

## ЗАСТОСУВАННЯ МЕТОДУ ГОЛОВНИХ КОМПОНЕНТ В ЗАДАЧІ АНАЛІЗУ СПЕКТРІВ ВІЛЬНИХ КОЛИВАНЬ

<sup>1</sup>Національний технічний університет України  
«Київський політехнічний інститут імені Ігоря Сікорського»

*Представлено основні методи, що застосовуються для аналізу, розподілу та класифікації простору вхідних даних. Проаналізовано і визначено головні аспекти використання вказаних методів. Визначено, що найприйнятнішим у цьому дослідженні є застосування методу головних компонент. Описано можливі алгоритми методу головних компонент і застосовано комбінацію цих алгоритмів для задачі розподілу вхідних даних під час аналізу сигналів та їх спектрів, отриманих в процесі неруйнівного контролю методом вільних коливань. Метою статті є дослідження можливості зменшення вектора інформативних ознак методом головних компонент без втрати якості розпізнавання стану об'єктів. Об'єктами дослідження можуть бути складові електродвигунів (шихтований магнітопривід), деталі літальних апаратів, виготовлені з композитних матеріалів та інші конструкції, що потребують аналізу неруйнівними методами контролю. Досліджено спектри, зняті під час неруйнівного контролю методом вільних коливань зразків вуглепластикових панелей з дефектною та бездефектною зон зразка. Визначено максимальну кількість головних компонент — дві головні компоненти, та значення максимальних дисперсій цих головних компонент для сформованого набору вхідних значень амплітуд для трьох, п'яти і десяти гармонік. Для аналізу якості розділення простору вхідних даних на класи (бездефектною та дефектною зони зразка) використано міру відстані Махаланобіса. Відмічено підвищення якості розділення значень з ушкодженої і неушкодженої зон зразка на два класи. Отже, застосування методу головних компонент, в цьому дослідженні, дозволило підвищити надійність розпізнавання стану об'єктів.*

**Ключові слова:** розподіл простору вхідних даних, інформативні параметри, метод головних компонент, спектральний аналіз, неруйнівний контроль.

### Вступ

В основі задачі неруйнівного контролю методом вільних коливань лежить порівняння спектрів сигналів, отриманих від досліджуваного об'єкта з дефектних і бездефектних зон. На сьогодні для порівняння і аналізу еталонного та дефектного спектрів існує багато критеріїв відмінності. Вибір інформативних параметрів, критеріїв його оцінювання залежить від багатьох вагомих величин: матеріалу, з якого виготовлений досліджуваний зразок, величини дефекту, глибини залягання дефекту, товщини дослідного зразка, зовнішніх впливів тощо. Найчастіше у вказаному методі контролю інформативними параметрами виступають гармоніки вільних коливань. Але кожна гармоніка по-різному реагує на той чи інший дефект, має різний ступінь зашумлення. Для коректного аналізу інформативних параметрів і визначення виду дефекту необхідно правильно розділити вихідні дані, пов'язані сильною кореляцією, на суттєво впливові, значення яких суттєво залежать від стану об'єкта контролю, і залишкові (несуттєві), які значною мірою не визначаються станом об'єкта.

*Мета роботи.* Дослідження можливості зменшення розмірності вектора інформативних ознак за допомогою методу головних компонент (МГК) зі збереженням заданої надійності розпізнавання стану об'єкта під час контролю методом вільних коливань.

Розподіл простору вхідних даних на декілька векторів меншої розмірності, що будуть відокремлювати неінформативні складові, якими можна знехтувати, та інформативні параметри, які ха-

рактизують діагностичні ознаки різних типів (в залежності від виду дефекта), є важливою задачею для контролю технічного стану різного роду електричних, енергетичних та технічних досліджуваних об'єктів. Це забезпечить зменшення витрат на час діагностування і підвищить точність потрапляння відповідних інформативних даних в простір вектора діагностичних ознак, що відповідають відповідному технічному стану об'єкта.

Як досліджувані об'єкти можуть бути використані прилади, обладнання та деталі, виготовлені з композитних матеріалів: деталі літаків, електродвигуни, енергетичні установки, силові конструкції апаратів, що піддаються нагріванню, тощо.

### Теоретичні відомості

Для розділення простору вхідних даних, класифікації вхідних значень за конкретними параметрами та виділення найінформативніших параметрів існує багато методів, до прикладу, факторний аналіз, дискримінантний аналіз, методи розпізнавання образів, метод головних компонент.

Факторний аналіз застосовується для дослідження взаємозв'язків між значеннями змінних, які впливають на систему або об'єкт дослідження. Метою аналізу головних факторів є пошук і виділення спільного найвагомішого для усіх змінних фактора, тоді як аналіз головних компонент враховує всі фактори, які впливають на значення вхідних змінних. Цей метод використовується переважно для визначення структури даних. Дискримінантний аналіз розв'язує задачу пошуку змінних або функцій, які розділяють масив вхідних даних. Він застосовується за умови нормального розподілу незалежних змінних.

Методи розпізнавання образів поділяються на основні чотири групи [1]. У методі порівняння зі зразком (з еталоном) застосовується геометрична нормалізація і знаходиться найменша відстань до прототипу. Тобто досліджуваний зразок масштабується до розмірів існуючих прототипів і відбувається порівняння відстаней зразка до цих прототипів. Метод нейронних мереж реалізується шляхом оптимізації вагових коефіцієнтів, що відповідають за вибір характеристичної властивості, до якої буде віднесено досліджуваний зразок. Цей метод вимагає або великої кількості прикладів правильно розпізнаних образів, або розробку структури нейронної мережі, що враховує специфіку поставленої задачі. У статистичному методі використовуються статистичні характеристики отриманих даних. Він дозволяє здійснювати навчання системи розпізнавання на основі вирішальних правил, побудованих шляхом інформаційного аналізу і синтезу наявного простору ознак. Структурні та синтаксичні методи реалізуються шляхом розбиття досліджуваного об'єкта на елементи і подальшої побудови правил, в залежності від входження окремих елементів та їх послідовностей. Отже, для використання будь-якого з методів розпізнавання образів необхідно мати еталонну навчальну вибірку, яка будується на основі великого масиву вхідних експериментальних даних. Чим більша навчальна вибірка, тим краще будуть розпізнаватися невідомі об'єкти.

У [2] описані методи розділення простору ознак, одним з яких є метод потенціальних функцій, який оснований на пошуку відокремлювальних функцій або площин. Як відокремлювальна функція в цьому методі використана лінійна комбінація виду  $f(x) = \lambda x = 0$ , де  $\lambda$  — ваговий коефіцієнт, а  $x$  — значення, яким характеризується об'єкт. Це рівняння свідчить про те, що ваговий вектор  $\lambda$  перпендикулярний до відокремлювальної прямої. Отже, метод потенціальних функцій покладений в основу методу головних компонент і таку відокремлювальну пряму можна назвати головною компонентою.

Метод головних компонент — це один з поширених методів обробки даних, метою якого є зменшення розмірності досліджуваних даних з мінімальними втратами корисної інформації. Цей метод полягає у пошуку додаткової системи прямих або площин простору змінних, які будуть відображати максимальну дисперсію вихідних значень — така пряма або площина і називається головною компонентою. Основною перевагою методу головних компонент є те, що він є єдиним математично обґрунтованим методом факторного аналізу [3].

Алгоритм методу головних компонент здійснюється ітераційно і пояснюється таким чином [4]. На першому етапі формується матриця початкових даних  $X$  з розмірністю  $p \times p$ , де рядки матриці — це аналізовані зразки ( $i = 1, \dots, p$ ), а стовпці — початкові змінні  $x_{ij}$  ( $j = 1, \dots, p$ ), що характеризують ці зразки. Далі початкові дані необхідно центрувати за формулою  $\tilde{x}_{ij} = x_{ij} - \bar{x}_j$ , де  $\tilde{x}_{ij}$  — центроване значення  $i$ -го рядка  $j$ -го стовпця, а  $\bar{x}_j$  — середнє значення даних  $j$ -го стовпця. Насту-

пний етап — це знаходження першої головної компоненти за лінійною комбінацією  $pc_i = \sum_{i=1}^p \alpha_k \tilde{x}_{ij}$ ,

де  $\sum_{k=1}^n \alpha_k^2 = 1$  — коефіцієнти факторного відображення, що задовольняють умові ортогональності.

Така лінійна комбінація повинна мати найбільшу дисперсію серед усіх інших нормовано-центрованих лінійних комбінацій. Тобто повинна виконуватися умова  $|pc_i|^2 = \sum_{i=1}^p \alpha_k x_{ij} \rightarrow \max_{\alpha_k}$ .

Мірою інформативності факторів слугує відношення дисперсій головних компонент і вхідних даних.

У [5], [6] описано застосування МГК для графічної інтерпретації розбіжності класів вхідних даних, оскільки головні компоненти ортогональні між собою і показують напрям найбільшої дисперсії. ГК є лінійними комбінаціями вхідних змінних  $x_1, x_2 \dots x_N$ , що мають спільний нормальний розподіл з вектором середніх  $\bar{m} = (m_1, m_2 \dots m_N)$  і коваріаційною матрицею  $S$ . ГК можна записати у вигляді лінійної комбінації вхідних даних [6]

$$Y_1 = \sum_{j=1}^N \alpha_{1j} \cdot x_j; \quad Y_2 = \sum_{j=1}^N \alpha_{2j} \cdot x_j, \dots, Y_N = \sum_{j=1}^N \alpha_{Nj} \cdot x_j \quad (1)$$

і 
$$\text{cov}(Y_i, Y_j) = 0, \text{ коли } i, j = \overline{1, N}, i \neq j, \quad (2)$$

тобто ці змінні некорельовані. Вони впорядковані за спаданням дисперсії  $D(Y_i), i = \overline{1, N}$ .

Вектори  $\vec{a}_i = (a_{i1}, a_{i2} \dots a_{iN})$ ,  $i = \overline{1, N}$ , що визначають головні компоненти, є власними векторами, яким відповідають власні значення  $L_i$ , що дорівнюють дисперсії  $D(Y_i)$ . Загальна дисперсія не змінюється в результаті переходу від змінних  $x_1, x_2 \dots x_N$ , до головних компонент  $Y_1, Y_2 \dots Y_N$ , тобто

$$D = \sum_{i=1}^N \sigma_{1i} = \sum_{i=1}^N D(Y_i). \quad (3)$$

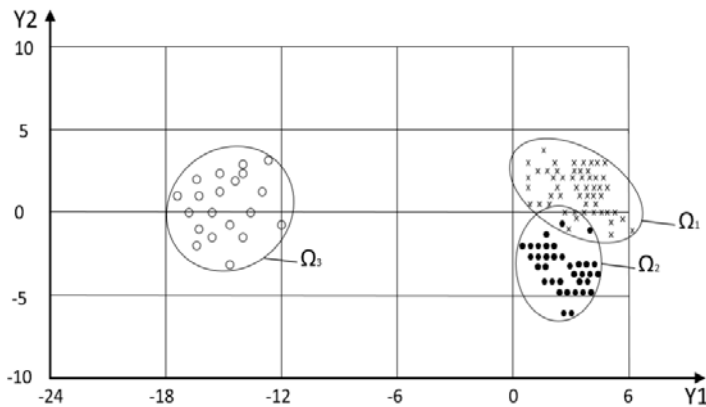


Рис. 1. Поділ вхідних даних на класи бездефектного ( $\Omega_3$ ) і дефектних ( $\Omega_1, \Omega_2$ ) сигналів

Кожна головна компонента змінних пояснює  $(100 \cdot D(Y_i)/D) \%$  загальної дисперсії. Тоді підмножина перших  $p < N$  змінних  $Y_i$  пояснює більшу частину загальної дисперсії і, таким чином виходить стислий опис структури залежності вхідних змінних [4]. Тому класи, які необхідно розділити, у [5] канонічно представлено за допомогою перших двох ГК  $Y_1$  і  $Y_2$  (що пояснюють найбільшу дисперсію усіх змінних), і проілюстровано графіком (рис. 1), оскільки початковий простір вхідних даних  $x_1, x_2 \dots x_N$ , як зазначено в літературі,

відносно високого порядку.

З рис. 1 випливає, що застосування МГК чітко дозволяє розділити класи даних бездефектного і дефектних сигналів. Але значення, отримані від об'єктів з дефектами різного роду, перетинаються.

У [6] описано алгоритм методу головних компонент, як знаходження декомпозиції вхідної матриці  $X$ , тобто розрахунок добутку двох матриць  $T$  і  $P$

$$X = T \cdot P^t + E = \sum_{a=1}^A t_a \cdot p_a^t + E, \quad (4)$$

де  $T$  — матриця рахунків,  $P$  — матриця навантажень,  $E$  — матриця залишків,  $t_a$  і  $p_a$  — число стовпців у матрицях  $T$  і  $P$  відповідно — дорівнює ефективному рангу матриці  $X$ . Ця величина позначається  $A$  і називається числом головних компонент, при чому  $A$  менше числа стовпців у вхідній

матриці  $X$ . Найбільші значення векторів навантажень відображають найвагомші ознаки вхідних даних, решта значень — залишкові, які в основному містять шуми. Залишкові дані відкидаються, зменшуючи розмірність простору вхідних даних, без втрати основних інформативних параметрів.

З вищеприведеного бачимо, що факторний аналіз застосовується для опису структури даних і виділення факторів, які впливають на формування такої структури. Дискримінантний аналіз використовується для визначення змінних, що впливають на розподіл даних і потребує виконання умови нормального розподілу незалежних змінних. Методи розпізнавання образів ефективні переважно за наявності великого масиву вхідних даних для побудови початкової навчальної вибірки. Тому в цій роботі пропонується застосування методу головних компонент для зменшення вектора вхідних даних і розподілу їх на класи (дефектний, бездефектний), оскільки він найбільше підходить для виконання поставленої задачі в умовах недостатньо великої кількості вхідних даних для формування еталонних навчальних вибірок.

На сьогоднішній день метод головних компонент застосовується в багатьох сферах науки і техніки, починаючи від визначення відхилень палеонапрямків у палеомагнетизмі [7] до оцінки стійкості перехідних процесів енергосистеми на основі кутової швидкості та активної потужності генератора [8].

Окрім того, метод головних компонент є основою в алгоритмі Eigenface [9] — розпізнавання облич за фотографією. У наведеній статті МГК використано для перенесення вхідних даних у новий простір, де найменш вагомі дані відкидаються, а впливовіші розподіляються на декілька базисів: зі спільними ознаками і з окремими специфічними ознаками.

### Застосування МГК у дослідженні спектрів вільних коливань

Спектри вільних коливань під час аналізу технічного стану об'єктів отримують шляхом перетворення Фур'є сигналу-відгука, що утворився під дією впливу різкого удару ззовні на об'єкт дослідження. Отже, досліджуваними сигналами можуть слугувати механічні коливання, що виникають при нанесенні удару бойком на контрольований об'єкт, які перетворюються в електричні. Також відгуки досліджуваного об'єкта на вплив коливання температури або зміни вібраційних параметрів чи хвильового опору матеріалів. Наприклад, сигнали, утворені під дією ударного вібраційного впливу на осердя асинхронного двигуна, що має шихтовану конструкцію, аналіз спектра таких сигналів дає інформацію про розшарування, ступінь пресування, пробій ізоляції між пластинами шихтованого магнітопроводу тощо.

Для реалізації дослідження використано сигнали, отримані під час неруйнівного контролю методом вільних коливань зразків вуглепластикових панелей. Вибрано два види сигналів: сигнали вільних коливань, знятих на ушкодженій зоні і на бездефектній зоні зразка. За допомогою перетворення Фур'є отримали спектри вказаних сигналів. Як інформативні параметри використано значення амплітуд гармонік спектрів зразків вуглепластикових панелей (рис. 2).

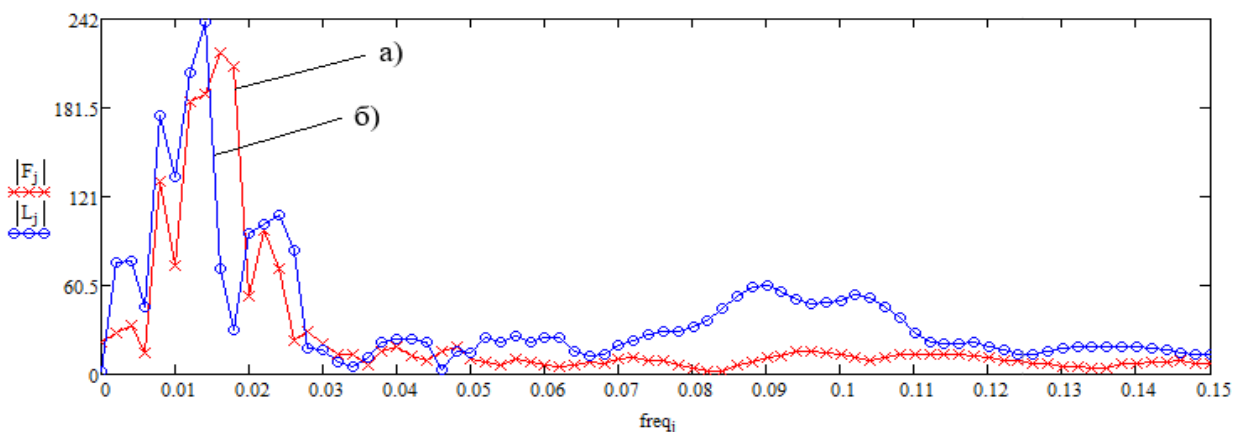


Рис. 2. Спектри сигналів з: а — бездефектної; б — дефектної зон зразків

За алгоритмом [10] першочергово необхідно сформувати матрицю вхідних даних. Для проведення дослідження сформовано матриці вхідних даних (табл. 1) з чотирьох значень для кожної амплітуди досліджуваних типів сигналів з бездефектного та пошкодженого зразків: 1) для 3-х гармонік; 2) для 5-ти гармонік; 3) для 10-ти гармонік.

Матриці вхідних даних

$a_{nm}$	$p = 3$		$p = 5$		$p = 10$	
	$i$	$j$	$i$	$j$	$i$	$j$
$a_{11}$	21,625	2,375	21,625	2,375	21,625	2,375
$a_{12}$	23,063	3,07	23,139	2,494	22,986	3,646
$a_{13}$	22,934	4,317	23,036	3,405	22,831	5,228
$a_{14}$	22,529	4,677	22,732	3,949	22,326	5,404
$a_{21}$	-10,834	-3,647	-10,834	-3,647	-10,834	-3,647
$a_{22}$	-10,623	-3,596	-11,592	-3,829	-9,653	-3,363
$a_{23}$	-10,579	-3,609	-11,085	-3,719	-10,072	-3,5
$a_{24}$	-8,934	-3,509	-9,82	-3,616	-8,049	-3,403
$a_{31}$	30,772	62,332	30,772	62,332	30,772	62,332
$a_{32}$	32,858	65,356	32,926	65,449	32,79	65,262
$a_{33}$	32,688	63,984	32,807	64,716	32,569	63,251
$a_{34}$	32,708	65,049	32,781	65,182	32,636	64,916
$a_{41}$			7,068	-12,032	7,068	-12,032
$a_{42}$			7,562	-12,634	7,207	-12,727
$a_{43}$			6,997	-13,012	5,866	-13,767
$a_{44}$			6,573	-13,276	5,585	-13,919
$a_{51}$			131,642	175,273	131,642	175,273
$a_{52}$			140,857	184,037	140,736	184,398
$a_{53}$			140,79	184,242	140,654	184,654
$a_{54}$			140,73	184,172	140,602	184,307
$a_{61}$					66,513	111,796
$a_{62}$					71,653	118,063
$a_{63}$					72,199	118,524
$a_{64}$					73,545	118,989
$a_{71}$					181,023	127,045
$a_{72}$					192,93	133,25
$a_{73}$					192,488	133,303
$a_{74}$					191,884	132,626
$a_{81}$					-78,432	-165,344
$a_{82}$					-82,612	-174,017
$a_{83}$					-83,514	-173,684
$a_{84}$					-81,21	-174,482
$a_{91}$					-49,657	-42,498
$a_{92}$					-52,274	-44,447
$a_{93}$					-51,645	-44,051
$a_{94}$					-52,865	-41,533
$a_{101}$					-136,795	-3,177
$a_{102}$					-148,234	-2,539
$a_{103}$					-148,123	-2,448
$a_{104}$					-147,766	-2,141

Матриця вхідних даних сформована зі значень:  $a_{nm}$ , де  $n = 1, 2, \dots, p$  — номер гармоніки ( $p = 3, 5$  і  $10$  — кількість гармонік),  $m = 1, 2, \dots, k$  — номер значення ( $k = 4$  — кількість значень), в залежності від впливових факторів: відсутність дефекту —  $i$  та наявність дефекту —  $j$ .

За алгоритмом знаходження ГК [10] після формування матриць вхідних даних проводимо центрування значень за формулою:  $\tilde{a}_{nm(i/j)} = a_{nm(i/j)} - \bar{a}_{nm(i/j)}$ .

За визначеними центрованими даними розраховуємо значення навантажень головних компонент, кількісний показник і власні значення ГК та значення дисперсій, які отримують після перенесення вхідних даних у простір визначених головних компонент.

### Результати дослідження

У проведеному експерименті визначено по дві головні компоненти (максимально можлива кількість компонент) для чотирьох значень за  $p$  кількістю гармонік. Для кожної кількості гармонік друга головна компонента пояснює максимальне значення дисперсій (табл. 2). Дисперсія, розрахована після перенесення вхідних даних у простір головних компонент, є наближеною оцінкою, що відображає найбільший розкид значень між двома класами.

Таблиця 2

## Значення дисперсій другої ГК для кожного дослідження

Кількість гармонік	3	5	10
Значення дисперсії за другою ГК, %	98,5	90,4	89,9

В розрахунку кожної головної компоненти кількість вихідних значень амплітуд зменшується вдвічі у порівнянні з кількістю вхідних значень. Це відображено у таблиці 3.

Таблиця 3

## Матриця вхідних даних, отриманих після застосування МГК

	$p = 3$		$p = 5$		$p = 10$	
	$i$	$j$	$i$	$j$	$i$	$j$
$a_{11}$	13,844	14,703	45,491	12,321	14,148	19,971
$a_{12}$	12,57	15,666	44,517	13,486	12,287	20,058
$a_{21}$	34,002	-11,441	69,203	-10,647	41,125	0,944
$a_{22}$	33,861	-11,276	69,791	-11,16	40,098	1,591
$a_{31}$	-43,69	-4,491	-8,679	-14,957	-35,156	-15,352
$a_{32}$	-47,331	-4,013	-12,468	-15,007	-38,662	-15,954
$a_{41}$			65,668	8,802	34,626	19,613
$a_{42}$			65,872	9,554	35,026	20,198
$a_{51}$			-159,173	1,835	-186,437	-22,023
$a_{52}$			-171,655	4,27	-199,319	-21,887
$a_{61}$					-95,52	-24,312
$a_{62}$					-103,595	-25,01
$a_{71}$					-186,401	47,002
$a_{72}$					-199,157	51,191
$a_{81}$					204,069	65,478
$a_{82}$					213,196	68,543
$a_{91}$					96,045	0,287
$a_{92}$					99,268	-0,226
$a_{101}$					128,753	-89,543
$a_{102}$					136,284	-98,177

Як міру відмінності між двома класами (бездефектного стану та дефектного стану) для аналізу інформативних параметрів найефективніше використати відстань Махаланобіса [4]. Чим більша відстань Махаланобіса (між значеннями), тим краще розрізняються дані, що належать різним класам. Цей показник розраховується за формулою

$$\mathbf{H} = (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2)^T \cdot \mathbf{M}^{-1} (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2), \quad (5)$$

де  $\boldsymbol{\mu}_1, \boldsymbol{\mu}_2$  — це вектори середніх значень змінних для класів  $X_1, X_2$ ;  $\mathbf{M}$  — об'єднана коваріаційна матриця для класів  $X_1, X_2$ .

В нашому випадку порівнювалися відстані між множинами гармонік вхідних даних і даних, отриманих після застосування МГК для сигналів отриманих в дефектній та бездефектній області зразка:

- 1) для 3-х гармонік: вхідні дані — 3,45; дані після застосування МГК —  $2,117 \cdot 10^9$ ;
- 2) для 5-ти гармонік: вхідні дані — 1,012; дані після застосування МГК —  $2,497 \cdot 10^{10}$ ;
- 3) для 10-ти гармонік: вхідні дані — 0,513; дані після застосування МГК —  $4,212 \cdot 10^{10}$ .

Таким чином, застосування методу головних компонент для опрацювання векторів інформативних параметрів дозволило збільшити відстань між значеннями, отриманими в дефектній і бездефектній області на декілька порядків. Отже, застосування МГК дозволяє підвищити надійність розпізнавання за амплітудними значеннями спектрів вільних коливань та зменшити розмірність простору вхідних даних (до трьох інформативних параметрів).

## Висновок

Проведене дослідження показує, що метод головних компонент можна застосовувати для аналізу спектрів складних сигналів з локально зосередженими параметрами, зокрема вільних коливань, як підготовчий етап аналізу вхідних даних. МГК дозволяє зменшити розмірність простору вхідних даних та водночас підвищити надійність розпізнавання стану об'єкта контролю.

## СПИСОК ВИКОРИСТАНОЇ ЛІТЕРАТУРИ

- [1] В. М. Коцовський, *Теорії розпізнавання образів*, лекції. Ужгород: *Методичні праці кафедри інформаційних управляючих систем та технологій*, 2014, 24 с. [Електронний ресурс]. Режим доступу: <https://dspace.uzhnu.edu.ua/jspui/handle/lib/16460>.
- [2] И. А. Биргер, *Техническая диагностика: Надежность и качество*. М.: Машиностроение, 1978, 240 с.
- [3] В. С. Бахрушин, *Методи аналізу даних: навчальний посібник для студентів*. Запоріжжя: КПУ, 2011, 268 с.
- [4] I. T. Jolliffe, and J. Cadima, "Principal component analysis: are view and recent developments," *Philosophical Transactions of the Royal Society*, A 374: 20150202, 2016. <https://doi.org/10.1098/rsta.2015.0202>.
- [5] З. П. Ненова, «Идентификация дефектов в ИИС вихретокового неразрушающего контроля.» дис., Киев, 1990, 282 с.
- [6] Steven M. Holland, *Principal components analysis (PCA)*. Department of Geology, University of Georgia, Athens, GA. 2019, 12 p. [Electronic resource]. Available: <http://strata.uga.edu/8370/handouts/pcaTutorial.pdf>.
- [7] Jeffrey J. Love, "Principal Component Analysis in Paleomagnetism," *Encyclopedia of Geomagnetism and Paleomagnetism*, no 1, pp. 845-850, 2007. [https://doi.org/10.1007/978-1-4020-4423-6\\_271](https://doi.org/10.1007/978-1-4020-4423-6_271).
- [8] Jingxuan Tang, and Huibin Sui, "Power System Transient Stability Assessment Based on PCA and Support Vector Machine," in *International Conference on Mechanical, Electrical, Electronic Engineering & Science MEEES2018*. Published by Atlantis Press, vol. 154, pp. 361-365, 2018. <https://doi.org/10.2991/meees-18.2018.63>.
- [9] Kevin Huang, "Principal Component Analysis in the Eigenface Technique for Facial Recognition," *Senior Theses and Projects*. Trinity College, Hartford, Connecticut. CT2012, 34 p. [Electronic resource]. Available: <https://digitalrepository.trincoll.edu/theses/216>.
- [10] L. I. Smith, "A tutorial on Principal Components Analysis," *Computer Science Technical Report*. University of Otago Library, Dunedin, New Zealand, no. OUCS-2002-12, 26 p., 2012. [Electronic resource]. Available: <http://hdl.handle.net/10523/7534>.

Рекомендована кафедрою метрології та промислової автоматики ВНТУ

Стаття надійшла до редакції 8.08.2022

**Еременко Володимир Станіславович** — д-р техн. наук, професор, завідувач кафедри інформаційно-вимірювальних технологій;

**Осинцева Марія Борисівна** — аспірантка кафедри інформаційно-вимірювальних технологій, e-mail: [mashaosintseva08@gmail.com](mailto:mashaosintseva08@gmail.com).

Національний технічний університет України «Київський політехнічний інститут імені Ігоря Сікорського», Київ

**V. S. Eremenko<sup>1</sup>**

**M. B. Osintseva<sup>1</sup>**

## Applying the Principal Component Analysis in the Problem of Analysis of the Specters of Free Oscillation

<sup>1</sup>National Technical University of Ukraine «Igor Sikorsky Kyiv Polytechnic Institute»

*This article addresses the main methods used for analysis, distribution, and classification of input data space. The main aspects of applying these methods are analyzed and determined. In this study it is found that the method of the principal components analysis (PCA) is the most suitable. Possible principal components analysis algorithms are described. A combination of these algorithms is used to distribute the input data in the analysis of signals and their spectra obtained by non-destructive testing after applying a free oscillations method. The article aims to study the possibility of reducing the vector of informative features after applying principal components analysis without losing the quality of recognition of the state of objects. The objects of study may be components of electric motors (shunt magnetic conductor), parts of aircraft made of composite materials and other structures that require analysis by non-destructive testing methods. Spectra taken during non-destructive testing by the method of free oscillations of samples of carbon fiber panels from defective and defect-free zones were studied. The analysis of three, five and ten harmonics was conducted and the maximum number of the principal components and the values of maximum dispersions of these components for the input values set of amplitudes were determined. The Mahalanobis distance was used to analyze the quality of the distribution of the input data space into classes (defect-free and defect zones of the sample). The quality of the separation of values from the damaged and undamaged zones of the sample into two classes was improved. Therefore, the use of principal component analysis, in this study, can increase the reliability of the recognition of the state of objects.*

**Keywords:** input data space distribution, informative parameters, principal component analysis, spectral analysis, non-destructive testing.

**Eremenko Volodymyr S.** — Dr. Sc. (Eng.), Professor, Head of the Chair of Information and Measuring Technologies;  
**Osintseva Mariia B.** — Post-Graduate Student of the Chair of Information and Measuring Technologies; e-mail: [mashaosintseva08@gmail.com](mailto:mashaosintseva08@gmail.com)