{(i,j) : \exists\ i'j'((i',j') \in Light(D)\ \&$$
$$[(i_v - i')(k-1)] < i_v - i \leq [(i_v - i')k]\ \&$$
$$[(j_v - j')(k-1)] < j_v - j \leq [(j_v - j')k]))\},$$

$$\Gamma_h(E_k(D)) = \{(i,j) : \exists\ i'j'((i',j') \in \Gamma_h(D)\ \&$$
$$[(i_v - i')(k-1)] < i_v - i \leq [(i_v - i')k]\ \&$$
$$[(j_v - j')(k-1)] < j_v - j \leq [(j_v - j')k]))\}.$$

Список литературы

[1] Косовская Т.М., Тимофеев А.В. Об одном новом подходе к формированию логических решающих правил – Вестник ЛГУ, 1985, №8. С. 22–27.

[2] Косовская Т.М. Доказательства оценок числа шагов решения некоторых задач распознавания образов, имеющих логические описания // Вестн. С.-Петербург. ун-та. Сер. 1. 2007. Вып. 4. С. 82–90.

[3] Косовская Т.М. Многоуровневые описания классов для уменьшения числа шагов решения задач распознавания образов, описываемых формулами исчисления предикатов // Вестн. С.-Петербург.ун-та. Сер. 10. 2008. Вып.1. С. 64–72.

Косовская Татьяна Матвеевна — д.ф.-м.н., доцент; ст. научн. сотр. лаборатории информационных технологий в управлении и робототехнике СПИИРАН, профессор кафедры математики Санкт-Петербургского государственного морского технического университета (СПбГМТУ), профессор кафедры прикладной кибернетики математико-механического факультета Санкт-Петербургского государственного университета (СПбГУ). Область научных интересов:

логический подход к решению задач искусственного интеллекта, теория сложности алгоритмов. Число научных публикаций - 64. kosovtm@gmail.com; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-0421.

Tatiana M. Kosovskaya — Doctor of Computer Science, Associate Professor; senior researcher of Laboratory of Information Technologies in Control and Robotechniques, SPIIRAS, Professor of Applied Cybernetics Chair, SPbSU, Professor of Mathematics Chair, SPbSMTU. Research area: logical approach to the solving of Artificial Intelligence problems, theory of complexity of algorithms. Number of publications — 64. kosovtm@gmail.com; SPIIRAS, 14-th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-0421.

Поддержка исследований. Работа выполнена при финансовой поддержке РФФИ, проект № 09-08-00767-а.

Рекомендовано ИТУР СПИИРАН, зав. лаб. Тимофеев А.В.

Статья поступила в редакцию 20.10.2011.

РЕФЕРАТ

Косовская Т.М. **Распознавание изображений на экране дисплея и оценки числа шагов алгоритмов в рамках логико-предметной распознающей системы**

Для изображений на экране дисплея задачи выделения и распознавания изображения отдельного объекта на сложной сцене сведены к доказательству логического следования вида $S(T) \Rightarrow \exists \overline{y}_{\neq} A_k(\overline{y})$, где T — упорядоченная по строкам и столбцам таблица яркостей (и, быть может, цветов) пикселей дисплея; $S(T)$ — описание дисплея в терминах выбранных признаков, характеризующих состояния пикселей и представляющих собой упорядоченный список постоянных атомарных формул; $A_k(\overline{y})$ — описание изображений из заданного k-го класса, представляющее собой дизъюнкцию элементарных конъюнций тех же признаков и удовлетворяющее условию: если описание области D таблицы T удовлетворяет условию $A_k(\overline{y})$, то эта область представляет собой изображение из k-го класса.

Доказаны оценки числа шагов решения рассмотренной задачи в двух постановках.

Первая постановка соответствует выделению и распознаванию эталонного объекта на сложной сцене. Для решения задачи распознавания в такой постановке доказано существование полиномиального разрешающего алгоритма, и получена оценка числа его шагов.

Задача идентификации изображения на сцене во второй постановке соответствует нахождению на сцене изображения, в котором присутствуют характерные черты изображений из заданного класса, отражённые в описании класса $A_k(\overline{x})$. Для решения задачи распознавания в такой постановке доказана её принадлежность классу $\mathbf{NP}$, и получена оценка числа шагов решающего её алгоритма.

Для уменьшения числа шагов решения рассмотренных задач предложено понятие размытого изображения, содержащего границу изображения и области наибольшей яркости изображения.

Доказаны оценки числа шагов решения задачи идентификации изображения на сложной сцене при использовании двухуровневого описания классов изображений, первым уровнем которого являются описания классов размытых изображений.

Рассмотрена задача инвариантного (относительно изменения масштаба) распознавания изображения. Для её решения предлагается использовать описания преобразований масштаба. Введённые в статье понятия геометрического и визуального центров изображения позволяют упростить решение задачи распознавания изображения с изменённым масштабом.

SUMMARY

Kosovskaya T.M. **Display screen image recognition and bounds of number of steps of an algorithm in the frameworks of logic-objective recognition system.**

For displainf of screen image the problem of location and recognition of an object situated on a complecated scene is reduced to the proof of a logical sequent of the form $S(T) \Rightarrow \exists \overline{y}_{\neq} A_k(\overline{y})$, where T is an ordered table of brightnesses (and, may be, of colours) of display pixels; $S(T)$ is a description of the display in the terms of the chosen characteristics of a pixel state, $S(T)$ is an ordered list of constant atomic formulas; $A_k(\overline{y})$ is a description of images from the k-th class represented by a disjunction of elementary conjunctions of the same characteristics and sitisfying the condition: if a description of the domain D from the table T satisfies the condition $A_k(\overline{y})$ then this domain represents an image from the k-th class.

Estimations of number of steps of some algorithms solving the considered problem with two settings are presented.

The first setting of the problem corresponds to the location and recognition of an etalon object situated on a complecated scene. For the problem with such a setting it is proved that there exists a polynomial solving algorithm and the upper bound of its steps is recieved.

The second setting of the image identification problem corresponds to the finding such a part of the complicated scene which contains main characteristics of an image from the done class fixed in the class description $A_k(\overline{x})$. For the problem with such a setting it is proved that it belongs to the class **NP** and the upper bound of the solving algorithm steps is recieved.

To decrease the number of steps of the described problems solving algorithms the notion of „fuzzy image“, containing frontier of an image and the domain of the maximal brightness of an image, is introduced.

Bounds of steps number of an algorithm using two-level description (the first one of which is a „fuzzy image“) for the identification problem of an image situated on a complecated scene are proved.

The problem of invariant (under rescaling) image recognition is regarded. Descriptions of rescaling are offered for solving such a problm. The introduced in rhe paper notions of geometrical and visual centers of an image allow to simplify the solution of the recognition problem for a rescaled image.