

Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»
Міністерство освіти і науки України

Національний аерокосмічний університет ім. М. Є. Жуковського
«Харківський авіаційний інститут»
Міністерство освіти і науки України

Кваліфікаційна наукова
праця на правах рукопису

КОВАЛЕНКО БОГДАН ВІТАЛІЙОВИЧ

УДК 004.932

ДИСЕРТАЦІЯ

**БАГАТОЕТАПНА ОБРОБКА ТА СТИСНЕННЯ ЗОБРАЖЕНЬ ВРС
КОДЕРОМ З ПРОГНОЗУВАННЯМ ПАРАМЕТРІВ**

172 Телекомунікації та радіотехніка
(шифр і назва спеціальності)

Електроніка та телекомунікації
(галузь знань)

Подається на здобуття наукового ступеня доктора філософії

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

_____ Б. В. Коваленко
(підпис)

Науковий керівник

Лукін Володимир Васильович, доктор технічних наук, професор.

Харків – 2025

АНОТАЦІЯ

Коваленко Богдан Віталійович. Багатоетапна обробка та стиснення зображень BPG кодером з прогнозуванням параметрів. – Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття наукового ступеня доктора філософії з галузі знань 17 Електроніка та телекомунікації за спеціальністю 172 «Телекомунікації та радіотехніка». – Національний аерокосмічний університет ім. М. Є. Жуковського «Харківський авіаційний інститут», Харків, 2025.

Дисертаційна робота присвячена удосконаленню та розробці нових ефективних методів стиснення з втратами та надання можливості прогнозувати параметри якості одно- та триканальних зображень різного походження. Об'єктом дослідження є процес стиснення з втратами та спотворення, які вносить даний тип стиснення. Предметом дослідження є методи прогнозування існування оптимальної робочої точки (ОРТ) та параметрів стисненого зображення в ОРТ та її околах для BPG (better portable graphics) кодера.

У роботі було проведено аналітичний аналіз основних додатків багатоетапної обробки зображень, включаючи етапи формування, пост-фільтрації та аналізу. Окрему увагу було приділено стисненню зображень з втратами. Були розглянуті основні вимоги, які висуваються до методів стиснення, як-то забезпечення досить великого коефіцієнта стиснення (КС), а також задовільно високої якості стиснених зображень, оскільки надалі вони підлягають подальшому аналізу та класифікації. Як результат виникає проблема з вибором оптимального значення параметру, що контролює стиснення (ПКС) для обраного кодера для забезпечення необхідного компромісу між КС і якістю стисненого зображення. Встановлено, що це завдання ускладнюється присутністю шуму на зображенні, що може значно вплинути на характеристики стиснення. Показано, що в умовах присутності шуму можна спостерігати оптимальну робочу точку, яка може забезпечити високу якість зображення із забезпеченням досить великого КС. Ця точка може спостерігатися для

різних значень ПКС залежно від характеристик зображень, шуму, а також попередньої обробки зображень (зокрема використання варіаційно-стабілізаційного перетворення). Виявлено, що наявні методи стиснення поступаються новим рішенням, зокрема BPG кодеру, який може забезпечити високий коефіцієнт стиснення, при цьому зберігши хорошу якість зображення. Показано необхідність у розробленні та вдосконаленні методів прогнозування оптимальної робочої точки для забезпечення ефективного стиснення з втратами в умовах ураження зображення шумом.

За результатами проведеного аналізу у роботі поставлено та вирішено **актуальне науково-прикладне завдання** розроблення методів обробки та стиснення з втратами та надання можливості прогнозувати параметри якості зображення. Досліджено характеристики BPG кодера під час його роботи з напівтоновими та кольоровими зображеннями у випадках наявності шуму на зображенні, продемонстровано можливість присутності оптимальної робочої точки (ОРТ) для стиснення BPG-кодером напівтонових та кольорових або інших триканальних зображень у випадках ураження цих зображень як адитивним, так і сигнально-залежним шумами (у випадках використання попередньої обробки зображень та без), розроблено метод прогнозування існування ОРТ в застосуванні до BPG-кодера для напівтонових і кольорових зображень, спотворених шумом, розроблено метод прогнозування середньоквадратичної похибки (СКП) при роботі з BPG-кодером.

Для вирішення наукових завдань використовувалися наступні методи: математичне та чисельне моделювання, методи машинного навчання, методи теорії ймовірностей і математичної статистики.

Мета дисертаційної роботи полягає в розробці ефективних методів обробки та стиснення з втратами та надання можливості прогнозувати параметри якості зображення.

Відповідно до поставленої мети в дисертаційній роботі сформульовано і вирішено такі основні завдання:

- аналіз особливостей і вимог до стиснення з втратами для BPG кодера;
- дослідження характеристик стиснення з втратами зображень, що спотворені

завадами, для BPG кодера;

- розробка методу прогнозування робочої точки для BPG кодера, а також характеристик зображень у цій точці та її околі;

- розроблення методу прогнозування характеристик спотворень під час стиснення з втратами для BPG кодера у застосуванні до зображень без завад;

Вперше досліджено характеристики BPG кодера під час його роботи в різних умовах, таких як робота з напівтоновими та кольоровими зображеннями у випадках наявності шуму на зображенні. Це дозволило сформулювати та надати рекомендації щодо стиснення зображень з втратами з використанням BPG кодера.

Отримало подальший розвиток використання оптимальної робочої точки (ОРТ) для різних кодерів, а саме вперше продемонстровано можливість присутності такої точки для стиснення BPG-кодером напівтонових та кольорових зображень у випадках ураження цих зображень як адитивним, так і сигнально-залежним шумами, показана залежність між положенням цієї точки та характеристиками зображення, шума чи методів обробки. Це дало змогу розрахувати ПКС для досягнення максимального ефекту фільтрації (максимальної якості стисненого зображення у відповідності до різних метрик).

Вдосконалено метод прогнозування існування ОРТ в застосуванні до BPG-кодера для напівтонових і кольорових зображень, спотворених шумом; основною різницею з попередніми методами є те, що запропонований метод вперше використовується для BPG-кодера і бере до уваги особливості стиснення для цього кодера; це дало змогу надати обґрунтовані рекомендації з вибору параметра Q для досягнення компромісу між якістю стиснутих зображень і коефіцієнтом стиснення.

Вдосконалено метод прогнозування середньоквадратичної похибки (СКП); головною різницею є попередній аналіз вхідних параметрів, що характеризують складність зображення, результатом цього аналізу став список параметрів, використання яких може забезпечити високу точність прогнозування, що дало змогу підвищити точність прогнозування без використання декількох ітерацій стиснення.

Практичне значення отриманих результатів полягає в тому, що: підтверджено існування ОРТ для BPG кодера як для напівтонових, так і для кольорових зображень,

у разі їх ураження адитивними або сигнально-залежними типами завад (розглянуто можливість використання варіаційно-стабілізаційного перетворення для попередньої обробки зображень), це дозволяє підвищити якість стисненого зображення у порівнянні з первинним до кількох дБ (згідно PSNR) і при цьому отримати задовільний коефіцієнт стиснення (від 12 до 51 в залежності від характеристик зображення), якщо стиснення буде проводитися в околі ОПТ; надано інформацію про поведінку спотворень, які вносить VPG кодер під час стиснення з втратами, а також, як ці спотворення ведуть себе в ситуації, коли на зображенні присутній шум, завдяки цьому отримана можливість враховувати особливості завад у подальших процедурах фільтрації або детектування об'єктів, що призводить до покращення ефективності цих процедур; запропоновано метод прогнозування існування робочої точки та параметрів у ній для VPG кодера, точність прогнозування становить 0,84-0,97 для різних метрик якості за критерієм R^2 , метод також може працювати з кольоровими зображеннями та кількома типами шуму (адитивний і сигнально-залежний), таке прогнозування дає змогу надати рекомендації щодо вставлення ПКС, і, як результат, отримати краще співвідношення між якістю стисненого зображення та отриманим КС; запропоновано метод прогнозування середньоквадратичної похибки спотворень для стиснення з втратами за допомогою VPG кодера, що враховує складність зображень, це дає змогу виставити необхідний параметр стиснення без необхідності проводити кілька ітерацій на стиснення з метою пошуку оптимального значення параметра та в разі прискорити стиснення з бажаною якістю.

Усі теоретичні розробки дисертації доведено до конкретних інженерних методик і алгоритмів у застосуванні до обробки та стиснення з втратами і прогнозування ефективності стиснення зображень.

Розроблені методи стиснення з втратами і прогнозування ефективності обробки та стиснення зображень були використані при виконанні науково-дослідної роботи у Інституті радіофізики і електроніки ім. О. Я. Усикова НАН України, а також впроваджені у навчальному процесі Національного аерокосмічного університету ім. М. Є. Жуковського "Харківський авіаційний інститут".

За матеріалами дисертації **опубліковано** 19 робот, з яких 6 – це статті у наукових фахових виданнях України, 1 з яких індексується в SCOPUS, 2 статті у наукових періодичних виданнях іншої держави (Швейцарія), які індексується в SCOPUS (обидві Q1), 7 – публікації у матеріалах конференцій, серед яких 6 у матеріалах міжнародних англомовних конференцій, 4 з яких включені до міжнародної наукометричної бази SCOPUS, та 4 – глави в книгах у наукових періодичних виданнях іншої держави.

Ключові слова: стиснення з втратами, багатоетапна обробка зображень, better portable graphics (BPG), оптимальна робоча точка (OPT), прогнозування OPT, зображення уражені шумом, прогнозування характеристик спотворень.

Список публікацій здобувача за темою дисертації:

Статті у наукових періодичних виданнях:

1. B. Kovalenko, V. Lukin, B. Vozel, “BPG-Based Lossy Compression of Three-channel Noisy Images with Prediction of Optimal Operation Existence and Its Parameters,” *Remote Sensing*, vol. 15, no. 6: 1669, 2023, doi: 10.3390/rs15061669 (Scopus).
2. B. Kovalenko, V. Lukin, S. Kryvenko, V. Naumenko, B. Vozel, “BPG-Based Automatic Lossy Compression of Noisy Images with the Prediction of an Optimal Operation Existence and Its Parameters,” *Applied Sciences*, vol. 12, no. 15: 7555, 2022, doi: 10.3390/app12157555 (Scopus).
3. B. Kovalenko, V. Lukin, “Analysis of distortions due to bpg-based lossy compression of noise-free and noisy images,” *Herald of Khmelnytskyi National University Technical sciences*, vol. 325, no. 5, pp. 128-135, 2023, doi: 10.31891/2307-5732-2023-325-5 (фаховий).
4. V. Naumenko, B. Kovalenko, V. Lukin, “BPG-based compression analysis of poisson-noisy medical images,” *Radioelectronic and computer systems*, no. 3, pp. 91-100, 2023, doi: 10.32620/reks.2023.3.08 (Scopus).

5. V. Makarichev, B. Kovalenko, V. Lukin, “Preliminary analysis of noisy image lossy compression by discrete atomic transform-based coder,” *Aerospace technic and technology*, vol. 186, no. 2, pp. 67-74, 2023, doi: 10.32620/akt.2023.2.07 (фаховий).
6. B. Kovalenko, V. Lukin, “Usage of different Chroma Subsampling Modes in Image Compression by BPG Coder,” *Ukrainian journal of remote sensing*, vol. 9, no. 3, pp. 11-16, 2022, doi: 10.36023/ujrs.2022.9.3.216 (фаховий).
7. B. Kovalenko, V. Lukin, S. Kryvenko, V. Naumenko, B. Vozel, “Prediction of Parameters in Optimal Operation Point for BPG-based Lossy Compression of Noisy Images,” *Ukrainian journal of remote sensing*, vol. 9, no. 2, pp. 4-12, 2022, doi: 10.36023/ujrs.2022.9.2.212 (фаховий).
8. Б. Коваленко, В. Лукін, “Використання візуальних метрик для аналізу стиснення з втратами зашумлених зображень,” *Авіаційно-космічна техніка і технологія*, вип. 6, с. 83-91, 2021, doi: 10.32620/akt.2021.6.09 (фаховий).
9. V. Lukin, B. Kovalenko, B. Vozel, “A Fast and Accurate Prediction of BPG Compression Parameters in Optimal Operation Point Neighbourhood for Three-channel Noisy Images,” *Fundamental Research and Application of Physical Science*, vol. 8, pp. 129–160, 2023, doi: 10.9734/bpi/fraps/v8/7110A.
10. V. Lukin, B. Kovalenko, S. Kryvenko, V Naumenko, B. Vozel, “Prediction of Optimal Operation Point Existence and Its Parameters in BPG-Based Automatic Lossy Compression of Noisy Images,” *Current Overview on Science and Technology Research*, vol. 9, pp. 1–36, 2022, doi: 10.9734/bpi/costr/v9/4316A.
11. V. Lukin, B. Kovalenko, S. Kryvenko, N. Ponomarenko, K. Egiazarian, J. Astola, “Compression of Noisy Images Taking into Account Visual Quality: A Comprehensive Study,” *Research Developments in Science and Technology*, vol. 8, pp. 89–109, 2022, doi: 10.9734/bpi/rdst/v8/2722B.
12. V. Lukin, B. Kovalenko, S. Kryvenko, V. Naumenko, B. Vozel, “Peculiarities of BPG-Based Automatic Lossy Compression of Noisy Images”, *Computational Techniques for Smart Manufacturing in Industry 5.0 Methods and Applications*, 1st Edition, ISBN 9781032506203 386 Pages 10 Color & 115 B/W Illustrations January 27, 2025 by CRC Press.

Наукові праці, які засвідчують апробацію матеріалів дисертації:

13. B. Kovalenko, V. Naumenko, S. Krivenko, V. Lukin, “Analysis of noisy image lossy compression by BPG using visual quality metrics,” *Advanced Trends in Information Theory (ATIT)*, Kiev, Ukraine, 2021, doi: 10.1109/ATIT54053.2021.9678575 (Scopus).
14. B. Kovalenko, V. Lukin, “Analysis of color image compression by BPG coder,” *2022 IEEE 3rd KhPI Week on Advanced Technology (KhPIWeek)*, Kharkiv, Ukraine, October 2022, 6 p., doi: 10.1109/KhPIWeek57572.2022.9916387 (Scopus).
15. B. Kovalenko, V. Lukin, “Prediction of optimal operating point for BPG-based lossy compression of images corrupted by Poisson noise,” *2023 IEEE 4th KhPI Week on Advanced Technology (KhPIWeek)*, Kharkiv, Ukraine, 2023, pp. 1-5, doi: 10.1109/KhPIWeek61412.2023.10312954 (Scopus).
16. B. Kovalenko, V. Lukin, “BPG-based compression of Poisson noisy images,” *2023 13th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Athens, Greece, 2023, pp. 1-7, doi: 10.1109/DESSERT61349.2023.10416544 (Scopus).
17. B. Kovalenko, V. Lukin, “Pre-requisites for Mean Square Error Prediction in Better Portable Graphics based Lossy Compression of Grayscale Images,” *2024 IEEE 42nd International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine, 2024, pp. 488-492, doi: 10.1109/ELNANO63394.2024.10756949.
18. B. Kovalenko, V. Lukin, B. Vozel, “MSE Prediction in BPG-based Lossy Compression of Grayscale Images,” *2024 IEEE 5th KhPI Week on Advanced Technology (KhPIWeek)*, Kharkiv, Ukraine, 2024, doi: 10.1109/KhPIWeek61434.2024.10877996.
19. Б. Коваленко, “Знаходження оптимального параметра для прогнозування оптимальної робочої точки BPG кодера,” *Матеріали XXII Всеукраїнської науково-технічної конференції молодих вчених, аспірантів та студентів*. Одеса, 21-22 квітня 2022 р. - Одеса, Видавництво ОНТУ, 2022 р. – 175-177 с.

ABSTRACT

Kovalenko Bohdan. Multi-stage image processing and compression by BPG coder with parameter prediction. – Qualifying scientific work, the manuscript.

The thesis for a degree of Doctor of Philosophy (PhD) in the field of knowledge 17 Electronics and Telecommunications in specialty 172 Telecommunications and radio engineering. – National Aerospace University “Kharkiv Aviation Institute”, Kharkiv, 2025.

The thesis is devoted to the improvement and development of new effective methods of lossy compression and the opportunity to predict the quality parameters of one- and three-channel images of different origin. The object of research is the process of lossy compression and the distortions introduced by this type of compression. The subject of the study is a method for predicting the existence of the optimal operation point (OOP) and the parameters of the compressed image in the OOP and its vicinity for a BPG (better portable graphics) coder.

The work provides an analytical analysis of the main applications of multistage image processing, including the stages of formation, post-filtering, and analysis. Special attention is paid to lossy image compression. The main requirements for compression methods are considered, such as ensuring a sufficiently large compression ratio (CR) and reasonably high quality of compressed images, since they are subject to further analysis and classification. As a result, there is a problem with choosing the optimal value of the parameter that control compression (PCC) for the selected encoder to ensure the necessary compromise between the CR and the quality of the compressed image. It has been found that this task is complicated by the presence of noise in the image, which can significantly affect the compression characteristics. It is shown that, in the presence of noise, an optimal operation point can be observed that can provide high image quality with a sufficiently large CR. This point can be observed for different values of the PCC depending on the characteristics of the images, noise, and image pre-processing (in particular, the usage of a variance-stabilizing transformation). It has been found that the existing compression methods are inferior to the new solution, in particular, the BPG encoder, which can provide a high

compression ratio while keeping good image quality. The necessity of developing and improving methods for predicting the optimal operation point to ensure efficient lossy compression in cases of noised image are shown.

In the work, the **actual scientific and applied task** of developing of pre-processing and lossy compression methods with the ability to predict image quality parameters are set and solved. The characteristics of the BPG coder for the cases of with grayscale and color images in the presence of noise in image are investigated, the possibility of the presence of an optimal operation point (OOP) for compression of grayscale and color or other three-channel images by the BPG coder in cases of these images corruption by both additive and signal-dependent noise (with and without image preprocessing) is demonstrated. The method for predicting the existence of OOP as applied to the BPG coder for grayscale and color images distorted by noise has been developed. The method for predicting the mean square error (MSE) when working with a BPG coder has been developed.

The following methods have been used to solve scientific problems: mathematical and numerical modeling, machine learning methods, methods of probability theory and mathematical statistics.

The aim of this thesis is to develop efficient methods of processing and lossy compression and to provide an opportunity to predict image quality parameters.

The scientific results are:

In accordance with the stated aim, the following main tasks have been formulated and solved in the thesis:

- analysis of the features and requirements for lossy compression via BPG encoder;
- analysis of the characteristics of lossy compression via BPG coder of noise-distorted images;
- development of a method for predicting the optimal operation point for a BPG encoder, as well as the characteristics of images at this point and its neighborhood;
- development of a method for predicting the distortion characteristics during lossy compression for a BPG coder in the application to noise-free images;

1) **For the first time**, the characteristics of the BPG coder have been investigated for its operation under various conditions, such as working with grayscale and color images in

cases of noise presence. This made possible to formulate and provide recommendations for lossy image compression using the BPG coder.

2) The use of the optimal operation point (OOP) for different encoders was **further developed**, namely, for the first time, the possibility of the presence of such point for compression of grayscale and color images by a BPG encoder in cases where these images are affected by both additive and signal-dependent noise has been demonstrated, and the dependence between the position of this point and the characteristics of the image, noise, or processing methods was shown. This made possible to calculate the PCC to achieve the maximum filtering effect (maximum quality of the compressed image in accordance with various metrics).

3) The method for predicting the existence of OOP in the application to the BPG coder for grayscale and color images corrupted by noise is **improved**; the main difference with the previous methods is that the proposed method is used for the first time for the BPG coder and takes into account the compression features of this coder; this made possible to provide reasonable recommendations for choosing the parameter Q to achieve a compromise between the quality of compressed images and the compression ratio.

4) The method of predicting the mean square error (MSE) **has been improved**; the main difference is a preliminary analysis of input parameters characterizing the complexity of the image, which resulted in a list of parameters that can provide high prediction accuracy, which made it possible to increase the prediction accuracy without using several compression iterations.

The **practical significance** of the obtained results are the following: the existence of the OOP for the BPG coder for both grayscale and color images, in case of their corruption by additive or signal-dependent types of noise, is confirmed (the possibility of using the variation-stabilization transform for image preprocessing is considered), this makes possible to improve the quality of the compressed image compared to the original one by up to several dB (according to PSNR) and at the same time obtain a satisfactory compression ratio (from 12 to 51, depending on the image characteristics) if the compression is performed in the OOP neighborhood. The information is provided on the behavior of distortions introduced by the BPG encoder in lossy compression, as well as how these distortions behave in a

situation where noise is present in the image, which makes it possible to take into account the characteristics of the distortions in subsequent filtering or object detection procedures, which leads to an improvement in the efficiency of these procedures; a method for predicting the existence of optimal operation point and parameters in it for a BPG coder is proposed, the prediction accuracy is 0.84-0.97 for various quality metrics according to the R^2 criterion, the method can also work with color images and several types of noise (additive and signal-dependent), such prediction allows providing recommendations for the configuration of PCC, and, as a result, to obtain a better balance between the quality of the compressed image and the resulting CR; a method for predicting the mean square error of distortion for lossy compression using a BPG encoder that takes into account the complexity of images is proposed, which makes it possible to set the required compression parameter without the need to perform several compression iterations to find the optimal value of the parameter and speed up compression with the desired quality.

All the theoretical developments of the thesis have been brought to specific engineering techniques and algorithms in their application to processing and lossy compression and prediction of image compression efficiency.

The results of the work have been used in research study at the O. Ya. Usikov Institute for Radio Physics and Electronics of the National Academy of Sciences of Ukraine and in the National aerospace university named after M. E. Zhukovsky “Kharkiv Aviation Institute”.

According to the thesis material, 19 papers **were published**, 6 of which were articles in journals of Ukraine, 1 of which is indexed in SCOPUS, 2 articles in foreign country periodical (both of them Q1, Switzerland) included in the international scientific metric database SCOPUS, 7 - publications in conference proceedings (including 6 in the proceedings of international conferences, 4 of which are included in the international scientific metric database SCOPUS), and 4 book's chapters in foreign country periodicals.

Key words: lossy compression, multi-stage image processing, better portable graphics (BPG), optimal operation point (OOP), OOP prediction, noise-affected images, distortion characteristics prediction.

ЗМІСТ

АНОТАЦІЯ	2
ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ	17
ВСТУП.....	19
РОЗДІЛ 1	27
ОБРОБКА ТА СТИСНЕННЯ ЗОБРАЖЕНЬ	27
1.1 Основні додатки оброблення зображень	27
1.1.1 Сфери формування зображень.....	27
1.1.2 Обробка та аналіз цифрових зображень	29
1.2 Стиснення зображень.....	31
1.2.1 Огляд методів стиснення.....	31
1.2.2 Стиснення без втрат.....	32
1.2.3 Стиснення із втратами	33
1.3 Вимоги до методів стиснення з втратами.....	35
1.3.1 Основні вимоги та їх пріоритет	36
1.3.2 Вимоги до метрик якості.....	37
1.3.3 Вимоги до швидкості та ступеня стиснення	42
1.4 Особливості стиснення з втратами зображень, що вражені шумом.....	43
1.4.1 Причини виникнення шуму	45
1.4.2 Типи шумів	46
1.4.3 Вплив шуму на стиснення з втратами	50
1.5 Існуючі кодери.....	52
1.5.1 Засновані на ДКП.....	53
1.5.2 Засновані на Вейвлетному перетворенні.....	57

1.5.3 Засновані на застосуванні машинного навчання	58
1.5.4 Порівняння ефективності кодерів	60
1.6 Висновки до розділу та формування цілей дослідження	62
РОЗДІЛ 2	64
ДОСЛІДЖЕННЯ ХАРАКТЕРИСТИК СТИСНЕННЯ З ВТРАТАМИ BPG КОДЕРА	64
2.1 Стиснення зображень в градаціях сірого.....	64
2.1.1 Вибір тестових зображень, метрик та моделі шуму.....	65
2.1.2 Отримані залежності.....	67
2.1.3 Існування робочої точки для кодера BPG	70
2.2 Стиснення кольорових зображень кодером BPG	78
2.2.1 Субдискретизація кольоровості.....	80
2.2.2 Результат використання різних режимів субдискретизації кольоровості ...	81
2.2.3 Ступінь стиснення.....	89
2.3 Стиснення зображень із пуасонівським шумом на основі BPG.....	91
2.3.1 Набір зображень і використані метрики.....	92
2.3.2 Аналіз отриманих результатів	93
2.3.3 Стиснення з використанням перетворення Анскомба	100
2.3.4 Порівняння двох підходів	102
2.4 Аналіз спотворень, що виникають під час стиснення з втратами на основі BPG зображень без шуму і з шумом.....	106
2.4.1 Характеристики шуму, внесеного кодером для зображень без шуму.....	108
2.4.2 Комбіновані шуми кодера.....	111
2.4.3 Характеристики шуму, що вноситься кодером	113
2.5 Висновки до розділу	116

РОЗДІЛ 3	118
РОЗРОБКА МЕТОДУ ПРОГНОЗУВАННЯ ОРТ І ПАРАМЕТРІВ У НІЙ ДЛЯ ВРГ КОДЕРА	118
3.1 Прогнозування ОРТ для випадку стиснення напівтонових зображень.....	118
3.1.1 Опис запропонованого методу	119
3.1.3 Деталі вписування.....	121
3.1.4 Фактори, що впливають на точність	128
3.1.5 Прийняття рішень та інші практичні випадки	129
3.2 Прогнозування ОРТ для випадку стиснення кольорових зображень.....	133
3.2.1 Відмінність від методу для напівтонових зображень	133
3.2.2 Прогнозування існування ОРТ і параметрів у ній.....	135
3.2.3 Перевірка прогнозу та алгоритм прийняття рішень.....	141
3.2.4 Прогнозування в околі ОРТ	144
3.2.5 Прогнозування за кількома вхідними параметрами.....	148
3.2.6 Обговорення практичних аспектів	150
3.3 Висновки по розділу	152
РОЗДІЛ 4	153
РОЗРОБКА МЕТОДУ ПРОГНОЗУВАННЯ СКП ДЛЯ ВРГ КОДЕРА.....	153
4.1 Передумови до прогнозування	153
4.1.1 Аналіз залежностей метрик якості від параметра контролю стиснення....	154
4.1.2 Залежність між ентропією та отриманим СКП.....	157
4.2 Прогнозування СКП на основі локальної активності в блоках зображення під час стиснення чорно-білих зображень із втратами за допомогою ВРГ.....	159
4.2.1 Опис запропонованого методу	160
4.2.2 Результати прогнозування	164

	16
4.3 Прогнозування на основі ентропії.....	165
4.3.1 Відмінність від оригінального методу.....	165
4.3.2 Результати та порівняння підходів прогнозування	167
4.4 Прогнозування на основі більшої кількості парметрів	169
4.4.1 Прогнозування з використанням зваженої суми	171
4.4.2 Прогнозування з використанням простої нейронної мережі	174
4.5 Висновки за розділом.....	181
ВИСНОВКИ.....	183
СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ	185
ДОДАТОК А.....	201
СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА.....	201
ДОДАТОК Б.....	204
АКТИ ВПРОВАДЖЕННЯ.....	204

ПЕРЕЛІК УМОВНИХ СКОРОЧЕНЬ

АБГШ – адитивний білий гаусів шум;
ВСП – варіаційно-стабілізуюче перетворення;
ДВП – дискретне вейвлет-перетворення;
ДЗЗ – дистанційне зондування Землі;
ДКП – дискретно-косинусне перетворення;
КС – коефіцієнт стиснення;
КТ – комп'ютерна томографія;
ЛО – локальні оцінки;
НМ – нейронна мережа;
ОРТ – оптимальна робоча точка;
ПКС – параметр, що контролює стиснення;
СКП – середньоквадратична похибка;
СФЗ – система формування зображень;
РСА – радіолокатори з синтезованою апертурою;
AWGN – Additive white Gaussian noise;
BPG – Better Portable Graphics;
CSF – contrast sensitivity function;
GLCM – gray Level Co-occurrence Matrices;
HEIF – High Efficiency Image File Format;
HEVC – High-Efficiency Video Coding;
HVS – human visual system;
MDSI – Mean Deviation Similarity Index;
MOS – mean opinion score;
MSE – mean square error;
MS-SSIM – Multi-Scale Structural Similarity;
OOP – optimal operation point;
PSNR – peak signal-to-noise ratio;
PSNR-HVS-M – PSNR based on HVS Masking effect;

RDC – rate-distortion curves;

RMSE – root mean square error;

SROCC – Spearman rank-order correlation coefficient;

SSIM – structural similarity index measure.

ВСТУП

Обґрунтування вибору теми дослідження. Стрімкий розвиток сенсорів [1] для фотозйомки вплинув на найрізноманітніші галузі, такі як дистанційне зондування [2], [3], медицина [4], [5] і соціальні мережі [6], [7], а поширення смартфонів і постійна конкуренція виробників у сфері якості фотозйомки надала нові можливості пересічним користувачам. Як результат це призвело до значного збільшення кількості одержуваних зображень і підвищення роздільної здатності [6] відповідних систем формування зображень (СФЗ).

Наразі використовується безліч систем і комплексів дистанційного зондування Землі (ДЗЗ) і реалізується безліч проектів [1], [2]. Основна причина полягає в тому, що системи і комплекси ДЗЗ можуть оперативно надавати цінні дані для екологічного моніторингу, сільського і лісового господарства та інших галузей. Очевидною перевагою ДЗЗ є можливість моніторингу територій великої площі, своєчасного оцінювання та контролю їхніх параметрів. Краща просторова роздільна здатність, а також більш часте спостереження за територіями є позитивними характеристиками ДЗЗ, що роблять його привабливим і корисним для сучасних застосувань [1], [2]. Багатоканальна зйомка є основним режимом роботи датчиків ДЗЗ на сьогоднішній день. Термін "багатоканальний" охоплює багатоспектральне фіксування, спільні оптичні та радіолокаційні спостереження, повнополяризаційне і гіперспектральне фіксування. Триканальні зображення є найпростішим прикладом таких даних ДЗЗ. Тенденції та спостереження, отримані для цих зображень, можуть бути згодом узагальнені на зображення з більшим числом компонент [8]. Слід зазначити, що багатоканальне спостереження відкриває нові можливості для дистанційного зондування та підвищує точність оцінки параметрів, класифікації та виявлення об'єктів.

Не залишилася осторонь і медична галузь [4]. Зі зростанням використання технологій цифрової візуалізації в медичній діагностиці та лікуванні збільшився і розмір наборів даних медичних зображень, що призвело до проблем зберігання, передавання й оброблення цих зображень [9].

Висока деталізація або ж більша роздільна здатність зображення допомагає краще зрозуміти контекст, що стоїть за цим зображенням, незалежно від того, чи то зображення із соцмережі, чи то зображення дистанційного зондування, однак, як згадувалося раніше, для кожної галузі тією чи іншою мірою постає питання зберігання й передавання цього масиву даних.

Вирішенням цих завдань може бути використання методів стиснення зображень. Як відомо, алгоритми стиснення зображень можна розділити на дві групи: стиснення з втратами і стиснення без втрат [10]. Стиснення з втратами призводить до втрати частини чи якості даних у процесі стиснення, а в разі стиснення без втрат декомпресовані дані повністю збігаються з відповідними вихідними даними. Іншими словами: стиснення без втрат зберігає всю інформацію, що міститься в даних, але коефіцієнт стиснення (КС), що досягається для використовуваних методів, зазвичай є невеликим і може бути непридатним на практиці. Тоді стиснення з втратами виявляється більш розумним або єдиним можливим рішенням, оскільки КС можна варіювати і контролювати. Втім, збільшення КС зазвичай призводить до погіршення якості стисненого зображення, і в кожному конкретному випадку необхідно знайти розумний компроміс, де пріоритет вимог залежить від конкретного застосунку.

Результат стиснення багато в чому залежить від ряду факторів, наприклад кодеру, що використовується. Зазвичай використовувані кодери можуть надати або кращу якість, або ж більший коефіцієнт стиснення. Різні зображення можуть по-різному «поводитися» при використанні одного і того ж значення параметра, що контролює стиснення (ПКС). Присутність шуму на зображенні відіграє не останню роль під час стиснення, адже він може істотно вплинути на результат стиснення з втратами. Зі свого боку, шум на зображенні може проявлятися по-різному: він може бути адитивним, залежним від сигналу або являти собою деяку комбінацію цих компонентів. На додаток, обрана метрика для оцінки якості стиснення також має свій вплив на оцінку ефективності стиснення.

Виходячи з сказаного вище, **актуальною науковою задачею**, що була розв'язана в даній дисертаційній роботі, є розробка ефективного методу стиснення з втратами,

що здатний не тільки забезпечити високий коефіцієнт стиснення, а й прийнятну якість стиснення головним чином для зображень, що спотворені завадами.

Об'єкт дослідження. Стиснення за втратами та спотворення, які вносить даний тип стиснення.

Предмет дослідження. Методи прогнозування існування оптимальної робочої точки (ОРТ) та параметрів стисненого зображення в ОРТ та її околах для BPG кодера. BPG (Better Portable Graphics) - це новий та перспективний кодер, який націлений на те, щоб стати новим стандартом в області стиснення. Для цього в нього є всі передумови, зокрема підтримка всіх поширених форматів представлення зображення, що робить доцільним його використання в різних галузях отримання зображень. Окрім цього цей кодер базується на підмножині відкритого стандарту стиснення відео HEVC, а також має високий ступінь стиснення, що робить файли набагато меншими за JPEG при аналогічній якості. Згадані особливості дають передумови вважати, що BPG кодер може також доволі ефективно впоратися із завданнями стиснення.

Мета і завдання дослідження. Метою дисертаційної роботи є розроблення ефективних методів стиснення з втратами та надання можливості прогнозувати параметри якості зображення.

Відповідно до поставленої мети дисертації було сформульовано та вирішено такі основні **завдання**:

- аналіз особливостей і вимог до стиснення з втратами для BPG кодера;
- дослідження характеристик стиснення з втратами зображень, що спотворені завадами, для BPG кодера;
- розробка методу прогнозування оптимальної робочої точки для BPG кодера, а також параметрів у цій точці;
- розроблення методу прогнозування характеристик спотворень під час стиснення з втратами для BPG кодера.

Методи дослідження. Для розв'язання поставлених завдань у дисертаційній роботі було використано такі методи: методи моделювання для побудови кривих швидкість-спотворення, вписування кривих (регресії) для прогнозування існування робочої точки і значень у ній для конкретного значення ПКС, машинного навчання

для побудови нейронної мережі, призначеної для прогнозування параметрів, теорія ймовірності та статистики для аналізу точності прогнозування.

Наукова новизна отриманих автором результатів полягає в такому:

– **Вперше досліджено** характеристики BPG кодера під час його роботи в різних умовах, таких як робота з напівтоновими та кольоровими зображеннями у випадках наявності шуму на зображенні. Це дозволило сформулювати та надати поради щодо стиснення зображень з втратами з використанням модифікацій BPG кодера.

– **Отримало подальший розвиток** використання оптимальної робочої точки (ОРТ) для різних кодерів, а саме вперше продемонстровано можливість присутності такої точки для стиснення BPG-кодером напівтонових та кольорових зображень у випадках ураження цих зображень як адитивним, так і сигнально-залежними шумами. Це дало змогу розрахувати ПКС для досягнення максимального ефекту фільтрації (максимальної якості стисненого зображення у відповідності до різних метрик).

– **Вдосконалено** метод прогнозування існування ОРТ в застосуванні до BPG-кодера для півтонових і кольорових зображень, спотворених шумом; основною різницею з попередніми методами є те, що запропонований метод вперше використовуються до BPG-кодера і бере до уваги особливості стиснення для цього кодера; це дало змогу надати обґрунтовані рекомендації з вибору параметра Q для досягнення компромісу між якістю стиснутих зображень і коефіцієнтом стиснення.

– **Вдосконалено** метод прогнозування середньоквадратичної похибки (СКП); головною різницею є попередній аналіз вхідних параметрів, результатом цього аналізу став список параметрів, використання яких може забезпечити кращу точність прогнозування (зокрема із застосуванням нейромережі), що дало змогу підвищити точність прогнозування без використання декількох ітерацій стиснення.

Особистий внесок здобувача. Усі основні результати дисертації отримані автором самостійно. У роботах, опублікованих у співавторстві, здобувачеві належать

такі результати. В роботах [1], [2], перевірено наявність робочої точки для напівтонових зображень використовуючи метрики візуальної якості зображення такі як пікове відношення сигналу до шуму або ж peak signal-to-noise ratio (PSNR) та її модифікація що враховує оцінки якості зображення людською зоровою системою та ефект маскування PSNR-HVS-M (обидві метрики оцінюються у децибелах, більші значення означають кращу якість). У роботі [3] були розглянуті випадки роботи з кольоровими зображеннями. Розглянуто роботу кодера з різною субдискретизацією кольоровості (4:2:0, 4:2:2, 4:4:4), а також випадки різної інтенсивності шуму, що присутній на зображеннях. У [4] показано можливість існування OPT для кодерів, заснованих на дискретному атомарному перетворенні. У роботі [5] оптимальна робоча точка знайдена для випадків ураження медичних зображення сигнально-залежним шумом. У роботах [6-8] здобувач запропонував модифікований метод прогнозування існування OPT за допомогою двох статистичних параметрів, розрахованих у блоках розміром 8x8 пікселів, що не перекриваються. У [9] цей метод було адаптовано для роботи з кольоровими зображеннями. У роботі [10] метод прогнозування для кольорових зображень було розширено з використанням додаткових параметрів. У роботі [11] був проведений аналіз спотворень, що вносяться стисненням із втратами BPG кодером, як у випадку незашумлених, так і зашумлених зображень. У роботі [12] були розглянуті особливості автоматичного стиснення з втратами на основі BPG зашумлених зображень. У [13] розглянуто можливість використання метрик, відмінних від тих, що засновані на PSNR, а саме багатомасштабна версія метрики індексів структурної подібності (structural similarity index measure - SSIM) або ж MS-SSIM (ця метрика, на відміну від PSNR, має межі варіювання від 0 до 1, але так само, як в PSNR, більші значення характеризують кращу якість зображення). У [14] проведено порівняльний аналіз для кольорних компонент (у системі RGB) для різних методів субдискретизації. В [15] було порівняно декілька вхідних параметрів для прогнозування OPT з метою виявлення того що може забезпечити кращі показники точності прогнозування. У роботах [16] оптимальна робоча точка знайдена для випадків ураження сигнально-залежним шумом зображень з різних областей. У [17] метод прогнозування OPT для кольорових зображень був

адаптований для роботи з сигнально-залежним типом шуму. У [18] було проаналізовано зв'язок між складністю зображення та ентропією для подальшого використання в методах прогнозування характеристик спотворень. У [19] метод прогнозування ОРТ для напівтонових зображень було модифіковано для підвищення точності прогнозування за рахунок використання більшої кількості вхідних параметрів, що характеризують особливості зображення.

Список публікацій здобувача за темою дисертації наведено в Додатку А.

Апробація матеріалів дисертації. Основні положення дисертації доповідалися й обговорювалися на «IEEE 3rd International Conference on Advanced Trends in Information Theory (ATIT)» (2021 Київ, Україна), «IEEE 3rd KhPI Week on Advanced Technology (KhPIWeek)» (2022, Харків, Україна), «IEEE 4th KhPI Week on Advanced Technology (KhPIWeek)» (2023, Харків, Україна), «13th International Conference on Dependable Systems, Services and Technologies (DESSERT)» (2023, Афіни, Греція), «IEEE 42nd International Conference on Electronics and Nanotechnology (ELNANO)» (2024, Київ, Україна), «IEEE 5th KhPI Week on Advanced Technology (KhPIWeek)» (2024, Харків, Україна).

Зв'язок роботи з науковими програмами, планами, темами, грантами. Дослідження, представлені в цій дисертації, проводилися на кафедрі інформаційно-комунікаційних технологій Національного аерокосмічного університету імені О.О. Зеленського. Були відображені в таких звітах: НДР Д504-1/2019-М (№ ДР 0119U100940) «Розробка системи обробки мультимедійних даних, її впровадження в існуючі та перспективні системи та комплекси», 2021 р.; НДР Д504-2/2021-Ф (№ ДР 0121U112176) "Методи багатоканального оброблення зображень у мобільних системах на основі прогнозування та машинного навчання", 2021-2023 роки; НДР Д504-3/2023-М (№ ДР 0123U101143) "Розробка методів попереднього аналізу фотоплетизмографічних сигналів з метою визначення інформативних ознак, що характеризують зміни артеріального тиску", 2023 р.

Практична значущість результатів:

– Підтверджено існування ОРТ для VPG кодера як для напівтонових, так і для кольорових зображень, у разі їхнього ураження адитивними або сигнально-

залежними типами завад. Це дозволяє підвищити якість стисненого зображення у порівнянні з первинним до кількох дБ (згідно PSNR) і при цьому отримати задовільний коефіцієнт стиснення, якщо стиснення буде проводитися в околі ОРТ.

– Надано інформацію про поведінку спотворень, які вносить BPG кодер під час стиснення з втратами, а також, як ці спотворення ведуть себе в ситуації, коли на зображенні присутній шум. Завдяки цьому отримана можливість враховувати особливості завад у подальших процедурах фільтрації або детектування об'єктів, що призводить до покращення ефективності цих процедур.

– Запропоновано метод прогнозування робочої точки та параметрів у ній для BPG кодера, точність прогнозування становить 0,84-0,97 для різних метрик якості за критерієм R^2 . Метод також може працювати з кольоровими зображеннями та кількома типами шуму (адитивний і сигнально-залежний). Таке прогнозування дає змогу надати рекомендації щодо вставлення ПКС, і, як результат, отримати краще співвідношення між якістю стисненого зображення та отриманим КС.

– Запропоновано метод прогнозування середньоквадратичної похибки спотворень для стиснення з втратами за допомогою BPG кодера, що враховує складність зображень. Це дає змогу виставити необхідний параметр стиснення без необхідності проводити кілька ітерацій на стиснення з метою пошуку оптимального значення параметра та в разі прискорити стиснення з бажаною якістю.

Публікації. Основні результати дисертації опубліковано в 19 роботах, у тому числі: 6 статей у наукових журналах України, 1 з яких індексується в SCOPUS, 2 статті у наукових періодичних виданнях іншої держави (Швейцарія), які індексується в SCOPUS (обидві Q1), 7 – публікації у матеріалах конференцій, серед яких 6 у матеріалах міжнародних англomовних конференцій, 4 з яких включені до міжнародної наукометричної бази SCOPUS, та 4 – глави в книгах у наукових періодичних виданнях іншої держави.

Структура та обсяг дисертації. Дисертація складається з анотації, змісту, списку скорочень, вступу, чотирьох розділів, висновків, списку використаних джерел і додатків. Повний обсяг роботи становить 206 сторінок друкованого тексту, з яких анотація — на 8 сторінки, зміст — на 4 сторінки, список скорочень — на 2 сторінки, основний текст — на 165 сторінок, список із 119 використаних джерел — на 15 сторінок, додатки — 4 сторінок. Дисертація містить 83 рисунків (5 з них на 3 окремих сторінках), 28 таблиць (1 з них на 1 окремій сторінці).

РОЗДІЛ 1

ОБРОБКА ТА СТИСНЕННЯ ЗОБРАЖЕНЬ

У цьому розділі розглядаються різні методи стиснення. Розглянуто особливості стиснення з втратами, а також вимоги, що висуваються до цього типу стиснення. Також розглянуто популярні кодери, що здійснюють стиснення з втратами, і проведено порівняльний аналіз їх ефективності.

1.1 Основні додатки оброблення зображень

Під зображеннями розумітимемо растрові зображення, що складаються з масивів даних, що зберігають у собі значення яскравості пікселів. В силу свого розмаїття дані масиви можуть бути як двовимірні (для напівтонових зображень), так і тривимірними (для кольорових, а також багатоканальних зображень).

Далі розглянемо докладніше особливості формування зображень, що належать до різних сфер діяльності, а також різні способи аналізу та обробки даних зображень.

1.1.1 Сфери формування зображень

Значну частину інформації людина отримує за допомогою візуального огляду і, як результат, візуальна інформація набула ролі однієї з найважливіших складових цифрового простору. Чимала частина візуальної інформації - це зображення.

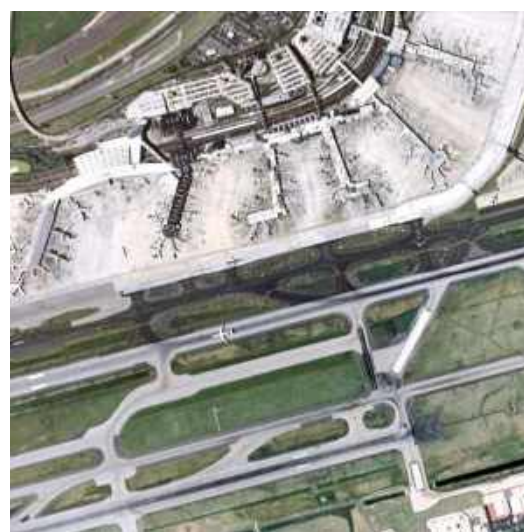
Зображення - один із найпоширеніших типів даних, що передаються каналами зв'язку, відображаються, обробляються та зберігаються в комп'ютерах та інших пристроях [1-10]. З огляду на свою поширеність дані зображення можуть мати різні характеристики, такі як яскравість, контрастність або наявність артефактів або інших спотворень, через особливості їх формування [11].

Спочатку варто детальніше розглянути області формування зображень та особливості, що властиві даним зображенням.

1) Дистанційне зондування. Засоби дистанційного зондування Землі (ДЗЗ) та інші види знімальних пристроїв в даний час виробляють величезну кількість даних, які можуть бути використані в різних додатках [1], [2]. Сучасні сенсори дають змогу оцінювати параметри (біомаса, ерозія тощо) територій великої площі, що знімаються, і контролювати їхню зміну в часі, визначати межі та форму об'єктів, а також здійснювати контроль якості поверхні. Очевидні тенденції пов'язані з підвищенням роздільної здатності сенсорів [1], більш частими спостереженнями, або отриманням більшої кількості зображень за певний час. Все це призводить до швидкого збільшення об'єму отриманих даних, де отримані зображення повинні передаватися лініями зв'язку, оброблятися з відповідною ефективністю, зберігатися на борту або на Землі, інтерпретуватися і поширюватися [12]. Приклади зображень дистанційного зондування приведено на рис 1.1.



а



б

Рисунок 1.1 – Приклади зображень дистанційного зондування

2) Медицина. Медична візуалізація включає різні методи створення візуальних уявлень людського тіла шляхом обробки біомедичних сигналів [4]. Отримані зображення різняться залежно від методу візуалізації та досліджуваної анатомічної частини, що призводить до появи цілого ряду зображень зі своїми особливостями. Наприклад, медичні системи візуалізації, такі як рентгенівська та комп'ютерна томографія (КТ), використовують рентгенівське випромінювання або комп'ютерну

томографію для отримання зображень, які реєструються шляхом вимірювання результуючого ослаблення [9]. Також зі зростанням використання технологій цифрової візуалізації в медичній діагностиці та лікуванні збільшився і розмір наборів даних медичних зображень, що призвело до проблем зберігання, передачі та обробки цих зображень.

3) Соцмережі. Розвиток фотосенсорів торкнувся і пересічних користувачів. Оскільки виробники смартфонів весь час борються за покращення сенсорів у своїх пристроях, це призвело до того, що користувачі смартфонів отримали можливість фотографувати з найкращою якістю та великою роздільною здатністю [7]. Також варто відзначити, що в сучасному світі у більшості людей є хоча б один обліковий запис в якій-небудь соцмережі, куди щодня завантажуються зображення від інших користувачів. Ці зображення можуть різнитися за змістом і якістю - це можуть бути фотографії тварин, текстів, текстур, будівель тощо, крім того, вони можуть бути зроблені в умовах недостатньої освітленості або під впливом інших видів завад або викривлень, що можуть вплинути на фінальну якість зображень. Цей тип зображень повинен ефективно передаватися та зберігатися на сервері ресурсу.

Виходячи з вище сказаного, кожна галузь дійсно має свої особливості, і, як результат, відрізняються не тільки контент на зображенні, а й проблеми, які притаманні тій чи іншій галузі, - це може бути як характерні викривлення, так і шуми. Дані особливості повинні бути враховані в подальших методах обробки зображення. Варто відзначити, що не дивлячись на свої особливості, зображення, як і раніше, необхідно ефективно зберігати і передавати, тому завдання по розробці методів, що вирішують дану проблему, є донині актуальним [13].

1.1.2 Обробка та аналіз цифрових зображень

Обробка зображень є одним із ключових елементів у процесі надання користувачеві задовільних даних [14], [15]. Серед завдань можуть бути такі напрямки: корекція кольору, зміна розміру [16], [17] або контрасту [18-20], а також більш

комплексні завдання, такі як фільтрація [21], розпізнавання образів [22] і стиснення [23]. Варто розглянути кожен із напрямів детальніше.

1) Фільтрація. Головне завдання - підвищення якості зображень [21], [24]. Фільтри використовують для посилення, придушення або виділення певних характеристик чи інформації в сигналі або наборі даних [25], [26]. Існує кілька типів, у тому числі лінійні фільтри, такі як низькочастотні, високочастотні, смугові та режекторні фільтри, які працюють із сигналами або даними, використовуючи лінійні операції, такі як згортка. Далі - нелінійна фільтрація, яка, на відміну від лінійних фільтрів, застосовує нелінійні операції для зміни або обробки даних. Крім того, фільтрація в області ортогональних перетворень, яка передбачає перетворення сигналу або зображення в частотну область за допомогою таких методів, як наприклад перетворення Фур'є. Як результат, фільтрація застосовується в частотній області, після чого виконується зворотне перетворення для повернення в часову або просторову область.

2) Розпізнавання образів. Даний напрямок набув найбільшого поширення в сучасний час у зв'язку з розвитком засобів машинного навчання та комп'ютерного зору, що дозволяє вирішувати поточні проблеми класифікації та розпізнавання за допомогою алгоритмів, заснованих на нейромережах. Але і до стрімкого розвитку нейромереж дана область була дуже популярна і використовувала інші підходи щодо детектування об'єктів на зображенні. Розпізнавання образів — це процес, який включає аналіз введених даних, вилучення з них шаблонів [27-29], у спробі порівняти їх з певними стандартами і використання отриманих результатів для визначення дій, які система буде робити в майбутньому [30]. Розпізнавання образів аналізує те, як комп'ютери стежать за навколишнім середовищем і можуть розпізнавати закономірності, що їх цікавлять, у навколишньому середовищі, а потім виносити логічні та прийнятні судження про групи, до яких ці закономірності належать.

3) Стиснення зображень. Скорочення об'єму даних, що використовуються для представлення зображення, шляхом компактного кодування або апроксимації [31-33].

Оскільки завдання зберігання та передавання великих обсягів зображень стоїть дуже гостро, то варто сконцентруватися на стисненні зображень, а саме розглянути

можливості та вимоги, що висуваються до цієї галузі обробки зображень. Нагадаємо також, що стиснення з втратами пов'язане з класифікацією та, за певних умов, може частково усувати шум, тобто виконувати специфічну фільтрацію [34].

1.2 Стиснення зображень

Виходячи з опису напрямків обробки зображення стає очевидним, що деякі з них пов'язані між собою або безпосередньо впливають один на одного [35-37]. Варто зазначити, що для розв'язання задачі ефективного зберігання та передавання зображень основну роль відіграють методи стиснення [38-40]. Однак оскільки якість фінального зображення, яке отримає користувач, відіграє не останню роль, стиснення повинно ефективно справлятися із завданням зменшення розміру зображення, водночас зберігаючи якість зображення [41-44]. Тут потрібно зазначити, що ці два параметри зазвичай взаємовиключні [32], тобто при збільшенні ступеню стиснення зменшується розмір зображення та якість зображення і, навпаки, при використанні малих значень коефіцієнту стиснення (КС) результуючий ступінь стиснення виявляється меншим, проте досягається краща якість зображення. Як результат, хоч і існують різні методи стиснення, проте позитивний ефект, за якого і якість, і коефіцієнт стиснення досягають задовільних рівнів, спостерігається не завжди. У зв'язку з цим необхідно вивчити сучасні методи стиснення і розглянути можливості прогнозування параметрів стиснення або ж оптимального параметра для стиснення для досягнення прийнятних для користувача коефіцієнта стиснення і якості зображення. У зв'язку з цим початковим завданням стає розгляд наявних методів стиснення та їх особливостей.

1.2.1 Огляд методів стиснення

Основне завдання методів стиснення - це зберегти якомога більше корисної інформації і водночас зменшити розмір даних, наскільки це можливо. Оскільки цифрове зображення - це масив значень пікселів, то часто сусідні пікселі корелюють

між собою і містять надлишкові біти. Ця надмірність і використовується в алгоритмах стиснення. Надлишкові біти видаляються із зображення, внаслідок чого розмір стисненого зображення виявляється меншим за вихідне.

Стиснення має два компоненти: це скорочення надмірності та скорочення нерелевантних даних. Різниця полягає в тому, що під час скорочення надмірності видаляють зайві або повторювані біти, водночас під час скорочення нерелевантних даних прибирають непотрібну або найменш важливу інформацію, яку не сприйматиме спостерігач. [45].

Своєю чергою існує кілька типів надмірності. Перша - це надмірність кодування, яка проявляється в неефективному представленні сірого на зображенні. Наприклад, якщо для представлення кожного рівня сірого використовується більше кодових символів, ніж необхідно, зображення містить надмірність кодування. Далі - міжпиксельна надмірність, що виникає через кореляцію пікселів зображення. Остання це психо-візуальна надмірність, що проявляється шляхом ігнорування надлишкових даних зоровою системою людини [46].

Як результат, стиснення зображень - це процес представлення зображення таким чином, щоб максимально видалити надмірність і при цьому, по можливості, зберегти (забезпечити бажану) якість результуючого зображення.

Різні методи стиснення по-різному справляються з вищеописаним завданням, проте всіх їх можна розділити на дві групи: стиснення без втрат і стиснення з втратами. Представники кожної з груп використовують різні алгоритми і підходять для вирішення різних завдань.

1.2.2 Стиснення без втрат

У схемі стиснення без втрат стиснене зображення збігається із вхідним [47], [48]. Методи стиснення зображень без втрат спочатку розбивають зображення на масиви пікселів. Потім проводиться обробка кожного окремого пікселя. Перший етап включає в себе прогнозування значення наступного пікселя зображення за сусідніми пікселями. На другому етапі різниця між прогнозуванням значенням та фактичною

інтенсивністю наступного пікселя кодується за допомогою різних методів кодування [45].

Існують такі методи кодування: кодування довжин серій і статистичне кодування. Перший метод є найпростішою технікою стиснення зображень, у якій послідовність однакових символів замінюється парою, що містить символ і довжину, з якою повторюється число. Другий враховує статистичні особливості зображення. У нього включені такі методи: кодування Хаффмана, арифметичне кодування, кодування області. Кожен із яких має свої особливості.

У кодуванні Хаффмана менший бітовий код присвоюється значенням пікселів, які зустрічаються часто, а більший бітовий код - значенням пікселів, що повторюються. Для стиснення всього зображення вхідне зображення розділяється на блоки 8x8, потім кожен блок кодується певним символом, після чого кодування Хаффмана застосовується до кожного блоку [46].

Арифметичне кодування націлюється на перший та останній символ, який кодується та декодується першим [45]. Інші особливості цього методу: алфавіт символів не повинен бути нескінченним, так само, як і всі послідовності символів, кількість дійсних чисел в інтервалі від 0 до 1 може призначити унікальний під інтервал для будь-якої заданої вхідної послідовності символів [46].

Кодування області є покращеною версією кодування довжин серій. Як результат, він є більш ефективним методом, однак має обмеження, які можуть бути застосовані до нелінійних перетворень [45].

Стиснення без втрат в основному використовується для чутливих до спотворень зображень, в яких втрата будь-яких даних є вкрай небажаною. До таких зображень можна віднести медичні зображення, архівні документи та інші зображення, що містять важливу інформацію.

1.2.3 Стиснення із втратами

Вище згадувалося, що не всі методи стиснення здатні одночасно надати великий коефіцієнт стиснення та відмінну (задовільну) якість зображення. До таких можна

віднести методи, засновані на стисканні з втратами, які звичайно можуть надати великий коефіцієнт стиснення (десятки і навіть сотні), проте це позначається на якості зображення. Хоча на даний момент дана область стиснення просунулась вперед, значно піднявши планку компромісу між якістю зображення і коефіцієнтом стиснення [40], [49], проте це завдання все ще залишається невирішеним і вимагає подальших поліпшень для досягнення бажаного компромісу. Відповідне рішення може бути знайдено шляхом вивчення роботи безпосередньо методу стиснення, що дозволить вибрати оптимальні параметри для стиснення. В свою чергу, це дозволить не тільки досягти бажаного результату, але і дозволить заощадити час і обчислювальні ресурси. Крім того, дане рішення може бути прийняте за допомогою прогнозування параметрів стисненого зображення, як результат на базі даного методу можна коригувати параметри стиснення, що так само дозволить зекономити ресурси. Однак розробка даного методу вимагає знань про особливості стиснення з втратами, які будуть викладені нижче.

Головною відмінністю стиснення з втратами від методів без втрат полягає в тому, що під час стиснення з втратами оригінальне і розтиснуте зображення не є ідентичними. Це пов'язано з втратами, що виникають під час додаткового кроку, який і відрізняє цей метод, а саме квантування. Під час квантування деяка інформація, що належить до психо-візуальної надмірності на зображенні, опускається. Загальна схема стиснення із втратами наведена на рисунку 1.2.



Рисунок 1.2 – Схема алгоритму стиснення із втратами

Алгоритми стиснення з втратами часто починається з поділу вхідного зображення на невеликі блоки меншого розміру. Потім для кожного блоку отримують відповідні коефіцієнти на основі певного перетворення, прикладом кодування з перетворенням є дискретне косинусне перетворення або ж вейвлет-перетворення. Результуючі коефіцієнти потім обчислюються за допомогою методів квантування, після чого вихідні дані з блоку квантування використовуються для методу кодування символів для отримання вихідних даних. У декодері відбувається зворотний процес для відновлення зображення.

Головною особливістю стиснення з втратами став великий коефіцієнт стиснення, і, як результат, даний тип стиснення отримав велике поширення у різних системах та додатках. Водночас вирішення проблеми погіршення якості, пов'язаної із втратами, стало основою появи нових методів і підходів. Однак для того, щоб розробити, покращити або оцінити метод стиснення, потрібно виділити вимоги, які даний метод повинен виконувати.

1.3 Вимоги до методів стиснення з втратами

Як вже говорилося вище, під час квантування деяка інформація втрачається через прирівнювання до нуля коефіцієнтів перетворення, це призводить до збільшення важливої міри, що характеризує ефективність стиснення, а саме КС. Ця величина може бути розрахована наступним чином:

$$CR = \frac{B_{in}}{B_c}, \quad (1.1)$$

де CR – коефіцієнт стиснення;

B_{in} – розмір вхідного зображення в байтах;

B_c – розмір стисненого файлу так само в байтах.

Чим більший даний коефіцієнт, тим менший розмір стисненого файлу до вхідного. У разі стиснення з втратами даний коефіцієнт може бути досить високий. У зв'язку з цим до стиснення із втратами можуть висуватися різні вимоги, які будуть розглянуті далі.

1.3.1 Основні вимоги та їх пріоритет

Збільшення КС зазвичай призводить до погіршення якості стисненого зображення, і в кожному конкретному випадку необхідно знайти розумний компроміс. Цей компроміс залежить від багатьох чинників: 1) яким є призначення зображення, що розглядається, і як його надалі обробляють (аналізують, інтерпретують); 2) який кодер (метод стиснення) використовують і наскільки легко варіювати й контролювати параметри стиснення; 3) якими є властивості (зміст, складність) стиснутого зображення; 4) якими є основні вимоги до стиснення та їхній пріоритет. [50].

Першою вимогою є те, що стиснення з втратами не повинно призводити до надто великого погіршення якості зображення з погляду його візуального сприйняття людиною, класифікації або виявлення меж чи об'єктів.

Друга вимога виходить з того факту, що різні методи стиснення використовують різні механізми зміни і контролю КС або якості стисненого зображення. Зважаючи на цю особливість, контроль стиснення повинен бути легко здійснюваним, а саме параметр контролю стиснення повинен легко налаштовуватися і бути зрозумілим для користувача, маючи просту структуру (наприклад більше значення параметра еквівалентно більшому коефіцієнту стиснення та меншій якості).

Зважаючи на те, що зображення можуть досить сильно відрізнятися одне від одного своїм наповненням і, як результат, вони відрізняються своєю складністю. Особливістю стиснення зображень різної складності є те, що складніше стискати зображення більш комплексної структури із забезпеченням належної якості для даного КС. Отже третьою вимогою є те, що метод стиснення повинен брати до уваги складність зображення та підстроюватись таким чином, щоб досягти кращого

результату. Окрім цього, наявність шуму на зображенні також може суттєво вплинути на результат стиснення [51]. Тому це також повинно бути враховано.

Остання вимога стосується забезпечення потрібного компромісу, який може різнитися залежно від додатку, в якому буде використовуватись метод стиснення, це може бути забезпечення якомога більш високої якості стиснених зображень для даного КС, виконати якнайшвидше або забезпечити максимальний КС при допустимій якості зображення.

Можна помітити, що деякі вимоги суперечливі, а деякі з них можуть бути складно здійсненими, однак, щоб розв'язати цю проблему, необхідно якось оцінити роботу стиснення з втратами. Крім уже знайомого коефіцієнта стиснення, який може сказати, наскільки зменшився розмір зображення, є низка метрик, які можуть розповісти про якість стисненого зображення. Правильний вибір потрібної метрики може вплинути на підсумковий результат по досягненню компромісу між великим коефіцієнтом стиснення і прийнятною якістю зображення.

1.3.2 Вимоги до метрик якості

Від вибору метрики якості зображення багато в чому залежить те, як буде оцінено фінальне зображення [52-54]. Мова йде про метрики з еталоном, бо маємо первинне зображення (яке стискається, I^{or}) та стиснене зображення I^c . Різні метрики беруть до уваги різні особливості зображення. Однак їх так само можна розділити на підгрупи. Спочатку розглянемо метрики, засновані на піковому відношенні сигналу до шуму або ж peak signal-to-noise ratio (PSNR).

Для отримання PSNR метрики необхідно розрахувати середньоквадратичну похибку, відому як mean square error (MSE) [55], оскільки обидві метрики пов'язані між собою варто зазначити, що обидві беруть до уваги особливості пікселів зображення. Для розрахунку цих метрик необхідно скористатися виразами (1.2) і (1.3).

$$MSE = \sum_{i=1}^I \sum_{j=1}^J (I_{ij}^{or} - I_{ij}^c)^2 / (IJ), \quad (1.2)$$

де i та j - індекси пікселів, а I і J - висота та ширина зображення. PSNR розраховується для спотвореного та еталонного зображень як

$$PSNR = 10 \log_{10} \left(\frac{255^2}{MSE} \right). \quad (1.3)$$

де 255 - це максимальне значення пікселя для 8-бітових беззнакових цілих типів даних.

Подальший розвиток метрики вилився в її модифікацію, названу PSNR-HVS, яка розширює звичайний PSNR завдяки використанню функції контрастної чутливості (Contrast Sensitivity Function - CSF), а також PSNR-HVS-M, що використовує принцип міжкоефіцієнтного контрастного маскування базисних функцій дискретного косинусного перетворення (ДКП) [56]. Метрику PSNR-HVS-M можна розрахувати так:

$$PSNR - HVS - M = 10 \log_{10} \left(\frac{255^2}{MSE - HVS - M} \right), \quad (1.4)$$

де MSE-HVS-M обчислюється в блоках 8x8 в області DCT, а потім агрегується.

Ще однією особливістю цих метрик є здатність враховувати особливості зорового сприйняття людини. Для досягнення оцінки, яка б відповідала результатам оцінки якості зображення людською зоровою системою (human visual system - HVS), деякі дослідники об'єднали характеристики HVS [57] із чистими математичними алгоритмами. До таких метрик, крім PSNR-HVS і PSNR-HVS-M, також відносяться подальші модифікації даних метрик PSNR-HA і PSNR-HMA [58]. Обидві метрики націлені перейняти принципи роботи MS-SSIM (представника іншого сімейства метрик), а саме дають змогу врахувати кілька аспектів HVS, зокрема таку психо-візуальну особливість, що розтягування зображення та зміна середнього (у деяких

межах) не чинять істотного впливу на сприйняття людиною зображень. Ця метрика розраховується як:

$$PSNR - HA = 10 \log_{10} \left(\frac{255^2}{MSE - HA} \right), \quad (1.5)$$

Друге сімейство метрик ґрунтуються на індексі структурної подібності (structural similarity index measure - SSIM). Основна ідея цієї метрики полягає в тому, на зображенні сусідні пікселі схожі між собою [59]. Цей взаємозв'язок містить інформацію про структуру об'єктів, розташованих на зображенні. Ще одна відмінність від інших метрик якості, таких як MSE або PSNR, полягає в тому, що ці підходи оцінюють абсолютні помилки:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \quad (1.6)$$

де x – еталонне зображення;

y – спотворене шумом (викривлене) зображення;

μ_x і μ_y – середні значення блоків еталонного і спотвореного шумом зображень, відповідно;

σ_x^2 і σ_y^2 – дисперсії блоків еталонного і спотвореного зображень, відповідно; σ_{xy} – коваріація блоків еталонного і спотвореного зображень.

Розвитком цієї метрики стала багатомасштабна версія метрики (Multi-Scale) SSIM, що отримала аббревіатуру MS-SSIM. Особливість цієї метрики полягає в тому, що вона розраховується в декількох масштабах через процес кількох етапів субдискретизації [60].

Варто відзначити, що деякі метрики можуть одночасно використовувати кілька підходів для оцінки якості зображення. Прикладом такої метрики може бути MDSI [61], яка дає змогу отримати ефективну та надійну модель оцінювання якості повного еталонного зображення на основі градієнтної та хроматичної подібності. Градієнтна

подібність використовується для вимірювання локальних структурних спотворень. На додаток до цього для вимірювання колірних спотворень використовується хроматична подібність. У цій метриці також використовується поєднання відхилень для обчислення оцінки якості за двома картами подібності. Згідно MDSI, зображення з ідеальною якістю оцінюється нульовим балом (значенням), оскільки в карті подібності немає жодних відхилень.

З огляду на викладене вище можна сформулювати такі вимоги: для об'єктивного оцінювання якості роботи стиснення з втратами необхідно використовувати метрики, які братимуть до уваги особливості зорової системи людини, а ще варто використовувати метрики з різних сімейств, щоб упевнитися в правдивості (точності) отриманих оцінок.

Однак все ще залишається відкритим питання, яку метрику варто використовувати або, іншими словами, яка метрика є найкращою. Одним із варіантів відповіді є MOS (mean opinion score), яка раніше не згадувалася через свою особливість, що не дає змоги використовувати дану метрику в автоматизованих системах оцінювання якості зображення, а саме в даній метриці використовується велика кількість спостерігачів для оцінювання обраного набору зображень, наприклад Tampere image database 2013 (TID2013). Не дивлячись на очевидний недолік, ця метрика може допомогти з вибором метрик, які можуть бути використані як найбільш точні.

Для того, щоб перевірити точність метрики, необхідно розрахувати коефіцієнт кореляції між метрикою, що розглядається, і MOS для одного набору зображень. Розраховуватися буде Коефіцієнт рангової кореляції Спірмена (вона ж SROCC - Spearman rank-order correlation coefficient), який описується таким чином:

$$SROCC(O, S) = 1 - \left[\frac{6 \sum_{i=1}^N d_i^2}{N(N^2 - 1)} \right], \quad (1.7)$$

де d_i позначає різницю між i -м рангом зображень за об'єктивними (O), і суб'єктивними (S) оцінками, тобто оцінками MOS.

Для цього коефіцієнта кореляції більші значення вказують на кращу відповідність розглянутої метрики до суб'єктивної оцінки, представленої MOS.

Особливістю даної бази даних зображень TID2013 є те, що вона містить велику кількість зображень, а саме 3000 тестових зображень, отриманих із 25 еталонних зображень, 24 типів викривлень для кожного еталонного зображення та 5 рівнів для кожного типу викривлень, що були оцінені в діапазоні від 0 до 9. Були проведені розрахунки згаданого вище коефіцієнту кореляції між MOS та різними метриками контролю якості для трьох типів спотворень, що пов'язані із стисненням з втратами. Результати приведені в таблиці 1.1.

Таблиця 1.1 – Результати розрахунку метрики SROCC для обраних метрик оцінки якості зображень

Назва метрики	Значення SROCC	Назва метрики	Значення SROCC
PSNR	0,914	WASH	0,389
MSE	-0,914	CWSSIM	0,862
SSIM	0,893	MSVD	-0,406
MSSSIM	0,934	IFC	0,875
PSNRHVS	0,964	UQI	0,824
PSNRHVSM	0,961	QILV	0,902
PSNRHA	0,966	VSNR	0,915
PSNRHMA	0,967	NQM	0,904
FSIM	0,961	UNIQUE	0,902
MDSI	-0,967	MSUNIQUE	0,9062

З таблиці видно, що низка метрик має доволі високі абсолютні значення коефіцієнта кореляції, що може слугувати обґрунтуванням до їх подальшого використання. Такого статусу набули такі метрики: PSNR-HVS, PSNR-HVS-M, PSNR-HA, PSNR-HMA, MS-SSIM і MDSI. Однак до цього набору варто також додати метрики MSE і PSNR, які, незважаючи на менший коефіцієнт кореляції, все ще залишаються стандартними та розповсюдженими критеріями оцінювання якості зображень, що, в свою чергу, може слугувати відправною точкою під час порівняння.

Важливим фактором, який варто відзначити, є те, як інтерпретувати результати обраних метрик, а саме якими є їх межі та яка закономірність (наприклад, більші

значення - краща якість зображень чи навпаки). Для набору обраних метрик закономірності такі: PSNR, та її модифікації, включно з PSNRHVS, PSNRHVSM, PSNRHA, PSNRHMA, обчислюються в децибелах (варто зазначити, що значення модифікацій мають дещо більші значення для того самого рівня якості в порівнянні з оригінальним PSNR), а більші значення цих метрик означають кращу якість зображень. Така сама ситуація і з MS-SSIM, для цієї метрики теж справедлива закономірність: більші значення - краща якість, однак, для цієї метрики значення вимірюються в межах від 0 до 1. Що стосується MDSI, то для цієї метрики менші значення характерні для кращої якості зображення.

1.3.3 Вимоги до швидкості та ступеня стиснення

Раніше згадувалося, що багато в чому вимоги до алгоритмів стиснення залежать від додатка, в якому вони будуть використовуватися. Для додатків, націлених на роботу в реальному часі, ключовим аспектом є час, витрачений на виконання алгоритму. Прикладом таких додатків можуть бути потокові сервіси, відеоконференції та онлайн-ігри, дані в таких додатках повинні стискатися і розтискатися досить швидко, щоб уникнути затримок. Ця вимога також справедлива під час роботи з великими масивами даних, де швидке стиснення дає змогу знизити навантаження на мережу, у такий спосіб знижуючи ризик її перевантаження, а також потенційні вразливості, які стосуються такого сценарію.

У додатках, не націлених на роботу в реальному часі, швидкість виконання алгоритму посідає не першу роль, що, своєю чергою, дає змогу застосувати складніші та ресурсоемніші алгоритми для досягнення більшого значення ступеня стиснення. [62]. Більший коефіцієнт стиснення означає, що дані можна зберігати або передавати швидше, що може бути особливо важливим у системах з обмеженою ємністю зберігання або пропускнуою здатністю. Звідси впливає наступна закономірність: простіші алгоритми можуть бути швидшими, але можуть не забезпечувати такого високого ступеня стиснення або такої високої якості, як складніші алгоритми. Чималу роль відіграють обчислювальні ресурси обладнання, на якому виконуються

алгоритми стиснення. Наприклад, обчислювальні ресурси призначеного для користувача настільного комп'ютера можуть значно перевершувати ті, що є на борту літального апарата, де розміри і ресурси вкрай обмежені.

Як підсумок швидкість стиснення і декомпресії має вирішальне значення, особливо в додатках, де потрібен швидкий доступ до даних або їх передача. Швидкість може залежати від складності алгоритму стиснення і доступних обчислювальних ресурсів. Що стосується ступеня стиснення, то вона так само залежить від додатка, проте, в даному випадку це завжди компроміс між якістю фінального зображення та його розміром. Наприклад, у додатках, де простір для зберігання даних вкрай обмежений, може знадобитися високий ступінь стиснення. З іншого боку, у додатках, де якість даних має першорядне значення, може бути прийнятним нижчий ступінь стиснення.

Щоб задовольнити ці вимоги, потрібно придивитися до використання сучасних методів стиснення, прикладом таких може послужити BPG кодер, який потенційно може задовольнити більшість з наведених вище вимог. Основою так вважати служить хоча б той факт, що він розроблений на основі підмножини HEVC (відкритий стандарт стиснення відео), що може свідчити про його швидкодію. Це також дає підґрунтя вважати, що він може надати великий коефіцієнт стиснення, а наявність параметра контролю стиснення (що має назву Q), який змінюється у межах від 1 до 51, дає можливість налаштування ступеня стиснення, що дасть змогу досягти потрібного компромісу між якістю стисненого зображення та КС.

1.4 Особливості стиснення з втратами зображень, що вражені шумом

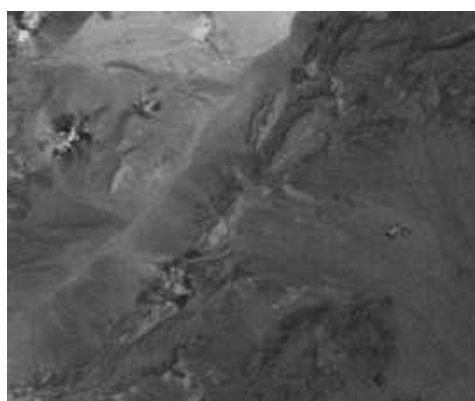
Крім уже згаданої залежності між контентом на зображенні (його складністю) і результатом стиснення з втратами, варто згадати ще не менш важливу характеристику зображення, а саме наявність на ньому шуму. Присутність шуму на зображенні може негативно позначитися на результаті стиснення.

Наявність шуму може проявлятися по-різному [63], [64]: він може бути видимим або ні. У разі, якщо шум є помітним на вхідних зображеннях, він може суттєво

вплинути на їхню якість [65], [66]. Приклади уражених шумом зображень приведені на рисунку 1.3.



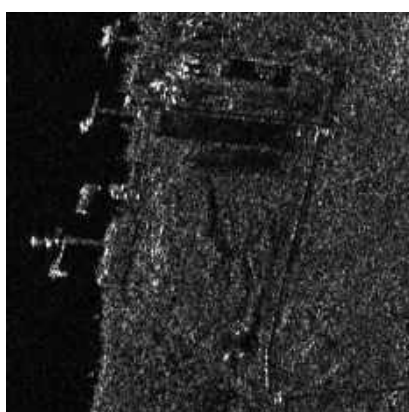
а



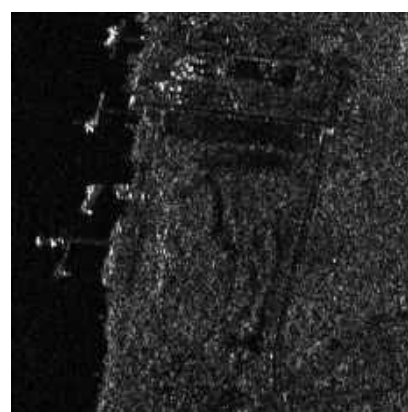
б

Рисунок 1.3 – Два приклади компонентних зображень, що вражені шумом у гіперспектральних даних AVIRIS

Ще одним прикладом можуть послужити зображення, отримані за допомогою радіолокаторів з синтезованою апертурою (РСА), які дають змогу отримати радіолокаційні зображення поверхні Землі незалежно від метеоумов і часу доби, однак ціною таких можливостей є наявність шуму на отриманих зображеннях. Приклад таких зображень представлено на рис.1.4.



а



б

Рисунок 1.4 – Приклади зображень, отриманих за допомогою двохполяризаційного РСА

Крім того, раніше згадувалося, що дистанційне зондування - не єдина сфера, де присутній шум на отриманих зображеннях: так, медичні зображення також страждають від наявності на них шуму. Приклади медичних зображень, на яких присутній видимий шум, представлено на рис. 1.5.

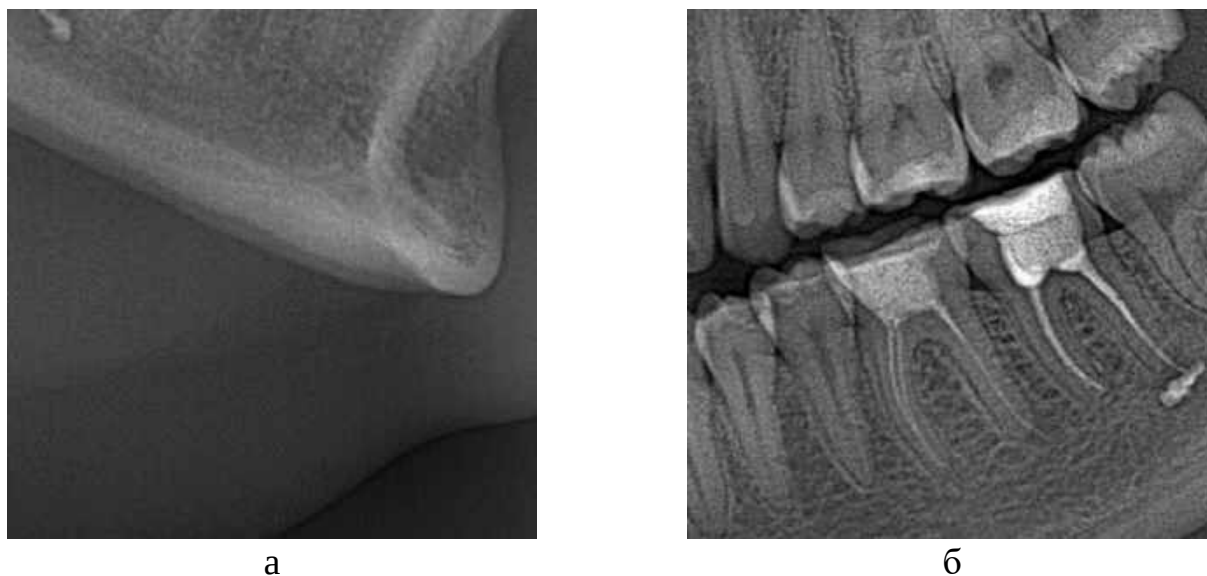


Рисунок 1.5 – Приклади медичних зображень, що вражені шумом

Такий шум маскує деталі зображень, тому виникає необхідність працювати з ним (враховувати та придушувати). Для ефективної боротьби з шумом знання про його тип може значно вплинути на результат стиснення. У правильній класифікації типу шуму може допомогти знання про причини його виникнення, оскільки кожен з таких типів має характерні особливості та закономірності.

1.4.1 Причини виникнення шуму

Шум на зображенні може виникати з низки причин [67-70], часто це пов'язано з електронним шумом, який може виникати в сенсорі, крім того, сенсор може вносити шум через нестачу освітлення або ж через температуру. Це якраз є одним із джерел появи адитивного білого гаусового шуму на зображенні. Крім причин, пов'язаних з апаратурою і навколишнім середовищем, шум може так само виникати через "програмні" причини, наприклад через помилки аналого-цифрового перетворення або

помилки бітів під час передачі. Варто так само зазначити, що залежно від області отримання зображення, шум, представлений на них, може мати свої характерні причини виникнення [71], [72]. У зв'язку з цим варто розглянути популярні напрямки, в яких часто використовуються зображення.

Почати можна із зображення дистанційного зондування, які часто можуть бути схильними до впливу шуму [67], [73]. Наведемо кілька прикладів з області цієї галузі, коли шум помітний на отриманих зображеннях і може бути інтенсивним. Перший приклад - спекл, який завжди присутній на зображеннях, отриманих радаром із синтезованою апертурою (рис. 1.4). Інший приклад - компоненти мультиспектральних і гіпер-спектральних даних ДЗЗ, для яких вхідний PSNR для деяких компонент може становити близько 20-30 дБ (рис. 1.3). Це також може бути характерно для так званих зображень нічного освітлення [8].

Системи медичної візуалізації, як-от рентген і комп'ютерна томографія (КТ), використовують рентгенівські промені або комп'ютерну томографію для отримання зображень ділянки інтересу пацієнта, які реєструються шляхом вимірювання результуючого загасання [69], [74]. Функція щільності ймовірності шуму у зображеннях, отриманих цими системами, часто моделюється за допомогою розподілу Пуассона і широко відома як шум пуассонівський, дробовий шум, фотонний шум, шум Шотта або квантовий шум [75].

1.4.2 Типи шумів

Шум має різні причини появи, а також по-різному проявляється на зображенні, проте його можна класифікувати, виходячи з його властивостей. Спочатку варто відзначити, які типи шумів існують, а також, які їх особливості. На рис 1.6 представлено схему класифікації типів шуму (деякі з яких були згадані раніше).



Рисунок 1.6 – Різновиди шуму

Шум можна розділити на кілька груп: адитивний, сигнально-залежний (зокрема, мультіплікативний) та імпульсний. У свою чергу імпульсний шум також можна розділити на статичний і динамічний [76].

Почнемо з адитивного типу шуму. Адитивний шум - це небажаний сигнал, який додається до деякого реального сигналу. Найпоширенішим представником цього класу є адитивний білий гаусів шум (АБГШ, англ. AWGN). Свою назву цей шум отримав від того, що його функція щільності ймовірності відповідає нормальному або гаусову розподілу [77]. Вважається однією з найпростіших моделей шуму, проте це хороша стартова точка в дослідженнях, оскільки навіть шум такої моделі може негативно вплинути на вихідне зображення. Даний тип шуму, як і будь-який інший, може значно знизити помітність деталей на зображенні, а в разі подальшого позбавлення від цього шуму фільтри можуть змазати дрібні ділянки і деталі, які були вражені цим шумом [78]. Оскільки шум адитивний, то зображення, уражене шумом, може бути представлено таким чином:

$$I_{ij}^{\text{ш}} = I_{ij}^{\text{іст}} + n_{ij}, \quad (1.8)$$

де $I_{ij}^{\text{ш}}$ – отримане спотворене AWGN зображення,

i, j – індекси пікселя,

$I_{ij}^{\text{іст}}$ – істинне неспотвореного шумом зображення,

n_{ij} – значення AWGN, що має нульове середнє значення та дисперсію σ^2 .

Мультиплікативний шум - це небажаний сигнал, який замість того, щоб додаватися, множиться на деякий справжній сигнал. Спекл-шум - найкращий приклад мультиплікативного шуму. Спекл-спекл виникає внаслідок інтерференції багатьох хвиль однієї й тієї самої частоти, кожна з яких має випадкові фази та амплітуди. Ці хвилі складаються, утворюючи результуючу хвилю з випадковою амплітудою (і, отже, інтенсивністю). У системах візуалізації спекл виникає внаслідок когерентного опрацювання зворотно розсіяних сигналів від кількох розподілених цілей. Цей тип шуму має таку модель:

$$I_{ij}^{\text{ш}} = I_{ij}^{\text{іст}} \cdot \mu_{ij}, \quad (1.9)$$

де $I_{ij}^{\text{ш}}$ – отримане спотворене спекл-шумом зображення;

μ_{ij} – випадкова мультиплікативна компонента, яка має математичне очікування m_μ , що дорівнює, 1 і дисперсію σ_μ^2 .

Імпульсний випадковий або піко-подібний шум - до такого типу можна віднести «Сіль і Перець». Він проявляється у вигляді випадкових чорних і білих пікселів, тобто в темних областях будуть світлі пікселі, а в світлих - темні. Звідси він і отримав свою назву, оскільки для напівтонового зображення значення уражених пікселів можуть набувати значення або 0, або 255 для чорного та білого відповідно [76].

Останній тип не є строго адитивним, мультиплікативним або імпульсним, а радше є формою невизначеності, пов'язаною з вимірюванням самого сигналу. Такий тип шуму називають сигнально-залежним [70], [79]. Прикладом такого шуму є дробовий шум або пуассонівський шум. Сила шуму Пуассона залежить від інтенсивності пікселів - пікселі з високою інтенсивністю створюють більшу дисперсію шуму. Іншими словами, це можна розглядати так: дисперсія шуму дорівнює середньому значенню зображення. Для моделювання цього шуму для досліджень можна скористатися Гауссовою апроксимацією розподілу Пуассона [80].

$$P(\mu) \approx N(\mu, \mu) \quad (1.10)$$

де μ – очікуване значення сигналу. Середнє значення і дисперсія розподілу Гауса рівні, і вони дорівнюють середньому значенню випадкової величини Пуассона.

Продемонстровані вище моделі можна використовувати для створення зображень, уражених шумом, для використання в подальших дослідженнях, наприклад ефективності роботи кодера і вплив шумів на його роботу. Приклади таких зображень наведено на рисунку 1.7.

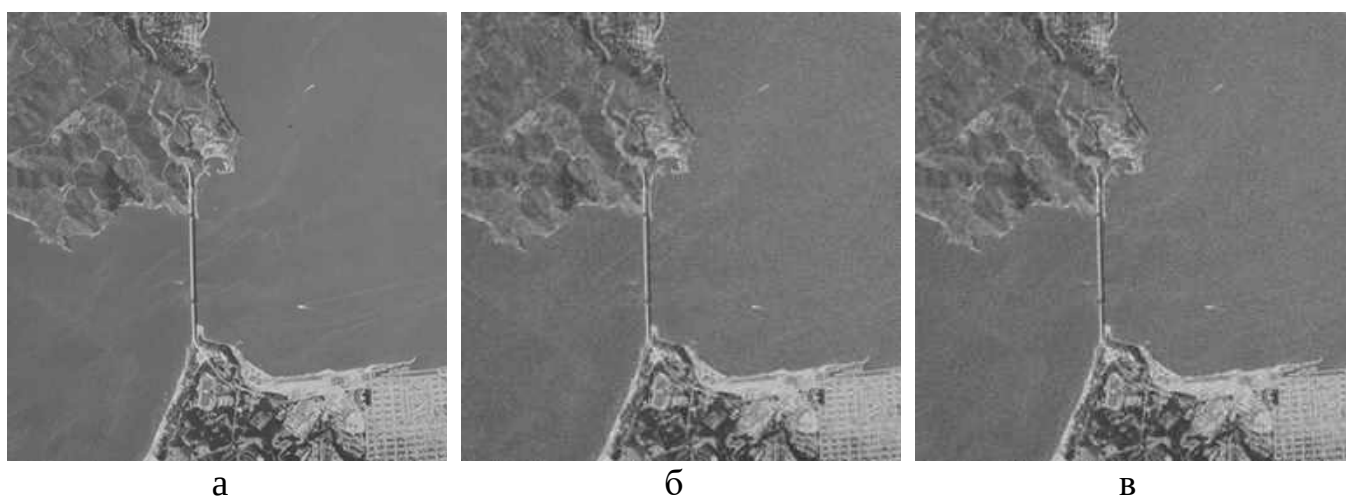


Рисунок 1.7 – Зображення дистанційного зондування Фріско: а) без шуму, б) з АБГШ із дисперсією 100, в) з шумом Пуассона

На цьому зображенні видно, як вплинули деякі зі згаданих типів шумів на вихідне зображення. Особливо добре помітний шум на однорідних ділянках зображення. Також варто відзначити той факт, що отримане погіршення якості може суттєво вплинути на сприйняття людиною зображення, а також детектування об'єктів на зображенні різними алгоритмами [81]. Крім того, це призводить до можливої необхідності фільтрації [82-84] і може позначитися на ефективності стиснення та подальшому зберіганні або передачі зображень.

1.4.3 Вплив шуму на стиснення з втратами

Раніше було згадано більшість негативних ефектів, що може принести наявність шуму на зображенні. Однак не було розглянуто, як саме він вплине на стиснення з втратами. Особливості стиснення з втратами для уражених шумом зображень були виявлені кілька десятиків років тому [85]. Найважливішими є дві особливості. По-перше, в результаті стиснення з втратами виникає ефект фільтрації шуму [86]. Фільтрація досягається через операцію квантування, оскільки за своєю суттю вона є наближенням до порогової операції, і відповідна схема квантування зможе забезпечити шумозаглушення. Порогові методи порівнюють вхідні дані із заданим пороговим значенням і встановлюють його рівним нулю, якщо його величина менша за порогове значення. Ідея полягає в тому, що коефіцієнти, незначущі відносно порога, ймовірно, зумовлені шумом, тоді як значущі коефіцієнти є важливими сигнальними структурами.

Друга особливість походить від згаданого ефекту фільтрації. Завдяки цьому ефекту для деяких зображень і властивостей шуму можна спостерігати оптимальну робочу точку [8]. Далі буде більш детально розглянуто її особливості, наразі достатньо знати, що під OPT розумітимуть параметр, який керує стисненням (ПКС), для кодера, використовуваного в такий спосіб, щоб спостерігалася мінімальна "відстань" між стиснутим і вільним від шуму зображеннями. Іншими словами, в OPT якість стисненого зображення краща, ніж поза областю даної точки і ближча за візуальною якістю до нестисненого не зашумленого (істинного) зображення. Ця відстань може бути охарактеризована різними способами, починаючи з традиційних показників, таких як середня квадратична похибка (MSE) і пікове відношення сигнал/шум (PSNR), і закінчуючи метриками візуальної якості, такими як, наприклад, PSNR-HVS-M або MS-SSIM.

Щоб побачити цю точку (значення), можна застосувати наступний метод: для симуляції необхідно до істинного зображення (не ураженого шумом) додати шум, після чого застосувати стиснення із втратами. Тоді можна обчислити показники

(значення метрик Metr_{ct}) для всього діапазону значень ПКС між стисненим зображенням і відповідним істинним зображенням.

Для того щоб продемонструвати приклади таких залежностей, було обрано шість напівтонових зображень з області дистанційного зондування, мініатюри цих зображень представлено на рисунку 1.8, і виконано операцію, описану вище.



Рис 1.8 – Мініатюри використаних зображень: а) diego б) frisco в) fr01 г) fr02 д) fr03 е) fr04

Приклади залежностей параметра контролю якості для шести тестових зображень представлені на рис. 1.9. У цих прикладах показано залежності, отримані для VP8 кодера для всіх значень ПКС (в даному кодері - це Q , яке може бути встановлене в межах від 1 до 51). Щодо моделі шуму, то використовувався згаданий раніше адитивний білий гаусів шум.

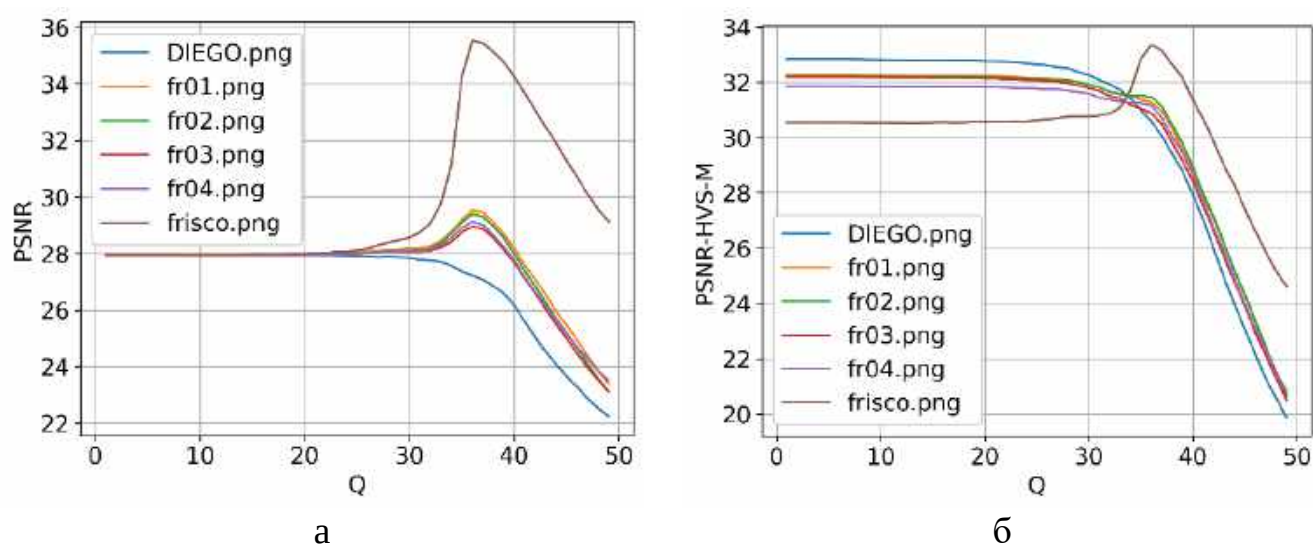


Рисунок 1.9 – Залежності $\text{PSNR}_{ct}(Q)$ (а) та $\text{PSNR-HVS-M}_{ct}(Q)$ (б) для дисперсії шуму, що дорівнює 100, для шести тестових зображень

Як видно із залежностей для двох метрик, спостерігається область ($Q \approx 35$), у якій для деяких зображень якість зростає і сягає максимуму. Це трапляється для нескладних зображень і частково для зображень середньої складності, тоді як для складних зображень (Diego) спостерігається монотонний спад у якості зображення. Для таких зображень оптимальним місцем для стиснення є точка перед цим спадом. Для решти зображень (простих і середніх) оптимальним параметром для стиснення є той, де спостерігається пік.

Оскільки в оптимальній робочій точці (ОРТ) можливо отримати не тільки великий КС і високу якість для випадку обговорених вище зображень, то бажано використовувати стиснення саме в ОРТ. Однак варто врахувати той фактор, що, не маючи істинного зображення, потрібно якось попередньо дізнатися про можливість існування ОРТ для конкретного кодера і обраного зображення, а також позицію цієї точки (або, іншими словами, значення ПКС для ОРТ). Для кількох кодерів (JPEG, JPEG2000, AGU [87], ADCT [88]) методи розрахунку (визначення) ПКС в ОРТ були раніше розроблені [89], але для кодеру BPG вони не були запропоновані, що обмежує можливості його використання для стиснення зображень, що спотворені шумом.

1.5 Існуючі кодери

Існує велика кількість кодерів і їхніх модифікацій, які використовують різні методи, а також перетворення, зокрема, згадані раніше. Залежно від області або ж додатка, в якому вони використовуються, характеристики даних кодерів можуть відрізнятися. Різнитися очевидно можуть і результати стиснення, наразі не існує ідеального кодера, який міг би надати великий коефіцієнт стиснення, але й не «зачепити» якість зображення. У зв'язку зі згаданим вище варто детальніше розглянути різні кодери, щоб виявити їхні сильні та слабкі місця, оскільки це може не тільки допомогти у виборі ефективного методу стиснення, а й дасть змогу окреслити області, які можуть бути поліпшені.

Як згадувалося раніше, якість роботи різних кодерів можна оцінити за коефіцієнтом стиснення, а також якістю стисненого зображення, які зазвичай представляють у вигляді так званих кривих швидкість-спотворення (rate-distortion curves - RDC) – дивись приклади на рис. 1.9. Саме їх головним чином будемо використовувати під час подальшого аналізу.

Загалом кодери стиснення з втратами можна розділити за перетворенням, які вони використовують. У такому разі можна, наприклад, отримати такі групи: ті, що використовують дискретне косинусне перетворення (ДКП), вейвлет-перетворення (ДВП) або ж методи машинного навчання.

1.5.1 Засновані на ДКП

В основі цих кодерів, як очевидно з назви, лежить дискретне косинусне перетворення. Це метод на основі ортогонального перетворення, який розбиває зображення на блоки та розраховує просторовий частотний спектр. Він виражає кінцеву послідовність точок даних у термінах суми косинусних функцій, що коливаються на різних частотах. Це перетворення використовується в стисненнях із втратами у випадках, коли високочастотні компоненти можуть бути відкинуті.

Дане перетворення розраховується зазвичай для зображень у блоках пікселів 8x8. Однак це не є непорушним законом і існують методи стиснення з іншими значеннями розміру блоків, наприклад можна зустріти розмір блоку 32x32 пікселів. Для розміру блоку NxN перетворення можна розрахувати таким чином [90]:

$$D(i, j) = \frac{1}{\sqrt{2N}} C(i) C(j) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p(x, y) \cos \left[\frac{(2x+1)i\pi}{2N} \right] \cos \left[\frac{(2y+1)j\pi}{2N} \right] \quad (1.11)$$

де $p(x, y)$ - є блок з x, y -ми елементами зображення, представленого матрицею p , N – розмір сторони блоку, в якому розраховуватиметься ДКП.

Зворотне перетворення у свою чергу може бути записано так:

$$p(i, j) = \frac{1}{\sqrt{2N}} C(i)C(j)D(x, y) \cos \left[\frac{(2x + 1)i\pi}{2N} \right] \cos \left[\frac{(2y + 1)j\pi}{2N} \right] \quad (1.12)$$

Після ДКП-перетворення перші частоти в наборі є зазвичай найбільш значущими; інші, залежно від допустимої втрати роздільної здатності, часто можуть бути видалені. Як правило, масив коефіцієнтів перетворення 8x8, в якому елемент (0,0) (верхній лівий) є ДК (нульовою) складовою, а записи зі зростаючими значеннями вертикального та горизонтального індексів представляють вищі вертикальні та горизонтальні просторові частоти.

Після цього отримана матриця ДКП коефіцієнтів готова до подальшої операції квантування. Квантування досягається шляхом ділення кожного елемента в трансформованій матриці (отриманій після ДКП) на елемент із матриці квантування, після чого отримане значення округлюється до найближчого цілого числа. В отриманій матриці значення, що знаходяться на її початку (верхньому лівому кутку), відносяться до чутливих для людського ока низьких частот. Так само варто зазначити, що ця матриця може містити велику кількість нулів, це відноситься до менш важливої інформації, яка може бути відкинута (ключовий момент стиснення з втратами) [90].

Фінальним етапом стиснення є кодування даних в отриманих матрицях. Від його ефективності суттєво залежать загальні характеристики стиснення [91]. В даний час широко застосовуються такі методи кодування: кодування Хаффмана, що використовується в таких форматах як JPEG, MPEG; арифметичне кодування, яке також використовується в JPEG; алгоритм Лемпеля-Зіва, що використовується в таких форматах як TIFF, GIF, PNG; кодування довжин серій, що також використовується в TIFF, а окрім цього і в BMP. [92].

Що стосується отримання розтиснутого зображення, то для цього всі вищеописані операції проводять у зворотному порядку, а саме декодування, де-квантування та застосування зворотного перетворення (для цього сімейства кодерів - це зворотне дискретне косинусне перетворення).

Цей докладний опис стиснення з втратами необхідний для того, щоб зрозуміти різницю між кодерами, що базуються на ДКП. Серед кодерів, що належать до цього

сімейства, можна виділити таких представників: JPEG, WebP, HEIF і BPG. Розглянемо особливості деяких із них.

Першим розглянемо JPEG як один із найпоширеніших і загальноприйнятий формат стиснення зображень. Цей кодер має низку особливостей. Наприклад, вже зрозуміло, що під час квантування багато коефіцієнтів прирівнюються до нуля, тому JPEG користується цим і кодує коефіцієнти квантування в зигзагоподібній послідовності. Перевага полягає в об'єднанні відносно великих серій нулів, які дуже добре стискаються. Крім того, присутні 4 варіанти кодування: послідовний, прогресивний, ієрархічний і без втрат. У першому варіанті кодування в послідовності проводиться зліва направо і зверху вниз, щоб гарантувати, що кожен піксель кодується тільки один раз. Прогресивне кодування зазвичай досягається шляхом багаторазового сканування. При ієрархічному кодуванні зображення стискається до декількох рівнів роздільної здатності. Очікується, що під час кодування без втрат зображення міститиме повну деталізацію під час розпакування.

HEIF або ж High Efficiency Image File Format (високоєфективний формат графічних файлів) [93]. Даний формат розроблений для зберігання і передачі зображень, який часто може забезпечити більший коефіцієнт стиснення при тій самій якості, що і JPEG. Досягається цей приріст завдяки використанню ефективного алгоритму кодування, іменованого HEVC (High-Efficiency Video Coding - високоєфективне кодування відео), так само відомий як H.265. HEVC досягає цього завдяки деяким особливостям, таким як більші розміри блоків, поліпшена компенсація руху об'єктів і режимів прогнозування. Завдяки цьому HEIF може використовувати різні розміри блоків для розрахунку ДКП від 4x4 до 32x32 пікселів у той час, як у JPEG це значення дорівнює 8x8 пікселів. Так само варто зазначити, що обчислення ДКП можуть не завжди підходити до систем з обмеженими ресурсами, а щоб розв'язати цю задачу, в HEIF використовують наближення, щоб знайти баланс між якістю і складністю обчислення.

Одним із відносно нових кодерів, що добре зарекомендували себе, є BPG кодер, який позиціонує себе як заміна загальноприйнятому JPEG. Заснований цей кодер на вже згаданій підмножині відкритого стандарту стиснення відео HEVC [94]. З

основних особливостей можна згадати таке: високий динамічний діапазон, що досягається вбудованою підтримкою 8-14 біт на канал; більший коефіцієнт стиснення порівняно з JPEG за таких самих показників якості стисненого зображення; підтримка найпоширеніших форматів кольоровості, а саме відтінки сірого, YCbCr 4:2:0, 4:2:2, 4:4:4:4. Крім того, в останніх дослідженнях цей кодер був порівняний з іншими представниками ДКП кодерів. [95]. Використовуючи метрику якості PSNR-HVS-M, автори продемонстрували криві залежності метрики від коефіцієнта стиснення.

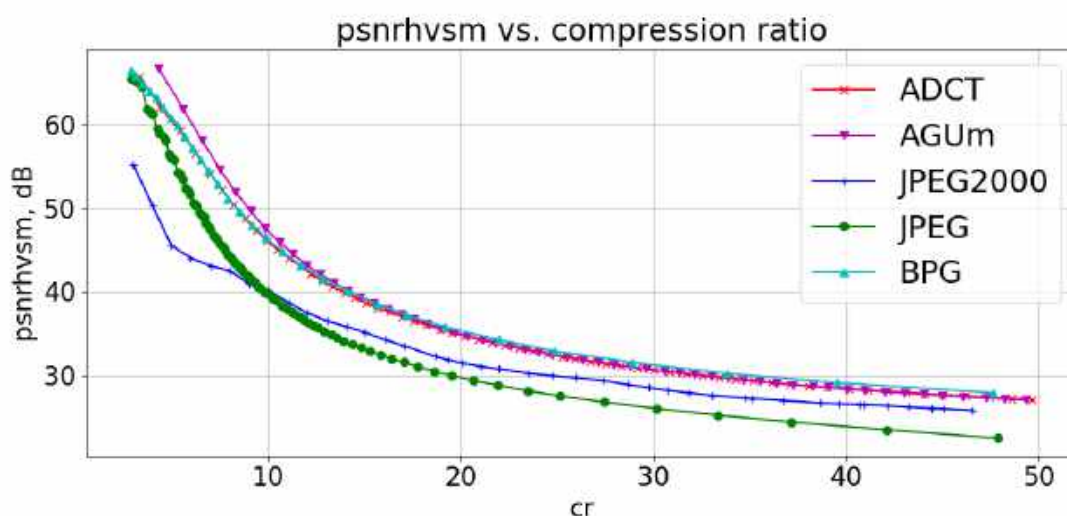


Рисунок 1.10 – Залежність PSNR-HVS-M від коефіцієнта стиснення для тестового зображення (фрагмент стоматологічного медичного зображення розміром 512×512 пікселів, як один із поширених прикладів зображень, що можуть бути схильні до шуму), стисненого різними кодерами

Як видно з отриманих залежностей (рис 1.10), BPG дійсно перевершує не тільки JPEG, а й інших представників ДКП кодерів (крім ділянки з малими значеннями коефіцієнта стиснення, в якій є лідером AGUm – варіант кодеру AGU, орієнтований на візуальну якість). Спираючись на дані залежності, можна зробити висновок, що BPG є перспективним кодером і варто детальніше дослідити його характеристики, а також можливість застосування методів визначення оптимальної робочої точки для даного кодера та її прогнозування.

1.5.2 Засновані на Вейвлетному перетворенні

Характерною особливістю кодерів, заснованих на ДКП, є блочні артефакти, які утворюються на стисненому зображенні. Як альтернативу було запропоновано використовувати дискретне вейвлетне (малохвильове) перетворення.

Однак перед тим, як заглиблюватися в стиснення, що ґрунтується на дискретному вейвлет-перетворенні, для початку варто розібратися, що собою являють вейвлети. Вейвлети - це сигнали, які локальні в часі та масштабі, а також мають хвилеподібну форму [96]. Тобто вейвлет - це сигнал ефективно обмеженої тривалості, середнє значення якого дорівнює нулю. Багато вейвлетів також мають властивість, ідеальну для компактного представлення сигналу: ортогональність. Ця властивість гарантує, що дані не будуть представлені надмірно. Сигнал можна розкласти на безліч зсунутих і масштабованих уявлень вихідного вейвлета. Вейвлет-перетворення можна використовувати для розкладання сигналу на складові вейвлети. Щойно це буде зроблено, коефіцієнти вейвлетів можна буде прорідити, щоб видалити деякі деталі. Вейвлети мають велику перевагу: вони здатні відокремлювати дрібні деталі сигналу. Маленькі вейвлети можуть використовуватися для виділення дуже дрібних деталей у сигналі, тоді як великі вейвлети можуть визначати грубі деталі.

Далі будуть розглянуті представники даного сімейства кодерів, що використовують у своїй основі дискретне вейвлет-перетворення. Серед таких кодерів можна виділити JPEG 2000, EZW, SPIHT і EBCOT.

JPEG 2000 [97] є не тільки продовженням і поліпшенням оригінального методу, але ще й використовує іншу архітектуру кодування на основі вейвлетів як основу для створюваної системи кодування. Початкове рішення, було натхненне алгоритмом вейвлетного квантування з трелліс-кодом (WTCQ), заснованим на об'єднанні вейвлетів і квантування з трелліс-кодом (TCQ), після чого його вдосконалили під час серії основних експериментів і як результат з'явилася рекомендація ISO 15444/ITU-T Recommendation T.8000, відома як стандарт JPEG2000. [98]. Основні особливості: прогресивна передача і декодування на основі масштабованості роздільної здатності

за рахунок точності пікселів; довільний доступ до просторових областей (або областей інтересу), а також до компонентів; стійкість до бітових помилок.

SPIHT [99], [100] — це ефективний алгоритм стиснення зображень на основі вейвлетів. Він використовує розрідженість вейвлет-коефіцієнтів для досягнення високих коефіцієнтів стиснення при збереженні хорошої візуальної якості. SPIHT працює шляхом ітеративного поділу вейвлет-коефіцієнтів на значущі та незначущі набори з подальшим ентропійним кодуванням. Він широко використовується в додатках стиснення зображень.

EZW [101] — ще один популярний алгоритм стиснення на основі вейвлетів. Він послідовно кодує вейвлет-коефіцієнти таким чином, щоб забезпечити прогресивне декодування. Він починається з грубого наближення зображення і уточнює його, додаючи детальну інформацію. EZW особливо корисний для прогресивної передачі та масштабованого кодування.

EBCOT [102] — це основний механізм кодування, який використовується в JPEG2000. Він поєднує в собі ДВП з контекстним моделюванням і арифметичним кодуванням. EBCOT ефективно кодує вейвлет-коефіцієнти шляхом адаптивного усічення бітових площин залежно від їхньої значущості. Він забезпечує чудову продуктивність стиснення і підтримує такі функції, як кодування області, що цікавить, і стиснення без втрат.

Варто зазначити, що кодери, засновані на вейвлетному перетворенні, хоч і мають перевагу над JPEG, проте все ще не можуть продемонструвати кращу якість за того самого коефіцієнта стиснення для інших методів що засновані на ДКП (дивись дані на рис. 1.10). Це твердження ґрунтується на роботі [95] де було показано нездатність кодерів заснованих на вейвлетному перетворенні перевершити вже згаданий раніше BPG кодер.

1.5.3 Засновані на застосуванні машинного навчання

У сучасному світі машинне навчання та нейронні мережі набули стрімкого розвитку і поступово впроваджуються в більшість галузей нашого життя. Не оминули

вони стороною і таку важливу галузь як обробка зображень, зокрема методи стиснення з втратами [103]. Однак початковий задум не змінився, методи, засновані на нейронних мережах, являючи собою нову віху розвитку технологій, прагнуть, як і раніше, перевершити своїх попередників, покращуючи фінальний результат або розв'язуючи проблеми, характерні для загальноприйнятих методів (блокові артефакти, тривалий час обчислень і т.д.).

Наприклад, у [104] використовується автокодувальник для стиснення зображення. Для порівняння, зазвичай характеристики на кривій швидкість-спотворення можна регулювати шляхом зміни кроку квантування, однак у випадку з автокодувальниками це потребуватиме навчання одного перетворення на кожну точку спотворення-швидкість за заданого розміру кроку квантування. Однак попри тривалий час на навчання результати можуть перевершувати JPEG, JPEG-2000. Так автори [105] представили підхід, який може змагатися зі згаданими раніше кодерами. Хоч автокодувальники можна використовувати для розв'язання проблеми в потребі в гнучкіших кодерах, однак, їх важко оптимізувати напряду через притаманну їм не диференційованість втрат під час стиснення.

В [106] автори використовують підхід, що складається з кодера з нелінійним перетворенням, рівномірного квантувальника, декодера з таким самим нелінійним перетворенням і модуля постобробки. Останній елемент потрібен для видалення артефактів стиснення і розмиття із зображення.

В [107] пропонується наскрізний підхід на основі згорткових нейронних мереж для стиснення мультиспектральних зображень. Мультиспектральне зображення вводиться в кодер у вигляді тривимірного тензора, що дає змогу вивчити особливості просторово-спектрального поєднання. Ентропійний кодер на основі арифметичного кодування використовується для подальшого зменшення обсягу даних, а декодовані проміжні ознаки дають розпаковане тривимірне гіперспектральне зображення. Автори стверджують, що результати експериментів показують значні поліпшення в MS-SSIM порівняно з традиційними методами за низьких швидкостей передачі даних.

Однак збільшення обчислювальної складності неминуче для методів, заснованих на нейронних мережах. Їх важко реалізувати в деяких додатках, таких як мобільні

платформи для вбудованих додатків і платформи з обмеженими обчислювальними ресурсами, наприклад, супутники дистанційного зондування.

1.5.4 Порівняння ефективності кодерів

Як видно, існує велике розмаїття в підходах до стиснення, це вилилося у великий вибір кодерів, які можна використовувати для стиснення зображення. Ще одним кроком до вибору правильного кодера, який підійде під конкретне завдання, є порівняння плюсів і мінусів представлених на даний момент рішень.

Почати порівняння можна з перетворень/підходів, що лежать в основі кодера. Виходячи з описаного вище видно, що загалом кодери можна розділити за підходами, що лежать в їхній основі. Вибір правильного напрямку (підходу) дасть змогу істотно скоротити різноманітність вибору, відсіявши варіанти, які не можуть задовольнити поставлені вимоги.

Першими порівнянню підлягають дві великі групи кодерів, засновані на ДКП і ДВП. Дискретне косинусне перетворення працює в частотній області, в той час як вейвлет-перетворення має додаткову властивість просторової локалізації, тобто, крім інформації про частоту, так само зберігається інформація про місцезнаходження. Обробка сигналу теж відбувається по-різному, у ДКП обробляється весь сигнал як єдине ціле, розділяючи сигнал на блоки для стиснення, у ДВП забезпечується просторова локалізація, що дає змогу вибрати особливості із сигналу різних масштабів і місць. Щодо обчислювальних здібностей, то вейвлет-перетворення потребує більших вільних ресурсів для обчислення, в той час як дискретно косинусне перетворення потребує значно менше. Не дивно, що сучасні методи стиснення відео базуються на ДКП [108].

Якщо додати до цього нові алгоритми, засновані на нейромережах, можна знайти такі відмінності. По-перше, замість трансформації в частотну область ці алгоритми використовують моделі глибокого навчання для вивчення компактного представлення зображення. Далі, що стосується самого стиснення, методи, що базуються на навчанні, можуть досягти більшого коефіцієнта стиснення і кращої

якості зображення, однак, це призводить до сильного ускладнення моделі. З цього випливає наступний пункт, а саме складність обчислення безпосередньо залежить від складності конкретної моделі, однак, зазвичай, методи, що базуються на нейронних мережах, потребують більших обчислювальних здібностей, особливо на етапі навчання. При аналізі літератури не знайдено публікацій щодо стиснення зображень, що спотворено завадами, такими методами.

Як результат можна зробити кілька висновків. ДКП є ефективним в обчислювальному відношенні, але може втратити деяку інформацію. ДВП інколи забезпечує кращий ступінь стиснення, але вимагає більшої обчислювальної потужності. Методи на основі НМ дають змогу домогтися вищої якості та меншого розміру файлів, але вони складніші й потребують більших обчислювальних ресурсів.

Отже, оптимальним варіантом для стиснення можна вважати кодери на основі ДКП, однак, залишилося не закритим питання, який все-таки кодер використовувати. Раніше вже коротко говорилося, що BPG має низку переваг щодо інших кодерів, так само згадувалося про особливості стиснення з втратами зображень, що вражені шумом. Виходячи з цього, ще одним пунктом порівняння можна виділити, як кодери справляються зі стисненням із втратами зображень що містять в собі шум. Автори [95] представили залежності метрики якості від коефіцієнта стиснення для різних зображень, уражених шумом, приклади отриманих залежностей подано на рисунку 1.11.

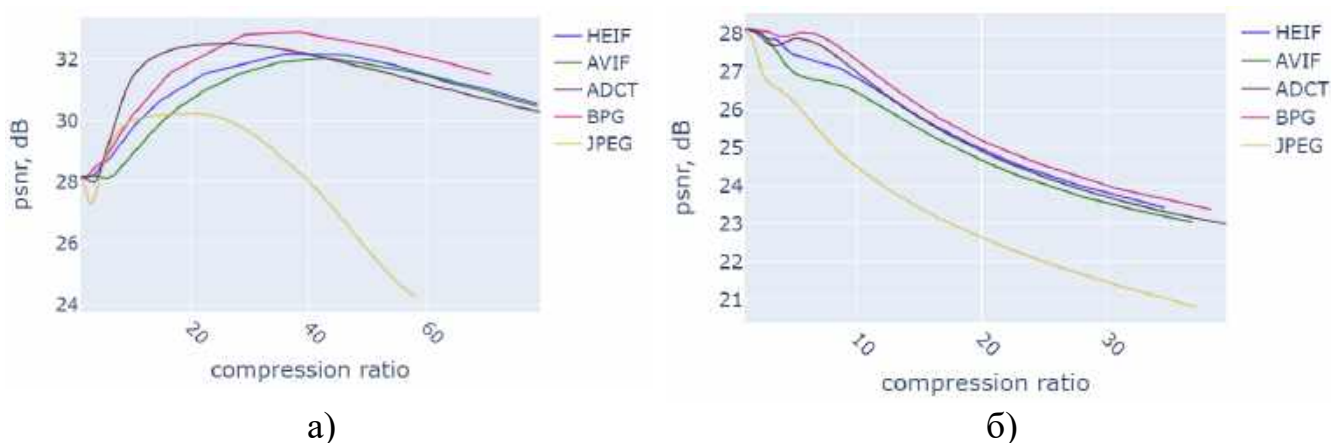


Рисунок 1.11 – Залежності $PSNR_{ct}$ від коефіцієнта стиснення для простого а) та складного б) зображень, уражених шумом стиснутих різними кодерами

Як видно із залежностей, BPG кодер може надати кращу якість при більшому коефіцієнті стиснення як для складних, так і для простих зображень в умовах ураження зображень шумом. Дана особливість цього кодера робить його чудовим вибором при використанні його у випадках роботи з зображеннями ураженими шумом. Залишається відкритим питання, чи можливе прогнозування ОРТ методами, аналогічними до інших ДКП кодерів [109], або ж потрібно розробляти нові підходи до розв'язання цієї задачі.

1.6 Висновки до розділу та формування цілей дослідження

Беручи до уваги зростаючу кількість зображень, особливо дистанційного зондування, яке створює велику кількість зображень, що можуть бути багатоканальними або мати досить-таки велику роздільну здатність, стає очевидним потреба в ефективних методах стиснення з втратами. Так само варто приділити окрему увагу шуму, що присутній на зображеннях ДЗЗ та інших типів, який має великий вплив на візуальну якість зображень і може вплинути на подальшу обробку. Як результат, потрібно брати до уваги не тільки особливості самого стиснення з втратами, а й особливості зображень, а отже, виникає запит на розробку ефективного методу стиснення, основою якого може стати як знання про ОРТ, так і можливості прогнозування такої точки, а використання таких сучасних кодерів, як BPG, може значно підвищити ефективність виконання цього завдання. Враховуючи зазначені особливості та тенденції в галузі стиснення зображень із втратами, основними завданнями дослідження є:

- дослідження маловивчених особливостей BPG кодера, зокрема, характеристик стиснення як напівтонових, так і кольорових зображень;
- дослідження здатності даного кодера до роботи із зображеннями, ураженими різними типами шумів включно із сигнально-залежними варіантами;
- розробка методів прогнозування існування ОРТ і параметрів у ній для BPG кодера;

– розроблення методів прогнозування значень метрик якості, що оцінюють якість зображення, стисненого BPG кодером.

РОЗДІЛ 2

ДОСЛІДЖЕННЯ ХАРАКТЕРИСТИК СТИСНЕННЯ З ВТРАТАМИ VPG КОДЕРА

У даному розділі детально розглянуті особливості роботи VPG кодера в режимі роботи стиснення з втратами в разі застосування його до напівтонових і кольорових зображень. Так само розглянуто випадки ураження зображення, призначених для стиснення, різними типами шуму, зокрема адитивним білим гаусовим і пуассонівським шумами. До всього іншого було розглянуто і продемонстровано таке поняття, як оптимальна робоча точка і який внесок вона вносить у процедуру стиснення з втратами VPG кодером. Окрема увага приділена аналізу спотворень, що вносить безпосередньо сам кодер, а також, як вони поведуться в ситуаціях, коли на зображенні вже є один із вищезгаданих типів шуму.

2.1 Стиснення зображень в градаціях сірого

Відправною точкою в дослідженні кодера VPG стане розгляд його характеристик при роботі з напівтоновими зображеннями.

В першому розділі згадувалась головна вимога до стиску зображень з втратами, а саме отримання задовільної якості зображення при максимально великому коефіцієнті стиснення, як результат приходимо до завдання знаходження компромісу. Для вирішення даної задачі, як вже згадувалось раніше (в першому розділі), використовують криву швидкість-спотворення (rate distortion curve) або її аналоги.

Загальноприйнятий спосіб побудови кривої швидкість-спотворення — це порівняння оригінального зображення (що підлягає стиску) з відповідним стисненим зображенням, використовуючи метрику для різного значення параметра, що контролює стиск (для випадку з VPG кодером це параметр Q). Однак окрему увагу варто приділити порівнянням між стиснутим зображенням, що уражене шумом (оскільки, як згадувалося в першому розділі, зображення з шумом - це один із найпоширеніших випадків) та істинним зображенням без шуму. Дані залежності хоч

і неможливо отримати за межами симуляції, але вони можуть дати цінні дані про наявність особливостей, характерних для VPG кодера, включаючи потенційну наявність оптимальної робочої точки.

2.1.1 Вибір тестових зображень, метрик та моделі шуму

У першому розділі було розглянуто як поняття зображення, так і області їх отримання. Так само побічно було порушено поняття як складність зображень. Однак не було розглянуто, як саме трактувати таке поняття, як складність зображення.

Загалом можна розділити зображення на три категорії: складні зображення, прості зображення та ті, що не можна віднести до жодної з двох згаданих – зображення середньої складності. Щоб віднести якесь із зображень до певної групи необхідно розрахувати його складність. Один варіант - це оцінити візуально, а саме який відсоток зображення має в собі однорідні ділянки і який насичений різноманітними дрібними деталями або текстурними об'єктами. Текстурними об'єктами на зображенні визначається певна ділянка щільно розташованих пікселів, яка демонструє регулярність у появі елементів цієї структури та менший ступінь відмінностей значень пікселя від середнього значення для всієї області текстури, ніж фоновий контраст малих об'єктів. Якщо однорідних ділянок більше, ніж неоднорідних то зображення можна віднести до простих зображень і навпаки якщо неоднорідних ділянок більше, то зображення можна вважати складним. Приклад складного та простого зображення представлені на рис. 2.1.



Рис 2.1 – а) приклад складного зображення; б) приклад простого зображення.

З рисунку видно, що для складного зображення (рис. 2.1, а) характерна наявність великої кількості текстурних або дрібних деталей, зібраних в одній області, характерної для міської інфраструктури, також на зображенні присутня велика кількість тонких ліній, характерних для доріг. У той же час у другому випадку (рис. 2.1, б) можна помітити, що більшу частину зображення займають однорідна водна поверхня, в той час, як згадана міська інфраструктура займає малу частину зображення. Зображення, що мають приблизно однакове співвідношення однорідних та неоднорідних об'єктів, можна віднести до зображення середньої складності.

Оскільки надалі будуть використовуватися назви використаних зображень під час аналізу залежностей, то логічно буде навести мініатюри використовуваних зображень. Також варто згадати, що всі зображення мають розмір 512x512 пікселів і представлені тільки яскравісною компонентою в градаціях сірого, тобто 8 біт на піксель. Перший набір зображень, який буде використовуватися, вже було продемонстровано в першому розділі (мініатюри показані на рис. 1.8), але оскільки для можливості порівняння отриманих даних необхідно використовувати більше одного набору зображень на рис. 2.2 наведено мініатюри другого набору зображень.



Рис 2.2 – Мініатюри використаних зображень у другому наборі: а) baboon; б) barbara; в) goldhill; г) lenna; д) peppers

У першому розділі було згадано різні метрики, які можуть оцінити якість зображення. У цьому пункті будуть виділені ті представники, які будуть використані для подальшого дослідження особливостей стиснення з втратами BPG кодера.

Для аналізу роботи кодера з одноканальними або напівтоновими зображеннями будуть використовуватися наступні метрики: PSNR-HVS-M, MS-SSIM і PSNR. Вибір даних метрик обумовлений тим, що перші дві метрики, як показано в першому розділі, мають досить великий коефіцієнт кореляції, тоді як PSNR є однією з найпоширеніших метрик і послужить доцільною точкою для порівняння. Оскільки дані метрики можуть бути розраховані по-різному, необхідно внести ясність у позначення даних метрик. Для першого випадку, коли метрика розраховується між стисненими зображенням I_{ij}^c $i = 1, \dots, I_{Im}$, $j = 1, \dots, J_{Im}$ та зображенням з шумом I_{ij}^n $i = 1, \dots, I_{Im}$, $j = 1, \dots, J_{Im}$ позначення буде мати такий вигляд $\text{Метр}^{cn}(Q)$ для конкретного значення Q . У той же час для другого випадку, коли метрика розраховується між стисненим і істинним зображенням I_{ij}^t , $i = 1, \dots, I_{Im}$, $j = 1, \dots, J_{Im}$, вона матиме позначення $\text{Метр}^{ct}(Q)$.

2.1.2 Отримані залежності

Для початку слід розглянути залежності, отримані для згаданих вище метрик для BPG-кодера для всіх значень параметра контролю стиснення (у межах від 1 до 51) для вибраного набору зображень (рис. 1.8). Ці залежності представлені на рис. 2.3.

Метрики на цих залежностях розраховані між стисненими та зображеннями, ураженими шумом.

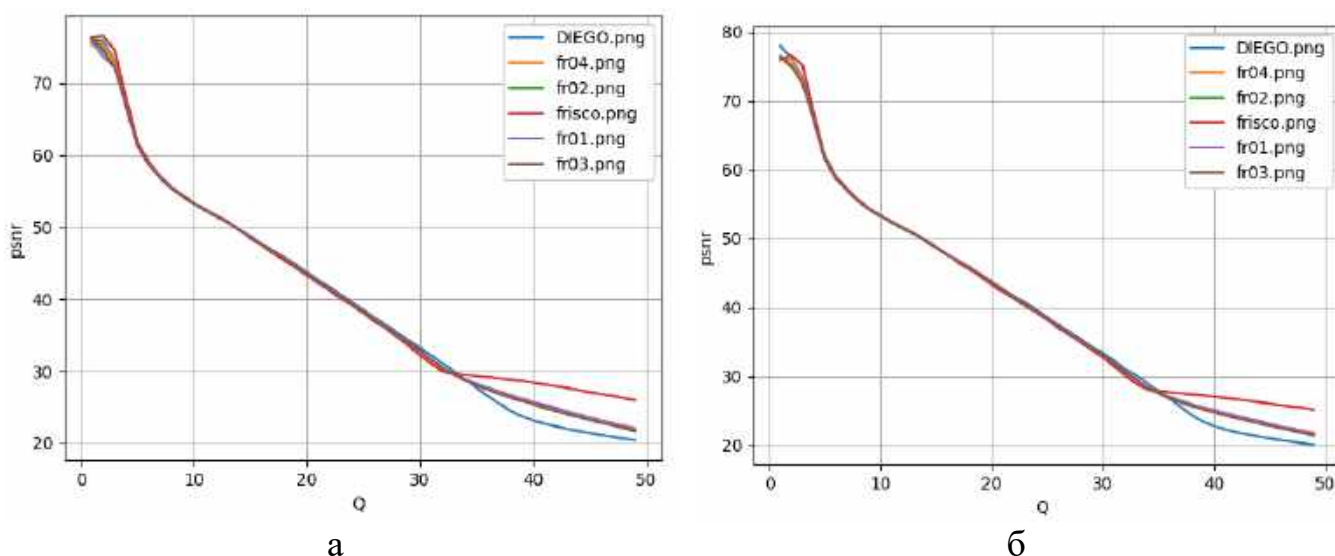


Рис 2.3 – Залежності аналізованих метрик $PSNR^{cn}$ від Q : для дисперсії шуму, що дорівнює 64 (а), для дисперсії шуму, що дорівнює 100 (б)

Як видно, на цих кривих можна виділити три основні зони. Перша - для $Q < 9$, де залежності нелінійні, а якість стиснених зображень характеризується дуже високими значеннями $PSNR^{cn}$, що означає майже повну відсутність спотворень. Потім, Q приблизно від 9 до $Q=33$, спостерігається практично лінійне зниження $PSNR^{cn}$ (приблизно як $63-Q$). Ця частина кривих відповідає невидимим спотворенням (при Q близько 28 і менше) та видимим спотворенням при $Q > 28$. Видимі спотворення пов'язані з фільтрацією шуму, і, частково, з погіршенням відображення змісту інформації. Третя область ($Q > 33$) характеризується відхиленням поведінки в залежності від складності зображення. Для більшої дисперсії шуму (рис. 2.3 б) третя область зміщена до області більш великих значень.

Виходячи з залежностей, можна зробити попередній висновок, що величина метрики, отриманої для стисненого зображення, може приблизно бути обчислена наступним чином:

$$PSNR(Q) \approx 63 - Q, \text{ dB.} \quad (2.1)$$

Якщо провести аналогічний аналіз з використанням інших метрик якості таких як PSNR-HVS-M та MS-SSIM, можна отримати наступні залежності, представлені на рис 2.4.

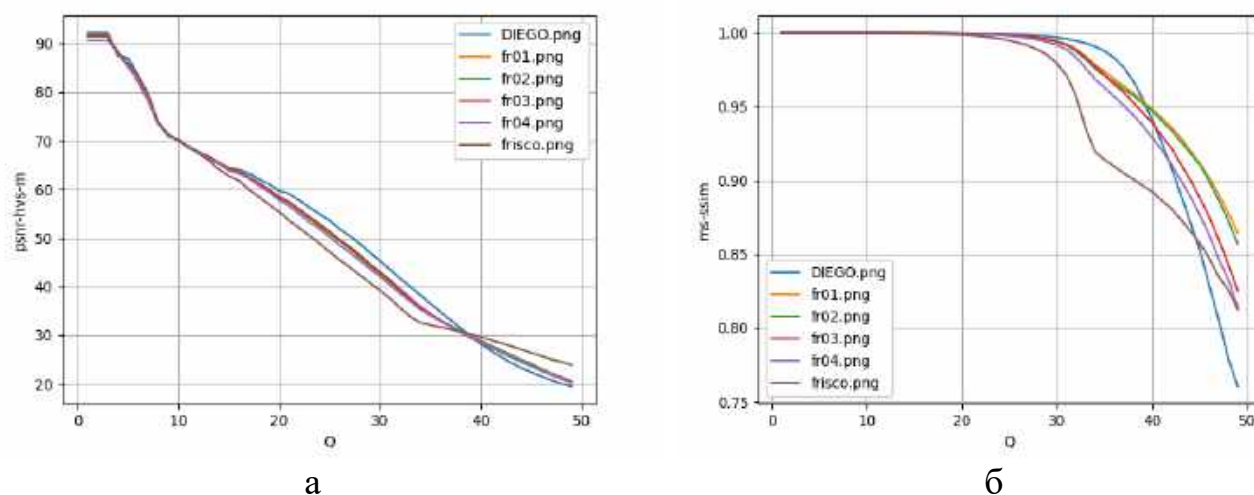


Рис 2.4 – Залежності аналізованих метрик PSNR-HVS-M^{cn} (а) та MS-SSIM^{cn} (б) від Q для дисперсії шуму, що дорівнює 64.

Залежності, отримані для PSNR-HVS-M^{cn} (рис. 2.4 а), в цілому схожі з залежностями, отриманими для PSNR^{cn} за винятком деяких відмінностей. На ділянці $Q > 13$ починається невелика розбіжність значень метрики, пов'язана зі складністю зображень, проте подальша поведінка залишається аналогічною PSNR^{cn} метриці.

Для MS-SSIM^{cn} (рис. 2.4 б) метрики залежності тримаються разом для всіх тестових зображень при $Q < 25$. Після цього починається спад у якості зображення, однак, для Q приблизно більшого, ніж 34, починаються відмінності в поведінці метрики знову-таки в залежності від складності зображення.

Спільний аналіз всіх чотирьох залежностей показує, що при $Q < 29$ спотворення, що вносяться, є непомітними. Це означає, що можна стиснути без візуальних втрат, а потім, після передачі (зберігання) і декомпресії зображення, виконати фільтрацію шуму [110].

Варто зазначити, що висновки, зроблені вище для залежностей PSNR^{cn} (Q) для зашумлених зображень, практично справедливі і для випадку стиснення безшумних

(істинних) зображень. Графіки наведено на рис. 2.5, з них видно, що апроксимація (1) справедлива для п'яти з шести тестових зображень, які розглядаються (винятком є зображення Фріско) при Q від 7 до $Q \approx 35$.

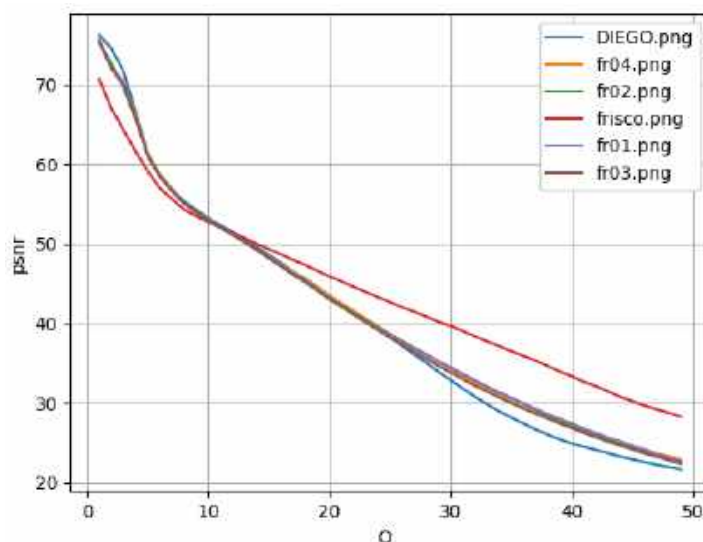


Рис 2.5 – Залежність PSNR(Q) для зображень без шумів

В цілому для напівтонових зображень залежності виявилися досить закономірними, а саме зі збільшенням параметра контролю стиснення якість зображення починає зменшуватися, проте самі залежності можуть дещо відрізнятися одна від одної в залежності від використаного зображення. Далі варто розглянути залежності, які можуть бути отримані лише в симуляції, оскільки зображення, які будуть стискатися здебільшого вражені шумом, як результат, безшумне (істинне) зображення відсутнє. Дані експерименти дозволять пролити світло на питання поведінки спотворень, отриманих під час стиснення з втратами.

2.1.3 Існування робочої точки для кодера BPG

Теоретично, тобто в симуляціях, коли є зображення без шуму, після чого додається до нього AWGN і застосовується стиск із втратами, можна обчислити метрики між стисненим зображенням і відповідним зображенням без шуму, іншими словами, обчислити згадану раніше $\text{Метр}^{\text{ct}}(Q)$. Приклади цих залежностей раніше

вже були частково показані в першому розділі для дисперсії шуму, що дорівнює 100. Тому варто розглянути випадки для інших значень дисперсії. Залежності для шести зображень дистанційного зондування, уражених шумом із дисперсією 64, продемонстровано на рисунку 2.6.

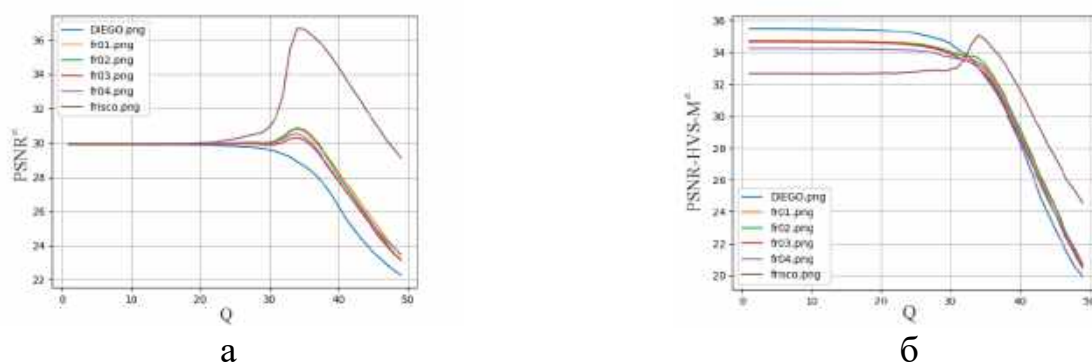


Рис 2.6 – Залежності PSNR^{ct} (а) та $\text{PSNR-HVS-M}^{\text{ct}}$ (б) від Q для кодера BPG для шести тестових зображень, спотворених AWGN з дисперсією, що дорівнює 64

Як бачимо з залежностей для PSNR^{ct} (рис. 2.6, а) оптимальна робоча точка спостерігається на п'яти з шести тестових зображень. Для найпростішого тестового зображення структури Frisco OPT є більш «очевидною», ніж для інших зображень. Для найбільш складного зображення Diego OPT відсутня і залежність для цього зображення залишається монотонною. Крім того, якщо OPT присутня на зображенні, то значення Q , для якого вона спостерігається, є практично однаковими для всіх зображень.

Аналіз залежностей модифікації метрики PSNR, а саме PSNR-HVS-M (рис. 2.6, б), показує, що OPT можна спостерігати і для метрик візуальної якості, хоча це відбувається і рідше. Максимум OPT не такий "очевидний", як для PSNR. Варто зазначити, що для деяких зображень все ж існує певний локальний максимум у тій самій точці, що й OPT. Також потрібно звернути увагу на той факт, що для цих значень дисперсій (які менші, ніж були продемонстровані у прикладі в першому розділі) характерним є те, що оптимальна робоча точка або локальні мінімуми для розглянутих зображень є менш очевидними, ніж для більших значень дисперсії шуму. Ця особливість стає ще більш помітною, якщо поглянути на отримані залежності для

ще більших значень дисперсії шуму, що дорівнює 196, для PSNR-HVS-M^{ct} і MS-SSIM^{ct}, наведені на рис. 2.7.

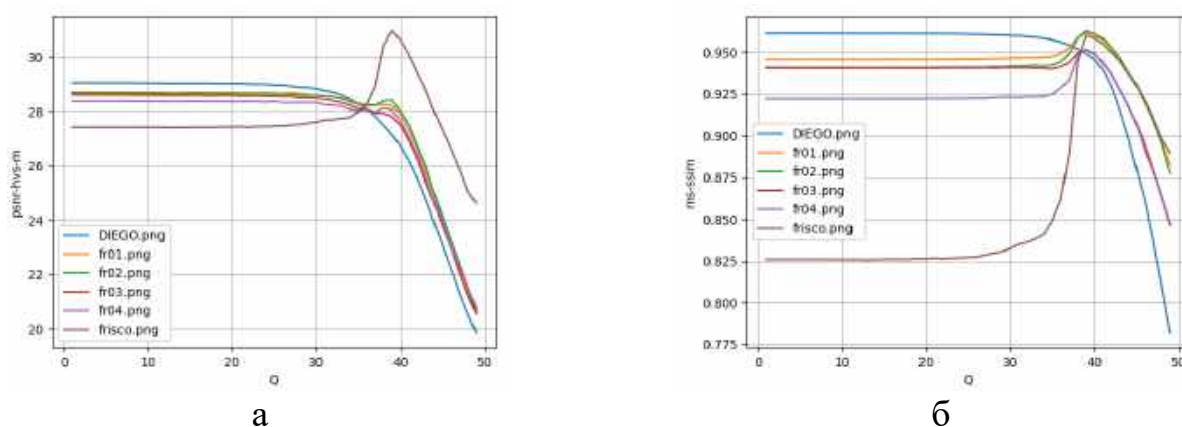


Рис 2.7 – Залежності розглянутих метрик (розрахованих між істинним і стисненим зображеннями) від Q : SNR-HVS-M^{ct} (Q) (а) та MS-SSIM^{ct} (Q) (б) за дисперсії шуму, що дорівнює 196

Виходячи з отриманих залежностей (рис 2.6, 2.7), можна зробити висновок, що при збільшенні Q можливі три варіанти: 1) якість зображення починає покращуватися до досягнення ОПТ (пов'язаного з максимумом залежності), після чого якість починає швидко знижуватися; 2) якість продовжує залишатися майже такою самою і можливі навіть локальні максимуми, а потім настає швидке зниження; 3) якість монотонно знижується.

Перша ситуація має місце для метрики MS-SSIM^{ct} та PSNR^{ct} для всіх п'яти зображень, крім тестового зображення diego, а також для метрики PSNR-HVS-M^{ct} для тестового зображення frisco. Друга ситуація спостерігається з метрикою PSNR-HVS-M^{ct} для деяких тестових зображень середньої складності. Третя ситуація має місце для зображення diego складної структури за всіма трьома метриками візуальної якості.

Очевидно, що для першої ситуації можна рекомендувати стиск в ОПТ, у другій доцільно стиск із втратами в околиці локальних максимумів, а в третій ситуації бажано проводити стиснення в точці перед початком різкого спаду візуальної якості зображення.

Щоб упевнитися в достовірності отриманих результатів, а саме існування робочої точки для VPG-кодера, був узятий інший набір зображень і виконана аналогічна процедура отримання залежностей (до оригінальних зображень додається шум, після чого стискають і порівнюють стиснуті зображення з оригінальними, тобто не стиснутими зображеннями). Результати представлені на рис 2.8 і 2.9. Залежності PSNR-HVS- M^{ct} від Q , побудовані для 5 нових тестових зображень для значень дисперсії шуму, рівної 64 (а), 100 (б) і 196 (в) представлені на рис. 2.8.

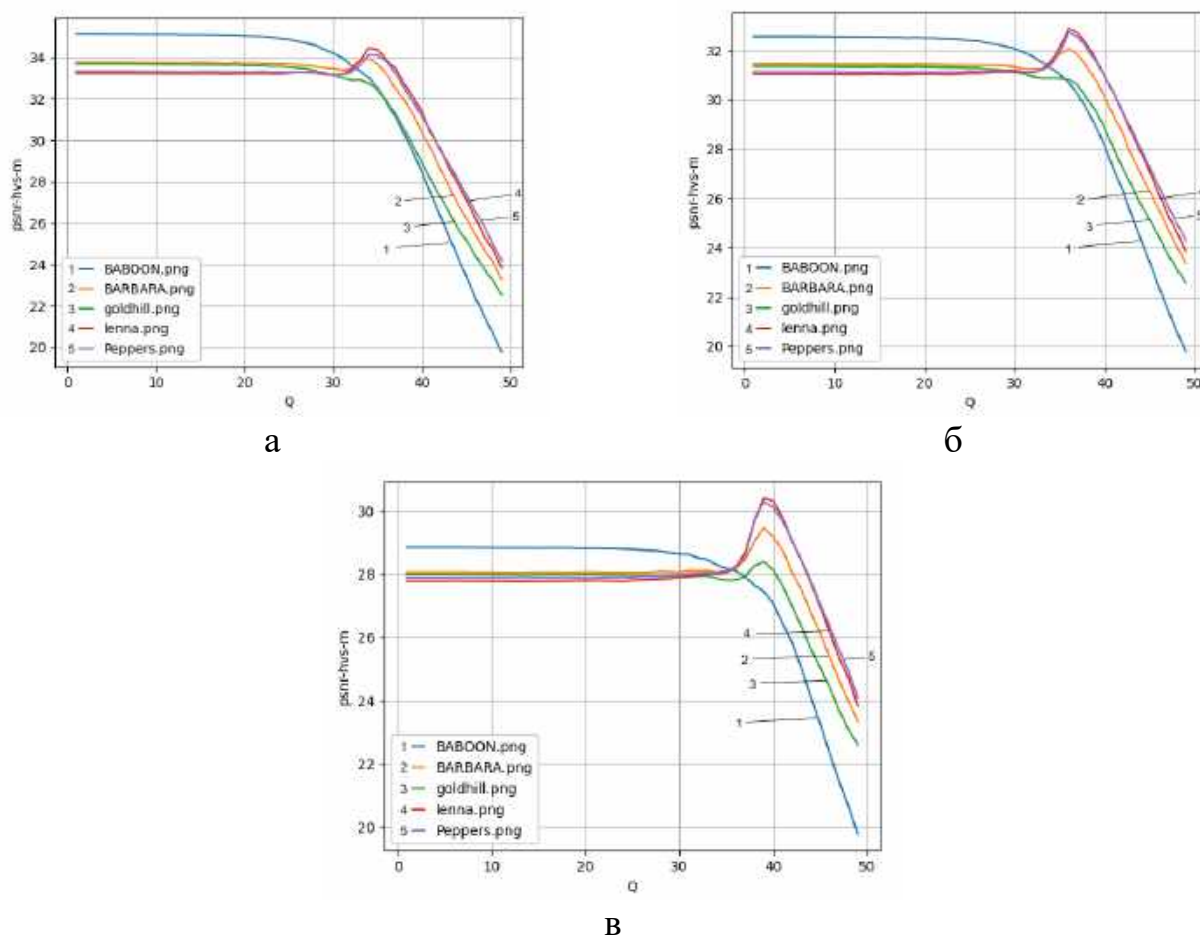


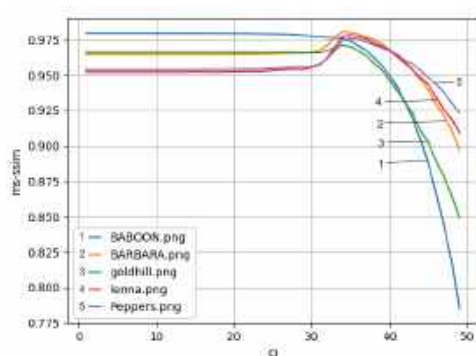
Рис 2.8 – Залежності PSNR-HVS- M^{ct} від Q для стисненого та істинного зображень для значень дисперсії шуму, рівною 64 (а), 100 (б) і 196 (в)

Для першого випадку (рис. 2.8, а) справедливим буде твердження, що робоча точка спостерігається для зображень Peppers, Lena, Barbara при значенні $Q = 34$. Для складного зображення Baboon залежність монотонно зменшується і не має OPT. Для тестового зображення Goldhill залежність також монотонно спадає.

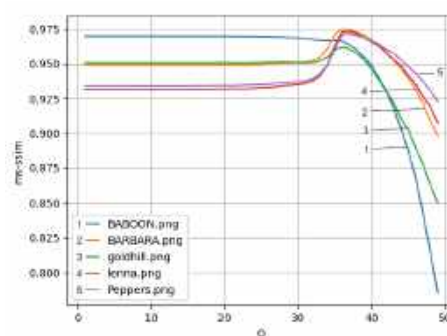
Наступний випадок, який буде проаналізовано - дисперсія шуму дорівнює 100 (рис. 2.8, б) має схожі з попереднім тенденції. Для зображень Peppers, Lena, Barbara так само спостерігається оптимальна робоча точка, проте при трохи більших значеннях ($Q = 36$), для зображення Baboon залежність монотонно зменшується, як і для дисперсії, рівної 64. Для зображення Goldhill є ділянка значень Q від 32 до 36, на якому значення PSNR-HVS- M^{ct} практично не змінюються, проте вони все ж таки нижче, ніж на початковій ділянці, тобто формально OPT не спостерігається.

Для останнього випадку (рис. 2.8, в) при дисперсії шуму, що дорівнює 196, OPT спостерігається для всіх зображень, крім зображення Baboon. Однак значення Q стали більше; для зображень, у яких можна спостерігати OPT, воно становить 38. Також варто окремо відзначити, що при такій дисперсії шуму OPT почала спостерігатися і для зображення Goldhill, яке не мало її при менших значеннях дисперсії шуму.

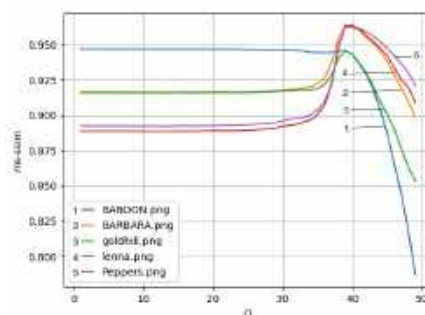
Залежності, що отримані для MS-SSIM^{ct} (між стисненим і істинним зображенням), представлені на рис. 2.9. Як і раніше, залежності були побудовані для значень дисперсії шуму, що дорівнюють 64 (а), 100 (б) і 196 (в).



а



б



в

Рис 2.9 – Залежності MS-SSIM^{ct} від Q для стисненого та істинного зображень для значень дисперсії шуму, рівної 64 (а), 100 (б) і 196 (в)

У першому випадку (рис. 2.9, а) ОРТ спостерігається для всіх зображень, крім зображення Baboon. Для нього, як і для PSNR-HVS-M^{cl}, після $Q = 28$ значення метрики починають плавно зменшуватися. Для решти зображень робоча точка має місце при значенні $Q = 34$.

Для значення дисперсії шуму, що дорівнює 100 (рис. 2.9, б), ОРТ так само спостерігається для чотирьох зображень, але при значенні $Q = 36$.

З отриманих залежностей (рис. 2.8, 2.9) видно, що загалом для нового набору зображень проявляються такі ж самі закономірності як і для першого набору. Йдеться про різні моделі поведінки зображень, пов'язаних з їх складністю: прості зображення мають оптимальну робочу точку в районі $Q = 34$, частина зображень мають локальний максимум у тому ж значенні параметра контролю стиснення, а зображення складної структури не мають ОРТ і після досягнення певного значення починають монотонно втрачати у візуальній якості зображення.

Вище було наведено багато графіків, які свідчать про те, що зображення (прості зображення), стиснуті в оптимальній робочій точці, отримали більшу оцінку метрик оцінювання якості або, інакше кажучи, стиснуті зображення, для яких спостерігається ОРТ, виявляються ближчими до істинного, що не містить шум варіанта, у зв'язку з чим логічним є продемонструвати, який візуальний вигляд має цей виграв у якості зображення. У зв'язку з чим зображення, для яких спостерігалася ОРТ, були стиснуті та порівняні з нестиснутим зашумленим варіантом. Приклади (порівняння) зашумлених зображень і стиснутих в ОРТ представлено на рис. 2.10 і 2.11.

На рис. 2.10 наведено приклад зашумленого стандартного тестового зображення (рис. 2.10, а), стисненого в околі ОРТ. Як видно з рисунка, шум добре помітний, особливо на однорідних ділянках. Аналіз стисненого зображення (рис. 2.10, б) показує, що шум помітно зменшується, проте контури і деякі дрібні деталі трохи розмиваються. Загалом, стиснене зображення виглядає краще, ніж з шумом.

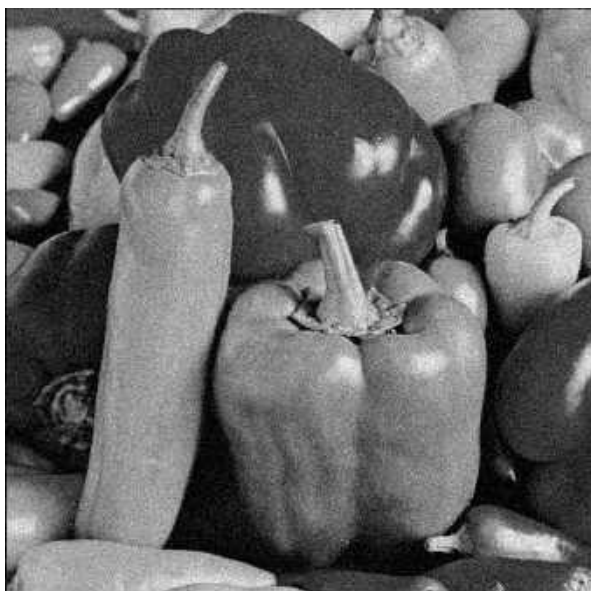


Рис 2.10 – Зображення з шумом Peppers (a) і те саме зображення, стиснене в ОПТ (б), дисперсія шуму 196

Аналогічна ситуація для зображень з області дистанційного зондування, наведеного на рисунку 2.11.



Рис 2.11 – Зашумлене зображення frisco (a) і те саме зображення, стиснене в ОПТ (б), дисперсія шуму дорівнює 196

Основна мета на цьому етапі – зрозуміти, що таке Q_{OPT} і як його можна визначити для даного зображення. Попередні спостереження, які впливають із графіків на рисунках 2.6-2.9, є наступними. Для всіх зображень, для яких спостерігаються OPT, вони мають місце для досить вузького інтервалу значень Q . Це хороша властивість, що дозволяє припустити, що OPT не залежить від наявного зображення і метрики, але, ймовірно, залежить від дисперсії шуму.

Варто також відзначити одну гіпотезу, засновану на аналізі результатів [111], [112]. В [112] виявлено, що для OPT (за PSNR) виконується умова:

$$\text{PSNR}^{\text{cn}}(Q) \approx \text{PSNR}^{\text{n}} \quad (2.2)$$

де PSNR^{n} – це метрика, отримана для зашумленого зображення.

PSNR^{cn} – це метрика, отримана для зашумленого стисненого зображення.

Це означає, що для $\sigma^2 = 64$ $\text{PSNR}^{\text{cn}}(Q) \approx 30$ дБ, для $\sigma^2 = 100$ $\text{PSNR}^{\text{cn}}(Q) \approx 28$ дБ и для $\sigma^2 = 196$ $\text{PSNR}^{\text{cn}}(Q) \approx 25$ дБ. З урахуванням виразу (2.1) це має статися за $Q=33, 35$ і 38 відповідно. Графіки представлені на рис. 2.9 для дисперсій шуму, рівних 64 і 100, показують, що це справді так. Додаткові дослідження, проведені для інших дисперсій та інших тестових зображень, показали, що вираз (2.2) дійсно справедливий для $Q=Q_{\text{OPT}}$. Підставивши (2.1) у (2.2) і провівши нескладні перетворення, також можна отримати (для 8-бітних зображень) наступну закономірність:

$$Q_{\text{OPT}} \approx 14.9 + 20 \log_{10}(\sigma). \quad (2.3)$$

Цей вираз може допомогти у виборі потрібної точки для досягнення OPT, а також може послужити опорою для подальшого розроблення методів прогнозування OPT.

Останнім із не розкритих питань, що залишились, це одержуваний коефіцієнт стиснення під час роботи з напівтоновими зображеннями. Залежності, представлені на рис 2.12 показують відношення коефіцієнта стиснення до параметра контролю стиснення VPG кодера.

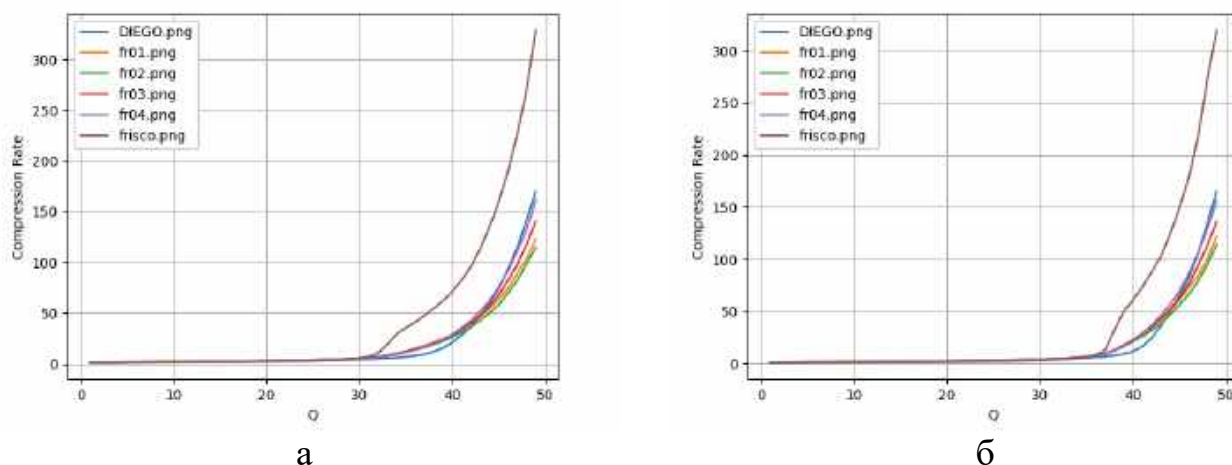


Рис 2.12 – Залежності КС від Q при дисперсії шуму дорівнює 64 (а) та 196 (б).

Висновки такі. По-перше, КС (при однакових значеннях Q) залежить від складності зображення. Наприклад, при $Q=28$ і дисперсії шуму, що дорівнює 64, досягнуті КС для найпростішого і найбільш складного тестових зображень frisco і diego неістотно різняться (вони становлять близько 4,3 і 3,7 відповідно, тобто йдеться про стиснення практично без втрат). Для $Q=35$ ситуація, наступна: КС для frisco становить близько 35,3 тоді як КС для diego становить близько 6,8, тобто, КС значно більше для зображення простої структури та значно відрізняється від попереднього випадку. По-друге, порівняння даних на рисунках 2.12 (а) і 2.12 (б) можна стверджувати, що КС для більш зашумлених зображень менше. Наприклад, для $\sigma^2=196$ і $Q=35$ значення КС для зображень frisco та diego дорівнюють 5,5 і 6,4 відповідно. Різне збільшення КС починається, коли Q виявляється більше $Q_{\text{орт}}$, особливо для зображень простих структур. По-третє, немає істотної різниці у загальних тенденціях до зображень різного походження. Результати для високотекстурних зображень diego і baboon дуже близькі, так само як і результати для зображень простої структури frisco і lena які теж схожі між собою.

2.2 Стиснення кольорових зображень кодером BPG

Вище були продемонстровані залежності, отримані для напівтонових стиснених зображень, які як містять, так і не містять шум, для різних метрик. Були продемонстровані випадки, коли оптимальна робоча точка присутня на даних залежностях, проте залишається відкритим питання, чи справедливі ці виводи і для стиснення кольорового зображення. У першому розділі згадувалося, що крім можливості стиснення напівтонових зображень, BPG також може працювати з кольоровими зображеннями, також було згадано про можливість вибору різних форматів кольоровості при роботі з кольоровими зображеннями. Робота даних форматів буде проаналізована за аналогією з варіантом для напівтонових зображень для того, щоб встановити, чи існує оптимальна робоча точка для кольорових зображень, і, в залежності від відповіді, дати рекомендації, як варто діяти (яку величину параметра контролю стиснення необхідно вибрати для досягнення кращого співвідношення коефіцієнта стиснення до якості стисненого зображення). Для аналізу роботи буде використовуватися два набори зображень: перший - це кольорові варіанти з наборів для напівтонових зображень, і другий містить зображення дистанційного зондування. Як і в минулий раз, зображення мають розмір 512×512 пікселів. Мініатюри зображень наведені на рисунках 2.13 та 2.14.

Набір кольорових версій зображень - з різних областей (рис. 2.13).



Рис 2.13 – Мініатюри використаних зображень у другому наборі: а) baboon б) diego в) frisco г) lenna д) peppers

Набір зображень, що містять виключно зображення дистанційного зондування (рис. 2.16).

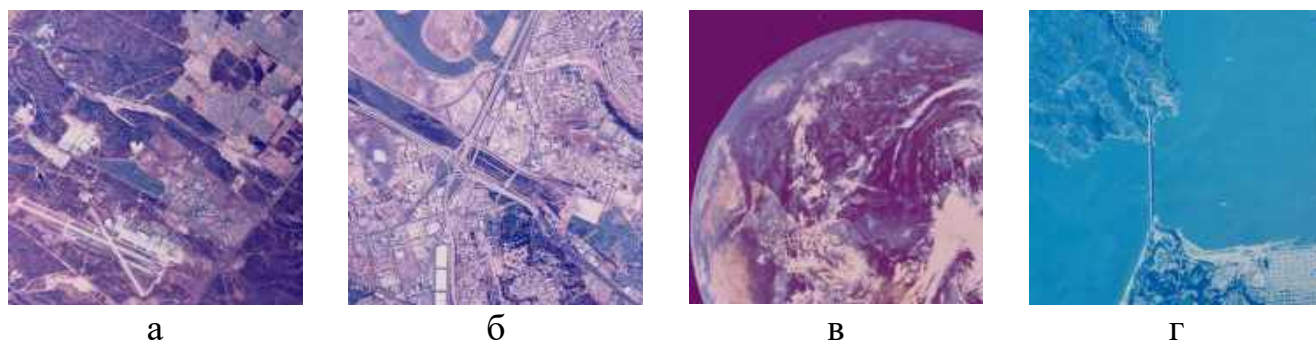


Рис 2.14 – Мініатюри використаних зображень у другому наборі: а) diego б) diego2 в) earth г) frisco

2.2.1 Субдискретизація кольоровості

Як уже згадувалося раніше, кодер VPG підтримує формати кольоровості YCbCr 4:2:0, 4:2:2, 4:4:4. YCbCr використовується для цифрового кодування колірної інформації, придатної для стиснення і передавання відео та зображень. Перетворення колірного простору YCbCr сильно декорелює колірні канали, тому їх можна кодувати незалежно без втрати ефективності. Перетворення зображень з RGB та інших колірних просторів в YCbCr не становить проблеми.

Оскільки зорова система людини менш чутлива до кольоровості, ніж до яскравості, канали Cb і Cr (що відповідають за колір) можна субдискретизувати в два рази як у горизонтальному, так і у вертикальному напрямках майже без втрати якості зображення, що сприймається. Це знижувальна множина дає вибірку YCbCr 4:2:0, де на кожен набір значень Cb і Cr припадає чотири вибірки Y. Схема підвиборки 4:2:2 означає, що кожен компонент Cb і Cr відбирається у підвибірку з коефіцієнтом 2 по горизонталі. При цьому враховуємо той факт, що кольорові та триканальні зображення ДЗЗ часто підлягають візуалізації та аналізу фахівцями.

Щоб оцінити та порівняти різні режими субдискретизації кольоровості при стисканні із втратами на основі VPG, необхідно використати декілька показників. Насамперед його коефіцієнт стиснення (КС), з іншого боку, необхідно оцінити візуальну якість зображення. Для оцінки візуальної якості стисненого зображення використовуватимуться метрики PSNR, PSNR-NA і MDSI. Незважаючи на те, що цього разу будуть використовуватися кольорові зображення, позначення метрик

будуть такими ж, як і для випадку з напівтоновими зображеннями. ($\text{Метр}^{\text{cn}}(Q)$ – метрика, розрахована між зашумленим зображенням та стиснутим, та $\text{Метр}^{\text{ct}}(Q)$ – метрика, розрахована між істинним зображенням та стиснутим).

2.2.2 Результат використання різних режимів субдискретизації кольоровості

Спочатку варто відзначити той факт, що варіанти представлення кольоровості 4:2:0 і 4:2:2 відносяться до субдискретизації колірних компонентів. З одного боку, їх використання дає змогу частково розв'язати завдання стиснення практично без втрат. З іншого боку, субдискретизація зображень колірних компонентів призводить до внесення додаткових і специфічних спотворень, які необхідно враховувати при застосуванні до стиснення з втратами зашумлених трьохканальних зображень.

Для початку розглянемо залежності метрики PSNR^{cn} , отриманої між стисненим зображенням і зашумленим зображенням, для всіх значень параметра контролю стиснення. Результати представлені на рисунку 2.15.

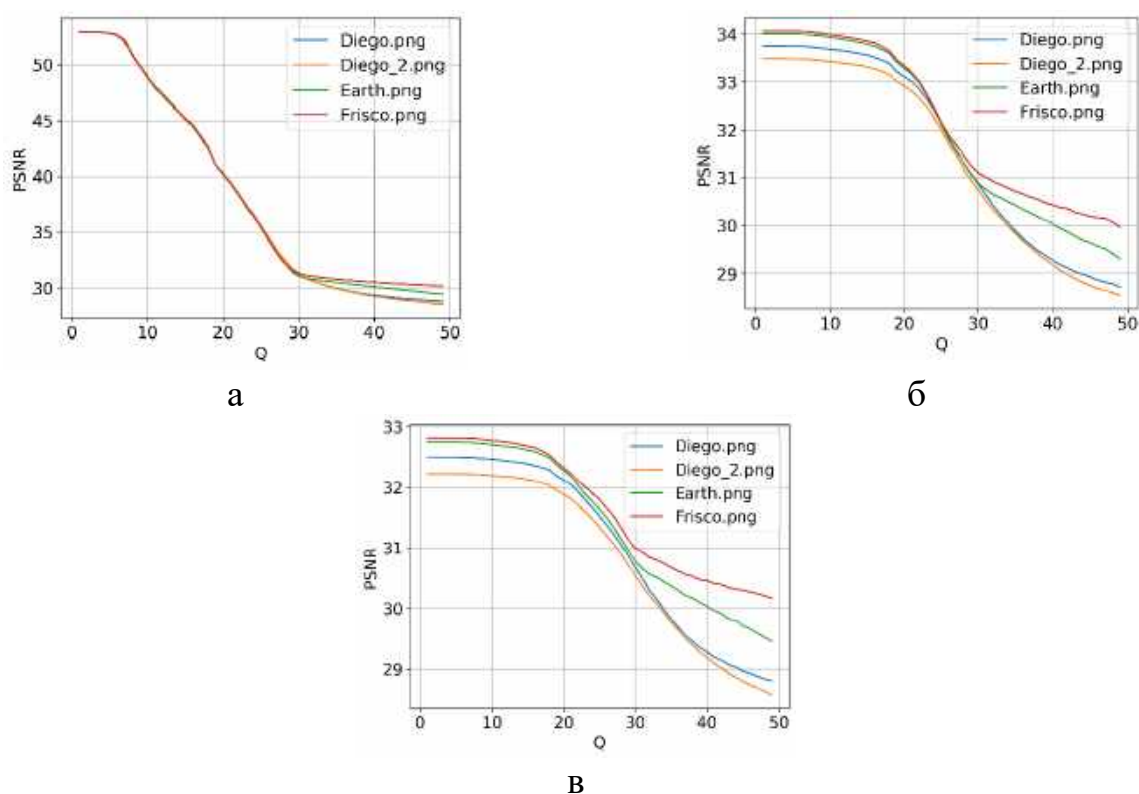


Рис 2.15 – Залежності $\text{PSNR}^{\text{cn}}(Q)$ для чотирьох розглянутих триканальних зображень у форматах 4:4:4 (а), 4:2:2 (б) і 4:2:0 (в)

Оскільки це типово для стиснення з втратами, якість стиснутих зображень погіршується при збільшенні КС. Це видно на рисунку 2.15, де подано графіки $PSNR^{cp}(Q)$ для всіх трьох можливих форматів представлення зашумлених зображень ($\sigma^2=49$). Для всіх форматів і для всіх тестових зображень розглянуті криві швидкість-спотворення монотонно зменшуються. Однак існують суттєві відмінності в їх поведінці залежно від формату стисненого зображення. Для формату 4:4:4 (рис. 2.15,а) можна спостерігати три очевидні інтервали різної поведінки. Перший - при $Q < 7$, $PSNR^{cp}(Q) \approx 53$ дБ, тобто втрати, що вносяться стисненням, незначні й неможливо помітити відмінності між вихідним і відповідним стисненим зображенням. Далі для $7 \leq Q \leq Q_{bx}$ практично існує лінійна залежність, яку можна апроксимувати як $PSNR^{cp}(Q) \approx 60 - Q$, де Q_{bx} приблизно відповідає вхідному PSNR, (близько 31,3 дБ для розглянутого випадку $\sigma^2=49$). Для двох вищезазначених інтервалів криві швидкість-спотворення для різних зображень практично збігаються. Для третього інтервалу $Q > Q_{bx}$ криві продовжують зменшуватися, хоча для простіших зображень структур (frisco і earth) зниження відбувається повільніше.

Особливістю залежностей $PSNR^{cp}(Q)$ для форматів 4:2:2 (рис. 2.15, б) і 4:2:0 (рис. 2.15, в) є те, що вони поведуться інакше порівняно з варіантом 4:4:4, але схожі між собою. Обидві залежності починаються ($Q=1$) з $PSNR^{cp}$ близько 34 дБ (рис. 2.15, б) і 32,5 дБ (рис. 2.15, в), це означає, що зображення після субдискретизації вже досить сильно відрізняються від початкових. За $Q < 15$ якість стиснених зображень залишається практично такою самою, як і за $Q=1$. Потім, при збільшенні Q , існує очевидна тенденція до більшої різниці між оригіналом і відповідними стиснутими зображеннями. Порівняння відповідних кривих (на рис. 2.15, б і 2.15, в) показує, що значення $PSNR^{cp}$ для формату 4:2:2 трохи більше, принаймні за малих Q .

Однак залежності $PSNR^{cp}(Q)$ (між зашумленим і стисненим) не показують, що відбувається з шумом, якщо до зображень з шумом застосовується стиснення з втратами. З цієї причини часто розглядають залежності $PSNR^{ct}(Q)$ між стисненим і відповідним істинним (безшумним) зображеннями. Так само, як і для випадку напівтонових зображень, на практиці отримати пару зображень, перше з яких це те

що зовсім не містять шуму, та таке саме, але з шумом неможливо. Однак це можливо під час аналізу змодельованих даних, коли є зображення без шуму, до нього додають шум, а потім застосовують стиснення з кількома значеннями параметра контролю стиснення. Такий аналіз корисний, оскільки дає змогу зрозуміти, як змінюються властивості стиснутих зображень у разі зміни параметра Q (у разі BPG кодера).

Як приклад таких залежностей, на рис 2.16 представлено графіки $PSNR^{ct}(Q)$ для всіх трьох розглянутих форматів.

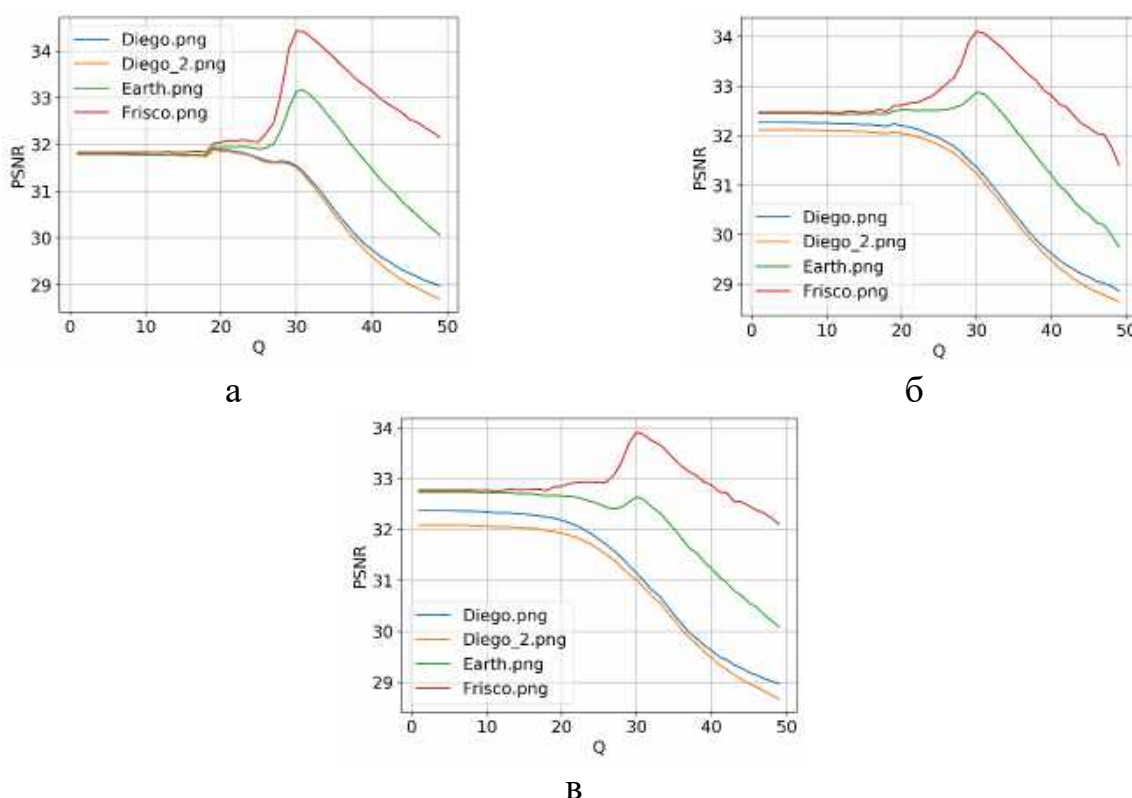


Рис 2.16 – Залежності $PSNR^{ct}(Q)$ для чотирьох розглянутих триканальних зображень у форматах 4:4:4 (а), 4:2:2 (б) і 4:2:0 (в)

Як видно з рисунка 2.16, залежності для зображень простої та середньої складності frisco і earth мають очевидні максимуми (оптимальні робочі точки), які можна побачити для $Q_{OPT}=30$. У свою чергу, залежності для зображень diego і diego2, які є представниками складної структури, або не мають явних максимумів, або монотонно спадають. Цей приклад показує, що в OPT варто стискати прості зображення, а ось для складних зображень варто застосувати інший підхід. Наразі

однією з рекомендацій можна назвати таку: варто використовувати рекомендований Q менший, ніж Q_{OPT} , щоб уникнути значних спотворень. Наведені приклади також підводять до думки, що розумно виробити підхід, здатний передбачити, чи існує OPT для даного зображення і дисперсії шуму.

Хоча залежності для одного і того ж зображення, здаються схожими, для різних форматів кольоровості, проте існують деякі відмінності. Головна з них полягає в тому, що значення $PSNR^{ct}$ в OPT для зображень простої структури серед усіх форматів є найбільшим для формату 4:4:4 (рис. 2.16, а). Крім того, варто зазначити, що в разі використання формату 4:4:4 для малих значень Q величина $PSNR$ є однаковою для всіх тестових зображень, то для форматів 4:2:2 і 4:2:0 для різних зображень значення $PSNR$ перестають бути однаковими (рис. 2.16, б і рис. 2.16, в).

Використовуючи ті самі зображення, було побудовано аналогічні залежності для інших значень дисперсій шуму. На рис. 2.17 показано графіки метрики MDSI для двох інших значень дисперсії шуму - $\sigma^2=64$ і $\sigma^2=196$ (тут варто нагадати, що для метрики MDSI, на відміну від інших, менші значення відповідають кращій візуальній якості).

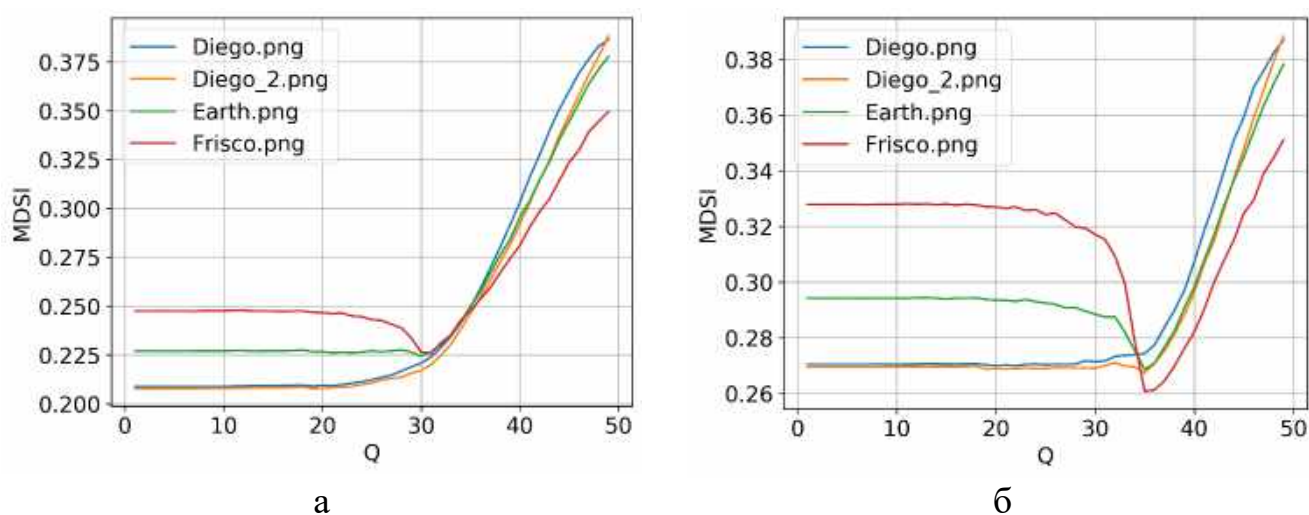


Рис 2.17 – Залежності $MDSI^{ct}$ від Q за дисперсії шуму, що дорівнює 64 (а) і 196 (б) для формату 4:4:4

Аналіз показує, що для зображень простих структур спостерігаються очевидні локальні мінімуми. За $\sigma^2=196$ "неочевидний" мінімум існує навіть для зображення

складної структури diego 2. За $\sigma^2=64$ значення Q_{OPT} дорівнюють 31 (рис. 2.17, а), а за $\sigma^2=196$, OPT спостерігається за $Q_{OPT}=36$.

Додатковий аналіз було проведено для іншого набору зображень, включно з diego і frisco (що були в початковому наборі зображень) для метрик MDSI, а також іншої модифікації PSNR, а саме PSNR-HA. Отримані залежності представлено на рис 2.18 і 2.19.

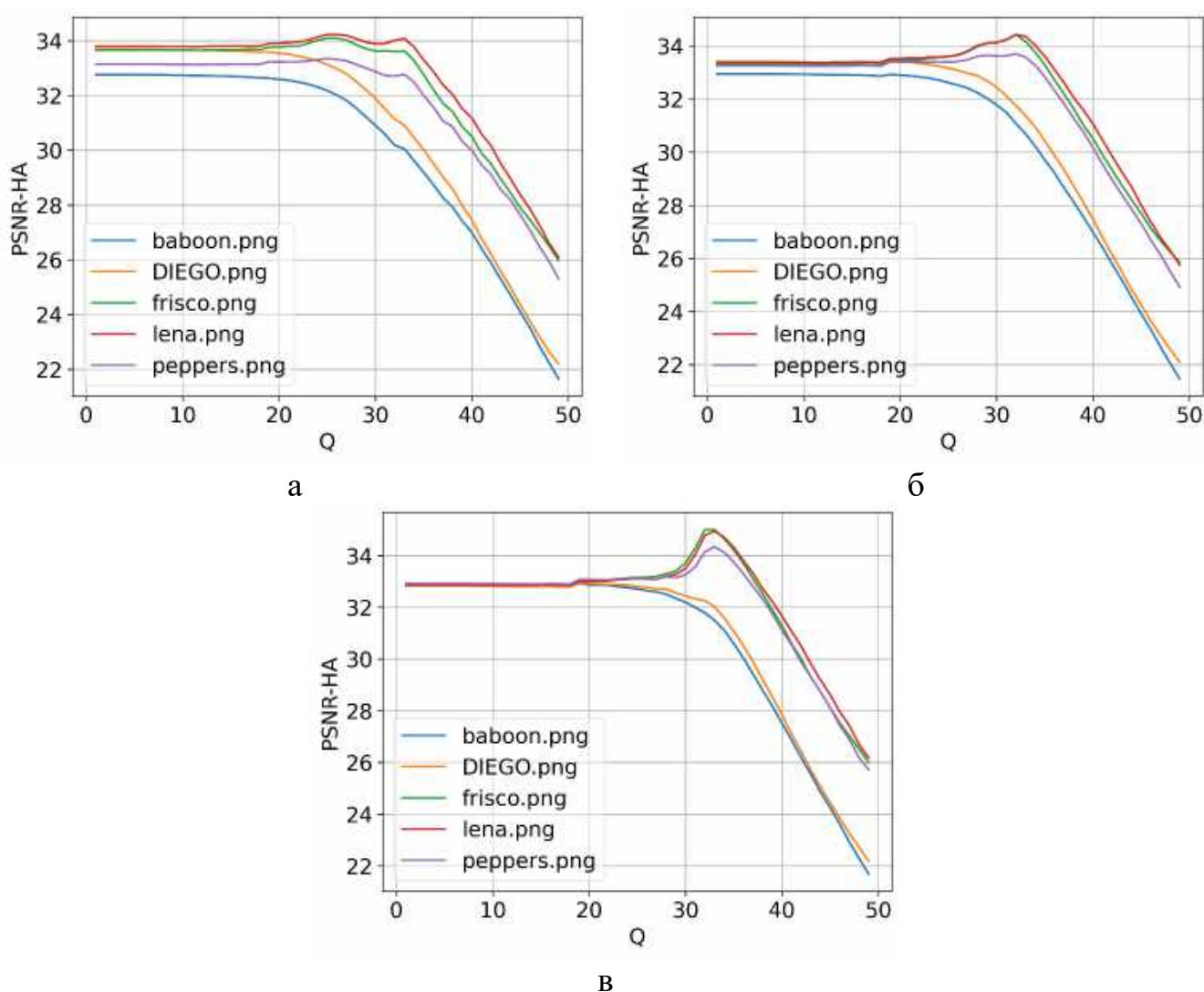


Рис 2.18 – Залежності $PSNR-HA^{ct}$, розраховані між стиснутими зашумленими зображеннями з дисперсією шуму, що дорівнює 100, і зображеннями без шумів для режимів субдискретизації кольоровості: 4:2:0 (а), 4:2:2 (б), 4:4:4 (в)

Для режиму 4:2:0 (рис. 2.18, а) для зображень складної структури збільшення параметра Q призводить до появи плоскої області з подальшим монотонним

зменшенням $PSNR-HA^{ct}$ за Q , що дорівнює 22 і більше. Для зображень середньої та простої структури ситуація інша: точка, за якої спостерігається зниження якості зображення, зміщується в бік більших значень Q , також можуть спостерігатися локальні максимуми.

З переходом у режим 4:2:2 (рис. 2.18, б) спостерігаються аналогічні результати, за винятком того, що початкова точка зниження якості зміщується в бік дещо більших значень для зображень складних структур. Для інших зображень згадані раніше локальні максимуми стають більш вираженими.

Останній можливий режим - 4:4:4 (2.18, в). Цього разу ОПТ для зображень простої та середньої структури стає найочевиднішою, відправна точка зниження якості для зображень складної структури перебуває практично в тому самому місці (за того самого Q), що й ОПТ для зображень простої структури ($Q=32$).

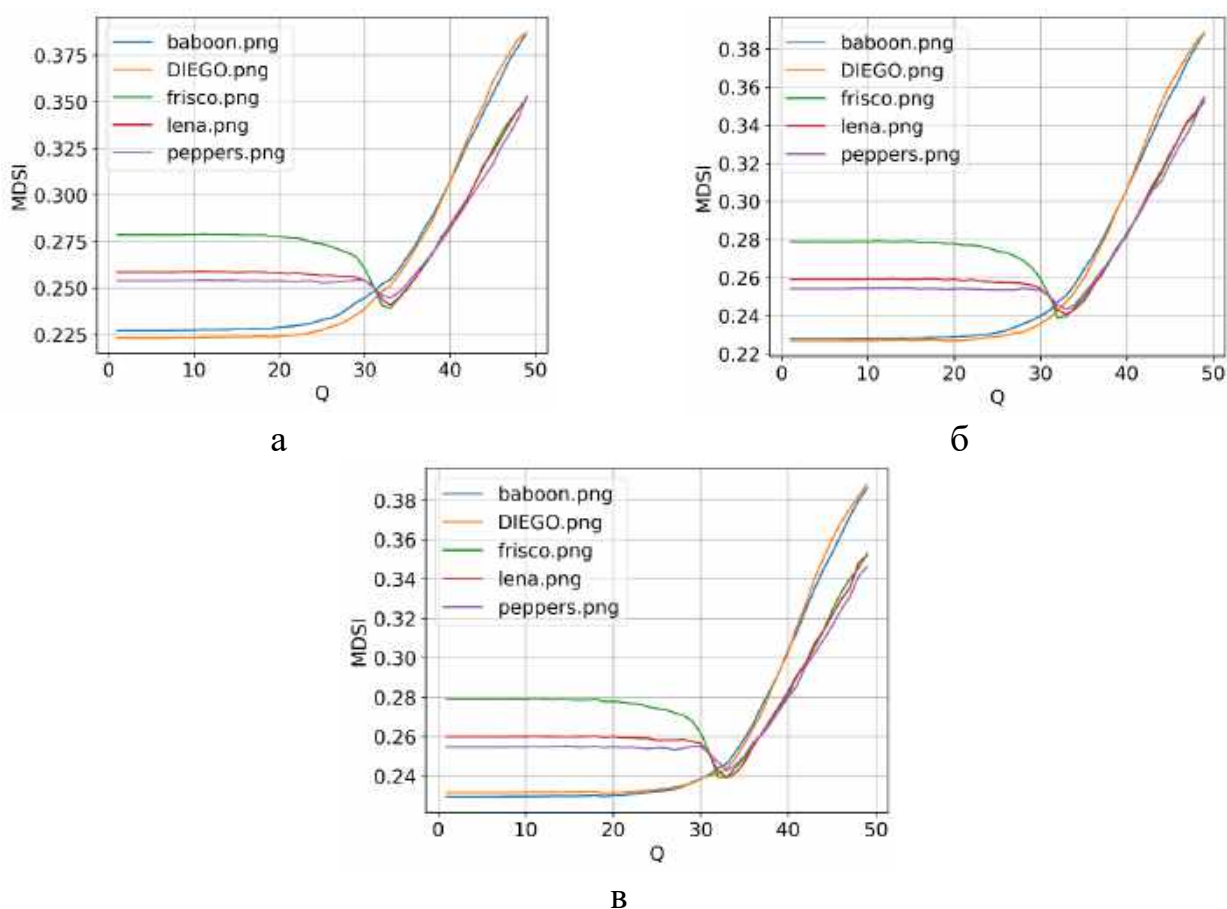


Рис 2.19 – Залежності $MDSI^{ct}$, розраховані між стиснутими зашумленими зображеннями з дисперсією, що дорівнює 100, і зображеннями без шумів для режимів кольорової субдискретизації: 4:2:0 (а), 4:2:2 (б), 4:4:4 (в)

Висновки, зроблені вище на основі аналізу залежностей метрики PSNR-НA^{ct}, підтверджуються результатами аналізу метрики MDSI^{ct} (рис. 2.19). Знову ж таки, не існує OPT для зображень зі складною структурою, проте OPT існують для зображень із простою та середньою структурою, де для всіх цих зображень поліпшення якості у OPT може бути помітним у всіх режимах роботи. Зображення складної структури стискаються з кращою якістю для режиму 4:4:4 за Q трохи меншого, ніж $Q_{OPT}=32$, що характерно для зображень середньої та простої структури.

Представлені результати показують кілька важливих тенденцій. По-перше, для заданої дисперсії шуму Q_{OPT} (якщо OPT існує) практично однаковий для різних метрик і різних зображень. Те саме спостерігалось під час стиснення з втратами зображень у відтінках сірого, спотворених AWGN. Однак для триканальних зображень Q_{OPT} трохи менше, ніж Q_{OPT} для зображень у відтінках сірого (одноканальних) за тієї ж дисперсії AWGN. Виходячи з цього, необхідно змінити (2.3) таким образом:

$$Q_{OPT} \approx 12.9 + 20 \log_{10}(\sigma). \quad (2.4)$$

Причина цього полягає в наступному: у разі триканальних зображень спочатку вони перетворюються в колірний формат YCbCr, а потім покомпонентно застосовується VP8-стиснення (розпакування виконується у зворотному порядку). Таке перетворення змінює дисперсію шуму і зменшує діапазон представлення даних зображення. Зокрема, дисперсія шуму в зображеннях компонентів стає $\sigma_Y^2 \approx 0.45\sigma^2$, $\sigma_{Cb}^2 \approx 0.39\sigma^2$ та $\sigma_{Cr}^2 \approx 0.43\sigma^2$, тобто менше, ніж у компонентах RGB. Тоді оптимальне Q також має бути меншим. По-друге, завдяки введеній вище формулі (2.4) стає можливим визначити Q_{OPT} .

Оскільки існування оптимальної робочої точки було підтверджено для випадку стиснення кольорових зображень, бажано також провести візуальний аналіз результатів стиснення зображень в OPT. Візуальні результати стиснення кольорових зображень аналогічні тим, що були отримані для напівтонових зображень, у тому

плані, що так само видно ефект фільтрації, а шум було придушено на однорідних ділянках для всіх варіантів подання кольоровості. Приклади отриманих зображень, стиснутих в ОРТ, представлені на рис. 2.20.



Рис 2.20 – Початкове (зашумлене) зображення (а) і зображення, стиснуті в ОРТ для форматів 4:4:4 (б), 4:2:2 (в) і 4:2:0 (г)

На рисунку 2.20 (а) видно шум, особливо в однорідних областях зображення. Зображення, стиснуті в ОПТ для вихідних форматів 4:4:4, 4:2:2 і 4:2:0, показані на рисунках 2.20 (б), 2.20 (в) і 2.20 (г) відповідно. Видно, що шум добре придушено, а корисна інформація (краї, текстури, дрібні об'єкти) зберігається добре для всіх трьох форматів. Значну різницю у візуальній якості зображень (на рис. 2.20, б-г) виявити складно, що наводить на думку, що стиснення в ОПТ виглядає розумним рішенням для забезпечення кращої якості стисненого зображення.

2.2.3 Ступінь стиснення

Ще одна не менш важлива мета в дослідження BPG кодера - це надати інформацію про вплив різної субдискретизації кольоровості на коефіцієнт стиснення. П'ять тестових зображень з вихідним розміром файлу 768 КБ були уражені AWGN з дисперсією, що дорівнювала 100, і стиснуті кодером BPG з початковим значенням параметра управління стисненням, що дорівнює 1. Результати представлені в таблиці 2.1.

Таблиця 2.1 – Початкові коефіцієнти стиснення для різних режимів вибірки кольоровості

Назва зображення	Q	Коефіцієнт стиснення		
		4:4:4	4:2:2	4:2:0
baboon	1	1,158966	1,766764	2,321141
DIEGO	1	1,208332	1,834994	2,383039
frisco	1	1,288324	2,013722	2,681773
lena	1	1,296848	2,03216	2,718279
peppers	1	1,279801	2,007273	2,689392

З даних цієї таблиці видно, що режим 4:2:0 перевершує режими 4:4:4 і 4:2:2 у плані забезпечення більших значень КС, режим 4:2:2 посідає проміжне місце між режимами 4:4:4 і 4:2:0. Наступним кроком буде порівняння двох залежностей КС від Q для режимів 4:2:0 (рис. 2.21, а) і 4:4:4 (рис. 2.21. б).

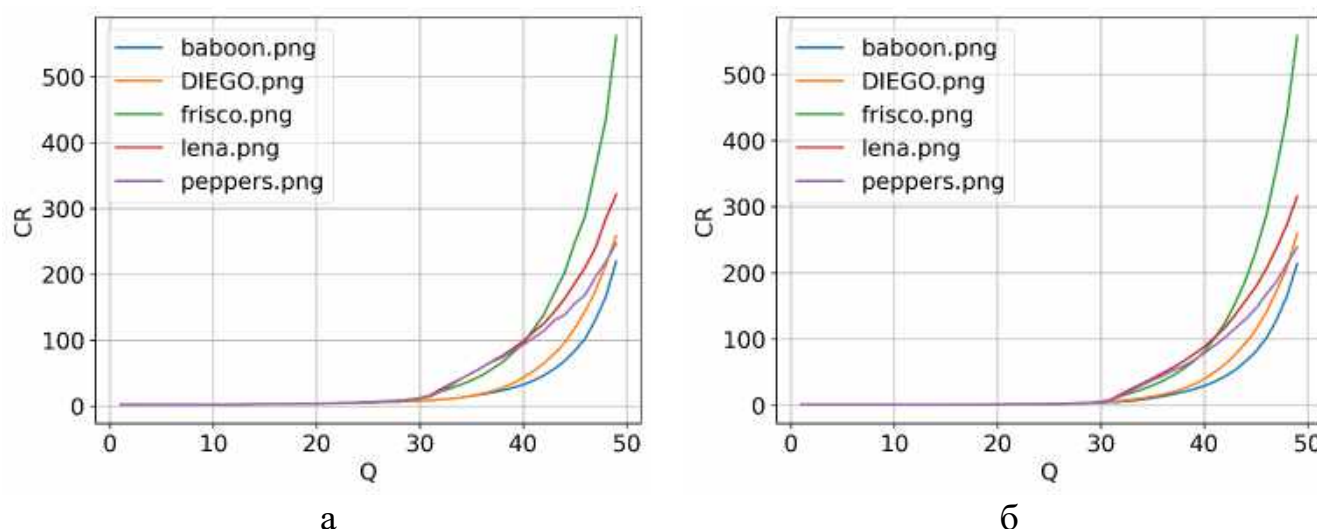


Рис 2.21 – Залежності ступеня стиснення зображень з шумом із дисперсією, що дорівнює 100, для режимів кольорової під-вибірки 4:2:0 (а) і 4:4:4 (б)

З огляду на отримані залежності можна зробити висновок, що криві поведуться аналогічно (порівняно між 4:2:0 і 4:4:4), але криві для режиму 4:2:0 починають "підйом" раніше за умови, що Q дорівнює близько 25.

На особливу увагу заслуговує область оптимальної робочої точки і значення КС у ній. Таблиця 2.2 містить інформацію (значення КС) для всіх трьох режимів ВРГ-кодера.

Таблиця 2.2 – Коефіцієнт стиснення в ОРТ для різних режимів вибірки кольоровості

Назва зображення	Q	Коефіцієнт стиснення		
		4:4:4	4:2:2	4:2:0
baboon	33	8,775278	12,59846	13,8833
DIEGO	33	12,10651	14,08063	14,68142
frisco	33	32,75609	42,07843	46,19015
lena	33	42,25247	51,71074	55,60385
peppers	33	35,29761	45,26773	48,78874

Перше спостереження полягає в тому, що коефіцієнт стиснення для режиму 4:2:2 перевершує режим 4:4:4. Зі свого боку, режим 4:2:0 перевершує режим 4:2:2, але цього разу різниця не така велика. Виходячи з цього спостереження, у сценарії ОРТ-стиснення краще використовувати 4:2:2 з двох причин. По-перше, ОРТ більш

очевидний, ніж для режиму 4:2:0, і майже такий самий, як і в режимі 4:4:4. По-друге, КС практично такий самий, як і в режимі 4:2:0.

На окрему увагу заслуговують залежності показника візуальної якості від ступеня стиснення. Результати порівняння представлено на рисунку 2.22.

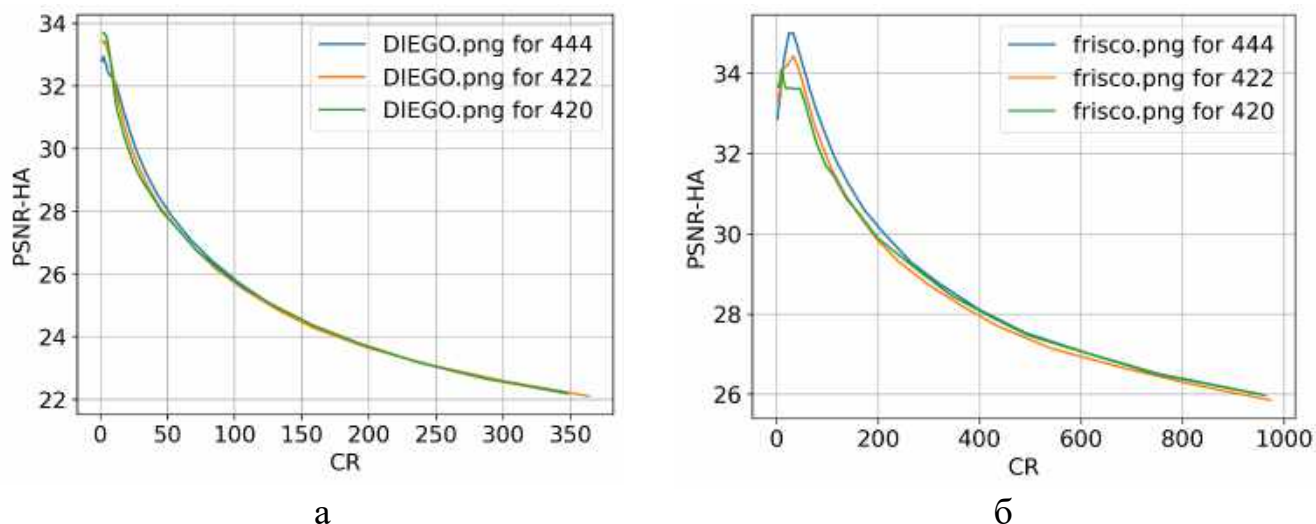


Рис 2.22 – Порівняння залежностей PSNR-НА від CR для тестових зображень: DIEGO (а) і frisco (б), спотворених AWGN з дисперсією, що дорівнює 100

З рисунка 2.24 видно, що однакові коефіцієнти стиснення для різних режимів субдискретизації призводять до різної якості зображення в діапазоні коефіцієнтів стиснення від 1 до 250 (у разі зображень простої структури). Інакше кажучи, при тому ж ступені стиснення режим 4:4:4 може забезпечити кращу якість порівняно з 4:2:2, а також з 4:2:0 у разі зображень із простою структурою. Для складних зображень цей ефект не спостерігається.

2.3 Стиснення зображень із пуасонівським шумом на основі VPG

Вище було показано залежності, отримані для моделі адитивного білого гаусового шуму як для кольорових, так і для напівтонових зображень, однак ця модель шуму є далеко не єдиною, що може бути присутня на зображеннях на практиці (детальніше про це йшлося в першому розділі). У зв'язку з цим виникає питання, як поведе себе VPG кодер під час роботи, наприклад, із сигнало-залежним типом шуму,

і чи покаже він результати, подібні до тих, які було отримано для попередньої моделі шуму, в плані наявності оптимальної робочої точки.

2.3.1 Набір зображень і використані метрики

Для отримання залежностей і подальшого аналізу даних залежностей буде використовуватися той самий набір зображень, що й під час аналізу напівтонових зображень. Крім уже згаданих зображень, введено новий набір зображень, що містять медичні зображення, оскільки ці зображення часто схильні до того, щоб бути ураженими сигнально-залежним типом шуму (детальніше про це йшлося в першому розділі). Мініатюри використаних зображень із нового набору наведено на рис. 2.23.

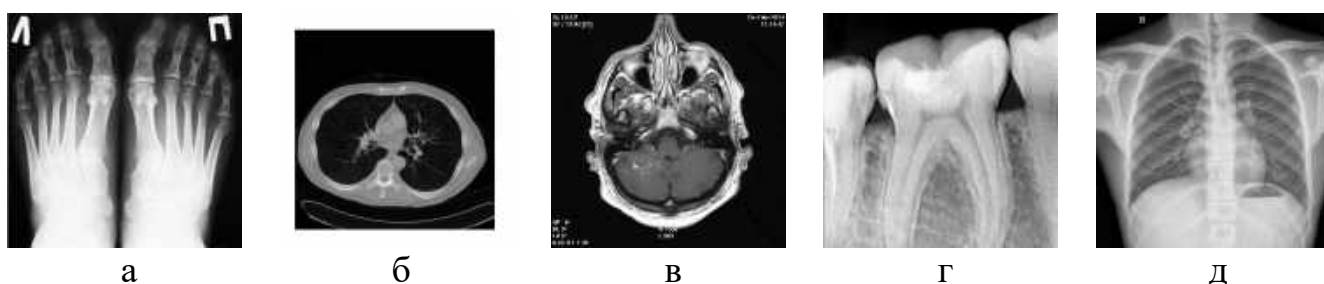


Рис 2.23 – Мініатюри використаних зображень у наборі з галузі медицини: а) foot б) kidney в) mri_head г) tooth д) x-ray

Варто нагадати, що сила шуму Пуассона залежить від інтенсивності пікселів: пікселі з більшою інтенсивністю створюють більшу дисперсію шуму. Оскільки в подальших симуляціях потрібно охопити ширший діапазон змін еквівалентної дисперсії шуму, необхідно було якось розширити наявний набір зображень. Для цієї мети нами були отримані додаткові зображення, в яких кожне значення зображення у вихідному зображенні було розділене на 2 (далі для цього типу зображень буде використовуватися аббревіатура "DIV-2"). Як приклад, якщо середнє значення вихідного зображення дорівнювало 100, середнє значення відповідного нового зображення стало рівним 50. На рисунку 2.24 представлено деякі приклади використаних зображень із нового набору.

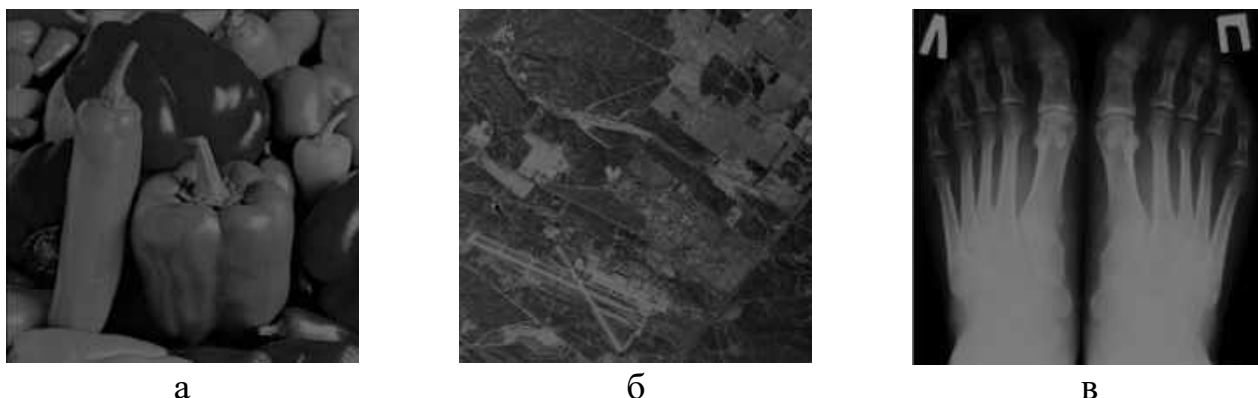


Рис 2.24 – Мініатюри використаних зображень у другому наборі: а) peppers-div2 б) diego-div2 в) foot-div2

Що стосується метрик якості, то буде використано той самий набір метрик, а саме загальноприйнята PSNR і зв'язка метрик, що оцінюють з перспективи людської системи зору PSNR-HVS-M і MS-SSIM. Найменування, що позначають між якими зображеннями було розраховано метрику, залишилось без змін.

2.3.2 Аналіз отриманих результатів

Залежності метрик отримано в той самий спосіб, що й ті, які було отримано для попередніх двох випадків (кольорові та напівтонові зображення, уражені адитивним гаусовим шумом) - до істинних безшумних зображень було додано шум згідно з розглянутою моделлю, після чого відбувається стиснення та порівняння отриманого зображення з істинним для всіх значень параметра контролю стиснення. Отримані залежності для трьох наборів зображень представлено на рис. 2.25.

Згідно з метрикою $PSNR^{ct}$ (рис. 2.25), майже всі зображення в трьох наборах даних мають ОПТ. Це відбувається для всіх зображень простої складності, тоді як зображення середньої складності (fr01, fr02, fr03 і fr04) можуть мати деякі локальні максимуми. Для складного зображення Diego ОПТ відсутня.

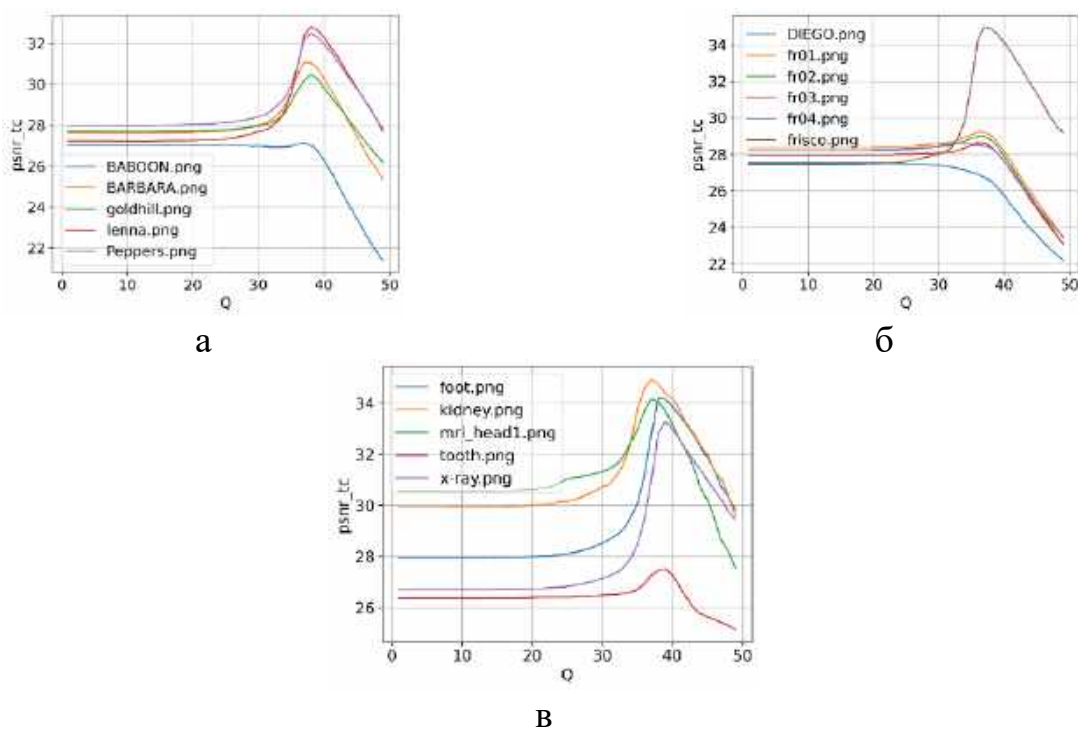


Рис 2.25 – Залежності $PSNR^{ct}(Q)$ для трьох наборів зображень: стандартного (а), дистанційного зондування (б), медичного (в)

Для тієї самої метрики версії зображень "DIV-2" представлені на рис. 2.26 і, на перший погляд, ведуть себе аналогічно своїм первинним версіям.

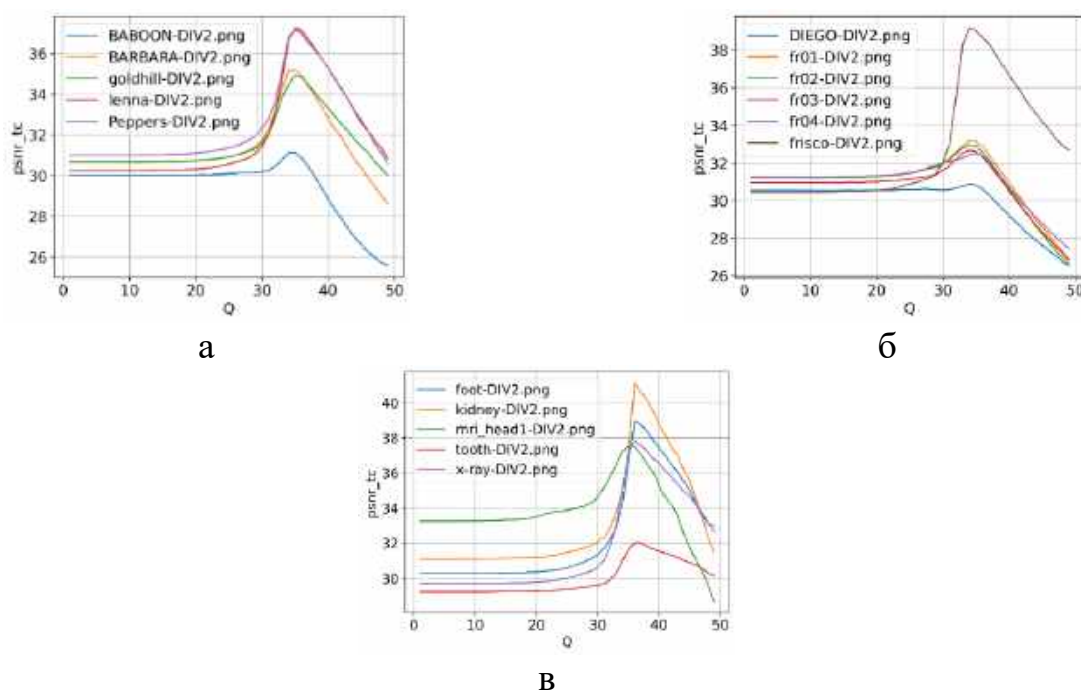


Рис 2.26 – Залежності $PSNR^{ct}(Q)$ для трьох наборів зображень "DIV-2": стандартного (а), дистанційного зондування (б), медичного (в)

Однак під час спільного представлення двох залежностей помітні кілька відмінностей. Порівняння наведено на рисунку 2.27.

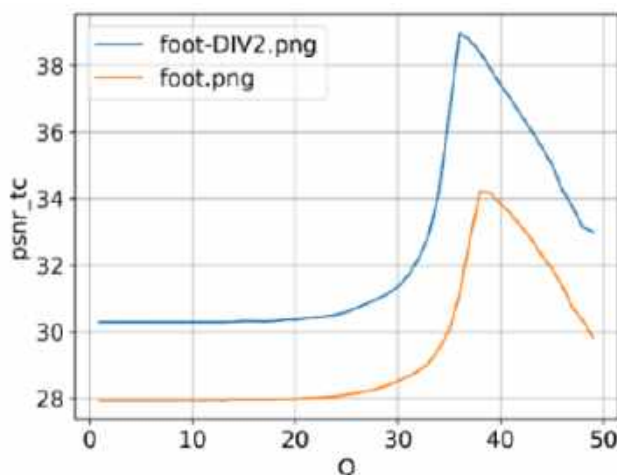


Рис 2.27 – Порівняння залежностей між первинним зображенням і "DIV2"-версією того ж зображення

Слід зазначити, що положення ОРТ змінилося для зображень типу "DIV2" у бік менших значень Q . Це стає зрозумілим з аналізу графіків для тієї самої метрики на рис. 2.27. Для вихідного зображення ОРТ має місце за Q , що дорівнює 38, але в разі "DIV2"-зображення ОРТ спостерігається за $Q = 35$. Таким чином, підтверджується той факт, що позиція ОРТ, у разі стиснення з втратами, яке застосовується до зображень, уражених пуасонівським шумом, залежить від властивостей зображення і використовуваної метрики, і цю залежність необхідно визначити.

Що стосується решти метрик, то для них так само було побудовано відповідні залежності, які наведено на рис. 2.28.

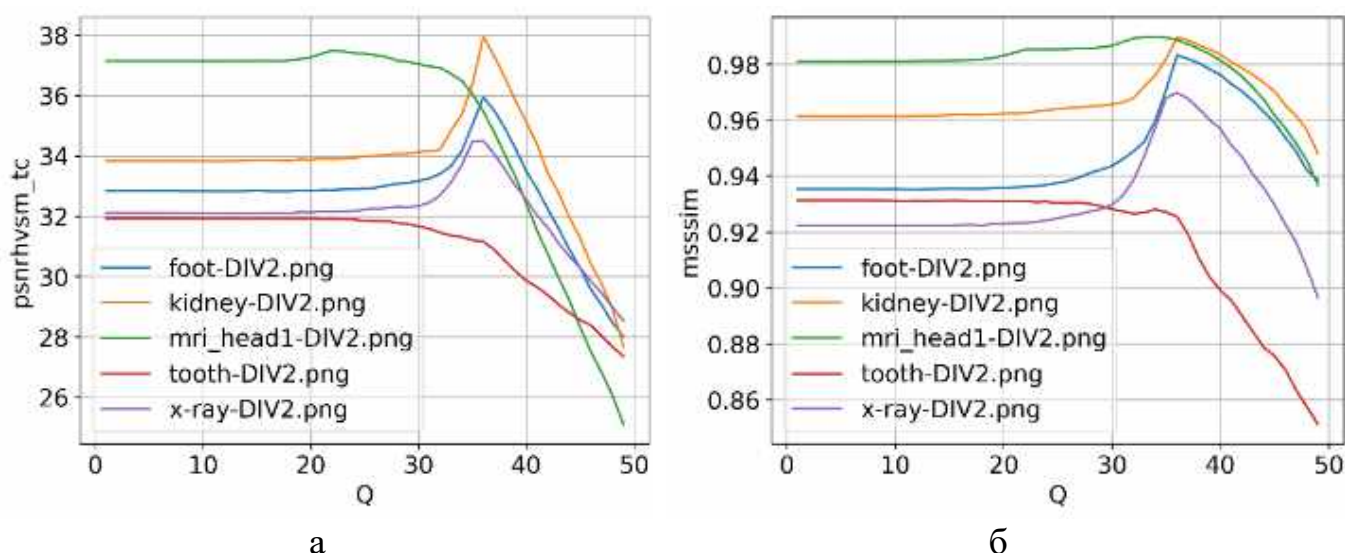


Рис 2.28 – Залежності PSNR-HVS-M^{ct} (Q) (а) та MS-SSIM^{ct} (Q) (б) для набору даних медичних зображень

Для обох метрик існують зображення, які мають оптимальну робочу точку (x-ray, kidney, і foot), і зображення, які не мають OPT, але можуть мати локальний максимум (зображення, що мають назву tooth і mri_head1).

Для кращого розуміння положення оптимальної робочої точки та її залежності від властивостей зображення (зокрема, його складності), для сигнально-залежних завад, отримані дані були структуровані в таблицях 2.3 і 2.4. У цих таблицях показано значення всіх трьох метрик, використаних у цьому дослідженні (PSNR, PSNR-HVS-M і MS-SSIM), і коефіцієнт стиснення, отримані для значення Q, у якому спостерігалася оптимальна робоча точка. В таблиці також є значення середнього значення для кожного зображення, виходячи з того, що ґрунтуючись на отриманих раніше залежностях, можна припустити, що позиція Q_{OPT} може залежати від середнього, розрахованого для конкретного зображення.

З таблиці 2.3 видно, що для метрики PSNR, за винятком зображення Diego, OPT присутня для всіх інших зображень. Для метрики PSNR-HVS-M розглянуті зображення ДЗЗ не мають OPT. Для метрики MS-SSIM OPT відсутня для таких зображень, як tooth, Diego і baboon.

Таблиця 2.3 – положення OPT для різних метрик розрахованих для зображень з оригінального набору

Назва зображення	Середнє для зображень	Q_{OPT} (PSNR)	Q_{OPT} (PSNR-HVS-M)	Q_{OPT} (MS-SSIM)	KC_{psnr}
BABOON.png	129,699	37	-	-	6,6
BARBARA.png	112,447	37	37	38	15,92
goldhill.png	112,049	38	-	38	12,61
lenna.png	124,102	38	38	38	34,42
Peppers.png	104,210	38	37	38	37,61
DIEGO.png	115,342	30	-	-	4,41
fr01.png	97,899	37	-	36	15,65
fr02.png	98,689	37	-	35	14,15
fr03.png	105,292	36	-	35	11,61
fr04.png	98,064	33	-	32	7,62
frisco.png	117,796	37	36	37	40,89
foot.png	122,608	38	38	39	67,86
kidney.png	102,264	37	37	37	44,43
mri_head1.png	61,999	37	26	35	18,39
tooth.png	156,203	39	-	-	18,58
x-ray.png	140,174	39	39	39	69,14

Для додаткового набору зображень (зображення зі зменшеним середнім) дані подано в таблиці 2.4.

Слід зазначити, що зображення DIV-2 для всіх метрик змістили позиції Q_{OPT} у бік менших значень. Значення зміщення варіюється від 2 до 3 залежно від метрики і зображення. Загальна різниця (зміщення) в положенні Q_{OPT} варіюється від 1 до 5 залежно від метрики, за винятком зображення mri_head1-DIV2, яке має локальний максимум, що його потенційно можна назвати OPT, і розташовується в Q , що дорівнює 22.

Варто зазначити, що значення коефіцієнта стиснення, наведеного в таблицях 2.3, 2.4, наведено для Q_{OPT} , отриманого для метрики PSNR, оскільки для цієї метрики Q_{OPT} з'являється частіше, якщо порівнювати з іншими розглянутими метриками.

З огляду на отримані дані, наведені в таблицях 2.3 і 2.4, видно, що позиція Q_{OPT} відрізняється залежно від зображення і, як результат, стає неможливим використовувати вираз (2.3) для визначення Q_{OPT} .

Таблиця 2.4 – положення OPT для різних метрик розрахованих для зображень із "DIV-2" набору

Назва зображення	Середнє для зображень	Q_{OPT} (PSNR)	Q_{OPT} (PSNR-HVS-M)	Q_{OPT} (MS-SSIM)	KC_{psnr}
BABOON-DIV2	64,595	35	-	34	9,09
BARBARA-DIV2	55,973	35	34	35	20,06
goldhill-DIV2.png	55,799	35	-	35	20,48
lenna-DIV2.png	61,802	35	35	35	29,01
Peppers-DIV2.png	51,854	35	35	35	30,18
DIEGO-DIV2.png	57,420	34	-	-	7,73
fr01-DIV2.png	48,700	34	-	34	9,79
fr02-DIV2.png	49,094	34	-	33	9,11
fr03-DIV2.png	52,396	34	-	33	9,62
fr04-DIV2.png	48,783	35	-	30	12,39
frisco-DIV2.png	58,648	34	33	34	23,38
foot-DIV2.png	61,091	36	36	36	44,95
kidney-DIV2.png	50,826	36	36	36	36,98
mri_head1-DIV2.png	30,715	35	22	34	13,7
tooth-DIV2.png	77,862	36	-	-	21,3
x-ray-DIV2.png	69,840	36	36	36	43,44

Оскільки вже було отримано залежності метрики від Q , то це дає змогу дізнатися, де саме знаходиться положення OPT. Для кожного зображення можливо обчислити середнє значення, припускаючи, що MSE залежить від цього параметра. Отримані значення були використані для отримання діаграми розсіювання і подальшого вписування кривих. Результат такої операції представлено на рис. 2.29.

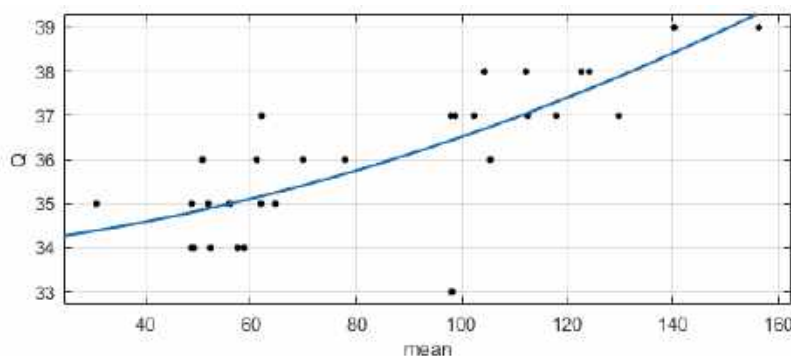


Рис 2.29 – Діаграма розсіювання Q_{OPT} проти середнього значення і вписана крива

Для кривої, продемонстрованої на рисунку 2.29, було отримано таке модифіковане рівняння для визначення Q_{OPT} :

$$Q_{OPT} = (11,85 \cdot x^2 + 864,4 \cdot x + 2,6 \cdot 10^6) / (x + 7,7 \cdot 10^4) \quad (2.5)$$

де x - середнє значення цільового зображення.

Таким чином, для кожного зображення, що підлягає стисненню і пошкодженого пуасонівським шумом, легко визначити Q_{OPT} .

Крім того, варто звернути увагу на ще одну цікаву особливість. Результати показують, що первинні зображення в основному мають вищі значення КС, ніж їх аналоги DIV-2. Однак аналіз також показує, що деякі зображення DIV-2 можуть мати значення КС більші, ніж вихідні зображення. Залежності КС від Q для пари таких зображень представлені на рис. 2.30.

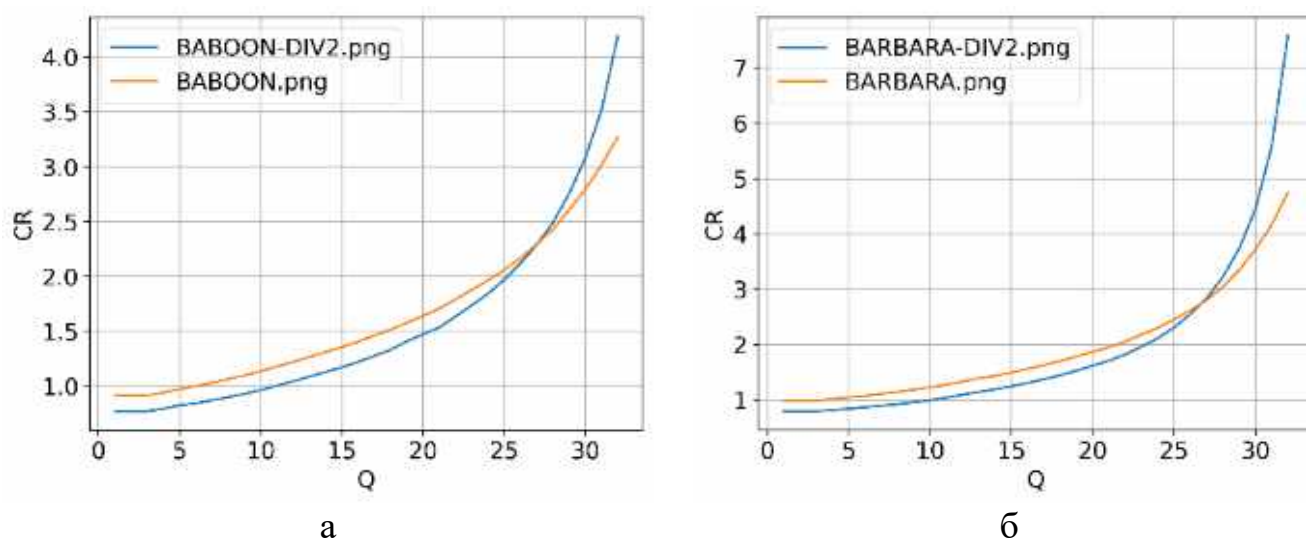


Рис 2.30 – Залежності КС(Q) для двох пар зображень

З рис. 2.30 видно, що обидва зображення мають певну точку ($Q = 27$), де ступінь стиснення для зображень DIV-2 стає більшим, порівнюючи з вихідними версіями відповідних зображень.

2.3.3 Стиснення з використанням перетворення Анскомба

Під час стиснення зображень, що містять сигнально-залежний шум, варто приділити увагу й іншому традиційному підходу, який може допомогти впоратися з даним типом шуму, - використанню варіаційно-стабілізуючого перетворення (ВСП). Варто розглянути, як використання цього підходу може вплинути на роботу VP8 кодера, іншими словами, як використання цього підходу спільно із згаданим кодером позначиться на фінальних характеристиках стисненого зображення.

У цьому випадку буде використано модифіковану версію перетворення Анскомба [113], що дає змогу замінити випадкову змінну з розподілом Пуассона на випадкову змінну зі стандартним розподілом Гауса. Модифіковане перетворення, що використовується, можна описати таким чином:

$$X_{ans} = [16 \times \sqrt{x_p + \frac{3}{8}}]. \quad (2.6)$$

де x_p — дані Пуассона, а $[]$ означає округлення до найближчого цілого числа.

Зворотне перетворення можна виразити як:

$$X_{inv} = \left(\frac{X_{ans}}{16}\right)^2 - \frac{3}{8}. \quad (2.7)$$

Для початку необхідно внести ясність, коли буде використовуватися пряме і зворотне перетворення: модифіковане перетворення Анскомба буде застосовуватися перед стисненням, а зворотне перетворення буде виконуватися після декомпресії зображення. В усьому іншому було використано ту саму процедуру для створення кривих швидкість-спотворення, що й раніше. Отримані залежності представлено на рисунку 2.31.

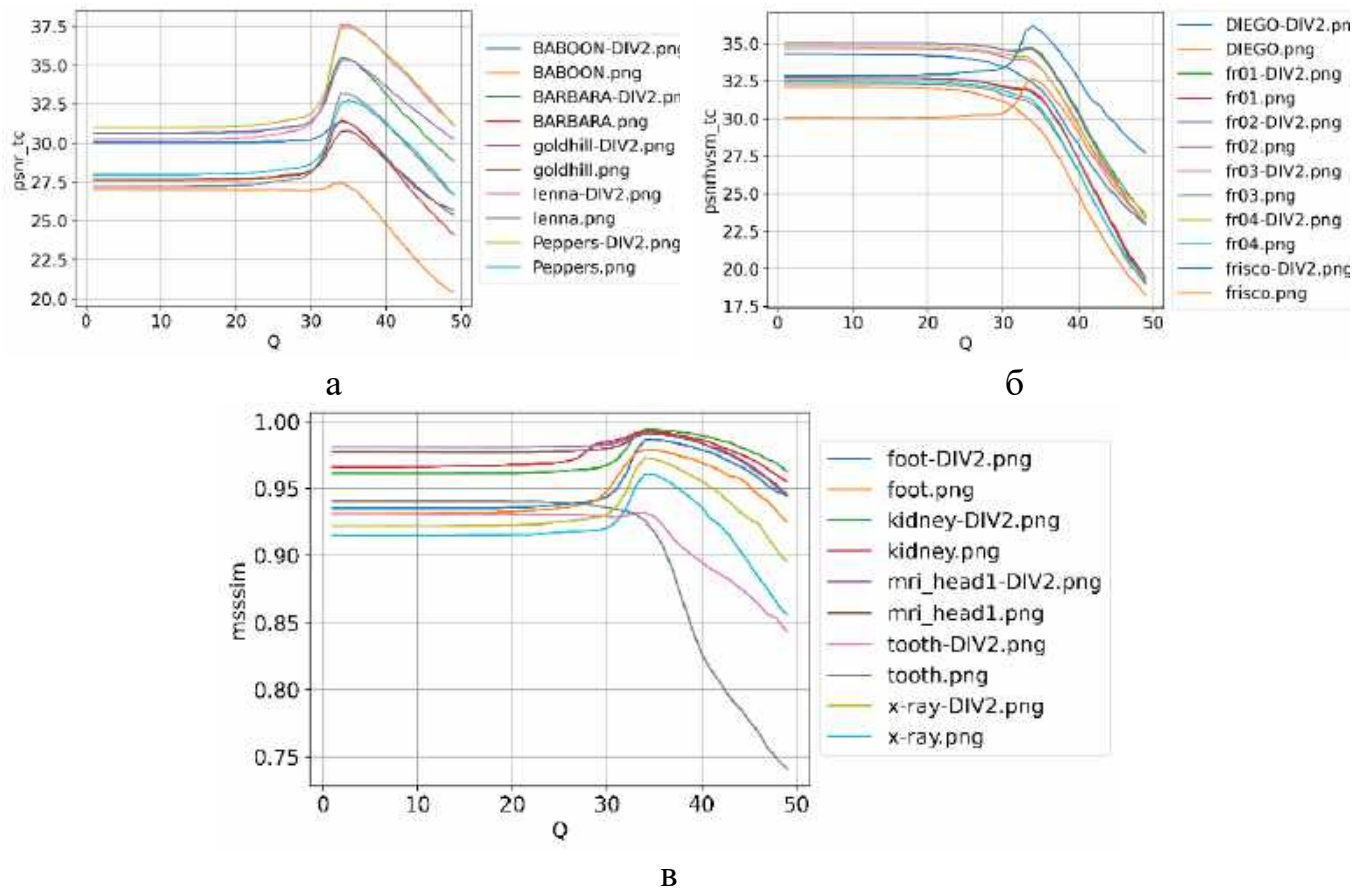


Рис 2.31 – Залежності PSNR^{ct} (а), PSNR-HVS-M^{ct} (б) та MS-SSIM^{ct} (Q) (в) для різних наборів зображень

Після модифікованого перетворення Анскомба дисперсія адитивного шуму в усіх зображеннях приблизно дорівнює 64. Якщо вихідні дані були 8-бітними, то отримані зображення після (2.6) також будуть 8-бітними. Крім того, з практичного погляду важливо, щоб усі позиції метричних піків перебували за Q , що дорівнює 34, що дає змогу встановити фіксоване Q за наявності ОПТ.

Тепер необхідно коротко проаналізувати кожну метрику. Перша аналізована залежність - PSNR^{ct} від Q (рис. 2.31, а). Прості та середньо-структурні зображення мають значні піки, що відображають суттєве підвищення якості. Приклади складних зображень, таких як Baboon і його розділена версія, можуть являти собою випадки невеликого збільшення якості в ОПТ, цей факт призводить до висновку, що робити стиснення в ОПТ складних зображень є правильним рішенням. Ще один аспект, який необхідно відзначити, це те, що оригінальні версії зображень мають менші значення PSNR порівняно з розділеними версіями.

У другому випадку (рис. 2.31, б) отримано залежності метрики PSNR-HVS-M^{ct} для зображень дистанційного зондування. Ці залежності характеризуються такою поведінкою: тільки для зображення простої структури Frisco і його розділеної версії спостерігаються ОПТ для одного і того ж $Q_{\text{OPT}}=34$. У свою чергу для того самого значення Q для частини зображень спостерігаються локальні максимуми, при чому для зображень із набору "DIV2" ці піки є більш вираженими. Для складних зображень не спостерігаються ні ОПТ, ні локальний максимум.

Остання ситуація - це метрика MS-SSIM^{ct} для медичних зображень (рис. 2.31, в), і вона також має властивості, аналогічні попереднім двом метрикам. Однак спад у якості після досягнення ОПТ є не таким стрімким (виняток - зображення з назвою tooth).

2.3.4 Порівняння двох підходів

Після того, як було підтверджено, що оптимальна робоча точка існує як для прямого стиснення кольорових зображень, так і для варіанту, який використовує перетворення Анскомба, виникає логічне запитання: "який із цих варіантів варто використовувати?", оскільки кожен із них має свої переваги та недоліки.

Щоб визначити, який підхід кращий, необхідно проаналізувати такі показники ефективності: отриману якість зображень і коефіцієнт стиснення, а також тривалість обчислень, необхідну для обох підходів.

Отримані значення метрик і коефіцієнтів стиснення подано в таблиці 2.5. У цій таблиці наведено отримані значення в оптимальній робочій точці, що відповідають двом підходам. Як уже зрозуміло із залежностей отриманих раніше для двох підходів, у разі модифікованого підходу на основі перетворення Анскомба положення ОПТ для різних зображень фіксоване, тоді як для прямого стиснення (без використання перетворення Анскомба) оптимальна робоча точка розташовується в різних місцях залежно від властивостей зображення (таблиця 2.4).

Таблиця 2.5 – Порівняння двох підходів в ОПТ для метрики PSNR

Назва зображення	Прямий підхід			Підхід на базі перетворення Анскомба		
	Q _{ОПТ}	PSNR	CR	Q _{ОПТ}	PSNR	CR
BABOON.png	37	27,12	6,6	34	27,42	6,39
BABOON-DIV2	35	31,13	9,06	34	31,39	6,58
lenna.png	38	32,79	34,42	34	33,19	23,24
lenna-DIV2.png	35	37,22	29,01	34	37,57	19,67
DIEGO.png	30	27,39	4,41	34	26,87	9,09
DIEGO-DIV2.png	34	30,88	7,73	34	30,92	6,51
frisco.png	37	34,97	40,89	34	34,99	39,02
frisco-DIV2.png	34	39,17	23,38	34	39,15	23,01
tooth.png	39	27,52	23,92	34	27,77	11,92
tooth-DIV2.png	36	32,03	21,3	34	32,13	11,42
kidney.png	37	34,92	44,43	34	34,74	21,32
kidney-DIV2.png	36	41,13	36,98	34	42,36	13,43

Таблиця 2.6 – Порівняння двох підходів в ОПТ для PSNR-HVS-M і MS-SSIM

Назва зображення	Прямий підхід			Підхід на базі перетворення Анскомба		
	Q _{ОПТ}	PSNR-HVS-M	MS-SSIM	Q _{ОПТ}	PSNR-HVS-M	MS-SSIM
BABOON.png	37	29,57	0,96	34	29,74	0,96
BABOON-DIV2	35	32,05	0,96	34	32,65	0,96
lenna.png	38	31,88	0,97	34	32,59	0,97
lenna-DIV2.png	35	35,35	0,98	34	36	0,98
DIEGO.png	30	31,56	0,97	34	29,72	0,97
DIEGO-DIV2.png	34	32,16	0,96	34	32,38	0,96
frisco.png	37	32,71	0,97	34	32,7	0,97
frisco-DIV2.png	34	36,31	0,98	34	36,2	0,98
tooth.png	39	27,54	0,92	34	28	0,93
tooth-DIV2.png	36	31,16	0,93	34	31,48	0,93
kidney.png	37	32,03	0,99	34	31,87	0,99
kidney-DIV2.png	36	37,95	0,99	34	39,99	0,99

У таблиці 2.5 наведено пари зображень (оригіналу і розділеного) для складної і простої структури з усіх трьох використаних вище наборів. По-перше, необхідно відзначити той факт, що отримані значення метрики якості зображення PSNR майже завжди вищі для підходу, заснованого на методі Анскомба. Водночас стандартний

(прямий) метод стиснення часто може забезпечити значно вищі коефіцієнти стиснення.

У таблиці 2.6 наведено значення двох інших метрик в OPT. Як видно, значення MS-SSIM практично однакові. Значення PSNR-HVS-M поводяться майже так само, як і в разі PSNR: підхід, заснований на перетворенні Анскомба, має невелику перевагу порівняно зі стандартним.

Останній аспект, який необхідно розглянути, - це час, необхідний для виконання стиснення в обох підходах. Спочатку визначимо час, необхідний для виконання процедур стиснення і декомпресії. Тут необхідно уточнити характеристики комп'ютера, що використовується для аналізу, оскільки час, що витрачається на виконання конкретного завдання, може варіюватися залежно від пристрою, на якому запускається алгоритм. У цьому експерименті було використано платформу на базі Windows, оснащену процесором Intel i5 9600k з 16 гігабайтами оперативної пам'яті. Також варто відзначити той факт, що в даному випадку вже відомо Q_{OPT} , і не використовується ніяких підходів для прогнозування Q_{OPT} ; іншими словами, Q фіксоване і дорівнює 34. Для медичних зображень середній час стиснення і розпакування становить 0,2547 с; для складних зображень, таких як "tooth", необхідний час становить 0,3703 с, а для простих зображень - близько 0,2 с. Ще один факт, який необхідно обговорити, - це те, що час, необхідний для DIV2-зображень, менший, ніж для звичайної версії зображення. Наприклад, DIV2-версія зображення tooth стискається за 0,2847 секунди.

Аналогічні тенденції спостерігаються і для зображень із двох інших наборів: складні зображення потребують більше часу на стиснення і розпакування; версії DIV 2 потребують менше часу, ніж оригінали. Варто також зазначити, що зображення з набору даних дистанційного зондування потребують більше часу: для зображення Diego потрібно 0,3981 с, а для Frisco - 0,3172 с. Аналогічна ситуація відбувається і зі стандартними тестовими зображеннями: Baboon стискається/розтискається за 0,3984 с, а Lenna - за 0,3354 с.

Крім того, необхідно дізнатися, чи має модифіковане перетворення Анскомба істотний вплив на час стиснення/розпакування даних. Для відповіді на це запитання

в таблиці 2.7 подано середній і сумарний час для трьох наборів зображень. Загальний (або сумарний) час - це час, необхідний для стиснення і розпакування всіх зображень у наборі, зі свого боку, середній час - це час, витрачений у середньому на одне зображення для обраного набору. Також варто нагадати той факт, що набір медичних зображень містить 10 зображень, набір зображень дистанційного зондування - 12 зображень, а стандартний набір - 10 зображень.

Таблиця 2.7 – Порівняння двох підходів в OPT для PSNR-HVS-M і MS-SSIM

Назва набору даних	сумарний час для прямого підходу	сумарний час для підходу на базі перетворення Анскомба	середній час для прямого підходу	середній час для прямого підходу на базі перетворення Анскомба
Medical	2,5467	2,5688	0,2547	0,2569
RS	3,7681	3,9829	0,3141	0,3319
Standard	3,0031	3,0934	0,3003	0,3093

Як видно, пряме і зворотне перетворення додають від 0,0221 с до 0,2148 с до сумарного часу і від 0,009 с до 0,0178 с до середнього значення часу. З цього можна зробити висновок, що перетворення Анскомба можна використовувати, у разі якщо потрібна більша якість зображення, а високий ступінь стиснення не є критичним. У випадках, коли потрібен високий ступінь стиснення, рекомендується стискати зображення без перетворення Анскомба, але в цьому разі для досягнення Q_{OPT} необхідно використовувати який-небудь алгоритм прогнозування OPT.

Важливо зазначити, що кодер BPG працює таким чином, що менші значення Q вимагають більше часу на стиснення і розпакування. Наприклад, для стиснення стандартного набору даних за $Q = 1$ в середньому потрібно 0,4452 с, а загальний час для всіх зображень у наборі даних становить 4,452 с. Той самий набір даних за $Q = 39$ можна стиснути і розпакувати за 0,1987 с у середньому і 1,987 с загалом для всіх зображень у наборі.

2.4 Аналіз спотворень, що виникають під час стиснення з втратами на основі ВРГ зображень без шуму і з шумом

Впродовж усього розділу було проаналізовано різні ситуації стиснення за допомогою ВРГ кодера зображень (як півтонових, так і кольорових), що могли як не містити шум, так і бути спотвореними різними типами шуму. Але відкритим залишилося питання, які саме завади вносить сам кодер, а також чи впливає наявність шуму на зображенні на ці завади. Крім того, бажано відповісти на запитання, якими є характеристики цих спотворень і, наприклад, чи є статистичні характеристики гаусовими або негаусовими.

Раніше вже згадувалося, що існують залежності між властивостями зображення і результатами стиснення. Виходячи з цього факту, для цього дослідження було використано два зображення, які можна розглядати як такі, що мають просту та складну структури, приклади даних зображень, представлено на рис. 2.32. Зображення Frisco є прикладом простого зображення, тоді як зображення fr01 є зображенням зі складною структурою (містить значно більше пікселів, що відповідають краям, дрібним деталям і текстурам).



а



б

Рис 2.32 – Використані тестові зображення Frisco (а) и fr01 (б)

Загальний метод аналізу спотворень для випадку відсутності шуму може бути виконаний таким чином: узяти істинне не уражене шумом 8-бітове напівтонове

зображення, яке можна позначити як I_{ij}^t , $i = 1, \dots, N_I$; $j = 1, \dots, N_J$ (N_I та N_J є розміром зображення); потім стиснути це зображення з різними значеннями параметра Q . Кожне стиснене зображення можна представити у вигляді $I_{ij}^c(Q)$, $i = 1, \dots, N_I$; $j = 1, \dots, N_J$, N_I та N_J . У результаті спотворення, отримані під час стиснення з втратами для конкретних значень Q , можуть бути описані різницею:

$$\Delta_{ij}(Q) = I_{ij}^t - I_{ij}^{ct}(Q),, i = 1, \dots, N_I; j = 1, \dots, N_J. \quad (2.8)$$

де I_{ij}^{ct} - стиснуте безшумне зображення.

Далі необхідно провести аналіз отриманого масиву різниць (2.8). Такий аналіз може бути виконано різними способами. Один із них - провести простий статистичний аналіз, наприклад, отримати гістограми відмінностей і перевірити, чи відповідають їх розподіли нормальному закону. Крім того, варто розрахувати статистичні параметри, щоб відповісти на деякі додаткові запитання: наскільки важким є хвіст розподілу, якщо він виявився негаусовим, і чи є розподіл симетричним? Для цього далі використовуються такі параметри, як дисперсія, середнє, ексцес і асиметрія. Два останні параметри мають бути близькими до нуля, якщо розподіл випадкової величини (випадкового процесу) близький до нормального.

Варто відзначити той факт, що далі буде використано кілька значень (певний діапазон) величин Q . Це пояснюється двома фактами. По-перше, на практиці може знадобитися забезпечити різні характеристики стиснених зображень залежно від вимог та їхнього пріоритету. По-друге, для випадку стиснення зашумлених зображень необхідно приділяти основну увагу околу ОРТ, оскільки вона має велике практичне значення в реальних умовах. Виходячи з цих двох чинників було обрано діапазон значення Q , які починаються з 15 і закінчуються 40 з кроком 5.

Наступний крок - перевірити, чи впливає на поведінку спотворень шум, який може бути присутнім на стискаємому зображенні (оскільки на практиці важко отримати зображення без шуму, за винятком випадків, коли зображення створюються комп'ютером). У разі зашумлених зображень можна використовувати два способи використання формули різниць (2.8). Перший - використовувати (2.8), як це було

раніше (для зображень без шуму). У цьому випадку вихідне зображення без шуму порівнюється із зображенням, яке було спотворено шумом, а потім стиснуто з деяким параметром Q . Це дає змогу досліджувати властивості залишкового шуму в стиснених зображеннях, що важливо для методів постфільтрації стиснених зображень [110]. Другий варіант - використовувати I_{ij}^n (зображення з шумом) замість I_{ij}^t (зображення без шуму); іншими словами, стиснене зображення порівнюється з відповідним зашумленим зображенням.

Перший варіант - перевірити, якими є властивості комбінованих спотворень, тобто спільного впливу шуму і стиснення з втратами. У другому випадку різницевий аналіз характеризує властивості спотворень, що вносяться кодером у зашумлені зображення, або те, як наявність шуму в зображенні впливає на розподіл спотворень.

Як модель шуму для цього аналізу було використано адитивний білий гаусовий шум (AWGN), який відомий як одна з найпростіших моделей шуму і хороша відправна точка для досліджень.

2.4.1 Характеристики шуму, внесеного кодером для зображень без шуму

Почнемо аналіз із простого зображення Frisco, де шум відсутній. Розраховані статистичні параметри, отримані для цього випадку, подано в таблиці 2.8.

Таблиця 2.8 – Статистичні характеристики спотворень для зображення Frisco

Q	Ексцес	Асиметрія	Середнє	Дисперсія
15	0,2003	-0,0021	-0,0026	0,5289
20	0,4136	0,0147	-0,0019	1,1677
25	0,9217	0,0125	-0,0204	2,5667
30	2,0129	0,0539	-0,0341	5,2719
35	3,4047	0,0835	-0,0880	10,6142
40	4,9193	0,2543	0,0062	22,5580

З даних таблиці 2.8 можна зробити такі висновки: за Q , що дорівнює від 15 до 20, значення ексцесу близькі до нуля, як і значення асиметрії, а отже, отримані спотворення можна розглядати як гаусові. Зі збільшенням Q починають зростати і

значення ексцесу. Для $Q = 25$ ексцес дорівнює приблизно одиниці, і кожен наступний крок збільшує значення ексцесу на одиницю. У результаті розподіл дедалі менше і менше схожий на гаусовий. Також необхідно зазначити, що значення середнього й асиметрії близькі до нуля для всіх розглянутих випадків Q . Таким чином, виходить, що розподіл внесених помилок симетричний і має важкий хвіст. Також варто зазначити, що дисперсія відмінностей (середня квадратична помилка (MSE) внесених викривлень) швидко збільшується зі зростанням Q , і викривлення стають помітними за $Q=40$ (зазвичай це відбувається, якщо MSE перевищує 20).

Наступна ситуація - стиснення складнішого зображення. Статистичні параметри отриманих відмінностей подано в таблиці 2.9.

Таблиця 2.9 – Статистичні характеристики спотворень для зображення fr01

Q	Ексцес	Асиметрія	Середнє	Дисперсія
15	-0,0226	0,0037	-0,0155	0,5855
20	0,1362	0,0277	-0,0139	1,8969
25	0,3321	0,0465	-0,0505	6,0132
30	0,8125	0,0902	-0,0359	16,4319
35	1,6626	0,1742	-0,0735	41,4552
40	2,9806	0,2448	-0,1909	93,4525

У цьому випадку ситуація аналогічна, за винятком таких моментів. Ексцес близький до нуля для широкого діапазону значень Q (від 15 до 25), і стає близьким до одиниці для $Q = 30$. З цих результатів можна зробити наступний висновок - розподіл спотворень зберігає свої гаусові властивості навіть за більших значень Q порівняно з випадком стиснення зображень простих стриктур. Ще одна відмінність між стисненням зображень простої та складної структури полягає в тому, що дисперсія відмінностей є вищою для складних зображень, особливо за $Q>25$.

Отримані висновки також можна підтвердити аналізом отриманих гістограм. Наприклад, для зображення frisco (рис. 2.33) за Q , що дорівнює 30 і більше, розподіл стає негаусівським (важкохвостим).

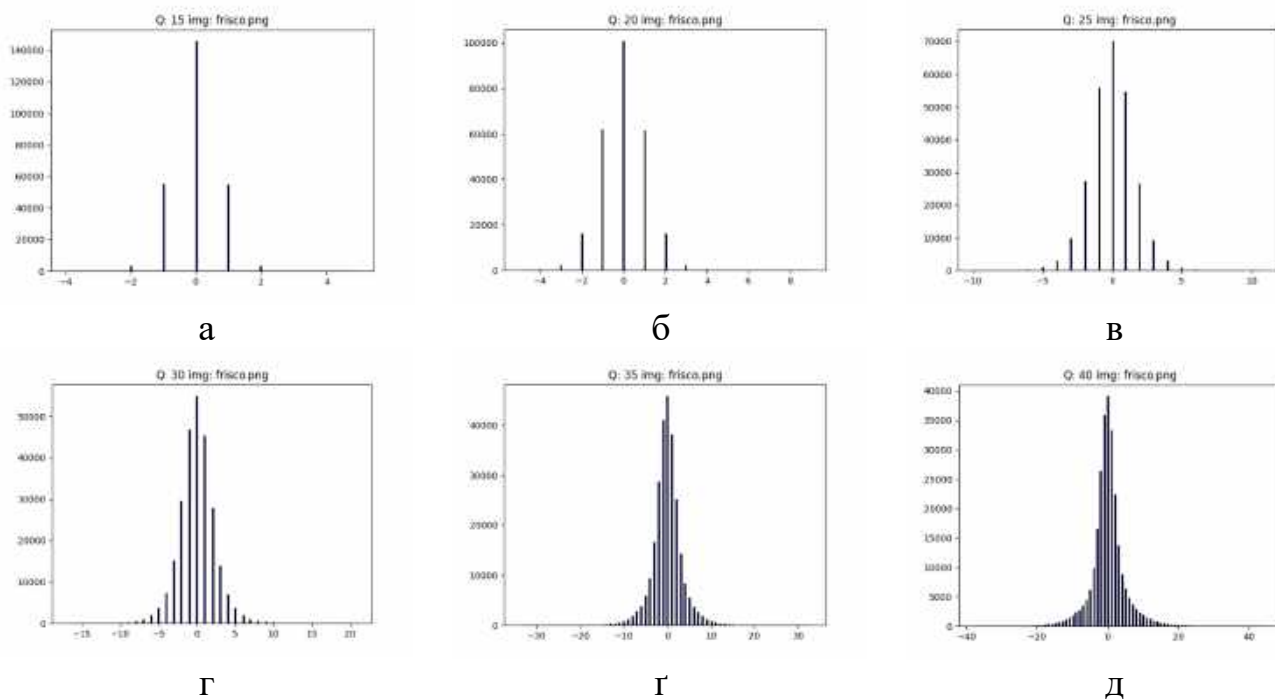


Рис 2.33 – Гістограма різниці зображення frisco для шести значень Q

Розподіли різниць (2.8), отримані під час стиснення зображення fr01, представлені на рис. 2.34.

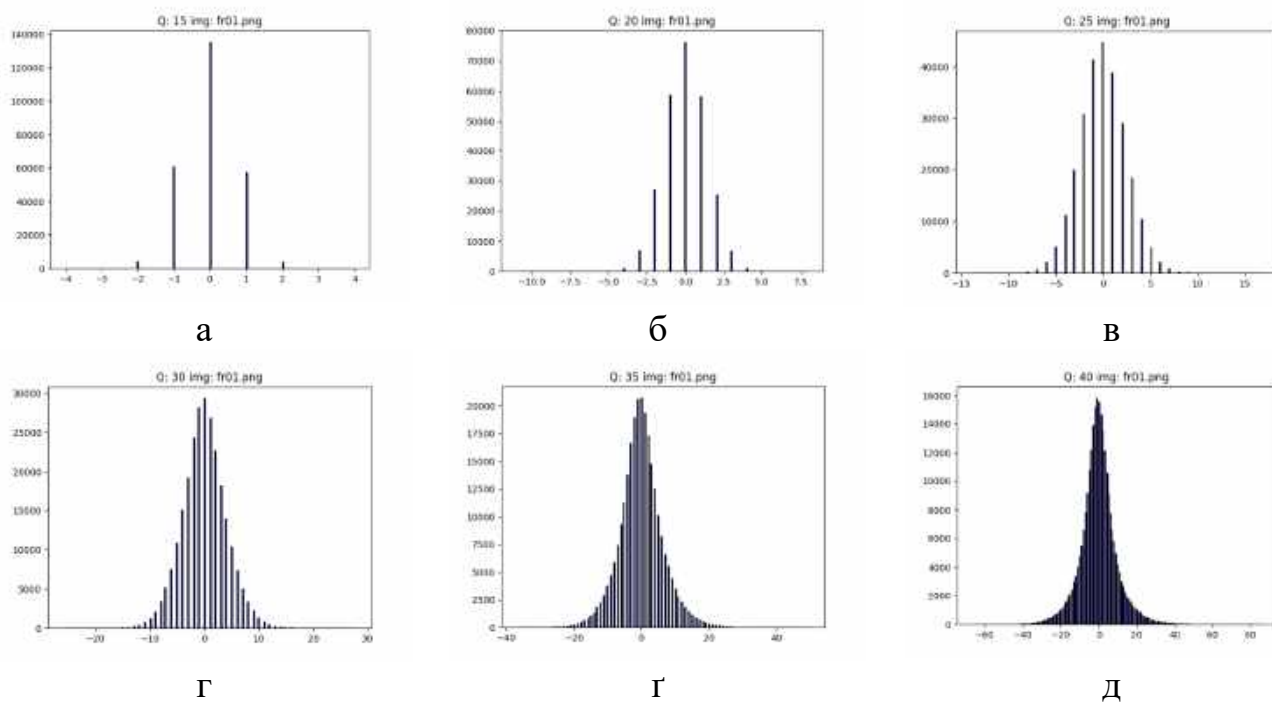


Рис 2.34 – Гістограми значень різниць зображення fr01 для шести значень Q

Водночас, аналізуючи дані на рис 2.34, можна зробити висновок, що загалом, існують аналогічні ситуації, як і для простого зображення. Розподіл стає менш гаусовим зі збільшенням Q . Але в цьому випадку розподіли явно негаусові за великих Q .

2.4.2 Комбіновані шуми кодера

Раніше було згадано, що у зв'язку з наявністю шуму на зображенні можливі два варіанти розгляду спотворень. У цьому пункті буде розглянуто випадок, коли до зображення додаватимуть шум і буде використовуватись формула (2.8), але в цьому разі стискатимуть зашумлене зображення. У результаті формулу (2.8) можна змінити таким чином:

$$\Delta_{ij}(Q) = I_{ij}^t - I_{ij}^{cn}(Q),, i = 1, \dots, N_i; j = 1, \dots, N_j. \quad (2.9)$$

де I_{ij}^{cn} — стиснуте зображення з шумом.

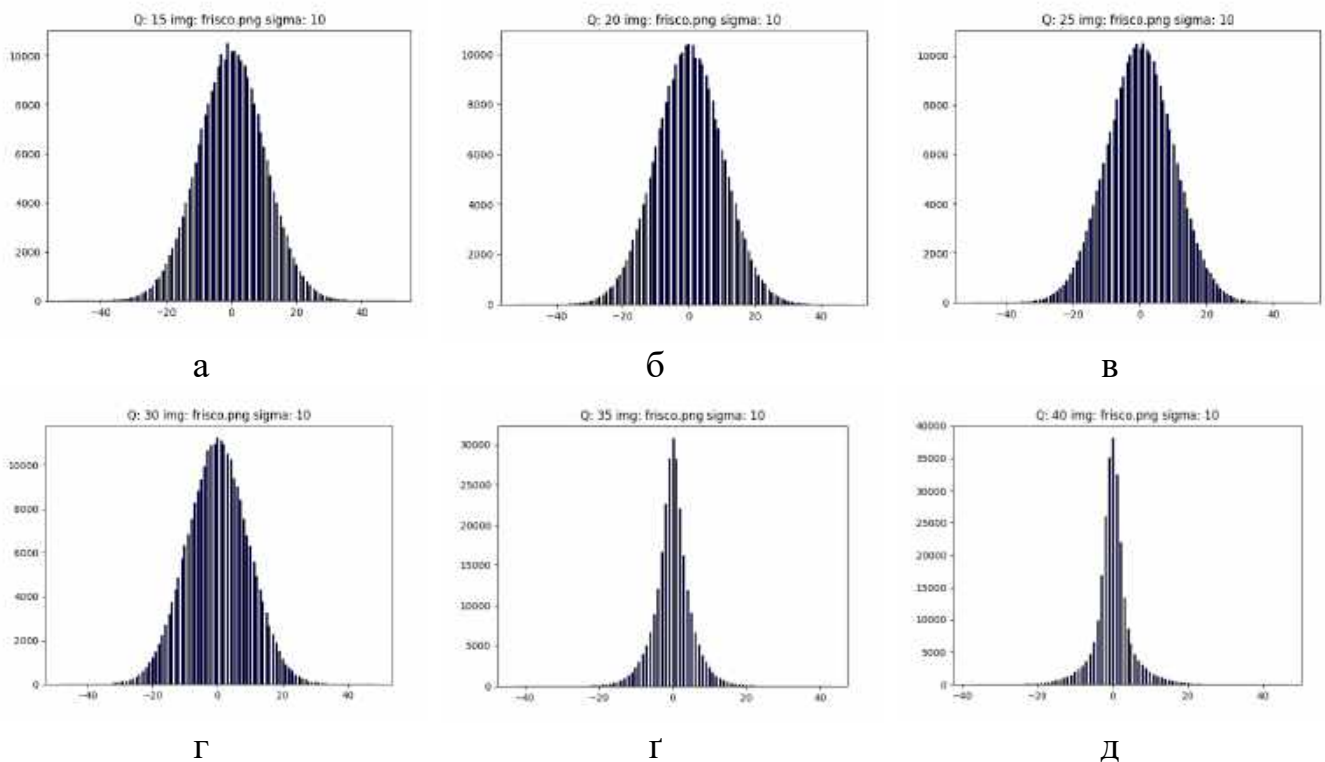


Рис 2.35 – Гістограми відмінностей між безшумними і стиснутими зашумленими зображеннями для зображення frisco

Спочатку розглянемо отримані для цього варіанта гістограми, представлені на рис. 2.35 та 2.36. Також важливо зазначити, що дисперсія AWGN дорівнює 100 (у цьому разі шум видно на більшості зашумлених зображень).

З аналізу даних рис. 2.35 можна бачити, що для Q у діапазоні від 15 до 30 розподіл має гаусову форму. За більшого Q складається враження, що залишковий шум у стисненому зображенні стає негаусовим. Ефект фільтрації шуму за рахунок стиснення з втратами спостерігається для $Q=35$ і $Q=40$, де $Q=35$ відповідає оптимальній робочій точці.

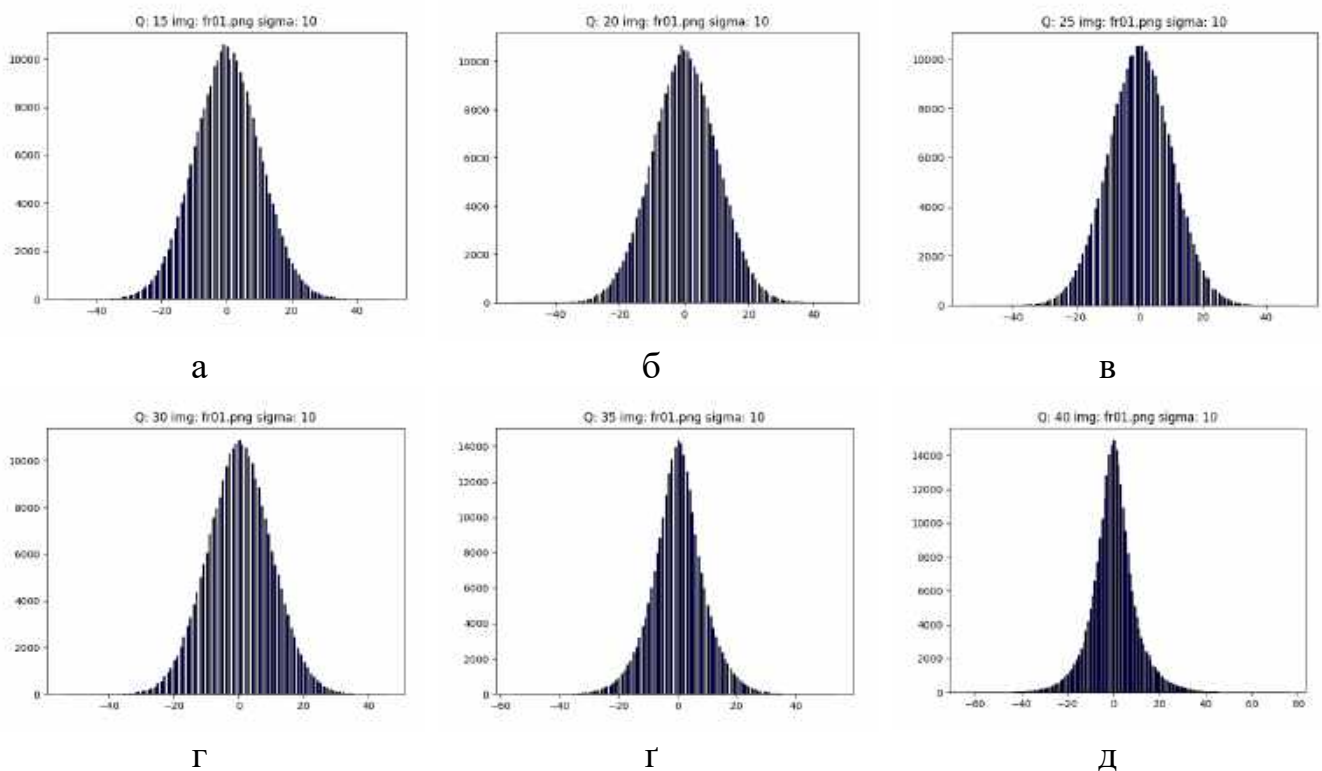


Рис 2.36 – Гістограми відмінностей між безшумними і стиснутими зашумленими зображеннями для зображення fr01

У випадку складнішого зображення (рис. 2.36) тенденції аналогічні. Цього разу гістограми можуть мати додаткові або локальні піки (вони не такі гладкі, як раніше). Розподіли на рисунках 2.36 (г) і 2.36 (д) здаються вужчими, ніж для інших гістограм. Ідеться про часткове придушення шуму, коли до зашумлених зображень застосовується стиснення з втратами.

Попередні висновки, що впливають з аналізу гістограм (рис. 2.35 і 2.36), підтверджуються статистичними даними, представленими в таблицях 2.11 і 2.12. Ексцес близький до нуля за $Q=30$ і менше. Але він не близький до нуля для $Q=35$ і 40 , що вказує на негаусовість властивостей залишкового шуму. Ще один цікавий момент полягає в тому, що дисперсія спотворень менша за дисперсію AWGN для $Q=30, 35$ і 40 - це пов'язано з вищезгаданим ефектом фільтрації шуму.

Таблиця 2.9 – Статистичні характеристики спотворень (залишкового шуму) зашумленого зображення frisco

Q	Ексцес	Асиметрія	Середнє	Дисперсія
15	0,0198	-0,0105	-0,0252	104,0136
20	0,0247	-0,0078	-0,0139	103,3174
25	0,0329	-0,0052	0,0046	100,1885
30	0,0419	-0,0079	0,0031	90,8979
35	2,6541	0,0054	-0,0148	24,9837
40	4,7843	0,2334	-0,0180	25,0086

Таблиця 2.10 – Статистичні характеристики спотворень (залишкового шуму) зашумленого зображення fr01

Q	Ексцес	Асиметрія	Середнє	Дисперсія
15	0,0382	-0,0074	0,0285	103,3563
20	0,0389	-0,0079	0,0303	103,0498
25	0,0558	-0,0095	0,0535	101,6223
30	0,0716	-0,0077	0,0331	98,9175
35	0,9200	-0,0102	0,0107	77,7847
40	2,6319	0,1937	-0,0498	98,4561

2.4.3 Характеристики шуму, що вноситься кодером

Ситуація, що залишилася, яку слід розглянути, - це коли відмінності, отримані в результаті стиснення з втратами для конкретних значень Q , можна описати як:

$$\Delta_{ij} = I_{ij}^n - I_{ij}^{cn}, i = 1, \dots, N_i; j = 1, \dots, N_j \quad (2.10)$$

де I_{ij}^n — зашумлене зображення (штучно зашумлене зображення у симуляціях), а I_{ij}^{cn} — стиснене зашумлене зображення.

Гістограми Δ_{ij} представлені на рисунках 2.37 і 2.38. Загалом, на перший погляд, для обох зображень здається, що спотворення є більш гаусовими, ніж ті, що були показані раніше.

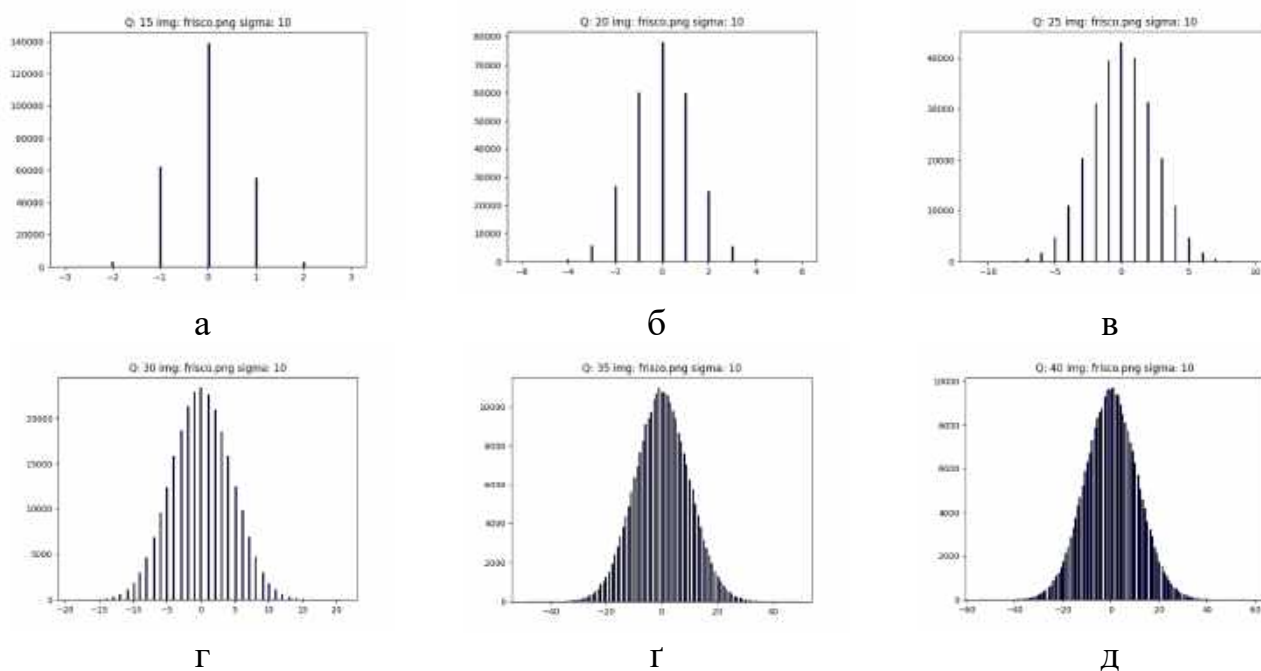


Рис 2.37 – Гістограма різниць (2.10) для зашумленого зображення frisco

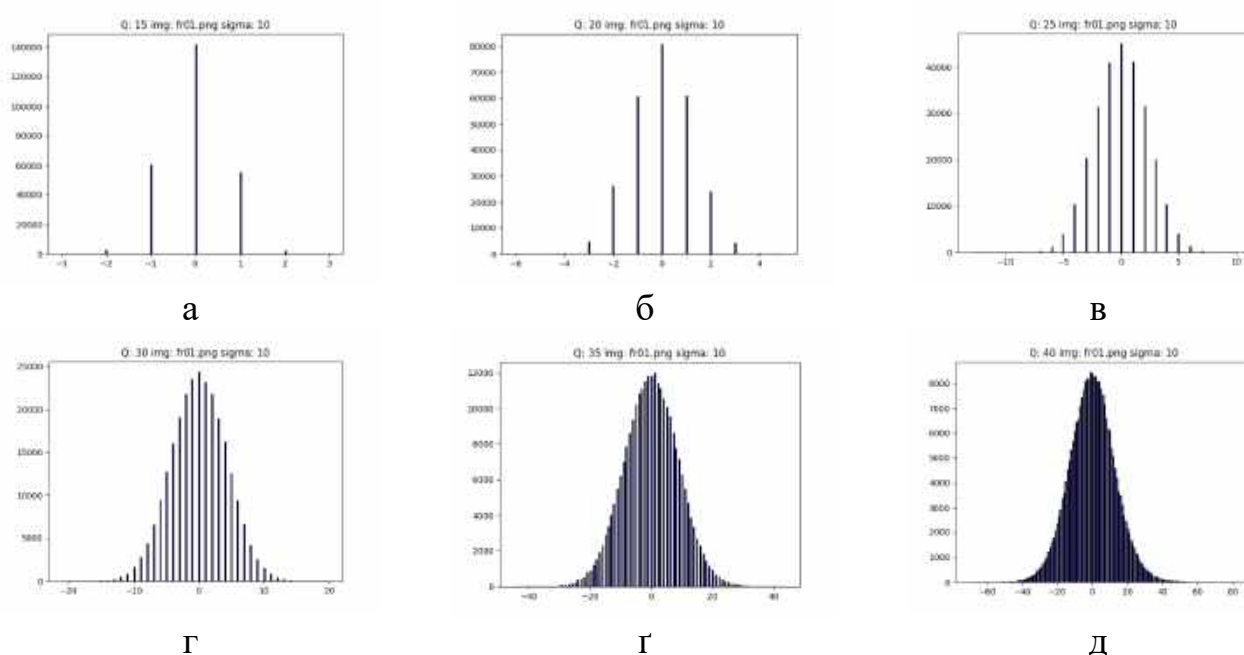


Рис 2.38 – Гістограма різниць (2.10) для зашумленого зображення fr01

Статистичні параметри (представлені в таблицях 2.11 і 2.12) підтверджують це твердження. Для зображення простої структури ексцес і асиметрія близькі до нуля для всіх значень Q. Зображення складної структури (таблиця 2.12) показує аналогічні результати, за винятком Q, що дорівнює 40.

Таблиця 2.11 – Статистичні характеристики спотворень для зашумленого зображення frisco

Q	Ексцес	Асиметрія	Середнє	Дисперсія
15	-0,1551	0,0260	-0,0276	0,5396
20	-0,1021	0,0022	-0,0162	1,7164
25	-0,0656	-0,0058	0,0023	5,6524
30	-0,0310	-0,0002	0,0008	19,8241
35	0,0331	0,0017	-0,0171	94,6806
40	0,1052	0,0188	-0,0204	120,9301

Таблиця 2.12 – Статистичні характеристики спотворень для зашумленого зображення fr01

Q	Ексцес	Асиметрія	Середнє	Дисперсія
15	-0,2062	0,0179	-0,0243	0,5137
20	-0,1956	-0,0134	-0,0225	1,5821
25	-0,1404	-0,0012	0,0007	5,1657
30	-0,0633	-0,0076	-0,0197	18,3098
35	0,0621	-0,0036	-0,0421	78,5642
40	0,5193	0,0802	-0,1026	169,9732

Виходячи з отриманих залежностей, стає зрозумілим, що присутність шуму на зображенні впливає на спотворення, що вносяться BPG кодером, у такий спосіб: отримані спотворення починають більше бути схожими на ті, що мають гаусовий розподіл, навіть для великих значень параметра контролю стиснення (порівняно з тим, коли шум не був присутній на зображенні, для Q, що перевищують 25, залежності дедалі менше були схожими на гаусові). Як результат, ця особливість свідчить про те, що під час стиснення зображень, спотворених шумом, наявність цього самого шуму необхідно враховувати.

2.5 Висновки до розділу

У цьому розділі за отриманими результатами можна зробити такі висновки:

- Зроблено аналіз ефективності стиснення за допомогою BPG напівтонових зображень, уражених адитивним шумом. Показано, що OPT може існувати для стандартних критеріїв оцінки якості зображення, таких як PSNR. Більше того, OPT може існувати для таких метрик візуальної якості, як MS-SSIM і PSNR-HVS-M, хоча це трапляється не так часто (у порівнянні з PSNR). Показано, що існування OPT залежить від дисперсії шуму та складності зображення, де OPT існує з більшою вірогідністю для зображень з більш простою структурою та/або більш інтенсивним шумом. Ще однією важливою особливістю є те, що зі збільшенням інтенсивності шуму, також збільшується і Q_{OPT} .

- Продемонстровано особливості стиснення на основі BPG у разі застосування стиснення до кольорових зображень з використанням різної дискретизації кольоровості. Показано можливе існування OPT за такими метриками якості зображення, як PSNR-NA і MDSI. Закономірності існування OPT аналогічні випадку з напівтоновими зображеннями: залежність від складності зображення та інтенсивності шуму.

- Проведено аналіз ефективності стиснення зображень, що містять пуасонівський шум, за допомогою BPG кодеру. Після проведених експериментів підтвердилося, що в разі ураження зображення сигнально-залежним шумом положення OPT може зміщуватися, ґрунтуючись на змісті (складності) зображення. Також порівняно роботу даного кодеру при використанні перетворення Анскомба та при прямому стисненні зображень, як результат в залежності від вимог варто використовувати тій чи інший спосіб.

- Показано, які спотворення вносить BPG кодер у разі стиснення зображення з наявністю на них шуму і без. Показано, що для безшумних зображень зі збільшенням Q внесені спотворення стають негаусовими. Це починає відбуватися за Q близько 30. За більшого Q розподіл спотворень стає негаусовим. Також показано, що наявність

шуму в стисненому зображенні впливає на спотворення, внесені розглянутим кодером. У проведених експериментах використовувалася модель шуму AWGN, внаслідок чого комбіновані спотворення здебільшого близькі до гаусових, особливо під час порівняння зашумленого і стисненого зображень.

РОЗДІЛ 3

РОЗРОБКА МЕТОДУ ПРОГНОЗУВАННЯ ОРТ І ПАРАМЕТРІВ У НІЙ ДЛЯ VP8 КОДЕРА

У цьому розділі розглядаються методи прогнозування оптимальної робочої точки під час роботи з VP8 кодером. Також проаналізовано фактори, що впливають на точність цього прогнозування. Окрім цього, метод прогнозування ОРТ був модифікований для роботи з кольоровими зображеннями, зважаючи на особливості, що виникають під час використання різних режимів кольоровості. До всього іншого, в даному розділі зосереджено увагу на практичних аспектах прийняття рішень, заснованих на даних прогнозування.

3.1 Прогнозування ОРТ для випадку стиснення напівтонових зображень

У попередньому розділі було розглянуто поняття оптимальної робочої точки, а також передумови до її появи. Ґрунтуючись на отриманих даних, можна зробити висновок про можливість прогнозування цієї точки.

Однак для можливості спостереження робочої точки необхідно мати як саме зашумлене зображення, так і його безшумну (істинну) версію, що найчастіше є неможливим (докладніше в другому розділі). У зв'язку з чим пропонується наступний спосіб вирішення цього завдання. Необхідно розрахувати нову метрику, яка називається $\Delta\text{Метр}$ і складається з різниці $\text{Метр}^{\text{ct}}(Q_{\text{орт}})$ і Метр^{n} . Для розрахунку першої частини (Метр^{ct}) необхідне значення $Q_{\text{орт}}$, яке може бути розраховане, маючи отриманий раніше вираз (2.3). Що стосується другої метрики, то для неї необхідне значення σ , яке апіорі відомо або попередньо оцінено з високою точністю. Маючи на руках дві ці метрики стає можливим розрахувати $\Delta\text{Метр}$.

Детальніше про сам підхід прогнозування буде розказано в наступних підрозділах. Забігаючи наперед, можна сказати, що спрогнозувавши метрику $\Delta\text{Метр}$, залежно від величини цієї метрики можливі такі результати: якщо прогнозоване $\Delta\text{Метр}$ позитивне, висновок полягає в тому, що ОРТ (згідно з розглянутою метрикою)

існує, і навпаки - якщо негативне, то оптимальна робоча точка відсутня для розглянутого зображення.

3.1.1 Опис запропонованого методу

Для початку варто детальніше розглянути особливості запропонованого методу, а також які початкові вимоги до нього висуваються. Щоб прогнозування було доцільним, до використання висуваються дві базові вимоги - це швидкість і точність. Очевидним критерієм за швидкістю є те, що розроблюваний алгоритм прогнозування має бути швидшим, ніж саме стиснення. Під точністю мається на увазі, що помилки прогнозування $\Delta\text{Метр}$ повинна мати досить малу середньоквадратичну помилку (RMSE).

Вище вже було згадано деякі деталі підходу до прогнозування. Загалом запропонований підхід ґрунтується на раніше запропонованому варіанті прогнозування для двох інших ДКП-кодерів [109]. У цьому підході припускають, що існує залежність між $\Delta\text{Метр}$ і деяким вхідним параметром (або ж кількома параметрами), який "характеризує" властивості зображення та шуму і може бути легко обчислений. Ця залежність розраховується заздалегідь і доступна до того моменту, коли необхідно спрогнозувати існування ОРТ для розглянутого зображення, що підлягає стисненню.

В якості вхідних параметрів у роботі [109] було запропоновано і випробувано кілька статистичних параметрів. Вони засновані на порівнянні амплітуд коефіцієнтів ДКП, обчислених у блоках 8×8 пікселів, з певними порогами, пов'язаними із середньоквадратичним відхиленням шуму. Простота отримання таких вхідних параметрів пов'язана з двома аспектами. По-перше, ДКП у блоках 8×8 пікселів є стандартною операцією в обробці зображень або відео [114] і може бути реалізована дуже швидко. По-друге, ДКП не потрібно проводити для всіх можливих позицій блоків даного зображення. Достатньо взяти 300...1000 випадково розташованих блоків, щоб оцінити вхідний параметр із достатньою точністю.

Більш детально необхідно розглянути два статистичні параметри, $P_{2\sigma}$ и $P_{2,7\sigma}$, які можна легко обчислити в наборі блоків 8x8 пікселів в області ДКП. Обидва параметри можливо використовувати для того, щоб характеризувати зображення, і їх буде використано для отримання залежності дельта метрики від вхідного параметра. Розглянемо випадок АБГШ.

Параметр $P_{2\sigma}$ розраховується як:

$$P_{2\sigma} = \sum_{m=1}^M \frac{P_{2\sigma}(m)}{M}, P_{2\sigma}(m) = ((\sum_{k=0}^7 \sum_{l=0}^7 \delta(k, l, m)) - 1)/63, \quad (3.1)$$

$$\delta(k, l, m) = 1, \text{ якщо } |D(k, l, m)| < 2\sigma$$

$$\text{інакше } 0.$$

де M - кількість блоків, що розглядаються, $D(k, l, m)$ - це ДКП-коефіцієнт у m -му блоці, $m=1, \dots, M$. Інакше кажучи, це оцінка ймовірності того, що абсолютні значення коефіцієнтів ДКП у блоках менші за 2σ , яка, як припускається, апіорно відома.

Аналогічно другий параметр $P_{2,7\sigma}$ розраховується як:

$$P_{2,7\sigma} = \sum_{m=1}^M \frac{P_{2,7\sigma}(m)}{M}, P_{2,7\sigma}(m) = ((\sum_{k=0}^7 \sum_{l=0}^7 \delta(k, l, m)) - 1)/63, \quad (3.2)$$

$$\delta(k, l, m) = 1, \text{ якщо } |D(k, l, m)| > 2,7\sigma$$

$$\text{інакше } 0.$$

Обидва статистичні параметри взято з теорії придушення шумів для фільтрів на основі ДКП. У роботі [115] було показано, що ці параметри здатні спільно характеризувати складність зображення та інтенсивність шуму. Наприклад, $P_{2\sigma}$ набуває малих значень для зображень зі складною структурою та/або низькою інтенсивністю шуму.

Наступне завдання - отримати залежності, що пов'язують метрики (насамперед $\Delta PSNR$ і $\Delta PSNR\text{-HVS-M}$) із вхідними параметрами ($P_{2\sigma}$ або $P_{2,7\sigma}$). Цей етап виконується заздалегідь, і основне завдання полягає в тому, щоб врахувати всі можливі ситуації, тобто широкий спектр властивостей зображення та інтенсивності шуму. У результаті необхідно побудувати діаграми розсіювання $\Delta \text{Метр}$ за вхідними параметрами. Кожен графік може бути отримано так: до тестового зображення, що

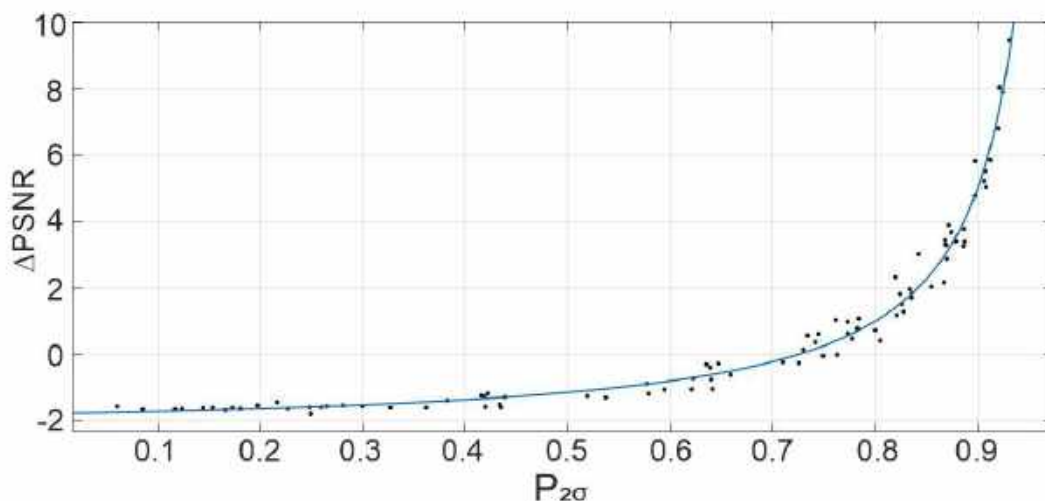
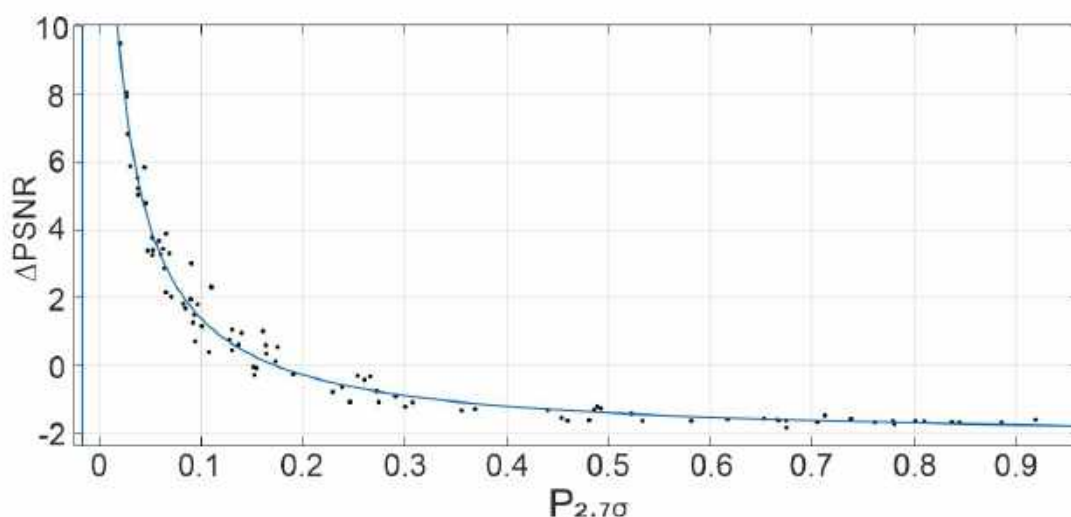
розглядається, додається шум із заданою дисперсією шуму, а потім воно стискається за допомогою Q_{OPT} (2.3), після чого необхідно визначити вхідні та вихідні параметри. Для того, щоб покрити більшу кількість всіляких варіацій, необхідно використовувати різні за складністю зображення і кілька значень інтенсивності властивого на зображенні шуму.

3.1.3 Деталі вписування

Для того, щоб мати можливість порівняти вхідні параметри зображення з отриманими заздалегідь значеннями, або, іншими словами, спрогнозувати $\Delta\text{Метр}$, необхідно вписати в отримані раніше діаграми розсіювання криві. Для того, щоб вписати криві в набір значень, скористаємося MATLAB Fitting Tool, що пропонує широкий вибір виразів для розрахунку вписаної кривої, включно з поліномами, експонентами або зваженою сумою експонент, раціональними функціями, рядами Фур'є, а також довільними (написаними самим користувачем) виразами.

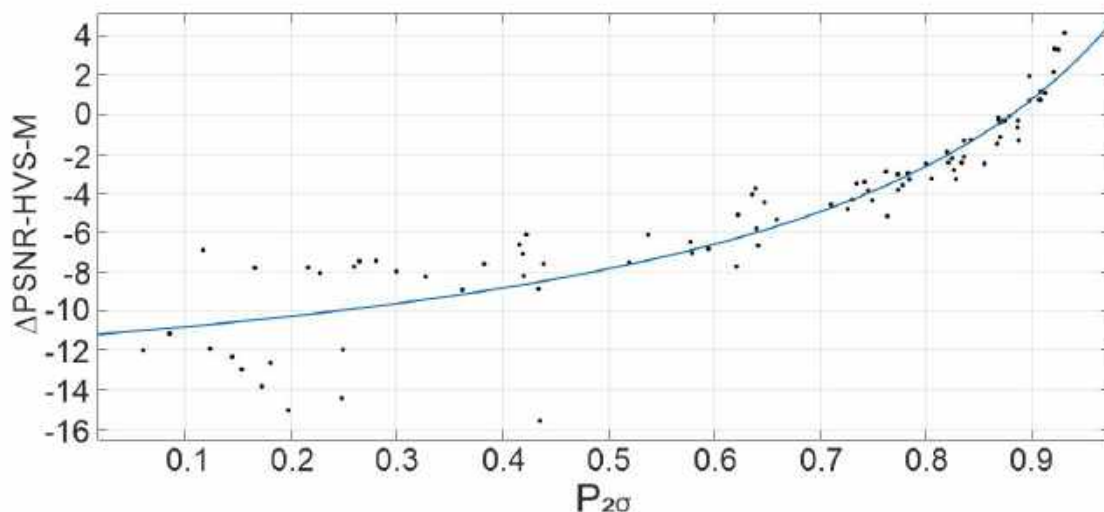
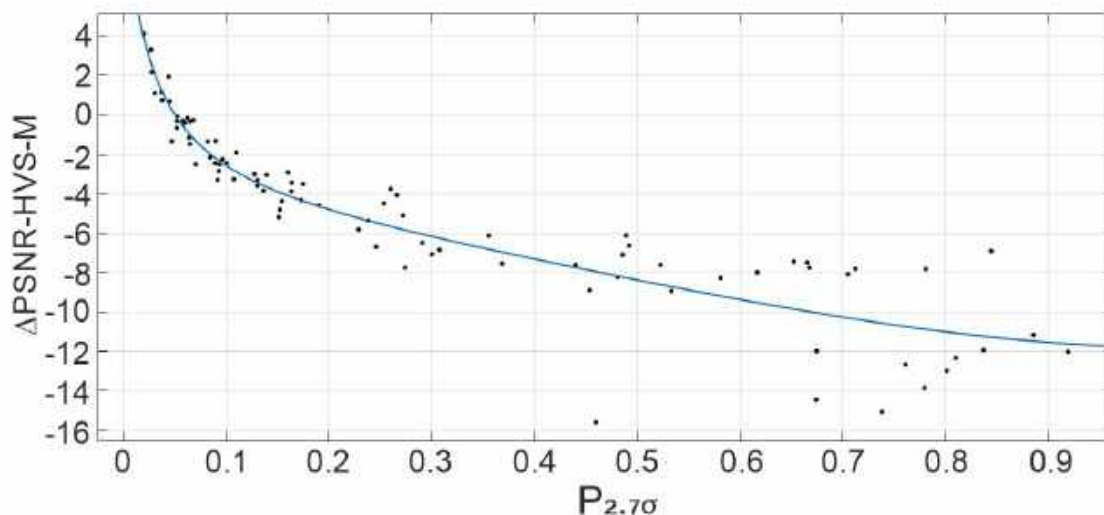
Маючи всю вищевикладену інформацію та засоби, стає можливим розрахувати вхідні параметри, отримати діаграми розсіювання і зробити вписування. Всього було використано одинадцять тестових зображень різної складності, дисперсія шуму варіювалася в межах від 0,25 до 400. Що стосується зображень, то використовується той самий набір напівтонових зображень, що й у другому розділі (рис. 2.2, 2.3). Як дельта метрики будуть використовуватися ΔPSNR , а також $\Delta\text{PSNR-HVS-M}$.

Спочатку розглянемо залежності між ΔPSNR і вхідними параметрами ($P_{2\sigma}$ або $P_{2,7\sigma}$), які наведено на рис. 3.1 и 3.2.

Рис 3.1 – Залежності між ΔPSNR та $P_{2\sigma}$ Рис 3.2 – Залежності між ΔPSNR та $P_{2,7\sigma}$

Точки діаграм розсіювання ΔPSNR до $P_{2\sigma}$ (рис. 3.1) і ΔPSNR до $P_{2,7\sigma}$ (рис. 3.2) розташовані компактно, і демонструють монотонне збільшення (рис. 3.1) або зменшення (рис. 3.2). Обидва вхідних параметра перебувають у межах від нуля до одиниці, для малих $P_{2\sigma}$ і великих $P_{2,7\sigma}$, ΔPSNR негативні й близькі до -2 дБ, у цьому разі ОПТ відсутній і не варто стискати зображення за допомогою $Q_{\text{ОПТ}}$ (2.3).

Для $\Delta\text{PSNR-HVS-M}$ метрики отримані діаграми представлено на рисунках 3.3 і 3.4.

Рис 3.3 – Залежності між $\Delta\text{PSNR-HVS-M}$ та $P_{2\sigma}$ Рис 3.4 – Залежності між $\Delta\text{PSNR-HVS-M}$ та $P_{2,7\sigma}$

Точки діаграм розсіювання $\Delta\text{PSNR-HVS-M}$ до $P_{2\sigma}$ (рис. 3.3) і $\Delta\text{PSNR-HVS-M}$ до $P_{2,7\sigma}$ (рис. 3.4) розташовані менш компактно (у порівнянні з даними на рис. 3.1 та 3.2); при цьому спостерігаються явні тенденції зростання $\Delta\text{PSNR-HVS-M}$ зі збільшенням $P_{2\sigma}$ і зменшенням метрики, коли $P_{2,7\sigma}$ зменшується; візуальний аналіз дає змогу припустити, що ці залежності монотонні; точки діаграми розсіювання розташовані компактніше для більших $P_{2\sigma}$ і менших $P_{2,7\sigma}$, що відповідає зображенням простішої структури та/або інтенсивнішому шуму.

Також варто зазначити, що для $\Delta\text{PSNR-HVS-M}$ і для ΔPSNR справедливим є те, що в обох випадках можливе досить точне вписування кривих, що дає змогу встановити приблизні аналітичні залежності між вихідними і вхідними параметрами. Крім цього, є точки, де ці криві перетинають нульовий рівень: $P_{2\sigma} \approx 0,73$ і $P_{2,7\sigma} \approx 0,16$ для ΔPSNR ; $P_{2\sigma} \approx 0,84$ і $P_{2,7\sigma} \approx 0,05$ для $\Delta\text{PSNR-HVS-M}$.

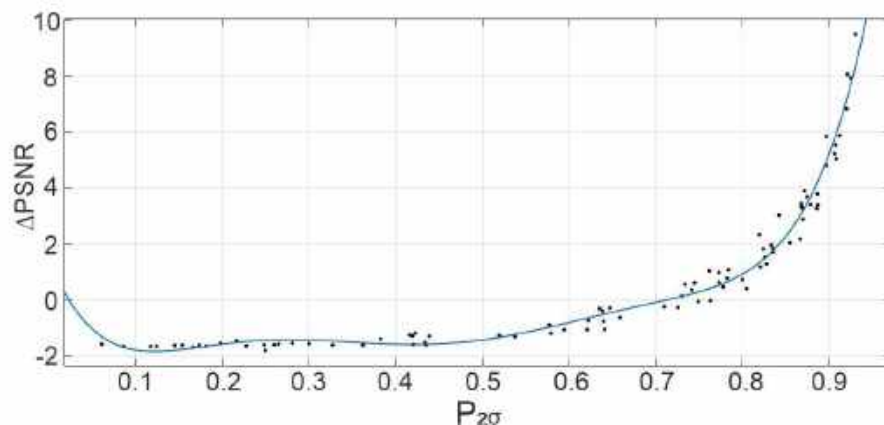
Рисунки 3.1 - 3.4 демонструють візуально високу точність вписування, однак, точність або якість вписування може бути охарактеризована кількома кількісними параметрами (критеріями), такими як R^2 , RMSE, Adjusted R^2 [116]. Для хорошого вписування значення параметрів R^2 і Adjusted R^2 мають бути якомога більшими (наближатися до одиниці), а RMSE - якомога меншими. Регресія вважається дуже хорошою, якщо R^2 і Adjusted R^2 перевищують 0,9. Точність вписування має велике значення, оскільки воно прямо впливає на точність прогнозування $\Delta\text{Метр}$.

Криві, вписані в діаграми розсіювання на рис. 3.1-3.4, були отримані за допомогою раціональних функцій. Для даних на рис. 3.1 RMSE становить близько 0,37, а R^2 і Adjusted R^2 - близько 0,98. Для діаграми розсіювання на рис. 3.2 RMSE дорівнює 1,72, а значення R^2 і Adjusted R^2 близькі до 0,85. Для даних на рис. 3.3 значення метрик наступні: R^2 і Adjusted R^2 - наближаються до 0,975, а RMSE дорівнює 0,43. Для діаграми розсіювання на рис. 3.4 RMSE становить близько 1,66, а R^2 і Adjusted R^2 - близько 0,86. З отриманих результатів можна зробити висновок, що вписування для ΔPSNR значно точніше, ніж для $\Delta\text{PSNR-HVS-M}$ (це пов'язано не з поганим методом вписування, а з більшим розкидом даних для $\Delta\text{PSNR-HVS-M}$). Краще використовувати $P_{2\sigma}$ для прогнозування ΔPSNR і $P_{2,7\sigma}$ для прогнозування $\Delta\text{PSNR-HVS-M}$. Вирази, а також параметри для залежностей, продемонстрованих на рисунках 3.1-3.4, наведено в таблиці 3.1.

Як уже було згадано, раціональні функції не єдиний варіант, як можна вписати криву в набір даних. У зв'язку з цим необхідно перевірити, чи здатні інші апроксимації продемонструвати аналогічні або, за можливістю, дати кращі результати в плані точності. Прикладом використання інших методів для вписування є моделі Фур'є, представленого на рисунку 3.6.

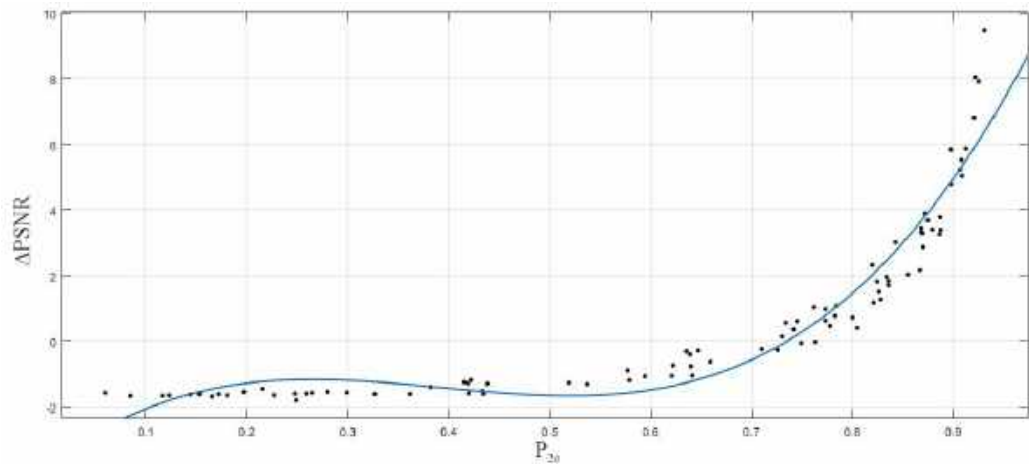
Таблиця 3.1 – Параметри вписаних кривих.

Залежність	Вираз	Параметри
ΔPSNR до $P_{2\sigma}$	$f(x) = (p_1 * x + p_2) / (x^3 + q_1 * x^2 + q_2 x + q_3)$	$p_1=1,533 \times 10^4$, $p_2=-1,112 \times 10^4$, $q_1=75,71$, $q_2=-6291$, $q_3=6139$
ΔPSNR до $P_{2,7\sigma}$	$f(x) = (p_1 x + p_2) / (x^2 + q_1 * x + q_2)$	$p_1=-6,162 \times 10^5$, $p_2=1,077 \times 10^5$, $q_1=2,843 \times 10^5$, $q_2=4501$
$\Delta\text{PSNR-HVS-M}$ до $P_{2\sigma}$	$f(x) = (p_1 x + p_2) / (x^2 + q_1 * x + q_2)$	$p_1=2,895 \times 10^5$, $p_2=-2,549 \times 10^5$, $q_1=-1,72 \times 10^4$, $q_2=2,263 \times 10^4$
$\Delta\text{PSNR-HVS-M}$ до $P_{2,7\sigma}$	$f(x) = (p_1 * x + p_2) / (x^3 + q_1 * x^2 + q_2 x + q_3)$	$p_1=-10,97$, $p_2=0,558$, $q_1=-1,99$, $q_2=1,82$, $q_3=0,048$

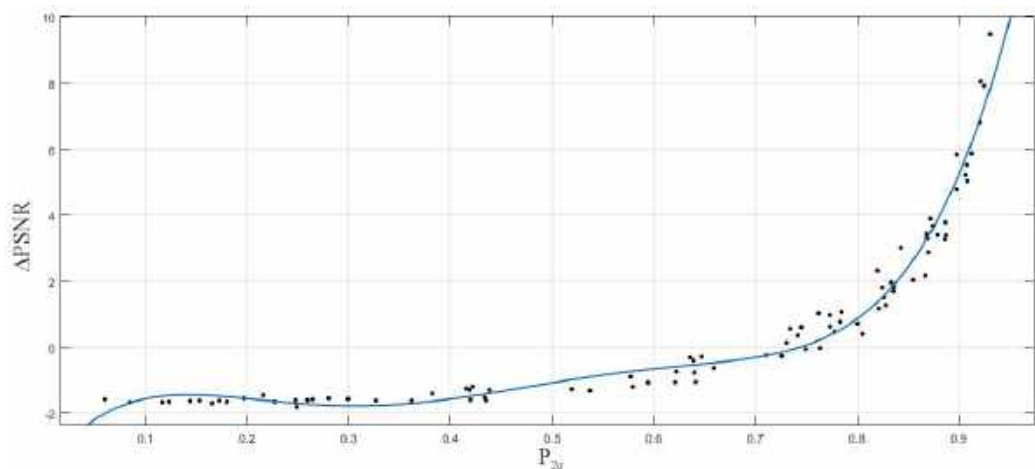
Рис 3.6 – Діаграма розсіювання ΔPSNR до $P_{2\sigma}$ і вписана крива моделі Фур'є.

З рисунка видно, що модель Фур'є 3, здатна забезпечити добру точність вписування в плані кількісних критеріїв (RMSE близько 0,42, R^2 і Adjusted R^2 близько 0,976), але отримані криві виявилися не монотонними. Те саме стосується і

поліноміальних апроксимацій. Для поліномів порядку 2 і 3 параметри точності вписування гірші, ніж зазначено вище, тоді як для поліномів порядку 4 і 5 криві стають не монотонними. Приклади отриманих залежностей для поліномів третього і п'ятого порядків наведено на рисунку 3.7.



а



б

Рис 3.7 – Приклади вписування поліномів ΔPSNR проти $P_{2\sigma}$ для поліномів третього (а) і п'ятого (б) порядків

Для даних на рис. 3.7 (а) отримано $R^2=0,93$, $\text{RMSE}=0,71$, $\text{Adjusted } R^2=0,93$, тобто результати доволі хороші. Для полінома п'ятого порядку (рис. 3.7, б) результати вписування ще кращі ($R^2=0,97$, $\text{RMSE}=0,47$, $\text{Adjusted } R^2=0,97$). При цьому підігнані криві не є монотонними, спостерігаються локальні максимуми.

Нарешті, хороші апроксимації можна отримати для зважених сум двох експонент. Навіть з урахуванням того, що для $\Delta\text{PSNR-HVS-M}$ точки на діаграмі розсіювання розташовані не так компактно (в порівнянні з ΔPSNR), що погіршує якість підсумкового вписування, проте все одно показує високі показники точності RMSE приблизно 1,64, R^2 і Adjusted R^2 приблизно 0,87. Сама вписана крива продемонстрована на рисунку 3.8

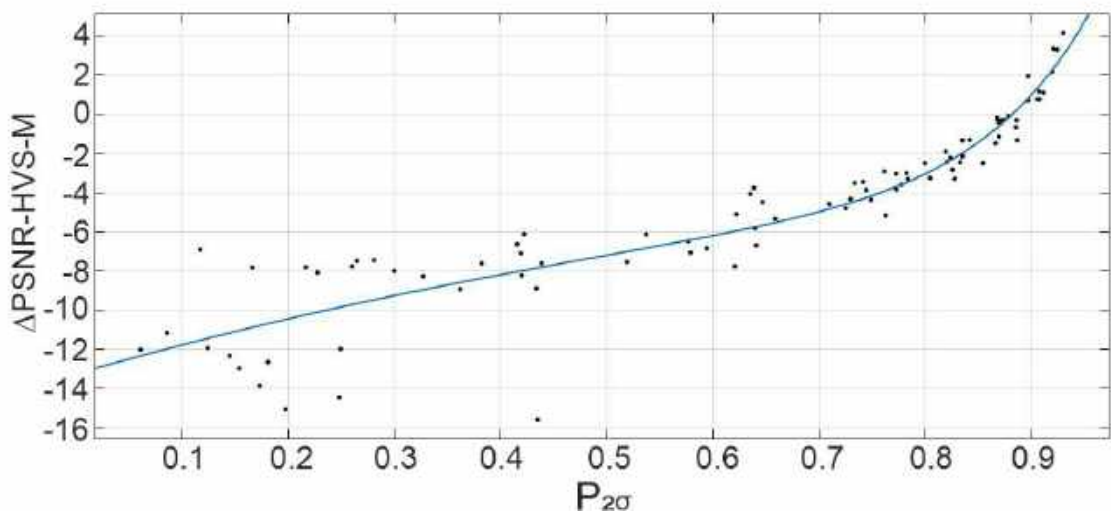


Рис 3.8 – Діаграма розсіювання $\Delta\text{PSNR-HVS-M}$ до $P_{2\sigma}$ і вписана крива, що отримана через зважену суму двох експонент.

Маючи під рукою вписані криві, можна спрогнозувати ΔPSNR і $\Delta\text{PSNR-HVS-M}$ для будь-якого зображення, що підлягає стисненню. Для цього необхідно розрахувати вхідний параметр і підставити його в апроксимуючу криву, визначену в таблиці 3.1. Наприклад, оціночне значення $P_{2\sigma}$ дорівнює 0,6. Тоді, згідно з даними на рис. 3.1 (або рис. 3.6), прогнозований ΔPSNR становить близько -0,8 дБ. Що стосується іншої метрики, то згідно з даними на рис. 3.2 і рис. 3.8 прогнозований $\Delta\text{PSNR-HVS-M}$ близько -6,5 дБ і близько -6,1 дБ відповідно. У будь-якому разі можна зробити висновок, що ОПТ для цього зображення відсутнє і стискати його за допомогою $Q_{\text{ОПТ}}$ не варто.

3.1.4 Фактори, що впливають на точність

Спираючись на отримані результати вище, можна зробити висновок, що вписування і, як результат, прогнозування для $\Delta\text{PSNR-HVS-M}$ гірше, ніж для ΔPSNR . Можна стверджувати, що прогнозування не є точним для $\Delta\text{PSNR-HVS-M}$, коли $P_{2\sigma}$ становить близько 0,2. Дійсно, у цьому випадку $\Delta\text{PSNR-HVS-M}$ може становити від -16 дБ до -7 дБ залежно від зображення, тоді як спрогнозовані значення становлять близько -10,5 дБ (рис. 3.8). Така точність прогнозування здається невисокою, оскільки помилки можуть сягати 5,5 дБ. Однак на практиці це не є проблемою з кількох причин. По-перше, настільки малі $P_{2\sigma}$ зазвичай відповідають зображенням, зіпсованим шумом малої інтенсивності (який є непомітним). Тому жодного позитивного ефекту від фільтрації шуму внаслідок стиснення зображення з втратами не спостерігатиметься. По-друге, такі значення $\Delta\text{PSNR-HVS-M}$ зазвичай відповідають PSNR-HVS-M^n близько 55...60 дБ. Тоді після стиснення виходить PSNR-HVS-M^c близько 45...50 дБ, тобто внесені спотворення непомітні.

Варто також розглянути можливість впливу реалізації шуму на точність прогнозування. Очевидно, що для заданого тестового зображення і дисперсії AWGN показники ΔPSNR , $\Delta\text{PSNR-HVS-M}$ і навіть Q_{OPT} можуть змінюватися залежно від реалізації шуму. Для вивчення цього аспекту було проаналізовано три зображення різної складності, що зазнали впливу AWGN з трьома різними значеннями. Були виміряні значення ΔPSNR і $\Delta\text{PSNR-HVS-M}$ для Q_{OPT} і обчислені дисперсії цих параметрів. Встановлено, що дисперсії цих параметрів становлять близько 0,001; навіть у найгіршому випадку максимальна дисперсія дорівнювала 0,0033 для ΔPSNR і 0,0046 для $\Delta\text{PSNR-HVS-M}$. Це означає, що впливом реалізації можна знехтувати, тут варто зазначити, що RMSE вписування становить близько 0,4 для ΔPSNR і близько 1,2 для $\Delta\text{PSNR-HVS-M}$ в області, яка нас цікавить (область OPT). Якщо Q_{OPT} відрізнявся для різних реалізацій, то відмінності дорівнюють 1 (наприклад, OPT спостерігався для Q , що дорівнював або 33, або 34), а різниця в значеннях ΔPSNR або $\Delta\text{PSNR-HVS-M}$ у сусідніх точках незначна. Таким чином, на практиці впливом реалізації шуму на параметри OPT можна знехтувати.

3.1.5 Прийняття рішень та інші практичні випадки

Розглянувши алгоритм прогнозування оптимальної робочої точки і його особливості, можна приступити до обговорення алгоритму дій, ґрунтуючись на результатах прогнозування. Раніше вже частково згадувалося, що OPT присутній, якщо $\Delta\text{Метр}$ позитивна. Виходячи з цього можна скласти наступний алгоритм дій у разі використання ΔPSNR і $\Delta\text{PSNR-HVS-M}$. Залежно від суми даних метрик можливо кілька варіантів вибору параметра для стиснення. Якщо $\Delta\text{PSNR} + \Delta\text{PSNR-HVS-M} > 1$ дБ, то можна вважати, що OPT існує, і стиснення зображення необхідно проводити за допомогою Q_{OPT} . Якщо ж значення метрик можна віднести до інтервалу $-1 \text{ дБ} \leq \Delta\text{PSNR} + \Delta\text{PSNR-HVS-M} \leq 1 \text{ дБ}$, у цьому разі OPT може існувати, і треба використовувати $Q_{\text{OPT}-1}$ (але не менше 28), щоб уникнути надмірного згладжування. Ще один випадок, коли $\Delta\text{PSNR} + \Delta\text{PSNR-HVS-M} \leq -1 \text{ дБ}$, у цьому разі OPT найімовірніше відсутній і необхідно використовувати $Q=28$, щоб забезпечити невидимість внесених спотворень і можливість видалення шуму в ньому шляхом застосування постфільтрації декомпресованих зображень [114].

Якщо ситуація 1 помилково розпізнається як ситуація 2 або навпаки, практично нічого не відбувається, оскільки рекомендовані значення Q відрізняються тільки на одиницю. Якщо ситуація 2 помилково розпізнається як ситуація 3, це не страшно, оскільки відбувається "акуратне" стиснення (за рахунок меншого Q). Якщо ситуація 3 помилково розпізнається як ситуація 2, то спостерігається трохи більше розмиття стисненого зображення (проте КС більший). Нарешті, ситуації 1 і 3 практично ніколи не помиляються в класифікації.

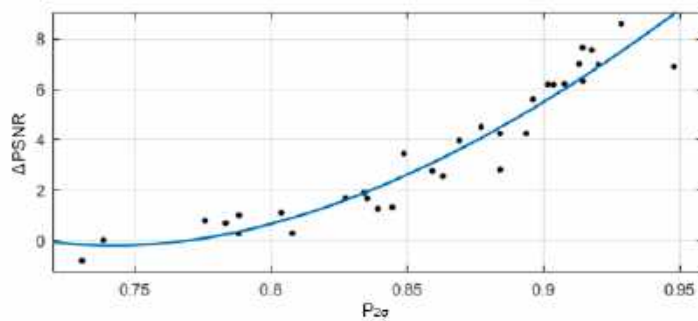
Як результат процедура стиснення з використанням прогнозування може мати такий вигляд. Для початку необхідно оцінити дисперсію шуму за допомогою будь-якого сліпого методу оцінювання дисперсії шуму (якщо дисперсія шуму не відома заздалегідь), далі необхідно розрахувати Q_{OPT} і $P_{2\sigma}$, після чого вже можна розрахувати ΔPSNR і $\Delta\text{PSNR-HVS-M}$, і, зважаючи на результати, здійснити стиснення з обраним Q .

Варто зазначити, що цей метод прогнозування також може бути модифіковано для роботи з іншими типами шуму. Прикладом складнішої моделі шуму може слугувати представник сигнально-залежного сімейства шумів - Пуассонівський шум.

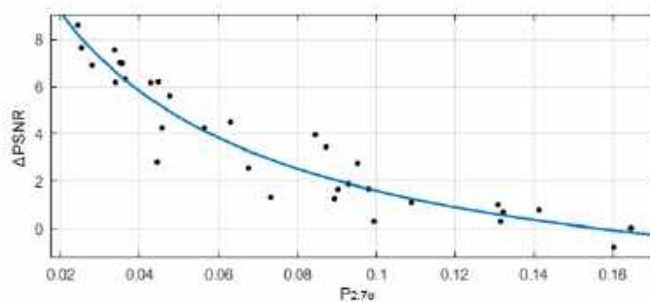
Основною зміною (модифікацією) є використання (2.5) замість (2.3) для визначення Q_{OPT} , щодо вхідних параметрів, то можуть бути використані ті самі вхідні параметри ($P_{2\sigma}$ і $P_{2,7\sigma}$), що й раніше, однак метод їх розрахунку зазнав змін. Оскільки ми маємо справу з пуассонівським шумом, дисперсію шуму можна вважати рівною середньому значенню зображення, тоді σ дорівнює квадратному кореню з цього середнього значення.

Щоб перевірити можливість такої модифікації, було проведено аналогічний експеримент, у якому було використано 32 зображення (включно з розділеними версіями - докладніше про них ішлося в другому розділі). Кожна точка цього графіка відповідає одному зображенню з розрахованими вхідними параметрами за формулами (3.1) і (3.2) з використанням середніх значень цього зображення і дельта-метрики для Q_{OPT} , розрахованої за (2.5). Для вписування було використано раціональні функції, як і для випадку ураження Гаусовим шумом. Приклади отриманих графіків подано на рисунку 3.9, 3.10.

Як видно з рис. 3.9, судячи із загальної компактності діаграми розсіювання, вписування виконано добре. Це підтверджують і метрики оцінки якості вписування - середньоквадратичне значення (RMSE) і R^2 . Варто нагадати, що RMSE має бути якомога меншим, а R^2 має наближатися до одиниці. Для випадку ΔPSNR до $P_{2\sigma}$ (рис. 3.9, а) RMSE дорівнює 0,8004, а R^2 - 0,9222. З іншого боку, для $P_{2,7\sigma}$ (рис. 3.9, б) точки на діаграмі розсіювання є дещо менш компактними, внаслідок чого RMSE дорівнює 0,9254, а R^2 - 0,896. Ситуація для $\Delta\text{PSNR-HVS-M}$ (рис. 3.10) не настільки сприятлива, але все ж добре вписується з використанням раціональних функцій.

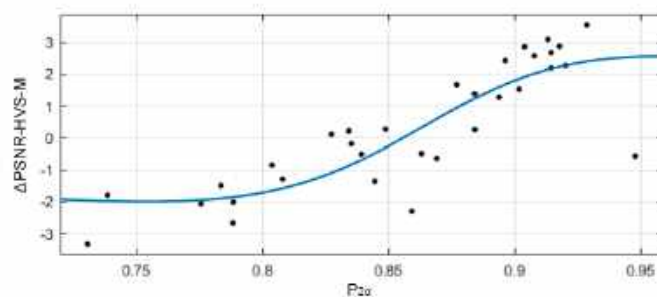


а

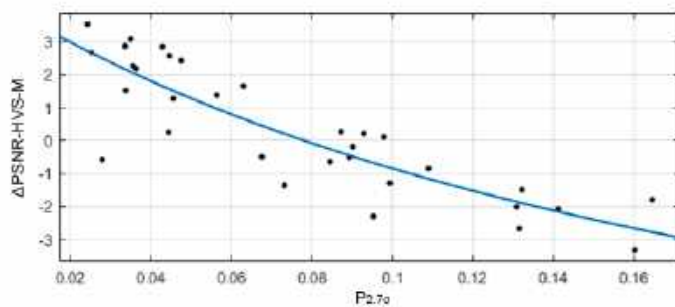


б

Рис 3.9 – Діаграма розсіювання ΔPSNR до $P_{2\sigma}$ (а) і ΔPSNR до $P_{2,7\sigma}$ (б) і вписана крива



а



б

Рис 3.10 – Діаграма розсіювання $\Delta\text{PSNR-HVS-M}$ до $P_{2\sigma}$ (а) і $\Delta\text{PSNR-HVS-M}$ до $P_{2,7\sigma}$ (б) та вписана крива

Вписування для метрики $\Delta\text{PSNR-HVS-M}$ до $P_{2\sigma}$ має дещо інакший вигляд і не такий монотонний, як для решти отриманих результатів. Це пов'язано з тим, що є точки, які сильно відхиляється від інших. Незважаючи на це, загальна точність вписування описується значеннями 0,7268 для R^2 і 1,068 для RMSE для $\Delta\text{PSNR-HVS-M}$ до $P_{2\sigma}$ (рис. 3.10, а) і 0,7509 для R^2 і 1,02 для RMSE для $\Delta\text{PSNR-HVS-M}$ до $P_{2,7\sigma}$ (рис. 3.10, б). Вирази та параметри для всіх отриманих вище кривих подано в таблиці 3.2.

Таблиця 3.2 – Параметри вписаних кривих.

Залежність	Вираз	Параметри
ΔPSNR до $P_{2\sigma}$	$f(x) = (p1 \cdot x^2 + p2 \cdot x + p3) / (x + q1)$	$p1 = 209,2$ $p2 = -310,7$ $p3 = 115,2$ $q1 = 0,01813$
$\Delta\text{PSNR-HVS-M}$ до $P_{2\sigma}$	$f(x) = (p1 \cdot x + p2) / (x^2 + q1 \cdot x + q2)$	$p1 = 0,4566$ $p2 = -0,3908$ $q1 = -1,737$ $q2 = 0,7645$
ΔPSNR до $P_{2,7\sigma}$	$f(x) = (p1 \cdot x + p2) / (x^2 + q1 \cdot x + q2)$	$p1 = -7,974 \cdot 10^4$ $p2 = 1,246 \cdot 10^4$ $q1 = 2,023 \cdot 10^4$ $q2 = 787$
$\Delta\text{PSNR-HVS-M}$ до $P_{2,7\sigma}$	$(x) = (p1 \cdot x + p2) / (x^2 + q1 \cdot x + q2)$	$p1 = -7,436 \cdot 10^5$ $p2 = 5,818 \cdot 10^4$ $q1 = 6,024 \cdot 10^4$ $q2 = 1,322 \cdot 10^4$

Отримані вирази дають змогу зробити прогнозування ΔPSNR і $\Delta\text{PSNR-HVS-M}$. Наприклад, для тестового зображення під назвою pointloma розраховані значення цих метрик дорівнюють 6,08 і 3,13 відповідно. Спрогнозовані значення для цього зображення для ΔPSNR з використанням $P_{2\sigma}$ і $P_{2,7\sigma}$ дорівнюють 6,76 і 7,60, а для метрики $\Delta\text{PSNR-HVS-M}$ вони дорівнюють 2,26 і 2,52. Отримані результати свідчать про досить хорошу точність прогнозування. Результати для інших прикладів зображень наведено в таблиці 3.3.

Таблиця 3.3 – Прогнозовані та розраховані значення метрики.

Назва зображення та вхідний параметр		Розраховані значення		Прогнозовані значення	
		ΔPSNR	$\Delta\text{PSNR-HVS-M}$	ΔPSNR	$\Delta\text{PSNR-HVS-M}$
Aerial	P_2	1,9	0,43	1,499	-1,161
	$P_{2,7}$			1,641	-0,801
Airplane	P_2	8,75	6,03	7,72	2,439
	$P_{2,7}$			9,11	3
Aerial-DIV2	P_2	3,74	0,06	3,293	0,346
	$P_{2,7}$			3,253	0,439
Airplane-DIV2	P_2	11	7,37	8,303	2,503
	$P_{2,7}$			10,266	3,313

Як результат, запропонований метод може бути використаний для визначення оптимальної робочої точки не тільки у випадках ураження Гауссовим шумом, а й у випадку Пуасонівського шуму.

3.2 Прогнозування OPT для випадку стиснення кольорових зображень

Вище було продемонстровано можливість прогнозування існування оптимальної робочої точки під час стиснення BPG кодером для напівтонових зображень, уражених різними видами шуму. Однак виходячи з даних, отриманих у другому розділі, що говорять про можливість спостереження OPT для кольорових зображень, варто розглянути можливість застосування згаданого раніше алгоритму прогнозування для випадку кольорових зображень. Необхідно провести аналіз за аналогією з попереднім підходом, і залежно від результатів дати рекомендації.

Варто зазначити, що почати необхідно з розгляду особливостей і відмінностей, які необхідно буде ввести до оригінального методу прогнозування, щоб мати можливість працювати з кольоровими зображеннями.

3.2.1 Відмінність від методу для напівтонових зображень

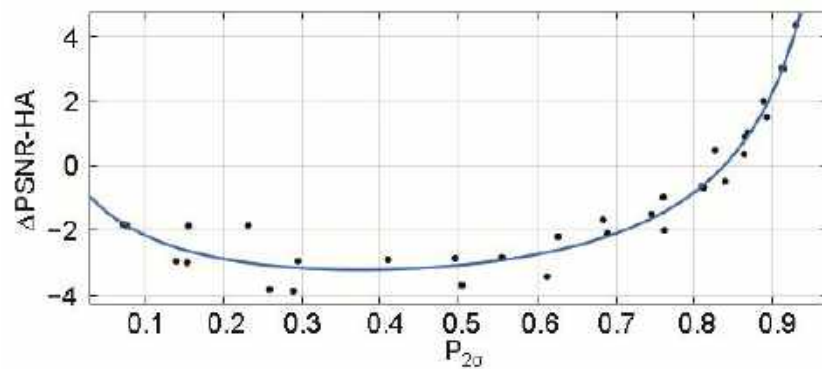
Першою відмінністю є те, що для розрахунку $Q_{\text{ОРТ}}$ буде використовуватися вираз (2.4) замість (2.3), це пов'язано з тим, що робоча точка для кольорових зображень зміщена порівняно з напівтоновими зображеннями (детальніше про це згадувалося в другому розділі). Друга відмінність полягає в тому, що вхідні параметри $P_{2\sigma}$ і $P_{2,7\sigma}$ розраховуватимуться дещо інакше. Оскільки формули (3.1) і (3.2) розраховують для одно-канальних (напівтонових) зображень, то у випадку з кольоровими зображеннями вхідні параметри будуть рахуватися для кожного каналу (червоний, зелений і блакитний), а після цього отримані значення необхідно усереднити.

Також варто зазначити, що в разі прогнозування оптимальної робочої точки для кольорових зображень буде використовуватися інший набір метрик якості. Для цього завдання обрано метрики PSNR-НА і MDSI, про які були наведена інформація в першому і другому розділі. Передбачається, що $\Delta\text{PSNR-НА}$ або ΔMDSI достатньо пов'язані з вхідними параметрами $P_{2\sigma}$ або $P_{2,7\sigma}$. Якщо таке припущення справедливе, то для заданого зашумленого зображення (з відомим стандартним відхиленням шуму σ), що підлягає стисненню, обчислюють $P_{2\sigma}$ або $P_{2,7\sigma}$ і, використовуючи апріорно відому залежність (функцію), прогнозують $\Delta\text{PSNR-НА}$ або ΔMDSI за аналогією, як це було зроблено для інших метрик у випадку з напів-тоновими зображеннями.

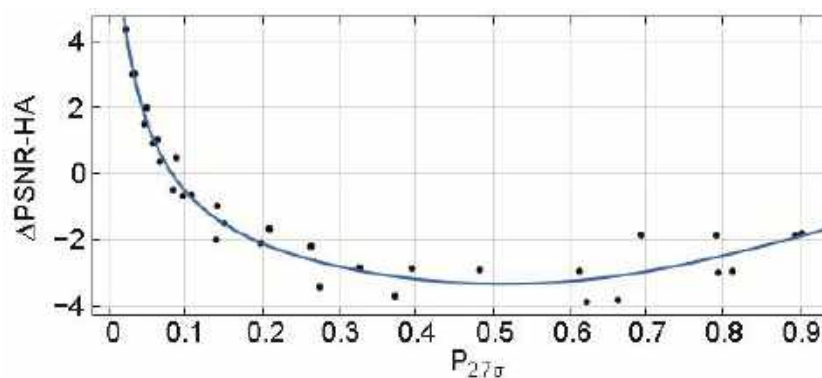
В усьому іншому метод прогнозування залишається аналогічним попередньому варіанту. Графіки розсіювання отримують один раз і заздалегідь (мається на увазі, що цю роботу виконують в автономному режимі, а залежність між вихідними і вхідними параметрами доступна в той момент, коли необхідно виконати прогнозування для нового зображення). Кожна точка діаграми розсіювання відповідає одному тестовому зображенню, спотвореному AWGN із заданим стандартним відхиленням шуму. Потім оцінюється вхідний параметр і зображення стискається за допомогою $Q_{\text{ОРТ}}$, визначеного в (2.4). Після декомпресії обчислюється $\Delta\text{Метр}$. На всіх діаграмах розсіювання вертикальна вісь відповідає вихідному параметру, а горизонтальна - вхідному. Для побудови діаграм розсіювання для випадку з кольоровими зображеннями було використано тестові зображення різної складності (набір, що використовувався в другому розділі), дисперсія шуму варіювалася в межах від 0,25 до 400.

3.2.2 Прогнозування існування ОРТ і параметрів у ній

У першому розділі згадувалося, що BPG кодер здатний працювати з декількома режимами кольоровості, у свою чергу в другому розділі було встановлено, що оптимальна робоча точка може спостерігатися для всіх режимів кольоровості. Ґрунтуючись на цьому необхідно перевірити можливість прогнозування для кожного режиму. Почнемо зі встановлення залежності між $\Delta\text{PSNR-НА}$ і $P_{2\sigma}$ або $P_{2,7\sigma}$ для формату 4:4:4. Для побудови діаграм використовувався той самий набір зображень, що й у другому розділі (рис. 2.13, 2.14). Отримані діаграми розсіювання і вписані в них криві представлені на рисунку 3.11.



а



б

Рис 3.11 – Діаграми розсіювання $\Delta\text{PSNR-НА}$ та $P_{2\sigma}$ (а) і $\Delta\text{PSNR-НА}$ та $P_{2,7\sigma}$ (б) та вписані криві (формат 4:4:4)

Варто нагадати, що малі значення $P_{2\sigma}$ і великі $P_{2,7\sigma}$ відповідають зображенням зі складною структурою або малим значенням дисперсії. PSNR-НА виражається в дБ, і його зміну на 0,5-1 дБ можна помітити під час візуального огляду. Повертаючись до отриманих залежностей, точки на діаграмі розсіювання розташовуються компактно, що дає змогу реалізувати вписування кривої з високою точністю. Після цього стає можливим досить точно прогнозування для обох вхідних параметрів.

З отриманих результатів можна виділити декілька спостережень: бувають ситуації, коли значення $\Delta\text{PSNR-НА}$ негативні та становлять близько -3 дБ, це відбувається для зображень зі складною структурою, спотворених AWGN з дисперсією менше 10. Аналогічна ситуація і для тестових зображень Diego та Diego 2 при $\sigma^2=25$. Існують також зображення і значення дисперсії шуму, коли $\Delta\text{PSNR-НА}$ негативний і його значення становить близько -1 дБ. Проте також існують зображення і значення дисперсії шуму, для яких $\Delta\text{PSNR-НА}$ позитивна. З великою ймовірністю це відбувається, якщо $P_{2\sigma}>0,84$ (рис. 3.11, а) або $P_{2,7\sigma}<0,08$ (рис. 3.11, б). Позитивні значення $\Delta\text{PSNR-НА}$ мають місце здебільшого для зображень із простою структурою, уражених досить інтенсивним шумом. У цьому разі доцільним є стиснення в OPT.

Для другої метрики (MDSI) результати аналогічні. Графіки розсіювання та вписані криві наведено на рис. 3.12. Варто зазначити, що MDSI змінюється в межах від 0 до 0,6, де можна помітити різницю приблизно в 0,02. Знову ж таки, точки на діаграмі розсіювання розташовані компактно, що дає змогу отримати хороші значення точності вписування кривої.

Висновки, які можна зробити з аналізу даних на рис. 3.12, такі. Існують позитивні значення ΔMDSI , які відповідають $P_{2\sigma}<0,85$ (рис. 3.12, а) або $P_{2,7\sigma}>0,08$ (рис. 3.12, б). Тоді стиснення потрібно проводити "обережно", тобто використовувати $Q<Q_{\text{OPT}}$, наприклад $Q_{\text{OPT}}=25$. У інших ситуаціях, тобто для $P_{2\sigma}\geq 0,85$ або $P_{2,7\sigma}\leq 0,08$, OPT спостерігаються з високою ймовірністю та варто стискати відповідні зображення в OPT. Що стосується самого вписування, то так само, як і у випадку $\Delta\text{PSNR-НА}$, внаслідок компактності розташованих на графіку точок, вписування зроблено досить-таки точно. Про це також свідчать значення метрик, що характеризують точність. Для вхідного параметра $P_{2\sigma}$ значення метрик дорівнюють $R^2 = 0,9154$, Adjusted $R^2 = 0,8991$

і RMSE дорівнює 0,007974. Для другого параметру ($P_{2,7\sigma}$) також характерна висока точність вписування 0,9036, 0,885 і 0,008512 для R^2 , Adjusted R^2 і RMSE відповідно.

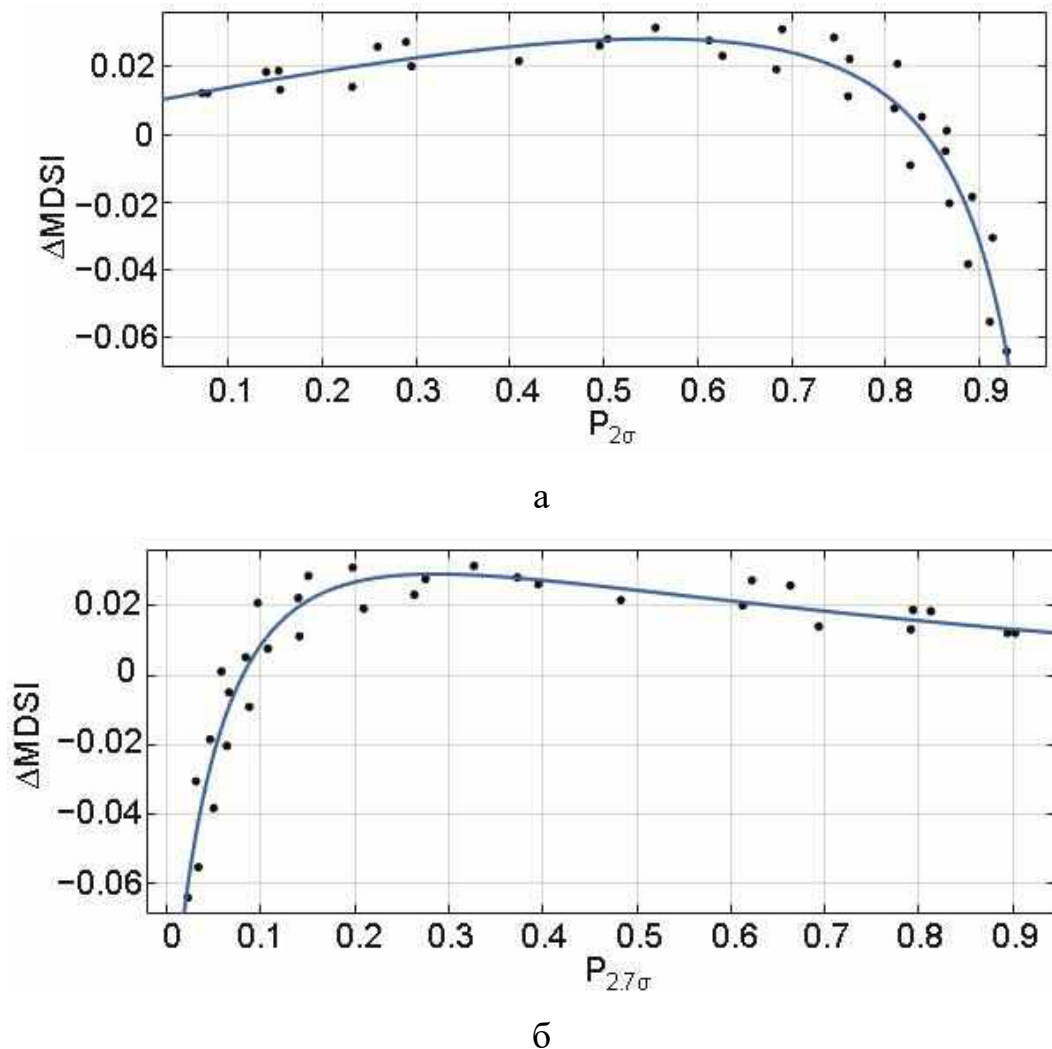


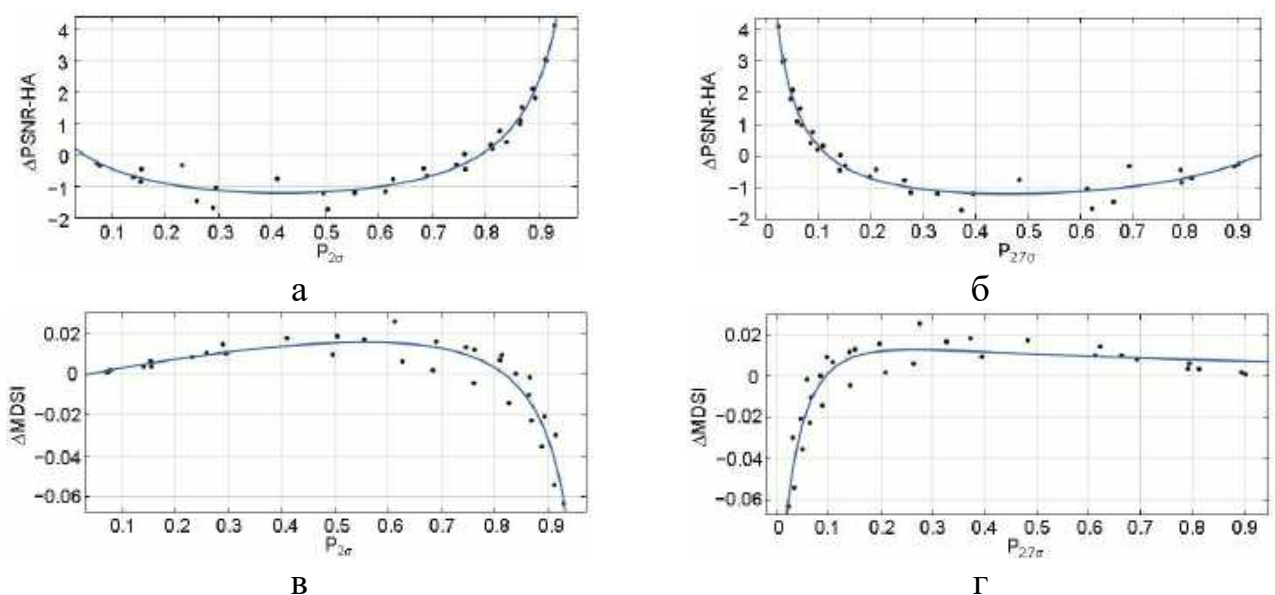
Рис 3.12 – Діаграми розсіювання $\Delta MDSI$ до $P_{2\sigma}$ (а) і до $P_{2,7\sigma}$ (б), а також вписані криві (формат 4:4:4)

Вирази, а також параметри для отриманих вписаних кривих для двох метрик наведено в таблиці 3.4.

Для режимів субдискретизації кольоровості (4:2:2 і 4:2:0) було проведено аналогічні процедури. Результати представлено на рисунках 3.13 і 3.14. Для початку розглянемо режим роботи 4:2:2 (рис. 3.13).

Таблиця 3.4 – Параметри вписаних кривих.

Залежність	Вираз	Параметри
$\Delta\text{PSNR-HA}$ до $P_{2\sigma}$ (4:4:4)	$f(x) = (p_1 \cdot x^2 + p_2 \cdot x + p_3) / (x^3 + q_1 \cdot x^2 + q_2 \cdot x + q_3)$	$p_1 = 1,195 \cdot 10^5$ $p_2 = -1,003 \cdot 10^5$ $p_3 = 147,4$ $q_1 = -1,92 \cdot 10^4$ $q_2 = 1,778 \cdot 10^4$ $q_3 = 2454$
$\Delta\text{PSNR-HA}$ до $P_{2,7\sigma}$ (4:4:4)	$f(x) = (p_1 \cdot x^2 + p_2 \cdot x + p_3) / (x^3 + q_1 \cdot x^2 + q_2 \cdot x + q_3)$	$p_1 = 3,114$ $p_2 = -4,159$ $p_3 = 0,3203$ $q_1 = -1,482$ $q_2 = 1,015$ $q_3 = 0,03138$
ΔMDSI до $P_{2\sigma}$ (4:4:4)	$f(x) = (p_1 \cdot x^2 + p_2 \cdot x + p_3) / (x^3 + q_1 \cdot x^2 + q_2 \cdot x + q_3)$	$p_1 = -36,59$ $p_2 = 25,2$ $p_3 = 4,732$ $q_1 = -59,71$ $q_2 = -478,2$ $q_3 = 547,8$
ΔMDSI до $P_{2,7\sigma}$ (4:4:4)	$f(x) = (p_1 \cdot x^2 + p_2 \cdot x + p_3) / (x^3 + q_1 \cdot x^2 + q_2 \cdot x + q_3)$	$p_1 = -61,94$ $p_2 = 110,5$ $p_3 = -8,54$ $q_1 = 2341$ $q_2 = 1243$ $q_3 = 71,58$

Рис 3.13 – Діаграми розсіяності $\Delta\text{PSNR-HA}$ до $P_{2\sigma}$ (а), $\Delta\text{PSNR-HA}$ до $P_{2,7\sigma}$ (б), ΔMDSI до $P_{2\sigma}$ (в), ΔMDSI до $P_{2,7\sigma}$ (г) і вписані криві (формат 4:2:2).

На перший погляд результати досить сильно схожі з тими, що були отримані для режиму роботи 4:4:4. Як результат висновки теж практично однакові. Згідно з метрикою $\Delta\text{PSNR-НА}$, ОПТ існує, якщо $P_{2\sigma} \geq 0,85$ (рис. 3.13, а) або $P_{2,7\sigma} \leq 0,1$ (рис. 3.13, б). Графік розсіювання на рис. 3.13 (в) аналогічний графіку на рис. 3.12 (а), а графік розсіювання на рис. 3.13 (г) - аналогічний графіку на рис. 3.12 (б). За метрикою ΔMDSI висновки практично ті самі, що й за $\Delta\text{PSNR-НА}$: ОПТ існує, якщо $P_{2\sigma} \geq 0,82$ (рис. 3.13, в) або $P_{2,7\sigma} \leq 0,1$ (рис. 3.13, г). Таким чином, результати для двох різних візуальних метрик перебувають у досить хорошій згоді. Параметри і самі вирази для режиму роботи 4:2:2 наведено в таблиці 3.5.

Таблиця 3.5 – Параметри вписаних кривих.

Залежність	Вираз	Параметри
$\Delta\text{PSNR-НА}$ до $P_{2\sigma}$ (4:2:2)	$f(x) = (p1 \cdot x^2 + p2 \cdot x + p3) / (x^3 + q1 \cdot x^2 + q2 \cdot x + q3)$	$p1 = 4,964 \cdot 10^4$ $p2 = -4,162 \cdot 10^4$ $p3 = 1942$ $q1 = -1,602 \cdot 10^4$ $q2 = 1,342 \cdot 10^4$ $q3 = 2861$
$\Delta\text{PSNR-НА}$ до $P_{2,7\sigma}$ (4:2:2)	$f(x) = (p1 \cdot x^2 + p2 \cdot x + p3) / (x^3 + q1 \cdot x^2 + q2 \cdot x + q3)$	$p1 = -5,772 \cdot 10^4$ $p2 = 6,093 \cdot 10^4$ $p3 = -6402$ $q1 = 2,003 \cdot 10^4$ $q2 = -2,481 \cdot 10^4$ $q3 = -717,6$
ΔMDSI до $P_{2\sigma}$ (4:2:2)	$f(x) = (p1 \cdot x^2 + p2 \cdot x + p3) / (x^3 + q1 \cdot x^2 + q2 \cdot x + q3)$	$p1 = -25,17$ $p2 = 21,54$ $p3 = -0,8092$ $q1 = -152$ $q2 = -256$ $q3 = 408,8$
ΔMDSI до $P_{2,7\sigma}$ (4:2:2)	$f(x) = (p1 \cdot x + p2) / (x^2 + q1 \cdot x + q2)$	$p1 = 0,008652$ $p2 = -0,0008153$ $q1 = 0,1452$ $q2 = 0,00661$

Щодо режиму роботи 4:2:0, то отримані залежності (рис. 3.14) аналогічно повторюють поведінку для минулих двох випадків.

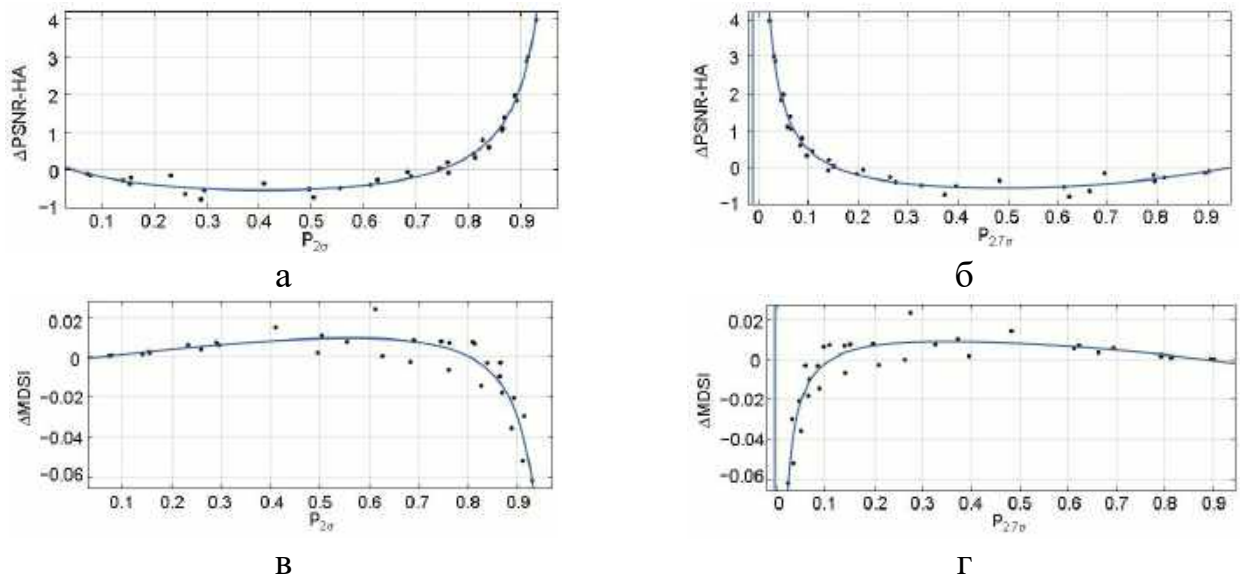


Рис 3.14 – Діаграми розсіяності $\Delta\text{PSNR-HA}$ до $P_{2\sigma}$ (а), $\Delta\text{PSNR-HA}$ до $P_{2,7\sigma}$ (б), ΔMDSI до $P_{2\sigma}$ (в), ΔMDSI до $P_{2,7\sigma}$ (г) і підігнані криві (формат 4:2:0).

Вирази для вписаної кривої режиму 4:2:0 для різних метрик і вхідних параметрів зазначено в таблиці 3.6.

Таблиця 3.6 – Параметри вписаних кривих.

Залежність	Вираз	Параметри	
$\Delta\text{PSNR-HA}$ до $P_{2\sigma}$ (4:2:0)	$f(x) = (p1 \cdot x^2 + p2 \cdot x + p3) / (x^3 + q1 \cdot x^2 + q2 \cdot x + q3)$	$p1 = 6922$ $p2 = -5483$ $p3 = 243,9$	$q1 = -4101$ $q2 = 3003$ $q3 = 1025$
$\Delta\text{PSNR-HA}$ до $P_{2,7\sigma}$ (4:2:0)	$f(x) = (p1 \cdot x^2 + p2 \cdot x + p3) / (x^3 + q1 \cdot x^2 + q2 \cdot x + q3)$	$p1 = 2,433$ $p2 = -2,668$ $p3 = 0,3562$	$q1 = -2,571$ $q2 = 2,324$ $q3 = 0,02283$
ΔMDSI до $P_{2\sigma}$ (4:2:0)	$f(x) = (p1 \cdot x^2 + p2 \cdot x + p3) / (x^3 + q1 \cdot x^2 + q2 \cdot x + q3)$	$p1 = -12,84$ $p2 = 11,17$ $p3 = -0,7154$	$q1 = -86,94$ $q2 = -255,8$ $q3 = 334,8$

Що стосується значень точності, то для режимів роботи 4:2:2 і 4:2:0 показники точності вписування наведено в таблиці 3.7.

Таблиця 3.7 – Точність вписування для режимів роботи 4:2:2 и 4:2:0

Залежність	R^2	Adjusted R^2	RMSE
$\Delta\text{PSNR-НА}$ до $P_{2\sigma}$ (4:2:2)	0,965	0,9582	0,2938
$\Delta\text{PSNR-НА}$ до $P_{2,7\sigma}$ (4:2:2)	0,9623	0,9551	0,3048
ΔMDSI до $P_{2\sigma}$ (4:2:2)	0,8778	0,8543	0,007885
ΔMDSI до $P_{2,7\sigma}$ (4:2:2)	0,853	0,8372	0,008334
$\Delta\text{PSNR-НА}$ до $P_{2\sigma}$ (4:2:0)	0,9869	0,9843	0,1482
$\Delta\text{PSNR-НА}$ до $P_{2,7\sigma}$ (4:2:0)	0,9844	0,9814	0,1615
ΔMDSI до $P_{2\sigma}$ (4:2:0)	0,8558	0,8281	0,007658
ΔMDSI до $P_{2,7\sigma}$ (4:2:0)	0,8431	0,813	0,007988

Аналіз даних, наведений раніше для всіх варіантів роботи, показує таке. По-перше, для метрики $\Delta\text{PSNR-НА}$ вписування виконано з хорошими показниками точності для всіх форматів. Значення R^2 становлять близько 0,97 (від 0,95 до 0,99), а RMSE - близько або менше 0,5. Це відповідні передумови для точного прогнозування. По-друге, вписування для метрики ΔMDSI характеризується R^2 близько 0,9 для формату 4:4:4 і близько 0,83 для двох інших форматів. Така точність прийнятна і на практиці. По-третє, використаний вхідний параметр практично не має значення, оскільки значення R^2 і RMSE (для одного й того самого формату і прогнозованої метрики) практично однакові.

Таким чином, результати для різних форматів трохи відрізняються, але загальні тенденції збігаються.

3.2.3 Перевірка прогнозу та алгоритм прийняття рішень

Упевнившись, що вписування можливе з досить-таки високими показниками точності, можна перейти безпосередньо до прогнозування. Варто нагадати, які

можливі результати під час прогнозування, іншими словами, як трактувати $\Delta\text{Метр}$ (у цьому разі $\Delta\text{PSNR-НА}$ або ж ΔMDSI). Розглянемо варіанти на прикладі $\Delta\text{PSNR-НА}$: перший випадок, коли $\Delta\text{PSNR-НА}$ більший за 1 дБ, у такому разі ОПТ присутнє з великою ймовірністю, у такому разі для вибору параметра контролю стиснення необхідно використовувати (2.4). Другий можливий підсумок прогнозування це $-1\text{дБ} \leq \Delta\text{PSNR-НА} \leq 1\text{дБ}$ у цьому разі ОПТ може існувати, і використовуйте $Q=Q_{\text{ОПТ}}-1$, це дасть змогу уникнути надмірного згладжування зображення. Останній із можливих результатів $\Delta\text{PSNR-НА} \leq -1\text{дБ}$, коли ОПТ відсутній, і для такого випадку необхідно використовувати $Q=Q_{\text{ОПТ}}-3$, але не більше 25. За такого значення Q спотворення непомітні або, принаймні, не заважають подальшому обробленню зображення.

Отримані результати свідчать що прогнозування може бути доволі точним, тому потрібно підтвердити це на практиці. Для цього було обрано трьохканальні зображення, які не використовувалися при отриманні діаграми розсіювання (рис. 3.15).

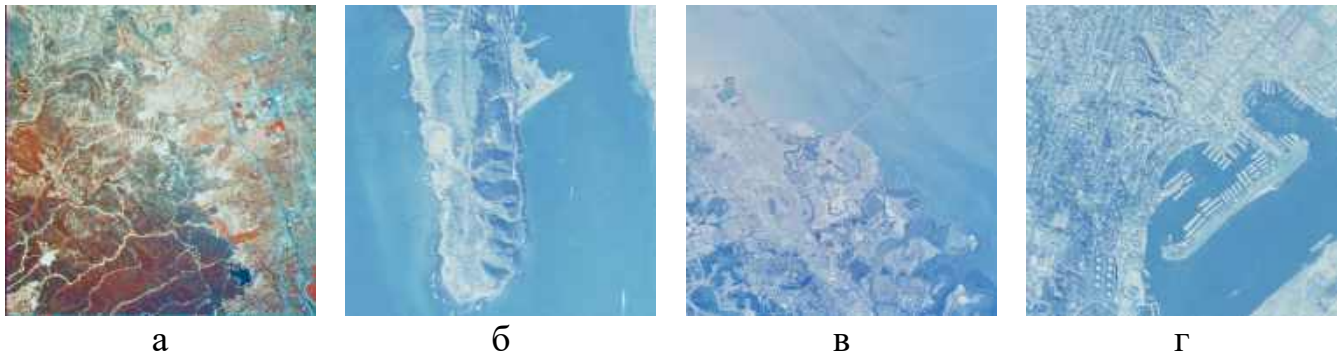


Рис 3.15 – Тестові трьох-канальні зображення Woodland Hills (а), Point Loma (б), Foster City (в) і Shelter Island (г).

Використовуючи отримані раніше вписані криві, було порівняно значення дельта-метрик. Порівнюються прогнозовані та реальні дані, отримані в ході розрахунку. Результати для набору зображень (рис. 3.15), уражених AWGN з дисперсією, що дорівнює 100, зібрано в таблиці 3.8.

Таблиця 3.8 – Порівняння прогнозованих та реальних значень дельта метрик для режимів роботи 4:2:2 і 4:2:0

Зображення	Формат	Вхідний параметр	Прогнозоване $\Delta PSNR - HA$	Розраховане $\Delta PSNR - HA$	Прогнозоване $\Delta MDSI$	Розраховане $\Delta MDSI$
Woodland Hills	4:4:4	$P_{2\sigma}$	-1,7	-1,8	0,021	0,033
		$P_{2,7\sigma}$	-1,64		0,023	
Point Loma	4:4:4	$P_{2\sigma}$	2,2	1,8	-0,031	-0,018
		$P_{2,7\sigma}$	2,52		-0,037	
Foster City	4:4:4	$P_{2\sigma}$	0,53	0,97	-0,006	-0,024
		$P_{2,7\sigma}$	0,53		-0,008	
Shelter Island	4:4:4	$P_{2\sigma}$	-0,29	-0,74	0,005	0,008
		$P_{2,7\sigma}$	-0,29		0,005	
Woodland Hills	4:2:2	$P_{2\sigma}$	-0,42	-0,52	0,01	0,019
		$P_{2,7\sigma}$	-0,43		0,01	
Point Loma	4:2:2	$P_{2\sigma}$	2,31	2,16	-0,031	-0,019
		$P_{2,7\sigma}$	2,6		-0,037	
Foster City	4:2:2	$P_{2\sigma}$	1,04	1,6	-0,011	-0,023
		$P_{2,7\sigma}$	1,08		-0,011	
Shelter Island	4:2:2	$P_{2\sigma}$	0,48	0,18	-0,002	0,01
		$P_{2,7\sigma}$	0,48		-0,001	
Woodland Hills	4:2:0	$P_{2\sigma}$	-0,02	-0,12	0,005	0,012
		$P_{2,7\sigma}$	-0,02		0,005	
Point Loma	4:2:0	$P_{2\sigma}$	2,18	2,09	-0,029	-0,016
		$P_{2,7\sigma}$	2,45		-0,033	
Foster City	4:2:0	$P_{2\sigma}$	1,06	1,57	-0,011	-0,021
		$P_{2,7\sigma}$	1,09		-0,01	
Shelter Islands	4:2:0	$P_{2\sigma}$	0,62	0,47	-0,004	0,007
		$P_{2,7\sigma}$	0,62		-0,004	

Аналіз даних показує таке: 1) прогнози на основі обох розглянутих вхідних параметрів дуже близькі для обох прогнозованих метрик; 2) прогнози досить близькі до отриманих (розрахованих) значень досліджуваних метрик; 3) прогнози показують, що OPT існує для тестового зображення Point Loma за обома метриками для всіх трьох форматів і для зображення Foster City, також за обома метриками та для всіх трьох форматів; OPT не існує для зображення Woodland Hills згідно з обома

метриками для всіх трьох форматів; 4) ситуація гранична для зображення Shelter Islands - OPT не існує для формату 4:4:4, він існує для форматів 4:2:2 і 4:2:0, але значення метрики дуже близькі до нуля.

Як результат для всіх режимів роботи спрогнозовані дані близькі до справжніх (розрахованих даних), що свідчить про те, що цей метод прогнозування оптимальної робочої точки для VPG кодера можна застосувати до кольорових зображень, уражених Гаусовим шумом.

3.2.4 Прогнозування в околі OPT

Виходячи із залежностей, отриманих у другому розділі, що відображають відношення якості стисненого зображення до збільшення параметра контролю стиснення, інтерес може становити не тільки сама OPT, а й її околиці, оскільки вони також можуть дати необхідну інформацію для ухвалення рішень щодо подальшої обробки стисненого зображення. У зв'язку з цим з'являється інтерес до прогнозування для околів оптимальної робочої точки.

Аналіз кривих у другому розділі (рис. 2.15) показує, що немає сенсу прогнозувати дельта-параметри в найближчому околі $Q = Q_{OPT}-1$ або $Q = Q_{OPT}+1$, оскільки криві плавні, а відмінності між значеннями $\Delta MDSI(Q_{OPT}-1)$, $\Delta MDSI(Q_{OPT})$ і $\Delta MDSI(Q_{OPT}+1)$ малі й неінформативні. У зв'язку з цим необхідно вивчити можливість прогнозування дельта-параметрів у ширших околах, наприклад, у $Q = Q_{OPT}-4$ і $Q = Q_{OPT}+4$.

Методика прогнозування та аналізу його властивостей така сама, як і раніше. По-перше, аналізуються графіки розсіювання і можливість вписування в них кривої. По-друге, вивчається, яка точність може бути забезпечена в цьому випадку.

Приклади отриманих графіків для режиму роботи 4:2:2 представлено на рисунку 3.16. У якості вхідних параметрів знову використовувалися $P_{2\sigma}$ і $P_{2,7\sigma}$. Попередній аналіз цих діаграм розсіювання показує, що точки розташовані компактно, що свідчить про наявність явної залежності між розглянутими параметрами.

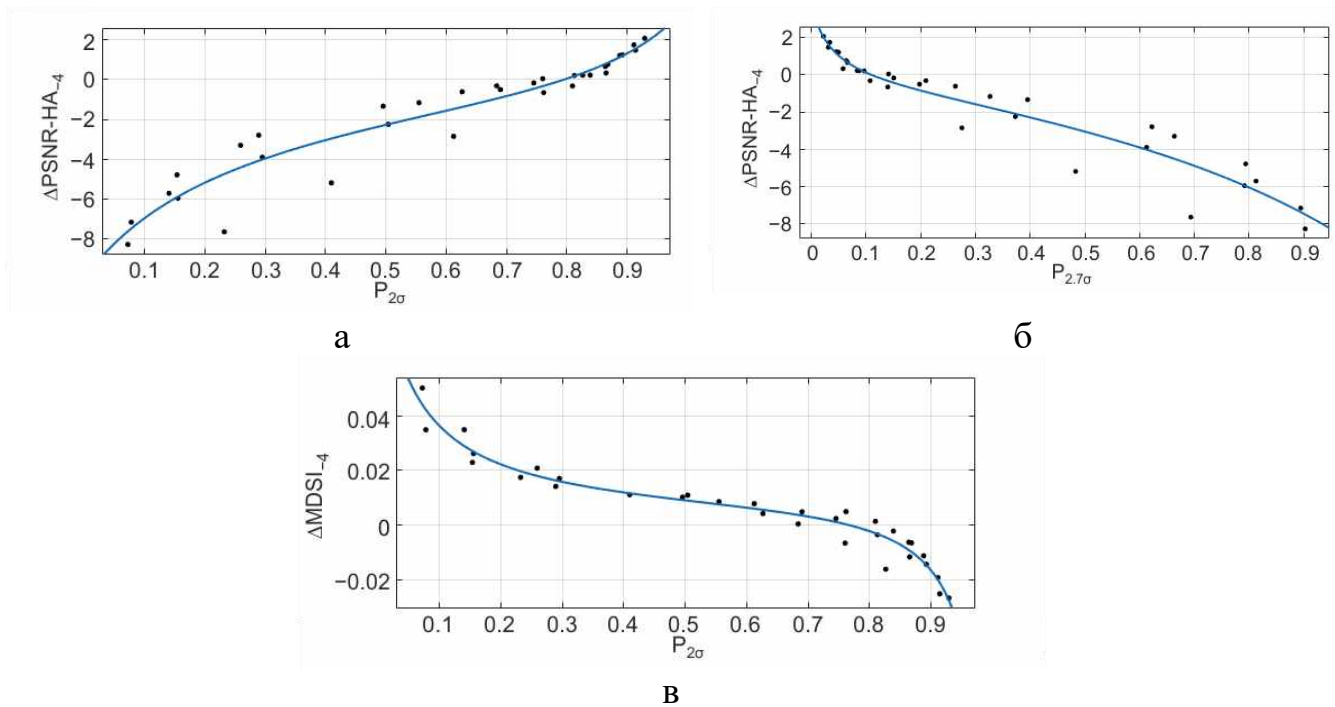


Рис 3.16 – Діаграми розсіювання для: $\Delta\text{PSNR-HA}_{-4}$ до $P_{2\sigma}$ (а), $\Delta\text{PSNR-HA}_{-4}$ до $P_{2,7\sigma}$ (б), ΔMDSI_{-4} до $P_{2\sigma}$ (в) і вписані криві

Для вписування всіх кривих було використано раціональні функції, що зарекомендували себе як такі, що можуть надати гарні показники точності вписування для таких завдань. Як результат, для $\Delta\text{PSNR-HA}_{-4}$ і вхідного параметра $P_{2\sigma}$ (рис. 3.16, а) величини, що характеризують точність, дорівнюють: $R^2=0,911$, Adjusted R^2 дорівнює 0,898, а $\text{RMSE}=0,941$. Як видно, отримана точність досить висока. Здебільшого $\Delta\text{PSNR-HA}_{-4}$ негативні, хоча є зображення, для яких спостерігаються невеликі позитивні значення $\Delta\text{PSNR-HA}_{-4}$. Це відбувається за $P_{2\sigma}>0,8$, тобто для зображень із простою структурою або високою інтенсивністю шуму.

Для тієї самої метрики, але для параметра $P_{2,7\sigma}$ (рис. 3.16, б) крива вписана також досить добре: $R^2=0,911$, Adjusted R^2 дорівнює 0,894, а $\text{RMSE}=0,960$. Ці характеристики дещо гірші, ніж для $P_{2\sigma}$, застосованого як вхідний параметр.

Якщо ж як дельта метрику використовувати ΔMDSI (рис. 3.16, в), то для неї отримуємо таке: $R^2=0,951$, Adjusted R^2 дорівнює 0,944, а $\text{RMSE}=0,0042$. Як видно, позитивний вплив стиснення з втратами на якість зображення спостерігається за $P_{2\sigma}>0,85$.

Що стосується протилежного околу оптимальної робочої точки $\Delta\text{Метр}_{+4}$, то ситуація схожа в плані хороших показників точності вписування кривих. Результати вписування показано на рисунку 3.17.

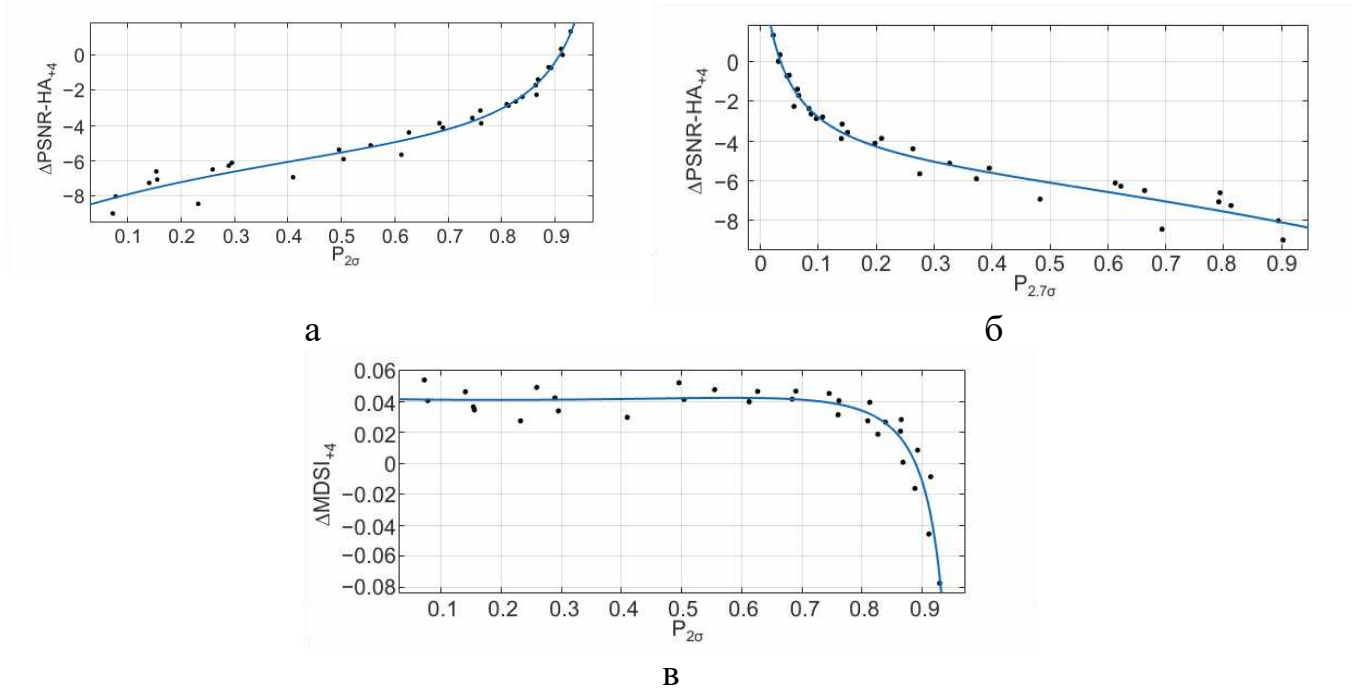


Рис 3.17 - Діаграми розсіювання для: $\Delta\text{PSNR-HA}_{+4}$ та $P_{2\sigma}$ (а), $\Delta\text{PSNR-HA}_{+4}$ та $P_{2,7\sigma}$ (б), ΔMDSI_{+4} та $P_{2\sigma}$ (в) і вписані криві

На рис. 3.17, (а) представлено графік розсіювання $\Delta\text{PSNR-HA}_{+4}$ проти $P_{2\sigma}$ для формату 4:2:2, де $R^2=0.967$, Adjusted R^2 дорівнює 0.963, а $\text{RMSE}=0.526$. Таким чином, отримана точність досить висока. Здебільшого $\Delta\text{PSNR-HA}_{+4}$ має негативні значення, хоча існують зображення, для яких невеликі позитивні значення $\Delta\text{PSNR-HA}_{+4}$ мають місце для $P_{2\sigma}>0,9$, тобто для зображень із дуже простою структурою або високою інтенсивністю шуму. Для другого випадку (рис. 3.17, б) крива вписана досить добре, де $R^2=0,965$, Adjusted R^2 дорівнює 0,960, а $\text{RMSE}=0,547$. Ці характеристики лише трохи гірші, ніж для $P_{2\sigma}$. Графік розсіювання ΔMDSI_{+4} до $P_{2\sigma}$ з вписаною кривою подано на рис.3.17, (в). Аналіз показує компактний характер даних, де $R^2=0.888$, Adjusted R^2 дорівнює 0.871 і $\text{RMSE}=0.010$. Позитивний вплив стиснення з втратами на візуальну якість зображення має місце за $P_{2\sigma}>0,89$. Варто зазначити, що R^2 і

Adjusted R^2 трохи менші за 0,9, і це може свідчати о високій точності вписування та передбачення.

Аналогічні операції було проведено для решти режимів роботи, точність вписування для отриманих вписаних кривих зазначено в таблиці 3.9, 3.10.

Таблиця 3.9 – Точність вписування Δ Метр-4 для режимів роботи 4:4:4 і 4:2:0

Формат	Дельта-метрика	R^2	Adjusted R^2	RMSE
4:2:0	Δ PSNR- HA_{-4} до $P_{2\sigma}$	0,931	0,9144	1,114
4:2:0	Δ MDSI- $_4$ до $P_{2\sigma}$	0,947	0,9343	0,0058
4:4:4	Δ PSNR- HA_{-4} до $P_{2\sigma}$	0,9447	0,934	0,2036
4:4:4	Δ MDSI- $_4$ до $P_{2\sigma}$	0,965	0,9549	0,0018
4:2:0	Δ PSNR- HA_{-4} до $P_{2,7\sigma}$	0,9309	0,9176	1,093
4:2:0	Δ MDSI- $_4$ до $P_{2,7\sigma}$	0,9493	0,9371	0,0057
4:4:4	Δ PSNR- HA_{-4} до $P_{2,7\sigma}$	0,9398	0,928	0,2124
4:4:4	Δ MDSI- $_4$ до $P_{2,7\sigma}$	0,9502	0,936	0,0022

Представлені результати (таблиця 3.9) показують, що дельта-параметри для $Q=Q_{OPT-4}$ можуть бути спрогнозовані з досить високою точністю для двох форматів, як і в разі формату 4:2:2. Як видно, у всіх випадках R^2 і Adjusted R^2 більші за 0,9, тобто вписування досить хороше, а отже, прогнозування досить точне.

Таблиця 3.10 – Точність вписування Δ Метр- $_{+4}$ для режимів роботи 4:4:4 і 4:2:0

Формат	Дельта-метрика	R^2	Adjusted R^2	RMSE
4:2:0	Δ PSNR- HA_{+4} до $P_{2\sigma}$	0,9689	0,9629	0,6853
4:2:0	Δ MDSI- $_{+4}$ до $P_{2\sigma}$	0,8841	0,8618	0,01216
4:4:4	Δ PSNR- HA_{+4} до $P_{2\sigma}$	0,9613	0,9538	0,3736
4:4:4	Δ MDSI- $_{+4}$ до $P_{2\sigma}$	0,9161	0,896	0,00837
4:2:0	Δ PSNR- HA_{+4} до $P_{2,7\sigma}$	0,9689	0,9643	0,673
4:2:0	Δ MDSI- $_{+4}$ до $P_{2,7\sigma}$	0,8726	0,8538	0,01251
4:4:4	Δ PSNR- HA_{+4} до $P_{2,7\sigma}$	0,9512	0,9418	0,4194
4:4:4	Δ MDSI- $_{+4}$ до $P_{2,7\sigma}$	0,9001	0,8761	0,009135

З даних наведених в табл. 3.10 можна зробити висновок, що ситуація схожа і для Δ Метр- $_{+4}$, а саме тенденції, що спостерігаються, справедливі для двох форматів (4:4:4 та 4:2:0). Результати вписування для Δ PSNR- HA_{+4} хороші як для $P_{2\sigma}$, так і для $P_{2,7\sigma}$ у

сенсі високого R^2 . Однак ті самі величини для ΔMDSI_{+4} не настільки хороші й можуть бути поліпшені. Тому в наступному підрозділі розглянемо один зі способів досягнення цієї мети.

3.2.5 Прогнозування за кількома вхідними параметрами

Одним зі способів підвищення точності є використання декількох вхідних параметрів, які можуть класифікувати зображення з різних боків. Варто нагадати, що маючи набір локальних оцінок (ЛО) імовірностей у блоках, легко обчислити не тільки їх середнє, а й дисперсію, асиметрію або ексцес. Ці параметри можуть бути використані як один із варіантів додаткових вхідних параметрів. Таким чином, використання двох вхідних параметрів не призводить до значного збільшення обсягу обчислень, необхідних для прогнозування, оскільки ці параметри розраховуються за локальною активністю, яка була розрахована заздалегідь для досліджуваного зображення. Один із прикладів використання кількох параметрів для вписування продемонстровано на рисунку 3.18.

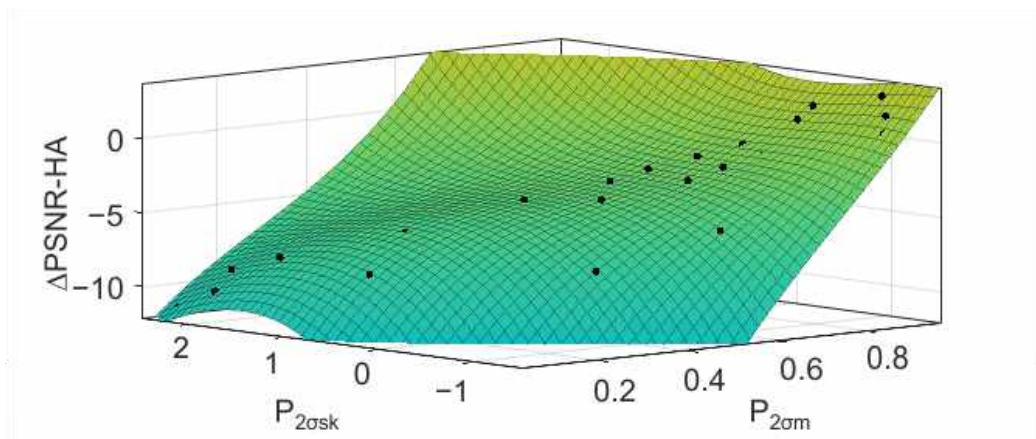


Рис 3.18 – Вписана площина для $\Delta\text{PSNR-HA}$, де вхідними параметрами є $P_{2\sigma}$ і асиметрія локальних оцінок цього параметра

Дані щодо використання інших параметрів наведено в таблиці 3.11, у ній також наведено значення метрик точності для пар вхідних параметрів, а також значення

RMSE для випадку використання лише одного параметра ($P_{2\sigma}$ або $P_{2,7\sigma}$) для можливості порівняння результатів вписування.

Таблиця 3.11 – Характеристики точності вписування для двох вхідних параметрів і порівняння з однопараметричним вписуванням

Прогнозований параметр	Перший вхідний параметр	Другий вхідний параметр	R^2	Adjusted R^2	RMSE	RMSE для одного вхідного параметра
$\Delta PSNR-HA_{-4}$	$P_{2\sigma}$	Асиметрія ЛО $P_{2\sigma}$	0,995	0,993	0,323	1,114
$\Delta PSNR-HA_{-4}$	$P_{2,7\sigma}$	Асиметрія ЛО $P_{2,7\sigma}$	0,994	0,991	0,359	1,093
$\Delta MDSI_{-4}$	$P_{2\sigma}$	Асиметрія ЛО $P_{2\sigma}$	0,973	0,961	0,004	0,006
$\Delta MDSI_{-4}$	$P_{2,7\sigma}$	Асиметрія ЛО $P_{2,7\sigma}$	0,966	0,952	0,005	0,006
$\Delta PSNR-HA$	$P_{2\sigma}$	Асиметрія ЛО $P_{2\sigma}$	0,99	0,986	0,459	0,875
$\Delta PSNR-HA$	$P_{2,7\sigma}$	Асиметрія ЛО $P_{2,7\sigma}$	0,987	0,981	0,521	0,88
$\Delta MDSI$	$P_{2\sigma}$	Дисперсія ЛО $P_{2\sigma}$	0,929	0,900	0,011	0,011
$\Delta MDSI$	$P_{2,7\sigma}$	Ексцес ЛО $P_{2,7\sigma}$	0,937	0,911	0,011	0,012
$\Delta PSNR-HA_{+4}$	$P_{2\sigma}$	Асиметрія ЛО $P_{2\sigma}$	0,985	0,978	0,525	0,685
$\Delta PSNR-HA_{+4}$	$P_{2,7\sigma}$	Асиметрія ЛО $P_{2,7\sigma}$	0,981	0,973	0,581	0,673
$\Delta MDSI_{+4}$	$P_{2\sigma}$	Ексцес ЛО $P_{2\sigma}$	0,941	0,917	0,009	0,012
$\Delta MDSI_{+4}$	$P_{2,7\sigma}$	Ексцес ЛО $P_{2,7\sigma}$	0,922	0,89	0,011	0,013

Як видно, у всіх розглянутих випадках спостерігається достатнє поліпшення RMSE. Для деяких комбінацій із двох вхідних параметрів RMSE порівняно з випадком одного вхідного параметра зменшився майже вчетверо. Поліпшення точності спостерігається і для параметрів R^2 , що характеризують вписування.

3.2.6 Обговорення практичних аспектів

Отримані дані показують, що точність прогнозування $\Delta\text{Метр}$ в околі OPT перебуває на такому самому високому рівні, як і для самої OPT, крім цього, існує можливість підвищити значення точності завдяки використанню кількох параметрів. Як результат, стає можливим спрогнозувати $\Delta\text{Метр}_{-4}$, $\Delta\text{Метр}_{\text{OPT}}$, і $\Delta\text{Метр}_{+4}$.

Варто нагадати, що, використовуючи один і той самий вхідний параметр або пару вхідних параметрів, можна одночасно обчислити дельта-метрику в $Q=Q_{\text{OPT}}-4$, $Q=Q_{\text{OPT}}$ і $Q=Q_{\text{OPT}}+4$. Це дає змогу ухвалювати рішення про те, яке значення Q слід встановити. Припустимо, що було обчислено $P_{2\sigma}$ і асиметрію локальних оцінок цього параметра. Потім, використовуючи отримані двовимірні апроксимації, було отримано спрогнозовані значення $\Delta\text{PSNR-NA}_{-4}$, $\Delta\text{PSNR-NA}_{\text{OPT}}$, і $\Delta\text{PSNR-NA}_{+4}$. На рис. 3.19 ці значення показано двома стовпами, де червоні та сині стовпи відповідають прогнозованому та істинному значенням, відповідно. На рис. 3.19, (а) показано випадок для зображення простої структури (тестове зображення Frisco, дисперсія шуму 196), а на рис. 3.19, (б) - дані для зображення складної структури (тестове зображення Diego, дисперсія шуму 64).

Аналіз даних на рис. 3.19 показує таке. По-перше, істинне і спрогнозовані значення дуже близькі. По-друге, якщо OPT існує, тобто $\Delta\text{PSNR-NA}$ є позитивним і більшим за $\Delta\text{PSNR-NA}_{-4}$ і $\Delta\text{PSNR-NA}_{+4}$ (рис. 3.19, а), то рішення має бути ухвалено на користь стиснення в OPT. Також можна стиснути зображення за допомогою $Q=Q_{\text{OPT}}+4$, якщо відповідне $\Delta\text{PSNR-NA}_{+4}$ є позитивним, це дає змогу забезпечити більший КС за рахунок дещо гіршої візуальної якості. Припустимо тепер, що всі дельта-метрики $\Delta\text{PSNR-NA}_{-4}$, $\Delta\text{PSNR-NA}_{\text{OPT}}$ і $\Delta\text{PSNR-NA}_{+4}$ негативні (як на рис. 3.19, б), де $\Delta\text{PSNR-NA}_{-4} > \Delta\text{PSNR-NA}_{\text{OPT}} > \Delta\text{PSNR-NA}_{+4}$. Тоді можна прийняти рішення про стиснення такого зображення з $Q=Q_{\text{OPT}}-4$.

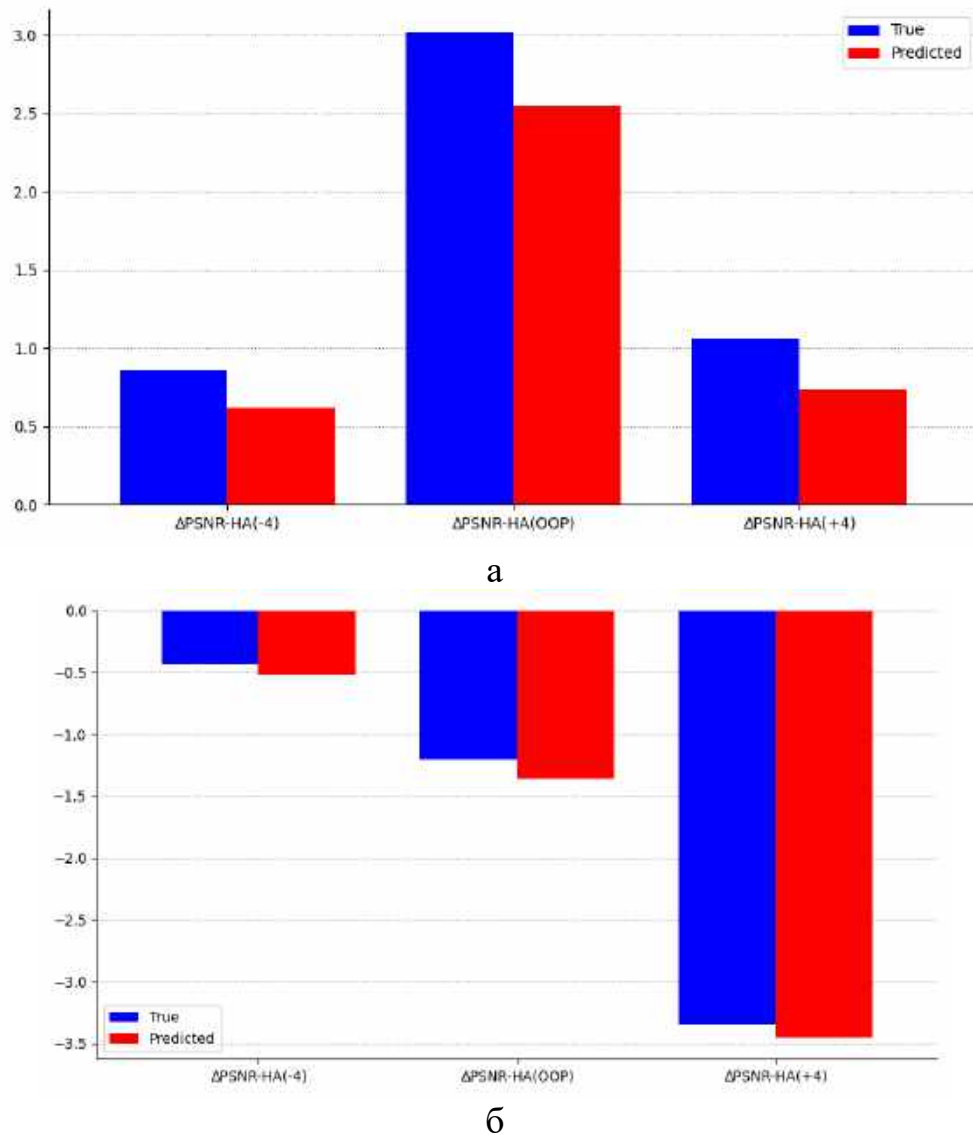


Рис 3.19 – Гістограма $\Delta\text{PSNR-HA}$ для OPT і його околу для зображення простої структури (а) і зображення складної структури (б)

Як підсумок, алгоритм прогнозування та ухвалення рішень наступний: перший, за аналогією з попереднім, оцінювання дисперсії шуму (цей крок необхідний, якщо дисперсія шуму не відома заздалегідь), далі потрібно розрахувати Q_{OPT} згідно з (2.4) і, потім, $Q_{\text{OPT}-4}$ та $Q_{\text{OPT}+4}$, після чого розраховувати вибрані вхідні параметри (прикладом є $P_{2\sigma}$ і асиметрію його локальної оцінки). Фінальними етапами залишаються розрахунок $\Delta\text{PSNR-HA}_{-4}$, $\Delta\text{PSNR-HA}_{\text{OPT}}$, і $\Delta\text{PSNR-HA}_{+4}$, і, ґрунтуючись на отриманих даних, ухвалити рішення про вибір параметра контролю стиснення. Ухвалення рішення схоже з попередніми методами (якщо $\Delta\text{Метр}$ від'ємна, то OPT відсутня і навпаки).

3.3 Висновки по розділу

У цьому розділі за отриманими результатами можна зробити такі висновки:

– Продемонстровано метод прогнозування оптимальної робочої точки для BPG кодера при роботі з напівтоновими зображеннями. Він заснований на попередньому отриманні апроксимуючих залежностей (за допомогою діаграм розсіювання і вписування кривих). Потім для заданого зображення необхідно оцінити простий статистичний параметр (аналіз коефіцієнтів ДКП), використати його значення як вхідні дані апроксиматора і виконати прогнозування з ухваленням рішень. Точність прогнозування проаналізовано і показано як придатну для практики. Також наведено можливість модифікації цього методу для роботи з сигнально-залежними типами шумів. Результати прогнозування було перевірено на деяких реальних прикладах і показано хорошу точність для прийняття рішень про існування ОРТ.

– Показано, що описаний метод прогнозування може бути модифікований для роботи з кольоровими зображеннями. Цей метод може бути застосований до різних форматів кольоровості (4:4:4, 4:2:2, 4:2:0). Запропоновану процедуру прогнозування протестовано для триканальних зображень, які не використовувалися для отримання діаграм розсіювання і вписування кривих. Показано, що прогнозування для них є досить точним. Крім того, розглянуто прогнозування для трьох точок в околах ОРТ, що дає змогу ухвалювати рішення про налаштування Q завдяки спільному аналізу прогнозів.

РОЗДІЛ 4

РОЗРОБКА МЕТОДУ ПРОГНОЗУВАННЯ СКП ДЛЯ VP8 КОДЕРА

У цьому розділі розглядаються методи прогнозування значень СКП для стиснених VP8 кодером зображень. Крім того, детальніше розглядаються передумови до прогнозування, а також особливості самого прогнозування під час роботи з напівтоновими зображеннями. Розроблено та продемонстровано методи, засновані на різних підходах, серед яких методи, що базуються на використанні локальної активності в блоках або ентропії. У даному розділі також приділено увагу завданню підвищення точності прогнозування СКП і проведено порівняння розроблених методів. Окрему увагу приділено корекції параметра контролю стиснення, ґрунтуючись на результатах прогнозування.

4.1 Передумови до прогнозування

Виходячи з попередніх розділів, стають очевидними позитивні ефекти від стиснення зображень за допомогою VP8 кодера в контексті як наявності оптимальної робочої точки, так і можливості її прогнозування. Однак можливі ситуації, коли необхідно контролювати якість стиснутих зображень, оскільки їх може бути проаналізовано на наступному етапі оброблення зображень, а спотворення, отримані під час процедури стиснення, мають не перевищувати певний рівень. Як результат, такий контроль може бути корисним у сценаріях, коли зображення має бути стиснуто без візуальних втрат або має бути досягнутий відповідний компроміс між досягнутим коефіцієнтом стиснення (КС) і якістю стисненого зображення. Для розв'язання цієї задачі в даному розділі запропоновано способи прогнозування середньої квадратичної помилки (СКП) внесених спотворень для кращого стиснення з втратами, використовуючи VP8 кодер.

Ґрунтуючись на результатах, отриманих у попередніх розділах, можна дійти висновку, що СКП і, відповідно, пікове відношення сигнал/шум (PSNR) добре контролюються параметром Q, що використовується в VP8-кодері для зміни

характеристик стиснення. Однак СКП також залежить від характеристик зображення, які можна описати різними методами, зокрема за допомогою ентропії. Таким чином, під час стиснення зображення необхідно враховувати його характеристики. У цьому розділі вивчається, наскільки добре різні методи можуть описувати характеристики зображення, а саме його складність, для майбутнього прогнозування СКП під час стиснення з втратами на основі BPG.

4.1.1 Аналіз залежностей метрик якості від параметра контролю стиснення

Варто почати з розгляду залежностей метрик якості стиснутих зображень для різних значень параметра контролю стиснення задля виявлення закономірностей, що дали б змогу визначити особливості поведінки зображень та виділити ділянки, на яких прогнозування метрик є найбільш раціональним (доцільним).

Для отримання кривих швидкість-спотворення було використано набір зображень, який складається з зображень на рис. 2.2, а також кількох нових, включно із зображеннями Beach14 і Beach16, які являють собою напівтонові версії реальних кольорових (трьох-канальних) знімків дистанційного зондування низької складності, крім того, одне штучно згенероване зображення RSA. Мініатюри даних зображень продемонстровано на рис. 4.1.

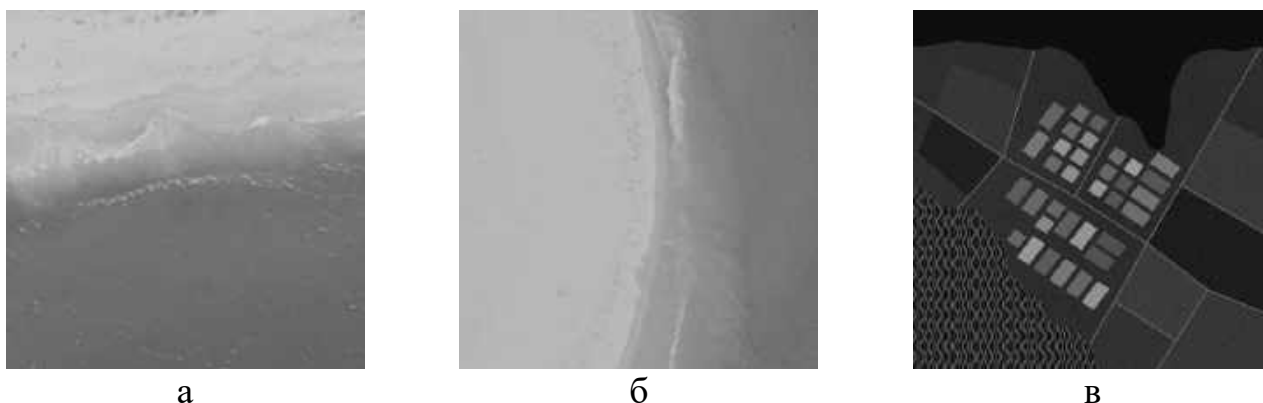


Рис 4.1 – Мініатюри використаних зображень: а) Beach14; б) Beach16; в) RSA

В якості метрик якості були обрані вже використані раніше PSNR і PSNR-HVS-M. Результати для цих метрик наведено на рис. 4.2. Оскільки мета -

продемонструвати залежність між поведінкою зображення та його складністю, лінії на отриманих графіках підписано не назвою зображень, а значеннями їх ентропії (як один зі способів кількісно охарактеризувати складність зображення). Її було визначено для всіх шести тестових зображень. Значення ентропії дорівнюють 7,36 для Baboon, 5,82 для Frisco, 5,18 для Goldhill, 4,93 для RSA, 5,91 і 5,11 для зображень Beach14 і Beach16, відповідно. Теоретично, найвищі значення ентропії відповідають зображенням із вищою складністю.

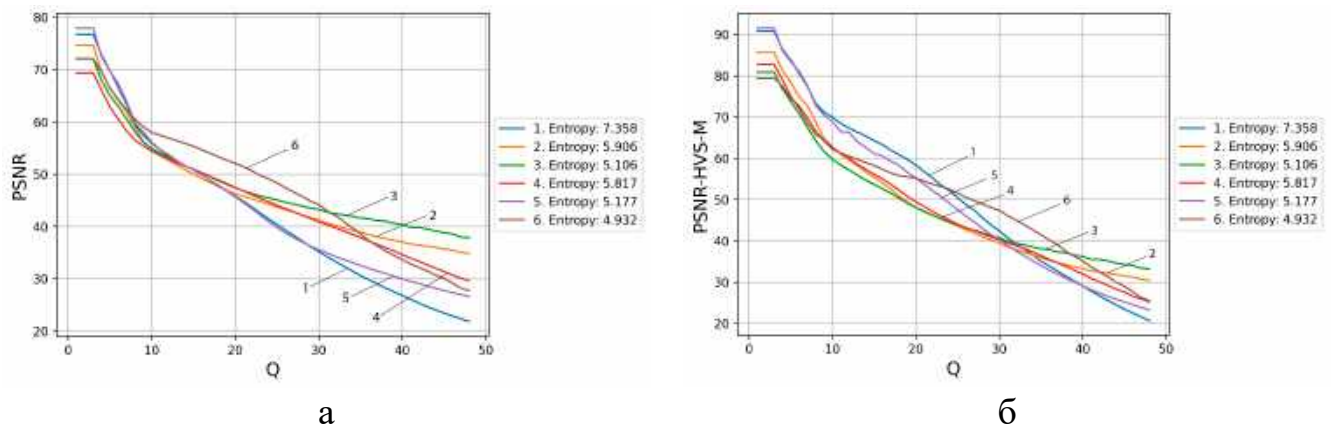


Рис 4.2 – Залежність PSNR (а) і PSNR-HVS-M (б) від Q для шести розглянутих зображень

На рис. 4.2 (а) представлено залежність PSNR від Q для шести тестових зображень. Можна виділити чотири інтервали з різною поведінкою представлених залежностей. За $Q < 4$ поведінка PSNR однакова для всіх зображень, хоча для різних зображень спостерігаються різні значення цієї метрики - PSNR більший для складніших зображень. Однак значення PSNR у будь-якому разі більше 69 дБ, тобто спотворення практично не вносяться. У зв'язку з цим немає необхідності прогнозувати СКП для $Q < 4$. Починаючи з $Q = 4$ і до $Q = 8$, спостерігається швидке зниження PSNR. Однак для $Q = 8$ PSNR перевищує 55 дБ. Як результат, спотворення, як і раніше, непомітні (варто нагадати, що спотворення починають бути помітними, якщо PSNR перебуває в межах від 29 до 45 дБ). Третій інтервал приблизно відповідає Q від 9 до 30. У цьому інтервалі PSNR зменшується приблизно на 1 дБ при збільшенні Q на 1. Саме цей інтервал становить практичний інтерес, особливо для Q від 20 до 30,

оскільки відповідне СКП для цієї ситуації перевищує 1 для $Q=20$ і перебуває в межах від 2,5 для тестового зображення RSA до 20 для тестового зображення Baboon. Це той випадок, коли стиснення без візуальних втрат може вимагати забезпечення бажаного СКП або іншої метрики. Потім, для $Q>30$, спостерігається четвертий інтервал, де помітна достатня різниця в поведінці кривих швидкість-спотворення. Дійсно, значення PSNR можуть відрізнятися більш ніж на 12 дБ (прикладом можуть слугувати дані для $Q\geq 40$), і тут потрібне прогнозування СКП і коригування Q . Найменший PSNR у цьому інтервалі має місце для найскладнішого зображення Baboon, тоді як для зображень із простою структурою значення PSNR на кілька дБ більші. Варто відзначити той факт, що цей інтервал в основному відповідає видимим спотворенням, а СКП може перевищувати 100.

Що стосується другої метрики, то результати наведено на рисунку 4.2 (б) і видно, що знову можна виділити чотири інтервали. За $Q<3$ значення PSNR-HVS-M постійні для кожного зображення і дуже великі, тобто стиснення відбувається практично без втрат. За Q від 4 до 8 відбувається досить швидке зменшення PSNR-HVS-M, хоча значення PSNR-HVS-M все ще великі і внесені спотворення згідно цієї метрики, як і раніше, непомітні. Починаючи з $Q=9$ і до $Q=30$ спостерігається майже лінійне зменшення PSNR-HVS-M, а спотворення стають ледь помітними. Нарешті, за $Q>30$ спотворення стають помітними, і якість зображення може змінюватися в доволі широких межах за одного й того ж Q (на 8 дБ для $Q=40$). У цьому інтервалі значення PSNR-HVS-M для зображень із найскладнішою структурою найменші. Це означає, що такі зображення мають найгіршу візуальну якість.

Узагальнюючи отримані спостереження, можна стверджувати, що немає необхідності прогнозувати СКП для перших двох інтервалів, тобто для $Q<9$. Більше того, немає гострої необхідності робити це і для $Q<20$, оскільки спотворення, що вносяться, не помітні. Водночас прогнозування якості для $Q\geq 20$ є доцільним, оскільки в цьому випадку спотворення можуть бути помітними і залежати від змісту (структури) зображення.

Крім того, отримані результати дають змогу припустити, що поведінка СКП може корелювати з ентропією зображення, хоча ця кореляція може бути нестрогою.

Таким чином, залежність між ентропією зображення і СКП стиснутих даних має бути вивчена ретельніше.

4.1.2 Залежність між ентропією та отриманим СКП

Щоб встановити залежність між ентропією і СКП, набір зображень (що складається із зображень, наведених на рис. 2.2, 2.25 і 4.2), було стиснено для різних значень параметра контролю стиснення в діапазоні значень від 15 до 40 з кроком 5. Для кожного використаного зображення було розраховано значення ентропії, а також значення СКП для кожного значення Q . Отримані результати представлені у вигляді діаграм розсіювання, де горизонтальна вісь відноситься до ентропії, а вертикальна вісь відповідає СКП, та наведені на рисунку 4.3.

З отриманих залежностей на рис. 4.3 видно, що ентропія для більшості тестових зображень перевищує 7, а СКП для них зосереджена навколо певних значень: 0,6 для $Q=15$, 2,0 для $Q=20$ і 6,0 для $Q=25$. Однак є зображення, для яких ентропія менша за 7 (наприклад, gray21.512, kidney, mri_head), і для яких СКП менша, ніж для більшості інших зображень. Водночас для таких зображень, як Goldhill і Foot, ентропія яких менша за 6,5, СКП перебуває на тому ж рівні, що й для більшості інших. Можна також констатувати, що спотворення для всіх розглянутих зображень непомітні за $Q \leq 25$.

Якщо розглядати великі значення параметра контролю стиснення, то видно, що для більшості зображень значення СКП зосереджені в районі 16 для $Q=30$, 40 для $Q=35$ і 100 для $Q=40$. Це ясно показує, що для $Q \geq 30$ внесені спотворення можуть бути помітні, а для $Q \geq 35$ їх видно з великою ймовірністю. Винятком можуть бути зображення з малою ентропією, такі як штучні зображення (RSA і Gray21.512) і зображення з великими однорідними областями (Frisko, mri_head, kidney). Це показує, що ентропія може бути використана для виявлення «відмінних» зображень, для яких СКП, імовірно, менше, ніж для більшості інших за $Q \leq 30$, і для використання регресії для прогнозування СКП.

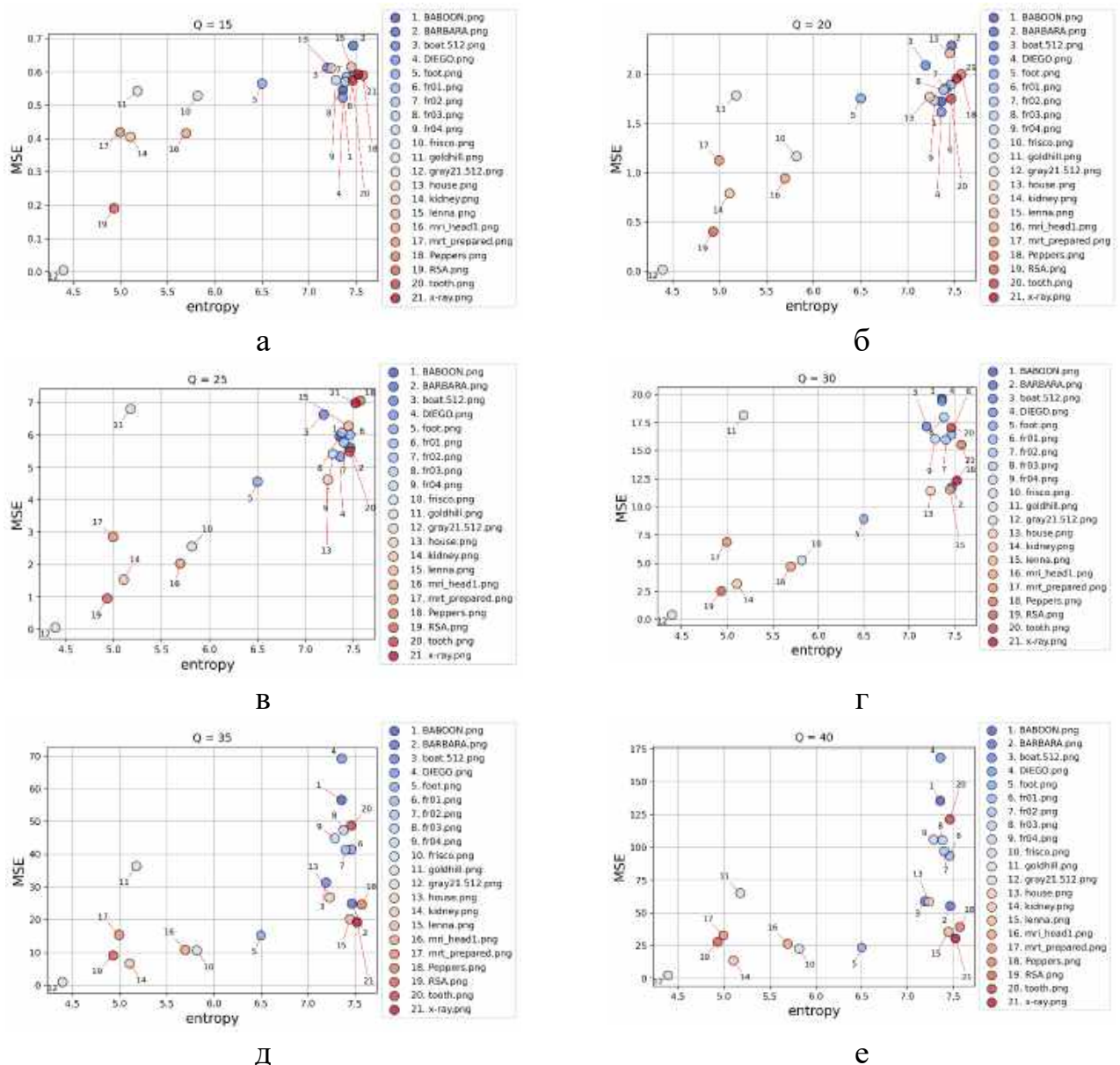


Рис 4.3 – Графіки залежності СКП від ентропії для $Q=15$ (а), 20 (б), 25 (в), 30 (г), 35 (д), і 40 (е)

Варто відзначити той факт, що значення СКП для зображень з ентропією, більшою за 7, концентруються близько певних значень за $Q \leq 30$, у той час як за $Q > 30$ вони можуть бути абсолютно різними. Наприклад, для $Q=40$ значення СКП може становити від 30 до 170. Це означає, що під час прогнозування СКП варто розглянути використання інших показників, крім Q та ентропії, для $Q > 30$.

Окрім цього, було розраховано коефіцієнти лінійної (Пірсона) і рангової (Спірмана) кореляції між ентропією і СКП, і, незважаючи на згаданий вище недолік, отримані значення кореляції можна охарактеризувати як задовільні. Для даних на

рисунку 4.3 вони становлять від 0,60 для $Q=40$ до 0,86 для $Q=15$. Коефіцієнти кореляції Спірмена трохи менші. Це означає, що існує суттєва кореляція між СКП та Q .

4.2 Прогнозування СКП на основі локальної активності в блоках зображення під час стиснення чорно-білих зображень із втратами за допомогою BPG

Раніше було сказано, що ентропія не єдиний метод чисельно відобразити складність зображення, а також було обговорено той факт, що не завжди можна використовувати тільки ентропію. У зв'язку з цим необхідно попередньо розглянути інші методи обчислення складності зображення, і як ці параметри можуть бути використані при прогнозуванні СКП.

Такими двома параметрами, що характеризують цільове зображення, яке підлягає стисненню, є два статистичні параметри [117], отримані за допомогою непересічних блоків розміром 8×8 пікселів.

$$\sigma_{im}^2 = \frac{1}{64} * \sum_{i=m}^{m+7} \sum_{j=n}^{n+7} (I(i,j) - \bar{I})^2, \quad (4.1)$$

де

$$\bar{I} = \frac{1}{64} * \sum_{i=m}^{m+7} \sum_{j=n}^{n+7} I(i,j)$$

де i та j - індекси пікселів вихідного зображення I , m і n позначають лівий верхній кут даного блоку.

Другий параметр характеризує спотворення, а формула для розрахунку їх локальної дисперсії має такий вигляд:

$$\sigma_{dist}^2 = \frac{1}{64} * \sum_{i=m}^{m+7} \sum_{j=n}^{n+7} (\Delta_p(i,j) - \bar{\Delta})^2, \quad (4.2)$$

де

$$\bar{\Delta} = \frac{1}{64} * \sum_{i=m}^{m+7} \sum_{j=n}^{n+7} \Delta_p(i, j)$$

де Δ_p - різницеве зображення, яке можна отримати, віднімаючи стиснене зображення від вихідного:

$$\Delta_p(i, j) = I(i, j) - I^c(i, j) + 128, i = 1, \dots, I_{im}, j = 1, \dots, J_{im}$$

Це різниця між двома зображеннями, або дельта-зображення, може бути використане для характеристики та візуалізації викривлень, що вносяться стисненням із втратами (детальніше про неї було розказано в другому розділі). Оскільки значення дельта-зображення можуть бути негативними та позитивними, тому у формулі вище використовується п'єдестал +128, щоб компенсувати негативні значення.

4.2.1 Опис запропонованого методу

Описані вище параметри можуть бути використані для опису зображення і можуть бути використані для прогнозування СКП. Для початку необхідно розрахувати їх для певного зображення. У даному випадку розраховуватись параметри будуть для двох зображень, використаних раніше, а саме Frisco і fr01.

Після розрахунку цих параметрів можна представити отримані дані у вигляді діаграм розсіювання, де кожна точка являє собою σ_{im} і σ_{dist} для горизонтальної та вертикальної вісей, відповідно. Ці діаграми показано на рисунку 4.4.

Спочатку проаналізуємо результати для зображення fr01 (рис. 4.4 а, б). Як видно, точки на діаграмі розсіювання загалом розташовані квазівипадково, середнє значення цих величин близьке до постійного значення за великих σ_{im} . Для $Q=20$ ця постійна величина становить близько 1,4 (рис. 4.4, а), тобто спотворення непомітні. Водночас для $Q=40$ воно становить близько 12, а це означає, що спотворення можуть бути помітними. Одна з тенденцій полягає в тому, що зі зростанням σ_{im} середнє значення σ_{dist} також збільшується, а потім насичується.

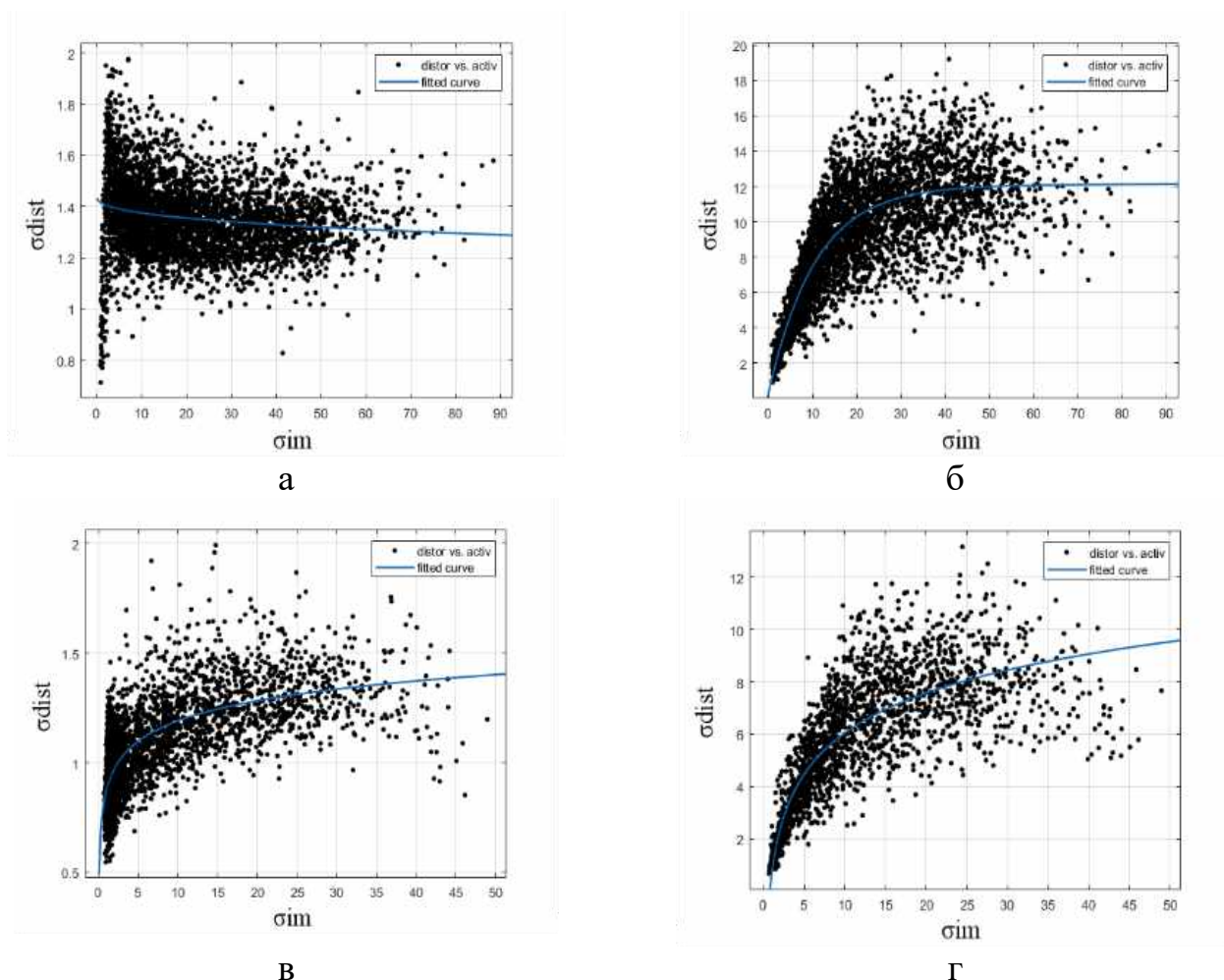


Рис 4.4 – Отримані діаграми розсіювання для зображення fr01 для $Q = 20$ (а) і $Q = 40$ (б) і для зображення Frisco для $Q = 20$ (в) і $Q = 40$ (г)

Для зображення Frisco (рис. 4.4 в, г) загальні тенденції аналогічні. Для $Q=20$ середнє значення σ_{dist} становить близько 1,4, принаймні за досить великих значень σ_{im} . У свою чергу, для $Q=40$ σ_{dist} збільшується, якщо σ_{im} стає більшим, і σ_{dist} становить близько 8 для $\sigma_{im} > 10$.

Спільний аналіз даних для двох зображень, показує, що σ_{dist} у блоках для зображення простої структури Frisco в середньому менший, ніж для зображення складної структури fr01. Це підтверджує припущення, що СКП спотворень, що вносяться, для того самого Q може бути меншим для зображень більш простої структури.

Для того, щоб підтвердити це твердження, необхідно отримати залежності СКП від Q для інтервалу значень від $Q=15$ (коли спотворення, що вносяться, незначні) до $Q=40$ (коли спотворення, що вносяться, починають вносити досить серйозні зміни). Цю залежність було отримано шляхом стиснення тестових зображень за допомогою заданого Q , декомпресії зображення і обчислення СКП між вихідним і стиснутим зображеннями. Отримавши такі дані, було створено діаграму розсіювання, наведену на рисунку 4.5, крім цього, в цю діаграму було вписано криву за допомогою відповідного інструменту в MATLAB, у результаті чого було отримано такий вираз:

$$f(x) = \text{СКП}(Q) = 0.0915 \times \exp(0.1692 \times Q) \quad (4.3)$$

Для отримання згаданої діаграми було використано 12 напівтонових зображень, що використовувалися раніше в розділах 2 і 3. Кожна точка діаграми розсіювання на рисунку 4.4 являє собою значення СКП і Q для тестового зображення, стисненого для певного Q .

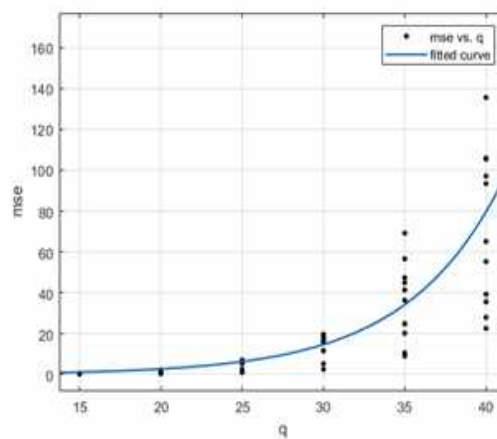


Рис 4.5 – Залежність СКП від Q

Виходячи з отриманої діаграми розсіювання (рис. 4.5) видно, що зі збільшенням Q значення СКП загалом також збільшуються, однак, стають менш компактними. Це особливо помітно для $Q=35$ і $Q=40$. Тоді, якщо використовувати вираз (4.3) для прогнозування СКП, помилка прогнозу може виявитися досить великою. Справжні значення СКП для конкретних зображень можуть бути в кілька разів більшими і

меншими за прогнозоване значення, особливо для $Q > 30$. Крім цього, це також підтверджується значеннями точності вписування: $R^2 = 0,6729$ і $RMSE = 19,7369$. Очевидно, що R^2 занадто малий і потребує поліпшення. Як результат стає зрозумілим той факт, що використання тільки формули (4.3) недостатньо для того, щоб здійснювати прогнозування СКП. Однак вона може допомогти у створенні (перерахунку значень) агрегованої діаграми розсіювання, за якою вже можна проводити прогнозування. Як підсумок цей нормалізований графік можна описати як залежність $((\ln(\sigma_{\text{dist}}/0.0915))/0.1692)/Q$ до σ_{im}/Q .

Для отримання цього агрегатного графіка будуть використовуватися ті самі два зображення fr01 і Frisco, для яких було отримано раніше діаграми розсіювання (представлені на рис. 4.4). Нову отриману діаграму розсіювання і вписану криву представлено на рис. 4.6.

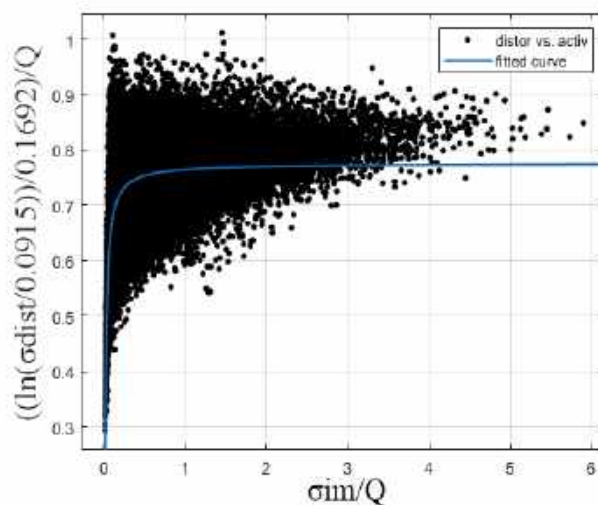


Рис 4.6 – Отримана агрегована діаграма розсіювання

Загальні властивості вписаної кривої полягають у тому, що вона швидко збільшується за σ_{im}/Q менше 0,5, а потім стає практично постійною. Вписана крива виражається як:

$$f(Q) = -0.0116 \times (\sigma_{\text{im}}/Q)^{-0.9785} + 0.7761 \quad (4.4)$$

Тепер, маючи вираз для залежності СКП від Q (4.3) і вписану криву для агрегованої діаграми розсіювання (4.4), можна спрогнозувати σ_{dist} для кожного блоку 8×8 пікселів для зображення, що підлягає стисненню, використовуючи відповідне значення σ_{im} , отримане для даного блоку:

$$\sigma_{\text{dist}} = \sqrt{(0.0915 \times \exp(0.1692 \times Q)) \times (-0.0116 \times (\sigma_{\text{im}}/Q)^{-0.9785} + 0.7761)} \quad (4.5)$$

Для всього зображення середнє квадратичне значення σ_{dist} , або СКП, може бути спрогнозовано таким чином:

$$\text{СКП} = [\sigma_{\text{dist}}^2] = [(\sqrt{(0.0915 \times \exp(0.1692 \times Q)) \times (-0.0116 \times (\sigma_{\text{im}}/Q)^{-0.9785} + 0.7761)})^2] \quad (4.6)$$

де $[]$ позначає середнє значення отриманих прогнозів для всіх розглянутих блоків.

4.2.2 Результати прогнозування

Маючи вищеописані вирази, стає можливим спрогнозувати значення і перевірити, чи збігаються спрогнозовані значення з реальними (розрахованими) значеннями. Прогнозовані значення були отримані за допомогою обумовленого вище методу, зокрема, за допомогою виразу (4.6), у якому як вхідні параметри використовується Q і σ_{im} , що стосується розрахованих значень, то вони були обчислені між вхідним зображенням і стисненим зображенням для відповідного значення Q . Дана пара значень СКП представлена в таблиці 4.1 для зображень Frisco і Fr01.

З наведених у таблиці даних видно, що для зображення Frisco прогнозовані значення є доволі точними, однак варто зазначити, що для більших значень Q дані відрізняються сильніше (приміром, для $Q=40$ різниця становить 5,75, в той час як для $Q=20$ така різниця становить 0,08), хоча всі вони є трохи більшими за відповідні істинні значення.

Таблиця 4.1 – результат прогнозування СКП

Назва зображення	СКП	Значення параметра контролю стиснення Q					
		15	20	25	30	35	40
Frisco	розрахований	0,53	1,17	2,57	5,27	10,62	22,56
	прогнозований	0,57	1,25	2,72	5,94	12,96	28,31
Fr01	розрахований	0,59	1,89	6,02	16,43	41,46	93,49
	прогнозований	0,67	1,53	3,52	8,08	18,56	42,65

Однак для зображення складної структури fr01 ситуація дещо інша, а саме: прогнозовані та істинні значення також доволі близькі для $Q=15$ і 20 , але за $Q \geq 25$ різниця між істинними та прогнозованими значеннями виявляється доволі великою. Цього разу прогнозовані значення менші за істинні.

Виходячи з викладеного вище, незважаючи на те, що було взято до уваги структуру зображення за допомогою σ_{im} в блоках, необхідно поліпшити прогнозування. У зв'язку з чим варто розглянути можливість використання або інших вхідних параметрів для опису зображення, або скористатися комбінацією з декількох таких параметрів.

4.3 Прогнозування на основі ентропії

Раніше, у цьому розділі, вже було згадано про те, що одним зі способів опису складності зображення є ентропія. У зв'язку з чим варто розглянути її використання як вхідного параметру для прогнозування СКП, і з'ясувати, чи дасть змогу використання цього параметра розв'язати описану вище проблему підвищення точності прогнозування.

4.3.1 Відмінність від оригінального методу

Оскільки початковий метод прогнозування проводитиметься з використанням формули (4.6), отриманої за допомогою вписування кривої в агрегатний графік, то необхідно внести зміни в цей крок. Замість використання цього будуть

використовуватися значення ентропії, отримані для того ж діапазону значень параметра контролю стиснення (від 15 до 40).

Отже, підхід до використання ентропії такий: необхідно замінити залежність СКП від Q у формулі (4.6) на залежність СКП як від ентропії, так і від Q . У результаті виникає завдання багато-параметричного вписування, у якому як вхідні параметри виступають ентропія і Q . Цю задачу вписування розв'язано за допомогою використання MATLAB, щодо функцій для вписування, то в цьому випадку нами використані поліноми. Варто зазначити, що для того, щоб покрити більшу кількість можливих варіантів значень ентропії, оригінальний набір зображень, який використовувався раніше, був розширений додатковими зображеннями з SIPI dataset [118], як підсумок, фінальний набір зображень містить 47 зображень різної складності різного походження.

Як уже згадувалося, для вписування використовувався поліноми, а саме поліном третього порядку як для Q , так і для ентропії, сам вираз має такий вигляд:

$$\begin{aligned} \text{СКП}(Q, E) = & 57.9488 + 6.5065Q - 51.0937E - 0.4792Q^2 + 0.8025QE + 7.4946E^2 \\ & + 0.0052Q^3 + 0.0328Q^2E - 0.1608QE^2 - 0.2677E^3 \end{aligned} \quad (4.7)$$

де E — це значення ентропії, розраховане для зображення, що підлягає стисненню. Отриману вписану площину представлено на рисунку 4.7.

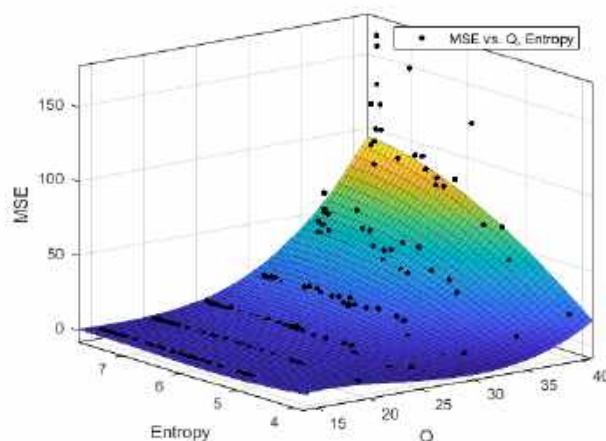


Рис 4.7 – Отримана вписана площина для вхідних параметрів: ентропія та Q

На перший погляд, точність вписування досить велика, крім цього це також підтверджують і показники якості вписування. Значення R^2 дорівнює 0,9308, а значення RMSE дорівнює 8,1252.

Вписування вже має кращий вигляд, ніж у разі прогнозування лише за параметром Q, однак варто так само розглянути результати прогнозування та порівняти їх як із розрахованими значеннями, так і з прогнозованими значеннями для попереднього підходу.

4.3.2 Результати та порівняння підходів прогнозування

Зробивши всі необхідні зміни до оригінального методу прогнозування, щоб мати можливість використовувати інформацію про значення ентропії зображення як вхідний параметр, стає можливим спрогнозувати та порівняти отримані значення з розрахованими значеннями. Отримані значення подано в таблиці 4.2.

Таблиця 4.2 – Порівняння реальних і прогнозованих значень СКП

Назва зображення	Q	Значення СКП	
		розрахований	прогнозований
frisco.png	25	2,57	2,59
BABOON.png	25	5,96	3,64
mri_head1.png	25	2,04	1,87
texture_1.1.08.png	25	6,43	3,12
frisco.png	30	5,27	5,87
BABOON.png	30	19,65	10,58
mri_head1.png	30	4,71	4,44
texture_1.1.08.png	30	10,81	7,46
frisco.png	35	10,62	12,14
BABOON.png	35	56,62	23,88
mri_head1.png	35	10,81	9,84
texture_1.1.08.png	35	18,3	16,36
frisco.png	40	22,56	22,18
BABOON.png	40	135,58	45,52
mri_head1.png	40	26,49	19,65
texture_1.1.08.png	40	33,9	31,64

Як видно з даних, представлених у таблиці 4.2, прогнозовані значення досить близькі до розрахованих при малих Q для всіх зображень. Однак зі збільшенням параметра контролю стиснення розбіжність між значеннями, для деяких зображень, стає більш помітною. Особливо це помітно для BABOON.png і mri_head1.png для значень $Q = 40$, однак варто зазначити, що настільки великі значення Q не часто використовуються на практиці через спотворення, що вносяться.

З результатів вище видно, що загалом прогнозовані значення доволі близькі до реальних (розрахованих) значень. Однак виглядає доцільним порівняти два підходи для того, щоб зрозуміти якому з підходів варто віддати перевагу.

Для такого порівняння було побудовано діаграми розсіювання, аналогічні тій, що була представлена раніше на рис. 4.5, проте в даному випадку замість розрахованих значень було використано спрогнозовані значення для двох підходів. Ці залежності представлено на рис. 4.8.

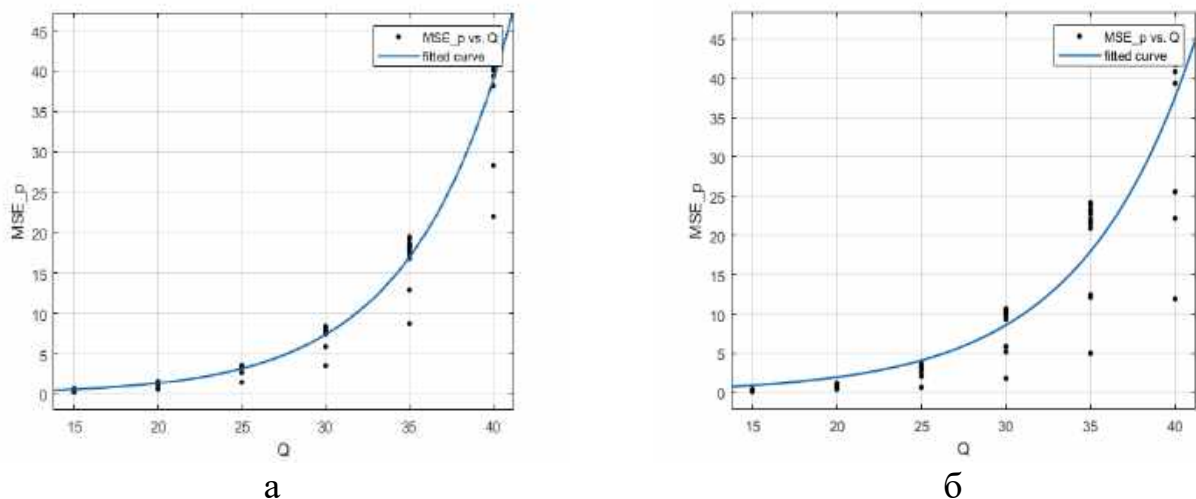


Рис 4.8 – Залежності СКП від Q , отримані за формулою (4.6) (а) і за формулою (4.7) для ентропійного підходу (б)

Для побудови діаграм розсіювання на рисунку 4.8 було використано ті ж самі 12 зображень, що й для отримання даних на рис. 4.5. Як видно, кожна залежність має схожий з отриманою раніше вигляд, але точки діаграми розсіювання на рис. 4.8, отримані за допомогою (4.6), компактніші, ніж оригінал. Це підтверджується і

точністю вписування: R^2 становить 0,9526, а RMSE дорівнює 3,0437. Водночас підхід, що ґрунтується на ентропії (рис. 4.8, б), має більше розкиданих точок, а щодо значень R^2 і RMSE то вони становлять 0,8672 і 5,2055 відповідно. У даному випадку розкид точок повторює той розкид, що можна було спостерігати для діаграми, які містять розраховані значення СКП.

Як підсумок, обидва підходи можна використовувати для оцінки того, як поводитиметься СКП для конкретного зображення зі збільшенням параметра контролю стиснення, проте прогнозовані значення за допомогою методу, заснованого на розрахунку ентропії, зазвичай ближчі до істинних (розрахованих) значень.

4.4 Прогнозування на основі більшої кількості параметрів

Реалізовані методи прогнозування з використанням двох параметрів (параметр, що характеризує зображення, та саме значення Q , з яким буде стискатися зображення) продемонстрували хороші результати, проте все ще існують ситуації, для яких можна розглянути можливість покращення точності прогнозування. Проблеми з точністю частково пов'язані з неправильною (неточною) класифікацією складності зображення, у зв'язку з чим варто розглянути варіант одночасного використання декількох параметрів для прогнозування. Далі буде розглянуто як самі параметри, так і варіанти їх використання, а саме: об'єднання даних параметрів за допомогою зваженої суми або використання простої нейронної мережі для самого прогнозування.

Для початку необхідно внести ясність, про які параметри буде йти мова. Уже не раз йшлося про великий вибір параметрів, які можуть характеризувати зображення з різних боків, проте для даного методу обрано такі: співвідношення країв, інформаційна ентропія, контраст, кореляція та енергія.

Перший параметр відношення країв може дати уявлення про два фактори: складність об'єкта на зображенні та кількість об'єктів на зображенні. Цей параметр може бути описаний як:

$$R = \frac{P_{edge}}{(I \times J)} \quad (4.8)$$

де P_{edge} - кількість пікселів, позначених як краї за допомогою будь-якого оператора детектора країв. У цьому випадку використовувався оператор Canny. I і J - кількість стовпців і рядків цільового зображення.

Варто також відзначити той факт, що на відміну від минулих використаних параметрів, решта входних параметрів розраховуються на основі матриці розбіжностей на рівні сірого (анг. Gray Level Co-occurrence Matrices - GLCM). Це матриця, яка визначається на зображенні як розподіл значень пікселів, що збігаються (градацій сірого або кольорів), за заданого зміщення. Елемент такої матриці позначається як $p(i,j)$.

Наступний параметр - енергія, яка характеризує рівномірність розподілу рівнів сірого в зображенні і може бути розрахована таким чином:

$$Enr = \sum_{i=1}^N \sum_{j=1}^N (p(i,j))^2 \quad (4.9)$$

де N - кількість рівнів сірого в зображенні.

Ще один параметр, який буде використовуватися, - кореляція, яку можна оцінити таким чином:

$$Cor = \frac{[\sum_{i=1}^N \sum_{j=1}^N ij * p(i,j) - \mu_x \mu_y]}{\sigma_x \sigma_y} \quad (4.10)$$

Цей параметр вимірює кореляцію між локальними рівнями сірого в зображенні.

Контрастність також є параметром, за яким можна судити про характеристики зображення, і буде використовуватися поряд з іншими. Цей параметр описує рівень чіткості зображення. Контрастність можна розрахувати таким чином:

$$C = \sum_{i=1}^N \sum_{j=1}^N (i-j)^2 p(i,j) \quad (4.10)$$

Останній параметр, що буде використовуватися, це вже використана раніше ентропія. Яка буде розраховуватися наступним чином:

$$Ent = - \sum_{i=1}^N \sum_{j=1}^N p(i,j) \log(p(i,j)) \quad (4.11)$$

Що стосується зображень, то буде використовуватися розширений набір зображень, який був задіяний нами під час аналізу прогнозування за допомогою ентропії.

4.4.1 Прогнозування з використанням зваженої суми

В основному процес прогнозування залишився тим самим. Перед процедурою стиснення для цільового зображення обчислюють кілька вхідних параметрів (наведених вище), значення яких вписують за допомогою деякої площини на основі еталонних даних, отриманих заздалегідь для набору різних зображень за певного значення параметра управління стисненням (він чинить найбільший вплив на отримані значення СКП, оскільки безпосередньо управляє ступенем стиснення). На основі цього вписування можна визначити значення СКП для використовуваного Q. У результаті ці знання допомагають встановити точне значення Q і стиснути зображення з необхідним рівнем якості.

Оскільки використовуватиметься велика кількість вхідних параметрів, необхідно якось розв'язати проблему їх вписування, як варіант можна використати зважену суму. Цей підхід використовувався в [119] і показав, що завдяки такому об'єднанню можна отримати один параметр, який зможе охарактеризувати зображення. Він являє собою зважену суму ентропії, відношення країв, контрасту,

кореляції та енергії, причому ваги перших трьох параметрів дорівнюють 1, а інших - 1. Новий параметр складності можна описати так:

$$Ws = Ent + R + C - Cor - Enr \quad (4.12)$$

Щоб мати можливість «порівняти» дійсні значення складності із зібраними даними, потрібно вписати поверхню в отримані дані. Для того щоб це зробити, як і раніше використовується MATLAB Fitting Tools. Результати вписування представлені на рис. 4.9.

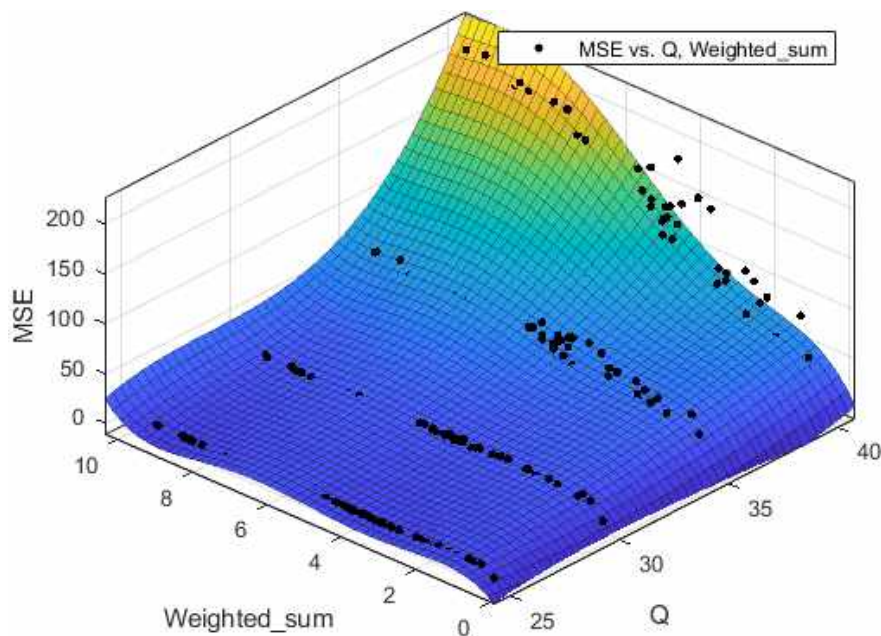


Рис 4.9 – Отримана вписана площина для вхідних параметрів: зважена сума та Q

Для отримання площини було використано вписування за допомогою поліномів і вираз має такий вигляд:

$$\begin{aligned} \text{СКП}(Q, Ws) = & 159.0726 - 9.0298*x - 136.8081*y + 0.0256*x^2 + 7.3241*x*y + \\ & 27.1570*y^2 + 0.0026*x^3 + 0.0413*x^2*y - 1.4895*x*y^2 - 2.2029*y^3 - \\ & 0.0025*x^3*y - 0.0486*x^2*y^2 + 0.4157*x*y^3 - 0.4167*y^4 + 0.0013*x^3*y^2 - \\ & 0.0046*x^2*y^3 - 0.0072*x*y^4 + 0.0239*y^5 \end{aligned} \quad (4.13)$$

Ґрунтуючись на даних, наведених на рис. 4.9, можна сказати, що результати вписані добре. Це важливо, оскільки чим точніша процедура вписування, тим точнішими будуть результати прогнозування. Для рис. 4.9 значення точності вписати такі: $R^2 = 0,8390$, а Adjusted R^2 дорівнює 0,8290; що стосується значень RMSE, то вони дорівнюють 19,9119.

Тепер, коли доступний вираз для СКП на Q і W_s (4.13), можна спрогнозувати значення СКП для обраного зображення, використовуючи описані вище параметри (4.8-4.11). Отримавши дані прогнозування, можливо провести порівняння прогнозованих значень і реальних (обчислених) значень СКП. Це порівняння представлено в таблиці 4.3.

Таблиця 4.3 – Порівняння реальних і прогнозованих значень СКП

Назва зображення	Q	Значення СКП	
		розрахований	прогнозований
aerial.png	25	5,39	8,14
BABOON.png	25	5,96	8,3
mri_head1.png	25	2,04	4,45
texture_1.1.08.png	25	6,43	4,98
aerial.png	30	14,81	17,31
BABOON.png	30	19,65	17,75
mri_head1.png	30	4,71	9,69
texture_1.1.08.png	30	10,81	9,54
aerial.png	35	38,62	37,16
BABOON.png	35	56,62	38,26
mri_head1.png	35	10,81	20,94
texture_1.1.08.png	35	18,3	17,32
aerial.png	40	90,06	74,41
BABOON.png	40	135,58	76,89
mri_head1.png	40	26,49	42,76
texture_1.1.08.png	40	33,9	28,77

Як видно, загалом прогнозовані значення СКП досить близькі до реальних. Також слід зазначити, що прогнозовані значення для Q , що дорівнює від 30 до 35, є трохи ближчими до розрахункових порівняно з іншими значеннями Q . Різниця між реальними і прогнозованими значеннями коливається від 0,97 до 50 (у випадку

BABOON.png з високим значенням $Q = 40$). Ці результати показують, що прогнозування СКП з такими параметрами не тільки можливе, а й може бути відносно точним.

4.4.2 Прогнозування з використанням простої нейронної мережі

Інший варіант використання такої кількості параметрів полягає не об'єднання їх в один, а використання нейронної мережі та подачі всіх параметрів, що характеризують зображення, на її вхід, таким самим вхідним параметром буде значення параметра контролю стиснення Q , що має найбільший вплив на одержуване в ході стиснення значення СКП.

Як і в попередньому випадку (прогнозування за допомогою вписування) використовується MATLAB, однак цього разу використовується модуль Neural Network. Як НМ-модель використовується двошарова мережа із сигмоїдальними прихованими нейронами та лінійними вихідними нейронами, що підходить для задач регресії. Загальну схему моделі наведено на рис. 4.10.

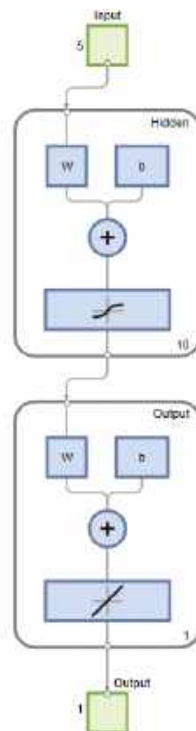


Рис 4.10 – Використовувана модель нейронної мережі для прогнозування СКП

Варто зазначити, що кількість входів моделі дорівнює кількості параметрів, що описують складність зображення. Сама кількість буде різнитися залежно від обраного набору вхідних параметрів, який може складатися з трьох, чотирьох або ж п'яти параметрів, щоб оцінити, чи збільшення кількості входів вплине на точність прогнозування. Точність прогнозування можна описати за допомогою діаграми розсіювання прогнозованих і реальних значень з апроксимуючими кривими, вписаними в них. Кореляція R між прогнозованими та істинними значеннями (в ідеалі дорівнює одиниці) також може характеризувати точність прогнозування.

Перед навчанням нейронних мереж зазвичай надають характеристики набору даних, для яких було підготовлено навчальні, тестові та перевіірочні дані. Для цього методу прогнозування було використано 73 зображення (ті що використовувалися раніше, а також частина текстурних зображень із SIPI dataset) і обчислено значення складності для всіх значень Q , а отримані дані поділено на три групи з таким відсотковим співвідношенням: навчальні дані: 70 %, дані перевірки: 15%, тестові дані: 15 %, а розмір шару: 10.

Почнемо з першого випадку з трьома вхідними параметрами: Q , а також параметрами, що характеризують зображення: відношення країв і кореляція. Результат представлено на рис. 4.11.

Перед тим, як почати аналізувати результати, необхідно зазначити, що горизонтальна вісь - це фактичне значення, а вертикальна вісь - вихід з нейронної мережі (прогнозоване значення). Також поруч із вертикальною віссю знаходиться функція для вписування, яка є лінійною. Над діаграмою розсіювання вказано дані, для яких її було отримано, а також значення коефіцієнта кореляції.

Повертаючись до отриманих даних на рисунку 4.11, можна зазначити, що є кілька спостережень. Перше: точки на діаграмі розсіювання розташовані з високою щільністю, особливо для діапазонів від 0 до 200. Друге спостереження - залежність між прогнозованим і розрахованими значеннями СКП майже лінійна. І останнє - значення для тестового набору близькі до тренувальних. Значення точності для тренувальних такі: R дорівнює 0,974, а для тестового набору - 0,972. Усі ці

спостереження дають змогу зробити висновок про те, що існує можливість робити прогнози з високою точністю.

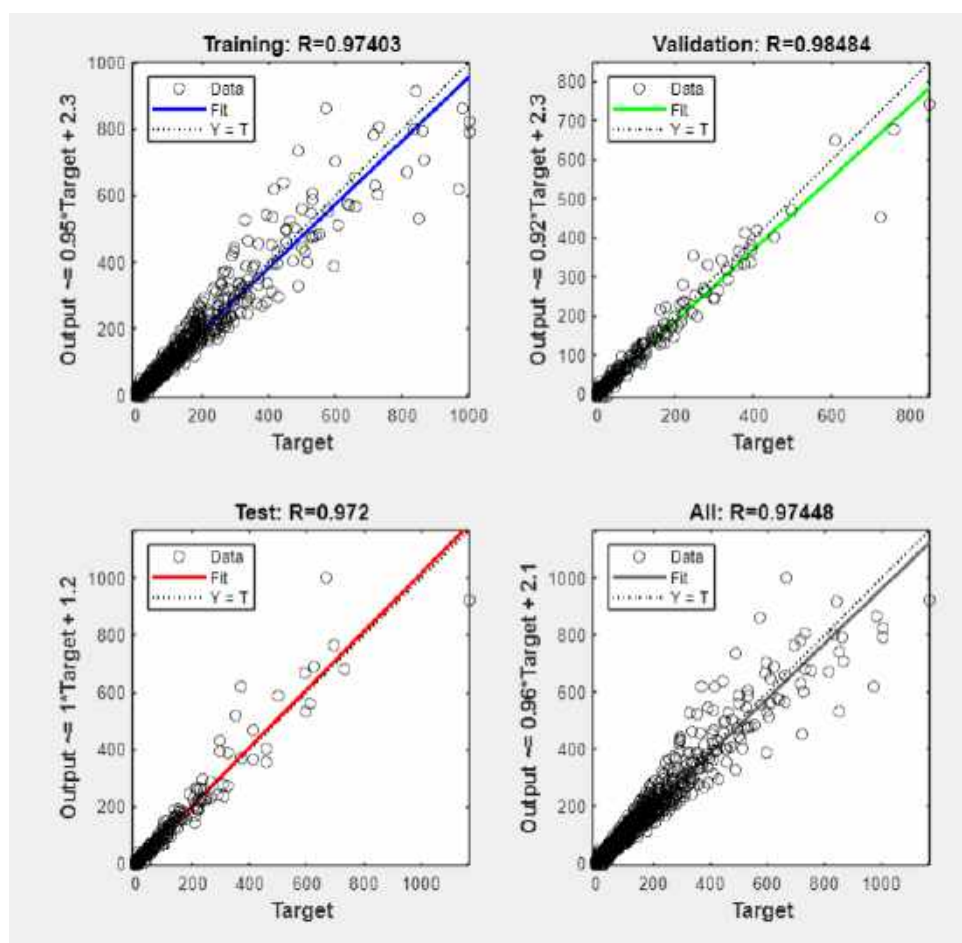


Рис 4.11 – Результати навчання і тестування нейронної мережі з трьома вхідними параметрами: Q , відношення країв і кореляція

Отримані раніше дані показують позитивні результати, однак, варто розглянути решту вхідних параметрів, щоб встановити можливість підвищення точності прогнозування за рахунок використання іншої комбінації вхідних параметрів. Виходячи з цього далі розглянуті такі параметри як Q , контраст і енергія, подані на вхід тієї самої нейронної мережі (модель якої подано на рисунку 4.10). Результати використання таких параметрів для навчання НМ подано на рис. 4.12.

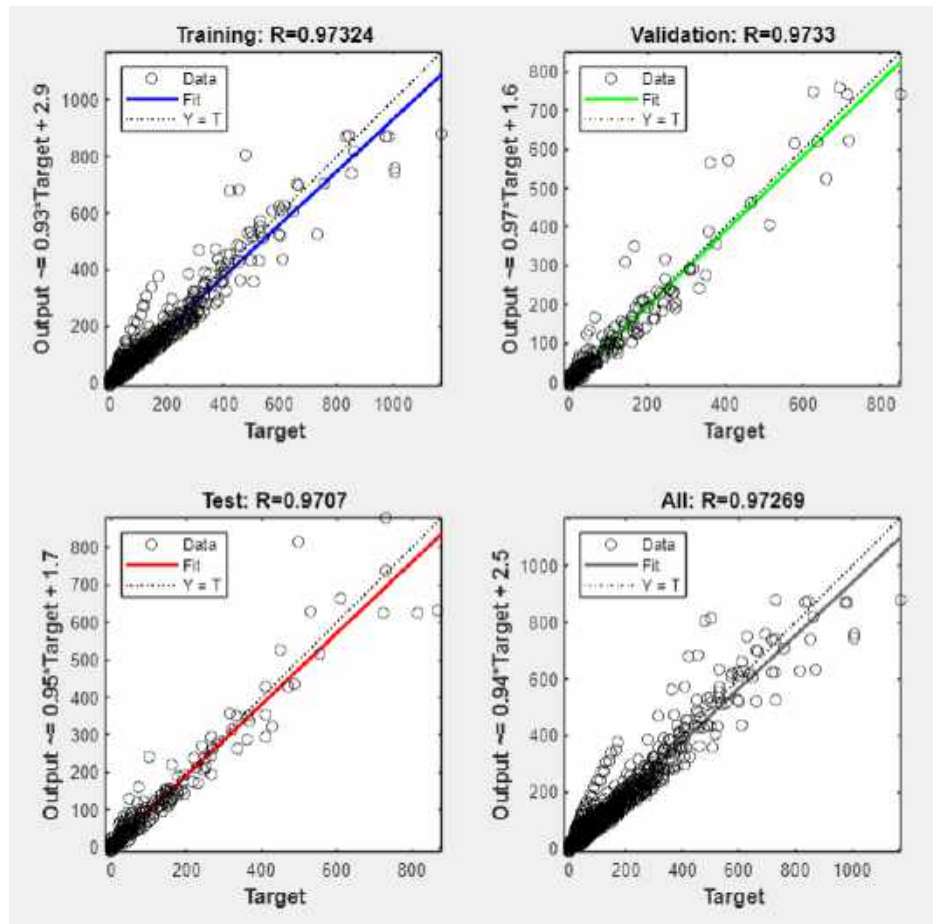


Рис 4.12 – Результати навчання і тестування нейронної мережі з трьома вхідними параметрами: Q, контраст, енергія

З рисунка 4.12 видно, що за великих значень СКП значення стають щільнішими (тобто прогнозовані значення стають ближчими до фактичних). З іншого боку, за малих значень СКП спостерігалось відхилення від ідеальної діагоналі. Значення R стають трохи меншими порівняно з попереднім набором вхідних параметрів. Прикладом є те, що для тестового набору даних для нових параметрів R дорівнює 0,971, тоді як для попередніх параметрів це значення дорівнює 0,972.

Грунтуючись на отриманих результатах видно, що зміна комбінації вхідних параметрів не дала позитивного результату (у питанні підвищення точності прогнозування). У зв'язку з чим варто розглянути можливість використання більшої кількості вхідних параметрів. Так, результати тестування та навчання чотирьох параметрів подано на рис. 4.13.

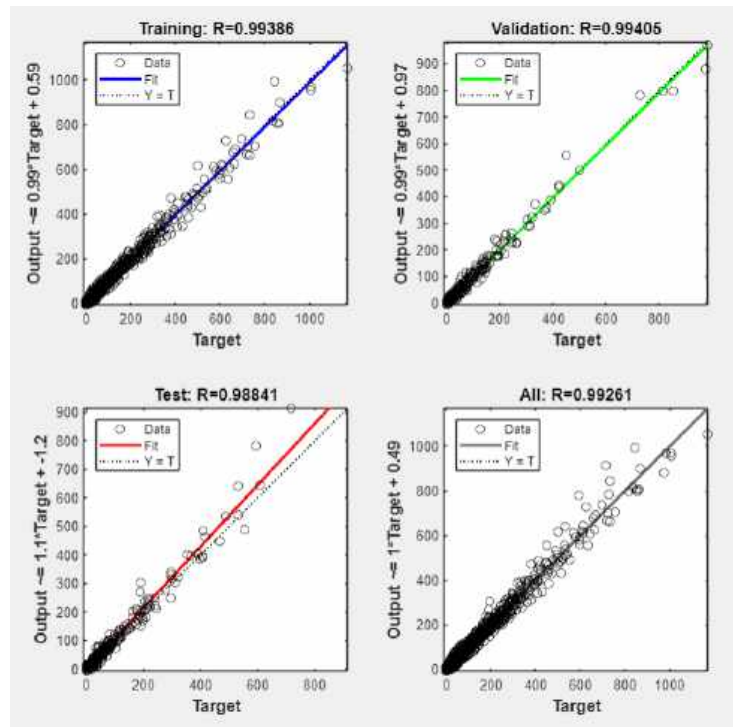


Рис 4.13 – Результати навчання і тестування нейронної мережі з чотирма вхідними параметрами: Q, відношення країв, кореляція, контраст

На перший погляд результати, представлені на рисунку 4.13, демонструють поліпшення точності прогнозування. Точки на діаграмі розсіювання розташовані ближче до ідеальної лінії, і в результаті прогнозовані значення стають ближчими до реальних. Це підтверджується і значеннями кореляції: для чотирьох параметрів. Значення R для набору, що включає всі групи зображень (навчання, підтвердження і тестування) дорівнюють 0,993, порівняно з 0,974 або 973 отриманих для попередніх двох варіацій прогнозування. З цього можна зробити висновок, що збільшення кількості вхідних параметрів, які характеризують складність зображення, відіграє позитивну роль у підвищенні точності прогнозування.

Очевидним наступним кроком для розгляду є випадок, у якому використовується чотири параметри характеристики зображення (які також використовувалися в підході з використанням зваженої суми) і параметра контролю стиснення Q. Результати представлено на рис. 4.14.

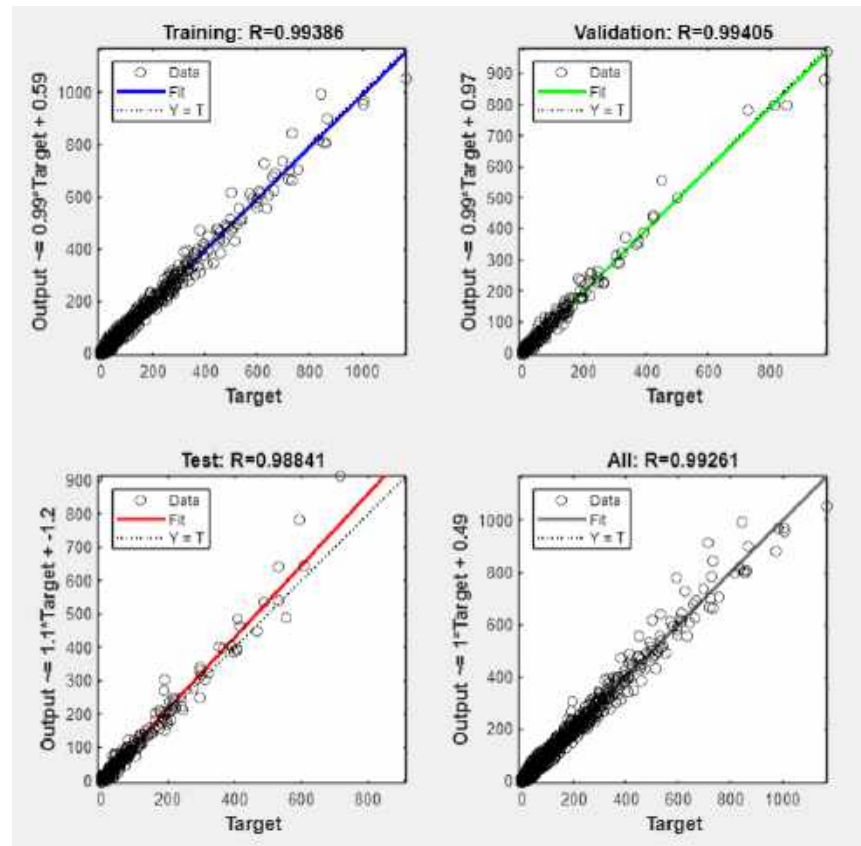


Рис 4.14 – Результати навчання і тестування нейронної мережі з п'ятьма вхідними параметрами: Q, крайовий коефіцієнт, кореляція, контраст, енергія

Отримані залежності (рис. 4.14) стали трохи точнішими, але збільшення точності не настільки велике (утім, це пов'язано з тим, що в попередньому випадку, у разі використання чотирьох параметрів, уже було досягнуто високих значень точності). Значення R дорівнюють 0,9932 для п'яти вхідних параметрів і 0,9926 для чотирьох параметрів.

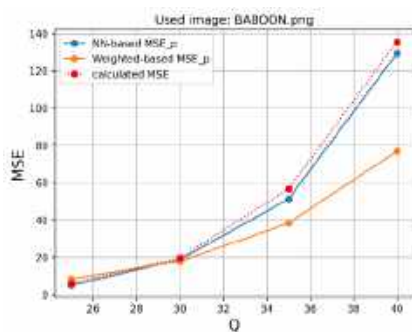
Далі необхідно трохи детальніше поглянути на отримані дані та порівняти їх із розрахованими заздалегідь значеннями. Отримані результати подано в таблиці 4.4.

Як і в попередніх підходах, значення СКП (прогнозованого і розрахованого) близькі одне до одного. Але очевидно, що для таких зображень, як baboon і mre_head1, точність прогнозування збільшилася, особливо для діапазону від 30 до 35. Значення для $Q = 40$ також стають більш точними, але ці значення Q вносять більше викривлень під час процедури стиснення і призводять до того, що на практиці їх застосовують рідше, але все ж таки можна використовувати в деяких додатках.

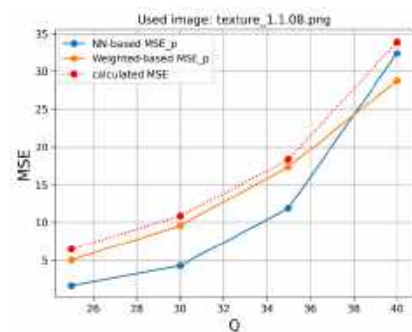
Таблиця 4.4 – Порівняння реальних і прогнозованих значень СКП

Назва зображення	Q	Значення СКП	
		розрахований	прогнозований
aerial.png	25	5,39	7,59
BABOON.png	25	5,96	4,89
mri_head1.png	25	2,04	0,42
texture_1.1.08.png	25	6,43	1,64
aerial.png	30	14,81	19,28
BABOON.png	30	19,65	18,68
mri_head1.png	30	4,71	5,77
texture_1.1.08.png	30	10,81	4,25
aerial.png	35	38,61	37,84
BABOON.png	35	56,62	51,2
mri_head1.png	35	10,81	10,29
texture_1.1.08.png	35	18,3	11,84
aerial.png	40	90,06	84,36
BABOON.png	40	135,58	129,44
mri_head1.png	40	26,49	35,35
texture_1.1.08.png	40	33,9	32,41

Якщо ж порівнювати отримані значення з методом прогнозування з використанням зваженої суми, то отримані значення можна представити у вигляді залежностей СКП від Q, в якому кожна з кривих належатиме або до істинних (розрахованих), або до прогнозованих значень за одним із методів. Чим ближче спрогнозована крива до істинної, тим точніше виконано прогнозування. Це порівняння представлено на рис. 4.15.



а



б

Рис 4.15 – Залежності прогнозованих і розрахункових значень СКП від Q для зображень BABOON і texture_1.1.08

Як видно з першого випадку (рис. 4.15, а), значення на основі нейронної мережі дуже близькі до реальних значень СКП, що означає, що цей тип прогнозування може бути використано в підходах, які потребують високої точності. Водночас для другого випадку (рис. 4.15, б) зважені суми дають кращі результати з позиції точності прогнозів СКП. Однак варто визнати, що в обох випадках підходи до прогнозування можуть надати загальну інформацію про поведінку зображення при збільшенні значення Q .

Отримані результати дають змогу зробити висновок, що нейронні мережі справді можуть бути використані для прогнозування СКП. Більше того, для досягнення високих значень точності прогнозування достатньо чотирьох або максимум п'яти входів.

4.5 Висновки за розділом

У цьому розділі за отриманими результатами можна зробити такі висновки:

- Продемонстровано залежність між складністю зображення й отриманим СКП, що стало передумовою до прогнозування. Незважаючи на те, що для всіх зображень спостерігалася монотонна поведінка (збільшення СКП зі зростанням Q), залежно від складності зображення, особливо для $Q > 30$, спостерігалася значна різниця в якості зображення. Також було показано, що не варто прогнозувати якість зображення для $Q \leq 20$, коли спотворення, що вносяться, незначні та візуально непомітні. Крім того, доведено, що ентропія досить надійна для використання як одного з параметрів класифікації складності зображення, необхідного для прогнозування СКП під час стиснення з втратами кодером BPG.

- Показано можливість використання оцінок локальної активності в блоках для прогнозування СКП. Показано, що для зображень із складнішою структурою істинні та прогнозовані значення СКП більші, ніж для зображень із простішою структурою. Крім цього, було розглянуто можливість використання ентропії зображення для класифікації складності зображення і, як результат, одним із вхідних параметрів для

подальшого прогнозування. Запропоновано підхід до прогнозування, і виконано порівняння двох підходів.

– Виконано аналіз можливості багато-параметричного прогнозування СКП для BPG кодера. Показано, що використання параметрів опису складності, таких як відношення країв, кореляція, контраст, енергія та ентропія, відкриває додаткові можливості для класифікації різних зображень. Спільне використання цих параметрів за допомогою зваженої суми дає змогу досягти високої точності вписування і, як наслідок, високої точності прогнозування.

– Також було розглянуто можливість використання нейронної мережі для розв'язання задачі прогнозування СКП. Результати показують, що навіть використання простої НМ із трьома вхідними параметрами дає змогу досягти високих значень точності. Для досягнення максимальної точності доцільно використовувати більшу кількість вхідних параметрів: відношення країв, кореляція, контраст, енергія та Q (основний параметр, що впливає на значення СКП). Обидва підходи показують близькі результати, але для деяких зображень підходи на основі нейронної мережі досягають кращих результатів з точки зору точності прогнозованих значень.

ВИСНОВКИ

У дисертаційному дослідженні розв'язано низку задач, пов'язаних з розробкою ефективних методів стиснення з втратами та надання можливості прогнозувати параметрів якості зображення. Зокрема, проаналізовані особливості та характеристики BPG кодера, що стало підставою до розробки методів прогнозування характеристик як самого стиснення з втратами, так і спотворень під час стиснення. В результаті була вирішена актуальна задача ефективного стиснення зображень ДЗЗ.

За результатами дисертаційної роботи можна зробити наступні висновки:

1. Вперше досліджено характеристики BPG кодера під час роботи в різних режимах, зокрема стиснення напівтонових, кольорових та триканальних зображень. Розглянуто випадки стиснення даних зображень в умовах, коли вхідне зображення спотворено шумом різних типів, зокрема адитивним або сигнально-залежним. Підтверджено можливе існування оптимальної робочої точки для цього кодера, а також позитивного результату під час стиснення з втратами зображень у ній або в її околі для обох типів завад. Ґрунтуючись на отриманих результатах, надано рекомендації щодо вибору необхідного ПКС для досягнення цієї оптимальної робочої точки. Крім цього, проведено аналіз спотворень, що вносить BPG кодер в результаті стиснення для випадків присутності та відсутності шуму.

2. Запропоновано метод та швидкий алгоритм прогнозування існування оптимальної робочої точки для BPG кодера, що ґрунтується на заздалегідь отриманих апроксимуючих залежностях. Також запропоновано модифікацію даного методу для випадків сигнально-залежного стиснення кольорових зображень. Можливість прогнозування продемонстрована для кількох сучасних метрик візуальної якості. Приблизно на 5% підвищено точність прогнозування завдяки розрахунку та спільній обробці кількох вхідних параметрів, що характеризують складність зображення та рівень завад.

3. Запропоновано та проаналізовано метод прогнозування СКП внаслідок стиснення з втратами, що базується на оцінці локальної активності в блоках. Крім того, розглянуто можливість використання ентропії як альтернативного вхідного

параметра. Також запропоновано модифікацію методу з обробкою кількох вхідних параметрів (таких як: відношення країв, кореляція, контраст, енергія) нейромережею, що дозволяє забезпечити прогнозування СКП з відносною похибкою не більше 27%.

4. Доведено, що запропоновані методи прогнозування та стиснення є універсальними та можуть бути використані при обробці зображень дистанційного зондування, медичної діагностики та звичайних зображень. Прогнозування зазвичай вимагає часу, що в 5,1 разів менше, ніж стиснення.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Kerekes, J. P. Optical Sensor Technology [Text] / J. P. Kerekes – The SAGE Handbook of Remote Sensing. – London, UK : SAGE Publications. – 2009. – P. 95 – 107. – doi: <https://doi.org/10.4135/9780857021052.N7>.
2. Khorram, S. Future Trends in Remote Sensing [Text] / S. Khorram, C.F. van der Wiele, F.H. Koch, S.A.C. Nelson, M.D. Pott // In Principles of Applied Remote Sensing; Springer International Publishing. – 2016. – P. 277–285. – doi: https://doi.org/10.1007/978-3-319-22560-9_9.
3. Marcuccio, S. Smaller Satellites, Larger Constellations: Trends and Design Issues for Earth Observation Systems [Text] / S. Marcuccio, S. Ullo, M. Carminati, O. Kanoun // IEEE Aerosp. Electron. Syst. Mag. – 2019. – Vol. 34. – P. 50–59. – doi: <https://doi.org/10.1109/MAES.2019.2928612>.
4. Prince, Jerry L. Medical Imaging Signals and Systems [Text] / Jerry L. Prince, Jonathan Links // Pearson Cloth. – 2005. – 496 pp. – ISBN: 978-0-13-214518-3
5. Eikelboom, R. H. Methods and Limits of Digital Image Compression of Retinal Images for Telemedicine [Text] / R. H. Eikelboom, K. Yogesan; C.J. Barry, I.J. Constable, M.–L. Tay–Kearney, L. Jitskaia, P.H. House// Investigative Ophthalmology & Visual Science. – 2000. – Vol. 41. – P. 1916-1924. – PMID: 10845617.
6. Sheikh, H. R. Image information and visual quality [Text] /H. R. Sheikh, A. C. Bovik// IEEE Trans. on Image Processing. – 2006. – Vol. 15, No. 2, P. 430-444. – doi: <https://doi.org/10.1109/TIP.2005.859378>.
7. Kozhemiakina, N. Means and results of efficiency analysis for data compression methods applied to typical multimedia data [Text] / N. Kozhemiakina, N. Ponomarenko, V. Lukin, K. Egiazarian, J. Astola // IEEE First International Scientific-Practical Conference Problems. – 2014. – P. 12-14. – doi: <https://doi.org/10.1109/INFOCOMMST.2014.6992281>.
8. Kovalenko, B. BPG-Based Lossy Compression of Three-Channel Noisy Images with Prediction of Optimal Operation Existence and Its Parameters [Electronic resource] / B. Kovalenko, V. Lukin, B. Vozel // Remote Sensing. – 2023. – Vol. 15, no. 6.

– Article No. 1669. – doi: <https://doi.org/10.3390/rs15061669>.

9. Naumenko, V. BPG-based compression analysis of Poisson-noisy medical images [Electronic resource] / V. Naumenko, B. Kovalenko, V. Lukin, // Radioelectronic and Computer Systems. – 2023. – No. 3. – P. 91-100. – doi: <https://doi.org/10.32620/reks.2023.3.08>.

10. Kovalenko, B. Usage of different Chroma Subsampling Modes in Image Compression by BPG Coder [Text] / B. Kovalenko, V. Lukin // Ukrainian journal of remote sensing. – 2022. – Vol. 9, no. 3. – P. 11-61. – doi: <https://doi.org/10.36023/ujrs.2022.9.3.216>.

11. Lukin, V. Compression of Noisy Images Taking into Account Visual Quality: A Comprehensive Study [Text] / V. Lukin, B. Kovalenko, S. Kryvenko, N. Ponomarenko, K. Egiazarian, J. Astola // Research Developments in Science and Technology. – 2022. – Vol. 8. – P. 89–109. – doi: <https://doi.org/10.9734/bpi/rdst/v8/2722B>.

12. Lukin, V. Prediction of Optimal Operation Point Existence and Its Parameters in BPG-Based Automatic Lossy Compression of Noisy Images [Text] / V. Lukin, B. Kovalenko, S. Kryvenko, V. Naumenko, B. Vozel // Current Overview on Science and Technology Research. – 2022. – Vol. 9. – P. 1-36. – doi: <https://doi.org/10.9734/bpi/costr/v9/4316A>.

13. Pillai, D. K. New Computational Models for Image Remote Sensing [Text] / D. K. Pillai // Big Data Analytics for Satellite Image Processing and Remote Sensing. – 2018. – P. 1-21. – doi: <https://doi.org/10.4018/978-1-5225-3643-7.ch001>.

14. Plataniotis K.N. Color Image Processing and Applications [Text] / K.N. Plataniotis, A.N. Venetsanopoulos // Springer-Verlag, NY. – 2000. – doi: <https://doi.org/10.1007/978-3-662-04186-4>.

15. Kharchenko, V. Concepts of green IT engineering: taxonomy, principles and implementation [Text] / V. Kharchenko, O. Illiashenko // Green IT Engineering: Concepts, Models, Complex Systems Architectures. – 2017. – Springer, Cham. – P. 3-19. – doi: https://doi.org/10.1007/978-3-319-44162-7_1.

16. Zheng, C. Qi A New Image Pre-processing for Improved Performance of Entropy Coding [Text] / Zheng, C. Qi, G. Wang // Proceedings of CCPR, China. – 2010. –

P. 1-6. – doi: <https://doi.org/10.1109/CCPR.2010.5659320>.

17. Lukin, V. Automation in remote sensing data pre-processing [Electronic resource] / V. Lukin, M. Uss, S. Abramov, I. Vasilyeva, G. Proskura, O. Ieremeiev, V. Abramova, O. Rubel, N. Kozhemiakina, V. Naumenko // Space Research in Ukraine. – Report to COSPAR, Periodyka, Kiev. – 2021. – P. 96-103. – Режим доступу: https://cosparhq.cnes.fr/assets/uploads/2021/02/Ukraine_2018-2020_compressed.pdf (дата звернення: 30.01.2025).

18. Soni, A. K. Green Computing in Image Processing: An Assessment [Text] / A. K. Soni, M. Dixit // International Journal . – May 2017. – Vol. 7, No. 5. – P. 523-527. – doi: <https://doi.org/10.23956/ijarcsse/SV7I5/0181>.

19. Pratt, W.K. Digital Image Processing. Fourth Edition [Text] / W.K. Pratt. – NY.: Wiley-Interscience. – USA. – 2007. – 1429 p. – doi: <https://doi.org/10.1117/1.2744044>.

20. Dougherty, G. Digital Image Processing for Medical Applications [Text] / G. Dougherty – Cambridge University Press. – 2009. – ISBN: 978-0-521-86085-7.

21. Milanfar, P. A Tour of Modern Image Filtering: New Insights and Methods, Both Practical and Theoretical [Text] / P. Milanfar. – IEEE Signal Processing Magazine. – 2013. – Vol. 30. – P. 106-128. – doi: <https://doi.org/10.1109/MSP.2011.2179329>.

22. Lukin, V.V. Processing and classification of multichannel remote sensing data [Text] / V.V. Lukin, N.N. Ponomarenko, A.A. Kurekin, O.B. Pogrebnyak // Proceedings of MICAI, Mexico. 2011. – Vol. 7095. – P. 487-498. – doi: https://doi.org/10.1007/978-3-642-25330-0_43.

23. Blau, Y. Rethinking Lossy Compression: The Rate-Distortion-Perception Tradeoff [Text] / Y. Blau, T. Michaeli // Proceedings of “the 36th International Conference on Machine Learning, PMLR 97. – May 2019. – P. 675-685. – doi: <https://doi.org/10.48550/arXiv.1901.07821>.

24. Lukin, V.V. Image filtering: Potential efficiency and current problems [Text] / V.V. Lukin, S.K. Abramov, N.N. Ponomarenko, K. Egiazarian, and J. Astola // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). – 2011. –

P.1433-1436. – doi: <https://doi.org/10.1109/ICASSP.2011.5946683>.

25. Oktem, R. Locally adaptive DCT filtering for signal-dependent noise removal [Text] / R. Oktem, K. Egiazarian, V.V. Lukin, N.N. Ponomarenko, O.V. Tsymbal // EURASIP Journal on Advances in Signal Processing, – 2007. – Vol. 2007. – 10 p. – doi: <https://doi.org/10.1155/2007/42472>.

26. Lukin V. Processing multichannel radar images by modified vector sigma filter for edge detection enhancement [Text] / V. Lukin, O. Tsymbal, B. Vozel, K. Chehdi // Proceedings of ICASSP. – 2006. – Vol. II. – P. 833-836. – doi: <https://doi.org/10.1109/ICASSP.2006.1660472>.

27. Marques, R. Target Detection in SAR Images Based on a Level Set Approach [Text] / R. Marques, F. Medeiros, D. Ushizima // IEEE Trans. on Systems, Man and Cybernetics. – 2009. – Vol. 39, No. 2. – P. 214 – 222. – Режим доступа: <https://escholarship.org/uc/item/6nq7157n> (дата звернення: 30.01.2025).

28. Kang, X. SAR image edge detection by ratio-based Harris Method [Text] / X. Kang, C. Han, Y. Yang, T. Tao // ICASSP 2006 Proceedings. – May 2006. – Vol. 2. – P. 837 -840. – doi: <https://doi.org/10.1109/ICASSP.2006.1660473>.

29. Naumenko, A. Texture detection in noisy images by combining several local parameters [Text] / A. Naumenko, S. Krivenko, N. Ponomarenko, A. Zelensky, V. Lukin // Problems of Infocommunications Science and Technology (PIC S&T). – Kharkiv, Ukraine. – 13-15 Oct. 2015. – P. 230 – 233. – doi: <https://doi.org/10.1109/INFOCOMMST.2015.7357321>.

30. Kussul, N. Parcel-Based Crop Classification in Ukraine Using Landsat-8 Data and Sentinel-1A Data [Text] / N. Kussul, G. Lemoine, F. J. Gallego, S. V. Skakun, M. Lavreniuk, A. Y. Shelestov // IEEE J. of Select. Topics in Appl. Earth Observ. and Rem. Sens. – 2016. – Vol. 9(6). – P. 2500–2508. – doi: <https://doi.org/10.1109/JSTARS.2016.2560141>.

31. Hussain, A. J. Image compression techniques: A survey in lossless and lossy algorithms [Text] / A. J. Hussain, A. Al-Fayadh, N. Radi // Neurocomputing. – 2018. – Vol. 300. – P. 44-69. – doi: <https://doi.org/10.1016/j.neucom.2018.02.094>.

32. Christophe, E. Hyperspectral Data Compression Tradeoff [Text] / E. Christophe – Optical Remote Sensing. Berlin: Springer. – 2011. – Vol. 8. – P. 9-29. – doi: https://doi.org/10.1007/978-3-642-14212-3_2.
33. Naït-Ali, A. Standards in Medical Image Compression [Text] / A. Naït-Ali, C. Cavarro-Ménard, B. Gibaud, J. Chabriaïs // Compression of Biomedical Images and Signals. – 2010. – doi: <https://doi.org/10.1002/9780470611159.ch4>.
34. Chang, S. G. Image denoising via lossy compression and wavelet thresholding [Text] / S. G. Chang, B. Yu, M. Vetterli // Proceedings of International Conference on Image Processing, Santa Barbara. – 1997. – Vol. 1. – P. 604-607. – doi: <https://doi.org/10.1109/ICIP.1997.647985>.
35. Zabala, A. Impact of lossy compression on mapping crop areas from remote sensing [Text] / A. Zabala, X. Pons // International Journal of Remote Sensing. – 2013. – Vol. 34, No. 8. – P. 2796-2813. – doi: <https://doi.org/10.1080/01431161.2012.750772>.
36. Ozah, N. Compression improves image classification accuracy [Text] / N. Ozah, A. Kolokolova // In: Canadian Conference on Artificial Intelligence. – 2019. – Springer, Heidelberg. – P. 525-530. – doi: https://doi.org/10.1007/978-3-030-18305-9_55.
37. Lam, K. W. The effects on image classification using image compression technique [Text] / K. W. Lam, W. Lau, Z. Li // Amsterdam, International Archives of Photogrammetry and Remote Sensing. – 2000. – Vol. 33 (Part B7). – P. 744–751. – doi: <https://doi.org/10.1080/01431160210142842>.
38. Баранник, В.В. Информационная технология сжатия изображений на основе двумерного плавающего полиадического кодирования трансформант Уолша [Текст] / В.В. Баранник, А.В. Яковенко // Системи обробки інформації. – 2008. – Вип. 3. – С. 13-17. – Режим доступу: http://nti.khai.edu/csp/nauchportal/Arhiv/OIKIT/2008/OIKIT39/p_278-282.pdf (дата звернення: 30.01.2025).
39. Бараник, В. В. Метод квадратурного стиснення трансформант вейвлет-перетворення у двовимірному поліадичному просторі [Електронний ресурс] / В. В. Бараник, А. В. Ширяев // Сучасна спеціальна техніка. – 2011. – № 2(25). – С. 73-80. –

Режим доступу: <http://elar.naiau.kiev.ua/jspui/handle/123456789/2208> (дата звернення: 30.01.2025).

40. Lukin, V. Lossy compression of images without visible distortions and its applications [Text] / V. Lukin, M. Zriakhov, S. Krivenko, N. Ponomarenko, Z. Miao // IEEE 10th International Conference on Signal Processing Proceedings. – 2010. – Beijing. – P. 694-697. – doi: <https://doi.org/10.1109/ICOSP.2010.5655751>.

41. Li, F. A Two-Step Procedure for Image Lossy Compression by ADCTC with a Desired Quality [Text] / F. Li, S. Krivenko, V. Lukin // IEEE 11th International Conference on Dependable Systems, Services and Technologies (DESSERT). – Ukraine, May 2020. – P. 307-312. – doi: <https://doi.org/10.1109/DESSERT50317.2020.9125000>.

42. Li, F. A Two-step Approach to Providing a Desired Visual Quality in Image Lossy Compression [Text] / F. Li, S.S. Krivenko, V.V. Lukin// IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET). – 2020. – 5 p. – doi: <https://doi.org/10.1109/TCSET49122.2020.235483>.

43. Kozhemiakin, R. Image quality prediction for DCT-based compression [Text] / R. Kozhemiakin, V. Lukin, B. Vozel // in 2017 14th International Conference The Experience of Designing and Application of CAD Systems in Microelectronics (CADSM). – 2017. – P. 225-228. – doi: <https://doi.org/10.1109/CADSM.2017.7916121>.

44. Krivenko, S. Prediction of Visual Quality Metrics in Lossy Image Compression [Text] / S. Krivenko, F. Li, V. Lukin, B. Vozel, O. Krylova// 2020 IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO). – Kyiv, Ukraine. – 2020. – P. 478-483. – doi: <https://doi.org/10.1109/ELNANO50318.2020.9088819>.

45. Parmar, C. A Review of Image Compression [Electronic resource] / C. Parmar // Journal of information, knowledge and research in electrical engineering. – 2012. – Vol. 2, iss. 2. – P. 281-284. – Режим доступу: https://www.researchgate.net/publication/257868354_A_Review_of_Image_Compression (дата звернення: 30.01.2025).

46. Sindhu, M. Images and Its Compression Techniques – A Review [Electronic

resource] / M. Sindhu, R. Rajkamal // International Journal of Recent Trends in Engineering. – 2009. – Vol. 2, no. 4. – P. 71-75. – Режим доступу: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=3becdbbebb8dd99a6df2c5922af94135c2ea0300> (дата звернення: 30.01.2025).

47. Ponomarenko, N. N. Lossy and Visually Lossless Compression of Single-Look SAR Images [Text] / N. N. Ponomarenko, V. V. Lukin, R. A. Kozhemiakin, K. O.Egiazarian, M. K. Chobanu // Telecommunications and Radio Engineering. – 2013. – Vol. 72, No 8. – P. 711-729. – doi: <https://doi.org/10.1615/TelecomRadEng.v72.i8.60>.

48. Magli E. Multiband lossless compression of hyperspectral images [Text] / E. Magli // IEEE Transactions on Geoscience and Remote Sensing. – 2009. – Vol. 47, No. – P. 1168-1178. – doi: <https://doi.org/10.1109/TGRS.2008.2009316>.

49. Krivenko, S. Smart Lossy Compression of Images Based on Distortion Prediction [Text] / S. Krivenko, O. Krylova, E. Bataeva, V. Lukin // Telecommunications and Radio Engineering. – 2018. – Vol. 77, No. 17. – P. 1535-1554. – doi: <https://doi.org/10.1615/TELECOMRADENG.V77.I17.40>.

50. Kovalenko, B. Prediction of parameters in optimal operation point for BPG-based lossy compression of noisy images [Text] / B. Kovalenko, V. Lukin et al. // Ukrainian Journal of Remote Sensing. – 2022. – Vol. 9, no. 2. – P. 4-12. – doi: <https://doi.org/10.36023/ujrs.2022.9.2.212>.

51. Al-Shaykh, O. K. Lossy compression of noisy images [Text] / O. K. Al-Shaykh, R. M. Mersereau // IEEE Transactions on Image Processing. – 1998. – Vol. 7, no. 12. – P. 1641-1652. – doi: <https://doi.org/10.1109/83.730376>.

52. Lin, W. Perceptual Visual Quality Metrics: A Survey [Text] / W. Lin, C-C. Jay Kuo // Journal of Visual Communication and Image Representation. – May 2011. – Vol. 22, No. 4. – P. 297-312. – doi: <https://doi.org/10.1016/j.jvcir.2011.01.005>.

53. Silpa, K. Comparison of image quality metrics [Electronic resource] / K. Silpa, S. A. Mastani // International Journal of Engineering Research and Technology (IJERT). – June 2012. – Vol. 1, No. 4. – P.1-5. – Режим доступу: <https://www.ijert.org/research/comparison-of-image-quality-metrics-IJERTV1IS4105.pdf> (дата звернення: 30.01.2025).

54. Inglada, J. Similarity measures for multisensor remote sensing images [Text] / J. Inglada // IEEE International Geoscience and Remote Sensing Symposium. – 2002. – Vol.1. – P. 104-106. – doi: <https://doi.org/10.1109/IGARSS.2002.1024955>.
55. Wang, Z. Mean squared error: love it or leave it? A new look at image fidelity measures [Text] / Z. Wang, A.C. Bovik // IEEE Signal Processing Magazine. – Jan. 2009. – P. 98-117. – doi: <https://doi.org/10.1109/MSP.2008.930649>.
56. Ponomarenko, N. On between-coefficient contrast masking of dct basis functions. In Proceedings of the Third International Workshop on Video [Electronic resource] / N. Ponomarenko, F. Silvestri, K. Egiazarian, M. Carli, J. Astola, V. Lukin, // Processing and Quality Metrics for Consumer Electronics. – 2007. – Vol. 4. – Режим доступа: https://ponomarenko.info/vpqm07_p.pdf (дата звернення: 30.01.2025).
57. Lukin, V. Analysis of HVS-Metrics' Properties Using Color Image Database TID2013 [Text] / V. Lukin, N. Ponomarenko, K. Egiazarian, J. Astola // Advanced Concepts for Intelligent Vision Systems (ACIVS). – Italy. – October 2015. – P. 613-624. – doi: https://doi.org/10.1007/978-3-319-25903-1_53.
58. Ponomarenko, N. Modified image visual quality metrics for contrast change and mean shift accounting [Electronic resource] / N. Ponomarenko, O. Ieremeiev, V. Lukin, K. Egiazarian, M. Carli, // Proceedings of the 11th International Conference the Experience of Designing and Application of CAD Systems in Microelectronics (CADSM). – 2011. – Режим доступа: <https://ponomarenko.info/papers/psnrhma.pdf> (дата звернення: 30.01.2025).
59. Wang, Z. Image quality assessment: from error visibility to structural similarity [Text] / Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli // IEEE Transactions on Image Processing. – 2004. – Vol. 13, no. 4. – P. 600-612. – doi: <https://doi.org/10.1109/TIP.2003.819861>.
60. Wang, Z. Multiscale structural similarity for image quality assessment [Text] / Z. Wang, E. P. Simoncelli, A. C. Bovik // The Thirty-Seventh Asilomar Conference on Signals, Systems & Computers. – 2003. – Vol. 2. – P. 1398-14021. – doi: <https://doi.org/10.1109/ACSSC.2003.1292216>.
61. Ziaei Nafchi, H. Mean Deviation Similarity Index: Efficient and Reliable Full-

Reference Image Quality Evaluator [Text] / H. Ziaei Nafchi, A. Shahkolaei, R. Hedjam, M. Cheriet // IEEE Access. – 2016. – Vol. 4. – P. 5579-5590. – doi: <https://doi.org/10.1109/ACCESS.2016.2604042>.

62. Elakkiya, S. Comprehensive Review on Lossy and Lossless Compression Techniques Evaluator [Text] / S. Elakkiya, K. S. Thivya // Journal of The Institution of Engineers (India): Series B. – 2022. – Vol. 13. – P. 1003-1012. – doi: <https://doi.org/10.1007/s40031-021-00686-3>.

63. Liu, C. Automatic Estimation and Removal of Noise from a Single Image [Text] / C. Liu, R. Szeliski, S.B. Kang Bing // Pattern Analysis and Machine Intelligence, IEEE Transactions. – 2008. – No. 30(2). – P. 299-314. – doi: <https://doi.org/10.1109/TPAMI.2007.1176>.

64. Vozel, B. Noise identification and estimation of its statistical parameters by using unsupervised variational classification [Text] / B. Vozel, K. Chehdi, L. Klaine, V.V. Lukin, S.K. Abramov // IEEE International Conference on Acoustics Speech and Signal Processing Proceedings. – 2006. – Vol 2. – P. 841-844. – doi: <https://doi.org/10.1109/ICASSP.2006.1660474>.

65. Colom, M. Analysis and Extension of the PCA Method, Estimating a Noise Curve from a Single Image [Text] / M. Colom, A. Buades // Image Processing On Line. – 2016. – Vol 6. – P. 365–390. – doi: <https://doi.org/10.5201/ipol.2016.124>.

66. Chatterjee, P. Practical Bounds on Image Denoising: From Estimation to Information [Text] / P. Chatterjee, P. Milanfar // Image Processing, IEEE Transactions on. – 2011. – No. 20(5). – P. 1221-1233. – doi: <https://doi.org/10.1109/TIP.2010.2092440>.

67. Abramova, V. Blind Estimation of Speckle Characteristics for Sentinel Polarimetric Radar Images [Text] / V. Abramova, S. Abramov, V. Lukin, K. Egiazarian // IEEE Microwaves, Radar and Remote Sensing Symposium. – 2017. – 4 p. – doi: <https://doi.org/10.1109/MRRS.2017.8075078>.

68. Zoran, D. Scale invariance and noise in natural images [Text] / D. Zoran, Y. Weiss // Computer Vision, IEEE 12th International Conference. – 2009. – P. 2209-2216. – doi: <https://doi.org/10.1109/ICCV.2009.5459476>.

69. Abramova, V. Analysis of Noise Properties in Dental Images [Text] / V.

Abramova, S. Krivenko, V. Lukin, O. Krylova // IEEE 40th International Conference on Electronics and Nanotechnology (ELNANO). – 2020. – P. 511-515. – doi: <https://doi.org/10.1109/ELNANO50318.2020.9088768>.

70. Rakhshanfar, M. Estimation of Gaussian, Poissonian and Gaussian, and Processed Visual Noise and Its Level Function [Text] / M. Rakhshanfar, M. A. Amer // IEEE Transactions on Image Processing. – 2016. – No. 25(9). – P. 4172-4185. – doi: <https://doi.org/10.1109/TIP.2016.2588320>.

71. Azzari, L. Gaussian-Cauchy mixture modeling for robust signal-dependent noise estimation [Text] / L. Azzari, A. Foi // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)– 2014. – P. 5357-5361. – doi: <https://doi.org/10.1109/ICASSP.2014.6854626>.

72. Methods for Blind Estimation of the Variance of Mixed Noise and Their Performance Analysis [Text] / S. Abramov, V. Zabrodina, V. Lukin, B. Vozel, K. Chehdi, J. Astola // In: Numerical Analysis – Theory and Applications, Ed. Jan Awrejcewicz. – InTech, Austria - 2011. – P. 49-70. - ISBN: 978-953-307-389-7.

73. Zhong, P. Multiple-Spectral-Band CRFs for Denoising Junk Bands of Hyperspectral Imagery [Text] / P. Zhong, R. Wang // IEEE Transactions on Geoscience and Remote Sensing. – 2013. – Vol. 51, No 4. – P. 2260-2275. – doi: <https://doi.org/10.1109/TGRS.2012.2209656>.

74. Ragesh, N. K. Digital image denoising in medical ultrasound images: a survey [Electronic resource] / N. K. Ragesh, A.R. Anil, R. Rajesh // ICGST AIML-11 Conference. – 2011. – P. 12-14. – Режим доступа: https://www.academia.edu/1332306/Digital_Image_Denoising_in_Medical_Ultrasound_images_A_Survey (дата звернення: 30.01.2025).

75. Prasath, V. S. Quantum Noise Removal in X-Ray Images with Adaptive Total Variation Regularization [Text] / V. S. Prasath // Informatica. – 2017. – Vol. 28, no. 3. – P. 505-515. – doi: <https://doi.org/10.15388/Informatica.2017.141>.

76. Owotogbe, J. S. A Comprehensive Review On Various Types of Noise in Image Processing [Electronic resource] / J. S. Owotogbe, T. S. Ibiyemi, B. A. Adu // International Journal of Scientific & Engineering Research. – 2019. – Vol. 10, iss. 11. – P.

388-393. – Режим доступа: <https://www.ijser.org/researchpaper/A-COMPREHENSIVE-REVIEW-ON-VARIOUS-TYPES-OF-NOISE-IN-IMAGE-PROCESSING.pdf> (дата звернення: 30.01.2025).

77. Ponomarenko, M. Blind estimation of white Gaussian noise variance in highly textured images [Text] / N. Ponomarenko, N. Gapon, V. Voronin, K. Egiazarian // IS&T Int'l. Symp. on Electronic Imaging: Image Processing: Algorithms and Systems XVI. – 2018. – 5 p. – doi: <https://doi.org/10.2352/issn.2470-1173.2018.13.ipas-382>.

78. Tallapragada, V. V. S. Mixed image denoising using weighted coding and non-local similarity [Text] / V. V. Satyanarayana Tallapragada, N. Alivelu Manga, G. V. Pradeep Kumar, M. Venkata Naresh // SN Applied Sciences. – 2020. – Vol. 2. – Article No. 997. – doi: <https://doi.org/10.1007/s42452-020-2816-y>.

79. Foi, A. Practical Poissonian-Gaussian Noise Modeling and Fitting for Single Image Raw Data [Text] / A. Foi, M. Trimeche, V. Katkovnik, K. Egiazarian // IEEE Transactions on Image Processing. – 2007. – No 10. P. 1737-1754. – doi: <https://doi.org/10.1109/TIP.2008.2001399>.

80. Azzari, L. Modeling and Estimation of Signal-Dependent and Correlated Noise [Text] / L. Azzari, L. R. Borges, A. Foi // Advances in Computer Vision and Pattern Recognition. – 2018. – P. 1-36. – doi: https://doi.org/10.1007/978-3-319-96029-6_1.

81. Proskura, G. Analysis of Improvement of Noisy Multichannel Image Controlled Pixel-by-Pixel Classification by Post-Classification Processing [Text] / G. Proskura, I. Vasilyeva, V. Lukin // IEEE 15th International Conference on Advanced Trends in Radioelectronics, Telecommunications and Computer Engineering (TCSET) – 2020. – 6 p. – doi: <https://doi.org/10.1109/TCSET49122.2020.235488>.

82. Chatterjee, P. Is Denoising Dead? [Text] / P. Chatterjee, P. Milanfar // IEEE Transactions on Image Processing. – 2010. – Vol. 19, No 4. – P. 895-911. – doi: <https://doi.org/10.1109/TIP.2009.2037087>.

83. Rubel, O. On Prediction of DCT-based Denoising Efficiency under Spatially Correlated Noise Conditions [Text] / O. Rubel, V. Lukin, K. Egiazarian // Proceedings of TCSET 2016. – 2016. – P. 750-754. – doi: <https://doi.org/10.1109/TCSET.2016.7452171>.

84. Lukin, V. V. Denoising Efficiency for Multichannel Images Corrupted by

Signal-dependent Noise [Text] / V. V. Lukin, S. K. Abramov, R. A. Kozhemiakin, M. L. Uss, B. Vozel, K. Chehdi // International Kharkov Symposium on Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves. – June 2013. – 3 p. – doi: <https://doi.org/10.1109/MSMW.2013.6622048>.

85. Zhang, J. Overlapped variable size block motion compensation [Text] / J. Zhang, M. O. Ahmad, M. N. S. Swamy // Proceedings of International Conference on Image Processing, Santa Barbara. – 1997. – Vol. 3. – P. 642-645. – doi: <https://doi.org/10.1109/ICIP.1997.632203>.

86. Odegard, J. E. Joint Compression and Speckle Reduction of SAR Images using Embedded Zerotree Models [Electronic resource] / J. E. Odegard, H. Guo, C. S. Burrus, R. G. Baraniuk // In Workshop on Image and Multidimensional Digital Signal Processing. – 1996. – P. 80–81. – Режим доступа: <https://hdl.handle.net/1911/20172> (дата звернення: 30.01.2025).

87. Ponomarenko, N. DCT Based High Quality Image Compression [Text] / N. Ponomarenko, V. Lukin, K. Egiazarian, J. Astola, // Image Analysis. – 2005. – Vol. 3540. – doi: https://doi.org/10.1007/11499145_119.

88. Ponomarenko, N. ADCTC: a new high quality DCT based coder for lossy image compression [Electronic resource] / N. Ponomarenko, V. Lukin, K. Egiazarian, J. Astola // CD ROM Proceedings of LNLA. – 2008. – P. 6. – Режим доступа: https://www.ponomarenko.info/lnla_adctc.pdf (дата звернення: 30.01.2025).

89. Ponomarenko, N. Lossy Compression of Noisy Images Based on Visual Quality: A Comprehensive Study [Text] / N. Ponomarenko, S. Krivenko, V. Lukin, K. Egiazarian, J. Astola // EURASIP Journal on Advances in Signal Processing. – 2010. – Article No. 976436. – doi: <https://doi.org/10.1155/2010/976436>.

90. Cabeen, K. Image compression and the Discrete Cosine Transform [Electronic resource] / K. Cabeen, P. Gent // Режим доступа: <https://www.math.cuhk.edu.hk/~lmlui/dct.pdf> (дата звернення: 30.01.2025).

91. Kavitha, P. A Survey on Lossless and Lossy Data Compression Methods / P. Kavitha [Electronic resource] // International Journal of Computer Science & Engineering Technology (IJCSET). – 2016. – Vol. 7, no. 3. – Режим доступа:

<https://ijcset.com/docs/IJCSET16-07-03-049.pdf> (дата звернення: 30.01.2025).

92. Jayasankar, U. A survey on data compression techniques: From the perspective of data quality, coding schemes, data type and applications [Text] / U. Jayasankar, V. Thirumal, D. Ponnurangam // Journal of King Saud University - Computer and Information Sciences. – 2021. – Vol. 33, iss. 2. – P. 119-140. – doi: <https://doi.org/10.1016/j.jksuci.2018.05.006>.

93. Lainema, J. HEVC still image coding and high efficiency image file format [Text] / J. Lainema, M. M. Hannuksela, V. K. M. Vadaikal, E. B. Aksu // IEEE International Conference on Image Processing (ICIP). – 2016. – P. 71-75. – doi: <https://doi.org/10.1109/ICIP.2016.7532321>.

94. Bellard F. BPG Image format. [Electronic resource] / F. Bellard. – 2018. Режим доступу: <http://bellard.org/bpg/> (дата звернення: 30.01.2025).

95. Kryvenko, S. Peculiarities of noisy image lossy compression [Text] / S. Kryvenko, V. Lukin // Herald of Khmelnytskyi National University Technical sciences. – 2024. – Vol. 333, no. 2. – P. 278-283. – doi: <https://doi.org/10.31891/2307-5732-2024-333-2-44>.

96. Mozammel, M. Image Compression Using Discrete Wavelet Transform [Electronic resource] / M. Mozammel, H. Chowdhury, A. Khatun // JCSI International Journal of Computer Science Issues. – 2012. – Vol. 9, iss. 4, no. 1. – P. 327-330. – Режим доступу: https://www.researchgate.net/publication/266485442_Image_Compression_Using_Discrete_Wavelet_Transform (дата звернення: 30.01.2025).

97. Skodras, A. The JPEG 2000 still image compression standard [Text] / A. Skodras, C. Christopoulos, T. Ebrahimi // IEEE Signal Processing Magazine. – 2001. – Vol.18, no. 5. – P. 36-58. – doi: <https://doi.org/10.1109/79.952804>.

98. Lawson, S. Image compression using wavelets and JPEG2000: a tutorial [Electronic resource] / S. Lawson, J. Zhu // Electronics and communication engineering journal. – 2002. – Режим доступу: https://engineering.purdue.edu/~ee538/Image_compression_wavelets_jpeg2000.pdf (дата звернення: 30.01.2025).

99. Khelifi, F. Joined Spectral Trees for Scalable SPIHT-Based Multispectral Image Compression [Electronic resource] / F. Khelifi, A. Bouridane, F. Kurugollu // IEEE Transactions on Multimedia. – Vol. 10. – P. 316-329. – doi: <https://doi.org/10.1109/TMM.2008.917357>.
100. Christophe E. Hyperspectral Image Compression: Adapting SPIHT and EZW to Anisotropic 3-D Wavelet Coding [Text] / E. Christophe, C. Mailhes, P. Duhamel // IEEE Transactions on Image Process. – 2008. – Vol. 17, No. 12. – P. 2334-2346. – doi: <https://doi.org/10.1109/TIP.2008.2005824>.
101. Shapiro, J. M. Embedded image coding using zerotrees of wavelet coefficients [Text] / J. M. Shapiro // IEEE Transactions on Signal Processing. – 1993. – Vol. 41, no. 12. – P. 3445-3462. – doi: <https://doi.org/10.1109/78.258085>.
102. Taubman, D. High performance scalable image compression with EBCOT [Text] / D. Taubman // IEEE Transactions on Image Processing. – 2000. – Vol. 9, no. 7. – P. 1158-1170. – doi: <https://doi.org/10.1109/83.847830>.
103. Li, X. Neural Image Compression and Explanation [Electronic resource] / X. Li, S. Ji // IEEE Access. – 2020. – Режим доступа: <https://arxiv.org/pdf/1908.08988> (дата звернення: 30.01.2025).
104. Dumas, T. Autoencoder Based Image Compression: Can the Learning be Quantization Independent? [Text] / T. Dumas, A. Roumy, C.M. Guillemot // IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). – 2018. – P. 1188-1192. – doi: <https://doi.org/10.1109/ICASSP.2018.8462263>.
105. Theis, L. Lossy imagecompression with compressive autoencoders [Text] / L. Theis, W. Shi, A. Cunningham, F. Huszár // International Conference on Learning Representations. – 2017. – doi: <https://doi.org/10.48550/arXiv.1703.00395>.
106. Zhou, L. Variational Autoencoder for Low Bit-rate Image Compression [Electronic resource] / L. Zhou, C. Cai, Y. Gao, S. Su, J. Wu, // CVPR Workshops. – 2018. – P. 2617-2620. – Режим доступа: https://openaccess.thecvf.com/content_cvpr_2018_workshops/papers/w50/Zhou_Variational_Autoencoder_for_CVPR_2018_paper.pdf (дата звернення: 30.01.2025).
107. Wei, K. Research on Convolutional Neural Network-Based Compression

Methods for Multispectral Images [Text] / K. Wei, H. Liu, P. Yu, Y. Ren, G. Gao, C. Wen // Journal of Physics: Conference Series. – 2024. – Vol. 2717. – doi: <https://doi.org/10.1088/1742-6596/2717/1/012006>.

108. Iwasokun, G. B. Discrete transformation technique for video compression [Text] / G. B. Iwasokun, M. O. Olaoye // Iran Journal of Computer Science. – 2021. – Vol. 4. – P. 281-292. – doi: <https://doi.org/10.1007/s42044-021-00085-3>.

109. Zemliachenko, A. N. Lossy compression of noisy remote sensing images with prediction of optimal operation point existence and parameters [Text] / A. N. Zemliachenko, S. K. Abramov, V. V. Lukin, B. Vozel, K. Chehdi // Journal of Applied Remote Sensing. – 2015. – Vol. 9, no. 1. – Article No. 095066. – doi: <https://doi.org/10.1117/1.JRS.9.095066>.

110. Kovalenko, B. Analysis of the potential efficiency of post-filtering noisy images after lossy compression [Text] / B. Kovalenko, V. Lukin, V. Rebrov // Ukrainian journal of remote sensing. – 2023. – Vol. 10, no. 1. – P. 11-16. – doi: <https://doi.org/10.36023/ujrs.2023.10.1.231>.

111. Ponomarenko, N. Estimation of accessible quality in noisy image compression [Electronic resource] / N. Ponomarenko, M. Zriakhov, V. V. Lukin, J. T. Astola, K. O. Egiazarian // Proceedings of the 14th European Signal Processing Conference, EUSIPCO. – 2006. – P. 4. – Режим доступа: <https://ieeexplore.ieee.org/document/7071332> (дата звернення: 30.01.2025).

112. Ponomarenko, N. Lossy Compression of Images with Additive Noise [Text] / N. Ponomarenko, V. Lukin, M. Zriakhov, K. Egiazarian, J. Astola // Advanced Concepts for Intelligent Vision Systems. – 2005. – Vol. 3708. – doi: https://doi.org/10.1007/11558484_48.

113. Anscombe, F. J. The transformation of poisson, binomial and negative-binomial data [Text] / F. J. Anscombe // Biometrika. – 1948. – Vol. 35, iss. 3-4. – P. 246-254. – doi: <https://doi.org/10.1093/biomet/35.3-4.246>.

114. Wei, Z. Spatio-Temporal Just Noticeable Distortion Profile for Grey Scale Image/Video in DCT Domain [Text] / Z. Wei, K. N. Ngan // IEEE Transactions on Circuits and Systems for Video Technology. – 2009. – Vol. 19, no. 3. – P. 337-346. – doi: <https://doi.org/10.1109/TCSVT.2009.2013518>.

115. Ieremeiev, O. Full-Reference Quality Metric Based on Neural Network to Assess the Visual Quality of Remote Sensing Images [Text] / O. Ieremeiev, V. V. Lukin, K. Okarma, K. O. Egiazarian, // Remote Sensing. – 2020. – Vol. 12. – Article No. 2349. – doi: <https://doi.org/10.3390/rs12152349>.
116. Cameron, A. C. An R-squared measure of goodness of fit for some common nonlinear regression models [Text] / A. C. Cameron, F. A. G. Windmeijer // Journal of Econometrics. – 1997. – Vol. 77, iss. 2. – P. 329-342. – doi: [https://doi.org/10.1016/S0304-4076\(96\)01818-0](https://doi.org/10.1016/S0304-4076(96)01818-0).
117. Abramova, V. A Fast and Accurate Prediction of Distortions in DCT-Based Lossy Image Compression [Text] / V. Abramova, V. Lukin, S. Abramov, S. Kryvenko, P. Lech, K. Okarma // Electronics. – 2023. – Vol. 12, no. 11. – Article No. 2347. – doi: <https://doi.org/10.3390/electronics12112347>.
118. USC-SIPI. Image Database. [Electronic resource] // University of Southern California Signal and Image Processing Institute (USC-SIPI). – Режим доступа: <http://sipi.usc.edu/database/> (дата звернення: 30.01.2025).
119. Zhang, Y. An Adaptive Infrared Image Preprocessing Method Based on Background Complexity Descriptors [Text] / Y. Zhang, Z. Zhang, X. Wang, X. Wang, J. Ge, F. Bian // Eighth International Conference on Instrumentation & Measurement, Computer, Communication and Control (IMCCC). – 2018. – P. 344-349. – doi: <https://doi.org/10.1109/IMCCC.2018.00079>.

ДОДАТОК А.

СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА

Статті у наукових періодичних виданнях:

1. Б. Коваленко, В. Лукін, “Використання візуальних метрик для аналізу стиснення з втратами зашумлених зображень,” *Авіаційно-космічна техніка і технологія*, вип. 6, с. 83-91, 2021, doi: 10.32620/akt.2021.6.09 (фаховий).
2. V. Lukin, B. Kovalenko, S. Kryvenko, N. Ponomarenko, K. Egiazarian, J. Astola, “Compression of Noisy Images Taking into Account Visual Quality: A Comprehensive Study,” *Research Developments in Science and Technology*, vol. 8, pp. 89–109, 2022, doi: 10.9734/bpi/rdst/v8/2722B.
3. B. Kovalenko, V. Lukin, “Usage of different Chroma Subsampling Modes in Image Compression by BPG Coder,” *Ukrainian journal of remote sensing*, vol. 9, no. 3, pp. 11-16, 2022, doi: 10.36023/ujrs.2022.9.3.216 (фаховий).
4. V. Makarichev, B. Kovalenko, V. Lukin, “Preliminary analysis of noisy image lossy compression by discrete atomic transform-based coder,” *Aerospace technic and technology*, vol. 186, no. 2, pp. 67-74, 2023, doi: 10.32620/akt.2023.2.07 (фаховий).
5. V. Naumenko, B. Kovalenko, V. Lukin, “BPG-based compression analysis of poisson-noisy medical images,” *Radioelectronic and computer systems*, no. 3, pp. 91-100, 2023, doi: 10.32620/reks.2023.3.08 (Scopus).
6. B. Kovalenko, V. Lukin, S. Kryvenko, V. Naumenko, B. Vozel, “Prediction of Parameters in Optimal Operation Point for BPG-based Lossy Compression of Noisy Images,” *Ukrainian journal of remote sensing*, vol. 9, no. 2, pp. 4-12, 2022, doi: 10.36023/ujrs.2022.9.2.212 (фаховий).
7. B. Kovalenko, V. Lukin, S. Kryvenko, V. Naumenko, B. Vozel, “BPG-Based Automatic Lossy Compression of Noisy Images with the Prediction of an Optimal Operation Existence and Its Parameters,” *Applied Sciences*, vol. 12, no. 15: 7555, 2022, doi: 10.3390/app12157555 (Scopus).
8. V. Lukin, B. Kovalenko, S. Kryvenko, V. Naumenko, B. Vozel, “Prediction of Optimal Operation Point Existence and Its Parameters in BPG-Based Automatic Lossy

Compression of Noisy Images,” *Current Overview on Science and Technology Research*, vol. 9, pp. 1–36, 2022, doi: 10.9734/bpi/costr/v9/4316A.

9. B. Kovalenko, V. Lukin, B. Vozel, “BPG-Based Lossy Compression of Three-channel Noisy Images with Prediction of Optimal Operation Existence and Its Parameters,” *Remote Sensing*, vol. 15, no. 6: 1669, 2023, doi: 10.3390/rs15061669 (Scopus).

10. V. Lukin, B. Kovalenko, B. Vozel, “A Fast and Accurate Prediction of BPG Compression Parameters in Optimal Operation Point Neighbourhood for Three-channel Noisy Images,” *Fundamental Research and Application of Physical Science*, vol. 8, pp. 129–160, 2023, doi: 10.9734/bpi/fraps/v8/7110A.

11. B. Kovalenko, V. Lukin, “Analysis of distortions due to bpg-based lossy compression of noise-free and noisy images,” *Herald of Khmelnytskyi National University Technical sciences*, vol. 325, no. 5, pp. 128-135, 2023, doi: 10.31891/2307-5732-2023-325-5 (фаховий).

12. V. Lukin, B. Kovalenko, S. Kryvenko, V. Naumenko, B. Vozel, “Peculiarities of BPG-Based Automatic Lossy Compression of Noisy Images”, *Computational Techniques for Smart Manufacturing in Industry 5.0 Methods and Applications*, 1st Edition, ISBN 9781032506203 386 Pages 10 Color & 115 B/W Illustrations January 27, 2025 by CRC Press.

Наукові праці, які засвідчують апробацію матеріалів дисертації:

13. B. Kovalenko, V. Naumenko, S. Krivenko, V. Lukin, “Analysis of noisy image lossy compression by BPG using visual quality metrics,” *Advanced Trends in Information Theory (ATIT)*, Kiev, Ukraine, 2021, doi: 10.1109/ATIT54053.2021.9678575 (Scopus).

14. B. Kovalenko, V. Lukin, “Analysis of color image compression by BPG coder,” 2022 *IEEE 3rd KhPI Week on Advanced Technology (KhPIWeek)*, Kharkiv, Ukraine, October 2022, 6 p., doi: 10.1109/KhPIWeek57572.2022.9916387 (Scopus).

15. Б. Коваленко, “Знаходження оптимального параметра для прогнозування оптимальної робочої точки BPG кодера,” *Матеріали XXII Всеукраїнської науково-технічної конференції молодих вчених, аспірантів та студентів*. Одеса, 21-22 квітня 2022 р. - Одеса, Видавництво ОНТУ, 2022 р. – 175-177 с.

16. B. Kovalenko, V. Lukin, "BPG-based compression of Poisson noisy images," *2023 13th International Conference on Dependable Systems, Services and Technologies (DESSERT)*, Athens, Greece, 2023, pp. 1-7, doi: 10.1109/DESSERT61349.2023.10416544 (Scopus).
17. B. Kovalenko, V. Lukin, "Prediction of optimal operating point for BPG-based lossy compression of images corrupted by Poisson noise," *2023 IEEE 4th KhPI Week on Advanced Technology (KhPIWeek)*, Kharkiv, Ukraine, 2023, pp. 1-5, doi: 10.1109/KhPIWeek61412.2023.10312954 (Scopus).
18. B. Kovalenko, V. Lukin, "Pre-requisites for Mean Square Error Prediction in Better Portable Graphics based Lossy Compression of Grayscale Images," *2024 IEEE 42nd International Conference on Electronics and Nanotechnology (ELNANO)*, Kyiv, Ukraine, 2024, pp. 488-492, doi: 10.1109/ELNANO63394.2024.10756949.
19. B. Kovalenko, V. Lukin, B. Vozel, "MSE Prediction in BPG-based Lossy Compression of Grayscale Images," *2024 IEEE 5th KhPI Week on Advanced Technology (KhPIWeek)*, Kharkiv, Ukraine, 2024, doi: 10.1109/KhPIWeek61434.2024.10877996.

ДОДАТОК Б.

АКТИ ВПРОВАДЖЕННЯ

ЗАТВЕРДЖУЮ

В.о. директора Інституту радіофізики
та електроніки ім. О.Я. Усикова НАН України



Юрій Погвінов

«26» грудня 2024 р.

АКТ

науково-технічної комісії щодо впровадження наукових положень і
результатів

дисертації на здобуття ступеня доктора філософії аспіранта
Національного аерокосмічного університету (ХАІ)

Коваленко Богдана Віталійовича

Науково-технічна комісія Інституту радіофізики та електроніки ім. О.Я. Усикова НАН України у складі голови комісії виконуючого обов'язки заступника директора з наукової роботи, канд. фіз.-мат. наук Соболяка Олександра Васильовича та членів комісії завідуючого відділом поширення радіохвиль в природних середовищах, канд. фіз.-мат. наук Роєнко Олександра Миколайовича та старшого наукового співробітника лабораторії моніторингу і спектроскопії, канд. фіз.-мат. наук Єгорова Вадима Анатолійовича

склала даний акт про те, що в результаті виконання спільних науково-дослідних робіт, були використані та впроваджені наступні положення та результати, що отримані аспірантом Коваленко Б.В.:

- метод прогнозування середньоквадратичної похибки внесених спотворень для VPG кодера при стисненні напівтонових зображень;
- удосконалений метод прогнозування середньоквадратичної похибки для VPG кодера з використанням більшої кількості вхідних параметрів.

Розроблені в дисертації методи та програмно-алгоритмічні засоби дозволили:

1) прогнозувати середньоквадратичну похибку для прийняття рішень стосовно параметрів кодера для забезпечення потрібного рівню завад/якості одноканальних зображень та багатоканальних даних під час їх покомпонентної обробки;

2) підвищити точність прогнозування середньоквадратичної похибки або пікового відношення сигнал-шум.

Голова комісії:

Члени комісії:

Олександр Соболяк

Олександр Роєнко

Вадим Єгоров

«Затверджую»

Проректор з НР Національного
аерокосмічного університету імені
М.С. Жуковського «Харківський авіаційний

інститут», д-р техн. наук, професор
Павліков Володимир Володимирович



2025 р.

АКТ

науково-технічної комісії щодо впровадження наукових положень і результатів
кандидатської дисертації аспіранта Національного аерокосмічного університету «ХАІ»

Коваленко Богдана Віталійовича

Науково-технічна комісія у складі: заступник завідувача кафедри інформаційно-комунікаційних технологій ім. О.О. Зеленського (504) к.т.н., доцент Васильєва Ірини Карлівна (голова комісії), та члени комісії к.т.н., доцент, доцент каф. 504 Хуторненко Сергій Володимирович, к.т.н., доцент, докторант каф. 504 Єремєєв Олег Ігорович склали даний акт про те, що в навчальний процес каф. 504 були впроваджені наступні наукові положення та результати, що отримані аспірантом Коваленко Б. В.:

- метод аналізу характеристик швидкість-спотворення для BPG та інших кодерів під час стиснення напівтонових зображень та метод прогнозування внесених спотворень (курс «Цифрова обробка даних»);

- удосконалений метод прогнозування OPT та в її околі для роботи з кольоровими зображеннями (курс «Обробка мультимедійних даних»).

Розроблені в дисертації методи та програмно-алгоритмічні засоби дозволили:

1) підвищити якість стиснутих зображень в разі виявлення та стиснення зображення в оптимальній робочій точці;

2) забезпечити прийняття рішення про налаштування параметрів кодера завдяки попередньому аналізу ентропії як характеристики складності зображень.

Голова комісії

Члени комісії

Ірина ВАСИЛЬЄВА

Сергій ХУТОРНЕНКО

Олег ЄРЕМЕЄВ