

Міністерство освіти і науки України
Харківський національний університет імені В.Н. Каразіна

Кваліфікаційна наукова
праця на правах рукопису

Берчанов Анатолій Андрійович

УДК 004.932.2 : 004.89

ДИСЕРТАЦІЯ
МЕТОД КОДУВАННЯ ІНФРАЧЕРВОНИХ ЗОБРАЖЕНЬ З
ВИКОРИСТАННЯМ МОДЕЛЕЙ ШТУЧНОГО ІНТЕЛЕКТУ В
ІНФОРМАЦІЙНИХ СИСТЕМАХ

Спеціальність 122 Комп'ютерні науки
Галузь знань 12 Інформаційні технології

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

_____ Берчанов А.А.

Науковий керівник: Бараннік Володимир Вікторович, доктор технічних наук,
професор

Харків – 2026

АНОТАЦІЯ

Берчанов А. А. Метод кодування інфрачервоних зображень з використанням моделей штучного інтелекту в інформаційних системах – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття ступеня доктора філософії за спеціальністю 122 Комп'ютерні науки (галузь знань 12 Інформаційні технології). Харківський національний університет імені В. Н. Каразіна, Харків, 2026.

У дисертаційній роботі вирішена актуальна *науково-прикладна задача*, що полягає у зниженні бітрейту інфрачервоного потоку даних для передачі в інформаційних системах з визначеним рівнем термальної цілісності.

Виявлено потребу зниження бітрейту інфрачервоних потоків даних під час їх передачі в інформаційних системах. Масштабування комплексів моніторингу на базі інфрачервоних сенсорів високої роздільної здатності створює надмірне бітове навантаження та, як наслідок, ускладнює своєчасну доставку даних. Це знижує їхню актуальність для систем прийняття рішень.

Проведено дослідження технічних характеристик сучасних інфрачервоних камер та параметрів генерації тепловізійного ПЧ-трафіку. Аналіз показав, що розширений динамічний діапазон та висока інтенсивність потоку даних призводять до суттєвого росту бітрейту. Водночас застосування існуючих універсальних стандартів стиснення з втратами знижує інформативність семантично важливих теплових сигнатур об'єктів інтересу.

Досліджено методи кодування інфрачервоних зображень, а також підходи до семантико-орієнтованої обробки з використанням моделей штучного інтелекту. Обґрунтовано доцільність застосування метричної оцінки інформативності сегментів у поєднанні з інтелектуальним виявленням об'єктів інтересу.

Сформульовано **мету дисертаційного дослідження**, яка полягає у розробці методу кодування інфрачервоних зображень (ПЧ-зображень) для

зниження бітрейту без втрат термальної цілісності на основі семантико-орієнтовної обробки з використанням моделей штучного інтелекту (ШІ).

Для досягнення поставленої мети, необхідно вирішити наступні завдання:

1. Розробити метод формування двошарової інформативної карти ПЧ-зображень для забезпечення умов щодо зменшення бітрейту з визначеною термальною цілісністю на основі використання моделей ШІ.

2. Створити метод спектрально-групового стиснення сегментів бітових пластів ПЧ-зображень для зменшення бітового об'єму без втрат термальної цілісності

3. Розробити метод семантико-орієнтованого кодування ПЧ-зображень для передачі в інформаційних системах.

4. Створити метод для підвищення термальної цілісності об'єктно-інформативних сегментів ПЧ-зображень без втрат бітрейту в інформаційних системах.

5. Здійснити порівняльну оцінку методів кодування ПЧ-зображень за абляційною методикою.

Об'єктом дослідження є процеси зменшення бітового об'єму ПЧ-кадрів в інформаційно-інтелектуальних системах.

Предметом дослідження є методи кодування інфрачервоних зображень для зменшення бітового навантаження каналів даних у інформаційно-інтелектуальних системах.

Проведено дослідження характеристик ПЧ-потоків, що генеруються сучасними пристроями. Виявлено недоліки при обробці ПЧ-кадрів існуючими методами кодування. Виявлені обмеження спеціальних ПЧ-кодеків та розповсюджених нейромережових підходів.

Розроблено метод біфуркації бітових пластів кадру із розширеним динамічним діапазоном. Метод дозволяє розділити інформаційну структуру зображення на значущі та фонові складові у рівнях окремих бітових пластів. Такий підхід дозволяє зменшити просторову надлишковість кадру.

Розроблено метод визначення метрики інформативності сегментів. Це дозволяє здійснювати кількісну оцінку інформативності просторових областей на основі їхньої текстурної насиченості.

Розроблено метод формування двопластової інформативної карти інфрачервоного кадру. Метод дозволяє виокремити пріоритетні області обробки у порогово-метричному просторі старшого та молодшого бітових пластів.

Розроблено метод виявлення об'єктно-інформативних сегментів на основі застосування моделей штучного інтелекту. Це дозволяє сегментувати об'єкти інтересу через кон'юнкцію результату нейронної мережі із сформованою картою інформативності. Такий підхід дозволяє обмежити рівень спотворень у пріоритетних областях кадру.

Розроблено метод різницевого кодування для формування мінісегментів зі зменшеним динамічним діапазоном. Це дозволяє трансформувати значення пікселів ІЧ-кадру у локальні температурні відхилення.

Розроблено метод корекції низькочастотних коефіцієнтів на основі високочастотних компонент вейвлет-перетворення. Метод дозволяє звужити динамічний діапазон низькочастотних значень при цьому не вносячи додаткових спотворень.

Розроблено метод адаптивної квантизації на основі рівня інформативності та бітового пласту сегментів. Метод дозволяє диференційовано розподіляти коефіцієнти квантизації залежно від пріоритетності сегментів. Це рішення дозволяє обмежувати бітрейт ІЧ-поток при встановленому рівні термальної цілісності.

Розроблено метод спектрально-групового кодування сегментів. Метод дозволяє здійснювати структурне кодування спектральних значень за типами частотних груп в умовах квадрореконструкції. Застосування цього підходу звужує обсяг бітового надлишку та зменшує інтенсивність потоку даних.

Проведено абляційний аналіз впливу розроблених методів на кодування ІЧ-кадрів. Відображено, що комплексна інтеграція запропонованих методів кумулятивно знижує бітовий об'єм даних зберігаючи цілісність інформативних областей.

Наукова новизна отриманих результатів досліджень.

1. *Удосконалено* метод побудови двопластової інформативної карти на основі порогових вирішальних правил. Відмінності методу полягають у визначенні адаптивних асоціативних залежностей в порогово-метричному просторі на основі агрегації та нормування інформації високочастотної області перетворення Хаара для сегментованих бітових пластів ІЧ-зображень. Це дозволяє створити умови для збільшення рівня стиснення при встановлених вимогах щодо термальної цілісності ІЧ-зображень.

2. *Отримано подальший розвиток* методу виявлення областей об'єктів інтересу на основі застосування моделі штучного інтелекту. Відмінності методу полягають у виявленні об'єктно-інформативних сегментів за рахунок кон'юнкції результату нейронної мережі із сформованою інформативною картою. Це дозволяє зберегти термальну цілісність об'єктів інтересу за умов обмеженого бітрейту.

3. *Вперше розроблено* метод спектрально-групового стиснення на основі структурних перетворень у хвильовому просторі Хаара. Відмінності методу полягають у: диференційованому структурному кодуванні хвильового простору сегментів за частотними областями в умовах квадрореконструкції та встановленням нижньої межі високочастотної області. Це дозволяє зменшити бітовий об'єм пластових сегментів ІЧ-зображень без втрат термальної цілісності.

4. *Отримано подальший розвиток* методу кодування ІЧ-зображень на основі семантико-орієнтовної обробки. Відмінності методу полягають у спектрально-груповому кодуванні з врахуванням наявності об'єктно-інформативних сегментів ІЧ-зображень в біфуркаційно-хвильовому просторі Хаара. Це дозволяє знизити бітрейт інфрачервоного потоку даних з

визначеним рівнем термальної цілісності в процесі обробки та передачі в інформаційних системах.

5. *Удосконалено* метод відновлення стиснутих ІЧ-зображень на основі використання бігармонічних функцій. Відмінності методу полягають у просторовому прорідженні низькоінформативних областей бітових пластів ІЧ-зображень з їх регенерацією за регулярно-просторовим шаблоном в процесі відновлення з використанням бігармонічних функцій. Це дозволяє підвищити рівень термальної цілісності інформативних сегментів ІЧ-зображень без втрат бітрейту.

Практичне значення отриманих результатів досліджень.

Впровадження розробленого методу кодування ІЧ-зображень у інформаційних системах дозволяє зменшити бітрейт потоку теплових кадрів в межах заданої термальної цілісності, тобто:

- на 8% для кадрів з переважним вмістом низькоінформативних фонових сегментів;
- на 10% для кадрів з переважним вмістом інформативних сегментів, що містять термальні контрасти;
- на 7% для кадрів з переважним вмістом об'єктно-інформативних сегментів, що насичені тепловими сигнатурами об'єктів інтересу.

Для розробленого методу відносно існуючих досягається перевага у застосуванні штучного інтелекту. Це дозволяє підвищити SSIM для об'єктних областей на 93% у порівнянні із існуючими методами. Що задовольняє вимогам по збереженню цілісності теплових полів об'єктів інтересу.

Використання штучного інтелекту у поєднанні із метричним визначенням класу інформативності сегментів дає змогу застосовувати при кодуванні бітових пластів селективні коефіцієнти квантизації. Це дозволяє регулювати PSNR від 24 дБ до 38 дБ, що надає перевагу у коефіцієнті стиснення у 4-10% над існуючими методами.

Бігармонічна реконструкція низькоінформативних областей ІЧ-кадру дозволяє збільшити коефіцієнт стиснення на 6%, при цьому зменшуючи середнє квадратичне відхилення у локальних областях.

Ключові слова: штучний інтелект, машинне навчання, нейронні мережі, інформаційна безпека, модель захищеного обміну даними, семантична сегментація, криптостійкість, виявлення об'єктів, вейвлет-перетворення, обчислювальна складність, кодування зображень, стиснення зображень, комп'ютерний зір, інфрачервоні зображення, цифрова обробка зображень.

ABSTRACT

Berchanov Anatolii. Method of encoding infrared images using artificial intelligence models in information systems – Qualification scientific work on the rights of the manuscript.

Thesis for the degree of Doctor of Philosophy in the specialty 122 Computer Science (field of knowledge 12 Information technologies). V. N. Karazin Kharkiv National University, Kharkiv, 2026.

In the dissertation, a *scientific and applied problem* is solved related to reducing the bit rate of the infrared data stream for transmission in information systems with a certain level of thermal integrity.

The need to reduce the bit rate of infrared data streams during their transmission in information systems has been identified. Scaling monitoring complexes based on high-resolution infrared sensors creates excessive bit load and, as a result, complicates timely data delivery. This reduces their relevance for decision-making systems.

A study of the technical characteristics of modern infrared cameras and parameters for generating thermal IR traffic has been conducted. The analysis has shown that the extended dynamic range and high intensity of the data stream lead to a significant increase in the bit rate. At the same time, the use of existing

universal lossy compression standards reduces the informativeness of semantically important thermal signatures of objects of interest.

Infrared image coding methods have been investigated, as well as approaches to semantically-oriented processing using artificial intelligence models. The feasibility of using a metric assessment of the informativeness of segments in combination with intelligent detection of objects of interest has been substantiated.

The goal of the dissertation research is formulated, which is to develop a method for encoding infrared images (IR images) to reduce the bit rate without losing thermal integrity based on semantic-oriented processing using artificial intelligence (AI) models.

To achieve this goal, it is necessary to solve the following tasks:

1. Develop a method for forming a two-layer informative map of IR images to ensure conditions for reducing the bit rate with a certain thermal integrity based on the use of AI models.

2. Create a method for spectral-group compression of segments of bit layers of IR images to reduce the bit volume without losing thermal integrity

3. Develop a method for semantic-oriented encoding of IR images for transmission in information systems.

4. Create a method for increasing the thermal integrity of object-informative segments of IR images without losing bit rate in information systems.

5. To carry out a comparative assessment of methods for encoding IR images using the ablation technique.

The object of the research is the processes of reducing the bit volume of IR frames in information and intelligent systems.

The subject of the research is methods of encoding infrared images to reduce the bit load of data channels in information and intelligent systems.

The characteristics of IR streams generated by modern devices have been studied. Deficiencies in processing IR frames using existing encoding methods have been identified. Limitations of special IR codecs and common neural network approaches have been identified.

A method of bifurcation of bit layers of a frame with an extended dynamic range has been developed. The method allows dividing the information structure of an image into significant and background components at the levels of individual bit layers. This approach allows reducing the spatial redundancy of the frame.

A method for determining the metric of segment information has been developed. This allows for a quantitative assessment of the information content of spatial regions based on their texture saturation.

A method for forming a two-layer information map of an infrared frame has been developed. The method allows for the identification of priority processing areas in the threshold-metric space of the senior and junior bit layers.

A method for detecting object-information segments based on the use of artificial intelligence models has been developed. This allows for segmenting objects of interest through the conjunction of the result of a neural network with the formed information content map. This approach allows for limiting the level of distortion in priority regions of the frame.

A method for differential coding has been developed for the formation of mini-segments with a reduced dynamic range. This allows for the transformation of the values of IR frame pixels into local temperature deviations.

A method for correcting low-frequency coefficients based on high-frequency components of the wavelet transform has been developed. The method allows to narrow the dynamic range of low-frequency values without introducing additional distortions.

An adaptive quantization method has been developed based on the level of information content and bit layer of segments. The method allows to differentially distribute quantization coefficients depending on the priority of segments. This solution allows to limit the bit rate of the IR stream at a set level of thermal integrity.

A method of spectral-group coding of segments has been developed. The method allows to carry out structural coding of spectral values by types of frequency groups under conditions of quadrodecomposition. The use of this

approach narrows the amount of bit excess and reduces the intensity of the data stream.

An ablative analysis of the impact of the developed methods on the coding of IR frames was carried out. It was shown that the complex integration of the proposed methods cumulatively reduces the bit volume of data while preserving the integrity of informative areas.

Scientific novelty of the main research results.

1. The method for constructing a two-layer informative map based on threshold decision rules was improved. The differences of the method lie in the definition of adaptive associative dependencies in the threshold-metric space based on the aggregation and normalization of information of the high-frequency region of the Haar transform for segmented bit layers of IR images. This allows creating conditions for increasing the compression level with the established requirements for the thermal integrity of IR images.

2. Further development of the method for detecting areas of objects of interest based on the application of an artificial intelligence model was obtained. The differences of the method are in the detection of object-informative segments due to the conjunction of the result of the neural network with the formed informative map. This allows to preserve the thermal integrity of the objects of interest under conditions of limited bitrate.

3. For the first time, a spectral-group compression method was developed based on structural transformations in the Haar wave space. The differences of the method are: differentiated structural coding of the wave space of segments by frequency domains under conditions of quadrodecomposition and setting the lower limit of the high-frequency domain. This allows to reduce the bit volume of layer segments of IR images without loss of thermal integrity.

4. Further development of the IR image coding method based on semantic-oriented processing was obtained. The differences of the method lie in spectral-group coding taking into account the presence of object-informative segments of IR images in the Haar bifurcation-wave space. This allows to reduce the bit rate of

the infrared data stream with a certain level of thermal integrity during processing and transmission in information systems.

5. The method of restoring compressed IR images based on the use of biharmonic functions has been improved. The differences of the method lie in the spatial thinning of low-informative areas of bit layers of IR images with their regeneration according to a regular spatial template in the process of restoration using biharmonic functions. This allows to increase the level of thermal integrity of informative segments of IR images without loss of bit rate.

Practical significance of the research results obtained.

Implementation of the developed method of encoding IR images in information systems allows to reduce the bit rate of the thermal frame stream within the given thermal integrity, i.e.:

- by 8% for frames with a predominant content of low-informative background segments;
- by 10% for frames with a predominant content of informative segments containing thermal contrasts;
- by 7% for frames with a predominant content of object-informative segments, saturated with thermal signatures of objects of interest.

For the developed method, an advantage is achieved in the use of artificial intelligence over existing ones. This allows to increase SSIM for object areas by 93% compared to existing methods. Which satisfies the requirements for preserving the integrity of thermal fields of objects of interest.

The use of artificial intelligence in combination with the metric definition of the information class of segments makes it possible to use selective quantization coefficients when encoding bit layers. This allows you to adjust the PSNR from 24 dB to 38 dB, which provides an advantage in the compression ratio of 4-10% over existing methods.

Biharmonic reconstruction of low-information regions of the IR frame allows to increase the compression ratio by 6%, while reducing the mean square deviation in local regions.

Keywords: artificial intelligence, machine learning, neural networks, information security, secure data exchange model, semantic segmentation, cryptoresistance, object detection, wavelet transform, computational complexity, image coding, image compression, computer vision, infrared images, digital image processing.

Список наукових публікацій здобувача

Публікації, в яких опубліковано основні наукові результати дисертації:

Статті у наукових фахових виданнях, що входять до міжнародних наукометричних баз:

1. Barannik V., Berchanov A., Barannik V., Dihtiar M., Osovytskyi M., Sushko A., Babenko Y. Positional coding method in differential wave space. *Informatyka, Automatyka, Pomiarы w Gospodarce i Ochronie Środowiska*. 2026. Vol. 16. No 1. P. 138–146. DOI: <https://doi.org/10.35784/iapgos.7776> (Scopus, Q4).

2. Бараннік В. В., Берчанов А. А., Бараннік В. В., Суханов О. Ю., Перцев П. Д., Лиходєєв О. С., Юдін О. К., Прокопенко Р. О., Чаун Ю. М. Метод семантично-орієнтованої обробки бітових площин інфрачервоних зображень. *Вісник НТУУ "КПІ". Серія Радіотехніка, Радіоапаратобудування*. 2025. № 102. С. 15–24. DOI: <https://doi.org/10.20535/RADAP.2025.102.15-24>. URL: <https://radap.kpi.ua/radiotechnique/article/view/2084> (дата звернення: 30.01.2026) (Web of Science, Q4).

3. Бараннік В. В., Берчанов А. А., Бараннік В. В., Шейгас О. М., Перцев П. Д., Вовк О. В., Бабенко Ю. М., Прокопенко Р. О., Ерошенко В. П. Метод групового кодування інфрачервоних зображень у спектрально-хвильовому просторі. *Вісник НТУУ "КПІ". Серія Радіотехніка, Радіоапаратобудування*.

2025. № 99. С. 24–34. DOI: <https://doi.org/10.20535/radap.2025.99.24-34> (Web of Science, Q4).

Статті у наукових фахових виданнях України:

4. Бараннік В. В., Берчанов А. А. Метод біфуркації бітових площин в системах обробки інфрачервоних зображень. *Наукоємні технології*. 2025. Т. 66, № 2. С. 208–213. DOI: <https://doi.org/10.18372/2310-5461.66.20325>.

5. Бараннік В., Берчанов А. Метод спектрально-групового кодування бітових шарів інфрачервоних зображень. *Інфокомунікаційні технології та електронна інженерія*. 2025. Вип. 5, № 2. С. 15–24. DOI: <https://doi.org/10.23939/ictee2025.02.015>.

6. Берчанов А. А., Бараннік В. В. Метод обробки інфрачервоних зображень у спектральному домені. *Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: Технічні науки*. 2025. Т. 36 (75), № 3. С. 55–60. DOI: <https://doi.org/10.32782/2663-5941/2025.3.2/08>.

Наукові праці, які засвідчують апробацію матеріалів дисертації:

7. Barannik V., Eliseev E., Barannik V., Berchanov A., Lenets V., Fedorovskiy O. Method of Cluster-Driven Spectral-Parametric Transformant Coding for Adaptive Compression of Segmented Video Data. *2026 IEEE 18th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)* (Lviv, Ukraine, 17–21 Feb., 2026). Lviv, 2026. P. 1–6. DOI: <https://doi.org/10.1109/TCSET65181.2026.11461085> (Scopus).

8. Barannik V., Berchanov A., Kurlan O., Osovytskyi M., Pertsev P., Prokopenko R. Method of multilevel spectral processing of infrared images in bit planes. *Cyber Hygiene & Conflict Management in Global Information Networks (CH&CMiGIN'25)*: Proceedings of the International Workshop (Kyiv, Ukraine,

20–22 June, 2025). 2025. Vol. 4024. P. 227–235. URL: <https://ceur-ws.org/Vol-4024/short07.pdf> (дата звернення: 30.01.2026) (Scopus).

9. Бараннік В. В., Берчанов А. А., Бараннік В. В., Перцев П. Д., Курлан О. О. Підвищення ефективності доведення інфрачервоних знімків методом спектрально-хвильового кодування. *Кіберборотьба: розвідка, захист та протидія: матеріали III Міжнар. наук.-практ. конф.* (Київ, 6–7 трав. 2025 р.). Київ, 2025. С. 14–15. DOI: <https://doi.org/10.61929/viti.mnpk.3.2025>.

10. Kurlan O., Prokopenko R., Pertsev P., Manakov V., Berchanov A., Shchyrba V. A Method of Effective Coding of a Dynamic Video Information Resource in an Information and Telecommunications Network. *2024 IEEE 5th International Conference on Advanced Trends in Information Theory (ATIT)* (Lviv, Ukraine, 21–23 Nov., 2024). Lviv, 2024. P. 51–57. DOI: <https://doi.org/10.1109/ATIT64324.2024.11222407> (Scopus).

11. Берчанов А. А., Бараннік В. В., Прокопенко Р. О., Бараннік Д. В., Перцев П. Д. Інтелектуалізовані методи обробки тепловізійних зображень. *Комп'ютерне моделювання в наукоємних технологіях (КМНТ-2024): матеріали 10-ї Міжнар. наук.-техн. конф.* (Харків, 27–29 листоп. 2024 р.). Харків, 2024. С. 24–26. URL: <https://csai.karazin.ua/news/uk/x-naukovo-tehnichna-mizhnarodna-28052025> (дата звернення: 05.02.2026).

12. Barannik V., Berchanov A., Barannik V., Babenko M., Onyshchenko R., Kolodiichuk L. Method of Mini Segments Encoding in Difference Space Using Haar Wavelet. *2023 IEEE 5th International Conference on Advanced Information and Communication Technologies (AICT)* (Lviv, Ukraine, 21–25 Nov., 2023). Lviv, 2023. P. 1–4. DOI: <https://doi.org/10.1109/AICT61584.2023.10452674> (Scopus).

13. Бараннік В. В., Берчанов А. А. Перспективи розвитку технологій обробки відеоінформації в умовах інформаційно-інтелектуального протистояння. *Наукоємні технології в інфокомунікаціях (НІСТ'2023): матеріали Міжнар. наук.-практ. конф.* (Кам'янець-Подільський, 1–3 черв. 2023 р.). Кам'янець-Подільський, 2023. URL:

<http://repositsc.nuczu.edu.ua/bitstream/123456789/18013/1/%D0%9D%D0%86%D0%A1%D0%A2%202023.pdf> (дата звернення: 30.01.2026).

14. Krasnorutsky A., Kolesnyk V., Berchanov A., Barannik V., Kharchenko N., Malko O. Method of Structural-Statistical Coding of Video Segments in Spectral-Cluster Space. *2022 IEEE 4th International Conference on Advanced Trends in Information Theory (ATIT)* (Kyiv, Ukraine, 15–17 Dec., 2022). Kyiv, 2022. P. 32–37. DOI: <https://doi.org/10.1109/ATIT58178.2022.10024240> (Scopus).

ЗМІСТ

ВСТУП	19
РОЗДІЛ 1. ОБҐРУНТУВАННЯ НЕОБХІДНОСТІ УДОСКОНАЛЕННЯ ІНФОРМАТИВНИХ ТЕХНОЛОГІЙ ОБРОБКИ ІЧ-ЗОБРАЖЕНЬ	28
1.1 Аналіз сфер застосування інфрачервоних зображень та їх ролі в інформаційних системах широкого призначення	29
1.2 Дослідження технічних характеристик інфрачервоних камер як базової компоненти формування даних	36
1.3 Аналіз можливостей і обмежень методів кодування інфрачервоних зображень	46
1.4 Аналіз доцільності застосування нейромережових технологій для підвищення ефективності обробки інфрачервоних зображень	60
Висновки за першим розділом	70
РОЗДІЛ 2. РОЗРОБКА МЕТОДУ ПОБУДОВИ ДВОПЛАСТОВОЇ ІНФОРМАТИВНОЇ КАРТИ ІНФРАЧЕРВОНОГО КАДРУ НА ОСНОВІ МОДЕЛЕЙ СТАТИСТИЧНО-ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ	73
2.1 Обґрунтування принципів семантико-орієнтованої обробки та підходів до сегментації інфрачервоних кадрів	74
2.2 Обґрунтування підходів до обробки 16-бітних даних та аналіз методів HDR-перетворень	77
2.3 Розробка методу біфуркації бітових пластів для підвищення структурної сумісності даних при подальшій обробці	79
2.4 Аналіз та обґрунтування вибору методу трансформації просторових даних у спектральний простір	84
2.5 Впровадження методу формування спектрального простору на основі вейвлет-перетворення Хаара	90

2.6 Розробка методу визначення метрики інформативності на основі структурно важливих ознак сегментів	94
2.7 Розробка методу формування двопластової інформативної карти інфрачервоного кадру	98
2.8 Розробка методу виявлення об'єктно-інформативних сегментів на основі застосування моделей штучного інтелекту	104
Висновки за другим розділом	109
РОЗДІЛ 3. ОБҐРУНТУВАННЯ ТА РОЗРОБКА МЕТОДУ КОДУВАННЯ ІНФРАЧЕРВОНИХ ЗОБРАЖЕНЬ У ІНФОРМАЦІЙНИХ СИСТЕМАХ	111
3.1 Розробка методу ієрархічної сегментації інфрачервоного кадру з урахуванням його структурно-статистичних характеристик	112
3.2 Розробка методу різницевого кодування для формування мінісегментів зі зменшеним динамічним діапазоном	115
3.3 Розробка методу перетворення сегментів у спектрально-різницевий простір	118
3.4 Розробка методу корекції низькочастотних коефіцієнтів на основі високочастотних компонент вейвлет-перетворення	124
3.5 Розробка методу адаптивної квантизації на основі рівня інформативності та бітового пласту сегментів	127
3.6 Розробка методу спектрально-групового кодування для зменшення бітової інтенсивності кадрів.....	132
3.7 Розробка методу кодування інфрачервоних зображень на основі семантико-орієнтованої обробки сегментів	139
Висновки за третім розділом.	144
РОЗДІЛ 4. ЕКСПЕРИМЕНТАЛЬНА ОЦІНКА ХАРАКТЕРИСТИК ЗАПРОПОНОВАНОГО МЕТОДУ КОДУВАННЯ.....	146

4.1 Розробка методу просторового прорідження малоінформативних областей та їх бігармонічної реконструкції.....	147
4.2 Розробка моделі оцінки ефективності методу кодування інфрачервоних зображень	160
4.3 Проведення абляційного аналізу впливу структурно-спектральних етапів кодування на бітову інтенсивність та якість відновлення інфрачервоних кадрів	166
4.4 Порівняльна оцінка бітрейту та характеристик стиснення ІЧ-зображень запропонованого методу кодування з відомими підходами....	171
Висновки за четвертим розділом.....	181
ВИСНОВКИ.....	183
ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ	188
Додаток А Список публікацій здобувача за темою дисертації та відомості про апробацію результатів дисертації	211
Додаток Б Акти реалізації науково-прикладних результатів досліджень	215
Додаток В Науково-дослідна робота	219

ВСТУП

Обґрунтування вибору теми дослідження. Сучасний етап розвитку інформаційних технологій характеризується інтеграцією сенсорних мереж у критично важливі сфери людської діяльності. Ця тенденція супроводжується переходом до методів збору та обробки даних у реальному часі. Важливим елементом таких систем стають сенсори, що працюють у інфрачервоному діапазоні спектра. Вони забезпечують всепогодну та цілодобову візуалізацію випромінювання теплових полів об'єктів та навколишнього середовища. Значного поширення ІЧ-камери здобули у наступних сферах: дистанційний моніторинг промислових та інфраструктурних об'єктів, енергоаудит, медична діагностика. Okремо варто звернути увагу на напрямок інтеграції ІЧ-сенсорів у системи «розумного міста» для забезпечення ситуаційної обізнаності. Проте найбільш перспективним є використання ІЧ-каналів у складі дистанційно керованих та автономних мобільних платформ. До них належать безпілотні літальні апарати (БПЛА) та роботизовані комплекси. У таких системах існує потреба у трансляції потоку даних в режимі реального часу. Це необхідно для: верифікації рішень бортової автоматики, дистанційного керування в умовах невизначеності, забезпечення ситуаційної обізнаності.

Сучасні тепловізійні матриці генерують потоки даних із розрядністю 14–16 біт на піксель. Це необхідно для фіксації широкого динамічного діапазону температурних градієнтів. Крім цього, зростає просторова роздільна здатність сенсорів. Відповідно, виникає збільшення бітрейту потоків даних. Така інтенсивність може перевищувати пропускну здатність каналів передачі інформації. Як наслідок, великі об'єми інформації можуть приходити із затримками або втрачатись. Це є неприпустимим при потребі обробки даних у реальному часі. Наприклад у режимах роботи наведення чи контролю.

В результаті аналізу виявлено науково-технічне протиріччя. Воно полягає у невідповідності між необхідністю зниження бітрейту потоку інфрачервоних даних та вимогами до збереження теплової семантики об'єктів інтересу. Отже, зниження бітрейту ІЧ-поток даних для передачі в інформаційних системах з визначеним рівнем термальної цілісності є **актуальною науково-прикладною задачею**.

Значний внесок у розвиток теорії і методів обробки видової інформації для зменшення часу доставки та об'єму даних зробили багато науковців, серед них: Бараннік В.В., Жураковський Б.Ю., Корольов А.В., Лукін В.В., Рубан І.В., Русин Б.П. [1-27, 29-32, 40-45, 52, 57-59, 62, 63, 69, 82, 98-109]. Із зарубіжних дослідників вагомий внесок зробили такі вчені як Anil K., Ballé J., Bovik A., Cover T., Gonzalez R., LeCun Y., Malik J., Shannon C., Szeliski R., Wiegand T. [39, 92, 113, 150].

Аналіз існуючих підходів до стиснення візуальних зображень виявляє їхні обмеження стосовно специфіки інфрачервоних даних. Такі стандарти стиснення з втратами базуються на психовізуальних моделях сприйняття, орієнтованих на людський зір. Вони видаляють високочастотні компоненти сигналу, які око сприймає як шум. Проте, для машинного алгоритму ці дані можуть містити критично важливі ознаки. Наприклад такі: деталі текстури, мікrokонтрасти, теплові сигнатури малорозмірних об'єктів. Застосування таких методів призводить до неконтрольованого спотворення термальної цілісності. Це може спричиняти наступні ефекти: зниження ймовірності правильного виявлення об'єктів, помилки класифікації, порушення роботи систем наведення або навігації.

Перспективним напрямом вважається використання методів глибокого навчання для побудови кодеків нового покоління. Зокрема, згорткові та варіаційні автокодувальники забезпечують ефективне зниження розмірності даних шляхом їх перетворення у компактне латентне представлення. Для відновлення високочастотних деталей текстур застосовуються генеративні мережі. В тому числі такі як: генеративно-змагальні мережі (GAN), дифузійні

моделі. Це дозволяє забезпечувати візуальну якість кадрів при суттєвому рівні стиснення. Однак інтеграція цих підходів у бортові системи стикається з обмеженнями. По-перше, генеративна природа відновлення зображень створює ризик виникнення «галюцинацій». Тобто синтезу правдоподібних, але фізично відсутніх об'єктів. Це є неприпустимим у системах відповідального призначення: оборона, безпека, промислова діагностика. По-друге, висока обчислювальна складність глибоких нейромереж, особливо генеративних, часто перевищує можливості мобільних енергоефективних процесорів безпілотних платформ.

Сформульовану задачу пропонується вирішувати шляхом переходу від енергетичного до семантико-орієнтованого підходу кодування. Інфрачервоні сцени характеризуються високою семантичною нерівномірністю. Це дозволяє перерозподілити бітовий ресурс на користь областей інтересу за рахунок агресивної компресії фонових ділянок. Для цього доцільне застосування моделей штучного інтелекту. Що дозволить сегментувати кадр на основі структурно-семантичного аналізу. Крім цього, статистично-енергетичні методи потребують удосконалення. Що дозволить застосовувати їх для семантичного аналізу сцени. Такий підхід забезпечує баланс між рівнем стиснення та збереженням термальної точності. Звідси **тематика дисертаційних досліджень**, що полягає в розробці методу кодування інфрачервоних зображень з використанням моделей штучного інтелекту в інформаційних системах, є актуальною.

Зв'язок роботи з науковими програмами, планами, темами.

Робота виконана у рамках: Розпорядження Кабінету Міністрів України «Про затвердження плану заходів з реалізації Концепції розвитку штучного інтелекту в Україні на 2025-2026 роки» від 9 травня 2025 р. № 457-р; розпорядження Кабінету Міністрів України «Про схвалення Концепції розвитку штучного інтелекту в Україні» від 2 грудня 2020 р. № 1556-р; Закону України «Про електронні комунікації» від 30.09.2020 р. № 1089-IX; Закону України «Про Національну програму інформатизації» від 01.12.2022

р. № 2807-IX; Розпорядження Кабінету Міністрів України «Про затвердження плану заходів на 2025 рік з реалізації Стратегії кібербезпеки України» від 7 березня 2025 р. № 204-р; Розпорядження Кабінету Міністрів України «Про затвердження плану заходів на 2023–2025 роки з реалізації Стратегії кібербезпеки України; пріоритетного напрямку розвитку науки і техніки «Інформаційні та комунікаційні технології». Дисертаційні дослідження проводились у відповідності із планами науково-технічної діяльності Харківського національного університету імені В. Н. Каразіна. Основні результати дисертаційної роботи відображені в проміжному звіті по НДР «Розробка математичних моделей обробки видової інформації для підвищення якості функціонування дистанційних сервісів» (№ 0125U004039), в якій автор є відповідальним виконавцем.

Мета і завдання досліджень. Мета досліджень полягає у розробці методу кодування інфрачервоних зображень (ІЧ-зображень) для зниження бітрейту без втрат термальної цілісності на основі семантико-орієнтовної обробки з використанням моделей штучного інтелекту (ШІ).

Для досягнення поставленої мети, необхідно вирішити наступні завдання:

1. Розробити метод формування двошарової інформативної карти ІЧ-зображень для забезпечення умов щодо зменшення бітрейту з визначеною термальною цілісністю на основі використання моделей ШІ.

2. Створити метод спектрально-групового стиснення сегментів бітових пластів ІЧ-зображень для зменшення бітового об'єму без втрат термальної цілісності

3. Розробити метод семантико-орієнтованого кодування ІЧ-зображень для передачі в інформаційних системах.

4. Створити метод для підвищення термальної цілісності об'єктно-інформативних сегментів ІЧ-зображень без втрат бітрейту в інформаційних системах.

5. Здійснити порівняльну оцінку методів кодування ПЧ-зображень за абляційною методикою.

Об'єкт дослідження. Процеси зменшення бітового об'єму ПЧ-кадрів в інформаційно-інтелектуальних системах.

Предмет дослідження. Методи кодування інфрачервоних зображень для зменшення бітового навантаження каналів даних у інформаційно-інтелектуальних системах.

Методи дослідження. Методи теорії інформації та кодування. Методи штучного інтелекту та машинного навчання. Методи цифрової обробки зображень. Методи штучного інтелекту та машинного навчання. Методи глибокого навчання. Методи згорткових нейронних мереж (CNN) та архітектури YOLO. Методи генеративно-змагальних мереж (GAN) та автокодувальників. Методи математичного моделювання та статистичного аналізу. Експериментальні методи та абляційний аналіз.

Наукова новизна отриманих результатів. Отримані результати дисертаційного дослідження мають наукову новизну:

1. Удосконалено метод побудови двопластової інформативної карти на основі порогових вирішальних правил. Відмінності методу полягають у визначенні адаптивних асоціативних залежностей в порогово-метричному просторі на основі агрегації та нормування інформації високочастотної області перетворення Хаара для сегментованих бітових пластів ПЧ-зображень. Це дозволяє створити умови для збільшення рівня стиснення при встановлених вимогах щодо термальної цілісності ПЧ-зображень.

2. Отримано подальший розвиток методу виявлення областей об'єктів інтересу на основі застосування моделі штучного інтелекту. Відмінності методу полягають у виявленні об'єктно-інформативних сегментів за рахунок кон'юнкції результату нейронної мережі із сформованою інформативною картою. Це дозволяє зберегти термальну цілісність об'єктів інтересу за умов обмеженого бітрейту.

3. Вперше розроблено метод спектрально-групового стиснення на основі структурних перетворень у хвильовому просторі Хаара. Відмінності методу полягають у: диференційованому структурному кодуванні хвильового простору сегментів за частотними областями в умовах квадродекомпозиції та встановленням нижньої межі високочастотної області. Це дозволяє зменшити бітовий об'єм пластових сегментів ІЧ-зображень без втрат термальної цілісності.

4. Отримано подальший розвиток методу кодування ІЧ-зображень на основі семантико-орієнтовної обробки. Відмінності методу полягають у спектрально-груповому кодуванні з врахуванням наявності об'єктно-інформативних сегментів ІЧ-зображень в біфуркаційно-хвильовому просторі Хаара. Це дозволяє знизити бітрейт інфрачервоного потоку даних з визначеним рівнем термальної цілісності в процесі обробки та передачі в інформаційних системах.

5. Удосконалено метод відновлення стиснутих ІЧ-зображень на основі використання бігармонічних функцій. Відмінності методу полягають у просторовому прорідженні низькоінформативних областей бітових пластів ІЧ-зображень з їх регенерацією за регулярно-просторовим шаблоном в процесі відновлення з використанням бігармонічних функцій. Це дозволяє підвищити рівень термальної цілісності інформативних сегментів ІЧ-зображень без втрат бітрейту.

Практичне значення отриманих результатів. Впровадження розробленого методу кодування ІЧ-зображень у інформаційних системах дозволяє зменшити бітрейт потоку теплових кадрів в межах заданої термальної цілісності, тобто:

- на 8% для кадрів з переважним вмістом низькоінформативних фонових сегментів;
- на 10% для кадрів з переважним вмістом інформативних сегментів, що містять термальні контрасти;

– на 7% для кадрів з переважним вмістом об’єктно-інформативних сегментів, що насичені тепловими сигнатурами об’єктів інтересу.

Для розробленого методу відносно існуючих досягається перевага у застосуванні штучного інтелекту. Це дозволяє підвищити SSIM для об’єктних областей на 93% у порівнянні із існуючими методами. Що задовольняє вимогам по збереженню цілісності теплових полів об’єктів інтересу.

Застосування штучного інтелекту у поєднанні із метричним визначенням класу інформативності сегментів дає змогу застосовувати при кодуванні бітових пластів селективні коефіцієнти квантизації. Це дозволяє регулювати PSNR від 24 дБ до 38 дБ, що надає перевагу у коефіцієнті стиснення у 4-10% над існуючими методами.

Бігармонічна реконструкція низькоінформативних областей ІЧ-кадру дозволяє збільшити коефіцієнт стиснення на 6%, при цьому зменшуючи середнє квадратичне відхилення у локальних областях.

Практична важливість отриманих результатів дисертації підтверджується актами впровадження результатів науково-прикладних досліджень на базі підприємств: ТОВ “Одеський авіаційний завод”; ТОВ “ЮЕЙ ДЕФЕНС” (Додаток Б).

Особистий внесок здобувача. Всі положення, які виносяться на захист, отримані автором особисто. У наукових працях, опублікованих в співавторстві, здобувачу належить наступне: у працях [1, 6] – запропоновано, розроблено та досліджено метод структурно-статистичної обробки ІЧ-зображень із аналізом впливу методу на бітрейт ІЧ-потоків даних; у працях [7, 14] – розроблено підхід до адаптивного кодування сегментів у спектрально-параметричному просторі. Запропоновано механізм диференційованої обробки сегментів на основі їх частотних характеристик у хвильовому просторі; у працях [13, 11] – проаналізовано фактори, що обмежують ефективність сучасних систем обміну видової інформації. Обґрунтовано необхідність застосування інтелектуальних методів обробки

для збереження термальної цілісності потоків даних. Визначено перспективи розвитку технологій кодування на основі семантичного аналізу сцени; у працях [3, 8] – розроблено та досліджено метод спектрально-групового стиснення на основі структурних перетворень у хвильовому просторі Хаара; у працях [4, 5] – розроблено та досліджено метод біфуркації бітових шарів для подальшої побудови двопластової інформативної карти та обробки ПЧ-даних із широким динамічним діапазоном; у працях [2, 12] – розроблено інтелектуалізований метод обробки ПЧ-зображень на основі порогових вирішальних правил для застосування семантико-орієнтовної обробки; у працях [9, 10] – запропоновано та обґрунтовано метод диференційовано-структурного кодування при квадродекомпозиції сегментів.

Апробація результатів дисертації. Теоретичні і практичні результати дисертаційної роботи, доповідалися і були схвалені на міжнародних і всеукраїнських науково-технічних конференціях: 2026 IEEE 18th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET) (Lviv, Ukraine, 17–21 Feb., 2026); Cyber Hygiene & Conflict Management in Global Information Networks (CH&CMiGIN'25) (Kyiv, Ukraine, 20–22 June, 2025); Кіберборотьба: розвідка, захист та протидія (Київ, 6–7 трав. 2025 р.); 2024 IEEE 5th International Conference on Advanced Trends in Information Theory (ATIT) (Lviv, Ukraine, 21–23 Nov., 2024); Комп'ютерне моделювання в наукоємних технологіях (КМНТ-2024) (Харків, 27–29 листоп. 2024 р.); 2023 IEEE 5th International Conference on Advanced Information and Communication Technologies (AICT) (Lviv, Ukraine, 21–25 Nov., 2023); Наукоємні технології в інфокомунікаціях (НІСТ'2023) (Кам'янець-Подільський, 1–3 черв. 2023 р.); 2022 IEEE 4th International Conference on Advanced Trends in Information Theory (ATIT) (Kyiv, Ukraine, 15–17 Dec., 2022).

Публікації результатів дисертації. Основні положення та результати дисертаційної роботи опубліковано у 14 наукових працях, серед яких: 3 статей опубліковано у науково-технічних журналах, які включені до

міжнародних наукометричних баз Scopus/Web of Science; 3 статей у наукових журналах, які включені в «Перелік наукових фахових видань України» та мають категорію Б. Апробація результатів дисертації відображена у 8 тезах доповідей на міжнародних та всеукраїнських науково-технічних та науково-практичних конференціях, серед яких 5 входять до IEEE/Scopus.

Структура та обсяг дисертації. Дисертація складається зі вступу, чотирьох розділів, висновків та додатків. Повний обсяг дисертації складає 220 сторінок, у тому числі 164 сторінок основного тексту. У дисертації наведено 52 рисунків (у тому числі 14 рисунків на 14 окремих сторінках), 7 таблиць (у тому числі 3 таблиці на 3 окремих сторінках), список використаних джерел із 163 найменувань на 23 сторінках, 3 додатки на 9 сторінках. Дисертація написана українською мовою.

Глибока вдячність висловлюється доктору технічних наук, професору Баранніку Володимирі Вікторовичу за професійний приклад провідного вченого в напрямі цифрової обробки та безпеки відеоінформації.

РОЗДІЛ 1. ОБҐРУНТУВАННЯ НЕОБХІДНОСТІ УДОСКОНАЛЕННЯ ІНФОРМАТИВНИХ ТЕХНОЛОГІЙ ОБРОБКИ ІЧ-ЗОБРАЖЕНЬ

Розділ присвячено аналізу специфіки генерації, обробки та передачі інфрачервоних даних у сучасних інформаційних системах.

Доведено актуальність задачі зниження бітрейту інфрачервоного потоку даних для передачі в інформаційних системах з визначеним рівнем термальної цілісності. Виявлено науково-технічну суперечність між необхідністю зниження бітрейту потоку інфрачервоних даних та вимогами до збереження теплової семантики об'єктів інтересу.

Аргументовано доцільність застосування семантико-орієнтованого підходу до кодування, що дозволяє знизити бітове навантаження із заданим рівнем цілісності. Сформовано набір вимог до перспективного методу кодування термальних кадрів із застосуванням методів штучного інтелекту.

Поставлено науково-прикладну задачу щодо зниження бітрейту інфрачервоного потоку даних для передачі в інформаційних системах з визначеним рівнем термальної цілісності. Визначено мету дослідження: розробити метод кодування ІЧ-зображень для зниження бітрейту без втрат термальної цілісності на основі семантико-орієнтовної обробки з використанням моделей ШІ.

1.1 Аналіз сфер застосування інфрачервоних зображень та їх ролі в інформаційних системах широкого призначення

У сучасному світі інформаційні технології стали основою для розвитку критичних сфер людської діяльності: оборони, промисловості, медицини, екологічного моніторингу та інтелектуальних міських систем. Зростаючі вимоги до безпеки, ефективності та автоматизації процесів стимулюють широке використання моніторингових систем [2]. Такі системи повинні забезпечувати безперервне спостереження і контроль у складних умовах експлуатації. Однак існуючі методи обробки візуальної інформації, що базуються на роботі у видимому спектрі світла, стикаються з істотними обмеженнями при функціонуванні в реальних умовах [38, 112, 128, 148].

Ефективність сучасних систем візуального моніторингу обмежується залежністю від умов зовнішнього середовища [38]. До основних дестабілізуючих факторів належать: низький рівень освітленості, складні погодні умови, аерозольні завади та засоби маскування. Вплив цих чинників обмежує інформативність зображень та, як наслідок, працездатність системи. Водночас оперативність прийняття рішень та ефективність заходів реагування безпосередньо корелюють із повнотою та актуальністю вхідної інформації.

Особливо гостро цей недолік проявляється в умовах, коли моніторингові системи повинні працювати в екстремальних ситуаціях: під час техногенних аварій, природних катастроф, військових дій чи інших надзвичайних обставин [35, 49]. У таких сценаріях відмова або неефективна робота систем візуального спостереження може призвести до критичних наслідків: загрози життю і здоров'ю людей, значних матеріальних збитків, порушення функціонування стратегічно важливих об'єктів інфраструктури. Тому актуальною стає потреба в застосуванні альтернативних технологічних рішень, здатних забезпечити надійне отримання видової інформації

мінімізуючи вплив зовнішнього середовища та можливих завадових факторів [126, 157, 158].

У цьому контексті інфрачервоні (ІЧ) технології посідають важливе місце серед перспективних напрямів розвитку інформаційних систем із компонентами моніторингу [112]. Перевагою інфрачервоного спектра є принципово відмінний механізм формування зображення, що базується на реєстрації власного теплового випромінювання об'єктів, яке визначається їх температурними характеристиками [54, 151]. Це забезпечує функціональну незалежність ІЧ-систем від зовнішніх джерел освітлення та дозволяє отримувати інформативні зображення в умовах повної відсутності видимого світла [95].

Додатковою перевагою інфрачервоного діапазону є його підвищена проникна здатність через низку оптичних перешкод, включаючи аерозольні утворення, легкі димові завади та деякі типи камуфляжних покриттів, що робить ІЧ-технології незамінними в умовах обмеженої видимості [38, 81]. Здатність реєструвати температурні контрасти між об'єктами та фоном створює унікальні можливості для виявлення цілей, що мають температурні характеристики, відмінні від навколишнього середовища, навіть за наявності візуальних перешкод [80, 116, 160].

Широкий спектр практичних застосувань інфрачервоних технологій демонструє їх універсальність та ефективність у різноманітних галузях. Особливе значення ІЧ-камери мають у складі бортових комплексів оборонного призначення, встановлених на безпілотних літальних апаратах, гелікоптерах, тактичних літаках, надводних та наземних платформах [55, 120, 121, 156]. У розвідувальних комплексах інфрачервоний канал використовується для пошуку, виявлення та супроводу цілей на великих відстанях, контролю протяжних ділянок території, моніторингу переміщення техніки та живої сили, а також для оцінки змін теплової обстановки в зоні інтересу. Ударні бортові комплекси додатково покладаються на ІЧ-зображення при цілевказанні, уточненні координат об'єктів, формуванні зон

ураження та постопераційному аналізі наслідків застосування засобів ураження. В обох класах систем інфрачервоні камери інтегруються з оптичними, радіолокаційними та навігаційними сенсорами у єдиний інформаційний контур, що забезпечує багатоканальну ситуаційну обізнаність екіпажу або автоматизованої системи керування [139].

Функціонування таких бортових комплексів супроводжується формуванням значних обсягів ПЧ-даних з високою просторовою і часовою роздільною здатністю [96, 162]. Потоки тепловізійних кадрів повинні передаватися в умовах обмеженої пропускну здатності, а також оброблятися в режимі реального часу [47, 78]. Це необхідно для підтримки систем автоматичного виявлення, класифікації та супроводу цілей. Як наслідок, формуються наступні вимоги до методів кодування: з одного боку, необхідно забезпечити зменшення бітового об'єму, а з іншого – зберегти теплові сигнатури об'єктів інтересу у зонах прийняття рішень [130].

У галузі пожежної безпеки інфрачервоні технології відіграють важливу роль у системах раннього виявлення загорань [54, 151]. ПЧ-сенсори здатні реєструвати теплові аномалії та локальні підвищення температури на початкових стадіях пожежі, коли видимі ознаки горіння ще відсутні [38, 157]. Це забезпечує часовий резерв для ініціації заходів пожежогасіння та евакуації, зменшуючи ризик катастрофічних наслідків [137].

У промисловому секторі тепловізійні технології інтегруються в системи технічної діагностики та превентивного обслуговування обладнання [5, 148]. Інфрачервоний контроль дозволяє виявляти дефекти ізоляції, перегрів електричних з'єднань, механічне зношування вузлів та інші аномалії на ранніх стадіях експлуатації, що забезпечує підвищення надійності виробничих процесів та зниження ризику аварійних ситуацій [2, 41, 44].

У рамках концепції «розумного міста» інфрачервоні камери забезпечують безперервне спостереження за громадськими місцями та транспортними вузлами [35, 49]. Інтеграція з системами штучного інтелекту дозволяє реалізувати автоматичне розпізнавання підозрілої поведінки та

класифікацію загроз на основі аналізу теплових сигнатур [88, 92]. Алгоритми машинного навчання ефективно відрізняють людей від тварин і транспортних засобів навіть в умовах обмеженої видимості, забезпечуючи точне виявлення потенційних порушень та миттєве сповіщення служб безпеки [14].

Супутникові ІЧ-системи дистанційного зондування Землі відкривають широкі можливості для великомасштабного моніторингу навколишнього середовища [34, 64]. Інфрачервоні знімки високої роздільної здатності використовуються для створення точних тематичних карт, що слугують для: моніторингу змін рослинного покриву, оцінки стану сільськогосподарських угідь, виявленню лісових пожеж, контролю забруднення водних об'єктів та прогнозування кліматичних змін на основі аналізу температурних полів поверхні Землі [10, 129, 139].

У медичній галузі термографічні методи діагностики базуються на аналізі розподілу температур на поверхні тіла людини, що відображає особливості кровопостачання тканин та метаболічні процеси [94, 95]. Медична термографія знаходить застосування в діагностиці судинних патологій, запальних процесів, ортопедичних порушень та інших станів, що супроводжуються локальними змінами температурного режиму тканин [137, 151, 157].

Комерційне впровадження ІЧ-технологій охоплює широкий спектр застосувань: системи охорони приватної власності, інтелектуальні будівельні систем управління енергоспоживанням [126, 148], автомобільні системи нічного бачення [38, 81, 162].

Наведені сфери застосування ІЧ-камер проілюстровано на схемі рисунку 1.1. Із чого можна зробити висновок, що технології фіксації ІЧ-зображень інтегровано у широкий спектр прикладних задач. Масштабування цих систем корелює зі зростанням обсягів даних [78]. Збільшення кількості інфрачервоних сенсорів та підвищення їх роздільної здатності призводить до підвищення обсягів генерованої інформації [96, 120, 121, 156].

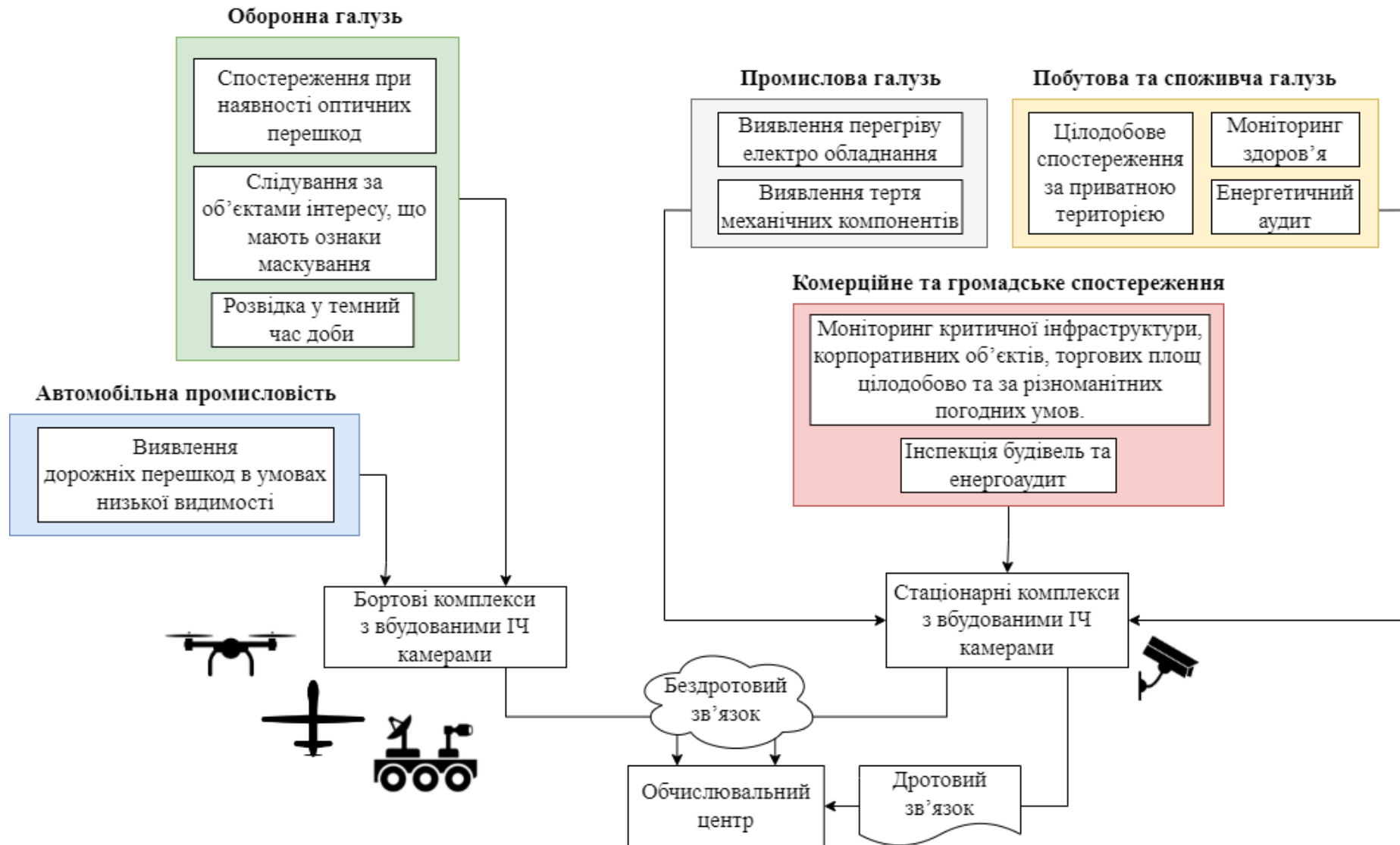


Рисунок 1.1 – Схема сфер застосування ІЧ-камер

Сучасні ПЧ камери високої роздільності можуть генерувати потоки даних обсягом від декількох мегабайт до сотень мегабайт на секунду [74, 78]. У багатосенсорних системах, які об'єднують десятки або сотні ПЧ-камер для моніторингу великих територій чи масштабних об'єктів, сукупний інформаційний потік може досягати гігабайт і навіть терабайт на годину [47, 149]. Це створює суттєве навантаження на системи передачі та обробки даних [130].

Обмеження можуть посилюватись через необхідність одночасної передачі різнотипної інформації крізь спільні канали зв'язку [12, 14]. Коли в єдиний канал надходять потоки від камер видимого діапазону, інфрачервоних сенсорів, радарних систем та інших джерел даних, сумарний бітовий потік може в рази перевищувати ресурс доступної пропускної здатності. Це призводить до формування черг на передачу, накопичення затримок та необхідності вибіркового відсікання частини інформації для забезпечення функціонування системи [57, 130, 158].

Важливим аспектом є забезпечення режиму реального часу для систем автоматичного моніторингу та прийняття рішень [158]. У системах безпеки затримки в передачі ПЧ-даних навіть на декілька секунд можуть призвести до втрати можливості своєчасного виявлення та нейтралізації загроз [120, 139]. Втрата кадрів або їх надмірна затримка порушує часову послідовність подій, що унеможлиблює коректне відстеження траєкторій рухомих об'єктів, оцінку швидкості їх переміщення та прогнозування подальшої поведінки [156, 162].

Крім цього, існує фактор нестабільності та обмеженості зв'язку у інформаційних системах. Як наслідок, доступний ресурс бітової швидкості може варіюватись протягом часу. Відповідно, методи кодування даних повинні забезпечувати адаптивність, однак забезпечуючи цілісність важливої інформації. Така потреба обґрунтована наступними практичними сценаріями:

- міські бездротові мережі стикаються з перевантаженням через велику кількість користувачів та інтерференцію сигналів [121];

- безпілотні літальні апарати, оснащені ІЧ-камерами, працюють в умовах нестабільної якості зв'язку [55, 96, 156]. Вона може змінюватися залежно від відстані до базової станції, погодних умов та рельєфу місцевості [107].

Окрім зазначених обмежень пов'язаних із обміном інформації, існують апаратні обмеження. Віддалені пристрої із вбудованими засобами спостереження зазвичай мають енергоефективне апаратне забезпечення [149]. Це зумовлює лімітовану обчислювальну потужність. Ситуація ускладнюється необхідністю функціонування бортових систем у обчислювально адаптивному режимі [60, 92]. Наприклад при переході пристрою в режим підвищеної економії енергії, при перегріві процесора або при необхідності звільнити обчислювальні потужності для пріоритетних задач. Це створює конкуренцію за обчислювальні потужності та обмежує застосування ресурсомістких алгоритмів первинної обробки, зокрема методів комп'ютерного зору [48, 111]. Відтак, методи кодування відеоінформації не повинні спиратися виключно на обчислювально місткі підходи. Натомість, архітектура системи обробки даних має забезпечувати гнучкість конфігурації. Дозволяючи адаптувати складність кодування до поточної наявності обчислювальних ресурсів [93].

Економічний аспект полягає в тому, що збільшення пропускної здатності каналів зв'язку або розгортання додаткової інфраструктури вимагає значних капітальних інвестицій, які можуть бути економічно недоцільними або технічно неможливими в певних умовах експлуатації [36, 74]. Тому, для підвищення якості функціонування ІЧ-інформаційних сервісів, практичним рішенням стає застосування покращених алгоритмів стиснення ІЧ-зображень. Вони повинні враховувати семантичні особливості даних та забезпечувати зменшення бітового обсягу при збереженні критично важливої інформації [72, 119, 132, 147].

1.2 Дослідження технічних характеристик інфрачервоних камер як базової компоненти формування даних

Така різноманітність сфер застосування свідчить про технологічну зрілість та економічну доступність інфрачервоних систем візуального моніторингу. Однак масштабне впровадження ІЧ-технологій у критично важливі сфери діяльності породжує низку серйозних технічних викликів, пов'язаних із обробкою, передачею та зберіганням великих обсягів візуальних даних.

Обсяги даних, які генерують сучасні інфрачервоні камери, суттєво залежать від спектрального діапазону роботи, роздільної здатності матриці, розрядності пікселів та частоти кадрів. Для багатокамерних систем безперервного моніторингу ці обсяги створюють критичні обмеження для інфраструктури зберігання, пропускну здатності мереж передачі даних та обчислювальних ресурсів для їх обробки в реальному часі.

Ефективне вирішення цих викликів вимагає розробки спеціалізованих методів стиснення, адаптованих до специфіки інфрачервоних зображень. Проте перед розробкою таких методів необхідно провести детальний аналіз технічних характеристик сучасних ІЧ-камер, особливостей формування зображень у різних спектральних діапазонах та закономірностей генерації візуальних даних [112, 157]. Таке розуміння дозволить визначити параметри, що мають зберігатися при стисненні, та обґрунтувати вимоги до методів кодування враховуючи різні способи застосувань [53, 61, 89].

Класифікація ІЧ-камер здійснюється за довжиною хвилі випромінювання, до якого чутливий сенсор. Спектральний діапазон безпосередньо визначає тип формованих даних, їхню структуру, бітову глибину, динамічний діапазон і придатність до конкретних способів обробки та кодування.

Відповідно до міжнародної класифікації, виділяють чотири основні інфрачервоні діапазони:

- ближній діапазон (Near-Infrared Range, NIR) – довжина хвилі 0.7–0.9 мкм;
- короткохвильовий діапазон (Short-Wave Infrared, SWIR) – довжина хвилі 0.9–1.7 мкм;
- середньохвильовий діапазон (Mid-Wave Infrared, MWIR) – довжина хвилі 3–5 мкм;
- довгохвильовий діапазон (Long-Wave Infrared, LWIR) – довжина хвилі 8–15 мкм.

Позиціонування ІЧ-діапазонів відносно інших видів випромінювання зображено на рисунку 1.2. Варто відзначити наближеність NIR та SWIR до візуального діапазону. Ця особливість полягає не лише у фізичних характеристиках випромінювання, а й у природі даних.

Камери ближнього інфрачервоного діапазону формують зображення відбивального типу. Тому вони близькі за структурою до видимого спектра. Тож, такі дані можуть оброблятися подібно до звичайних RGB або монохромних кадрів.

Тоді як MWIR та LWIR камери фіксують теплове випромінювання об'єктів [3, 53]. Саме ця відмінність зумовлює різні статистичні властивості пікселів, що безпосередньо впливає на ефективність алгоритмів стиснення [34, 64, 109].

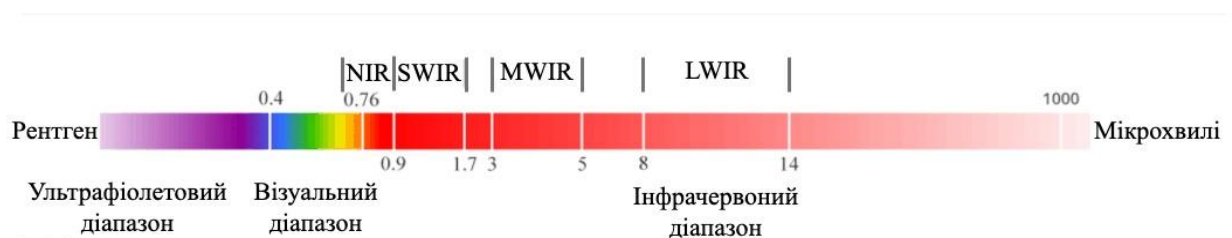


Рисунок 1.2 – Електромагнітний спектр із позначенням піддіапазонів інфрачервоного випромінювання (SWIR, MWIR, LWIR)

Варто звернути увагу на характерну особливість NIR камер. Вони мають вбудоване кодування відеопотоків із застосуванням розповсюджених

відеокодеків (H.264, H.265, MJPEG) [28, 89]. Розповсюджене застосування цих кодеків обґрунтоване схожістю NIR каналу на оптичний відео канал.

Приклади NIR пристроїв:

- Hikvision KS-LSD8040 – роздільна здатність 1920x1080, 8-бітна глибина, кількість кадрів в секунду 30. Використовує кодеки H.264 або H.265 для кодування відео;

- e-con Systems See3CAM_CU55M – роздільна здатність 1920x1080, 8-12-бітна глибина, кількість кадрів в секунду 60. Передає дані у форматах UVC Y8/Y12 з підтримкою MJPEG для стиснення.

Приклад порівняння NIR кадру та оптичного можна побачити на рисунку 1.3. Де обидва знімки зроблені у нічний час доби. Оптичний знімок на рисунку 1.3 а) містить ледь помітну інформацію. Тоді як за тих самих умов ІЧ знімок на рисунку 1.3 б) відображає більш повне представлення сцени. Наявний ефект підсвітки, присутній на рисунку 1.3 б). Що супроводжується підвищеною яскравістю по центру, та її затуханням ближче до меж кадру. Даний ефект утворюється через NIR підсвітку камери [46, 56, 135].



а)

б)

Рисунок 1.3 – Порівняння кадрів у видимому діапазоні та NIR: а) зображення отримане з оптичної камери, б) зображення отримане з ІЧ NIR камери See3CAM_CU27

Камери SWIR-діапазону забезпечують реєстрацію відбитого випромінювання, здатного проникати у поверхневі шари матеріалів. Отримані зображення відображають структурні та матеріальні особливості об'єктів, включно з вологою, типом тканин або композицією речовини [112, 128].

Приклади SWIR пристроїв:

- GH-SW130T-GigE – роздільна здатність 1280x1024, 8-10-12-бітова глибина, кількість кадрів в секунду 89;

- LyTID SIRIS – роздільна здатність 640x512, 16-бітова глибина, кількість кадрів в секунду 200.

Безпосередньо камери LyTID SIRIS та GH-SW130T-GigE не використовують вбудованих кодеків, а видають потік у нестиснутому (RAW) форматі за протоколами Camera Link або GigE Vision, з можливістю подальшої компресії зовнішніми пристроями [33, 75, 123, 124, 140].

Приклад порівняння оптичного кадру та SWIR можна бачити на рисунку 1.4. Варто звернути увагу на проникні можливості SWIR, що дозволяють отримати інформацію за димовими перешкодами оточуючого середовища [35, 38, 148].



а)

б)

Рисунок 1.4 – Порівняння кадрів у видимому діапазоні та SWIR: а) зображення отримане з оптичної камери, б) зображення отримане з ІЧ SWIR камери

Камери MWIR-діапазону формують термальні дані, де кожен піксель містить значення, пропорційне температурі поверхні. Ці зображення є важливими для наукових, оборонних і промислових систем, де необхідна висока точність вимірювання теплових характеристик.

MWIR-сенсори – це, як правило, охолоджені фотонні детектори, що забезпечують поглиблену бітову розрядність та чутливість менше 20 мК (0.02 °С). Що зумовлює застосування широких динамічних діапазонів при формуванні термальних кадрів.

Приклади MWIR пристроїв:

- FLIR 715CTZ – роздільна здатність 640x512, 12-14-бітова глибина, кількість кадрів в секунду 30;

- Teledyne FLIR Neutrino SX12 ISR1200 – роздільна здатність 1280x1024, 16-бітова глибина, кількість кадрів в секунду 60.

Обидві наведені камери використовують пропріетарні закриті кодеки або передачу нестиснутих даних через протоколи Camera Link, GigE Vision, HD-SDI. Оскільки компресійні кодеки не застосовуються безпосередньо камерою, їх може використовувати зовнішнє обладнання. Це типово для камер такого класу, орієнтованих на технічні і оборонні застосування [26, 40, 58, 69, 107].

На рисунку 1.5 відображене порівняння оптичних кадрів та MWIR. Для MWIR кадрів характерне «підсвічування» об'єктів інтересу, що мають теплову сигнатуру.



а)

б)

Рисунок 1.5 – Порівняння кадрів у видимому діапазоні та MWIR: а) зображення отримане з оптичної камери, б) зображення отримане з ІЧ MWIR камери

Камери LWIR-діапазону реєструють власне теплове випромінювання об'єктів, не потребуючи зовнішнього освітлення, як і MWIR камери. Розрядність даних зазвичай становить 14 біт, у професійних моделях – до 16 біт. Довші довжини хвиль LWIR-випромінювання означають, що на нього менше впливають атмосферні частинки, і воно може ефективно проникати крізь дим, туман і пил. Це робить камери LWIR ідеальними для зовнішнього застосування та застосування на великих відстанях, а також для застосувань, що потребують роботи в складних атмосферних умовах [53, 78].

Інформаційно LWIR-зображення відзначаються високою просторовою кореляцією пікселів, малою варіативністю шуму та значною міжкадровою надлишковістю, що створює передумови для застосування спектральних і семантично-орієнтованих методів кодування [20, 61, 88, 136, 143].

Приклади LWIR пристроїв:

- InfraTec VarioCAM HD Z – роздільна здатність 1024×768, 16-бітова глибина, кількість кадрів в секунду 30;
- FLIR A655sc – роздільна здатність 640×480, 16-бітова глибина, кількість кадрів в секунду 50.

Так само, як і у MWIR камерах, LWIR камери не виконують стиснення даних. Потоки інформації передаються по GigE Vision протоколу у нестиснутому RAW форматі [32, 52, 101].

На рисунку 1.6 зображено порівняння кадрів MWIR та LWIR. Оскільки LWIR спектр має більшу проникну здатність крізь дим, туман і пил, то на великій відстані кадр відображає інформацію із меншими завадами, у порівнянні із MWIR.



Рисунок 1.6 – Порівняння ІЧ кадрів у діапазонах MWIR та LWIR: а) зображення отримане з ІЧ MWIR камери, б) зображення отримане з ІЧ LWIR камери

У таблиці 1.1 наведені характеристики згаданих ІЧ камер класів NIR, SWIR, MWIR, LWIR. Варто відзначити варіативність роздільної здатності та кількості кадрів у секунду. Виробники реалізують у своїх виробках різноманітні значення цих параметрів, незалежно від класу камери. Однак можна прослідкувати чітке збільшення бітової глибини зі збільшення спектрального діапазону, від NIR до LWIR. Відмінність спектрального діапазону впливає на формування семантичних особливостей кадру. Як наслідок, тепловізійні MWIR та LWIR діапазони мають відмінну інформаційну надлишковість у порівнянні із діапазонами NIR та SWIR, що наближені до візуального спектру [1, 4, 5].

З огляду на виявлені відмінності спектральних діапазонів, саме дані MWIR та LWIR характеризуються специфічною інформаційною надлишковістю, відмінною від візуального спектра. Це обумовлює можливості для підвищення їх ефективності стиснення та визначає пріоритетність дослідження цих діапазонів у контексті розробки спеціалізованих методів кодування [3, 16].

Таблиця 1.1 – Характеристики ІЧ камер класів NIR, SWIR, MWIR, LWIR

Модель	Клас	Роздільна здатність	Бітова глибина	Кількість кадрів в секунду
Hikvision KS-LSD8040	NIR	1920x1080	8	30
e-con Systems See3CAM_CU55M	NIR	1920x1080	8	60
GH-SW130T-GigE	SWIR	1280x1024	8	89
LyTID SIRIS	SWIR	640x512	16	200
FLIR 715CTZ	MWIR	640x512	14	30
Teledyne FLIR Neutrino SX12 ISR1200	MWIR	1280x1024	16	60
InfraTec VarioCAM HD Z	LWIR	1024x768	16	30
FLIR A655sc	LWIR	640x480	16	50

Для оцінки об'ємів даних, що потенційно можуть генеруватись ІЧ-камерами необхідно перейти до показників бітової швидкості (бітрейту). Бітрейт визначає утворений бітовий об'єм на одиницю часу. Зазвичай оцінюється бітове навантаження на секунду.

Бітрейт R_{raw} нестисненого потоку кадрів визначається роздільною здатністю ($W \times H$) сенсора, бітовою глибиною D пікселів та кадровою частотою F в секунду:

$$R_{\text{raw}} = W \cdot H \cdot D \cdot F. \quad (1.1)$$

Значення бітрейту R_{raw} , що генерують MWIR та LWIR ІЧ-камери наведені у таблиці 1.2.

Таблиця 1.2 – Значення бітрейту R_{raw} для ІЧ-камер

Модель ІЧ-камери	R_{raw} (Мбіт/с)
FLIR A655sc	24.576
FLIR 715CTZ	137.626
InfraTec VarioCAM HD Z	377.487
Teledyne FLIR Neutrino SX12 ISR1200	1258.291

На рисунку 1.7 відображено графік залежності бітрейту від характеристик ІЧ-камери. Можна помітити ріст бітрейту наближений до експоненційного при збільшенні роздільної здатності та частоти фіксації кадрів.

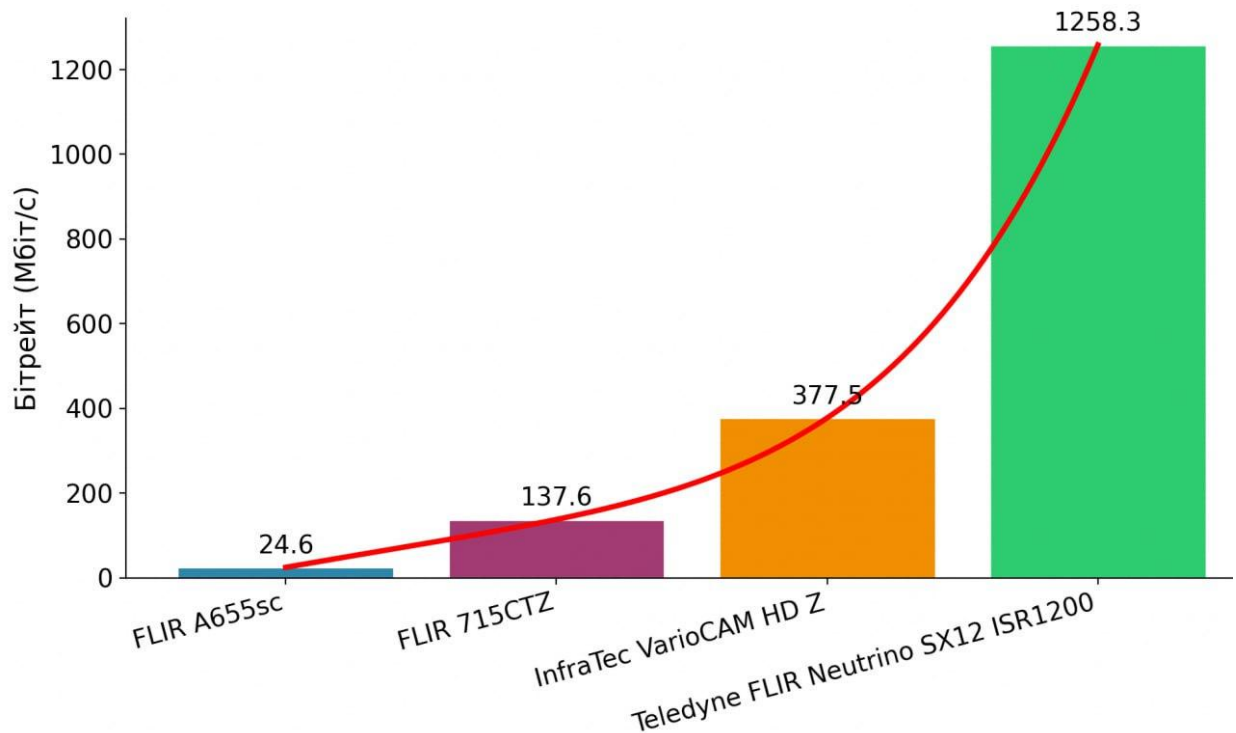


Рисунок 1.7 – Графік залежності бітрейту від характеристик ІЧ-камери

Аналіз отриманих значень бітрейту R_{raw} (Табл. 1.2) свідчить про те, що сучасні тепловізійні ІЧ-сенсори генерують інформаційні потоки значної інтенсивності. Навіть при стандартній частоті кадрів 30–60 FPS бітовий об'єм нестисненого потоку варіюється в межах від десятків мегабіт до понад 1 Гбіт/с.

Така інтенсивність генерації даних створює навантаження на архітектуру інформаційних систем.

По-перше, виникають обмеження при обробці. Бортові обчислювальні модулі автономних платформ (БПЛА, роботизованих комплексів) мають обмеження щодо доступних обчислювальних ресурсів. Обробка, аналіз та фільтрація сирого 16-бітного потоку високої роздільності у режимі реального часу призводять до суттєвого навантаження на процесор, збільшення затримок та зниження загальної швидкодії системи прийняття рішень [14, 101, 143].

По-друге, існують обмеження при передачі. Пропускна здатність каналів зв'язку часто є нижчою за бітрейт вихідного потоку даних. Спроба

передачі таких обсягів без попередньої обробки призводить до переповнення буферів, втрати пакетів та порушення хронології потоку інформації, що знижує актуальність даних для оператора або автоматичної системи аналізу [52, 82].

Для роботи у цих обмеженнях застосовуються методи обробки даних, що зменшують інформаційну надмірність. Проте з'являється ризик спотворення семантично важливих ознак: теплових контрастів та сигнатур об'єктів інтересу.

Зазначені фактори формують науково-технічне протиріччя. Воно полягає у невідповідності між необхідністю зниження бітрейту потоку інфрачервоних даних та вимогами до збереження теплової семантики об'єктів інтересу.

Унаслідок цього, актуальна **науково-прикладна задача** полягає у зниженні бітрейту інфрачервоного потоку даних для передачі в інформаційних системах з визначеним рівнем термальної цілісності.

1.3 Аналіз можливостей і обмежень методів кодування інфрачервоних зображень

Як було показано у попередньому підрозділі, інфрачервоні камери, особливо MWIR та LWIR діапазонів, формують дані з підвищеною бітовою глибиною та широким динамічним діапазоном. Для збереження повної температурної інформації та службових параметрів зйомки провідні виробники тепловізійних систем розробили низку спеціалізованих пропрієтарних форматів та кодеків. До таких форматів належать R.JPEG, .IS2, .SEQ, .FFF, .IRB, .IMG/.SIT/.IRI, .DAT, .ANA тощо, кожен з яких реалізує власну схему кодування й структуру метаданих. Наприклад, формат R.JPEG компанії FLIR поєднує JPEG-подібне стиснення з вбудованими термальними параметрами, тоді як контейнер .SEQ призначений для зберігання послідовностей термографічних кадрів, включно з RAW-

компонентами. Аналогічні підходи реалізують формати .IS2 (InfraTec), .IRI (Testo) та інші, орієнтовані на відтворення температурних полів без втрати метричної точності [16, 62, 142].

Попри високий рівень достовірності, наявні такі недоліки: обмежені специфікації зберігання даних, алгоритми кодування та службові структури доступні лише через пропрієтарні SDK або взагалі не документуються публічно. Це ускладнює їх інтеграцію в багатосенсорних інформаційних системах, обмежує можливості модифікації та незалежної верифікації характеристик і стримує розробку уніфікованих методів обробки ІЧ-даних [31, 43].

У зв'язку з цим виникає потреба перейти від аналізу окремих виробничих форматів до розгляду загальних принципів кодування зображень, які можуть бути застосовані до інфрачервоних кадрів незалежно від конкретного постачальника обладнання. З позицій теорії обробки даних методи кодування доцільно поділяти на два основні класи: безвтратні, що забезпечують повну відтворюваність вихідних даних, та з втратами, що досягають більш високих коефіцієнтів стиснення за рахунок контрольованих спотворень.

У рамках безвтратного підходу одним із найпоширеніших форматів є PNG, який широко використовується для зберігання зображень з гарантованим збереженням усіх бітів вхідного сигналу. PNG базується на поетапному усуненні статистичної надлишковості. Алгоритм стиснення включає два етапи: попередню фільтрацію, яка знижує ентропію шляхом запису різницевих значень, та ентропійне кодування за допомогою алгоритму DEFLATE, що є комбінацією LZ77 і кодів Хаффмана [37, 122, 161]. Завдяки підтримці 16-бітного динамічного діапазону, цей формат забезпечує точне збереження температурних градієнтів, що дозволяє коректно представляти дані інфрачервоної зйомки [94, 155].

Водночас, використання даного підходу для передачі поточкових даних в умовах, наближених до реального часу, виявляє низку обмежень. Основним

недоліком є недостатній рівень компресії, зумовлений тим, що принципи безвтратного стиснення застосовуються рівномірно до всього масиву пікселів кадру без урахування його змісту. Такий метод ігнорує семантичні закономірності сцени, внаслідок чого система кодування позбавлена адаптивності: вона витрачає значний обсяг даних на детальне збереження фонових ділянок нарівні з цільовими об'єктами, що є неефективним в умовах обмеженої пропускної здатності каналу зв'язку [50, 72, 78, 119, 145].

Таким чином, універсальний безвтратний підхід, реалізований у PNG, не забезпечує балансу між ступенем стиснення та збереженням інформативності, що зумовлює необхідність застосування методів, що враховують не тільки статистичні, а і семантичні властивості кадру.

Такі методи кодування застосовуються у кодеках із втратами. Кодеки із втратами ґрунтуються на принципі часткового видалення інформації, яка має незначний вплив на сприйняття зображення людиною чи технічною системою. На відміну від безвтратних методів, де відновлене зображення повністю ідентичне вихідному, у випадку стиснення з втратами допускається контрольована деградація даних задля суттєвого зменшення обсягу інформації. Основна ідея таких методів полягає у використанні психовізуальних та статистичних моделей для виявлення і вилучення надлишкових або малозначущих компонент сигналів. Для цього виконуються етапи перетворення зображення у спектральний простір, де енергія сигналу зосереджується в окремих коефіцієнтах (наприклад, дискретне косинусне чи вейвлет-перетворення), подальшого квантування цих коефіцієнтів із заданою точністю, а також ентропійного кодування для ефективного подання отриманих даних. Втрати інформації переважно виникають саме на етапі квантування, коли незначні деталі або шумові компоненти сигналу відкидаються. У результаті досягається високе співвідношення стиснення при мінімальному впливі на візуальну якість, що робить такі кодеки особливо ефективними для передачі або зберігання великих обсягів мультимедійних даних, зокрема відео та фотозображень, де важливим є

компроміс між якістю реконструкції та швидкістю передачі інформаційними каналами [36, 74, 132, 133, 135].

Існує велика кількість популярних кодеків із втратами. Одним із найпопулярніших є JPEG, який став базовим стандартом цифрової обробки та зберігання зображень. У процесі обробки зображення конвертується з колірного простору RGB у простір YCbCr, де компонент Y відповідає яскравості, а компоненти Cb та Cr - кольоровим відхиленням. Така конверсія мотивована фізіологічними особливостями зору людини, який значно чутливіший до змін яскравості, ніж до кольорових відтінків. Це дозволяє застосувати субколірну дискретизацію - тобто зменшення просторової роздільності кольорових компонентів (наприклад, схема 4:2:0) без помітної втрати якості сприйняття. Основний принцип JPEG полягає у використанні дискретного косинусного перетворення (ДКП), що дозволяє представити блоки зображення у вигляді набору частотних коефіцієнтів, де більшість енергії сигналу зосереджується у низькочастотних складових. Після цього виконується квантування коефіцієнтів, при якому високочастотні компоненти, що відповідають дрібним деталям або шумам, відкидаються або грубо апроксимуються, що забезпечує значне зменшення обсягу даних. На завершальному етапі результати квантування піддаються ентропійному кодуванню, зазвичай із використанням алгоритму Хаффмана, що забезпечує додаткову економію простору. Завдяки простоті реалізації, високій ефективності стиснення та прийнятній візуальній якості реконструйованих кадрів, стандарт JPEG став універсальним рішенням для фото- та відеозастосувань [28, 32, 56, 135, 152].

Після широкого впровадження стандарту JPEG розвиток технологій стиснення зображень продовжився у напрямку підвищення ефективності, гнучкості та якості реконструкції. JPEG, попри свою простоту та універсальність, має низку обмежень, серед яких – помітні артефакти блокування при високих коефіцієнтах стиснення, фіксований розмір блоків 8x8, а також відсутність масштабованості та підтримки розширеної глибини

кольору. Ці недоліки стимулювали появу низки сучасних кодеків, спрямованих на підвищення продуктивності при збереженні прийнятної сумісності з існуючими системами [61, 73, 83, 134, 152].

Одним із перших рішень нового покоління став формат WebP, розроблений компанією Google. Його принципи ґрунтуються на внутрішньокадровому кодуванні відеостандарту VP8, що використовує механізми просторового передбачення макроблоків (зокрема, алгоритми TrueMotion, H_PRED, V_PRED) та подальше стиснення різницевого сигналу адаптивним бінарним арифметичним кодером. WebP підтримує як режими з втратами, так і без втрат, що робить його зручним для веб-застосувань, де важливими є швидкість завантаження та мінімальний розмір файлів. Однак через обмеження глибини кольору та відсутність підтримки 16-бітних форматів з розширеним динамічним діапазоном (HDR) цей кодек рідко використовується в технічних або наукових системах, зокрема при обробці інфрачервоних (ІЧ) кадрів [46, 53, 64, 86, 127].

Для підтримки кодування HDR зображень та забезпечення зворотної сумісності із кодеком JPEG було розроблено JPEG XT. Цей метод кодування складається із базового (Base Layer) та розширювального (Extension Layer) шарів. Процес обробки інформації розпочинається з отримання вхідного зображення з підвищеною бітовою глибиною, яке підлягає первинній трансформації за допомогою операторів тонального відображення (Tone Mapping Operators). Метою цього етапу є генерація стандартного 8-бітного зображення (Low Dynamic Range, LDR), яке слугуватиме базовим шаром. Це зображення кодується згідно з алгоритмом JPEG. Включаючи перетворення колірного простору, субдискретизацію, блочне розбиття, дискретне косинусне перетворення, квантування та ентропійне кодування. Формуючи таким чином основний потік даних, що доступний для інтерпретації стандартним JPEG декодером [46, 87, 89, 127, 138].

Ключовим етапом, що відрізняє JPEG XT від попередників, є процедура обчислення залишкового сигналу (residual signal), яка виконується

безпосередньо в процесі кодування [46, 86, 87, 127, 138]. Для цього система емулює процес декодування: щойно стиснений базовий шар деквантується та проходить через обернене DCT. Після чого реконструйоване LDR-зображення трансформується назад у простір вихідного HDR-сигналу. Шляхом віднімання реконструйованої апроксимації від оригінального вхідного зображення виділяється різницевий сигнал. Він містить інформацію про деталі, що були втрачені під час тонального відображення та квантування базового шару.

Отриманий залишковий сигнал розглядається як окреме зображення, що формує розширювальний шар, і піддається власному циклу кодування, який також базується на ДКП-трансформації. Проте використовує специфічні матриці квантування, оптимізовані для збереження енергії різницевого сигналу. Фінальна стадія обробки даних полягає у мультиплексуванні двох отриманих бітових потоків в єдиний файловий контейнер. Потік базового шару записується як стандартне JPEG зображення. Тоді як потік розширювального шару інкапсулюється у спеціальні маркери метаданих, що дозволяє приховувати додаткову інформацію від застарілого програмного забезпечення. Водночас надаючи сумісним декодерам можливість повної реконструкції вихідного HDR-зображення шляхом сумування базового та розширювального сигналів. Схема основних етапів кодування JPEG XT зображена на рисунку 1.7.



Рисунок 1.7 – Схема основних етапів кодування JPEG XT

Основним недоліком JPEG XT є його обчислювальна складність та менша ефективність стиснення порівняно із більш сучасними методами. Необхідність декодування у процесі кодування створює додаткове навантаження на процесор, що сповільнює передачу даних. Крім того, так як стандарт базується на застарілій архітектурі JPEG, це часто призводить до того, що підсумковий файл стає більшим за аналоги при тій самій візуальній якості [65, 66, 68, 70, 154].

Суттєвий крок уперед у розвитку кодеків став можливим завдяки формату HEIC (High Efficiency Image Coding). Процес перетворення графічної інформації у форматі HEIC базується на архітектурі стандарту вискоефективного кодування HEVC (H.265), використовуючи алгоритми внутрішньокадрового (Intra) стиснення. На відміну від методів з фіксованою блоковою структурою, в HEIC застосовується адаптивна сегментація зображення на дерева кодування (CTU), розмір яких зазвичай становить 64x64 пікселі. Залежно від локальних характеристик текстури, кожний CTU піддається рекурсивному розбиттю за принципом квадродерева на менші одиниці кодування (CU), що дозволяє ефективно обробляти як однорідні ділянки, так і зони з високою деталізацією аж до блоків розміром 4x4 пікселі.

Ключовим механізмом усунення просторової надмірності виступає етап внутрішньокадрового передбачення (Intra-Prediction), де замість безпосереднього збереження значень пікселів виконується їх екстраполяція на основі реконструйованих значень сусідніх блоків. Цей процес оперує розширеним набором із 35 режимів передбачення, що включає планарний (Planar) та постійний (DC) режими для градієнтних і однорідних областей відповідно, а також 33 режими спрямованого передбачення для точного відтворення геометричних структур під різними кутами.

Після формування прогнозного зображення обчислюється різницевий сигнал (residual), який відображає відхилення оригіналу від передбачення, і цей сигнал піддається частотному розкладанню. На даному етапі застосовується дискретне косинусне перетворення (DCT) для більшості блоків, проте для блоків розміром 4x4 впроваджено дискретне синусне перетворення (DST), що забезпечує коректнішу репрезентацію країв об'єктів. Отримані спектральні коефіцієнти проходять процедуру квантування, яка є основним етапом втрати інформації, оскільки точність коефіцієнтів знижується задля зменшення ентропії даних.

Сформований масив квантованих даних та синтаксичних елементів структури квадродерева надходить на вхід ентропійного кодера CABAC (Context-Adaptive Binary Arithmetic Coding). Цей алгоритм виконує завершальне стиснення без втрат. Для покращення якості даних під час декодування застосовуються контурні фільтри. Зокрема фільтр деблокінгу для згладжування артефактів на межах блоків, що дозволяє мінімізувати спотворення, формуючи остаточне реконструйоване зображення. Схема основних етапів перетворення даних при кодуванні HEIC зображена на рисунку 1.8.

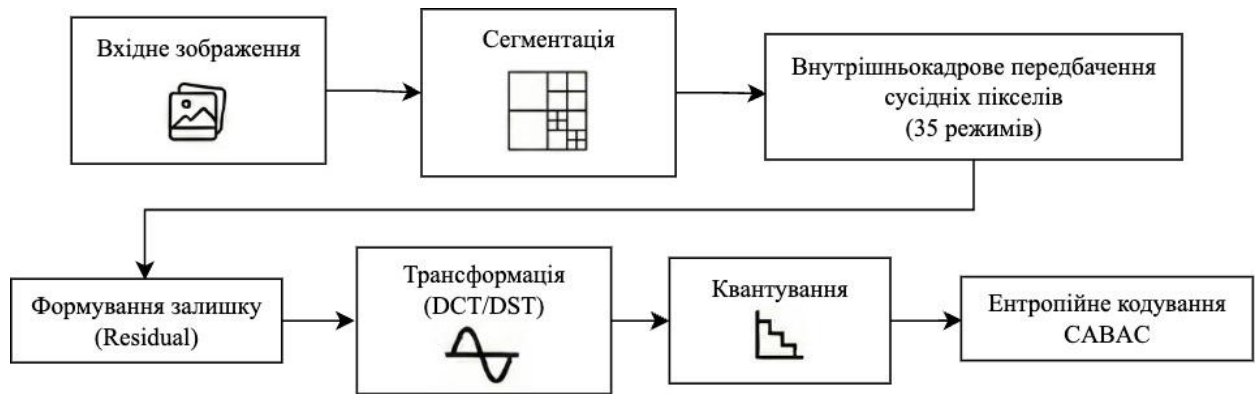


Рисунок 1.8 – Схема основних етапів кодування HEIC

Попри ефективність компресії, HEIC має низку обмежень. Серед яких першочерговим є складна патентно-ліцензійна політика, що ускладнює інтеграцію формату у наукові або промислові рішення. Окрім того, специфіка застосовуваних механізмів пост-фільтрації може негативно впливати на фінальну візуальну якість, спричиняючи втрату високочастотних текстурних деталей та виникнення артефактів неприродної розмитості внаслідок надмірного згладжування корисного сигналу разом із шумом [85].

Як відкрита альтернатива HEIC і подальший крок у розвитку форматів, з'явився JPEG XL. Цей кодек створено з метою забезпечення високої ефективності стиснення, підтримки широкого динамічного діапазону та глибини кольору до 16 біт. При цьому він зберігає зворотну сумісність із стандартним JPEG. JPEG XL поєднує можливості безвтратного та втратного стиснення, й забезпечує підвищену якість реконструкції при низькому бітрейті. Завдяки відкритій ліцензії та сучасній архітектурі він позиціонується як універсальне рішення для майбутніх мультимедійних стандартів. Водночас поширення JPEG XL наразі залишається обмеженим через неповну підтримку з боку браузерів, а також через обмежену кількість готових програмних реалізацій, здатних повною мірою використовувати його потенціал. Додатковим фактором є новизна цього формату, через що його ефективність та доцільність застосування у спеціалізованих галузях поки що не досліджені достатньо глибоко. Зокрема, залишається відкритим питання

практичного використання JPEG XL для обробки інфрачервоних зображень, де висуваються підвищені вимоги до збереження точності термальних сигнатур об'єктів інтересу [51, 163].

Попри наявність нових рішень, одним із найбільш поширених залишається JPEG 2000. На відміну від своїх попередників, він ґрунтується не на ДКП, а на вейвлет-перетворенні. Що зменшує прояв артефактів блокування, вищу якість при великих ступенях стиснення та можливість прогресивного декодування. Крім того, JPEG 2000 додатково підтримує безвтратний режим та 16-бітну глибину. Завдяки цим властивостям JPEG 2000 став поширеним стандартом для зберігання та передачі інфрачервоних зображень. У таблиці 1.2 наведені характеристики загальних кодеків.

Ефективність застосування існуючих методів стиснення для інфрачервоних зображень детально досліджено у роботі K. Mantel та S. Forchhammer (DTU, 2017) [53], де було проведено порівняльний аналіз кодеків JPEG 2000, JPEG XT та HEVC. Експерименти виконувались на 16-бітних теплових зображеннях діапазону LWIR, отриманих із кількох відкритих баз даних. Результати показали, що кодеки JPEG 2000 і HEVC забезпечують значно кращу якість відновлення порівняно з JPEG XT. Зокрема, JPEG 2000 виявився ефективнішим при низьких бітрейтах (до 1,4 bpp), забезпечуючи стабільну якість навіть за сильного стиснення, тоді як HEVC продемонстрував найвищу ефективність у середньому та високому діапазонах бітових швидкостей (1,4–6,5 bpp). Водночас, незважаючи на переваги у якості, застосування HEVC у наукових і промислових системах є обмеженим через його патентну закритість і необхідність ліцензування, що робить JPEG 2000 більш придатним для офіційного та стандартизованого використання в інфрачервоній візуалізації.

Таблиця 1.2 – Характеристики загальних кодеків

Кодек	Бітова глибина	Використані методи обробки	Доцільність кодування ІЧ кадрів
JPEG	8 біт	Трансформація: ДКП. Ентропійне кодування: Хаффман.	Низька доцільність. Придатний лише для 8-бітних кадрів.
WebP	8 біт	Трансформація: ДКП 4×4 або 8×8 . Ентропійне кодування: інтервальне кодування.	Низька доцільність. Придатний лише для 8-бітних кадрів. Ефективний лише для веб-застосувань.
JPEG XL	До 16 біт	Трансформація: ДКП. Ентропійне кодування: асиметричні системи числення (ANS).	Висока потенційна доцільність. Підтримка 16 біт, потенційно висока ефективність. Недостатньо досліджений для ІЧ-застосувань.
JPEG XT	До 16 біт	Трансформація: ДКП (LDR і HDR). Ентропійне кодування: Хаффман.	Обмежена доцільність. Підтримує високу бітову глибину але недостатньо ефективний.
HEVC	До 16 біт	Трансформація: ДКП/ДСП 4×4 – 32×32 . Ентропійне кодування: CABAC.	Обмежена доцільність. Висока ефективність, але обмежена практична придатність через регуляторні обмеження.
JPEG 2000	До 16 біт	Трансформація: вейвлет-перетворення (CDF $9/7$, $5/3$). Ентропійне кодування: EBCOT; Додаткові функції: ROI, прогресивне декодування.	Найвища доцільність серед універсальних кодеків. Забезпечує високу ефективність і розповсюджений для ІЧ-зображень.

Важливим критерієм при оцінці якості тепловізійних ІЧ кадрів є температурна достовірність. Температурна достовірність – це міра відповідності значень інтенсивності у пікселях стиснутого чи обробленого зображення їхнім реальним фізичним значенням. Оскільки, кожне значення пікселя представляє температуру у відповідній позиції кадру, то спотворення пікселів впливає і на температурну оцінку.

Для оцінки відхилення піксельних значень зазвичай використовується середнє квадратичне відхилення (СКВ) [115]. Воно дає усереднене уявлення про відхилення пікселів у всьому зображенні. Однак, для тепловізійних ІЧ зображень важлива температурна достовірність саме тих областей інтересу, які представляють термальні сигнатури об'єктів.

СКВ може маскувати великі локальні викиди похибки. Якщо група пікселів об'єкта сильно спотворені, а велика площа фону відновлена точно, показник СКВ буде низьким, хоча фактична інформативність кадру для задач виявлення та аналізу буде погіршена.

Наочним підтвердженням цієї проблеми є результати порівняльного кодування інфрачервоного зображення популярними алгоритмами, наведені на рисунку 1.9. На оригінальному кадрі (рис. 1.9 а) чітко виділяється теплова сигнатура об'єкта на фоні неба, що характеризується високим контрастом та чіткими межами. При застосуванні стандартних методів кодування JPEG 2000, JPEG XT та HEVC (рис. 1.9 б, в, г відповідно) спостерігається втрата інформативності у семантично важливих областях.

Аналіз метричних показників, наведених на рисунку 1.9, демонструє, що при коефіцієнтах стиснення (КС) від 5.3 до 7.8, значення похибки СКВ сягає рівня 511–528 одиниць. При цьому, у семантично значимих областях СКВ може досягати 1000 і більше. Для 16-бітних термальних даних така похибка може відповідати відхиленню вимірної температури на кілька градусів, що є неприпустимим для задач автоматизованого моніторингу або високоточного наведення.

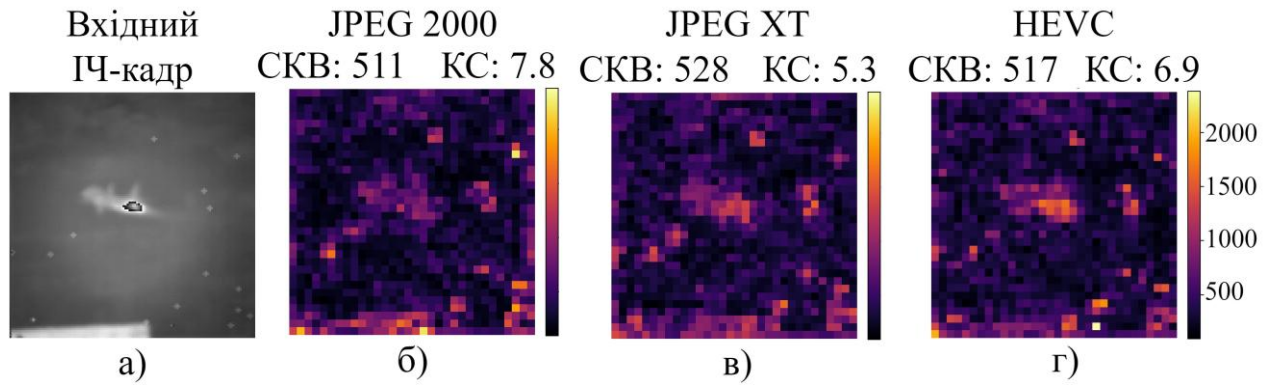


Рисунок 1.9 – Візуалізація просторового розподілу помилки СКВ при стисненні ІЧ-кадру універсальними кодеками: а) вхідний ІЧ-кадр, б) карта СКВ JPEG 2000, в) карта СКВ JPEG XT, г) карта СКВ HEVC

Для оцінки ефективності методів кодування в інформаційних системах необхідно перейти до аналізу бітрейту стисненого потоку кадрів при різних характеристиках генераційного ІЧ-пристрою.

Обернено пропорційний бітрейт R_{raw} сирого потоку даних до коефіцієнту стиснення CS дорівнює бітрейту R_{comp} стисненого потоку кадрів:

$$R_{\text{comp}} = \frac{R_{\text{raw}}}{CS}. \quad (1.2)$$

На рисунку 1.10 зображено графік залежності бітрейту R_{comp} від роздільної здатності та кадрової частоти генераційного ІЧ-пристрою, при кодуванні методом JPEG 2000 з коефіцієнтом стиснення $CS=7.8$. Наведені результати підтверджують, що зростання просторових і часових параметрів сенсорів формує потік, інтенсивність якого залишається високою навіть після застосування стиснення. Спроби збільшення рівня стиснення стандартними методами будуть призводити до більшої деградації локальних контрастів об'єктів інтересу. Такий вплив погіршує термальну цілісність семантично важливих сегментів кадру.

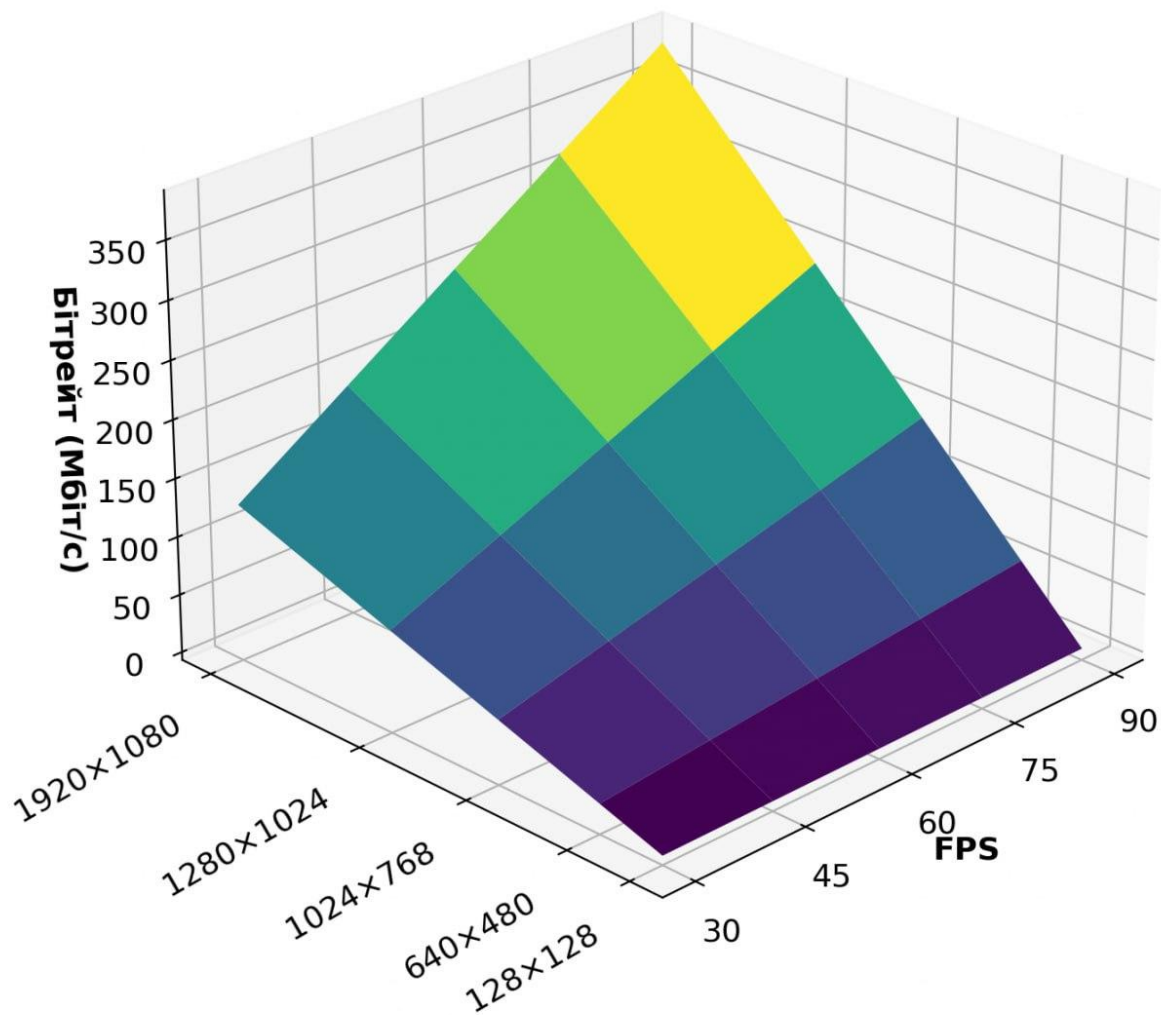


Рисунок 1.10 – Графік залежності бітрейту R_{comp} від роздільної здатності та кадрової частоти генераційного ПЧ-пристрою

Із наведених даних полягає висновок: вимогою до розроблюваного методу є забезпечення бітрейту, що не поступається показникам існуючих аналогів. При цьому метод повинен забезпечувати збереження термальної цілісності важливих сегментів у кадрі.

Узагальнюючи проведений аналіз методів кодування ПЧ-кадрів, можна стверджувати, що наявні рішення не повною мірою задовольняють вимоги сучасних інформаційних систем. Пропріетарні спеціалізовані формати забезпечують необхідну термальну точність [108]. Проте обмеженість інформації щодо їх реалізації та відсутність стандартизації ускладнюють

інтеграцію в гетерогенні системи обробки даних. Методи безвтратного стиснення зберігають метричну достовірність, але не забезпечують достатнього рівня стиснення. Універсальні стандарти стиснення з втратами, хоча й підтримують необхідну бітову глибину (до 16 біт), оптимізовані насамперед під особливості зорового сприйняття людини. Як показав проведений аналіз, вони схильні згладжувати високочастотні деталі. Це призводить до внесення спотворень у теплові сигнатури об'єктів інтересу навіть за низьких показників середньої квадратичної похибки по всьому кадру.

Для розв'язання цієї проблеми необхідно перейти до адаптивних методів, які здатні враховувати структуру та зміст сцени [39, 67]. Перспективним інструментом для реалізації такої адаптивності є методи штучного інтелекту та машинного навчання, здатні виявляти складні залежності в даних [76, 91].

1.4 Аналіз доцільності застосування нейромережових технологій для підвищення ефективності обробки інфрачервоних зображень

На відміну від кодеків із жорстко фіксованою структурою перетворень, нейронні моделі у задачах навченого стиснення (learned image compression, LIC) формують параметричне представлення даних шляхом навчання на репрезентативних вибірках. У процесі навчання моделей виконується оптимізація за критеріями, що включають бітовий об'єм та спотворення. Такі моделі здатні наближено відтворювати складні статистичні, просторові та навіть перцептивні закономірності вхідних зображень. Це відкриває можливість побудови кодеків, у яких структура простору ознак та схема кодування формуються адаптивно, а не задаються наперед жорстко фіксованими перетвореннями. Нейронні мережі дозволяють виявляти закономірності не через явне аналітичне моделювання сигналу чи шуму, а через навчання на даних, що відображають температурні розподіли та типові

текстури. Завдяки цьому можливе формування представлень, які краще узгоджуються з структурою ІЧ-даних, ніж існуючі перетворення загального призначення.

Ще однією важливою перевагою глибинних архітектур є здатність до автоматичного виділення семантично значущих областей [131, 153]. Нейронні моделі здатні враховувати просторові залежності на різних масштабах, виявляти контури об'єктів, зони локального перегріву, теплові сигнатури цілей. Це дозволяє реалізовувати неоднорідне кодування. Забезпечуючи більшу точність реконструкції в об'єктах інтересу та агресивніше стиснення у фонових областях. Додатково, нейронні декодери можуть виконувати постобробку: зменшення артефактів, відновлення згладжених контурів, уточнення температурних градієнтів [71, 79].

Сучасні нейромереві підходи до кодування можна умовно розділити на дві групи. Перша – це використання нейронних мереж як повноцінної заміни існуючих кодеків. У цьому підході основний процес кодування й декодування моделюється єдиною нейронною системою, яка навчається представляти дані у стислому вигляді. Такі моделі зазвичай побудовані на основі архітектури автокодувальників. Вона складаються з:

- енкодера, який перетворює зображення у компактне латентне представлення;
- декодера, який реконструює кадр із латентного представлення мінімізуючи втрати.

Типову архітектуру згорткового автокодувальника зображено на схемі рисунку 1.11. Де схематично зображено, як розмірність даних зменшується, при наближенні до шару латентного представлення. На відміну від існуючих кодеків, у навченому стисненні більшість етапів обробки кадру оптимізуються в межах однієї моделі за допомогою градієнтного навчання.

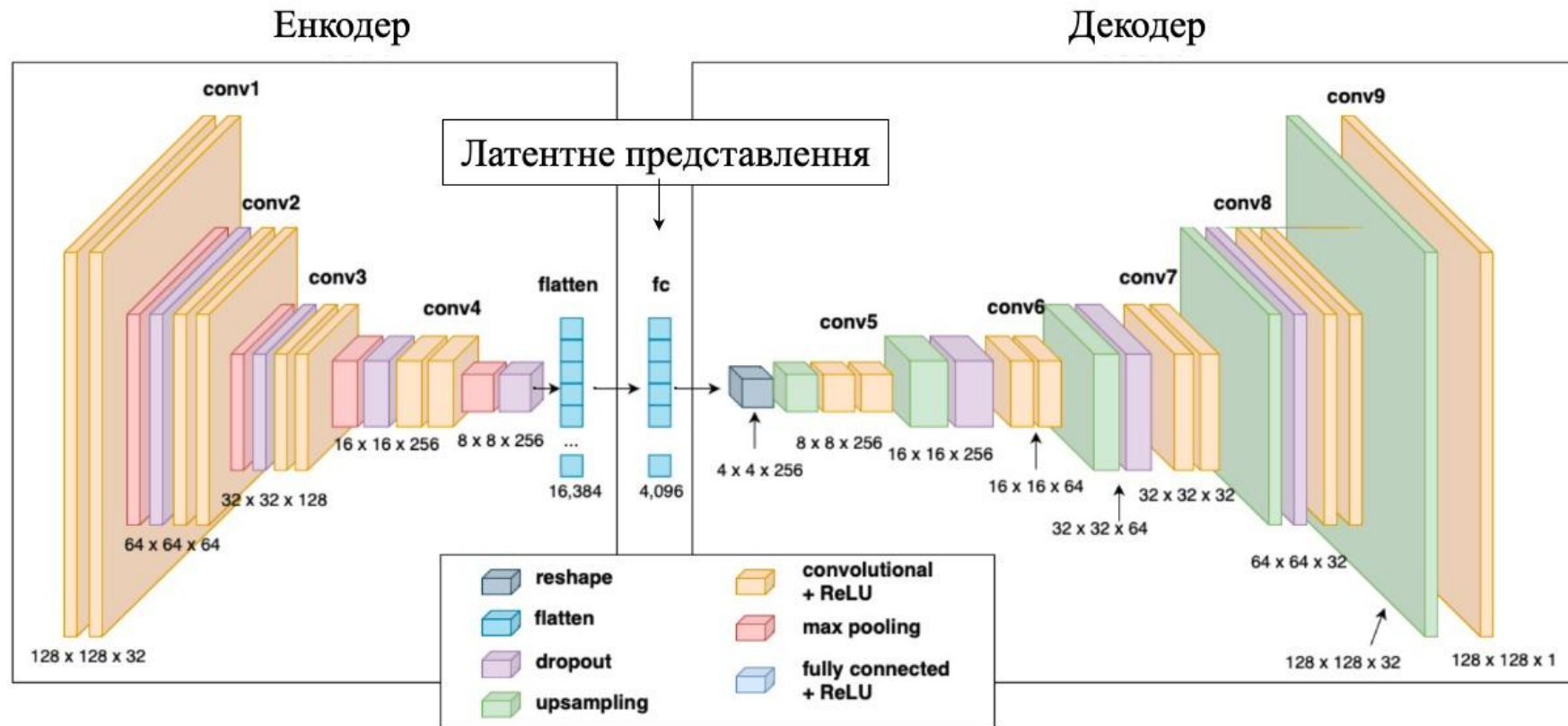


Рисунок 1.11 – Схема архітектури згорткового автокодувальника

Суттєвий розвиток забезпечили моделі Ballé et al. (Google Research), які використали концепцію варіаційного автокодувальника (Variational Autoencoder, VAE). На відміну від звичайного автокодувальника, VAE не просто стискає дані до латентного простору, а навчається ймовірнісному опису цього простору. Енкодер у такій системі не повертає фіксований вектор ознак, а формує параметри розподілу (середнє і дисперсію), з якого вибірка латентних змінних здійснюється стохастично. Це дозволяє моделі краще узагальнювати закономірності сигналу та передбачати невизначеність при реконструкції.

Щоб підвищити ефективність ентропійного кодування, Ballé запровадив гіперапріорну модель (hyperprior model) – додаткову нейронну мережу, яка оцінює статистику латентних змінних (їх дисперсію або шумовий рівень) і передає ці оцінки декодеру як допоміжну інформацію [113]. Такий підхід дає змогу більш точно моделювати розподіл ймовірностей, що безпосередньо визначає ступінь стиснення, зменшуючи надлишковість у представленнях.

Подальші дослідження, зокрема [150], розвинули цю ідею, запровадивши контекстне моделювання (context modeling), яке враховує просторові залежності між сусідніми латентними елементами. Якщо у попередніх підходах ймовірність кожного коефіцієнта розглядалася незалежно, то в контекстних моделях вона обчислюється з урахуванням уже закодованих сусідів, що підвищує точність передбачення і зменшує обсяг переданої інформації. Додатково використовуються мультишарові гіперпріори, що дозволяють декодеру отримувати ієрархічну інформацію про структуру зображення – від глобальних контрастів до локальних деталей [159].

Слід відзначити, що у відповідь на потребу стандартизації кодеків на основі нейронних мереж, вже з'явилися перші ініціативи. Найбільш показовим прикладом є JPEG AI, який розробляється як новий міжнародний стандарт для стиснення зображень із використанням глибинного навчання.

Його принцип роботи ґрунтується на автокодувальній архітектурі, де квантування та ентропійне кодування інтегруються у навчальний процес. У результаті кодування формується латентне представлення, яке дозволяє стискати інформацію більш ефективно. Перевершуючи існуючі кодеки JPEG чи подібні, і водночас забезпечуючи збереження якості. Аналогічні ідеї реалізуються в експериментальних системах для відео, зокрема у Deep Contextual Video Compression (DCVC), що поєднує варіаційні автокодувальники з механізмами уваги для ефективного моделювання просторово-часових залежностей. Варто підкреслити, що спільною рисою цих рішень є орієнтація на практичне використання, зокрема шляхом поступового формування єдиних технічних вимог, які б забезпечували сумісність між різними реалізаціями.

У результаті такі нейронні системи не лише перевершують існуючі кодеки за перцептивною якістю при однаковому бітрейті, але й демонструють здатність до адаптації під специфіку даних. Для інфрачервоних зображень це означає можливість перенавчання моделей на даних з різними спектральними характеристиками, що відкриває перспективу створення навчених ІЧ кодеків (learned IR-codecs).

Інший напрям застосування нейронних мереж полягає не у повній заміні існуючих кодеків, а у їх доповненні спеціалізованими нейронними модулями, здатними покращити результати існуючих алгоритмів стиснення. Такий підхід є еволюційним продовженням розвитку систем кодування, оскільки дозволяє інтегрувати досягнення штучного інтелекту у вже існуючі стандарти без потреби повного оновлення інфраструктури чи зміни формату даних. Нейронні моделі можуть бути включені на різних етапах процесу кодування – від попередньої обробки сигналу до післядекодуювальної реконструкції, виконуючи роль інтелектуальних фільтрів або адаптивних оптимізаторів [97].

Одним із найпоширеніших напрямів є використання мереж для постобробки зображень. Такі нейромережеві моделі застосовуються після

декодування для усунення артефактів, спричинених сильним стисненням. Одні з відомих прикладів таких мереж:

- ARCNN (Artifact Reduction Convolutional Neural Network) - створена для усунення артефактів JPEG-стиснення, таких як блоковість і розмиття, які з'являються після збереження зображень у форматі JPEG з низькою якістю;

- DnCNN (Denoising Convolutional Neural Network) - універсальніша модель, ніж ARCNN. Вона призначена для усунення різних типів шумів, включаючи гаусовий шум, JPEG-артефакти.

На рисунку 1.12 зображено приклад застосування DnCNN. Де відновлене зображення (Рис. 1.12 в) має високу ступінь подібності до оригінального зображення (Рис. 1.12 а).

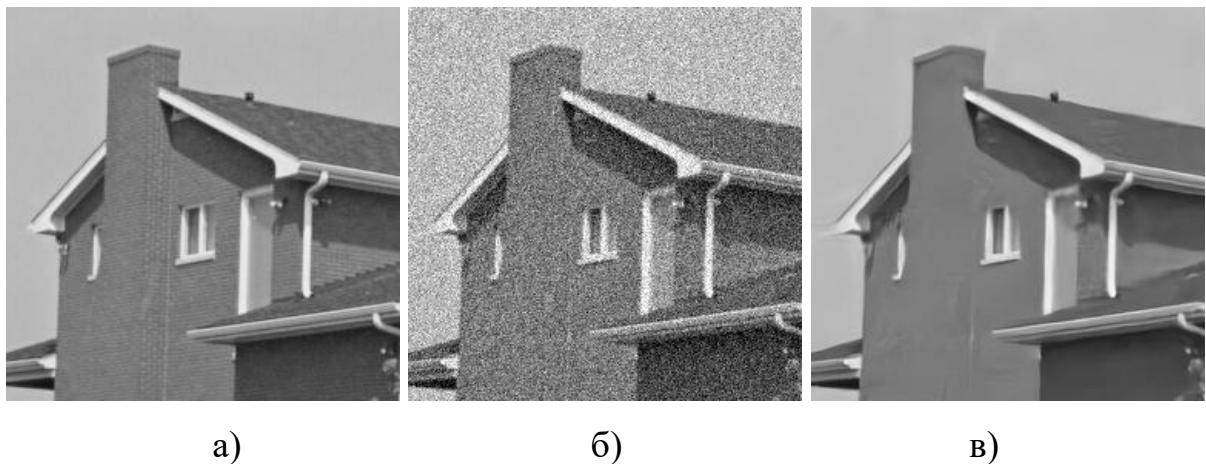


Рисунок 1.12 – Приклад застосування DnCNN: а) оригінальне зображення, б) зображення із шумовими спотвореннями, в) відновлене зображення застосовуючи DnCNN

Подібним чином нейронні фільтри інтегруються у сучасні стандарти, зокрема VVC/H.266. Вони замінюють існуючі методи постобробки більш складними моделями. Це дозволяє краще відновлювати текстури та деталі при зменшеному бітрейті.

Важливим практичним напрямом є також супер-роздільна здатність (super-resolution), де нейронні мережі (наприклад, EDSR або SRGAN)

використовуються для підвищення просторової роздільності кадру після декодування. Це особливо актуально для ІЧ-камер із обмеженою роздільною здатністю. Де апаратне обмеження роздільності компенсується програмною реконструкцією високочастотних деталей. Це не виключає додаткового застосування denoising-мереж, які адаптивно пригнічують сенсорний шум або перешкоди, спричинені компресією. При застосуванні до ІЧ зображень, такі моделі здатні враховувати тип шуму, що притаманний певному сенсору або умовам зйомки, і відновлювати структуру температурних полів, зменшуючи втрати термальної достовірності.

Не менш перспективним напрямом є реконструкція зображень (image inpainting). Це метод відновлення втрачених або відкинутих при стисненні ділянок кадру. Може бути особливо корисним у випадках, коли у процесі кодування відбувається агресивне видалення неінформативних областей, наприклад фонових зон. Серед моделей, що можуть відновлювати такі зони розповсюдженими є: моделі на основі генеративно змагальних мереж (Generative Adversarial Networks, GAN), дифузійні моделі.

Ще одним перспективним напрямом розвитку є застосування нейронних мереж для виявлення об'єктів та семантичної сегментації. Це відкриває можливість переходу до контекстно-орієнтованого або семантичного кодування. У цьому підході зображення попередньо аналізується нейронною моделлю, яка виконує розподіл кадру на області різної інформативності. Наприклад об'єкти, що мають високу діагностичну або аналітичну цінність, та фонові зони, які не несуть суттєвої інформації. Для інформативних областей застосовується високоякісне або майже безвтратне кодування. Тоді фонові зони можуть стискатися більш агресивно. Такий підхід не лише підвищує ефективність розподілу бітового об'єму, а й дозволяє забезпечити збереження найцінніших температурних і структурних деталей. У контексті ІЧ-зображень це особливо важливо для систем моніторингу. Де точне відтворення меж і температурних контрастів об'єктів є критичним, тоді як фонові області можуть бути спрощені без шкоди для

подальшої інтерпретації. Для реалізації таких механізмів застосовуються сучасні архітектури нейронних мереж, наприклад: U-Net, Mask R-CNN, SegFormer. Вони здатні виконувати точну сегментацію навіть у семантично насичених сценах. Таким чином, поєднання сегментаційних моделей з існуючими кодами створює передумови для розроблення семантично орієнтованих методів стиснення. У таких методах якість реконструкції керується не лише статистичними критеріями, а й можливостями до виявлення окремих областей зображення [77].

Отже, модульне доповнення існуючих методів стиснення нейронними мережами створює ефективний компроміс між стабільністю перевірених алгоритмів і гнучкістю сучасних інтелектуальних систем. Це дозволяє поступово інтегрувати досягнення штучного інтелекту у існуючі стандарти кодування, розвиваючи їх у бік більш адаптивних систем без радикального перегляду існуючих технологічних основ.

Однак, широке впровадження нейромережових моделей у галузь ІЧ-обробки супроводжується низкою обмежень. Насамперед це значні обчислювальні витрати, які виникають як на етапі навчання моделей, так і під час їх функціонування у режимі реального часу. Що ускладнює використання таких моделей на бортових платформах реального часу, характерних для дистанційних ІЧ-систем (БПЛА, мобільні роботизовані комплекси). Крім того, важливим недоліком є низький рівень інтерпретованості результатів, що ускладнює процес налагодження та обмежує їх використання у сферах, де потрібна прозорість прийняття рішень. Не менш суттєвим недоліком є залежність від якості та обсягу навчальних даних, адже упередженість або обмеженість вибірки безпосередньо відображається на результатах роботи моделі. Існує дефіцит відкритих високоякісних баз ІЧ-зображень, які б охоплювали різні спектральні діапазони у різних умовах оточуючого середовища. Висока варіативність ІЧ-даних - відмінності у сенсорах, оптиці, температурному фоні та рівнях шуму - ускладнює узагальнення нейронних моделей.

До цього слід додати наявність ще одного недоліку – можливість виникнення так званих галюцинацій. На відміну від існуючих кодеків, де після стиснення основним наслідком є деградація якості (розмиття, блокові артефакти), нейромережеві методи іноді можуть породжувати у процесі декодування фрагменти з недостовірною або вигаданою інформацією. Це пов'язано з генеративною природою моделей, які здатні «домислювати» відсутні деталі на основі статистики навчальних даних. Така особливість є критичною у застосуваннях, де від збереження достовірності даних залежить безпека чи точність прийняття рішень.

Зокрема, відомо, що JPEG AI вносить суттєві спотворення для регулярних структур та тексту на кадрі. На рис. 1.13 зображено такий приклад. Можна побачити зміну семантичної структури білої ділянки радіаторної сітки автомобіля. На оригінальному зображенні (Рис. 1.13 а) вона була вертикальною, а на декодованому зображенні (Рис. 1.13 б) стала горизонтальною.



а)

б)

Рисунок 1.13 – Приклад галюцинації JPEG AI спричинений регулярною структурою радіаторної сітки автомобіля: а) оригінальне зображення, б) закодоване та декодоване зображення із використанням JPEG AI

На іншому прикладі (Рис. 1.14), можна побачити спотворення кадру після використання JPEG AI, що призвело до суттєвої втрати

інформативності. Наведене спотворення викликане наявністю регулярної структури LED екрану на зображенні.

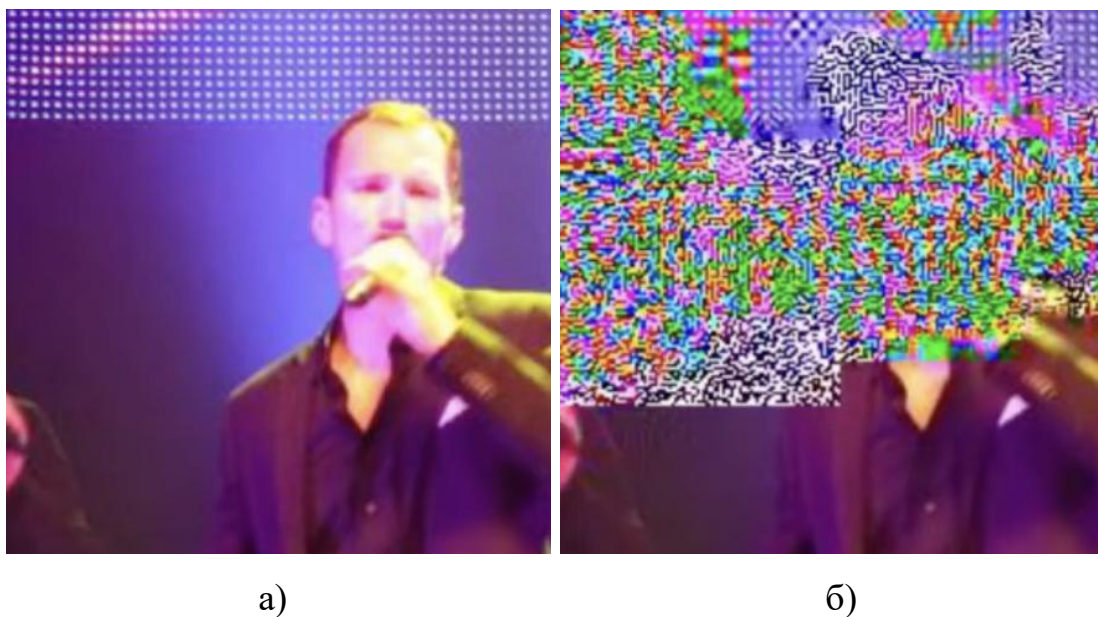


Рисунок 1.14 – Приклад галюцинації JPEG AI спричинений регулярною структурою LED екрану на зображенні: а) оригінальне зображення, б) закодоване та декодоване зображення із використанням JPEG AI

Аналіз сучасних підходів до стиснення інфрачервоних зображень виявив обмеженість наявних рішень. Стандартизовані та нейромережеві методи не забезпечують необхідного балансу між ступенем стиснення, точністю реконструкції та універсальністю. Зокрема, застосування нейронних мереж ускладнюється такими факторами: значні обчислювальні витрати, дефіцит навчальних даних та нестабільність результатів. Це зумовлює дисбаланс між вимогами до точності й компактності бітового представлення даних та фактичними можливостями існуючих методів.

Вказані недоліки обґрунтовують доцільність розробки спеціалізованого методу кодування ІЧ-кадрів. Метод повинен інтегрувати спектрально-ентропійний аналіз та адаптивні механізми штучного інтелекту. Такий підхід забезпечить оптимізацію параметрів стиснення на основі статистичних, спектральних і семантичних характеристик кадру. Отже, *мета досліджень*

полягає у розробці методу кодування ІЧ-зображень для зниження бітрейту без втрат термальної цілісності на основі семантико-орієнтовної обробки з використанням моделей штучного інтелекту.

Для досягнення поставленої мети необхідно вирішити наступні завдання:

1. Розробити метод формування двопластової інформативної карти ІЧ-зображень для забезпечення умов щодо зменшення бітрейту з визначеною термальною цілісністю на основі використання моделей ШІ.

2. Створити метод спектрально-групового стиснення сегментів бітових пластів ІЧ-зображень для зменшення бітового об'єму без втрат термальної цілісності

3. Розробити метод семантико-орієнтовного кодування ІЧ-зображень для передачі в інформаційних системах.

4. Створити метод для підвищення термальної цілісності об'єктно-інформативних сегментів ІЧ-зображень без втрат бітрейту в інформаційних системах.

5. Здійснити порівняльну оцінку методів кодування ІЧ-зображень за абляційною методикою.

Висновки за першим розділом

1. Проведений аналіз сучасних технологій стиснення та кодування інфрачервоних зображень показав зростання ролі інтелектуальних інформаційних систем у задачах дистанційного моніторингу, медичної діагностики та промислового контролю. У зв'язку з підвищенням вимог до точності вимірювань, швидкодії передачі та збереження достовірності об'єктів інтересу, питання ефективного кодування ІЧ-даних набуває особливої наукової і практичної актуальності.

2. Встановлено, що серед загальних методів стиснення найбільш розповсюдженими для ІЧ-зображень залишаються JPEG 2000, JPEG XT та

HEVC. Однак, їхня архітектура орієнтована переважно на обробку сигналів у видимому спектрі. Такі кодеки не враховують специфіку теплових зображень та потребу у збереженні значень важливих термальних сигнатур, що призводить до втрат інформативності при сильному стисненні.

3. Виявлено, що пропрієтарні ПЧ-кодеки (R.JPEG, IS2, SEQ, FFF, IRB, IMG, IRI, DAT, ANA), хоч і забезпечують термальну цілісність кадрів, є несумісними між собою та не стандартизованими. Крім того, інформація обмежена про деталі їх реалізації та експериментальні дані. Це обмежує їхнє широке використання у відкритих інформаційних системах, наукових дослідженнях і промислових комплексах.

4. Розглянуті нейронні методи кодування зображень продемонстрували значний потенціал у підвищенні ефективності стиснення та адаптації до властивостей даних. Показано, що глибокі нейронні мережі здатні аналізувати складні семантичні залежності, виконувати адаптивне кодування та відновлення.

5. Узагальнено, що існує два основні напрями інтеграції штучних нейронних мереж у систему кодування:

1) доповнення існуючих методів нейронними модулями для підвищення якості реконструкції (denoising, super-resolution, inpainting, виявлення об'єктів інтересу);

2) створення повністю нейронних схем кодування (learned codecs), у яких процеси аналізу, стиснення та реконструкції реалізуються єдиною мережею або групою мереж.

6. Встановлено основні виклики впровадження нейронних методів у кодування ПЧ-зображень: висока обчислювальна складність, дефіцит великих тепловізійних баз даних, складність інтерпретації роботи моделей. Ці чинники поки що обмежують практичне використання таких рішень у реальних інформаційних системах.

7. Виявлене протиріччя між вимогами до збереження термальної цілісності та зменшенням бітрейту передачі інфрачервоних зображень. Для

його усунення необхідна розробка нового методу кодування, який би поєднував переваги спектральних та структурно-статистичних підходів із інтелектуальними можливостями нейромережових моделей, забезпечуючи збереження температурної достовірності, структурної цілісності та ефективність стиснення даних.

Результати досліджень щодо матеріалів, викладених в першому розділі, опубліковані в таких наукових працях як: [6, 7, 8, 9].

РОЗДІЛ 2. РОЗРОБКА МЕТОДУ ПОБУДОВИ ДВОПЛАСТОВОЇ ІНФОРМАТИВНОЇ КАРТИ ІНФРАЧЕРВОНОГО КАДРУ НА ОСНОВІ МОДЕЛЕЙ СТАТИСТИЧНО-ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ

В розділі розроблено метод побудови двошарової інформативної карти інфрачервоного кадру на основі моделей статистично-інтелектуального аналізу.

Обґрунтована потреба та принципи семантико-орієнтованого підходу до кодування ІЧ-кадрів, що засновуються на сегментації та визначенні рівня інформативності сегментів.

Розглянуто методи обробки зображень із розширеним динамічним діапазоном з позиції кодування термальних кадрів глибиною 16 біт.

Розроблено метод біфуркації бітових пластів. Метод дозволяє підвищити структурну сумісність даних, відділяючи семантично значущий старший бітовий пласт від молодшого бітового пласту, що описує малопомітні термальні флуктуації.

Розроблено метод визначення метрики інформативності на основі структурно важливих ознак сегментів. Метод базується на нормованій агрегації високочастотних спектральних коефіцієнтів сегментів.

Розроблено метод формування двошарової інформативної карти. Метод враховує рівень інформативності бітових пластів, що визначений на основі адаптивних порогових правил.

Розроблено метод виявлення об'єктно-інформативних сегментів на основі застосування моделей штучного інтелекту. Метод дозволяє інтелектуально визначити області кадру, що відповідають об'єктам інтересу.

2.1 Обґрунтування принципів семантико-орієнтованої обробки та підходів до сегментації інфрачервоних кадрів

Під семантичними ознаками інфрачервоної сцени у контексті даної роботи розуміються такі елементи зображення, які безпосередньо використовуються оператором або автоматизованою системою для прийняття рішень [143]: теплові сигнатури цілей, ділянки перегріву обладнання, дефектні області в матеріалах [136], контури конструктивних елементів, температурні аномалії на фоні [147]. На відміну від фонових зон з відносно однорідними температурними полями, ці області характеризуються локально підвищеною структурною і температурною інформативністю [85]. Саме вони визначають корисність ІЧ-кадру в задачах виявлення, діагностики, моніторингу та контролю [129]. При кодуванні сегменти, що містять семантичні ознаки, повинні розглядатися як значно цінніші, ніж статистично подібні фонові структури.

Розповсюджені підходи до оцінювання якості відновлених зображень базуються переважно на глобальних статистичних метриках, таких як середньоквадратична похибка (RMSE), пікове відношення сигнал/шум (PSNR) [115] або перцептивно наближені показники на кшталт SSIM [10]. Ці метрики оцінюють помилки рівномірно по всьому полю зображення, не враховуючи, у якій саме області вони виникають. Для інфрачервоних даних такий підхід є обмеженим. Однакова за величиною похибка у фоні та на межі теплової сигнатури має зовсім різний вплив з погляду прикладної задачі. Для оператора або алгоритму виявлення критичною є коректність локального контрасту, форми та положення об'єкта інтересу [88], тоді як несуттєве спотворення фону зазвичай не має впливу. Отже, орієнтація лише на глобальні статистичні критерії може призводити до нераціонального розподілу бітового об'єму [20].

Семантично неконтрольовані спотворення негативно впливають на результативність прикладних задач. У системах спостереження та

моніторингу зниження локального контрасту в області цілі безпосередньо зменшує дальність виявлення та ймовірність коректної ідентифікації об'єкта, особливо за умов маскування або близьких до фонових температур. У технічній діагностиці обладнання спотворення величини і форми теплової плями може призвести до хибної оцінки рівня нагріву, неправильного визначення площі перегрітої ділянки чи ігнорування ранніх стадій дефектів [101]. У задачах автоматизованого аналізу (детекція, сегментація, трекінг) додаткові артефакти або втрата дрібномасштабних структур здатні суттєво погіршити роботу алгоритмів, навіть якщо глобальний PSNR залишається прийнятним [22]. Це підтверджує, що мінімізація усередненої помилки не є достатньою умовою збереження корисного навантаження ІЧ-кадрів [19].

З урахуванням наведеного можна сформулювати принцип семантично орієнтованого кодування [9]: ступінь допустимих спотворень і щільність бітового представлення мають визначатися не лише просторовими параметрами сигналу, а й семантичною значущістю відповідних областей для цільової задачі. Для семантично важливих регіонів необхідно забезпечити підвищену термальну цілісність [104], що включає: мінімальні похибки значень, відтворення локальних градієнтів і контурів, відсутність штучних високочастотних артефактів, які можуть бути хибно інтерпретовані. Натомість для фонових областей допускається агресивніше стиснення, більш сильне згладжування, за умови збереження глобальної узгодженості температурного поля [24].

Така диференціація забезпечується за рахунок сегментації кадру [17] та наявності формалізованої процедури оцінки інформативності сегментів [100]. У процесі сегментації ІЧ-кадр повинен бути розділений на непересічні сегменти пікселів [118]. Для зменшення складності методу, пропонується виконувати рівномірну сегментацію [27], що забезпечує однаковий розмір блоків, які у сукупності утворюють сегментну сітку [29].

Механізм сегментації та оцінки інформативності сегментів відкриває можливості для підвищення інформаційної безпеки обміну даними.

Адаптивне застосування криптографічних методів на основі рівня інформативності сегментів може дозволити підвищити криптостійкість даних.

Вибір розміру сегменту впливає на такі характеристики подальшого кодування, як:

- складність обробки – чим більшу кількість даних вміщує сегмент, тим більша кількість операцій та апаратної пам'яті необхідні для його обробки;
- деталізованість кодування – чим менший розмір сегментів, тим точніша диференціація кадру, що дозволяє ретельніше опрацьовувати дрібні семантичні структури. Однак при цьому можуть не враховуватись більш глобальні характеристики сигналу.

Експериментальним чином було визначено, що для ІЧ-кадрів збалансованим розміром сегменту є 8x8 [59, 99, 102, 110]. Розглянемо вхідне інфрачервоне зображення з шириною K та висотою L , як двомірну матрицю $A^{(K, L)}$. На основі цієї матриці формується сітка сегментів, кожен з яких описується як $A^{(k, l)}$, що відповідає блоку розміром 8x8 елементів. Де (k, l) визначають положення сегмента у загальній структурі кадру. Отже усі сегменти $A^{(k, l)}$ представляють сукупний кадр:

$$A^{(K, L)} = \left\{ A^{(k, l)} \right\}. \quad (2.1)$$

Узагальнюючи, можна констатувати, що врахування семантичних ознак інфрачервоної сцени при формуванні бітового потоку є необхідною умовою для узгодження вимог до зменшення бітрейту з реальними задачами моніторингу, виявлення та діагностики [21]. Застосування семантично орієнтованих механізмів дозволяє забезпечити пріоритетне збереження найбільш важливих для аналізу областей без надмірних витрат ресурсів на фонові зони.

2.2 Обґрунтування підходів до обробки 16-бітних даних та аналіз методів HDR-перетворень

Важливою особливістю даних, отриманих від сучасних інфрачервоних сенсорів, є їхній широкий динамічний діапазон. На відміну від типових зображень видимого спектра, що зазвичай представляються 8 бітами на піксель (256 рівнів яскравості), тепловізійні ІЧ-дані оперують з глибиною до 16 біт (65 536 рівнів) [144]. Цей розширений діапазон несе фундаментальний фізичний сенс: він безпосередньо корелює з температурною чутливістю сенсора. Кожен крок у динамічному діапазоні відповідає мінімальній вимірюваній різниці температур, що дозволяє фіксувати дрібні теплові флуктуації. Тож ІЧ-зображення з глибиною 16 біт містять як метрично значущу температурну інформацію [106], так і деталізуючі та шумові компоненти. Вимогою до спеціалізованого методу кодування ІЧ-зображень є здатність коректно обробляти та зберігати дані з глибиною 16 біт на піксель [37].

Пряме опрацювання 16-бітних даних у єдиному просторі ускладнює побудову гнучких схем стиснення [31], які могли б по-різному обробляти більш і менш значущі складові сигналу. Крім цього, існує потреба у забезпеченні апаратної підтримки 8-бітних операцій. Для ефективної роботи за умов обмеженої обчислювальної інфраструктури [62]. Оскільки зазвичай шини даних та процесорні команди розраховані на 8 бітні потоки даних [16].

Розповсюдженими є підходи до масштабування 16-бітних даних у 8-бітні. Як наслідок такої операції, утворені 8-бітні дані мають втрачену інформацію, тобто створення. Для компенсації цих спотворень відбувається розрахунок залишкової матриці, яка представляє із себе різницю між оригінальним кадром, та зворотно масштабованим.

Застосовуються різні методи масштабування. Найпростішим є лінійне, яке рівномірно відображає вхідний 16-бітний діапазон у цільовий [50, 142]. Перевагою є обчислювальна простота та збереження лінійності даних, однак

недоліком є вкрай низька точність, якщо значуща інформація сконцентрована лише у вузькій частині повного діапазону. Це призводить до неоптимального розподілу у рамках бітового об'єму.

Більш поширеним для візуалізації є логарифмічне або степеневе кодування, що імітує нелінійність людського зору, надаючи більше біт для темних ділянок і менше – для яскравих. До цього підходу відносяться методи тонального відображення (Tone Mapping Operators, ТМО). Вони адаптуються до локального вмісту зображення і забезпечують перцептивну якість для 8-бітних дисплеїв, а не збереження метричних даних.

Підходи із нелінійним масштабуванням корисні для перцептивного відображення, однак вони спотворюють вихідні термальні величини. Крім того, побудована залишкова матриця не відображає семантики зображення. Натомість, вона містить лише статистичну інформацію щодо втрат, що зменшує потенціал до її адаптивної обробки. Наведені недоліки обмежують використання цих розповсюджених підходів до обробки ІЧ-кадрів [103].

У зв'язку з цим доцільним є введення етапу перетворення, який би забезпечив відокремлення структурно важливих і залишкових компонент ІЧ-сигналу при збереженні точності та оборотності даних.

Для вирішення цієї задачі в даній роботі пропонується метод біфуркації бітових пластів [6], який виконує безвратне розділення 16-бітного кадру на два 8-бітні пласти: пласт старших бітів (MSB) та пласт молодших бітів (LSB) [15]. Цей підхід спирається на відому властивість двійкової системи числення [7]: старші біти визначають основну частину амплітуди сигналу та глобальну структуру сцени, тоді як молодші біти відповідають за тонкі варіації, текстури та шум. Реалізація цієї ідеї дозволяє отримати дві окремі 8-бітні матриці, що є структурно сумісними з подальшими етапами кодування та можуть оброблятися диференційовано [25].

2.3 Розробка методу біфуркації бітових пластів для підвищення структурної сумісності даних при подальшій обробці

В ході методу біфуркації бітових пластів виконується екстракція старших бітових значень та молодших [124]. Виразимо кожне піксельне значення як $A^{(k,l)}(q,s)$, де (q, s) є координатами значення в рамках сегменту. Для виконання екстракції старших біт, пропонується представити кожне 16-бітне значення вхідного зображення $A^{(k,l)}(q,s)$ у вигляді бітової послідовності. До даної бітової послідовності необхідно застосувати операцію побітового зсуву вправо на 8 значень. Таким чином молодші бітові значення відсікаються та зберігаються лише старші. Це зображено у наступній формулі:

$$\text{Amsb}^{(k,l)}(q,s) = \text{MSB}\left(A^{(k,l)}(q,s)\right) = A^{(k,l)}(q,s) \gg 8. \quad (2.2)$$

Де:

- $\text{Amsb}_{r,t}^{(k,l)}(q,s)$ – результат екстракції старших біт;
- MSB – процедура екстракції старших біт;
- \gg – операція побітового зсуву вправо.

Застосовуючи MSB до усіх значень з сегменту $A^{(k,l)}$ отримаємо сегмент старших бітових значень $\text{Amsb}^{(k,l)}$:

$$\text{Amsb}^{(k,l)} = \text{MSB}\left(A^{(k,l)}\right) = \left\{ \text{MSB}\left(A^{(k,l)}(q,s)\right), A^{(k,l)}(q,s) \in A^{(k,l)} \right\}. \quad (2.3)$$

Для виконання екстракції молодших біт необхідно відкинути старші бітові значення та зберегти лише молодші. Для цього пропонується

застосувати побітову кон'юнкцію молодшої бітової маски LSB_M зі значенням $A^{(k,l)}(q,s)$:

$$LSB_M = 0000000011111111;$$

$$Alsb^{(k,l)}(q,s) = LSB\left(A^{(k,l)}(q,s)\right) = A^{(k,l)}(q,s) \& LSB_M. \quad (2.4)$$

Де:

- $Alsb^{(k,l)}(q,s)$ – результат екстракції молодших біт;
- LSB – процедура екстракції молодших біт;
- LSB_M – молодша бітова маска;
- $\&$ – операція кон'юнкції.

Застосовуючи LSB до усіх значень з сегменту $A^{(k,l)}$ отримаємо сегмент молодших бітових значень $Alsb^{(k,l)}$:

$$Alsb^{(k,l)} = LSB\left(A^{(k,l)}\right) = \left\{ LSB\left(A^{(k,l)}(q,s)\right), A^{(k,l)}(q,s) \in A^{(k,l)} \right\}. \quad (2.5)$$

Приклад застосування бітової біфуркації для значення $A^{(k,l)}(q,s) = 33123$ та екстракції значень $Amsb^{(k,l)}(q,s) = 129$ та $Alsb^{(k,l)}(q,s) = 99$ можна побачити на рисунку 2.1.

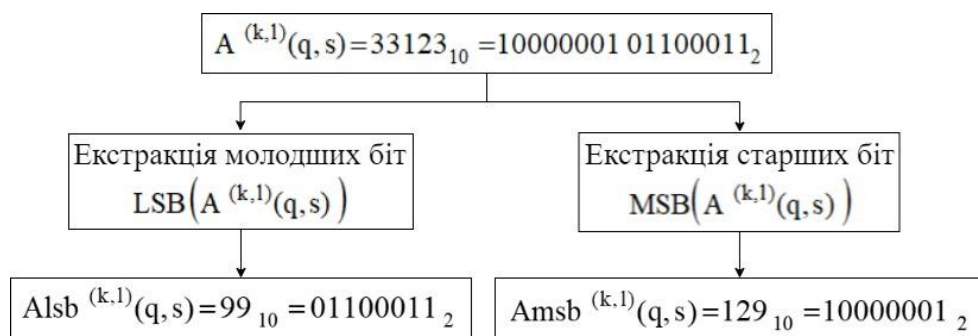


Рисунок 2.1 – Схема екстракції значень $Amsb^{(k,l)}(q,s)$ та $Alsb^{(k,l)}(q,s)$

Якщо застосувати бітову біфуркацію до кожного сегменту $A^{(k,l)}$ з вхідного зображення A , то можна утворити пласт старших бітових значень A_{msb} та пласт молодших бітових значень A_{lsb} . Виразимо це наступними формулами:

$$\begin{aligned} A_{lsb} &= \text{LSB} \left(A^{(k,l)} \right), A^{(k,l)} \in A; \\ A_{msb} &= \text{MSB} \left(A^{(k,l)} \right), A^{(k,l)} \in A. \end{aligned} \quad (2.6)$$

Схема утворення бітових пластів A_{msb} та A_{lsb} зображена на рисунку 2.2.

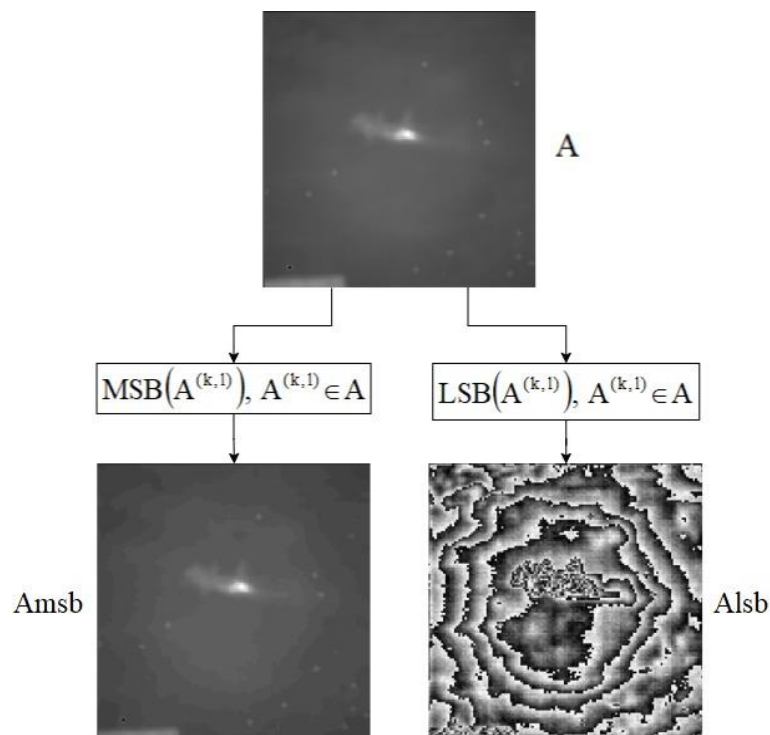


Рисунок 2.2 – Схема утворення бітових пластів A_{msb} та A_{lsb}

Візуальний аналіз A_{msb} та A_{lsb} демонструє, що початкове зображення A та площина старших бітових значень A_{msb} мають значну структурну схожість. Контури об'єктів, глобальний розподіл інтенсивностей та температурні градієнти відображаються у A_{msb} практично без істотної деформації, хоча кількість доступних рівнів інтенсивності зменшується до

256. Це природно, оскільки старші біти формують основну частину числового значення пікселя та, відповідно, несуть основне навантаження щодо семантики сцени та структури теплових полів. Відкинувши молодші біти, зображення втрачає дрібні деталі, переважно непомітні для людського ока.

У той же час площина молодших бітів $Alsb$ за візуальним виглядом часто наближається до шумового поля [109, 140, 141]: на ній відсутні чіткі великомасштабні структури, а домінують дрібні флуктуації, пов'язані з шумом сенсора, низькоамплітудними текстурями та залишковими варіаціями температури. Вони можуть не утворювати цілісних семантичних візуальних структур.

Розуміння особливостей бітових репрезентацій $Amsb$ та $Alsb$ важливе для подальших перетворень. Адже спираючись на це, методи обробки та кодування можуть бути селективно застосовані до відповідного типу бітової площини.

Щоб формалізувати ці спостереження та оцінити потенційні переваги кодування кожного з пластів, було виконано ентропійний аналіз пластів на рівні сегментів. Для кожного сегмента з пласту старших бітів $Amsb^{(k,l)}$ та пласту молодших бітів $Alsb^{(k,l)}$ обчислюється ентропія. З утворених значень побудовані карти ентропії, що відображені на рисунку 2.3.

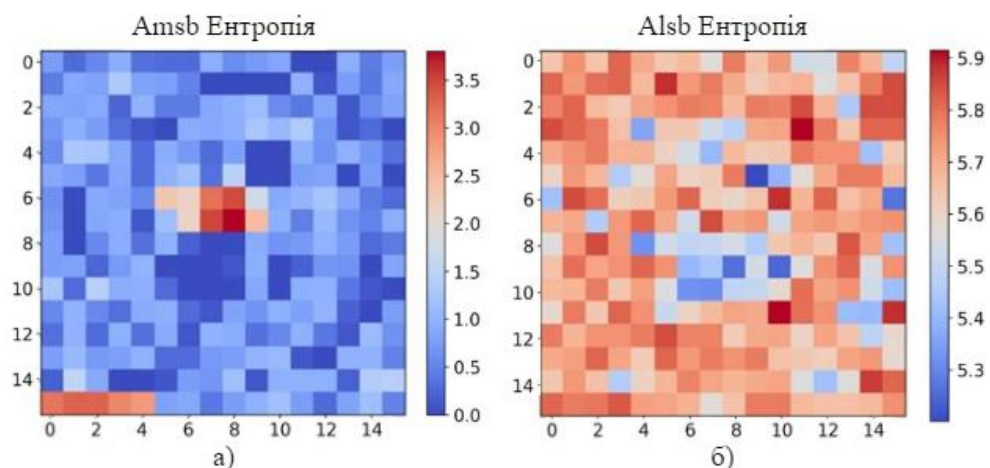


Рисунок 2.3 – Карта ентропії:

а) пласту старших бітів A_{msb} ; б) пласту молодших бітів A_{lsb}

Отримані карти ентропії показують, що для більшості сегментів пласту старших бітів A_{msb} значення ентропії є низькими – особливо у фонових та однорідних областях сцени [134]. Підвищені значення ентропії спостерігаються переважно у сегментах, що містять теплові сигнатури об'єктів, різкі переходи або структурно насичені деталі. Це означає, що більша частина A_{msb} є статистично однорідною й добре придатною до стиснення.

Навпаки, для площини молодших бітів A_{lsb} характерні вищі значення ентропії. Що обмежує потенціал стиснення. Разом із тим, саме ця площина містить інформацію, менш критичну з точки зору візуальної якості та метричної інтерпретації, що створює можливість використання більш агресивних методів кодування з втратами. Особливо у сегментах $A_{lsb}^{(k,l)}$, що позиційно знаходяться за межами важливих теплових сигнатур.

Таким чином, метод біфуркації бітових пластів не лише переводить 16-бітне зображення до двох 8-бітних пластів, але й забезпечує можливість для аналізу даних з точки зору їх семантичного навантаження. Площина старших бітів A_{msb} формує стабільний, слабо ентропійний підпростір, у якому зручно застосовувати підходи, орієнтовані на виявлення структур об'єктів інтересу. Площина молодших бітів A_{lsb} виділяє деталізуючу й шумову складову в окремий канал, для якого можуть бути використані відмінні, потенційно більш агресивні стратегії стиснення, з урахуванням позиційного розташування сегментів. Такий поділ дозволяє будувати подальшу систему кодування як композицію диференційованих перетворень, адаптованих до статистичних і семантичних властивостей кожної з площин, що є ключовою передумовою подальших етапів розроблюваного методу.

2.4 Аналіз та обґрунтування вибору методу трансформації просторових даних у спектральний простір

Запропонований у попередньому підрозділі підхід до біфуркації бітових пластів дозволив здійснити первинну сегрегацію інформаційного потоку, виділивши семантично насичені старші біти (MSB). Однак, з позиції семантико-орієнтованого аналізу, представлення даних виключно у просторовій області є обмеженим, оскільки воно не дозволяє явно ідентифікувати та ранжувати структурні елементи зображення за ступенем їхньої значущості.

Перехід у спектральний простір [8] трансформує масив пікселів у впорядкований набір частотних коефіцієнтів. Таке перетворення забезпечує перерозподіл і концентрацію енергії, що дозволяє виокремити найбільш інформативні складові сигналу.

Під час обробки візуальних зображень основним критерієм виступає збереження перцептивної подібності. Натомість обробка теплових кадрів вимагає збереження висококонтрастних температурних градієнтів. Зазначені структури формують теплові сигнатури цільових об'єктів. Спотворення температурних градієнтів знижує інформативність даних та ускладнює їх подальший аналіз.

Відповідно, спектральна трансформація у цьому контексті виконує такі взаємопов'язані завдання:

- ущільнення енергії для зменшення загального обсягу даних;
- відображення семантичного змісту сигналу, необхідного для його аналітичного використання.

Методи спектрального перетворення, які лежать в основі сучасних алгоритмів кодування зображень, розвиваються навколо ідеї відокремлення суттєвих енергетичних компонент сигналу від його надлишкових або малозначущих складових. Залежно від математичних принципів і типу

використовуваних базисів ці методи можна умовно поділити на дві групи: гармонічні перетворення та вейвлет-перетворення. Кожна з них відображає різні підходи до представлення сигналу у частотному просторі та має власний баланс між енергетичною ефективністю, складністю реалізації і точністю відновлення.

До першої групи належать дискретне косинусне (DCT) та дискретне синусне (DST) перетворення. Їхня сутність полягає у представленні зображення як сукупності гармонік, які описують поведінку сигналу у вигляді синусоїдальних або косинусних хвиль. Такі перетворення забезпечують сильну енергетичну концентрацію в низькочастотній області спектра, оскільки природні зображення зазвичай складається з плавно змінних компонент. Унаслідок цього значна частина енергії сигналу може бути представлена невеликою кількістю коефіцієнтів, що робить ці методи особливо ефективними для задач стиснення [61]. Їхньою важливою перевагою є наявність швидких алгоритмів обчислення та широка апаратна підтримка, що сприяє використанню DCT і DST у таких стандартах, як JPEG [73], HEVC [83] та AVIF [152].

Однак гармонічні перетворення мають обмеження, пов'язані з їхньою глобальною природою. Рівномірний розподіл косинусних або синусних базисів по всьому простору обмежує можливість одночасної локалізації сигналу за частотою та простором. Для інфрачервоних зображень, де важливо відтворювати локальні зміни температурних полів без розмиття контурів або втрати локальної структури, це обмеження стає суттєвим. У таких випадках доцільно застосовувати методи, які поєднують переваги частотного аналізу з можливістю просторової локалізації – саме ці властивості забезпечує друга група методів.

До другої групи належать вейвлет-перетворення (DWT), які ґрунтуються на принципі багаторівневої частотної декомпозиції сигналу. На відміну від гармонічних методів, де базисні функції мають фіксовану частоту, у DWT базиси масштабуються і зсуваються, що дає змогу одночасно

аналізувати структуру сигналу як у частотному, так і у просторовому вимірі. Завдяки цьому вейвлети здатні детально описувати як великомасштабні теплові контури, так і дрібні локальні відхилення температури.

Ефективність вейвлет-перетворення залежить від вибору базисних функцій. Вони визначають просторово-частотну локалізацію енергії сигналу та впливають на параметри декомпозиції і ступінь збереження інформативності. Для інфрачервоних зображень характерне одночасне існування плавних фонових областей та різких контурів теплових об'єктів. Неоптимальний вибір базису призводить до надмірного згладжування температурних контрастів або генерує артефакти на межах об'єктів.

Важливим параметром виступає довжина фільтра, яка регулює рівень згладжування та стійкість алгоритму до шуму. Короткі базиси, зокрема сімейства Хаар (Haar) [75], відзначаються низькою обчислювальною складністю та точно відтворюють різкі переходи. Проте застосування короткого фільтра генерує сходинкові артефакти в умовах агресивного квантування, що обмежує його ефективність у висококонтрастних локальних областях. З іншого боку, збільшена довжина фільтрів підвищує ефективність аналізу плавних сцен.

Іншим визначальним параметром базисної функції є симетричність, яка відповідає за мінімізацію фазових зсувів під час аналізу та синтезу сигналу. Класичні ортогональні базиси сімейства Добеші (Daubechies) забезпечують високу енергетичну компактність і достатню гладкість. Водночас їхня несиметричність спричиняє зсуви контурів. Вирішенням цієї проблеми виступає застосування модифікованих сімейств. Функції Сімлет (Symlet) та Коіфлет (Coiflet) є менш асиметричною версією фільтрів Добеші. Таке наближення до симетрії дозволяє значно точніше зберігати просторові структури на зображенні.

Згідно з математичною теорією вейвлетів, досягнення повної симетрії при збереженні компактного носія є неможливим для ортогональних фільтрів, за винятком базису Хаара. Це обмеження зумовлює

розповсюдження біортогональних базисів (Biorthogonal), наприклад, сімейства Коена–Добеші–Фово (CDF). Такі перетворення дозволяють незалежно формувати фільтри для розкладу (аналізу) та відновлення (синтезу) сигналу. Зазначена особливість забезпечує наявність точної лінійної фази та знижує ймовірність просторових викривлень на межах об'єктів. Порівняно з ортогональними аналогами, біортогональні вейвлети забезпечують точніше відновлення інформативності під час стиснення з втратами, завдяки чому вони інтегровані як базові алгоритми у стандарт JPEG 2000 [33, 84].

Однак, при порівнянні із базисом Хаара, роздільна архітектура фільтрів аналізу та синтезу у біортогональних базисах підвищує вимоги до апаратного забезпечення. Підвищене навантаження на обчислювач обмежує швидкодію вбудованих систем [80] під час потокової обробки даних.

На рисунку 2.4. зображені візуалізації розповсюджених вейвлет базисів. На них чітко проглядається локальна обмеженість вейвлет базисів. Незважаючи на їх різноманітні форми.

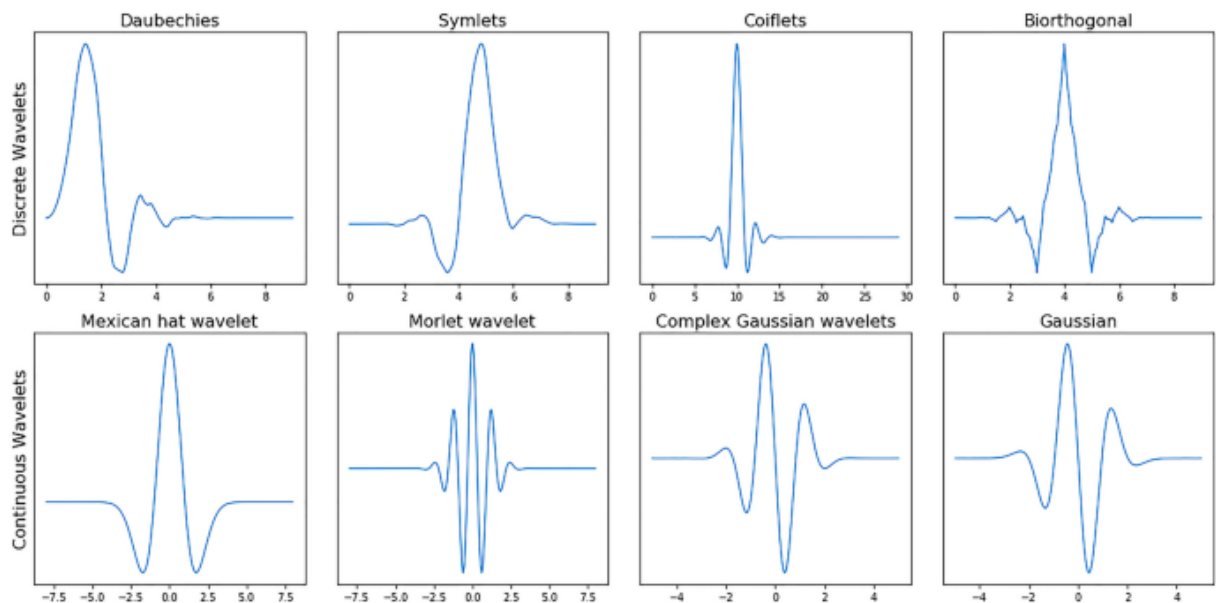


Рисунок 2.4 – Візуалізація розповсюджених базисних функцій для вейвлет перетворення

Гнучкість у виборі базису є однією з переваг вейвлет-перетворення. Вона дозволяє адаптувати алгоритм під конкретний тип сцени або сенсор, вибираючи фільтри, що краще відображають особливості статистики сигналу. Це відкриває перспективу створення адаптивних або гібридних вейвлет-структур, у яких параметри базису підбираються автоматично відповідно до характеристик вхідного зображення. Такий підхід здатен забезпечити потенціал для подальшого розвитку методу кодування на основі вейвлет перетворення.

У таблиці 2.1 наведена порівняльна характеристика методів спектрального перетворення DCT, DST та DWT за наступними критеріями: тип базисних функцій, енергетична компактність, можливість локального аналізу, гнучкість вибору базису, обчислювальна складність, застосування у кодеках.

У розроблюваному методі пропонується застосовувати вейвлет-перетворення на базисі Хаара. Вибір цього базису обґрунтовується двома основними факторами:

- низька обчислювальна складність: реалізація прямого та оберненого перетворень виключно через базові арифметичні операції (додавання, віднімання, ділення на константу) знижує навантаження на обчислювач та забезпечує можливість потокової обробки ІЧ-даних у реальному часі на вбудованих системах;

- локалізація температурних градієнтів: ступінчаста форма базисної функції Хаара просторово збігається з характером різких переходів. Це дозволяє чітко розділяти усереднений фоновий рівень (низькочастотна компонента) та локальні контури (високочастотна компонента).

Таблиця 2.1 – Порівняльна характеристика методів спектрального перетворення

Критерій	DCT (Discreet Cosine Transform)	DST (Discreet Sine Transform)	DWT (Discrete Wavelet Transform)
Тип базисних функцій	Косинусні базиси з фіксованими частотами	Синусні базиси з нульовими граничними умовами	Масштабовані й зсунуті вейвлет базиси
Енергетична компактність	Висока – більшість енергії сигналу зосереджена у низькочастотних коефіцієнтах	Помірна – енергія розподілена більш рівномірно	Висока – локалізована в низькочастотних піддіапазонах з багаторівневою деталізацією
Можливість локального аналізу	Обмежена – однакові частотні базиси для всієї області	Обмежена – однакові частотні базиси для всієї області	Висока – одночасний аналіз у просторі та частоті
Гнучкість вибору базису	Відсутня – базис фіксований	Відсутня – базис фіксований	Висока – вибір серед різних вейвлет-сімейств (Haar, Daubechies, Symlet, і т.д.)
Обчислювальна складність	Низька – наявні швидкі алгоритми (FFT-похідні)	Низька – наявні швидкі алгоритми (FFT-похідні)	Залежить від глибини частотної декомпозиції та базису
Застосування у кодеках	JPEG, HEVC, AVIF, JPEG XL	HEVC (в окремих режимах)	JPEG 2000

2.5 Впровадження методу формування спектрального простору на основі вейвлет-перетворення Хаара

Вейвлет-перетворення застосовується не безпосередньо до вихідних 16-бітних значень, а до сегментів з 8-бітних пластів $A_{msb}^{(k,l)}$ та $A_{lsb}^{(k,l)}$. Позначимо застосування вейвлет-перетворення функціоналом W , тоді його застосування виглядатиме наступним чином:

$$W\left(A_{msb}^{(k,l)}\right); W\left(A_{lsb}^{(k,l)}\right). \quad (2.7)$$

Для зручності опису деталей вейвлет-перетворення позначимо сегмент, до якого воно застосовується як G . Елемент такого сегменту можна отримати за допомогою одного позиційного індекса: $G(x)$, де x – позиційний індекс у послідовності. Індксація може бути задана построчною нумерацією елементів сегменту. Такий підхід дозволяє реалізувати одномірне дискретне вейвлет-перетворення Хаара [117], не втрачаючи при цьому інформації про структуру сегменту.

Застосування вейвлету Хаара дозволяє перейти від однієї послідовності вихідних значень до двох нових послідовностей – низькочастотної та високочастотної, які утворюються попарною обробкою сусідніх елементів. Для кожної пари значень з вихідної послідовності обчислюється:

$$\begin{aligned} \ell_{p+1} &= \frac{G(2p+1) + G(2p+2)}{2}; \\ h_{p+1} &= \frac{G(2p+2) - G(2p+1)}{2}. \end{aligned} \quad (2.8)$$

Де:

- ℓ_{p+1} – низькочастотне значення із пари $G(2p+1)$ та $G(2p+2)$;

- h_{p+1} – високочастотне деталізуюче значення із пари $G(2p+1)$ та $G(2p+2)$;
 - p – індекс за допомогою якого відбувається звернення до значень з G .
- Індексація починається з 0.

$$p = \{0, \dots, P-1\};$$

$$P = \frac{\text{size}(G)}{2}. \quad (2.9)$$

Де:

- $\text{size}(G)$ – кількість значень у сегменті G .
- P – кількість пар сусідніх елементів у мінісегменті.

Після застосування даної операції для всіх значень p утвориться дві групи значень:

$$L_i = \{ \ell_1^{(i)}, \dots, \ell_P^{(i)} \};$$

$$H_i = \{ h_1^{(i)}, \dots, h_P^{(i)} \}. \quad (2.10)$$

Де:

- L_i – група отриманих низькочастотних значень, що описують локально усереднені значення (апроксимацію) сигналу;
- H_i – група високочастотних значень, що характеризують локальні перепади між сусідніми елементами та відповідають деталізуючим складовим;
- i – рівень розрахунку вейвлету.

Подальші перетворення виконуються ітеративно на основі низькочастотної послідовності L_{i-1} попереднього рівня. До того моменту, коли у групі низькочастотних значень L_i буде лише одне значення, тобто: $\text{size}(L_i) = 1$. Отже провести попарний розрахунок буде неможливо.

Оскільки, розрахунок значень L_i та H_i проводиться на основі низькочастотних значень L_{i-1} попереднього рівня, то кількість пар сусідніх елементів P буде дорівнювати:

$$P = \frac{\text{size}(L_{i-1})}{2}. \quad (2.11)$$

Де $\text{size}(L_{i-1})$ – кількість значень з попереднього низькочастотного рівня L_{i-1} .

Для групи сегментних значень G кількість рівнів розрахунку I дорівнює:

$$I = \log_2(G_{\text{size}}). \quad (2.12)$$

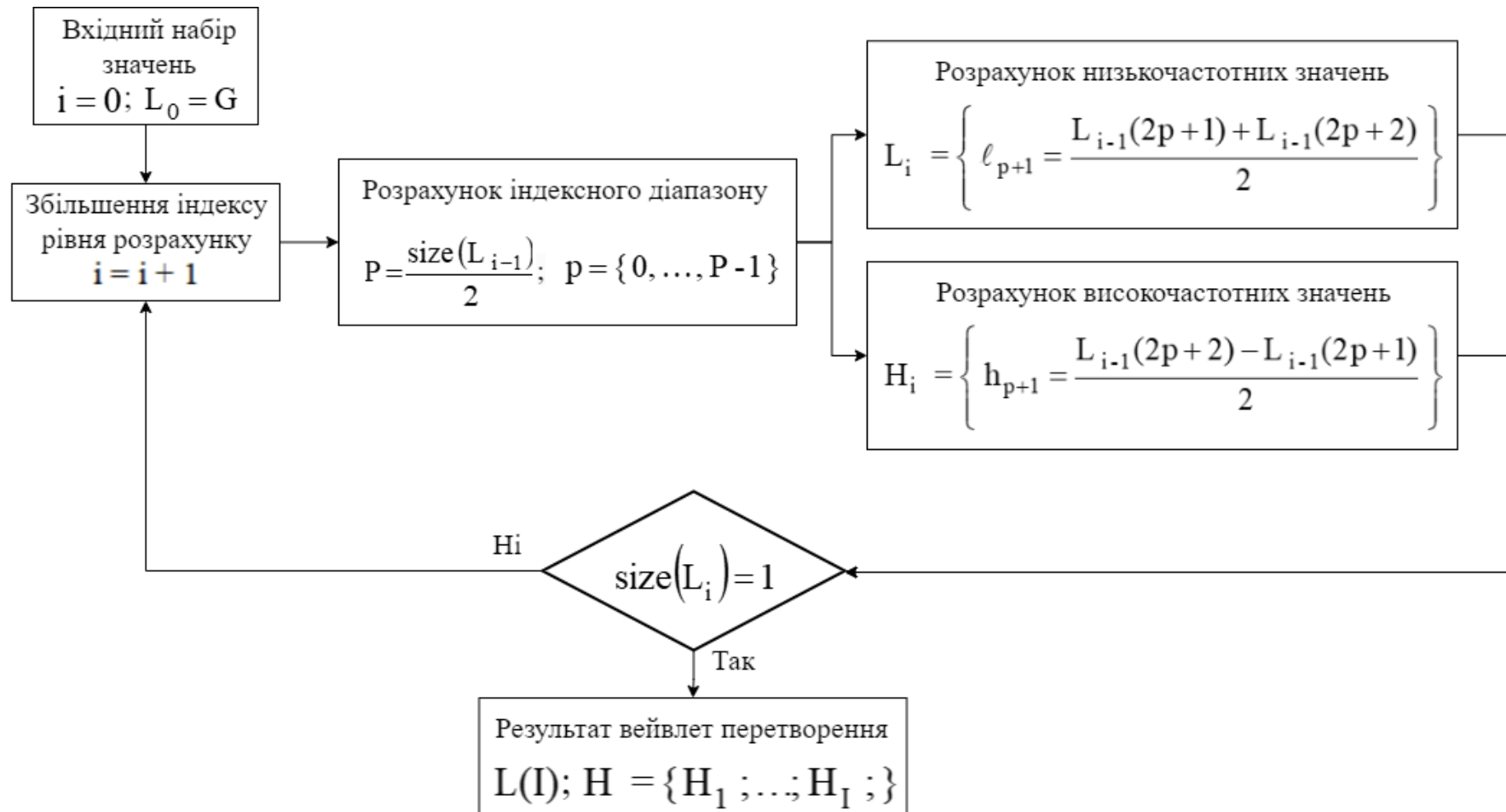
Для початкової групи низькочастотних значень L_0 , де $i=0$, правдивий наступний вираз: $L_0 = G$. Тобто початкова група низькочастотних значень L_0 представляє сегментні значення G .

У результаті ітеративного розрахунку вейвлету Хаара формується:

- один кінцевий низькочастотний коефіцієнт L_I , який є узагальненою апроксимацією сегменту G . У подальшому будемо звертатись до нього як $L(I)$;

- набір високочастотних коефіцієнтів $H = \{H_1; \dots; H_I\}$, де кожне значення містить високочастотні компоненти відповідного масштабу, отримані на ітераціях $i = [1..I]$. Кінцеве високочастотне значення позначимо як $H(I)$.

На рисунку 2.5 зображено схему розрахунку вейвлету Хаара.

Рисунок 2.5 – Схема застосування вейвлету Хаара до групи сегментних значень G

Застосування вейвлету Хаара до сегментної групи значень G утворює сукупність значень $(L(I), H)$. Отже:

$$\begin{aligned} (L(I), H) &= W(G); \\ (L^{(k,l)}(I), H^{(k,l)}) &= W(A^{(k,l)}). \end{aligned} \quad (2.13)$$

Властивістю вейвлет перетворення Хаара є його оборотність. Маючи L_{i+1} та H_{i+1} , можна відновити попередній рівень:

$$\begin{aligned} L_i(2p) &= L_{i+1}(p) + H_{i+1}(p); \\ L_i(2p + 1) &= L_{i+1}(p) - H_{i+1}(p). \end{aligned} \quad (2.14)$$

Послідовний підйом від рівня I до рівня 0 дозволяє точно відновити вихідну послідовність L_0 , а отже й сегмент G . Таким чином, при збереженні всіх високочастотних коефіцієнтів та кінцевого низькочастотного значення перетворення є строго безвтратним. Усі похибки можуть виникати лише на подальших етапах квантування.

Таким чином, формування спектрального простору на основі вейвлет-перетворення Хаара забезпечує перехід від просторового представлення до енергетично та структурно впорядкованого набору коефіцієнтів, який є сприятливим для подальших етапів аналізу та кодування. Це створює узгоджену основу для наступних процедур аналізу та зменшення бітової інтенсивності ІЧ-кадрів [116].

2.6 Розробка методу визначення метрики інформативності на основі структурно важливих ознак сегментів

Для диференціювання сегментів кадру, необхідно ввести формальну метрику, яка кількісно описує рівень інформативності кожного сегмента та дозволяє керувати подальшими режимами обробки.

Кожній сегментній групі значень G ставиться у відповідність скалярна величина G_M , яка інтерпритується як метрика інформативності:

$$G_M = \text{Metr}(G). \quad (2.15)$$

Де $\text{Metr}(G)$ – функціонал визначення метрики інформативності. Спираючись на значення G_M можна диференціювати сегменти за рівнем інформативності. Для цього використаємо порогове значення T . Значення T може задаватися відповідно до вимог до якості та компресії. Якщо $G_M \geq T$, то сегмент G вважається інформативним, інакше малоінформативним. У залежності від рівня інформативності сегменту, при подальшій обробці можуть використовуватись різні методи кодування чи різні конфігурації одного методу. Назвемо це семантико-орієнтованим кодуванням SAP та позначимо обробку малоінформативного сегменту як $SAP_{\text{low}}(G)$, а високоінформативного $SAP_{\text{high}}(G)$. Описаний метод семантико-орієнтованої обробки зображено на схемі рисунку 2.6.

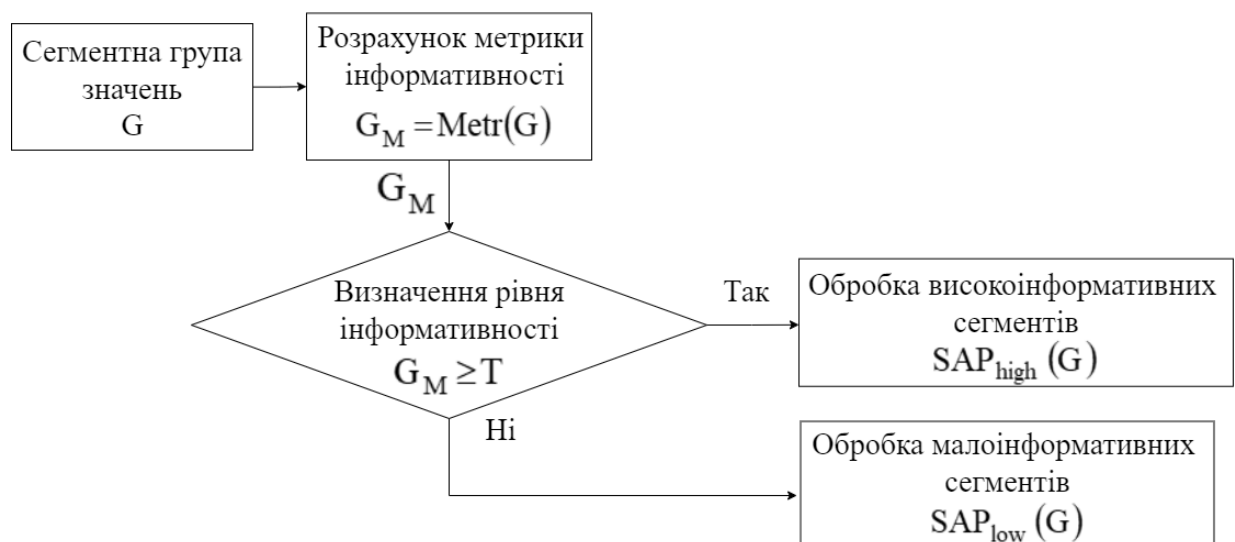


Рисунок 2.6 – Схема методу семантико-орієнтованої обробки сегментної групи значень G

Специфіка ІЧ-зображень полягає в тому, що семантично значущі об'єкти, дефекти, аномальні теплові зони або маскувальні структури проявляються через локальні зміни температурних рівнів, тобто у вигляді контурів і різких градієнтів. У спектральному представленні, сформованому на основі попередньо описаного вейвлет-перетворення Хаара, такі структури породжують підвищені високочастотні компоненти. Тому для побудови метрики інформативності доцільно використовувати саме високочастотні коефіцієнти як основне джерело структурної інформації.

Для розрахунку метрики інформативності G_M визначимо суму високочастотних значень H . Оскільки коефіцієнти можуть мати як додатні, так і від'ємні значення, їх пряме підсумовування потенційно призводило б до взаємної компенсації та заниження оцінки. Для уникнення цього використовується сума модулів високочастотних значень, яка відображає енергію високих частот у межах сегмента:

$$HSum = \sum |H_i|, \quad H_i \in H. \quad (2.16)$$

Де:

- $|H|$ - абсолютне високочастотне значення;
- $HSum$ - сума абсолютних високочастотних значень для сегменту G .

Величина $HSum$ інтерпретується як локальна міра структурної насиченості: що більше різких переходів, контурів і текстурних деталей містить G , то вищим є значення цієї енергії. Вірне і зворотне, значення $HSum$ залишається низьким для сегментів з майже однорідними значеннями.

Для забезпечення коректної порівнюваності між різними сегментами у кадрі необхідно виконати нормалізацію отриманих значень $HSum$. Нормалізація дозволяє привести енергію високих частот до шкали $[0; 1]$, незалежної від абсолютних рівнів сигналу та масштабу сцени.

Нехай серед усіх сегментів кадру значення сукупної енергії $HSum$ мають мінімум $HSum_{\min}$ та максимум $HSum_{\max}$. Тоді нормалізована оцінка для сегмента визначається як:

$$\text{Norm}(HSum) = \begin{cases} \frac{HSum - HSum_{\min}}{HSum_{\max} - HSum_{\min}}, & \text{якщо } HSum_{\max} > HSum_{\min}, \\ 0, & \text{якщо } HSum_{\max} = HSum_{\min}. \end{cases} \quad (2.17)$$

Враховуючи, що метрика інформативності G_M для сегменту G розраховується на основі високих частот H , можна виразити це наступною рівністю:

$$G_M = \text{Metr}(G) = \text{Metr}(H) = \text{Norm}(HSum), \text{ де } G \rightarrow H. \quad (2.18)$$

Тобто G_M є нормалізованою сумарною енергією високих частот сегмента. Така форма забезпечує зв'язок між кількістю високочастотних компонент та інформативністю: чим більш структурно насичений сегмент G , тим ближче G_M до 1. Для однорідних фонів значення метрики G_M наближається до 0.

На рисунку 2.7 зображена схема розрахунку метрики інформативності G_M для сегменту G .



Рисунок 2.7 – Схема методу розрахунку метрики інформативності G_M

Запропонована метрика має низку важливих властивостей. По-перше, вона безпосередньо пов'язана з температурною структурою сцени, оскільки будується на високочастотних компонентах, що генеруються саме локальними градієнтами температури. По-друге, з обчислювальної точки зору метрика є економною, оскільки базується на коефіцієнтах вейвлет-перетворення, які будуть використані в межах основних перетворень при подальшому кодуванні. Її обчислення зводиться до операцій взяття модуля, підсумовування та лінійної нормалізації, що добре узгоджується з вимогами до швидкодії в умовах можливих обмежень апаратного забезпечення інформаційних систем реального часу.

Таким чином, розроблений метод визначення метрики інформативності сегментів забезпечує формальний перехід від якісного поняття важливості ділянки ІЧ-кадру до кількісної шкали, на основі якої може будуватися адаптивна, семантико-орієнтована стратегія кодування.

2.7 Розробка методу формування двошарової інформативної карти інфрачервоного кадру

У попередніх підрозділах було показано, що розділення сегментів 16-бітного ІЧ-зображення на пласти старших $A_{msb}^{(k,1)}$ та молодших $A_{lsb}^{(k,1)}$

бітів дає змогу виокремити дві принципово різні за структурою складові сигналу. Пласт старших бітів відображає глобальну просторову організацію сцени, основні теплові градієнти та контури об'єктів, тоді як пласт молодших бітів містить дрібні деталізуючі компоненти та значну частку сенсорного шуму. У контексті семантико-орієнтованого кодування це означає, що одні й ті самі просторові ділянки сцени можуть проявляти свою інформативність у різному ступені в кожному з бітових пластів. Відповідно постає задача побудови узгодженого механізму класифікації сегментів, який, з одного боку, враховував би специфіку кожного пласту, а з іншого – формував би єдине рішення щодо важливості сегмента для подальшого кодування.

Для кожного сегмента у старшому пласті $A_{msb}^{(k,1)}$ та відповідного за положенням сегмента у молодшому пласті $A_{lsb}^{(k,1)}$ обчислюються значення метрики інформативності $A_{msb}_M^{(k,1)}$ та $A_{lsb}_M^{(k,1)}$:

$$A_{msb}_M^{(k,1)} = \text{Metr}(A_{msb}^{(k,1)}); \quad A_{lsb}_M^{(k,1)} = \text{Metr}(A_{lsb}^{(k,1)}). \quad (2.19)$$

Обчислення здійснюється за єдиною методикою, наведеною у підрозділі 2.6, яка ґрунтується на спектральному аналізі та агрегуванні енергії високочастотних компонент. Таким чином, інформативність кожної сегментної позиція $(k,1)$ у біфуркованому кадрі описується парою чисел $(A_{msb}_M^{(k,1)}, A_{lsb}_M^{(k,1)})$, що відображають структурну насиченість відповідної ділянки сцени в обох бітових пластах A_{msb} та A_{lsb} .

Подальша класифікація інформативності сегментів починається з аналізу площини старших бітів $A_{msb}^{(k,1)}$, оскільки саме вона несе основний внесок у відтворення теплової структури сцени. На основі вибірки значень $\{A_{msb}_M^{(k,1)}\}$ визначається порогове значення T_{MSB} , яке розділяє сегменти з виразними тепловими структурами та сегменти, що відповідають однорідному фону. Якщо для певної позиції $(k,1)$ виконується нерівність

$A_{MSB}^{(k,l)} \geq T_{MSB}$, то відповідна позиція вважається інформативною. У цьому випадку приймається рішення, що дана ділянка сцени є семантично важливою, і формальне значення єдиної карти інформативності встановлюється як $M^{(k,l)} = 1$.

Важливо, що статус інформативності у термінах семантики сцени поширюється на обидві бітові площини, тобто як $A_{MSB}^{(k,l)}$, так і $A_{LSB}^{(k,l)}$ у цій позиції (k,l) трактуються як такі, що потребують підвищеного рівня збереження інформації у подальших етапах кодування.

Однак існують ситуації, коли у старшому бітовому пласті сегмент має низький рівень структурної насиченості, тобто $A_{MSB}^{(k,l)} < T_{MSB}$. Це може відповідати як справді однорідному фону, так і випадкам, коли основна інформативність сегмента зосереджена у слабких, малоконтрастних сигнатурах, представлених переважно в молодших бітах. Наприклад, віддалені об'єкти або слабке теплове випромінення. У такій ситуації аналіз переноситься до пласту молодших бітів A_{LSB} , де для позиційно того ж самого сегмента $A_{LSB}^{(k,l)}$ використовується метрика $A_{LSB}^{(k,l)}$. На основі її розподілу визначається поріг T_{LSB} , який враховує підвищену зашумленість A_{LSB} пласту.

Якщо для сегмента, який не був визнаний інформативним за MSB-критерієм, виконується нерівність $A_{LSB}^{(k,l)} \geq T_{LSB}$, це означає, що у площині молодших бітів спостерігається наявність виражених високочастотних сигналів, які можуть відповідати корисним деталям сцени. У цьому випадку сегмент вважається інформативним в єдиній карті інформативності встановлюється значення $M^{(k,l)} = 1$.

При цьому, як і у випадку з порогом T_{MSB} , наявність високої інформативності у пласті A_{LSB} позиційно відповідає сегменту у пласті A_{MSB} . Тобто, якщо хоча б одна з бітових площин вказує на наявність структурно

значущих компонент, ділянка сцени в цілому вважається інформативною, і обидві відповідні бітові репрезентації мають оброблятися у режимі із підвищеним збереженням точності.

Якщо для сегментної позиції виконуються одночасно дві умови:

$$\text{Amsb}_M^{(k,1)} < T_{\text{MSB}}; \text{Alsb}_M^{(k,1)} < T_{\text{LSB}}; \quad (2.20)$$

то це свідчить про відсутність помітних теплових контрастів як на рівні основної структури, так і на рівні дрібних деталей. Такий сегмент інтерпретується як низькоінформативний (фоновий). Для нього у єдиній карті інформативності встановлюється значення $M^{(k,1)} = 0$.

Отже, карта $M^{(k,1)}$ є бінарною маскою, яка ділить усі сегменти кадру на два класи: інформативні (значення 1) та низькоінформативні (значення 0) з урахуванням обох бітових пластів. Схема побудови карти інформативності зображена на рисунку 2.8.

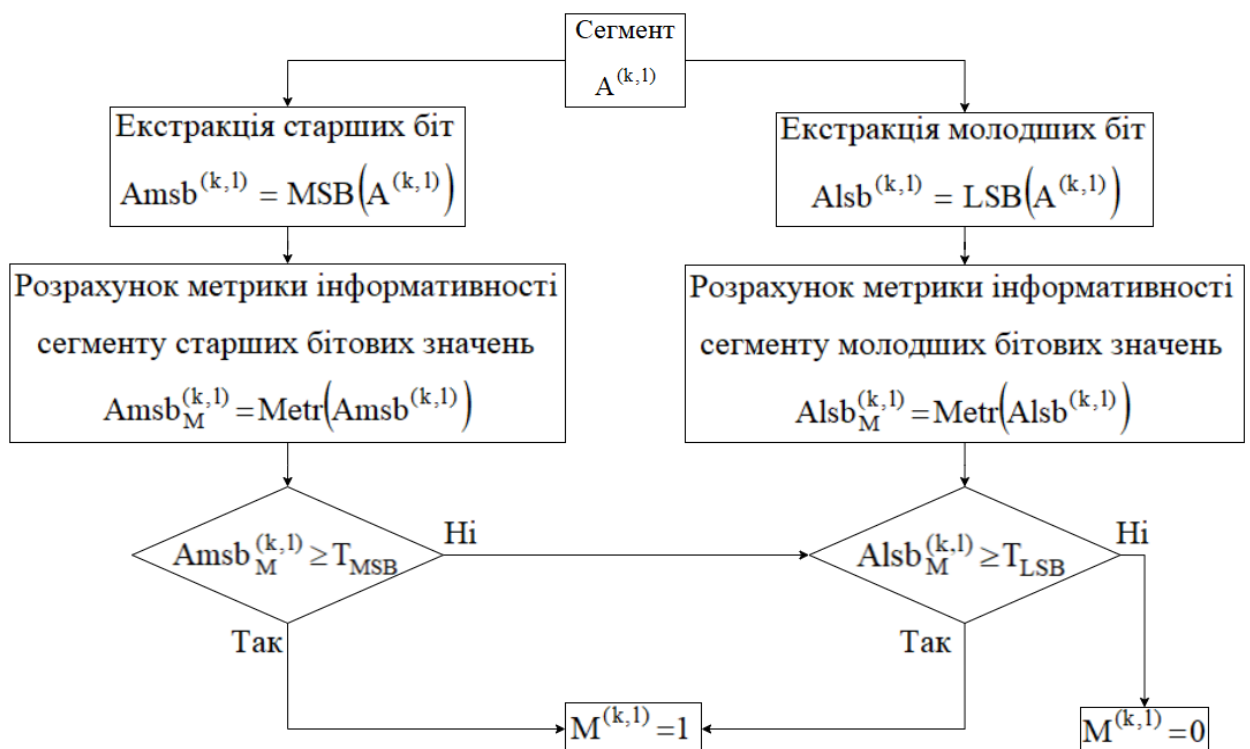


Рисунок 2.8 – Схема побудови карти інформативності

Запропонований механізм фактично реалізує логіку диз'юнкції щодо виявлення семантики: сегмент $A^{(k,1)}$ вважається інформативним, якщо він проявляє значущу структурну активність хоча б в одній із площин A_{msb} чи A_{lsb} . При цьому для подальшого кодування використовується саме єдина карта $M^{(k,1)}$, яка виступає вказівником наявності семантичної важливості та вибору параметрів кодування. У сегментах з $M^{(k,1)}=1$ застосовуються більш помірні режими обробки, тоді як для сегментів з $M^{(k,1)}=0$ допускається агресивніше стиснення, спрямоване на зменшення бітового об'єму.

На рисунку 2.9 візуально зображені теплові карти метрик інформативності $A_{msb}_M^{(k,1)}$ та $A_{lsb}_M^{(k,1)}$, утворені карти інформативних сегментів на основі порогових значень T_{MSB} та T_{LSB} , та результуюча карта інформативних сегментів $M^{(k,1)}$.

Важливо підкреслити, що, попри використання єдиної карти інформативності, параметри кодування для площин старших і молодших бітів можуть відрізнятися, відображаючи їх різну чутливість до спотворень. Проте сам статус сегмента як інформативного чи низькоінформативного визначається один раз і застосовується як єдине джерело правди при кодуванні обох бітових репрезентацій. Такий підхід забезпечує узгодженість трактування семантичної важливості сегмента на всіх рівнях обробки і створює основу для подальших етапів семантико-орієнтованого кодування [98].

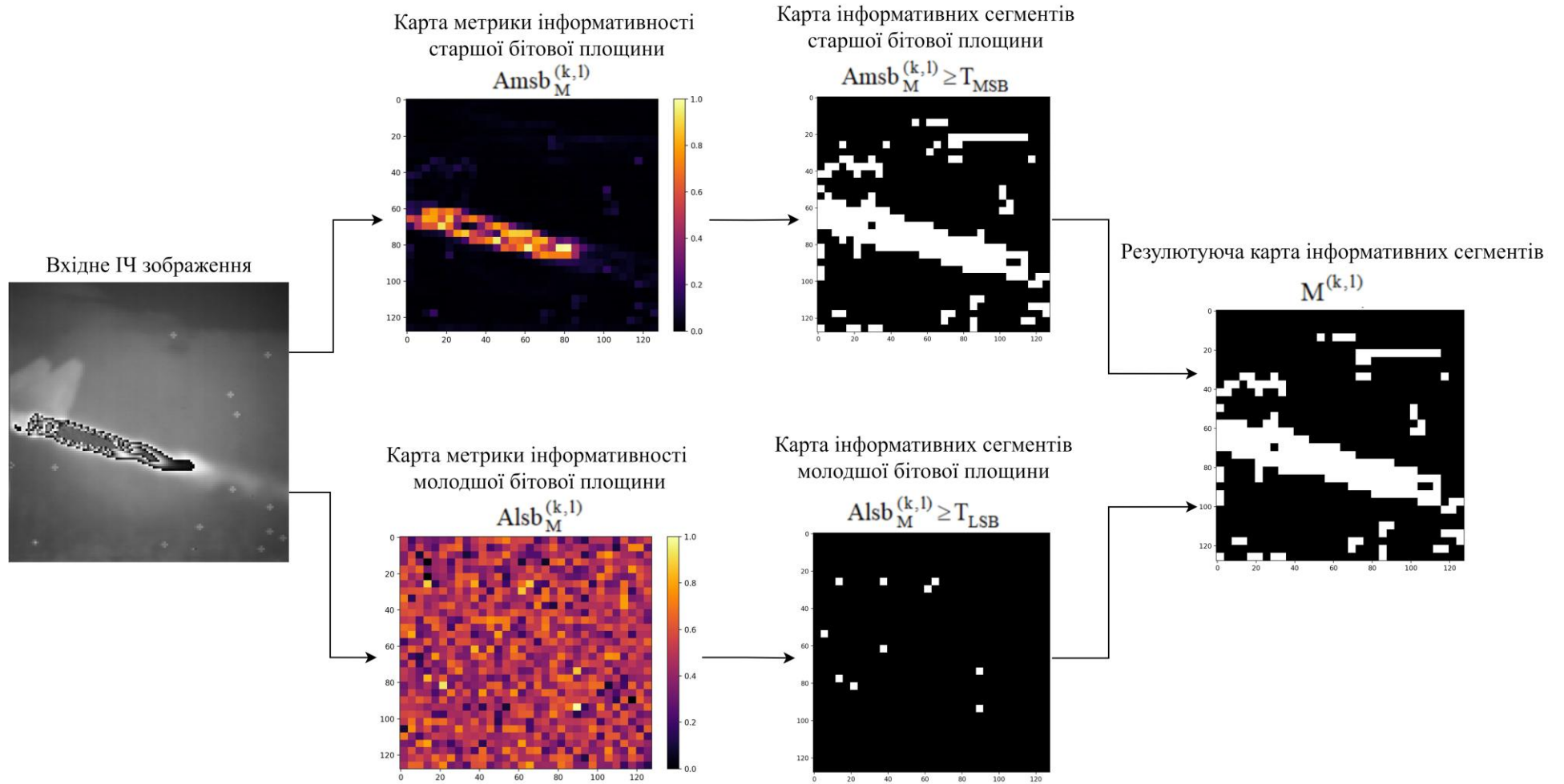


Рисунок 2.9 – Візуалізація основних етапів утворення карти інформативних сегментів $M^{(k,l)}$

2.8 Розробка методу виявлення об'єктно-інформативних сегментів на основі застосування моделей штучного інтелекту

Методи оцінки інформативності, що базується на спектральних характеристиках сигналу, незважаючи на свої переваги, має обмеження: він не дозволяє розрізнити природу об'єктів. Внаслідок цього ділянки з неоднорідними текстурними поверхнями або шумами інтерпретуються з таким рівнем інформативності, що і ділянки які відносяться до реальних об'єктів інтересу. Це призводить до нераціонального розподілу бітового ресурсу. Для усунення цього недоліку пропонується метод виявлення об'єктно-інформативних сегментів, що базується на розпізнаванні образів методами штучного інтелекту (ШІ) [112].

В основі розробленого підходу лежить поняття об'єктної інформативності, яка визначається належністю сегмента зображення до заданого класу цільових об'єктів (наприклад техніка, люди). Таким чином використання моделей ШІ дозволяє покращити побудову інформативної карти.

Важливим є вибір архітектури нейронної мережі, що забезпечувала б баланс між точністю виявлення та обчислювальною складністю в умовах потенційно обмежених апаратних ресурсів. Аналіз існуючих рішень дозволяє виділити два основні підходи: двостадійні та одностадійні детектори.

Двостадійні детектори (Two-Stage Detectors), такі як сімейство R-CNN (Faster R-CNN, Mask R-CNN), характеризуються високою точністю завдяки розділенню процесу на етап генерації регіонів-кандидатів та етап їх класифікації. Однак, значна обчислювальна складність та низька швидкість інференсу роблять їх малопридатними для обробки відеопотоку в реальному часі. Тому для подальшого аналізу доцільніше розглядати клас одностадійних детекторів.

У цьому класі було проведено порівняння архітектури YOLOv5 із методом SSD (Single Shot MultiBox Detector) [35]. Аналіз виявив

архітектурні обмеження SSD при роботі з ІЧ-зображеннями. Незалежна детекція на різних рівнях піраміди ознак створює умови, коли нижні шари мережі мають недостатню насиченість ознаками для локалізації дрібних об'єктів. Таке обмеження за умов розповсюдженого використання ІЧ-камер в аерофотозйомці є суттєвим, оскільки в таких кадрах об'єкти зазвичай малорозмірні.

Натомість архітектура YOLOv5 вирішує ці проблеми завдяки використанню структури PANet (Path Aggregation Network), яка забезпечує двонаправлений потік інформації: ознаки глибоких шарів передаються на нижні рівні для покращення локалізації. Базуючись на цьому, для реалізації методу було обрано архітектуру YOLOv5 [160], яка має високий рівень апробації, в тому числі і у задачах виявлення ІЧ-об'єктів.

З метою обмеження обчислювальної складності у запропонованому методі на вхід нейромережі подається старший бітовий пласт A_{msb} , який містить основну структурну інформацію про сцену. У результаті роботи методу детектування D одержується набір виявлених об'єктних областей R , кількості n . Кожна окрема область $r_i \in R$ містить індекс класу $Class_i$ та коефіцієнт впевненості $Conf_i \in [0, 1]$. Математично це можна виразити наступним чином:

$$\begin{aligned} R &= D(A_{msb}); \\ R &= \{r_1, r_2, \dots, r_n\}; \\ r_i &= (Class_i, Conf_i). \end{aligned} \quad (2.21)$$

Певні області r_i можуть мати низький коефіцієнт впевненості. Що може означати хибне виявлення об'єктів інтересу. Тому вплив таких областей на процес кодування необхідно обмежувати. Для цього введено порогове значення T_{conf} . Якщо $Conf_i \geq T_{conf}$, то область r_i береться до уваги для виявлення об'єктно-інформативних сегментів, інакше – пропускається.

Якщо значення $M^{(k,l)} = 1$ з двопластової інформативної карти позиційно пересікається із областю r_i , то сегментна позиція (k, l) вважається об'єктно-інформативною. Тоді до інформативної карти вноситься значення $M^{(k,l)} = 2$. Інакше, значення $M^{(k,l)}$ лишається без змін. Ця логіка представлена наступними виразом:

$$M^{(k,l)} = \begin{cases} 2, & \text{якщо } \text{Conf}_i \geq T_{\text{conf}}, M^{(k,l)} = 1, M^{(k,l)} \in r_i \\ M^{(k,l)} & \end{cases} \quad (2.22)$$

Таким чином, завдяки виявленню об'єктів із використанням ШІ, сегменти, що раніше були визначені інформативними та позиційно входять до об'єктних областей, вважаються об'єктно-інформативними.

На рисунку 2.10 зображено приклад формування об'єктно-інформативної карти. Де у вхідного зображення (Рис. 2.10 а), було виявлено об'єкти автомобілів та людини (Рис. 2.10 в) та побудовано двопластову інформативну карту (Рис. 2.10 б). На їх основі було сформовано об'єктно-інформативну карту (Рис. 2.10 в), де червоним кольором позначені об'єктно-інформативні сегменти, білим – інформативні, а чорним – малоінформативні. Завдяки виявленню інформативних сегментів в рамках об'єктних областей досягається реалізація механізму семантичної сегментації.

У наведеному прикладі кількість об'єктно-інформативних сегментів складає 19.2% від загальної кількості інформативних сегментів. Досягнуте підвищення точності дозволяє ефективніше перерозподіляти бітовий об'єм для усунення втрат інформативності об'єктів інтересу.

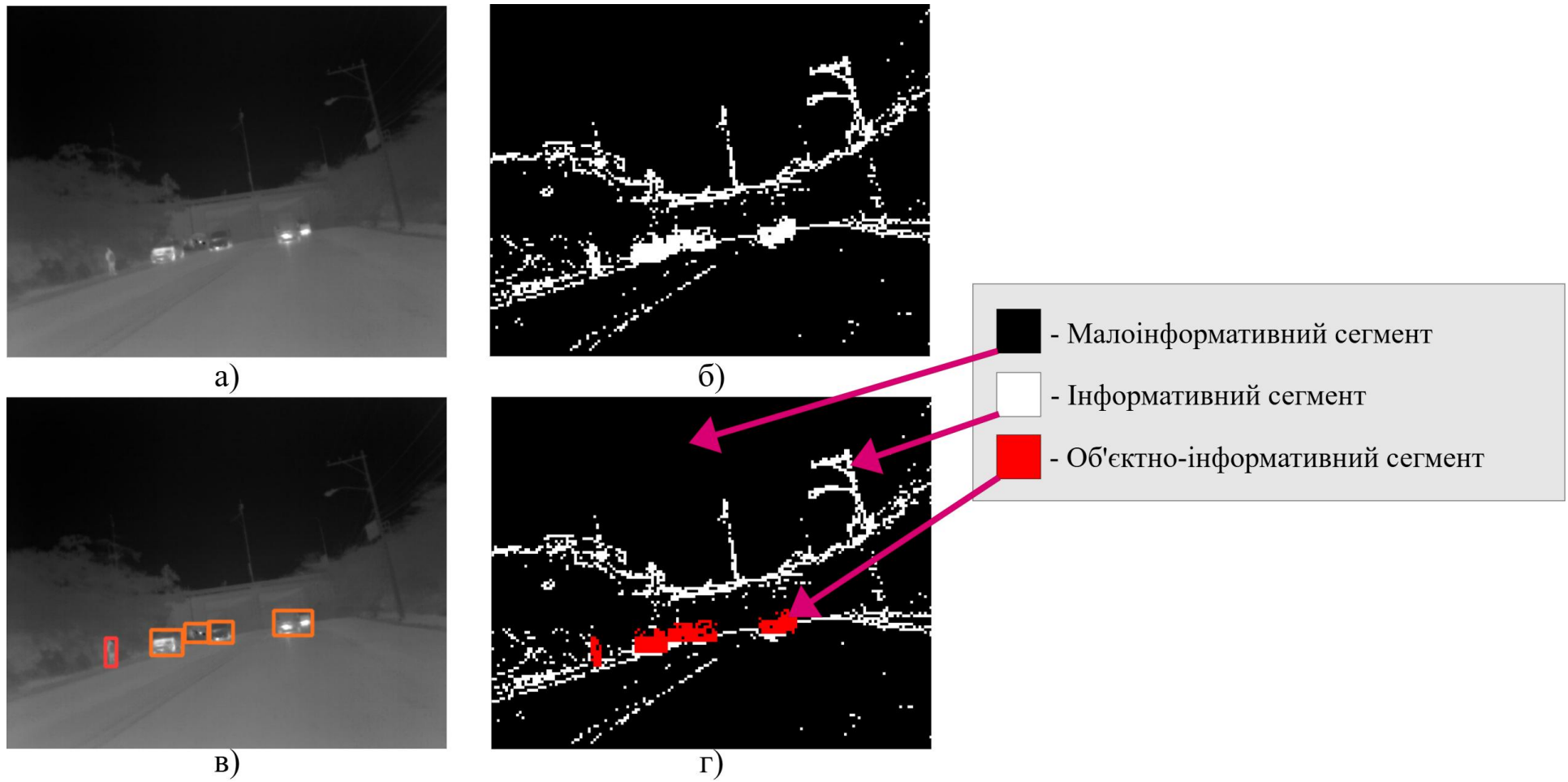


Рисунок 2.10 – Приклад побудови об'єктно-інформативної карти а) вхідне ІЧ-зображення; б) інформативна карта без об'єктних сегментів; в) виявлені об'єктні області; г) інформативна карта з об'єктно-інформативними сегментами

Важливою архітектурною особливістю запропонованого методу є можливість адаптивного керування обчислювальними ресурсами, де застосування блоку ШІ є опціональним рівнем покращення якості, а не критичною залежністю. Це забезпечує масштабованість методу для роботи в умовах суттєвих апаратних обмежень. На практиці це дозволяє системі динамічно відключати нейромережевий модуль у сценаріях дефіциту ресурсів: наприклад, при переході пристрою в режим підвищеної економії енергії, при перегріві процесора або при необхідності звільнити обчислювальні потужності для пріоритетних задач (шифрування каналу зв'язку, навігація). У таких випадках формування карти інформативності автоматично переводиться на спрощений алгоритм на основі метрики інформативності, що підвищує стабільність процесу кодування.

Підсумовуючи, аналіз інформативності сегментів на основі спектрального підходу обмежує можливість розпізнавання семантичної природи структур. Для усунення цього обмеження розроблено метод виявлення об'єктно-інформативних сегментів із застосуванням нейромережевої моделі, який обробляє старший бітовий пласт ІЧ-зображення та ідентифікує об'єкти інтересу. Така ідентифікація дозволяє виявити об'єктно-інформативні сегменти серед структурно значущих інформативних сегментів. Це дозволяє забезпечити збереження теплових сигнатур під час подальшої квантизації та стиснення. Водночас запропонований підхід передбачає опціональність використання інтелектуального модуля. Це дозволяє динамічно адаптувати обчислювальне навантаження бортових систем в умовах апаратних обмежень без порушення стабільності базового процесу кодування [11, 30, 105].

Висновки за другим розділом

1. Обґрунтовано підхід до виявлення термальних контрастів шляхом аналізу високочастотних складових інфрачервоного сигналу. Наявність локальних температурних градієнтів збільшує амплітуду високих частот у спектральному просторі та підвищує сумарну енергію високочастотної компоненти. Такий перерозподіл енергії дозволяє кількісно оцінити структурну насиченість кадру і розрахувати метрику інформативності сегментів.

2. Обґрунтовано вибір вейвлет-перетворення Хаара для трансформації просторових даних у спектральний домен. Ступінчаста форма базисної функції Хаара узгоджується із профілем різких температурних сигнатур, що відповідають об'єктам інтересу та текстурам середовища.

3. **Наукова новизна.** Удосконалено метод побудови двопластової інформативної карти на основі порогових вирішальних правил. Відмінності методу полягають у визначенні адаптивних асоціативних залежностей в порогово-метричному просторі на основі агрегації та нормування інформації високочастотної області перетворення Хаара для сегментованих бітових пластів ІЧ-зображень. Це дозволяє створити умови для збільшення рівня стиснення при встановлених вимогах щодо термальної цілісності ІЧ-зображень.

4. **Наукова новизна.** Отримано подальший розвиток методу виявлення областей об'єктів інтересу на основі застосування моделі штучного інтелекту. Відмінності методу полягають у виявленні об'єктно-інформативних сегментів за рахунок кон'юнкції результату нейронної мережі із сформованою інформативною картою. Це дозволяє зберегти термальну цілісність об'єктів інтересу за умов обмеженого бітрейту.

5. Обґрунтована необхідність можливості опціонального застосування моделі штучного інтелекту для виявлення об'єктно-інформативних сегментів. Робота нейромережових архітектур створює інтенсивне обчислювальне

навантаження. На платформах із обмеженими апаратними ресурсами цей фактор може негативно впливати на швидкість передачі даних та роботу платформи. Можливість відключення інтелектуального аналізу з переходом на базову спектральну метрику підтримує стабільне функціонування системи кодування за умов апаратного дефіциту.

Результати досліджень щодо матеріалів, викладених в другому розділі, опубліковані в таких наукових працях як: [11, 15, 21, 25].

РОЗДІЛ 3. ОБҐРУНТУВАННЯ ТА РОЗРОБКА МЕТОДУ КОДУВАННЯ ІНФРАЧЕРВОНИХ ЗОБРАЖЕНЬ У ІНФОРМАЦІЙНИХ СИСТЕМАХ

У розділі розроблено метод кодування інфрачервоних зображень у інформаційних системах.

Розроблено метод різницевого кодування, що дозволяє зменшити динамічний діапазон мінісегментів.

Розроблено метод перетворення сегментів у спектрально-різницевої простір. Метод оснований на застосуванні вейвлет перетворенні Хаара до різницевої значень мінісегментів.

Розроблено метод корекції низькочастотних коефіцієнтів, що дозволяє зменшити низькочастотний динамічний діапазон.

Розроблено метод адаптивної квантизації високочастотних коефіцієнтів на основі рівня інформативності та бітового пласту сегментів.

Розроблено метод спектрально-групового кодування. Метод дозволяє зменшити бітову інтенсивність сегментів за рахунок представлення високочастотних та низькочастотних спектральних груп у вигляді кодових значень.

3.1 Розробка методу ієрархічної сегментації інфрачервоного кадру з урахуванням його структурно-статистичних характеристик

Для кращого представлення внутрішньосегментної неоднорідності пропонується виконати квадродекомпозицію сегментів [17]. Внаслідок чого, утвориться ієрархічна сегмента структура.

У середині кожного сегмента $A^{(k,l)}$ виділяються чотири мінісегменти розміром 4×4 , позначені $A_{r,t}^{(k,l)}$, де (r, t) – координати мінісегменту у середині сегменту $A^{(k,l)}$. У свою чергу значення з мінісегменту $A_{r,t}^{(k,l)}$ адресуються як $A_{r,t}^{(k,l)}(q,s)$, де (q, s) – координати значення у мінісегменті $A_{r,t}^{(k,l)}$. Тобто: $q, s \in \{1..4\}$.

Таке формальне визначення структури кадру дозволяє встановити відповідність між координатами пікселів та координатами сегментів і мінісегментів, забезпечуючи основу для подальших операцій [19, 100, 118].

Описану ієрархію сегментів можна зобразити наступним схемою на рисунку 3.1. Запровадження внутрішньої структури у вигляді чотирьох мінісегментів $A_{r,t}^{(k,l)}$ із розміром 4×4 забезпечує додатковий рівень деталізації. Мінісегменти дають змогу точніше визначати локальні структури та гнучкіше адаптувати подальші перетворення до реального розподілу даних [29].

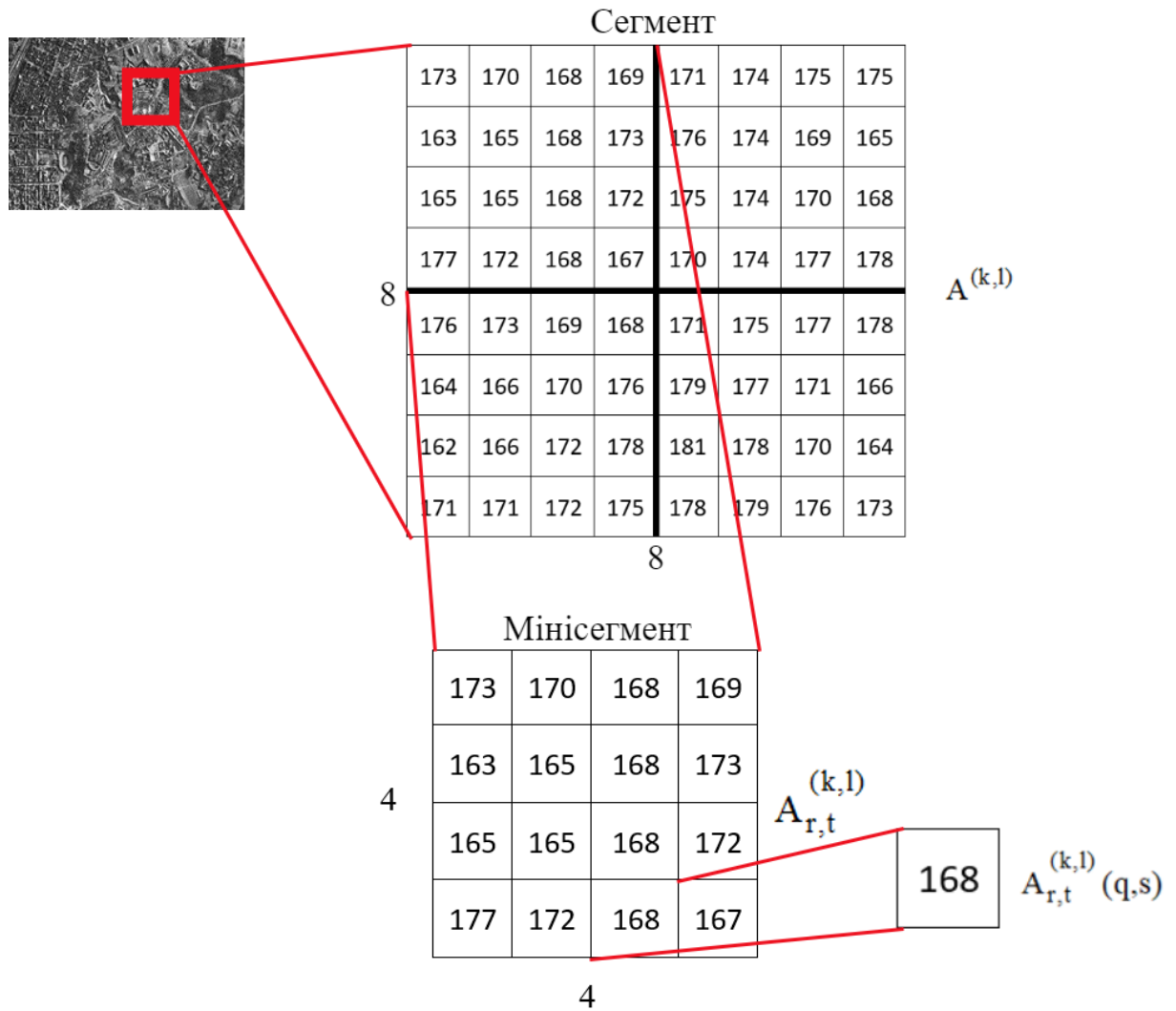


Рисунок 3.1 – Схема сегментації зображення

Додаткову увагу слід приділити обробці меж кадру. Якщо розміри зображення не кратні 8, застосовується механізм доповнення країв шляхом реплікації граничних рядків і стовпців [135]. Це забезпечує коректне формування сегментів без порушення структури даних та гарантує, що після декодування оригінальні розміри кадру відновлюються без спотворень [32, 42, 56]. Такий підхід відповідає усталеним практикам блокових кодеків [28].

Для ілюстрації роботи сегментації наведено приклад реального сегменту на рисунку 3.2.

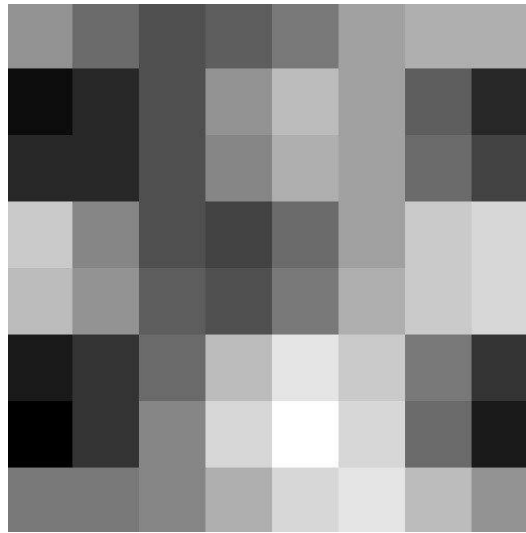


Рисунок 3.2 – Піксельне відображення сегменту $A^{(1,1)}$, для прикладів

Представимо піксельне відображення у вигляді матриці із наступними значеннями:

$$A^{(1,1)} = \begin{pmatrix} 173 & 170 & 168 & 169 & 171 & 174 & 175 & 175 \\ 163 & 165 & 168 & 173 & 176 & 174 & 169 & 165 \\ 165 & 165 & 168 & 172 & 175 & 174 & 170 & 167 \\ 177 & 172 & 168 & 167 & 170 & 174 & 177 & 178 \\ 176 & 173 & 169 & 168 & 171 & 175 & 177 & 178 \\ 164 & 166 & 170 & 176 & 179 & 177 & 171 & 166 \\ 162 & 166 & 172 & 178 & 181 & 178 & 170 & 164 \\ 171 & 171 & 172 & 175 & 178 & 179 & 176 & 173 \end{pmatrix}. \quad (3.1)$$

Відповідно у ролі мінісегменту $A_{1,1}^{(1,1)}$ виступатиме наступна матриця:

$$A_{1,1}^{(1,1)} = \begin{pmatrix} 173 & 170 & 168 & 169 \\ 163 & 165 & 168 & 173 \\ 165 & 165 & 168 & 172 \\ 177 & 172 & 168 & 167 \end{pmatrix}. \quad (3.2)$$

Значення з мінісегменту $A_{1,1}^{(1,1)}$, відповідають наступним виразам:

$$\begin{aligned}
 A_{1,1}^{(1,1)}(1,1) &= 173; \quad A_{1,1}^{(1,1)}(2,1) = 170; \\
 A_{1,1}^{(1,1)}(1,2) &= 163; \quad A_{1,1}^{(1,1)}(2,2) = 165.
 \end{aligned}
 \tag{3.3}$$

Таким чином, запропонований метод ієрархічної сегментації кадру забезпечує структуроване представлення інфрачервоного зображення. Воно формалізоване математичними виразами, що складають основу для опису подальших етапів перетворень [18, 23, 45]. Ієрархічність моделі дозволяє адаптувати подальші алгоритмічні процедури до властивостей даних, що формує можливості для підвищення ефективності наступних етапів кодування [98].

3.2 Розробка методу різницевого кодування для формування мінісегментів зі зменшеним динамічним діапазоном

Інфрачервоні зображення із високою просторовою кореляцією містять значну частку надлишкової інформації. Вона пов'язана з локальними фоновими областями сцени. Навіть після біфуркації бітових пластів, коли структурно значуща інформація виділена в окремий 8-бітний пласт старших бітів, локальний динамічний діапазон значень у сегментах $A^{(k,1)}$ і відповідних мінісегментах $A_{r,t}^{(k,1)}$ може залишатися суттєво широким. Це може зменшувати ефективність наступних перетворень та кодування в цілому. Оскільки частина бітового ресурсу витрачається на опис рівня фону, а не відносних варіацій, які несуть основну інформативність [13].

Для зменшення цього локального надлишку запропоновано метод різницевого кодування мінісегментів. За цим методом абсолютні значення інтенсивності замінюються на відносні відхилення від локального базового рівня. Ідея полягає у переході до різницевого простору, де описуються саме локальні перепади, що відповідають температурним градієнтам і контурам об'єктів [59].

Для того щоб перетворити мінісегмент $A_{r,t}^{(k,l)}$ у різницевий простір необхідно визначити мінімальний елемент $\mu_{r,t}^{(k,l)}$ у мінісегменті:

$$\mu_{r,t}^{(k,l)} = \min_{A_{r,t}^{(k,l)}(q,s) \in A_{r,t}^{(k,l)}} \{A_{r,t}^{(k,l)}(q,s)\}. \quad (3.4)$$

Далі здійснюється перехід у різницевий простір шляхом віднімання мінімуму $\mu_{r,t}^{(k,l)}$ від кожного елемента $A_{r,t}^{(k,l)}(q,s)$:

$$A'_{r,t}{}^{(k,l)}(q,s) = A_{r,t}^{(k,l)}(q,s) - \mu_{r,t}^{(k,l)}. \quad (3.5)$$

Де $A'_{r,t}{}^{(k,l)}(q,s)$ – отриманий різницевий елемент. Враховуючи приналежність значень до мінісегментів, формалізуємо утворення різницевих мінісегментів наступним виразом:

$$A_{r,t}^{(k,l)}(q,s) \in A_{r,t}^{(k,l)} \Rightarrow A'_{r,t}{}^{(k,l)} = \{A'_{r,t}{}^{(k,l)}(q,s)\}. \quad (3.6)$$

Де $A'_{r,t}{}^{(k,l)}$ - це мінісегмент у різницевому просторі. У результаті такого перетворення всі елементи різницевого мінісегмента невід'ємні, а мінімальне значення завжди дорівнює нулю:

$$A'_{r,t}{}^{(k,l)} \in \left[0, v_{r,t}^{(k,l)} - \mu_{r,t}^{(k,l)}\right]. \quad (3.7)$$

Де $v_{r,t}^{(k,l)}$ – максимальне значення у вихідному мінісегменті $A_{r,t}^{(k,l)}$.

Отже, локальний динамічний діапазон у мінісегменті зменшується з $\left[v_{r,t}^{(k,l)}, \mu_{r,t}^{(k,l)}\right]$ до $\left[0, v_{r,t}^{(k,l)} - \mu_{r,t}^{(k,l)}\right]$, а фоновий рівень винесено у скалярний параметр $\mu_{r,t}^{(k,l)}$.

Перетворення є оборотним за умови збереження значення $\mu_{r,t}^{(k,l)}$. На етапі декодування відновлення вихідного мінісегмента $A_{r,t}^{(k,l)}$ здійснюється за формулою:

$$A_{r,t}^{(k,l)} = A'_{r,t}{}^{(k,l)} + \mu_{r,t}^{(k,l)}. \quad (3.8)$$

Це забезпечує точне відтворення даних за відсутності додаткових втрат на наступних етапах обробки [37, 122, 161]. Таким чином, різницеве кодування вводить для кожного мінісегмента $A_{r,t}^{(k,l)}$ компактну параметричну модель: скалярний базовий рівень $\mu_{r,t}^{(k,l)}$ та матрицю варіацій $A'_{r,t}{}^{(k,l)}$, які разом задають локальну структуру сигналу.

Розглянемо приклад перетворення у різницевий простір мінісегменту $A_{1,1}^{(1,1)}$ (Рис. 3.3). Мінімальним значенням мінісегменту $A_{1,1}^{(1,1)}$ є $\mu_{1,1}^{(1,1)} = 163$. Тоді значеннями різницевого мінісегменту будуть:

$$\begin{aligned} A'_{1,1}{}^{(1,1)}(1,1) &= 10; \quad A'_{1,1}{}^{(1,1)}(2,1) = 7; \quad A'_{1,1}{}^{(1,1)}(3,1) = 5; \quad A'_{1,1}{}^{(1,1)}(4,1) = 6; \\ A'_{1,1}{}^{(1,1)}(1,2) &= 0; \quad A'_{1,1}{}^{(1,1)}(2,2) = 2; \quad A'_{1,1}{}^{(1,1)}(3,2) = 5; \quad A'_{1,1}{}^{(1,1)}(4,2) = 10; \\ A'_{1,1}{}^{(1,1)}(1,3) &= 2; \quad A'_{1,1}{}^{(1,1)}(2,3) = 2; \quad A'_{1,1}{}^{(1,1)}(3,3) = 5; \quad A'_{1,1}{}^{(1,1)}(4,3) = 9; \\ A'_{1,1}{}^{(1,1)}(1,4) &= 14; \quad A'_{1,1}{}^{(1,1)}(2,4) = 9; \quad A'_{1,1}{}^{(1,1)}(3,4) = 5; \quad A'_{1,1}{}^{(1,1)}(4,4) = 4. \end{aligned} \quad (3.9)$$

Із утворених різницевих значень формується мінісегмент у різницевому просторі $A'_{1,1}{}^{(1,1)}$. На рисунку 3.3 зображена схема різницевого кодування мінісегменту.



Рисунок 3.3 – Схема різницевого кодування мінісегменту

Отже, розроблений метод різницевого кодування мінісегментів виконує зменшення динамічного діапазону без втрати семантично важливих температурних градієнтів. В результаті, утворені дані представляють семантично значущу інформацію, готову до подальшої обробки.

3.3 Розробка методу перетворення сегментів у спектрально-різницевий простір

При застосуванні вейвлет перетворення W до різницевого мінісегменту $A'_{r,t}^{(k,l)}$ буде здійснено перетворення у спектрально-різницевий простір:

$$\left(H'_{r,t}^{(k,l)}, L'_{r,t}^{(k,l)}(I) \right) = W \left(A'_{r,t}^{(k,l)} \right). \quad (3.10)$$

Де:

- $H'_{r,t}^{(k,l)}$ – високочастотні коефіцієнти у спектрально-різницевому просторі;
- $L'_{r,t}^{(k,l)}(I)$ – низькочастотний коефіцієнт у спектрально-різницевому просторі.

Розглянемо детальніше принцип застосування вейвлет перетворення до різницевого мінісегменту $A'_{1,1}^{(1,1)}$. Нехай початкова група низькочастотних значень L_0 для цього мінісегменту має наступні значення:

$$A'_{1,1}^{(1,1)} = L_0 = \begin{pmatrix} 10 & 7 & 0 & 1 \\ 0 & 2 & 0 & 5 \\ 0 & 0 & 1 & 5 \\ 12 & 7 & 1 & 0 \end{pmatrix}. \quad (3.11)$$

Розглянемо розрахунок проміжкових частотних значень L_1 та L_2 для мінісегменту $A'_{1,1}^{(1,1)}$. На основі вихідного значення L_0 розрахунок L_1 та H_1 зображено на рисунку 3.4.

На схемі відображений етап розрахунку індексного діапазону:

$$p = \frac{G_{\text{size}}}{2} = \frac{16}{2} = 8;$$

$$p = \{0, \dots, 7\}. \quad (3.12)$$

Використовуючи утворені значення p відбувається індексація між парами значень $L_0(2p+1)$ та $L_0(2p+2)$.

В результаті першого рівня обрахунку вейвлету Хаара [102] утворюються групи спектральних коефіцієнтів L_1 та H_1 :

$$L_1 = \{8.5; 5.5; 1; 7.5; 2; 7; 11.5; 4.5\};$$

$$H_1 = \{-1.5; 0.5; 1; 2.5; 0; 2; -2.5; -0.5\}. \quad (3.13)$$

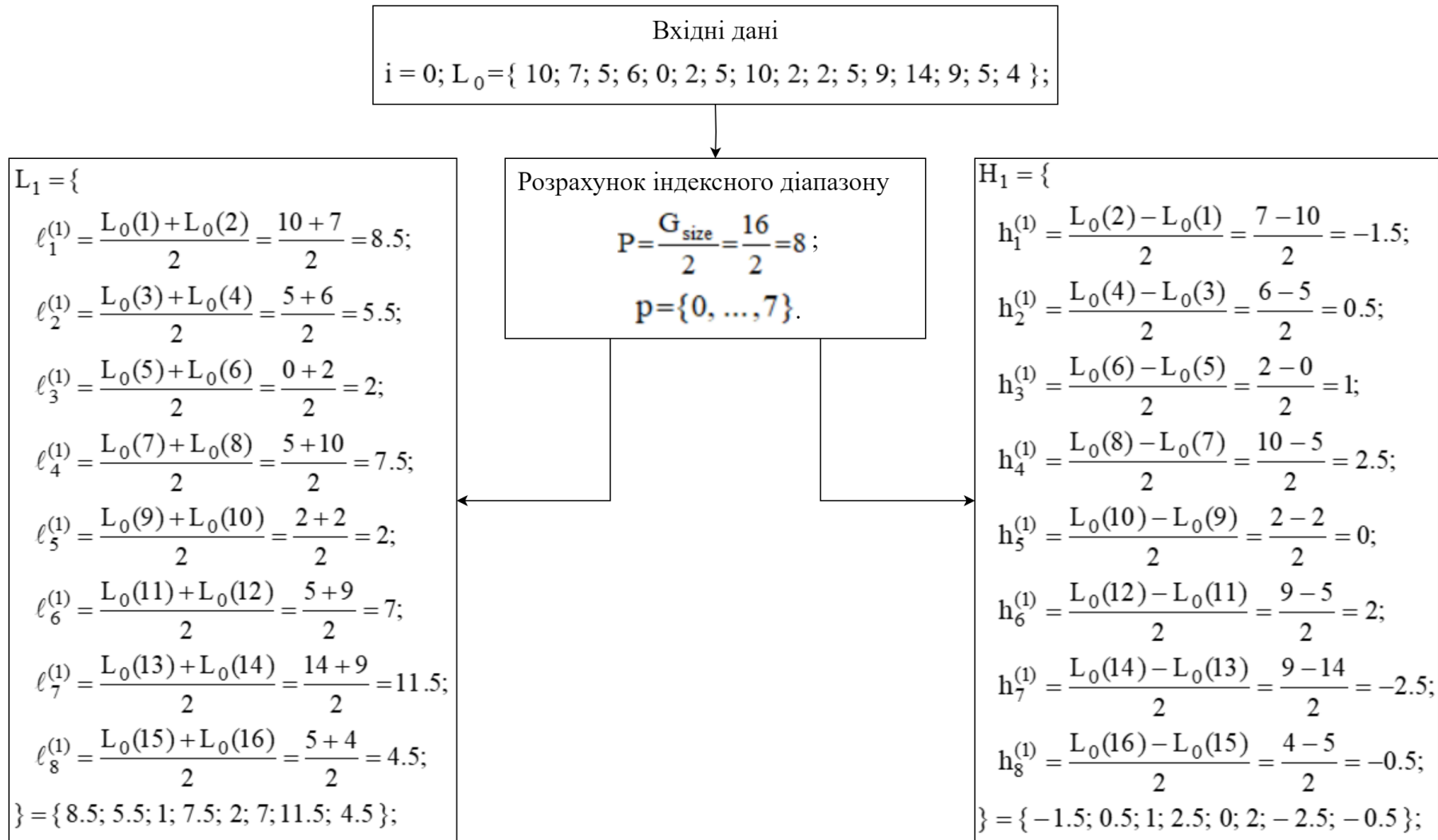


Рисунок 3.4 – Схема розрахунку проміжкових частотних значень L_1 та H_1

Отримавши значення L_1 та H_1 можна переходити до наступного рівня обрахунку вейвлету Хаара, коли $i = 2$. Він полягає у розрахунку L_2 та H_2 на основі отриманих низькочастотних значень L_1 . На рисунку 3.5 зображена схема утворення груп значень L_2 та H_2 .

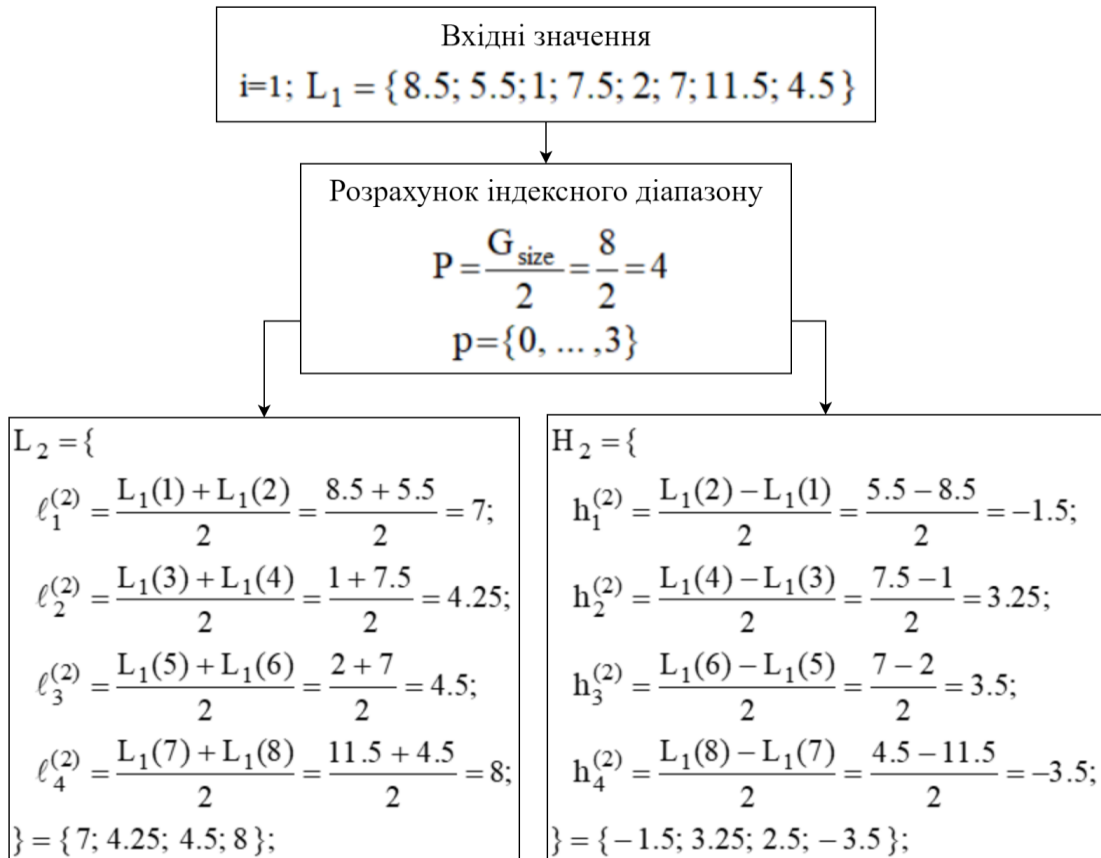


Рисунок 3.5 – Приклад розрахунку L_2 та H_2

Продовжуючи таким чином розрахунки для наступних ітерацій, результатом застосування вейвлету Хаара до різницевого мінісегменту $A'_{1,1}^{(1,1)}$ буде:

$$L_{1,1}^{(1,1)}(I) = 3.1875;$$

$$H_{1,1}^{(1,1)} = \left\{ \begin{array}{l} 0.3125; -1.375; 1.75; -1.5; \\ 3.25; 2.5; -3.5; -1.5; 0.5; 1.0; \\ 2.5; 0.0; 2.0; -2.5; -0.5; \end{array} \right\}. \quad (3.14)$$

де:

- $L_{1,1}^{(1,1)}(I)$ – крайнє низькочастотне значення після застосування вейвлету Хаара до різницевого мінісегменту $A'_{1,1}^{(1,1)}$;
- $H_{1,1}^{(1,1)}$ – набір високочастотних значень після застосування вейвлету Хаара до різницевого мінісегменту $A'_{1,1}^{(1,1)}$.

Таким чином відбувається перетворення у спектрально-різницевий простір кожного мінісегменту з сегменту $A^{(1,1)}$. В результаті сегмент представляють групи спектральних коефіцієнтів [146]. На рисунку 3.6 зображено схему утворених спектральних значень. На даній схемі позначені:

- $A'_{1,1}^{(1,1)}$, $A'_{1,2}^{(1,1)}$, $A'_{2,1}^{(1,1)}$, $A'_{2,2}^{(1,1)}$ – різницеві мінісегменти утворені із сегменту $A^{(1,1)}$;
- $H_{1,1}^{(1,1)}$, $H_{1,2}^{(1,1)}$, $H_{2,1}^{(1,1)}$, $H_{2,2}^{(1,1)}$ – групи високочастотних значень після застосування вейвлету Хаара до різницевих мінісегментів;
- $L_{1,1}^{(1,1)}(I)$ – отримане низькочастотне значення після застосування вейвлету Хаара до різницевого мінісегменту $A'_{1,1}^{(1,1)}$;
- $L_{1,2}^{(1,1)}(I)$ – отримане низькочастотне значення після застосування вейвлету Хаара до різницевого мінісегменту $A'_{1,2}^{(1,1)}$;
- $L_{2,1}^{(1,1)}(I)$ – отримане низькочастотне значення після застосування вейвлету Хаара до різницевого мінісегменту $A'_{2,1}^{(1,1)}$;
- $L_{2,2}^{(1,1)}(I)$ – отримане низькочастотне значення після застосування вейвлету Хаара до різницевого мінісегменту $A'_{2,2}^{(1,1)}$.

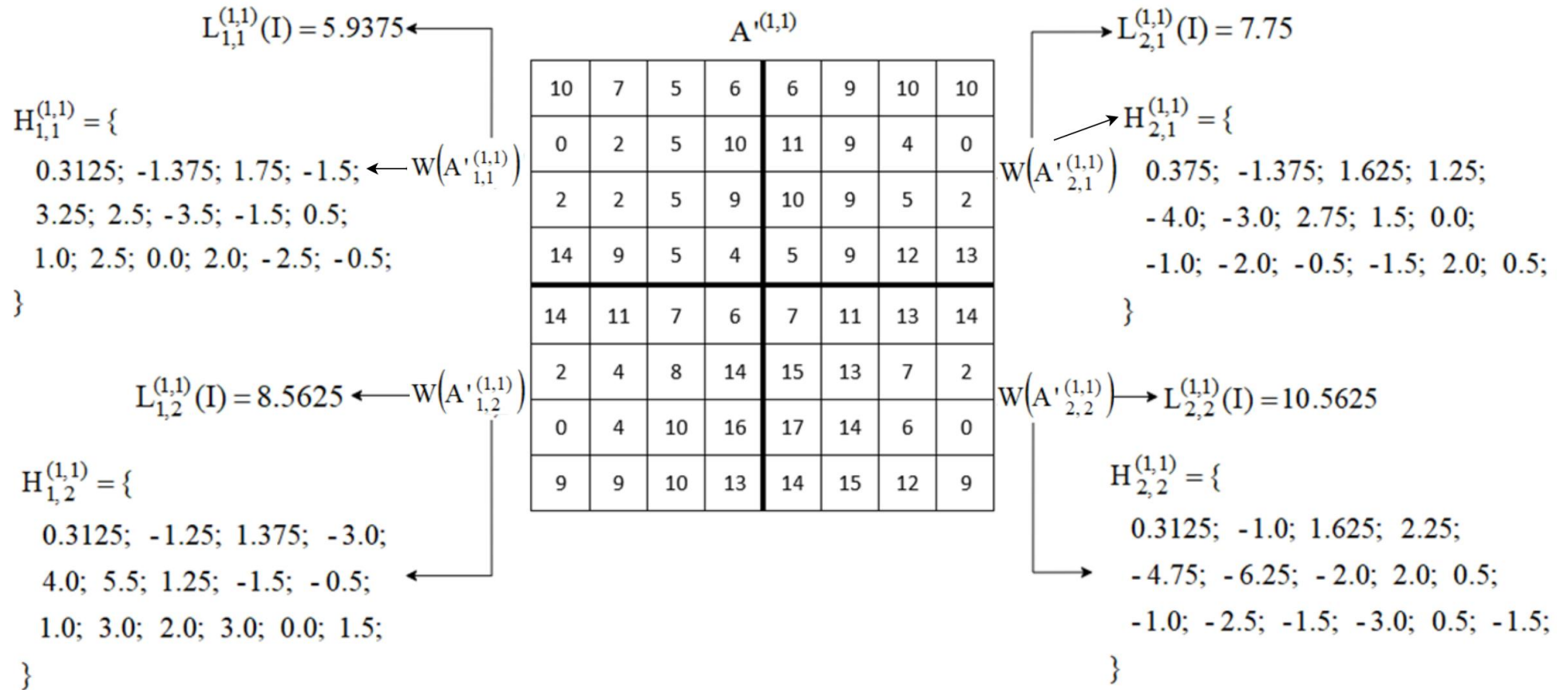


Рисунок 3.6 – Схема застосування вейвлету Хаара до мінісегментів з різницевого сегменту $A'_{r,t}^{(k,l)}$

Отримане спектральне представлення характеризується компактним розподілом енергії. Значна її частина концентрується у невеликій кількості низькочастотних коефіцієнтів, тоді як амплітуди високочастотних компонентів є меншими. Це підвищує потенційну ефективність подальших операцій квантування і кодування, оскільки дозволяє гнучко розподіляти бітовий ресурс між апроксимаційною та деталізуючими складовими, а також застосовувати різні режими точності залежно від семантичної важливості.

Додатковою перевагою є те, що вейвлет Хаара застосовується до різницевих мінісегментів $A'_{r,t}^{(k,l)}$, у яких базовий рівень уже винесений у скалярний параметр $\mu_{r,t}^{(k,l)}$. Це зменшує надлишкову низькочастотну складову у сигналі, що позитивно впливає на ефективність подальшої обробки.

3.4 Розробка методу корекції низькочастотних коефіцієнтів на основі високочастотних компонент вейвлет-перетворення

Утворене спектральне представлення мінісегментів на основі вейвлет-перетворення Хаара містить низькочастотний коефіцієнт $L_{r,t}^{(k,l)}(I)$ та відповідний кінцевий високочастотний коефіцієнт $H_{r,t}^{(k,l)}(I)$ [125]. Ці коефіцієнти отримуються з розрахунку напівсум та на піврізниць пар сусідніх елементів x_1 та x_2 :

$$L = \frac{x_1 + x_2}{2}; H = \frac{x_1 - x_2}{2}. \quad (3.15)$$

З огляду на те, що вихідні значення різницевого мінісегмента є невід'ємними (вони формуються відніманням локального мінімуму), для будь-якої пари $x_1 \geq 0$. Виконується нерівність:

$$|x_1 - x_2| \leq x_1 + x_2. \quad (3.16)$$

Ділячи обидві частини на 2, отримуємо фундаментальну для подальшого аналізу залежність:

$$\left| \frac{x_1 - x_2}{2} \right| \leq \frac{x_1 + x_2}{2} \Rightarrow |H| \leq L. \quad (3.17)$$

Отже, кінцеве низькочастотне значення завжди не менше за абсолютне значення відповідного високочастотного коефіцієнта:

$$\left| H_{r,t}^{(k,l)}(I) \right| \leq L_{r,t}^{(k,l)}(I). \quad (3.18)$$

Запропонований метод корекції низькочастотних коефіцієнтів використовує цю властивість для додаткового зменшення динамічного діапазону низькочастотних значень без внесення спотворень у дані [3]. Ідея полягає в тому, щоб частину величини, яка завідомо надлишкова з точки зору нерівності $|H| \leq L$, вилучити з низькочастотного коефіцієнта, не змінюючи при цьому високочастотну компоненту. Формально вводиться модифіковане низькочастотне значення \tilde{L} :

$$\tilde{L} = L - |H|. \quad (3.19)$$

Оскільки $|H| \leq L$, гарантовано, що $\tilde{L} \geq 0$, а отже, новий коефіцієнт належить до вузького, зміщеного до нуля діапазону. При цьому інформація не втрачається: під час декодування початкове значення L може бути однозначно відновлене за відомими \tilde{L} та H за простою оборотною формулою:

$$L = \tilde{L} + |H|. \quad (3.20)$$

Таким чином, перехід від пари (L, H) до пари (\tilde{L}, H) є перетворенням у просторі коефіцієнтів, що не створює жодних додаткових вимог до службової інформації [90].

Зменшення динамічного діапазону низькочастотних коефіцієнтів має прямий вплив на ефективність подальшого кодування [114]. Оскільки саме низькочастотні значення мають більшу амплітуду і визначають верхню межу діапазону для спектральних коефіцієнтів. Операція корекції низькочастотних значень призводить до зменшення максимальних амплітуд. Це є причиною зниження кількості біт, необхідних для представлення значення. Отже, підвищується бітова насиченість представлення, дозволяючи ту саму інформацію описати меншою кількістю біт без втрат точності [50].

Для кількісної оцінки ефекту від використання запропонованої корекції вводиться показник бітового об'єму. V_{enc} - це кількість біт, необхідних для збереження закодованого кадру без низькочастотної корекції, а V_{cor} - з нею. Тоді різниця між ними відображає скорочену кількість біт:

$$\Delta V = V_{enc} - V_{cor}. \quad (3.21)$$

Показники ΔV можна побачити на рисунку 3.7, які були отримані у наслідок кодування інфрачервоних зображень із тепловізійної камери FLIR. Роздільна здатність зображень становить 640 на 512 пікселів.

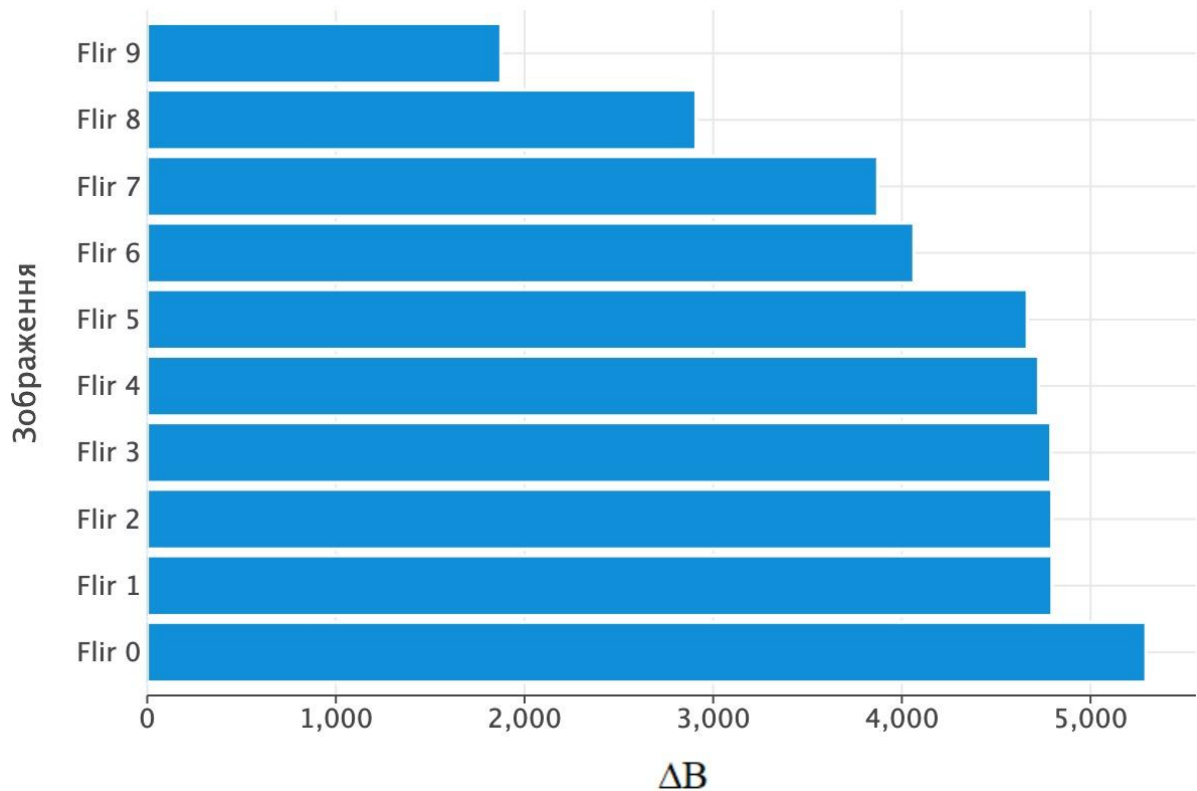


Рисунок 3.7 – Графік залежності зміни бітового об'єму ΔB від насиченості ІЧ-кадру

Результати, відображені на рисунку 3.7, показують стабільне додаткове зменшення бітового об'єму закодованих кадрів при збереженні повної відтворюваності початкових даних. Це означає, що використання аналітичної залежності між кінцевими низько- та високочастотними коефіцієнтами вейвлет-перетворення Хаара дозволяє підвищити ефективність кодування.

3.5 Розробка методу адаптивної квантизації на основі рівня інформативності та бітового пласту сегментів

Для зменшення бітового об'єму сегментів на основі їх рівня інформативності $M^{(k,l)}$ пропонується застосувати квантизацію. Квантизація є однією з базових операцій стиснення, що полягає у заміні вихідних значень сигналу обмеженим набором рівнів, унаслідок чого зменшується розрядність подання даних за рахунок контрольованої втрати точності. У контексті

інфрачервоних зображень доцільно застосовувати квантизацію до високочастотних коефіцієнтів $H_{r,t}^{(k,l)}$. Оскільки їх спотворення впливає на інтерпретацію деталей, тоді як спотворення низькочастотної компоненти може впливати на загальний рівень температурного поля.

Для врахування семантичних відмінностей бітових пластів, позначимо високочастотні коефіцієнти з пласту A_{msb} як $H_{msb,r,t}^{(k,l)}$, а з пласту A_{lsb} як $H_{lsb,r,t}^{(k,l)}$.

Операцію квантизації задамо у вигляді оператора Q :

$$Q(h, q) = \text{floor}\left(\frac{h}{q}\right). \quad (3.22)$$

Де h – вхідне значення для квантизації, $q > 0$ - коефіцієнт квантизації. Очевидно, що за збільшення q зменшується бітовий обсяг для зберігання значень, але водночас зростає рівень похибки відновлення. Таким чином, вибір q визначає компроміс між якістю реконструкції та ступенем стиснення.

З урахуванням двох бітових пластів та трьох рівнів інформативності доцільно ввести коефіцієнти квантизації, що дозволяють адаптивно враховувати цю семантичну специфіку. У таблиці 3.1 наведено позначення адаптивних коефіцієнтів квантизації.

Таблиця 3.1 – Позначення коефіцієнтів квантизації у залежності від семантичної специфіки даних

Рівень інформативності	Старший бітовий пласт A_{msb}	Молодший бітовий пласт A_{lsb}
Об'єктно-інформативний	$q_{obj}^{(MSB)}$	$q_{obj}^{(LSB)}$
Інформативний	$q_{inf}^{(MSB)}$	$q_{inf}^{(LSB)}$
Малоінформативний	$q_{low}^{(MSB)}$	$q_{low}^{(LSB)}$

Таким чином параметри $q_{obj}^{(MSB)}$, $q_{inf}^{(MSB)}$, $q_{low}^{(MSB)}$ задають інтенсивність квантування високочастотних коефіцієнтів у пласті старших бітів A_{msb} для об'єктно-інформативних, інформативних, низькоінформативних сегментів відповідно. Наступні коефіцієнти $q_{obj}^{(LSB)}$, $q_{inf}^{(LSB)}$, $q_{low}^{(LSB)}$ є аналогічними параметрами для пласту молодших бітів A_{lsb} .

Виходячи зі специфіки інфрачервоних зображень, формується така ієрархія коефіцієнтів квантизації. По-перше, для об'єктно-інформативних сегментів $M^{(k,1)} = 2$, коефіцієнти квантизації мають бути мінімальними, оскільки саме в цих областях зосереджені теплові сигнатури об'єктів інтересу. Для інформативних сегментів $M^{(k,1)} = 1$, допускаються вищі значення коефіцієнтів квантизації. Однак рівень спотворень повинен бути контрольованим, оскільки такі сегменти містять контури об'єктів середовища та температурні градієнти. Для малоінформативних сегментів $M^{(k,1)} = 0$, допускаються високі значення коефіцієнтів квантизації. Наведені правила відображаються нерівностями:

$$\begin{aligned} q_{obj}^{(MSB)} &\leq q_{inf}^{(MSB)}, q_{inf}^{(MSB)} \leq q_{low}^{(MSB)}, q_{inf}^{(MSB)} \leq q_{low}^{(MSB)}; \\ q_{obj}^{(LSB)} &\leq q_{inf}^{(LSB)}, q_{inf}^{(LSB)} \leq q_{low}^{(LSB)}, q_{inf}^{(LSB)} \leq q_{low}^{(LSB)}. \end{aligned} \quad (3.23)$$

По-друге, необхідно врахувати семантику бітових пластів. Квантування може бути агресивнішим для кожного рівня інформативності у молодшому бітовому пласті. Тобто:

$$q_{obj}^{(MSB)} \leq q_{obj}^{(LSB)}, q_{inf}^{(MSB)} \leq q_{inf}^{(LSB)}, q_{low}^{(MSB)} \leq q_{low}^{(LSB)}. \quad (3.24)$$

Таке співвідношення узгоджується з тим, що старші біти несуть основну структурну інформацію, тоді як спотворення молодших бітів менш помітні на рівні семантичної інтерпретації сцени.

Підсумовуючи, ієрархію коефіцієнтів квантизації можна описати сукупністю умов:

$$q_{obj}^{(MSB)} \leq q_{obj}^{(LSB)} \leq q_{inf}^{(MSB)} \leq q_{inf}^{(LSB)} \leq q_{low}^{(MSB)} \leq q_{low}^{(LSB)}. \quad (3.25)$$

Ця нерівність відображає одночасно пріоритет збереження інформативних сегментів та пріоритет збереження MSB пласту порівняно з LSB пластом.

На основі визначених правил для кожного сегмента у позиції (k, l) визначаються коефіцієнти квантизації:

$$q_{MSB}(k, l) = \begin{cases} q_{obj}^{(MSB)}, & M^{(k, l)} = 2 \\ q_{inf}^{(MSB)}, & M^{(k, l)} = 1; \\ q_{low}^{(MSB)}, & M^{(k, l)} = 0 \end{cases} \quad (3.26)$$

$$q_{LSB}(k, l) = \begin{cases} q_{obj}^{(LSB)}, & M^{(k, l)} = 2 \\ q_{inf}^{(LSB)}, & M^{(k, l)} = 1. \\ q_{low}^{(LSB)}, & M^{(k, l)} = 0 \end{cases} \quad (3.27)$$

Квантування високочастотних коефіцієнтів у пласті старших бітів виконується за правилом:

$$Q\left(H_{msb}_{r,t}^{(k,l)}, q_{MSB}(k, l)\right). \quad (3.28)$$

Аналогічно у пласті молодших бітів:

$$Q\left(H_{lsb}_{r,t}^{(k,l)}, q_{LSB}(k, l)\right). \quad (3.29)$$

Таким чином, сформований підхід адаптивної квантизації, де рівень інформативності сегмента впливає на допустимий рівень спотворень у кожному бітовому пласті. Це дозволяє зосередити бітовий ресурс на відтворенні важливих ділянок кадру, одночасно агресивно стискаючи малозначущі фонові області.

На рисунку 3.8 а) відображено вхідний ІЧ-кадр, а на рисунку 3.8 б) – той самий кадр після кодування із застосуванням розробленого методу адаптивної квантизації. Видно, що цілісність об'єкта інтересу та інших висококонтрастних елементів сцени збережена. Водночас на фонових ділянках кадру спостерігаються регулярні артефакти у вигляді сітчастої структури, що є наслідком застосування агресивнішої квантизації в малоінформативних областях. Для порівняння, зображено кадр (рис 3.8 в) до якого була застосована квантизація із тими самим коефіцієнтами квантизації, однак без адаптивності. Помітна очевидна втрата термальної цілісності об'єкту.

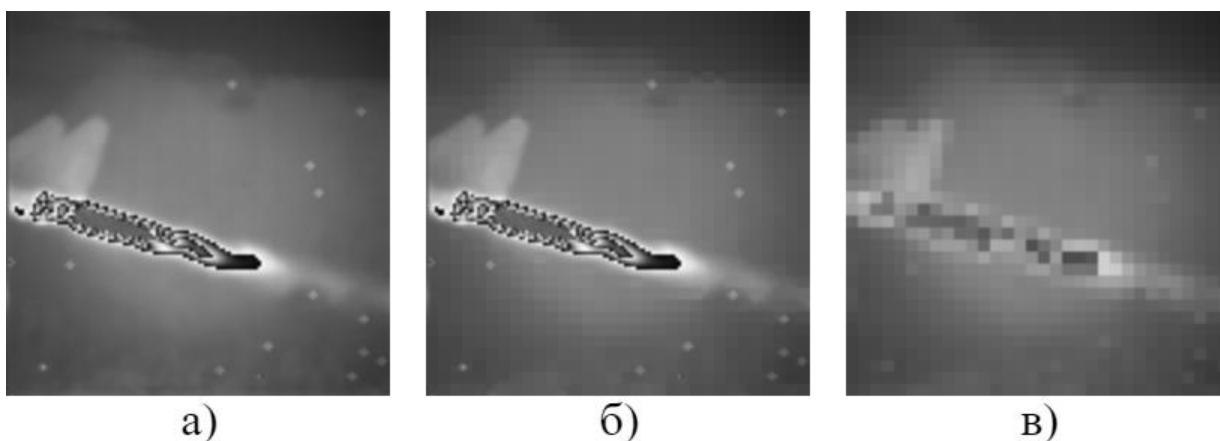


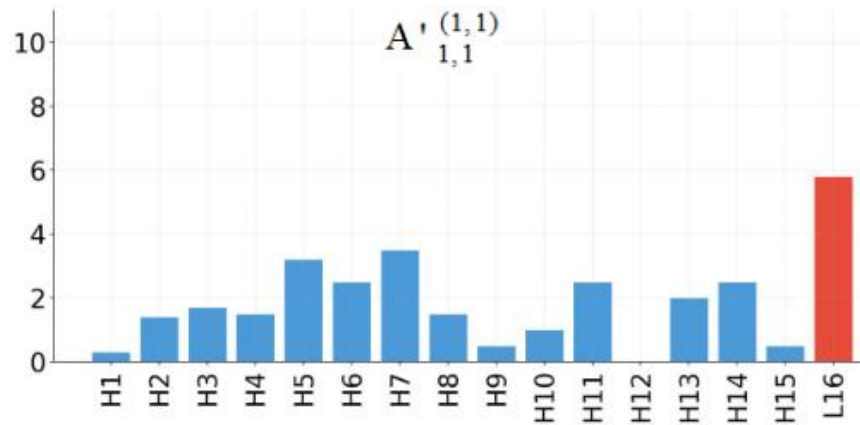
Рисунок 3.8 – Приклад застосування адаптивної квантизації на основі виявлення теплової сигнатури: а) вхідне ІЧ-зображення; б) зображення після адаптивної квантизації; в) зображення після загальної квантизації

Розроблений метод адаптивної квантизації дозволяє визначити інтенсивність стиснення високочастотних коефіцієнтів відповідно до рівня

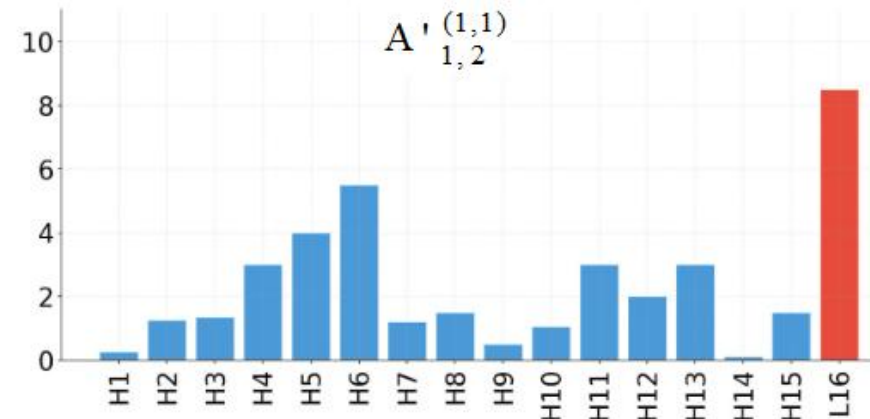
інформативності сегмента та його приналежності до старшого або молодшого бітового пласту. Такий підхід забезпечує концентрацію бітового ресурсу на інформативних ділянках кадру та дозволяє зменшувати загальний обсяг даних без порушення цілісності теплових сигнатур. Водночас застосування високих коефіцієнтів квантизації для малоінформативних областей зменшує їх просторову деталізацію та формує артефакти. Цей вплив зменшує цілісність фону, проте дозволяє зберігати семантичну цілісність об'єктів інтересу.

3.6 Розробка методу спектрально-групового кодування для зменшення бітової інтенсивності кадрів

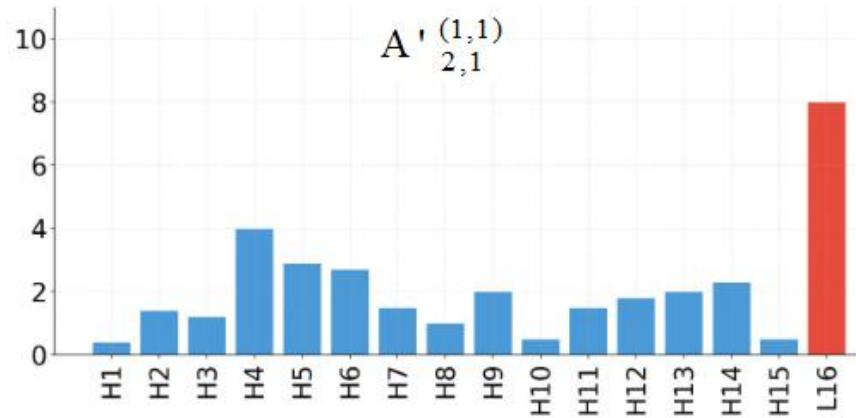
Для подальшого стиснення спектральних коефіцієнтів $L_{r,t}^{(k,l)}$ (I) та $H_{r,t}^{(k,l)}$ необхідно врахувати їх статистичні відмінності. Низькочастотні коефіцієнти формуються як напівсуми елементів (апроксимаційна складова), тоді як високочастотні — як напіврізниці (деталізуюча складова). Внаслідок цього їхні динамічні діапазони істотно відрізняються. На рисунку 3.9 зображені гістограми, що побудовані для низькочастотних значень $L_{r,t}^{(k,l)}$ (I) та високочастотних коефіцієнтів $H_{r,t}^{(k,l)}$ у різницевих мінісегментах. Їх аналіз показує, що середній рівень низькочастотних значень приблизно на 40-45 % перевищує максимальні за модулем високочастотні значення $|H_{r,t}^{(k,l)}|$. Це емпірично доводить, що низькочастотна компонента сконцентрована в ширшому діапазоні, тоді як високочастотні коефіцієнти частіше зосереджені поблизу нуля. На гістограмах (рис. 3.9) стовпчики H1..H15 відповідають високочастотним значенням $|H_{r,t}^{(k,l)}|$, та стовпчики L16 відповідають низькочастотним значенням $L_{r,t}^{(k,l)}$ (I).



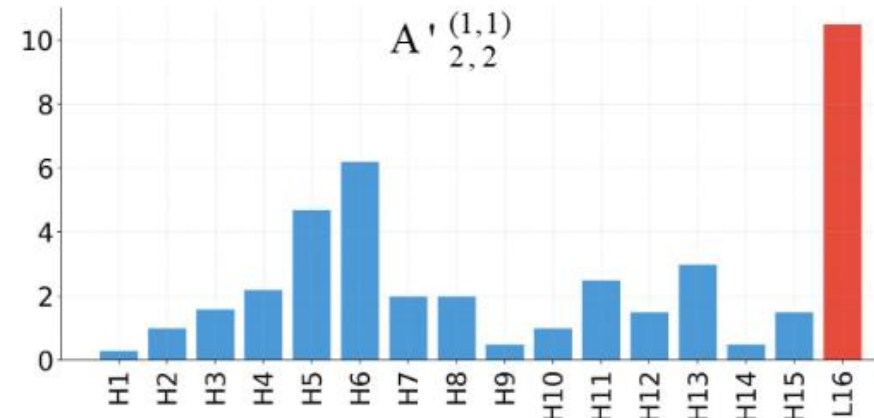
a)



б)



в)



г)

Рисунок 3.9 – Гістограми залежності типу частоти від частотних значень різницевої мінісегментів а) для мінісегменту

$A'_{1,1}^{(1,1)}$ б) для мінісегменту $A'_{1,2}^{(1,1)}$ в) для мінісегменту $A'_{2,1}^{(1,1)}$ г) для мінісегменту $A'_{2,2}^{(1,1)}$

Якщо кодувати низько- та високочастотні коефіцієнти в одному спільному числовому діапазоні, то низькочастотні значення будуть недостатньо точно відновлені, або високочастотні коефіцієнти будуть кодуватися надмірною кількістю біт. Це знижує ефективність подальшої обробки. Тому пропонується виконувати роздільне групування та кодування низько- та високочастотних компонент, з урахуванням специфіки їхніх динамічних діапазонів.

На першому етапі формуються групи низькочастотних значень для кожного сегмента. Сегмент $A^{(k,l)}$ складається з чотирьох різницевих мінісегментів $A'_{r,t}{}^{(k,l)}$, до яких застосовано вейвлет-перетворення Хаара. Отже, для кожного з цих мінісегментів маємо по одному кінцевому низькочастотному коефіцієнту $L_{r,t}{}^{(k,l)}(I)$. Тоді для сегмента $A^{(k,l)}$ формується група низькочастотних значень:

$$L^{(k,l)} = \{ L_{1,1}{}^{(k,l)}(I), L_{1,2}{}^{(k,l)}(I), L_{2,1}{}^{(k,l)}(I), L_{2,2}{}^{(k,l)}(I) \}. \quad (3.30)$$

Для відновлення низькочастотного мінісегменту $A'_{r,t}{}^{(k,l)}$ серед низьких частот необхідне лише кінцеве низькочастотне значення $L_{r,t}{}^{(k,l)}(I)$. Тому формування групи низькочастотних значень $L^{(k,l)}$ для сегменту $A^{(k,l)}$ включає тільки кінцеві низькочастотні значення.

Розглянемо приклад формування сукупності низькочастотних значень $L^{(1,1)}$ для сегменту $A^{(1,1)}$. На рисунку 3.10 зображена схема формування $L^{(1,1)}$ із різницевого сегменту $A^{(1,1)}$. Різницевий сегмент $A^{(1,1)}$ утворений шляхом трансформації у різницевий простір усіх його мінісегментів: $A'_{1,1}{}^{(1,1)}$, $A'_{1,2}{}^{(1,1)}$, $A'_{2,1}{}^{(1,1)}$, $A'_{2,2}{}^{(1,1)}$.

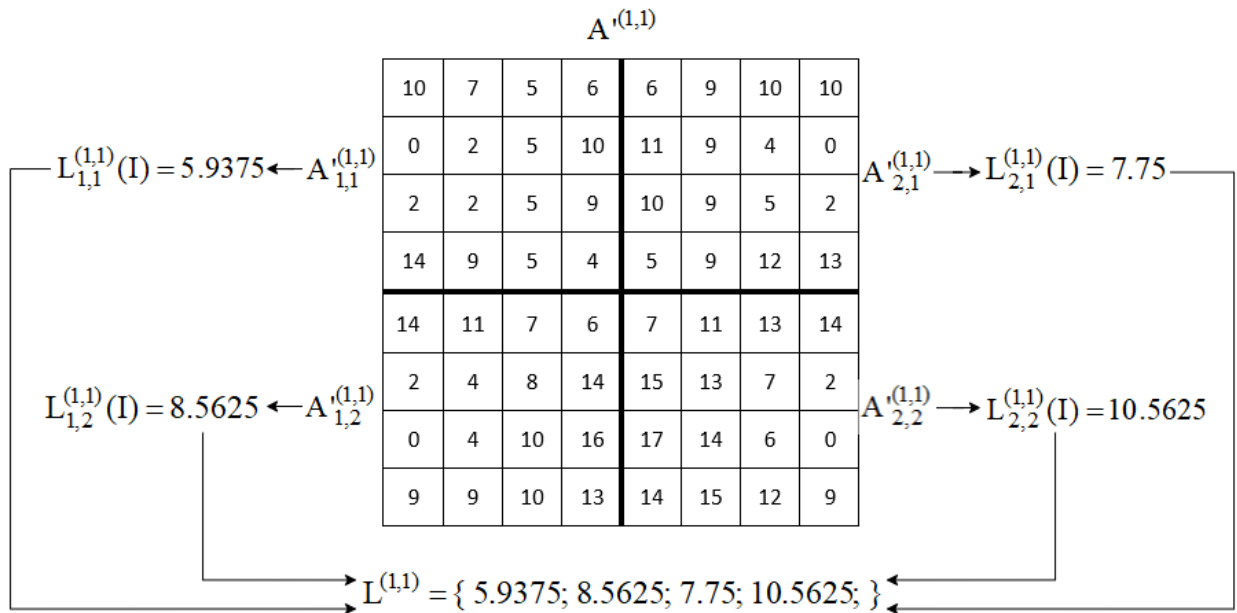


Рисунок 3.10 – Схема утворення групи низькочастотних значень $L^{(1,1)}$

Наступним кроком є перехід від групових наборів коефіцієнтів до компактних кодових представлень. Для групи низькочастотних значень $L^{(k,1)}$ визначається максимальний елемент λ :

$$\lambda = \max_{L^{(k,1)}(g) \in L^{(k,1)}} \{ L^{(k,1)}(g) \}. \quad (3.31)$$

Закодуємо групу низькочастотних значень $L^{(k,1)}$ використовуючи наступну формулу:

$$CL^{(k,1)} = \sum_{g=1}^{\text{len}(L^{(k,1)})} \text{floor}(L^{(k,1)}(g)) \cdot \text{floor}(\lambda)^{\text{len}(L^{(k,1)})-g}. \quad (3.32)$$

Де:

- $L^{(k,1)}(g)$ – елемент із групи низькочастотних значень $L^{(k,1)}$ з порядковим номером g ;

- floor – операція відокремлення цілого числа від залишку. Наприклад: $\text{floor}(5.9) = 5$;
- $\text{len}(L^{(k,1)})$ – кількість елементів у групі низькочастотних значень $L^{(k,1)}$. Відповідає кількості мінісегментів у сегменті;
- $CL^{(k,1)}$ – кодове значення, утворене із сукупності низькочастотних значень $L^{(k,1)}$.

Аналогічний підхід застосовується для груп високочастотних коефіцієнтів $H_{r,t}^{(k,1)}$. Для кожної групи $H_{r,t}^{(k,1)}$ визначається максимальне за модулем значення γ :

$$\gamma = \max_{H_{r,t}^{(k,1)}(g) \in H_{r,t}^{(k,1)}} \left\{ \left| H_{r,t}^{(k,1)}(g) \right| \right\}. \quad (3.33)$$

Після чого всі елементи високочастотної групи $H_{r,t}^{(k,1)}$ кодуються за наступною формулою:

$$CH_{r,t}^{(k,1)} = \sum_{g=1}^{\text{len}(H_{r,t}^{(k,1)})} \text{floor}\left(\left| H_{r,t}^{(k,1)}(g) \right|\right) \cdot \text{floor}(\gamma)^{\text{len}(H_{r,t}^{(k,1)})-g}. \quad (3.34)$$

Де:

- $H_{r,t}^{(k,1)}(g)$ – елемент у сукупності високочастотних значень $H_{r,t}^{(k,1)}$ з порядковим номером g ;
- $\text{len}(H_{r,t}^{(k,1)})$ – кількість елементів у групі високочастотних значень $H_{r,t}^{(k,1)}$ мінісегменту;
- $CH_{r,t}^{(k,1)}$ – кодове значення, утворене із набору значень $H_{r,t}^{(k,1)}$.

В результаті кодування кожен сегмент $A^{(k,1)}$ буде представлений кодом низькочастотних значень $CL^{(k,1)}$ та кодами високочастотних значень $CH_{r,t}^{(k,1)}$ відповідно до кожного мінісегменту $A_{r,t}^{(k,1)}$.

Розглянемо роботу кодування набору низькочастотних значень на прикладі. Припустимо, що група низькочастотних значень $L^{(1,1)}$ має наступні значення:

$$L^{(1,1)} = \{ 5.9375; 8.5625; 7.75; 10.5625; \}. \quad (3.35)$$

Тоді розрахунок коду $CL^{(1,1)}$ для низькочастотних значень $L^{(1,1)}$ з сегменту $A^{(1,1)}$ виглядатиме так:

$$\begin{aligned} \lambda &= 10.5625; \\ CL^{(1,1)} &= \text{floor}(5.9375) \cdot \text{floor}(10.5625)^3 + \text{floor}(8.5625) \cdot \text{floor}(10.5625)^2 + \\ &\quad \text{floor}(7.75) \cdot \text{floor}(10.5625)^1 + \text{floor}(10.5625) \cdot \text{floor}(10.5625)^0 = \quad (3.36) \\ &\quad 5 \cdot 10^3 + 8 \cdot 10^2 + 7 \cdot 10^1 + 10 \cdot 1 = 5880. \end{aligned}$$

На рисунку 3.11 зображено схематичним способом утворення низькочастотного спектрально-групового коду $CL^{(1,1)}$.

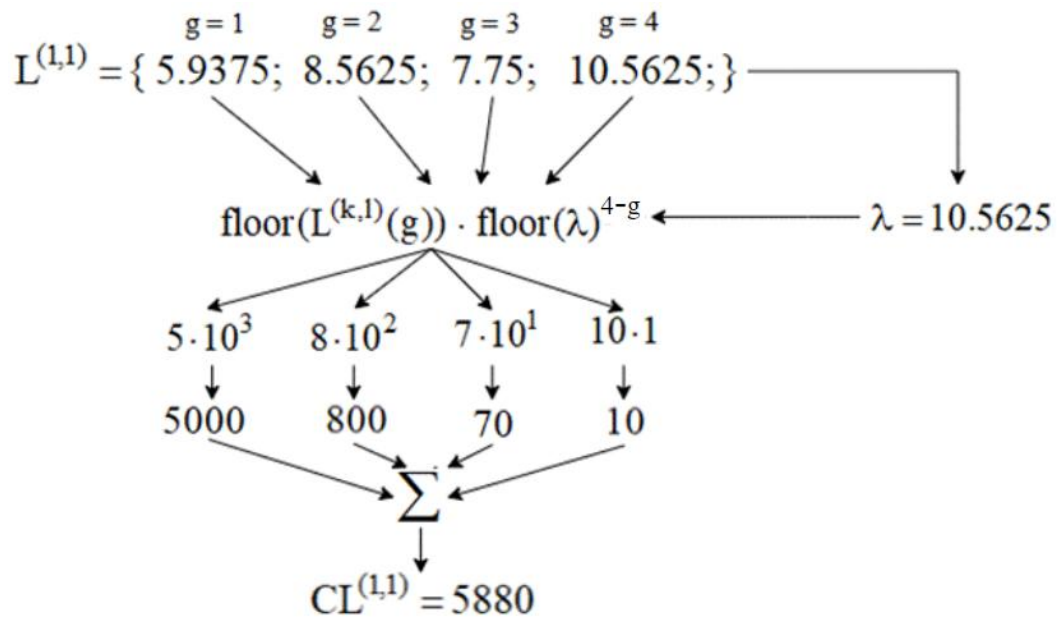


Рисунок 3.11 – Схема утворення низькочастотного спектрально-групового коду $CL^{(1,1)}$

Застосуємо високочастотне спектрально-групове кодування до груп високочастотних значень $H_{1,1}^{(1,1)}$, $H_{1,2}^{(1,1)}$, $H_{2,1}^{(1,1)}$, $H_{2,2}^{(1,1)}$ з прикладу рисунку 3.6.

Будуть отримані наступні коди:

- $CH_{1,1}^{(1,1)} = 2\,541\,723$;
- $CH_{1,2}^{(1,1)} = 1\,660\,630\,326$;
- $CH_{2,1}^{(1,1)} = 93\,210\,136$;
- $CH_{2,2}^{(1,1)} = 16\,269\,331\,141$.

У результаті застосування спектрально-групового кодування сегмент $A^{(1,1)}$ замість повного набору числових спектральних коефіцієнтів може бути представлений кодovими значеннями:

$$\left\{ CL^{(1,1)}, CH_{1,1}^{(1,1)}, CH_{1,2}^{(1,1)}, CH_{2,1}^{(1,1)}, CH_{2,2}^{(1,1)} \right\}. \quad (3.37)$$

Такий підхід обмежує кількість надлишкової інформації та знижує загальний бітовий обсяг представлення даних. Процес кодування потребує збереження допоміжних параметрів для точної зворотної реконструкції інфрачервоного сигналу. До основних допоміжних параметрів належать:

- мінімальні значення мінісегментів;
- максимальні значення групових спектральних коефіцієнтів;
- відкинуті залишки спектральних коефіцієнтів;
- інформація про знакові компоненти високочастотних коефіцієнтів.

Підсумовуючи, розроблений метод спектрально-групового кодування забезпечує роздільне формування компактних кодових представлень для низькочастотних і високочастотних спектральних коефіцієнтів. Запропоноване перетворення дозволяє зберігати інформативність сегментів при зменшенні бітрейту потоку даних, що створює умови для підтримки актуальності передачі інфрачервоних кадрів в інформаційних системах.

3.7 Розробка методу кодування інфрачервоних зображень на основі семантико-орієнтованої обробки сегментів

У розробленому методі кодування ІЧ-зображень обробка кадру починається із ряду структурно-статистичних перетворень. В першу чергу відбувається сегментація ІЧ-кадру $A^{(K,L)}$, у наслідок чого утворюються неперекривні сегменти $A^{(k,l)}$ розміром 8×8 пікселів.

Для кожного сегменту $A^{(k,l)}$ виконується бітова біфуркація $MSB(A^{(k,l)})$ та $LSB(A^{(k,l)})$. Це дозволяє утворити сегменти старшого бітового пласту та молодшого: $A_{msb}^{(k,l)}$ та $A_{lsb}^{(k,l)}$.

У подальшому, для кожного сегменту бітового пласту виконується ієрархічна сегментація за принципом квадродекомпозиції. В результаті

утворюються мінісегменти $Amsb_{r,t}^{(k,l)}$ для старшого бітового пласту та мінісегменти $Alsb_{r,t}^{(k,l)}$ для молодшого бітового пласту.

Для зменшення динамічного діапазону утворених мінісегментів застосовується різницеве перетворення. В результаті цієї обробки, старший бітовий пласт може бути представленим у вигляді різницевих мінісегментів $Amsb'_{r,t}^{(k,l)}$. А молодший бітовий пласт у вигляді різницевих мінісегментів $Alsb'_{r,t}^{(k,l)}$.

Для переходу у спектральний простір застосовується вейвлет перетворення. Воно дозволяє представити різницеві мінісегменти у вигляді наборів спектральних коефіцієнтів. Таким чином, при обробці мінісегментів старшого бітового шару $W(Amsb'_{r,t}^{(k,l)})$ утворюються високочастотні значення $Hmsb'_{r,t}^{(k,l)}$ та кінцеве низькочастотне значення $Lmsb'_{r,t}^{(k,l)}(I)$. При обробці мінісегментів молодшого бітового шару $W(Alsb'_{r,t}^{(k,l)})$ утворюються високочастотні значення $Hlsb'_{r,t}^{(k,l)}$ та кінцеве низькочастотне значення $Llsb'_{r,t}^{(k,l)}(I)$.

Наступним етапом, для утворених низькочастотних значень $Lmsb'_{r,t}^{(k,l)}(I)$ та $Llsb'_{r,t}^{(k,l)}(I)$ виконується корекція коефіцієнтів. Це дозволяє зменшити динамічний діапазон низькочастотних значень, не призводячи до втрат цілісності. Унаслідок чого, утворюються скореговані низькочастотні коефіцієнти $\tilde{Lmsb}'_{r,t}^{(k,l)}(I)$ та $\tilde{Llsb}'_{r,t}^{(k,l)}(I)$.

На основі передньо визначених високочастотних коефіцієнтів можна визначити метрику інформативності сегментів. Для цього необхідно виконати агрегацію абсолютних високочастотних коефіцієнтів із кожного мінісегменту та нормалізувати утворене значення. Враховуючи розділення сегментів на бітові пласти, дані розрахунки треба провести для сегментів

старшого бітового пласту $\text{Metr}(\text{Amsb}'^{(k,l)})$ та молодшого $\text{Metr}(\text{Alsb}'^{(k,l)})$.

Утворенні значення метрики інформативності позначаються як $\text{Amsb}'_M^{(k,l)}$ та $\text{Alsb}'_M^{(k,l)}$.

На основі значень метрики інформативності сегментів бітових пластів можна побудувати карту інформативності застосовуючи порогові правила. Для кожного бітового пласту застосовується окремі порогові значення, що дозволяє враховувати відмінну насиченість пластів. Врешті, для кожного сегменту визначається рівень інформативності $M^{(k,l)}$, спільний для сегментної позиції (k, l) в обох пластах Amsb та Alsb . В сукупності значення $M^{(k,l)}$ утворюють двопластову інформативну карту кадру.

Для більш точного визначення сегментів, що відповідають об'єктам інтересу може бути застосована модель штучного інтелекту. Модель ШІ дозволяє виявити область розташування об'єкта інтересу, а інформативні сегменти у цій області визначаються як об'єктно-інформативні. Для виявлення об'єктів інтересу модель ШІ повинна обробляти старший бітовий пласт Amsb , оскільки він відображає 8-бітну семантику ІЧ-кадру. Позначимо детектування об'єктів як $D(\text{Amsb})$. Важливо звернути увагу, що застосування ШІ є опціональним. Розроблений метод може кодувати ІЧ-кадри за умов обмежених апаратних ресурсів без виконання ШІ.

Маючи карту інформативності із значеннями $M^{(k,l)}$, що позиційно визначають рівень інформативності сегментів, можна застосувати просторове прорідження низькоінформативних областей. Цей етап дозволяє збільшити стиснення неінформативних областей. Однак для відновлення проріджених сегментів необхідно виконати їх реконструкцію на етапі декодування.

На основі рівня інформативності сегментів $M^{(k,l)}$ виконується адаптивна квантизація високочастотних коефіцієнтів. Для цього необхідно визначити коефіцієнти квантизації, що залежать від бітового пласту та рівня

інформативності. Таким чином виконання квантизації визначається як $Q\left(\text{Hmsb}'_{r,t}^{(k,l)}, q_{\text{MSB}}(k,l)\right)$ для мінісегментів старшого бітового пласту. Для мінісегментів молодшого бітового застосування квантизації визначається як $Q\left(\text{Hlsb}'_{r,t}^{(k,l)}, q_{\text{LSB}}(k,l)\right)$. В результаті квантизації утворюються квантовані високочастотні значення $Q\text{Hmsb}'_{r,t}^{(k,l)}$ та $Q\text{Hlsb}'_{r,t}^{(k,l)}$.

Для представлення спектральних коефіцієнтів у більш компактній бітовій формі виконується спектрально-групове кодування. В результаті його виконання групи однорідних спектральних значень описуються один кодовим значенням. Високочастотні спектральні значення вже сформовані у групи відповідно до кожного мінісегменту. Застосування спектрально-групового кодування до них матиме вигляд $\text{SGC}\left(Q\text{Hmsb}'_{r,t}^{(k,l)}\right)$ для старшого бітового пласту та $\text{SGC}\left(Q\text{Hlsb}'_{r,t}^{(k,l)}\right)$ для молодшого бітового пласту. Із кінцевих низькочастотних значень $\tilde{\text{Lmsb}}'_{r,t}^{(k,l)}(I)$ та $\tilde{\text{Llsb}}'_{r,t}^{(k,l)}(I)$, що відносяться до мінісегментів, сформовано сегментні групи низькочастотних значень $\tilde{\text{Lmsb}}'^{(k,l)}$ та $\tilde{\text{Llsb}}'^{(k,l)}$ для старшого та молодшого бітового пласту відповідно. Застосування спектрально-групового кодування до них має вигляд $\text{SGC}\left(\tilde{\text{Lmsb}}'^{(k,l)}\right)$ та $\text{SGC}\left(\tilde{\text{Llsb}}'^{(k,l)}\right)$.

В результаті спектрально-групового кодування утворюються кодові значення відповідно до кожної групи значень:

- $\text{CHmsb}'_{r,t}^{(k,l)}$ та $\text{CHlsb}'_{r,t}^{(k,l)}$ – кодові значення для високочастотних коефіцієнтів бітових пластів.
- $\text{CLmsb}'^{(k,l)}$ та $\text{CLlsb}'^{(k,l)}$ – кодові значення для низькочастотних коефіцієнтів бітових пластів.

На рисунок 3.12 зображена схема методу кодування ІЧ-зображень на основі семантико-орієнтованої обробки сегментів.

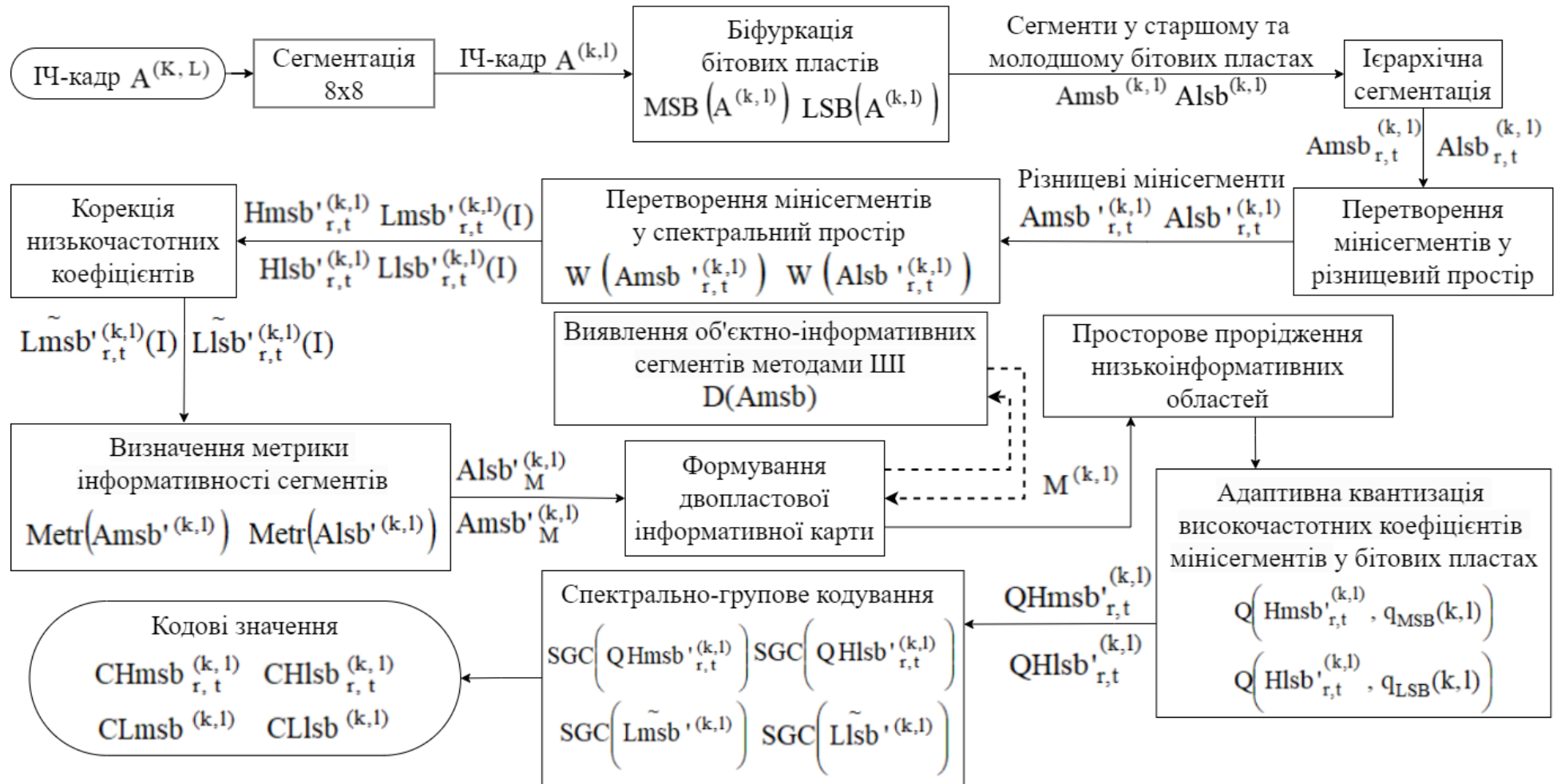


Рисунок 3.12 – Схема методу кодування ІЧ-зображень на основі семантико-орієнтованої обробки сегментів

В результаті описаних структурно-статистичних перетворень розроблений метод дозволяє представити вхідний ПЧ-кадр у вигляді компактних кодових значень. Описані семантико-орієнтовані підходи виконуються як на основі метричних показників інформативності, так і на опційному застосуванні ШІ. Що дозволяє селективно обробляти області ПЧ-кадру із різним рівнем інформаційної насиченості. Для подальшого розвитку розробленого методу кодування виявлені семантичні неоднорідності можуть бути використані для побудови моделі захищеного обміну даними, шляхом адаптивного засхоствування криптографічних методів.

Висновки за третім розділом.

1. Розроблено метод різницевого кодування мінісегментів інфрачервоного зображення. Метод базується на відніманні локального мінімуму всередині кожного мінісегмента. Це формує різницевий мінісегмент, що предсталає матрицю варіацій. Такий підхід дозволяє звужити динамічний діапазон даних та зберегти температурні градієнти.

2. Розроблено метод корекції низькочастотних коефіцієнтів спектрального простору. Корекція здійснюється відніманням кінцевого абсолютного високочастотного значення від кінцевого низькочастотного. Відповідна операція додатково звужує динамічний діапазон спектрального представлення. Водночас повна оборотність перетворення зберігає цілісність даних при декодуванні.

3. **Наукова новизна.** Вперше розроблено метод спектрально-групового стиснення на основі структурних перетворень у хвильовому просторі Хаара. Відмінності методу полягають у: диференційованому структурному кодуванні хвильового простору сегментів за частотними областями в умовах квадрореконструкції та встановленням нижньої межі високочастотної області. Це дозволяє зменшити бітовий об'єм пластових сегментів ПЧ-зображень без втрат термальної цілісності.

4. **Наукова новизна.** Отримано подальший розвиток методу кодування ІЧ-зображень на основі семантико-орієнтовної обробки. Відмінності методу полягають у спектрально-груповому кодуванні з врахуванням наявності об'єктно-інформативних сегментів ІЧ-зображень в біфуркаційно-хвильовому просторі Хаара. Це дозволяє знизити бітрейт інфрачервоного потоку даних з визначеним рівнем термальної цілісності в процесі обробки та передачі в інформаційних системах.

5. Розроблено метод адаптивної квантизації високочастотних коефіцієнтів. Адаптивний підхід враховує рівень інформативності сегмента та його відношення до бітового пласту. Об'єктно-інформативні сегменти обробляються мінімальними коефіцієнтами квантизації. Це дозволяє зберігати теплові сигнатури об'єктів інтересу. Водночас малоінформативні фонові області квантуються з вищими коефіцієнтами. Такий розподіл ресурсів дозволяє зменшити бітрейт ІЧ-потoku із визначеним рівнем термальної цілісності

Результати досліджень щодо матеріалів, викладених в третьому розділі, опубліковані в таких наукових працях як: [30, 98, 102].

РОЗДІЛ 4. ЕКСПЕРИМЕНТАЛЬНА ОЦІНКА ХАРАКТЕРИСТИК ЗАПРОПОНОВАНОГО МЕТОДУ КОДУВАННЯ

Проводиться експериментальна оцінка характеристик запропонованого методу кодування.

Розроблено метод просторового прорідження малоінформативних областей та їх бігармонічної реконструкції.

Розроблено модель оцінки ефективності методу кодування інфрачервоних зображень.

Проведено абляційний аналізу впливу структурно-спектральних етапів кодування на бітову інтенсивність та якість відновлення інфрачервоних кадрів.

Проведено порівняльну оцінку бітрейту та характеристик стиснення ІЧ-зображень запропонованого методу кодування з відомими підходами.

4.1 Розробка методу просторового прорідження малоінформативних областей та їх бігармонічної реконструкції

Після класифікації сегментів кадру на високо- та низькоінформативні постає необхідність перейти від локального аналізу окремих блоків до розгляду ширших просторових структур. Хоча інформативність кожного сегмента вже визначена, і враховується при його кодуванні, необхідно враховувати масштабніші просторові закономірності, які характерні для інфрачервоних зображень. Зазвичай низькоінформативні сегменти утворюють протяжні однорідні області, що відповідають фоновим ділянкам із низькою динамікою температури. Саме ця структурна впорядкованість відкриває можливість переходу від обробки сегментів до аналізу надмірності в межах цілісних областей ІЧ-сцени.

Ефективність виявлення низькоінформативних областей кадру можуть підтвердити показники ентропії. Для цього розглянемо приклад для ІЧ-кадру на рисунку 4.1 а). На карті метрики інформативності (рис 4.1 б) чітко видно відповідність областей із підвищеними значенням 0.8-1.0 метрики до позиціонування об'єкту інтересу. Ентропійна карта (рис 4.1 в) демонструє локальну інформаційну насиченість. Ця карта формується шляхом ковшного аналізу вікон 8×8 , у межах яких обчислюється гістограма сигналу та відповідна ентропія Шеннона [123]. Низькоінформативні регіони виявляються ділянками з низькою ентропією, де локальний розподіл інтенсивностей звужений, варіативність у межах сегментного просторового вікна є низькою. У таких зонах нормалізоване значення ентропії залишається у діапазоні 0.5-0.0, що відображає низький рівень невизначеності у локальному фрагменті зображення. Області з насиченою тепловою структурою створюють розширений розподіл інтенсивностей і характеризуються підвищеною ентропією, що означає збільшений обсяг локальної інформації, яку необхідно зберегти під час кодування [64].

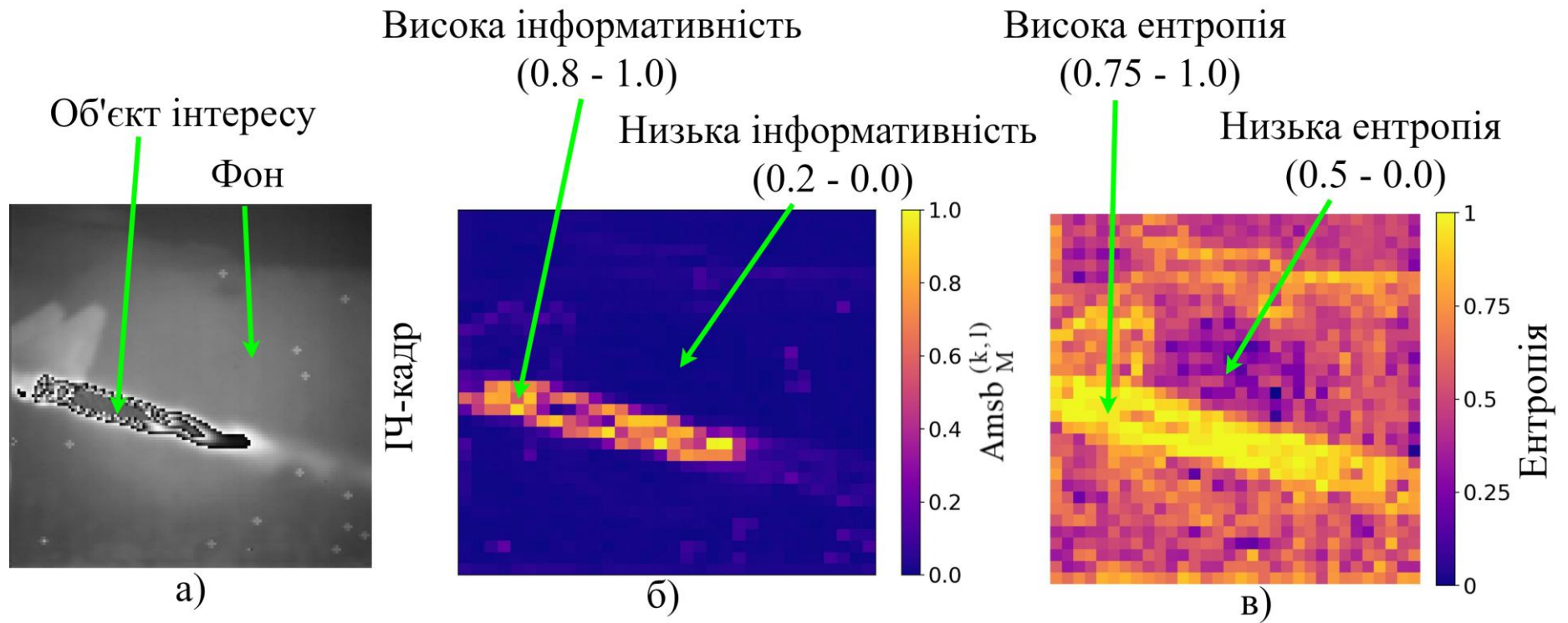


Рисунок 4.1 – Вплив інформаційної насиченості на метрику інформативності та ентропію: а) ІЧ-кадр із об'єктом інтересу
б) карта метрики інформативності ІЧ-кадру в) карта ентропії ІЧ-кадру

Такий підхід дозволяє оцінити інформаційну насиченість локального регіону. Низька ентропія означає, що значення інтенсивностей у локальній області маю високу ступінь подібності. Области з насиченою структурою об'єкта демонструють високі значення ентропії, що свідчить про складні локальні зміни, які не можна ефективно апроксимувати на основі навколишніх опорних сегментів.

Отже, ентропійна просторова закономірність створює сприятливі умови для відкидання частини фонових сегментів. Низька ентропія є прямим математичним маркером надмірності. У таких областях сусідні блоки демонструють високий ступінь подібності, а відмінності між ними визначаються переважно шумом сенсора або малозначущими флуктуаціями. Така однорідна структура областей створює передумови для їхнього прорідження, оскільки повне кодування кожного сегмента лише дублює інформацію, вже наявну в сусідніх блоках. Отже, частина сегментів у межах низькоінформативних областей може бути відкинута, а обсяг переданих даних – істотно зменшений без ризику втрати інформативності об'єктів інтересу.

Для забезпечення можливості відновлення пропущених сегментів у низькоінформативних областях необхідно реалізувати механізм опорних сегментів. Ці сегменти не підлягають відкиданню та зберігають ключові параметри області: просторову структуру, характер температурних градієнтів, середній рівень сигналу. Вони виконують роль локальних репрезентативних точок, на основі яких під час декодування може бути реконструйована решта сегментів, вилучених на етапі кодування. Їхній вибір має бути достатнім, щоб забезпечити збереження внутрішньої організації області, але водночас достатньо економним, щоб підсилити результат прорідження.

Запропонованим способом упорядкування таких опорних сегментів є використання регулярної просторової маски $p(k,l)$, що забезпечує рівномірний розподіл опорних точок по всій низькоінформативній області.

Підхід, структурно подібний до сітки «шахової дошки», дозволяє визначати опорні сегменти із високою щільністю та стабільністю реконструкції. Регулярність такої структури забезпечує оптимальне просторове покриття: кожен пропущений сегмент у межах області має сусідні опорні блоки, що зменшує похибку під час відновлення значень. Крім того, фіксований характер розташування опорних сегментів усуває потребу в додатковій передачі карти їхніх позицій, що обмежує збільшення бітового об'єму службової інформації [133, 138].

Математично метод прорідження $\text{thin}(A^{(k,l)})$ можна описати наступним чином:

$$p(k,l) = \begin{cases} 0, & \text{якщо } (k+1) \text{ є парним} \\ 1, & \text{якщо } (k+1) \text{ є непарним} \end{cases};$$

$$\tilde{A}^{(k,l)} = \text{thin}(A^{(k,l)}) = \begin{cases} A^{(k,l)}, & \text{якщо } M^{(k,l)} = [1, 2] \text{ або } p(k,l) = 1. \\ 0 & \end{cases} \quad (4.1)$$

Де:

- $\tilde{A}^{(k,l)}$ – одержаний сегмент після прорідження, що може дорівнювати початковому сегменту $A^{(k,l)}$ або бути повністю обнуленим;
- $p(k,l)$ – бінарне значення з регулярної просторової маски, що визначає чи є неінформативний сегмент $A^{(k,l)}$ опорним.

На рисунку 4.2 а) зображена карта інформативності ГЧ-кадру. Де білим позначені високоінформативні сегменти, а чорним низькоінформативні. На основі цієї карти інформативності формується матриця прорідження ГЧ-кадру, відповідний приклад на рисунку 4.2 б). Синім позначені сегменти, що мають бути прорідженими, тобто їх значення прирівнюються до 0. На рисунку 4.2 в) зображений результуючий ГЧ-кадр, до якого була застосована операція прорідження. Пікселі у чорних сегментах мають значення 0.

температури мають семантичне значення та інтерпретуються як теплові аномалії або сигнатури реальних об'єктів.

Бігармонічна реконструкція ґрунтується на розв'язанні диференційного рівняння:

$$\Delta^2 \tilde{A}_{r,t}^{(k,l)}(q,s) = 0. \quad (4.2)$$

У цьому виразі символ Δ^2 позначає бігармонічний оператор. Зазначений оператор утворюється внаслідок подвійного застосування оператора Лапласа Δ [68]. Він дорівнює сумі четвертих змішаних похідних за просторовими параметрами функції. У нашому випадку диференціювання має відбуватись за параметрами (q,s) , що вказують на координати пікселів у прорідженому сегменті $\tilde{A}_{r,t}^{(k,l)}$, який необхідно реконструювати. Шляхом алгебраїчного розкриття квадратів суми це рівняння подається через частинні похідні:

$$\frac{\partial^4 \tilde{A}_{r,t}^{(k,l)}(q,s)}{\partial q^4} + 2 \frac{\partial^4 \tilde{A}_{r,t}^{(k,l)}(q,s)}{\partial q^2 \partial s^2} + \frac{\partial^4 \tilde{A}_{r,t}^{(k,l)}(q,s)}{\partial s^4} = 0. \quad (4.3)$$

У дискретній формі бігармонічне рівняння вирішується наступним шляхом. Ядро згортки перетворення Лапласа, яке було двічі застосоване до себе (бо бігармонічне перетворення є двічі застосованим оператором Лапласа) утворює бігармонічне ядро згортки ВС:

$$BC = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 2 & -8 & 2 & 0 \\ 1 & -8 & 20 & -8 & 1 \\ 0 & 2 & -8 & 2 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix}. \quad (4.4)$$

Значення ваг у ядрі ВС позиційно відповідають пікселям, які застосовуються для відновлення центрального елемента. Пікселі, що відносяться до більших за модулем ваг, мають більший вплив. Оскільки невідомі пікселі потрібно відновити на основі опорних відомих, то ядро ВС необхідно застосувати для побудови рівнянь. Згідно з математичною моделлю, зважена сума пікселів у прорідженому сегменті $\tilde{A}^{(k,l)}$ має дорівнювати нулю. Це зумовлює необхідність виконання операції дискретної згортки:

$$\tilde{A}^{(k,l)} * BC = 0. \quad (4.5)$$

Зазначена операція формує систему рівнянь, де кожен окремий вираз базується на невідомому значенні цільового пікселя, обчислюючи його через сусідні елементи. Оскільки дванадцять елементів бігармонічного ядра ВС дорівнюють нулю, значення відповідних пікселів не впливають на загальну суму та вилучаються з обчислень. Решта тринадцять ненульових просторових позицій генерують члени лінійного алгебраїчного рівняння.

У залежності від віддалення пікселів від центру застосовуються відповідні вагові коефіцієнти. До основних груп пікселів належать:

- x_1 – позначає центральний невідомий піксель, що підлягає реконструкції. Позиційно відповідає ваговому коефіцієнту 20;
- $x_2 \dots x_5$ – відповідають чотирьом найближчим сусідам по вертикалі та горизонталі. Множаться на вагові коефіцієнти -8;
- $x_6 \dots x_9$ – ідентифікують чотири діагональні пікселі локального околу. Множаться на вагові коефіцієнти 2;
- $x_{10} \dots x_{13}$ – позначають найбільш віддалені пікселі на зовнішніх межах ядра згортки. Відповідають ваговим коефіцієнтам 1.

Застосування такої індексації трансформує двовимірну операцію згортки у систему лінійних рівнянь. Де кожне рівняння у системі є обчисленням кожного невідомого елементу у прорідженій області:

$$\begin{cases} \dots \\ 20x_1 - 8x_2 - 8x_3 - 8x_4 - 8x_5 + 2x_6 + 2x_7 + 2x_8 + 2x_9 + x_{10} + x_{11} + x_{12} + x_{13} = 0 \\ \dots \end{cases} \quad (4.6)$$

Розв'язком описаної системи рівнянь, є визначення невідомих пікселів. У процесі декодування невідомими пікселями вважаються ті, що дорівнюють нулю. З них складається карта проріджених сегментів $\tilde{A}^{(k,l)}$ і застосовується для виконання реконструкції.

На рисунку 4.3 відображена схема застосування просторового прорідження та бігармонічної реконструкції в процесі кодування та декодування ІЧ-кадру. Основні етапи прорідження та реконструкції виділено червоним кольором. На схемі блок «Структурно-статистична обробка кадру» є узагальненням етапів: сегментації, різницевого кодування, вейвлет-перетворення, низькочастотної корекції. Блок «Формування спектрально-групових кодів» включає в себе адаптивну квантизацію та спектрально-групове кодування. При декодуванні блок «Зворотня спектрально-структурна обробка» відображає усі зворотні перетворення закодованого кадру, для відновлення його просторової структури. Із відновленої структури можна виявити проріджені сегменти та застосувати до них бігармонічну реконструкцію, як фінальний етап декодування [155].



Рисунок 4.3 – Схема застосування просторового прорідження та бігармонічної реконструкції в процесі кодування та декодування ІЧ-кадру

Наведені ілюстрації (рис. 4.4) демонструють, як бігармонічне перетворення відновлює сегменти різного рівня інформативності у фрагменті ІЧ-кадру. На вхідному фрагменті (рис. 4.4 а) виділено червоним кольором інформативний сегмент, а зеленим – малоінформативний.

У реконструйованому фрагменті (рис. 4.4 б) помітно, що малоінформативний сегмент має високу подібність. Тоді як інформативний сегмент надмірно згadzений і втратив високочастотні термальні структури. Це підтверджується показниками RMSD:

- RMSD = 3 440 – для неінформативного сегменту;
- RMSD = 16 517 – для інформативного сегменту.

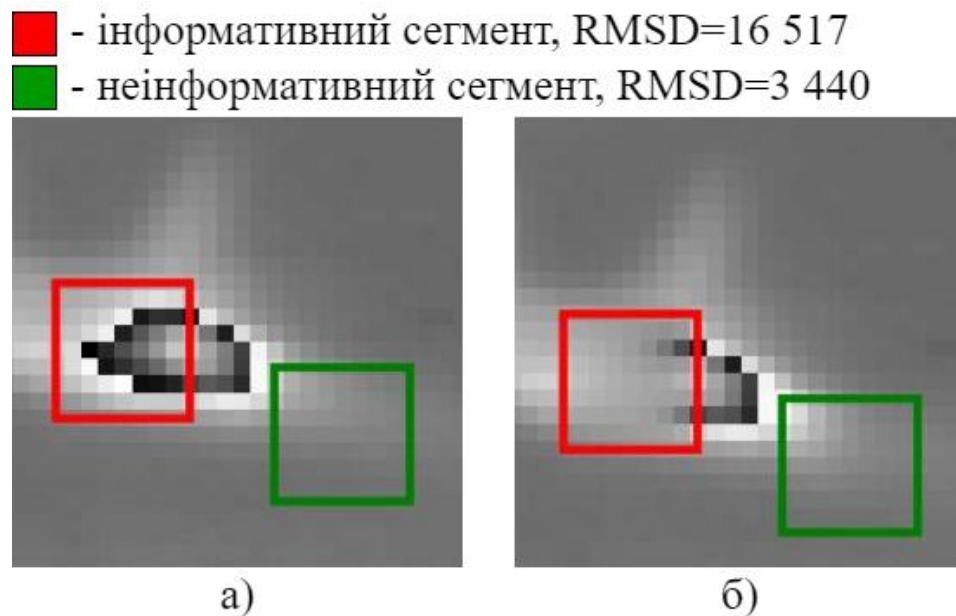


Рисунок 4.4 – Приклад бігармонічної реконструкції сегментів ІЧ-кадру: а) вхідне зображення б) реконструйоване зображення

Бігармонічну реконструкцію наведених сегментів зображено у 3D репрезентації (рис 4.5). Можна помітити згadzення локальних піків реконструйованого інформативного сегменту (рис 4.5 в). Водночас, рівномірний градієнт (рис 4.5 г) неінформативного сегменту (рис 4.5 д) відновлено із високою подібністю (рис 4.5 д).

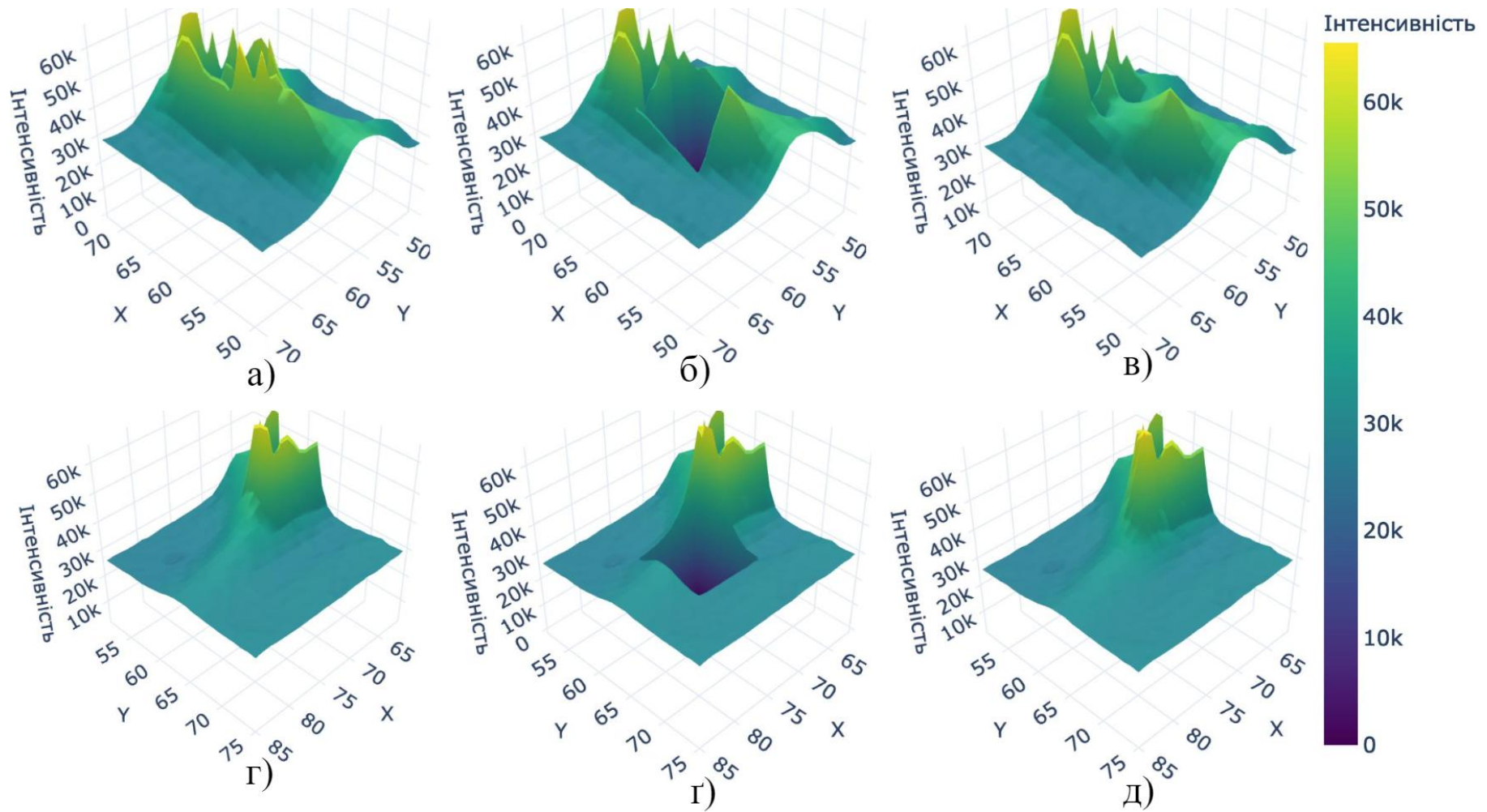


Рисунок 4.5 – 3D репрезентація бігармонічної реконструкції сегментів ІЧ-кадру: а) інформативний сегмент б) вилучений інформативний сегмент в) реконструйований інформативний сегмент г) неінформативний сегмент г) вилучений неінформативний сегмент д) реконструйований неінформативний сегмент

Завдяки застосуванню бігармонічної реконструкції вдається досягти не тільки зменшення бітового об'єму, а і локального покращення якості кадру. При кодуванні, у кадр вносяться спотворення на етапі адаптивної квантизації. Для таких спотворень характерна сітчаста структура квантизованих областей. Спотворення, що перцептивно схожі на сітку, насправді є різкими переходами інтенсивності між сегментами [86].

Властивість бігармонічного перетворення зменшувати кривину проріджених сегментів виступає у ролі фільтру згляджування. Такий фільтр обмежує високочастотні артефакти, що утворені після квантизації.

На рисунку 4.6 зображене порівняння внесених спотворень із застосуванням бігармонічної реконструкції та без. До вхідного ПЧ-кадру (рис 4.6 а) було застосовано адаптивну квантизацію. Внаслідок чого на декодованому кадрі (рис 4.6 б) спостерігаються внесені спотворення, а $RMSD = 423$. На рисунку 4.6 г) зображена посегментна карта $RMSD$ для цієї обробки.

У наступному експерименті вхідний ПЧ-кадр було закодовано із застосуванням прорідження та реконструкції. Декодований кадр зображено на рисунку 4.6 в), $RMSD = 435$. Однак його карта $RMSD$ на рисунку 4.6 г) відображає затухання $RMSD$ в певних областях. Отже, ступінь втрат у них зменшилась. Хоча помітно, що у обмеженій кількості сегментів відбулись значні викиди $RMSD$.

Для кращого розуміння впливу реконструкції на цілісність кадру, на рисунку 4.6 д) зображена карта впливу реконструкції прорідження на $RMSD$. При її побудові, для кожного сегмента була обрахована різниця $RMSD$ при кодуванні із реконструкцією та без. Значення в червоних сегментах означають збільшення втрат, коли зелені сегменти відображають зменшення.

Статистично серед сегментів, що зазнали спотворень, 25% з них було покращено, тобто $RMSD$ зменшився після реконструкції. Коли 18% сегментів зазнали більших втрат. Для більшості сегментів 56% $RMSD$ не змінився.

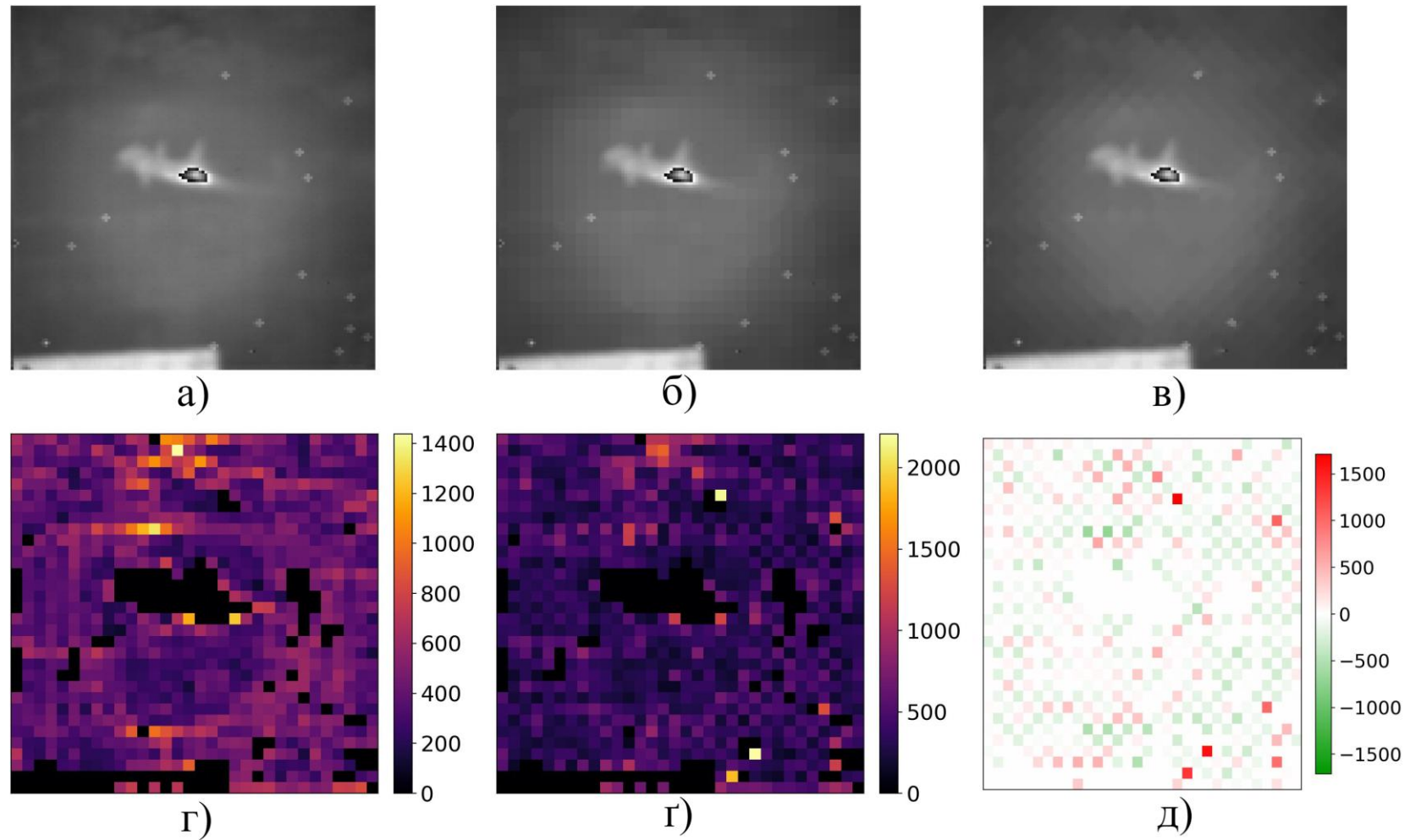


Рисунок 4.6 – Візуалізація спотворень за наявності та відсутності прорідження ПЧ-кадру: а) вхідний кадр; б) декодований кадр без прорідження; в) декодований кадр із прорідженням; г) карта RMSD без прорідження; г) карта RMSD із прорідженням; д) карта впливу реконструкції прорідження на RMSD

Серед покращених сегментів, у середньому RMSD зменшився на 164. Однак, для погіршених сегментів RMSD у середньому збільшився на 193. Це пов'язано із тим, що низькоінформативні проріджені сегменти мали сусідні високоінформативні. Тому при реконструкції інтенсивність прорідженого сегменту було наближено до інформативних. Що утворило локальні викиди втрат.

Отже, була обґрунтована доцільність застосування просторового прорідження та бігармонічної реконструкції винятково до низькоінформативних областей ІЧ-зображення. Це обмеження зумовлене тим, що реконструкція високоінформативних сегментів призводить до надмірного згладжування та втрати високочастотних термальних структур.

Також виявлено умови локального покращення якості кадру за рахунок використання етапу реконструкції. Здатність бігармонічного перетворення зменшувати кривину відтворених сегментів виконує функцію фільтра згладжування. Такий фільтр обмежує високочастотні артефакти, утворені після адаптивної квантизації.

4.2 Розробка моделі оцінки ефективності методу кодування інфрачервоних зображень

Метод оцінки ефективності запропонованого підходу до кодування інфрачервоних зображень повинен включати використання кількісних показників бітової ефективності та якості відновлення, із врахуванням семантичної значущості різних областей кадру. Такий підхід дає змогу не лише порівняти запропонований метод з відомими рішеннями, але й коректно інтерпретувати отримані результати з позицій практичного застосування в інформаційних системах [72].

Важливим етапом є оцінка бітового об'єму та коефіцієнту стиснення. Нехай початкове інфрачервоне зображення має просторову розмірність $K \times L$

та бітову глибину D на піксель. Тоді сирий бітовий об'єм такого кадру дорівнює:

$$V_{\text{raw}} = K \cdot L \cdot D. \quad (4.7)$$

Після кодування кадру утворюється бітовий потік довжиною V_{enc} біт. Тоді коефіцієнт стиснення [122] визначається як:

$$C_{\text{cmp}} = \frac{V_{\text{raw}}}{V_{\text{enc}}}. \quad (4.8)$$

Для аналізу бітового об'єму, утвореного розробленим методом кодування, пропонується враховувати внесок окремих кодових компонент: низькочастотних та високочастотних спектрально-групових кодів. Сумарний бітовий об'єм кадру у цьому випадку подається у вигляді:

$$V_{\text{enc}}^{(\text{IR})} = \sum_{k,l} \left(\text{BC} \left(\text{CLmsb}^{(k,l)} \right) + \text{BC} \left(\text{CLlsb}^{(k,l)} \right) + \sum_{r,t} \left(\text{BC} \left(\text{CHmsb}_{r,t}^{(k,l)} \right) + \text{BC} \left(\text{CHlsb}_{r,t}^{(k,l)} \right) \right) \right) + \text{ServInf};$$

$$\text{BC}(C) = \log_2 C. \quad (4.9)$$

Де:

- $\text{CLmsb}^{(k,l)}$, $\text{CLlsb}^{(k,l)}$ - значення низькочастотних спектрально-групових кодів з площин старших біт MSB та молодших LSB;

- $\text{CHmsb}_{r,t}^{(k,l)}$, $\text{CHlsb}_{r,t}^{(k,l)}$ - значення високочастотних спектрально-групових кодів з площин старших біт MSB та молодших LSB;

- BC – функція визначення кількості біт, необхідних для представлення кодового значення C ;

- ServInf – службова інформація, що необхідна для декодування.

До службової інформації ServInf відноситься: мінімальні значення $\mu_{r,t}^{(k,l)}$ мінісегментів, максимальні значення γ та λ частотних груп, знакові компоненти та дробові частки частотних коефіцієнтів, рівень інформативності сегментів. ServInf кодується без втрат [97].

Наступним елементом оцінки ефективності є оцінка втрат якості реконструйованих кадрів. Для кожного кадру початкове зображення $A^{(K,L)}$ порівнюється з відновленим $\tilde{A}^{(K,L)}$. Базовим показником відхилення є середньоквадратична похибка RMSD (Root Mean Square Deviation) [115]:

$$\text{RMSD} = \sqrt{\frac{1}{K \cdot L} \cdot \sum_{x=1}^K \sum_{y=1}^L (A(x,y) - \tilde{A}(x,y))^2}. \quad (4.10)$$

Де $A(x,y)$ та $\tilde{A}(x,y)$ - значення пікселів з відповідних зображень із координатами (x, y) .

На основі RMSD обчислюється пікове відношення сигнал/шум [91]:

$$\text{PSNR} = 20 \cdot \log_{10} \left(\frac{\text{MAX}}{\text{RMSD}} \right);$$

$$\text{MAX} = 2^D - 1. \quad (4.11)$$

Де MAX – максимально можливе значення пікселя із бітовою глибиною D.

PSNR та RMSD дозволяють оцінити статистичні відхилення даних. Для оцінки структурних спотворень, що впливають на перцептивні якості використовується індекс структурної подібності SSIM [10], який можна описати як:

$$\text{SSIM} = \ell(A, \tilde{A}) \cdot c(A, \tilde{A}) \cdot s(A, \tilde{A}). \quad (4.12)$$

Де:

- $\ell(A, \tilde{A})$ - відмінність яскравості між A та \tilde{A} ;
- $c(A, \tilde{A})$ - відмінність контрасту між A та \tilde{A} ;
- $s(A, \tilde{A})$ - відмінність структури між A та \tilde{A} .

Оскільки запропонований метод кодування є семантично орієнтованим, оцінка якості повинна відображати оцінку втрат у семантично важливих областях зображення, окрім загальних показників по всьому кадру. Для цього, з піксельних значень $A_{r,t}^{(k,l)}(q,s)$ вхідного зображення $A^{(K,L)}$, які позиційно відносяться до високоінформативних сегментів $M^{(k,l)}=[1,2]$, сформуємо групу високоінформативних значень A_{\uparrow} . Позначимо це наступним виразом:

$$A_{\uparrow} = \left\{ A_{r,t}^{(k,l)}(q,s) \in A^{(K,L)} \mid M^{(k,l)}=[1,2] \right\}. \quad (4.13)$$

Подібним чином сформуємо групу низькоінформативних значень A_{\downarrow} :

$$A_{\downarrow} = \left\{ A_{r,t}^{(k,l)}(q,s) \in A^{(K,L)} \mid M^{(k,l)}=0 \right\}. \quad (4.14)$$

За наявності закодованого зображення $\tilde{A}^{(K,L)}$ можна сформувати для нього групи інформативних значень \tilde{A}_{\uparrow} та неінформативних \tilde{A}_{\downarrow} , аналогічним способом. Маючи $(A_{\uparrow}, A_{\downarrow})$ та $(\tilde{A}_{\uparrow}, \tilde{A}_{\downarrow})$ розрахуємо середньоквадратичну похибку для інформативних значень RMSD_{\uparrow} та неінформативних RMSD_{\downarrow} :

$$\text{RMSD}_{\uparrow} = \sqrt{\frac{1}{n(A_{\uparrow})} \cdot \sum_{(x,y) \in A_{\uparrow}} (A_{\uparrow}(x,y) - \tilde{A}_{\uparrow}(x,y))^2}; \quad (4.15)$$

$$\text{RMSD}_{\downarrow} = \sqrt{\frac{1}{n(A_{\downarrow})} \cdot \sum_{(x,y) \in A_{\downarrow}} (A_{\downarrow}(x,y) - \tilde{A}_{\downarrow}(x,y))^2}. \quad (4.16)$$

Де:

- $n(A_{\uparrow})$ - кількість значень у високоінформативній групі;
- $n(A_{\downarrow})$ - кількість значень у низькоінформативній групі.

Маючи RMSD_{\uparrow} та RMSD_{\downarrow} , закономірним є розрахунок PSNR_{\uparrow} та PSNR_{\downarrow} :

$$\text{PSNR}_{\uparrow} = 20 \cdot \log_{10} \left(\frac{\text{MAX}}{\text{RMSD}_{\uparrow}} \right); \quad (4.17)$$

$$\text{PSNR}_{\downarrow} = 20 \cdot \log_{10} \left(\frac{\text{MAX}}{\text{RMSD}_{\downarrow}} \right). \quad (4.18)$$

Для проведення експериментальної оцінки запропонованого методу кодування ІЧ зображень була використана підмножина набору даних Open Turbulent Image Set (OTIS). OTIS містить монохромні зображення з різними розмірами (від 112×112 до 520×520 пікселів), що характеризуються різноманітною структурною складністю. Незважаючи на те, що OTIS не містить ІЧ кадрів, його особливістю є турбулентні спотворення [34]. Вони візуально подібні до термальних флуктуацій ІЧ кадрів. На рисунку 4.7 наведено приклад зображення з набору OTIS і різною інтенсивністю турбулентних спотворень. Крім характерних спотворень, зображення з набору OTIS мають семантичні структури, які зручні для аналізу стабільності методу кодування.

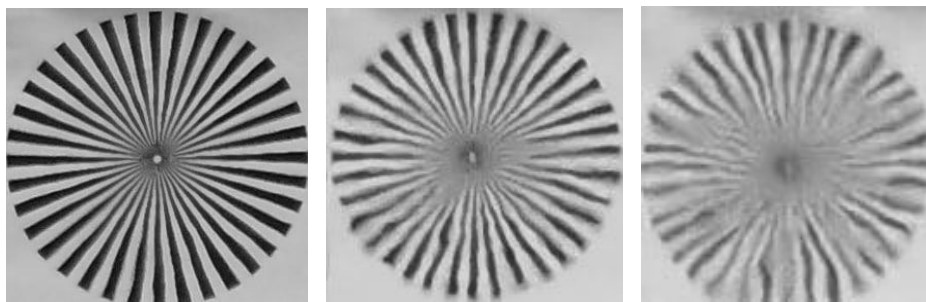


Рисунок 4.7 – Приклад зображення з набору OTIS і різною інтенсивністю турбулентних спотворень

Для оцінки ефективності обробки ІЧ кадрів був використаний набір з понад 100 інфрачервоних кадрів, що представляють 16-бітні зображення з роздільною здатністю від 128x128 до 512x640 пікселів. До набору включені сцени з однорідним фоном, середньою структурною насиченістю та складними багатокомпонентними тепловими структурами. Метричні та шумові характеристики кадрів відповідають сучасним неохолоджуваним тепловізійним сенсорам. Тип зафіксованого ІЧ випромінювання є порівнянним із продукцією промислових камер класу FLIR Tau 2, FLIR A655sc, Seek Thermal CompactPRO тощо, що формують 14–16-бітні кадри з подібними діапазонами і контрастністю. Це забезпечує релевантність отриманих оцінок реальним умовам експлуатації бортових систем.

Реалізація запропонованого методу кодування виконана мовою Kotlin, із застосуванням фреймворку Kotlin Multiplatform, що дає змогу виконувати один і той самий код у середовищах Linux, Windows, macOS, на JVM-платформах, а також в операційних системах Android та iOS. Операції над обробкою матриць реалізовані із застосуванням числової бібліотеки Multik. Для проведення експериментів та побудови графіків, використовується мова Python з бібліотеками NumPy та Matplotlib [47].

4.3 Проведення абляційного аналізу впливу структурно-спектральних етапів кодування на бітову інтенсивність та якість відновлення інфрачервоних кадрів

Для кількісної оцінки внеску окремих етапів обробки проведено абляційний аналіз розробленого методу кодування. У межах цього аналізу послідовно досліджується вплив структурно-спектральних етапів на бітову інтенсивність та якість відновлення інфрачервоних кадрів.

Для кожної конфігурації кодування оцінювалися:

- коефіцієнт стиснення C_{cmp} ;
- середньоквадратичне відхилення RMSD, показники PSNR та SSIM для всього кадру;
- спеціальні метрики похибки обчислені лише в межах неінформативних сегментів: RMSD_{\downarrow} , PSNR_{\downarrow} .

Для абляційного аналізу метод сконфігуровано таким чином, щоб до інформативних сегментів не вносились спотворення. Отже, буде явно відображена різниця між спеціальними метриками похибки та загальними. Тому з причини відсутності втрат інформативних областей, метрики RMSD_{\uparrow} та PSNR_{\uparrow} в даному аналізі не застосовуються.

Аналіз проводився для послідовності конфігурацій кодування, у яких поетапно вмикаються запропоновані модулі. За базову конфігурацію C_0 прийнято варіант, у якому використовується ієрархічна сегментація та вейвлет-перетворення Хаара з подальшим формуванням спектрально-групових кодів без будь-яких додаткових перетворень динамічного діапазону чи семантично орієнтованих операцій. Наступні конфігурації отримуються шляхом послідовного додавання таких етапів:

- C_1 – бітова біфуркація (розділення 16-бітного сигналу на пласти старших та молодших бітів MSB/LSB);

- C_2 – перехід у різницевий простір на рівні мінісегментів;
- C_3 – корекція низькочастотних значень з урахуванням залежності між кінцевими низькими та високими частотами;
- C_4 – семантично орієнтована адаптивна квантизація високочастотних коефіцієнтів на основі метрики інформативності;
- C_5 – просторове прорідження низькоінформативних областей з наступним бігармонічним відновленням.

Для кожної з цих конфігурацій були обчислені усереднені значення метрик оцінювання ефективності на наборі даних ІЧ зображень. Результати наведено у таблиці 4.1. Вона демонструє динаміку зміни показників при послідовному нарощуванні етапів кодування.

Таблиця 4.1 – Показники ефективності етапів перетворення $C_0..C_5$ розробленого методу кодування ІЧ зображень

Конфігурація	C_{cmp}	C_{cmp} % змін	PSNR	RMSD	SSIM	RMSD_{\downarrow}	PSNR_{\downarrow}
C_0	1.835	-	∞	0	1	0	∞
C_1	2.263	23.319	∞	0	1	0	∞
C_2	2.36	4.297	∞	0	1	0	∞
C_3	2.37	0.419	∞	0	1	0	∞
C_4	7.661	223.223	45.453	390.769	0.97	437.006	44.487
C_5	8.111	5.868	44.086	526.226	0.97	588.565	43.12

У базовій конфігурації C_0 досягається коефіцієнт стиснення $C_{\text{cmp}} \approx 1.84$ при повністю безвтратній реконструкції ($\text{RMSD} = 0$, $\text{PSNR} \rightarrow \infty$, $\text{SSIM} = 1$). Фактично усувається частина статистичної надмірності, але не враховується специфіка бітового представлення, чи семантичні властивості ІЧ-сцени. Дана конфігурація слугує референтною точкою для подальших порівнянь.

Застосування бітової біфуркації на етапі C_1 (розділення 16-бітних значень на окремі MSB- та LSB-пласти з подальшим незалежним кодуванням) забезпечує істотний приріст ефективності без введення втрат. Коефіцієнт стиснення зростає до $C_{\text{cmp}} \approx 2.27$, що відповідає збільшенню більш ніж на 23% відносно базового варіанту C_0 . Тоді як RMSD залишається нульовим. Окрім покращення показників кодування, цей етап важливий тим, що розділення бітових площин покращує структурну організованість даних для наступних етапів обробки.

У подальшій конфігурації C_2 додається перехід мінісегментів у різницевий простір. Обчислення значень відносно локального мінімуму зменшує динамічний діапазон в межах мінісегментів $A_{r,t}^{(k,l)}$. При цьому коефіцієнт стиснення підвищується до $C_{\text{cmp}} \approx 2.37$, що дає ще близько 4% приросту, зберігаючи повну безвтратність ($\text{RMSD} = 0$). Даний етап працює як локальне вирівнювання, яке зменшує вплив абсолютної температурної складової і залишає відхилення від локального фону, що кодуються компактніше.

Наступна конфігурація C_3 – корекція низькочастотних значень із використанням аналітичної залежності між кінцевими низькочастотними коефіцієнтами та відповідними високочастотними коефіцієнтами, що отримані при застосуванні вейвлету Хаара. У середньому коефіцієнт стиснення зростає до $C_{\text{cmp}} \approx 2.38$, що еквівалентно 0,4% додаткового виграшу порівняно з попереднім етапом C_2 . Хоча абсолютний приріст на цьому кроці незначний, він досягається без будь-яких спотворень сигналу і створює більш щільну структуру спектрально групових кодів.

Найсуттєвіший вплив на бітову інтенсивність має впровадження семантично орієнтованої адаптивної квантизації високочастотних коефіцієнтів. На цьому етапі C_4 значення коефіцієнту квантизації для

високочастотних компонентів визначається з урахуванням метрики інформативності сегмента: для інформативних сегментів використовується менший коефіцієнт, для низькоінформативних – більший. У результаті середній коефіцієнт стиснення різко зростає до $C_{\text{cmp}} \approx 7.66$, що означає більш ніж трикратне поліпшення відносно безвтратної конфігурації з корекцією низьких частот C_3 . При цьому глобальні показники якості становлять $\text{PSNR} \approx 45.453$ дБ та $\text{SSIM} \approx 0.97$, що відповідає високій візуальній якості реконструйованих кадрів [83].

На рисунку 4.8 зображений графік залежності коефіцієнту стиснення C_{cmp} від етапу обробки $C_0 \dots C_5$.

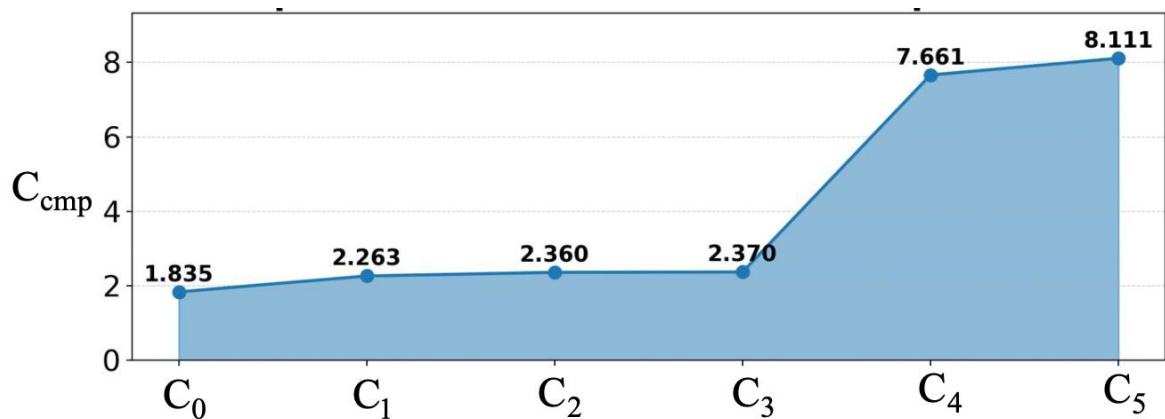


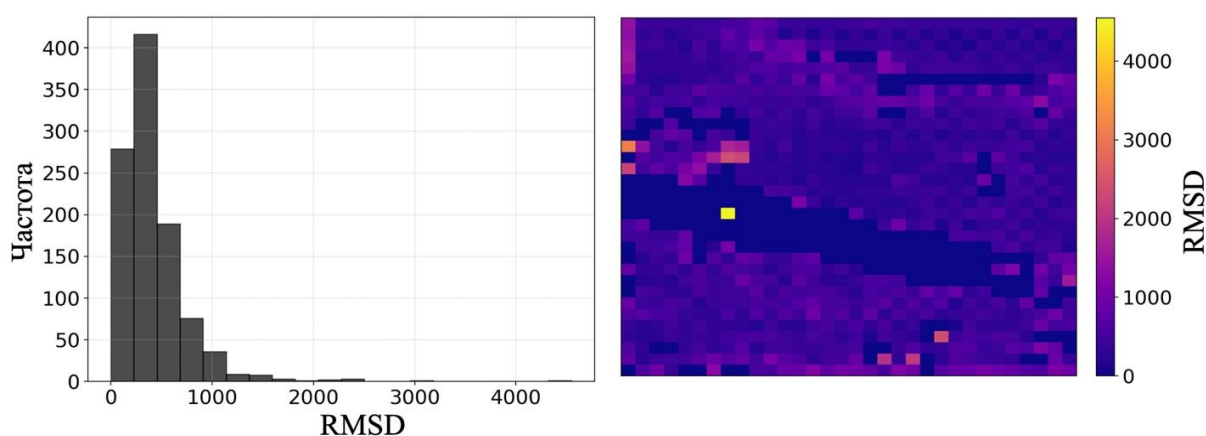
Рисунок 4.8 – Графік залежності коефіцієнту стиснення C_{cmp} від етапу обробки $C_0 \dots C_5$

Завершальною конфігурацією C_5 є просторове прорідження низькоінформативних областей із подальшою бігармонічною реконструкцією відкинутих сегментів. Дозволяє додатково знизити бітову інтенсивність за рахунок виключення з кодування частини фонового контенту, який може бути відновлений на основі опорних сегментів. У середньому коефіцієнт стиснення зростає до $C_{\text{cmp}} \approx 8.11$, що дає ще близько 6% виграшу відносно конфігурації C_4 з адаптивною квантизацією. При цьому глобальний PSNR

зменшується несуттєво – до $PSNR \approx 44.086$ дБ, а SSIM залишається на рівні 0.97, тобто спад якості візуального сприйняття є помірним.

Ключовим є той факт, що похибка в інформативних сегментах фактично відсутня. А значення втрат $RMSD_{\downarrow}$ у неінформативних областях, більші за середні по кадру. Це підтверджує, що адаптивна квантизація, керована картою інформативності $M^{(k,l)}$, не вносить спотворень у сегменти, де присутні об'єкти інтересу. Втрати концентруються в неінформативних, фонових областях. Таким чином, досягається компроміс між зменшенням бітового об'єму та збереженням цілісності семантично важливих ділянок ІЧ-зображення [76].

Цю особливість зображено на карті значень RMSD для декодованого ІЧ-зображення, рисунок 4.9 б). Карта побудована на основі значень RMSD для кожного сегменту, що розташовані відповідно до позиції на кадрі. А на рисунку 4.9 а) зображений їх кількісний розподіл на гістограмі. Він відображає те, що кількість суттєвих викидів RMSD (2000+) обмежена. Отже спотворення контрольовані.



а)

б)

Рисунок 4.9 – Розподілення втрат при кодуванні ІЧ-кадру розробленим методом: а) графік залежності сегментних значень RMSD від їх кількості; б) карта RMSD для кожного сегменту

Результати абляційного аналізу демонструють послідовний характер накопичення ефекту від запропонованих етапів. Структурні перетворення – бітова біфуркація, перехід у різницевий простір і корекція низькочастотних коефіцієнтів – формують компактне та статистично зручне представлення. Воно забезпечує безвтратне зниження бітової інтенсивності. Подальше введення семантично орієнтованої адаптивної квантизації та просторового прорідження низькоінформативних областей дає значний приріст у коефіцієнті стиснення. Спотворення концентруються у малозначущих зонах кадру.

4.4 Порівняльна оцінка бітрейту та характеристик стиснення ІЧ-зображень запропонованого методу кодування з відомими підходами

Проведемо порівняльну оцінку розробленого методу із існуючими методами обробки зображень. Для цього розділимо ІЧ-зображення на 3 категорії за їх рівнем інформативності:

- об'єктно-інформативні – зображення, що містять виявлений об'єкт інтересу (рис 4.10 а);
- інформативні – зображення, що насичені температурними контрастами але відсутній об'єкт інтересу (рис 4.10 б);
- неінформативні – зображення, де малопредставлені виражені термальні області. Основну частину кадру складають фонові ділянки (рис 4.10 в).

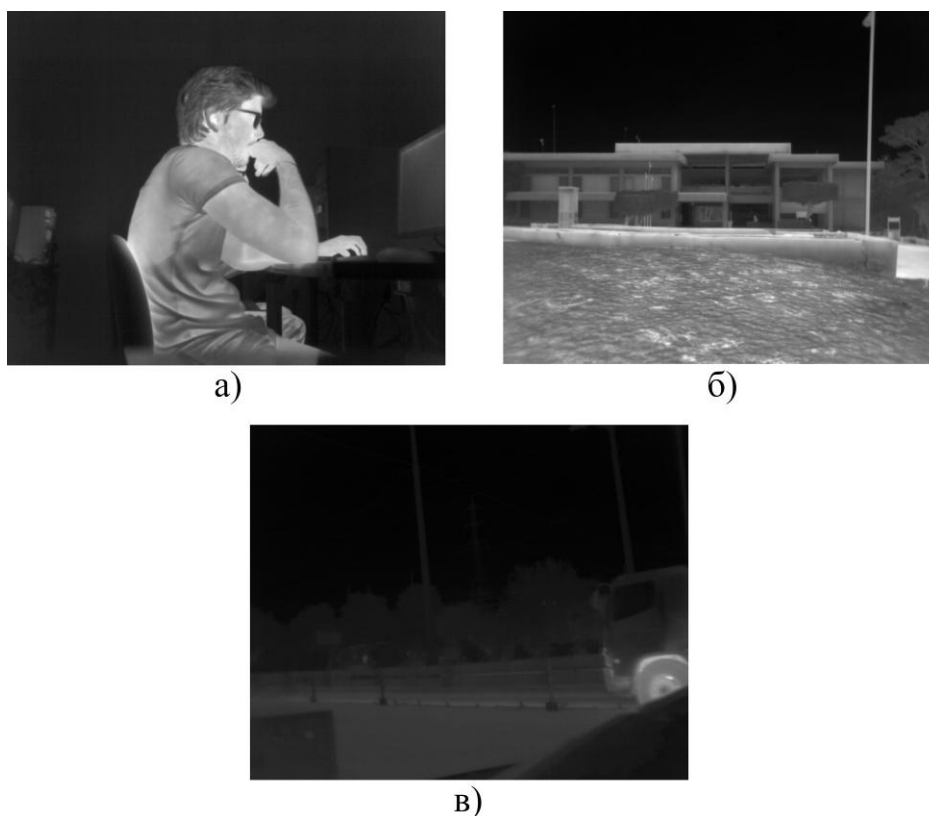


Рисунок 4.10 – Приклади ІЧ-кадрів з різним рівнем інформативності: а) об'єктно-інформативний кадр «Людина» б) інформативний кадр «Будівля» в) неінформативний кадр «Дорога»

У залежності від категорії інформативності кадру, можуть застосовуватись різні конфігурації втрат ІЧ-кадру.

Як було зазначено, при обробці ІЧ-кадрів розробленим методом формується карта інформативності сегментів. Її можна використати для аналізу областей інформативності кадру. На рисунку 4.11 а) відображена карта інформативності кадру «Дорога». З графіку 4.11 б) можна побачити, що 95.02% кадру складають малоінформативні сегменти. Тоді як лише 4.98% відноситься до інформативних сегментів. Таке співвідношення інформативності кадру дозволяє ефективно його стискати, при збереженні цілісності інформативних сегментів. За рахунку агресивного стиснення суттєво домінуючої неінформативної області.

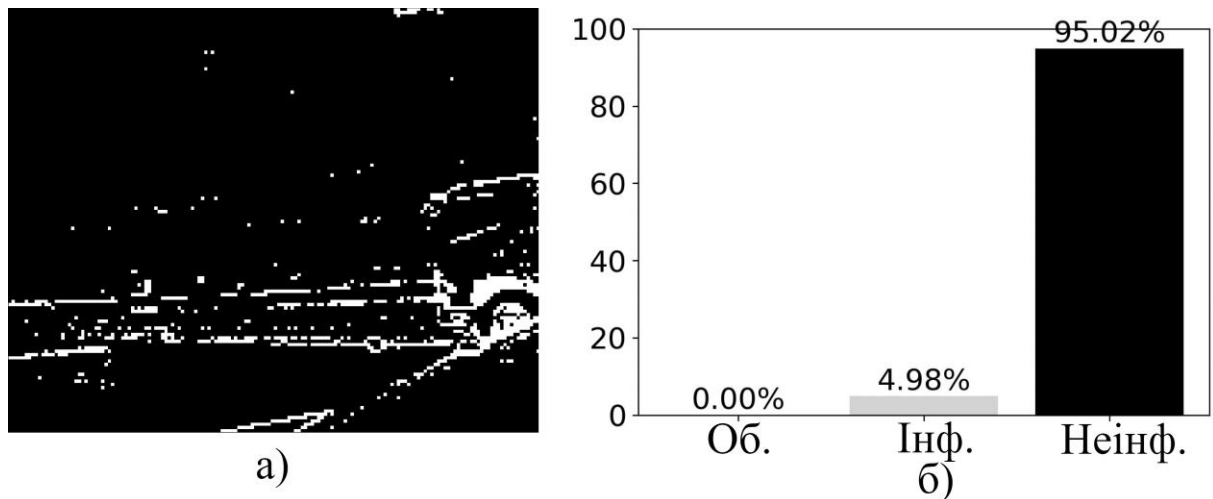


Рисунок 4.11 – Розподіл інформативності у ІЧ-кадрі «Дорога»: а) карта інформативності сегментів кадру б) графік залежності класу інформативності сегментів від відсоткового співвідношення у кадрі

Підхід до внесення втрат може варіюватись для інформативних кадрів. На рисунку 4.12 а) зображена карта інформативності ІЧ-кадру «Будівля». Через високу текстурність сцени, інформативні сегменти складають 20.86% кадру (рис 4.12 б). Для більшого стиснення такого кадру може допускатись внесення контрольованих спотворень у інформативні області. Крім цього, підвищення порогу інформативності зменшить кількість інформативних сегментів, лишивши найбільш пріоритетні.

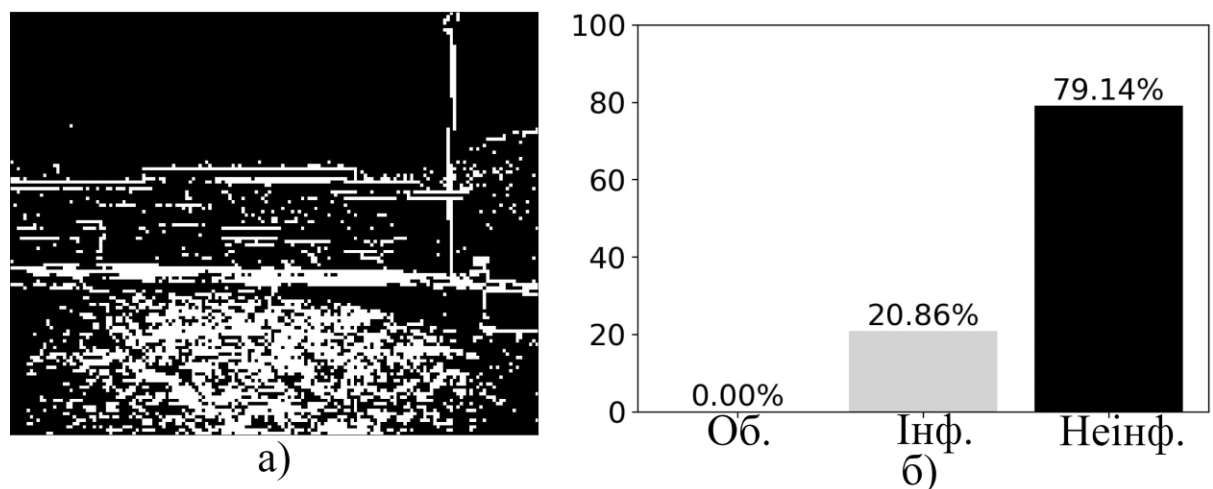


Рисунок 4.12 – Розподіл інформативності у ІЧ-кадрі «Будівля»: а) карта інформативності сегментів кадру б) графік залежності класу інформативності сегментів від відсоткового співвідношення у кадрі

На рисунку 4.13 а) зображено карту інформативності кадру «Людина». При кодуванні об'єктно-інформативних кадрів важливо зберегти інформативність сегментів, що відносять до об'єкту інтересу. На даному кадрі таких сегментів 16.7% (рис 4.13 б). Тому, для більшого стиснення може знадобитись внесення спотворень до інформативних сегментів. Вони містять термальні контури, що не відносяться до об'єктів інтересу.

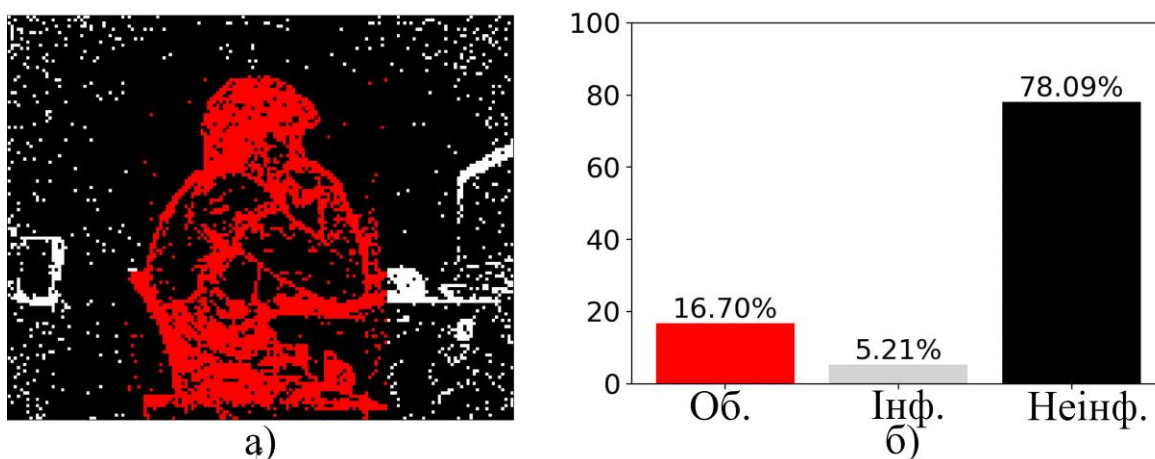


Рисунок 4.13 – Розподіл інформативності у ІЧ-кадрі «Людина»: а) карта інформативності сегментів кадру б) графік залежності класу інформативності сегментів від відсоткового співвідношення у кадрі

Розглянемо показники втрат RMSD при кодуванні об'єктно-інформативного кадру розробленим та існуючими методами: JPEG 2000 та HEVC [89, 134]. Методи кодування було сконфігуровано таким чином, щоб загальнокадровий RMSD був наближено-співмірним. Показники кодування для ІЧ-кадру на рисунку 4.14 а):

- розробленим методом: $RMSD = 1730$, $C_{cmp} = 11.03$ (рис 4.14 б);
- JPEG 2000: $RMSD = 1760$, $C_{cmp} = 10.6$ (рис 4.14 в);

- HEVC: $RMSD = 1798$, $C_{cmp} = 10.01$ (рис 4.14 г).

Отже при співмірних загальних втратах, розроблений метод має перевагу у коефіцієнті стиснення C_{cmp} на 4.06% – 10.19%.

Окрім переваги у стисненні, варто проаналізувати карти просторового розповсюдження $RMSD$. З них видно, що кодеки HEVC та JPEG 2000 вносять суттєві спотворення у область об'єктів інтересу. Тоді як розроблений метод не вносить спотворень у об'єктно-інформативні сегменти. Однак, вносить більші спотворення у фонові структури. Внаслідок цього, максимальне $RMSD$ у сегменті може досягати 12 000. Тоді як сегментне $RMSD$ для HEVC та JPEG 2000 обмежуються 8000.

Для аналізу ефективності кодування окремих сегментів за їх рівнем інформативності у таблиці 4.2 наведені показники для застосованих методів обробки зображень.

Варто звернути увагу на показники втрат для об'єктно-інформативних сегментів. Методи JPEG 2000 та HEVC вносять у них найбільші втрати. Показник $RMSD$ для них у діапазоні 2992 – 3279. Коли розроблений метод дозволяє кодувати ці сегменти безвтратно, $RMSD = 0$.

Натомість, кількісно відображено, що при обробці розробленим методом інформативні сегменти мають найбільші втрати: $RMSD = 3740$. Для експериментального кадру (рис 4.14 а) інформативними сегментами є області кадру із високочастотною термальною семантикою, що не входять до об'єктної області: дерева, стовпи, дорога.

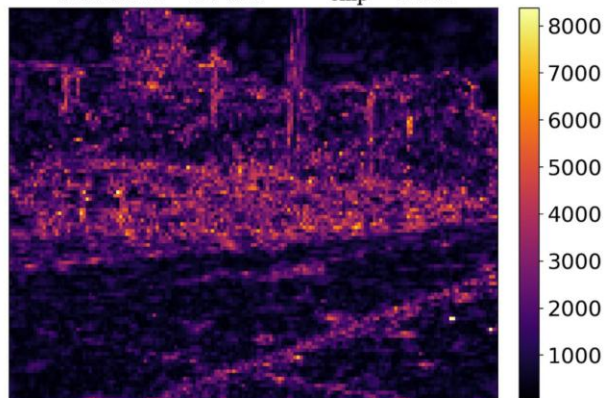
Показники для неінформативних сегментів відображають те, що всі розглянуті методи вносять у них найменші втрати. PSNR не менше 36 дБ. При цьому коефіцієнт стиснення для них найбільший: $C_{cmp} = 37.93$. Це пояснюється високою подібністю даних областей та обмеженістю високочастотних деталей.

Об'єктно-інформативний ІЧ-кадр

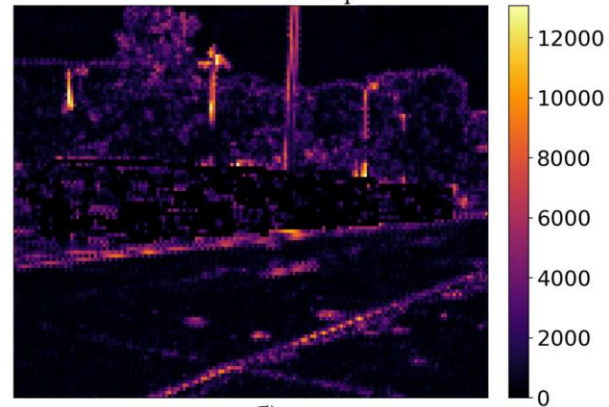


а)

JPEG 2000

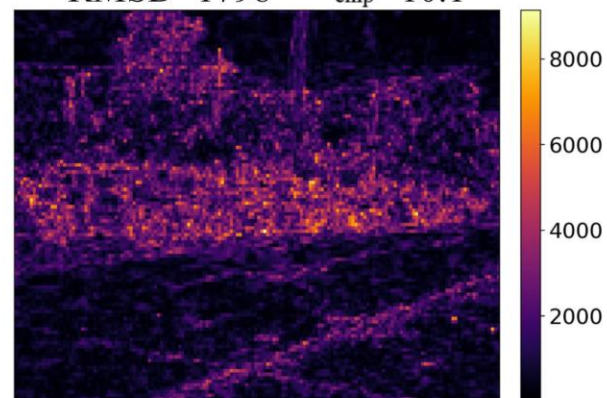
RMSD=1760 $C_{\text{cmp}}=10.6$ 

в)

Розроблений метод
RMSD=1730 $C_{\text{cmp}}=11.03$ 

б)

HEVC

RMSD=1798 $C_{\text{cmp}}=10.1$ 

г)

Рисунок 4.14 – Порівняння карт RMSD після кодування ІЧ-кадру: а) об'єктно-інформативний кадр; б) карта RMSD після кодування розробленим методом; в) карта RMSD після кодування JPEG 2000; г) карта RMSD після кодування HEVC

Таблиця 4.2 – Показники ефективності кодування окремих сегментів за їх рівнем інформативності

Метод обробки зображень	Показник	Рівень інформативності сегменту		
		Неінформативний	Інформативний	Об'єктно-інформативний
Розроблений метод	C_{cmp}	37.93	9.23	1.91
	RMSD	783.41	3740.79	0
	PSNR	38.45 дБ	24.87 дБ	∞
JPEG 2000	RMSD	995.53	2654.95	2992.71
	PSNR	36.37 дБ	27.85 дБ	26.81 дБ
HEVC	RMSD	986.99	2570.54	3279.17
	PSNR	36.44 дБ	28.13 дБ	26.01 дБ

Показники з таблиці 4.2 відображають суттєві відмінності у підході до кодування та внесенні спотворень між розробленим та існуючими методами. Для кількісного підтвердження семантичної цілісності кадрів було розраховано індекс структурної цілісності SSIM між оригінальним кадром (рис 4.14 а) та декодованими. Наведені показники SSIM (рис 4.15) відображають співмірний рівень цілісності, а незначні відхилення узгоджуються із відхиленнями RMSD.

При цьому, було визначено показник SSIM для об'єктно-інформативних сегментів, за умови внесення спотворень у службову інформацію та втрати залишку високочастотних коефіцієнтів. Даний експеримент показав значення SSIM більші для розробленого методу у порівнянні із існуючими на 93%. Що задовольняє вимогам по збереженню цілісності теплових полів об'єктів інтересу [151].

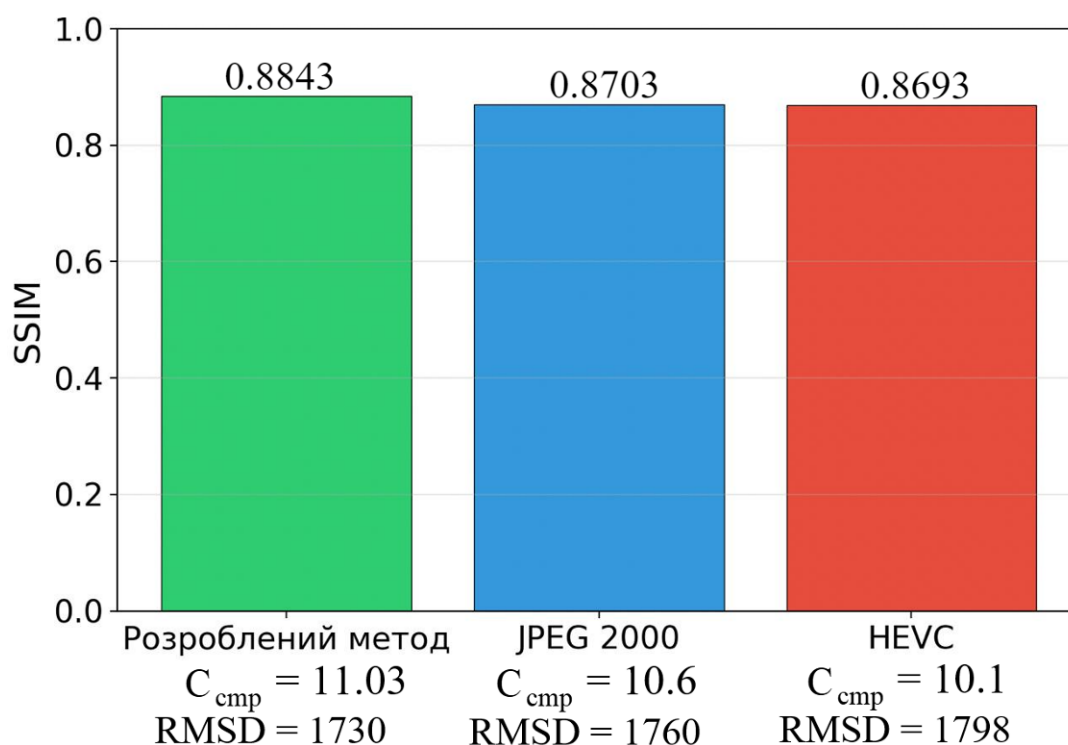


Рисунок 4.15 – Графік залежності SSIM від методу кодування

Для здійснення порівняльного аналізу впливу методу кодування на потокову передачу даних необхідно розглянути показники бітрейту. Результати експериментального вимірювання бітрейту за частоти кадрів 30 FPS та ідентичного значення середньоквадратичного відхилення (RMSD = 1500) наведено у таблиці 4.3.

Порівняно з іншими типами інформативності, об'єктно-інформативні кадри піддаються меншому стисненню за допомогою розробленого методу кодування. Це зумовлено наявністю значної частки сегментів, для яких необхідно зберегти підвищену цілісність даних. Зазначений фактор звужує можливості щодо зменшення бітрейту. Незважаючи на це, бітрейт об'єктно-інформативних кадрів для розробленого методу виявляється на 7% меншим порівняно зі стандартом JPEG 2000 та на 11% меншим порівняно з HEVC.

Таблиця 4.3 – Показники бітрейту отримані експериментальним шляхом при 30 FPS методами обробки сконфігурованими на однаковий RMSD = 1500

Метод обробки зображень		Роздільна здатність		
		480p 720x480	720p 1280x720	1080p 1920x1080
Розроблений метод	Малоінформативне	9.9 Мбіт/с	25.74 Мбіт/с	59.4 Мбіт/с
	Інформативне	15.3 Мбіт/с	39.78 Мбіт/с	91.8 Мбіт/с
	Об'єктно-інформативне	16.4 Мбіт/с	42.38 Мбіт/с	98.4 Мбіт/с
JPEG 2000	Малоінформативне	10.7 Мбіт/с	28.5 Мбіт/с	64.2 Мбіт/с
	Інформативне	16.8 Мбіт/с	44.8 Мбіт/с	100.8 Мбіт/с
	Об'єктно-інформативне	17.5 Мбіт/с	46.6 Мбіт/с	104.9 Мбіт/с
HEVC	Малоінформативне	13.2 Мбіт/с	35.1 Мбіт/с	79.2 Мбіт/с
	Інформативне	17.9 Мбіт/с	47.6 Мбіт/с	107.4 Мбіт/с
	Об'єктно-інформативне	18.2 Мбіт/с	48.4 Мбіт/с	109.2 Мбіт/с

Для існуючих стандартів кодування не фіксується принципової різниці у бітрейті між інформативними та об'єктно-інформативними кадрами. Причиною цього є відсутність механізмів об'єктного розпізнавання. Відповідно, способи обробки об'єктних областей не відрізняються від інших областей кадру.

На противагу існуючим рішенням, для розробленого методу різниця між інформативними та об'єктними типами кадрів є суттєвою. Перевага розробленого методу під час обробки інформативних кадрів складає 10% над JPEG 2000 та 17% над HEVC. На відміну від об'єктних сегментів, кількість інформативних сегментів піддається регулюванню за допомогою порогового параметру. Такий підхід забезпечує відсікання менш виражених високочастотних структур. Водночас, допустиме внесення обмежених спотворень безпосередньо до інформативних областей. Зазначені підходи створюють умови для утримання балансу між бітовим об'ємом та термальною інформативністю зображення.

Для малоінформативних кадрів перевага розробленого методу є менш вираженою, ніж для інформативних. Це пояснюється низькою насиченістю таких кадрів високочастотними деталями. Як для запропонованого, так і для існуючих методів, відсутність високочастотних складових обмежує загальний потенціал стиснення даних. Проте навіть за таких умов розроблений алгоритм дозволяє досягти бітрейту, який на 8% менший за показники JPEG 2000 та на 34% менший за HEVC. Метод HEVC показав низьку ефективність обробки малоінформативних ІЧ-кадрів.

На рисунку 4.16 зображено графік залежності бітрейту від роздільної здатності ІЧ-кадру, на основі табличних даних. На ньому видно на скільки суттєвою є відмінність бітрейтів для кадрів із високою роздільною здатністю 1920x10080.

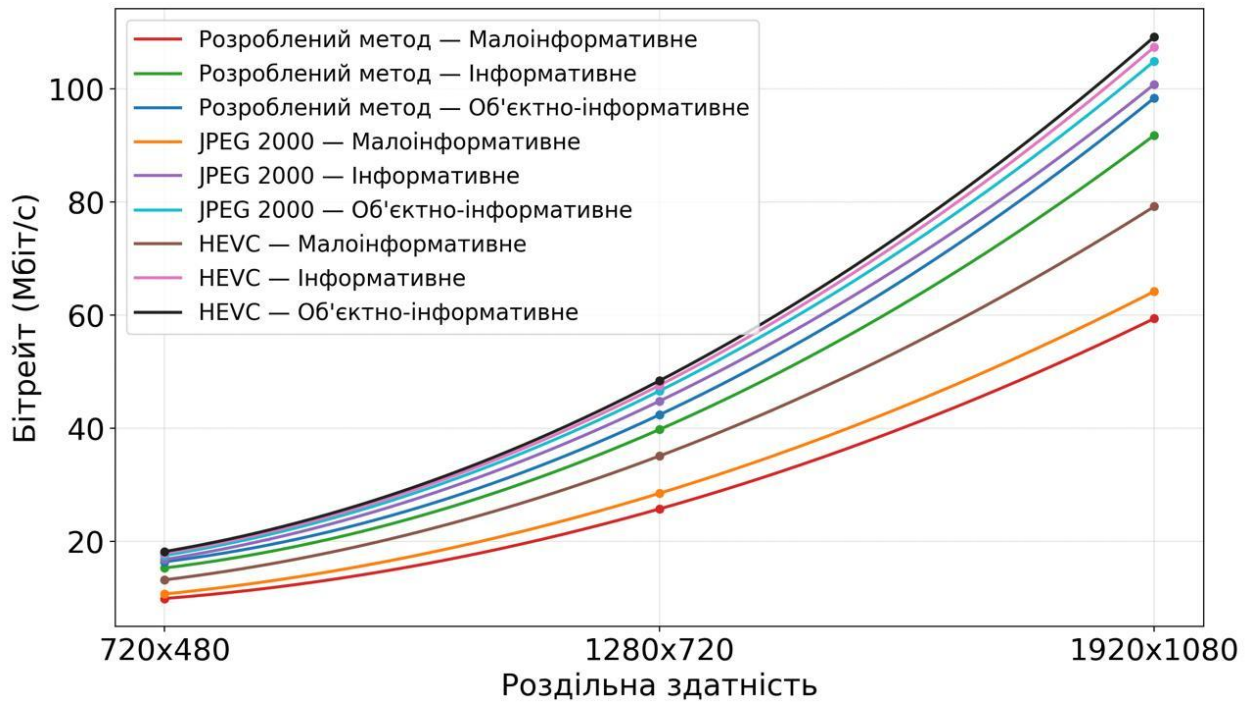


Рисунок 4.16 – Графік залежності бітрейту від роздільної здатності ІЧ-кадру

Підсумовуючи, існуючі методи кодування орієнтовані на зниження психовізуальної надмірності за однаковим принципом по всьому кадру. Водночас розроблений метод кодування ІЧ-кадрів дозволяє адаптувати параметри стиснення до визначеного рівня інформативності локальних областей. Такий підхід дозволяє зберегти встановлений рівень цілісності термальних сигнатур та зменшити бітрейт на 7–10% у залежності від рівня інформативності сегментів.

Висновки за четвертим розділом

1. Наукова новизна. Удосконалено метод відновлення стиснутих ІЧ-зображень на основі використання бігармонічних функцій. Відмінності методу полягають у просторовому прорідженні низькоінформативних областей бітових пластів ІЧ-зображень з їх регенерацією за регулярно-просторовим шаблоном в процесі відновлення з використанням бігармонічних функцій. Це дозволяє підвищити рівень термальної цілісності інформативних сегментів ІЧ-зображень без втрат бітрейту.

2. Проведено абляційний аналіз впливу структурно-спектральних етапів кодування на бітову інтенсивність та точність відновлення інфрачервоних кадрів. Результати аналізу підтверджують наступне:

- послідовне застосування бітової біфуркації, різницевого кодування та корекції низькочастотних коефіцієнтів усуває статистичну надмірність та, як наслідок, підвищує коефіцієнт стиснення до 2.37 без внесення спотворень;

- застосування семантично орієнтованої адаптивної квантизації зосереджує втрати переважно в низькоінформативних сегментах, що збільшує коефіцієнт стиснення до 7.22;

- інтеграція просторового прорідження з бігармонічною реконструкцією дозволяє досягти підсумкового коефіцієнта стиснення 7.66.

Така архітектура перетворень забезпечує індекс структурної подібності (SSIM) на рівні 0.97 для всього кадру та забезпечує цілісність теплових сигнатур об'єктів інтересу.

3. Здійснено порівняльну оцінку характеристик розробленого методу з існуючими стандартами кодування. Експериментальні дані підтверджують наступне:

- для об'єктно-інформативних сегментів індекс структурної подібності перевищує показники існуючих аналогів на 93%, що зберігає заданий рівень термальної цілісності об'єктів інтересу;

- досягається зниження бітрейту на 7-10% у порівнянні із існуючими методами, за умови ідентичного середньоквадратичного відхилення.

Результати досліджень щодо матеріалів, викладених в четвертому розділі, опубліковані в таких наукових працях як: [103, 105, 117].

ВИСНОВКИ

У дисертаційній роботі вирішена *науково-прикладна задача*, яка полягає у зниженні бітрейту інфрачервоного потоку даних для передачі в інформаційних системах з визначеним рівнем термальної цілісності. Розроблено метод формування двопластової інформативної карти, яка застосовується для семантико-орієнтованого кодування сегментів. За наявності достатніх апаратних ресурсів метод дозволяє задіяти штучну нейромережеву модель для виявлення об'єктно-інформативних сегментів.

Подальший розвиток отримало застосування бігармонічної реконструкції для відновлення проріджених низькоінформативних областей ІЧ-кадру на етапі декодування.

При аналізі сформульованої задачі було виявлене науково-технічне протиріччя. Воно полягає у невідповідності між необхідністю зниження бітрейту потоку інфрачервоних даних та вимогами до збереження теплової семантики об'єктів інтересу.

Звідси, *мета дисертаційних досліджень* полягає у розробці методу кодування інфрачервоних зображень для зниження бітрейту без втрат термальної цілісності на основі семантико-орієнтовної обробки з використанням моделей штучного інтелекту.

Основні наукові результати:

1. Проведено дослідження характеристик ІЧ-потоків, що генеруються сучасними пристроями. Виявлено недоліки при обробці ІЧ-кадрів існуючими методами кодування. Виявлені обмеження спеціальних ІЧ-кодеків та розповсюджених нейромережевих підходів.

2. Розроблено метод біфуркації бітових пластів кадру із розширеним динамічним діапазоном. Метод дозволяє розділити інформаційну структуру зображення на значущі та фонові складові у рівнях окремих бітових пластів. Такий підхід дозволяє зменшити просторову надлишковість кадру.

3. Розроблено метод визначення метрики інформативності сегментів. Це дозволяє здійснювати кількісну оцінку інформативності просторових областей на основі їхньої текстурної насиченості.

4. Розроблено метод формування двопластової інформативної карти інфрачервоного кадру. Метод дозволяє виокремити пріоритетні області обробки у порогово-метричному просторі старшого та молодшого бітових пластів.

5. Розроблено метод виявлення об'єктно-інформативних сегментів на основі застосування моделей штучного інтелекту. Це дозволяє сегментувати об'єкти інтересу через кон'юнкцію результату нейронної мережі із сформованою картою інформативності. Такий підхід дозволяє обмежити рівень спотворень у пріоритетних областях кадру.

6. Розроблено метод різницевого кодування для формування мінісегментів зі зменшеним динамічним діапазоном. Це дозволяє трансформувати значення пікселів ІЧ-кадру у локальні температурні відхилення.

7. Розроблено метод корекції низькочастотних коефіцієнтів на основі високочастотних компонент вейвлет-перетворення. Метод дозволяє звужити динамічний діапазон низькочастотних значень при цьому не вносячи додаткових спотворень.

8. Розроблено метод адаптивної квантизації на основі рівня інформативності та бітового пласту сегментів. Метод дозволяє диференційовано розподіляти коефіцієнти квантизації залежно від пріоритетності сегментів. Це рішення дозволяє обмежувати бітрейт ІЧ-поток при встановленому рівні термальної цілісності.

9. Розроблено метод спектрально-групового кодування сегментів. Метод дозволяє здійснювати структурне кодування спектральних значень за типами частотних груп в умовах квадродекомпозиції. Застосування цього підходу звужує обсяг бітового надлишку та зменшує інтенсивність потоку даних.

10. Проведено абляційний аналіз впливу розроблених методів на кодування ІЧ-кадрів. Відображено, що комплексна інтеграція запропонованих методів кумулятивно знижує бітовий об'єм даних зберігаючи цілісність інформативних областей.

Основні результати дисертаційної роботи відображені в проміжному звіті по НДР «Розробка математичних моделей обробки видової інформації для підвищення якості функціонування дистанційних сервісів» (№ 0125U004039), в якій автор є відповідальним виконавцем.

Результати досліджень прийняті до впровадження на базі підприємств: ТОВ «Одеський авіаційний завод»; ТОВ «ЮЕЙ ДЕФЕНС» (Додаток Б).

Таким чином, мета дослідження щодо розробки методу кодування інфрачервоних зображень для зниження бітрейту без втрат термальної цілісності на основі семантико-орієнтовної обробки з використанням моделей штучного інтелекту досягнута і всі поставлені в ході дослідження завдання вирішені повністю. Наукові результати є внеском у розвиток методів семантико-орієнтованої обробки даних для передачі у інформаційних системах.

Отримані результати можуть бути використані у розробці апаратно-програмних комплексів передачі інфрачервоних тепловізійних даних в інформаційних системах. Розроблений метод адаптовано для інтеграції в бортові системи з обмеженими апаратними ресурсами. Впровадження цієї технології знижує бітове навантаження потоків даних та забезпечує актуальність інформації під час динамічного моніторингу.

Основні практичні результати дослідження:

Застосування розробленого методу кодування інфрачервоних зображень із застосуванням моделей штучного інтелекту в інформаційних системах дозволяє отримати наступні переваги:

- зменшення бітрейту на 7-10% у залежності від інформаційної насиченості ІЧ-кадрів при порівнянні із існуючими методами;

- збереження термальної цілісності об'єктів інтересу, що підтверджується збільшеним SSIM на 93% для об'єктно-інформативних сегментів при порівнянні із існуючими методами.

Достовірність результатів щодо характеристик розробленого методу підтверджується наступним:

- дотриманням принципів теорії кодування інформації. Сформовані у дослідженні математичні моделі виступають закономірним продовженням відомих підходів. Завдяки цьому виконана аналітична робота та опсиані перетворення характеризуються системною узгодженістю;

- результатами проведених експериментів. Проведені експерименти охоплюють широкий діапазон варіативних вхідних даних. Задані параметри тестування цілеспрямовано адаптовано до характеристик реальних інформаційних систем. Такий підхід дозволяє об'єктивно екстраполювати отримані лабораторні показники на умови експлуатації діючих систем;

- апробацією та впровадженням здобутків роботи. Теоретичні висновки пройшли незалежне рецензування шляхом опублікування у виданнях, що індексуються наукометричними базами Scopus та Web of Science, а також через представлення на профільних науково-технічних конференціях. Водночас прикладна цінність методу документально засвідчується офіційними актами його впровадження у робочі процеси спеціалізованих підприємств.

Результати дисертаційної роботи доцільно використати:

- при проектуванні бортових комплексів, зокрема для безпілотних літальних апаратів та автономних роботизованих платформ. Застосування розробленого методу обмежує бітрейт інфрачервоного потоку даних та, як наслідок, дозволяє підвищити актуальність інформації для систем автоматичного прийняття рішень або віддаленого оператора в режимі реального часу;

- для розвитку систем промислової діагностики та віддаленого моніторингу. Отримані результати формують підґрунтя для розвитку

масштабованих багатосенсорних мереж контролю об'єктів критичної інфраструктури та інформаційних систем розумного міста. Застосування розробленого методу дозволяє знизити бітове навантаження ІЧ-потоків даних, що відкриває можливості для розширення кількості ІЧ-сенсорів в рамках обмеженої пропускної здатності інформаційних систем;

- при реалізації освітньо-професійних та наукових програм. Залучення розроблених методів розширить навчально-методичну базу. Як наслідок, це адаптує теоретичну підготовку студентів спеціальності «Комп'ютерні науки» до сучасних вимог розробки та експлуатації інформаційних систем.

Основними цільовими користувачами результатів дослідження можуть бути:

- підприємства оборонно-промислового комплексу та приладобудівної галузі, зокрема ТОВ «Одеський авіаційний завод» та ТОВ «ЮЕЙ ДЕФЕНС», де результати досліджень вже прийняті до впровадження;

- інженерні підрозділи, що створюють спеціалізовані апаратно-програмні комплекси для обробки видової інформації. Розроблений метод надає інструментарій, який дозволяє зберігати встановлений рівень термальної цілісності забезпечуючи зменшений бітрейт інформаційного потоку;

- наукові колективи, які досліджують семантико-орієнтовані методи обробки інформації та алгоритми комп'ютерного зору. Результати дослідження розширюють можливості застосування методів кодування та нейромережових підходів для класифікації інформативності сегментів. Це актуально для інформаційних систем, де точність відтворення локальних структур виступає визначальним фактором ефективності виявлення цілей.

ПЕРЕЛІК ДЖЕРЕЛ ПОСИЛАННЯ

1. Баранник В. В. Кодирование трансформированных изображений в инфокоммуникационных системах / В. В. Баранник, В. П. Поляков. Харьков : ХУПС, 2010. 234 с.
2. Баранник В. В. Метод повышения информационной безопасности в системах видеомониторинга кризисных ситуаций : монография / В. В. Баранник, О. С. Кулица. Черкассы, 2015. 143 с.
3. Баранник В. В. Метод сжатия изображений на основе неравновесного позиционного кодирования битовых плоскостей / В. В. Баранник, Н. К. Гулак, Н. А. Королева. *Радіоелектронні і комп'ютерні системи*. 2009. № 1. С. 84–92. URL: http://nbuv.gov.ua/UJRN/recs_2009_1_13 (дата звернення: 31.03.2026).
4. Баранник В. В. Рельефное представление изображений пирамидальным кодированием / В. В. Баранник. *Інформаційно-керуючі системи на залізничному транспорті*. 2001. № 1. С. 17–25.
5. Баранник В. В. Структурно-комбинаторное представление данных в АСУ / В. В. Баранник, Ю. В. Стасев, Н. А. Королева. Харків : ХУПС, 2009. 252 с.
6. Бараннік В. В. Метод біфуркації бітових площин в системах обробки інфрачервоних зображень / В. В. Бараннік, А. А. Берчанов. *Наукоємні технології*. 2025. Т. 66, № 2. С. 208–213. DOI: <https://doi.org/10.18372/2310-5461.66.20325>.
7. Бараннік В. В. Перспективи розвитку технологій обробки відеоінформації в умовах інформаційно-інтелектуального протиборства / В. В. Бараннік, А. А. Берчанов. *Наукоємні технології в інфокомунікаціях (НІСТ'2023): матеріали Міжнар. наук.-практ. конф. (Кам'янець-Подільський, 1–3 черв. 2023 р.)*. Кам'янець-Подільський, 2023. URL: <http://repositsc.nuczu.edu.ua/bitstream/123456789/18013/1/%D0%9D%D0%86%D0%A1%D0%A2%202023.pdf> (дата звернення: 30.01.2026).

8. Бараннік В. Метод спектрально-групового кодування бітових шарів інфрачервоних зображень / В. Бараннік, А. Берчанов. Інфокомунікаційні технології та електронна інженерія. 2025. Вип. 5, № 2. С. 15–24. DOI: <https://doi.org/10.23939/ictee2025.02.015>.

9. Берчанов А. А. Метод обробки інфрачервоних зображень у спектральному домені / А. А. Берчанов, В. В. Бараннік. Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: Технічні науки. 2025. Т. 36 (75), № 3. С. 55–60. DOI: <https://doi.org/10.32782/2663-5941/2025.3.2/08>.

10. Ієремєєв О. І. Комбінована метрика візуальної якості зображень дистанційного зондування на основі нейронної мережі / О. І. Ієремєєв, В. В. Лукін, К. Окарма. Радіоелектронні і комп'ютерні системи. 2020. № 4. С. 4–15. DOI: <https://doi.org/10.32620/reks.2020.4.01>.

11. Інтелектуалізовані методи обробки тепловізійних зображень / А. А. Берчанов, В. В. Бараннік, Р. О. Прокопенко та ін. Комп'ютерне моделювання в наукоємних технологіях (КМНТ-2024): матеріали 10-ї Міжнар. наук.-техн. конф. (Харків, 27–29 листоп. 2024 р.). Харків, 2024. С. 24–26. URL: <https://csai.karazin.ua/news/uk/x-naukovo-tehnichna-mizhnarodna-28052025> (дата звернення: 05.02.2026).

12. Критерії вибору спектрально-ефективних сигналів у бездротових інформаційних мережах / В. Козловський, О. Толстікова, А. Савченко та ін. Наукоємні технології. 2022. № 4(56). С. 273–286. DOI: <https://doi.org/10.18372/2310-5461.56.17125>.

13. Метод адаптивного цілісного арифметичного кодування з врахуванням RLE-перетворення / В. В. Бараннік, Ю. В. Цімура, Д. С. Гаврилов та ін. Системи і технології зв'язку, інформатизації та кібербезпеки. 2023. № 3. С. 5–13. DOI: <https://doi.org/10.58254/viti.3.2023.01.05>.

14. Метод аналізу взаємодії параметрів QOE та QOS на основі алгоритмів керування машинами / Р. Одарченко, М. Рябенко, М. Іванова та

ін. Наукоємні технології. 2022. № 4(56). С. 305–316. DOI: <https://doi.org/10.18372/2310-5461.56.17130>.

15. Метод групового кодування інфрачервоних зображень у спектрально-хвильовому просторі / В. В. Бараннік, А. А. Берчанов, В. В. Бараннік та ін. Вісник НТУУ «КПІ». Серія Радіотехніка, Радіоапаратобудування. 2025. № 99. С. 24–34. DOI: <https://doi.org/10.20535/RADAR.2025.99.24-34>.

16. Метод зберігання елементів службових складових криптокомпресійних кодограм відеозображень / В. В. Бараннік, С. О. Сідченко, Д. В. Бараннік та ін. Вісник НТУУ «КПІ». Серія Радіотехніка, Радіоапаратобудування. 2023. № 92. С. 28–40. DOI: <https://doi.org/10.20535/RADAR.2023.92.28-40>.

17. Метод кластеризації послідовності трансформант за структурними ознаками їх спектрально-параметричного опису / В. В. Бараннік, Є. С. Єлісеєв, Ю. Цімура та ін. Наукоємні технології. 2024. № 2(62). С. 185–192. DOI: <https://doi.org/10.18372/2310-5461.62.18712>.

18. Метод кодування відеоданих в спектрально-параметричному просторі / Ю. Цімура, А. Костромицький, О. Суханов та ін. Інфокомунікаційні технології та електронна інженерія. 2024. Т. 4, № 1. С. 61–69. DOI: <https://doi.org/10.23939/ict2024.01.061>.

19. Метод кодування відеосегментів в спектрально-кластерному просторі з виявленням структурних особливостей / В. В. Бараннік, А. О. Красноручський, В. О. Колесник та ін. Вісник НТУУ «КПІ». Серія Радіотехніка, Радіоапаратобудування. 2022. № 90. С. 21–30. DOI: <https://doi.org/10.20535/RADAR.2022.90.21-30>.

20. Метод кодування значимих за впливом на семантичну цілісність відеосегментів для забезпечення доступності / В. В. Бараннік, Ю. М. Бабенко, В. В. Бараннік та ін. Наукоємні технології. 2022. Т. 54, № 2. С. 118–126. DOI: <https://doi.org/10.18372/2310-5461.54.16749>.

21. Метод семантично-орієнтованої обробки бітових площин інфрачервоних зображень / В. В. Бараннік, А. А. Берчанов, В. В. Бараннік та ін. Вісник НТУУ «КПІ». Серія Радіотехніка, Радіоапаратобудування. 2025. № 102. С. 15–24. DOI: <https://doi.org/10.20535/RADAP.2025.102.15-24>.

22. Метод стиснення відеосегментів з забезпеченням їх достовірності в спектрально-кластерному просторі / В. В. Бараннік, А. О. Красноручський, В. О. Колесник та ін. Вісник НТУУ «КПІ». Серія Радіотехніка, Радіоапаратобудування. 2023. № 92. С. 41–53. DOI: <https://doi.org/10.20535/RADAP.2023.92.41-53>.

23. Метод стиснення кластеризованих трансформант на основі блочного кодування з локально-монотонним визначенням довжини / В. В. Бараннік, Є. С. Єлісеєв, Ю. Цімура та ін. Наукоємні технології. 2024. № 3(63). С. 274–281. DOI: <https://doi.org/10.18372/2310-5461.63.18971>.

24. Модель інформативного опису спектрального простору відеосегментів за діагонально-нерівномірною текстурою / В. В. Бараннік, С. С. Шульгін, Р. С. Онищенко та ін. Наукоємні технології. 2022. № 4(56). С. 259–267. DOI: <https://doi.org/10.18372/2310-5461.56.17124>.

25. Підвищення ефективності доведення інфрачервоних знімків методом спектрально-хвильового кодування / В. В. Бараннік, А. А. Берчанов, В. В. Бараннік та ін. Кіберборотьба: розвідка, захист та протидія: матеріали III Міжнар. наук.-практ. конф. (Київ, 6–7 трав. 2025 р.). Київ, 2025. С. 14–15. DOI: <https://doi.org/10.61929/viti.mnpk.3.2025>.

26. Теоретичні основи створення методів протидії прихованим інформаційним атакам в сучасній гібридній війні / А. М. Алімпієв, В. В. Бараннік, Т. В. Белікова та ін. Системи обробки інформації. 2017. Вип. 4(150). С. 113–121.

27. Технологія ковзного кодування нерівномірних діагональних послідовностей в двовимірному спектральному просторі трансформанти / В. В. Бараннік, С. С. Шульгін, Ю. М. Бабенко та ін. Вісник НТУУ «КПІ». Серія

Радіотехніка, Радіоапаратобудування. 2023. № 94. С. 13–23. DOI: <https://doi.org/10.20535/RADAP.2023.94.13-23>.

28. A Fast JPEG Image Compression Algorithm Based on DCT / W. Xu, N. Wang, A. Hussain et al. IEEE International Conference on Smart Cloud (SmartCloud). Washington, DC : IEEE, 2020. P. 106–110. DOI: <https://doi.org/10.1109/SmartCloud49737.2020.00028>.

29. A method for forming a truncated positional code system for transformed video images / V. Barannik, R. Onyshchenko, G. Pris et al. Informatyka, Automatyka, Pomiar y w Gospodarce i Ochronie Środowiska. 2024. Vol. 14, No. 3. P. 56–60. DOI: <https://doi.org/10.35784/iapgos.6131>.

30. A Method of Effective Coding of a Dynamic Video Information Resource in an Information and Telecommunications Network / A. Berchanov, R. Prokopenko, P. Pertsev et al. IEEE 5th International Conference on Advanced Trends in Information Theory (ATIT). Lviv : IEEE, 2024. P. 51–57. DOI: <https://doi.org/10.1109/ATIT64324.2024.11222407>.

31. A Method of Scrambling for the System of Cryptocompression of Codograms Service Components / V. Barannik, S. Sidchenko, D. Barannik et al. Emerging Networking in the Digital Transformation Age. TCSET 2022 / eds. M. Klymash et al. Cham : Springer, 2023. P. 444–459. (Lecture Notes in Electrical Engineering ; vol. 965). DOI: https://doi.org/10.1007/978-3-031-24963-1_26.

32. A method to control bit rate while compressing predicted frames / V. Barannik, N. Kharchenko, O. Y. Othman Shadi et al. The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM). Lviv : IEEE, 2015. P. 36–38. DOI: <https://doi.org/10.1109/CADSM.2015.7230789>.

33. A Novel Image Compression Method Using Wavelet Coefficients and Huffman Coding / S. Thomas, A. Krishna, S. Govind et al. Journal of Engineering Research. 2025. Vol. 13, No. 1. P. 361–370. DOI: <https://doi.org/10.1016/j.jer.2023.08.015>.

34. Altamimi A. Lossless and near-lossless compression algorithms for remotely sensed hyperspectral images / A. Altamimi, B. Ben Youssef. *Entropy*. 2024. Vol. 26, No. 4. Article 316. DOI: <https://doi.org/10.3390/e26040316>.
35. Altay F. The use of thermal cameras for pedestrian detection / F. Altay, S. Velipasalar. *IEEE Sensors Journal*. 2022. Vol. 22, No. 12. P. 11489–11498. DOI: <https://doi.org/10.1109/JSEN.2022.3172386>.
36. An Extensive Analysis of Digital Image Compression Techniques Using Different Image Files and Color Formats / Fauziah, D. A. Lantana, Nurhayati et al. *International Journal on Advanced Science, Engineering and Information Technology*. 2023. Vol. 13, No. 5. P. 1971–1977. DOI: <https://doi.org/10.18517/ijaseit.13.5.19319>.
37. An improved lossless image compression algorithm based on Huffman coding / X. Liu, P. An, Y. Chen et al. *Multimedia Tools and Applications*. 2021. Vol. 81. P. 4781–4795. DOI: <https://doi.org/10.1007/s11042-021-11017-5>.
38. Analysis of thermal imaging performance under extreme foggy conditions: Applications to autonomous driving / J. M. Rivera Velázquez, L. Khoudour, G. Saint Pierre et al. *Journal of Imaging*. 2022. Vol. 8, No. 11. Article 306. DOI: <https://doi.org/10.3390/jimaging8110306>.
39. Balle J. End-to-end Optimized Image Compression / J. Balle, V. Laparra, E. P. Simoncelli. *International Conference on Learning Representations (ICLR)*. Toulon : OpenReview.net, 2017. DOI: <https://doi.org/10.48550/arXiv.1611.01704>.
40. Barannik D. Steganographic Coding Technology for Hiding Information in Infocommunication Systems of Critical Infrastructure / D. Barannik, V. Barannik. *IEEE 4th International Conference on Advanced Trends in Information Theory (ATIT)*. Kyiv : IEEE, 2022. P. 88–91. DOI: <https://doi.org/10.1109/ATIT58178.2022.10024185>.
41. Barannik V. Architectural presentation of isotopic levels of relief of images / V. Barannik, A. Hahanova, A. Slobodyanyuk. *IEEE 10th International Conference on Experience of Designing and Application of CAD Systems in Microelectronics (CADSM)*. Polyana : IEEE, 2009. P. 385–387.

42. Barannik V. Indirect information hiding technology on a multiadic basis / V. Barannik, N. Barannik, O. Slobodyanyuk. *Informatyka, Automatyka, Pomiar w Gospodarce i Ochronie Środowiska*. 2021. Vol. 11, No. 4. P. 14–17. DOI: <https://doi.org/10.35784/iapgos.2812>.

43. Barannik V. Method of indirect information hiding in the process of video compression / V. Barannik, V. Khimenko, N. Barannik. *Radioelectronic and Computer Systems*. 2021. No. 4. P. 119–131. DOI: <https://doi.org/10.32620/reks.2021.4>.

44. Barannik V. Method of the 3-D image processing / V. Barannik, S. Karpenko. *Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET)*. Lviv-Slavsko : IEEE, 2008. P. 378–380.

45. Barannik V. Quadrature compression of images in polyadic space / V. Barannik, A. Shiryaev. 11th International Conference on Modern Problems of Radio Engineering, Telecommunications and Computer Science (TCSET). Lviv-Slavsko : IEEE, 2012. P. 422.

46. Bayer CFA Pattern Compression With JPEG XS / T. Richter, S. Föbel, A. Descampe et al. *IEEE Transactions on Image Processing*. 2021. Vol. 30. P. 6557–6569. DOI: <https://doi.org/10.1109/TIP.2021.3095421>.

47. Bilal A. Efficient cooperative image transmission in one-way multi-hop sensor network / A. Bilal, I. Hacı. *The International Journal of Electrical Engineering & Education*. 2020. Vol. 57, No. 4. P. 321–339. DOI: <https://doi.org/10.1177/0020720918816009>.

48. Breaking the Barriers of One-to-One Usage of Implicit Neural Representation in Image Compression: A Linear Combination Approach With Performance Guarantees / S. Sanjeet, S. Hosseinalipour, J. Xiong et al. *IEEE Internet of Things Journal*. 2025. Vol. 12, No. 7. P. 8947–8957. DOI: <https://doi.org/10.1109/JIOT.2024.3502690>.

49. Chidurala V. Occupancy estimation using thermal imaging sensors and machine learning algorithms / V. Chidurala, X. Li. *IEEE Sensors Journal*. 2021. Vol. 21, No. 6. P. 8627–8638. DOI: <https://doi.org/10.1109/JSEN.2021.3049311>.

50. Comparing Freeman Chain Code 4 Adjacency Algorithm and LZMA Algorithm in Binary Image Compression / R. Resdiansyah, J. D. Adrian, A. H. Wijaya et al. *Journal of Physics: Conference Series*. 2021. Vol. 1783. Article 012045. DOI: <https://doi.org/10.1088/1742-6596/1783/1/012045>.

51. Compressed Image Generation with Denoising Diffusion Codebook Models / G. Ohayon, H. Manor, T. Michaeli et al. *ArXiv*. 2025. DOI: <https://doi.org/10.48550/arXiv.2502.01189>.

52. Compression method in terms of ensuring the fidelity of video images in infocommunication networks / V. Barannik, A. Krasnorutsky, V. Kolesnik et al. *Radioelectronic and Computer Systems*. 2022. No. 4(100). P. 10–24. DOI: <https://doi.org/10.32620/reks.2022.5/09>.

53. Compression of Infrared Images / C. Mantel, S. Forchhammer. *Proc. IS&T Int'l. Symp. on Electronic Imaging: Visual Information Processing and Communication VIII*. 2017. P. 21–26. DOI: <https://doi.org/10.2352/ISSN.2470-1173.2017.2.VIPC-401>.

54. Compression of Pulsed Infrared Thermography Data With Unsupervised Learning for Nondestructive Evaluation of Additively Manufactured Metals / X. Zhang, J. Saniie, S. Bakhtiari et al. *IEEE Access*. 2022. Vol. 10. P. 9094–9107. DOI: <https://doi.org/10.1109/ACCESS.2022.3141654>.

55. Cooperative multiagent deep reinforcement learning for reliable surveillance via autonomous multi-UAV control / W. J. Yun, S. Park, J. Kim et al. *IEEE Transactions on Industrial Informatics*. 2022. Vol. 18, No. 10. P. 7086–7096. DOI: <https://doi.org/10.1109/TII.2022.3143175>.

56. Deep Residual Learning-Based Enhanced JPEG Compression in the Internet of Things / H. Qiu, Q. Zheng, G. Memmi et al. *IEEE Transactions on Industrial Informatics*. 2021. Vol. 17, No. 3. P. 2124–2133. DOI: <https://doi.org/10.1109/TII.2020.2994743>.

57. Description of the OFDM symbol with the help of mathematical laws. Analysis of technologies that were used in this case / V. Barannik, M. Lytvinenko, D. Okladnoy et al. *2nd International Conference on Advanced Information and*

Communication Technologies (AICT). Lviv : IEEE, 2017. P. 183–187. DOI: <https://doi.org/10.1109/AIACT.2017.8020095>.

58. Detections of sustainable areas for steganographic embedding / V. Barannik, A. Alimpiev, D. Barannik et al. 14th International Conference on the Experience of Designing and Application of CAD Systems in Microelectronics (CADSM). Polyana : IEEE, 2017. P. 110–112. DOI: <https://doi.org/10.1109/EWDTS.2017.8110028>.

59. Development of Adaptive Arithmetic Coding Method to the Sequence of Bits / V. Barannik, D. Havrylov, V. Barannik et al. Engineer of the XXI Century / eds. S. Zawiślak, J. Rysiński. Cham : Springer, 2020. (Mechanisms and Machine Science ; vol. 70). DOI: https://doi.org/10.1007/978-3-030-13321-4_18.

60. Developments in Image Processing Using Deep Learning and Reinforcement Learning / J. Valente, J. António, C. Mora et al. Journal of Imaging. 2023. Vol. 9, No. 10. Article 207. DOI: <https://doi.org/10.3390/jimaging9100207>.

61. Developments in international video coding standardization after AVC, with an overview of versatile video coding (VVC) / B. Bross, J. Chen, J.-R. Ohm et al. Proceedings of the IEEE. 2021. Vol. 109, No. 9. P. 1463–1493. DOI: <https://doi.org/10.1109/JPROC.2020.3043399>.

62. Devising a conceptual method for generating cryptocompression codograms of images without loss of information quality / V. Barannik, S. Sidchenko, D. Barannik et al. Eastern-European Journal of Enterprise Technologies. 2021. Vol. 4, No. 2(112). P. 6–16. DOI: <https://doi.org/10.15587/1729-4061.2021.237359>.

63. Devising a method for complex steganographic embedding of information in the structural-psychovisual space / V. Barannik, D. Barannik, M. Babenko et al. Eastern-European Journal of Enterprise Technologies. 2026. No. 1/9(139). P. 19–30. DOI: <https://doi.org/10.15587/1729-4061.2026.351181>.

64. Dua Y. Comprehensive Review of Hyperspectral Image Compression Algorithms / Y. Dua, V. Kumar, R. S. Singh. Optical Engineering. 2020. Vol. 59, No. 9. Article 090902. DOI: <https://doi.org/10.1117/1.OE.59.9.090902>.

65. Efficient Learned Image Compression Through Knowledge Distillation / F. Allemand, A. Fiandrotti, S. Chaudhuri et al. ArXiv. 2025. DOI: <https://doi.org/10.48550/arXiv.2509.10366>.
66. Efficient Lightweight Attention Based Learned Image Compression / Z. He, L. Luo, L. Zhang, H. Guo, C. Zhu. 2023 IEEE International Conference on Visual Communications and Image Processing (VCIP). Jeju : IEEE, 2023. P. 1–5. DOI: <https://doi.org/10.1109/VCIP59821.2023.10402612>.
67. ELIC: Efficient Learned Image Compression With Unevenly Grouped Space-Channel Contextual Adaptive Coding / D. He, Z. Yang, W. Peng et al. Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans : IEEE, 2022. P. 5718–5727. DOI: <https://doi.org/10.48550/arXiv.2203.10886>.
68. End-to-End Optimized Versatile Image Compression With Wavelet-Like Transform / H. Ma, D. Liu, N. Yan et al. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2022. Vol. 44, No. 3. P. 1247–1263. DOI: <https://doi.org/10.1109/TPAMI.2020.3026003>.
69. Evaluation of Effectiveness of Masking Methods of Aerial Photographs / V. Barannik, D. Barannik, V. Fustii et al. 3rd International Conference on Advanced Information and Communications Technologies (AICT). Lviv : IEEE, 2019. P. 415–418. DOI: <https://doi.org/10.1109/AIACT.2019.8847820>.
70. Extreme Compression of Adaptive Neural Images / L. Hoshikawa, M. V. Conde, T. Ohashi et al. 2025 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW). Honolulu : IEEE, 2025. P. 4042–4052. DOI: <https://doi.org/10.1109/ICCVW69036.2025.00420>.
71. Fan K. Infrared Image Super-Resolution via Progressive Compact Distillation Network / K. Fan, K. Hong, F. Li. Electronics. 2021. Vol. 10, No. 24. Article 3107. DOI: <https://doi.org/10.3390/electronics10243107>.
72. Garg G. Analysis of image types, compression techniques and performance assessment metrics: a review / G. Garg, R. Kumar. Journal of Information and Optimization Sciences. 2022. Vol. 43, No. 3. P. 429–436. DOI: <https://doi.org/10.1080/02522667.2022.2037282>.

73. General video coding technology in responses to the joint call for proposals on video compression with capability beyond HEVC / B. Bross, J. Chen, C. Kounnas et al. *IEEE Transactions on Circuits and Systems for Video Technology*. 2020. Vol. 30, No. 5. P. 1226–1240. DOI: <https://doi.org/10.1109/TCSVT.2019.2949619>.

74. Hazowary P. A Comparison of Image Compression Techniques Over Wireless Fading Channel / P. Hazowary, S. K. Dutta, R. Subadar. *Contemporary Issues in Communication, Cloud and Big Data Analytics*. Singapore : Springer, 2022. (Lecture Notes in Networks and Systems ; vol. 281). DOI: https://doi.org/10.1007/978-981-16-4244-9_19.

75. Hemalatha K. Image compression using HAAR discrete wavelet transform / K. Hemalatha, V. Muneeswaran. *5th International Conference on Devices, Circuits and Systems (ICDCS)*. Coimbatore : IEEE, 2020. P. 271–274. DOI: <https://doi.org/10.1109/ICDCS48716.2020.243596>.

76. High-Fidelity Generative Image Compression / F. Mentzer, G. Toderici, M. Tschannen et al. *34th Conference on Neural Information Processing Systems (NeurIPS)*. Vancouver : Curran Associates, 2020. P. 11913–11924. DOI: <https://doi.org/10.48550/arXiv.2006.09965>.

77. High-precision dichotomous image segmentation with frequency and scale awareness / Q. Jiang, J. Cheng, Z. Wu et al. *IEEE Transactions on Neural Networks and Learning Systems*. 2025. Vol. 36, No. 5. P. 8619–8631. DOI: <https://doi.org/10.1109/TNNLS.2024.3426529>.

78. Image Compression Techniques in Wireless Sensor Networks: A Survey and Comparison / B. A. Lungisani, C. K. Lebekwe, A. M. Zungeru et al. *IEEE Access*. 2022. Vol. 10. P. 82511–82530. DOI: <https://doi.org/10.1109/ACCESS.2022.3195891>.

79. Infrared Image Super-Resolution via Transfer Learning and PSRGAN / Y. Huang, Z. Jiang, R. Lan et al. *IEEE Signal Processing Letters*. 2021. Vol. 28. P. 982–986. DOI: <https://doi.org/10.1109/LSP.2021.3077801>.

80. Infrared small and dim target detection with transformer under complex backgrounds / F. Liu, C. Gao, F. Chen et al. *IEEE Transactions on Image Processing*. 2023. Vol. 32. P. 5921–5932. DOI: <https://doi.org/10.1109/TIP.2023.3326396>.

81. InfraredStereo3D: Breaking night vision limits with perspective projection positional encoding and groundbreaking infrared dataset / Y. Niu, Z. Yin, W. He et al. *Remote Sensing*. 2025. Vol. 17, No. 12. Article 2035. DOI: <https://doi.org/10.3390/rs17122035>.

82. Integration of Video Image Decryption Coding into a Remote Video Information Service / A. Krasnorutsky, V. Barannik, S. Shulgin et al. *IEEE 5th International Conference on Advanced Information and Communication Technologies (AICT)*. Lviv : IEEE, 2023. P. 1–6. DOI: <https://doi.org/10.1109/AICT61584.2023.10452679>.

83. Intra prediction and mode coding in VVC / J. Pfaff, T. D. T. Nguyen, H. Schwarz et al. *IEEE Transactions on Circuits and Systems for Video Technology*. 2021. Vol. 31, No. 10. P. 3834–3847. DOI: <https://doi.org/10.1109/TCSVT.2021.3072430>.

84. Iwahashi M. Non Separable Two Dimensional Discrete Wavelet Transform for Image Signals / M. Iwahashi, H. Kiya. *Discrete Wavelet Transforms*. London : IntechOpen, 2013. DOI: <https://doi.org/10.5772/51199>.

85. Jamil S. Review of Image Quality Assessment Methods for Compressed Images / S. Jamil. *Journal of Imaging*. 2024. Vol. 10, No. 5. Article 113. DOI: <https://doi.org/10.3390/jimaging10050113>.

86. JPEG XS – A New Standard for Visually Lossless Low-Latency Lightweight Image Coding / A. Descampe, P. Barten, A. Desmit et al. *Proceedings of the IEEE*. 2021. Vol. 109, No. 9. P. 1559–1577. DOI: <https://doi.org/10.1109/jproc.2021.3080916>.

87. JPEG-XR-GCP: Promoting JPEG-XR Compression by Gradient-Based Coefficient Prediction / Y. Tang, T. Xiang, Y. Yang et al. *12th International*

Conference on Advanced Computational Intelligence (ICACI). Dali : IEEE, 2020. P. 51–58. DOI: <https://doi.org/10.1109/ICACI49185.2020.9177623>.

88. Kountouris M. Semantics-empowered communication for networked intelligent systems / M. Kountouris, N. Pappas. IEEE Communications Magazine. 2021. Vol. 59, No. 6. P. 96–102. DOI: <https://doi.org/10.1109/MCOM.001.2000604>.

89. Kumar S. N. Performance Comparison of Jpeg, Jpeg XT, Jpeg LS, Jpeg 2000, Jpeg XR, HEVC, EVC and VVC for Images / S. N. Kumar, M. V. V. Bharadwaj, S. Subbarayappa. International Conference for Convergence in Technology. Pune : IEEE, 2021. P. 1–8. DOI: <https://doi.org/10.1109/i2ct51068.2021.9418160>.

90. Lakshmi T. N. Performing Image Compression and Decompression Using Matrix Substitution Technique / T. N. Lakshmi, S. Jyothi. Learning and Analytics in Intelligent Systems. Cham : Springer, 2020. Vol. 15. P. 25–36. DOI: https://doi.org/10.1007/978-3-030-46939-9_3.

91. Learning End-to-End Lossy Image Compression: A Benchmark / Y. Hu, W. Yang, Z. Ma et al. IEEE Transactions on Pattern Analysis and Machine Intelligence. 2022. Vol. 44, No. 8. P. 4194–4211. DOI: <https://doi.org/10.1109/TPAMI.2021.3065339>.

92. LeCun Y. Deep learning / Y. LeCun, Y. Bengio, G. Hinton. Nature. 2015. Vol. 521. P. 436–444. DOI: <https://doi.org/10.1038/nature14539>.

93. Lee S. Energy-Constrained Implicit Neural Representations for Deep Image Compression / S. Lee, J.-B. Jeong, E.-S. Ryu. IEEE Signal Processing Letters. 2023. Vol. 30. P. 663–667. DOI: <https://doi.org/10.1109/LSP.2023.3279780>.

94. Lossless medical image compression by using difference transform / R. Rojas-Hernández, J. L. Díaz-de-León-Santiago, G. Barceló-Alonso et al. Entropy. 2022. Vol. 24, No. 7. Article 951. DOI: <https://doi.org/10.3390/e24070951>.

95. Manda M. P. A Fast Image Thresholding Algorithm for Infrared Images Based on Histogram Approximation and Circuit Theory / M. P. Manda, H. S. Kim.

Algorithms. 2020. Vol. 13, No. 9. Article 207. DOI: <https://doi.org/10.3390/a13090207>.

96. Medeiros I. Swarm-Based and Energy-Aware Unmanned Aerial Vehicle System for Video Delivery of Mobile Objects / I. Medeiros, A. Boukerche, E. Cerqueira. IEEE Transactions on Vehicular Technology. 2022. Vol. 71, No. 1. P. 766–779. DOI: <https://doi.org/10.1109/TVT.2021.3126229>.

97. Mentzer F. Learning Better Lossless Compression Using Lossy Compression / F. Mentzer, L. Van Gool, M. Tschannen. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle : IEEE, 2020. P. 6637–6646. DOI: <https://doi.org/10.1109/CVPR42600.2020.00667>.

98. Method of Cluster-Driven Spectral-Parametric Transformant Coding for Adaptive Compression of Segmented Video Data / V. Barannik, A. Berchanov, V. Barannik et al. IEEE 18th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET). Lviv : IEEE, 2026.

99. Method of Coding Subbands of Non-Homogeneous Spectrum of Video Segments in Uneven Diagonal Space / V. Barannik, S. Shulgin, N. Barannik et al. IEEE 4th International Conference on Advanced Trends in Information Theory (ATIT). Kyiv : IEEE, 2022. P. 72–75. DOI: <https://doi.org/10.1109/ATIT58178.2022.10024236>.

100. Method of Coding Video Images Based on Meta-Determination of Segments / V. Barannik, V. Barannik, Y. Babenko et al. Digital Ecosystems: Interconnecting Advanced Networks with AI Applications. TCSET 2024 / eds. A. Luntovskyy et al. Cham : Springer, 2024. P. 566–589. (Lecture Notes in Electrical Engineering ; vol. 1198). DOI: https://doi.org/10.1007/978-3-031-61221-3_27.

101. Method of compression and ensuring the fidelity of video images in infocommunication networks / V. Barannik, A. Krasnorutsky, V. Kolesnyk et al. Radioelectronic and Computer Systems. 2022. Vol. 4. P. 129–142. DOI: <https://doi.org/10.32620/reks.2022.4.10>.

102. Method of Mini Segments Encoding in Difference Space Using Haar Wavelet / V. Barannik, A. Berchanov, V. Barannik et al. IEEE 5th International Conference on Advanced Information and Communication Technologies (AICT). Lviv : IEEE, 2023. P. 1–4. DOI: <https://doi.org/10.1109/AICT61584.2023.10452674>.

103. Method of multilevel spectral processing of infrared images in bit planes / V. Barannik, A. Berchanov, O. Kurlan et al. Cyber Hygiene & Conflict Management in Global Information Networks (CH&CMiGIN'25): Proceedings of the International Workshop (Kyiv, Ukraine, 20–22 June, 2025). 2025. Vol. 4024. P. 227–235. URL: <https://ceur-ws.org/Vol-4024/short07.pdf> (accessed: 30.01.2026).

104. Method of recurrent truncated-positional coding video segments in uneven diagonal space / V. Barannik, S. Shulgin, V. Kozlovskiy et al. Radioelectronic and Computer Systems. 2023. No. 2. P. 129–142. DOI: <https://doi.org/10.32620/reks.2023.2.11>.

105. Method of Structural-Statistical Coding of Video Segments in Spectral-Cluster Space / A. Krasnorutsky, V. Kolesnyk, A. Berchanov et al. IEEE 4th International Conference on Advanced Trends in Information Theory (ATIT). Kyiv : IEEE, 2022. P. 32–37. DOI: <https://doi.org/10.1109/ATIT58178.2022.10024240>.

106. Method Taking into Account Level of Structural and Statistical Saturation of Video Segments in the Coding Process / V. Barannik, Y. Babenko, V. Barannik et al. IEEE 4th International Conference on Advanced Trends in Information Theory (ATIT). Kyiv : IEEE, 2022. P. 66–71. DOI: <https://doi.org/10.1109/ATIT58178.2022.10024193>.

107. Methodological fundamentals of deciphering coding of aerophotography segments on special equipment of unmanned complex / V. Barannik, S. Shulgin, A. Krasnorutsky et al. 2nd International Conference on Advanced Trends in Information Theory (ATIT). Kyiv : IEEE, 2020. P. 38–43. DOI: <https://doi.org/10.1109/ATIT50783.2020.9349257>.

108. MISC: Ultra-low Bitrate Image Semantic Compression Driven by Large Multimodal Model / C. Li, G. Lu, D. Feng et al. *IEEE Transactions on Image Processing*. 2024. Vol. 34. P. 335–349. DOI: <https://doi.org/10.1109/TIP.2024.3515874>.
109. Mishra D. Wavelet-based Deep Auto Encoder-Decoder (WDAED) based Image Compression / D. Mishra, A. V. Paliy, V. V. Barannik. *IEEE Transactions on Circuits and Systems for Video Technology*. 2020. Vol. 31, No. 4. P. 1452–1462. DOI: <https://doi.org/10.1109/TCSVT.2020.3010627>.
110. Modification of a gray-level dynamic range based on a number of binary bit representation for image compression / A. Putra, S. Supriadi, A. Wibawa et al. *Science in Information Technology Letters*. 2020. Vol. 1, No. 1. P. 9–16. DOI: <https://doi.org/10.31763/sitech.v1i1.17>.
111. Neural Image Compression: Generalization, Robustness, and Spectral Biases / K. Lieberman, J. Diffenderfer, C. Godfrey et al. *37th Conference on Neural Information Processing Systems (NeurIPS)*. New Orleans : Curran Associates, 2023. DOI: <https://doi.org/10.48550/arXiv.2307.08657>.
112. Nguyen T. X. B. A review of modern thermal imaging sensor technology and applications for autonomous aerial navigation / T. X. B. Nguyen, K. Rosser, J. Chahl. *Journal of Imaging*. 2021. Vol. 7, No. 10. Article 217. DOI: <https://doi.org/10.3390/jimaging7100217>.
113. Nonlinear Transform Coding / J. Balle, V. Barannik, O. Khimenko et al. *IEEE Journal of Selected Topics in Signal Processing*. 2021. Vol. 15, No. 2. P. 310–321. DOI: <https://doi.org/10.1109/JSTSP.2020.3034501>.
114. Omar H. M. Image Compression using Principal Component Analysis / H. M. Omar, M. Morsli, S. Yaichi. *2nd International Conference on Mathematics and Information Technology (ICMIT)*. Adrar : IEEE, 2020. P. 226–231. DOI: <https://doi.org/10.1109/ICMIT47780.2020.9047014>.
115. On the Computation of PSNR for a Set of Images or Video / O. Keleş, M. A. Yılmaz, A. M. Tekalp et al. *Picture Coding Symposium (PCS)*. Bristol : IEEE, 2021. P. 1–5. DOI: <https://doi.org/10.1109/PCS50896.2021.9477470>.

116. PBT: Progressive background-aware transformer for infrared small target detection / H. Yang, Q. Zhang, J. Wang et al. *IEEE Transactions on Geoscience and Remote Sensing*. 2024. Vol. 62. P. 1–13. DOI: <https://doi.org/10.1109/TGRS.2024.3415080>.

117. Positional coding method in differential wave space / V. Barannik, A. Berchanov, V. Barannik et al. *Informatyka, Automatyka, Pomiary w Gospodarce i Ochronie Środowiska*. 2026. Vol. 16, No. 1. P. 138–146. DOI: <https://doi.org/10.35784/iapgos.7776>.

118. Processing Marker Arrays of Clustered Transformants for Image Segments / V. Barannik, A. Krasnorutsky, V. Barannik et al. *Emerging Networking in the Digital Transformation Age. TCSET 2022* / eds. M. Klymash et al. Cham : Springer, 2023. P. 428–443. (Lecture Notes in Electrical Engineering ; vol. 965). DOI: https://doi.org/10.1007/978-3-031-24963-1_25.

119. Qasim A. Review on techniques and file formats of image compression / A. Qasim, R. Din, F. Q. A. Alyousuf. *Bulletin of Electrical Engineering and Informatics*. 2020. Vol. 9, No. 2. P. 602–610. DOI: <https://doi.org/10.11591/eei.v9i2.2085>.

120. Qi Y. HD video transmission of multi-rotor Unmanned Aerial Vehicle based on 5G cellular communication network / Y. Qi, Y. Jong Hoon. *Computer Communications*. 2020. Vol. 160. P. 688–696. DOI: <https://doi.org/10.1016/j.comcom.2020.07.024>.

121. Qin C. Coordination of drones at scale: Decentralized energy-aware swarm intelligence for spatio-temporal sensing / C. Qin, E. Pournaras. *Transportation Research Part C*. 2023. Vol. 157. Article 104387. DOI: <https://doi.org/10.1016/j.trc.2023.104387>.

122. Rahman M. A. The Impact of State-of-the-Art Techniques for Lossless Still Image Compression / M. A. Rahman, M. Hamada, J. Shin. *Electronics*. 2021. Vol. 10, No. 3. Article 360. DOI: <https://doi.org/10.3390/electronics10030360>.

123. Ranjan R. An Improved Image Compression Algorithm Using 2D DWT and PCA with Canonical Huffman Encoding / R. Ranjan, P. Kumar.

Entropy. 2023. Vol. 25, No. 10. Article 1382. DOI: <https://doi.org/10.3390/e251001382>.

124. Ranjan R. Canonical Huffman Coding Based Image Compression using Wavelet / R. Ranjan. *Wireless Personal Communications*. 2020. Vol. 117, No. 3. P. 2193–2206. DOI: <https://doi.org/10.1007/s11277-020-07967-y>.

125. Rasheed M. H. Image Compression Based on 2D Discrete Fourier Transform and Matrix Minimization Algorithm / M. H. Rasheed, V. V. Barannik, A. V. Paliy. *Array*. 2020. Vol. 6. Article 100024. DOI: <https://doi.org/10.1016/j.array.2020.100024>.

126. Real-time implementation of hybrid visible light/infrared communications supporting full-range dynamic dimming control / S. Liu, X. You, J. Chen et al. *IEEE Photonics Journal*. 2024. Vol. 16, No. 1. P. 1–14. DOI: <https://doi.org/10.1109/JPHOT.2024.3353776>.

127. Reduction of Artefacts in JPEG-XR Compressed Images / K.-L. Hou, M.-S. Li, Y.-C. Lee et al. *Sensors*. 2019. Vol. 19, No. 5. Article 1214. DOI: <https://doi.org/10.3390/s19051214>.

128. Review on Infrared Imaging Technology / F. Hou, Y. Zhang, Y. Zhou et al. *Sustainability*. 2022. Vol. 14, No. 18. Article 11161. DOI: <https://doi.org/10.3390/su141811161>.

129. Romualdas B. Qualitative Rating of Lossy Compression for Aerial Imagery by Neutrosophic WASPAS Method / B. Romualdas, K. Giruta. *Symmetry*. 2021. Vol. 13, No. 2. Article 273. DOI: <https://doi.org/10.3390/sym13020273>.

130. Security key indicators assessment for modern cellular networks / R. Odarchenko, S. Gnatyuk, V. Gnatyuk et al. *IEEE First International Conference on System Analysis & Intelligent Computing (SAIC)*. Kyiv : IEEE, 2018. P. 1–7. DOI: <https://doi.org/10.1109/SAIC.2018.8516889>.

131. SFHN: Spatial-frequency domain hybrid network for image super-resolution / Z. Wu, W. Liu, J. Li et al. *IEEE Transactions on Circuits and Systems*

for Video Technology. 2023. Vol. 33, No. 11. P. 6459–6473. DOI: <https://doi.org/10.1109/TCSVT.2023.3271131>.

132. Shah T. J. A Review of Contemporary Image Compression Techniques and Standards / T. J. Shah, M. T. Banday. Examining Fractal Image Processing and Analysis / eds. S. R. Nayak, J. Mishra. Hershey : IGI Global, 2020. P. 121–157. DOI: <https://doi.org/10.4018/978-1-7998-0066-8.ch006>.

133. Shihab H. S. Image Compression Techniques on-Board Small Satellites / H. S. Shihab. Iraqi Journal of Science. 2023. Vol. 64, No. 3. P. 1518–1534. DOI: <https://doi.org/10.24996/ijs.2023.64.3.40>.

134. Shirvaikar M. V. A comparative study of the AV1, HEVC, and VVC video codecs / M. V. Shirvaikar, C. Grecos. Real-Time Image Processing and Deep Learning 2025. Orlando : SPIE, 2025. Vol. 13458. P. 12. DOI: <https://doi.org/10.1117/12.3053634>.

135. Siddeq M. M. A Novel Method for Image and Video Compression Based on Two-Level DCT with Hexadata Coding / M. M. Siddeq, M. A. Rodrigues. Sensing and Imaging. 2020. Vol. 21. Article 36. DOI: <https://doi.org/10.1007/s11220-020-00302-6>.

136. Significant Microsegment Transformants Encoding Method to Increase the Availability of Video Information Resource / V. Barannik, Yu. Babenko, O. Kulitsa et al. 2nd International Conference on Advanced Trends in Information Theory (ATIT). Kyiv : IEEE, 2020. P. 52–56. DOI: <https://doi.org/10.1109/ATIT50783.2020.9349256>.

137. Son D.-M. Visible and Near Infrared Image Fusion Using Base Tone Compression and Detail Transform Fusion / D.-M. Son, H.-J. Kwon, S.-H. Lee. Chemosensors. 2022. Vol. 10, No. 4. Article 124. DOI: <https://doi.org/10.3390/chemosensors10040124>.

138. Tahar B. An efficient JPEG-2000 based multimodal compression scheme / B. Tahar, K. Fouad, K. Abdellah. Multimedia Tools and Applications. 2021. Vol. 80. P. 21241–21260. DOI: <https://doi.org/10.1007/s11042-021-10776-5>.

139. Task-Oriented Image Transmission for Scene Classification in Unmanned Aerial Systems / K. Xu, S. Bin, G. Jie et al. *IEEE Transactions on Communications*. 2022. Vol. 70, No. 8. P. 5181–5192. DOI: <https://doi.org/10.1109/TCOMM.2022.3182325>.

140. The Application of Discrete Wavelet Transform for Digital Image Compression / K. Ahmad, T. Pukky, A. Aris et al. *Jurnal Matematika, Sains, dan Teknologi*. 2024. Vol. 25, No. 1. P. 1–8. DOI: <https://doi.org/10.33830/jmst.v25i1.3955.2024>.

141. The Applications of Discrete Wavelet Transform in Image Processing / M. Adnan, Q. Diyar, A. Diloan et al. *Journal of Soft Computing and Data Mining*. 2020. Vol. 1, No. 2. P. 31–43. DOI: <https://doi.org/10.30880/jscdm.2020.01.02.004>.

142. The method of masking overhead compaction in video compression systems / V. Barannik, S. Sidchenko, N. Barannik et al. *Radioelectronic and Computer Systems*. 2021. No. 2. P. 51–63. DOI: <https://doi.org/10.32620/reks.2021.2.05>.

143. The Methods of Intellectual Processing of Video Frames in Coding Systems in Progress Aeromonitor to Increase Efficiency and Semantic Integrity / A. Krasnorutsky, R. Onyshchenko, D. Barannik et al. *IEEE 4th International Conference on Advanced Trends in Information Theory (ATIT)*. Kyiv : IEEE, 2022. P. 53–56. DOI: <https://doi.org/10.1109/ATIT58178.2022.10024208>.

144. The technology of the video stream intensity controlling based on the bit-planes recombination / V. Barannik, M. Karpinski, V. Tverdokhleba et al. *4th International Symposium on Wireless Systems within the International Conferences on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS-SWS)*. Lviv : IEEE, 2018. P. 25–28. DOI: <https://doi.org/10.1109/IDAACS-SWS.2018.8525560>.

145. The video information resources integrity concept by using binomial slots / A. Alimpiev, V. Barannik, S. Podlesny et al. *XIIIth International Conference on Perspective Technologies and Methods in MEMS Design (MEMSTECH)*. Lviv

: IEEE, 2017. P. 193–196. DOI:
<https://doi.org/10.1109/MEMSTECH.2017.7937564>.

146. Tree-Encoded Bitmaps / H. Lang, A. Beischl, V. Leis et al. Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data. Portland : ACM, 2020. P. 937–967. DOI:
<https://doi.org/10.1145/3318464.3380588>.

147. Ungureanu V.-I. Image-Compression Techniques: Classical and "Region-of-Interest-Based" Approaches Presented in Recent Papers / V.-I. Ungureanu, P. Negirla, A. Korodi. Sensors. 2024. Vol. 24, No. 3. Article 791. DOI: <https://doi.org/10.3390/s24030791>.

148. Usage of IR sensors in the HVAC systems, vehicle and manufacturing industries: a review / M. A. Altaf, J. Ahn, D. Khan et al. IEEE Sensors Journal. 2022. Vol. 22, No. 10. P. 9164–9176. DOI:
<https://doi.org/10.1109/JSEN.2022.3166190>.

149. Uthayakumar J. Highly Reliable and Low-Complexity Image Compression Scheme Using Neighborhood Correlation Sequence Algorithm in WSN / J. Uthayakumar, M. Elhoseny, K. Shankar. IEEE Transactions on Reliability. 2020. Vol. 69, No. 4. P. 1398–1423. DOI:
<https://doi.org/10.1109/TR.2020.2972567>.

150. Variational image compression with a scale hyperprior / J. Balle, D. Minnen, S. Singh et al. International Conference on Learning Representations (ICLR). Vancouver : OpenReview.net, 2018. DOI:
<https://doi.org/10.48550/arXiv.1802.01436>.

151. Viability of Image Compression in Vibrothermography / X. Chi, Y. Zhang, D. Di Maio et al. Experimental Techniques. 2021. Vol. 45. P. 345–362. DOI: <https://doi.org/10.1007/s40799-020-00395-4>.

152. Video compression using generalized binary partitioning, trellis coded quantization, perceptually optimized encoding, and advanced prediction and transform coding / J. Pfaff, H. Schwarz, T. D. T. Nguyen et al. IEEE Transactions

on Circuits and Systems for Video Technology. 2020. Vol. 30, No. 5. P. 1281–1295. DOI: <https://doi.org/10.1109/TCSVT.2019.2945918>.

153. VIF-Net: An Unsupervised Framework for Infrared and Visible Image Fusion / R. Hou, D. Zhou, R. Nie et al. IEEE Transactions on Computational Imaging. 2020. Vol. 6. P. 640–651. DOI: <https://doi.org/10.1109/TCI.2020.2965304>.

154. Wang J. Learned image compression with adaptive channel and window-based spatial entropy models / J. Wang, Q. Ling. IEEE Transactions on Consumer Electronics. 2024. Vol. 70, No. 4. P. 6430–6441. DOI: <https://doi.org/10.1109/TCE.2024.3485179>.

155. Xin G. Soft compression for lossless image coding based on shape recognition / G. Xin, P. Fan. Entropy. 2021. Vol. 23, No. 12. Article 1680. DOI: <https://doi.org/10.3390/e23121680>.

156. Xue R. Information transmission schemes based on adaptive coded modulation for UAV surveillance systems with satellite relays / R. Xue, M. Zhao, H. Tang. IEEE Access. 2020. Vol. 8. P. 191355–191364. DOI: <https://doi.org/10.1109/ACCESS.2020.3027737>.

157. Yang B. Evaluation of low-cost smartphone-based infrared cameras to assess the cooling and refrigerated storage temperatures of fresh produce / B. Yang, G. Dev Kumar, K. Mis Solval. Foods. 2022. Vol. 11, No. 21. Article 3440. DOI: <https://doi.org/10.3390/foods11213440>.

158. Yang S.-H. Quality Control for Hybrid Unicast and Multicast Video Transmission Systems / S.-H. Yang, T.-W. Liu. IEEE International Conference on Consumer Electronics – Taiwan (ICCE-Taiwan). Taoyuan : IEEE, 2020. P. 1–2. DOI: <https://doi.org/10.1109/ICCE-Taiwan49838.2020.9258044>.

159. Yang Y. Advances in Diffusion-Based Generative Compression / Y. Yang, S. Mandt. ArXiv. 2026. DOI: <https://doi.org/10.48550/arXiv.2601.18932>.

160. YOLO-FIRI: Improved YOLOv5 for Infrared Image Object Detection / S. Li, Y. Li, Y. Li et al. IEEE Access. 2021. Vol. 9. P. 141861–141875. DOI: <https://doi.org/10.1109/ACCESS.2021.3120870>.

161. Yousif R. I. Image Compression Based on Arithmetic Coding Algorithm / R. I. Yousif, N. H. Salman. Iraqi Journal of Science. 2021. Vol. 62, No. 1. P. 329–334. DOI: <https://doi.org/10.24996/ij.s.2021.62.1.31>.

162. Zhang L. Enhanced High-Definition Video Transmission for Unmanned Driving in Mining Environments / L. Zhang, W. Yang. Applied Sciences. 2024. Vol. 14, No. 10. Article 4296. DOI: <https://doi.org/10.3390/app14104296>.

163. Zhang X. Learning Optimal Lattice Vector Quantizers for End-to-end Neural Image Compression / X. Zhang, X. Wu. ArXiv. 2024. DOI: <https://doi.org/10.48550/arXiv.2411.16119>.

Додаток А

Список публікацій здобувача за темою дисертації та відомості про апробацію результатів дисертації

Список публікацій здобувача, в яких опубліковані основні наукові результати дисертації:

Статті в журналах Scopus/Web of Science:

1. Barannik V., Berchanov A., Barannik V., Dihtiar M., Osovytskyi M., Sushko A., Babenko Y. Positional coding method in differential wave space. *Informatyka, Automatyka, Pomiaru w Gospodarce i Ochronie Środowiska*. 2026. Vol. 16. No 1. P. 138–146. DOI: <https://doi.org/10.35784/iapgos.7776>. **(Scopus, Q4)**

2. Бараннік В. В., Берчанов А. А., Бараннік В. В., Суханов О. Ю., Перцев П. Д., Лиходєєв О. С., Юдін О. К., Прокопенко Р. О., Чаун Ю. М. Метод семантично-орієнтованої обробки бітових площин інфрачервоних зображень. *Вісник НТУУ "КПІ". Серія Радіотехніка, Радіоапаратобудування*. 2025. № 102. С. 15–24. DOI: <https://doi.org/10.20535/RADAP.2025.102.15-24>. URL: <https://radap.kpi.ua/radiotechnique/article/view/2084> (дата звернення: 30.01.2026). **(Web of Science, Q4)**

3. Бараннік В. В., Берчанов А. А., Бараннік В. В., Шейгас О. М., Перцев П. Д., Вовк О. В., Бабенко Ю. М., Прокопенко Р. О., Ерошенко В. П. Метод групового кодування інфрачервоних зображень у спектрально-хвильовому просторі. *Вісник НТУУ "КПІ". Серія Радіотехніка, Радіоапаратобудування*. 2025. № 99. С. 24–34. DOI: <https://doi.org/10.20535/radap.2025.99.24-34>. **(Web of Science, Q4)**

Статті в журналах категорії Б:

4. Бараннік В. В., Берчанов А. А. Метод біфуркації бітових площин в системах обробки інфрачервоних зображень. *Наукоємні технології*. 2025. Т.

66, № 2. С. 208–213. DOI: <https://doi.org/10.18372/2310-5461.66.20325>.

(Категорія Б)

5. Бараннік В., Берчанов А. Метод спектрально-групового кодування бітових шарів інфрачервоних зображень. *Інфокомунікаційні технології та електронна інженерія*. 2025. Вип. 5, № 2. С. 15–24. DOI: <https://doi.org/10.23939/ict2025.02.015>. **(Категорія Б)**

6. Берчанов А. А., Бараннік В. В. Метод обробки інфрачервоних зображень у спектральному домені. *Вчені записки Таврійського національного університету імені В. І. Вернадського. Серія: Технічні науки*. 2025. Т. 36 (75), № 3. С. 55–60. DOI: <https://doi.org/10.32782/2663-5941/2025.3.2/08>. **(Категорія Б)**

Список публікацій, які засвідчують апробацію матеріалів дисертації:

7. Barannik V., Eliseev E., Barannik V., Berchanov A., Lenets V., Fedorovskiy O. Method of Cluster-Driven Spectral-Parametric Transformant Coding for Adaptive Compression of Segmented Video Data. *2026 IEEE 18th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET)* (Lviv, Ukraine, 17–21 Feb., 2026). Lviv, 2026. **(Scopus)**

8. Barannik V., Berchanov A., Kurlan O., Osovytskyi M., Pertsev P., Prokopenko R. Method of multilevel spectral processing of infrared images in bit planes. *Cyber Hygiene & Conflict Management in Global Information Networks (CH&CMiGIN'25): Proceedings of the International Workshop* (Kyiv, Ukraine, 20–22 June, 2025). 2025. Vol. 4024. P. 227–235. URL: <https://ceur-ws.org/Vol-4024/short07.pdf> (дата звернення: 30.01.2026). **(Scopus)**

9. Бараннік В. В., Берчанов А. А., Бараннік В. В., Перцев П. Д., Курлан О. О. Підвищення ефективності доведення інфрачервоних знімків методом спектрально-хвильового кодування. *Кіберборотьба: розвідка, захист та*

проти́дія: матеріали III Міжнар. наук.-практ. конф. (Київ, 6–7 трав. 2025 р.). Київ, 2025. С. 14–15. DOI: <https://doi.org/10.61929/viti.mnpg.3.2025>.

10. Kurlan O., Prokopenko R., Pertsev P., Manakov V., Berchanov A., Shchyrba V. A Method of Effective Coding of a Dynamic Video Information Resource in an Information and Telecommunications Network. *2024 IEEE 5th International Conference on Advanced Trends in Information Theory (ATIT)* (Lviv, Ukraine, 21–23 Nov., 2024). Lviv, 2024. P. 51–57. DOI: <https://doi.org/10.1109/ATIT64324.2024.11222407>. **(Scopus)**

11. Берчанов А. А., Бараннік В. В., Прокопенко Р. О., Бараннік Д. В., Перцев П. Д. Інтелектуалізовані методи обробки тепловізійних зображень. *Комп'ютерне моделювання в наукоємних технологіях (КМНТ-2024)*: матеріали 10-ї Міжнар. наук.-техн. конф. (Харків, 27–29 листоп. 2024 р.). Харків, 2024. С. 24–26. URL: <https://csai.karazin.ua/news/uk/x-naukovo-tehnichna-mizhnarodna-28052025> (дата звернення: 05.02.2026).

12. Barannik V., Berchanov A., Barannik V., Babenko M., Onyshchenko R., Kolodiichuk L. Method of Mini Segments Encoding in Difference Space Using Haar Wavelet. *2023 IEEE 5th International Conference on Advanced Information and Communication Technologies (AICT)* (Lviv, Ukraine, 21–25 Nov., 2023). Lviv, 2023. P. 1–4. DOI: <https://doi.org/10.1109/AICT61584.2023.10452674>. **(Scopus)**

13. Бараннік В. В., Берчанов А. А. Перспективи розвитку технологій обробки відеоінформації в умовах інформаційно-інтелектуального протиборства. *Наукоємні технології в інфокомунікаціях (НІСТ'2023)*: матеріали Міжнар. наук.-практ. конф. (Кам'янець-Подільський, 1–3 черв. 2023 р.). Кам'янець-Подільський, 2023. URL: <http://repositsc.nuczu.edu.ua/bitstream/123456789/18013/1/%D0%9D%D0%86%D0%A1%D0%A2%202023.pdf> (дата звернення: 30.01.2026).

14. Krasnorutsky A., Kolesnyk V., Berchanov A., Barannik V., Kharchenko N., Malko O. Method of Structural-Statistical Coding of Video Segments in Spectral-Cluster Space. *2022 IEEE 4th International Conference on Advanced*

Trends in Information Theory (ATIT) (Kyiv, Ukraine, 15–17 Dec., 2022). Kyiv, 2022. P. 32–37. DOI: <https://doi.org/10.1109/ATIT58178.2022.10024240>.
(Scopus)

Додаток Б

Акти реалізації науково-прикладних результатів досліджень

ЗАТВЕРДЖУЮ

В.о. головного інженера

ТОВ "Одеський авіаційний завод"

Микола ГОНЧАРЕНКО

_____ 2025 року

А К Т

впровадження результатів науково-прикладних досліджень
Берчанова Анатолія Андрійовича

Комісія у складі: В.о. головного технолога Олега Маревського, провідного інженера технолога Юрія Ведмідя, провідного інженера Анатолія Якимчука склала дійсний акт, який полягає в тому, що при виконанні дослідно-конструкторських робіт використані наступні результати науково-прикладних досліджень Берчанова Анатолія Андрійовича:

Метод семантично-орієнтованої обробки із застосуванням бітової біфуркації інфрачервоних зображень, що заснований на спектральному аналізі інформативності бітових площин, дозволяє локалізувати семантично значущі сегменти зображення за допомогою структурного розділення старших та молодших бітів, з подальшою їх адаптивною компресією відповідно до виявленого рівня теплової варіативності. З метою підвищення оперативності обміну інформацією. Метод включає:

1) оцінювання інформативності мінісегментів на основі спектральних характеристик, зокрема енергії та ентропії вейвлет-компонент, з метою класифікації областей зображення на семантично важливі та фонові для подальшої адаптивної компресії;

2) бітову біфуркацію, яка передбачає поділ інфрачервоного зображення із широким динамічним діапазоном на окремі бітові площини старших та молодших бітових значень для подальшої диференційованої обробки з урахуванням їхньої семантичної значущості;

3) адаптивне кодування, яке полягає у виборі стратегії стиснення залежно від інформативності сегментів: для важливих областей використовується компресія з мінімальними втратами, для малозначущих — більш агресивне стиснення.

Впровадження результатів досліджень Берчанова А.А. в автоматизованій системі відображення інформації на основі програмно-апаратних реалізацій дозволило забезпечити:

- більшу точність при відновленні теплової структури зображень, що підтверджується отриманим значенням PSNR на рівні 32 дБ. Це на 12 % перевищує показники, притаманні існуючим методам обробки інфрачервоних даних і дозволяє зберегти температурну інформацію об'єктів інтересу;

- покращення результатів у задачах зменшення обсягу даних. У ході експериментальних випробувань було отримано коефіцієнт стиснення, що дорівнює 9.16, що перевищує показники актуальних алгоритмічних рішень приблизно на 7 % за умови збереження однакового загального рівня точності відновлення зображень.

Голова комісії



Олег МАРЕВСЬКИЙ

Члени комісії:



Юрій ВЕДМІДЬ

Анатолій ЯКИМЧУК

ЗАТВЕРДЖУЮ

Директор ТОВ «ОБІД ДІФЕНС»

Дмитро ЛІСОЦЬКИЙ

« _____ » _____ 2025 р.

**А К Т**

впровадження результатів науково-прикладних досліджень

Берчанова Анатолія Андрійовича

Комісія у складі: голови начальника конструкторського відділу-керівника конструкторської групи Кукобко Сергія, членів Купрія Олександра, Кравідза Романа склала дійсний акт. Він стосується використання в підсистемах обробки та передачі інформації результатів науково-прикладних досліджень Берчанова Анатолія Андрійовича, а саме:

Метод спектрально-групового кодування бітових шарів інфрачервоних зображень, що базується на поетапному багаторівневому кодуванні зображень з високою розрядністю з метою зниження обсягу переданих даних при збереженні термальної семантики. Метод включає: представлення 16-бітного інфрачервоного зображення у вигляді окремих старших та молодших бітових шарів, що дозволяє здійснювати диференційовану обробку з урахуванням семантичної значущості піксельних рівнів; перетворення у різницевий простір, що забезпечує усунення постійної складової, зниження ентропії та концентрацію градієнтної інформації; групове кодування, що передбачає стиснення: кожного набору високочастотних коефіцієнтів мінісегменту до одного групового коду; низькочастотних коефіцієнтів усіх мінісегментів сегменту до одного зведеного коду.

Впровадження отриманих Берчановим Анатолієм Андрійовичем наукових результатів на дослідних макетах підсистем обробки та передачі інформації дозволяє забезпечити:

- збільшення точності реконструкції термального зображення, що підтверджується значенням пікового відношення сигнал/шум (PSNR), яке становить 38 дБ, що є на 9% вищим, ніж у сучасних відомих підходів до стиснення інфрачервоних зображень. Такий рівень достовірності забезпечує надійне збереження температурної семантики сцен, особливо в умовах оперативного моніторингу об'єктів у складному фоні чи при маскуванні.

- підвищення ефективності стиснення даних. Експериментально отримано коефіцієнт стиснення 8.63, що є на 6% вищим, ніж у сучасних алгоритмах при аналогічному рівні якості реконструкції.

Голова комісії

Члени комісії:



Кукобко Сергій

Кравідз Роман

Купрій Олександр

Додаток В

Науково-дослідна робота

Додаток 6

АНОТОВАНИЙ ЗВІТ

**перехідної науково-дослідної роботи,
яка виконувалась в межах кафедральної тематики
за 2026 рік**

1. **Номер державної реєстрації НДР:** 0125U004039
2. **Назва теми НДР:** «Розробка математичних моделей обробки видової інформації для підвищення якості функціонування дистанційних сервісів»
3. **Вид дослідження:** фундаментальна чи прикладна (потрібне вибрати).
4. **Науковий керівник НДР:** Бараннік Володимир Вікторович
5. **Номер телефону керівника НДР:** +380 63 521 90 89
6. **E-mail керівника НДР:** vvbar.off@gmail.com
7. **Терміни виконання етапу:** початок - 03.12.2025 , закінчення - 28.02.2026 .
8. **Науковий результат** (отримані наукові і (або) науково-технічні результати, їх наукова новизна, переваги над вітчизняними або зарубіжними аналогами) _____
Було розроблено метод семантично-орієнтованої обробки із застосуванням бітової біфуркації інфрачервоних зображень, що заснований на спектральному аналізі інформативності бітових площин, дозволяє локалізувати семантично значущі сегменти зображення за допомогою структурного розділення старших та молодших бітів, з подальшою їх адаптивною компресією відповідно до виявленого рівня теплової варіативності. В ході експериментів було одержане покращення результатів у задачах зменшення обсягу даних. У ході експериментальних випробувань було отримано коефіцієнт стиснення, що дорівнює 9.16, що перевищує показники актуальних алгоритмічних рішень приблизно на 7% за умови збереження однакового загального рівня точності відновлення зображень.
Практична цінність (короткий опис можливого практичного застосування отриманого результату) Розроблений метод може бути застосований для удосконалення інформаційних систем обробки видової інформації. Метод дозволяє зменшити бітрейт потоку даних при збереженні діагностичної точності теплових сигнатур об'єктів інтересу.

9. Впровадження/використання результатів у навчальному процесі.

Теоретичні положення та алгоритми роботи можуть бути впроваджені в навчальний процес з таких дисциплін: цифрова обробка зображень, теорія інформації та кодування, методи штучного інтелекту в інформаційних системах.

10. Бібліографічний перелік монографій, підручників, посібників, наукових статей, інших публікацій, дисертацій, магістерських робіт, дипломних проектів.

1. Barannik V., Berchanov A., Barannik V., Dihtiar M., Osovyytskiy M., Sushko A., Babenko Y. Positional coding method in differential wave space. *Informatyka, Automatyka, Pomiarы w Gospodarce i Ochronie Środowiska*. 2026. Vol. 16. No 1. P. 138–146. DOI: <https://doi.org/10.35784/iapgos.7776> (Scopus).
2. Бараннік В. В., Берчанов А. А., Бараннік В. В., Суханов О. Ю., Перцев П. Д., Лиходєєв О. С., Юдін О. К., Прокопенко Р. О., Чаун Ю. М. Метод семантично-орієнтованої обробки


Додаток 6

бітових площин інфрачервоних зображень. Вісник НТУУ "КПІ". Серія Радіотехніка, Радіоапаратобудування. 2025. № 102. С. 15–24. DOI: <https://doi.org/10.20535/RADAP.2025.102.15-24> (Web of Science).

3. Бараннік В. В., Берчанов А. А. Метод біфуркації бітових площин в системах обробки інфрачервоних зображень. Наукоємні технології. 2025. Т. 66, № 2. С. 208–213. DOI: <https://doi.org/10.18372/2310-5461.66.20325>.

11. Рішення вченої ради факультету від 17 березня 2026 р. протокол № 8 про виконання роботи.

Зав. кафедрою


(підпис)к.т.н., доцент, Струков В. М.
(вч. звання, наук. ст., ПІБ)

Науковий керівник


(підпис)д.т.н., професор, Бараннік В. В.
(вч. звання, наук. ст., ПІБ)

Онлайн сервіс створення та перевірки кваліфікованого та удосконаленого електронного підпису

ПРОТОКОЛ

створення та перевірки кваліфікованого та удосконаленого електронного підпису

Дата та час: 19:15:12 06.05.2026

Назва файлу з підписом: Berchanov_diss.pdf.asice

Розмір файлу з підписом: 5.8 МБ

Перевірені файли:

Назва файлу без підпису: Berchanov_diss.pdf

Розмір файлу без підпису: 6.9 МБ

Результат перевірки підпису: Підпис створено та перевірено успішно. Цілісність даних підтверджено

Підписувач: Берчанов Анатолій Андрійович

П.І.Б.: Берчанов Анатолій Андрійович

Країна: Україна

РНОКПП: 3578403475

Час підпису (підтверджено кваліфікованою позначкою часу для підпису від Надавача): 19:15:04 06.05.2026

Сертифікат виданий: "Дія". Кваліфікований надавач електронних довірчих послуг

Серійний номер: 514B5C86A1E5DA110400000028A80800DBB13A05

Тип носія особистого ключа: ЗНКІ криптомодуль ІІТ Гряда-301

Алгоритм підпису: ДСТУ 4145

Тип підпису: Кваліфікований

Тип контейнера: Підпис та дані в архіві (розширений) (ASiC-E)

Формат підпису: З повними даними для перевірки (XAdES-B-LT)

Сертифікат: Кваліфікований

Версія від: 2026.04.06 13:00