

ИНДЕКСАЦИЯ ДАКТИЛОСКОПИЧЕСКИХ ИЗОБРАЖЕНИЙ**В.Ю. Гудков****FINGERPRINT INDEXING****V.U. Gudkov**

В работе предлагается метод расчета индексов дактилоскопического изображения на основе векторов состояния для частных признаков. Расчет одного индекса опирается на параметры векторов состояния, увязанных в цепь. Набор индексов служит ссылками в коллекции изображений.

Ключевые слова: отпечаток пальца, частный признак, вектор состояния, индексация.

This paper presents a method of fingerprint indexing based on the state vectors of minutiae. The parameters of state vectors fitted in chain are used to calculate the indexes. Data base references are the same as indexes.

Keywords: fingerprint, minutiae, state vector, indexing.

Введение

Изображения отпечатков пальцев или ладоней относятся к дактилоскопическим изображениям (ДИ). Они характеризуются частными признаками в виде окончаний и разветвлений линий [1]. Обычная модель частных признаков формируется в виде

$$L_m = \{M_i = \{(x_i, y_i), \alpha_i, t_i\} | i \in 1 \dots n\}, \quad (1)$$

где M_i - частный признак; мощность $|L_m| = n$; i - номер частного признака; (x_i, y_i) , α_i и t_i - координаты, направление и тип M_i . Частные признаки детектируют по скелету изображения [1]. Координаты (x_i, y_i) определяются координатами вершины скелета. Направление α_i , как угол определяется цепью вершин скелета [2]. Тип $t_i \in \{0, 1\}$ отражает валентность вершины скелета [2].

Идентификация ДИ на основе модели (1) неэффективна [1]. Поэтому вводят специальные метрики. Наиболее удачные из них отражают структуру узора, например в виде топологических векторов [3], уменьшающих ошибки распознавания. Однако сложный анализ, который необходимо выполнять при сопоставлении топологических векторов, снижает производительность процедуры идентификации личности. Первая попытка для преодоления этого недостатка сделана в [4] на основе векторов состояния, которые повышают производительность процедуры идентификации и практически не увеличивают ошибки распознавания. Но желаемого эффекта не достигается. На базах данных объемом в несколько миллионов дактилокарт установление личности может выполняться несколько часов.

Ключевым решением обозначенной проблемы является индексация изображений, что актуально в эпоху развития интернета с внушительным количеством фотографий, картин, документов и других источников данных. Индексация ДИ тоже актуальна, поскольку указывает магистральное направление для повышения производительности дактилоскопических систем. Однако в настоящее время неизвестно эффективных алгоритмов индексации ДИ [1].

1. Вектор состояния частного признака

Для каждого частного признака определяют один вектор состояния. Их множество записывают в виде

$$L_c = \{S_i = \{(c_j, n_j)\} | i \in 1 \dots n, j \in 1 \dots m_i\}, \quad (2)$$

где S_i - вектор состояния частного признака M_i ; мощность $|L_c| = |L_m| = n$; i - номер вектора состояния, совпадающий с номером частного признака; j - номер связи в векторе состояния; c_j - состояние, сформированное частным признаком с номером n_j , m_i - количество связей, зависящее от типа частного признака t_i по (1).

Опишем процедуру построения векторов состояния.

В области ясного ДИ выделяют линии и формируют стилизованное изображение в виде скелета [1]. По скелету детектируют окончания и разветвления. Их направления показаны на рис. 1. Направление указывает на область увеличения числа линий и параллельно касательной к скелету линии в малой окрестности частного признака.

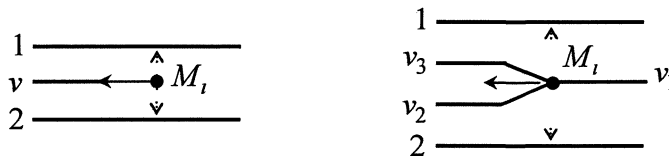


Рис. 1. Направление и две проекции от окончания и разветвления

Каждый частный признак M_i по (1) нумеруют и задают координатами, направлением как углом и типом. Затем от каждого частного признака фиксируют две проекции: вправо и влево перпендикулярно направлению частного признака на смежные скелетные линии номер 1 и 2 (на рис. 1 проекции на вершины скелета изображены пунктиром).

Выберем частный признак M_i и проведем через координаты (x_i, y_i) вправо и влево сечение на глубину нескольких линий m перпендикулярно касательным к пересекаемым линиям. Пронумеруем по спирали, разворачивающейся по часовой стрелке, рассеченные линии, которые назовем связями. Сечение проходит по кривой линии, параллельно направлению кривизны линий. Глубина сечения m варьируется: $m \in 1...8$ вправо и также влево. Одна линия в сечении образует две связи. Центральная линия - одну или три связи по формуле

$$m_t = \begin{cases} 4m + 1, & \text{если } t_i - \text{окончание,} \\ 4m + 3, & \text{если } t_i - \text{разветвление.} \end{cases} \quad (3)$$

Вектор состояния определяют по сечению, изображенному на рис. 2 или рис. 3 пунктиром. Для этого прослеживают ход каждой связи от сечения до встречи на связи с другим частным признаком или с проекцией от него. На связях детектируют состояния:

1 - на связи частный признак или проекция от него, направление частного признака ориентировано навстречу ходу связи;

0 - на связи частный признак или проекция от

него, направление частного признака ориентировано по ходу связи.

С состоянием, детектированным на связи, ассоциируют номер частного признака, инициирующего это состояние. Так формируется вектор состояния, показанный в виде табл. 1 для окончания и табл. 2 для разветвления [4]. Здесь символ $_$ - код пропуска.

На рис. 2 в сечении для окончания 19 пронумерованы связи 0-16 и частные признаки 20-28; вектор состояния представлен в табл. 1. На рис. 3 в сечении для разветвления 19 пронумерованы связи 0-18 и частные признаки 20-27; вектор состояния представлен в табл. 2. Код пропуска указывает на обрыв связи и запоминается в маске вектора состояния. Каждый бит состояния располагается в позиции, определяемой номером связи. При мутациях частных признаков вектор состояния меняется, но измененные номера связей рассчитываются [3].

Определение 1. Мутацией первого рода называется изменение типа частного признака без его перехода через линию.

Определение 2. Мутацией второго рода называется переход частного признака через линию с сохранением его типа.

Пусть S'_i и S''_i - два вектора состояний для частных признаков M'_i и M''_i с различных образцов ДИ. Быструю и грубую проверку степени подобия M'_i и M''_i выполняют сравнением S'_i и S''_i операцией ИСКЛЮЧАЮЩЕЕ ИЛИ и подсчетом числа совпадающих бит с учетом масок векторов

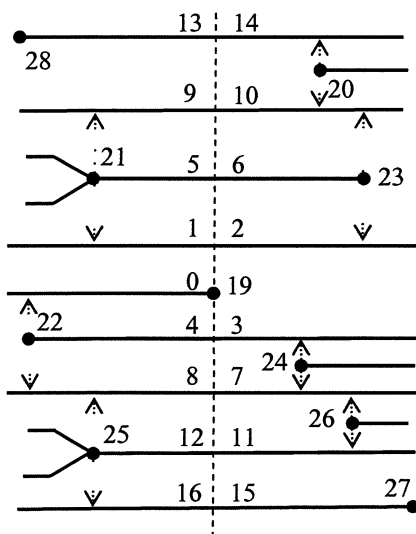


Рис. 2. Сечение для окончания

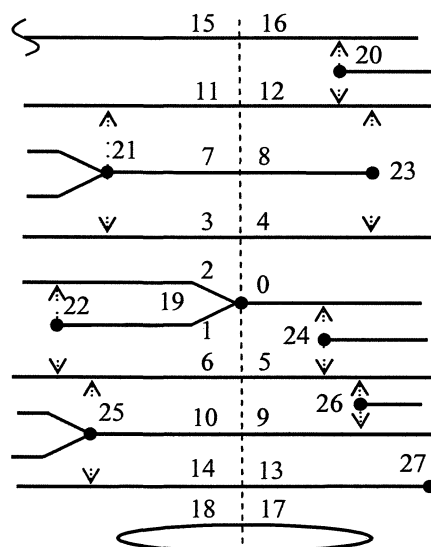


Рис. 3. Сечение для разветвления

Таблица 1
Вектор состояния для окончания

| Номер связи | Индекс | Состояние |
|-------------|--------|-----------|
| 0 | 22 | 1 |
| 1 | 21 | 0 |
| 2 | 23 | 1 |
| 3 | 24 | 0 |
| 4 | 22 | 1 |
| 5 | 21 | 0 |
| 6 | 23 | 1 |
| 7 | 24 | 0 |
| 8 | 25 | 0 |
| 9 | 21 | 0 |
| 10 | 20 | 0 |
| 11 | 26 | 0 |
| 12 | 25 | 0 |
| 13 | 28 | 1 |
| 14 | 20 | 0 |
| 15 | 27 | 1 |
| 16 | 25 | 0 |

Таблица 2
Вектор состояния для разветвления

| Номер связи | Индекс | Состояние |
|-------------|--------|-----------|
| 0 | 24 | 0 |
| 1 | 22 | 1 |
| 2 | 22 | 1 |
| 3 | 21 | 0 |
| 4 | 23 | 1 |
| 5 | 24 | 0 |
| 6 | 25 | 0 |
| 7 | 21 | 0 |
| 8 | 23 | 1 |
| 9 | 26 | 0 |
| 10 | 25 | 0 |
| 11 | 21 | 0 |
| 12 | 20 | 0 |
| 13 | 27 | 1 |
| 14 | 25 | 0 |
| 16 | 20 | 0 |
| 15, 17, 18 | — | — |

состояний (биты, отмеченные кодом пропуска, не сравнивают). Это число коррелирует со степенью подобия сравниваемых частных признаков. Данная операция позволяет ускорить процедуру идентификации ДИ в десятки раз [4].

Теорема 1. Вектор состояния устойчив по Ляпунову.

Доказательство. Выберем частный признак M_i . Для одного M_i рассчитывается один вектор состояния S_i . При мутации первого рода связи вектора S_i перенумеровываются. Поскольку мутация может быть правой или левой, для вектора S_i рассчитываются S_i^r и S_i^l . Однако их номера связей можно восстановить для сравнения с S_i . При этом код пропуска формируется на двух связях, образованных левой или правой крайней линией в сечении. Далее, при мутации второго рода рассчитываются другие векторы состояния S_j^r и S_j^l , для которых код пропуска формируется на четырех связях, образованных двумя левыми или правыми крайними линиями в сечении. Итак, для мутации, не превышающей род $\delta(\epsilon)$, можно указать количество кодов пропуска, не превышающих величину ϵ . Выполняется устойчивость по Ляпунову [5].

2. Индексация изображения

Для вектора состояния S_i , на примере окончания (см. рис. 2) и разветвления (см. рис. 3) вычислим скалярное число переднего фронта F_i и заднего фронта B_i . Для этого по сечению выделим передний и задний фронт связей. Выписывая по порядку означенные связи, на основе табл. 1 построим два фронта для окончания в виде табл. 3 и табл. 4, а на основе табл. 2 - для разветвления в виде табл. 5 и табл. 6.

Выписывая сверху вниз из табл. 3-6 состояния, получим записи переднего фронта $F^a = 00110001$ и заднего фронта $B^a = 100011000$ для окончания, а также записи переднего фронта $F^b = 00110001$ и заднего фронта $B^b = 00011000$ для разветвления. В записях объединим состояния, ассоциированные с одинаковым номером частного признака. Тогда $F^c = 01001$ и $B^c = 1010$, а $F^d = 01001$ и $B^d = 1010$. Припишем к записям лидирующую единицу: $F^e = 101001$ и $B^e = 11010$, а $F^f = 101001$ и $B^f = 1010$.

Запись есть число. Если встречается код пропуска, то число обнуляется: $F^g = 101001$ и $B^g = 11010$, а $F^h = 0$ и $B^h = 0$. Код пропуска формируется обрывом или замыканием связи. Тогда для каждого вектора состояния S_i рассчитывают ненулевое скалярное число переднего фронта F_i и заднего фронта B_i . Для увеличения вероятности расчета ненулевых значений F_i и B_i применяют:

- выбор изображения с большим числом частных признаков;
- увеличение глубины сечения при построении вектора состояния S_i ;
- уменьшение глубины сечения для расчета переднего фронта F_i и заднего фронта B_i по сравнению с глубиной сечения вектора состояния S_i

Определение 3. Векторы состояния S_j и S_k называются центрально связанными, если состояния, ассоциированные с частными признаками M_j и M_k , располагаются на связях с минимальными номерами.

Таблица 3
Передний фронт для окончания

| Номер связи | Индекс | Состояние |
|-------------|--------|-----------|
| 14 | 20 | 0 |
| 10 | 20 | 0 |
| 6 | 23 | 1 |
| 2 | 23 | 1 |
| 3 | 24 | 0 |
| 7 | 24 | 0 |
| 11 | 26 | 0 |
| 15 | 27 | 1 |

Таблица 4
Задний фронт для окончания

| Номер связи | Индекс | Состояние |
|-------------|--------|-----------|
| 13 | 28 | 1 |
| 9 | 21 | 0 |
| 5 | 21 | 0 |
| 1 | 21 | 0 |
| 0 | 22 | 1 |
| 4 | 22 | 1 |
| 8 | 25 | 0 |
| 12 | 25 | 0 |
| 16 | 25 | 0 |

Таблица 5
Передний фронт для разветвления

| Номер связи | Индекс | Состояние |
|-------------|--------|-----------|
| 16 | 20 | 0 |
| 12 | 20 | 0 |
| 8 | 23 | 1 |
| 4 | 23 | 1 |
| 0 | 24 | 0 |
| 5 | 24 | 0 |
| 9 | 26 | 0 |
| 13 | 27 | 1 |
| 17 | — | — |

Таблица 6
Задний фронт для разветвления

| Номер связи | Индекс | Состояние |
|-------------|--------|-----------|
| 15 | — | — |
| 11 | 21 | 0 |
| 7 | 21 | 0 |
| 3 | 21 | 0 |
| 2 | 22 | 1 |
| 1 | 22 | 1 |
| 6 | 25 | 0 |
| 10 | 25 | 0 |
| 14 | 25 | 0 |
| 18 | — | — |

Построим простую цепь центрально связанных векторов состояния $C = S_i, \dots, S_j$ такую, что скалярные числа F_k и B_k , или только F_k , или только B_k ненулевые, где $k \in i \dots j$ и $i, j \in 1 \dots n$. Вычислим хеш-функцию h вида $h(F_i, \dots, F_j)$, $h(B_i, \dots, B_j)$ или $h(F_i, \dots, F_j, B_i, \dots, B_j)$, значение z которой поместим в список индексов изображения I_m . Фиксируя длину различных простых цепей, вычислим возможные значения z хеш-функции.

Определение 4. Индексом изображения называется скалярная величина, вычисляемая на основе ненулевых чисел переднего и заднего фронтов, определяемых в цепи центрально связанных векторов состояния.

Таким образом, при индексировании ДИ выполняется отображение в виде

$$\Gamma: \{S_k \in L_c \mid k \in i \dots j\} \rightarrow I_m = \{z\}, \quad (4)$$

где S_k – вектор состояния частного признака M_k по (1) в модели L_c по (2); z – значение хеш-функции для ненулевых скалярных F_k и B_k в цепи длиной $|C|$.

Заключение

В работе предложены векторы состояния, приближенно характеризующие области частных признаков. Доказывается устойчивость векторов состояния при мутациях частных признаков, что

обеспечивает возможность расчета устойчивых скалярных значений переднего и заднего фронтов векторов состояния, увязывание которых в цепь определяет индекс ДИ. Для одного ДИ вычисляется несколько таких индексов, играющих роль ссылок в коллекции изображений. Ссылки используются процедурой идентификации ДИ для ее ускорения.

Литература

1. Maltoni, D. *Handbook of fingerprint recognition* / D. Maltoni, D. Maio, A. K. Jain. - New York: Springer-Verlag, 2003. - 348p.
2. Новиков, Ф.А. *Дискретная математика для программистов: учебник для вузов* / Ф.А. Новиков. - СПб.: Питер, 2008. - 384 с.
3. Пат. 2185660 Российская Федерация, МПК G 06 K 9/52. *Способ кодирования отпечатка папиллярного узора* / В.Ю. Гудков, А. С. Боков. - № 2000118065/09; заявл. 07.07.2000; опубл. 20.07.2002; Бюл. № 20. - 13 с.
4. Пат. 2331108 Российская Федерация, МПК G 06 K 9/62. *Способ сравнения отпечатков папиллярных узоров* / Д.И. Аркабаев, В.Ю. Гудков. - № 2006140939/09; заявл. 20.11.2006, опубл. 10.08.2008; Бюл. № 22. - 17 с.
5. Афанасьев, В.Н. *Математическая теория конструирования систем управления* / В.Н. Афанасьев, В.Б. Колмановский, В.Р. Носов. - М.: Высшая школа, 1998. - 574 с.

Поступила в редакцию 21 декабря 2009 г.