

ПРИСТРІЙ ГОЛОСОВОЇ ІДЕНТИФІКАЦІЇ ОСОБИ В СИСТЕМАХ КОНТРОЛЮ ДОСТУПУ

Метою роботи є створення пристрою який дозволить підвищити безпеку, буде мати низьку вартість та задовільну точність.

В роботі представлено мікропроцесорний пристрій голосової ідентифікації особи в системах контролю доступу. Приведена структурна схема пристрою, та процес вилучення векторів ознак з вхідного мовного сигналу для побудови відбитку.

Ключові слова: голосова біометрія, біометрична система, розпізнавання голосу, система контролю доступу.

V.O. KLEKOTIUK, L.P. DYUZHAYEV

National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute"

DEVICE VOICE IDENTIFICATION IN ACCESS CONTROL SYSTEMS

Abstract — The aim is to create a device that will improve safety, will have low cost and satisfactory accuracy.

The paper presents microprocessor unit voice identification in access control systems. Present a block diagram of the device and the process of extracting feature vectors of input speech signal for the construction footprint.

Keywords: voice biometrics, biometric system, voice recognition, access control system.

Вступ

В сучасному світі все більшого розвитку набувають біометричні системи (БС) контролю доступу користувача до певних ресурсів чи об'єктів. Головним показником БС є індивідуальна характеристика користувача, за якою надається право доступу. Одним із напрямів біометричної ідентифікації є ідентифікація людини за голосом.

Мова - це послідовність звуків. Звук в свою чергу - це суперпозиція (накладення) звукових коливань (хвиль) різних частот. Хвиля характеризується двома атрибутами - амплітудою і частотою. Для того, що б зберегти звуковий сигнал на цифровому носії, його необхідно розбити на безліч проміжків і взяти деяке «усереднене» значення на кожному з них.

Таким чином, механічні коливання перетворюються в набір чисел, придатний для обробки на сучасних ЕОМ.

Звідси випливає, що завдання розпізнавання мови зводиться до «співставлення» безлічі чисельних значень (цифрового сигналу)[1].

Постановка проблеми

На даний момент основним засобом контролю доступу є пропускні пункти, обладнані різними засобами контролю доступу. Але більшість із засобів контролю доступу мають високу ціну. Причому більша частина витрат припадає виділення персонального засобу ідентифікації кожному користувачеві. Рішенням даної проблеми може стати голосова ідентифікація. Використання біометрії дозволяє відмовитися від карт доступу, виключити втрату засоби ідентифікації і його крадіжку. А використання голосу дозволить відмовитися від дорогого устаткування для зчитування даних.

Пристрій голосової ідентифікації

Початковим етапом голосової ідентифікації є отримання голосу користувача. Для цього необхідний мікрофон, фільтр і аналого-цифровий перетворювач (АЦП), для подальшої роботи з цифровим записом голосу.

У загальному вигляді процес введення мовних повідомлень представлено на рисунку 1.

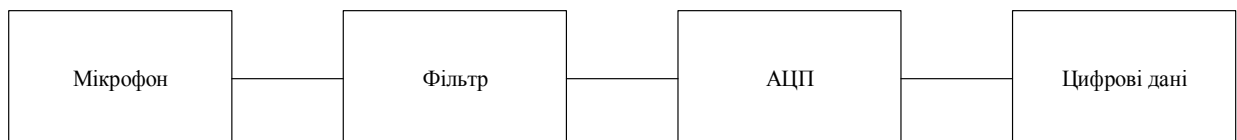


Рис. 1. Схема запису голосу

З виходу мікрофона сигнал подається на вхід блоку фільтрації. Наступним етапом є проходження АЦП.

Далі оцифрований сигнал потрапляє в блок цифрової обробки. В блоці цифрової обробки сигнал фільтрується і перетворюється в вектор, з яким в подальшому буде працювати мікропроцесор.

Також, отриманий вектор зберігається в енергонезалежній пам'яті. Це необхідно для подальшого порівняння з отриманим відбитком. Після порівняння відбитка в пам'яті з отриманим відбитком,

мікроконтролер подає команду на блок управління зовнішнім пристроєм, наприклад магнітний замок. Загальна схема пристрою представлена на рис. 2.

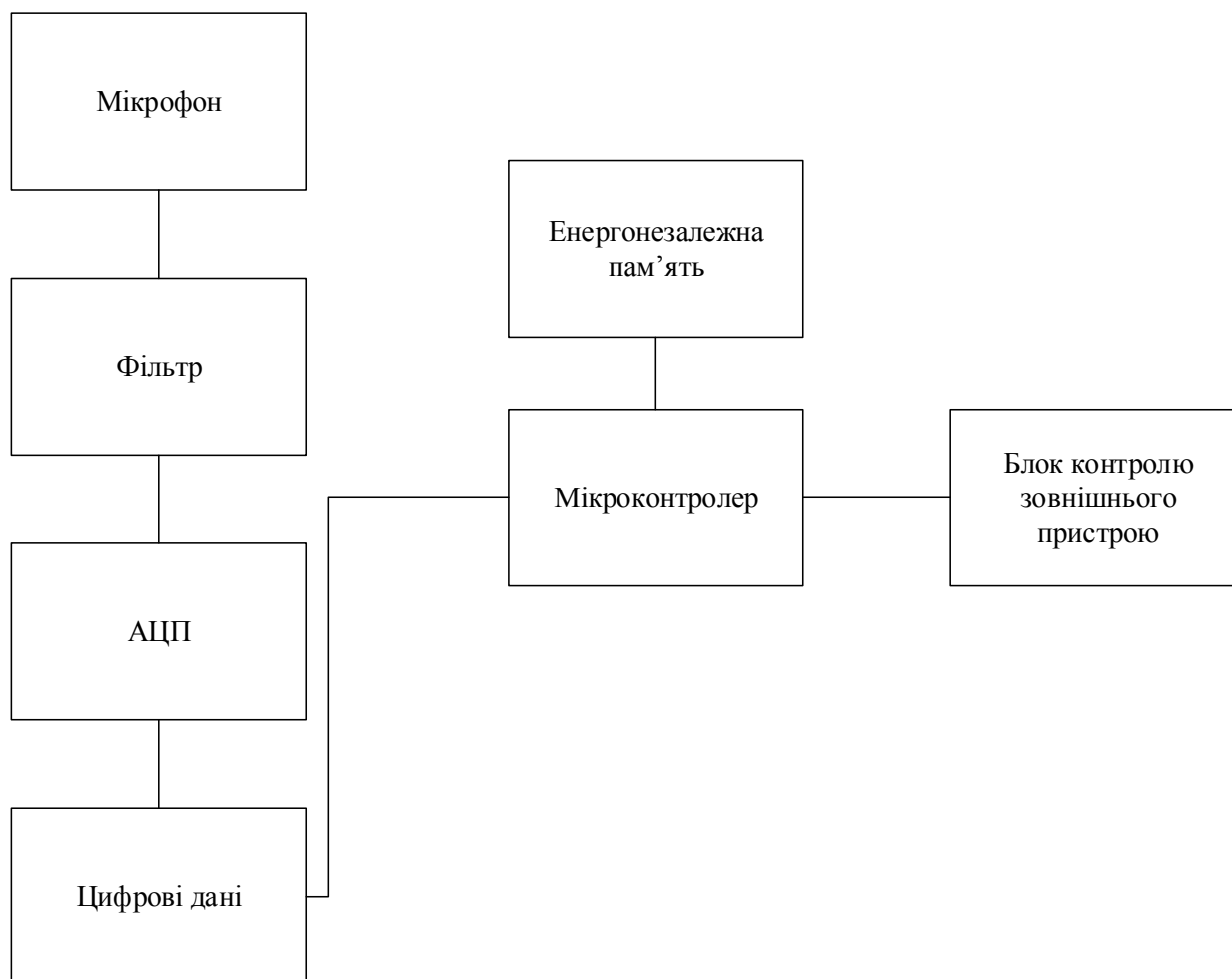


Рис. 2. Схема пристрою

Сам процес голосової ідентифікації не вимогливий до ресурсів, і складається з двох етапів. Перший крок — отримання голосових відбитків другий крок є порівняння голосових відбитків[2].

Отримання голосового відбитку

На етапі вилучення ознак відбувається перетворення вхідного мовного сигналу в послідовність векторів ознак.

На рисунку 3 представлено алгоритм вилучення ознак з вхідного мовного сигналу.

Перший крок — перетворення аналогового сигналу в цифровий. Цей процес розбивається на два етапи: дискретизація і квантування. При дискретизації сигнал розбивається на окремі вимірювання амплітуди через деякий часовий інтервал. Число вимірювань в секунду називається частотою дискретизації. Процес подання дійсних значень цілими називається квантуванням. Надалі відцифрований сигнал буде позначатися як $x[t]$, де t деякий момент часу.

Після чого відбувається розбиття відцифрованого сигналу на невеликі фрагменти (фрейми), що характеризують окремі елементи мовного сигналу, для яких можна зробити припущення, що сигнал є стаціонарним. Розбиття на фрейми проводиться таким чином: вибирається віконна функція, яка приймає нульове значення всередині деякого часового відрізка, а за його межами — нульове. Далі вона послідовно застосовується до сигналу x , таким чином, «втягується» інформація всередині фрейму. Вилучення сигналу проводиться шляхом перемноження значення сигналу $x[t]$ в момент часу t зі значенням віконної функції $w[t]$ в момент часу t :

$$y[t] = w[t] \cdot x[t],$$

де $x[t]$ — значення сигналу; $w[t]$ — значення віконної функції.

Віконна функція характеризується трьома величинами: шириною (в мілісекундах), зміщенням (число мілісекунд між границями послідовних вікон) і формою.

Наступний крок — витяг спектральної інформації для сигналу, отриманого на попередніх етапах. На цьому етапі необхідно з'ясувати, яка кількість енергії міститься в кожному частотному діапазоні. Для

вилучення спектральної інформації застосовується дискретне перетворення Фур'є.

На вхід подається розбитий на фрейми сигнал, а на виході для кожного з T частотних діапазонів – комплексне число $X[k]$, яке є амплітудою і фазою вихідного сигналу:

$$X[k] = \sum_{t=0}^{T-1} x[t] e^{-i \frac{2\pi}{T} kt},$$

де $x[t]$ — значення сигналу; T — кількість частотних діапазонів.

Потім здійснюється перехід від величини частоти звуку до значення висоти (мел) за формулою:

$$mel(f) = 1127 \ln\left(1 + \frac{f}{700}\right),$$

де f — частота.

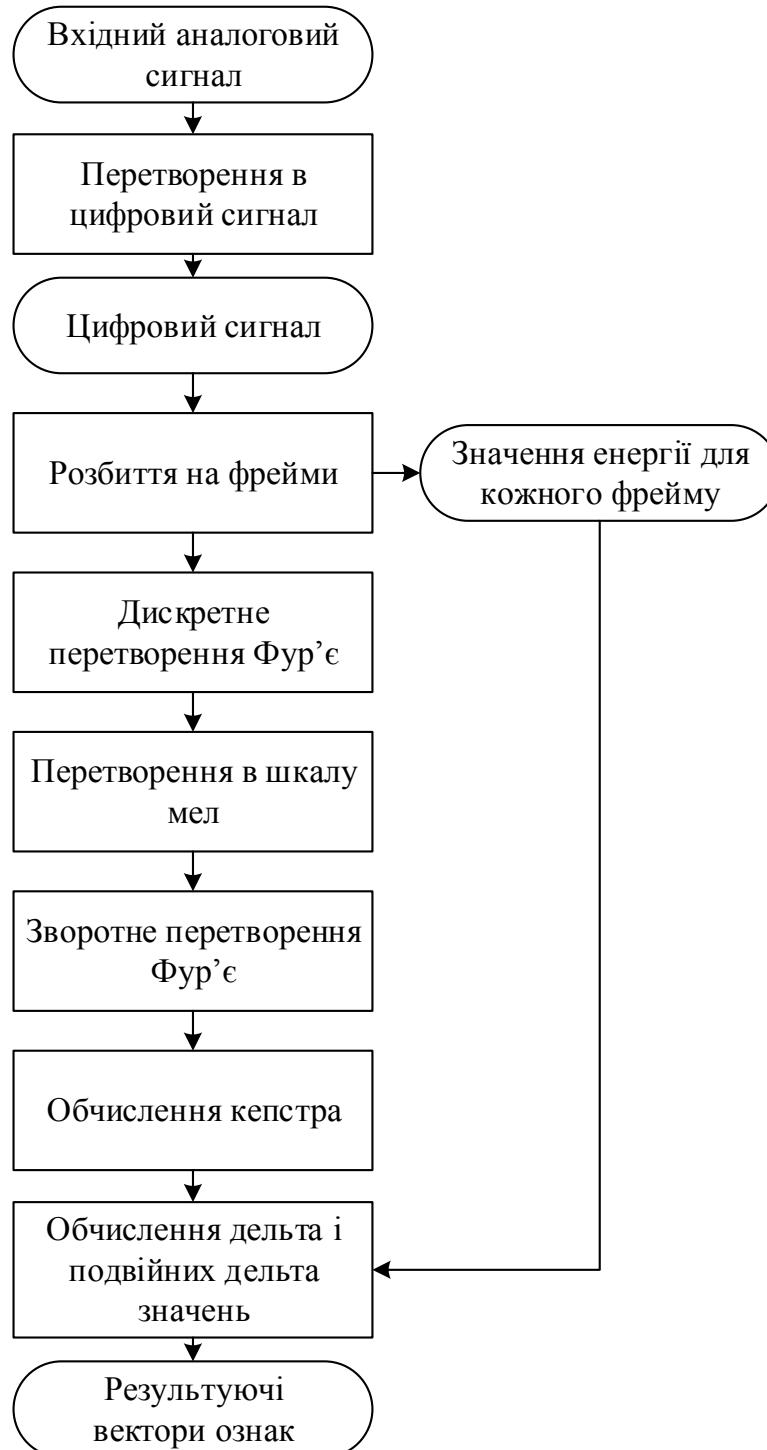


Рис. 3. Алгоритм вилучення ознак з вхідного мовного сигналу

Це необхідно для врахування значущого чинника: людський слух має неоднакову чутливість для різних частотних діапазонів. Далі формуються трикутні фільтри, в яких накопичується 20 значень енергії для кожного з частотних діапазонів (10 фільтрів розподілені лінійно нижче 1000 Гц, а решта – логарифмічно вище 1000Гц) і береться логарифм кожного отриманого значення мел. Використання логарифма робить оцінки ознак менш чутливими до відмінностей у способі подачі вхідного сигналу.

Наступним кроком є обчислення кепстра. Кепстр дозволяє відокремити джерело звукової хвилі від фільтра, властивості якого проявляються при проходженні вихідної хвилі по голосовому каналу при утворенні звуку [3]. При цьому велика частина корисної інформації міститься в фільтрі.

Формально кепстр можна визначити наступним чином:

$$c[t] = \sum_{t=0}^{T-1} \log \left(\sum_{t=0}^{T-1} x[t] e^{-i \frac{2\pi}{T} kt} \right) e^{-i \frac{2\pi}{T} kt},$$

де $x[t]$ — значення сигналу; T — кількість частотних діапазонів;

Крім того, для кожного фрейму додається ще одна ознака – енергія сигналу:

$$E = \sum_{t=1}^{t2} x^2[t].$$

Останнім етапом є обчислення дельта значень і подвійних дельта значень для отриманих на попередніх етапах ознак, які також додаються в результуючий вектор ознак. Дельта значення обчислюються за формулою:

$$d(t) = \frac{c(t+1) - c(t-1)}{2},$$

де $c(t)$ — кепстральний коефіцієнт.

Подвійні дельта значення обчислюються аналогічно, з тією лише різницею, що замість кепстральних коефіцієнтів використовуються обчислені дельта значення [4].

Висновки

В ході роботи було розроблено функціональну схему пристрою голосової ідентифікації в системах контролю доступом, та алгоритм за яким він буде працювати. Даний програмний комплекс дуже гнучкий і має великий простір для подальшого удосконалення і додавання нових функцій, що робить його не тільки вигідним програмним продуктом, але і перспективним проектом для розвитку і отримання прибутку

Література

1. Розпізнавання голосу — Режим доступу: <https://habrahabr.ru/post/226143/> — Назва з екрану
2. Голосова ідентифікація — Режим доступу: [https://libelloc.bsuir.by/bitstream/Голосова ідентифікація. PDF](https://libelloc.bsuir.by/bitstream/Голосова%20ідентифікація.PDF) — Назва з екрану
3. Дюжаєв Л. П. Аналіз впливу параметрів обробки звукового сигналу на якість розпізнавання голосових команд / Л. П. Дюжаєв, В. Ю. Коваль // Вісник НТУУ «КПІ». Сер. Радіотехніка. Радіоапаратобудування. — 2014. — №56. — С. 34—41.
4. Розпізнавання голосу — Режим доступу: https://ru.wikipedia.org/wiki/Распознавание_голоса — Назва з екрану

References

1. Rozpiznavannia holosu — Rezhim dostupu: <https://habrahabr.ru/post/226143/> — Nazva z ekranu
2. Holosova identifikatsiia — Rezhim dostupu: [https://libelloc.bsuir.by/bitstream/Holosova identifikatsiia. PDF](https://libelloc.bsuir.by/bitstream/Holosova%20identifikatsiia.PDF) — Nazva z ekranu
3. Diuzhaiev L. P. Analiz vplyvu parametriv obrobky zvukovoho sygnalu na iakist rozpiznavannia holosovyh komand / L. P. Diuzhaiev, V. Yu. Koval // Visnyk NTUU «KPI». Ser. Radiotekhnika. Radioaparotobuduvannia. — 2014. — №56. — S. 34—41.
4. Rozpiznavannia Holosu — Rezhim dostupu: [https://ru.wikipedia.org/wiki/Raspiznavaniie golosa](https://ru.wikipedia.org/wiki/Raspiznavaniie_golosa) — Nazva z ekranu

Рецензія/Peer review : 26.1.2017 р. Надрукована/Printed :30.3.2017 р.
Стаття рецензована редакційною колегією