

ОБ ОДНОМ МЕТОДЕ ПОИСКА ПО ШАБЛОНУ НА ТОПОГРАФИЧЕСКИХ КАРТАХ

Е. Е. ИВАНКО

Институт математики и механики УрО РАН, Екатеринбург, Россия

e-mail: ivanko@ural.ru

An original method of the pattern matching is discussed. Suggested method is based on the comparison of the particular weighted graphs, called as semantic networks. This work continues the author's series of publications on information processing with the help of semantic networks.

Введение

Всякий алгоритм поиска по шаблону нацелен на решение определенного класса задач, обусловленного спецификой исследуемых изображений [1]. Предлагаемая методика ориентирована на работу с бинарными контурными изображениями. Данная статья освещает случай изображений, имеющих высокую плотность линий на единицу площади. Наиболее эффективен метод на изображениях, средняя толщина линий которых равняется одному пикселю. Среди типичных классов подобных изображений, встречающихся на практике, можно выделить, например, бинаризованные топографические карты или дактилоскопические узоры. Некоторые результаты по построению и обработке таких изображений можно найти в [2, 3].

Предлагаемый алгоритм показал высокую скорость работы и устойчивость к искажениям входного изображения. Описанный в статье подход сочетает в себе черты классического сопоставления участков карты с образцом и анализа статистических характеристик изображения.

1. Метод

Метод состоит из двух последовательных шагов.

1. Формирование взвешенного графа по шаблонному изображению. Вершинами такого графа служат небольшие части изображения — знаки. Число случаев, в которых два знака оказались в одной окрестности наперед заданного размера, определит вес ребра между знаками. Построенный указанным способом граф назовем *семантической сетью (SN)* [4–8]. Каждый знак в сети выражен через некоторые другие знаки характерным для шаблона образом. В такой семантической сети отражены локальные особенности формы линий и характерное взаимное расположение этих особенностей на шаблоне. Структура,

основанная на подобной идеологии, была предложена в 1973 году Хараликом для анализа текстур [9].

2. Последовательный перебор знаков основного изображения. Для каждого знака будем сравнивать список его соседей в окрестности на основном изображении со списком соседей этого знака в семантической сети шаблона. Если списки соседей окажутся близки в некотором смысле, определенном ниже, значит, текущий знак, возможно, находится на месте искомого шаблона в исследуемом изображении и должен быть отмечен. Высокая плотность отмеченных знаков на изображении покажет предположительное местонахождение искомого шаблона.

Входными параметрами для метода являются $n \times n$ — размер знака и m — размер окрестности. В случае, когда линейный размер знака n достаточно велик ($n \geq 10$), а величина m -окрестности при этом мала ($m = 0, 1$), метод имеет общие черты с классическим сопоставлением участка изображения с шаблоном. В этом случае большую роль играет сходство изображений знака и шаблона. Такой подход обеспечивает высокую точность, но неустойчив к искажениям изображений. Пусть шаблон имеет размеры $N \times N$. Если размер знака n достаточно мал ($n = 1 \dots 5$), а m -окрестность сравнима по размеру с шаблоном ($m \geq N/10$), то метод схож со статистическим расчетом особенностей шаблона. В таком расчете главную роль играют количественные характеристики некоторых простых параметров изображения, например плотность пикселей одного цвета. Здесь высокая устойчивость к искажениям достигается за счет низкой точности поиска.

В промежуточных случаях в методе сочетаются два процесса: выявление присущих шаблону частей изображения и отражение статистического взаиморасположения этих частей. На практике исследован случай при небольшом размере знака ($n = 2, 3$) и небольшой окрестности ($m = 3 \dots 6$), позволяющий отражать статистические особенности взаимного расположения локальных форм линий на шаблоне.

2. Алгоритм

Определение 1. Изображением $I_{[N,M]}$, где $N > 2$, $M > 2$, называется матрица размера $N \times M$, элементами которой являются числа $\{0, 1\}$.

Определение 2. Пусть задано изображение $I_{[N,M]}$, тогда шаблоном $I_{[N_s, M_s]}^s$ называется произвольная матрица размера $N_s \times M_s$, где $1 < N_s < N$, $1 < M_s < M$, элементами которой являются числа $\{0, 1\}$.

Определение 3. Пусть задано изображение $I_{[N,M]} = (a_{ij})$, $n, x, y \in \mathbb{N}$, $1 < n \leq \min\{N, M\}$, тогда знаком $S_{(x,y)}^n$ на $I_{[N,M]}$, где $x + n - 1 \leq N$, $y + n - 1 \leq M$, называется следующая подматрица размера $(n \times n)$ матрицы $I_{[N,M]}$:

$$\begin{pmatrix} a_{x,y} & \cdots & a_{x+n-1,y} \\ \vdots & \ddots & \vdots \\ a_{x,y+n-1} & \cdots & a_{x+n-1,y+n-1} \end{pmatrix}.$$

Под знаком S^n понимается произвольная матрица размера $(n \times n)$, элементами которой являются числа $\{0, 1\}$. Всего таких знаков $2^{n \times n}$.

Определение 4. Пусть задано изображение $I_{[N,M]} = (b_{ij})$, $n, m, x, y \in \mathbb{N}$, $1 < n, m \leq \min\{N, M\}$, задан знак $S_{(x,y)}^n$ на $I_{[N,M]}$: $1 \leq x \leq N - n + 1$, $1 \leq y \leq M - n + 1$. Пусть

задана матрица

$$K = \begin{pmatrix} a_{x-m,y-m} & \cdots & \cdots & \cdots & \cdots & \cdots & a_{x+n-1+m,y-m} \\ \vdots & \ddots & & & & & \vdots \\ \vdots & & a_{x,y} & \cdots & a_{x+n-1,y} & \cdots & \vdots \\ \vdots & & \vdots & & \vdots & & \vdots \\ \vdots & \cdots & a_{x,y+n-1} & \cdots & a_{x+n-1,y+n-1} & & \vdots \\ \vdots & & & & & \ddots & \vdots \\ a_{x-m,y+n-1+m} & \cdots & \cdots & \cdots & \cdots & \cdots & a_{x+n-1+m,y+n-1+m} \end{pmatrix},$$

где $a_{ij} = b_{ij}$, если $1 \leq i \leq N$, $1 \leq j \leq M$; $a_{ij} = 0$, иначе. Все знаки $S_{(x',y')}^n$ на K , где $(x-x')(y-y') \neq 0$ и $1 \leq x' \leq N-n+1$, $1 \leq y' \leq M-n+1$, считаются принадлежащими выколотой m -окрестности $\check{O}_m(S_{(x,y)}^n)$ знака $S_{(x,y)}^n$ на изображении $I_{[N,M]}$.

Определение 5. Семантической сетью **SN** изображения $I_{[N,M]}$ называется взвешенный граф, вершинами которого являются знаки размера $(n \times n)$; функция весов на ребрах $w(S_i^n, S_j^n) \rightarrow \mathbb{N} \cup \{0\} \forall i, j$.

Алгоритм 1. Построение семантической сети по шаблонному изображению $I_{[N_s, M_s]}^s$.

1. Определить размер знака n , размер окрестности m ($n, m \in \mathbb{N}$, $1 < n$, $m \leq \min\{N, M\}$). Вершинами семантической сети будут считаться всевозможные знаки S^n размера $(n \times n)$, функция весов изначально всюду нулевая: $w(S_i, S_j) = 0 \forall i, j$.

2. Пусть на вход подано шаблонное изображение $I_{[N_s, M_s]}^s$.

3. Перебирать последовательно всевозможные значения $x, y \in \mathbb{N}$: $1 \leq x \leq N_s - n + 1$, $1 \leq y \leq M_s - n + 1$.

4. Для каждой позиции (x, y) :

а) прочесть знак $S_{(x,y)}^n$ на $I_{[N_s, M_s]}^s$;

б) для всякого знака $S_{(x',y')}^n \in \check{O}_m(S_{(x,y)}^n)$ функция весов увеличивается на 1:

$$w(S_{(x,y)}^n, S_{(x',y')}^n) = w(S_{(x,y)}^n, S_{(x',y')}^n) + 1.$$

Если представить семантическую сеть **SN** матрицей инцидентности, то каждому знаку S_i^n будет соответствовать вектор длины $2^{n \times n}$:

$$\vec{S}_i^n = (l_{S_1^n}, l_{S_2^n}, \dots, l_{S_{2^{n \times n}}^n}), \quad (1)$$

где $l_{S_j^n}$ — вес ребра (S_i^n, S_j^n) в семантической сети **SN**.

Аналогичный вектор может быть построен по выколотой m -окрестности любого вхождения знака S_i^n в исследуемое изображение:

$$\vec{S}_i^{n'} = (h_{S_1^n}, h_{S_2^n}, \dots, h_{S_{2^{n \times n}}^n}), \quad (2)$$

где $h_{S_j^n}$ — количество вхождений знака S_j^n в данную окрестность.

Близость векторов \vec{S}_i^n и $\vec{S}_i^{n'}$ покажет, насколько список соседей данного знака, полученный по окрестности в исследуемом изображении, похож на список соседей этого знака в семантической сети шаблона. Высокая степень близости векторов позволит считать текущий знак находящимся в области исследуемого изображения, похожей на шаблон.

В общем случае величина шаблона может существенно отличаться от величины m -окрестности, поэтому, возможно, длины \vec{S}_i^n и $\vec{S}_i^{n'}$ также существенно различны. В такой ситуации в качестве меры близости имеет смысл использовать величину, отражающую согласованность направлений векторов и не зависящую от их длин.

Подходящей мерой является, например:

$$\cos(\widehat{\vec{S}_i^n \vec{S}_i^{n'}}) = \frac{(\vec{S}_i^n, \vec{S}_i^{n'})}{|\vec{S}_i^n| |\vec{S}_i^{n'}|}. \quad (3)$$

Таким образом, знак на изображении будет отмечен, если $\cos(\widehat{\vec{S}_i^n \vec{S}_i^{n'}})$ не будет меньше некоторого наперед выбранного порога P . Это ограничение позволит отбирать лишь те векторы $\vec{S}_i^{n'}$, направление которых близко к направлению \vec{S}_i^n .

Предложение. Значение порога P определяется следующим образом.

1. Пусть по шаблонному изображению $I_{[N_s, M_s]}^s$ построена семантическая сеть **SN**.
2. Перебирая последовательно всевозможные значения x, y на $I_{[N, M]}$, найдем $M_c = \max(\cos(\widehat{\vec{S}_{(x,y)}^n \vec{S}_{(x,y)}^{n'}}))$ и $m_c = \min(\cos(\widehat{\vec{S}_{(x,y)}^n \vec{S}_{(x,y)}^{n'}}))$, где $x, y \in \mathbb{N}: 1 \leq x \leq N - n + 1, 1 \leq y \leq M - n + 1$. Учитывая, что по построению все координаты $\vec{S}_{(x,y)}^n$ и $\vec{S}_{(x,y)}^{n'}$ неотрицательны, имеем $M_c, m_c \in [0, 1]$.

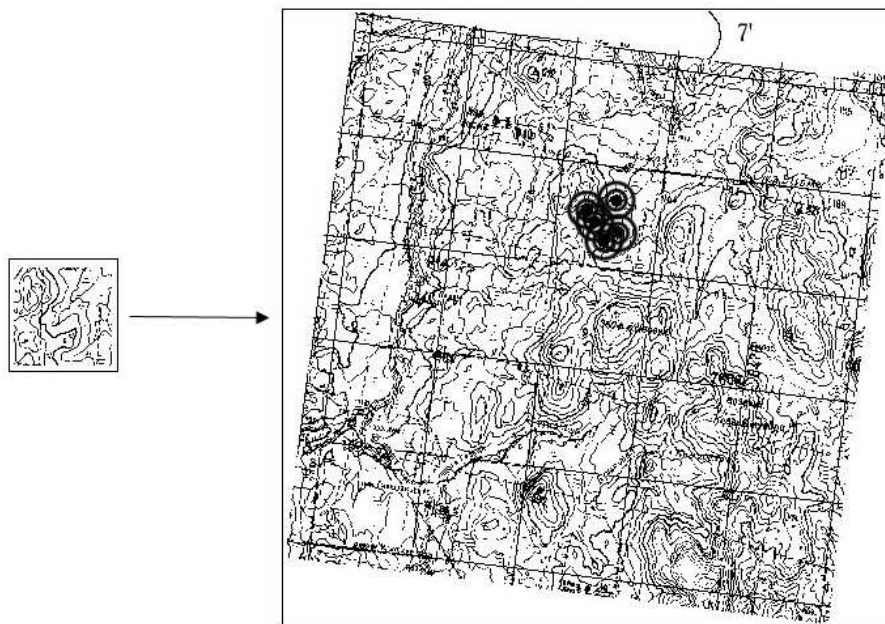
3. Пусть $S_0 = N_s \times M_s$ — площадь шаблона, $S = N \times M$ — площадь исследуемого изображения, тогда S_0/S — доля знаков шаблона от общего числа знаков изображения. Теоретически каждый из знаков на месте шаблона даст близкое к единице значение косинуса в формуле (3), а знак вне шаблона даст несколько меньшее значение. Опираясь на это соображение, будем считать, что знак находится на шаблоне, если выражение для косинуса (3) на нем принимает значение не менее

$$P = M_c - \frac{S_0}{S}(M_c - m_c). \quad (4)$$

На практике для бинаризованных топографических карт часто $M_c = 1, m_c = 0$.

Алгоритм 2. Поиск шаблона $I_{[N_s, M_s]}^s$ в исследуемом изображении $I_{[N, M]}$.

1. Используя *алгоритм 1* на шаблоне $I_{[N_s, M_s]}^s$, построить семантическую сеть **SN**.
2. Выбрать порог $P \in \mathbb{R} \cap (0, 1)$, пользуясь *Предложением*.
3. На исследуемом изображении $I_{[N, M]}$ перебирать последовательно всевозможные значения $x, y: 1 \leq x \leq N - n + 1, 1 \leq y \leq M - n + 1$.
4. Для всякой позиции (x, y) :
 - а) прочесть знак $S_{(x,y)}^n$ на $I_{[N, M]}$;
 - б) прочесть вектор $\vec{S}_{(x,y)}^n$ из **SN**, согласно формуле (1);
 - в) построить вектор $\vec{S}_{(x,y)}^{n'}$ по $\check{O}_m(S_{(x,y)}^n)$ на $I_{[N, M]}$, согласно формуле (2);
 - г) рассчитать $D = \cos(\widehat{\vec{S}_{(x,y)}^n \vec{S}_{(x,y)}^{n'}})$ по формуле (3);
 - д) если $D \geq P$, отметить знак $S_{(x,y)}^n$ на $I_{[N, M]}$.
5. Участки изображения, на которых плотность отмеченных знаков существенно превосходит среднюю плотность знаков, отмеченных на изображении, считаются участками, похожими на шаблон.



Пример поиска по шаблону на повернутой топографической карте. Размер карты 450×450 пикселей, шаблона 75×75 пикселей. Разрешение 72×72 dpi. Поворот выполнен в PhotoShop 6.0.

3. Результаты

Предлагаемый метод опробован на бинаризованных топографических картах. Для тестирования на исходной неискаженной карте выбирался и сохранялся участок — шаблонное изображение. После этого карта подвергалась искажению, на искаженной карте выполняли поиск шаблона.

Тестирование проводилось на трех исходных участках черно-белой топографической карты разрешением 72×72 dpi и размером 450×450 пикселей каждый (см. рисунок). Исходные изображения топографических карт были бинаризованы в PhotoShop 6.0 с помощью метода 50%-го порога. Затем карта масштабировалась так, чтобы линии уровня в среднем оказались не более одного пикселя толщиной. Для тестирования на каждом из трех участков карты выбирались по три различных шаблона. Таким образом, для каждого из описанных ниже видов искажений проведена серия из девяти экспериментов по поиску шаблона на искаженной карте с выбором трех различных исходных участков карт и трех различных шаблонов на каждом участке. При тестировании размер знака и размер окрестности принимались равными 3 ($n = m = 3$).

В таблице отражены ограничения на величину искажений, при которых шаблон устойчиво выделялся на изображении во всех девяти экспериментах (в скобках указаны величины искажений, при которых отношение плотности отмеченных знаков на месте шаблона в изображении к средней плотности отмеченных знаков недостаточно велико для очевидного визуального выделения шаблона; для выявления шаблона в этом случае следует использовать несложный анализ плотности меток на карте).

Пропорциональные по осям растяжение и сжатие (масштабирование), а также пово-

Растяжение	Сжатие	Поворот CW/CCW	Шум	Потери
10(15) %	10(13) %	6(7)°	8(8) %	10(15) %

роты изображения производились в PhotoShop 6.0. Шум и потери указаны в процентах случайным образом измененных пикселей черного или белого цвета соответственно от общего числа пикселей на изображении.

4. Некоторые особенности чувствительности предложенного метода к искажениям

Предложенный алгоритм нечувствителен к параллельным переносам и “легким” искажениям изучаемого изображения. Под “легкими” здесь понимаются в первую очередь искажения, приведенные в таблице разд. 3. В более общем случае такими искажениями можно считать произвольные искажения, при которых плотность отмеченных знаков на месте шаблона в изображении превышает плотность отмеченных знаков на всем изображении хотя бы в два раза. Интуитивно эти искажения слабо влияют на форму и взаимное расположение знаков.

В практических задачах встречается ситуация, когда шаблон ищется на изображении, склеенном из множества подызображений с небольшими относительными смещениями этих подызображений друг относительно друга.

В качестве примера возникновения таких входных изображений можно предложить ситуацию, в которой робот, перемещаясь, фотографирует окружающее пространство по частям, а затем ищет на общей фотографии предмет по шаблону. В силу ряда физических и конструктивных причин робот не всегда способен собрать целое изображение без относительных смещений его частей.

В случае, когда размер окрестности m и размер знака n существенно меньше размеров подызображений, на которые разбито изучаемое изображение, предлагаемый метод практически нечувствителен к описанной деформации.

Можно предположить, что метод обладает большим потенциалом для специализированной работы с конкретными искажениями. Такая настройка достигается за счет группировки знаков в классы. Знаки одного класса получают из некоего формирующего класса знака путем применения к нему заданного искажения (например, поворота на различные углы). Существенное сокращение алфавита знаков в этом случае можно компенсировать увеличением размера знака n . Изложенная гипотеза находится в стадии исследования.

5. Время работы алгоритма

Пусть n — размер знака, m — размер окрестности, N_s , M_s — ширина и высота шаблонного изображения соответственно, а N , M — ширина и высота исследуемого изображения. Алгоритм содержит три операции с плавающей точкой на $n^2(2m+1)^2$ операций ввода-вывода и порядка 10^1 целочисленных операций на одну операцию ввода-вывода. Принимая во внимание указанные соотношения и тот факт, что операция ввода-вывода обычно занимает существенно больше времени, чем арифметическая операция, в настоящей оценке предлагается пренебречь арифметическими операциями и измерять время работы алгоритма количеством операций ввода-вывода.

Предложенный алгоритм состоит из двух последовательных частей: сканирование шаблонного изображения с целью построения шаблонной семантической сети и сканирование основного изображения для выявления аналогично “определенных” знаков.

Согласно алгоритму 1, для количества итераций T_1 при построении семантической сети по шаблонному изображению $I_{[N_s, M_s]}^s$ справедлива следующая оценка:

$$T_1 \leq n^2(2m+1)^2(N_s - n + 1)(M_s - n + 1),$$

где n^2 — количество операций ввода-вывода, требующихся для считывания знака; $(2m+1)^2$ — количество знаков в m -окрестности; $(N_s - n + 1)(M_s - n + 1)$ — число m -окрестностей на шаблоне.

Количество итераций T_2 в ходе работы алгоритма 2 на изображении $I_{[N, M]}$ можно оценить аналогичным образом:

$$T_2 \leq n^2(2m+1)^2(N - n + 1)(M - n + 1).$$

Оценки приведены для двух последовательных частей алгоритма. Таким образом, суммируя T_1 и T_2 , можно оценить общее число операций ввода-вывода T :

$$T = T_1 + T_2 \leq n^2(2m+1)^2((N - n + 1)(M - n + 1) + (N_s - n + 1)(M_s - n + 1)). \quad (5)$$

Подобный подход к оценке времени работы алгоритма подтвержден следующим вычислительным экспериментом: на персональном компьютере Pentium 4 1.7 ГГц, RAM 512 Мбайт, Video NVIDIA GeForce2 MX400 32 Мбайт проводился поиск шаблона размером 75×75 пикселей на изображении размера 450×450 пикселей. Размер знака $n = 3$, размер окрестности $m = 3$ (изображения соответствуют рисунку, приведенному в разд. 3). Время работы алгоритма составило 125 с. Согласно оценке (5) количество операций ввода-вывода составило порядка 10^8 . Для определения корректности приведенной в данном разделе оценки было замерено время работы тестовой программы, осуществившей только 10^8 аналогичных операций ввода-вывода. Время работы этой программы составило 120 с. Полученный результат показывает целесообразность выбора одной операции ввода-вывода как единицы измерения времени работы алгоритма.

Заключение

В настоящей статье рассмотрен способ поиска по шаблону, опирающийся на сопоставление между собой особых информационных структур, названных семантическими сетями. В данной работе такую информационную структуру удобно рассматривать как взвешенный граф. Однако идеологическая особенность, лежащая в основе предлагаемого метода, состоит в выражении, определении знака через другие знаки, что присуще именно семантическим сетям.

Ранее автором опубликован ряд работ [10–12], посвященных классификации контурных бинаризованных изображений, трехмерных воксельных объектов и звукозаписей слов русского языка. Основой всех работ, включая настоящую статью, является переход от сопоставления собственно информационных образов к сопоставлению графов, специальным образом построенных из частей информационных образов.

Единый подход к обработке информации позволил получить хорошие результаты в области классификации информационных образов различных типов. Можно предположить, что и описанный метод также окажется эффективным для поиска по шаблону в воксельных объектах и звукозаписях. Автором ведутся исследования в этом направлении.

Актуальным применением предложенной технологии может оказаться поиск по базе данных топографических карт или дактилоскопических узоров.

Список литературы

- [1] ФОРСАЙТ Д. А., ПОНС Ж. Компьютерное зрение. Современный подход: Пер. с англ. М.: Изд. дом "Вильямс", 2004.
- [2] YAMADA H., YAMAMOTO K., HOSOKAWA K. Directional mathematical morphology & reformalized Hough Transformation for the analysis of topographic maps // IEEE Trans. on PAMI. 1993. Vol. PAMI-15, N 4. P. 380–387.
- [3] JANG K.S., YI J., JUNG J.Y., KIM J., CHANG K.H. A recognition of map using the geometric relations between lines and the structural informations of objects // Intern. Conf. on Image Proc. (ICIP'97). 1997. 3-Vol. Set-Vol. 3. P. 150–154.
- [4] ПОСПЕЛОВ Д.А. Логико-лингвистические модели в системах управления. М.: Энергоиздат, 1981.
- [5] ПОСПЕЛОВ Д.А. Моделирование рассуждений. Опыт анализа мыслительных актов. М.: Радио и связь, 1989.
- [6] РУВАШКИН В.Ш. Представление и анализ смысла в интеллектуальных информационных системах. М.: Наука, 1989.
- [7] ЗАГОРУЛЬКО Ю.А. Методы представления и обработки знаний: Семантические сети и системы продукций: Метод. пособие. Новосибирск: Изд-во НГУ, 1996.
- [8] SOWA J.H. Semantic Networks. <http://www.jfsowa.com/pubs/semnet.htm>, 2004.
- [9] HARALICK R.M., SHANMUGAM K., DINSTEN I. Textural features for image classification // IEEE Trans. on SMC. 1973. SMC-3(6). P. 610–621.
- [10] ИВАНКО Е.Е., ПЕРЕВАЛОВ Д.С. Об одном знаковом методе обработки информации и его применении для анализа текстов и контурных изображений // Проблемы теоретической и прикладной математики: Тр. 35-й регион. молодеж. конф. Екатеринбург, УрО РАН, 2004. С. 269–273.
- [11] ИВАНКО Е.Е., ПЕРЕВАЛОВ Д.С. Знаковый метод обработки информации и его применение к анализу текстов, классификации изображений и звукозаписей // Наука и будущее: идеи, которые изменяют мир: Матер. междунар. конф. (тез. докл.). М., 2004. С. 78–80.
- [12] IVANKO E., PEREVALOV D. On using sign method for 3d images recognition and classification // Intern. Conf. on Computing, Communications and Control Technologies: CCCT'04, Austin, Texas USA. 2004. Vol. V. P. 248.

*Поступила в редакцию 31 августа 2004 г.,
в переработанном виде — 15 ноября 2004 г.*