

ВИБІР МЕТОДИКИ ТА РОЗРАХУНОК КОЕФІЦІЄНТІВ ПОМИЛКОВОГО ПРОПУСКУ ТА ПОМИЛКОВОЇ ВІДМОВИ ДОСТУПУ У СИСТЕМАХ БІОМЕТРИЧНОЇ ІДЕНТИФІКАЦІЇ

У практиці обробки зображень завдання пошуку відповідності за біометричними даними отримало велике поширення та відоме як проблема «пошуку за зразком». Формально проблема розглядається як процес ототожнення еталонного зображення з одним з множини образів фрагментів, що лежать у заданій області другого зображення. Алгоритми встановлення подібності в основних варіантах тісно чи іншою мірою пов'язані з отриманням характеристик стохастичного взаємозв'язку порівнюваних фрагментів зображень. Усі вони ґрунтуються на ідеях кореляційної та спектральної теорії сигналів. Однією із мір подібності зображень, яка характеризує якість роботи біометричної системи ідентифікації, є визначення коефіцієнтів помилкового пропуску та помилкової відмови доступу. Одна з таких меток розглядається у статті.

FAR, FRR, БІОМЕТРИЧНА ІДЕНТИФІКАЦІЯ, ЗОБРАЖЕННЯ, МІРА ПОДІБНОСТІ, СТАТИСТИЧНА ВЛАСТИВІСТЬ

Вступ

Завдання аналізу та розпізнавання зображень у системах доступу до технологічних систем та мереж на сьогодні не має достатньо ефективних рішень: точність розпізнавання варіюється від 60 до 70%. Це потребує розробки та синтезу нових методів та алгоритмів аналізу зображень складних об'єктів на основі сучасного програмно-математичного підходу. Для систем ідентифікації, які розпізнають об'єкти на цифровому зображенні, найбільш корисною інформацією є відомості про контури, тобто про лінії, що проходять на границях однорідних областей. Вважається, що такими областями є об'єкти, для яких різниця яскравостей будь-яких двох елементів зображення (пікселів, групи пікселів), не перевищує певного порогу. Тому, по завершенні попередньої обробки зображення, система розпізнавання, в першу чергу, робить пошук контурів зображення. За результатами пошуку послідовним порівнянням отриманих результатів з інформацією, яка є у відповідних базах даних, визначається ідентичність двох зображень. На основі цього приймається управляюче рішення щодо надання дозволу на доступ об'єкта ідентифікації до систем з обмеженим доступом.

На жаль, біометричні технології ідентифікації, наприклад, по обличчю або по його термограмі, надзвичайно чутливі до зовнішніх умов, тобто освітленості, поворотів голови, кутів її нахилу і т.п. Це призводить до того, що така технологія характеризується низьким відсотком успішного розпізнавання користувачів та найвищим відсотком помилкових спрацьовувань. У цьому випадку для якості її роботи використовують два критерії:

1. FAR (англ.: *False Acceptance Rate*, FAR) – коефіцієнт помилкового доступу, який є процентним показником випадків, при яких перевірка особи виявилася помилково успішною.

2. FRR (англ.: *False Rejection Rate*, FRR) – коефіцієнт помилкової відмови в доступі, який є процентним показником випадків, при яких перевірка особи помилково завершилася невдачею.

Т.ч., виходячи зі сказаного, вибір методики та розрахунок FAR та FRR для систем біометричної ідентифікації, є актуальною науковою та практичною проблемою.

Проблема автоматичного аналізу форми та стану просторових об'єктів, інформація про які представлена у вигляді пласких зображень, є актуальною для систем забезпечення превентивної безпеки. Зазначеній проблемі достатньо приділяли увагу зарубіжні вчені, включаючи вчених країн СНД, а саме: Т. Kanade та В. Lucas (Канада), У. Претт, М. Каас, А. Witkin, D. Terzopoulos, S. Arulampalam, R. Gonzalez, R. Woods (США), М. Rachuta, В. Wilamowski, А. Malinowski (Польща), Б. Залеський, О. Ферцев, О. Кравченко, Я. Фурман, А. Кравецький, Р. Хафізов, Н. Соловійов, О. Сергеев, М. Красильников (Росія) та ін. у роботах який не останню позицію займали питання визначення та розрахунку розрахунок FAR та FRR.

В Україні багато організацій також працює в області цифрової обробки біометричних сигналів. Серед цих організацій є три потужних колективи, що проводять повний цикл робіт у галузі розпізнавання, починаючи від комплексних досліджень проблеми та закінчуючи виготовленням дослідних і серійних зразків спеціальної апаратури. До таких колективів відносяться Інститут кібернетики та Фізико-механічний інститут ім. Г. В. Карпенка НАН України, Інститут математики РАН та Інститут технічної кібернетики АН Білорусії. В першу чергу, необхідно відзначити дослідників Фізико-механічного інституту ім. Г. В. Карпенка НАН України, та серед них – дослідження Б. Русина, які присвячені розробці нових підходів до побудови біометричних систем аутентифікації та криптографічного захисту, розробці нової системи інформативних ознак зображень людини та методи їх ідентифікації.

Виходячи з того, що алгоритми встановлення подібності зображень, які в основних варіантах пов'язані з отриманням характеристик стохастичного взаємозв'язку порівнюваних фрагментів зображень, поточною *метою* є показ однієї з методик визначення коефіцієнтів помилкового

пропуску та помилкової відмови доступу, що характеризують якість роботи біометричної системи ідентифікації.

1. Передумови та вибір методики розрахунку FAR та FRR

Як відзначено у численних літературних джерелах, включаючи [1], процес біометричної ідентифікації складається з комбінації викликів двох базисних функцій нижнього рівня програмного інтерфейсу:

1) створення біометричного шаблону з біометричного зразка;

2) порівняння біометричних шаблонів.

Проведені попередні дослідження дозволили встановити, що результатом біометричного порівняння є число, тобто міра подібності: чим вищою вона є, тем вищою є ймовірність того, що пред'явлені для порівняння зразки належать одній людині. Згідно до [1], вхідними даними біометричного порівняння є пара цифрових біометричних шаблонів і міра подібності шаблонів. Залежність міри подібності від вхідних шаблонів і біометричних зразків – величина детермінована. У реальній експлуатації систем розпізнавання процес отримання зразків біометричних характеристик залежить від зовнішніх факторів, що привносить свої погрешності (наприклад, забруднені ділянки, вологість, температура, дефекти поверхні пристрою сканування, різноманітні шуми і т.д.). Отже, виходячи з цього, для моделювання невизначеності таких факторів доцільним є використання ймовірнісно-статистичних методів. При цьому вихідний сигнал біометричного порівняння повинен бути випадковим.

В [1] показано, що біометричне порівняння характеризується двома розподілами міри подібності $f^{gen}(x)$ в «своїх» порівняннях, коли отримані зразки біометричної характеристики однієї людини, і $f^{imp}(x)$ – в «чужих» порівняннях, коли порівнюються біометричні зразки різних людей.

Як вже зазначалося, показники ефективності біометричних систем, це FAR та FRR. При цьому FAR є ймовірністю помилки 2-го роду, а FRR – ймовірністю помилки 1-го роду. Помилки FAR та FRR в ймовірнісній інтерпретації при фіксованому порозі t ухвалення рішення можна обчислити по формулах:

$$FRR(t) = \int_{-\infty}^t f^{gen}(s) ds ; FAR(t) = \int_t^{+\infty} f^{imp}(s) ds .$$

Якщо розглядати біометричну систему цілком, включаючи спотворюючі фактори, то вона є стохастичною, але в ній існує детермінована підсистема.

Розглянемо загальний випадок мультибіометричної ідентифікації та можливості її використання для розв'язку завдань, винесених у заголовки роботи.

При мультибіометричній ідентифікації вихідні сигнали біометричних порівнянь інтегруються у мультибіометричну міру подібності λ , яку, як правило, вважають

функцією G відгуків біометричних порівнянь, тобто $\lambda = G(x_1, \dots, x_n)$, де x_i – результат i -го біометричного порівняння. Функція G повинна бути скалярною, тому що мультибіометрична система підпадає під вимоги до біометричних систем у цілому. Відповідно, при ідентифікації можна порівнювати λ з деяким порогом і ухвалювати рішення про ідентифікацію.

Якщо входи мультибіометричного порівняння x_i мають стохастичний характер, то це вірно й для результату $\lambda = G(x_1, \dots, x_n)$. Т.ч., мультибіометричну ідентифікацію можна розглядати як систему, де за вхідні сигнали приймаються відгуки біометричного порівняння, а за вихідні – результуюча міра подібності. Вагова функція G повністю визначає властивості такої моделі.

При розробці технології мультибіометричної ідентифікації слід накласти певні обмеження на G з метою максимізації цільового критерію якості мультибіометричної ідентифікації. Типовими критеріями якості можна обрати різні співвідношення помилок 1-го, 2-го роду та часу мультибіометричного порівняння.

2. Визначення статистичних властивостей біометричних порівнянь

Якщо розподіли біометричних порівнянь відомі та визначені спільними щільностями $f^{gen}(x_1, \dots, x_n)$ та $f^{imp}(x_1, \dots, x_n)$, то статистика $\lambda = \ln f^{gen} - \ln f^{imp}$ дає мультибіометричну міру подібності з мінімальними помилками ідентифікації 1-го та 2-го роду [2, 3]. Відповідно, λ також максимізує довільний функціонал якості, монотонний по помилках 1-го та 2-го роду. Розв'язок завдання може бути знайдено у такий спосіб [1]:

$$G(x_1, \dots, x_n) = \ln f^{gen}(x_1, \dots, x_n) - \ln f^{imp}(x_1, \dots, x_n) .$$

Т.ч., завдання визначення оптимальної міри подібності можна звести до завдання оцінювання щільностей «своїх» та «чужих» біометричних порівнянь. Але на етапі навчання мультибіометричної системи (тобто на етапі визначення розподілів біометричних порівнянь), як правило, доступна досить обмежена інформація для оцінки статистичних властивостей біометричних систем. Це пояснюється наступним:

1) При використанні емпіричних частот як оцінок дійсних функцій розподілу, спостерігається сильна залежність від навчальної вибірки та значна дисперсія результатів навчання.

2) Слід враховувати, що мультибіометричні технології застосовують для отримання прийнятної якості роботи системи розпізнавання на дуже низьких рівнях помилки 2-го роду – від 10^{-6} . Відповідно, досить непростою проблемою є верифікація результатів. При цьому дуже важливим частковим завданням інтеграції вважається екстраполяція помилок 1-го та 2-го роду на значення, які неможливо перевірити в ході випробувань

технічних пристроїв розпізнавання. Наприклад, у системах з великим числом користувачів необхідно забезпечити помилку 2-го роду на рівні 10^{-9} і менше. У такому випадку для перевірки результатів необхідна база даних обсягом приблизно 10^9 записів, якої на сьогоднішній день у відкритому доступі не знайдено.

3) Навіть при прийнятті певних допущень про динаміку помилок розпізнавання, дисперсія прогнозу зі зменшенням рівня FAR росте неприйнятними темпами. Аналогічна ситуація спостерігається для більшості мультибіометричних технологій [1, 15]. Основна причина цього полягає у нездатності емпіричних щільностей до узагальнення на генеральну сукупність. Зі зменшенням навчальної вибірки довірчий інтервал для щільностей розширюється. Відповідно, якість ідентифікації буде менш передбачуваною, що показано у роботі [2].

4) Для багатьох комбінацій біометричних характеристик недоступні мультибіометричні дані, що принципово не дозволяє проводити оцінку багатомірних щільностей. На сьогодні доступні кілька десятків одноmodalних біометричних баз. При цьому існуючих мультибіометричних баз даних явно недостатньо для дослідження якості ідентифікації та навчання у мультибіометричних технологія. Тому важливою проблемою є використання навчальної інформації з окремих каналів. Ця ідея була покладена в основу отримання результатів при розрахунках FAR та FRR. Тут відзначимо, що одноmodalних біометричних баз, які містять термограми людей, у відкритому доступі не знайдено. З метою моделювання та отримання практичних результатів використано штучно створену базу даних обсягом 2^{16} об'єктів, що при розрахунках дозволило забезпечити помилку 2-го роду на рівні 10^{-5} .

5) Використання емпіричних частот приводить до виникнення простору, який не параметризується, що підвищує ймовірність неадекватного навчання. З врахуванням викладеного в п. 4, відзначимо, що результати випробувань, які виконувалися при проведенні досліджень (числові та графічні дані наведені далі), не дозволяють з повною впевненістю стверджувати про статистичну незалежність результатів біометричних порівнянь. Для підтвердження або спростування даного положення необхідно мати величезний тестовий масив біометричних вимірів, якого на даний момент не має жодна організація. Однак можна припустити, що біометричні термохарактеристики людини, які досліджувалися, є незалежними і, отже, біометричні порівняння також є незалежними. Т.ч., якщо результати незалежні, то щільності розподілів у «своїх» та «чужих» порівняннях факторизуються у такий спосіб:

$$f^{gen}(x) = f_1^{gen}(x_1) \cdot \dots \cdot f_n^{gen}(x_n);$$

$$f^{imp}(x) = f_1^{imp}(x_1) \cdot \dots \cdot f_n^{imp}(x_n),$$

де $f_i^{gen}(x_i)$ та $f_i^{imp}(x_i)$ – щільності розподілу i -го біометричного тесту.

Вираз для оптимальної міри подібності для незалежних біометричних характеристик можна представити у такій формі:

$$G(x) = \ln f^{gen}(x) - \ln f^{imp}(x) =$$

$$= \sum_{i=1}^n (\ln_i^{gen}(x_i) - \ln_i^{imp}(x_i)) = \sum_{i=1}^n g_i(x_i),$$

де $g_i(x_i) = \ln_i^{gen}(x_i) - \ln_i^{imp}(x_i)$.

З наведеного виразу видно, що підсумкову міру подібності можна обчислити через функцію, що залежить тільки від властивостей рівно одного біометричного порівняння. Отримавши цей результат, його можна застосувати для розрахунків FAR та FRR. При цьому врахуємо рекомендації [4], де показано, що якість розпізнавання може бути поліпшена з ростом числа зразків, взятих при навчанні, а це – в ідеалі – може привести до досягнення деякої граничної якості розпізнавання, що задається параметром r_{nop} . Останнє пов'язане з тим, що при зростанні числа зразків, особливо отриманих у різний час, компенсується негативний вплив зовнішніх факторів [5]. З погляду правила ухвалення рішення, зразки рівноцінні. Отже, на щільності розподілу та на вирішальні правила слід накласти умову симетричності по змінним, тобто:

$$\forall i, j \quad f^{gen}(x_1 \dots x_i \dots x_j \dots x_n) = f^{gen}(x_1 \dots x_j \dots x_i \dots x_n),$$

$$\forall i, j \quad f^{imp}(x_1 \dots x_i \dots x_j \dots x_n) = f^{imp}(x_1 \dots x_j \dots x_i \dots x_n).$$

3. Методика розрахунку FAR та FRR

FRR (помилкова тривога) – імовірність помилок 1-го роду – не настільки критична для системи безпеки, хоча і створюють незручності, оскільки доводиться проходити верифікацію вразі. Від FAR (пропуск події) – імовірність помилок 2-го роду, – в результаті яких зловмисник може отримати доступ до системи, залежить надійність системи захисту від несанкціонованого доступу. Поява помилок FRR та FAR визначається такими характеристиками, як якість і роздільна здатність системи реєстрації, розмір області фіксації (сканування) об'єкта, математичні алгоритми, використовувані для порівняння зображень (контурів), кількість деталей, які застосовуються для порівняння.

Як правило, частота виникнення помилок FRR вище частоти виникнення FAR-помилки. Так, згідно [6], ймовірність виникнення FRR-помилки, у середньому, становить менше 2%, а ймовірність FAR-помилки – менше 0,0001%.

З метою перевірки результатів щодо розпізнавання біометричних контурів термограми обличчя людини на предмет відповідності зазначеному, покажемо методику проведення розрахунків так, як це зазначено нижче.

Дотримання умов вірогідності виявлення співпадіння потребує встановлення порогу для величини взаємної кореляції $\max \tilde{r}(k,l)$. Якщо $\max \tilde{r}(k,l) \geq r_{\text{пор}}$, то з заданою ймовірністю гарантується дійсна подібність знайденої пари фрагментів. Отже, у якості змінної величини можна обрати $r_{\text{пор}} \in [\Delta_i; \Delta_{i+1}; \dots; \Delta_n]$, де $n=10$. Крок зміни Δ можна встановити рівним 0,1: у багатьох публікаціях зазначається, що така величина є достатньою для оцінки динаміки змін FAR та FRR при зміні $r_{\text{пор}}$ та дозволяє з достатньою точністю візуально відобразити отримані результати. Точно розраховані значення FAR та FRR можуть бути відображені у вигляді таблиць. Зазначимо, що разом зі знаходженням значень FAR та FRR, доцільним є визначення значення FTE. При цьому, відповідно до [7], FTE – це «Помилка реєстрації» (англ.: *Failure to Enroll Rate*, FTE). У числовому значенні – це відсоток об'єктів, які не володіють можливістю зареєструватися в системі. FTE може бути визначена для кожного об'єкта окремо. Якщо об'єкт

взагалі не може бути зареєстрований у системі або після реєстрації не може бути розпізнаний, такий випадок відноситься до ймовірності помилки збору даних (англ.: *Failure to Acquire Rate*, FTR).

FTE розраховується для кожного об'єкта окремо – FTE(n) – як кількість неуспішних спроб реєстрації стосовно загального числа спроб реєстрації. Для отримання загального FTE всі персональні показники можуть бути зведені до середнього значення:

$$\overline{\text{FTE}} = \frac{1}{N} \sum_{n=1}^N \text{FTE}(n).$$

З метою моделювання та отримання практичних результатів була використана штучно створена база даних [FTE(n)], де $n = 2^{16}$ об'єктів. Вмістом [FTE(n)] були штучно спотворені бінарні зображення нормовані у базисі площини 512×512 пікселів. Еталонні зображення об'єктів приведено на рис. 1. Спотворення моделювалися афінними перетвореннями [8].

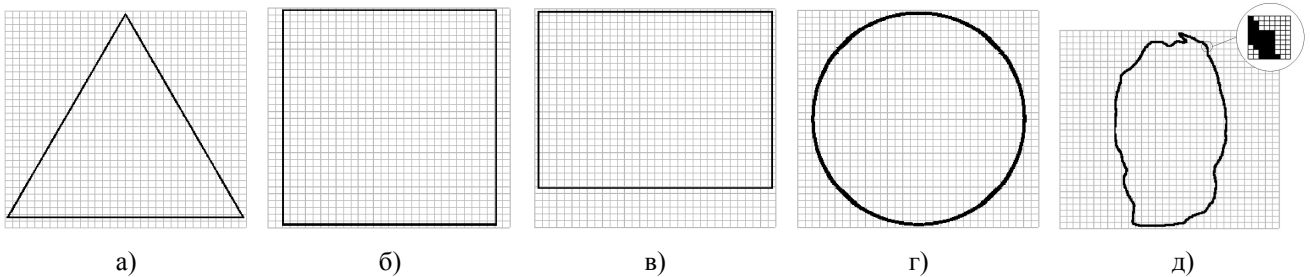


Рис. 1. Приклади фігур у вигляді бінарних зображень нормованих у базисі площини 512×512 пікселів (тут обмежено до площини 256×256 пікселів): а) трикутник; б) квадрат; в) прямокутник; г) коло; д) нормована фігура (контур термозображення обличчя людини)

Отримання кожного значення FTE(n) проводилося методом визначення усередненого показника взаємної кореляції кожного еталонного об'єкту заданого типу з кожним афінно-спотвореним значенням з ансамблю фігур [FTE(n)] при кількості повторів $K_{\text{повт}} = 4$. При цьому було показано, що використання $K_{\text{повт}} > 4$ не веде до суттєвого покращення отриманих даних (рис. 2). Повтором вважалася дія порівняння еталонного зразка з відповідним об'єктом з загальної бази даних.

Кожен об'єкт, отриманий методом афінних спотворень, перед занесенням у базу даних, був оброблений методом компенсації спотворення геометрії зображення. При ідентифікації позитивним вважалася таке значення, яке відповідало вимогам, що наведені у строчці 5 табл. 1. У такому разі FTE(n) вважалася таким, що дорівнює «1».

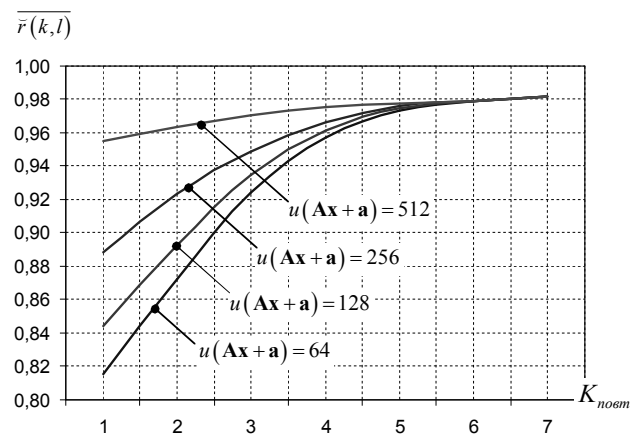


Рис. 2. Залежність усередненого показника взаємної кореляції ансамблів довільних фігур у залежності від кількості повторів $K_{\text{повт}}$

Таблиця 1 [9]

Словесний опис величини коефіцієнта кореляції

№	Значення коефіцієнта кореляції $\bar{r}(k, l)$	Інтерпретація
1	$0 < \bar{r}(k, l) \leq 0,2$	Надзвичайно слабка кореляція
2	$0,2 < \bar{r}(k, l) \leq 0,5$	Слабка кореляція
3	$0,5 < \bar{r}(k, l) \leq 0,7$	Середня кореляція
4	$0,7 < \bar{r}(k, l) \leq 0,9$	Сильна кореляція
5	$0,9 < \bar{r}(k, l) \leq 1,0$	Дуже сильна кореляція

З метою повного розуміння суті питання та для пояснення застосування вище використаних позначень, відзначимо, що еталонне зображення A (або його вибраний фрагмент) представлявся матрицею U_0 розміром $n \times n$, та порівнювався з зображенням B (або його вибраним фрагментом) в «зоні пошуку» Ω розміром $L \times L$, де $L = m + n$. Перекриття між зображеннями визначалися кроком h дискретних ґрат hZ^2 у площині P^2 на яких були задані спостережувані змінні $\{u_0(\mathbf{x}), \mathbf{x}=(x, y)\}$ на зображенні A або $\{u(\mathbf{x})\}$ на зображенні B .

У процесі ковзного пошуку, коли кожний черговий фрагмент отримувався з попереднього простим зрушенням на один дискрет (піксель), обчислювалася «функція подібності» між зображенням еталонного фрагмента $\{u_0(\mathbf{x}), \mathbf{x} \in \Gamma_A\}$ та зображеннями поточних (контрольованих) фрагментів $\{u(\mathbf{x}), \mathbf{x} \in \Gamma_B\}$. Метою було знаходження функції подібності, яка б з максимально можливою точністю та вірогідністю дозволяла локалізувати фрагмент, що відповідає еталонному фрагменту, фіксуючи в такий спосіб сумісні точки на зображеннях.

Взаємно відповідні елементи одного об'єкта на зображеннях задовольняли співвідношенню:

$$u_0(x, y) = (au(x+k, y+l) + b) \text{rect}\left(\frac{x}{n}, \frac{y}{n}\right) + \varepsilon(x, y), \quad (1)$$

де a та b – параметри контрасту та освітленості; k та l – параметри відносного зрушення зразка та його аналога на контрольованому зображенні; $\varepsilon(x, y)$ – шум;

$$\text{rect}\left(\frac{x}{n}, \frac{y}{n}\right) = \begin{cases} 1 & \text{при } x \leq n, y \leq n; \\ 0 & \text{в інших випадках.} \end{cases}$$

У такому формулюванні, як це показано у вигляді (1), процедура селекції зразка повинна була знайти параметри k та l , які характеризують зрушення реперних фрагментів. Для простоти вважали, що параметр b не міняється по полю зображень, що дозволило перейти до центрованих змінних:

$$\tilde{u}(x, y) = u(x, y) - \bar{u}, \bar{u} = \frac{1}{L^2} \sum_{(x, y) \in \Omega} u(x, y);$$

$$\tilde{u}_0(x, y) = u_0(x, y) - \bar{u}_0, \bar{u}_0 = \frac{1}{n^2} \sum_{x, y=1}^n u_0(x, y).$$

У якості міри відмінності в точці (k, l) було обрано середньоквадратичну помилку:

$$\varepsilon_a^2(k, l) = \sum_x \sum_y [\tilde{u}_0(x, y) - a\tilde{u}(x+k, y+l)]^2, \quad (2)$$

яка мінімувалася прямим перебором усіх можливих зрушень еталона по заданій області контрольованого зображення. Вважалося, що в точці екстремума реалізується подібність, якщо $\varepsilon_a^2(k, l) \leq \lambda$, де λ – деякий установленний поріг.

З вимоги мінімуму помилки $\varepsilon_a^2(k, l)'_a = 0$ була знайдена оцінку a і, після підстановки її у формулу (2), з [1] отримано вираз:

$$\varepsilon_a^2(k, l) = \sum_x \sum_y [\tilde{u}_0(x, y)]^2 - \frac{\left[\sum_x \sum_y [\tilde{u}_0(x, y)\tilde{u}(x, y)] \right]^2}{\sum_x \sum_y [\tilde{u}(x, y)]^2}. \quad (3)$$

Перший член з (3) – «енергія» еталонного сигналу. Він є величиною постійною, яка не залежить від параметрів зрушення (k, l) . Т.ч., точка екстремуму не змінювалася, якщо виконувалося нормування середньоквадратичної помилки до енергії еталона

$$\varepsilon_a^2(k, l) = 1 - \frac{\sum_x \sum_y [\tilde{u}_0(x, y)\tilde{u}(x, y)]^2}{\sum_x \sum_y [\tilde{u}(x, y)]^2 \sum_x \sum_y [\tilde{u}_0(x, y)]^2},$$

і замість мінімуму нормованої середньоквадратичної помилки знаходився максимум коефіцієнта кореляції поточного фрагмента з еталонном, тобто:

$$\bar{r}(k, l) = \frac{\sum_x \sum_y \tilde{u}_0(x, y)\tilde{u}(x, y)}{\left\{ \sum_x \sum_y [\tilde{u}(x, y)]^2 \sum_x \sum_y [\tilde{u}_0(x, y)]^2 \right\}^{0,5}}.$$

Було підтверджено відоме положення про те, що дотримання умов вірогідності виявлення співпадіння приводить до необхідності встановлення порогу для величини взаємної кореляції $\max \bar{r}(k, l)$: якщо $\max \bar{r}(k, l) \geq r_{\text{пор}}$, то з заданою ймовірністю гарантується дійсна подібність знайденої пари фрагментів. Величина порогу визначена функцією розподілу коефіцієнта кореляції (при випадкових вибірках) та заданою довірчою ймовірністю ухвалення рішення про дійсну подібність фрагментів.

До цього моменту з'ясувався лише сам факт існування статистичної залежності між двома ознаками. Далі з'ясуємо, які висновки можна зробити про силу чи слабкість цієї залежності, а також про її вигляд та спрямованість.

Враховуючи вище зазначене, спочатку були проведені розрахунки \overline{FTE} . Втім, при встановлених обмеженнях, було отримано значення $\overline{FTE} = 0$. Зміна r_{nop} у сторону його більш м'яких вимог надала можливості встановити, що \overline{FTE} дійсно змінюється та не дорівнює «0». Було встановлено, що візуальне відображення залежності \overline{FTE} від r_{nop} у вигляді графіку не має достатньої наочності. Відповідно, для отримання підтвердження методики розрахунку \overline{FTE} , за допомогою графічного редактора в ансамблі $[FTE(n)]$ були штучно

спотворені $n_{шт.спотв} = 29$ об'єктів, які за попередньо отриманими даними, мали параметр $\overline{r}(k,l) \geq 0,9$. Результати розрахунку $FTE(n)$, що були отримані для всіх еталонних об'єктів, показаних на рис. 1, наведено у табл. 2. При цьому, з врахуванням штучних спотворень, було отримане значення $\overline{FTE} \approx 3$ по всьому ансамблю $[FTE(n)] = 65536$. Це означає, що при виконанні процедури ідентифікації, у середньому 3 об'єктам було відмовлено в доступі як таким, для яких виникла помилка реєстрації, а загальна кількість об'єктів, яким було відмовлено у доступі, відображена у табл. 2, та відповідає $n_{шт.спотв} = 29$, що свідчить про дієздатність технології ідентифікації.

Таблиця 2

Розрахунок $\overline{FTE}(n)$ для фігур рис. 1 з врахуванням компенсації спотворення геометрії зображення

r_{nop}	Рис. 1-а		Рис. 1-б		Рис. 1-в		Рис. 1-г		Рис. 1-д	
	$\overline{r}(k,l)$	$\overline{FTE}(n)$	$\overline{r}(k,l)$	$\overline{FTE}(n)$	$\overline{r}(k,l)$	$\overline{FTE}(n)$	$\overline{r}(k,l)$	$\overline{FTE}(n)$	$\overline{r}(k,l)$	$\overline{FTE}(n)$
0,10	0,9859	1	0,9861	2	0,9865	2	0,9861	3	0,9107	0
0,20	0,9130	3	0,9136	5	0,9134	6	0,9137	5	0,8332	2
0,30	0,7651	0	0,7657	0	0,7658	0	0,7654	0	0,7113	0
0,40	0,7386	0	0,7394	0	0,7390	0	0,7390	0	0,6834	0
0,50	0,6061	0	0,6063	0	0,6065	0	0,6068	0	0,5665	0
0,60	0,5238	0	0,5245	0	0,5241	0	0,5245	0	0,4345	0
0,70	0,3482	0	0,3486	0	0,3486	0	0,3484	0	0,2811	0
0,80	0,2857	0	0,2864	0	0,2865	0	0,2859	0	0,2343	0
0,90	-0,0597	0	-0,0590	0	-0,0594	0	-0,0589	0	-0,1281	0
1,00	-0,0923	0	-0,0916	0	-0,0917	0	-0,0915	0	-0,1134	0

FRR для зареєстрованих у системі об'єктів, так само як і FTE, визначалася для кожного об'єкта окремо, оскільки вона може суттєво різнитися у різних користувачів. Більше того, FRR залежить не тільки від користувача, але й може мінятися з часом: як правило [7], цей показник зменшується в міру того, як користувач навчається працювати з системою. При цьому під поняттям «навчання роботи з системою» мається на увазі стабільне зростання випадків ідентифікації та аутентифікації об'єкта в системі доступу. У зв'язку з цим, при практичному використанні системи ідентифікації на основі біометричних даних, рекомендується використовувати значення FRR для окремо «навчених» і окремо для «ненавчених» користувачів.

Перший етап – розрахунок FRR для кожного користувача окремо при якому $FRR(n)$ обчислювалася, як відношення кількості відмов у доступі до загальної кількості спроб. При цьому кількість відмов регулювалася величиною r_{nop} .

Другий етап – усереднення FRR для розрахунків загального значення, тобто $\overline{FRR} = \frac{1}{N} \sum_{n=1}^N FRR(n)$. З цієї

формули видно, що очікувані значення будуть тим більше точними, чим більше об'єктів було притягнуто до процедури тестування системи ідентифікації. Відзначимо, що в реальних умовах існує багато причин, якими може бути обумовлене значення FRR: неправильне положення об'єкта перед пристроєм реєстрації, забруднення самого пристрою і т.д.

На FAR, так само як для FRR та FTE, при розрахунках було поширено поняття персонального FAR. Через статистичну природу цього параметра для отримання статистично надійного результату повинно бути зроблено багато спроб проходів. Як слідує з рис. 2, їх ефективна кількість $K_{ном} = 4$. Імовірність неправильного розпізнавання об'єкта $FAR(n)$, дані якого є в базі $[FAR(n)]$, розраховувалася за аналогією з $FRR(n)$, як відношення кількості успішних незалежних спроб розпізнавання

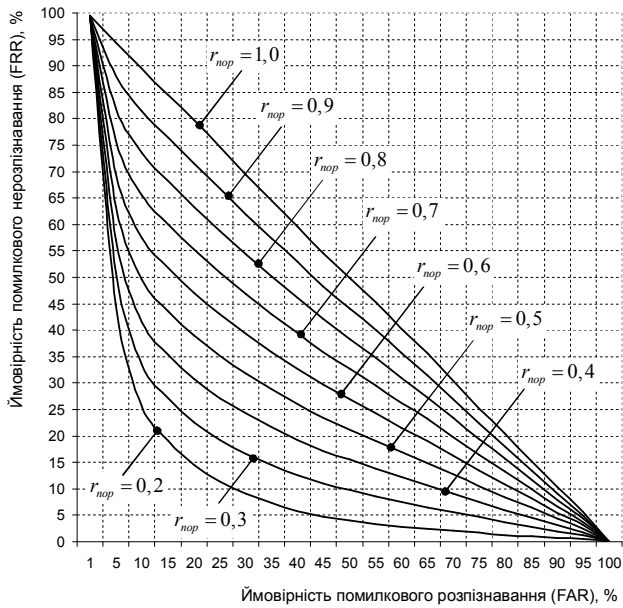


Рис. 3. Отримана залежність між помилками 1-го та 2-го роду при різних $r_{пор}$ для досліджуваних зображень (рис. 1)

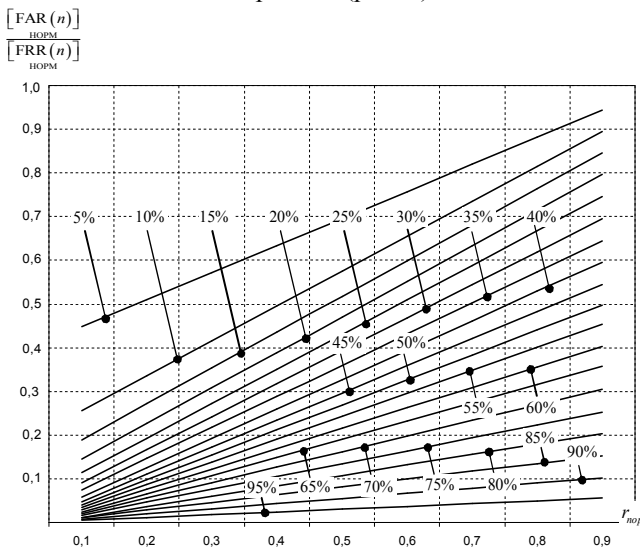


Рис. 4. Отримана нормована залежність

співвідношення $\frac{[FAR(n)]_{НОРМ}}{[FRR(n)]_{НОРМ}}$ від значення $r_{пор}$ для досліджуваних зображень (рис. 1)

Як видно, ROC – це робоча характеристика системи, а більш точно – відносна робоча характеристика. Графік ROC – це візуалізація компромісу між характеристиками FAR та FRR. У загальному випадку порівняльний алгоритм приймає рішення на підставі порогу, який визначає, наскільки близько повинен бути вхідний зразок до шаблону, щоб вважати це збігом. Якщо поріг зменшується, то кількість помилкових розбіжностей також буде зменшуватися, але при цьому буде більше по-

милкових прийомів. Відповідно, високий поріг зменшить FAR, але збільшить FRR. Лінійний графік залежності свідчить про відмінності для високої продуктивності: менше помилок – рідше виникають помилки.

Як видно з рис. 3, обмеження, які відповідають верхній характеристичній кривій, є більш ефективними, ніж ті, які притаманні іншим характеристичним кривим.

При аналізі та порівнянні ROC-кривих важливим принципом є розуміння методики тестування, у результаті якої вони отримані. Зокрема:

- при яких умовах та у яких обставинах проводилося тестування;
- який був сценарій використання системи;
- яка вихідна сукупність об'єктів ідентифікації як по кількості, так і по складу була використана, і т.д.

Згідно до [10], отримані результати можна віднести до однієї з відомих методик, а саме – до методики технологічного тестування. Слід відмітити, що також існують методики сценарного та операційного тестування.

Результати, отримані при різних методиках тестування, можуть сильно різнитися для однієї й тієї ж системи. Зазвичай, для будь-якого конкретного додатка, є доцільною фіксація припустимого значення FAR, і тоді значення FRR буде інтегральним критерієм точності для даної системи.

З використанням ряду публічних звітів про результати тестування систем біологічної ідентифікації [11...13, 16-18], приблизні значення точності верифікації в режимі операційного тестування для основних біометричних методів та отримані дані для однофакторної біометричної технології забезпечення превентивної безпеки у системах управління доступом, приведено у табл. 4.

При достатньо великих обсягах моделювання було помічено, що конкретні показники сильно варіюються залежно від використовуваного обладнання та погіршеності тестування. Але при цьому встановлено те, що однофакторна біометрична технологія забезпечення превентивної безпеки у порівнянні з такими методами розпізнавання, як по відбитковій пальця, по тривимірному зображенню особи та по райдужній оболонці ока, мають порівнянну точність, але розроблена технологія дозволяє отримати більш надійні дані щодо встановлення ідентичності об'єктів розпізнавання, поступаючи лише надзвичайно складним методам, які базуються на 3D-технологіях розпізнавання об'єктів.

Виходячи з даних, приведених у табл. 4, зробимо оцінку того, чи забезпечується достатня точність для рішення прикладних задач.

Як відмічається у [10], слід звернути увагу на те, що зазначена ймовірність неправильного розпізнавання FAR відповідає випадку *верифікації*, тобто порівнянню двох біометричних шаблонів між собою. Для більшості практичних завдань точність, що досягається в цьому випадку, при використанні кожного із методів, зазначених у табл. 4, цілком достатня.

Таблиця 4

Точність верифікації для різних біометричних методів

Метод →	3D технологія	2D технологія	Відбиток пальця	Райдужна оболонка ока	Термограма обличчя
FAR, % ↓	FRR алгоритм A4Vision	FRR найкращий 2D-алгоритм	FRR стандартний сканер	FRR найкращий сканер	FRR однофакторна біометрична технологія
	%				
0,1	0,2	19	0,4	4,7	0,338
0,01	1,0	28	1,0	5,3	1,154
0,001	1,5	–	1,3	6,0	1,681
0,0001	–	–	–	–	3,015

У випадку ідентифікації імовірність неправильного розпізнавання FAR збільшується пропорційно кількості об'єктів у базі даних системи при тій же чутливості (FRR). Так, наприклад, якщо при FRR, яка дорівнює 1,3%, пальцевий стандартний сканер у режимі верифікації забезпечує FAR=0,001% – один шанс зі ста тисяч. В режимі ідентифікації при тому ж FRR та базі даних в 10000 об'єктів FAR=10% – один шанс із десяти [10]. У практичному використанні така ситуація вже є неприпустимою для більшості додатків.

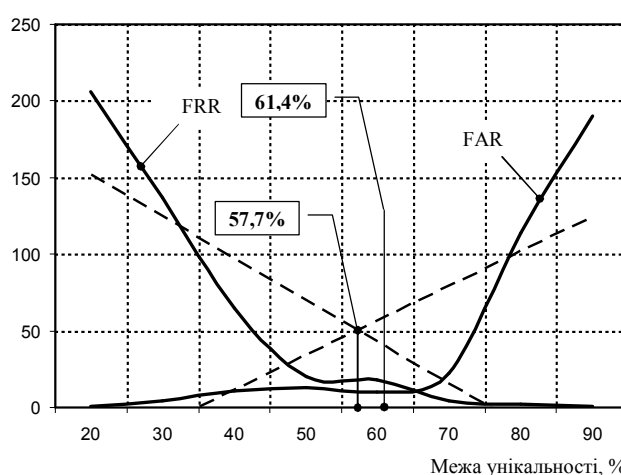
Т.ч., у режимі ідентифікації при базах даних на 1000...2000 об'єктів деякі існуючі методи (по райдужці, пальцеві, 3D-фото, а також однофакторна біометрична технологія) забезпечують прийнятну точність для систем контролю доступу. При базах даних більше, ніж 2000 об'єктів, жоден з біометричних методів «у чистому вигляді» не може бути застосований для більшості завдань. Для деяких завдань прийнятні напівавтоматичні розв'язки, коли людина-оператор отримує список найбільш схожих людей і на цій основі ухвалює остаточний розв'язок.

За даними [14] та з посиланнями на *Electronic Privacy Information Center*, максимальна ступінь розпізнавання у найкращій системі, яка нині знаходиться у експлуатації, досягає значення 61,4%. Ця цифра була обрана у якості опорної при порівнянні ефективності методів забезпечення спостереженості у технологічних системах спеціального призначення. Один з аналогічних результатів щодо ступеню розпізнавання на основі однофакторної біометричної технології забезпечення превентивної безпеки, який було отримано під час проведення численних математичних експериментів та який є характерним для всіх результатів, приведено у табл. 5 та візуально відображено на рис. 5. Для розрахунків використано методику, яка викладена у [10].

Висновки

Для збільшення точності в режимі ідентифікації доцільне використання декількох біометричних методів одночасно. Виходячи з аналогічних міркувань, Міжнародний підкомітет по стандартизації в області біометрії (ISO/IEC JTC1/SC37 Biometrics) розпочав розробку єдиного формату даних для автоматичного розпізнавання осіб, що включає мультибіометричні методи та технології.

Значення FRR та FAR, од

Рис. 5. Відповідність розрахованих співвідношень множин $[FRR(n)]$ та $[FAR(n)]$ загальновідомим даним

Таблиця 5

Значення FRR та FAR, які отримані при аналізі ступеню унікальності множин $[FRR(n)]$ та $[FAR(n)]$

Межа унікальності, %	FRR	FAR
20	206	1
30	137	4
40	65	11
50	20	13
60	18	10
70	4	22
80	2	114
90	1	190

Деякі виробники обладнання у галузі систем доступу, зважаючи на перспективу впровадження зазначеного стандарту, вже почали об'єднання кількох біометричних методів в один. Найімовірніше, незабаром розпізнавання особи з використанням кількох джерел інформації буде розглядатися як один біометричний ме-

тод. Об'єднання окремих біометричних методів та технологій розпізнавання особи в єдиний мультибіометричний метод приведе до істотного поліпшення точності в порівнянні з тією, що можуть дати системи, у яких використовується тільки один з методів, а також дозволить об'єднати переваги окремих методів за іншими критеріями.

Список літератури: 1. *Сесин, Е. М.* Построение моделей идентификации личности, основанных на сравнении множества физических или поведенческих характеристик человека [Текст] / Е. М. Сесин, В. М. Белов // Вестник СибГУТИ. — 2011. — №4. — С. 41-50. — ISSN 1998-6920. 2. *Ушмаев, О. С.* Информационная технология интеграции идентификации по изображению лица для ускорения автоматической дактилоскопической идентификации [Текст] / О. С. Ушмаев // Информатика и ее применения. — 2008. — №4. — Т. 2. — С. 66-73. — ISSN 1992-2264. 3. *Neuman, J.* On the problem of the most efficient tests of statistical hypotheses [Текст] / J. Neuman, E. S. Pearson // Philos. Trans. R. Soc. — 1993. — №231. — P. 289 -337. — ISSN 0962-8436. 4. *Phillips, P.* Facial recognition vendor test 2002 : Evaluation report, March 2003 [Електронний ресурс] / P. Phillips, P. Grother, R. Micheals, D. Blackburn, E. Tabassi, M. Bone // Портал : frvt.org. — Режим доступу \www/ URL: <http://www.frvt.org>. — Заголовок з екрана, доступ вільний, 16.02.2014. 5. *Ushmaev, O. S.* Problems of automatic fusion of biometric identifiers [Текст] / O. S. Ushmaev // Pattern Recognition and Image Analysis. — 2009. — V.19. — №3. — P. 534-538. — ISSN 1054-6618. 6. *Пахомов, С.* Отпечаток пальца вместо пароля [Текст] / С. Пахомов // Компьютер-Пресс. — 2004. — №4. — ISSN 0868-6157 ; [Електронний ресурс] // Портал : Компьютер-Пресс. — Режим доступу \www/ URL: <http://compress.ru/Archive/CP/2004/4/11/>. — Заголовок з екрана, доступ вільний, 17.02.2014. 7. *Гусев, Г. А.* Под мультибиометрическим контролем [Текст] / Г. А. Гусев // Системы безопасности. — 2009. — №3. — С. 134-136. — ISSN відсутній. 8. *Грузман, И. С.* Цифровая обработка изображений в информационных системах : навчальний посібник / И. С. Грузман, В. С. Киричук, В. П. Косых [та ін.] // Новосибирск : НГТУ, 2002. — 352 с. 9. *Казакова, Н. Ф.* Синтез методу виділення контурів у системах ідентифікації на основі усереднення перепадів яркості [Текст] / Н. Ф. Казакова, О. О. Фразе-Фразенко // Інформаційна безпека. — 2013. — №2(10). — С.48-57. — ISSN 2224-9613. 10. *Вакуленко, А.* Биометрические методы идентификации личности: обоснованный выбор и внедрение [Електронний ресурс] / А. Вакуленко, А. Юхнин // Портал : НПО ИНФОРМ: Биометрические системы безопасности. — Режим доступу \www/ URL: <http://www.npo-inform.com/press/biovybor/>. — Заголовок з екрана, доступ вільний, 17.02.2014. 11. *Phillips, P.* Facial recognition vendor test 2002: Evaluation report [Електронний ресурс] / P. Phillips, P. Grother, R. Micheals, D. Blackburn [и др.] // Портал : frvt.org. — Режим доступу \www/ URL: <http://www.frvt.org/>. — Заголовок з екрана, доступ вільний, 17.02.2014. 12. *Ushmaev, O. S.* Problems of automatic fusion of biometric identifiers [Текст] / O. S. Ushmaev // Pattern Recognition and Image Analysis. — 2009. — Т. 19. — № 3. — С. 534-538. 13. *Griffin, P.* Topics for multi-biometric research «MMUA-2003» : Panel Discussion [Електронний ресурс] / P. Griffin // Портал : без назви. — Режим доступу \www/ URL: <http://mmaaxs.ucsb.edu/>. — Заголовок з екрана, доступ вільний, 19.02.2014. 14. *Bonsor, K.* How Facial Recognition Systems Work [Електронний

ресурс] / K. Bonsor, R. Johnson // Портал : Howstuffworks — Режим доступу \www/ URL: <http://electronics.howstuffworks.com/gadgets/high-tech-gadgets/facial-recognition.htm>. — Заголовок з екрана, доступ вільний, 15.01.2013. 15. *Скопа О. О., Фразе-Фразенко О. О.* Анізотропна фільтрація зображень у системах аугментифікації // Матеріали II-ої Міжнарод. наук.-техн. конф. «Захист інформації і безпека інформаційних систем», Львів, 30 травня — 01 червня 2013 р. — Львів : НУ «Львівська політехніка», 2013. — С. 156-158. 16. *Фразе-Фразенко, А. А.* Система текстурных признаков, основанных на измерении пространственных частот [Текст] / А. А. Фразе-Фразенко // Технологічний аудит та резерви виробництва. — 2013. — №5/5(13). — С. 60-62. — ISSN 2226-3780. 17. *Казакова, Н. Ф.* Исследование и применение в системах защиты информации корреляционного критерия сходства графических структур [Текст] // Информационные системы в управлении, образовании, промышленности : монография / Н. Ф. Казакова, А. А. Фразе-Фразенко [и др.] ; под ред. В. С. Пономаренко. — Х. : ТОВ «Щедра садиба плюс», 2014. — 498 с. (Русск. яз.). — ISBN 978-617-7188-50-5. 18. *Фразе-Фразенко, О. О.* Огляд та аналіз поточного стану технологій розпізнавання образів та перспективи їх використання у системах захисту інформації // Удосконалення принципів та методів інформаційного забезпечення, інформаційної та фінансово-економічної безпеки підприємств та організацій сфери економіки, бізнесу та фінансів [Звіт про НДР] : (проміжн.) / О. О. Скопа, Н. Ф. Казакова, О. В. Орлик, Ю. В. Щербина, А. О. Петров, С. Л. Волков, О. І. Мацків, О. Г. Єсіна, А. Ю. Вакула, О. О. Фразе-Фразенко, А. В. Мінін, О. О. Йона, Є. В. Вавілов, К. Б. Айвазова ; кер. О. О. Скопа. — Одеса : ОНЕУ, 2013. — 0112U007713. — 236 с. — С. 38-69.

Поступила до редколегії 25.03.2014

УДК 004.738.5; 519.6 : 616-073.75; 004.932.2

Выбор методики и расчет коэффициентов ложного пропуска и ложного отказа в доступе в системах биометрической идентификации / А. А. Скопа, А. А. Фразе-Фразенко // Бионика интеллекта : науч.-техн. журнал. — 2014. №1 (82). — С. 80-89.

Алгоритмы установления сходства изображений связанные с получением характеристик стохастической взаимосвязи сравниваемых фрагментов. Все они основываются на идеях корреляционной и спектральной теории сигналов. Одной из мер сходства изображений, характеризующей качество работы биометрической системы идентификации, является определение FAR и FRR. Одна из таких методик рассматривается в статье.

UDK 004.738.5; 519.6 : 616-073.75; 004.932.2

The choice of methods and the calculation of FAR and FRR access to the biometric identification systems / O. O. Skopa, O. O. Frazе-Frazenko // Bionica Intellecta: Sci. Mag. — 2014. №1 (82). — С. 80-89.

Algorithms for establishing similarity of images associated with obtaining the stochastic characteristics of the relationship of the compared items. They are based on the ideas of correlation and spectral theory of signals. One of the measures of similarity of images, which characterizes the quality of biometric identification system is to determine the FAR and FRR. One such technique is discussed in the article.

Fig. 6. Ref.: 16 items.