

А. В. ОГНЕВ, А. П. ТИПИКИН

АЛГОРИТМ ИНВАРИАНТНОГО РАСПОЗНАВАНИЯ ОТПЕЧАТКОВ ПАЛЬЦЕВ ПО КЛЮЧЕВЫМ ТОЧКАМ

Рассматривается алгоритм, позволяющий существенно снизить зависимость вероятности распознавания отпечатков пальцев (ОП) от смещений, поворотов, фрагментации и дефектов поверхности папиллярного узора. Скорость распознавания увеличивается за счет введения этапа центрирования ОП с поиском наиболее достоверных пар базовых отрезков в сопоставляемых ОП, а также за счет сравнительной простоты распараллеливания алгоритмов распознавания, основанных на метрике Хаусдорфа, позволяющей находить расстояния между неравномощными множествами признаков.

Ключевые слова: идентификация отпечатка пальца, инвариантность, оцифровка, центрирование.

В современной криминалистике широко используются программы распознавания отпечатков пальцев (ОП). Для достижения высокой степени вероятности распознавания как на начальных, так и на завершающих стадиях работы программы требуется участие человека, и чем больше база данных ОП и выше вероятность повреждения отпечатков, тем чаще. Рассматриваемый алгоритм позволяет повысить степень автоматизации процесса распознавания, а также скорость распознавания на основе привлечения больших вычислительных ресурсов и ускорения отдельных процедур путем параллельной обработки информации.

Разработка алгоритмов распознавания ОП идет очень активно [1], но любой из существующих имеет свои недостатки: возможны ошибки false-reject rate (ошибочный отказ) и false-alarm rate (ошибочное подтверждение), возникающие из-за образующихся при сканировании смещений, поворотов, фрагментации, а также дефектов поверхности папиллярного узора. Избежать этих недостатков можно, если применять алгоритмы, инвариантные к названным негативным факторам. При решении задачи верификации (подтверждения личности) требуются мобильные датчики и вычислительные устройства (ВУ) со сравнительно небольшой тактовой частотой процессоров и объемом памяти. При решении задачи идентификации (определения личности) используются большие базы данных (БД), содержащие информацию об ОП и их обладателях, многократно повторяется процедура распознавания, поэтому возрастают требования к повышению производительности ВУ. В последнем случае целесообразно использовать параллельные алгоритмы сравнения и выполнять их на многопроцессорных ВУ. Так как в обоих случаях время принятия решения ограничено, вычислительная сложность инвариантного алгоритма не должна быть слишком большой, а для снижения требуемой емкости памяти желательно не хранить полное описание изображения ОП, а оперировать только с его моделями (множествами ключевых точек). Этого требуют соображения безопасности для того, чтобы при хищении из БД отпечатков ими нельзя было воспользоваться.

Известны два основополагающих алгоритма распознавания отпечатков пальцев: по отдельным деталям (ключевым точкам, КТ — развилкам и окончаниям, называемым минуциями) и по рельефу всей поверхности пальца [2]. В первом случае устройство регистрирует только некоторые точки, уникальные, свойственные конкретному отпечатку, и определяет их взаимное расположение. Во втором случае обрабатывается изображение всего отпечатка.

Алгоритм инвариантного распознавания отпечатков пальцев по ключевым точкам состоит из трех этапов: предварительная обработка изображения, центрирование, распознавание.

На этапе предварительной обработки каждое изображение ОП должно быть бинаризовано и отфильтровано. На этом этапе в черно-белом пиксельном изображении ОП устраняется часть ложных пропечаток, папиллярные линии утончаются до значения в один пиксел, логически устраняются некоторые их короткие ложные ответвления и небольшие разрывы. В этом изображении должны быть найдены граничные точки, определяющие контур информативной области ОП, его центр, а также выделены и описаны КТ в первичной (декартовой) системе координат относительно центра.

В настоящее время при использовании многих известных алгоритмов сравнения ОП сталкиваются с проблемой не вполне точного центрирования отпечатков [3]. Центрирование необходимо для ускорения процесса сравнения. Известно множество методов центрирования, но все они уязвимы к аффинным трансформациям (поворотам и смещениям) и дефектам папиллярного рисунка (фрагментации, порезам, повреждениям и т.д.) и не обеспечивают высокой точности [4]. Можно выделить два основных подхода: нахождение пиксельного центра тяжести и определение центра информативной области ОП. Если за центры сравниваемых ОП принимаются центры тяжести всех пикселей папилляров каждого отпечатка, то вследствие неправильного расположения пальца на считывающем устройстве при снятии отпечатка возникают неточности при фрагментации ОП.

Выделение КТ и вычисление их координат выполняется в первичной системе координат относительно неточного центра ОП путем просмотра каждого пиксела отпечатка и его ближайшей окрестности. Если в этом пикселе папилляр разветвляется или оканчивается, он принимается за КТ.

Широко применяемым инвариантным алгоритмом распознавания ОП является метод корреляционного сравнения [4]. Сопоставление двух ОП основано на многократных сравнениях множеств КТ и выполняется со скоростью 200—3000 сравнений в секунду [5], что приводит к значительным затратам времени. При сопоставлении перебираются пары КТ по принципу „каждая с каждой“. В окрестности каждой КТ проводится поиск ближайшей КТ другого отпечатка. Если она расположена на допустимом удалении, то эти две точки считаются совпавшими. В качестве меры близости двух ОП принимается число пар их КТ, признанных совпавшими. При сопоставлении двух ОП за достоверный центр одного из ОП принимается поочередно каждый пиксел его центральной области [6], размер которой определяется величиной возможного смещения ОП при сканировании и может составить до 10 % от длины рамки изображения ОП.

Перед каждым сопоставлением вносятся небольшой угол поворота одного из ОП и смещение его центра в другой пиксел центральной области и вновь вычисляется мера близости двух ОП. Повторные сопоставления двух ОП продолжаются до тех пор, пока не будут перебраны все возможные сочетания пикселей центральной области и возможных углов поворота одного из ОП, например, с шагом в 1° в диапазоне $\pm 15^\circ$. Решение об идентификации принимается по наибольшему значению меры близости из всех вариантов сопоставлений, если оно не меньше заданного порога. Число парных сравнений КТ двух ОП, требуемое для принятия решения, равно

$$N_{\text{ср}} = N_d N_{\text{ЦО}} N_{\text{КТ}}^2,$$

где N_d — число шагов поворота одного из ОП; $N_{\text{ЦО}}$ — число пикселей центральной области; $N_{\text{КТ}}$ — число КТ в одном из ОП. $N_d = 30$ при шаге 1° и диапазоне $\pm 15^\circ$; $N_{\text{ЦО}} = 60 \cdot 40 = 2400$ пикселей при растре 600×400 пикселей; $N_{\text{КТ}} = 70$ точек. Поэтому требуемое число парных сравнений точек $N_{\text{ср}} \approx 3,5 \cdot 10^8$. Это приводит к большим затратам времени при реализации данного способа сопоставления на современных компьютерах.

Рассматриваемый алгоритм позволяет существенно снизить зависимость вероятности распознавания ОП от смещений, поворотов, фрагментации и дефектов поверхности папиллярного узора, а также распараллелить их многовариантное сопоставление для повышения

скорости распознавания. Применяемые в алгоритме процедуры составления избыточных векторных описаний изображения ОП и параллельного их сопоставления позволяют автоматизировать формирование БД эталонов ОП и повысить вероятность их инвариантного распознавания. Инвариантность по отношению к смещениям, поворотам, фрагментации и наиболее вероятным дефектам изображений ОП достигается путем многовариантного сопоставления с использованием метрики Хаусдорфа [1] их избыточных векторных описаний в полярной системе координат относительно базовых отрезков (БО).

Скорость инвариантного распознавания ОП увеличивается за счет дополнительного предварительного этапа поиска наиболее достоверных пар базовых отрезков в сопоставляемых отпечатках. После ранжирования пар базовых отрезков можно сократить требуемое число вариантов сопоставления векторных описаний множеств минутий до трех, а число шагов попарной обработки минутий двух отпечатков — до 10^4 — 10^5 [6]. Дополнительным резервом повышения скорости является сравнительная простота распараллеливания алгоритмов распознавания, основанных на метрике Хаусдорфа, позволяющей находить расстояния между неравномоощными множествами признаков [7].

Требуемое число вариантов сравнения ОП по ключевым точкам может быть уменьшено на основе следующей методики центрирования [6]. Для достижения желаемой инвариантности необходимо перейти от „грубой“ (декартовой) к „новой“ (полярной) системе координат, центром которой будет являться центр достоверного базового отрезка. Для повышения эффективности работы алгоритма из рассмотрения исключаются КТ, близкие к границе информативной области („зашумленность“ изображения ОП при приближении к его краям повышается), и к центру ОП (из-за неточности расположения центра ОП). БО формируется только из оставшихся внутренних точек окончаний и разветвлений папилляров.

Минимального числа БО с сохранением единообразия их выбора в паре сравниваемых ОП можно достичь по следующему правилу обхода точек. Вращая по часовой стрелке радиус-вектор из центра первичной системы координат, последовательно соединить отрезками прямых встречающиеся КТ. Если несколько КТ встречаются в одном направлении радиуса-вектора, то предыдущая точка соединяется с самой удаленной от центра. При этом остальные КТ соединяются между собой последовательно по мере приближения к центру. Таким образом, формируется замкнутый контур БО, представляющий собой ломаную с вершинами в КТ.

После составления контуров БО необходимо в обоих ОП выделить пары БО, близкие по длине и взаимной ориентации. Последняя определяется значениями углов с четырьмя соседними БО в контуре (см. рисунок). Так как могут выпадать соседние БО или появляться ложные отрезки, то определить достоверную пару БО можно следующим образом. Сначала определим, в каком из ОП имеется меньшее число БО, входящих в ломаные. Этот отпечаток A используется для минимизации числа проверяемых пар БО. Для каждого БО отпечатка A необходимо выбрать такой парный ему БО из второго отпечатка B , для которого будет минимальной мера C :

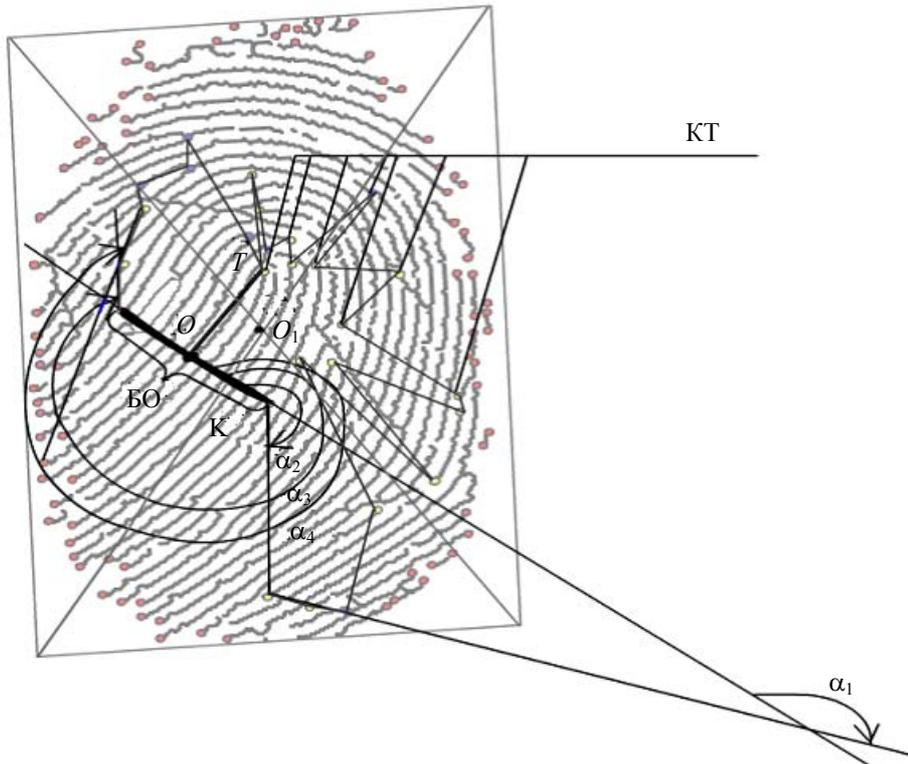
$$\begin{aligned}
 C_1 &= \frac{l_A - l_B}{L} + \frac{a_{A_1} - a_{B_1}}{\Lambda}, \\
 C_2 &= \frac{l_A - l_B}{L} + \frac{a_{A_2} - a_{B_2}}{\Lambda}, \\
 C_3 &= \frac{l_A - l_B}{L} + \frac{a_{A_3} - a_{B_3}}{\Lambda}, \\
 C_4 &= \frac{l_A - l_B}{L} + \frac{a_{A_4} - a_{B_4}}{\Lambda}, \\
 C &= \min\{C_1, C_2, C_3, C_4\},
 \end{aligned} \tag{1}$$

где l — длина БО; a — угол с соседним БО в пределах одного и того же ОП; L — максимальная длина БО в обоих ОП; Λ — максимально возможная величина угла, равная 2π .

В качестве центра O новой полярной системы координат в одном из ОП принимается центр достоверного БО (см. рисунок) той пары БО, для которой будет значение C наименьшим. Достоверная пара БО выбирается по следующей формуле:

$$C_{\min} = \min_i \left\{ \min_j \left\{ \min_K \{C\}_{K=1}^{K=4} \right\}_{j=1}^{j=N_B} \right\}_{i=1}^{i=N_A},$$

где N_A, N_B — число БО в отпечатках A и B соответственно; K — число соседних БО, используемых при нахождении меры C (1), где $K = 4$.



Учитывая возможную близость между собой по значению C нескольких первых проанжированных пар БО, в дальнейшем величины метрики Хаусдорфа целесообразно вычислять для первых трех пар БО, а расстояния между ними выбирать минимальными.

Векторные описания множеств КТ формируются в полярной системе координат с учетом расстояния от центра O достоверного БО до данной КТ, например до точки T (см. рисунок), и угла между лучом, проведенным из центра O в первый (K) конец БО, и лучом, проведенным из центра БО в данную КТ. Перевод из первичной (декартовой, O_1 — ее центр) в новую полярную систему координат осуществляется по следующим формулам:

$$L = \sqrt{(X_T - X_O)^2 + (Y_T - Y_O)^2},$$

$$\alpha = \arccos \frac{(X_T - X_O)^2 + (Y_T - Y_O)^2 + (X_K - X_O)^2 + (Y_K - Y_O)^2 - (X_K - X_T)^2 - (Y_K - Y_T)^2}{2\sqrt{(X_T - X_O)^2 + (Y_T - Y_O)^2} \sqrt{(X_K - X_O)^2 + (Y_K - Y_O)^2}},$$

где X_T, Y_T — абсцисса и ордината данной КТ в первичной системе координат; X_O, Y_O — абсцисса и ордината центра достоверного БО в первичной системе координат.

После того как сформированы описания множеств КТ относительно центров достоверных БО в обоих ОП, для каждого элемента множеств КТ, соответствующих отпечаткам A и B , вычисляются минимальные расстояния от всех элементов другого множества:

$$e_k^A = \min_t \left\{ d \left((\alpha_k^A, l_k^A), (\alpha_t^B, l_t^B) \right) \right\}_{t=1}^{t=N_B}, \quad k=1, \dots, N_A,$$

$$e_t^B = \min_k \left\{ d \left((\alpha_t^B, l_t^B), (\alpha_k^A, l_k^A) \right) \right\}_{k=1}^{k=N_A}, \quad t=1, \dots, N_B, \quad (2)$$

где $d(x, y)$ — расстояние между элементами множеств x и y ; t, k — номер КТ в множествах A и B соответственно.

Затем вычисляются максимальные значения из множеств полученных минимумов:

$$e_{\max}^A = \min_k \{ e_k^A \},$$

$$e_{\max}^B = \min_t \{ e_t^B \}. \quad (3)$$

После этого вычисляются максимум и минимум из двух полученных величин (3):

$$e_{\max i,j} = \max (e_{\max}^A, e_{\max}^B),$$

$$e_{\min i,j} = \min (e_{\max}^A, e_{\max}^B), \quad (4)$$

где i, j — номера выбранных достоверных БО в A и B соответственно.

Минимум (см. выражение (4)) находится для того, чтобы определить степень включения одного множества в другое. Такая модификация алгоритма сопоставления неравномоощных множеств на основе метрики Хаусдорфа позволяет численно определить степень идентичности фрагментов пары ОП.

Решение об идентичности ОП принимается на основе следующего анализа значений $e_{\max i,j}$ и $e_{\min i,j}$ (4). Если $e_{\max i,j} \leq h$ и $e_{\min i,j} \leq h$, то ОП идентичны. Если $e_{\max i,j} > h$ и $e_{\min i,j} > h$, то ОП принадлежат разным людям. Если $e_{\max i,j} > h$ и $e_{\min i,j} \leq h$, то фрагменты сопоставляемых ОП близки. Величина порога h выбирается пропорциональной среднему значению меры C (1) и может быть подобрана опытным путем на тестовых парах сопоставляемых ОП. Вычислительная сложность описанного алгоритма инвариантного распознавания ОП определяется в основном необходимостью трехкратного повторения расчетов по формулам (2)—(4). Затраты времени на остальные описанные выше процедуры составляют не более 4 %. Сопоставление двух ОП выполняется примерно за $3,5 \cdot 10^4$ шагов, каждый из которых эквивалентен вычислению одного расстояния между двумя КТ и по длительности не превышает затрат времени на 10 парных сравнений КТ в корреляционном методе. Таким образом, скорость инвариантного распознавания возрастает практически в 1000 раз.

Преимуществами данного алгоритма являются повышение степени инвариантности к поворотам, смещениям, фрагментации и дефектам ОП, а также возможность ускорения процесса распознавания за счет сравнительной простоты распараллеливания вычислений при центрировании ОП (1) и нахождении расстояний между сравниваемыми ОП по формулам (2)—(4).

СПИСОК ЛИТЕРАТУРЫ

1. Биометрия. Отпечаток пальца [Электронный ресурс]: <<http://www.biometrics.ru>>.
2. Рябов Г. В. Современные технологии идентификации личности по отпечатку пальца с использованием емкостных датчиков [Электронный ресурс]: <http://www.radioradar.net/articles/scientific_technical>.

3. Яковлев А. В., Румянцев Р. В. Разработка алгоритма идентификации личности по изображению отпечатка пальца // Методы и системы обработки информации. Сб. науч. ст. Ч. 2 / Под ред. С. С. Садыкова, Д. Е. Андрианова. М.: Горячая линия – Телеком, 2004. С. 41—48.
4. Кухарев Г. А. Биометрические системы: Методы и средства идентификации личности человека. СПб: Политехника, 2004.
5. Пат. № 2310910 РФ. Способ верификации и идентификации отпечатков папиллярных узоров / А. А. Борейшо, Д. А. Фокин, С. Я. Чакчир, А. Ю. Хозин, Е. М. Орлов. 16.05.2006.
6. Огнев А. В., Типикин А. П. Центрирование отпечатков пальцев при инвариантном распознавании на основе метрики Хаусдорфа // Оптико-электронные приборы и устройства в системах распознавания образов, обработки изображений и символической информации. Сб. мат. VIII Междунар. конф. „Распознавание – 2008“. Ч. 2. Курск: КурскГТУ, 2008. С. 34—35.
7. Ognev A. V., Tipikin A. P. Method of invariable recognition of fingerprints based on metric of Hausdorff // Proc. 6th Int. Conf. „Information and telecommunication technologies in intelligent systems“. Crete, Greece. June 2—6, 2008. P. 78—82.

Сведения об авторах

Антон Вадимович Огнев

— магистрант; Курский государственный технический университет, кафедра вычислительной техники; E-mail: light@kursknet.ru

Александр Петрович Типикин

— д-р техн. наук, профессор; Курский государственный технический университет, кафедра вычислительной техники; E-mail: cct@cafct.kstu.kursk.ru

Рекомендована кафедрой
вычислительной техники

Поступила в редакцию
12.09.08 г.