

ПРОГРАМНО-АПАРАТНИЙ КОМПЛЕКС ДЛЯ ЗАДАЧІ АУТЕНТИФІКАЦІЇ ОСОБИСТОСТІ ЗА ГОЛОСОВИМ СИГНАЛОМ**О. Ю. Мешков**

Херсонський національний технічний університет

вул. Бериславське шосе, 24, м. Херсон, 73008, Україна. E-mail: alexunder.meshkov@gmail.com

Голосова аутентифікація є одним із актуальних напрямів біометричної аутентифікації. Автором розроблено ряд методів та алгоритмів аналізу голосового сигналу людини та виділення інформаційно значущих параметрів голосу для задачі аутентифікації особистості. Також досліджено можливість використання різних типів мікрофонів та попередніх підсилювачів для акустичного запису голосового сигналу належної якості. З метою поєднання процедури акустичного запису, аналізу голосового сигналу та прийняття аутентифікаційного рішення, розроблено єдиний програмно-апаратний комплекс. Подано структуру розробленого комплексу та обґрунтовано доцільність використання кожної його складової. Подано узагальнений алгоритм роботи програмної частини комплексу у вигляді блок-схеми для скрипт-файлу для пакету прикладних математичних програм Sci-Lab 6.0.1.

Ключові слова: голосовий сигнал, аутентифікація, програмно-апаратний комплекс.

ПРОГРАМНО-АПАРАТНИЙ КОМПЛЕКС ДЛЯ ЗАДАЧІ АУТЕНТИФІКАЦІЇ ЛИЧНОСТІ ПО ГОЛОСОВОМУ СИГНАЛУ**А. Ю. Мешков**

Херсонский национальный технический университет

ул. Бериславское шоссе, 24, г. Херсон, 73008, Украина. E-mail: alexunder.meshkov@gmail.com

Голосовая аутентифікація являється одним із актуальних напрямів біометричної аутентифікації. Автором розроблено ряд методів та алгоритмів аналізу голосового сигналу людини та виділення інформаційно значущих параметрів голосу для задачі аутентифікації особистості. Також досліджено можливість використання різних типів мікрофонів та попередніх підсилювачів для акустичного запису голосового сигналу належної якості. З метою поєднання процедури акустичного запису, аналізу голосового сигналу та прийняття аутентифікаційного рішення, розроблено єдиний програмно-апаратний комплекс. Приведена структура розробленого комплексу та обґрунтовано доцільність використання кожної його складової. Подано узагальнений алгоритм роботи програмної частини комплексу в вигляді блок-схеми для скрипт-файлу для пакету прикладних математичних програм Sci-Lab 6.0.1.

Ключевые слова: голосовой сигнал, аутентифікація, програмно-апаратний комплекс.

АКТУАЛЬНІСТЬ РОБОТИ. Актуальність теми аутентифікації особистості обумовлена активним розвитком сучасних програмних технологій, значною інформатизацією суспільства та автоматизацією більшості сфер життя та діяльності людини. При цьому сучасний стан технологій аутентифікації свідчить про значний ухил в бік біометричних технологій. У першу чергу це пояснюється їх зручністю, надійністю та достовірністю.

Одним із доволі поширених методів біометричної аутентифікації є голосова аутентифікація. Голосовий сигнал людини – це природний сигнал, який формується голосовим апаратом людини. Він формується за рахунок модуляції повітряного потоку, що виходить з легень, за рахунок роботи ряду органів. З огляду на те, що фізіологія та анатомія голосового апарату у кожної людини є унікальною, голосовий сигнал людини також набуває індивідуальних особливостей. Тому використання цього виду сигналу для задачі аутентифікації є актуальним.

Дослідженню процедур голосової аутентифікації та аналізу голосового сигналу приділяли значну увагу вітчизняні та зарубіжні вчені, зокрема: Я.П. Драган, А.П. Кравченко, Е.Г. Жилияков, Є.А. Первушин, Д.С. Лавров, В.Н. Сорокін, Л.Р. Рабінер та інші [1–10]. У більшості випадків їх дослідження стосуються розробки нових методів аутентифікації особистості з використанням різних характеристик голосового сигналу. Серед найбільш поширених методів, використовуваних дослідниками, слід виділити

модернізовані методи, які базуються на використанні мел-частотних спектральних коефіцієнтів [1], тривалості фонем [2], вейвлет-аналізу [3], спектральні методи [4], аналіз сонограм та спектрограм голосу [5], а також методи кластерного аналізу та нейронних мереж [6–10]. Однак у розглянутих дослідженнях не відображено динамічні особливості голосового сигналу, які значно повинні впливати на актуальність та адекватність персоніфікованих голосових еталонів. Ця особливість голосового сигналу однозначно повинна бути врахована при побудові системи аутентифікації. До того ж у вказаних роботах не розглядається механізм технічної чи алгоритмічної реалізації розроблених методів. Тому розробка єдиного програмно-апаратного комплексу для задачі аутентифікації особистості за голосовим сигналом також є актуальною проблемою.

Мета роботи полягає в розробці нових методів аналізу голосу та створенні на їх основі програмно-апаратного комплексу для аутентифікації людини.

МАТЕРІАЛ І РЕЗУЛЬТАТИ ДОСЛІДЖЕНЬ. Навіть для сучасного рівня технологій аналізувати весь голосовий сигнал. Тому для задач автоматичної аутентифікації голосовий сигнал у більшості випадків параметризується, тобто подається у вигляді малої кількості інформаційно значущих параметрів. Найчастіше вони визначаються для вокалізованих ділянок голосу.

У попередніх роботах автора [11, 12] було розроблено методи виділення вокалізованих ділянок з

поток мовлення, визначення основних характеристик цих ділянок та побудови локалізованих структур голосових ділянок у спеціальному просторі характеристик «структура-частота». Також розроблено критерій аутентифікації особистості за характеристиками даних локалізованих структур.

Для кожної із сформованих таким чином локалізованих структур визначається середньозважений центр та деяка область простору навколо нього. Координати центру даної структури визначаються як середнє арифметичне значень частоти фреймів та коефіцієнтів K_T відповідно:

$$K_{Tj} = \frac{\sum_{i=1}^{n_j} K_{Ti}}{n_j}, \quad (1)$$

$$F_{0j} = \frac{\sum_{i=1}^{n_j} F_{0i}}{n_j},$$

де K_{Tj}, F_{0j} – координати центру j -ї структури; K_{Ti}, F_{0i} – координати i -ї точки j -ї структури; n_j – кількість точок в j -й структурі; $j = 1..m$ – кількість аналізованих звуків і відповідних локалізованих структур голосового сигналу у просторі характеристик.

Область простору, що виділяється навколо середньозваженого центру, має форму прямокутника. Це пов'язано з необхідністю швидкого та простого визначення меж координат точок простору характеристик, які потрапляють чи не потрапляють у дану область. Сторони кожного з цих прямокутників визначаються середньоквадратичним відхиленням точок простору, що відповідають фреймам досліджуваних сигналів:

$$\overline{S_{K_T}} = \frac{\sqrt{\sum_{i=1}^{n_j} (K_{Ti} - K_{Tj})^2}}{n_j}, \quad (2)$$

$$\overline{S_{F_0}} = \frac{\sqrt{\sum_{i=1}^{n_j} (F_{0i} - F_{0j})^2}}{n_j},$$

де $\overline{S_{K_T}}, \overline{S_{F_0}}$ – відповідні середньоквадратичні відхилення коефіцієнту структури та частоти.

Сторона кожного з утворених прямокутників приймається вдвічі більшою, ніж отримане значення середньоквадратичного відхилення відповідної характеристики голосового сигналу, оскільки відхилення враховується як в більшу, так і в меншу сторону (рис. 1.). Для кожного окремо взятого диктора система формує описані локалізовані структури голосового сигналу на основі даних його власних акустичних записів, які несуть індивідуальну інформацію про цього диктора.

При виконанні процедури аутентифікації диктора виконується одиничний запис паролльної фрази. Отриманий акустичний матеріал обробляється аналогічними алгоритмами і для нього формується такі ж локалізовані структури із середньозваженими центрами.

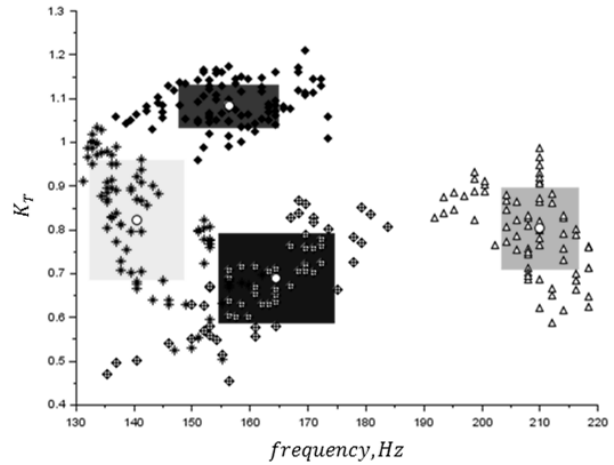


Рисунок 1 – Формування локалізованих структур голосового сигналу у просторі характеристик «структура-частота»

Критерій точності аутентифікації для такої процедури визначається як середнє значення часток площі локалізованих структур усіх фонем, які потрапили у відповідні базові локалізовані структури:

$$square_fit = \frac{\sum_{i=1}^n \frac{S_i}{S_{0i}}}{m} * 100\%, \quad (3)$$

де S_i – площа локалізованої структури аналізованої фонемі, яка потрапила до базової структури даної фонемі; S_{0i} – загальна площа локалізованої структури аналізованої фонемі; m – кількість аналізованих фонем.

Для визначення загальної площі локалізованої структури аналізованої фонемі S_0 достатньо перемножити сторони відповідного прямокутника у просторі характеристик голосового сигналу. Для визначення площі S_i необхідно спочатку перевірити, які вершини локалізованої структури аналізованої фонемі потрапляють у межі базової локалізованої структури даної фонемі. Якщо усі 4 вершини знаходяться у межах базової локалізованої структури, то коефіцієнт відношення площ для даної фонемі автоматично приймається рівним 1. Якщо ж деякі вершини виходять за межі локалізованої структури, програмний алгоритм визначає точки перетину меж базової та аналізованої структур. Ці точки обираються як вершини області перекриття і на їх основі визначаються довжини сторін області перекриття. За відомими сторонами визначається площа даної області.

Якщо отримані середньозважені центри разом із локалізованими структурами фонемі потрапляють у межі аналогічних базових структур для даного диктора, вважається, що диктор пройшов аутентифікацію. Якщо ж центр аналізованої локалізованої структури хоча б однієї фонемі не потрапляє до базової структури, або критерій точності аутентифікації має значення менше деякого визначеного рівня помилки (порогу аутентифікації), то вважається, що диктор не проходить процедуру аутентифікації.

З метою технічної та програмної реалізації даного методу аутентифікації необхідно розробити про-

грамні засоби реалізації авторських алгоритмів та забезпечити необхідний рівень якості акустичного запису голосового сигналу. З цією метою розроблено програмно-апаратний комплекс, який включає в себе апаратну та програмну частини. Кожна з цих частин має у своєму складі декілька основних бло-

ків, які пов'язані між собою комутаційним чи програмним способом. Загалом структуру розробленого програмно-апаратного комплексу можна подати у вигляді схеми на рис. 2.

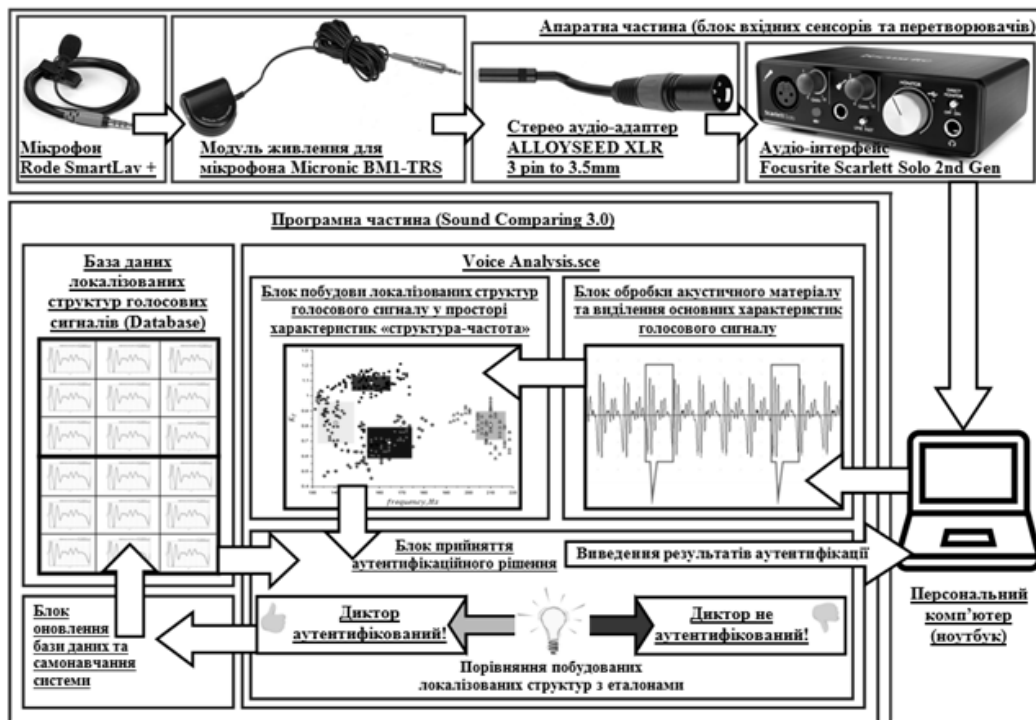


Рисунок 2 – Загальна структура програмно-апаратного комплексу аутентифікації особистості за голосовим сигналом

Апаратна частина комплексу являє собою блок входних сенсорів і перетворювачів, до якого входять петличний мікрофон Rode SmartLav +, модуль живлення для мікрофона Micronic BM1-TRS, стерео аудіо-адаптер ALLOYSEED XLR 3 pin to 3.5mm та аудіо-інтерфейс Focusrite Scarlett Solo 2nd Gen.

Петличний мікрофон Rode SmartLav + оснащений вбудованою активною електронікою у вигляді JFET-перетворення, що дозволяє йому повністю перекривати увесь голосовий частотний діапазон людини та гасити шумові складові сигналу у випадку їх наявності. Його було обрано серед інших типів

мікрофонів на основі ряду характеристик та експериментальних досліджень.

У ході експериментального дослідження автором використовувався ряд мікрофонів, серед яких:

- динамічний мікрофон Philips SBC MD 140.
- гарнітура Sven AP-525MV.
- петличний мікрофон Sennheiser ME 2-US.
- петличний мікрофон Rode SmartLav+.
- стійковий мікрофон Plantronics Audio 300.

Порівняння основних характеристик даних мікрофонів подано у табл. 1.

Таблиця 1 – Порівняння основних характеристик типів мікрофонів, що використовувались під час дослідження

№	Модель мікрофона	Тип мікрофона	Чутливість, дБ	Направленість	Частотний діапазон, Гц
1	Philips SBC MD 140	Динамічний мікрофон	-30	Кардіоїдний	20-20000
2	Sven AP-525MV	Гарнітура	-52	Однонаправлений	20-16000
3	Sennheiser ME 2-US	Петличний мікрофон	-31	Ненаправлений	40-18000
4	Rode SmartLav+	Петличний мікрофон	-35	Ненаправлений	20-20000
5	Plantronics Audio 300	Стойковий мікрофон	-58	Ненаправлений	100-8000

Як бачимо, за технічними характеристиками усі дані мікрофони практично відповідають вимогам дослідження. Тим не менше, при проведенні реальних записів голосових сигналів людини, результати записів різними мікрофонами суттєво відрізнялись.

При записі за допомогою гарнітури Sven AP-525MV спостерігалось наведення сигналу з системи живлення ноутбука. При цьому значно спотворювався корисний сигнал, знижувався його рівень, а на записі з'являлись побічні акустичні ефекти (луна, повільне згасання тощо). Тому для запису голосових сигналів людини даний мікрофон не придатний.

Записи, отримані за допомогою студійного мікрофона Philips SBC MD 140, були позбавлені описаних вище недоліків. Це може бути пов'язане з наявністю екранованого кабелю у конструкції даного мікрофона. При зміні орієнтації мікрофона на малі кути відносно ротового отвору диктора чутливість мікрофона змінювалась незначно. Аналогічна ситуація спостерігалась для петличного мікрофона Sennheiser ME 2-US.

При використанні петличного мікрофона Rode SmartLav+ якість записів після обробки була найкращою. Це пояснюється ідеальним підбором характеристик – частотного діапазону мікрофона та його рівномірною чутливістю у всіх напрямках. До того ж, у даному мікрофоні конструкційно передбачено пониження рівня власних шумів, а також активна електроніка у вигляді JFET-перетворення. Трішки гірші показники мав стаціонарний мікрофон Plantronics Audio 300, що пояснюється його більш вузьким частотним діапазоном. Однак його перевагою перед мікрофоном типу Rode SmartLav+ є значно більша зручність у використанні при виконанні людиною певних дій чи професійних обов'язків. Тим не менше, за необхідності мікрофон Rode SmartLav+ можна закріпити і в безпосередній близькості до ротового отвору у якості гарнітури, тому цей недолік можна технічно нівелювати. З урахуванням вищенаведених фактів для подальшого дослідження акустичні записи виконувались за допомогою петличного мікрофона Rode SmartLav+, закріпленого у безпосередній близькості до ротового отвору диктора.

Дискретизація та оцифрування аналогового сигналу, записаного за допомогою мікрофону виконується за допомогою аудіо-інтерфейсу Focusrite Scarlett Solo 2nd Gen. Оскільки даний аудіо-інтерфейс не передбачений для безпосереднього підключення мікрофонів з роз'ємами 3.5mm, яким оснащений мікрофон Rode SmartLav +, для комутації між ними використовується стерео аудіо-адаптер ALLOYSEED XLR 3 pin to 3.5mm.

Аудіо-інтерфейс Focusrite Scarlett Solo 2nd Gen має вбудований мікрофонний передпідсилювач з фантомним живленням 48 В, що дозволяє підключати професійні конденсаторні мікрофони, яким є Rode SmartLav +. Розрядність в 24 біти та максимальна частота дискретизації до 192 кГц дозволяють виконувати записи з дуже високою якістю. Також даний аудіо-інтерфейс було обрано за результатами первинних експериментальних досліджень, оскільки досить часто при записі на вбудовані звукові карти персональних комп'ютерів чи ноутбуків спостерігалось наведення сигналів з мережі. У даному ж аудіо-інтерфейсі наявний вбудований захист аналогових схем на вході та виході від перепадів напруги та наведень мережі.

Однак при безпосередньому підключенні даного мікрофону з вказаним аудіо-адаптером сигнал отримувався на рівні загального шуму. Інколи сигнал відгуку системи не спостерігався взагалі. Причиною цього є те, що RODE SmartLav + є мікрофоном з підтримкою від живлення пристрою, тобто він потребує живлення близько 5 В. З цієї причини він

адекватно працює зі смартфонами, ноутбуками та більшістю камер, і не працює з передпідсилювачем Focusrite Scarlett Solo 2nd Gen, який має лише 48 В фантомного живлення. З метою повноцінної комутації мікрофона та аудіо-інтерфейсу використовується модуль живлення для мікрофона Micronic BM1-TRS, який підключається до мікрофона.

Аудіо-інтерфейс підключається до персонального комп'ютера чи ноутбука через USB-з'єднання, і використовується як зовнішнє джерело звукозапису. При проведенні досліджень запис проводився у цифровому форматі у файлі формату *.wav.

Усі подальші операції відбуваються у програмній частині комплексу Sound Comparing 3.0. Вони виконуються за допомогою скрипт-файлу 'Voice Analysis.sce' для пакету прикладних математичних програм SciLab 6.0.1. Загальний скрипт-файл може бути умовно розглянутий як поєднання декількох основних блоків:

- блок обробки акустичного матеріалу та виділення основних характеристик голосового сигналу;
- блок побудови локалізованих структур голосового сигналу у просторі характеристик «структура-частота»;
- блок прийняття аутентифікаційного рішення;
- блок оновлення бази даних та самонавчання розробленої системи;
- база даних локалізованих структур голосових сигналів.

Робота програмної частини розробленого програмно-апаратного комплексу відбувається за розробленим авторським алгоритмом.

На першому етапі відбувається зчитування запису голосу з акустичного файлу та визначення максимумів спектру у кожному вікні сигналу. Результатом є масив даних амплітуди сигналу та масив максимумів спектру вікон сигналу.

На другому етапі відбувається пошук локальних максимумів сигналу. Результатом є масив положень локальних максимумів у сигналі.

На третьому етапі проводиться виділення меж кожної вокалізованої ділянки. У межах кожної ділянки виконується процедура фреймування та визначення динаміки основної частоти голосу. У результаті утворюється масив основних частот та амплітуд кожного фрейму для кожної вокалізованої ділянки.

На четвертому етапі виконується масштабування отриманих фреймів вокалізованих ділянок та визначення коефіцієнту середнього квадратичного відхилення для кожного фрейму на основі ряду базових сигналів. Визначається оптимальний базовий сигнал для даного диктору. Результатом є масив коефіцієнтів середнього квадратичного відхилення для кожного фрейму досліджуваного голосового сигналу.

На п'ятому етапі будуються локалізовані структури голосового сигналу у просторі характеристик і визначаються їх параметри – середньозважені центри та середньоквадратичні відхилення по всіх.

На шостому етапі визначаються критерії аутентифікації шляхом порівняння отриманих хмарних структур з базовими для даного диктора. Визначається, чи знаходяться середньозважені центри дос-

ліджуваних хмарних структур у межах відповідних базових структур, а також критерій *square_fit*.

На заключному сьомому етапі система приймає аутентифікаційне рішення. У випадку, якщо середньозважені центри усіх хмарних структур знаходяться у межах відповідних базових структур, а критерій *square_fit* не перевищує встановленого порогового значення, диктор вважається таким, що пройшов аутентифікацію. Якщо хоча б один із визначених критеріїв не відповідає вищевказаним вимогам система приймає рішення про відмову в аутентифікації.

У випадку, якщо диктор успішно пройшов аутентифікацію, дані про хмарні структури досліджуваних вокалізованих ділянок дозаписуються до відповідних даних бази. Тим самим реалізується додатковий етап роботи системи, який передбачає уточнення та самонавчання розроблених алгоритмів.

У ході дослідження для 60 дикторів, частина з яких була занесена до бази, було сформовано голосові еталони на основі різної кількості акустичних записів паролльної фрази. Запис проводився за технологією, описаною вище з використанням розробленого програмно-апаратного комплексу. Диктор вимовляв паролльну фразу у довільному стані, сидячи на стільці, тримаючи голову рівно. Запис паролльної фрази проводився у 5 сесій по дві вимови паролльної фрази послідовно. Інтервал між сесіями складав 15-20 хвилин. Дослідження проводились у різних приміщеннях з різним рівнем зашумленості, реверберації, а також з використанням різних ноутбуків у якості бази програмної частини комплексу.

Аналіз акустичних записів, отриманих за допомогою запропонованого програмно-апаратного комплексу, показує високу точність передачі акустичного матеріалу. Обране поєднання апаратних компонентів комплексу дозволяє виконувати записи з різним рівнем шуму, частково нівелюючи їх апаратним чином. Програмна реалізація авторських алгоритмів дозволяє проводити аутентифікацію особистості за час в середньому від 1 до 3 секунд, що дозволяє відносити дану систему до систем, що працюють в режимі реального часу.

ВИСНОВКИ. У роботі розв'язано актуальну науково-практичну задачу аутентифікації особистості шляхом аналізу її голосового сигналу. Дана задача була вирішена за допомогою розроблених математичних моделей з використанням методів обробки вихідного акустичного матеріалу, виділення характеристик голосового сигналу з потоку мови, побудови локалізованих структур голосового сигналу у просторі характеристик «структура-частота», формуванні на їх основі персоніфікованого голосового еталону особистості та подальшого порівняння голосових сигналів людини з даними еталонами. Усі розроблені методи та моделі об'єднано у єдиний програмно-апаратний комплекс для аутентифікації особистості за аналізом голосового сигналу.

Основні наукові та практичні результати роботи полягають у наступному:

1. Наведено авторську математичну модель голосового сигналу людини у вигляді набору даних про форму та взаємне розміщення локалізованих

структур голосового сигналу у просторі характеристик голосового сигналу.

2. Запропоновано спосіб аутентифікації диктора на основі порівняння вхідного акустичного матеріалу з персоніфікованим голосовим еталоном диктора та критерій аутентифікації за даним методом.

3. Розроблено програмне забезпечення для обробки голосового сигналу та прийняття аутентифікаційного рішення, що дозволило скоротити час прийняття аутентифікаційного рішення з високою точністю.

4. Досліджено можливість використання ряду технічних пристроїв запису та первинної обробки голосових сигналів людини в аналоговій та цифровій формі і визначено частоту дискретизації, адекватну для повноцінного перетворення аналогового голосового сигналу на цифровий без втрат інформації.

5. Розроблено програмно-апаратний комплекс для аутентифікації особистості у вигляді поєднання апаратної та програмної підсистем, основними складовими і блоками яких є блоки: вхідних сенсорів і перетворювачів, обробки голосового сигналу; виділення основних характеристик голосового сигналу; систему прийняття аутентифікаційного рішення, які в сукупності з програмним та інформаційним забезпеченням реалізують розроблені методи і моделі.

6. Досліджено можливість використання розробленого програмно-апаратного комплексу для аутентифікації особистості для дикторів різної статі, в умовах різного рівня зашумленості.

ЛІТЕРАТУРА

1. Первушин Е. А. Обзор основных методов распознавания дикторов. *Математические структуры и моделирование*. 2011. Вып. 24. С. 41–54.
2. Булгакова Е. В., Шолохов А. В., Томашенко Н. А. Метод идентификации дикторов на основе сравнения статистик длительностей фонем. *Научно-технический вестник информационных технологий, механики и оптики*. 2015. Том 15. № 1. С. 70–77.
3. Вишнякова О. А., Лавров Д. Н. Алгоритм фонемной сегментации на основе анализа скорости изменения энергии дискретного вейвлет-преобразования. *Вестник Омского университета*. 2011. № 4. С. 146–152.
4. Семенов Г. А. Выявление порога чувствительности спектрального метода к поиску периодического сигнала в нормальном шуме. *Известия АлтГУ*. 2011. №1-1. С. 192–196.
5. Цыплихин А. И., Сорокин В. Н. Сегментация речи на кардинальные элементы. *Информационные процессы*. 2006. Том 6, № 3. С. 177–207.
6. Венедиктова Е. В., Лавров Д. Н. Идентификация диктора по фиксированному набору частот с помощью линейного классификатора. *Математические структуры и моделирование*. 2008. №1 (18). С. 108–115.
7. Федоров Е. Е. Методика идентификации диктора на основе модифицированной вероятностной нейронной сети. *Наукові праці Донецького національного технічного університету*. Сер.: Інформатика, кібернетика та обчислювальна техніка. 2011. Вип. 13. С. 186–191.

8. Гапочкин А. В. Нейронные сети в системах распознавания речи. *Science Time*. 2014. №1 (1). С. 29–36.

9. Васильев Р. А. Исследование особенностей идентификации дикторов по голосу. *Известия ТулГУ. Технические науки*. 2013. № 3. С. 246–252.

10. Шевченко И. В., Шкарупа Н. С., Гончар М. В. Модель и метод построения многоуровневой системы распознавания. *Вісник Кременчуцького національного університету імені Михайла Остроградського*. Вип. 6/2017(107). Част. 1. С. 54–63.

11. Мешков О. Ю., Новіков О. О., Злепко С. М. Метод локальних максимумів для виділення вокалізованих ділянок голосового сигналу людини. *Вісник Хмельницького національного університету*. 2018. № 6. С. 197–210.

12. Мешков О. Ю., Новіков О. О., Новіков В. О. Аналіз голосових сигналів людини та аутентифікація особистості за голосом: Монографія. Херсон: вид-во ФОП Вишемирський В.С., 2018. 168 с. (9,17 друк. арк.).

SOFTWARE-HARDWARE COMPLEX FOR THE TASK OF PERSONAL VOICE AUTHENTICATION

O. Mieshkov

Kherson national technical university

vul. Beryslavske shose, 24, Kherson, 73008, Ukraine. E-mail: alexander.meshkov@gmail.com

Purpose. During the pre-study mathematical models of human voice signal as a localized human voice structures has been developed. To approbate and fulfill mathematical models of human voice for the task of personal authentication in practice a software-hardware complex has been developed. **Methodology.** The mathematical modeling of human voice taking into account the individual aspects of the signal has been performed in the specified domain of human voice characteristics such as fundamental frequency and human voice amplitude structure. For the task of human voice signal recording a set of hardware has been chosen. Each recording is processed by author's algorithms, vocal fragments of the voice are extracted, each vocal fragment is divided into separate frames. For every frame the chosen characteristics are defined. on the basis of these characteristics localized human voice structures are built. **Results.** The method of personal authentication by human voice analysis has been developed taking into account the form and location of localized human voice structures in specified human voice characteristics domain "structure-frequency". A special criterion for the task of personal authentication by this method is defined. Software-hardware complex for the task of personal authentication in the form of a combination of hardware and software subsystems, the main components and blocks of which are: input sensors and converters block, voice signal processing block; allocation of the main characteristics of the voice signal block; a system of acceptance of an authentication decision, which together with the software and information software implement the developed methods and models. **Originality.** For the first time, the system for the task of personal authentication has been formed as a combination of software and hardware parts. The software part implements human voice analysis and personal authentication algorithms, developed by author. **Practical value.** The developed complex can be used in different acoustic environments for different speakers with the required accuracy and reliability. References 11, tables 1, figures 2.

Key words: voice signal, authentication, software-hardware complex.

REFERENCES

1. Pervushin, E. A. (2011), "Overview of the basic recognition of speakers", *Mathematical structures and modeling*, vol. 24, pp. 41-54.

2. Bulgakova, E. V., Sholokhov, A. V., Tomashenko, N. A. (2015), "Method of speakers identification based on a comparison of statistics of phoneme durations", *Scientific-Technical Bulletin of Information Technologies, Mechanics and Optics*, vol. 15, no. 1, pp. 70-77.

3. Vishnyakova, O. A., Lavrov, D. N. (2011), "Phoneme segmentation algorithm based on the analysis of the rate of change of energy of a discrete wavelet transform", *Vestnik of Omsk University*. no. 4, pp. 146-152.

4. Semenov, G. A. (2011), "Identification of the threshold of sensitivity of the spectral method to the search for a periodic signal in a normal noise", *Izvestia AltGU*, no. 1-1, pp. 192-196.

5. Tsyplihin, A. I., Sorokin, V. N. (2006), "Speech segmentation into cardinal elements", *Information processes*, Volume 6, no. 3, pp. 177-207.

6. Venediktova, E. V., Lavrov, D. N. (2008), "Identification of the speaker by a fixed set of frequencies using the linear classifier", *Mathematical structures and modeling*, no. 1 (18), pp. 108-115.

7. Fedorov, E. E. (2011), "Method of speaker identification based on a modified probabilistic neural network", *Naukovi Pratsi Donetsk National Technical University. Ser.: Informatics, Cybernetics and obscuration technology*, Vol. 13. pp. 186-191.

8. Gapochkin, A. V. (2014), "Neural networks in speech recognition systems", *Science Time*, no. 1(1), pp. 29-36.

9. Vasiliev, R. A. (2013), "Investigation of features of speaker identification by voice", *Izvestiya of TSU. Technical science*, no. 3, pp. 246-252.

10. Shevchenko, I. V., Shkarupa, N. S., Gonchar, M. V. (2017) "Model and method of a multi-level recognition system construction", *Transactions of Kremenchuk Mykhailo Ostrohradskyi National University*, iss. 6(107) part 1, pp. 54-63.

11. Mieshkov, O. Yu, Novikov, O. O., Zlepko, S. M. (2018), "Method of local maximums for the task of vocal vocalization of the voice signal of the people", *Transactions of Kremenchuk Mykhailo Ostrohradskyi National Universit*, no. 6, pp. 197-210.

12. Mieshkov, O. Yu., Novikov, O. O., Novikov, V. O. (2018), *Analysis of voice signals of people and personal authentication by voice*, Kherson: Vishimirsky V.S.

Стаття надійшла 01.10.2018.