

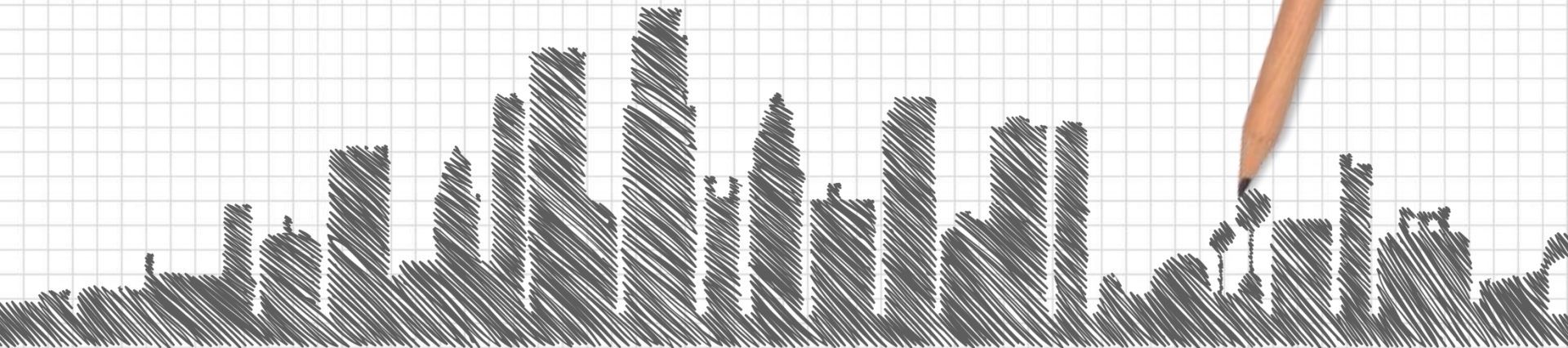


Вибір оптимального метода та оптимальних параметрів голосової біометрії. Розробка сервісу голосової біометрії

ВИКОНАВ: **ГРУШКО ЯРОСЛАВ**

КА-45, ІПСА

НАУК. КЕРІВНИК: Д.Т.Н., ПРОФЕСОР **ДАНИЛОВ ВАЛЕРІЙ ЯКОВИЧ**





МЕТА РОБОТИ

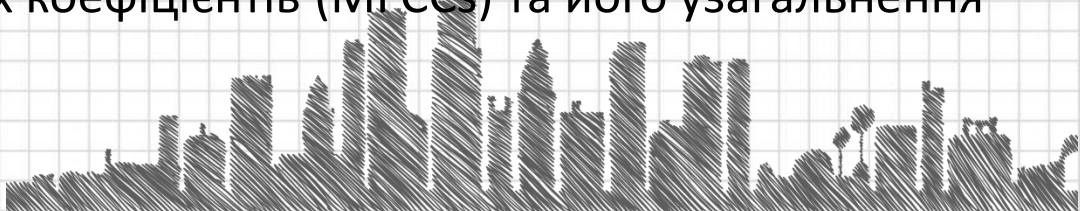
Розробка методики оцінки ефективності систем голосової біометрії.
Вибір оптимального метода, встановлення оптимальних параметрів та
розробка веб-сервісу голосової біометрії

ОБ'ЄКТ ДОСЛІДЖЕННЯ

Методи дикторозалежного та дикторонезалежного голосового
розпізнавання

ПРЕДМЕТ ДОСЛІДЖЕННЯ

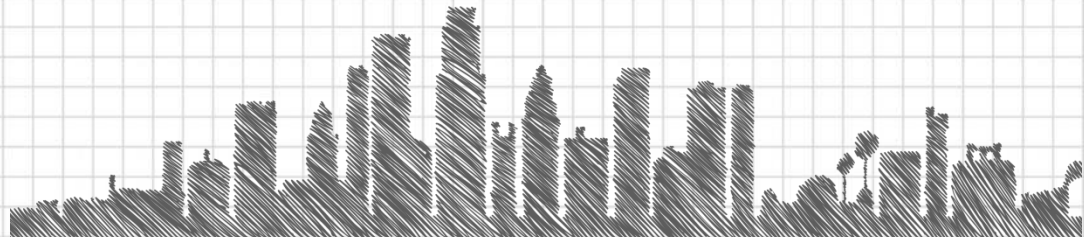
Метод мел-кепстральних коефіцієнтів (MFCCs) та його узагальнення





АКТУАЛЬНІСТЬ РОБОТИ

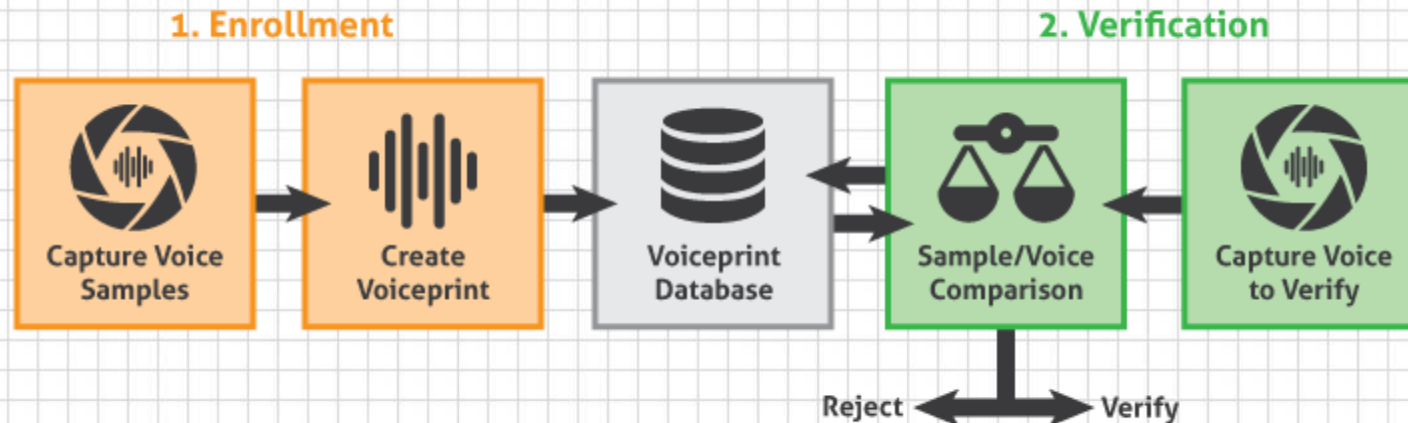
Останнім часом поняття голосової біометрії отримало неабияку популярність. Це пов'язано з тим, що даний підхід за останні декілька років перейшов на інший якісний рівень і став конкурентно спроможним до інших підходів, наприклад, таких, як паролі, відбитки пальців, ідентифікація по сітківці.





ГОЛОСОВА БІОМЕТРІЯ

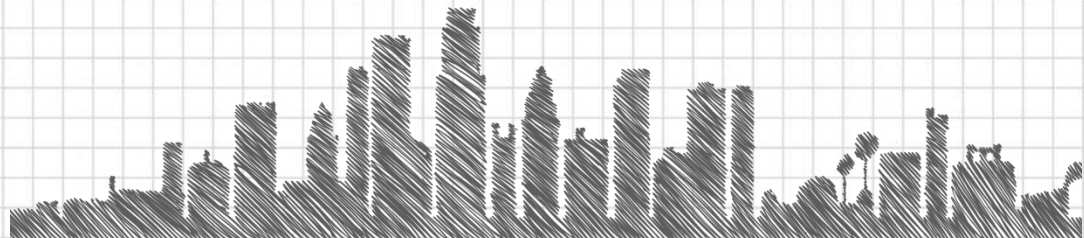
Процес отримання параметрів (відбитку) зразку голосу задля подальшої аутентифікації особи називається голосовою біометрією або голосовою ідентифікацією





ПОСТАНОВКА ЗАДАЧІ

- Вибрати оптимальний метод розпізнавання людини по голосу (голосова біометрія)
- Вибрати параметри ідентифікації
- Розробити методику оцінки ефективності систем голосової біометрії, яка буде в себе включати розроблені таблиці “значимості” ідентифікації (похибка першого роду)
- Знайти оптимальні значення параметрів системи
- Розробити десктопний додаток та веб-сервіс голосової біометрії

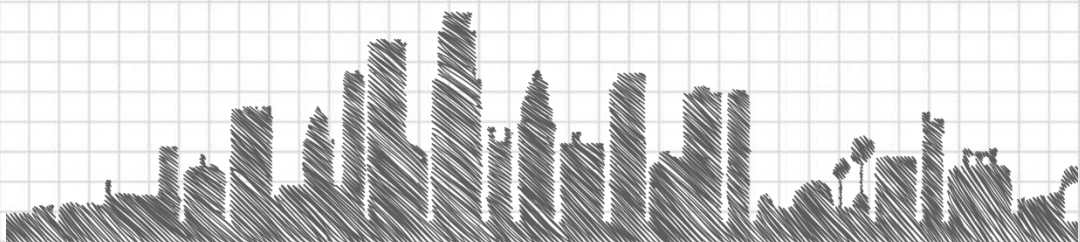




ПОПУЛЯРНІ МАТЕМАТИЧНІ МОДЕЛІ ГОЛОСОВОГО РОЗПІЗНАВАННЯ

Example system	Number of Speakers	Number of Languages	Size of Vocabulary	Hardware Resources	Development Time / Money	Recommended Setup(s)
1	Single	Single	< 50 words	Limited	Limited	LPC or LPCC with an HMM
2	Single	Single	≈ 200 - 500 Words	Limited	Limited	PLP or MFCC with an HMM
3	Single	Single	Full Language	Moderate	Limited	MFCC with an HMM
4	Multiple	Single	Full Language	Moderate	Moderate	MFCC with an HMM and Training Algorithm
5	Multiple	Multiple	Full Languages	Moderate	Abundant	RASTA-PLP with an HMM and Training Algorithm
6	Multiple	Multiple	Full Language	Abundant	Abundant	DNN (Most Accurate)

- Dynamic Time Warping
- Hidden Markov Model
- Vector Quantization
- Support Vector Machine
- Gaussian Mixture Model
- Mel-frequency cepstrum coefficients (MFCCs)





Існуючі системи голосового розпізнавання

- MARF
- VoiceVault
- Agnitio
- Nuance
- Voice Security Systems

The screenshot displays a forensic software interface with a table of assessments and a voice recognition application window.

Timestamp / Name	Suspect	Ref populations	Background Model	Reports
17/03/2014 19:07 Investigation 450	[Image]	English_America...	TI_IVEC_4	[Icon]
17/03/2014 19:05 Investigation 073	[Image]	English_America...	TI_IVEC_4	[Icon]
17/03/2014 19:03 Female Suspect As...	[Image]	English_America...	TI_IVEC_4	[Icon]
13/03/2014 14:37 Test Assessment ...				
11/03/2014 08:33 Brett Test 02				
07/03/2014 19:07 Brett Test				
07/03/2014 18:42 Assessment05				
07/03/2014 18:41 Assessment03				
07/03/2014 18:40 Assessment02				

The voice recognition application window shows a welcome message: "Welcome nik2!". It includes a microphone icon and a text prompt: "Press the button and say the phrase: '音乐家们经常成群结队地演奏乐器。'". The interface also displays a progress bar and a "VoiceVault Data Collection" header.

Additional windows visible include "Choose a Voice" (with "Female Voice (Default)" selected), "Copyrights & License" (with a license agreement text), and "Agnitio - Beta 1.01" (showing a list of commands).

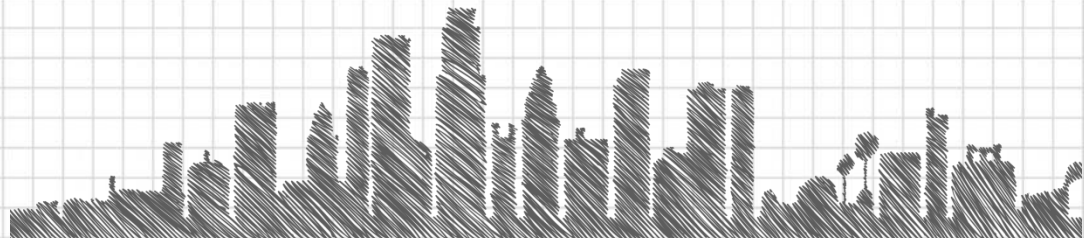


НЕДОЛІКИ ІСНУЮЧИХ СИСТЕМ ГОЛОСОВОГО РОЗПІЗНАВАННЯ

- Вони не безкоштовні

	Free Trial	Developer License	Production License
Cost	Free	\$750	Per User

- Метод ідентифікації приховується
- Українських аналогів небагато





MFCCs

- Метод голосової біометрії мел кепстральних коефіцієнтів
- Заснований на базі аналізу спектра спектру (кепстра)

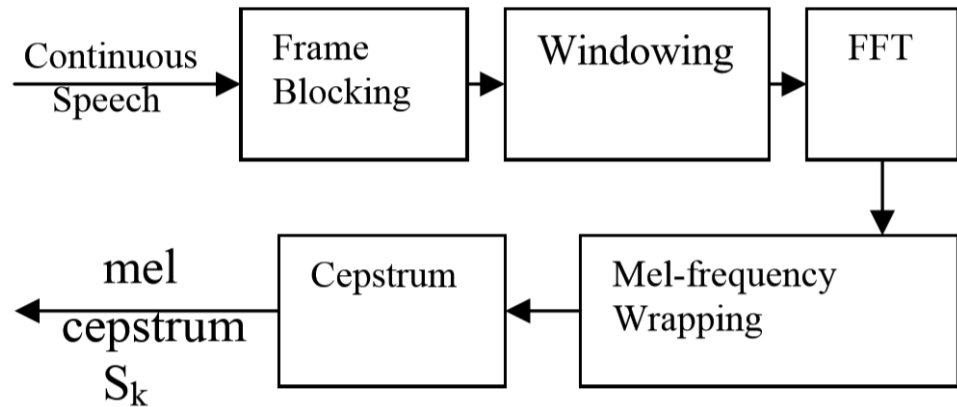
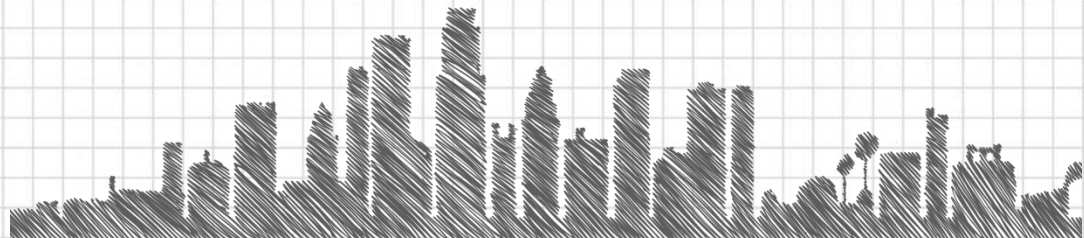


Рис 2.4 Блок-схема процесора MFCC



МЕТОДИКА ОЦІНКИ ЕФЕКТИВНОСТІ СИСТЕМ ГОЛОСОВОЇ БІОМЕТРІЇ:

- Запис по два зразки (одного і того ж слова) голосів кожним з досліджуваних осіб на носій інформації
- Отримання мел-кепстральних коефіц. 1-х зразків
- Отримання мел-кепстральних коефіц. 2-х зразків
- Попарне порівняння всіх зразків методом міських кварталів (манхеттенська відстань)
- Створення таблиці відстаней між зразками
- На діагоналі таблиці повинні бути мінімуми (відстані між зразками одних і тих же осіб), які залежать від к-ті мел-кепстр. коеф., діапазону частот (розмір вікна), метод обробки і т.п.
- Мінімізація відстаней визначається оптимальним набором параметрів





Метод різницевиx матриць

City Block Distance

Effectiveness = 100,00 %

Evaluation Error = 8,33 %

Matrix D:

	Алена1	Віталік1	Владимир1	Евгеній1	Ігорь1	Ната1	Наталья1	Нина1	Олег1	Саша1	Сереза1	Ярослав1
<u>True</u> Алена2	[19,08]	34,89	40,62	28,67	30,81	32,90	39,48	41,42	26,86	35,53	35,06	29,60
<u>True</u> Віталік2	30,38	[9,77]	19,75	21,30	22,75	22,07	42,69	31,47	26,89	18,38	46,27	15,98
<u>True</u> Владимир2	40,24	21,29	[8,00]	35,19	31,21	31,69	41,59	31,99	32,31	19,68	50,08	24,00
<u>True</u> Евгений2	27,91	22,21	27,89	[8,25]	28,29	27,97	34,48	26,99	29,73	24,90	45,10	25,41
<u>True</u> Игорь2	31,28	23,61	23,51	27,94	[16,87]	25,03	42,39	28,97	21,53	28,61	42,77	20,33
<u>True</u> Ната2	29,83	29,75	28,03	29,42	28,31	[11,12]	29,63	24,54	31,03	27,31	44,22	23,60
<u>True</u> Наталья2	32,24	41,06	38,57	33,65	37,21	31,22	[9,49]	28,81	38,25	33,45	45,60	33,08
<u>True</u> Нина2	39,13	31,59	33,06	26,40	28,40	23,11	27,44	[9,44]	33,66	28,80	42,12	24,33
<u>True</u> Олег2	25,62	31,14	28,95	28,54	27,92	36,06	39,09	34,38	[16,50]	28,58	42,07	26,70
<u>True</u> Саша2	32,77	27,43	24,18	31,09	34,62	29,02	35,64	30,60	31,55	[6,69]	51,72	19,57
<u>True</u> Сереза2	29,85	44,23	47,07	45,37	36,39	44,19	38,03	42,45	40,04	49,36	[7,61]	40,17
<u>True</u> Ярослав2	33,07	20,11	21,04	28,30	23,49	23,60	39,08	23,28	27,72	22,57	39,93	[18,88]



Різницеві матриці D





ПРОГРАМНИЙ ПРОДУКТ "Voice Recognition Laboratory"

Voice Recognition Laboratory

For identification please say:
"Sound,...sound,...sound,...sound!"

Voice Identification

Adding a new Client

ANALYSIS

Name

Effectiveness of recognition methods

MFCCs: <i>Name</i>	difference: %	nearest: %
MFSCs: <i>Name</i>	difference: %	nearest: %

Voice Recognition Laboratory

For identification please say:
"Sound,...sound,...sound,...sound!"

Voice Identification

Adding a new Client

ANALYSIS

Cawa

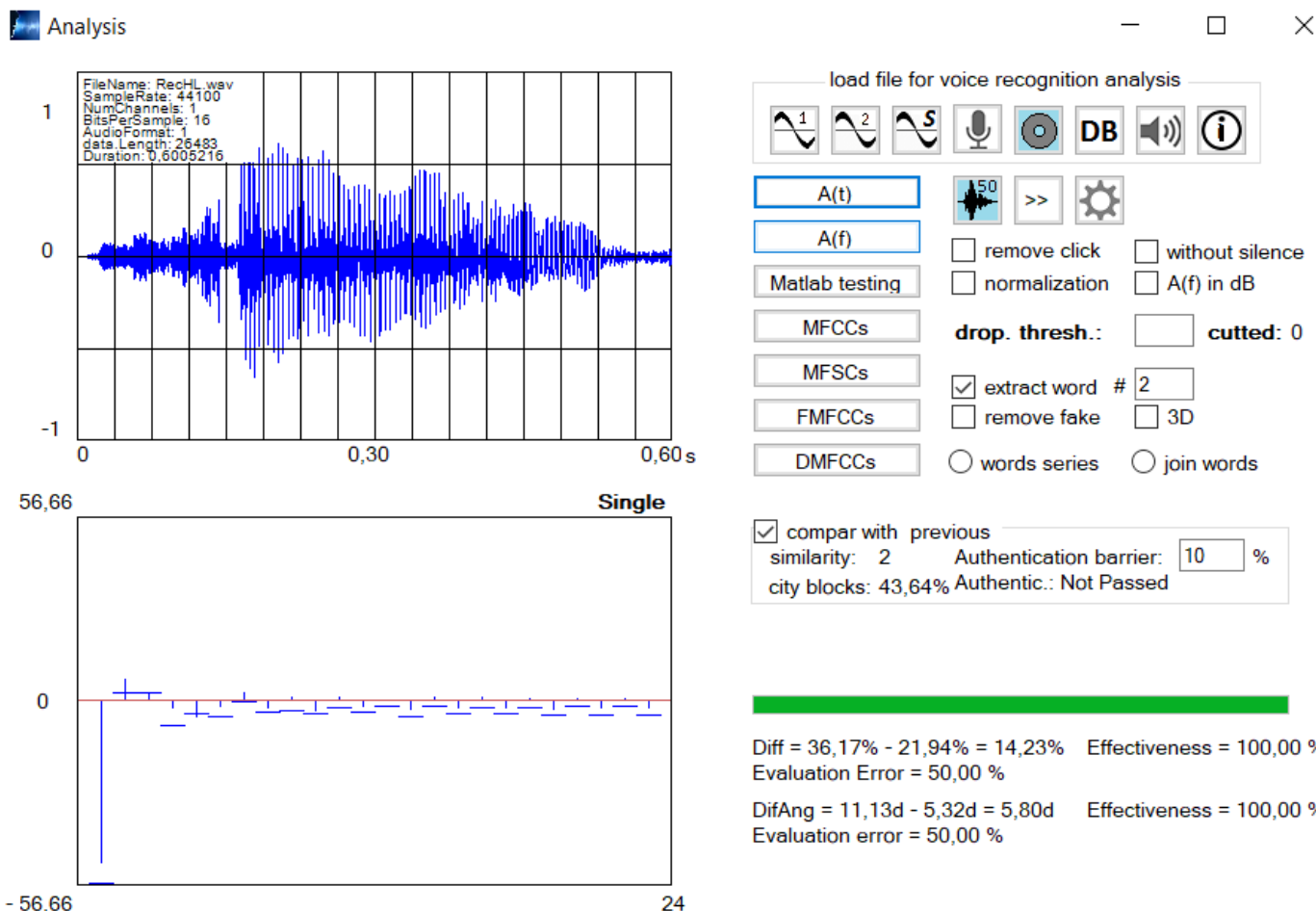
Effectiveness of recognition methods

MFCCs: <i>Cawa</i>	difference: 7,39%	nearest: 13,93%
MFSCs: <i>Cawa</i>	difference: 8,64%	nearest: 35,49%

- Форма для тестування найкращих методів ідентифікації



ПРОГРАМНИЙ ПРОДУКТ (Voice Recognition Laboratory)



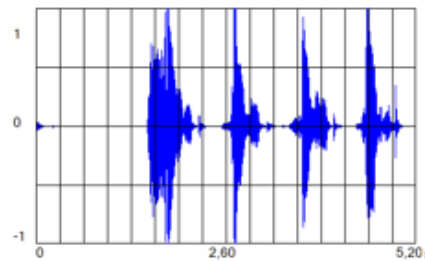
Форма Analysis для обробки, фільтрації, аналізу звукових сигналів і визначення оптимального методу ідентифікації



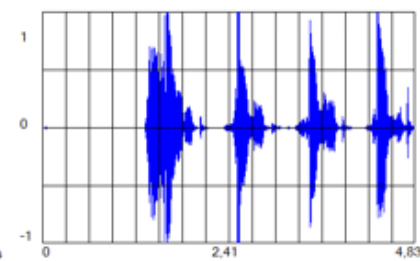
ПРОГРАМНИЙ ПРОДУКТ (Voice Recognition Laboratory)

Обробка звуку на формі "Analysis":

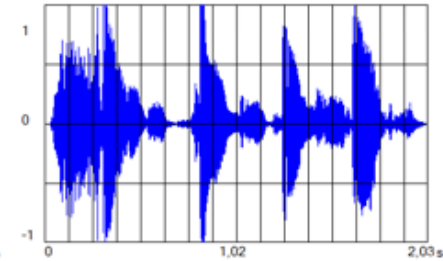
- (a) нормалізація
- (b) очистка від кліків
- (b) видалення непотрібної тиші в сигналі
- (e) розбивка фрази на слова;
- (g) видалення сильно зашумлених слів і з'єднання тих, що залишились в одну фразу



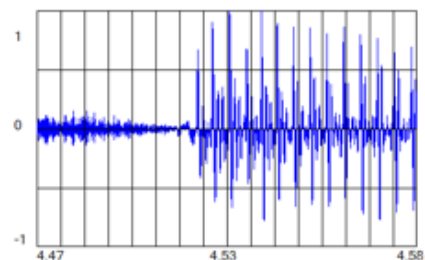
(a)



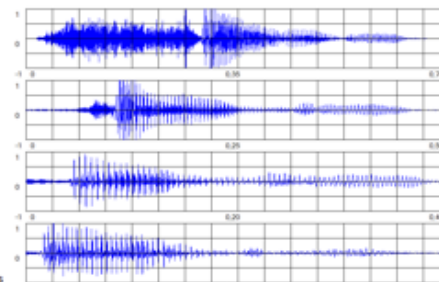
(b)



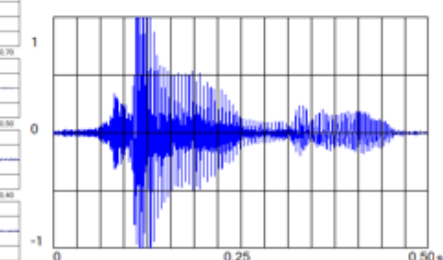
(c)



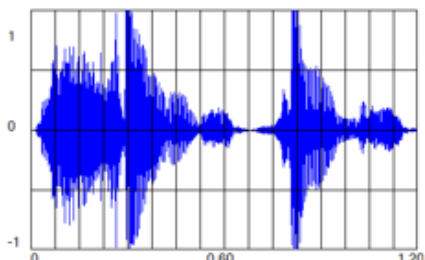
(d)



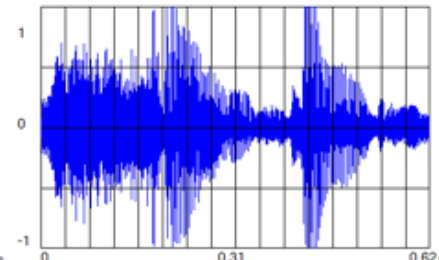
(e)



(f)



(g)



(h)

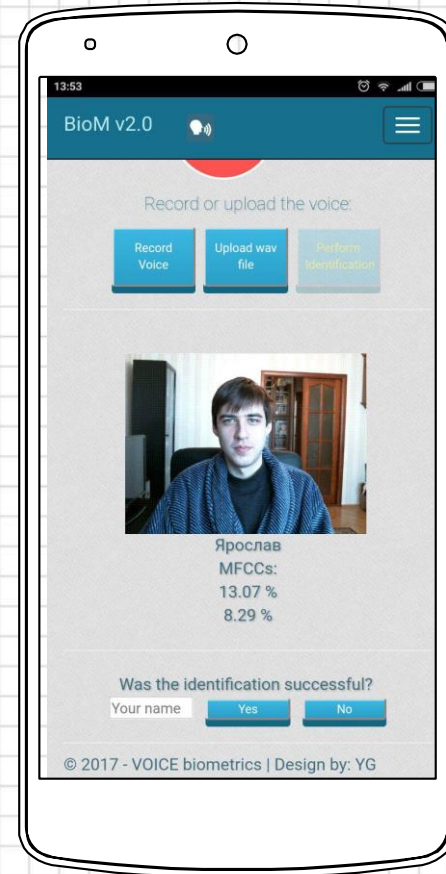
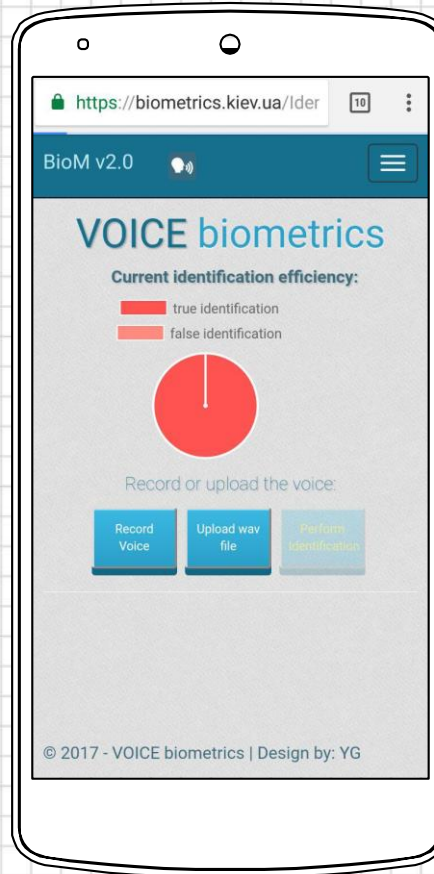


ПРОГРАМНИЙ ПРОДУКТ (Voice Biometrics)

- Веб-сайт
(<https://biometrics.kiev.ua>)
- 100% ідентифікація на 12-ти голосових зразках
- На базі новітньої технології Asp.net MVC (Microsoft)
- Має базу даних MySQL
- Сайт адаптований під мобільні

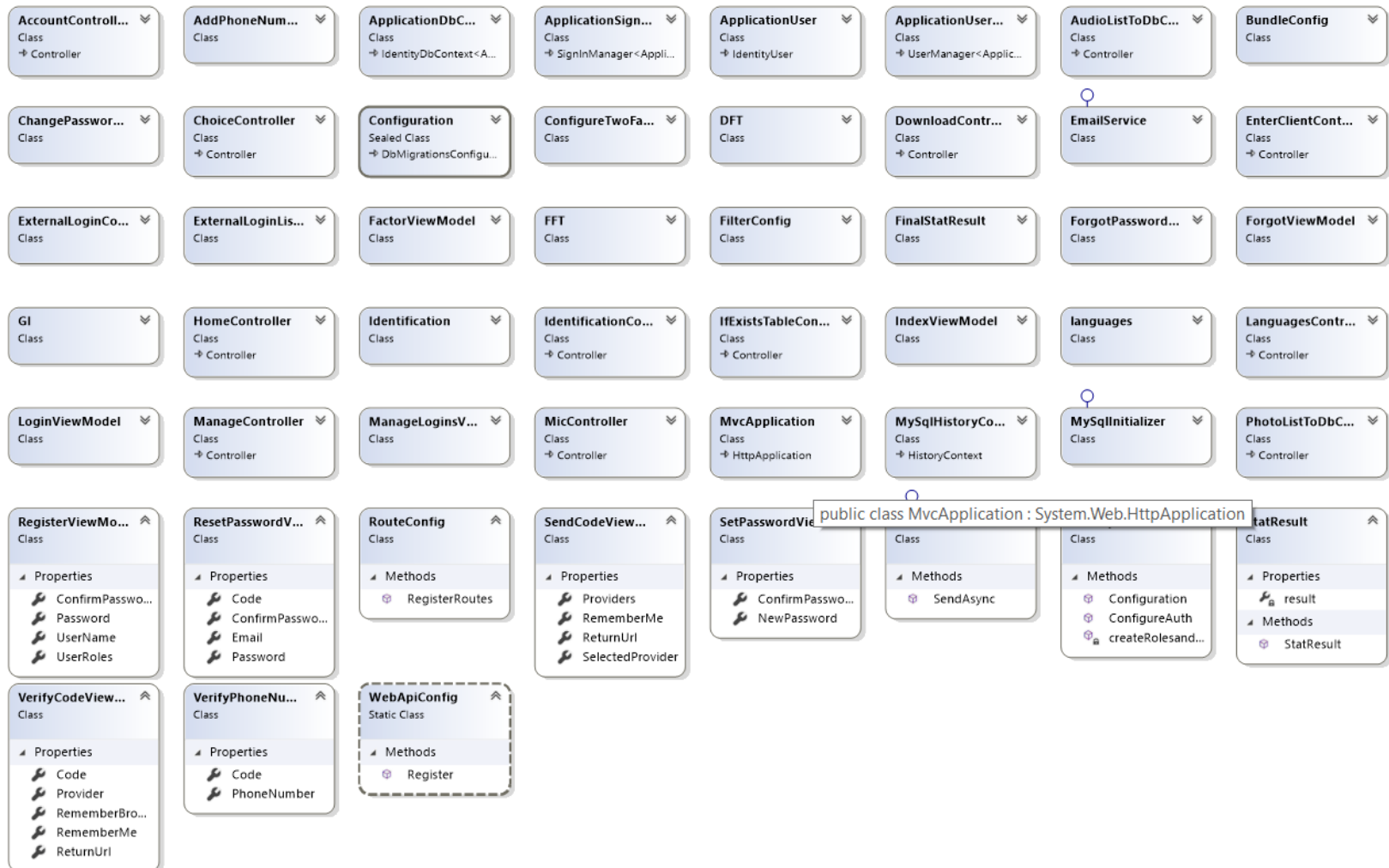
Екрани

- Амплітудно-часова візуалізація голосу
- Створений на базі десктоп додатку "Voice Recognition Laboratory"





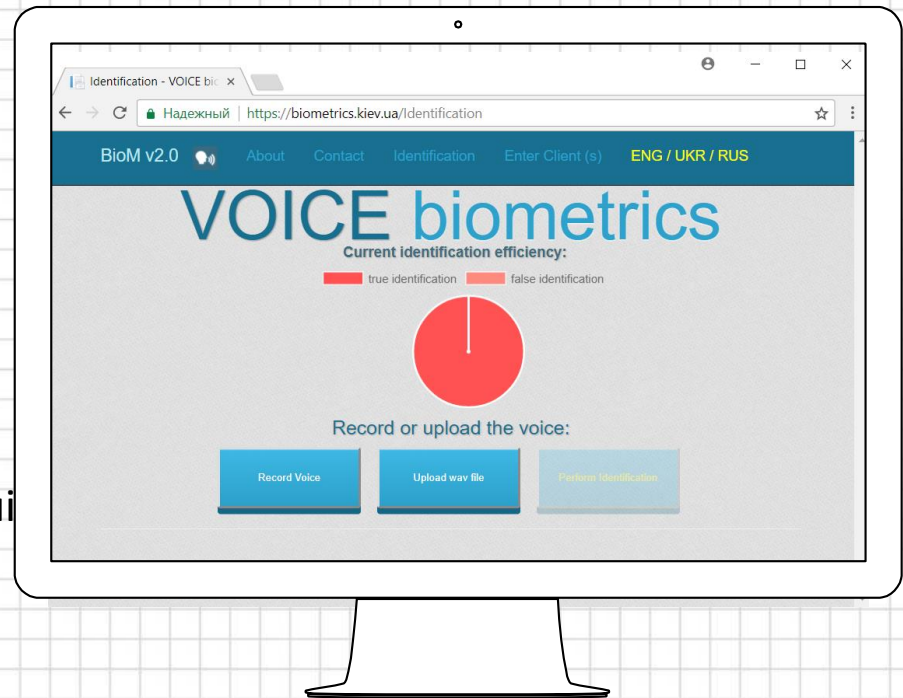
Діаграма класів програми "Voice Biometrics"





В ХОДІ РОЗРОБКИ ВЕБ САЙТУ ВИНИКЛИ НАСТУПНІ СКЛАДНОЩІ:

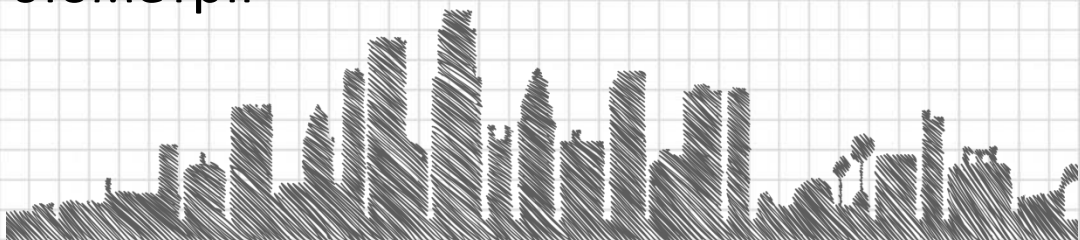
- Залежність ефективності розпізнавання особи від звукозаписуючого приладу, якості запису, наявності шумів і т.п.
- Ефективність розпізнавання диктора залежить від тестової фрази
- Застосування криптографічних протоколів SSL, задля надійної передачі звукових даних голосу
- Орендування бистродійного веб хостингу





ВИСНОВКИ

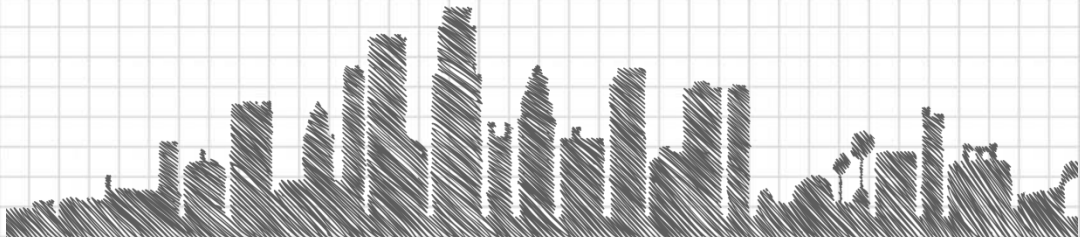
- За проведеним аналізом, було зроблено висновок про те, що метод мел-кепстральних коефіцієнтів (MFCCs) є одним з найбільш ефективних (ймовірність розпізнавання може бути краща за 99%)
- Розроблено власний метод оцінки ефективності систем голосової біометрії (метод різницевих матриць), запропоновано декілька альтернативних методів ідентифікації (MFSCs, FMFCs)
- Розроблено лабораторію та веб-сервіс для аналізу та використання голосової біометрії





ПОДАЛЬШЕ ПОКРАЩЕННЯ

- Досягти розпізнавання краще 90% для вибірки із 100 та більше осіб
- Застосувати комбінацію найбільш популярних методів ідентифікації задля покращення ефективності розпізнавання
- Надати JavaScript вбудовану службу перевірки автентичності для веб-сайтів





МОЇ ПУБЛІКАЦІЇ

- Спектральний і кепстральний аналіз звука для ідентифікації голосу
- Ефективний статистичний алгоритм знаходження ключових слів у текстових файлах

УДК 004.56.5

СПЕКТРАЛЬНЫЙ И КЕПСТРАЛЬНЫЙ АНАЛИЗ ЗВУКА ДЛЯ ИДЕНТИФИКАЦИИ ГОЛОСА

Я. В. Грушко^{1,а}, В. Я. Данилов^{1,б}

¹ Учебно-научный комплекс "ИИПА" НТУУ "КПИ", кафедра ММСА

Аннотация

В статье представлены алгоритмы спектрального и кепстрального анализа для идентификации мужского и женского голосов. Создан прототип системы распознавания, основанной на использовании мел-кепстрального анализа и энергетического распределения спектра голосового сигнала. Приведены вероятности дикторозависимого распознавания с использованием обоих подходов.

Ключевые слова: идентификация голоса, спектральный анализ, кепстральный анализ, мел-кепстральные коэффициенты, энергетическое распределение.

Введение

Голос — такая же неотъемлемая черта каждого человека, как и его лицо или отпечатки пальцев. Широкое распространение средств связи (стационарные и мобильные телефонные сети, интернет-телефония и т.д.) открывают большие возможности для применения данного идентификатора; кроме того, распознавание по голосу весьма удобно для пользователей и требует от них минимум усилий.

Технологии и средства идентификации по голосу применяются в ряде областей, непосредственно связанных с обработкой обращений пользователей по телефону (колл-центры и т.п.), что позволяет ускорить обслуживание абонентов и разгрузить операторов. В более значимых проектах (особенно связан-

методов распознавания, который учитывает голосовые признаки говорящего. Для анализа голосового сигнала в дикторозависимых методах обычно применяют кепстральный анализ [2, 3] который, по-сути, представляет собой анализ спектра анализируемого сигнала, называемый «кепстром» [4].

Популярным справочным руководством по идентификации человеческого голоса является работа И. Клементя, С. Мохова, Д. Николакопулуса, С. Синклара и др. «Modular Audio Recognition Framework v.0.3.0.6 (0.3.0 final) and its Applications» [5], где изложены основные подходы записи и анализа звуковых данных. В НТУУ «КПИ» проблемой анализа и параметризации речевых сигналов занимались Данилов В. Я. и Побережний Г. А. [6, 7, 8], которым разработаны

Матеріали міжнародної наукової конференції
Материалы международной научной конференции
Conference proceedings

ІНТЕЛЕКТУАЛЬНІ СИСТЕМИ ПРИЙНЯТТЯ РІШЕНЬ ТА
ПРОБЛЕМИ ОБЧИСЛЮВАЛЬНОГО ІНТЕЛЕКТУ

ИНТЕЛЛЕКТУАЛЬНЫЕ СИСТЕМЫ ПРИНЯТИЯ РЕШЕНИЙ И
ПРОБЛЕМЫ ВЫЧИСЛИТЕЛЬНОГО ИНТЕЛЛЕКТА

INTELLECTUAL SYSTEMS FOR DECISION MAKING AND
PROBLEMS OF COMPUTATIONAL INTELLIGENCE



● Дякую за увагу! ●

