

НАЦІОНАЛЬНИЙ ТЕХНІЧНИЙ УНІВЕРСИТЕТ УКРАЇНИ
«КИЇВСЬКИЙ ПОЛІТЕХНІЧНИЙ ІНСТИТУТ»

Факультет прикладної математики

Кафедра прикладної математики

«На правах рукопису»
УДК 656.073

«До захисту допущено»

Завідувач кафедри

_____ О. Р. Чертов

« ____ » _____ 2015 р.

Магістерська дисертація
на здобуття ступеня магістра

зі спеціальності 8.04030101 «Прикладна математика»

на тему: Пошук шаблонів у цифровому сигналі за допомогою поліномів
Кунченка

Виконав: студент 2 курсу, групи КМ-31М

Погода Михайло Володимирович _____

Науковий керівник завідувач кафедри, д-р техн. наук, доцент _____
Чертов О. Р.

Консультант із
нормоконтролю старший викладач Мальчиков В. В. _____

Рецензент професор, д-р техн. наук, проф. Бідюк П. І. _____

Засвідчую, що у цій магістерській
дисертації немає запозичень з праць інших
авторів без відповідних посилань.

Студент _____

Київ — 2015 року

**Національний технічний університет України
«Київський політехнічний інститут»**

Факультет прикладної математики

Кафедра прикладної математики

Рівень вищої освіти – другий (магістерський)

Спеціальність 8.04030101 «Прикладна математика»

ЗАТВЕРДЖУЮ
Завідувач кафедри
_____ О. Р. Чертов
(підпис)

«___» _____ 2015 р.

**ЗАВДАННЯ
на магістерську дисертацію студенту
Погоді Михайлу Володимировичу**

1. Тема дисертації: «Пошук шаблонів у цифровому сигналі за допомогою поліномів Кунченка»,
науковий керівник дисертації Чертов Олег Романович, д-р. техн. наук,
затверджені наказом по університету від «20» березня 2015 року № 785-С.
2. Термін подання студентом дисертації: «18» червня 2015 р.
3. Об'єкт дослідження: методи обробки цифрових сигналів, методи пошуку шаблонів у цифрових сигналах, метод Кунченка, методи покращення пошуку шаблонів у цифрових сигналах, модель обробки цифрового аудіосигналу, що містить мовлення, програмні продукти для розробки програмного модуля.
4. Предмет дослідження: вдосконалена модель пошуку шаблонів у цифровому сигналі за допомогою поліномів Кунченка.
5. Перелік завдань, які потрібно розробити:
 - проаналізувати існуючі методи для пошуку шаблонів у цифрових сигналах і обґрунтувати вибір методу пошуку шаблонів на основі поліномів Кунченка.;
 - розробити алгоритм пошуку шаблонів на основі поліномів Кунченка;
 - проаналізувати шляхи прискорення роботи та покращення результатів роботи розробленого алгоритму;

- дослідити методології з обробки цифрових сигналів, зокрема аудіосигналів, що містять мовлення;
- розробити програмне забезпечення для пошуку шаблонів у цифровому сигналі;
- проаналізувати вплив зміни параметрів алгоритму на швидкість та ефективність пошуку шаблонів;
- провести статистичний експеримент для порівняння ефективності роботи розробленого програмного забезпечення з існуючими методами.

6. Орієнтовний перелік ілюстративного матеріалу:

- таблиця порівняння існуючих методів;
- блок-схема роботи алгоритму пошуку інтервалу з екстремумом;
- графіки з прикладами обчислення коротко-часової енергії;
- таблиця порівняльної характеристика математичних пакетів;
- блок-схема роботи алгоритму пошуку шаблону;
- таблиця результатів пошуку шаблону без модифікації;
- графіки сигналу з шаблоном без модифікації;
- графіки пошук шаблону без модифікацій в синтетичному сигналі;
- графіки застосування пірамідального пошуку;
- графічне зображення запису мовлення й шаблону для пошуку;
- результати пошуку шаблону використовуючи прямокутне вікно, вікно Геммінга, вікно Гана;
 - графічне зображення короткочасної енергії запису мовлення й шаблону для пошуку використовуючи прямокутне вікно та мінімальний крок;
 - результати пошуку шаблону в енергії, використовуючи прямокутні вікна, вікна Гана, вікна Геммінга;
 - графічне зображення короткочасної енергії запису мовлення й шаблону для пошуку використовуючи прямокутне вікно та крок середнього розміру;
 - графічне зображення короткочасної енергії мовлення й шаблону для пошуку використовуючи вікно Гана та крок мінімального розміру;
 - графічне зображення короткочасної енергії запису мовлення й шаблону для пошуку використовуючи вікно Геммінга та крок мінімального розміру.

7. Орієнтовний перелік публікацій:

- VII наукова конференція магістрів та аспірантів «Прикладна математика та комп'ютинг – ПМК'2015».

8. Дата видачі завдання «25» жовтня 2013 р.

Календарний план

№ з/п	Назва етапів виконання магістерської дисертації	Термін виконання етапів магістерської дисертації	Примітка
1	Вибір напрямку дослідження та узгодження тематики МД з керівником	15 вересня–30 жовтня 2013	
2	Грунтовне ознайомлення з предметною областю	30 жовтня 2013–15 лютого 2014	
3	Проведення порівняльного аналізу методів пошуку шаблонів	15 лютого–1 вересня 2014	
4	Підготовка матеріалів першого та другого розділів МД	1 вересня–1 жовтня 2014	
5	Проведення досліджень з обробки аудіосигналів, що містять мовлення	1 жовтня–1 листопада 2014	
6	Розробка математичної моделі пошуку шаблонів за допомогою поліномів Кунченка	1 листопада–1 грудня 2014	
7	Підготовка матеріалів третього розділу МД	20 січня–1 лютого 2015	
8	Розробка програмного забезпечення для пошуку шаблонів	1 лютого–1 березня 2015	
9	Переддипломна практика, робота над публікацією	1 березня–1 квітня 2015	
10	Тестування програмного забезпечення	1 квітня–30 квітня 2015	
11	Завершення роботи над основною частиною МД	5 травня–1 червня 2015	
12	Попередній захист МД	1 червня–15 червня 2015	

Студент

Науковий керівник дисертації

М. В. Погода

О. Р. Чертов

(підпис)

(підпис)

РЕФЕРАТ

Метою цієї магістерської дисертації є розробка та оптимізація програмного комплексу для пошуку шаблонів у цифровому сигналі. Комплекс, що розробляється, призначений для виділення ділянок цифрового сигналу, що містять деякий шаблон, можливо у модифікованому вигляді.

Об'єктом дослідження є процеси аналізу цифрових сигналів та методи їх обробки.

Предметом дослідження є вдосконалення пошуку шаблонів за допомогою методу поліномів Кунченка у цифрових сигналах, зокрема в аудіосигналах, що містять мовлення.

Розглянуто існуючі методи обробки цифрових сигналів, пошуку шаблонів у цифрових сигналах, запропонований новий метод і проаналізовані шляхи покращення його роботи.

Результатом роботи є програмний комплекс, який був протестований на низці тестових даних, зокрема на аудіосигналах, що містили записане мовлення.

Наукова новизна полягає в тому, що були запропоновані модифікації методу пошуку шаблонів на базі поліномів Кунченка (пірамідальний пошук, застосування віконної функції, відмінної від прямокутної), які дозволили суттєво збільшити його швидкодію та ефективність.

Основні положення роботи доповідалися на конференції «ПМК'2015»

Робота складається з вступу, 5 розділів та висновків і налічує 94 сторінок. Містить 28 ілюстративних матеріалів, 3 таблиці, 2 додатки та посиляється на 15 літературних джерел.

Ключові слова: пошук шаблонів, цифровий сигнал, обробка сигналів, розпізнавання, математичний метод, поліноми Кунченка.

ABSTRACT

The thesis is dedicated to researching, designing and developing program solution for template matching in digital signal. Program was designed to correctly match subintervals of signal to some template that may be altered in some way.

Digital signal processing and analysis are the main objects of research.

Research object: improvement of template matching in digital signals using Kunchenko polynomials, especially in audio signals that contain speech.

Also described existing methods for template matching in digital signals, new method was proposed. Ways for optimisation for new method were observed.

Software for pattern matching was developed and tested alongside with existing methods on speech signal.

Science innovation of this works consists of modifications (coarse search, application of windowing functions) for template matching method that is based on Kunchenko polynomials that have increased method's performance significantly.

The main results of this work were presented at conference «PMK'15».

The work consists of an introduction, 5 sections, includes conclusions and 94 pages. Contains 28 illustrative materials, 3 tables, 2 appendices and has 15 references.

Keywords: template matching, digital signal, signal processing, recognition, mathematical method, Kunchenko polynomials.

Зміст

Перелік умовних позначень, скорочень і термінів	9
Вступ	10
1 Постановка задачі	11
2 Огляд існуючих методів	12
2.1 Опис предметної області	12
2.1.1 Метод ковзаючого вікна	13
2.1.2 Віконні функції	14
2.2 Опис існуючих методів	16
2.2.1 Взаємкореляція	16
2.2.2 Нормалізована взаємкореляція	18
2.2.3 Сума квадратів відстаней	19
2.2.4 Нормалізована сума квадратів відстаней	20
2.3 Порівняльний аналіз методів	20
2.4 Висновки	21
3 Метод Кунченка	22
3.1 Лінійний простір Кунченка	22
3.2 Застосування розкладення для пошуку шаблону в сигналі	24
3.3 Методи покращення пошуку шаблонів	25
3.3.1 Пірамідальний пошук	26
3.3.2 Метод прискореного ковзання	27
3.4 Пошук шаблонів в мовленнєвих сигналах	29
3.5 Висновки	32
4 Проектування програмних засобів	33
4.1 Вибір програмного середовища	33
4.2 Проектування архітектури програмної реалізації	34
4.3 Алгоритм роботи програми	36
4.4 Висновки	37
5 Математичне моделювання та статистичний експеримент	38
5.1 Пошук у синтетичному сигналі	38
5.2 Пошук шаблону у мовленнєвому сигналі	43
5.2.1 Пошук шаблону в аудіосигналі	44

	8
5.2.2 Пошук шаблону в коротко-часовій енергії	49
5.3 Висновки	66
Висновки	67
Перелік посилань	69
Додаток А Лістинг реалізації алгоритму пошуку шаблонів	71
Додаток Б Ілюстративні матеріали	82



ПЕРЕЛІК УМОВНИХ ПОЗНАЧЕНЬ, СКОРОЧЕНЬ І ТЕРМІНІВ

CC — (з англ. Cross-correlation) взаємо-кореляція.

SSD — (з англ. Sum of Squared Differencies) сума квадратів відстаней.

NCC — (з англ. Normalised Cross-correlation) нормалізована взаємо-кореляція.

NSSD — (з англ. Normalised Sum of Squared Differencies) нормалізована сума квадратів відстаней.

Корелянта — скалярний добуток у функціональному просторі.

ПЗ — програмне забезпечення.



ВСТУП

Інтелектуальна обробка інформації вже дуже давно зарекомендувала себе в сучасному світі. Людство генерує незчисленні обсяги інформації щосекунди. Багато згенерованої інформації (поки що) не зберігається, а з того, що зберігається, багато ніяк не аналізується.

Але з поширенням інформаційних технологій та з розповсюдженням пристроїв, що можуть ефективно обробляти великі обсяги інформації, знаходиться все більше сфер, де можливо отримати користь з раніш не оброблюваного інформаційного потоку.

Наприклад, аналізуючи зміни погодних умов упродовж дня, метеорологи мають змогу прогнозувати погоду наперед; аналізуючи різні показники життєдіяльності людини, лікарі знаходять прояви хвороб ще до перших її симптомів.

Задачі пошуку окремих наперед визначених властивостей чи образів у деякому масиві даних набули великого поширення в галузях, пов'язаних з обробкою цифрової інформації.

Співставлення з еталоном (template matching) широко застосовується в аналізі як одновимірних, так і двовимірних сигналів [1].

Зазвичай, постановка такої задачі включає в себе дані для аналізу, деякий шаблон для пошуку та, можливо, область, на якій потрібно здійснювати такий пошук. При цьому потрібно як правильно обрати представлення вхідних даних, так і визначити критерії, за якими відбуватиметься зіставлення шаблону з сигналом [2].

Існує розмаїття підходів та метрик, які можна використовувати для вирішення цієї задачі.

У даній роботі для цього пропонується використовувати поліноми Кунченка.

1 ПОСТАНОВКА ЗАДАЧІ

В магістерській дисертації ставиться задача пошуку шаблонів у цифрових сигналах за допомогою поліномів Кунченка. Дана задача включає наступні задачі:

1. Проаналізувати існуючі методи для пошуку шаблонів у цифрових сигналах і обґрунтувати вибір методу пошуку шаблонів на основі поліномів Кунченка.
2. Розробити алгоритм пошуку шаблонів на основі поліномів Кунченка.
3. Проаналізувати шляхи прискорення роботи та покращення результатів роботи розробленого алгоритму.
4. Дослідити методи з обробки цифрових сигналів, зокрема аудіосигналів, що містять мовлення.
5. Розробити програмне забезпечення для пошуку шаблонів у цифровому сигналі.
6. Проаналізувати вплив зміни параметрів алгоритму на швидкість та ефективність пошуку шаблонів.
7. Провести статистичний експеримент для порівняння ефективності роботи розробленого програмного забезпечення з існуючими методами.

2 ОГЛЯД ІСНУЮЧИХ МЕТОДІВ

Оскільки задача пошуку шаблонів у сигналах була поставлена дуже давно, існує багато методів пошуку шаблону в сигналі. Ці методи різняться за областю застосування, швидкістю, чутливістю до різних перетворень шаблону.

В цьому розділі розглянуті основні існуючі методи:

- взаємокореляція;
- нормалізована взаємокореляція;
- сума квадратів відстаней;
- нормалізована сума квадратів відстаней.

Проведено порівняльний аналіз методів й обґрунтована доцільність дослідження обраного методу.

2.1 Опис предметної області

В цьому підрозділі описані загальні терміни, що необхідні при описі методів.

Ми будемо використовувати в якості сигналу довільну функцію

$$f : X \rightarrow Y,$$

де множина вихідних значень $Y \in \mathbb{R}$, а множина вхідних значень $X \in \mathbb{R}$ для одновимірного сигналу, $X \in \mathbb{R}^2$ — для двовимірного сигналу тощо.

Під цифровим сигналом будемо розуміти функцію f^* , що задана спів-

відношеннями:

$$\begin{aligned} f^*(x_0) &= y_0, \\ f^*(x_1) &= y_1, \\ &\dots \\ f^*(x_n) &= y_n. \end{aligned}$$

Такі функції ще називають таблично-заданими, оскільки вони можуть бути записаними у вигляді таблиці.

Наприклад, для цифрового звукового сигналу, множина $\{x_i\}$ представляє собою моменти часу, в яких визначена амплітуда звукового сигналу $\{y_i\}$. Для цифрового чорно-білого зображення з розмірами 600×400 пікселів, областю визначення цифрового сигналу буде множина пар

$$\{(a, b)\}, a \in \{1, 2, \dots, 599, 600\}, b \in \{1, 2, \dots, 399, 400\}.$$

Областю значень сигналу може бути, наприклад, $\{0, 1, 2, \dots, 254, 255\}$, де значення 0 означає чорний піксель, а 255 — білий.

З кожної функції $f(x)$ можна утворити безліч таблично-заданих функцій, вибравши з області визначення цієї функції скінченну підмножину $\{x_i\}$ й зіставивши цієї підмножині відповідні значення функції $\{f(x_i)\}$. Такий процес ще називається дискретизацією сигналу. Якщо в якості вхідного сигналу виступає сигнал, що залежить від часу (наприклад, аудіо-сигнал), за умови, що $|x_i - x_{i-1}| = \delta$, то частотою дискретизації називають співвідношення $\frac{1}{\delta}$.

2.1.1 Метод ковзаючого вікна

Оскільки задача пошуку шаблону в цифровому сигналі ставиться як пошук наперед визначеного шаблону (або його трансформацій) в дискре-

тньому сигналі, що має (набагато) більшу довжину, то в методах пошуку шаблонів використовується метод ковзаючого вікна [3].

Цей метод полягає в розбитті вхідного сигналу на «вікна» — неперервні ділянки, кожна з яких має одну й ту саму довжину. Як правило, довжина вікна дорівнює довжині шаблону, а самі вікна накладаються одне на одного. Для кожного вікна шукається міра подібності вмісту вікна до шаблону.

Таким чином будується функція $e(x)$, яка характеризує подібність вікна, що задається параметром x до шаблону. За допомогою цієї функції можна оцінювати найвірогідніші позиції шаблону.

Така оцінка, як правило, робиться шукаючи локальні екстремуми, які вище/нижче якогось порогу e_0 .

2.1.2 Віконні функції

При обробці сигналів часто використовуються так звані віконні функції. Такі функції приймають нульове значення за межами певного інтервалу.

Такі функції множаться на значення сигналу для отримання частини сигналу, що складається лише з вікна.

Розглянемо три віконні функції, які найчастіше застосовуються при обробці цифрових сигналів, що містять аудіозаписи мовлення.

а) Прямокутне вікно (також відоме як вікно Діріхле) є найбільш простим з усіх вікон, — всі значення, окрім N значень вікна, дорівнюють нулю:

$$w(n) = 1.$$

Використання такого вікна має вигляд миттєвого «ввімкнення» сигналу на початку вікна та миттєвого «вимкнення» в кінці.

Інші вікна якимось чином прим'якують цей перехід, оскільки такий різ-

кий перехід призводить до небажаних ефектів при використанні таких технік, як дискретне перетворення Фур'є.

Результат застосування прямокутного вікна до (ко)синусоїди наведено на рисунку 2.1.

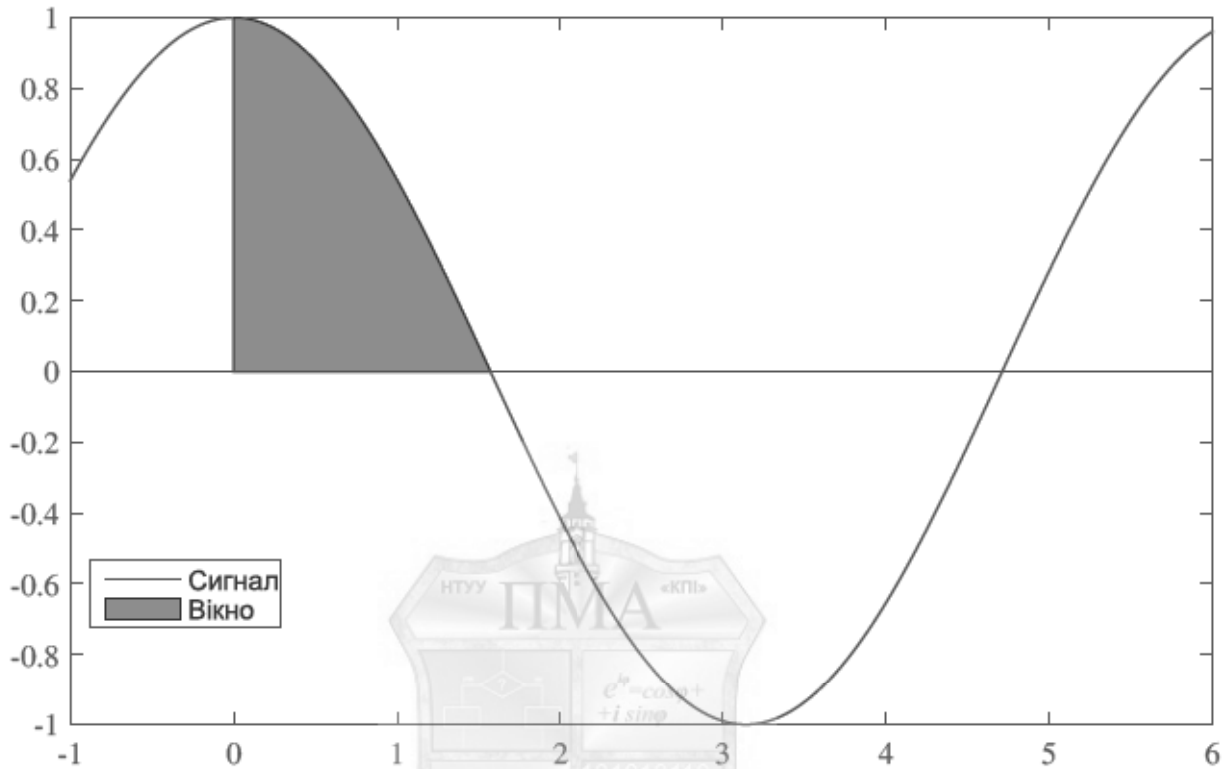


Рисунок 2.1 – Результат застосування прямокутного вікна

б) Вікно Гана (Hann window). Ця функція названа на честь Жуліуса фон Гана (Julius von Hann) й має наступний аналітичний запис [4]:

$$w(n) = \frac{1}{2} \left(1 - \cos \frac{2\pi n}{N-1} \right). \quad (2.1)$$

Ця функція приймає нульові значення на кінцях вікна [5].

Вигляд вагової функції-вікна зображено на малюнку 2.2, а результат застосування до (ко)синусоїди — на рисунку 2.3.

в) Вікно Геммінга [6]. Це вікно має наступний аналітичний запис:

$$w(n) = \frac{25}{46} - \frac{21}{46} \cos \frac{2\pi n}{N-1}. \quad (2.2)$$

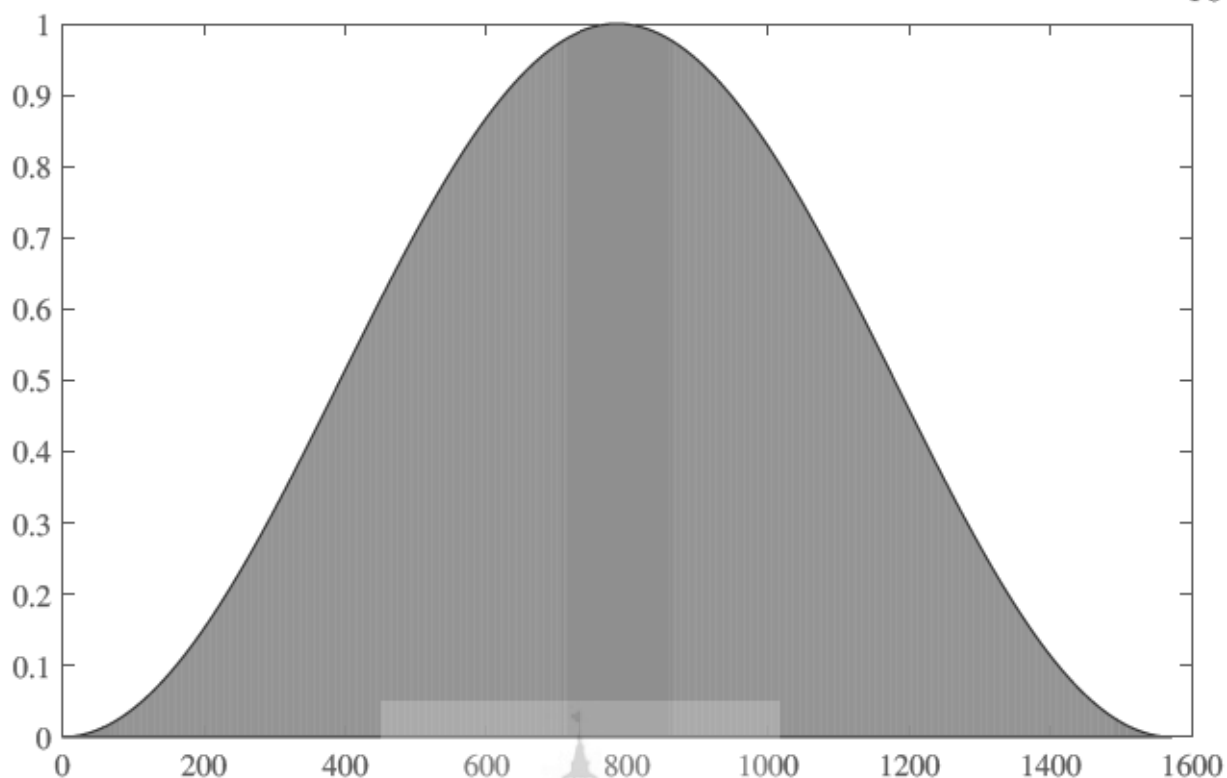


Рисунок 2.2 – Вигляд вікна Гана

Це вікно значно зменшує рівень бокових пелюсток при перетворенні Фур'є [7]. Вигляд вагової функції-вікна зображено на малюнку 2.4, а результат застосування до (ко)синусоїди — на рисунку 2.5.

2.2 Опис існуючих методів

2.2.1 Взаємкореляція

Взаємкореляція (cross-correlation) являє собою міру подібності двох функцій (сигналів), одна з яких зсувається відносно іншої.

Для дискретного сигналу взаємкореляція має вигляд:

$$CC(f,g)[n] \equiv (f \star g)[n] \stackrel{\text{def}}{=} \frac{1}{k} \sum_{m=1}^k f^*[m]g[m+n]. \quad (2.3)$$

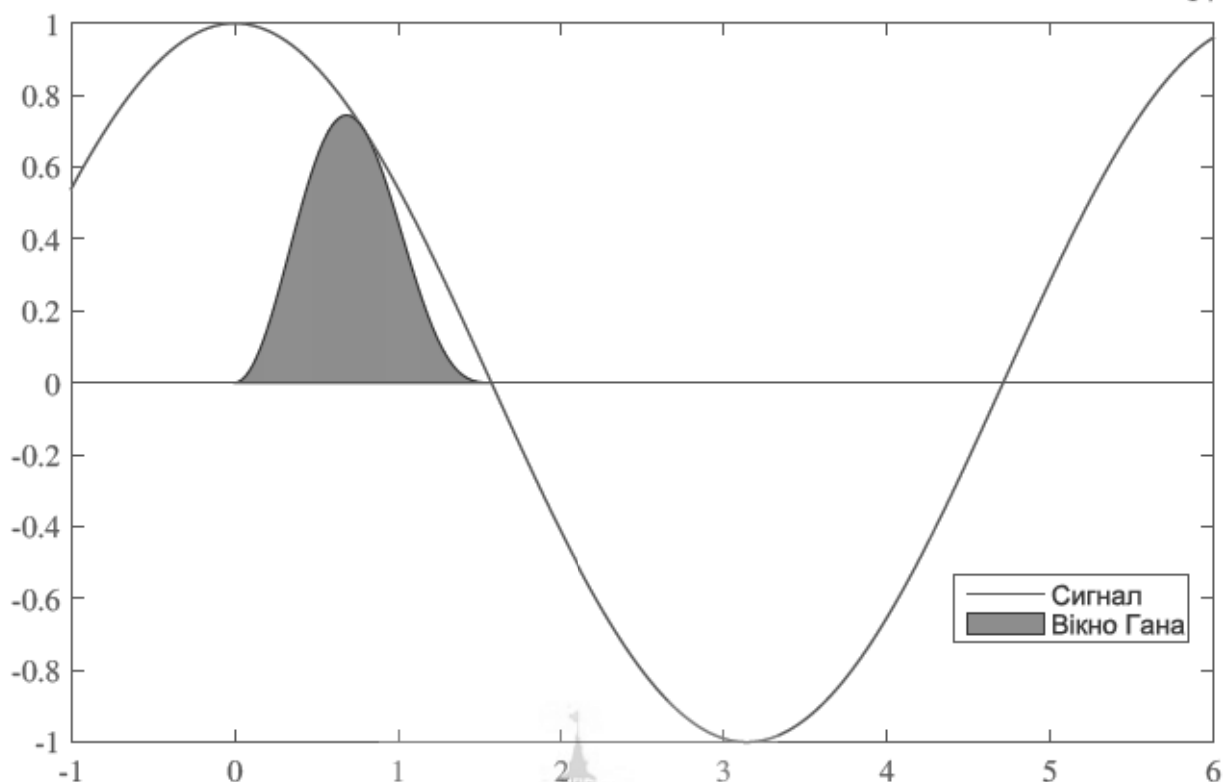


Рисунок 2.3 – Результат застосування вікна Гана

Для використання цього методу необхідно взяти в якості сигналу f вхідний сигнал, g — шаблон, що шукається; в якості k необхідно взяти довжину (кількість значень) шаблону g .

Ідея полягає в тому, що сума добутків значень шаблону на вхідний сигнал буде тим більше, чим більше вхідний сигнал схожий на шаблон на обраному проміжку.

Значним недоліком цього методу є залежність від амплітуди сигналу: якщо сигнал має на деякому проміжку амплітуду, значно більшу за середню амплітуду сигналу та шаблону, то значення взаємкореляції між сигналом та шаблоном на цьому проміжку буде значно більше, ніж значення взаємкореляції між шаблоном та шаблоном (автокореляція).

Саме тому більш практичним вважається використання нормалізованої взаємкореляції.

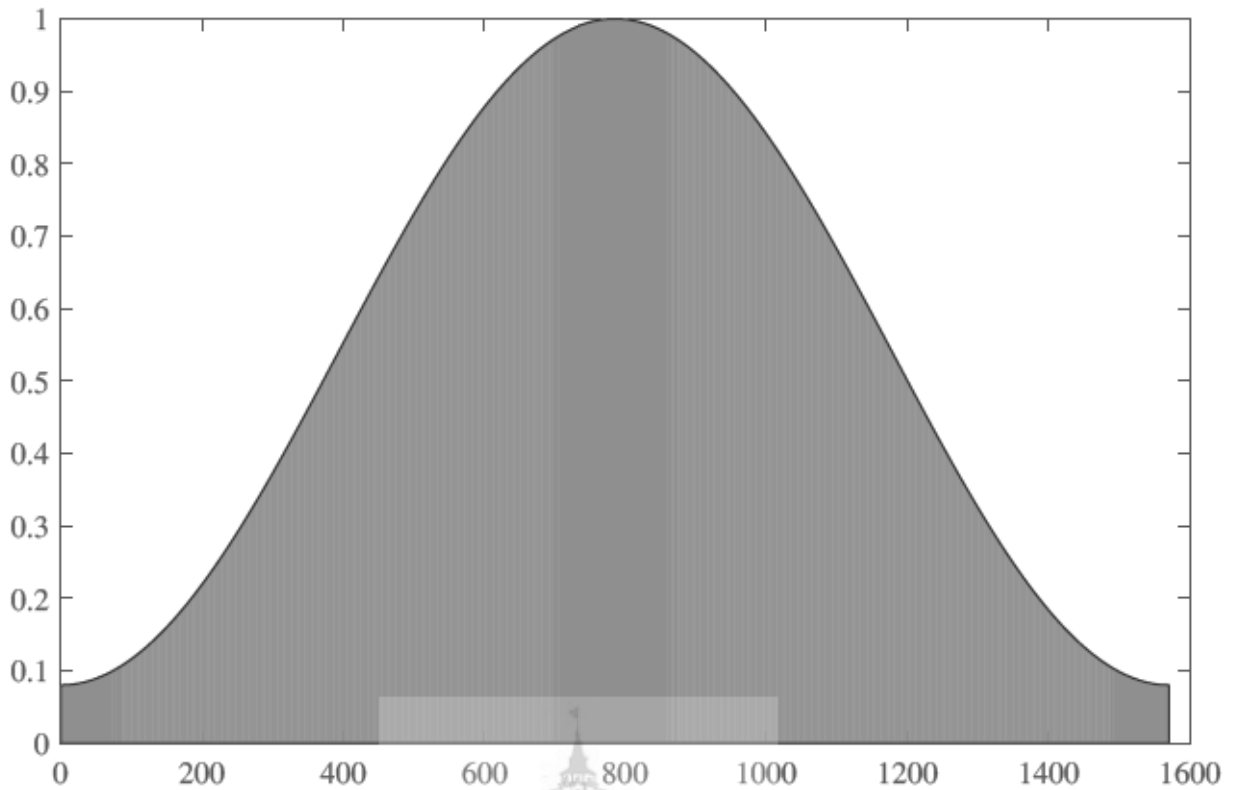


Рисунок 2.4 – Вигляд вікна Геммінга

2.2.2 Нормалізована взаємкореляція

Нормалізована взаємкореляція шукається як взаємкореляція між сигналами, що нормалізовані. Як правило, під нормалізацією розуміється віднімання від сигналу середнього значення на проміжку й ділення на середньо-квадратичне відхилення на цьому проміжку.

Для дискретного сигналу нормалізована взаємкореляція має вигляд:

$$NCC(f,g)[n] \stackrel{\text{def}}{=} \frac{1}{k} \sum_{m=1}^k \tilde{f}_n^*[m] \tilde{g}[m+n], \quad (2.4)$$

$$\tilde{f}_n[m] \stackrel{\text{def}}{=} \frac{f[m] - \bar{f}_n}{\sigma_{f_n}}, \quad (2.5)$$

$$\tilde{g}[m] \stackrel{\text{def}}{=} \frac{g[m] - \bar{g}}{\sigma_g}. \quad (2.6)$$

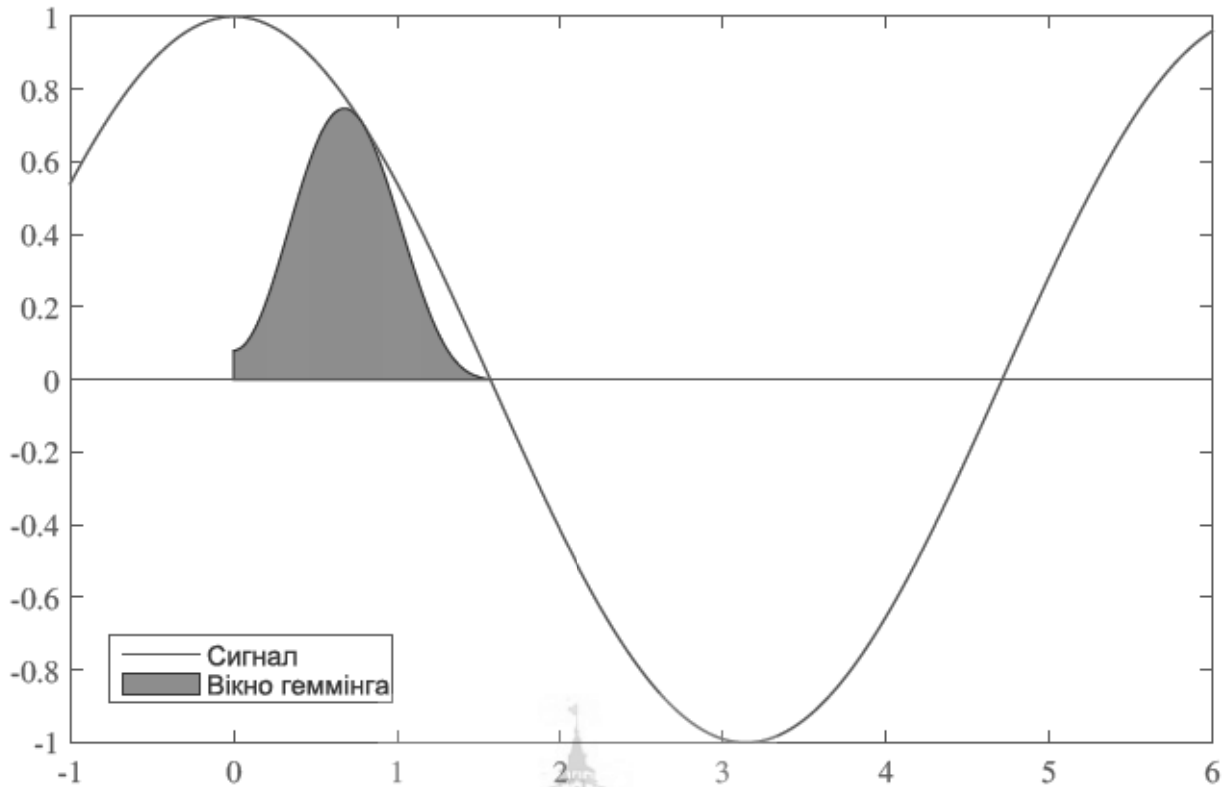


Рисунок 2.5 – Результат застосування вікна Геммінга

де σ_{f_n} – середньо-квадратичне відхилення вікна вхідного сигналу f , σ_g – середньо-квадратичне відхилення шаблону g , \bar{f}_n – середнє значення вхідного сигналу на проміжку вікна, а \bar{g} – середнє значення шаблону.

Цей метод дозволяє знаходити шаблон, навіть якщо вхідний сигнал має значні «спалахи» амплітуди. Також нормалізація шаблону дозволяє знаходити в сигналі шаблон із зміненою амплітудою.

Завдяки тому, що значення цієї метрики лежать в межах від -1 до 1 , задача пошуку порогу e_0 значно спрощується.

2.2.3 Сума квадратів відстаней

Метод полягає в пошуку квадрату відстані між k -вимірному вектору шаблону та обраним вікном сигналу.

Для дискретного сигналу це буде наступна залежність:

$$\text{SSD}(f, g)[n] \stackrel{\text{def}}{=} \frac{1}{k} \sum_{m=1}^k (f[m] - g[m])^2. \quad (2.7)$$

Отримана функціональна залежність характеризує відстань кожного вікна вхідного сигналу до шаблону. Таким чином, чим менше ця відстань — тим більше сигнал у обраному вікні схожий до шаблону.

На відміну від методу взаємокореляцій, на ефективність методу пошуку суми квадратів відстаней не впливає значення амплітуди сигналу.

2.2.4 Нормалізована сума квадратів відстаней

Аналогічно до нормалізованої взаємокореляції можна визначити нормалізовану суму квадратів відстаней:

$$\text{NSSD}(f, g)[n] \stackrel{\text{def}}{=} \frac{1}{k} \sum_{m=1}^k (\tilde{f}[m] - \tilde{g}[m])^2. \quad (2.8)$$

2.3 Порівняльний аналіз методів

В таблиці 2.1 підсумовані основні характеристики розглянутих методів. Слід зазначити, що підрахування взаємокореляції можна значно прискорити завдяки (швидкому) перетворенню Фур'є (FFT).

Таблиця 2.1 – Порівняння існуючих методів

	Складність обчислень	Стійкість до зміни амплітуди шаблону	Стійкість до артефактів вхідного сигналу
CC	$O(kn)$	–	–
NCC	$O(2kn)$	+	+
SSD	$O(kn)$	–	+
NSSD	$O(2kn)$	+	+

Як можна побачити, найпоширеніші методи, що застосовуються для сигналів загального виду, є дуже чутливими до зміни шаблону: тільки методи із нормалізацією сигналів дозволяють коректно знаходити шаблон із лінійно-зміненою амплітудою [8].

Експерименти показали, що ці методи не знаходять шаблон в сигналі, якщо амплітуда шаблону змінювалася не сталим чином.

2.4 Висновки

Було розглянуто предметну область поставленої задачі, описані метод ковзного вікна та віконні функції. Також було розглянуто найпоширеніші існуючі методи пошуку шаблонів й порівняно їх характеристики.

За результатом цього порівняння було визначено, що розглянуті методи при досить гарній швидкодії знаходять такі шаблони, що були змінені майже пропорційно. Тобто шаблони, що деформують з різною силою будуть, погано знаходитись цими методами. Саме тому в наступному розділі буде запропоновано використання розкладення функції в поліноми Кунченка як метод пошуку шаблонів.

3 МЕТОД КУНЧЕНКА

Відомо, що поліном Тейлора є одним з найкращих способів апроксимації функції на деякому інтервалі, проте задля використання цього методу необхідно забезпечити n -раз гладкість функції.

Кунченком Ю. П. було запропоновано використовувати для апроксимації поліном, побудований з неортонормованого базису функцій, що були отримані шляхом певних перетворень з шаблону [9].

В цьому розділі буде запропонований метод пошуку шаблонів за допомогою апроксимації поліномів у просторі Кунченка. Для цього буде розглянутий лінійний простір Кунченка, метод розкладання довільної функції в цьому просторі й застосування цього розкладення в проблемі пошуку шаблонів. Також будуть розглянуті методи, за допомогою яких можна поліпшити результати алгоритму.

3.1 Лінійний простір Кунченка

Нехай $f(x)$ — так звана породжуюча функція, що визначена на проміжку $[a, d]$, а $\{\phi_v(x)\}$ — довільна множина функціональний перетворень. Тоді, можна породити множину функцій

$$u_v(x) = \phi_v(f(x)). \quad (3.1)$$

З цієї множини можна вибрати підмножину лінійно-незалежних функцій. Лінійний простір, утворений на такій множині, називатимемо лінійним простором над незалежними породженими функціями, або лінійний простором Кунченка (LFKu).

На цьому просторі можна звичайним чином визначити скалярний до-

буток:

$$\Psi_{v,k} = (u_v(x), u_k(x)) = \int_a^d u_v(x)u_k(x) dx, \quad (3.2)$$

та відстань між двома елементами:

$$\rho_{v,k}^2 = \|u_v(x) - u_k(x)\|^2 = \int_a^d (u_v(x) - u_k(x))^2 dx. \quad (3.3)$$

Виберемо деяку породжену функцію $u_b(x)$, яку будемо називати основною. Тоді множину функцій $u_v(x)$, $v \neq b$ будемо називати доповнювальними функціями.

Можна утворити поліном Кунченка з доповнювальних функцій:

$$P_t(x) = \sum_{k=1, k \neq b}^t a_k u_k(x). \quad (3.4)$$

Відстань між основною функцією $u_b(x)$ та утвореним поліномом за визначенням буде дорівнювати:

$$\rho_{b,P}^2 = \int_a^d (u_b(x) - P_t(x))^2 dx. \quad (3.5)$$

Можна показати, що коефіцієнти a_k , $k \neq 0$ знаходяться з системи лінійних рівнянь:

$$\sum_{k=1, k \neq b}^t a_k F_{v,k} = F_{v,b}, \quad v = \overline{0, t}, v \neq b, \quad (3.6)$$

де

$$F_{v,k} \equiv \Psi_{v,k} - \Psi_v \Psi_k \|u_0(x)\|^2, \quad (3.7)$$

$$\Psi_v \equiv \frac{\Psi_{v,0}}{\|u_0(x)\|^2}, \quad (3.8)$$

$$a_0 = \Psi_b - \sum_{v=1, v \neq b}^t a_v \Psi_v. \quad (3.9)$$

Спростив вираз, отримаємо:

$$\rho_{bP}^2 = F_{b,b} - J_t, \quad (3.10)$$

де J_t — інфоркуна полінома, що має вигляд:

$$J_t \equiv \sum_{v=1, v \neq b}^t a_v F_{v,b}. \quad (3.11)$$

Тоді мірою апроксимації поліномом будемо вважати наступну величину:

$$e_t = \frac{J_t}{\int_a^d (u_b(x) - \Psi_b u_0(x))^2 dx}. \quad (3.12)$$

3.2 Застосування розкладення для пошуку шаблону в сигналі

Візьмемо в якості породжуючої функції шаблон $f(x)$, що є визначеним на проміжку $\forall x_i \in [a, d]$. Тоді, нехай $\{u_v(x)\}$, також визначені $\forall x_i \in [a, d]$ — породжені функції від шаблону, що утворюють лінійний простір Кунченко [10].

Якщо взяти за основну функцію $u_b(x)$ вхідний сигнал на проміжку

$[a, d]$, за допомогою (3.4) можна побудувати поліном наближення сигналу $P_t(x)$.

Оскільки поліном, визначений таким чином буде являти собою наближення вхідного сигналу за допомогою певних перетворень з шаблону, то можна вважати величину (3.12) мірою схожості вхідного сигналу до шаблону на обраному проміжку.

Таким чином, скомбінувавши метод ковзаючого вікна (пункт 2.1.1) з розкладенням сигналу в лінійному просторі Кунченко, можна отримати ефектограму, тобто функцію $e_t(x)$, що буде характеризувати схожість вхідного сигналу зі шаблоном у певному вікні.

Оскільки ефектограма може приймати значення від -1 до 1 , можливо визначити поріг e_0 , такий, що значення більше (або менше при умові його від'ємності) буде характеризувати відповідне вікно як таке, що містить модифікований шаблон.

Цей алгоритм потребує наступні обчислення:

- $O(t^2)O(k^2)$ обчислень для підрахунку попарних центрованих корелянтів (3.7) всіх породжених функцій (кожна з яких має k значень). Ця дія має бути виконаною лише один раз на початку пошуку, оскільки при ковзанні вікна вони не будуть змінюватись;

- $O(t)O(k)$ обчислень для підрахунку корелянтів із основною функцією (вхідним сигналом) $F_{v,b}$ на кожному кроці, оскільки значення основної функції змінюється із зсувом вікна;

- $O(t^3)$ обчислень для пошуку коренів системи лінійних рівнянь (3.6);

- $O(t)+O(k^2)$ для обчислення ефективності наближення на кожному кроці.

3.3 Методи покращення пошуку шаблонів

Як видно з попередніх обчислень, запропонований метод має більшу обчислювальну складність, ніж метод із використанням нормалізованої

взаємкореляції (розділ 2.3).

В цьому підрозділі будуть запропоновані методи прискорення роботи алгоритму.

3.3.1 Пірамідальний пошук

Оскільки основною складовою кількості обчислень можна вважати $O(k^2)$, було запропоновано використання методу так званого пірамідального пошуку [11, 12].

Ідея цього методу полягає в тому, що спочатку виконується певне масштабування (стиснення) шаблону та сигналу. Після цього виконується пошук стисненого шаблону в стисненому сигналі. Якщо знаходяться екстремуми, що характеризують знахідку шаблону, то на відповідній області нестисненого сигналу виконується пошук нестисненого шаблону [13].

Таким чином, якщо сигнал та шаблон були промасштабовані із співвідношенням 1 : 10, то головна складова кількості обчислень зміниться до $\frac{1}{100}O(k^2)$. Як правило, після першого етапу залишається значно менше областей, в яких потрібно уточнювати результат.

Також варто зазначити, що можливі два підходи щодо стиснення сигналів:

- Зміна частоти дискретизації (resampling). В цьому випадку можна використовувати, наприклад фільтр Ланцоша, якщо оригінальна частота дискретизації не ділиться без залишку на нову. Проте значно ефективніше обирати таку нову частоту дискретизації, що ділить оригінальну частоту націло. В такому випадку можна вибрати кожне m -те значення сигналу та шаблону, де m — співвідношення частот [8];

- Брати середнє значення від деякого околу точки.

3.3.2 Метод прискореного ковзання

Оскільки в процесі пошуку шаблону отримуємо функціональну залежність $e_t(x)$, що характеризує ступінь наближення вхідного сигналу певними модифікаціями шаблону, можна розглядати задачу цього пошуку як задачу оптимізації.

Найбільш ефективним застосуванням методів оптимізації буде використання алгоритм пошуку інтервалів, що можуть містити екстремум функції ефективності наближення. Такий алгоритм можна зображено на малюнку 3.1.

Використовуючи цей алгоритм можна знайти інтервал, що містить певний екстремум, а потім уточнювати отримане значення. Цього можна досягти як повним обчисленням всіх значень функції $e_t(x)$ на отриманому проміжку, так і використанням методів уточнення інтервалів, таких як методи дихотомії та золотого перетину.

Методи дихотомії та золотого перетину дозволяють уточнити (звужити) інтервал, що містить екстремум. Алгоритм методу золотого перетину має наступний вид:

а) Задати

$$x_0 = a,$$

$$x_1 = b.$$

б) Обчислити

$$x_l = x_1 - \frac{x_1 - x_0}{\phi},$$

$$x_r = x_0 + \frac{x_1 - x_0}{\phi},$$

де $\phi = \frac{1+\sqrt{5}}{2}$ — пропорція золотого перетину.

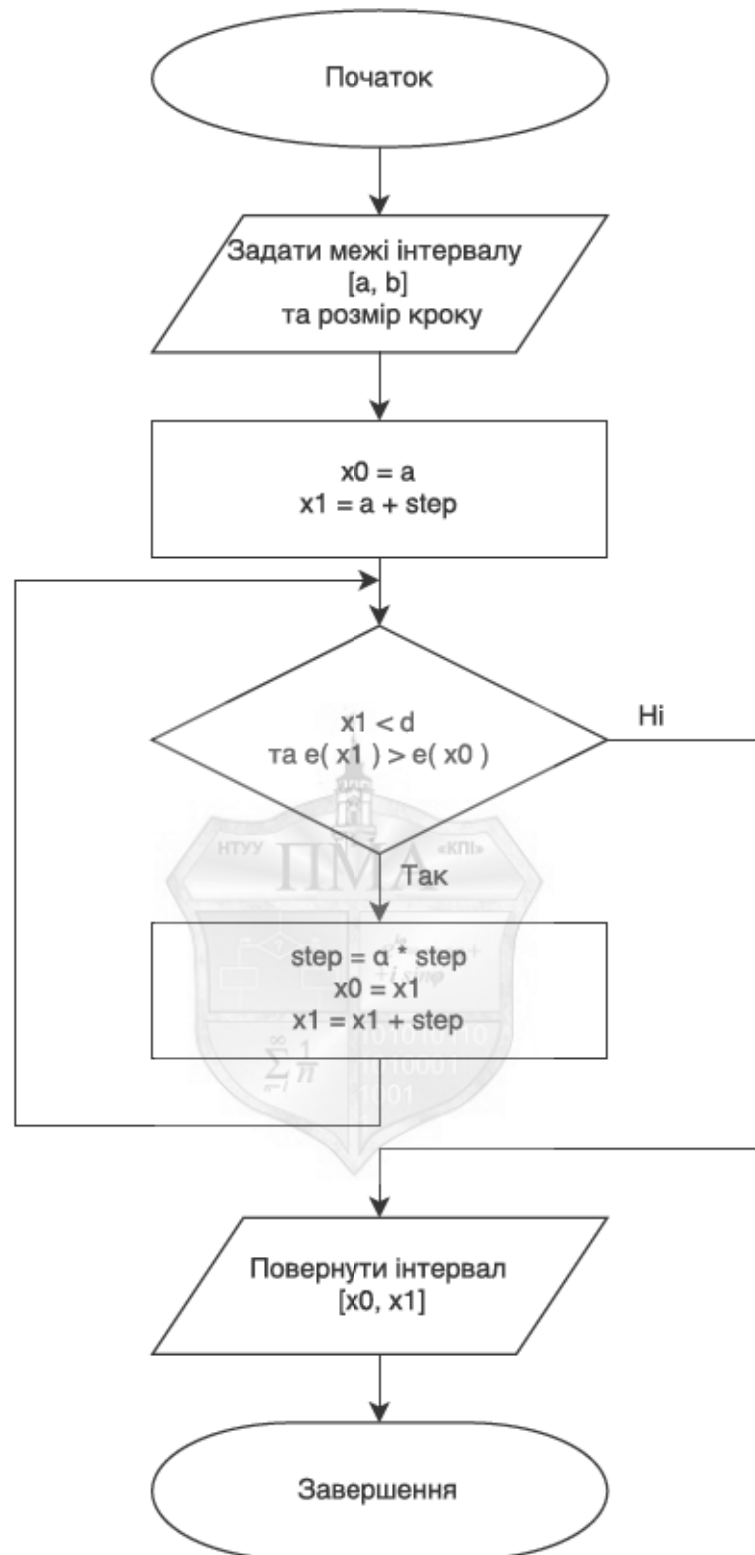


Рисунок 3.1 – Блок-схема роботи алгоритму пошуку інтервалу з екстремумом

в) Якщо $f(x_l) < f(x_r)$, то задати $x_0 = x_l$, інакше $x_1 = x_r$.

г) Перевірити умови зупинки (наприклад, $\|x_1 - x_0\| < \varepsilon$). Якщо умова не виконується, то перейти до кроку в), інакше повернути інтервал $[x_0, x_1]$ або $\frac{x_0+x_1}{2}$ в залежності від потреб задачі.

Слід зазначити, що в залежності від природи сигналу та шаблону, функція ефективності наближення може мати різкі спади, а тому є сенс обмежити максимально можливий крок, який можна пропустити при пошуку інтервалу, що може містити шаблон.

3.4 Пошук шаблонів в мовленнєвих сигналах

Обраний метод пошуку шаблонів має меншу чутливість до зміни шаблону, ніж інші методи. Розглянемо застосування цього методу до задачі пошуку шаблону в мовленнєвому сигналі, оскільки шаблони в сигналах такого типу (деяка мовленнєва одиниця) зазвичай мають змінну у часі амплітуду та енергію, а тому класичні методи дають гірші результати на сигналах такого типу.

Мовлення отримується за рахунок зміни в часі голосового тракту. Через це мовленнєвий сигнал має нестационарну природу. Оскільки при обробці сигналів найчастіше вважається, що амплітуда коливань є більш-менш незалежною від часу, то використання цих методів стає неприпустимим при обробці мовленнєвих сигналів. Використання таких методів до мовленнєвих сигналів буде мати дуже мало практичного значення [14]. Наприклад, обчислення енергії сигналу є фундаментальним інструментом при обробці сигналів:

$$E_T = \sum_{n=-\infty}^{\infty} s^2(n). \quad (3.13)$$

Енергія сигналу є корисним інструментом у випадку, колу сигнал має обмежену енергію. Обчислення енергії мовленнєвого сигналу буде мати ду-

же малу практичну цінність, адже мовленнєвий сигнал за природою має енергію та амплітуду такі, що змінюються в часі. Саме через це при обробці мовленнєвих сигналів є необхідність у інструментах, що дозволяють отримувати інформацію про змінну в часі енергію.

Було запропоновано використовувати існуючі методи, але в дещо зміненому форматі: використовувати ці інструменти лише тоді, коли можна вважати, що енергія сигналу є стаціонарною [15]. Мовленнєвий сигнал можна вважати стаціонарним, якщо розглядати його блоками (вікнами) по 10–30 мілісекунд. Така техніка має назву коротко-часова обробка сигналів (short term processing).

Коротко-часова енергія сигналу полягає в обчисленні енергії для кожного блоку (вікна) в 10–30 мілісекунд. Вікна можна брати як такі, що перетинаються, так і такі, що не перетинаються. Також можна зазначити, що енергія вікна, що містить певний мовленнєвий сигнал, зазвичай є значно вищою за енергію вікна без мовленнєвого сигналу.

Коротко-часова енергія в загальному випадку має наступний вигляд:

$$e[n] = \sum_{m=-\infty}^{\infty} (s(m)w(n-m))^2, \quad (3.14)$$

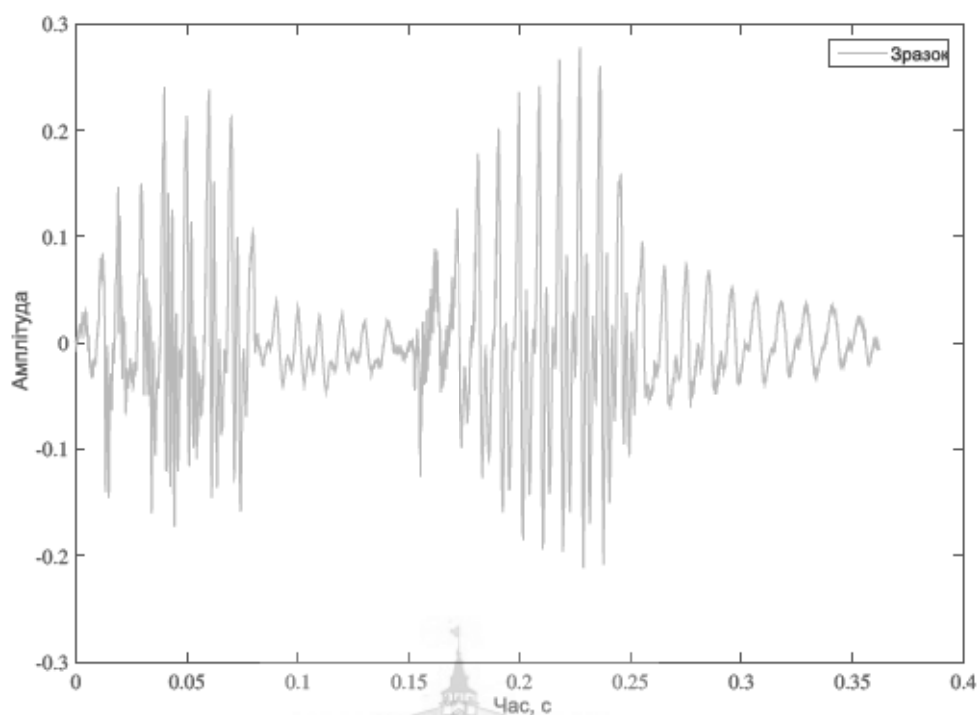
де w — аналітичний запис функції-вікна (пункт 2.1.2).

На рисунку 3.2 наведено приклад коротко-часової енергії мовленнєвого сигналу.

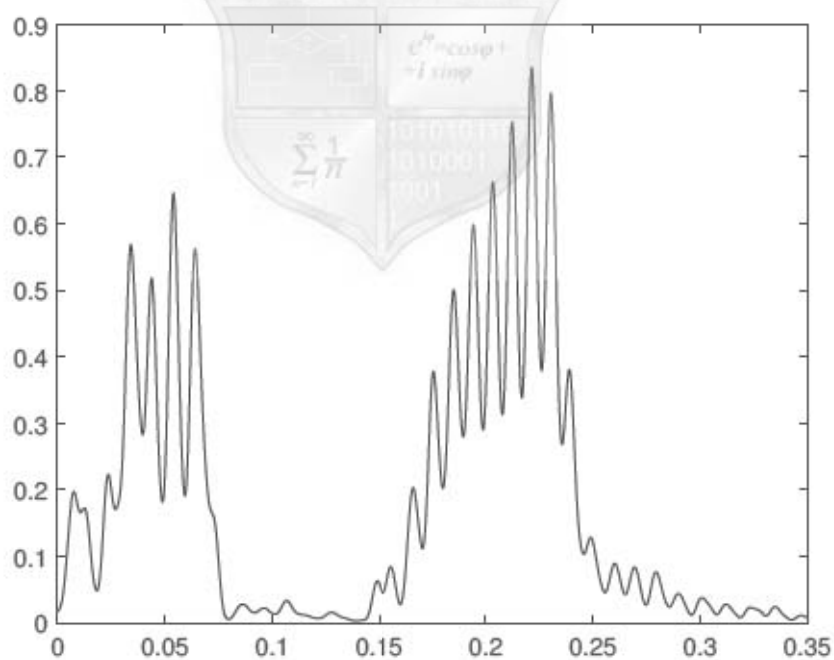
Як було зазначено, розмір вікна зазвичай обирається 10–30 мілісекунд, оскільки на таких проміжках енергія мовленнєвого сигналу вважається статичною. Якщо обирати розмір вікна більше за 40 мілісекунд, то отримана коротко-часова енергія буде більш гладкою.

Гладкість та означеність екстремумів коротко-часової енергії також залежить від вибору вікна.

В розділі 5 будуть розглянуті впливи обраної форми та розміру вікна на коротко-часову енергію.



(а) Сигнал



(б) Коротко-часова енергія

Рисунок 3.2 – Приклад обчислення коротко-часової енергії

3.5 Висновки

В цьому розділі був визначений лінійний простір над породжуючою функцією, або лінійний простір Кунченко. Було показано як, використовуючи наближення функції елементами базису цього простору, можна вирішити задачу пошуку шаблону в цифровому сигналі. Обраний метод має більш нечітку оцінку присутності шаблону в сигналі, ніж розглянуті в попередньому розділі. Також були розглянуті методики, завдяки яким можна прискорити роботу методу.



4 ПРОЕКТУВАННЯ ПРОГРАМНИХ ЗАСОБІВ

В цьому розділі будуть сформульовані вимоги до програмного середовища, в якому необхідно реалізувати метод пошуку шаблонів за допомогою поліномів Кунченко. Будуть наведені можливі варіанти вибору й обґрунтований вибір. Буде спроектована архітектура програмної реалізації та наведений алгоритм роботи цієї реалізації.

4.1 Вибір програмного середовища

Для полегшення розробки та забезпечення високої якості роботи програмних засобів, було вирішено використати один з існуючих математичних пакетів для роботи з цифровими сигналами.

Серед інших, розглядалися наступні математичні пакети: Mathworks[®] MATLAB[®] 2014b, GNU Octave, Python+Numpy. Порівняльна характеристика цих пакетів наведена в таблиці 4.1.

Таблиця 4.1 – Порівняльна характеристика математичних пакетів

	Кросплатформенність	Мінімальний розмір	Ліцензія
Matlab	+	1200 MiB	Propr.
Octave	+	650 MiB	GPL
Numpy	+	112 MiB	BSD-new

Усі ці середовища зарекомендували себе як якісне математичне забезпечення для обробки сигналів.

Matlab є найстарішим математичним пакетом із пакетів, що розгляда-

лися. Для цієї середи було написано безліч модулів для обробки сигналів і не тільки. Також, із усіх пакетів, що розглядалися, тільки в цьому пакеті є вбудована можливість апаратного прискорення обчислень на відеочіпі, що дає значний приріст продуктивності при обробці цифрових сигналів.

Середовище Octave було розроблено як альтернатива Matlab. Воно підтримує синтаксис вбудованої мови Matlab, і саме тому безліч модулів, що були написані для Matlab, працюють і на Octave. Цей пакет є програмним забезпеченням із відкритим кодом, що є значною перевагою над Matlab. Також варто зазначити, що цей пакет розвивається набагато швидше за Matlab, а тому, можливо, в майбутньому в ньому будуть функції, що зараз є лише в Matlab (наприклад, апаратне прискорення обчислень).

Numpy, на відміну від Matlab та Octave, є набором модулів/бібліотек для мови програмування Python. З усіх систем, Numpy розвивається найактивніше. До переваг Numpy можна віднести те, що програми на Numpy — це програми на мові Python, яка є набагато поширенішою за мови Matlab/Octave. Також ліцензія, за якою розповсюджується Numpy є більш вільною за GPL (за якою розповсюджується Octave), що дозволяє використовувати це середовище для комерційної розробки.

Для початкового проектування й дослідження було обрано середовище Matlab, оскільки знайти необхідний модуль для цього середовища значно легше, а можливість використовувати апаратне прискорення також є дуже бажаною характеристикою.

4.2 Проектування архітектури програмної реалізації

Оскільки було обрано математичне середовище Matlab, програмна реалізація методу пошуку шаблонів за допомогою розкладення функції поліномами Кунченка складається із наступних функцій-скриптів:

- `match` — функція-інтерфейс, точка входу до модуля пошуку шаблонів. Отримує параметрами цифровий сигнал, шаблон, функціональні

перетворення, за допомогою яких будується лінійний простір над породжуючою функцією. Саме в цій функції виконується «ковзання» вікна. Ця функція, в свою чергу, викликає `buildGeneratedFunctionsSystem`, `createIterator`, `approximate`;

- `buildGeneratedFunctionsSystem` — функція, що будує базис лінійного простору Кунченко та обчислює центровані скалярні добутки між ними;

- `createIterator` — створює вікно й дозволяє рухати це вікно заданим чином;

- `approximate` — обчислює апроксимацію обраного вікна вхідного сигналу базисом лінійного простору Кунченко. В цій функції будується система лінійних рівнянь, з якої знаходяться коефіцієнти при базисних функціях лінійного простору, рахується коефіцієнт наближення вхідного сигналу за допомогою породжених функцій.

Таким чином було створено наступні рівні абстракції:

- базис лінійного простору, що містить наступні поля:

- породжуюча функція (шаблон);
- породжені функції (функціональні перетворення шаблону);
- посилання на функцію скалярного добутку;
- наперед підраховані значення скалярного добутку для усіх пар породжених функцій.

Посилання на функцію скалярного добутку в цій структурі дозволяє змінювати метод пошуку корелянти між двома функціями. За замовчанням це є чисельне інтегрування методом трапецій від попарного добутку значень аргументів;

- ітератор в сигналі, який дозволяє певним чином обходити довільний сигнал. Змінюючи реалізацію цього ітератору можна змінювати вигляд вікна вхідного сигналу, змінювати шаг між вікнами тощо.

4.3 Алгоритм роботи програми

На рисунку 4.1 наведено блок-схему алгоритму пошуку шаблонів.

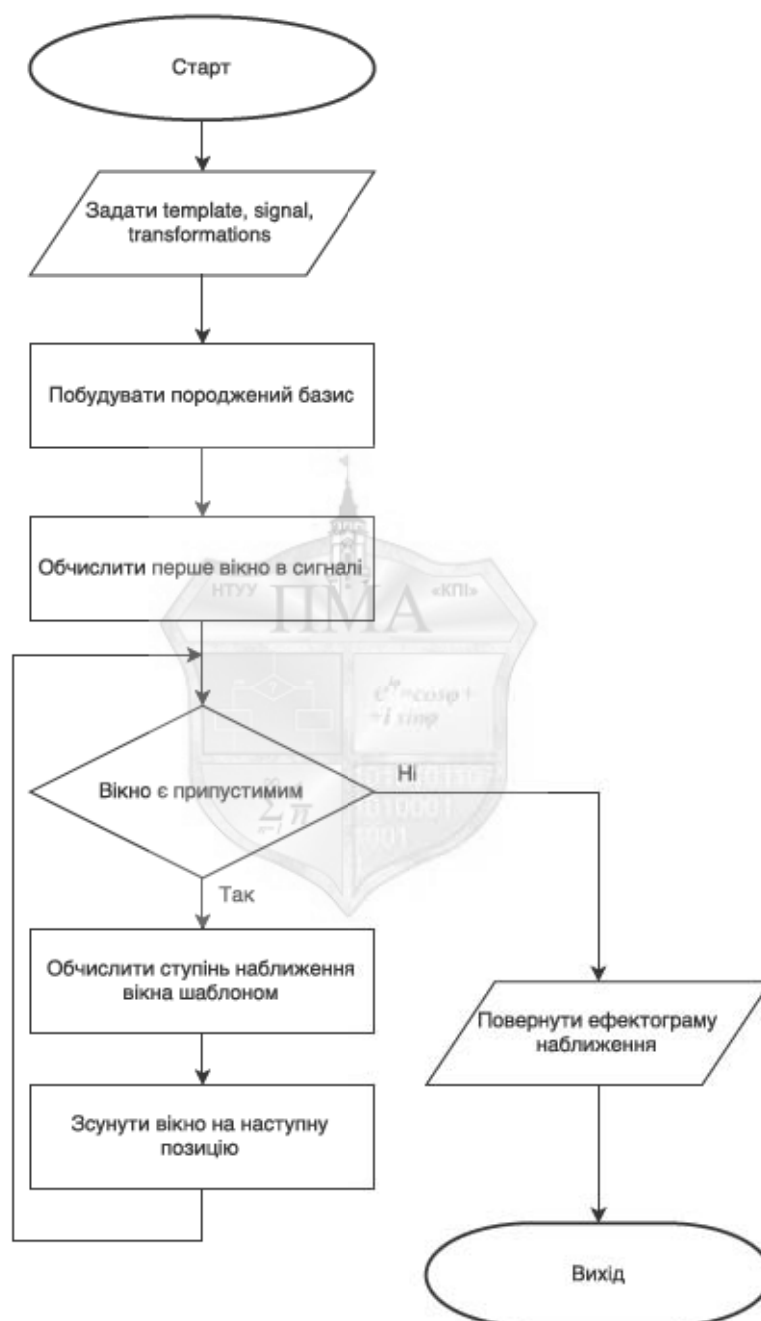


Рисунок 4.1 – Блок-схема роботи алгоритму пошуку шаблону

4.4 Висновки

В цьому розділі було обґрунтовано вибір середовища Matlab як середовища з великою кількістю модулів для обробки цифрових сигналів.

Завдяки абстрагуванню спроектована система має достатню гнучкість, щоб змінювати метод пошуку наступного вікна та обчислення скалярного добутку. Завдяки цьому отримана система дозволяє шукати шаблон не лише в одновимірному сигналі, а й в двовимірному.

Також можна зазначити, що після дослідження методу пошуку шаблонів за допомогою розкладення функції в лінійному просторі Кунченка, систему можна буде перевести в середовище NumPy, щоб можна було користуватися багатьма системними бібліотеками, що розроблені для мови програмування Python. Це дозволить програмному продукту мати, наприклад, графічний інтерфейс користувача.

Також має сенс мати сумісність з математичним середовищем Octave, адже останнє, на відміну від Matlab, є безкоштовним ПЗ з відкритим вихідним кодом.



5 МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА СТАТИСТИЧНИЙ ЕКСПЕРИМЕНТ

В цьому розділі буде проведений статистичний експеримент для того, що порівняти запропонований метод із існуючими класичними методами пошуку шаблонів.

Усі розрахунки проводились у математичному середовищу MathWorks MATLAB 2014b x64, що працювала на Intel™ Core i5-4590 (3,3 ГГц) та 16 ГіБ оперативної пам'яті під керування операційної систему GNU/Linux.

5.1 Пошук у синтетичному сигналі

В цьому підрозділі розглянемо пошук шаблону в сигналі, що був згенерований як певна композиція неперервних функцій.

Для початку порівняємо методи пошуку шаблонів, використовуючи сигнал із шаблоном, котрий був доданий без будь-яких видозмінень (окрім адитивного шуму/трансформації).

В рамках експерименту на проміжку $[0; 10]$ із частотою дискретизації 10 КГц будувався сигнал, що складався з 8–16 гаусіан нормального розподілу $\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ (кількість гаусіан і їх параметри обиралися випадковим чином). До цього сигналу (бази) у випадкову позицію додавався шаблон виду $\sin \alpha x^2$. Після накладання шуму, в отриманому сигналі за допомогою розглянутих методів шукалося місцезнаходження шаблону, яке порівнювалося з визначеною на попередньому кроці позицією. Приклад такого сигналу наведений на рисунку 5.1.

Результати статистичного експерименту наведено в таблиці 5.1.

На рисунку 5.2 наведені відповідні ефектограми нормалізованої взаємкореляції, нормалізованої суми квадратів відстаней та методу на основі

поліномів Кунченка. Як видно з цих ефектограм, всі розглянуті методи досить точно знаходять шаблон у штучному сигналі. Але, як видно з таблиці результатів, середній час роботи методу поліномів Кунченка значно більший за час роботи існуючих методів.

На рисунку 5.3 наведені проміжні ефектограми поліномів Кунченка при застосуванні пірамідального пошуку для прискорення роботи алгоритму з трьома проміжними шарами: 100, 1000 та 10 000 Гц відповідно.

На прикладі, що зображений на цьому рисунку видно, що на першому шарі було знайдено 4 екстремуми, що були більшими за поріг 0,3. Окопи отриманих екстремумів були розглянуті на наступному шарі, де було знайдено лише один екстремум, більший за поріг 0,01. На наступному кроці в околі цього екстремум було уточнене значення ефектограми.

Таким чином, результати статистичного експерименту показали, що якщо використовувати штучні сигнали для оцінки, то метод поліномів Кунченка має приблизно таку ж саму точність пошуку, що й розглянуті існуючі методи, а використання пірамідального пошуку дозволяє отримувати той самий результат навіть швидше за них.

Таблиця 5.1 – Результати пошуку шаблону без модифікації

	Час, с	Абсолютна похибка	Екстремум
Kunchenko	120,01	-0,02	0,05
NCC	5,95	-0,02	0,24
NSSD	5,78	-0,02	1,52
Kunchenko + Пірамідаль- ний пошук	0,81	-0,02	0,05

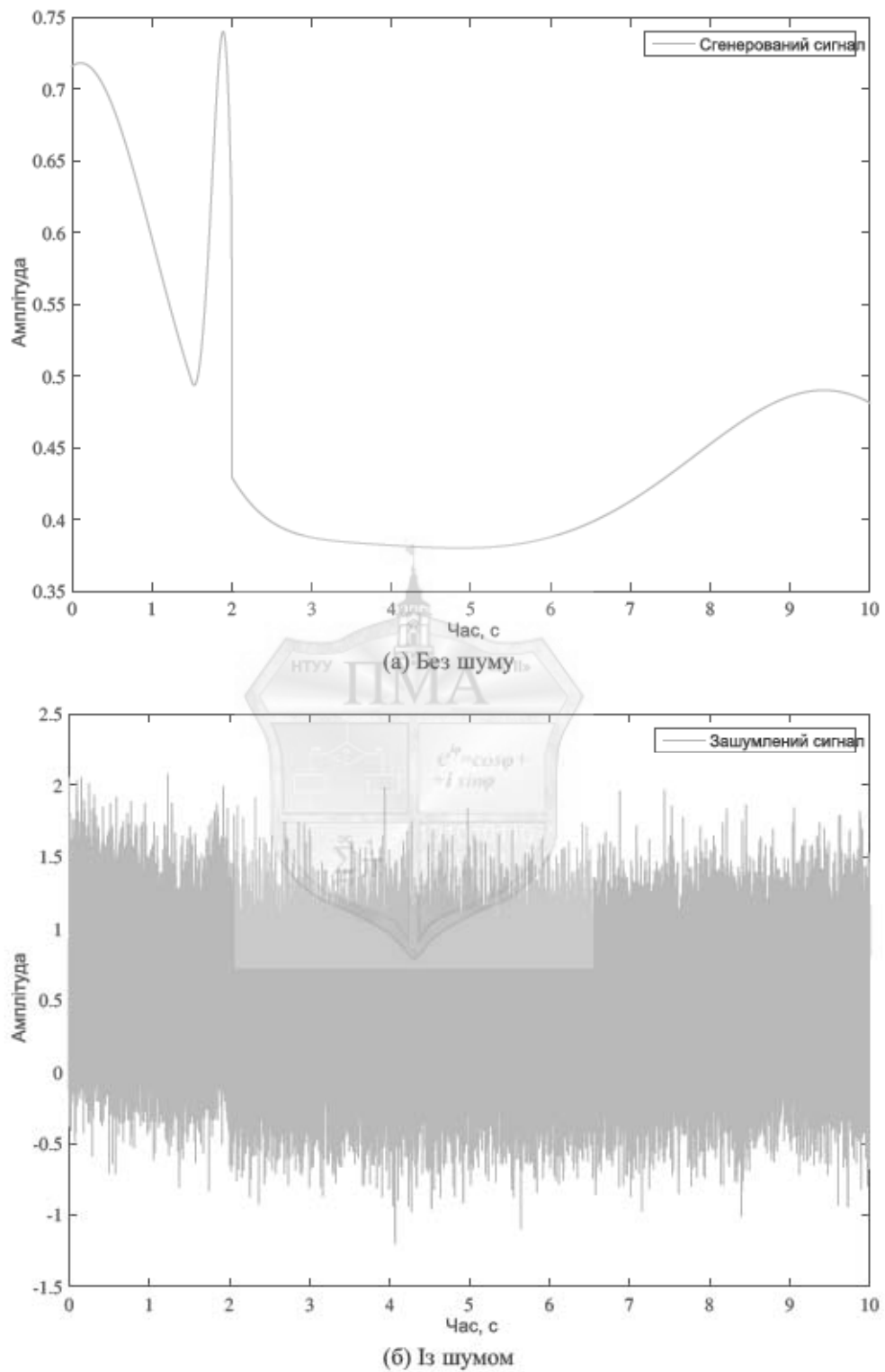
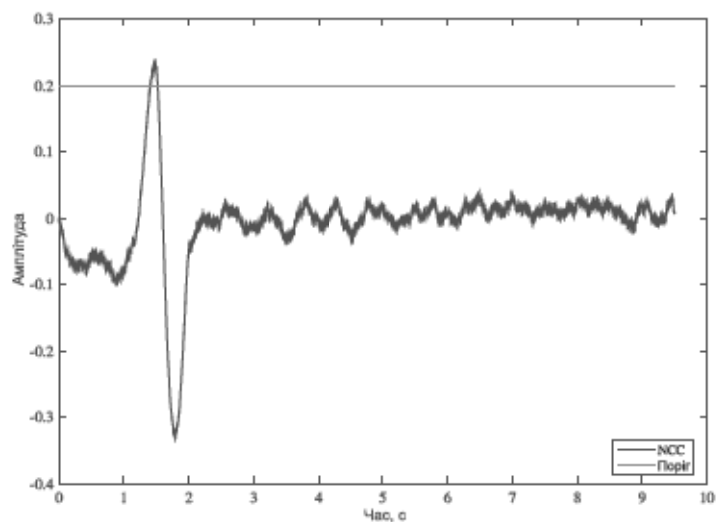
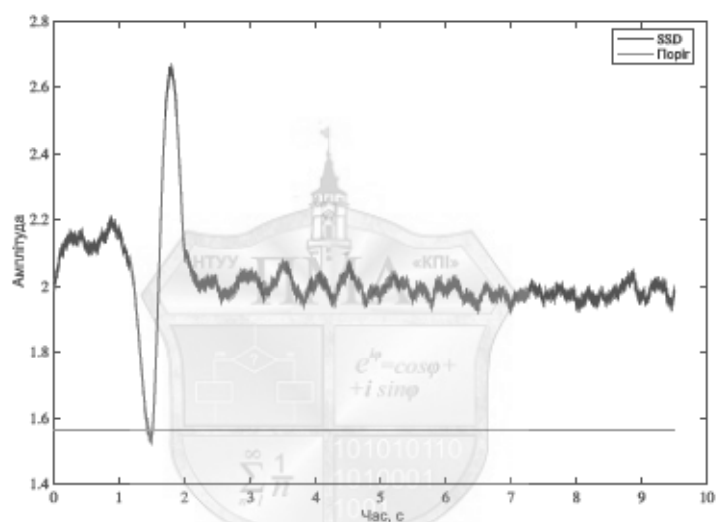


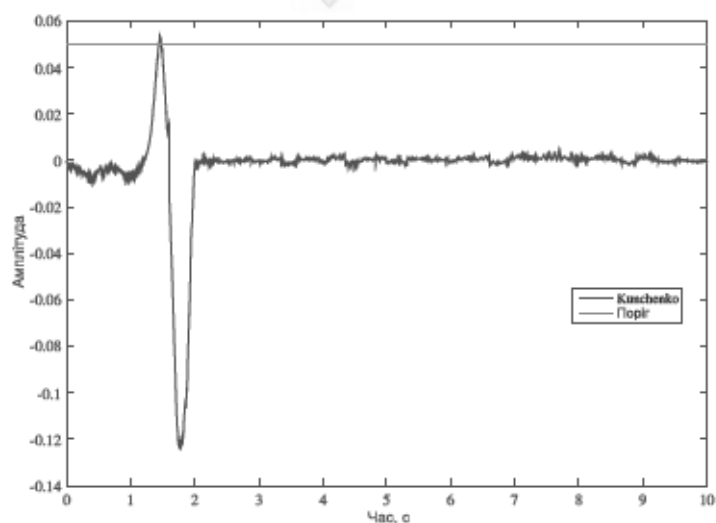
Рисунок 5.1 – Приклад сигналу із шаблоном без модифікації



(a) NCC

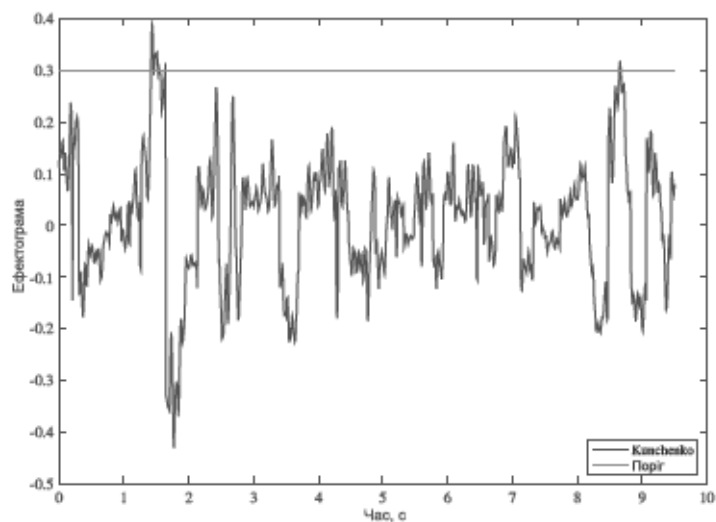


(б) NSSD

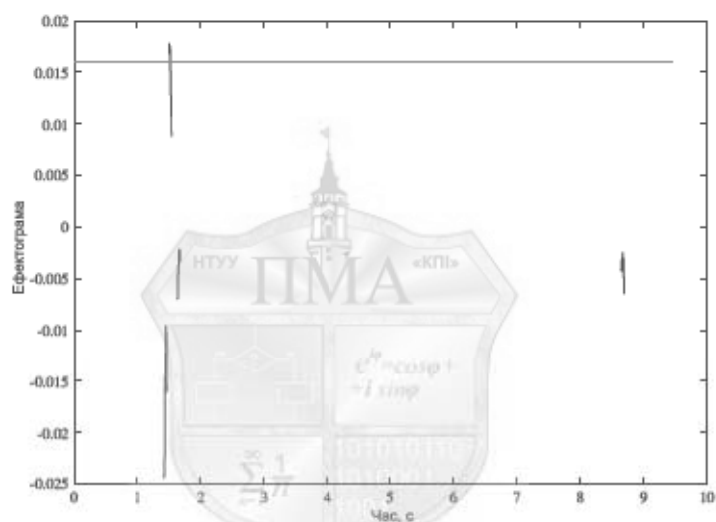


(в) Kunchenko

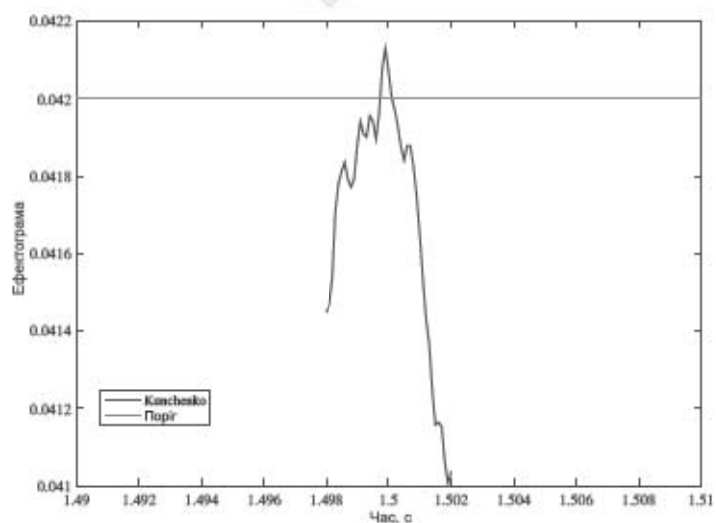
Рисунок 5.2 – Пошук шаблону без модифікацій в синтетичному сигналі



(a) Крок 0,01



(б) Крок 0,001



(в) Крок 0,0001

Рисунок 5.3 – Застосування пірамідального пошуку

5.2 Пошук шаблону у мовленнєвому сигналі

Протестуємо метод Кунченко для пошуку шаблону в мовленнєвому сигналі.

Як тестові дані було використано заздалегідь підготовлені акустичні файли, що містили запис озвучення ряду цифр. Як шаблон використовувався записаний аналогічно файл, що містив записи озвучення окремих цифр. Акустичний файл був оцифрований із частотою 8 000 Гц.

Таким чином, пошук шаблону повинен був виявити моменту часу в вхідних даних, коли було озвучено ту ж саму цифру, що й в шаблоні.

Як було розглянуто в розділі 3, при обробці мовленнєвих сигналів було запропоновано використовувати коротко-часову енергію.

При обчисленні коротко-часової енергії отримане значення буде залежати від довжини вікна, віконної функції та кроку між вікнами. В експериментах довжина вікна була обрана сталою та рівною 20 мілісекунд.

В експерименті буде проаналізовано залежність ефективності пошуку шаблону розглянутими методами від

- обраного представлення сигналу та шаблону (акустичний сигнал чи коротко-часова енергія); для коротко-часової енергії ще будуть розглянуті наступні параметри:

- віконна функція, що використовується при обрахунку енергії;
- крок між вікнами (мінімальний чи половина вікна);
- віконної функції, що використовується у методі ковзного вікна.

Розглянемо результати тестування для сигналу, що містить озвученим наступний ряд: 1; 2; 3; 4; 5; 6; 1; 2; 3; 1; 2; 3. Як шаблон розглядатимемо аудіо-запис, що містить озвучення цифри 1. В сигналі, що розглядається, цифра 1 вимовляється приблизно на 0,5, 5 та 7 секундах.

5.2.1 Пошук шаблону в аудіосигналі

На рисунку 5.4 наведено вигляд аудіосигналів.

На рисунках 5.5, 5.6 та 5.7 зображено результати пошуку шаблону в аудіосигналі.

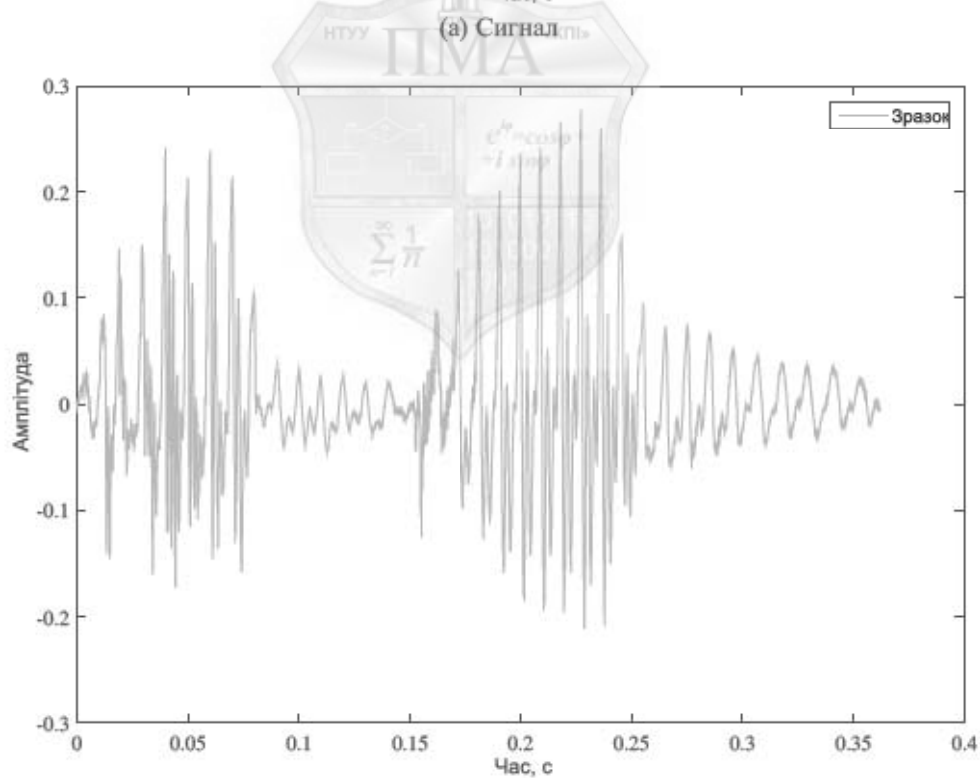
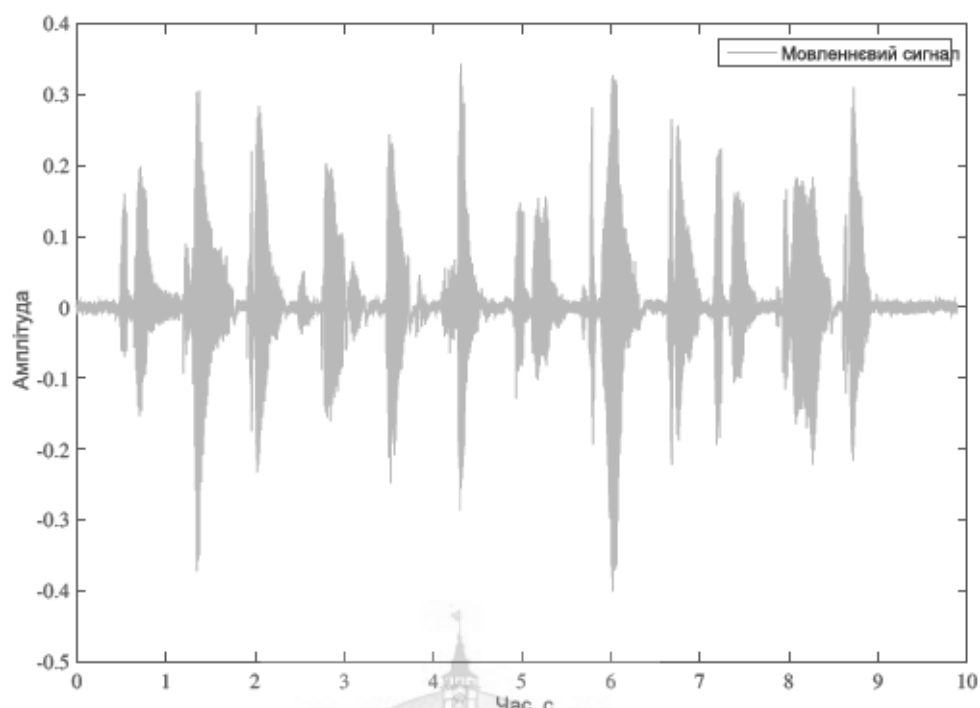
З цих результатів видно, що використання жодного з методів не дозволило знайти усі входження шаблонів в сигналі. З цих результатів можна зробити наступні спостереження:

- екстремуми, що відповідають коректним позиціям шаблону в сигналі, при використанні методу поліномів Кунченка, як правило, є більш вираженими ніж у інших методів;

- використання вікон, відмінних від прямокутного значно зменшує амплітуду ефектограм в зонах, що не містять шаблону; В цих зонах все ще є певні екстремуми, але вони стають більш вираженими.

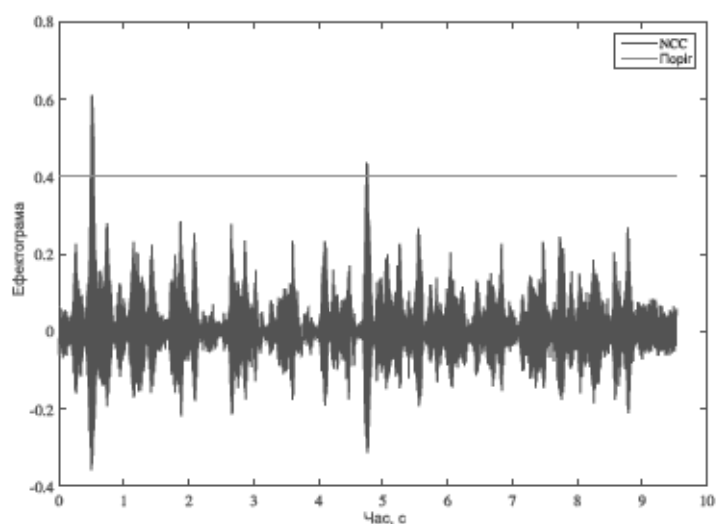
- метод сум квадратів відстаней при використанні вікон, відмінних від прямокутних, має тільки один коректний екстремум, менший за поріг.



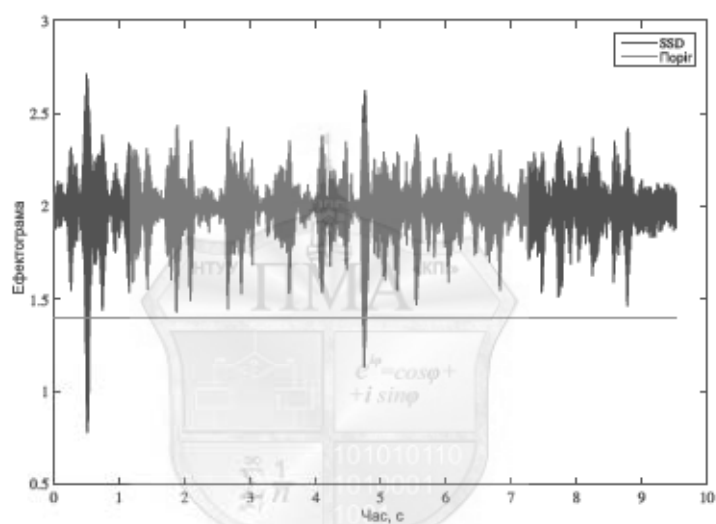


(б) Зразок

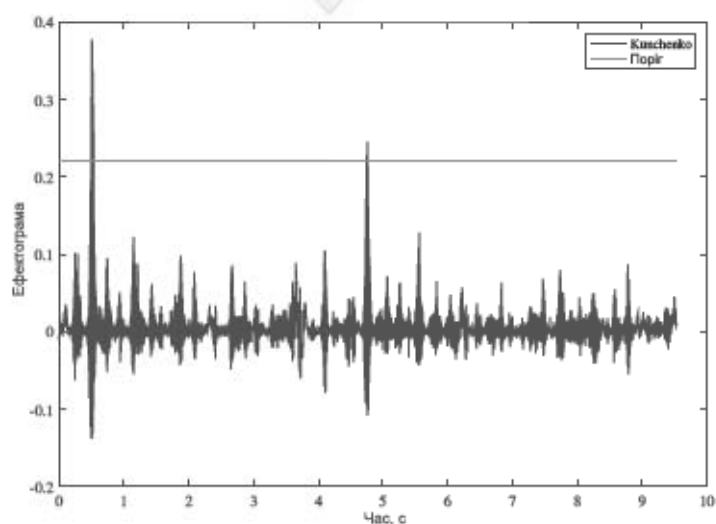
Рисунок 5.4 – Запис мовлення й шаблон для пошуку



(a) NCC

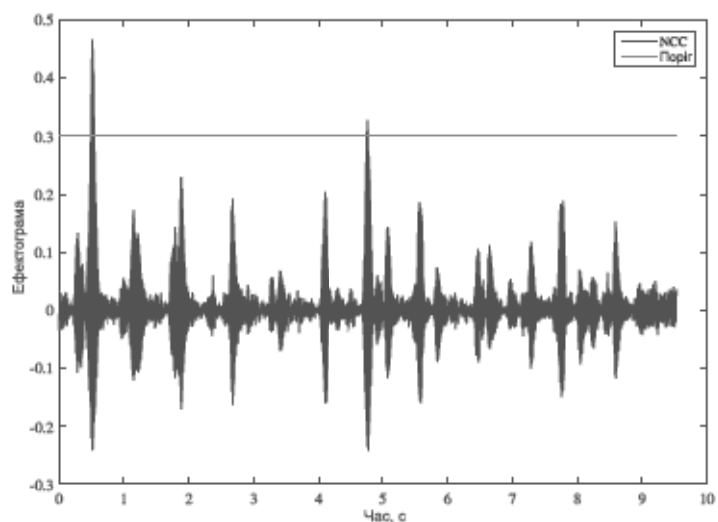


(б) NSSD

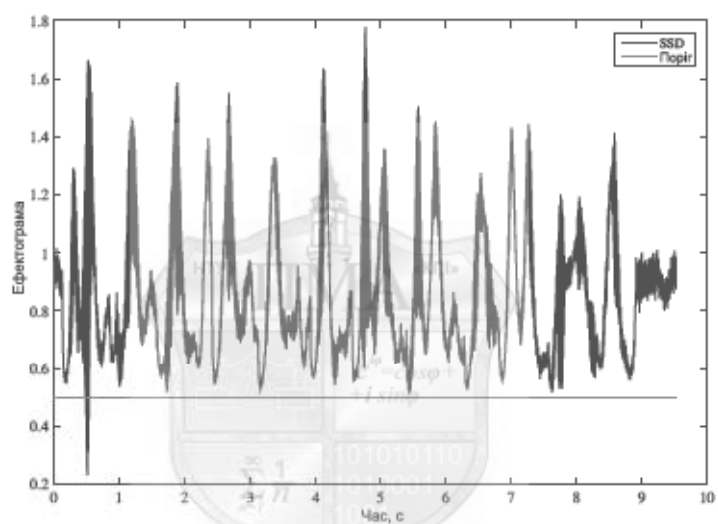


(в) Kunchenko

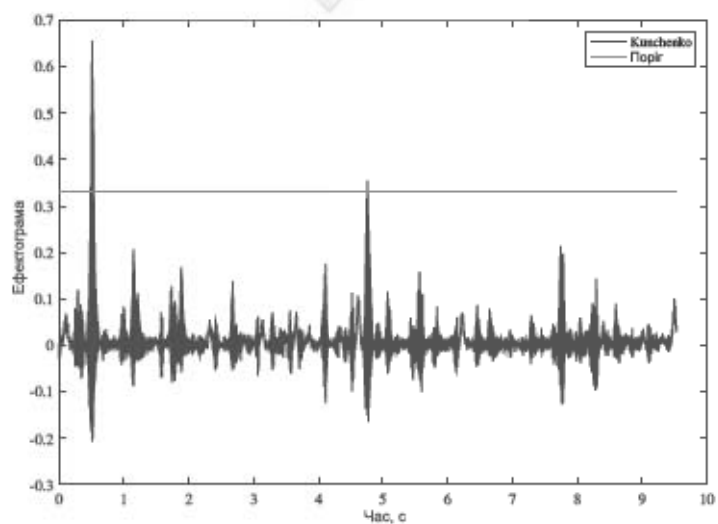
Рисунок 5.5 – Результати пошуку шаблону, використовуючи прямокутне
ВІКНО



(a) NCC

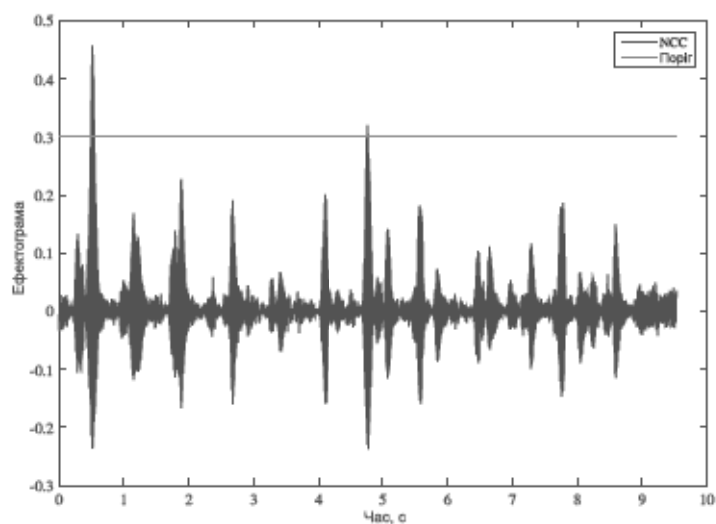


(б) NSSD

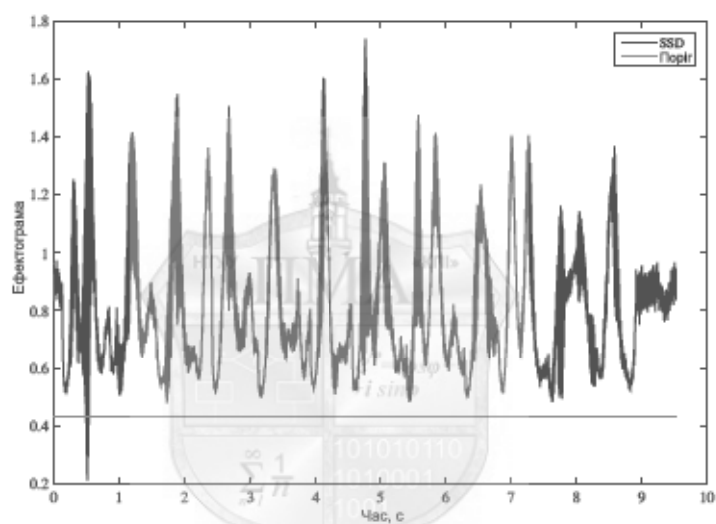


(в) Kunchenko

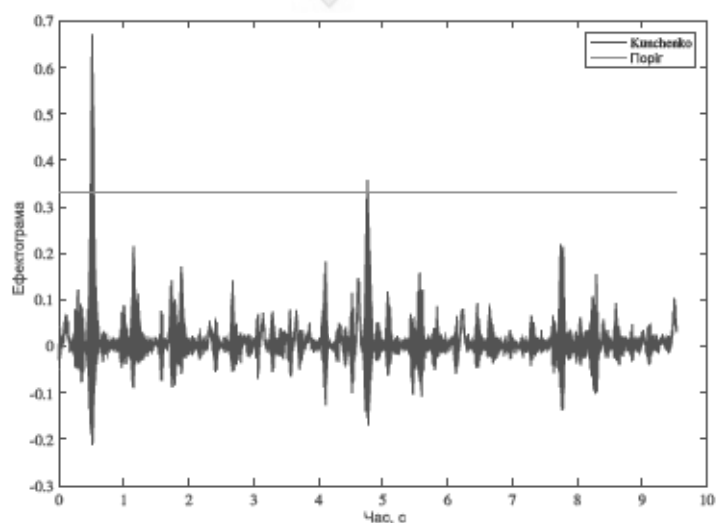
Рисунок 5.6 – Результати пошуку шаблону, використовуючи вікно Геммінга



(a) NCC



(б) NSSD



(в) Kunchenko

Рисунок 5.7 – Результати пошуку шаблону, використовуючи вікно Гана

5.2.2 Пошук шаблону в коротко-часовій енергії

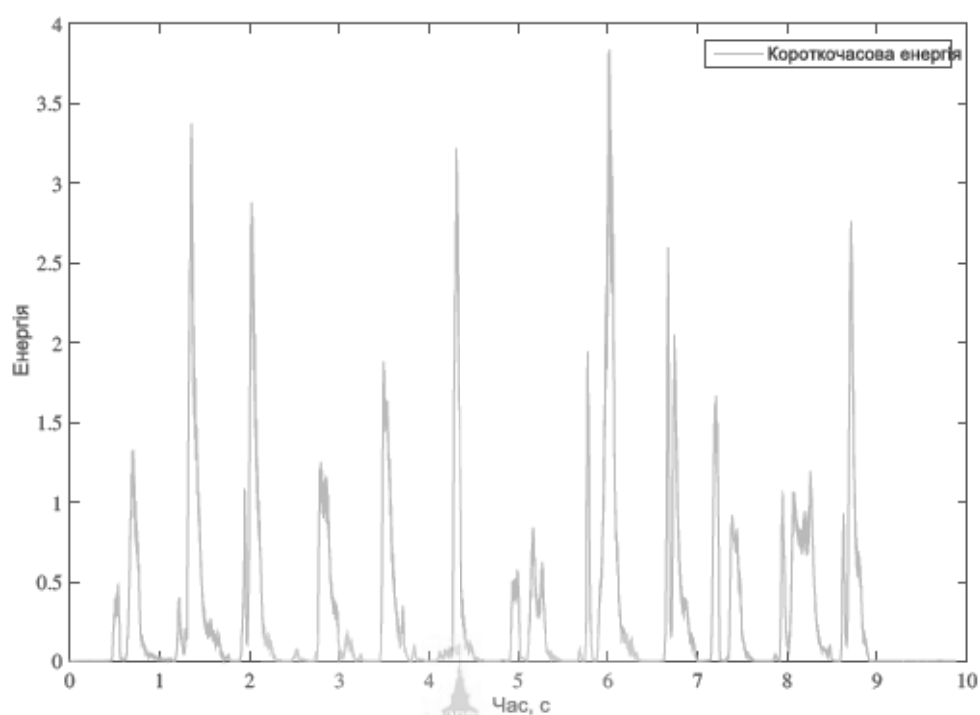
Розглянемо поведінку розглянутих алгоритмів для пошуку шаблонів в коротко-часовій енергії сигналу, що була отримана з використанням найпростішого й найшвидшого методу виділення вікон — прямокутні вікна. Такі вікна бралися з мінімально можливим кроком — кроком дискретизації.

На рисунку 5.8 наведений вигляд коротко-часової енергії шаблону та сигналу, що були отримані вказаним способом.

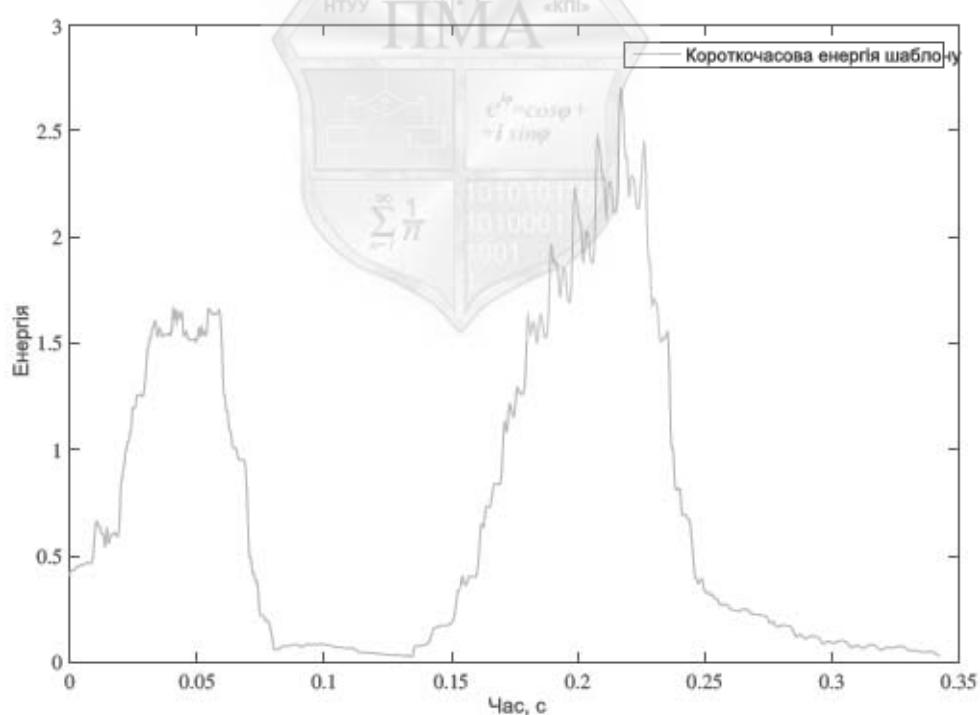
На рисунках 5.9, 5.10 та 5.11 зображено результати пошуку шаблону в аудіосигналі.

З цих результатів видно, що усі з розглянутих методів мали дуже багато хибних спрацювань, а отже використання прямокутного вікна з мінімальним кроком не є доцільним при підрахунку коротко-часової енергії.



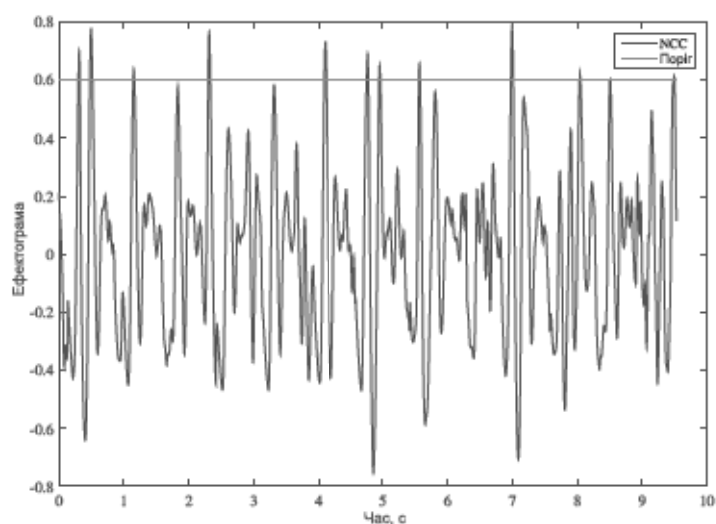


(а) Сигнал

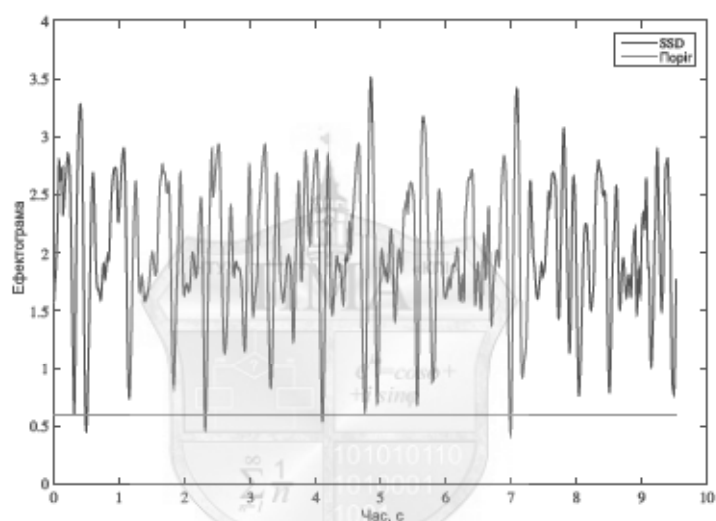


(б) Зразок

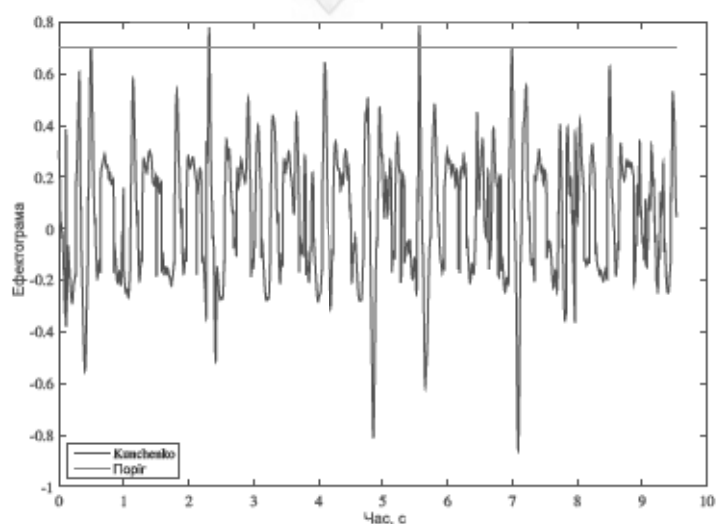
Рисунок 5.8 – Короткочасна енергія запису мовлення й шаблону для пошуку використовуючи прямокутне вікно та мінімальний крок



(a) NCC

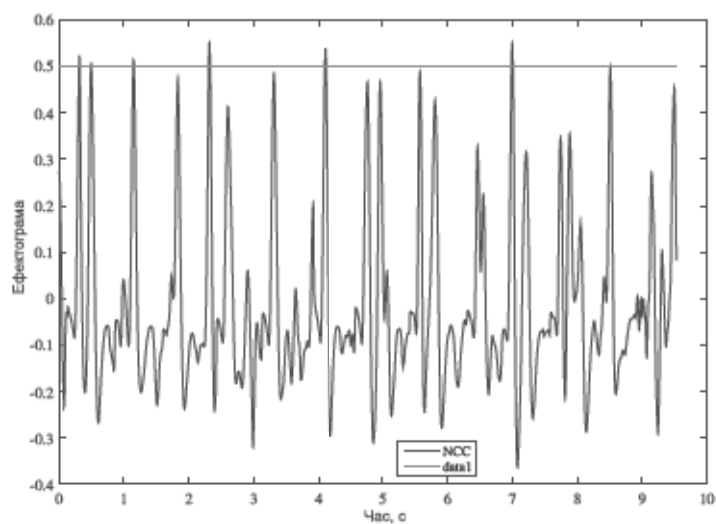


(б) NSSD

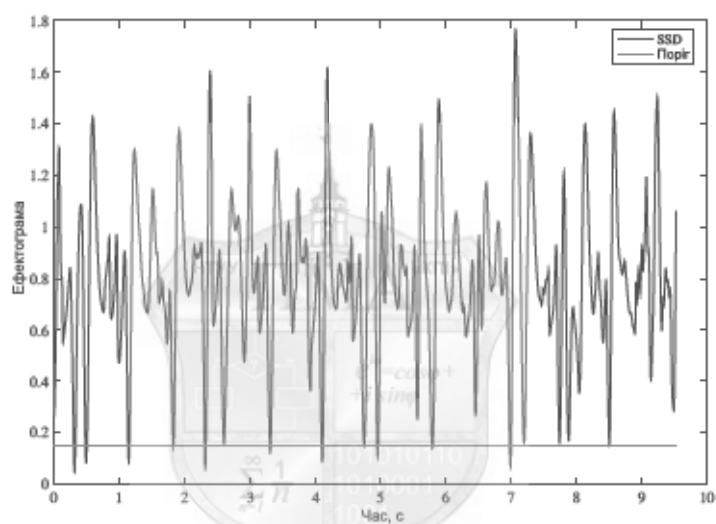


(в) Kunchenko

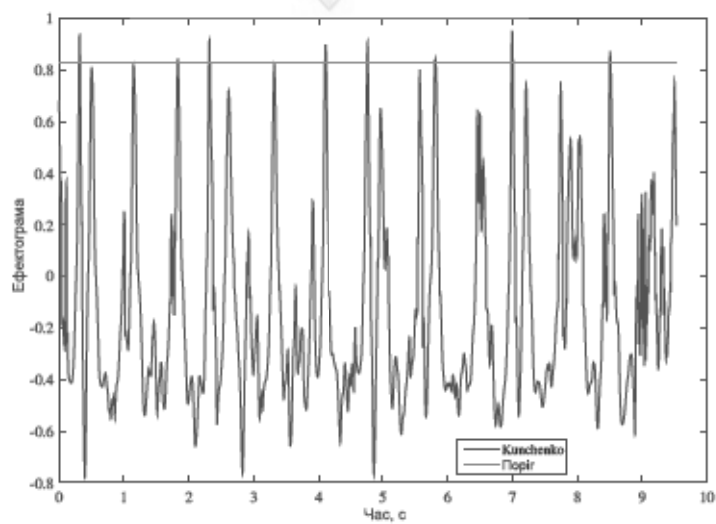
Рисунок 5.9 – Знайдений шаблон в енергії, наведеної на рисунку 5.8, використовуючи прямокутні вікна



(a) NCC

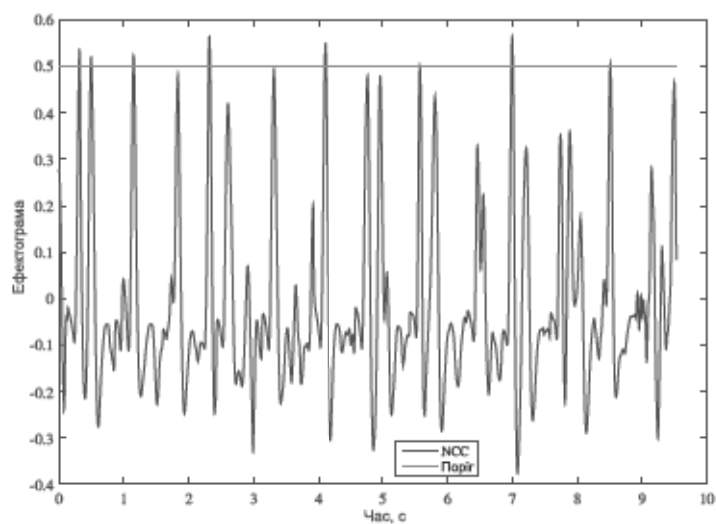


(б) NSSD

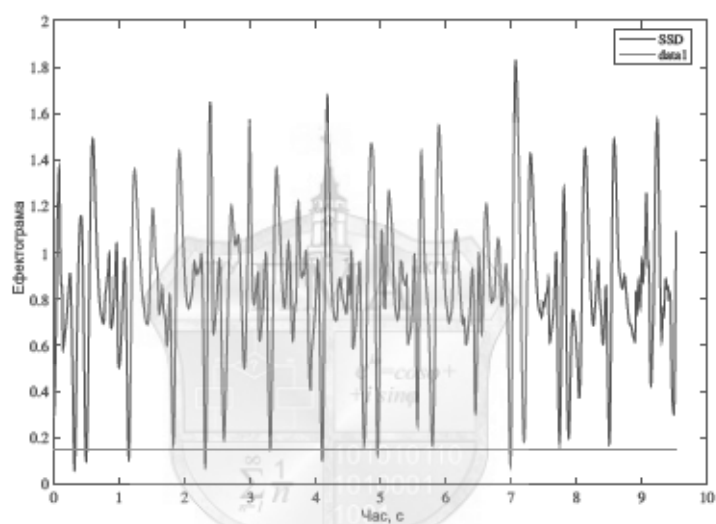


(в) Kunchenko

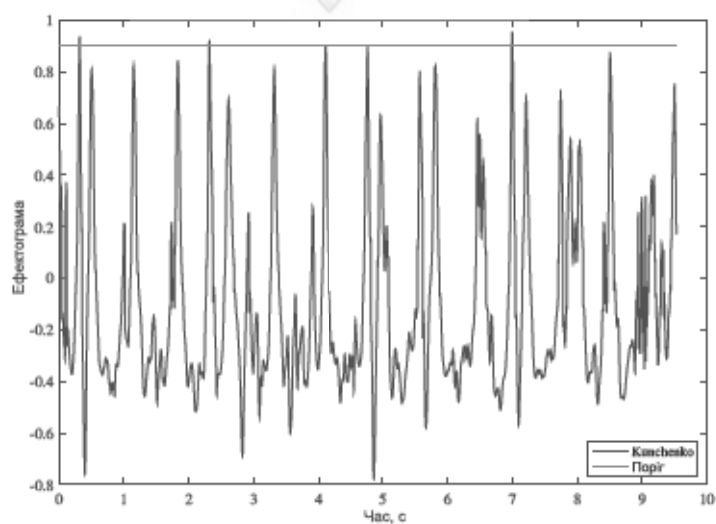
Рисунок 5.10 – Знайдений шаблон в енергії, наведеної на рисунку 5.8, використовуючи вікна Гана



(a) NCC



(б) NSSD



(в) Kunchenko

Рисунок 5.11 – Знайдений шаблон в енергії, наведеної на рисунку 5.8, використовуючи вікна Геммінга

Тепер будемо брати прямокутні вікна з кроком, що дорівнює половині вікна.

На рисунку 5.12 наведений вигляд коротко-часової енергії шаблону та сигналу, що були отримані вказаним способом. Як видно з цього рисунку, отримані коротко-часові енергії мають значно менше значень, а тому мають менш деталізований вигляд.

На рисунках 5.13 та 5.14 зображено результати пошуку шаблону в аудіосигналі.

З цих результатів видно, що використання великого інтервалу між вікнами дозволяє отримувати дуже схожі результати, що й при використанні менших проміжків, але значно швидше (адже сигнал та шаблон складається зі значно меншої кількості значень). Проте, так само, як і в попередньому випадку, результати пошуку шаблонів в коротко-часовій енергії, що була отримана з використанням прямокутних вікон мають невисоку точність.



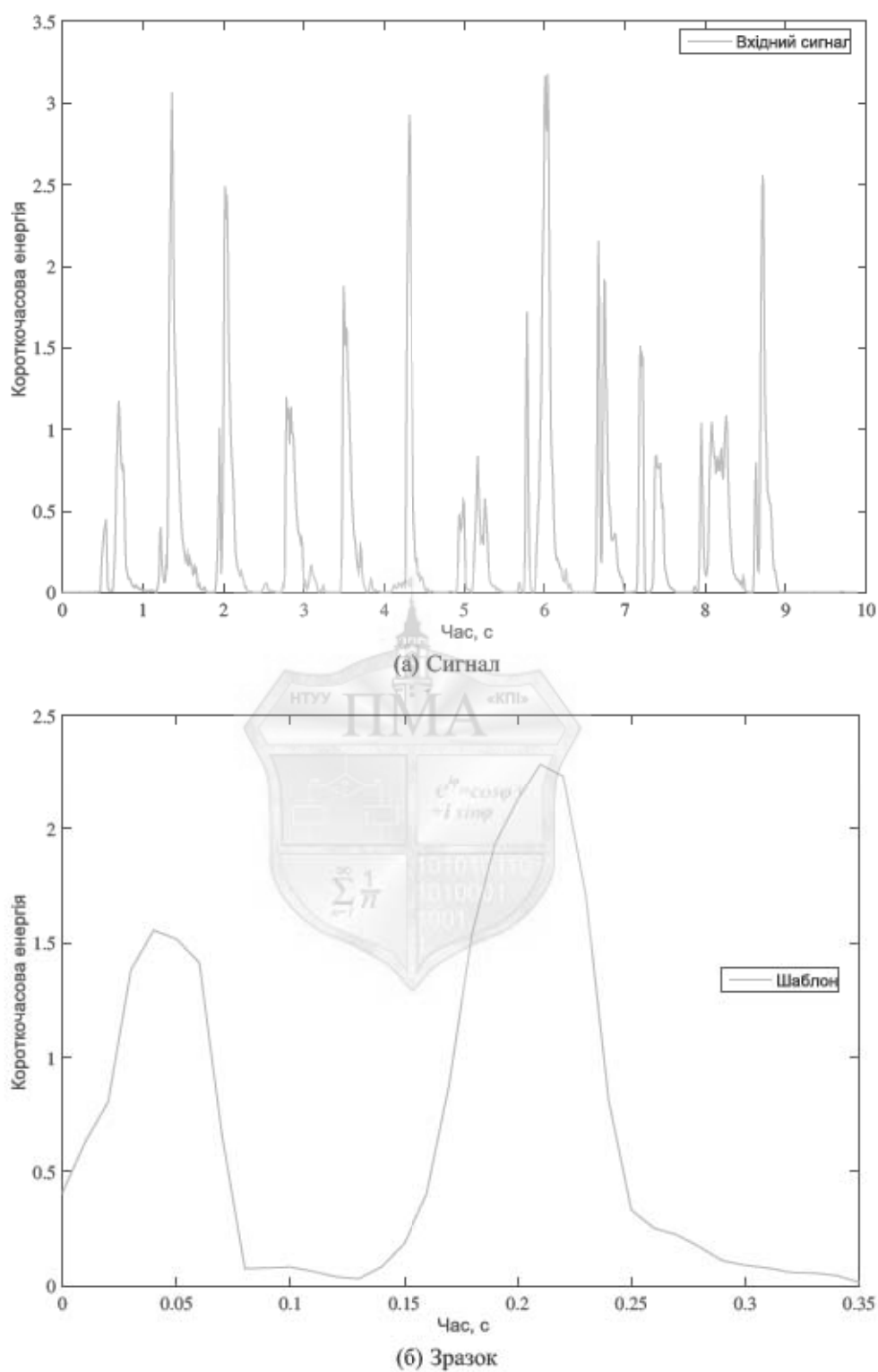
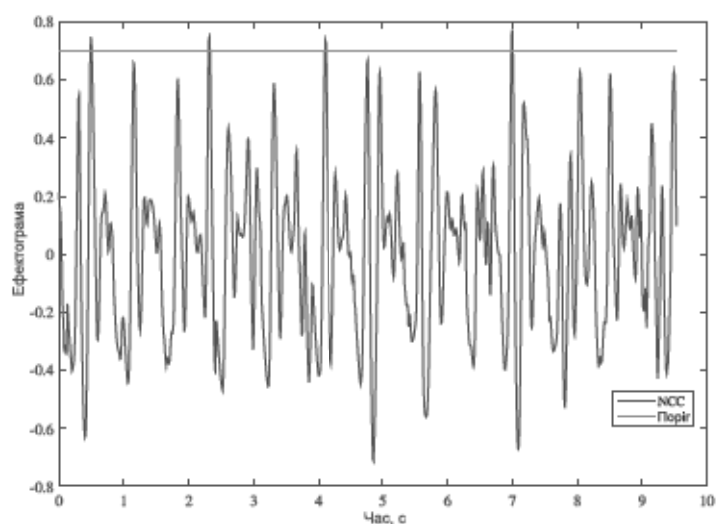
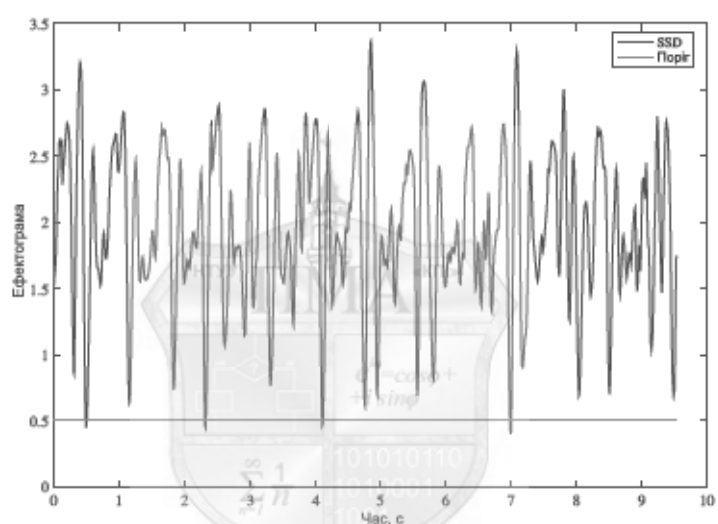


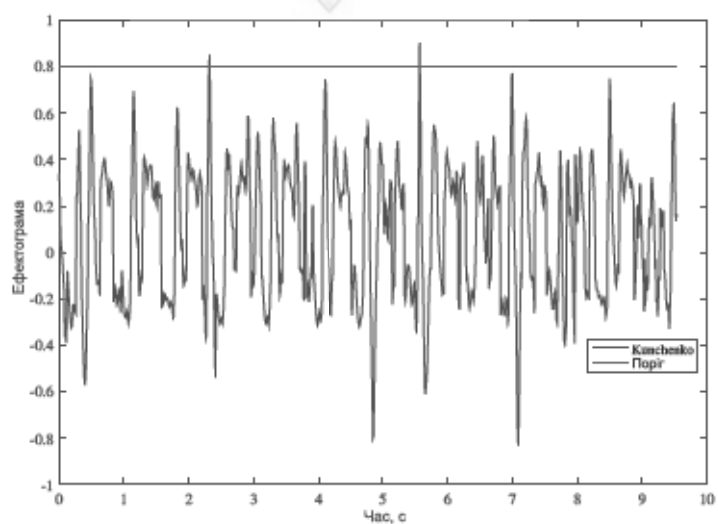
Рисунок 5.12 – Короткочасна енергія запису мовлення й шаблону для пошуку, використовуючи прямокутне вікно та крок середнього розміру



(a) NCC

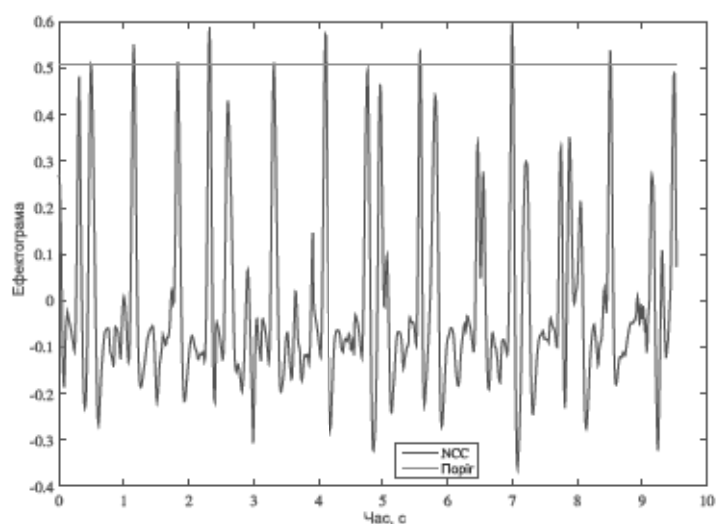


(б) NSSD

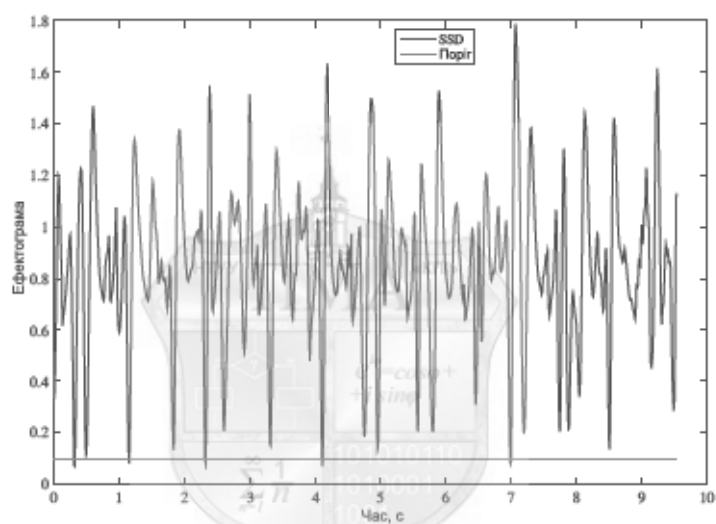


(в) Kunchenko

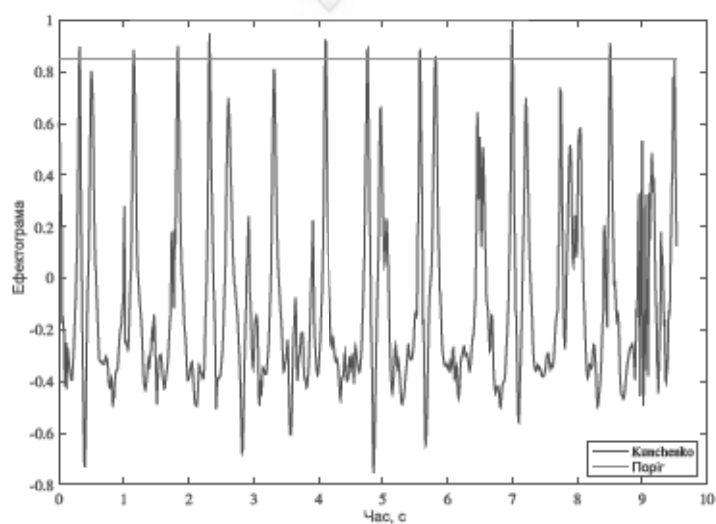
Рисунок 5.13 – Знайдений шаблон в енергії, наведеної на рисунку 5.12, використовуючи прямокутні вікна



(a) NCC



(б) NSSD



(в) Kunchenko

Рисунок 5.14 – Знайдений шаблон в енергії, наведеної на рисунку 5.12, використовуючи вікна Геммінга

Розглянемо поведінку розглянутих алгоритмів для пошуку шаблонів в коротко-часовій енергії сигналу, що була отримана з використанням віконної функції Гана. Такі вікна бралися з кроком, що дорівнює кроку дискретизації, тобто з мінімально можливим кроком.

На рисунку 5.15 наведений вигляд коротко-часової енергії шаблону та сигналу, що були отримані вказаним способом. Як видно з цього рисунку, отримані коротко-часові енергії мають більш гладку форму, ніж коротко-часова енергія, що була отримана з використанням прямокутного вікна.

На рисунках 5.16 та 5.17 зображено результати пошуку шаблону в аудіосигналі.

З цих результатів видно, що використання вікна Гана для обчислення коротко-часової енергії призвело до того, що кількість хибних спрацювань алгоритмів значно зменшилась, але тим не менше алгоритми мають невелику точність.



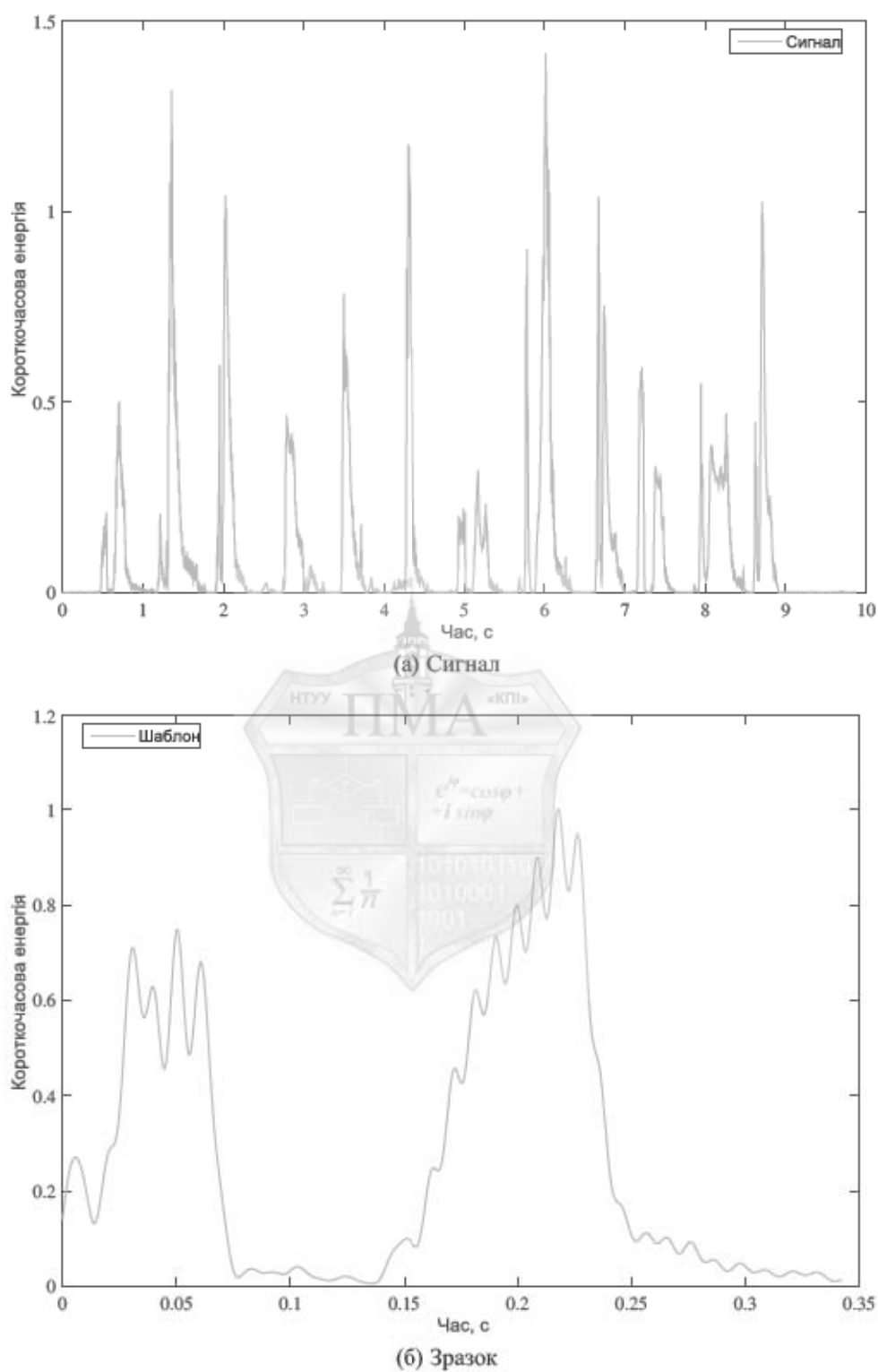
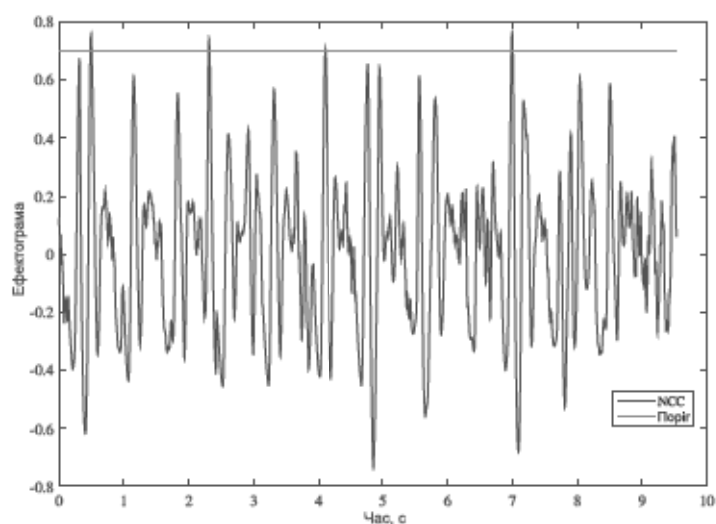
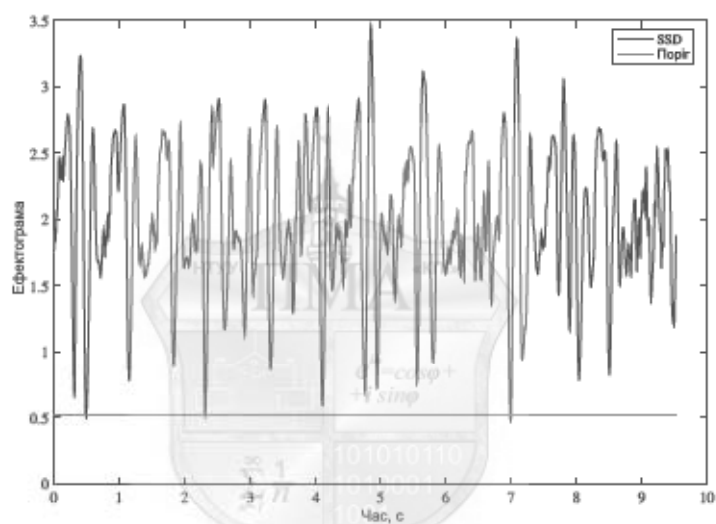


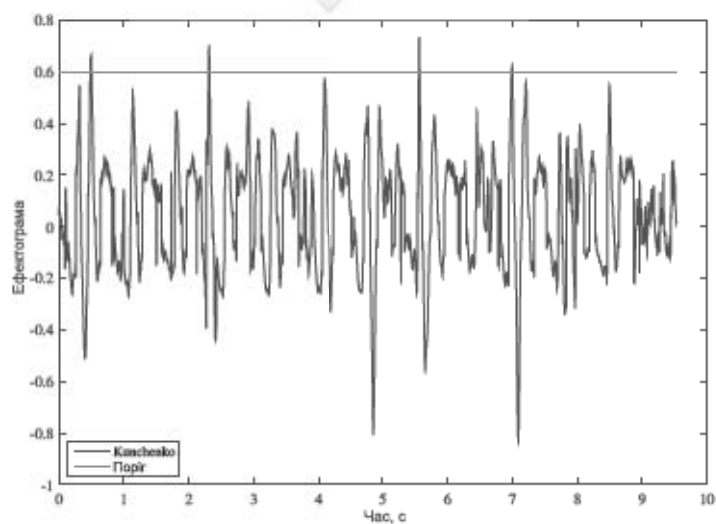
Рисунок 5.15 – Короткочасна енергія запису мовлення й шаблону для пошуку, використовуючи вікно Гана та крок мінімального розміру



(a) NCC

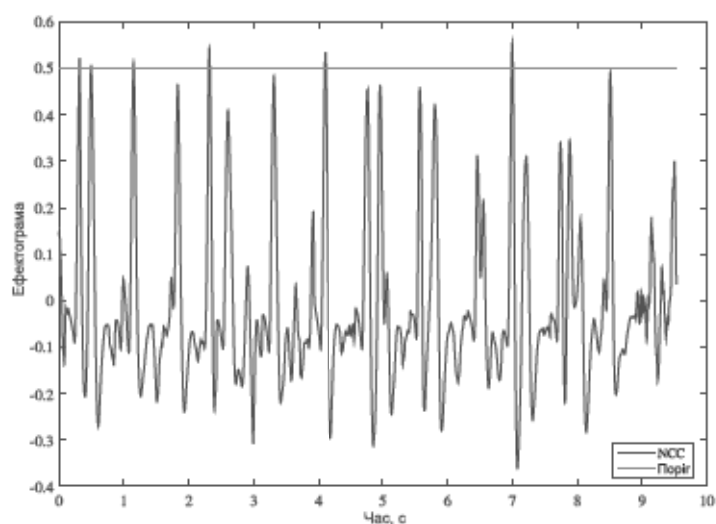


(б) NSSD

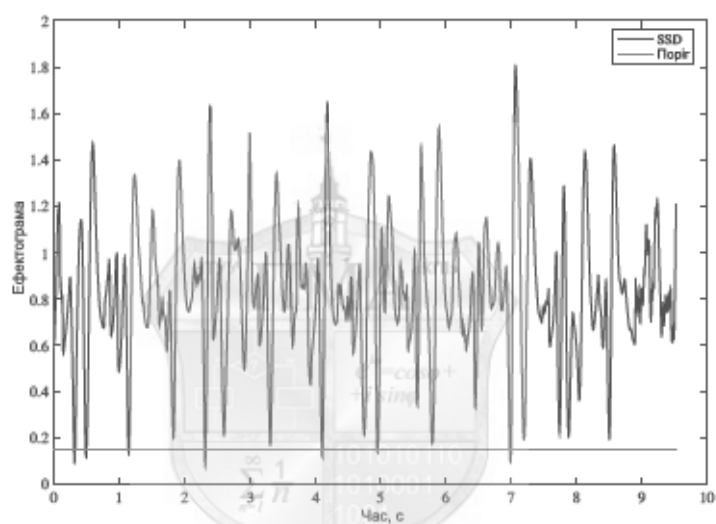


(в) Kunchenko

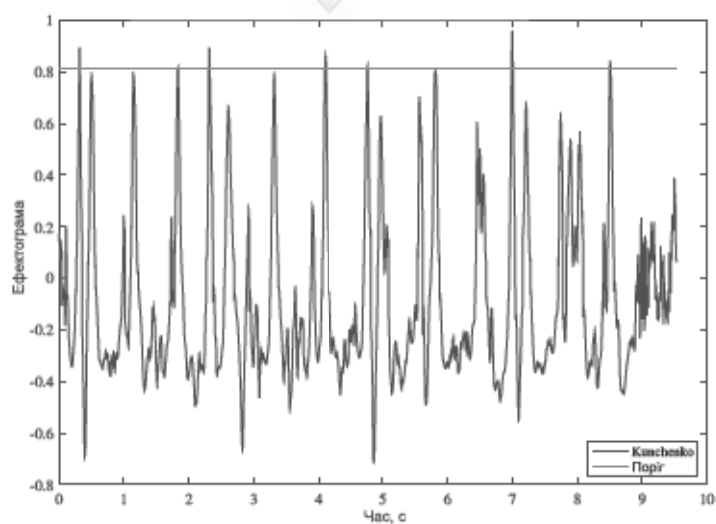
Рисунок 5.16 – Знайдений шаблон в енергії, наведеної на рисунку 5.15, використовуючи прямокутні вікна



(a) NCC



(б) NSSD



(в) Kunchenko

Рисунок 5.17 – Знайдений шаблон в енергії, наведеної на рисунку 5.15, використовуючи вікна Геммінга

Розглянемо поведінку розглянутих алгоритмів для пошуку шаблонів в коротко-часовій енергії сигналу, що була отримана з використанням віконної функції Геммінга. Такі вікна бралися з кроком, що дорівнює кроку дискретизації, тобто з мінімально можливим кроком.

На рисунку 5.18 наведений вигляд коротко-часової енергії шаблону та сигналу, що були отримані вказаним способом. По відношенню до коротко-часової енергії, що була сгенерована з використанням вікна Гана, отримана енергія має менш різкі перепади на графіку.

На рисунках 5.19 та 5.17 зображено результати пошуку шаблону в аудіосигналі.

З наведених графіків ефектограм видно, що для коротко-часової енергії, що була згенерована саме таким чином, метод поліномів Кунченка знайшов усі входження шаблону до сигналу. Також необхідно зазначити той факт, що використання в методі ковзного вікна вікон, відмінних від прямокутного, на відміну від попередніх ситуацій, не призводило до зменшення амплітуди ефектограми для хибних спрацювань.



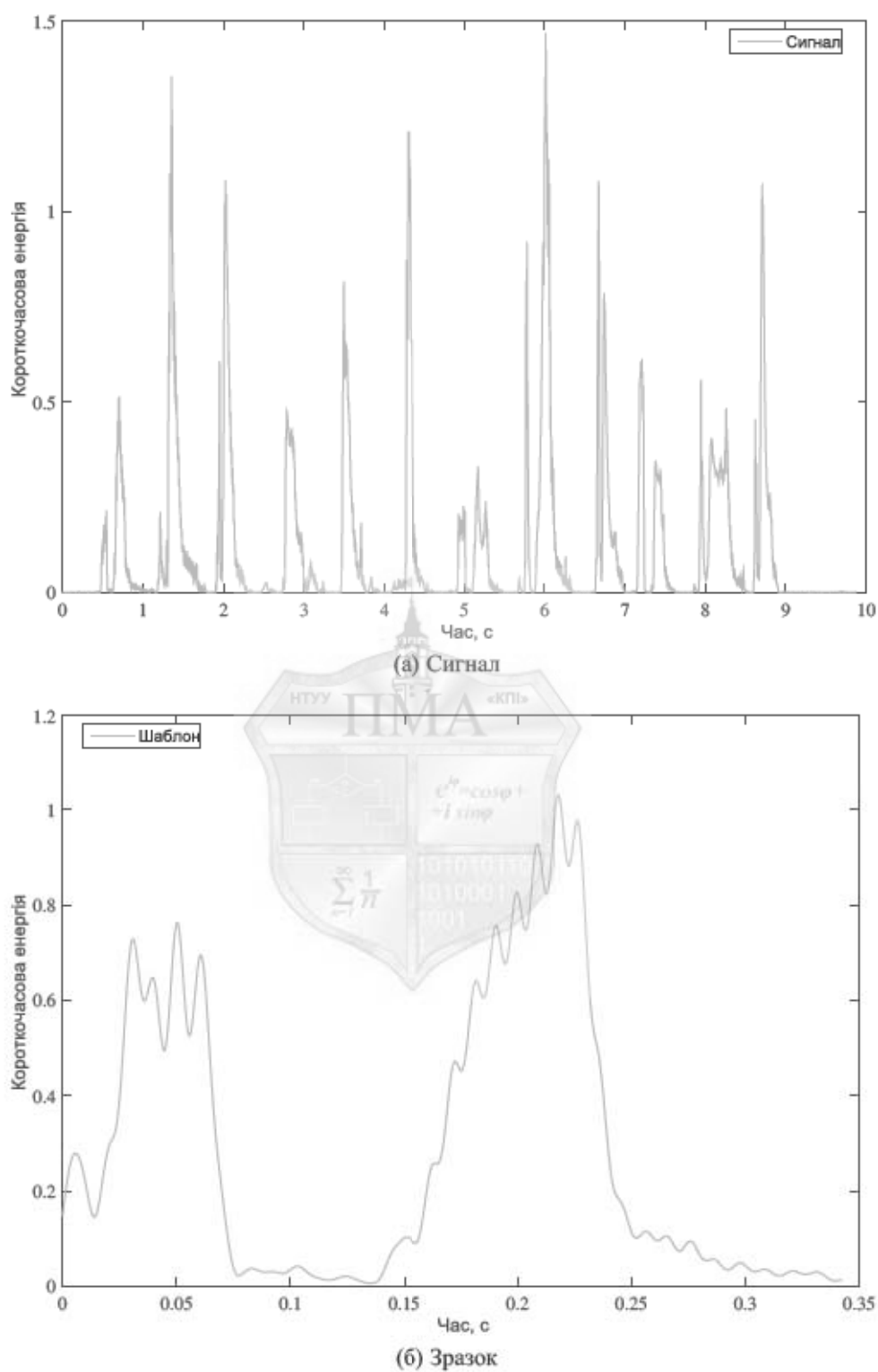
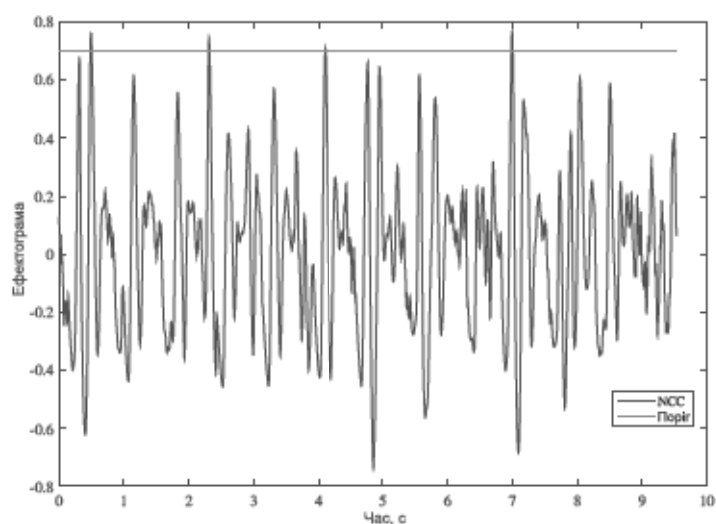
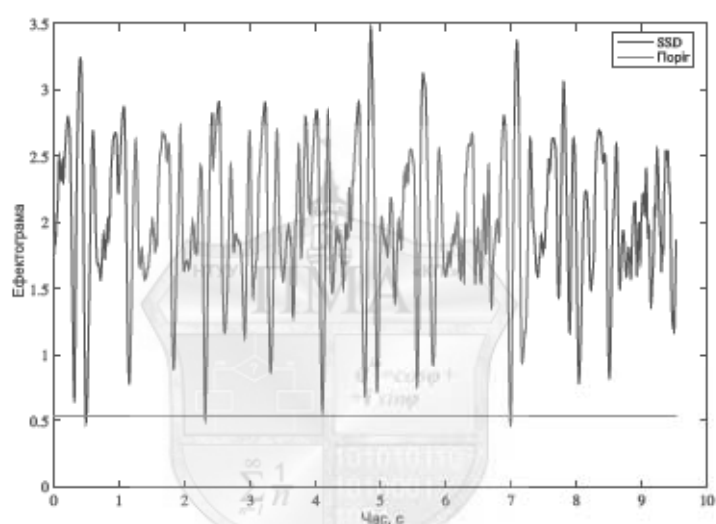


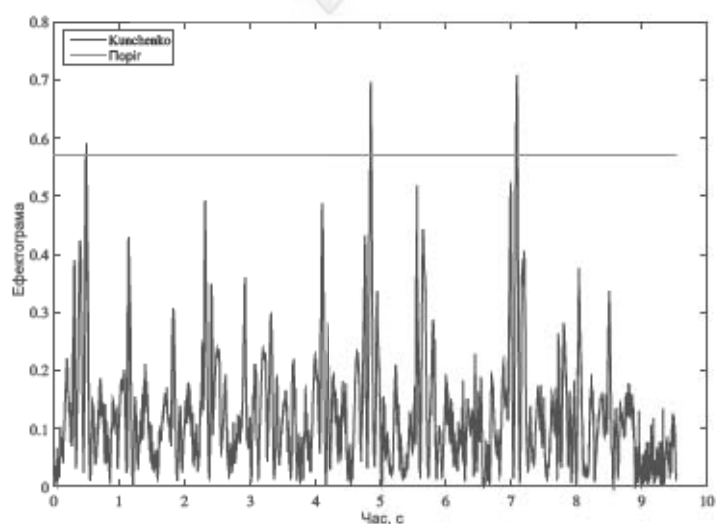
Рисунок 5.18 – Короткочасна енергія запису мовлення й шаблону для пошуку, використовуючи вікно Геммінга та крок мінімального розміру



(a) NCC

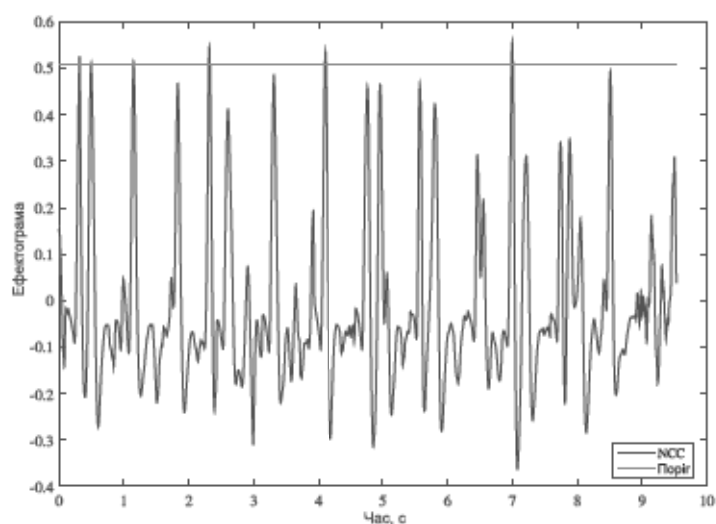


(б) NSSD

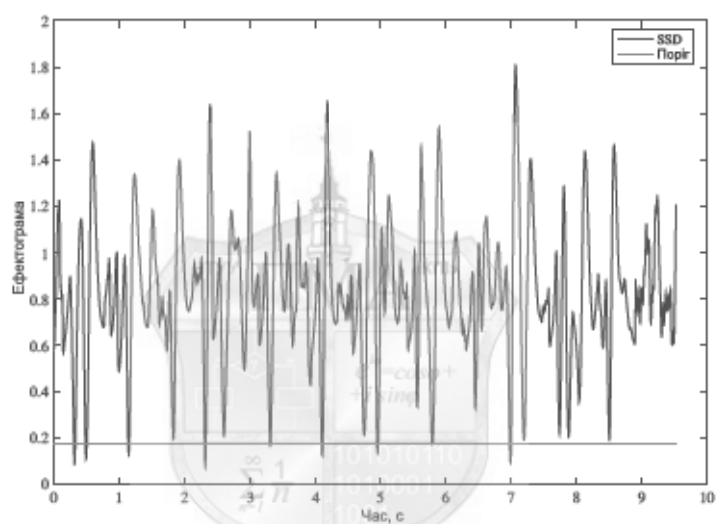


(в) Kunchenko

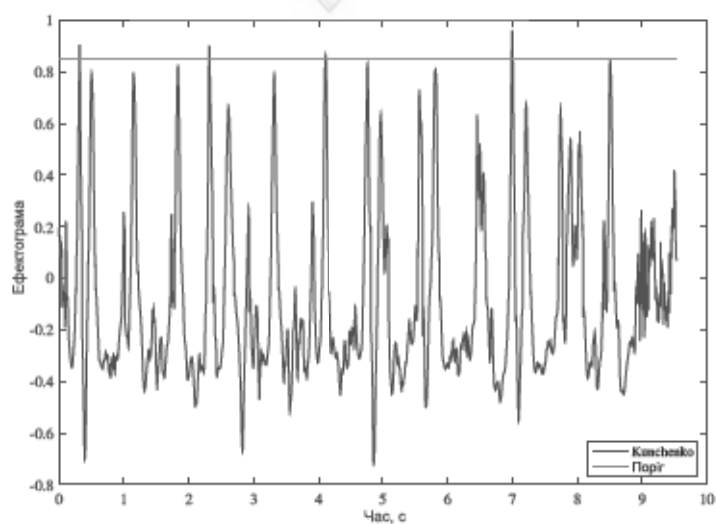
Рисунок 5.19 – Знайдений шаблон в енергії, наведеної на рисунку 5.18, використовуючи прямокутні вікна



(a) NCC



(б) NSSD



(в) Kunchenko

Рисунок 5.20 – Знайдений шаблон в енергії, наведеної на рисунку 5.18, використовуючи вікна Геммінга

5.3 Висновки

В цьому розділі був описаний статистичний експеримент для дослідження поведінки методів пошуку шаблонів, розглянутих в попередніх розділах, як на штучному сигналу, який був згенерований аналітично, так і на аудіозаписах мовлення.

В результаті тестування було виявлено, що:

- метод поліномів Кунченка дозволяє отримати з точністю, не гіршою за розглянути найпоширеніші методи, місцезнаходження шаблону в сигналі, навіть якщо шаблон зазнав змін;
- в залежності від типу сигналу, можливо збільшити точність пошуку завдяки використанню віконних функцій у методі ковзного вікна;
- завдяки використанню коротко-часової енергії метод поліномів Кунченка дозволяє з високою точністю знаходити шаблон в аудіосигналі, що містить у собі запис мовлення.



ВИСНОВКИ

В магістерській дисертації була розглянута задача пошуку шаблонів у цифрових сигналах. В ході дослідження були проаналізовані існуючі методи для пошуку шаблонів, зіставлені обмеження, що на них накладаються.

Було з'ясовано, що у випадку внесення значних змін в шаблон класичні методи (кросс-кореляція, сума квадратів відстаней тощо) суттєво втрачають свої розпізнавальні можливості. Тому для пошуку шаблонів був вибраний метод, що базується на розкладанні функції в базисі лінійного функціонального простору Кунченка, який в таких випадках працює значно краще.

В дисертації були запропоновані шляхи покращення роботи даного методу. Зокрема, застосування пірамідального пошуку дозволило значно прискорити роботу методу, оскільки такий пошук дозволяє відфільтрувати значну частину сигналу як таку, що не містить шаблону, за досить короткий проміжок часу. Використання методу прискореного ковзання також дозволяє значно прискорити роботу алгоритму.

Був розроблений та оптимізований програмний комплекс в математичному середовищі MathWorks MATLAB 2014b.

Під час статистичного експерименту було порівняно ефективність роботи запропонованого методу з існуючими методами як на штучно згенерованому сигналі, так і для аудіозапису мовлення. Як результат цього тестування можна зазначити, що запропонований метод дозволяє отримати аналогічну ступінь розпізнавання на штучно згенерованих сигналах і значно вищу якість на аудіозаписах мовлення.

Для пошуку шаблону в аудіозаписі з мовленням було використано визначення коротко-часової енергії й віконних функцій. Було проаналізовано поведінку алгоритму в залежності від обраних параметрів методу та вибрано оптимальний набір — віконна функція Гемінга з розміром вікна в 20 мілісекунд та кроком між вікнами в 1,25 мілісекунди.

Як перспективні напрями подальшого розвитку запропонованого методу можна зазначити такі: дослідження методів інтелектуального вибору параметрів алгоритму в залежності від характеристик сигналу, апаратне прискорення для більш швидкої роботи алгоритму та апробація методу на даних з іншої предметної області.



ПЕРЕЛІК ПОСИЛАНЬ

1. Brunelli R. Template matching techniques in computer vision: theory and practice. — Chippenham: Wiley, 2009. — 348 p.
2. Lewis J.P. Fast Template Matching. — *Vision Interface*, 1995. — P. 120–123.
3. Fredrikssona K., Navarrob G., Ukkonenc E. Sequential and indexed two-dimensional combinatorial template matching allowing rotations — Amsterdam: Elsevier, 2005. — P. 239–275.
4. Harris F.J. On the use of windows for harmonic analysis with the discrete Fourier transform. — *Proceedings of the IEEE* 66, 1978. — 51 p.
5. Oppenheim A.V., Schaffer R.W., Buck J.R. Discrete-Time Signal Processing. — Upper Saddle River, NJ: Prentice Hall, 1999. — 468 p.
6. Enochson L.D., Otnes R.K. Programming and Analysis for Digital Time Series Data. — U.S. Dept. of Defense, Shock and Vibration Info. Center, 1968. — 142 p.
7. Nuttall A.H. Some windows with very good sidelobe behavior. — *IEEE Transactions on Acoustics, Speech, Signal Processing*, 1981. — P. 84–91.
8. Turkowski K., Gabriel S. Filters for Common Resampling Tasks. — *Graphics Gems I*. Academic Press, 1990. — P. 147–165.
9. Кунченко Ю.П. Полиномы приближения в пространстве с порождающим элементом. — К.: Наук. думка, 2003. — 243 с.
10. Chertov O., Slipets T. Kunchenko's Polynomials for Template Matching. — Sarajevo: IEEE, 2013. — P. 47–49.
11. Vanderbrug G.J., Rosenfeld A. Two-stage Template Matching. — *IEEE Transactions on Computers*, C-26, 1977. — P. 102–103.
12. Чертов О.Р. Метод прискорення пошуку шаблонів за допомогою поліномів Кунченка / Чертов О.Р., Погода М.В. // Прикладна Математика та Комп'ютинг. ПМК, 2015: сьома наук. конф. магістрантів та аспірантів, Київ, 15–17 квіт. 2015 р.: зб. тез. доп. / [редкол.: Дичка І.А. та ін.]. — К.:

Просвіта, 2015. — с. 241–246

13. Rosenfeld A., Vanderbrug G.J. Coarse-Fine Template Matching. — IEEE Transactions on Systems, Man and Cybernetics, 7, 1977. — P. 104–107.

14. Theodoridis S., Koutroumbas K. Pattern Recognition, 4th edition. — Academic Press, 2009. — 254 p.

15. Theodoridis S. Introduction to pattern recognition: a MATLAB[®] approach. — Academic Press, 2010. — 214 p.

