

Параллельное Программирование Алгоритма Идентификации Дактилоскопических Изображений

Гудков Владимир Юльевич, Челябинский государственный университет, Миасс, Россия
diana@sonda.ru

Лепихова Дарья Николаевна, Южно-Уральский государственный университет, Челябинск, Россия
daria.lepikhova@yandex.ru

Аннотация

В статье рассматривается алгоритм идентификации дактилоскопических изображений на основе модели топологических векторов. Сравнение выполняется путем оценки степени подобия векторов как отдельно по топологии и по геометрии, так и совместно. Предлагается вариант параллельной реализации алгоритма сравнения, основанный на технологии OpenMP. Приводятся результаты вычислительных экспериментов.

Ключевые слова: отпечаток пальца, сравнение шаблонов, параллельный алгоритм, OpenMP.

1. ВВЕДЕНИЕ

Идентификация личности по отпечаткам пальцев является наиболее распространенным методом биометрической идентификации [2]. В компьютеризированных системах такую идентификацию обычно выполняют по шаблону, вычисляемому при обработке дактилоскопических изображений (ДИ) [3]. Современные автоматизированные дактилоскопические системы насчитывают до сотни миллионов изображений отпечатков пальцев и их шаблонов, а идентификация ДИ типа 1:1 требует такого же числа сравнений различных пар шаблонов [2]. Поэтому фактор времени является решающим для обеспечения конкурентоспособности больших компьютеризированных систем [5]. В связи с этим задача сокращения времени работы алгоритма сравнения шаблонов ДИ актуальна.

Одним из способов повышения реактивности системы является параллельное программирование базы данных, при котором задача сравнения шаблонов остается однопоточной. Такой способ решения актуальной проблемы не требует внесения изменений в алгоритм сравнения шаблонов ДИ. При всей привлекательности, однако, такой подход не позволяет ускорить задачи верификации ДИ типа 1:1.

Задача заключается в выделении и адаптации функций и данных алгоритма сравнения шаблонов ДИ с целью его параллельного программирования.

2. СРАВНЕНИЕ ШАБЛОНОВ ДИ

Оценку степени сходства двух ДИ выполняют методом сравнения их шаблонов. Под шаблоном понимают математическую модель, которая содержит описание признаков изображения. В статье рассматривается представление ДИ на основе топологических векторов [4]. Тогда математическую модель представляют в виде

$$T = \{(m_i, V_i)\}, m_i = \{x_i, y_i, \theta_i, t_i\}, V_i = \{e_l, n_l\}, i \in I = 1..n, \quad (1)$$

где m_i – частный признак, а V_i – топологический вектор, соответствующий m_i ; (x_i, y_i) , θ_i и t_i – координаты, вектор направления как угол и тип частного признака, $t_i \in \{0; 1\}$ (окончание и разветвление); e_l – событие, сформированное частным признаком номер n_l на связи номер l в векторе и

$l \in 0..w$; w – число связей в векторе (обычно до 35); n – число частных признаков изображения [3, 4].

При сравнении шаблонов T_1 и T_2 рассчитывают оценку подобия топологических векторов $V_i \in T_1$ и $V_j \in T_2$ и формируют упорядоченный набор этих оценок в виде

$$S = \{S_p(V_i, V_j) | V_i \in T_1, V_j \in T_2\} \quad (2)$$

где p – номер элемента ряда, причем меньшим номерам соответствуют лучшие оценки [3].

В работе оценки подобия по (2) представляют композицией оценок векторов по топологии и по геометрии.

Рассмотрим два частных признака m_i и m_j и соответствующие им топологические векторы V_i и V_j с числом связей, равным w . Степень подобия векторов по топологии вычисляют на основе функции $\lambda(e_l^i, e_l^j)$, где $e_l^i \in V_i$ и $e_l^j \in V_j$ – события, располагаемые на одноименных связях в двух векторах, а l – номер связи, в виде

$$\Lambda(V_i, V_j) = \frac{\sum_{l=0}^w \lambda(e_l^i, e_l^j)}{w}. \quad (3)$$

Аналогично степень подобия векторов по геометрии вычисляют по формуле

$$\Psi(V_i, V_j) = \frac{\sum_{l=0}^w \psi(\omega(m_i, m_k), \omega(m_j, m_g))}{w}, \quad (4)$$

где w – число связей в векторе; ψ – функция расчета степени подобия двух пар частных признаков; ω – функция расчета разностей геометрических характеристик для двух частных признаков; $m_i, m_k \in T_1$ и $m_j, m_g \in T_2$, задаваемых их номерами i, k и j, g на одноименных связях, причем номера частных признаков k, g определяют по номеру связи l : $k = n_l^i$ в векторе V_i и $g = n_l^j$ в векторе V_j .

В целом оценку подобия двух топологических векторов в (2) оценивают на основе выражений (3) и (4) в виде полинома

$$S_p(V_i, V_j) = \sum_{l=1}^w (c_l \Lambda^{a_l}(V_i, V_j) \Psi^{b_l}(V_i, V_j)) \quad (5)$$

где c_l – коэффициенты; a_l и b_l – значения степеней, w – число связей. Здесь для простоты изложения векторы V_i и V_j предполагают равномошными. В частном случае оценка аналогична вероятности произведения или суммы событий. Собственно оценку подобия двух шаблонов T_1 и T_2 вычисляют в виде

$$S(T_1, T_2) = \frac{\sum_{p=1}^n S_p}{\min(n_1, n_2)} \quad (6)$$

где n – число частных признаков в шаблонах T_1 и T_2 , которое при изложении сути работы предполагают равным.

3. ПАРАЛЛЕЛЬНОЕ ПРОГРАММИРОВАНИЕ

Алгоритм вычисления степени подобия шаблонов ДИ состоит из следующих основных этапов:

- выбор одноименных связей;
- оценка степени подобия векторов по геометрии;
- оценка степени подобия векторов по топологии;
- комплексная оценка степени подобия векторов;
- оценка степени подобия шаблонов.

На рис. 1 представлена общая схема оценки степени подобия векторов по геометрии, а на рис. 2 – по топологии.

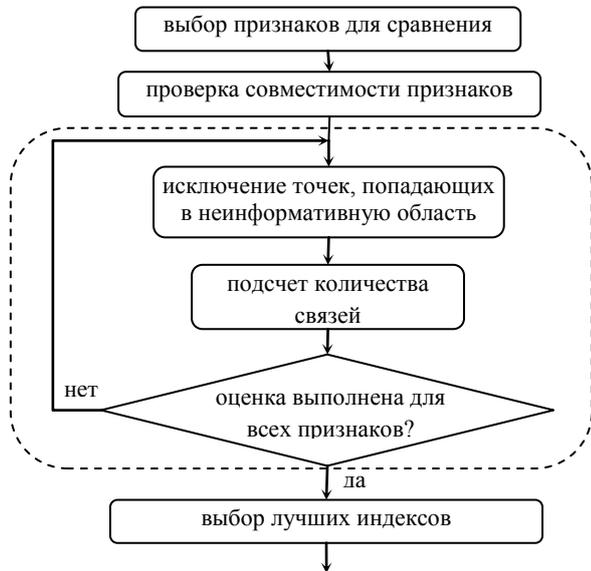


Рис. 1: Алгоритм оценки векторов по геометрии

На этапах, выделенных пунктиром, выполняется параллельная обработка данных, содержащих частные признаки изображения и связи между ними, причем каждый шаг вычислений выполняется независимо от других. Параллельные вычисления обеспечиваются технологией OpenMP [6].

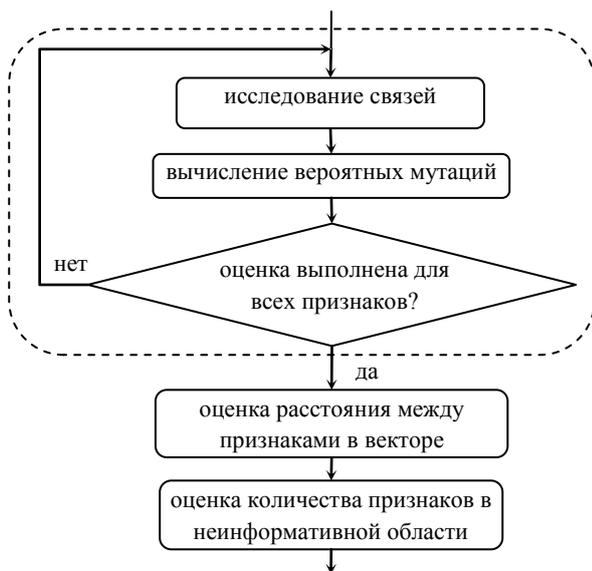


Рис. 2: Алгоритм оценки векторов по топологии

При оценке степени подобия по геометрии и по топологии применяется директива `omp for`. Каждый из параллельных процессов работает с собственным набором данных.

Проверка эффективности параллельных вычислений по серии вычислительных экспериментов показана на рис. 3.

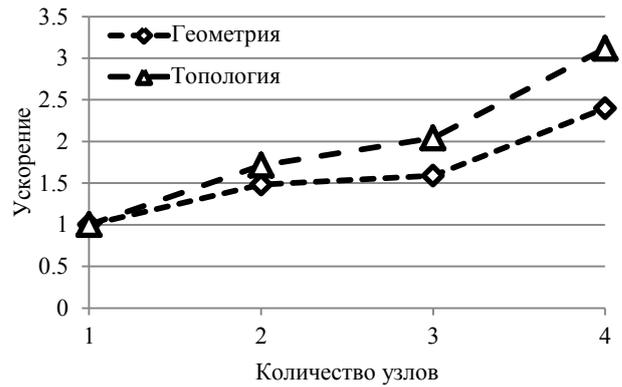


Рис. 3: Графики производительности алгоритмов

Общая схема параллельного алгоритма, который применялся в вычислительных экспериментах, представлена на рис. 4.

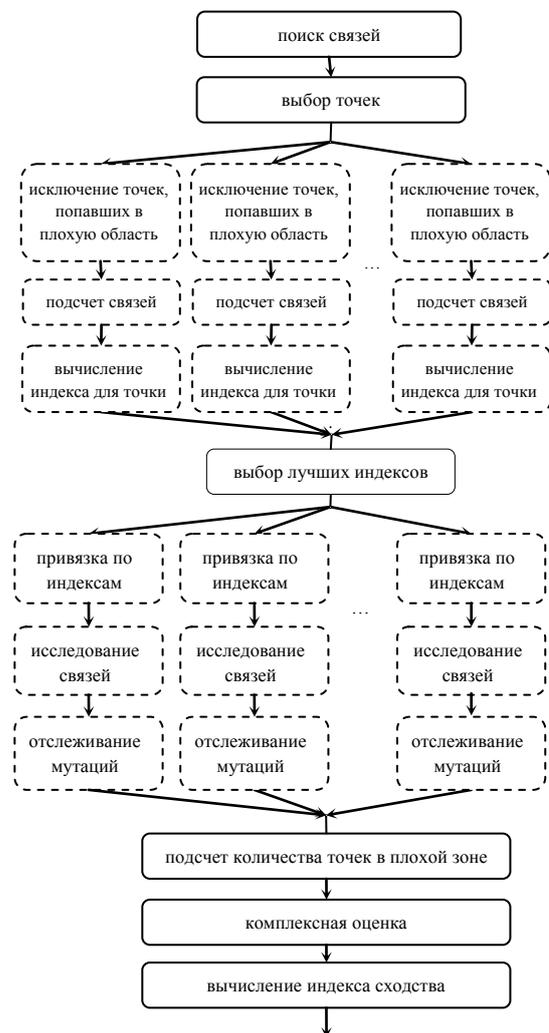


Рис. 4: Общая схема параллельного алгоритма

На графиках видно, что использование OpenMP позволяет сократить время оценки по геометрии в 1,8 раза, а по топологии – в 1,5 раза при организации 2 параллельных процессов. Здесь пунктиром отмечены параллельные секции, в которых возможно возникновение узких мест при оценке степени подобия. Универсальное решение для преодоления таких узких мест неизвестно [1]. Однако программирование позволяет выполнить следующую оптимизацию алгоритма.

4. МОДИФИКАЦИЯ АЛГОРИТМА

Общая оценка по топологии выполняется на основе вычисленных оценок подобия по топологии для каждой пары событий e_i^i и e_i^j . После оценки по топологии выполняется привязка по индексам точек группы. Чтобы избежать образования узких мест применяют:

- организацию параллельной обработки на узком участке при помощи технологии OpenMP;
- привязку по индексам совмещают с оценкой по топологии.

На рис. 5 показан псевдокод модифицированного алгоритма. Такая реализация алгоритма хотя и приводит к дублированию данных, однако значительно уменьшает время оценки.

```
#pragma omp parallel
{
  #pragma omp for
  for (i=0; i<n; i++)
    оценка_по_геометрии();
  #pragma omp single nowait
  выбор_индексов();
  #pragma omp for
  for (j=0; j<m; j++)
    #pragma omp single
      привязка_по_индексам();
  оценка_по_топологии();
}
```

Рис. 5: Алгоритм параллельных вычислений

Модифицированный алгоритм показывает лучшую производительность, но из-за неравномерного распределения циклов и нециклических разделов кода между потоками наблюдается и снижение производительности. Для преодоления указанных недостатков необходимо использовать основное преимущество OpenMP: единое адресное пространство для всех потоков, задействованных в одном процессе.

5. ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ

На рис. 6 приведен график зависимости среднего времени оценки подобия по топологии от числа параллельных процессов до преобразования алгоритма и после. Видно, что после преобразования алгоритм демонстрирует выигрыш по времени. Обе версии показывают некоторое снижение производительности при запуске 3 параллельных процессов, что объясняется необходимостью обмена приватными данными. Оптимальным является запуск алгоритма на $2n$ процессах, где n – количество ядер процессора.

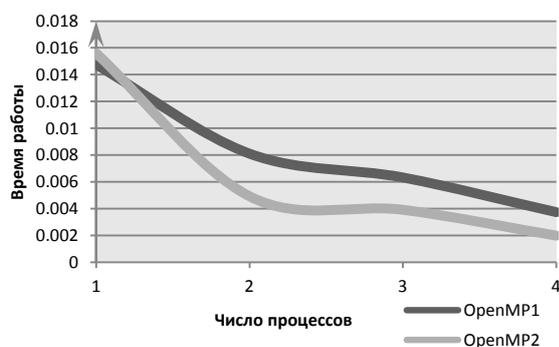


Рис. 6: Сравнение времени оценки по топологии

6. ЗАКЛЮЧЕНИЕ.

В работе рассматривается вариант ускорения алгоритма сравнения дактилоскопических изображений за счет организации параллельных вычислений на этапах выполнения оценок степени подобия векторов по топологии и по геометрии. Также предложен вариант оптимизации узких мест алгоритма, основанный на модификации параллельной версии этапа оценки по топологии. Рассмотренный вариант предполагает преобразование алгоритма таким образом, что привязка по индексам выполняется в том же цикле, что и оценка степени подобия векторов по топологии. При помощи директив OpenMP выполняется равномерное распределение нагрузки между параллельными потоками.

Предложенная реализация позволяет снизить время выполнения оценок по топологии и по геометрии в среднем в 1,8 раза при вычислении на двух потоках и в 3,5 раза на 4 потоках.

В данной работе технология OpenMP является базисом для построения параллельных вычислений. В качестве направления дальнейших исследований может быть выбрана реализация нового уровня параллелизма на основе графического процессора.

7. ЛИТЕРАТУРА

- [1] Андреев Н.Е. *Методы автоматизированного анализа производительности параллельных программ* // Вестник Новосибирского государственного университета. – 2009. – Т. 7, вып. 1. – С. 16-25.
- [2] Арутюнян А.Р. *Методы и алгоритмы анализа и синтеза деформаций дактилоскопических изображений: автореферат дис. ... канд. техн. наук* / М: Цифровичок, 2010. – 18 с.
- [3] Гордеева Д.Н., Гудков, В.Ю. *Распознавание дактилоскопических изображений* // Вестник МГТУ. Серия «Приборостроение». Специальный выпуск «Биотехнологии». – 2011. – С. 47–58.
- [4] А.С. Боков, В.Ю. Гудков. *Пат. 2185660 Российская Федерация, МПК G06K 9/52. Способ кодирования отпечатка папиллярного узора* – № 2000118065/09; заявл. 07.07.2000; опубл. 20.07.2002; Бюл. № 20. – 13 с.
- [5] Урмаев О.С. *Проблемы распараллеливания биометрических вычислений в крупномасштабных информационных системах* // Информатика и ее применения, т.3, вып. 1. – 2009. – С. 8-18.
- [6] OpenMP Tutorial. URL: <https://computing.llnl.gov/tutorials/openMP/>

Об авторах

Гудков Владимир Юльевич – доцент кафедры «Прикладной математики» Миасского филиала ЧелГУ
e-mail: diana@sonda.ru

Область научных интересов: обработка изображений, распознавание образов, дактилоскопические технологии.

Лепихова Дарья Николаевна – инженер-программист ООО «Сонда-Технологии».
e-mail: daria.lepikhova@yandex.ru

Область научных интересов: идентификация отпечатков пальцев, параллельные вычисления.