

УДК 654:679.76 (078)

Tatiana M. Tkachenko¹, D.S., Professor
ORCID ID 0000-0003-2105-5951 *e-mail*: tkachenkoknuba@gmail.com

Yulia H. Pilkevich¹, assistant of the department
ORCID ID 0000-0002-7618-0250 *e-mail*: y.rozorinova@gmail.com

Heorhii M. Rozorinov², D.S., Professor
ORCID ID 0000-0002-6095-7539 *e-mail*: grozoryn@gmail.com

¹ Kyiv National University of Construction and Architecture, Kyiv, Ukraine

² National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", Kyiv, Ukraine

RECOGNITION OF FISH ACOUSTIC PATTERNS AT MONITORING OF FRESHWATER ECOSYSTEMS

***Abstract.** The basic sources of contamination and obstruction of reservoirs are cleared not enough sewer water of industrial and communal enterprises, large stock-raising complexes, wastes of production; upcast of water and railway transport; wastes of roughing-out of flax, pesticides and other. Contaminents, getting in natural reservoirs, result in the quality changes of water, that, mainly, appear in the change of physical properties of water, in the change of chemical composition of water, in a presence floating substances on the surface of water and laying of them on the bottom of reservoirs. The increases of population, expansion of old and origin of new cities considerably increased entering of domestic flows internal reservoirs. Synthetic cleansers that is widely used in the way of life contaminate reservoirs in a yet greater degree. In the total the capacity of waters goes down for oxigenating, activity of bacteria that mineralize organic substances is paralysed. The unfavorable ecological state of many freshwater ecosystems inflicts substantial harm to the fish resources of reservoirs and puts under a threat possibility not only to develop fish industry, conducting fish artificially, but also simply to catch her. All of it stimulate to do events in relation to the improvement of the ecological state of fresh reservoirs. Voice vibrations are the important constituent of the ecological monitoring of the biota state of fresh reservoirs. Information is about formation of sound in a reservoir part of that is activity of fishes turns out by means of acoustic sensors, that farther yields to computer treatment. The modern methods of recognition of fish acoustic patterns are based on the standards of signals, with properties of average estimations, or on comparisons of acoustic signals with a standard. It is shown that for creation of standards, as a rule, executed: previous signal processing, extraction of features of acoustic signal. Acoustic signals that act from movable objects – fishes can change depending on objective external terms and physical state of reservoirs. The hard algorithms of recognition of acoustic patterns are characterized high probability of error. In this connection repressing are adaptive algorithms of recognition of acoustic patterns. In the process of forming of standards clarification of software comes true according to the features of acoustic signal. Realization of process of creation of standards allows to determine the measure of functional readiness of parameters and knowledge base for the decision of recognition tasks of acoustic signals. In the process of recognition the probability terms of the correct comparing are set to the standard, on default of that an algorithm stops to be executed and requires additional studies. It requires creation of standards that reflect the*

characteristic features of fish signals. Presently for authentication mostly choose such pattern of acoustic signals, as period length of signal fundamental wave. It can be determined or by the search of maximal value in an autocorrelation function, or by the search of minimum value in the function of mean value of difference of signal amplitudes, or by the search of difference of two maximal values in the sequence of going into detail wavelet-coefficients. It is shown that for the tasks of recognition of fish acoustic patterns, most exact and requiring the least studies there is presentation of acoustic signal as a set of sign vectors of frames. In detail methodologies of the period selection of fundamental wave of acoustic signal were analysed: SIFT, EFT-A and EFT-WT. Methodology of EFT-WT is characterized absence of the thresholds set in good time; by the rapid search of period of fundamental wave; by absence of dependence on a noise-level, as a certain range of frequencies is investigated. At the same time calculable complication of wavelet transform is relatively high, in this connection it is necessary optimization of calculation algorithms.

Keywords: *acoustic pattern; ecosystem; standard of signal; fundamental wave; period; recognition; frame*

Т.М. Ткаченко¹, Ю.Г. Пількевич¹, Г.М. Розорінов²

¹ Київський національний університет будівництва і архітектури, м. Київ, Україна

² Національний технічний університет України "Київський політехнічний інститут імені Ігоря Сікорського", м. Київ, Україна

РОЗПІЗНАВАННЯ АКУСТИЧНИХ ОБРАЗІВ РИБ ПРИ МОНІТОРИНГУ ПРІСНОВОДНИХ ЕКОСИСТЕМ

Анотація. *Основними джерелами забруднення і засмічення водойм є недостатньо очищені стічні води промислових і комунальних підприємств, великих тваринницьких комплексів, відходи виробництва; скидання водного і залізничного транспорту; відходи первинної обробки льону, пестициди та ін. Забруднюючі речовини, потрапляючи в природні водойми, призводять до якісних змін води, які, в основному, виявляються в зміні фізичних властивостей води, у зміні хімічного складу води, в наявності плаваючих речовин на поверхні води і відкладанні їх на дні водойм. Зростання населення, розширення старих і виникнення нових міст значно збільшили надходження побутових стоків у внутрішні водойми. Ще більшою мірою забруднюють водойми миючі синтетичні засоби, що широко використовуються в побуті. У результаті знижується здатність вод до насичення киснем, паралізується діяльність бактерій, що мінералізують органічні речовини. Неприятливий екологічний стан багатьох прісноводних екосистем завдає суттєвої шкоди рибним ресурсам водойм і ставить під загрозу можливість не тільки розвивати рибну галузь, розводячи рибу штучно, але і просто її виловлювати. Все це змушує вживати заходів щодо поліпшення екологічного стану прісних водойм. Звукові коливання є важливою складовою екологічного моніторингу стану біоти прісних водойм. Інформація про звукоутворення у водоймі, частиною якої є діяльність риб, отримується за допомогою акустичних датчиків і далі піддається комп'ютерній обробці. Сучасні методи розпізнавання акустичних образів риб базуються або на еталонах сигналів, з властивостями середньостатистичних оцінок, або на зіставленнях акустичних сигналів з еталоном. Показано, що для створення еталонів, як правило, виконується: попередня обробка сигналу, виділення ознак акустичного сигналу. Акустичні сигнали, що надходять від рухливих об'єктів – риб, можуть змінюватися залежно від об'єктивних зовнішніх умов*

і фізичного стану водойм. Жорсткі алгоритми розпізнавання акустичних образів характеризуються високою вірогідністю помилки. У зв'язку з цим кращими є адаптивні алгоритми розпізнавання акустичних образів. В процесі формування еталонів здійснюється уточнення програмного забезпечення згідно з особливостями акустичного сигналу. Реалізація процесу створення еталонів дозволяє визначати міру функціональної готовності параметрів і бази знань для вирішення завдань розпізнавання акустичних сигналів. В процесі розпізнавання задаються умови вірогідності правильного порівняння з еталоном, при невиконанні яких алгоритм перестав виконуватися і потребує додаткового навчання. Це вимагає створення еталонів, що відображають характерні особливості сигналів риб. Нині для ідентифікації найчастіше вибирають таку ознаку акустичних сигналів, як довжина періоду основної гармоніки сигналу. Її можна визначати або шляхом пошуку максимального значення в автокореляційній функції, або шляхом пошуку мінімального значення у функції середнього значення різниці амплітуд сигналу, або шляхом пошуку різниці двох максимальних значень в послідовності деталізуючих вейвлет-коефіцієнтів. Показано, що для завдань розпізнавання акустичних образів риб найбільш точним і вимагаючим найменшого навчання є представлення акустичного сигналу у вигляді набору векторів ознак фреймів. Детально було проаналізовано методики виділення періоду основної гармоніки акустичного сигналу: SIFT, BOT-A і BOT-BP. Методика BOT-BP характеризується відсутністю заздалегідь заданих порогів; швидким пошуком періоду основної гармоніки; відсутністю залежності від рівня шуму, оскільки досліджується конкретний діапазон частот. В той же час обчислювальна складність вейвлет-перетворення відносно висока, у зв'язку з чим потрібна оптимізація алгоритму обчислень.

Ключові слова: акустичний образ; екосистема; еталон сигналу; основна гармоніка; період; розпізнавання; фрейм

DOI: 10.32347/2411-4049.2021.1.20-34

Вступ

Несприятливий екологічний стан багатьох прісноводних екосистем завдає суттєвої шкоди рибним ресурсам водойм і ставить під загрозу можливість не тільки розвивати рибну галузь, розводячи рибу штучно, але і просто її виловлювати. Незаперечний факт, що в багатьох рибних водоймах кількість водних гідробіонтів скоротилася, а деякі види зникли зовсім. Все це змушує вживати заходів щодо поліпшення екологічного стану прісних водойм [1].

Відомо, що звукові коливання є важливою складовою екологічного моніторингу стану біоти прісних водойм. Інформація про звукоутворення у водоймі отримується за допомогою акустичних датчиків і далі піддається комп'ютерній обробці.

Сучасні методи розпізнавання акустичних образів риб базуються або на еталонах сигналів, що формуються в процесі попереднього дослідження, з властивостями середньостатистичних оцінок, або на зіставленнях акустичних сигналів з еталоном.

Для створення еталонів, як правило, виконується:

- попередня обробка сигналу (перетворення способу представлення інформації в акустичному сигналі, пригнічення шуму, скорочення об'єму інформації) [2–8];
- виділення ознак акустичного сигналу (наприклад, визначення меж інформаційної частини сигналу) [9–12].

Метою роботи є класифікація і порівняння методів розпізнавання акустичних образів риб прісних водойм, визначення їх достоїнств і недоліків.

Узагальнений алгоритм розпізнавання акустичних образів

Акустичні сигнали, що надходять від рухливих об'єктів – риб, можуть змінюватися залежно від об'єктивних зовнішніх умов і фізичного стану водойм. Жорсткі алгоритми розпізнавання акустичних образів характеризуються високою вірогідністю помилки. У зв'язку з цим кращими є адаптивні алгоритми розпізнавання акустичних образів.

Узагальнений алгоритм розпізнавання акустичних образів включає три основні операції, які можуть виконуватися або в режимі розпізнавання, або в режимі навчання, і взаємозв'язані між собою [12–16] (рис. 1).

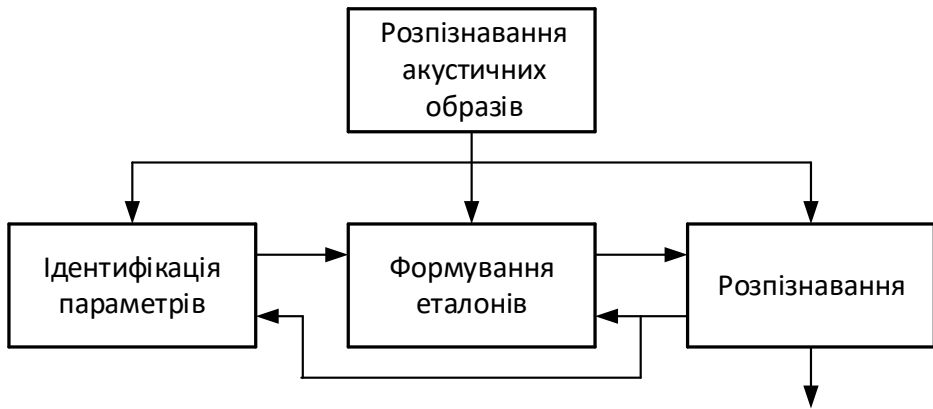


Рис. 1 – Узагальнений алгоритм розпізнавання акустичних образів

В процесі ідентифікації параметрів визначаються умови підготовки і налаштування параметрів, використовуваних при формуванні еталонів і порівнянні з ними. В процесі формування еталонів здійснюється персоніфікація програмного забезпечення згідно з особливостями акустичного сигналу. Реалізація перших двох процесів дозволяє визначати міру функціональної готовності параметрів і бази знань для вирішення завдань розпізнавання акустичних сигналів. В процесі розпізнавання задаються умови вірогідності правильного порівняння з еталоном, при невиконанні яких алгоритм перестав виконуватися і вимагає додаткового навчання, тобто перемикається в режим формування еталонів сигналів. Операція ідентифікації параметрів виконується при припиненні виконання двох інших операцій в тих випадках, коли міняється досліджуваний об'єкт або міняються технічні засоби і стандартне програмне забезпечення.

Формування еталонів

Алгоритм розпізнавання акустичного сигналу ґрунтується на базі цих еталонів. Це вимагає створення еталонів, що відображають характерні особливості сигналів риб.

На рис. 2 схематично показана операція формування еталонів для розпізнавання акустичних образів [12–16].

При очищенні від шуму акустичного сигналу використовуються методичні правила очищення, що враховують чисельні характеристики сигналу.

При визначенні меж корисної частини акустичного сигналу використовуються методичні правила знаходження меж інформаційної частини сигналу, що враховують чисельні характеристики сигналу.

При створенні еталонів використовуються методичні правила утворення словника еталонів сигналів, акустичні характеристики яких визначаються методами цифрової обробки сигналів.

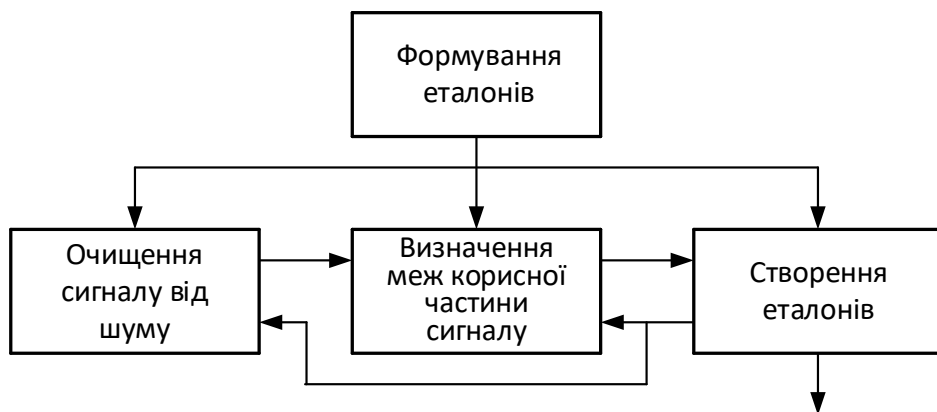


Рис. 2 – Операція формування еталонів

Розпізнавання

Розпізнавання акустичних сигналів базується на їх порівнянні з раніше створеними еталонами. На рис. 3 схематично показана операція розпізнавання акустичних образів.

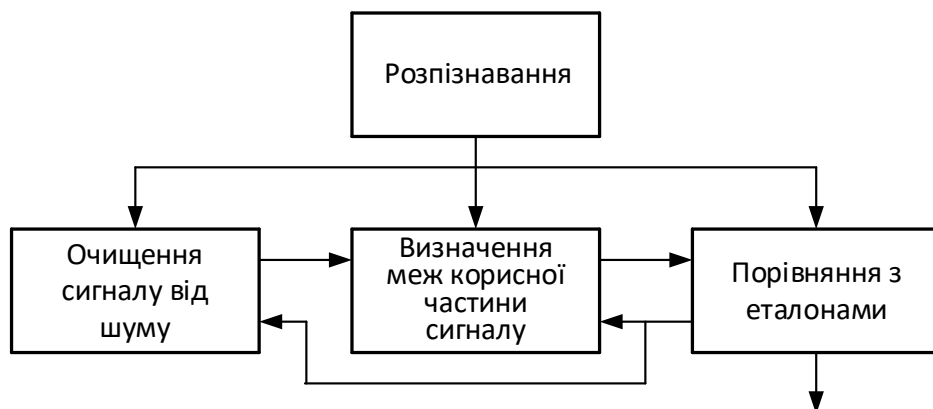


Рис. 3 – Операція розпізнавання акустичних сигналів

Спочатку використовуються методичні правила очищення сигналу від шуму, що враховують чисельні характеристики сигналу.

Далі використовуються методичні правила визначення меж корисної частини сигналу, що враховують чисельні характеристики сигналу.

На кінцевому етапі використовуються методичні правила порівняння розпізнаваних сигналів з еталонами, акустичні характеристики яких визначаються методами цифрової обробки.

Відповідно до структури систем розпізнавання акустичних образів необхідно розглянути методичні положення, на яких базуються блоки формування еталонів і розпізнавання цих систем.

Ідентифікація акустичних сигналів

Нині для ідентифікації найчастіше вибирають такі ознаки акустичних сигналів:

- енергетичні (відношення енергій двох ділянок сигналів, нормована енергія, енергія смуг спектру) [3, 13];
- основний тон (метод SIFT [3, 13], ґрунтований на пошуку максимуму в автокореляційній функції, ВОТ-метод [14]);
- кількість нуль-перетинів (чи кількість строгих мінімумів) [3];
- параметри лінійного передбачення [3, 6, 13];
- форманти (мовні сигнали) [3, 6, 13];
- кепстр [3, 15, 16, 18];
- кількість імпульсів рівної довжини [19];
- тривалість досліджуваної ділянки сигналу [20].

Для виділення цих ознак зазвичай використовують наступні методи цифрової обробки сигналів: фільтрацію [3, 4]; швидке перетворення Фур'є (ШПФ) [3, 4]; вейвлет-перетворення (ВП) [21–23]; кепстральний аналіз [3, 4]; кодування з лінійним передбаченням [3, 4].

З виділених ознак акустичного сигналу формують вектори ознак фреймів (блоків даних, ділянок рівної довжини).

Формальне представлення еталонів акустичного сигналу

Еталони акустичних сигналів найчастіше представляють у виді: 1) наборів векторів ознак фреймів; 2) кусочно-постійних моделей; 3) наборів векторів кодової книги; 4) станів прихованих марківських моделей (ПММ); 5) станів відповідної нейромережі; 6) гібриду ПММ і нейромережі.

Форма 1 є найпростішою і є еталоном сигналу $E_k = (e_{k1}, \dots, e_{ks}, \dots, e_{kN})$, де e_{ks} – вектор ознак s-го фрейма сигналу.

Форма 2 є пара – еталон сигналу $E_k = (e_{k1}, \dots, e_{ks}, \dots, e_{kQ})$ і темпоральна транскрипція сигналу $\tau_k = ((m_{k1}, M_{k1}), \dots, (m_{ks}, M_{ks}), \dots, (m_{kQ}, M_{kQ}))$, де m_{ks}, M_{ks} – мінімальна і максимальна повторюваність e_{ks} у сигналі, тобто декілька сусідніх фреймів сигналу можуть описуватися одним вектором ознак e_{ks} [14]. Зазвичай такі еталони представляють у вигляді детермінованого кінцевого автомата (рис. 4).

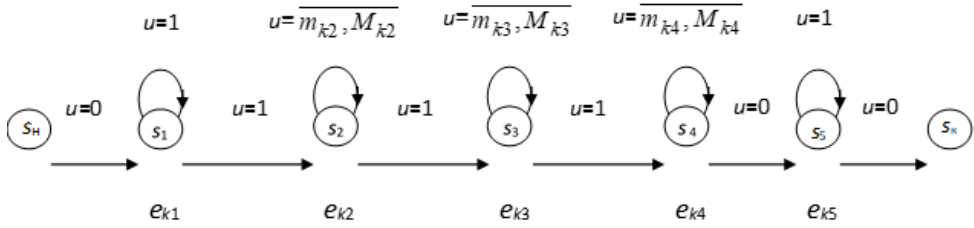


Рис. 4 – Детермінований кінцевий автомат

Ця форма дозволяє економніше представити сигнал, ніж перша [14], але у разі неточного зіставлення фреймів і вектора ознак e_{k_s} може дати меншу точність розпізнавання.

Форма 3 є еталонном сигналу $E_k = (e_{k1}, \dots, e_{k_s}, \dots, e_{kN})$, де e_{k_s} – вектор з кодової книги, замінюючий початковий вектор ознак фрейма сигналу. Ця форма дозволяє економніше представити сигнал, але дає меншу точність розпізнавання, ніж перша [4].

Форма 4 є безперервною, дискретною або напівбезперервну ПММ, що описує сигнал, і подібна до форми 2, що враховує повторюваність фреймів. Ця форма дозволяє економніше представити сигнал, чим перша. У разі використання безперервної ПММ, по точності розпізнавання порівняння з першою, але вимагає великої кількості повчальних даних [4].

Форма 5 є нейромережевою моделлю сигналу у вигляді системи рівнянь. Можливі два випадки:

а) на вхід послідовно подаються вектори ознак фреймів сигналу, на виході – маркер поточного фрейма;

б) на вхід подається набір векторів ознак усього сигналу, на виході – номер розпізаного образу.

У разі а) використовуються статичні мережі: багатошаровий перцептрон (MLP), радіально-базисна мережа (RBF), імовірнісна мережа (PNN), мережа Кохонена (SOM), аналіз головних компонент (PCA), аналіз незалежних компонент (ICA), багатошарова мережа прямого поширення (MLFFNN), машина опорних векторів (SVM), ієрархічне змішування експертів (HME) та ін. У разі б) частіше використовуються динамічні мережі: мережі каскадної кореляції (Cascade Correlation), мережі з затримкою за часом (TDNN) та ін. Ці мережі дають невисоку точність розпізнавання із-за нестационарності та нелінійності акустичного сигналу.

Форма 6 є гібридом ПММ і нейромережевої моделі [24]. Параметри ПММ коригуються за допомогою нейромережі. У режимі розпізнавання поточний фрейм розпізнається за допомогою нейромережі, а результат обробляється в ПММ. Ця форма за точністю порівняння з першою, але вимагає значно більше часу на навчання. Виходячи з цього, для завдань розпізнавання акустичних образів найбільш точним і вимагаючим найменшого навчання являється представлення акустичного сигналу у вигляді набору векторів ознак фреймів.

Методика SIFT виділення періоду основної гармоніки акустичного сигналу

Довжину періоду основної гармоніки можна визначати або шляхом пошуку максимального значення в автокореляційній функції, або шляхом пошуку мінімального значення у функції середнього значення різниці амплітуд сигналу, або шляхом пошуку різниці двох максимальних значень в послідовності деталізуючих вейвлет-коефіцієнтів. Для визначення довжини періоду основної гармоніки T_{0T} шляхом пошуку максимального значення в автокореляційній функції використовується методика SIFT [3, 13]. Алгоритм знаходження періоду показаний на рис. 5.

На першому етапі для підвищення надійності визначення частоти і зменшення обчислювальної складності за допомогою еліптичного фільтра нижніх частот з частотою зрізу $f_{зп} = 1000$ Гц виділяється частотний діапазон, що містить частоту основної гармоніки. Замість еліптичного фільтра, раціонально застосувати послідовне обчислення ДПФ, виділення нижніх частот і обчислення зворотного ДПФ для кожного фрейма сигналу завдовжки ΔN . Це дозволяє швидше і точніше виділити низькочастотну складову сигналу, тобто:

$$X(k) = \sum_{n=0}^{\Delta N-1} x(n)e^{-j(2\pi/\Delta N)nk}, \quad k \in \overline{0, \Delta N-1}; \quad (1)$$

$$X_H(k) = \begin{cases} X(k), & 0 \leq k \leq k_{зп} \\ 0, & k_{зп} < k \leq \Delta N-1 \end{cases}, \quad k_{цт} = (f_{зп} \cdot \Delta N / f_d); \quad (2)$$

$$y(n) = \frac{1}{\Delta N} \sum_{k=0}^{\Delta N-1} X_H(k)e^{j(2\pi/\Delta N)nk}, \quad n \in \overline{0, \Delta N-1}; \quad (3)$$

де f_d – частота дискретизації.

На другому етапі виконується децимація частоти дискретизації до значення $f_{1d} = 2000$ Гц, тобто вилучаються проміжні відліки сигналу:

$$s(n) = y(n \cdot \Delta n), \quad n \in \overline{0, \Delta N / \Delta n - 1}, \quad (4)$$

де $\Delta n = f_d / f_{1d}$ – коефіцієнт децимації.

На етапі 2 також обчислюються зважені різниці двох сусідніх відліків децимованого сигналу:

$$s_{\Delta}(n) = \begin{cases} s(n)w(n), & n = 0 \\ [s(n) - s(n-1)]w(n), & n > 0 \end{cases}, \quad n \in \overline{0, \Delta N / \Delta n - 1}, \quad (5)$$

де $w(n) = 0,54 + 0,46 \cos \frac{2\pi n}{\Delta N}$ – вікно Хеммінга.

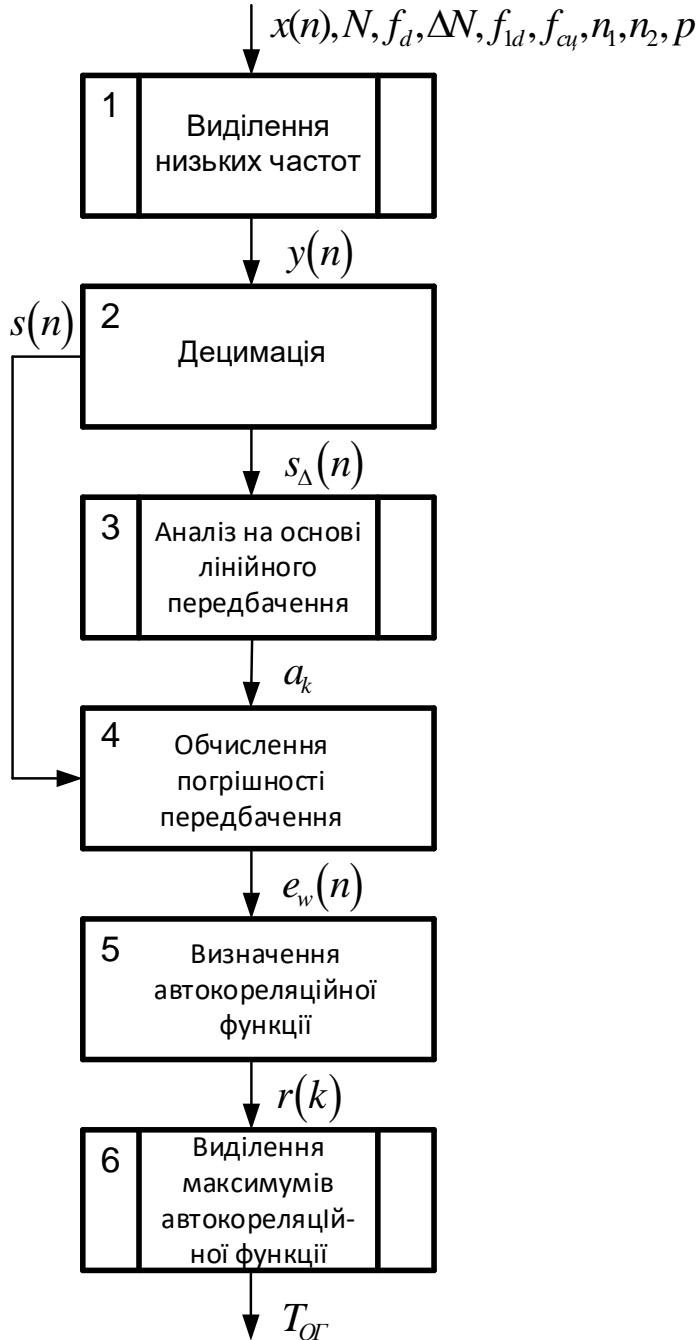


Рис. 5 – Алгоритм пошуку періоду основної гармоніки за методикою SIFT

На етапах 3-4 для відвертання появи побічних піків в автокореляційній функції виконується їх відсікання за методом КЛП (обчислення коефіцієнтів лінійного передбачення) і аналіз децимованого сигналу. Результатом аналізу є визначення коефіцієнтів a_k за алгоритмом Дарбіна (процедура $DARBIN(R(k), p)$), тобто:

$$R(k) = \sum_{n=0}^{\Delta N / \Delta n - 1 - k} s_{\Delta}(n) s_{\Delta}(n+k), \quad k \in \overline{0, p}, \quad (6)$$

$$\bar{a} = \text{DARBIN}(R(k), p), \quad \bar{a} = (a_1, \dots, a_p), \quad (7)$$

де $R(k)$ – автокореляційна функція, p – порядок лінійного передбачення.

Зазвичай порядок лінійного передбачення вибирають рівним $p=4, 5$.

На етапі 4 обчислюється погрішність передбачення $e(n)$ з використанням КЛП:

$$e(n) = \begin{cases} s(n), & n < p \\ s(n) - \sum_{k=1}^p a_k s(n-k), & n \geq p \end{cases}, \quad n \in \overline{0, \Delta N / \Delta n - 1}. \quad (8)$$

На цьому ж етапі виконується зважування обчисленої погрішності передбачення за допомогою вікна Хеммінга:

$$e_w(n) = e(n)w(n), \quad n \in \overline{0, \Delta N / \Delta n - 1}. \quad (9)$$

На етапі 5 обчислюється автокореляційна функція погрішності передбачення:

$$r(k) = \sum_{n=0}^{\Delta N / \Delta n - 1 - k} e_w(n) e_w(n+k), \quad k \in \overline{0, \Delta N / \Delta n - 1}. \quad (10)$$

На етапі 6 визначається, при яких значеннях n_0 , $n_0 \in [n_1, n_2]$ автокореляційна функція погрішності передбачення $r(k)$ максимальна, що відповідає виділенню максимумів (піків) в спектрі акустичного сигналу:

$$r_M = \max_{n_0} r(n_0), \quad n_0 \Delta n \in [n_1, n_2], \quad (11)$$

де n_1 – мінімальна довжина періоду основної гармоніки $n_1 = \inf T_{\text{ог}}$; n_2 – максимальна довжина періоду основної гармоніки $n_2 = \sup T_{\text{ог}}$.

Отримане значення визначається як \tilde{n} :

$$\tilde{n} = \arg \max_{n_0} r(n_0). \quad (12)$$

В результаті довжина періоду основної гармоніки може бути визначена у вигляді:

$$T_{\text{ор}} = \begin{cases} \tilde{n}\Delta n, & r_M \geq \gamma, \\ 0, & r_M < \gamma \end{cases} \quad (13)$$

де γ – порогове значення, визначуване експериментальним шляхом.

На рис. 6 показаний зразковий вид автокореляційної функції погрішності передбачення.

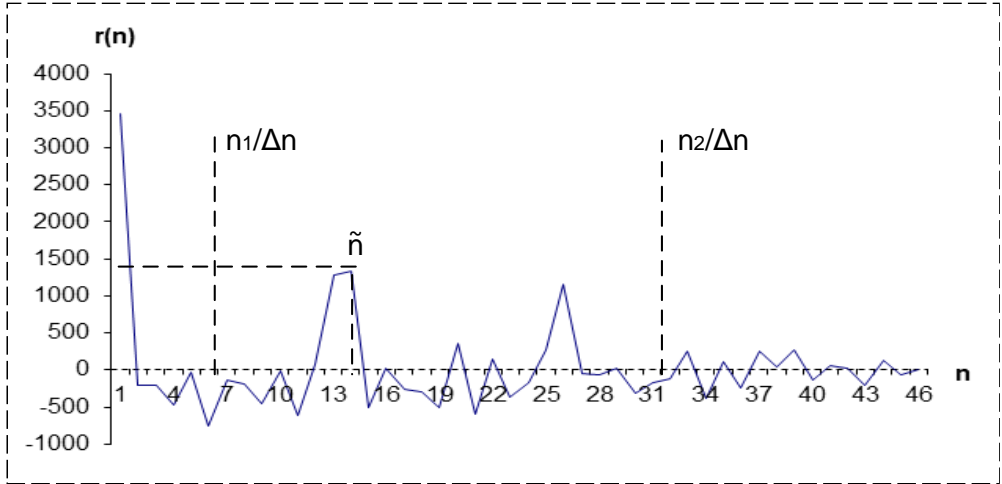


Рис. 6 – Автокореляційна функція погрішності передбачення

Методика ВОТ-А виділення періоду основної гармоніки акустичного сигналу

Для визначення довжини періоду основної гармоніки $T_{\text{ор}}$ шляхом пошуку мінімального значення у функції середнього значення різниці амплітуд сигналу використовується методика ВОТ-А [7, 14]. Згідно з цією методикою на першому етапі для кожного s-го фрейма сигналу завдовжки ΔN обчислюється функція середнього значення різниці

$$v(k) = \left| \frac{\Delta N - 1}{\sum_{m=0}^{\Delta N - 1} x(m) - x(m+k)} \right|, \quad k \in \overline{0, \Delta N - 1}. \quad (14)$$

На другому етапі визначається, при яких значеннях n_0 , $n_0 \in [n_1, n_2]$ функція середнього значення різниці $v(k)$ мінімальна, що відповідає виділенню періодів в акустичному сигналі:

$$\hat{v} = \min_{n_0} v(n_0), \quad n_0 \in [n_1, n_2], \quad (15)$$

де n_1 – мінімальна довжина періоду основної гармоніки, $n_1 = \inf T_{\text{ор}}$; n_2 – максимальна довжина періоду основної гармоніки, $n_2 = \sup T_{\text{ор}}$.

Отримане значення визначається як:

$$\tilde{n} = \arg \min_{n_0} \hat{v}(n_0). \quad (16)$$

Період основної гармоніки знаходиться як:

$$T_{\text{ог}} = \begin{cases} \tilde{n}, & n_1 \leq \tilde{n} \leq n_2 \\ 0, & \text{інакше} \end{cases}. \quad (17)$$

Методика ВОТ-ВП виділення періоду основної гармоніки акустичного сигналу

Визначити довжину періоду основної гармоніки $T_{\text{ог}}$ акустичного сигналу можна, застосувавши вейвлет-перетворення (методика ВОТ-ВП) [12, 14].

На першому етапі обчислюється пряме безперервне вейвлет-перетворення, яке можна апроксимувати відповідно до формули прямокутників у вигляді:

$$d_{\mu l} = \sum_{n=0}^{N-1} x(n) a_0^{-\mu/2} \overline{\psi(a_0^{-\mu} n - b_0 l)} \Delta t, \quad l \in \overline{0, N-1}, \quad \Delta t = 1/f_d, \quad (18)$$

де μ – рівень розкладання, на якому досягається гладка гармоніка, N – довжина сигналу, Δt – крок квантування.

При виконанні вейвлет-перетворення використовується базисна функція Морле:

$$\psi(\xi) = (2\pi)^{-1/2} \cos(k_0 \xi) e^{-\xi^2/2}, \quad k_0 = 5, \quad \xi = a_0^{-m} n - b_0 l. \quad (19)$$

Оскільки послідовність $d_{\mu l}$ є гладкою гармонікою, відпадає необхідність використання автокореляційної функції і функції середнього значення різниці амплітуд сигналу, які складно обчислювати. Замість цього на другому етапі в послідовності $d_{\mu l}$ визначаються два максимуми, що йдуть підряд, і обчислюється різниця між ними \tilde{n} , у вигляді

$$(d_{\mu j-1} \leq d_{\mu j} \geq d_{\mu, j+1} \wedge d_{\mu, m-1} \leq d_{\mu m} \geq d_{\mu, m+1} \wedge d_{\mu, k-1} \leq d_{\mu k} \geq d_{\mu, k+1} \wedge (j < k < m)) \rightarrow \tilde{n} = m - j. \quad (20)$$

Період основної гармоніки знаходиться як:

$$T_{\text{ог}} = \begin{cases} \tilde{n}, & n_1 \leq \tilde{n} \leq n_2 \\ 0, & \text{інакше} \end{cases}, \quad (21)$$

де n_1 – мінімальна довжина періоду основної гармоніки, $n_1 = \inf T_{\text{ог}}$; n_2 – максимальна довжина періоду основної гармоніки, $n_2 = \sup T_{\text{ог}}$.

Висновки

Головною перевагою методики SIFT є вирівнювання спектра сигналу, що полегшує оцінювання основної гармоніки, а до недоліків можна віднести: використання порогу, що на деяких фреймах акустичного сигналу може призвести до помилки; значну обчислювальну складність; залежність від рівня шуму.

Достоїнствами методики VOT-A є: відсутність заздалегідь заданих порогів; нескладність обчислень. При цьому, проте, зберігається залежність від рівня шуму.

Методика VOT-ВП характеризується відсутністю заздалегідь заданих порогів; швидким пошуком періоду основної гармоніки; відсутністю залежності від рівня шуму, оскільки досліджується конкретний діапазон частот. В той же час обчислювальна складність вейвлет-перетворення відносно висока, у зв'язку з чим потрібна оптимізація алгоритму обчислень. За сукупністю характеристик при обчисленні довжини періоду основної гармоніки акустичного сигналу кращою є методика VOT-ВП.

СПИСОК ЛІТЕРАТУРИ

1. Орлов Д.С. Экология и охрана гидросферы при химическом загрязнении: Учеб. пособие / Орлов Д.С., Садовникова Л.К., Лозановская И.Н. – М.: Высшая школа, 2012. – 167 с.
2. Оппенгейм А.В., Шафер Р.В. Цифровая обработка сигналов. – М.: Связь, 1979. – 416 с.
3. Рабинер Л.Р., Шафер Р.В. Цифровая обработка речевых сигналов. – М.: Радио и связь, 1981. – 495 с.
4. Rabiner L.R., Jang B.H. Fundamentals of speech recognition. – New Jersey: Prentice Hall PTR, Englewood Cliffs, 1993. – P. 507.
5. Солонина А.И., Улахович Д.А., Яковлев Л.А. Алгоритмы и процессоры цифровой обработки сигналов. – СПб.: БХВ-Петербург, 2001 – 464 с.
6. Сеунов Н.Ю. Обработка звука на РС. – СПб.: БХВ – Санкт-Петербург, 2001. – 1248 с.
7. Федоров Е.Е. Модели и методы преобразования речевых сигналов – Донецк: изд-во «Норд-Пресс», 2006. – 260 с.
8. Акустика: Справочник / Е.П. Ефимов, А.В. Никонов, М.А. Сапожков, В.И. Шоров / Под ред. М.А. Сапожкова. – М.: Радио и связь, 1989. – 336 с.
9. Федоров Е.Е., Шелепов В.Ю. Автоматическое определение начала и конца записи речи // Искусственный интеллект. – 2002. – № 4. – С. 295–298.
10. Freeman D., Sonthcott C., Boyd I.A. Voice activity detector for the Pan-European digital cellular mobile telephone service // IEE Colloquium «Digitized Speech Communication via Mobile Radio». – London. – 1988. – P. 61–65.
11. Карпов А.А. Робастный метод определения границ речи на основе спектральной энтропии // Искусственный интеллект. – 2004. – № 4. – С. 606–612.
12. Федоров Е.Е. Методика формирования акустических характеристик эталонов речи. – Донецк: изд-во «Вебер», 2008. – 282 с.
13. Маркел Д.Д., Грэй А.Х. Линейное предсказание речи. – М.: Связь, 1980. – 308 с.
14. Винцюк Т.К. Анализ, распознавание и интерпретация речевых сигналов. – К.: Наук. думка, 1987. – 261 с.
15. Shannon B.J., Paliwal K.K. A comparative study of filter bank spacing for speech recognition // Proc. of Microelectronics engineering research conference. – Brisbane. – 2003. – 310–312 pp.

16. Kocharov D.A. Sonority Measure for Automatic Speech Recognition // Труды междунар. конф. SPECOM'2006. – Санкт-Петербург. – 2006. – С. 245–250.
17. Атал Б.С. Автоматическое опознавание дикторов по голосам // ТИИЭР. – 1976. – Т. 64, №4. – С. 48–66.
18. Davis S.B., Mermelstein P. Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences // IEEE Trans. on Acoustic, Speech and Signal Processing. – 1980. – Vol. 28, №4. – 357–366 pp.
19. Дорохин О.А., Засыпкин А.В., Червин Н.А., Шелепов В.Ю. О некоторых подходах к проблеме компьютерного распознавания устной речи // Труды Междунар. конф. «Знание – Диалог – Решение» (KDS 97). – Т. 1. – Ялта. – 1997. – С. 234–240.
20. Молдокулова Н.В., Трунин-Донской В.Н. Лингво-акустические проблемы создания системы распознавания слитной речи на ЭВМ / Ответ. ред. Журавлев Ю.И.; АН КиргССР, Вычислительный центр АН СССР. – Фрунзе: Илим. – 1989. – 136 с.
21. Добеши И. Десять лекций по вейвлетам. – М.: РХД, 2004. – 464 с.
22. Чуи К. Введение в вэйвлеты. – М.: Мир, 2001. – 412 с.
23. Малла С. Вэйвлеты в обработке сигналов. – М.: Мир, 2005. – 671 с.
24. Bourlard H., Morgan N. Connectionist Speech Recognition: A Hybrid Approach. – Boston: Kluwer Academic Publishers, 1994. – 312 p.

Стаття надійшла до редакції 10.11.2020 і прийнята до друку після рецензування 08.02.2021

REFERENCES

1. Orlov, D.S., Sadovnikova, L.K., & Lozanovskaya, I.N. (2012). *Ekologiya i ohrana hidrosfery pri khimicheskom zagryaznenii* [Ecology and protection of the hydrosphere with chemical pollution]. Moscow: Vysshaya shkola (in Russian).
2. Oppengeym, A.V., & Shafer, R.V. (1979). *Tsyfrovaya obrabotka signalov* [Digital signal processing]. Moscow: Svyaz (in Russian).
3. Rabiner, L.R., & Shafer, R.V. (1981). *Tsyfrovaya obrabotka rechevykh signalov* [Digital processing of speech signals]. Moscow: Radio i Svyaz' (in Russian).
4. Rabiner, L.R., & Jang, B.H. (1993). *Fundamentals of speech recognition*. New Jersey: Prentice Hall PTR, Englewood Cliffs.
5. Solonina, A.I., Ulakhovich, D.A., & Yakovlev, L.A. (2001). *Algoritmy i protsessory tsyfrovoy obrabotki signalov* [Algorithms and processors for digital signal processing]. Sankt Peterburg: BKHV-Peterburg (in Russian).
6. Sekunov, N.Yu. (2001). *Obrabotka zvuka na PC* [Sound processing on PC]. Sankt Peterburg: BKHV-Peterburg (in Russian).
7. Fyedorov, Ye.Ye. (2006). *Modeli i metody preobrazovaniya rechevykh signalov* [Models and methods of converting speech signals]. Donetsk: Nord-Press (in Russian).
8. Yefimov, Ye.P., Nikonov, A.V., Sapozhkov, M.A., & Shorov, V.I. (1989). *Akustika: Spravochnik* [Acoustics: Handbook]. (M.A. Sapozhkov, Ed.). Moscow: Radio i Svyaz' (in Russian).
9. Fyedorov, Ye.Ye., & Sheleпов, V.Yu. (2002). Automatic detection of the beginning and end of speech recording. *Iskusstvennyi intellekt*, 4, 295-298 (in Russian).
10. Freeman, D., Sonthcott, C., & Boyd, I.A. (1988). Voice activity detector for the Pan-European digital cellular mobile telephone service. In IEE Colloquium *Digitized Speech Communication via Mobile Radio* (pp. 61-65). London.
11. Karpov, A.A. (2004). Robust method for determining the boundaries of speech based on spectral entropy. *Iskusstvennyi intellekt*, 4, 606-612 (in Russian).
12. Fyedorov, Ye.Ye. (2008). *Metodika formirovaniya akusticheskikh kharakteristik etalonov rechi* [Methodology for the formation of acoustic characteristics of speech standards]. Donetsk: Veber (in Russian).

13. Markel, D.D., & Grey, A.Kh. (1980). *Lineynoe predskazaniye rechi* [Linear speech prediction]. Moscow: Svyaz' (in Russian).
14. Vintsyuk, T.K. (1987). *Analiz, raspoznavaniye i interpretatsiya rechevykh signalov* [Analysis, recognition and interpretation of speech signals]. Kyiv: Naukova dumka (in Russian).
15. Shannon, B.J., & Paliwal, K.K. (2003). A comparative study of filter bank spacing for speech recognition. In *Proc. of Microelectronics engineering research conference* (pp. 310-312). Brisbane.
16. Kocharov, D.A. (2006). Sonority Measure for Automatic Speech Recognition. In *Trudy mezhd. konf. SPECOM'2006* (pp. 245-250). Sankt Peterburg.
17. Atal, B.S. (1976). Avtomaticheskoye opoznavaniye diktovor po golosam [Automatic recognition of speakers by voices]. In *Proceedings of IEEE*, 64(4), 48-66 (in Russian).
18. Davis, S.B., & Mermelstein, P. (1980). Comparison of Parametric Representations for Monosyllabic Word Recognition in Continuously Spoken Sentences. In *IEEE Trans. on Acoustic, Speech and Signal Processing*, 28(4), 357-366.
19. Dorokhin, O.A., Zasyplin, A.V., Chervin, N.A., & Shyelyepov, V.Yu. (1997). O nekotorykh podkhodakh k probleme komp'yuternogo raspoznavaniya ustnoy ryechi [Some approaches to the problem of computer speech recognition]. In *Trudy Myezhdunarodnoy konferentsiyi "Znaniye – Dialog – Ryeshyeniye" (KDS 97)* (Vol. 1. 234-240). Yalta (in Russian).
20. Moldokulova, N.V., & Trunin-Donskoy, V.N. (1989). *Lingvo-akusticheskiye problemy sozdaniya sistemy raspoznavaniya slitnoy ryechi na EVM* [Linguo-acoustic problems of creating a continuous speech recognition system on a computer]. Frunzye: Ilim (in Russian).
21. Dobyeshi, I. (2004). *Dyesyat' lektsiy po veyvletam* [Ten lectures on wavelets]. Moscow: RKhD (in Russian).
22. Chui, K. (2001). *Vvedenie v veyvlety* [Introduction to wavelets]. Moscow: Mir (in Russian).
23. Malla, S. (2005). *Veyvlety v obrabotkye signalov* [Wavelets in signal processing]. Moscow: Mir (in Russian).
24. Bourlard, H., & Morgan, N. (1994). *Connectionist Speech Recognition: A Hybrid Approach*. Boston: Kluwer Academic Publishers.

The article was received 10.11.2020 and was accepted after revision 08.02.2021

Ткаченко Тетяна Миколаївна

доктор технічних наук, професор, професор кафедри охорони праці та навколишнього середовища Київського національного університету будівництва і архітектури

Адреса робоча: 03037, Україна, м. Київ, Повітрофлотський просп., 31

ORCID ID 0000-0003-2105-5951 **e-mail:** tkachenkoknuba@gmail.com

Пількевич Юлія Георгіївна

Начальник відділу МТЗ Київського національного університету будівництва і архітектури

Адреса робоча: 03037, Україна, м. Київ, Повітрофлотський просп., 31

ORCID ID 0000-0002-7618-0250 **e-mail:** y.rozorinova@gmail.com

Розорінов Георгій Миколайович

доктор технічних наук, професор, професор кафедри акустичних мультимедійних електронних систем Національного технічного університету України "Київський політехнічний інститут імені Ігоря Сікорського"

Адреса робоча: 03056, Україна, м. Київ, просп. Перемоги, 37

ORCID ID 0000-0002-6095-7539 **e-mail:** grozoryn@gmail.com