

Вінницький національний технічний університет  
Факультет комп'ютерних систем та автоматики  
Кафедра комп'ютерних систем управління

# **Нейромережева система автентифікації суб'єкта за візуальним представленням його голосу**

Виконала: Воловодівська О.О.

Керівник: Ковтун В.В.

Вінниця ВНТУ - 2019 р.

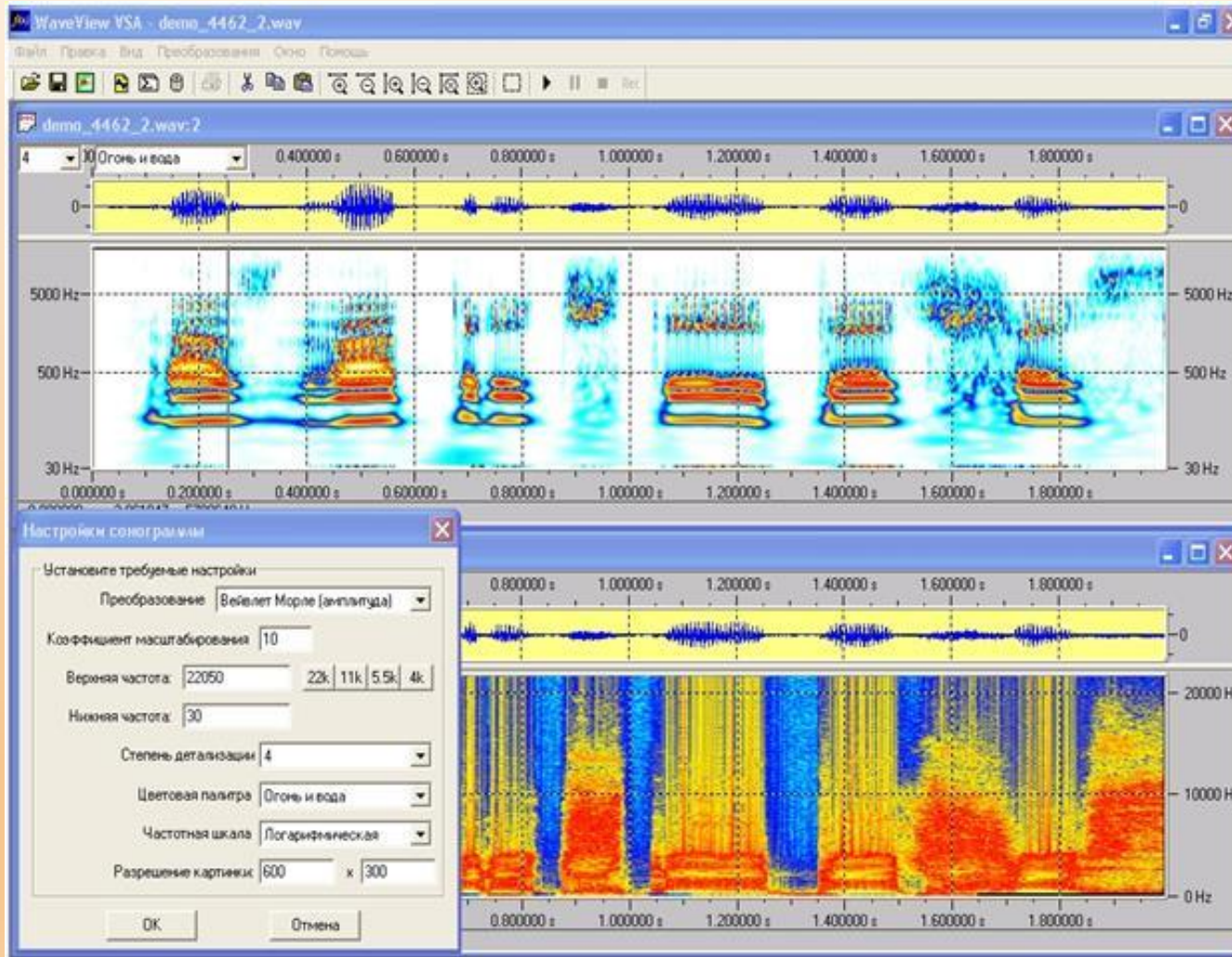
# Актуальність роботи

Можливість автентифікації суб'єкта за голосом засновано на аналізі унікальних характеристик мовного сигналу, обумовлених анатомічними властивостями (ротової порожнини, горла, голосових зв'язок тощо) та набутими мовленнєвими звичками (манера, гучність, швидкість мови).

Перевагою технології автентифікації суб'єкта за голосом забезпечується, насамперед, природністю мовного спілкування та невіддільністю мовної інформації від людини.

Складність такого підходу до автентифікації є те, що основним призначенням мовного спілкування є комунікація, отже, у мовному сигналі інформація про індивідуальні особливості голосу мовця накладаються на змістовну інформацію і для відокремлення цих складових мовний сигнал зазнає фільтрації та параметризації із представленням у вигляді спектрограми або вейвлет-діаграми.

# Актуальність роботи



Мовний  
сигнал

Вейвлет-  
діаграма

Мовний  
сигнал

Спектро-  
грама

# Об'єкт і предмет дослідження

**Об'єкт дослідження** - процес мовотворення людини.

**Предметом дослідження** - методи візуалізації мовних сигналів, математичні методи оброблювання цифрових рядів, методи цифрового оброблювання сигналів, методи теорії прийняття рішень і машинного навчання.

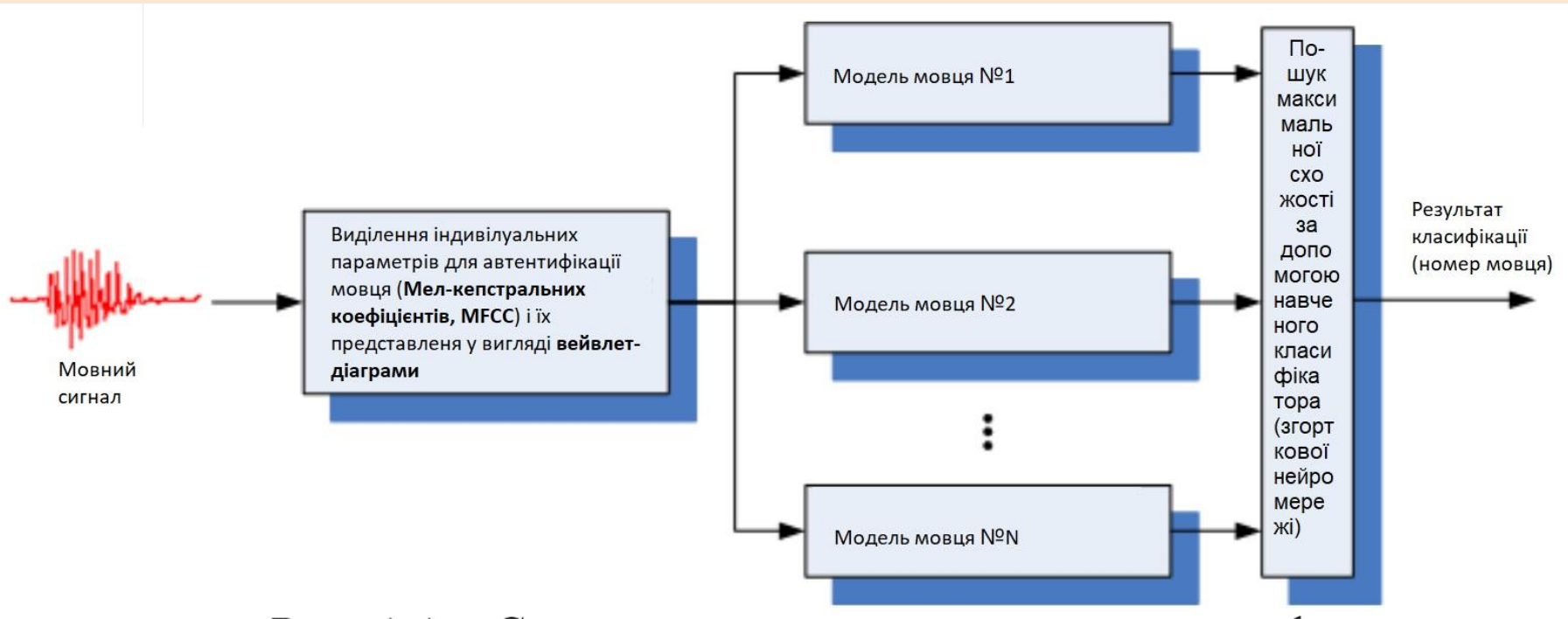
# Мета роботи

Метою є розроблення нейромережевої системи автентифікації суб'єкта за візуальним представленням його голосу

## Задачі:

1. Здійснення аналітичного огляду і аналізу досягнень в галузі систем автентифікації суб'єкта за голосом;
2. Розроблення методу візуалізації мовних сигналів в задачі автентифікації суб'єкта за голосом;
3. Дослідження ефективності розробленого методу за допомогою відповідної програмної системи для автентифікації особи суб'єкта за голосом.

# Технології розробки



Створена на основі огляду аналогів структурна схема системи автентифікації мовця за візуалізованим представленням його мовного сигналу

# Технології розробки

## Недоліки Фур'є-представлення (у вигляді спектрограми) мовного сигналу:

- втрата інформації в часі існування частотних компонентів сигналу;
- вихідний сигнал замінюється на періодичний, з періодом рівним тривалості досліджуваного зразка;
- ПФ не дає інформацію про динаміку зміни спектрального складу сигналу;
- ПФ погано працює при зміні параметрів процесу згодом (нестационарності), оскільки дає усереднені коефіцієнти для всього досліджуваного сигналу;
- ПФ вимагає знання сигналу не тільки в минулому, але й у майбутньому, що є неможливим для систем реального часу;
- погана пристосованість до сигналів з локальними особливостями.

# Технології розробки

**Вейвлет-перетворення** дозволяє судити не тільки про частотний спектр сигналу, але також про те, у який момент часу з'явилася та або інша гармоніка. З їхньою допомогою можна легко аналізувати переривчасті сигнали, або сигнали з гострими сплесками.

## **Переваги вейвлет-представлення мовного сигналу:**

локалізація в часовій і частотній області, що дозволяє робити ефективний частотно-часовий аналіз нестационарних сигналів, прикладом яких є мова;

- можливість масштабного перетворення й зсувів (стиск-розтягання досліджуваного сигналу за допомогою хвильових функцій різної періодичності);
- математичний апарат, який було розроблено для локалізації й класифікація особливих точок сигналу;
- відмінно відображають динаміку зміни сигналу уздовж „осі масштабів”, локалізацію різномасштабних деталей (спектр мови людини радикально змінюється з часом, а характер цих змін - дуже важлива інформація).

**У роботі для аналізу мовних сигналів використано вейвлет Хаара**



# Технології розробки

Після фільтрації шумів для виділення із мовного сигналу окремих елементів застосовується аналіз **енергії сигналу** на протязі кожних 10-20 мс. Обробка вхідного оцифрованого сигналу з метою зменшення обсягу вхідних даних полягає у застосуванні різних методів спектрального аналізу.

**Спектральний аналіз** даних реалізується або за допомогою віконного дискретного перетворення Фур'є або за допомогою дискретних вейвлет-перетворень. На сьогодні загальноприйнято проводити **стиснення спектру** за допомогою **процедури визначення мел-кепстральних коефіцієнтів**. Її результатом являється 16-24 коефіцієнтів, які достатньо повно характеризують весь діапазон звукових частот, які відчуває людина (від 16 Гц до 20000 Гц).

**Обчислення мел-частотних кепстральних коефіцієнтів** для обраного фрагменту мовного сигналу передбачає:

- пропускання фрагменту мовного сигналу крізь гребінку смугових фільтрів, центральні частоти яких розташовані за Мел-шкалою;
- обчислення мел-частотного кепстрального коефіцієнту всіх синалів, отриманих на виході гребінки фільтрів.

# Технології розробки

**Згорткова нейрона мережа** – це різновид **глибоких нейромереж**, яка за рахунок застосування спеціальної операції – власне, згортки – дозволяє водночас зменшити кількість інформації, що зберігається в пам'яті, і виділити опорні ознаки на досліджуваному зображенні, такі як ребра, контури або грані. На наступному рівні обробки з цих ребер і граней можна розпізнати повторювані фрагменти текстур, які далі можуть скластися в фрагменти зображення.

По суті кожен шар багат шарової згорткової нейромережі використовує власне перетворення. Якщо на перших шарах мережа оперує такими поняттями як "ребра", "грані" і т.п, то на останніх шарах використовуються поняття "текстура", "частини об'єктів".

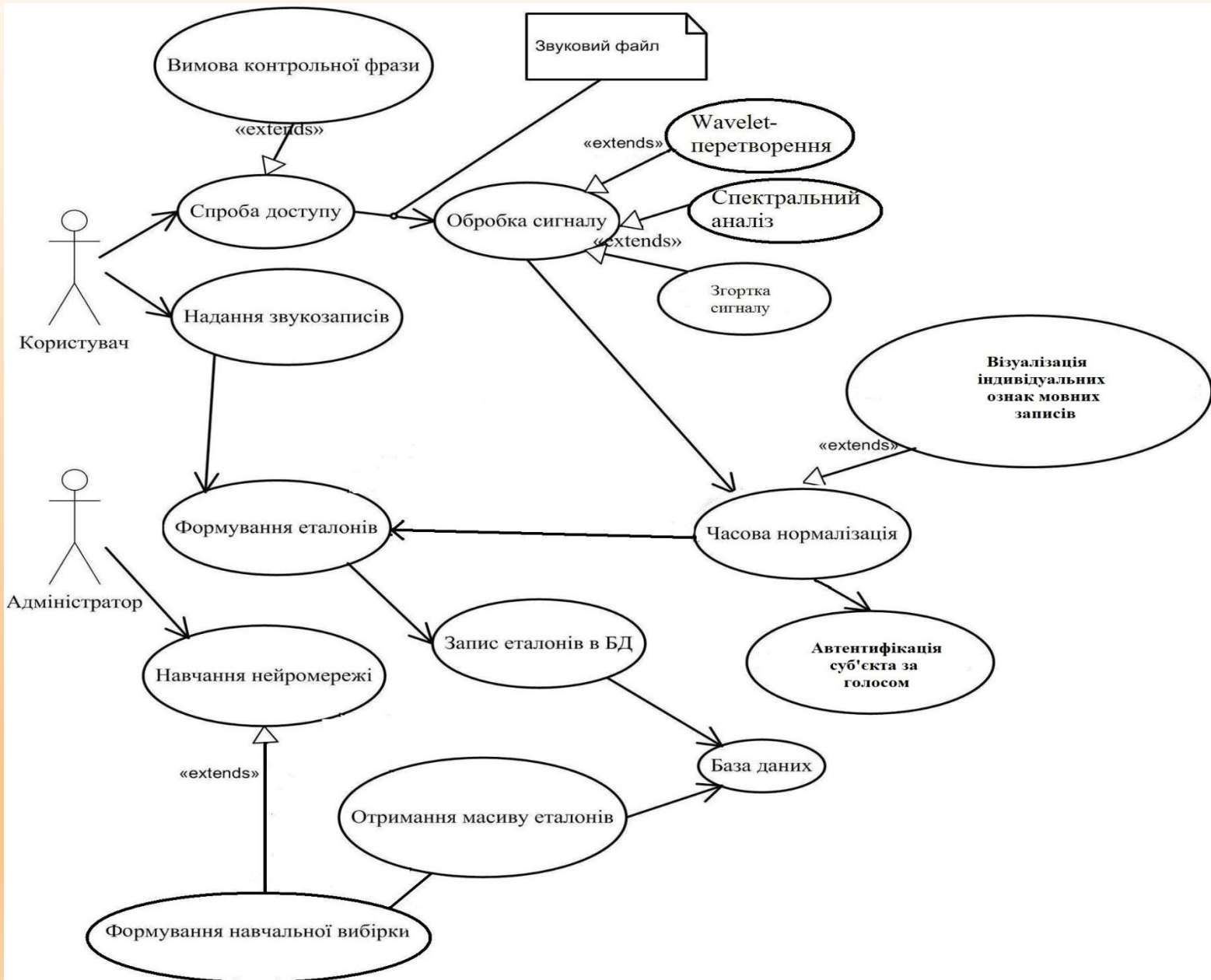
**В результаті такого опрацювання ми можемо виділити потрібний об'єкт на зображенні – індивідуальний рисунок енергії мел-частотних кепстральних коефіцієнтів, властивих конкретному мовцю.**

# Технології розробки

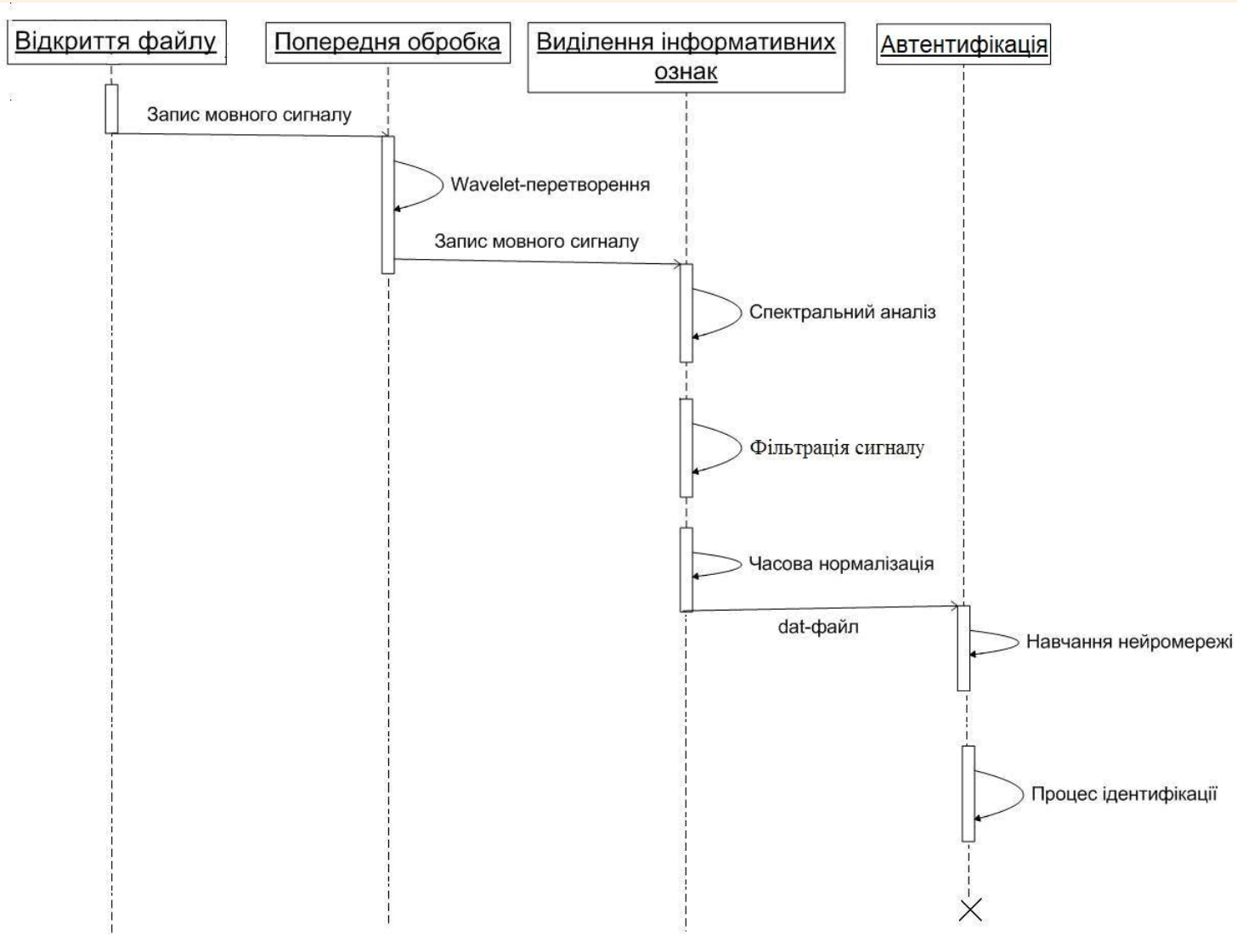
Для реалізації створеного ПЗ обрано програмне середовище **Matlab** виходячи із таких міркувань:

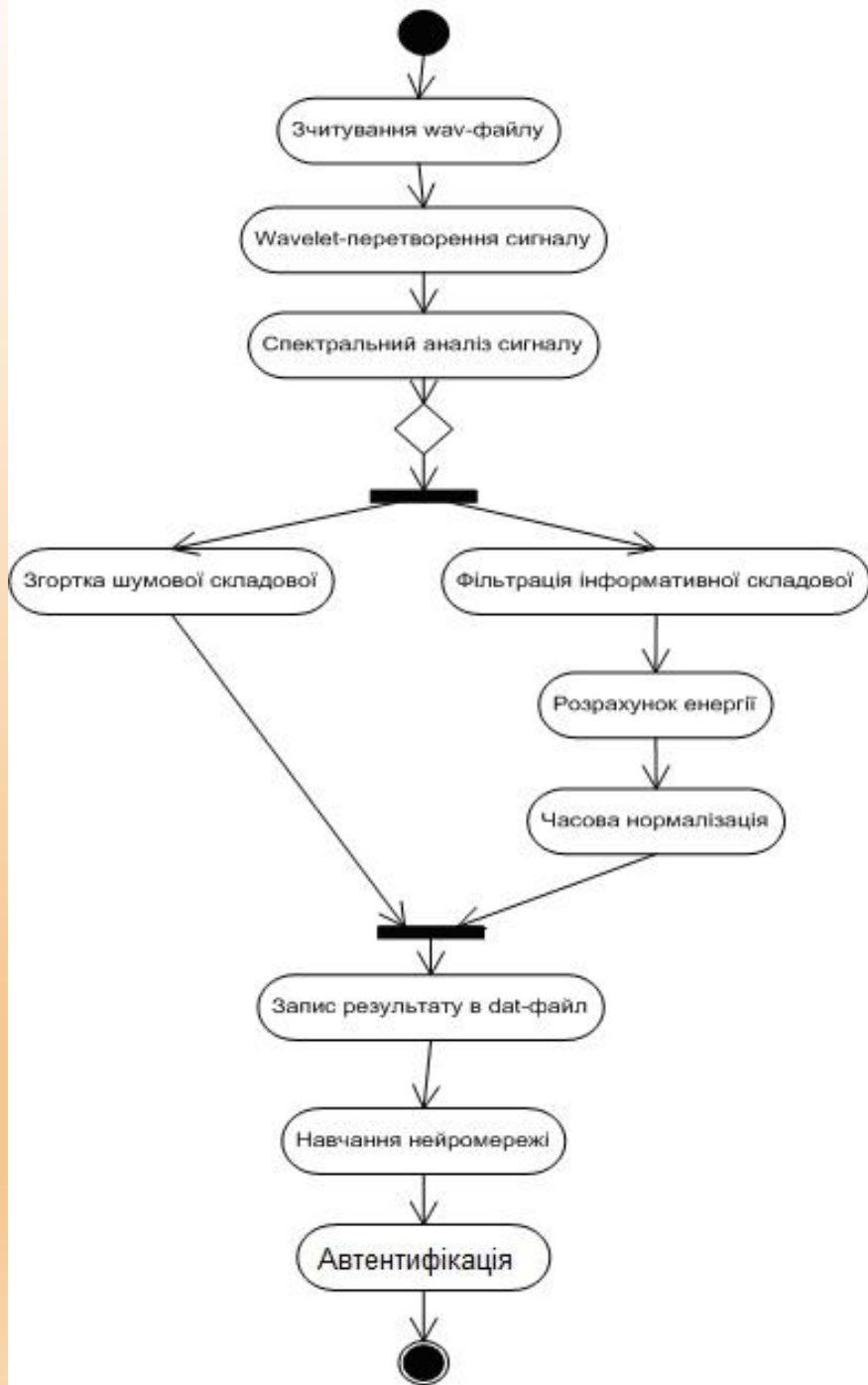
1. Система Matlab спеціально створена для проведення саме інженерних обчислювань: математичний апарат, що використовується в ній, гранично наближений до сучасного математичного апарата інженера, ученого й спирається на роботу з матрицями, векторами й комплексними числами; графічне подання функціональних залежностей тут організовано у формі як вимагає саме інженерна документація;
2. Мова програмування системи Matlab близька до актуальних мов програмування, але має свій синтаксис та особливості;
3. На відміну від більшості систем комп'ютерної математики, Matlab є відкритою системою; це означає, що практично всі процедури й функції Matlab доступні не тільки для використання, але й для модифікації; користувач може розширювати базу процедур та функцій створюючи власні підпрограми у тому числі за допомогою вбудованого графічного редактора користувача GUIDE; це робить Matlab незамінним засобом наукових обчислювальних робіт та імітаційного моделювання фізичних процесів;
- 4) останні версії Matlab дозволяють інтегрувати її з текстовим редактором Word, що робить зручним оформлення науково-технічних звітів, статей, дипломних та курсових проектів із включенням у них складних обчислень та виведенням графіків у текст.

# UML-діаграма варіантів використання



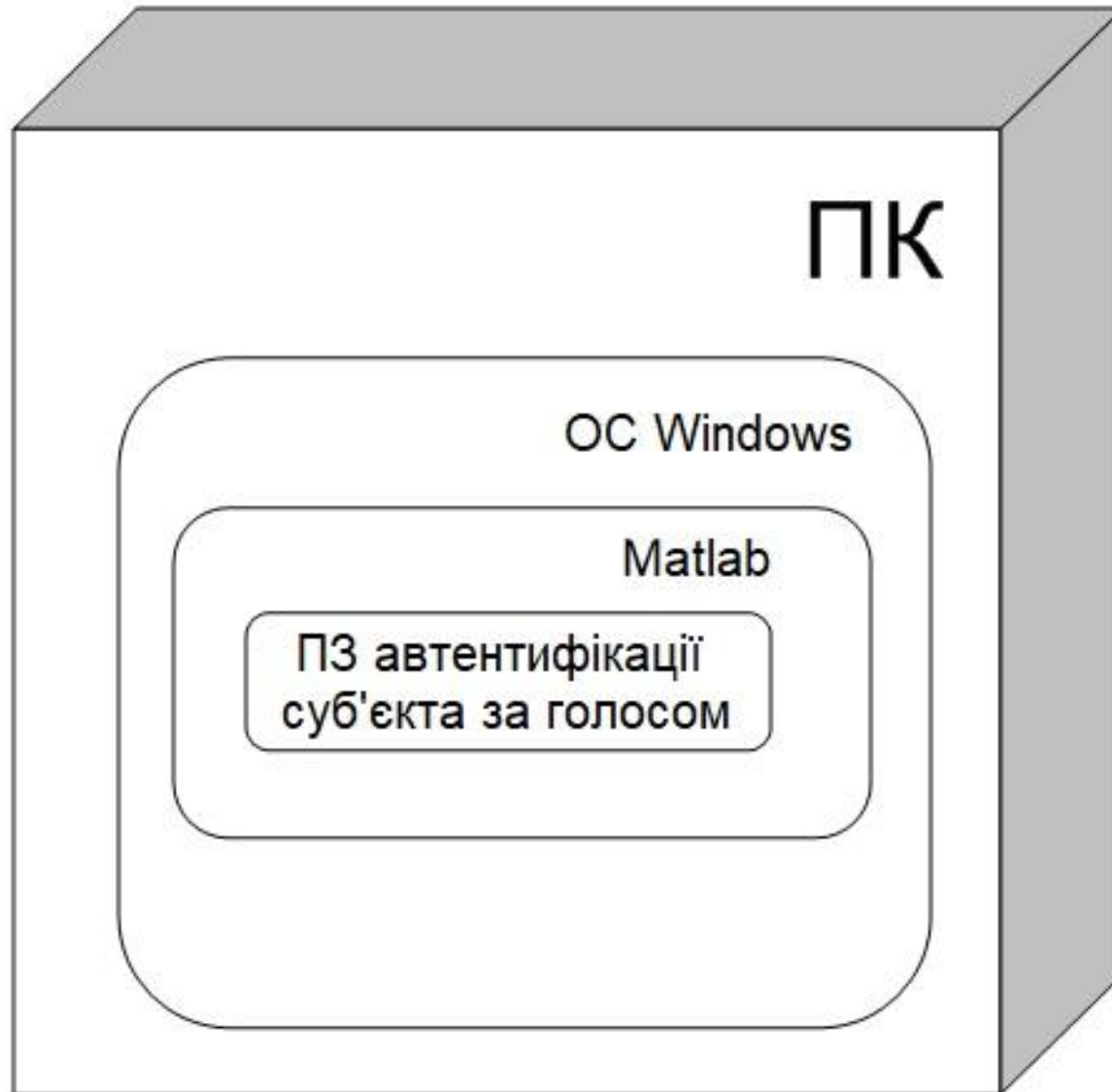
# UML-діаграма послідовності



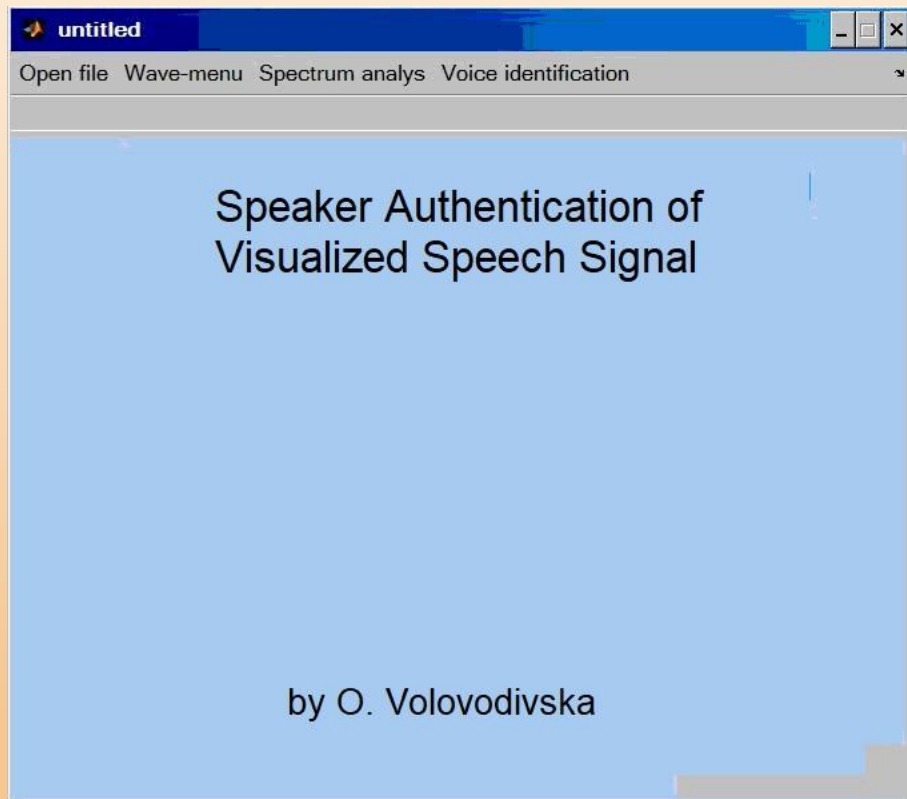


UML-діаграма  
активності

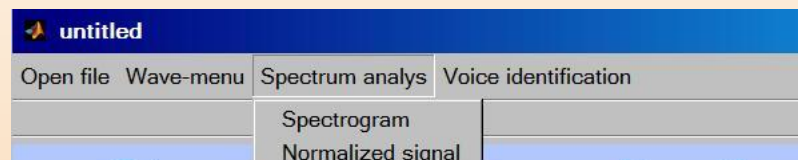
# UML-діаграма розгортання



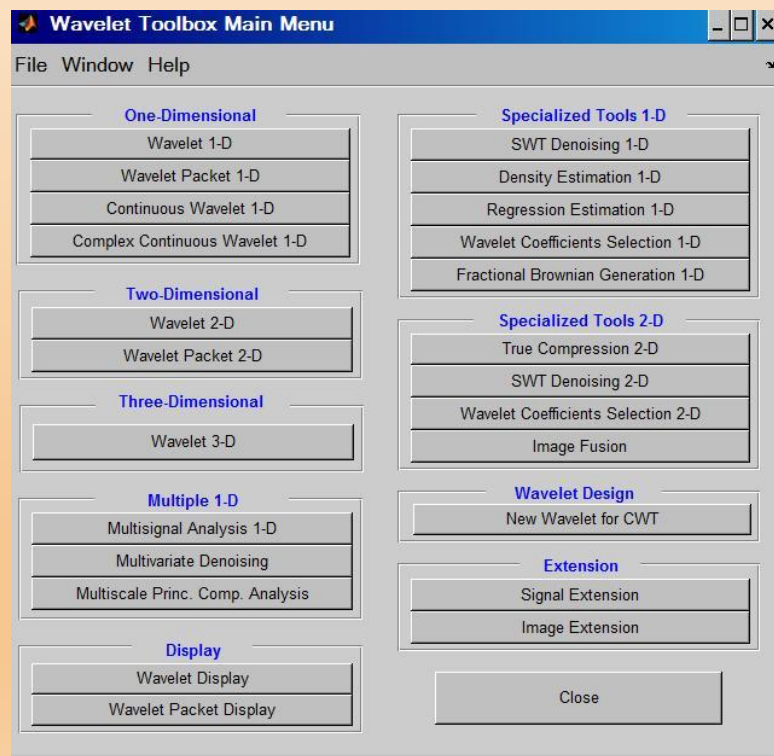
# Екранні форми та результати тестування створеного ПЗ



Стартове вікно ПЗ



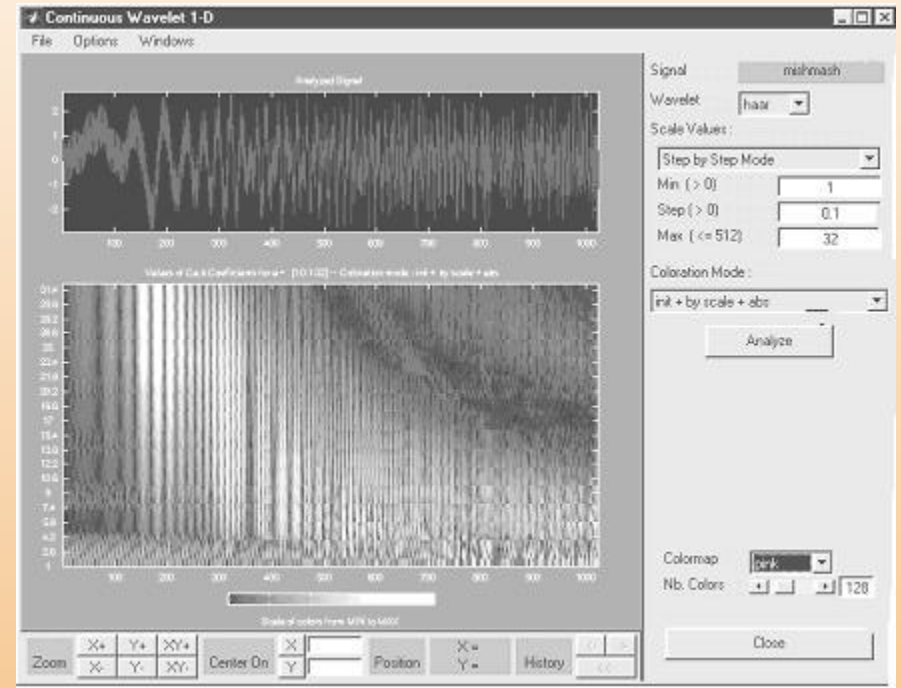
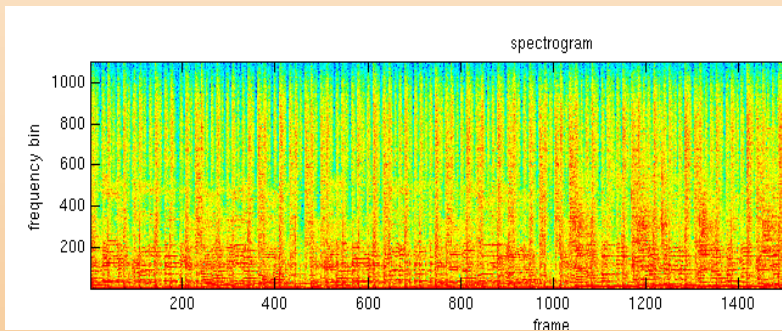
Отримання спектрограми



Отримання вейвлет-перетворення



# Екранні форми та результати тестування створеного ПЗ



Приклад отриманих за допомогою створеного ПЗ:  
-спектрограми (зліва),  
-вейвлет-діаграми Хаара (справа),  
для мовного сигналу, візуалізованого справа вверху.

Належність мовного сигналу	Результат автентифікації	
	Спектрограма	Вейвлет-грама
speaker1	Вірно	Вірно
speaker2	Вірно	Вірно
speaker3	Не вірно	Не вірно
speaker4	Вірно	Вірно
speaker5	Вірно	Вірно
speaker6	Вірно	Вірно
speaker7	Вірно	Вірно
speaker8	Не вірно	Вірно
speaker9	Вірно	Вірно
speaker10	Вірно	Вірно
speaker11	Вірно	Вірно
speaker12	Вірно	Вірно
speaker13	Вірно	Вірно
speaker14	Не вірно	Вірно
speaker15	Вірно	Вірно
speaker16	Вірно	Вірно
speaker17	Вірно	Вірно
speaker18	Вірно	Вірно
speaker19	Вірно	Вірно
speaker20	Вірно	Вірно

**Результати  
автентифікації  
20 мовців  
згортковою  
нейромережею  
на основі  
спектрограми або  
вейвлет-грами  
вхідного  
мовного сигналу.**

Для навчання  
використовувалися 1500  
зображень (30 с мовного  
матеріалу).  
Для тестування – 150  
Зображень (3 парольних  
фрази привалістю по 1 с)

# ВИСНОВКИ

1. Розглянуто і проаналізовано за обраними критеріями основні методи аналізу та оброблення мовних сигналів
2. Здійснено порівняння і аналіз актуальних існуючих систем автентифікації суб'єктів за голосом.
3. Викладена суть проблеми автентифікації суб'єктів за голосом та отримано узагальнену постановку задачі синтезу нейромережевої системи автентифікації суб'єкта за візуальним представленням його голосу.
4. Розроблено структурну схему та UML-діаграми нейромережевої системи автентифікації суб'єкта за візуальним представленням його голосу.
5. Викладено математику процедур вейвлет-перетворення мовних сигналів, їх фільтрації та часової нормалізації тривалості їх звування, в т.ч. алгоритму нелінійної часової нормалізації.
6. У програмному середовищі MATLAB розроблено програмне забезпечення для автентифікації суб'єкта за візуальним представленням його голосу.
7. Проведено дослідження доцільності використання часової нормалізації мовних сигналів у задачі автентифікації суб'єктів за голосом та встановлено, що використання процедури нормалізації є доцільним.

Дякую за увагу