DOI: 10.18372/2410-7840.18.11094
УДК 621.39

КРИТЕРИЙ ОБНАРУЖЕНИЯ КАК ВЛИЯЮЩИЙ ФАКТОР ОБЪЕМА БАЗЫ ДАННЫХ БИОМЕТРИЧЕСКИХ СИСТЕМ КОНТРОЛЯ ДОСТУПА

Андрей Фесенко

В статье для повышения быстродействия обработки радужной оболочки глаза и уменьшения объемов баз данных, для хранения изображения радужной оболочки, предлагается предварительная обработка входных изображений DOG-фильтром или фильтром Габора, последующим вычислением расстояния Хемминга по коду радужной оболочки глаза и ее статистических параметров. Оценка влияния искажений на изображение радужной оболочки глаза, на статистические параметры (плотность распределения вероятности) расстояния Хэмминга. Сравнение расстояния Хэмминга двух изображений радужной оболочки и принятия решения их идентичности. Выбраны статистические критерии обнаружения для уменьшения объема баз данных, предварительно обработанных изображений радужной оболочки глаза и повышения быстродействия идентификации-аутентификации системы контроля и управления доступом по радужной оболочке глаза.

Ключевые слова: *радужная оболочка, расстояние Хэмминга, статистические критерии, DOG-фильтр, плотность распределения.*

Вступление. Одним из самых популярных и, наверное, самым надежным биометрическим признаком для автоматического распознавания личности в настоящее время является радужная оболочка глаза (радужка). Вероятность ложного допуска системы распознавания по радужке менее 0,001% (фактически система не совершает ни одной ошибки ложного узнавания при более чем 2 млн элементарных сравнений эталонов). Вероятность ложного не допуска при этом составляет около 2%. Для всех прочих биометрических систем такая вероятность ложного допуска достигается лишь при неприемлемых значениях не допуска в десятки процентов, причина столь высоких характеристик систем распознавания радужки состоит в том, что структура радужки – устойчивый, хорошо выраженный и высокоинформативный биометрический признак.

Работа систем контроля доступа по радужной оболочке глаза основана на сравнении эталонного черно-белого изображения с изображением, полученным от идентифицируемой личности. Размер эталонного изображения отвечает следующим требованиям: размерность 320x200 пикселей, оттенки серого с минимальной градацией 256 цветов (8 бит на пиксель) [1], тогда объём эталонного изображения в базе данных будет занимать 500 кбайт на устройствах хранения. При большом количестве сотрудников, например 1000 человек, эталонные изображения будут занимать 500 Мбайт, а при 10000 – 5 Гбайт, тогда время на идентификацию одного человека, при таких объемах базы данных будет значительным и будет расти с увеличением базы данных эталонов, что не приемлемо по требованиям к СКУД [1].

Цель работы состоит в уменьшении времени идентификации личности в СКУД по радужной оболочке глаза.

Решение задачи. Для решения поставленной задачи предполагается в качестве параметра отличия двух изображений использовать расстояние Хемминга (HD) кода радужки (КР), а в качестве критерия принятия решения использовать статистические критерии обнаружения [2].

В системе идентификации КР обрабатывается следующим образом:

1. В процессе регистрации код радужки обработанный фильтром Габора или DOG-фильтром и сохраняется в базе данных для последующего сравнения.

2. При попытке распознавания, когда в систему поступает изображение радужки, для нее вычисляется код радужки, который сравнивается с каждым кодом в базе данных.

В качестве меры сходства двух радужек используется расстояние Хемминга между N-разрядными бинарными кодами идентифицируемой (I) радужки и регистрированными (R) значениями кодов, хранимых в базе данных [2, 3, 4].

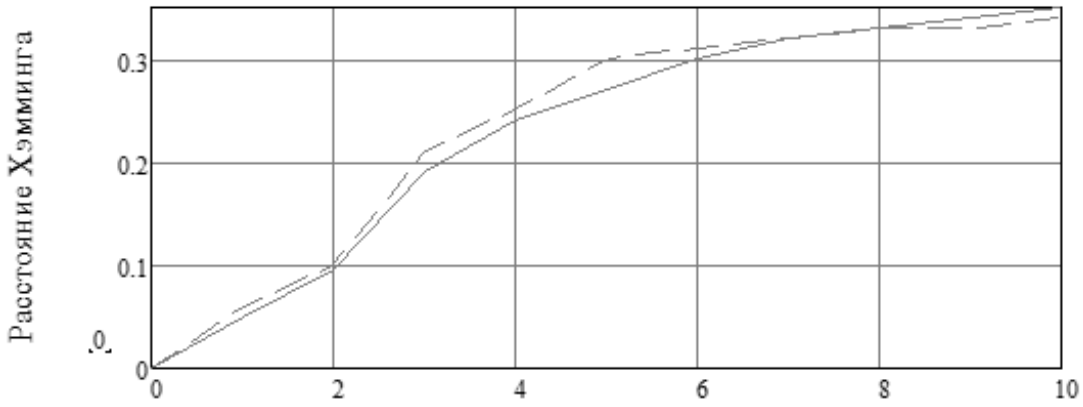
$$HD(IC_I, IC_R) = \frac{1}{N}\sum_{i=1}^{N} IC_{Ii} \oplus IC_{Ri},$$

где IC_I, IC_R – коды радужек, IC_{Ii}, IC_{Ri} – i-й бит кода IC_A.

Для полностью совпадающих КР расстояние Хэмминга будет равно 0. Максимальное значение HD – 1.

При вводе идентифицируемого изображения радужки неизбежно возникают искажения,

вызванные изменением условий освещения (изменения яркости и контраста), поворотом головы, что сопровождается поворотом полученного изображения и его деформацией вдоль осей координат, изменением расстояния до камеры. Алгоритмы принятия решений должны быть устойчивыми к этим искажениям, разумеется, в определенных пределах. Выбором системы признаков и нормализацией изображения радужки удалось компенсировать возможные искажения, кроме искажений поворота. На рис. 1 представлена зависимость расстояния Хэмминга для КР от угла наклона головы, обработанных фильтром Габора и DOG-фильтром.



– фильтр Габора – – DoG-фильтр

Рис. 1. Зависимость расстояния Хэмминга к наклону головы для фильтров Габора $(\omega_0 = \pi/8, \theta = 0)$, и DoG-фильтра $(\sigma = 5)$. Угол показан в градусах

Если предположить, что биты признаков статистически независимы, то математическое ожидание расстояния Хэмминга для различных радужек должно быть равным 0,5. Для одинаковых радужек в отсутствии помех математическое ожидание расстояния Хэмминга равно 0. Под действием помех математическое ожидание расстояния Хэмминга для одинаковых радужек увеличивается и в пределе достигает 0,5.

Выбор углового шага при формировании шаблона радужки является компромиссом. С одной стороны, за счет снижения влияния поворота изображения при уменьшении углового шага уменьшается вероятность принятия ошибочного решения, поскольку увеличивается разность математических ожиданий расстояний Хэмминга для одинаковых и различных радужек. С другой стороны, это ведет к увеличению размера кода радужки, и, как следствие, к увеличению времени идентификации.

Выбор углового шага $\Delta\alpha$ будем рассматривать как оптимизационную задачу, имеющую своей целью минимизацию следующей функции:

$$G(\Delta\alpha) = c_{AR} P_{AR}(\Delta\alpha) + c_{IA} P_{IA}(\Delta\alpha) + c_{ICC} V_{DB} \frac{2\alpha_{\max}}{\Delta\alpha},$$

где $P_{AR}(\Delta\alpha)$ – вероятность запрещения доступа зарегистрированному пользователю, $P_{IA}(\Delta\alpha)$ – вероятность разрешения доступа незарегистрированному пользователю, c_{AR} – стоимость потерь при запрете доступа зарегистрированному пользователю, c_{IA} – стоимость потерь при доступе незарегистрированного пользователя, c_{ICC} – стоимость однократного вычисления расстояния Хэмминга, V_{DB} – нормированное количество пользователей (отношение количества зарегистрированных к максимально возможному количеству пользователей), зарегистрированных в БД.

Зависимости $P_{AR}(\Delta\alpha)$ и $P_{IA}(\Delta\alpha)$ могут быть определены следующим образом. Плотность вероятностей расстояния Хэмминга для КР незарегистрированных лиц остается неизменной, что вытекает из предположения о статистической независимости битов КР. Плотность вероятностей для расстояния Хэмминга зависит от $\Delta\alpha$. При этом к случайной величине HD – расстоянию Хэмминга, для зарегистрированных радужек, прибавляется величина $\Delta HD(\Delta\alpha)$, распределение которой определяется приведенным выше графиком чувствительности. Закон распределения результирующей случайной величины $HD + \Delta HD(\Delta\alpha)$ получается сдвигом исходного закона на величину $\Delta HD(\Delta\alpha)$.

Зная законы распределения вероятностей расстояния Хэмминга для зарегистрированных и незарегистрированных пользователей, можно определить $P_{AR}(\Delta\alpha)$ и $P_{IA}(\Delta\alpha)$.

Принятие решения основано на статистической теории, иллюстрируемой рис. 2, где левая кривая соответствует зарегистрированным радужкам, правая – нарушителям. Каждая из них отображает данные, усредненные по всем экспериментам. В существующих системах порог *C* выбирается одинаковым для всех радужек. Если HD < C, фиксируется совпадение, иначе – несовпадение.

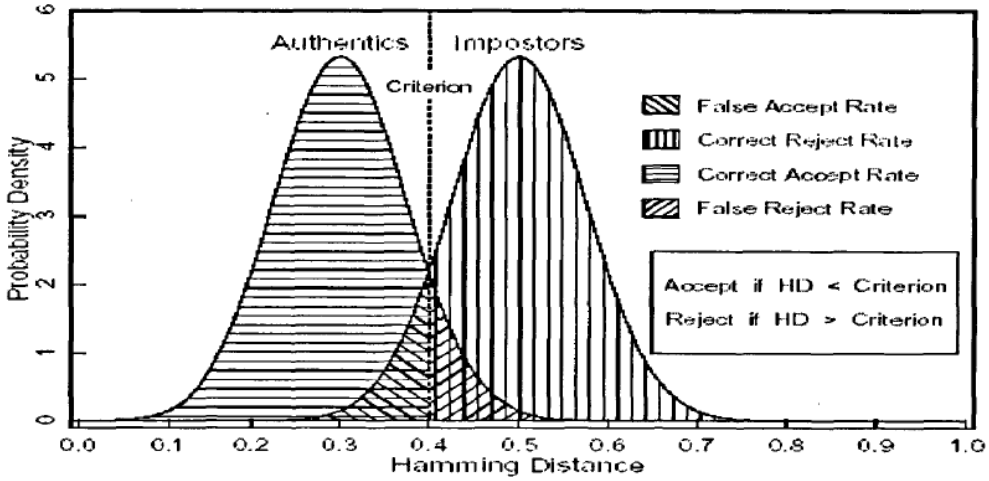


Рис. 2. Условные плотности распределения вероятностей значений РХ

Вероятность принятия ошибочного решения складывается из двух вероятностей:

$$P_{Error} = P_{AR} + P_{IA},$$

где $P_{AR} = \int_c^1 p_{Au}(HD_{\min})dHD_{\min}$ – вероятность запрещения доступа зарегистрированному пользователю, $P_{IA} = \int_0^c p_{\mathrm{Im}}(HD_{\min})dHD_{\min}$ – вероятность разрешения доступа незарегистрированному пользователю, C – значение порога, p_{Au} – вероятность появления на входе системы аутентичной радужки, p_{Im} – вероятность появления на входе системы неаутентичной радужки, $p_{Au}(HD_{\min})$ – условная плотность распределения Хэмминга для аутентичных радужек, $p_{\mathrm{Im}}(HD_{\min})$ – условная плотность распределения Хэмминга для неаутентичных радужек.

При принятии решения о совпадении двух кодов возможно четыре исхода. В двух случаях ответ верный, в двух – нет. Два правильных решения – это разрешение доступа авторизованному человеку и отказ в доступе неавторизованному. Неправильные решения – это разрешение доступа неавторизованному человеку и отказ авторизованному. Рис. 2 является иллюстрацией того, как различные решения связаны между собой. Существуют области, в которых распределения вероятностей РХ для совпадающих и несовпадающих КР перекрываются. В этой области при распознавании могут происходить ошибки [5,6].

Из рис. 2 видно, что качество принятия решения определяется величиной перекрытия распределений вероятностей. Качество распознавания улучшается, если расстояние между средними значениями РХ увеличивается, или дисперсии РХ уменьшаются.

Для того чтобы количественно охарактеризовать разделимость двух классов может быть использован Критерий разделимости *d*. Если μ_1 и μ_2 – математические ожидания, а σ_1, и σ_2 среднеквадратичные отклонения, то

$$d = \frac{|\mu_1 - \mu_2|}{\sqrt{(\sigma_1^2 + \sigma_2^2/2}}.$$

Критерий *d* может использоваться для оценки качества признаков. Чем лучше разделяющая способность признака, тем большее значение принимает *d*.

Методика выбора порога подробно разработана в статистической теории принятия решений. В [7] описаны критерии выбора порога:

1) критерий Байеса;
2) минимаксный;
3) критерий Неймана-Пирсона.

Критерий Байеса обеспечивает минимум среднего риска при принятии решения. Применение критерия Байеса целесообразно, когда система распознавания многократно осуществляет распознавание в условиях неизменного признакового пространства, при стабильном описании классов и неизменной платежной матрице. Решение об идентичности радужек при использовании Байесовской стратегии принимается, если

$$c_{IA}P_{Im}\,p_{Im}\left[HD_{min}(IC_1,IC_2)\right] \succ c_{AR}P_{Au}\,p_{Au}\left[HD_{min}(IC_1,IC_{2)}\right],$$

где $p_{\mathrm{Im}}\left[HD_{\min}(IC_1,IC_2)\right]$ – условная плотность вероятностей расстояния Хэмминга для различных радужек, $p_{Au}\left[HD_{\min}(IC_1,IC_{2)}\right]$ – условная плотность вероятностей расстояния Хэмминга для одинаковых радужек, c_{AR} – стоимость потерь при запрете доступа зарегистрированному пользователю, c_{IA} – стоимость потерь при доступе незарегистрированного пользователя, P_{Au} – вероятность того, что радужки идентичны, P_{Im} – вероятность того, что радужки различны.

Для применения критерия Неймана-Пирсона требуется только определение максимальной вероятности разрешения системой доступа для незарегистрированного лица – $P_{IA\max}$. Применяя критерий Неймана-Пирсона, значение порога *C* может быть определено из уравнения:

$$\int_0^C p_{\mathrm{Im}}(HD_{\min})dHD_{\min} = P_{IA\max}\,.$$

Если априорные вероятности событий, заключающихся в сопоставлении идентичных и различных радужек оценить невозможно, то может быть применен минимаксный критерий. Минимаксный критерий обеспечивает минимум максимальных потерь. Пороговое значение расстояния Хэмминга при использовании минимаксного критерия находится из уравнения

$$C_{AR}\int_C^1 p_{Au}(HD_{\min})dHD_{\min} = C_{IA}\int_0^C p_{lm}(HD_{\min})dHD_{\min}.$$

Все три критерия в зависимости от ситуации могут быть использованы в системе идентификации личности. Исследования законов распределения расстояний Хэмминга в присутствии случайных искажений показали, что для различных радужек законы распределения могут отличаться. На рис. 3 приведены полигоны частот для различных радужек. Полигоны построены следующим образом. Из исходного изображения с помощью модели случайных искажений формировалось множество искаженных изображений. Определение расстояний Хэмминга для каждой пары искаженных изображений одной и той же радужки дает оценку закона распределения расстояний Хэмминга для радужек одного человека. Сравнение каждого из искаженных изображений с множеством изображений других радужек дает оценку закона распределения расстояний для радужек нарушителей.

При этом использовалась следующая модель искажений:

$$I(x,y) = C(S(R(I(x,y))) + N(x,y),$$

где *S* – преобразование масштабирования, *R* – поворот, $I(x,y)$ – изображение, N – гауссов шум, C – преобразование яркости и контраста.

Как видно из рис. 3 распределения вероятностей для различных радужек существенно отличаются. Причина этого в свойствах структуры изображения радужной оболочки (наличие ярко выраженных особенностей, их количество и т.д.). Повышение эффективности может быть достигнуто путем определения индивидуального порога для каждой радужки из базы эталонов. Определение индивидуального порога происходит в измененной процедуре регистрации. Машинными методами моделируется ввод большого числа измененных радужек, а обработанные данные могут быть представлены на графике, аналогичном рис. 2, но обе кривые соответствуют не всем, а единственному классу. Полученное по результатам моделирования значение порога сохраняется в базе данных вместе с кодом и используется в процедуре идентификации.

Таким образом, теперь в базе данных будут храниться (при критерии Байеса) эталонное изображение радужки и расстояние Хэмминга, и порог обнаружения. При критерии Неймана-Пирсона расстояние Хэмминга и порог обнаружения. При минимаксном критерии минимальный и максимальный порог обнаружения и расстояния Хэмминга.

Если порог обнаружения и расстояние Хэмминга кодировать в формате с плавающей запятой, то на одно значение понадобится 4 байта (32 бита).

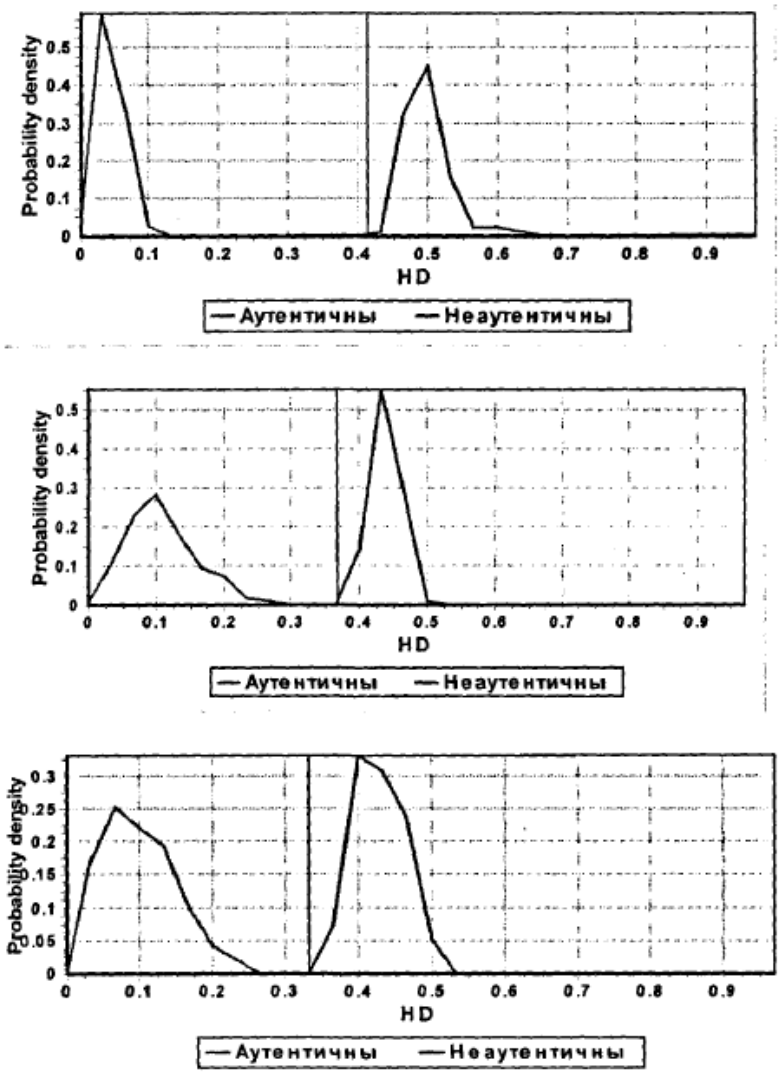


Рис. 3. Полигоны частот для радужек принадлежащих различным людям (значения порогов, определенные по критерию Неймана-Пирсона показаны черной сплошной линией)

Подсчитаем объем данных при использовании каждого критерия в СКУД на одного идентифицируемого.

1. Критерий Байеса – эталонное изображение радужки 500 кб, а также порог обнаружения 4 байта, 4 байта расстояние Хемминга, итого примерно 500 кб (8 байтов по сравнению с 500 кб незначительный прирост объема данных).

2. Критерий Неймана-Пирсона – 4 байта расстояние Хэмминга и 4 байта порог обнаружения (эталонное изображение радужки иметь при использовании этого критерия не нужно), итого 8 байт.

3. Минимаксный критерий- два порога 8 байтов и 4 байта расстояние Хэмминга всего 12 байт.

Из подсчетов можно сделать следующие заключение, что использование только двух последних критериев обнаружения дает значительный выигрыш в объемах баз данных (табл. 1), а значит и во времени обработки на идентификацию одного человека.

Таблица 1

Зависимость объема базы данных от статистического критерия

Критерий	Объем базы данных на 1 человека
Байеса	512008 байт
Неймана-Пирсона	8 байт
Минимаксный	12 байт

Выводы. В работе проанализированы статистические характеристики кода радужки в зависимости от способа обработки фильтром Габора и DOG-фильтром. Результаты обработки и вычисления расстояния Хэмминга практически идентичны (рис. 1). Получены полигоны частот (плотности распределения) расстояния Хэмминга для различных людей. Рассмотрены и опробованы статистические критерии обнаружения, повышающие скорость обработки изображений радужки и уменьшающих объем баз данных систем контроля и управления доступом.

Все известные исследования не ставили цель уменьшить объем базы данных и повысить

быстродействие работы с ней. Полученные результаты показывают, что возможны направления по уменьшению объема базы данных что влечет за собой одно из направлений – повышение быстродействия СКУД(табл. 1).

ЛИТЕРАТУРА

[1]. Автоматическая идентификация. Идентификация биометрическая. Форматы обмена биометрическими данными: ГОСТ Р ИСО/МЭК 19794-6-2006-2007, часть 6. Данные изображения радужной оболочки глаза. – 28 с.

[2]. Christel-Loic Tisse, Lionel Martin, Lionel Torres, Michel Robert. Person Identification Technique Using Human Iris Recognition. Vision Interface (VI2002), Canadian Image Processing and Pattern Recognition Society (CIPPRS), 2002, 15th International Conference on Vision Interface, pp. 294-299.

[3]. Orval E Phelps, Captain, USAF, Information Security: securing smart cards with iris Recognition, Department of the air force Air university. Air Force Institute Of Technology, March 2001.

[4]. John Daugman, Demodulation by Complex-valued Wavelets for Stochastic Pattern Recognition, The Computer Laboratory, University of Cambridge, Cambridge, UK, January 14, 2003.

[5]. John Daugman, Biometric Decision Landscapes: technical report, Number 482, University of Cambridge, The Computer Laboratory, 2000, 13 p. - Electronic Resource: https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-482.pdf

[6]. Горелик А. Л., Скрипкин В. А. Методы распознавания: Учеб. пособие для вузов.- 3-е изд., перераб. и доп. - М.: Высш. шк., 1989. – 232 с

[7]. Тихонов В. И. Статистическая радиотехника: Учебное пособие / В. И. Тихонов. – М.: Сов. Радио, 1966. – 678 с.

REFERENES

[1]. Automatic identification. The biometric identification. Exchange formats biometric data: GOST R ISO/IEC 19794-6-2006, 2007, part 6. Image data iris, 28 p.

[2]. Christel-Loic Tisse, Lionel Martin, Lionel Torres, Michel Robert. Person Identification Technique Using Human Iris Recognition. Vision Interface (VI2002), Canadian Image Processing and Pattern Recognition Society (CIPPRS), 2002, 15th International Conference on Vision Interface, pp. 294-299.

[3]. Orval E Phelps, Captain, USAF, Information Security: securing smart cards with iris Recognition, Department of the air force Air university. Air Force Institute Of Technology, March 2001.

[4]. John Daugman, Demodulation by Complex-valued Wavelets for Stochastic Pattern Recognition, The Computer Laboratory, University of Cambridge, Cambridge, UK, January 14, 2003.

[5]. John Daugman, Biometric Decision Landscapes: technical report, Number 482, University of Cambridge, The Computer Laboratory, 2000, 13 p. - Electronic Resource: https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-482.pdf

[6]. Gorelik, A.L. Skripkin, V.A. Recognition methods: tutorial for universities, 3rd ed., rev. and add., M .: Higher. wk., 1989, 232 p.

[7]. Tikhonov, V. I. Statistical radio engineering: textbook, M.: Soviet Radio, 1966, 678 p.

КРИТЕРІЙ ВИЯВЛЕННЯ ЯК ВПЛИВОВИЙ ФАКТОР ОБСЯГУ БАЗИ ДАНИХ БІОМЕТРИЧНИХ СИСТЕМ КОНТРОЛЮ ДОСТУПУ

В статті для підвищення швидкодії обробки райдужної оболонки ока і зменшення обсягів баз даних, для зберігання зображення райдужної оболонки, пропонується попередня обробка вхідних зображень DOG-фільтром або фільтром Габора, з наступним обчисленням відстані Хеммінга за кодом райдужної оболонки ока і її статистичних параметрів. Оцінка впливу спотворень зображення райдужної оболонки на статистичні параметри (щільність розподілу ймовірності) відстані Хеммінга. Порівняння відстані Хеммінга двох зображень райдужної оболонки і прийняття рішення їх ідентичності. Обрані статистичні критерії виявлення для зменшення обсягу баз даних, попередньо оброблених зображень райдужної оболонки ока, і підвищення швидкодії ідентифікації і аутентифікації системи контролю та управління доступом по райдужній оболонці ока.

Ключові слова: райдужна оболонка, відстань Хеммінга, статистичні критерії, DOG-фільтр, щільність розподілу.

DETECTION CRITERION AS THE INFLUENCING FACTOR OF THE DATABASE SIZE FOR BIOMETRIC ACCESS CONTROL SYSTEMS

In this study for improving performance of processing the iris and decreasing database capacity, for storing iris images proposed pre-processing input images by DOG-filter or Gabor filter followed by calculation of the Hamming distance by code of the iris and its statistical parameters. Evaluation of distortion impact on the iris image of the eye, on the statistical parameters (density of probability distribution) Hamming distance. Comparisons Hamming distance between two iris images and the decision of their identity. Statistical detection criteria are chosen to reduce the database capacity, pre-processed images of the iris, and speed of identification and authentication controls and access control iris.

Keywords: iris, Hamming distance, statistical criteria, DOG-filter distribution density

Фесенко Андрей Алексеевич, ассистент кафедры кибербезопасности и защиты информации КНУ имени Тараса Шевченко.
E-mail: a.fesenko@meta.ua

Фесенко Андрій Олексійович, асистент кафедри кібербезпеки та захисту інформації КНУ імені Тараса Шевченко.

Fesenko Andrey, Assistant Teacher of Cybersecurity and Information Security, Taras Shevchenko National University of Kyiv