DOI: https: //doi.org/10.32353/khrife.1.2019.018
УДК 343.982.32

Х. В. Луценко,
науковий співробітник Харківського НДІСЕ,
м. Харків, Україна,
ORCID: https://orcid.org/0000–0001–7883–3764
e-mail: about-work@ukr.net

К. В. Нікулін,
науковий співробітник Харківського НДІСЕ,
м. Харків, Україна,
ORCID: https://orcid.org/0000–0002–0665–6313
e-mail: knikulin87@gmail.com

ГОЛОСОВА ІДЕНТИФІКАЦІЯ ДИКТОРА ЯК ОДИН ІЗ СУЧАСНИХ БІОМЕТРИЧНИХ МЕТОДІВ ІДЕНТИФІКАЦІЇ ОСОБИ

У статті розглянуто найбільш розповсюджені біометричні системи ідентифікації осіб, у тому числі голосової ідентифікації диктора на відео-, звукозаписах.

Приведені галузі використання біометричних технологій та їх загальні характеристики. Наведений огляд використання груп ознак при ідентифікації, які характеризують голос.

Розглянуто основні переваги голосової біометрії, такі як простота реалізації системи; низька вартість (найнижча серед усіх біометричних методів); вона не потребує контакту і дозволяє здійснювати верифікацію на великій відстані на відміну від інших біометричних технологій.

Проведено аналіз існуючих методів розпізнавання мовної інформації, які ідентифікують особу за сукупністю унікальних характеристик голосу.

Ключові слова: біометричні технології ідентифікації, голосова ідентифікація диктора, верифікація особи, автоматичні методи розпізнавання диктора, експертиза відео-, звукозапису.

Постановка наукової проблеми. Біометричні технології — це технології, основані на вимірюванні унікальних характеристик певної особи. Це можуть бути як унікальні характеристики, отримані при народженні (ДНК, відбитки пальців, радужна оболонка ока тощо), так і характеристики, набуті із часом, або здатні змінюватись із віком або зовнішнім впливом (почерк, голос, хода тощо).

Біометричні технології активно використовуються в багатьох галузях, які пов’язані із забезпеченням безпеки доступу до інформації та матеріальних об’єктів, а також в унікальній ідентифікації особи. Це може бути доступ до робочих місць та мережевих ресурсів, охорона правопорядку, сфера соціальних послуг до захисту інформації, безпека банківських та інших фінансових операції тощо.

Аналіз основних досліджень і публікацій. Перший міжнародний патент на систему ідентифікації за голосом був поданий у 1983 р. дослідницьким телекомунікаційним центром CSELT (Італійська Республіка) за авторством Michele Cavazza та Alberto Ciaramella. У травні

2013 р. банківські підрозділи Barclays почали використовувати систему ідентифікації клієнтів по телефону протягом перших 30 секунд звичайної розмови. Система була розроблена кампанією Nuance.

Сьогодні відомі такі розробники систем ідентифікації за голосом: Nuance, США; Nok Nok Labs; VoiceVault, американська компанія з центром досліджень та розробок у Сполученому Королівстві Великої Британії та Північної Ірландії; Sensory, Inc, Сполучені Штати Америки; компанії ЦМТ, Російська Федерація; Іноваційний технологічний центр «Система-Саров», Російська Федерація; BioLink, Російська Федерація; АСМ Решения, Російська Федерація; ValidSoft; Auraya Systems; Authentify; KeyLemon; Verint Systems; VoiceTrust.

Загально визнаним лідером ринку є компанія Nuance, їх рішення використовує «Аерофлот», розпізнавання мовлення Siri (Speech Interpretation and Recognition Interface) — персональний помічник та система питання-відповідь, розроблена для iOS. Цей додаток використовує обробку природного мовлення, щоб відповідати на запитання і давати рекомендації. Siri пристосовується до кожного користувача індивідуально, вивчаючи його переваги протягом тривалого часу. Однак оскільки голос особи може змінюватись залежно від ряду чинників, даний метод не є абсолютно точним[1].

Мета статті — загальний огляд та аналіз існуючих методів біометричної ідентифікації особи, зокрема, голосової ідентифікації диктора.

Викладення основного матеріалу дослідження. В основі науки про ідентифікацію особи лежать ідеї виміру тіла людини та його частин. Ці ідеї вперше сформулював французький криміналіст Альфонс Бертильон (Alphonse Bertillon) (1853 — 1914) — співробітник паризької префектури, який займався реєстрацією злочинців. У 1879 році він представив систему ідентифікації злочинців, яка отримала назву «антропометрія» та включала в себе: виміри росту, довжини та об'єму голови, довжини рук, пальців, стоп тощо, а також словесний портрет злочинця, фотопортрет в анфас і в профіль а також опис особливих прикмет. Сучасна криміналістика досі використовує таку систему, доповнивши її антропоскопією, дактилоскопією, фотороботами, новими методами опису особливих прикмет на обличчі або тілі людини та технологіями їх реалізації. Однак поняття біометрії сформовано десятиріччям пізніше. У витоків ранньої біометрії стояв англійський дослідник Френсис Гальтон (Francis Galton).

Уперше термін «біометрія» з'явився близько 1980 р. В енциклопедії сучасної України існує таке тлумачення терміна «біометрія» (від «*біо*» і «*метрія*») — наука про застосування методів математичної статистики для вивчення явищ життя. Інша назва — варіаційна статистика[2].

---

[1] Биометрия от «А» до «Я»: полное руководство биометрической идентификации и аутентификации. URL : https://securityrussia.com/blog/biometriya.html (дата звернення 04.06.2019).

[2] Енциклопедія сучасної України. URL : http://esu.com.ua/search_articles.php?id=35327 (дата звернення 06.05.2019).

У міжнародному стандарті **ISO/IEC 2382–37:2012** Information technology — Vocabulary — Part 37: Biometrics дано таке визначення біометричних систем. **Біометрична система** — це система, призначена для автоматичного розпізнавання індивіда (особистості людини), заснованого на його поведінкових і біологічних характеристиках[3].

У галузі інформаційних технологій застосовується таке тлумачення: *біометрія* — сукупність автоматизованих методів і засобів ідентифікації людини, заснованих на її фізіологічній або поведінковій характеристиці.

Розгляд цієї теми необхідно розпочати з визначення певних понять. У сучасних системах та технологіях, що основані на біометрії, слід розрізняти біометричні верифікацію, ідентифікацію та автентифікацію.

Відповідно до **ISO/IEC 2382–37–2016 37.08.03 біометрична верифікація** (biometric verification) — процес підтвердження біометричної заяви при порівнянні. Використання терміна **«автентифікація»** замість терміна **«біометрична верифікація» —** неприпустимо[4].

Згідно з тим же **ISO/IEC 2382–37–2016 37.08.02 біометрична ідентифікація** (biometric identification) — процес пошуку по базі даних біометричних реєстрацій, який направлений на пошук та повернення ідентифікатора(ів) біометричного контрольного шаблона, зв'язаного з одним індивідом. Використання терміна **«автентифікація»** замість **«біометричної ідентифікації» —** неприпустимо.

У цьому ж міжнародному стандарті **ISO/IEC 2382–37–2016 37.08.01 автентифікація** (authentication) — дія, що доводить або показує безперечне походження або достовірність. Термін використовується у біометрії в якості синоніма для додатку **біометричної верифікації**, функції біометричної верифікації, також використовується в якості синоніма для додадків біометричної ідентифікації та функції біометричної ідентифікації.

Таким чином, *біометрична ідентифікація* — це спосіб ідентифікації особи за окремими специфічними біометричними ознаками (ідентифікаторами), які властиві конкретній особі. А *біометрична автентифікація* — це розпізнавання індивідуума на основі фізіологічних характеристик та поведінки. Автентифікація проводиться шляхом перевірки приналежності певній особі представленої ідентифікаційної ознаки.

У таблиці 1 наведено приклади використання біометричних технологій у різноманітних галузях використання[5].

---

[3] ISO/IEC 2382–37:2012. («Information technology — Vocabulary — Part 37:Biometrics», IDT). URL : https://www.iso.org/standard/55194.html (дата звернення 13.05.2019).

[4] ISO/IEC 2382–37–2016. [ISO/IEC 2382–37:2012] («Information technology — Vocabulary — Part 37:Biometrics», IDT). URL : https://meganorm.ru/Data2/1/4293747/4293747566.pdf (дата звернення 13.05.2019).

[5] Биометрические системы идентификации и аутентификации. URL : http://mirznanii.com/a/309729/biometricheskie-sistemy-identifikatsii-i-autentifikatsii (дата звернення 04.06.2019).

Таблиця 1

Галузі використання біометричних технологій та їх загальні характеристики

№ з/п	Галузь використання	Основні характеристики
1	2	3
1	Комп'ютерна безпека	Біометрія використовується для заміни (інколи посилення) стандартної процедури входу в різноманітні програми за паролем, старткартці, таблетці touch-memory тощо. Найрозповсюдженішим рішенням на базі біометричних технологій є ідентифікація (або верифікація) за біометричними характеристиками в корпоративній мережі або при вході на робочу станцію (персональний комп'ютер, ноутбук тощо)
2	Торгівля	Основні напрямки: — в магазинах, ресторанах, кафе біометричні ідентифікатори використовуються або безпосередньо як засіб ідентифікації покупця для зняття грошей з рахунка, або для підтвердження права покупця на будь-які знижки або інші пільги; — у торгових автоматах і банкоматах як засіб ідентифікації особи замість магнітних карток або доповнення до них; — в електронній комерції біометричні ідентифікатори використовуються як засоби віддаленої ідентифікації через мережу Інтернет, що значно надійніше аніж паролі, а у поєднанні із засобами кріптографії дає електронним трансакціям дуже високий рівень захисту
3	Системи контролю та управління доступом	У мережевій архітектурі, коли у будівлі є кілька входів, які обладнані біометричними замками, шаблони біометричних характеристик усіх співробітників зберігаються централізовано разом із інформацією про те, кому та куди (та, можливо, коли) дозволений вхід. У таких системах реалізуються такі технології розпізнавання: відбиток пальця, обличчя, форма руки, радужна оболонка ока, голос тощо
4	Автоматизовані	Основним призначенням такої системи є

	дактилоскопічні інформаційні системи	управління правами, які надані державою громадянам та іноземцям. Права громадян, голосування, місця проживання, праці іноземців, право отримувати соціальне забезпечення тощо визнаються та підтверджуються за допомогою документів та різноманітних карт. Нині такі системи набули широкого розповсюдження через те, що деякі країни почали використовувати їх для перевірки особи тих, хто в'їжджає
5	Комплексні системи	До систем даного типу належать рішення, які поєднують у собі системи перших трьох класів. Співробітник компанії реєструється у адміністратора систем усього один раз, надалі йому автоматично назначаються всі необхідні привілеї як на вхід до приміщення, так і на роботу в корпоративній мережі з її ресурсами

Окрім цих основних секторів користування біометричними технологіями, зараз починається активне використання біометрії в інших галузях, таких як:

— гральний бізнес. Біометрія використовується у двох напрямках: перевірка всіх за «чорними списками» (аналог масової ідентифікації за обличчями, яка використовується в аеропортах), а також система ідентифікації та платіжний засіб постійних клієнтів;

— ідентифікація у мобільних пристроях, таких як мобільні телефони, планшети, ноутбуки тощо;

— транспортна галузь як платіжний засіб;

— електронні системи голосування (використовується замість карток);

— медицина. Біометрія використовується для ідентифікації медичних працівників при отриманні доступу до закритих даних і для електронного підпису в історії хвороби.

Сучасна біометрична автентифікація базується на двох основних методах:

— статичний метод автентифікації — розпізнає фізичні параметри людини, якими вона володіє протягом усього життя: від свого народження до самої смерті (відбитки пальців, характеристики радужної оболонки ока, малюнок сітківки ока, термограма, геометрія обличчя, геометрія кисті рук, фрагмент генетичного коду тощо);

— динамічний метод автентифікації — аналізує характерні риси, особливості поведінки користувача, які демонструються у момент виконання будь-якої звичайної повсякденної дії (підпис, клавіатурний почерк, голос тощо).

Основним на світовому ринку біометричного захисту завжди був статичний метод. Динамічна автентифікація та комбіновані системи захисту інформації займали лише 20 % ринку. Однак в останні роки спостерігається активний розвиток динамічних методів захисту[6].

Ідентифікація за голосом відома досить давно, людина за відсутності будь-яких технічних засобів ідентифікувала іншу людину за трьома можливими ознаками — голосом, підписом та зовнішністю. Отже голосова ідентифікація — це один із найстаріших методів біометричної ідентифікації. Кожна людина має свій унікальний голос, який відрізняється від усіх інших певними ознаками.

Голосова ідентифікація — одна із найпривабливіших систем для ідентифікації, однак існуючі на даний момент проблеми у даному виді біометричних систем повинні бути, як мінімум, враховані у працюючих системах. Наприклад, розпізнавання за голосом може ефективно використовуватись як додатковий метод, наприклад, до розпізнавання за обличчям, оскільки ймовірність помилки самостійного розпізнавання за голосом складає 2–5 %. Сьогодні напрямок ідентифікації особи за голосом активно розвивається. Перевагою голосової біометрії є простота реалізації системи, яка, зазвичай, складається із голосового приймача, диктофона, голосового модулятора, біометричного програмного забезпечення та бази даних голосів. На відміну від інших біометричних технологій, голосова біометрія дозволяє здійснювати верифікацію на великій відстані. Одним з перспективних шляхів підвищення надійності голосової ідентифікація є залучення характеристик динаміки підсвідомих рухів, що активно використовується при ідентифікації по підпису. З іншого боку, існують галузі застосування, в яких голосова ідентифікація є найбільш зручною, наприклад, віддалений доступ до телекомунікаційних каналів зв'язку з аналізу голосових даних[7].

Метод розпізнавання за голосом ідентифікує особу за сукупністю унікальних характеристик голосу. Алгоритми аналізують основні ознаки, за якими приймається рішення про особу диктора: голосового джерела, резонансних частот мовленнєвого тракту, їх затухань, а також динаміку управління артикуляцією.

Спираючись на багатий досвід, сучасні науковці при ідентифікації за голосом використовують дві групи ознак, які характеризують голос.

Перша група — це фізіологічні (анатомічні) ознаки, які пов'язані з особливостями механізму мовотворення людини. Друга група — це так звані артикуляційні ознаки, які засновані на особливостях роботи нервової системи людини, яка визначає характер використання фізіологічних ознак.

Фізіологічні ознаки. Фізіологічні ознаки засновані на моделі мовленнєвого тракту. В даному випадку як основні ознаки виступають декілька параметрів, які характеризують голос:

---

[6] Биометрия от «А» до «Я»...

[7] *Царьов Р. Ю., Лемеха Т. М.* Біометричні технології. Навч. посіб. Одеса, 2016. С. 140.

— енергія мовного сигналу;

— частотний діапазон мовного сигналу;

— основна частота — визначає довжину мовного тракту;

— форманти — визначають концентрацію мовного сигналу за частотою та характеризують голосні звуки.

Артикуляційні ознаки. Якщо фізіологічні ознаки відображають статистичні властивості мовного апарату, то артикуляційні ознаки дозволяють здійснити опис поведінки мовного апарату у часі, тобто відобразити артикуляційну динаміку мови. Головним фактором, який впливає на цю групу ознак, є соціально обумовлені мовленнєві навики людини, її індивідуальний опит, темперамент та особливості характеру. Артикуляційні ознаки враховують інтонацію мовлення, ритм, наголоси, гучність. Для того, щоб отримати ці характеристики, використовується поняття синтагма.

Синтагма — це ритмічно-мелодична одиниця мови, граматично оформлена та визначена у межах більш складної цілої структури (наприклад, речення) із закінченою думкою. У межах синтагми відокремлюють сегменти характеристики мови та інтонаційні характеристики мови, а саме: інтенсивність голосу; мелодійність голосу; система наголошень; часові характеристики — довжина сегментів та пауз; темп мовлення; тон мовлення.

Слід зазначити, що сучасні системи ідентифікації за голосом можуть одночасно використовувати фізіологічні й артикуляційні ознаки.

Системи голосової ідентифікації можна поділити на наступні класи: текстозалежні, текстонезалежні, дикторозалежні, дикторонезалежні.

Дикторозалежні системи — це системи, які орієнтовані на ознаки мовлення певної людини або групи осіб, тому вони можуть використовуватися для ідентифікації тільки цієї особи (групи осіб). При зміні диктора (особи, яка ідентифікується системою) необхідно налаштовувати систему знову з використанням голосових ознак нового диктора.

Дикторонезалежні системи — це системи, які не прив’язані до голосових ознак певної особи, та можуть використовуватися для ідентифікації будь-якої особи. Такі системи самі виділяють необхідні ознаки голосу та порівнюють їх з еталоном з бази.

Текстозалежні системи — це системи голосової ідентифікації, які здійснюють ідентифікацію особи за допомогою певного ключового слова або ключової фрази, яку повинна вимовити особа, що проходить ідентифікацію, наприклад, проголошення парольної фрази, яка кожного разу генерується випадковим чином. Використання індивідуальних ознак і збіг згенерованої та розпізнаної фраз підвищує надійність.

Текстонезалежні системи — це системи, які здійснюють ідентифікацію особи за допомогою голосу без прив’язки до будь-яких ключових слів. У даному випадку важливе значення мають артикуляційні ознаки голосу людини, саме вони використовуються як головні ознаки, а фізіологічні ознаки виступають як вторинні. Текстонезалежна

ідентифікація має на увазі використання тільки індивідуальних ознак. Важливою характеристикою системи голосової ідентифікації є швидкість (швидкодія) визначення особистості. Підвищення швидкодії може бути досягнуто за рахунок використання нових швидких алгоритмів обробки даних.

За прогнозами Adweek, в 2019 році ринок платформ розпізнавання голосу зросте до 601 млн доларів. Все тому, що людям простіше розмовляти, ніж набирати текст. Необхідні такі голосові помічники, які будуть підтримувати звичне спілкування.

На ринку вже є багато помічників: Amazon Alexa, Google Assistant, Cortana, Bixby, «Алиса», SoundHound, Apple Siri, X.ai та інші.

Упровадження пристроїв голосового керування в автомобілі — одна з тенденцій, яка приведе до глобальних змін в автомобільному секторі. Такі пристрої зможуть централізовано керувати більшістю функцій автомобіля за допомогою людського мовлення, виключаючи необхідність використання кнопок, циферблатів, перемикачів тощо. Використовуючи пристрої розпізнавання голосу, користувачі зможуть легко керувати цілим рядом функціональних можливостей автомобіля, що для людини є більш комфортним та дозволяє не відволікатись від безпосереднього керування автомобілем, концентруючи увагу на водінні. Упровадження таких технологій буде зростати в найближчому та середньотривалому періоді[8].

Підвищення надійності голосової ідентифікації є важливим не тільки для такого напрямку, як розмежування доступу до фізичних та інформаційних об'єктів, наприклад, доступу до операційної системи персонального комп'ютера або віддаленого доступу до телекомунікаційних каналів зв'язку з аналізу голосових даних. Певний інтерес є і для суміжних напрямів мовних технологій: розпізнавання усного мовлення, управління голосовими командами тощо. Сьогодні широкого поширення набув електронно-цифровий підпис для захисту конфіденційних документів у вигляді захищеного електронного пристрою (token), у зв'язку з цим перспективним напрямком є розробка захисту конфіденційних документів на основі мовного підпису.

Крім того, практичні застосування таких досліджень корисні для правоохоронних органів, наприклад, ототожнення особи за фізичними параметрами голосу.

Однак розвиток технологій та технічний прогрес несе в собі не лише позитивні моменти, а й розширює можливості злочинців. Одним із основних правопорушень у телекомунікаційному середовищі є телефонне шахрайство, яке стрімко набирає популярності та перетворюється у справжню епідемію. Жертвами злочинців стають усі без винятку — це і бізнесмени, і чиновники, і зірки шоу-бізнесу, і звичайні громадяни. Нижче наведені основні види телефонного шахрайства та засоби боротьби із ними.

[8] Биометрия от «А» до «Я»...

Найрозповсюджений вид телефонного шахрайства — так званий «Родич у біді». Як це організовано? Людині дзвонять із невідомого номера. Злочинець представляється родичем або знайомим та схвильованим голосом повідомляє, що він затриманий співробітниками поліції та звинувачується у скоєнні того чи іншого злочину. Це може бути як дорожньо-транспортна пригода, так і зберігання зброї, наркотичних засобів, нанесення тілесних пошкоджень та навіть вбивство. Далі в розмову вступає так званий співробітник поліції, який впевненим тоном повідомляє, що неодноразово допомагав таким людям. Для вирішення питання необхідна певна сума грошей, яку слід привезти в умовлене місце та передати якомусь чоловіку. Ціна питання зазвичай складає від однієї до кількох тисяч доларів.

Схожі голоси, голоси однієї групи, які не мають відмінностей, можуть кодуватися в системах сотового зв'язку приблизно однаково, тому виявлятимуться подібними до ступеня змішування при слуховому сприйнятті та порівнянні інтегральних акустичних параметрів. Саме в такій особливості передачі мовлення по сотовому зв'язку лежать передумови здійснення телефонного шахрайства, коли при зверненні по сотовому зв'язку досить висока ймовірність помилкового впізнання чужого голосу як знайомого.

Як відомо, емоційний стан людини суттєво впливає на характеристики голосу, манеру розмови тощо. Траплялися випадки, коли, отримавши таке повідомлення, людина піддається на обман, навіть якщо особа, про яку йдеться у повідомленні, знаходиться поруч. Так, схвильовані батьки: вони завжди переймаються своїми дітьми, і реакція на можливу загрозу для них — дуже сильна. Проста, ефективна та нахабна схема, яка використовує сильні почуття та базові інстинкти. На такому ж принципі оснований вид телефонного шахрайства — «Мамо, у мене проблеми, не дзвони, перекажи гроші на цей рахунок».

Подібно відбиткам пальців у судовій криміналістичній експертизі слідів рук (дактилоскопічна експертиза), у експертизі відео-, звукозапису використовують свої об'єкти судової експертизи, а саме відео-, звукозаписи, зафіксовані на носіях інформації. У зв'язку з цим для судової криміналістичної експертизи використовують свої специфічні методи і технічні засоби.

Фізичною основою верифікації за голосом служить анатомія мовленнєвого тракту, властивості системи управління артикуляцією і особливості голосового джерела. Анатомія тракту визначає спектральні характеристики звуків мовлення, система управління артикуляцією впливає на темп мовлення, швидкість перехідних процесів і тривалість мовленнєвих сегментів, а голосове джерело визначає частоту основного тону і тембральні характеристики мовленнєвого сигналу. Досліджуються тільки такі ознаки, які можуть бути безпосередньо виміряні в мовленнєвому сигналі. Разом з тим, як показують результати досліджень, верифікація дикторів у просторі акустичних параметрів забезпечує

характеристики, що задовольняють та найбільше підходять до оточення реального застосування.

При діагностиці використовують відомі ознаки класу або групи об'єктів і порівнюють їх з ознаками конкретного об'єкта, в результаті чого визначається приналежність його до даного класу або групи. Для досягнення мети експерти вирішують різні види завдань.

Важливою характеристикою системи голосової ідентифікації є стійкість. Під перешкодами розуміються спотворення, шуми, імпульсні перешкоди тощо. Сучасні методи класифікації, що використовуються в системах голосової ідентифікації, дуже чутливі до шуму, що призводить до зниження надійності при впливі шуму. Прикладом спотворення каналу може бути реверберація звуку, тобто звук багато разів відбивається від предметів у приміщенні. Навколишній фон, в деяких випадках будучи перешкодою, може мати значний вплив на голосову ідентифікацію. Він може мати «значний» рівень сигналу (наприклад, звуки транспортних засобів; звуки, вироблені пристроями та засобами побутового призначення; звуки, властиві механізмам, приборам, апаратам, пристроям та засобам, які супроводжують роботу цих джерел; звуки живої природи; звуки явищ природи; звукові сигнали механізмів; звуки приборів, апаратів та пристроїв, спеціально призначених для створення, посилення та випромінювання звуку тощо) та «перекривати» діапазон мовленнєвого сигналу. У телефонному каналі перешкодами можуть бути клацання, перевантаження, музичні сигнали (тональні сигнали) тощо.

Оскільки сучасні цифрові мобільні пристрої зазвичай мають вбудований мікрофон і продуктивні апаратні засоби, то створення системи автентифікації за голосом із залученням більш витратних за обчисленнями методів цілком вирішуване завдання для мобільних платформ. Проте забезпечити мінімальні обчислювальні витрати при збереженні точності, завадостійкості до різних видів перешкод і достатню надійність при поширених апаратних засобах все ж необхідно.

Ідентифікація диктора — процес, за допомогою якого система може визначити, хто є диктором на основі інформації з мовленнєвого сигналу.

Останнім часом в інтегрованих системах безпеки, системах відеоспостереження, системах охоронного телебачення (СОТ) широко застосовується також звукозапис.

Інформація, що отримується, використовується і для виявлення порушників, і для аналізу стану аудіообстановки з метою контролю дій персоналу та охорони. Аудіоінформація використовується також у системах передачі інформації (СПІ) телефонних переговорів, у системах оповіщення, тривожного виклику тощо. У зв'язку з цим актуальним стає вирішення завдань, що пов'язані з аналізом звукової інформації, яка отримана під час запису в системах безпеки і яка може використовуватися надалі для аналізу в спеціалізованих лабораторіях

правоохоронних органів, лабораторіях і центрах судової експертизи, науково-дослідних і навчальних центрах із метою:

— ідентифікації особи за допомогою записаної фонограми;

— аналізу шумового фону, діагностики акустичної обстановки й умов проведення звукозапису;

— ідентифікації засобів звукозапису;

— підвищення якості та розбірливості у вже існуючих записів;

— захисту мовного сигналу від несанкціонованого доступу;

— стискання мовних повідомлень;

— встановлення дослівного змісту низькоякісних записів.

Під час вирішення завдань охорони фізичних об'єктів та інформаційних ресурсів від кримінальних і терористичних загроз дуже цікавим є використання аудіоінформації (звукових голосових записів) у системах контролю та управління доступом (СКУД).

Особливість таких систем полягає в тому, що вони допускають віддалену (за допомогою телефону) та приховану автентифікацію, що інколи є єдиним можливим засобом встановлення особистості співрозмовника.

Розглянемо найпоширеніші сучасні автоматичні методи розпізнавання диктора, що лежать в основі голосових біометричних систем.

Метод порівняння статистик основного тону[9]. Основний тон є одним з базових параметрів мовленнєвого сигналу і при цьому не дуже залежить від умов запису і типу каналу. У зв'язку з цим метод розпізнавання дикторів, заснований на параметричній статистики основного тону (ОТ), є одним з базових серед автоматичних методів розпізнавання диктора. Перевага спектрального методу, який використовується для обчислення частоти основного тону, полягає в тому, що він дозволяє оцінити частоту основного тону, використовуючи всю доступну частотну смугу сигналу. Амплітудно-частотна характеристика (АЧХ) каналу запису фонограми завжди значно впливає на форму спектра. Даний вплив необхідно враховувати, оскільки в іншому випадку АЧХ каналу, з одного боку, може замаскувати індивідуальні параметри голосу диктора, а з іншого, замаскувати частину спектра мовленнєвого сигналу і зробити його недоступним для подальшого біометричного розпізнавання. До ключових факторів, що впливають на ефективність автентифікації методом на основі аналізу статистики основного тону, слід віднести наступні:

— обсяг досліджуваного мовленнєвого матеріалу та обсяг мовленнєвого матеріалу в зразках голосу й мовлення диктора (ефективність системи помітно вище, якщо використовується запис

[9] *Матвеев Ю. Н., Симончик К. К.* Система идентификации дикторов по голосу для конкурса NIST SRE 2010 // Труды 20-й межд. конф. по компьютерной графике и зрению «ГрафиКон'2010». Санкт-Петербург, 2010.

тривалості мовлення, що є достатнім для отримання достовірних статистик основного тону);

— наявність (як для досліджуваних, так і для відео-, фонограм зі зразками усного мовлення) отриманих при одному і тому ж емоційному стані диктора);

— наявність (як для досліджуваних, так і для відео-, фонограм зі зразками усного мовлення), отриманих при одному і тому ж фоновому оточенні (в разі, коли вплив оточення настільки великий, що змінюється стиль мовлення диктора);

— відношення сигнал / шум;

— відсутність реверберації на записах.

Метод спектрально-формантного аналізу[10]. Метод ідентифікації дикторів на основі спектрально-формантних ознак здійснює порівняння досліджуваних записів природного мовлення з аналогічними зразками із мовленнєвої бази даних еталонів шляхом аналізу положення формант.

Основні етапи роботи:

— передобробка вхідного звукового файлу, що включає: видалення пауз; нормалізацію на канал; побудову моделей дикторів з використанням формантних векторів як вхідних ознак; побудову SVM (Support Vector Machine — метод опорних векторів) моделі дикторів;

— порівняння моделей дикторів.

Алгоритм методу ідентифікації дикторів на основі порівняння спектрально-формантних уявлень складається з наступних блоків:

— нормалізація на канал;

— обчислення ідентифікаційних моделей дикторів;

— обчислення індивідуальних порогів прийняття рішення «свій/чужий» шляхом порівняння отриманої ідентифікаційної моделі зі стандартним набором еталонних моделей свідомо «чужих» дикторів;

— безпосереднє порівняння отриманої ідентифікаційної моделі з моделлю з бази даних і прийняття рішення «свій/чужий» відповідно до індивідуальних порогів і заданими ймовірностями помилкової тривоги і пропуску цілі.

Метод аналізу суміші гаусових розподілів[11]. Суміші гаусових розподілів (СГР) сьогодні є одним з основних підходів при вирішенні завдання голосової біометричної автентифікації. Модель гаусової суміші голосу диктора забезпечує вірогідну модель основних звуків, що містяться в мовленні диктора. Для уявлення великого об'єму експериментальних розподілів, як базису, використовується лінійна комбінація гаусових функцій. Основною перевагою даного методу є

[10] *Koval S. L.* Formants matching as a robust method for forensic speaker identification. SPECOM'2006 Proceedings. XI International Conference «Speech and Computer», (SPECOM'2006), 25–29 June 2006. St. Petersburg, Russia. Pp. 125–128.

[11] Building Synthetic Voices, Festvox, Carnegie Mellon University. URL : http://www.festvox.org/flite/doc/index.html (дата звернення 08.05.2019).

можливість формування гладких апроксимацій експериментальних розподілів компоненту акустичного простору, форма яких може мати довільну форму.

Основною складністю при вирішенні завдання біометричної автентифікації для систем, заснованих на методі СГР, є нівелювання впливу неузгодженості, внесеного перешкодами, що містяться в каналі, що використовується під час запису. Причинами цієї неузгодженості можуть бути: шуми навколишнього середовища під час запису, спотворення в каналах запису і передачі мовного сигналу, мінливість голосу диктора з плином часу.

Під каналом у даному випадку розуміється така сукупність ефектів: спотворення, що вносяться записуючою апаратурою; вплив мікрофона пристрою, що використовується для отримання індивідуальних голосових біометричних характеристик диктора; вплив АЧХ каналу з'єднання. Модель гаусової суміші голосу диктора зручна для моделювання характеристик голосу диктора, каналу звукозапису, навколишнього середовища. Кожен з компонентів моделі відображає деякі загальні, але індивідуальні для кожного диктора особливості голосу. Саме тому даний підхід можна успішно застосовувати для вирішення задачі ідентифікації диктора.

Для того щоб побудувати модель диктора, необхідно точно оцінити її параметри, які найбільш точно відповідають розподілу векторів ознак навчального висловлювання. Існує певний ряд методів для оцінки параметрів моделі. Одним з найбільш популярних і таким, що добре себе зарекомендував, є метод оцінки максимальної правдоподібності. Мета оцінки за даним методом полягає у визначенні параметрів моделі, які максимально підвищать ймовірність правдоподібності моделі при заданих даних для навчання.

Модель являє собою ефективний алгоритм, який дає змогу проводити ідентифікацію з високою точністю розпізнавання. Однак виникає ряд проблем, які пов'язані з вибором числа компонентів моделі та ініціалізацією її початкових параметрів.

Метод спільного факторного аналізу[12]. Дане рішення полягає в застосуванні спільного факторного аналізу (Joint Factor Analysis, JFA), який дозволяє в окремому проголошенні диктора ефективно відокремлювати канальну інформацію від дикторської інформації. Це дозволяє будувати каналонезалежну GMM-модель мовлення диктора і пригнічувати ефекти каналу звукових даних, за якими відбувається побудова моделі диктора.

Метод матриці повної мінливості[13]. Одним з варіантів вирішення проблеми великого розміру моделей диктора є застосування

[12] *Щемелинин В. Л.* Методика и комплекс средств оценки эффективности аутентификации голосовыми биометрическими системами : дис. … канд. техн. наук. Санкт-Петербург, 2015. С. 22.

[13] Там само. С. 23.

низькорозмірних векторів ознак. Так, в модифікованої версії JFA для генерації векторів ознак використовується матриця повної мінливості (Total Variability, TV). Дана модифікована версія JFA часто називається TV-методом автоматичного розпізнавання дикторів.

Даний метод отримав широке поширення як найбільш перспективний метод розпізнавання дикторів, що забезпечує приведення великорозмірних вхідних даних до низькорозмірного вектору ознак, із забезпеченням збереження більшої частини корисної інформації. Це дозволяє знизити обсяг моделі диктора до 2–3 кбайт, що найчастіше є прийнятним при вирішенні завдання ідентифікації диктора на великий збір або побудову системи, що вимагає передачі моделей диктора вільними каналами зв'язку.

Метод імовірнісного лінійного дискримінантного аналізу[14]. Другим варіантом вирішення проблеми неузгодженості є модифікація JFA методу, що містить додаткову операцію на основі імовірнісного лінійного дискримінантного аналізу (Probabilistic Linear Discriminative Analysis, PLDA). Це дозволяє більш ефективно нівелювати ефект канальних перекручувань при вирішенні завдання голосового розпізнавання диктора.

Метод динамічної трансформації часової шкали Dynamic Time Warping (DTW)[15]. Цей метод дає змогу знайти близькість між двома послідовностями вимірювань за деякий проміжок часу. Вперше цей метод був застосований у розпізнаванні мови для визначення того, як два мовних сигнали представляють одну і ту ж вихідну виголошену фразу. У загальному випадку ці послідовності можуть бути різної довжини, і вимірювання може проводитися з різною швидкістю. Основною перевагою алгоритму DTW є простота реалізації. Проте, даний алгоритм непридатний для вирішення завдання текстонезалежної ідентифікації диктора.

Особливе місце ідентифікації особистості за голосом займає під час розслідування злочинів. Визначення таких характеристик за голосом диктора, як стать, вік, національність, діалект, емоційне забарвлення мовлення, також важливі в галузі криміналістики і антитерористичних діях. Результати ідентифікації важливі при проведенні експертиз відео-, звукозапису, при здійсненні експертного криміналістичного дослідження на основі теорії криміналістичної ідентифікації.

У сфері розпізнавання мовлення здійснено багато досліджень, але досі системи розпізнавання мовлення не є на сто відсотків точними. У

[14] *Захаров В. П., Рудешко В. І.* Біометричні технології в XXI столітті та їх використання правоохоронними органами : посіб. Львів, 2015. С. 492.

[15] Зб. наук.праць Центру воєнно-стратегічних досліджень Національного університету оборони України імені Івана Черняховського. 2018. № 3(64). С. 146. URL : https://nuou.org.ua/assets/journals/zbirnyk_cvsd/zb–3–64–2018.pdf (дата звернення 08.05.2019).

майбутньому системи розпізнавання мовлення повинні бути вільними від цих обмежень.

Зростання світового ринку розпізнавання голосу залежить від багатьох факторів. Одним з основних є збільшення попиту на послуги голосової біометрії. Високий попит голосової біометрії, яка є унікальною для будь-якої людини, має вирішальне значення у встановленні особи людини.

Біометрія й надалі буде дедалі більше застосовуватися у системах контролю за доступом, зокрема в різних освітянських і медичних закладах, біометричними все частіше будуть ставати ідентифікаційні карти, різні документи, системи голосування. Біометрія дасть змогу почуватися безпечніше, оскільки її основне завдання — боротьба зі злочинністю, тероризмом, способом збереження особистої інформації, підвищення комфорту громадян. Зручність для користувачів, простота, здатність до інтегрування з іншими методами розпізнання індивідуумів — це також досить суттєві чинники, що підтверджують доцільність застосування голосових технологій у біометричних системах як відокремлено від інших методів верифікації та ідентифікації особи, так і в комплексі з ними.

Висновки. Отже, розвиток біометричних систем — це ще один із способів йти в ногу з часом. На даний час існує певна кількість методів, що дають змогу вирішувати завдання ідентифікації диктора за голосом, причому кожен із наведених методів має свої переваги та недоліки. Проте найбільш поширеним методом є модель гаусових сумішей. Моделі гаусових сумішей добре себе зарекомендували як стохастична модель для побудови систем розпізнавання. Вони зручні не тільки для моделювання характеристик голосу диктора, але і каналу звукозапису, навколишнього середовища. Окремі компоненти моделі можуть моделювати окрему множину акустичних ознак. Кожний з компонентів моделі відображає як загальні, так і індивідуальні для кожного диктора особливості голосу. Саме тому даний підхід може бути успішно застосований для вирішення завдання текстонезалежної ідентифікації диктора**.**

Таким чином, голосова ідентифікація особи в певних умовах має істотні переваги, які необхідно розвивати, особливо в Україні. Біометрія дасть змогу почуватися безпечніше, оскільки її основне завдання — боротьба зі злочинністю, тероризмом та сприяння збереженню особистої інформації, підвищенню комфорту громадян.

References

Biometricheskie sistemyi identifikatsii i autentifikatsii. URL: http://mirznanii.com/a/309729/biometricheskie-sistemy-identifikatsii-i-autentifikatsii (data zvernennia 04.06.2019) [in Russian].

Biometriya ot «A» do «Ya»: polnoe rukovodstvo biometricheskoy identifikatsii i autentifikatsii. URL: https://securityrussia.com/blog/ biometriya.html (data zvernennia 23.05.2019) [in Russian].

Building Synthetic Voices, Festvox, Carnegie Mellon University. URL: http://www.festvox.org/flite/doc/index.html (data zvernennia 08.05.2019).

Entsyklopediia suchasnoi Ukrainy. URL: http://esu.com.ua/search_articles.php?id=35327 (data zvernennia 06.05.2019) [in Ukranian].

Huang, X., Acero, A., Hon, H. (2001). *Spoken language processing: a guide to theory, algorithm, and system development.* Prentice Hall PTR, 2001.

ISO/IEC 2382–37:2012. (Information technology — Vocabulary — Part 37:Biometrics, IDT) URL: https://www.iso.org/standard/55194.html (data zvernennia 13.05.2019).

ISO/IEC 2382–37–2016 [ISO/IEC 2382–37:2012] (Information technology — Vocabulary–Part37:Biometrics, IDT) URL: https://meganorm.ru/Data2/1/4293747/ 4293747566.pdf (data zvernennia 13.05.2019).

Koval, S. L. (2006). *Formants matching as a robust method for forensic speaker identification.* SPECOM'2006 Proceedings. XI International Conference «Speech and Computer», (SPECOM'2006), 25–29 June 2006. St. Petersburg, Russia, 125–128.

Matveev, Yu.N., Simonchik, K.K. (2010). *Sistema identifikatsii diktorov po golosu dlya konkursa NIST SRE 2010*. Trudyi 20-y mezhd. konf. po kompyuternoy grafike i zreniyu GrafiKon'2010' Sankt-Peterburg [in Russian].

Shchemelynyn, V. L. (2015). *Metodika i kompleks sredstv otsenki effektivnosti autentifikatsii golosovyimi biometricheskimi sistemami*: diss. na soisk. nauch. stepeni kand. tehn. nauk), (18–23) Sankt-Peterburg [in Russian].

Tsarov, R. Iu., Lemekha, T. M. (2016). *Biometrychni tekhnolohii'*: navch. posib. Odessa [in Ukranian].

Zakharov, V. P., Rudeshko, V. I. (2015). *Biometrychni tekhnolohii v XXI stolitti ta yikh vykorystannia pravookhoronnymy orhanamy*: posib, Lviv [in Ukranian].

Zbirnyk Naukovykh prats Tsentru voienno-stratehichnykh doslidzhen Natsionalnoho universytetu oborony Ukrainy imeni Ivana Cherniakhovskoho. URL: https://nuou.org.ua/assets/journals/zbirnyk_cvsd/zb–3–64–2018.pdf (data zvernennia 12.06.2019) [in Ukranian].

К. В. Луценко, К. В. Никулин

ГОЛОСОВАЯ ИДЕНТИФИКАЦИЯ ДИКТОРА КАК ОДИН ИЗ СОВРЕМЕННЫХ БИОМЕТРИЧЕСКИХ МЕТОДОВ ИДЕНТИФИКАЦИИ ЛИЧНОСТИ

В статье рассмотрены наиболее распространенные биометрические системы идентификации лиц, в том числе голосовой идентификации диктора на видео, звукозаписи.

Приведены области использования биометрических технологий и их общие характеристики. Приведенный обзор использования групп признаков при идентификации, характеризующих голос. Также в статье приведено распределение систем голосовой идентификации на соответствующие классы.

Рассмотрены основные преимущества голосовой биометрии такие, как простота реализации системы, низкая стоимость (самая низкая среди всех биометрических методов) не требует контакта, голосовая биометрия позволяет осуществлять верификацию на большом расстоянии в отличие от других биометрических технологий.

Сделан анализ существующих методов распознавания речевой информации, которые идентифицируют личность по совокупности уникальных характеристик голоса, определение их слабых и сильных сторон, на основании которых осуществлен выбор наиболее целесообразного метода для решения задачи текстонезависимого распознавания, а именно модель гауссовых смесей.

Предпосылкой развития голосовых технологий является значительное увеличение вычислительных возможностей, объема памяти при значительном

уменьшении габаритов компьютерных систем. Следует также отметить развитие математических методов, позволяющих выполнить необходимую обработку аудиосигнала путем выделения из него информативных признаков.

Развитие информационных технологий и множество практических применений, в которых используются технологии идентификации по голосу, делают эту область актуальной для дальнейшего теоретического и практического исследования.

Ключевые слова: биометрические технологии идентификации, голосовая идентификация диктора, верификация личности, автоматические методы распознавания диктора, экспертиза видео-, звукозаписи.

K. Lutsenko, K. Nikulin

VOICE SPEAKER IDENTIFICATION AS ONE OF THE CURRENT BIOMETRIC METHODS OF IDENTIFICATION OF A PERSON

The article deals with the most widespread biometric identification systems of individuals, including voice recognition of the speaker on video and sound recordings. The urgency of the topic of identification of a person is due to the active informatization of modern society and the increase of flows of confidential information.

The branches of the use of biometric technologies and their general characteristics are given. Here is an overview of the use of identification groups that characterize the voice. Also in the article the division of voice identification systems into the corresponding classes is given.

The main advantages of voice biometrics such as simplicity of system realization are considered; low cost (the lowest among all biometric methods); No need for contact, the voice biometry allows for long-range verification, unlike other biometric technologies.

The analysis of existing methods of speech recognition recognition identifying a person by a combination of unique voice characteristics, determining their weak and strong points, on the basis of which the choice of the most appropriate method for solving the problem of text-independent recognition, Namely the model of Gaussian mixtures, was carried out.

The prerequisite for the development of speech technologies is a significant increase in computing capabilities, memory capacity with a significant reduction in the size of computer systems. It should also be Noted the development of mathematical methods that make it possible to perform the Necessary processing of an audio signal by isolating informative features from it.

It has been established that the development of information technologies, and the set of practical applications, which use voice recognition technologies, make this area relevant for further theoretical and practical research.

Key words: biometric identification technologies, Voice Identification Speaker, person verification, automatic speech recognition methods, examination of video, audio recordings.

Надійшла до редколегії 10.06.2019