



УДК 519.6

Комбинаторные задачи высокой сложности и анализ плоских контурных изображений

М. Д. Каташевцев, В. И. Мартьянов
Иркутский государственный университет

Аннотация. В статье рассмотрена задача анализа плоских контурных изображений в рамках логико-эвристической парадигмы. Введена модель представления бинарного растрового изображения в виде многоосновной алгебраической системы, построенной на множествах дуг и связей между дугами. Описаны алгоритмы распознавания изображения в рамках введенной модели, ранее успешно используемые для решения задач анализа генетического кода и составления расписания.

Ключевые слова: анализ изображений; распознавание образов; логико-эвристический; растровое изображение; алгебраические системы.

1. Введение

Важнейшим направлением развития современных информационных технологий (и в целом прогресса, так как невозможно представить современную науку и технику без использования компьютеров) является создание математических методов быстрого решения информационных задач.

В идеале, эти методы должны обеспечивать скорость решения ряда важных информационных задач вне зависимости от объема данных. И действительно, есть ряд важных информационных задач, где это возможно.

Наглядным (и очень важным) примером этого являются реляционные базы данных (БД), где вычислимость запросов определенных типов не зависит от объема данных, а только линейно от сложности проекта самой БД.

В основе этого лежит следующий, довольно просто доказываемый, факт — сложность проверки принадлежности кортежа (a_1, a_2, \dots, a_n) , где $a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n$, отношению H , определенному на конечных множествах A_1, A_2, \dots, A_n , линейна от n (т.е. не зависит от числа кортежей, составляющих отношение H).

При компьютерной реализации конечные множества A_1, A_2, \dots, A_n являются доменами, а отношение H — таблицей реляционной БД.

Соответственно, вопрос о принадлежности кортежа (a_1, a_2, \dots, a_n) отношению H , сводится к проверке не пустоты курсора, полученного после выполнения данного SQL-запроса:

```
SELECT *
```

```
FROM H
```

```
WHERE  $A_1 = a_1$  AND  $A_2 = a_2$  AND ... AND  $A_n = a_n$ ,
```

что, при наличии общего индекса по столбцам соответствующим A_i , выполняется со скоростью линейно зависящей от n . Если искать истоки данных результатов, то, наверное, следует обратиться к Д.Кнуту [1], где представлен древовидный способ организации данных.

Практическим подтверждением этого является скорость работы банковских систем, использующих сетевые реляционные БД, с их мировыми сетями терминалов и банкоматов, где можно проводить операции с вкладами и денежными средствами в любой точке мира за считанные секунды.

Отметим, что практически к этому классу примыкают задачи поиска данных по ключевым словам в интернет-пространстве (поисковые системы «Google» [2], «Яндекс» и др.), где скорость поиска не замедляется из-за экспоненциального роста информации в глобальной сети.

Одним из важнейших классов информационных задач являются задачи распознавания образцов (образов), где также могут быть получены результаты независимости скорости распознавания образцов от их количества [3], вернее, верхней границы сложности распознавания одного образца с добавкой только количества образцов.

Результаты работы [3] получены для простой организации данных, практически не отличающейся (если абстрагироваться от специфики предметных задач сетевого планирования) от рассмотрения проверки принадлежности кортежа (a_1, a_2, \dots, a_n) отношениям H_1, H_2, \dots, H_k , которые определены на конечных множествах A_1, A_2, \dots, A_n .

Для такой постановки соответствующий результат работы [3] может быть сформулирован следующим образом

Теорема 1. *Верхняя граница сложности проверки принадлежности кортежа (a_1, a_2, \dots, a_n) отношениям H_1, H_2, \dots, H_k не превышает $O(n + k)$.*

Отметим, что другие формы организации данных для ряда предметных задач рассмотрены в [4]. В настоящей статье предложен определенный способ формального (математического) описания плоских связных контурных изображений и в рамках данной формализации доказано, что сложность распознавания образцов не превышает $O(n) + O(k)$, где $O(n)$ — верхняя граница сложности распознавания одного образца, k — количество образцов.

Возможно, данный результат до некоторой степени объясняет феномен мгновенного распознавания практически неограниченного количества образов зрением живых организмов.

Замечание 1. Очень важно учитывать, что графы, представляющие плоские контурные изображения, не являются произвольными, так как дуги, исходящие из одной вершины, располагаются по окружности 360 градусов и, следовательно, вполне упорядочены, а также имеют минимальный угол пересечения (т. е. количество дуг, исходящих из одной вершины графа ограничено). Да и сами дуги имеют ограниченное количество численных характеристик, см. замечание 2. Таким образом, не надо трактовать результаты статьи, как решение известных NP-полных задач для графов [5][6].

2. Формализация описания плоских контурных изображений

Составляющими элементами плоских контурных изображений будем считать дуги и связи дуг. Дуга arc основной количественной характеристикой имеет сектор окружности, измеряемый в градусах (точнее, в количестве минимальных шагов возрастания градусной меры дуги, что обеспечивает конечность количественных характеристик в некоторой шкале или масштабе).

Замечание 2. Отметим, что любые две несовпадающие точки a и b на плоскости (задающие луч \overline{ab}) можно соединить дугой заданной градусной меры α ($0 \leq \alpha \leq 360$) ровно двумя способами, в первом случае все точки дуги будут лежать справа от луча \overline{ab} , будем говорить что дуга обходится по часовой стрелке, во втором случае все точки дуги будут лежать слева от луча и речь будет идти об обходе против часовой стрелки. Для дуг градусной меры $\alpha \in \{0, 360\}$ направление обхода не определено.

Связь дуг rel основной количественной характеристикой имеет угол между дугами, измеряемый в градусах (точнее, в количестве минимальных шагов возрастания углов, что обеспечивает конечность количественных характеристик в некоторой шкале или масштабе).

Основными математическими моделями для данного подхода будут трехосновные алгебраические системы [7][8] вида

$$M = \langle Arc, Rel, V; Sector, Angle, R \rangle \quad (2.1)$$

где основное множество Arc – совокупность дуг; основное множество Rel – совокупность связей дуг; основное множество V – некоторый начальный отрезок натуральных чисел (представляет сектора дуг и углы связей дуг в некоторой шкале); одноместная функция $Sector$:

$Arc \rightarrow V$, т. е. определяет градусную меру дуги; одноместная функция $Angle : Rel \rightarrow V$, т. е. определяет угол связи дуг; трехместное отношение R соединяет связь дуг rel с соответствующими дугами, т. е. R – подмножество декартова произведения $Rel \times Arc \times Arc$.

Замечание 3. Для наших целей важно всегда работать только с конечными множествами, что достигается рассмотрением конечных множеств Arc , Rel , а также предположением о наличии минимального шага возрастания количественных характеристик дуг и связей дуг, т. е. конечное множество V имеет минимальное ненулевое значение, соответствующее минимальному шагу, и максимальное, соответствующее 360 градусам.

3. Преобразование растрового изображения

Построение формальной модели на основе растрового изображения состоит из двух этапов. Первый этап – волновая скелетизация. С помощью скелетизации на основе растрового изображения строится граф (скелет), который визуальнo адекватно соответствует исходному изображению.

Определение 1. *Бинарное растровое изображение есть функция*

$$f : X \times Y \rightarrow \{0, 1\}, X \in \mathbb{Z}, Y \in \mathbb{Z}$$

Определение 2. *Заполненная область есть множество*

$$F = \{(x, y) | f(x, y) = 1\}$$

Определение 3. *Точку $q(x_1, y_1)$ будем называть соседом точки $p(x, y)$ если $|x - x_1| \leq 1$ и $|y - y_1| \leq 1$ и $p \neq q$. Введем отношение соседства $N(p, q)$, которое истинно если p сосед q .*

Очевидно что точка p не может иметь более 8 соседей. Обозначим через N_F^p множество всех соседей точки $p(x, y)$ лежащих в заполненной области F :

$$N_F^p = \{q | q \in F \wedge N(p, q)\}$$

Далее будем считать что заполненная область F не имеет изолированных точек т.е.

$$\forall p \exists q : N(p, q) \\ p, q \in F$$

Замечание 4. Скелетом изображения f будем называть граф $G(V, E)$, «интуитивно адекватно отражающий» исходное изображение.

Определение 4. Волной w будем называть конечное множество точек $\{p_j\}$.

Определение 5. Множество волн $\{w_1, w_2, \dots, w_n\}$ будем называть подволнами волны w если:

$$\bigcup_{i=1, n} w_i = w$$

$$w_i \cap w_j = \emptyset, \quad i, j = \overline{1, n}, \quad i \neq j;$$

и для любых двух точек $p \in w_i, g \in w_j$, где $i \neq j$ верно $\neg N(p, g)$.

Введем функцию вычисляющую центр масс точек волны

$$g(w) = \frac{\sum_{p \in w} p}{|w|}$$

Опишем алгоритм построения скелета изображения f с непустой, связной заполненной областью F

Алгоритм 1. Зададим начальные условия. В качестве начальной волны подойдет любая точка из F . Имеем следующую начальную конфигурацию:

$w_0^0 = \{p\}, p \in F$ – начальная волна,

$W_0 = \{w_0^0\}$ – множество волн,

$F_0 = F$ – состояние заполненной области,

$G_0(V_0, E_0), V_0 = \{p\}, E_0 = \emptyset$ – начальное состояние скелета.

Определим n -й шаг итерации следующим образом. Для всякой i -й волны $w_i^{k_i-1}$ из W_{n-1} (k_i – соответствует k_i -й итерации w_i):

$$w_i^{k_i} = \bigcup_{p \in w_i^{k_i-1}} N_{F_{n-1}}^p \setminus \bigcup_{j < i} w_j^{k_j-1} \quad (3.1)$$

Если u_1, \dots, u_m есть подволны волны $w_i^{k_i}$, тогда

$$W_n^i = \{w_{l+1}, \dots, w_{l+m}\}$$

где

$$w_{l+j} = u_j, \quad j = \overline{1, m}, \quad l = |W_{n-1}| + \sum_{j < i} |W_n^j|.$$

Ребра в графе образуют вектора, связывающие центры масс полученных подволн с центром массы $w_i^{k_i-1}$

$$E_n^i = \{(v_i^{k_i-1}, v_{l+j}^{k_i})\}, \quad j = \overline{1, m}$$

$$V_n^i = \{v_{l+j}^{k_i}\}, j = \overline{1, m}$$

$$v_i^j = g(w_i^j)$$

Если же волна $w_i^{k_i}$ не имеет разрывов и $w_i^{k_i} \neq \emptyset$, то

$$W_n^i = \{w_i^{k_i}\}, E_n^i = \{(v_i^{k_{i-1}}, v_i^{k_i})\}, V_n^i = \{v_i^{k_i}\}$$

Если $w_i^{k_i} = \emptyset$

$$W_n^i = \emptyset, E_n^i = \emptyset, V_n^i = \emptyset$$

Таким образом, при $s_n = |W_{n-1}|$:

$$W_n = \{W_n^i\}, i = \overline{1, s_n}$$

$$V_n = V_{n-1} \bigcup_i V_n^i, i = \overline{1, s_n}$$

$$E_n = E_{n-1} \bigcup_i E_n^i, i = \overline{1, s_n}$$

$$F_n = F_{n-1} \setminus P_{n-1},$$

$$P_{n-1} = \{p : p \in w, w \in W_{n-1}\}$$

Если $|F_n| = 0$ то алгоритм прекращает цикл итераций, а граф

$$G = (V_n, E_n)$$

является скелетом исходного изображения f .

Утверждение 1. Алгоритм «волновая скелетизация» остановится на любой заполненной области размерности m , удовлетворяющей условиям алгоритма.

Доказательство. Пусть $m = 1$, тогда согласно алгоритму

$$|F_1| = |F_0 \setminus \{p : p \in w, w \in W_0\}| = |\{p\} \setminus \{p : p \in w_0\}| = |\{p\} \setminus \{p\}| = 0$$

следовательно алгоритм прекращает свою работу а граф $G = (V_1, E_1) = (\{p\}, \emptyset)$ является скелетом изображения.

Пусть $m > 2$ и существует такое l что для всякого $k < l$, $|F_k| < |F_{k-1}|$ и $|F_l| = |F_{l-1}|$, тогда $P_{n-1} = \emptyset$, отсюда следует, что $\{p : p \in w, w \in W_{n-1}\} = \emptyset$, а это возможно только в двух случаях:

1. если W_{n-1} пусто, тогда, в силу 3.1 и в силу отсутствия изолированных точек, $F_{n-2} = \emptyset$, получаем противоречие с условием остановки.

2. если $\forall w \in W : |w| = 0$, тогда, опять же в силу 3.1 и в силу отсутствия изолированных точек, получаем что $F_{n-2} = \emptyset$, снова получаем противоречие с условием остановки.

Следовательно такого l не существует, и алгоритм сходится для всякой непустой заполненной области без изолированных точек. \square

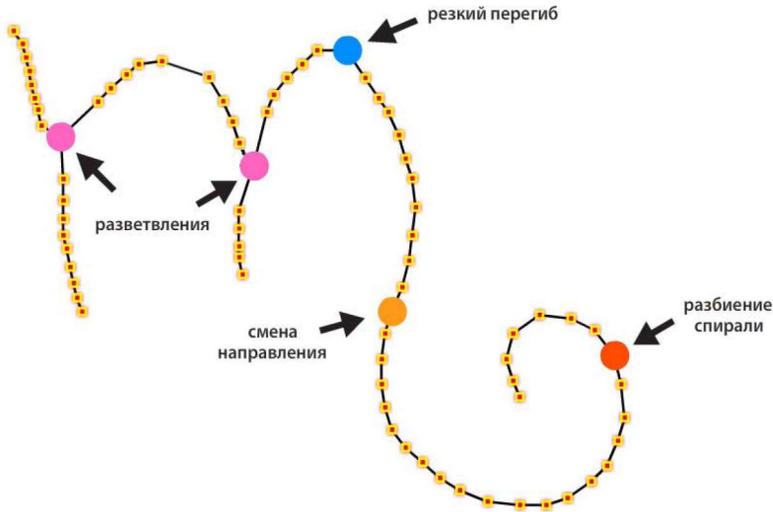


Рис. 1. Схема разбиения графа «особыми» точками

Второй этап – интерполяция графа дугами. Граф разбивается на простые пути. Разбиение графа на простые пути идет по вершинам имеющих количество соседей не равное 2.

Определение 6. Будем говорить что в узле v_i меняется направление обхода простого пути

$$\{v_1, \dots, v_{i-1}, v_i, v_{i+1}, \dots, v_n\}$$

в окрестности $\epsilon \in \{2, 3, \dots\}$, если

$$\text{Sign}\left(\sum_{j=i-\epsilon} r_{i,j}^\epsilon\right) \neq \text{Sign}\left(\sum_{j=i+\epsilon} r_{i,j}^\epsilon\right)$$

где $r_{i,j}^\epsilon$ – расстояние со знаком от точки определяемой узлом v_j до вектора образуемого точками узлов v_i и $v_{i-\epsilon}$ Если соотношение выполняется для $\epsilon = 2$ будем говорить, что в точке v_i меняется направление обхода.

1. Для каждого простого пути выполняется:

- a) Разбиение пути по точкам смены направления обхода
 - b) Для каждого разбиения выполняются
 - i) Разбиение спиралей. Чтобы определить закручен ли путь по спирали, надо проверить пересекает ли хорда путь. Если пересечение есть, то необходимо разбить путь точками пересечения
 - ii) Для каждого разбиения выполняется:
 - A) Разбиение по точкам перегиба. Точками перегиба считаются образующие две дуги отклонившиеся от угла идеального соединения. Угол γ идеального соединения двух дуг градусной меры α и β : $\gamma = 2\pi - (\alpha + \beta)$
2. Результатом п. 1 является множество подпутей, каждый из которых переводится в дугу. Градусная мера дуги вычисляется с использованием формулы Гюйгенса. Определив наиболее удаленную точку пути от хорды стягивающей путь, и вычислив её положение относительно хорды направленной от начала к концу пути, мы определяем направление обхода. Если точка слева, то обход ведется по часовой стрелке, если точка справа — обход ведется против часовой стрелки, если же точка лежит на прямой, то верны оба утверждения.
3. Расчет связей дуг. Связь между двумя дугами существует, если пути образующие дуги имели общие вершины. Угол соединения между дугами рассчитывается, как угол между стягивающими их хордами

Замечание 5. Стоит отметить что разбиение на дуги, как правило, выполняется на интерполированном графе в котором часть узлов удаленно в силу их избыточности.

Замечание 6. Хотя разбиение и может быть использовано как есть, на практике полезнее добавлять возможность вариации параметров, например, допускать возможность некоторого отклонения от угла идеального соединения или для точек смены направления расширять область проверки на смену направления обхода.

4. Постановка задачи анализа плоских контурных изображений

Общую схему решения комбинаторных задач высокой сложности логико-эвристическими методами [4] можно трактовать как преобразо-

вание начальной (инициальной) многоосновной алгебраической системы [6, 7] (например 2.1)

$$M_{ini} = \langle A_1, \dots, A_s; f_1, \dots, f_n; p_1, \dots, p_k \rangle \quad (4.1)$$

где A_i – основные множества, f_i – операции (функции) на основных множествах, p_i – предикаты (отношения) на основных множествах, в конечную (финальную) M_{fin} , удовлетворяющую ограничениям R_1, R_2, \dots, R_m .

Если искать аналоги, то последовательность таких преобразований можно считать допустимым (без оптимизации значений целевого функционала) управлением для задачи динамического программирования [9], где R_1, R_2, \dots, R_m фазовые ограничения.

При программной реализации многоосновная а.с. 4.1 становится реляционной БД, поиск последовательности преобразований для построения финальной а.с., удовлетворяющей ограничениям – комбинаторной задачей высокой сложности [4].

Для рассматриваемой здесь предметной области (анализ изображений) общая схема решения пока не может быть применена в полном объеме из-за начального этапа исследований (с точки зрения логико-эвристических методов) и отсутствия конкретных постановок прикладных задач (ближайшие планы применения логико-эвристических методов обсуждаются в заключении).

Поэтому ограничимся исследованием сложности проверки выполнимости ограничений R_1, R_2, \dots, R_m на математических моделях вида 2.1 с позиции обеспечения независимости скорости проверки от числа ограничений.

Проверка выполнимости ограничений сводится к проверке вложенности обобщенных изображений (образцов) в анализируемое изображение 2.1.

Уточним формализацию описания плоских контурных изображений для данного варианта проверки выполнимости ограничений.

Составляющими элементами образцов и анализируемого изображения будут также дуги и связи дуг. Численными характеристиками которых будут количество минимальных шагов возрастания для дуги (связи дуг) – градусные меры дуг, имеющие минимальное и максимальное значения.

Таким образом, одноместная функция $Sector : Arc \rightarrow V$ преобразуется в одноместную же функцию $Sector : Arc \rightarrow V \times V$, соответственно, одноместная функция $Angle : Rel \rightarrow V$ преобразуется в одноместную же функцию $Angle : Rel \rightarrow V \times V$.

Замечание 7. В принципиальном плане ввод минимальных и максимальных значений ничего нового не дает (от минимального до максимального всего конечное множество значений), но позволяет более

компактно задавать искомые образцы. Для упрощения технических деталей будем считать, что для анализируемых изображений функции *Sector* и *Angle* имеют одинаковое минимальное и максимальное значение

Пусть многоосновные а.с.

$$\begin{aligned} R_1 &= \langle Arc_1, Rel_1, V; Sector, Angle, R \rangle \\ R_2 &= \langle Arc_2, Rel_2, V; Sector, Angle, R \rangle \\ &\dots \\ R_m &= \langle Arc_m, Rel_m, V; Sector, Angle, R \rangle \end{aligned} \quad (4.2)$$

задают искомые образцы в анализируемом изображении 2.1.

Анализ изображения 2.1 состоит в поиске всех изоморфных вложений $\mu_{i,j}$ многоосновных а.с. R_1, R_2, \dots, R_m в многоосновную а.с. M 2.1, т.е. [6, 7] изоморфное вложение $\mu_{i,j} : R_i \rightarrow M$ состоит из инъективных отображений

$$\mu_{i,j} : Arc_i \rightarrow Arc; \mu_{i,j} : Rel_i \rightarrow Rel \quad (4.3)$$

такие, что:

- а) если $\mu_{i,j}(Ar) = Arr$, где $Ar \in Arc_i, Arr \in Arc, Sector(Ar) = (v_1, v_2), Sector(Arr) = (v_3, v_4)$, то $v_1 \leq v_3 \leq v_4 \leq v_2$;
- б) если $\mu_{i,j}(Re) = Ree$, где $Re \in Rel_i, Ree \in Rel, Angle(Re) = (v_1, v_2), Angle(Ree) = (v_3, v_4)$, то $v_1 \leq v_3 \leq v_4 \leq v_2$;
- в) если $R(Re, Ar_1, Ar_2)$, где $Re \in Rel_i, Ar_1 \in Arc_i, Ar_2 \in Arc_i$, то $R(\mu_{i,j}(Re), \mu_{i,j}(Ar_1), \mu_{i,j}(Ar_2))$.

5. Оценка сложности анализа изображений

Для облегчения понимания идеи доказательства основного результата, рассмотрим доказательство теоремы 1. Основой его является представление декартова произведения конечных множеств

$$A_1 \times A_2 \times \dots \times A_n$$

в древовидной форме Tree [1].

Определим более точно конечные множества

$$\begin{aligned} A_1 &= \{a_{1,1}, a_{1,2}, \dots, a_{1,m_1}\}; \\ A_2 &= \{a_{2,1}, a_{2,2}, \dots, a_{2,m_2}\}; \\ &\dots \\ A_n &= \{a_{n,1}, a_{n,2}, \dots, a_{n,m_n}\}; \end{aligned} \quad (5.1)$$

Понятно, что таблица 1 является универсумом для любых таблиц реляционной БД с доменами A_1, A_2, \dots, A_n . Ее представление отношения H в таблице состоит в пометке вершин n -го этажа, если путь от корня дерева до этой вершины n -го этажа, дает кортеж из отношения H .

Проверка принадлежности кортежа (a_1, a_2, \dots, a_n) , где $a_1 \in A_1, a_2 \in A_2, \dots, a_n \in A_n$, отношению H производится за n шагов в таблице 1 (этот процесс в дальнейшем будем называть интерпретацией кортежа (a_1, a_2, \dots, a_n) на дереве $Tree$). Действительно, a_1 позиционируется на 1-м этаже за 1 шаг, a_2 позиционируется на 2-ом этаже за 1 шаг и так далее, a_n позиционируется на n -м этаже за 1 шаг, где и определяется принадлежность кортежа (a_1, a_2, \dots, a_n) отношению H .

Таким образом, проверка осуществляется за n шагов. Для доказательства теоремы достаточно пометить вершины n -го этажа на принадлежность отношениям H_1, H_2, \dots, H_k . Тогда в результате интерпретации кортежа (a_1, a_2, \dots, a_n) за n шагов на дереве $Tree$ получим вершину n -го этажа пометки которой покажут принадлежность (или не принадлежность) отношениям H_1, H_2, \dots, H_k .

Оценка $O(n + k)$, а не $O(n)$, получается из-за необходимости пройти по списку отметок n -го этажа, что и дает добавку $O(k)$.

Замечание 8. Результат теоремы 1 типичный, так называемый обмен памяти на эффективность [5]. Конечно, задание декартово произведения деревом увеличивает необходимый объем памяти, но скорость выполнения операций предельно ускоряется. Следует отметить также, что на практике универсум (рис. 1) не строится, а строится только его часть, состоящая из кортежей отношений H_1, H_2, \dots, H_k . Вообще говоря, это замедляет скорость интерпретации, но незначительно не более, чем на $\ln(m)$, где $m = \max\{m_1, m_2, \dots, m_n\}$ (это связано с необходимостью перебора узлов «частичного» дерева, что, в силу упорядоченности, можно реализовать с помощью бинарного поиска).

Прежде, чем перейти к изложению основного результата, определим универсум для изображений, имеющих не более n дуг, и k вариантов дуг и связей дуг, т. е. множество V имеет k элементов, а минимальный сектор дуги или угол пересечения дуг равен $(360/k)$ градусов.

#	Значения вершин дерева												
0	Корень дерева												
1	$a_{1,1}$				$a_{1,2}$...	a_{1,m_1}			
2	$a_{2,1}$	$a_{2,2}$...	a_{2,m_2}	$a_{2,1}$	$a_{2,2}$...	a_{2,m_2}	...	$a_{2,1}$	$a_{2,2}$...	a_{2,m_2}
	...												
n	$a_{n,1}$	$a_{n,2}$...	a_{n,m_n}	$a_{n,1}$	$a_{n,2}$...	a_{n,m_n}	...	$a_{n,1}$	$a_{n,2}$...	a_{n,m_n}

Рис. 2. Дерево Tree

Пусть $Arc_1, Arc_2, \dots, Arc_n$ – множества дуг всех характеристик (образцов), т.е.

$$\begin{aligned} Arc_1 &= \{ar_{1,1}, ar_{1,2}, \dots, ar_{1,m_1}\}; \\ Arc_2 &= \{ar_{2,1}, ar_{2,2}, \dots, ar_{2,m_2}\}; \\ &\dots \\ Arc_n &= \{ar_{n,1}, ar_{n,2}, \dots, ar_{n,m_n}\}; \end{aligned} \quad (5.2)$$

Далее пусть $Rel_1, Rel_2, \dots, Rel_{n-1}$ – множества связей дуг всех характеристик, т.е.

$$\begin{aligned} Rel_1 &= \{re_{1,1}, re_{1,2}, \dots, re_{1,k_1}\}; \\ Rel_2 &= \{re_{2,1}, re_{2,2}, \dots, re_{2,k_2}\}; \\ &\dots \\ Rel_{n-1} &= \{re_{n-1,1}, re_{n-1,2}, \dots, re_{n-1,k_{n-1}}\}; \end{aligned} \quad (5.3)$$

$$Angle(re_{i,j}) = (j, j)$$

Дерево $TreeImage$ (универсум (рис. 2) для всех изображений, имеющих не более n дуг, и k вариантов дуг и связей дуг) строится по аналогии с деревом $Tree$ для декартова произведения

$$Arc_1 \times Rel_1 \times Arc_2 \times Rel_2 \times \dots \times Rel_{n-1} \times Arc_n$$

Соглашения по представлению элементов рис. 2 следующие:

1. для элемента $\alpha_{\mu,\chi}^\theta$ – число θ является позицией на этаже схемы (номер клетки в строке); $\alpha_{\mu,\chi}^\theta$ является χ -ым элементом из множества Arc_μ или Rel_μ ;
2. числа m_i , где i – номер этажа, равны $k^i - k$; число $t = k^{2n-1} - k$, отметим, что данные числа имеют чисто технический характер и уменьшают громоздкость выражений, стоящих в конце строк рис. 2.

Очевидно, что рис. 2 содержит все изображения, имеющие не более n дуг, и k вариантов дуг и связей дуг. Структуру дерева на рис. 2 будем задавать отношениями $Par_{arc}(x, y)$, $Brot_{arc}(x, y)$ для дуг, и $Par_{rel}(x, y)$ для связей дуг.

#	Значения вершин дерева												
0	Корень дерева												
1	$ar_{1,1}^1$			$ar_{1,2}^2$...	$ar_{1,m_1}^{m_1+k-1}$					
2	$re_{1,1}^1$	$re_{1,2}^2$...	$re_{1,k}^k$	$re_{1,1}^{k+1}$	$re_{1,2}^{k+2}$...	$re_{1,k}^{2k}$...	$re_{1,1}^{m_2}$	$re_{2,2}^{m_2+1}$...	$a_{2,k}^{m_2+k-1}$
	...												
$2n-1$	$ar_{n,1}^1$	$ar_{n,2}^2$...	$ar_{n,k}^k$	$ar_{n,1}^{k+1}$	$ar_{n,2}^{k+2}$...	$a_{n,k}^{2k}$...	$ar_{n,1}^t$	$a_{n,2}^{t+1}$...	$a_{n,k}^{t+k-1}$

Рис. 3. Дерево $TreeImage$

Отношение $Par_{arc}(x, y)$ задает отношение «родитель – потомок» на декартовом произведении $Arc \times Rel$, например, $Par_{arc}(ar_{i,j}^2, re_{i,w}^{k(j-1)+w})$, где $1 \leq w \leq k$. Отношение $Par_{arc}(x, y)$ связывает элементы, расположенные на соседних этажах, и может быть определено строго математически, а именно, $Par_{arc}(\alpha_{\mu,\chi}^\theta, \beta_{\zeta,\eta}^\gamma)$ тогда и только тогда, когда

1. $\mu = \zeta$ или $\mu + 1 = \zeta$
2. $(\theta - 1)k \leq \zeta \leq (\theta - 1)k + k - 1$

Отношение $Brot_{arc}(x, y)$ задает отношение «быть братом» на декартовом произведении $Arc \times Arc$. Отношение $Brot_{arc}(x, y)$ связывает элементы, расположенные на одном этаже и связанные с одним элементом верхнего этажа отношением «родитель-потомок», и может быть определено строго математически, а именно, $Brot_{arc}(\alpha_{\mu,\chi}^\theta, \beta_{\zeta,\eta}^\gamma)$ тогда и только тогда, когда

1. $\alpha = \beta$;
2. $\mu = \zeta$;
3. $\theta < \zeta$ и $\theta - \zeta < k$, а также $[\theta/k] > 0$, где операция $[]$ остаток от деления.

Отношение Par_{rel} задается по аналогии, на декартовом произведении $Rel \times Arc$.

Интерпретация ξ произвольного связного изображения 2.1 (в дальнейшем термин «изображение» будет означать только «изображение, имеющие не более n дуг, и k вариантов дуг и связей дуг», если, конечно, не оговорено противное), где

$$Arc = \{ar_1, ar_2, \dots, ar_w\}, w \leq n, Rel = \{re_1, re_2, \dots, re_t\} \quad (5.4)$$

на дереве $TreeImage$ производиться по следующей схеме

Основание индукции. Пусть $i = 1$ и $Sector(ar_1) = (j, j)$ и

$$Rrel_1 = \{re_i \mid R(re_i, Ar_1, Ar_2), Ar_1 = ar_1 \vee Ar_2 = ar_1\}$$

$$Arr_1 = \{aar_i \mid R(re_i, Ar_1, Ar_2), re_i \in Rrel_1, (Ar_1 = ar_1 \wedge Ar_2 = aar_i) \vee (Ar_2 = ar_1 \wedge Ar_1 = aar_i)\}.$$

Тогда полагаем $\xi(ar_1) = ar_{1,j}^j$, $\xi(re_i) = re_{1,v}^{(j-1)*k+v}$, где $re_i \in Rrel_1$, $Angle(re_i) = Angle(re_{1,v}) = v$. Если $aar_i \in Arr_1$, то $\xi(aar_i) = ar_{2,e}^{(d-1)*k+e}$, где d – позиция элемента $\xi(re_i)$ (т.е. $d = (j-1)*k + v$, $Sector(aar_i) = (e, e)$).

Отметим, что $Par_{arc}(\xi(ar_1), \xi(re_i))$, $Par_{rel}(\xi(re_i), \xi(aar_i))$, а также для любых aar_i и aar_j из Arr_1 выполняется $Brot_{arc}(\xi(aar_i), \xi(aar_j))$.

Интерпретация ξ продолжается для множества дуг

$$Arc_1 = Arc \setminus (\{ar_1\} \cup Arr_1),$$

множества связей дуг

$$Rel_1 = Rel \setminus Rrel_1.$$

Замечание 9. Важнейшим моментом построения отношений

$$Par_{arc}(x, y), Brot_{arc}(x, y) \text{ и } Par_{rel}(x, y)$$

на рис. 2 является их конструктивизм (эффе́ктивная вычислимость за один шаг) и это свойство сохраняется при построении интерпретации ξ , как для основания индукции (так и для индукционного шага, что будет показано ниже).

Индукционный шаг. Пусть после i -го шага получены непустые множества дуг $Arr_i = \{ar_{i_1}, ar_{i_2}, \dots, ar_{i_w}\}$, $Arc_i = Arc_{i-1} \setminus Arr_i$, множества связей дуг $Rrel_i, Rel_i = Rel_{i-1} \setminus Rrel_i$, причем по аналогии с основанием индукции, дуги из множества $\xi(Arr_i)$ располагаются на $2 * i + 1$ этаже рис. 2, связи дуг из множества $\xi(Rrel_i)$ располагаются на $2 * i$ этаже рис. 2.

По алгоритму основания индукции будем проводить построения для каждой дуги $ar_\alpha \in Arr_i$ такой, что ar_α не принадлежит объединению $\{ar_1\} \cup Arr_1 \cup \dots \cup Arr_{i-1}$. Пусть $Sector(ar_\alpha) = (j, j)$ и

$$Rrel_{i+1,\alpha} = \{re_s \mid R(re_s, Ar_1, Ar_2), re_s \in Rel_i, Ar_1 = ar_\alpha \vee Ar_2 = ar_\alpha\}$$

$$Arr_{i+1,\alpha} = \{aar_u \mid R(re_s, Ar_1, Ar_2), \\ (Ar_1 = ar_\alpha \wedge Ar_2 = aar_u) \vee (Ar_2 = ar_\alpha \wedge Ar_1 = aar_u)\}$$

Пусть $\xi(ar_\alpha) = ar_{i+1,j}^\beta$. Тогда $\xi(re_s) = re_{i+1,v}^{(\beta-1)*k+v}$, где $re_s \in Rrel_{i+1,\alpha}$,

$$Angle(res) = Angle(re_{i+1,v}^{(\beta-1)*k+v}) = v.$$

Если $aar_u \in Arr_{i+1,\alpha}$, то $\xi(aar_u) = ar_{2,e}^{(d-1)*k+e}$, где d – позиция элемента $\xi(re_s)$ (т.е. $d = (\beta - 1) * k + v$), $Sector(aar_u) = (e, e)$

Отметим, что $Par_{arc}(\xi(ar_\alpha), \xi(re_s))$, $Par_{rel}(\xi(re_s), \xi(aar_u))$, а также для любых aar_i и aar_j из $Arr_{i+1,\alpha}$ выполняется $Brot_{arc}(\xi(aar_i), \xi(aar_j))$.

Полагаем множество дуг Arr_{i+1} равным объединению всех $Arr_{i+1,\alpha}$, где ar_α произвольная дуга из множества Arr_i такая, что ar_α не принадлежит объединению $\{ar_1\} \cup Arr_1 \cup \dots \cup Arr_{i-1}$, также множество связей дуг $Rrel_{i+1}$ полагаем равным объединению всех $Rrel_{i+1,\alpha}$, соответствующие произвольным дугам ar_α из множества Arr_i (смотри, выше).

Интерпретация ξ продолжается для множества дуг

$$Arc_{i+1} = Arc_i \setminus Arr_i,$$

и для множества связей дуг

$$Rel_{i+1} = Rel_i \setminus Rrel_{i+1}.$$

Так как по условию интерпретируемое изображение 5.4 является связным, то процесс интерпретации ξ будет закончен не более, чем за w шагов индукции, где w – количество дуг.

Отметим, что соответствие при построении интерпретации ξ для любой дуги или связи дуг изображения 5.4 производится за один шаг, так как «связывание», соответствующего элемента рис. 2, производится вычислением одной арифметической формулы. Таким образом, доказана

Лемма 1. *Верхняя граница сложности построения интерпретации ξ для связного изображения 2.1 не превышает $O(w + t)$, где w – количество дуг, t – связей дуг.*

Теорема 2. *Пусть каждая из многоосновных а.с. R_1, R_2, \dots, R_m 4.2 имеет не более n дуг и представляет связное изображение. Тогда анализ связного изображения 2.1 имеет верхнюю границу сложности не превышающую $O(((w + t) * w) + m)$, , где w – количество дуг (t – количество связей дуг) изображения 2.1, причем множества дуг и связей дуг представлены выражениями 5.4.*

Доказательство. Построим интерпретации всех многоосновных а.с. R_1, R_2, \dots, R_m на универсуме рис. 2 для всех изображений, имеющих не более n дуг, и k вариантов дуг и связей дуг (сложность этой процедуры, конечно, не входит в оценку доказываемой теоремы).

Далее, пометим все вершины рис. 2 номерами многоосновных а.с., чьи элементы соответствуют этим вершинам. Каждой многоосновной а.с. R_i сопоставим пару чисел (a_i, b_i) , где a_i – количество помеченных вершин 2, соответствующих дугам, b_i – количество помеченных вершин рис. 2, соответствующих связям дуг (конечно, помеченных номером i).

Построим совокупность интерпретаций $\xi_1, \xi_2, \dots, \xi_w$ на рис. 2 (с помеченными вершинами), которые отличаются выбором первой дуги для основания индукции. А именно, интерпретация ξ_1 начинается традиционно с дуги ar_1 , интерпретация ξ_2 начинается с дуги ar_2 , и так далее. Последняя интерпретация ξ_w начинается, соответственно, с дуги ar_w .

Введем для каждой интерпретации ξ_i множество пар

$$(a_{i_1}, b_{i_1}), (a_{i_2}, b_{i_2}), \dots, (a_{i_m}, b_{i_m}) \tag{5.5}$$

где $a_{i_j}(b_{i_j})$ - количество помеченных j вершин рис. 2, соответствующих дугам (соответственно, связям дуг), полученных для интерпретации ξ_i . Если пара (a_{i_j}, b_{i_j}) равна паре (a_j, b_j) , то, таким образом, найдено изоморфное вложение j -го изображения (образца) в анализируемое изображение .

В силу леммы 1, построение каждого отображения ξ_i требует не более $w + t$ шагов и, таким образом, верхняя граница сложности поиска всех изоморфных вложений не более $O(((w+t)*w) + m)$, где «добавка» $O(m)$ возникает из-за необходимости сравнивать пары 5.5 с парами (a_j, b_j) . \square

Замечание 10. Как и в случае для базовой постановки задачи, на практике универсум (рис. 2) не строится а строится только его часть, состоящая из дуг и связей дуг многоосновных а.с. R_1, R_2, \dots, R_m 4.2.

6. Заключение

1. Адекватность предложенной формализации представления плоских контурных изображений общей задаче распознавания изображений, конечно, требует отдельного тщательного обсуждения, что не входит в цели данной статьи. Но отметим некоторые «генетические» достоинства данного подхода, например, устойчивость к масштабированию, гладкой деформации (растягиванию или сжатию), вращению образцов и анализируемого изображения. Наиболее популярные на сегодняшний день алгоритмы, реализованные на основе нейронных сетей [10], как правило, теряют свою работоспособность при простом масштабировании (равномерном сжатии, увеличении) или вращении изображения. Разработчики решают данную проблему, предлагая конечному пользователю в мануальном режиме повернуть, увеличить (уменьшить) изображение.
2. Следующим этапом развития данного подхода будет ввод относительных размеров дуг и проведения операций сжатия изображений, когда дуги, меньшие определенных размеров, стягиваются в точки, что «убирает» детали и позволяет проводить предварительный анализ для более простых изображений. А результаты анализа сжатых изображений могут быть использованы для анализа исходных изображений.
3. Как уже отмечалось выше, для рассматриваемой предметной области (анализ изображений) общая схема решения комбинаторных задач в логико-эвристическом подходе не используется. Предполагается ее применение при решении задач поиска «партизана и его собаки в лесу» и подобных визуальных головоломок.

Список литературы

1. Кнут Д. Искусство программирования для ЭВМ. Сортировка и поиск / Д. Кнут. – М. : Мир, 1978. – 848 с.

2. BigTable [Электронный ресурс]. — URL: <http://ru.wikipedia.org/wiki/BigTable>.
3. Мартьянов В. И. Логико-эвристические методы сетевого планирования и распознавание ситуаций / В. И. Мартьянов // Труды Междунар. конф. «Проблемы управления и моделирования в сложных системах». — Самара. 2001. — С. 203-215.
4. Обзор приложений логико-эвристических методов решения комбинаторных задач высокой сложности / В. И. Мартьянов, В. В. Архипов, М. Д. Каташевцев, Д. В. Пахомов // Современные технологии. Системный анализ. Моделирование / ИРГУПС. — 2010. — №4(28). — С. 61-67.
5. Лорьер Ж.-Л. Системы искусственного интеллекта / Ж.-Л. Лорьер. — М. : Мир, 1991.
6. Гери М. Вычислительные машины и труднорешаемые задачи / М. Гери, Д. Джонсон. — М. : Мир, 1982.
7. Мальцев А. И. Алгебраические системы / А. И. Мальцев. — М. : Наука, 1967. — 324 с.
8. Кокорин А. И. Вопросы разрешимости расширенных теорий / А. И. Кокорин, А. Г. Пинус // Успехи мат. наук. — 1978. — Т.33, вып.2. — С. 49-84.
9. Беллман Р. Динамическое программирование / Р. Беллман. — М. : Изд-во иностр. лит., 1960.
10. Еремин Д. М. Искусственные нейронные сети в интеллектуальных системах управления / Д. М. Еремин, И. Б. Гарцев. — М. : МИРЭА, 2004. — 75 с.

V. I. MARTYANOV, M. D. KATASHEVTSEV

Combinatorial problems of high complexity and analyse of sketch images

Abstract. The article is dedicated to the problem of recognition of sketch images. The problem is reviewed under logic-heuristic paradigm. The model of binary raster image representation like a system of arcs and relates between arcs is reviewed. The key algorithms of recognition for images representation according to the introduced model, which were found useful at solving scheduling and genetic analyse problems, is described.

Keywords: computer graphics; image recognition; logic-heuristic; raster; algebraic system.

Мартьянов Владимир Иванович, доктор физико-математических наук, профессор кафедры автомобильных дорог, Иркутский государственный технический университет, 664074, Иркутск, Лермонтова, 83 тел.: 8-914-899-15-12 (mmail@mail.ru)

Каташевцев Михаил Дмитриевич, аспирант НИ ИРГТУ Иркутский государственный технический университет, 664074, Иркутск, Лермонтова, 83 тел.: 8-914-923-24-63 (vlad.martyanov8@yandex.ru)

Martyanov Vladimir, Irkutsk State Technical University, 83, Lermontova St., Irkutsk, 664074 doctor of physical and mathematical science, professor Phone: 8-914-899-15-12 (vlad.martyanov8@yandex.ru)

Katashevtsev Mikhail, Irkutsk State Technical University, 83, Lermontova St., Irkutsk, 664074 postgraduate student, Phone: 8-914-923-24-63 (mmail@mail.ru)