УДК 004.056.5

О.А. Нємкова

Львівський інститут банківської справи університету банківської справи Національного банку України (Київ), Львів

БІОМЕТРИЧНА ІДЕНТИФІКАЦІЯ У КІБЕРПРОСТОРІ

Робота присвячена актуальній тематиці ідентифікації людини у кіберпросторі шляхом порівняння біометричних шаблонів з зображення, отриманого за допомогою відеоапаратури. Запропонована структура біометричного шаблону дозволяє пришвидшити пошук у великих масивах даних і може бути використана для автентифікації користувача мобільного банкінгу, або автентифікації при користуванні банкоматом.

Ключові слова: *біометричний шаблон, перцептивний хеш, мобільний банкінг, банкомат.*

Актуальність тематики

У сучасному інформаційному суспільстві актуальним для кожної країни є прагнення загальної цифрової ідентифікації громадян. Впровадження для кожної людини обов'язкових атрибутів - ідентифікаційного коду, електронного цифрового підпису, паспортних даних, номеру соціального страхування – потребує надійної ідентифікації, пов'язаної з біологічними особливостями людини (статичними та динамічними), на основі яких можна встановити її особистість у розподілених інформаційних мережах. Біометричні розробки впроваджувались вже не один десяток років, але до цього часу застосовувалися переважно в критичних інформаційних структурах, доступ до яких мала обмежена кількість людей. Це пояснює факт наявності пристроїв для проведення біометричної ідентифікації, які є надійними, але доволі складними та повільними, до того жті пристрої мають високу вартість, тому не можуть бути застосовані для великої кількості осіб в межах всього кіберпростору нашої планети.

На сьогодні все більше людей потребує надійної біометричної ідентифікації для можливості роботи з різноманітними інформаційними системами. Можна назвати багато таких систем:мобільний банкінг, банкомати,система відеоспостереження на автобанах, система електронного уряду та інші.Таки системи характеризуються великою кількістю учасників, так що їх можна називати масовими. Якщо розглядати міста, де можуть мешкати десятки мільйонів жителів (у Токіо на серпень 2014 року мешкало 37,5 млн. жителів, розмір міста визначався за світловою плямою вночі), то для інформаційних систем такого міста потрібна потужна система біометричної ідентифікації. При впровадженні масових систем біометричної ідентифікації доводиться йти на компроміс між швидкодією та результативністю (якістю розпізнавання). Таким чином, є необхідність у розробці масових систем біометричної ідентифікації, які б відповідали сучасним вимогам швидкодії та результативності.

Сучасний стан досліджень

Сучасний стан біометричних технологій характеризується наявністю великої кількості типів технологій, таких як розпізнавання за голосом, обличчям, райдужкою, сітківкою ока, рукою, геометрією мочки вуха, відбиткам пальців, перевіркою підпису. Всі види біометричних технологій можна стандартизувати за наступною моделлю. Спочатку виробляється початковий біометричний шаблон людини шляхом обробки її біометричних даних. Модель біометричного шаблону вимагає не відновлюваності біометрик людини з його даних. Початковий біометричний шаблон разом з іншою інформацією про людину зберігається у базах даних. Далі, коли людина має пройти ідентифікацію або автентифікацію, вимірюються біометричні дані та знову утворюється біометричний шаблон, вже вторинний. Далі відбувається процедура порівняння первинного та вторинного біометричних шаблонів, за результатами порівняння приймається рішення про ідентифікацію людини. Якщо між часом утворення первинного та вторинного біометричних шаблонів є проміжок, то, як правило, шаблони не співпадають, хоча належать тій самій людині. Тому для прийняття рішення про ідентифікацію потрібно мати критерій, за яким два неоднакових біометричних шаблони будуть вважатись такими, що належать одній людині. Деякі дослідники наполягають на ймовірнісному характері ідентифікації при порівнянні біометричних шаблонів.

Потрібно відмітити, що сучасні інформаційні системи, як правило, є розподіленими і охоплюють велику територію. Передача даних може відбуватись по незахищених каналах. Тому є потреба у шифруванні даних при передачі вторинного біометричного шаблону.

Основний матеріал

Розглянемо спосіб утворення біометричного шаблону за допомогою фотографії людини. В багатьох випадках фотографування є найзручнішим способомутворення біометрики. Відомо, що при

обробці великих масивів фотографій для пошуку конкретного фото використовується хешування, в основному застосовуються перцептивні хеші. На сьогоднішній день відомо декілька алгоритмів утворення перцептивного хешу з зображень, наприклад Simple Hash [1], DCT Based Hash [2, 3], Radial Variance Based Hash [2, 4] та Marr-Hildreth Operator Based Hash[2, 5].

Основні етапи алгоритму пошуку аналогічної фотографії в базах даних, включаючи утворення перцептивного хешу, є такими.

1. Зменшення зображення до потрібного розміру в пікселях. Результатом може бути картинка 8*8, або 16 × 16 пікселів. При цьому утворюється матриця 24 × 8, або 48 × 16, елементами якої є значення яскравості монохромних компонент зображення. При зменшенні зображення може використовуватися метод найближчого сусіда, або метод швидкого перетворення Фур'є. Від способу зменшення зображення залежить час роботи перцептивного хешу. Відновити з такою картинки початкову фотографію неможливо. Цей факт дає підстави для виконання умови не відновлюваності біометрики, в даному випадку фотографії, якщо прийняти хеш в якості біометричного шаблону.

2. Зменшення інформативності зображення (перехід до зображення в градаціях сірого) та утворення хешу. Для переходу до градацій сірого можна використати одну з монохромних компонент зображення: синю, зелену або червону. В результаті отримаємо матрицю 8*8, або 16*16. Далі для утворення хешу за середнім потрібно розрахувати середнє значення інтенсивності всіх комірок матриці і замінити значення в комірках на одиницю, якщо воно більше середнього, або на нуль, якщо воно менше середнього. Отримана бінарна матриця (містить тільки нулі або одиниці) є перцептивним хешем фотографії. Виявляється, що даний хеш не змінюється при зміні пропорцій, наприклад, при розтягуванні фото, або зміні відтінків, але хеші фотографій двох різних людей доволі сильно відрізняються.

3. Порівняння хешей двох фото шляхом обчислення відстані між ними (відстань Хемінга). Вважають, що хеші фото однієї людини розрізняються не більш ніж на 8 для матриці 8*8. Хеші фото двох різних людей для розміру бінарної матриці 8*8 відрізняються більше, ніж на 8. Цей результат був перевірений та підтверджений для людей одного віку, раси, статі, соціального положення, кольору волосся. Порівнювались тільки обличчя. Те, що візуально особи відрізнялись, відобразилось в їх хешах розміру 8*8 як відстань за Хемінгом більша за 8.

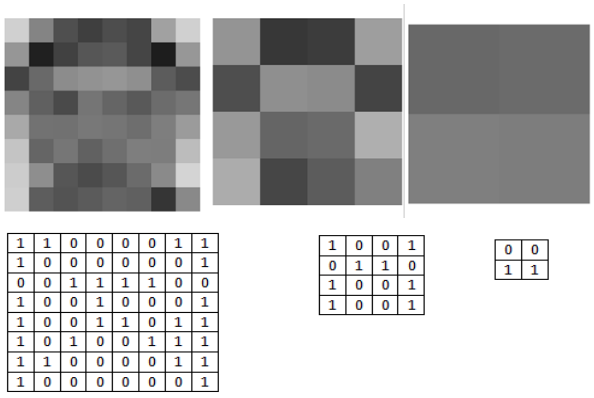
Таким чином, можна запропонувати перцептивні хеші за середнім в якості біометричного шаблону. Зберігаються хеші не в вигляді матриць, а у вигляді послідовності шістнадцяткових цифр.

Для задач ідентифікації в кіберпросторі важлива швидкість пошуку хешу з бази даних, а головне – однозначність розпізнавання. Якщо апріорі відомо, що вторинний біометричний шаблон може відрізнятись від первинного за відстанню Хемінгу, потрібно проводити порівняння повної довжини хешу, тому що невідомо, де можуть міститись різні біти.

Для оптимізації порівняння запропоновано таку процедуру.

Фотографія людини зменшується до трьох розмірів: 8 × 8 пікселів, 4 × 4 пікселів та 2 × 2 пікселів. З кожної матриці будується перцептивнийхеш за середнім. Отримаємо три різних за розмірністю бінарних матриці. Зменшення кількості пікселів в розмірі фото відповідає за укрупнення контурів зображення. Біометричний шаблон складається з трьох частин різної довжини. Перша відповідає найменшій матриці 2 × 2, друга – матриці 4 × 4, і найдовша відповідає матриці 8 × 8.

На рис. 1 – 4 зображено зменшені до різних розмірів фото чотирьох різних осіб,які умовно названі Особа1, Особа2, Особа3, Особа4, з відповідними бінарними матрицями, а в табл. 1 записані біометричні шаблони даних осіб.

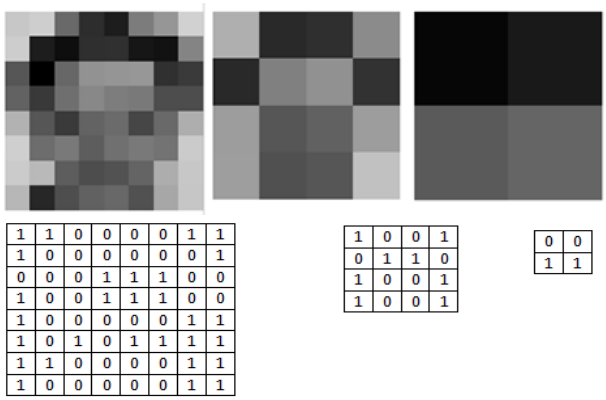


1	1	0	0	0	0	1	1
1	0	0	0	0	0	0	1
0	0	1	1	1	1	0	0
1	0	0	1	0	0	0	1
1	0	0	1	1	0	1	1
1	0	1	0	0	1	1	1
1	1	0	0	0	0	1	1
1	0	0	0	0	0	0	1

1	0	0	1
0	1	1	0
1	0	0	1
1	0	0	1

0	0
1	1

Рис. 1. Особа1 – зменшене фото до розміру 8 × 8 пікселів, 4 × 4 пікселів, 2 × 2 пікселів та відповідні бінарні матриці



1	1	0	0	0	0	1	1
1	0	0	0	0	0	0	1
0	0	0	1	1	1	0	0
1	0	0	1	1	1	0	0
1	0	0	0	0	0	1	1
1	0	1	0	1	1	1	1
1	1	0	0	0	0	1	1
1	0	0	0	0	0	1	1

1	0	0	1
0	1	1	0
1	0	0	1
1	0	0	1

0	0
1	1

Рис. 2. Особа2 – зменшене фото до розміру 8 × 8 пікселів, 4 × 4 пікселів, 2 × 2 пікселів та відповідні бінарні матриці

Таблиця 1

Біометричні шаблони, що утворені з бінарних матриць

№	I	II				III															
Особа1	3	**9**	**6**	**9**	**9**	C	3	8	1	3	C	9	1	9	B	A	7	C	3	8	1
Особа2	3	**9**	**6**	**9**	**9**	C	7	8	1	1	C	1	C	8	1	A	F	C	3	8	3
Особа3	3	**9**	**6**	**F**	**F**	C	3	8	1	7	0	5	A	F	F	F	F	F	F	3	C
Особа4	C	**6**	**7**	**7**	**0**	8	1	3	E	3	E	7	F	7	F	4	2	3	C	0	0

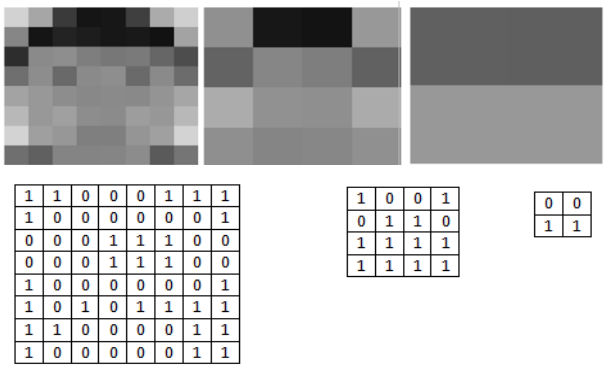


Рис. 3. Особа3 – зменшене фото до розміру 8 × 8 пікселів, 4 × 4 пікселів, 2 × 2 пікселів та відповідні бінарні матриці

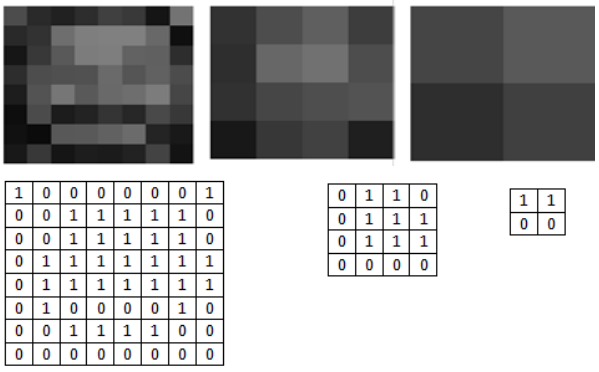


Рис. 4. Особа4 – зменшене фото до розміру 8 × 8 пікселів, 4 × 4 пікселів, 2 × 2 пікселів та відповідні бінарні матриці

Як видно з наведених зображень, за ними неможливо відновити фотографії і навіть здогадатись, хто на них зображений. Тим більш неможливо це зробити при використанні бінарних матриць. Але можна побачити досить велику різницю між самими бінарними матрицями. Для Особи1 та Особи2 є багато спільного в зображенні, тому їх біометричні шаблони починають відрізнятись тільки на рівні матриць 8*8. Біометричний шаблон Особи3 відрізняється від перших двох вже на рівні матриці 4 × 4, хоча початок матриці 8 × 8 в них однаковий. Біометричний шаблон Особи4 відрізняється від всіх інших шаблонів на рівні матриці 2 × 2.

Послідовність нулів та одиниць групується по чотири та записується у вигляді шістнадцяткової послідовності. Починається біометричний шаблон зі значення, яке утворюється з матриці 2 × 2, далі записані чотири значення матриці 4 × 4, а в кінці записані 16 значень матриці 8 × 8. У такому вигляді біометричний шаблон зберігається в базах даних. Його розмір складає 84 біти.

У табл. 1 значення шістнадцяткової послідовності об'єднані у групи I, II та III. Група I утворюється з бінарної матриці 2 × 2, група II утворюється з бінарної матриці 4 × 4, група III – відповідно з матриці 8 × 8.

При пошуку в базі даних первинного біометричного шаблону людини за вторинним шаблоном потрібно послідовно порівнювати значення груп I, II та III. Різні значення в групах I та II однозначно свідчать на користь того, що порівнювані шаблони належать різним людям. Такі шаблони можна далі не порівнювати.

Якщо в групі III різниця невелика (не більше 5), це означає, що шаблони належать одній людині. Всі вищенаведені міркування вірні для випадку, коли обличчя людини при утворенні вторинного біометричного шаблону не змінено настільки, що навіть за повноформатною фотографією його неможливо ідентифікувати. Наприклад, коли людина переносить операцію по кардинальній зміні форми обличчя, потрібно заново записувати первинний біометричний шаблон, щоб інформаційна система могла її ідентифікувати.

Виділення груп I, II та III у біометричному шаблоні дозволяє суттєво зменшити час пошуку ідентичного шаблону в базах даних. Під ідентичністю розуміємо відмінність не більш ніж на 8 позицій з 64 для групи III. Якщо вважати, що реально порівняння відбувається за бінарними значеннями, то «відкидання» суттєвої частини шаблонів відбувається вже на I групі, для цього достатньо перевірити не більш ніж чотири двійкові позиції. Це пришвидшує «відкидання» непідходящих шаблонів у двадцять разів. Перевірка групи II дає змогу «відкинути» непотрібні шаблони після порівняння не більш ніж 16 двійкових позицій, це пришвидшує «відкидання» у чотири рази.

Трохи довше будуть перевірятись шаблони дуже подібних людей, але відомо, що таких людей є небагато, тому в цілому пришвидшення пошуку має бути суттєвим, за оцінками орієнтовно у два рази.

Дійсно, якщо з трьох тисяч біометричних шаблонів тисяча відкидається після перевірки максимум 4 двійкових позицій, ще тисяча – після додаткових 12 позицій (разом 16), а решта перевіряється для ще 64 позицій (разом 80), то сумарно витрати часу будуть пропорційні коефіцієнту 100.

Якщо біометричний шаблон записати тільки для бінарної матриці 8×8, то три тисячі шаблонів будуть перевірятись з витратами часу, пропорційними коефіцієнту 192, тобто майже у два рази довше.

Висновки

Запропоновано у якості біометричного шаблону використовувати послідовність перцептивних хешей, отриманих для зображення, яке зменшене до 8×8 пікселів, 4×4 пікселів та 2×2 пікселів. Довжина шаблону складає 84 біти. Структура біометричного шаблону дозволяє проводити швидкий пошук серед великого масиву шаблонів. Перцептивні хеші інваріантні відносно багатьох спотворень та змін зображення, також є зручний спосіб визначення відповідності вторинного біометричного шаблону первинному за допомогою обмежень, що накладаються на відстань за Хемінгом.

Наявність великої кількості відеокамер з тенденцією до збільшення дає технічну можливість для ідентифікації людей в багаточисельних інформаційних системах з перспективою до можливості ідентифікації у всьому кіберпросторі. Можна відмітити перспективність таких біометричних шаблонів для застосування у платіжних системах.

Наприклад, кожний банкомат у цілях безпеки вже оснащений відеокамерою, яку можна використати для біометричної автентифікації. Дана технологія підходить також для мобільного банкінгу на основі смартфонів.

Список літератури

1. Simple and DCT perceptual hash-algorithms. - [Електронний ресурс]. - Режим доступу до ресурсу: http://www.hackerfactor.com/blog/index.php?/archives/432-Looks-Like-It.html.

2. Egmont-Petersen M. Image processing with neural net works - a review / M. Egmont-Petersen, D. de Ridder, H. Handels // Pattern Recognition. – 2002. – 35 (10). – P. 2279-2301.

3. Christoph Zauner. Implementation and Bench marking of Perceptual Image Hash Functions [Електронний ресурс] / Christoph Zauner. – 2010. – Режим доступу до ресурсу: http://www.phash.org/docs/pubs/thesis_zauner.pdf.

4. Practical evaluation of a radial soft hash algorithm / F.X. Standaert, F. Lefebvre, G. Rouvroy, B.M. Macq, J.J. Quisquater, J.D. Legat // Proceedings of the International Symposiumon Information Technology: Codingand Computing (ITCC), vol. 2, pp. 89-94. IEEE, Apr. 2005.

5. Zeng Jie. A Novel Block-DCT and PCA Based Image Perceptual Hashing Algorithm / Jie Zeng // IJCSI International Journal of Computer Science Issues. – January 2013. – Vol. 10, Issue 1, No 3. – P. 44-52.

Надійшла до редколегії 14.05.2015

Рецензент: д-р екон. наук, доц. С.В. Кавун, Харківський інститут банківської справи Університету банківської справи НБУ (Київ), Харків.

БИОМЕТРИЧЕСКАЯ ИДЕНТИФИКАЦИЯ В КИБЕРПРОСТРАНСТВЕ

А.А. Немкова

Работа посвящена актуальной тематике идентификации человека в киберпространстве путем сравнения биометрических шаблонов с изображения, полученного с помощью видеоаппаратуры. Предложенная структура биометрического шаблона позволяет ускорить поиск в больших массивах данных и может быть использована для аутентификации пользователя мобильного банкинга или аутентификации при пользовании банкоматом.

Ключевые слова: *биометрический шаблон, перцептивный хэш, мобильный банкинг, банкомат.*

BIOMETRIC IDENTIFICATION IN CYBERSPACE

A.A. Nemtcova

The work is devoted to actual topics of human identification in cyberspace by comparing the biometric template with the image obtained by the video equipment. The proposed structure of the biometric template allows you to speed up the search in large data sets and can be used to authenticate the user of mobile banking, or authentication when using an ATM.

Keywords: *biometric template, perceptual hash, mobile banking, ATM.*