

- знизити значення частинних коефіцієнтів прискорень ($K_n^T, K_n^R, K_n^f, K_n^U$) та підвищити вірогідність процесу форсованих випробувань, наблизивши ідентичність процесів, що відбуваються в РЕК (фізико-хімічних процесів), до звичайних умов експлуатації.

Список використаної літератури: 1. Вишинівський В. В., Кузавков В. В., Редзюк С. В. Аналіз методів форсованих випробувань для отримання залежності зміни діагностичного параметра від часу напрацювання напівпровідникових РЕК // Збірник наукових праць Військового інституту Київського національного університету імені Тараса Шевченка. – К.: ВІКНУ, 2013. – Вип. №43. 2. Волков А. Н. Роль ускоренных испытаний в определении надежности интегральных схем [Текст] / А. Н. Волков // Молодой ученый. — 2012. — №10. — С. 41-52. 3. Строгонов А. В. «Долговечность интегральных схем и производственные методы ее прогнозирования»: <http://www.chipinfo.ru/literature/chipnews/200206/8.html>. 4. Горлов М. И., Емельянов В. А., Строгонов А. В. Геронтология кремниевых интегральных схем. М: Наука. 2004. 5. Stojadinovic N. D. Failure physics of integrated circuits—a review // Microelectron. Reliab. 1983. Vol. 23. N4. P. 609-707. 6. Stojadinovic N. D., Ristic S. D. Failure physics of integrated circuits and relationship to reliability // Phys. Stat. Sol. (a). 1983. Vol. 75. P. 11-47. 7. Wurnik F., Pelloth W. Zuver lassigkeit von integrierten schaltungen // Nachrichtennische zeltschrift//1984. Vol. 37. N U.S. 710-712, 714-716. 8. ADI Reliability Handbook 2001. www.analoQ.com 9. Atmel corporation. Quality & reliability handbook 2001-2002. Rev. 09/01// www.atmel.com 10. РД 11 0755-90. Микросхеми інтегральні. Методи ускоренних испытаній на безотказність і довговечність. РНИИ «Електронстандарт». 1990. 11. Bowles J. B. A survey of reliability prediction procedures for microelectronics devices // IEEE Trans. Reliab. 1992. Vol. 41. N 1. P. 2-12. 12. Горлов М. И., Королев С. Ю., Кулаков А. В., Строгонов А. В. Расчет надежности интегральных схем по конструктивно-технологическим данным. Воронеж: Изд-во Воронежского университета. 1996.

Евгений Подгорный, Любовь Рябова, Владимир Темников

Национальный авиационный университет

УДК 004.62

СПОСОБ ПОВЫШЕНИЯ БЫСТРОДЕЙСТВИЯ СИСТЕМЫ КОНТРОЛЯ ДОСТУПА ПО РАДУЖНОЙ ОБОЛОЧКЕ ГЛАЗА

Анотація: Проаналізовано різні підходи до прискорення процедури пошуку еталонів в базах даних зображень райдужної оболонки ока автоматизованих систем розпізнавання осіб. На основі проведеного аналізу запропоновано спосіб підвищення швидкодії системи контролю доступу на об'єкти інформаційної діяльності.

Summary: The article analyzes different approaches to accelerate the search procedure of standards in databases iris image. Based on this analysis, the authors proposed a method for increasing the speed of the access control system at the facility of information activities.

Ключові слова: Система контролю доступу, райдужна оболонка ока, кластеризація.

I Введение

Одними из составных частей систем технической защиты информации на объектах информационной деятельности являются автоматизированные системы контроля доступа (СКД) на эти объекты, работа которых основана на сканировании, обработке и анализе биометрических признаков человека. Наиболее эффективным, как показал анализ литературных источников и исследования авторов статьи, является применение таких признаков человека, как радужная оболочка и сетчатка глаза, голос, геометрия лица и некоторых других [1, 2].

При этом СКД строятся по типовой схеме распознавания образов, включающей в себя подсистемы предварительной обработки изображений или сигналов, параметризации, классификации и принятия решения. Важными элементами СКД являются базы данных (БД) зарегистрированных пользователей.

Быстродействие СКД, являющееся важнейшим показателем эффективности ее работы, существенно зависит от быстродействия подсистемы классификации, которое в значительной степени определяется временем взаимодействия с БД и, в первую очередь, – продолжительностью поиска эталонных в базах данных. Это обуславливает высокую актуальность решения проблемы ускорения этого процесса.

Значимость решения указанной проблемы в настоящее время повышается из-за наблюдающегося увеличения размеров БД в различных применениях: база данных системы контроля на проходной в офис компании может включать до десяти тысяч эталонных, система контроля въезда-выезда в страну – порядка

1 млн. эталонов, база данных государственной системы идентификации – порядка 100 млн. эталонов. С ростом размеров БД продолжительность поиска эталонов в них существенно увеличивается. При этом заметим, что в процессе эксплуатации систем распознавания человека периодически возникает задача проверки консистентности БД, то есть проверка наличия в ней совпадающих эталонов (такая проверка требует проведения поиска по всей БД для каждого эталона).

В настоящей статье проанализированы известные способы уменьшения времени, необходимого для распознавания человека биометрическими системами, построенными на основе сканирования, обработки и анализа изображений радужной оболочки глаза (РОГ), указаны их недостатки и приведен разработанный авторами на основе проведенных исследований способ минимизации времени поиска эталонов в БД, применение которого позволяет значительно ускорить процесс распознавания людей по РОГ. Выбор РОГ для построения систем распознавания обусловлен тем, что распознавание людей по этому биометрическому признаку является одним из самых точных.

II Анализ современного состояния проблемы. Существующие способы уменьшения времени распознавания человека по радужной оболочке глаза

Описанные в литературе алгоритмы распознавания человека по изображению РОГ построены на основе теории распознавания образов и включают в себя следующие основные этапы:

1. Сканирование изображения РОГ.
2. Локализация РОГ (определение границ РОГ и зрачка).
3. Нормализация изображения РОГ (отображение кольца РОГ в прямоугольную область фиксированного размера).
4. Параметризация РОГ (представление изображения РОГ в виде информативных параметров - кодов).
5. Сравнение контрольных кодов РОГ с эталонными, хранящимися в БД.

Впервые этот широко известный в настоящее время алгоритм выделения параметров РОГ был предложен Дж. Даугманом [3]. Алгоритм Даугмана основан на модуляции сигнала изображения РОГ с применением вейвлетов Габора. Информативными параметрами РОГ являются их двоичные коды.

Время, необходимое для осуществления первых четырех этапов, не зависит от размера БД. Последний этап в большинстве существующих алгоритмов состоит в сравнении контрольного кода РОГ пользователя последовательно с каждым эталоном, хранящимся в БД, которое заключается в вычислении расстояния между ними по определенным правилам и сравнении его с пороговым значением. Способ вычисления расстояния или меры схожести выбирается в зависимости от способа вычисления эталона – например, для IrisCode [3] Даугман предложил использовать относительное расстояние Хэмминга:

$$HD(C_1, C_2) = \frac{1}{N} \sum_{i=1}^N (C_{1i} \oplus C_{2i}), \text{ где } C_1, C_2 - \text{коды РОГ, } N - \text{количество бит в коде.}$$

Вследствие необходимости проведения перебора всех эталонов, содержащихся в БД, процесс сравнения контрольных кодов РОГ с эталонными занимает относительно много времени. Попытки ускорить процесс сравнения были предприняты в [4, 5].

В [4] для систем распознавания РОГ, работающих на основе IrisCode, был разработан способ быстрого поиска FFS. Он основывается на точном совпадении кодов, что применимо для IrisCode, но не подходит для других, более эффективных, алгоритмов, таких, как преобразование Хаара. Кроме того, необходимость полного совпадения кодов снижает вероятность того, что код РОГ зарегистрированного пользователя будет найден в БД. Для повышения качества работы системы можно проводить поиск несколько раз, однако это приводит к еще большим затратам времени на распознавание человека.

Способ, описанный в [5], основан на вычислении чисел Эйлера. Применение этого способа при правильном распознавании 97,2% эталонов позволяет уменьшить время поиска в 3 раза по сравнению с оригинальным алгоритмом Дж. Даугмана. Однако этого явно недостаточно.

Как видно из проведенного анализа, решение проблемы повышения быстродействия биометрических систем распознавания по-прежнему остаётся актуальным.

III Основная часть

Анализ литературных источников и проведенные авторами исследования показали, что минимизировать продолжительность поиска соответствующего эталона в БД (а, следовательно, и время, необходимое для распознавания человека) можно двумя путями:

- 1) за счет уменьшения размера кодового слова;

2) за счет уменьшения необходимого количества процедур сравнения при применении методов кластеризации вместо часто применяющегося на практике простого перебора эталонов.

Методы уменьшения размера кодового слова

Продолжительность указанной процедуры можно уменьшить путем снижения размеров кодового слова, которым изображение РОГ описывается в процессе его параметризации.

В [6] описан способ уменьшения размеров кодового слова из-за отказа при его формировании от использования секторов по 45° в нижней и верхней частях глаза. Это делается для исключения влияния на результат распознавания человека век, ресниц, а также накапливающейся влаги и бликов от сканера в нижней части глаза (анализ показал, что на борьбу с негативным влиянием указанных артефактов расходуется значительное время).

Авторы [7] предлагают дополнительно уменьшить размер кодового слова, описывающего РОГ (а с ним и продолжительность поиска эталона в базе данных), путем уменьшения размера считывающей матрицы изображений РОГ. Так, например, вместо исходной матрицы, включающей 100 ключевых точек ($S*S$) изображения, предлагается брать укороченную матрицу размером ($S/3*S/3$). Размер кода заметно уменьшается (он становится равным 61), а количество неверно определенных изображений РОГ составляет лишь 1,5%. Малый процент ошибок объясняется содержанием на радужной оболочке глаза значительной информационной избыточности в виде большого количества уникальных для каждого человека ключевых элементов; при этом, для надежного распознавания человека достаточно совпадения 40% кода РОГ.

Очевиден тот факт, что при сканировании с последующим кодированием и сравнении единичного кода изображения выигрыш по времени будет относительно мал, но при проведении большого количества процедур суммарное уменьшение временных затрат может быть значительным.

Разработанный способ уменьшения необходимого количества процедур сравнения при применении кластеризации

Исключить простой перебор эталонов предлагается на основе кластеризации всего множества эталонов РОГ с последующим проведением поиска только в некотором подмножестве эталонов, хранящихся в БД.

Процесс поиска при применении разработанного способа состоит из двух этапов:

- на первом этапе все изображение РОГ разбивается на компактные подмножества (кластеры);
- на втором этапе из выбранного на первом этапе кластера выделяется искомым эталон.

Для разбиения множества эталонов на кластеры авторы применили один из иерархических агломеративных методов кластеризации [8], а именно – метод одиночной связи. Иерархические агломеративные методы различаются, главным образом, по правилам построения кластеров. Известны, по крайней мере, двенадцать различных методов кластеризации, наиболее распространенными из которых являются методы одиночной связи, полной связи, средней связи и метод Уорда. Главное преимущество выбранного метода заключается в том, что результаты его применения инвариантны к монотонным преобразованиям матрицы сходства. Это означает, что метод одиночной связи является одним из методов, оставляющих без изменения порядок очередного объединения элемента генеральной выборки с последним выделенным кластером, при относительном упорядочении элементов матрицы сходства.

При применении метода одиночной связи кластеры образуются по следующему правилу: объект будет присоединен к уже существующему кластеру, если, по крайней мере, один из его элементов находится на том же уровне сходства, что и объект, претендующий на включение в этот кластер. Таким образом, присоединение объекта к определенному кластеру обусловлено лишь наличием единственной связи между объектом и кластером.

Для определения расстояния между кластерами по результатам экспериментов была выбрана метрика «по дальнему соседу» (когда расстоянию между кластерами присваивается максимальное значение попарных расстояний между элементами двух кластеров). Этот подход позволяет получить более сбалансированное распределение объектов по кластерам, чем, например, использование метрики «по ближнему соседу» (когда расстоянию между кластерами присваивается минимальное значение попарных расстояний между элементами двух кластеров), применение которого ведет к укрупнению кластеров.

В процессе исследования был разработан алгоритм кластеризации. Количество кластеров в некоторых случаях известно заранее, однако чаще ставится задача определить оптимальное число кластеров с точки зрения того или иного критерия качества кластеризации. Решение задачи кластеризации принципиально неоднозначно, и тому есть несколько причин (как считает ряд авторов), так как не существует однозначно наилучшего критерия качества кластеризации. Известен целый ряд эвристических критериев, а также ряд алгоритмов, не имеющих четко выраженного критерия, но осуществляющих достаточно разумную кластеризацию «по построению». Все они могут давать разные результаты.

Алгоритм иерархической кластеризации включает в себя этапы последовательного (пошагового) объединения кластеров и определения расстояний между объединяемыми на каждом шаге кластерами.

Ниже приведены результаты исследований, проведенных с изображениями РОГ, содержащимися в базе изображений CASIO-Iris [9] на выборке из 124 человек по 20 различных изображений РОГ для каждого человека. Все пространство выбранных изображений было разбито на 18 кластеров; при этом 90% элементов содержалось в 8 кластерах, а наиболее представительный кластер содержал 31% элементов.

С применением алгоритма иерархической кластеризации была рассчитана зависимость количества кластеров M от расстояния между кластерами при пошаговой кластеризации генеральной совокупности данных РОГ. Начальный этап кластеризации представляет собой монотонный убывающий процесс, относительно численности некластеризованных объектов. Завершающий этап процесса кластеризации генерального множества изображений РОГ характеризуется формированием ряда крупных кластеров и дальнейшее объединение в кластер возможно только за счет резкого увеличения расстояния между ними, что объясняется слиянием крупных или разнородных кластеров. Другими словами, все методы кластеризации опираются на исходную «гипотезу компактности»: в пространстве объектов все близкие объекты должны относиться к одному кластеру, а все далекие объекты соответственно должны находиться в различных кластерах.

Анализ показал, что выбор меры сходства несущественно влияет на быстродействие систем распознавания. Поэтому в процессе исследований использовалась только одна мера сходства - евклидово расстояние.

В табл. 1 представлены значения евклидова расстояния между кластерами при их различном количестве.

Таблица 1 – Евклидово расстояние между кластерами при их различном количестве

Количество кластеров	10	9	8	7	6	5	4	3	2	1
Евклидово расстояние между кластерами	0,23	0,34	0,37	0,43	0,68	0,71	0,74	1,75	1,97	2,23

Из сравнения результатов расчетов, приведенных в табл. 1, видно, что значение евклидова расстояния при уменьшении количества кластеров с четырех до трех увеличивается скачком. Это позволяет сделать вывод о целесообразности принятия для дальнейшего анализа количества кластеров равным четырем.

Сложность выделения нужного скачка обусловлена наличием относительно большого количества малых скачков значений расстояний между кластерами – исходя лишь из простого визуального анализа, невозможно указать, какой из этих скачков “правильный”. Можно лишь утверждать, что скачок – свидетельство объединения очередного кластера с “посторонним” объектом или большим кластером. Поэтому для практического применения предлагается использовать простейшую формулу линейного прогноза значения случайной величины монотонного процесса. Эксперименты показывают, что количество кластеров, в которые попадают изображения радужной оболочки одного и того же глаза, существенно меньше общего количества кластеров и количества изображений РОГ одного человека.

Для того, чтобы оценить, насколько хорошо первоначально была проведена кластеризация и насколько быстро «стабилизируются» кластеры на каждом шаге кластеризации методом k -средних, вычислялась доля элементов, поменявших свой кластер. Для каждого зарегистрированного в БД пользователя можно вычислить частоту распределения элементов по кластерам и затем упорядочить кластеры по количеству входящих элементов в различные кластеры. Например, один человек встречается в 3 кластерах, причем в одном из них 14 раз, во втором – 5, в третьем – 1, что соответствует значениям частоты (0,7; 0,25 и 0,05); второй человек – 20 раз в одном кластере (соответствует частота – 1), а третий – 16 раз в одном и ещё по одному разу в 4 кластерах (соответствуют частоты – 0,8; 0,05; 0,05; 0,05 и 0,05). Далее усреднением этих распределений по всем пользователям получим среднюю частоту p_i – вероятность того, что эталон будет найден в i -ом по порядку перебора кластере. В нашем примере: в первом кластере $p_1 = (0,7+1+0,8)/3 = 0,833$, во втором – $p_2 = (0,25 + 0 + 0,05)/3 = 0,1$ и т. д.

Оценка продолжительности поиска

Для оценки продолжительности поиска исходные изображения РОГ были распределены по кластерам. По количеству входящих в тот или иной кластер изображений РОГ каждого пользователя определялась вероятность того, что изображение РОГ пользователя в процессе его распознавания будет обнаружено в этом кластере. Далее для каждого эталона кластеры ранжировались по возрастанию расстояния до центра кластера. После этого рассчитывалось суммарное время поиска с учётом вероятностей того, что изображение РОГ нужного пользователя будет обнаружено в каждом кластере. По продолжительности поиска для различных эталонов строится распределение и вычисляется среднее время поиска.

Общее время поиска равно сумме следующих слагаемых:

- продолжительность времени перебора всех k -элементов i -го кластера, помноженная на вероятность обнаружения в i -кластере;
- произведение среднего времени поиска в ближайшем $(i - 1)$ кластере из j -элементов на вероятность того, что объект будет найден в этом же кластере;
- средняя продолжительность поиска в следующем $(i - 2)$ кластере из l -элементов, умноженная на вероятность того, что объект будет найден в данном кластере и так далее.

Такой расчёт можно провести для каждого эталона в процессе поиска, зная какой размер у каждого следующего по удалённости кластера.

В предположении, что для каждого эталона сначала придется перебирать наиболее крупные кластеры, среднее время поиска для приведенных выше данных составит

$$T \leq \frac{M}{8}.$$

В то же время, полный перебор базы данных M с учётом равной вероятности того, что каждый эталон БД может оказаться искомым элементом, будет произведен за время, равное в среднем $M/2$.

Таким образом, разбиение выборки M на группы схожих объектов позволяет упростить дальнейшую обработку данных и принятия решений, применяя к каждому кластеру свой метод анализа. В качестве меры различия целесообразно использовать квадратичное евклидово расстояние, которое способствует увеличению контрастности кластеров.

Проведенный анализ показал, что использование статистических методов, в частности, методов кластеризации, позволяет существенно сократить время поиска соответствующего эталона в базе данных. Но возникает вопрос устойчивости принятого кластерного решения. По сути, проверка устойчивости кластеризации сводится к проверке её достоверности. Здесь существует эмпирическое правило — устойчивая типология сохраняется при изменении методов кластеризации. Результаты иерархического кластерного анализа можно проверить кластерным анализом по методу k -средних. Если сравниваемые классификации групп респондентов имеют долю совпадений более 70 % (более 2/3 совпадений), то кластерное решение принимается.

IV Выводы

В статье проанализированы методы уменьшения продолжительности поиска эталонов в больших базах данных и приведен разработанный авторами способ уменьшения продолжительности процедуры сравнения контрольного изображения РОГ с эталонами, содержащимися в БД, применение которого позволяет существенно повысить эффективность функционирования крупных систем контроля доступа при распознавании человека по радужной оболочке глаза.

Список использованной литературы: 1. Ворона В. А., Тихонов В. А. Системы контроля и управления доступом. – М.: Горячая линия-Телеком, 2010. – 272 с. 2. Гончар В. К. Биометрические системы и их применение // Бизнес и безопасность. – 2002. – № 6. – С. 30-31. 3. Daugman J. How Iris Recognition Works // IEEE Trans. – 1993. – CSVT 14(1). – P.21–30. 4. Hao F., Daugman J., Zielinski P. A Fast Search Algorithm for a Large Fuzzy Database // IEEE Trans. Information Forensics and Security. – 1994. – №3(2). – P.203–212. 5. Бериков В.С., Лбов Г.С. Современные тенденции в кластерном анализе // Всероссийский конкурсный отбор обзорно-аналитических статей по приоритетному направлению «Информационно-телекоммуникационные системы», 2008. – 26 с. 6. Павельева Е. А., Крылов А. С. Алгоритмы предобработки радужной оболочки глаза // Труды конференции «GraphiCon». – М., 2008. – С.314. 7. Павельева Е. А., Крылов А. С., Урмаев О.С. Развитие информационной технологии идентификации человека по радужной оболочке глаза на основе преобразования Эрмита // Системы высокой доступности. – 2009. – № 1. – С.36–42. 8. Факторный, дискриминантный и кластерный анализ / Дж.-О. Ким, Ч. У. Мьюллер, У. Р. Клекка, М. С. Олдендерфер, Р. К. Блэифилд; под ред. И. С. Енюкова. – М.: Финансы и статистика, 1989. – 215 с. 9. База данных CASIA-IrisV3; <http://www.cbsr.ia.ac.cn/IrisDatabase.htm>.