

ОПТИМІЗАЦІЯ ПОШУКУ БАЗОВИХ МАЛОХВИЛЬОВИХ ФУНКЦІЙ З ВИКОРИСТАННЯМ ГЕНЕТИЧНОГО АЛГОРИТМУ

© Лагун І. І., Наконечний Р. А., 2015

Запропоновано використання генетичного алгоритму для пошуку оптимальних базових малохвильових функцій у процесі очищення сигналів від шуму. Для реалізації генетичного алгоритму використана фітнес-функція у вигляді мінімального середньоквадратичного відхилення. Наведено результати досліджень очищення сигналу від шуму на основі вибраних оптимальних параметрів базових функцій для тестових сигналів пакета Matlab.

Ключові слова: малохвильовий базис, материнська малохвильова функція, оптимальна малохвильова функція, вейвлет перетворення, хвильове перетворення.

The genetic algorithm has been used to select the optimal wavelet basis functions which lead to effective denoising of signals. The mean square error (MSE) was used as fitness function for the implementation of genetic algorithm. The results of researches show efficient denoising for test signals of package Matlab based on selected optimal parameters.

Key words: basis wavelet, mother wavelet, optimal wavelet.

Вступ

Відомо, що представлення сигналів у часо-частотній, малохвильовій області дає змогу отримати всебічну інформацію про сигнал. Однак ефективність подання сигналів та їх аналіз у такій області суттєво залежить від вибору базових функцій, які при цьому використовуються. Оптимальний вибір таких функцій забезпечує необхідну точність апроксимації інформативних сигналів у часо-частотній області, а також дає змогу досягати високої швидкодії перетворення та мінімізувати необхідний обсяг пам'яті.

Оптимізація вибору базових малохвильових функцій відіграє важливе значення при очищенні сигналів від завад та шумів і у багатьох випадках визначає кількість рівнів розкладу сигналу та спосіб пороговання малохвильових коефіцієнтів.

У низці праць [3–5] для оптимізації вибору базових малохвильових функцій пропонується використовувати критерій мінімального середньоквадратичного відхилення відтвореного сигналу від оригінального

$$MSE(X, Y) = \frac{1}{N} \sum_{i=1}^N (x_i - y_i)^2, \quad (1)$$

де N – кількість відліків сигналу; X – вихідний сигнал та Y – відтворений сигнал.

Однак такий метод оптимізації у багатьох випадках не задовольняє споживачів, особливо тоді, коли аналізовані неперіодичні сигнали містять швидкоплинні складові.

У поданій праці пропонується використовувати генетичний алгоритм (ГА), який дає змогу отримати оптимальні параметри для послаблення шуму, що, своєю чергою, передбачає використання оптимальних малохвильових функцій для забезпечення високої якості реконструйованого сигналу.

1. Фільтрування сигналів на основі їх малохвильового перетворення

Методи малохвильової порогової обробки для видалення шумів та завад вперше розглянув Донохо у 1993 році. Ці методи не вимагають особливих припущень про природу сигналу, можливі

розриви та просторові зміни в ньому і використовують просторово-адаптивне багатомасштабне малохвильове перетворення.

Серед існуючих найвідоміших методів послаблення шумів та завад із малохвильовим перетворенням є методи, як запропонували Маллат і Донохо, а також метод інваріантної граничної фільтрації TIS [5]. Метод Маллата ґрунтується на очищенні від білого шуму добуванням квадратного кореня з максимальних значень малохвильових коефіцієнтів для різних масштабів. Такий підхід дає змогу максимально точно зберігати основні особливості сигналу, однак він вкрай складний для апаратної реалізації [2]. У методі Донохо використовують лише пірамідальні фільтри та один рівень порогоування, і в ньому не розглядаються різні характеристики частотних смуг [5].

Усі алгоритми малохвильової фільтрації побудовано на таких властивостях малохвильового розкладу:

- оскільки шуми та завади мають низький рівень енергії, то інформація про інформативну складову сигналу $x(n)$ утримується лише у невеликій кількості коефіцієнтів малохвильового розкладу, інші коефіцієнти прирівнюються до нуля;

- шумові складові рівномірно “перерозподіляються” за усіма рівнями малохвильового розкладу, зокрема якщо значення шумових складових є некорельовані й мають однакову дисперсію, то коефіцієнти їх розкладу відносно базису ортогональних малохвильових функцій також некорельовані між собою і мають однакову дисперсію.

Якщо вважати x вхідним інформативним сигналом, а y – функцією n -зашумлених вибірок, то:

$$y_i = x_i + s \cdot e_i, \quad i = 1 \dots n, \quad (2)$$

де e_i – гауссівський білий шум з $N(0,1)$; s – рівень шуму, який може бути відомим або невідомим.

Мета фільтрування полягає в тому, щоб відтворити вхідну інформативну функціональну залежність x з зашумлених даних, $y = (y_1, \mathbf{K}, y_n)$ з мінімальною похибкою. Переважно припускають, що вибірккові точки рівномірно розподіляються і $n = 2^j$ для деякого $j \in N$. Такі припущення дають змогу реалізувати дискретне малохвильове перетворення (ДМП) і обернену його форму з використанням швидкого алгоритму Маллата.

Ортогональність ДМП дає змогу отримати функцію перетворення білого шуму. Отже, якщо y_{jk} (j – рівень розкладу, k – індекс коефіцієнта на цьому рівні) є малохвильовими коефіцієнтами перетворення (2) у малохвильовому домені, то:

$$y_{jk} = w_{jk} + s \cdot e_{jk}, \quad (3)$$

де w_{jk} – малохвильові коефіцієнти інформативного сигналу $x(n)$, що складаються з апроксимувальних та деталізувальних коефіцієнтів, а $e_{jk} - N(0,1)$.

Загалом деталізувальні коефіцієнти з малими значенням можна розглядати як шум з рівномірним спектром, і вони можуть бути відкинуті. Такий підхід, за яким кожний коефіцієнт порівнюють із порогом, належить до порогової малохвильової обробки.

Порогову обробку малохвильових коефіцієнтів застосовують лише до коефіцієнтів деталізації d_{jk} , а не до масштабувальних коефіцієнтів. За допомогою порогоування виділяються значимі коефіцієнти, абсолютні значення яких вищі за поріг певного рівня l , який може бути функцією рівня розкладу j й індексу k , тобто: $l = l(j, k)$, однак в переважній більшості це лише функція j : $l = l(j)$. В останньому випадку цей поріг залежить лише від рівня. Результуючі малохвильові коефіцієнти, отримані в процесі “жорсткого” або “м’якого” порогоування, представляються відповідно виразами:

$$d_l^H(d_{jk}) = \begin{cases} 0, & \text{якщо } |d_{jk}| \leq l \\ d_{jk}, & \text{якщо } |d_{jk}| > l \end{cases}, \quad (4)$$

$$d_l^S(d_{jk}) = \begin{cases} 0, & \text{якщо } |d_{jk}| \leq l \\ d_{jk} - l, & \text{якщо } d_{jk} > l \\ d_{jk} + l, & \text{якщо } d_{jk} < -l \end{cases} \quad (5)$$

Від вибору рівня граничного шуму (дисперсії шуму) залежать якісні характеристики, які оцінюються у вигляді відношення сигнал/шум. Задання малих значень порогу λ зберігає інформацію про шумову складову в коефіцієнтах деталізації, а тому приводить до незначного збільшення відношення сигнал/шум. За високих значень порогу втрачається частина коефіцієнтів, які несуть корисну інформацію. Пошук оптимального значення рівня порогоування означає визначення такого порогу, за якого забезпечується найбільше значення відношення сигнал/шум. Крім того, якість усунення впливу шуму сигналу залежить також і від способу мультиплікативного порогового перемасштабування [10].

Отже, якість очищення сигналів від шумів з використанням ДМП залежить від таких основних параметрів:

- типу вибраної базової малошвилювальної функції w ;
- рівня розкладу j ;
- функції порогу λ ,
- правила вибору порогу β ;
- способу перемасштабування порогу μ . У запропонованому методі усунення шуму, для вибору оптимальних значень малошвилювальних характеристик фільтрування пропонується використання генетичного алгоритму, який забезпечує максимальні якісні характеристики очищення сигналів.

2. Використання генетичного алгоритму для розв'язання задач оптимізації

Генетичний алгоритм (ГА) належить до класу стохастичних алгоритмів глобальної оптимізації. Як і інші алгоритми цієї групи, він орієнтований на пошук найоптимальнішого розв'язку в умовах неможливості повного перебору варіантів.

Відмінності ГА від інших алгоритмів випадкового пошуку полягають у такому [1]:

- ГА використовує не самі характеристики, а їхні коди, які формуються в ланцюжки кінцевої довжини; у процесі ітерацій ці ланцюжки тестуються й видозмінюються;
- пошук відбувається не в точках простору, а на основі популяції точок;
- ГА використовує інформацію про минулі отримані розв'язки, що забезпечує концентрацію розв'язків у найефективніших областях простору;
- ГА використовує лише оцінки розв'язків, а не їхні похідні або інші параметри.

Загалом ГА використовує три фундаментальні генетичні оператори: вибору, схрещування (кросингвера) і мутації [2].

ГА побудований на правилі “виживає найпридатніший” і дає змогу отримувати оптимальне рішення після серії повторних обчислень з використанням генетичних операторів. Розмір популяції та ймовірність норми для схрещування і мутації належать до контрольних параметрів генетичного алгоритму. Алгоритм генерує чергові популяції додаткових рішень, які представлені хромосомою, тобто вирішення завдання триває до отримання прийнятних результатів, продуктивність яких оцінюється у фітнес-функції [7].

Для розглядуваного випадку основною метою оптимізації є усунення шуму та зведення до мінімуму середньоквадратичної похибки (MSE) між вихідним сигналом і очищеною версією цього сигналу. З огляду на це як фітнес-функцію доцільно використовувати значення середньоквадратичної похибки.

Ймовірність схрещування зазвичай вибирають доволі високою: 80–95 %. Однак, у деяких завданнях найкращого результату досягають при кросингвері з ймовірністю 60 % [7].

Ймовірність мутації вибирають малою: 0,5–1 %, оскільки дуже великий розмір популяції зазвичай не призводить до високих результатів (зростає швидкість збіжності алгоритму). Проведені дослідження показують, що оптимальний розмір популяції залежить від розміру кодових

ланцюжків (хромосом). Оптимальним є розмір популяції 20–30 особин, проте для розв’язання окремих задач потрібно 50–100 особин [7]. Вибір способу кодування зумовлений поставленим завданням і розміром об’єкта пошуку. Оператори життєвого циклу (схрещування і мутації) визначаються обраним кодуванням.

3. Застосування генетичного алгоритму для пошуку оптимальних базових малохвильових функцій у процесі очищення сигналу

Враховуючи усі необхідні вимоги, для досліджень вибрано сімейства ортогональних малохвильових функцій із компактним носієм: Добеші (db1...db20), Койфлети (coif1...coif5), Симлети (sym1...sym8).

Дослідження проводили на основі тестових сигналів пакета MATLAB та інструментарію “Genetic Algorithm and Direct Search Toolbox”. Шумові складові отримано моделюванням в пакеті MATLAB білого шуму з різними значеннями рівня шуму S та з кількістю відліків, що дорівнює кількості відліків корисного сигналу.

На рис.1 наведено узагальнену блок-схему пошуку оптимальних параметрів очищення сигналу від шуму з використанням ГА.

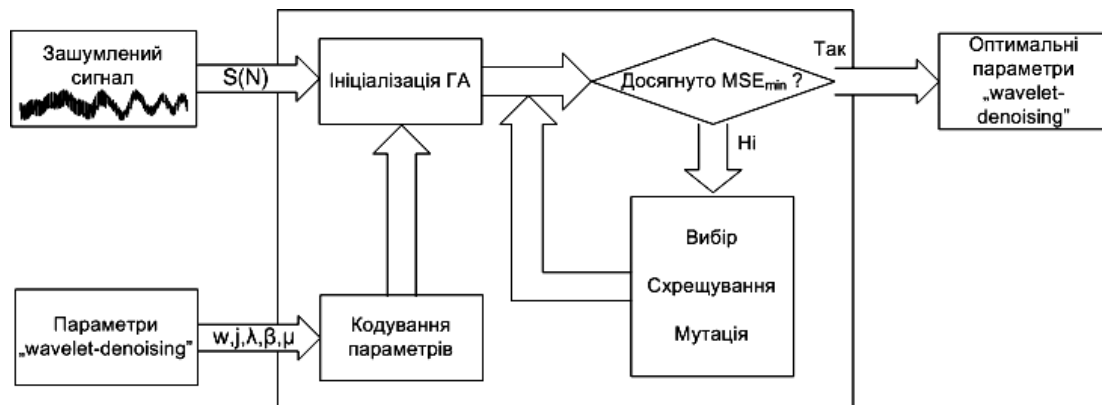


Рис. 1. Узагальнена блок-схема пошуку оптимальних параметрів очищення сигналу від шуму з використанням генетичного алгоритму

На перший вхід схеми надходить зашумлений сигнал $S(N)$, а на другий – набір параметрів компенсації шуму ($w, j, \lambda, \beta, \mu$). Оптимізація згаданих характеристик реалізується за допомогою ГА. Коли задовольняється критерій оптимізації, завершується робота ГА.

У результаті оптимізації отримують масив характеристик компенсації шуму ($w, j, \lambda, \beta, \mu$). Тобто процес очищення сигналів передбачає такі етапи:

1. ДМП зашумленого сигналу з використанням отриманого масиву оптимальних малохвильових коефіцієнтів.
2. Порогування шумових коефіцієнтів сигналу з використанням оптимального порогу та одержання модифікованих компонент сигналу.
3. Реконструкція сигналу з використанням ОДМП.

Для проведення досліджень використовували такі тестові сигнали пакета MATLAB: ‘brkintri’, ‘cuspamax’, ‘freqbrk’, ‘leleccum’, ‘linchirp’, ‘mfrqbrk’, ‘nearbrk’, ‘qdchirp’, ‘scddvbrk’, ‘sinfract’, ‘sinper8’, ‘sumlichr’, ‘sumsin’, ‘trsin’, ‘koch’, ‘wcantor’, ‘wntrsin’, ‘wstep’.

У процесі досліджень стаціонарний ГА з рівнями популяції від 40 до 100 осіб еволюціонував протягом 60, 80 та 100 поколінь з використанням коефіцієнта кросовера 90% та рівнем мутації 1%.

Значення параметрів для компенсації шуму вибирали такі:

- малохвильова функція w : Добеші (db1...db20), Койфлети (coif1...coif5), Симлети (sym1...sym8);
- рівень розкладання j : 1-10;

- функція порогоування λ : “soft”, “hard”;
- правило вибору порогу β : “heursure”, “rigsure”, “sqtwolog”, “minimax”;
- метод перемасштабування порогу μ : по scaling (one), single level (sln), multiple level (mln).

Найкращі показники було отримано для характеристик генетичного алгоритму PopulationSize=60, Generations=60. У табл. 1 наведено оптимальні характеристики компенсації шуму для рівнів 0,5 та 0,6.

Таблиця 1

**Оптимальні значення параметрів компенсації шуму
для $s = 0,5$ та $s = 0,6$ при PopulationSize=60, Generations=60**

Тестовий сигнал	MSE		Малохвильова функція w		Рівень розкладання j		Функція порогоування λ		Правило вибору порогу β		Метод перемасштабування порогу μ	
	0,5	0,6	0,5	0,6	0,5	0,6	0,5	0,6	0,5	0,6	0,5	0,6
‘brkintri’	0,00064	0,00114	sym8	coif5	7	7	s	h	heursure	sqtwolog	one	one
‘cuspamax’	0,00243	0,00308	sym4	coif5	7	6	s	h	heursure	sqtwolog	sln	one
‘freqbrk’	0,01996	0,02477	sym14	sym14	6	6	h	s	rigsure	heursure	one	sln
‘leleccum’	0,25637	0,34859	db1	db1	1	1	s	h	rigsure	heursure	one	one
‘linchirp’	0,15595	0,23109	db10	db10	7	7	s	h	rigsure	rigsure	mln	sln
‘mfrqbrk’	0,35761	0,44169	db11	db3	7	1	s	h	rigsure	rigsure	mln	sln
‘nearbrk’	0,02295	0,02265	sym2	coif1	6	5	s	h	rigsure	sqtwolog	sln	sln
‘qdchirp’	0,134947	0,15961	db11	db10	5	6	s	s	rigsure	rigsure	mln	mln
‘scddvbrk’	0,00048	0,00094	sym8	sym11	7	7	s	s	sqtwolog	heursure	sln	sln
‘sinfract’	0,00408	0,00464	coif4	sym7	7	7	h	h	rigsure	minimaxi	one	one
‘sinper8’	0,02624	0,05164	db1	sym14	6	7	h	h	heursure	rigsure	one	one
‘sumlichr’	0,22101	0,29901	db11	sym6	1	1	h	s	rigsure	rigsure	sln	sln
‘sumsin’	0,52876	0,53679	sym14	sym14	7	6	h	s	heursure	rigsure	one	sln
‘trsin’	0,02722	0,03440	sym14	sym14	7	7	h	h	heursure	heursure	one	one
‘vonkoch’	0,00123	0,00202	db9	coif4	7	7	s	s	heursure	sqtwolog	sln	mln
‘wcantor’	0,00053	0,00138	sym7	sym13	7	7	s	s	heursure	heursure	one	mln
‘wntrsin’	0,08859	0,10443	sym13	sym13	7	7	s	s	rigsure	heursure	sln	sln
‘wstep’	0,00095	0,02221	db1	db2	7	7	h	s	sqtwolog	heursure	mln	sln

У роботі [5] було запропоновано метод визначення ефективності базових функцій для опрацювання різних типів одновимірних сигналів за критерієм середньоквадратичної похибки. Для порівняння ефективності методів сигнали очищували від шуму з використанням малохвильових функцій, отриманих у результаті застосування ГА та на основі методу визначення ефективності за критерієм MSE при незмінних інших параметрах. Як критерій оцінювання ефективності усунення шуму використано відношення сигнал/шум SNR (*Signal-to-noise ratio*):

$$SNR = 10 \lg \left\{ \sum_{i=1}^N [x_i]^2 / \sum_{i=1}^N [x_i - y_i]^2 \right\}, \quad (6)$$

де N – кількість відліків сигналу; X – вихідний сигнал та Y – відтворений сигнал.

Таблиця 2

Результати очищення сигналів від шуму для $s = 0,6$

Тестовий сигнал	SNR, dB		Малохвильова функція w	
	Метод ефективності за MSE	GA	Метод ефективності за MSE	GA
1	2	3	4	5
‘brkintri’	69,42521	75,51659	sym8	coif5
‘cuspamax’	65,80537	72,6669	sym4	coif5
‘freqbrk’	56,66054	62,97335	sym14	sym14
‘leleccum’	52,40212	52,65604	db1	db1
‘linchirp’	51,53863	52,18064	db10	db10

1	2	3	4	5
'mfrqbrk'	51,53396	51,62043	db11	db3
'nearbrk'	60,49516	62,70283	sym2	coif1
'qdchirp'	52,89921	55,27089	db11	db10
'scddvbrk'	70,04649	74,1715	sym8	sym11
'sinfract'	70,99597	71,22968	coif4	sym7
'sinper8'	55,46316	58,24364	db1	sym14
'sumlichr'	51,61896	51,51368	db11	sym6
'sumsin'	49,20199	50,64084	sym14	sym14
'trsins'	53,8926	61,4708	sym14	sym14
'vonkoch'	74,25888	72,16366	db9	coif4
'wcantor'	73,68025	69,18272	sym7	sym13
'wntrsins'	54,4144	57,50216	sym13	sym13
'wstep'	67,39967	62,57543	db1	db2

Висновки

Проведені дослідження дають підстави стверджувати, що застосування генетичного алгоритму для визначення оптимальних базових малохвильових функцій у процесі фільтрації сигналів є ефективнішим, що підтверджується отриманими результатами і для SNR перевищення становить 5–10 %. Крім того, правильний вибір основних параметрів компенсації шуму, таких як рівень розкладання, тип функції порогоування, правило вибору порогу та його рівня та метод перемасштабування порогу мають важливе значення для підвищення ефективності процесу очищення сигналу від шуму.

Подальшим напрямом роботи передбачено дослідження ефективності інших критеріїв вибору базових функцій із застосуванням генетичного алгоритму.

1. Бураков М. В. Генетический алгоритм: теория и практика: учеб. пособие / М. В. Бураков. – СПб.: ГУАП, 2008. – 64 с.: ил. 2. Дьяконов В. MATLAB. Обработка сигналов и изображений: Специальный справочник / В. Дьяконов, И. Абраменкова. – СПб.: Питер, 2002. – С. 602. 3. Лагун І. І. Основні підходи до вибору оптимального малохвильового базису / І. І. Лагун, А. Й. Наконечний // Вісник Нац. ун-ту “Львівська політехніка”. – Сер.: Автоматика, вимірювання та керування. – 2009. – № 639. – С. 71. 4. Лагун І. І. Оцінка якості зображень при фільтрації сигналів різними базисними малохвильовими функціями / І. І. Лагун, А. Й. Наконечний // Збірник тез доповідей XI Міжнародної наукової конференції з контролю і управління в складних системах “КУСС-2012”. – Вінниця, 2012. – С. 101. 5. Лагун І. І. Вибір малохвильових базових функцій для опрацювання одновимірних сигналів / І. І. Лагун, А. Й. Наконечний // Вісник Нац. ун-ту “Львівська політехніка”. – Сер.: Автоматика, вимірювання та керування. – 2014. – № 802. – С. 3–8. 6. Наконечний А. Й. Цифрова обробка сигналів / А. Й. Наконечний, Р. А. Наконечний, В. А. Павлич. – Львів: Видавництво Львівської політехніки, 2010. – С. 360. 7. Панченко Т. В. Генетические алгоритмы: учебно-методическое пособие / под ред. Ю. Ю. Тарасевича. – Астрахань: Издательский дом “Астраханский университет”, 2007. – 87 с. 8. Смоленцев Н. К. Основы теории вейвлетов. Вейвлеты в Matlab. – М.: ДМК Пресс, 2005. – 304 с. 9. Donoho D. L. De-noising by soft thresholding // IEEE Transactions on Information Theory, 41, 1995. – P. 613–627. 10. Donoho D. L. & Johnstone I. M. Ideal spatial adaptation via wavelet shrinkage. Biometrika, 81, 1994. – P. 425–455.