

**Бараннік Д.В.**

Харківський національний університет радіоелектроніки

**Сідченко Є.С.**

Харківський національний університет імені В.Н. Каразіна

**Лиходєєв О.С.**

Харківський національний університет Повітряних Сил

**Помазуєв В.В.**

Харківський національний університет Повітряних Сил

**Федоровський О.В.**

Харківський національний університет Повітряних Сил

## МЕТОД ЗАСТОСУВАННЯ АПЕРТУР ДО АСИМЕТРИЧНИХ СИСТЕМ КОДУВАННЯ ВІДЕОДАНИХ

*У статті розглянуто актуальні проблеми стиснення відеозображень для інформаційних систем, що функціонують в умовах обмежених або нестабільних каналів зв'язку. Проаналізовано недоліки класичних і сучасних форматів кодування, зокрема їхню низьку стійкість до помилок та неадаптивність до структурної складності зображень, зокрема показано на практичному прикладі слабкість класичної сегментації  $8 \times 8$ , яка не враховує особливості зображення. Обґрунтовано доцільність використання асиметричних систем числення (ANS), зокрема варіанта  $rANS$ , як ефективного інструменту ентропійного кодування, що поєднує високу швидкодію з наближенням до теоретичної межі стиснення. Розглянуто недоліки ANS, а саме завадостійкість та залежність ефективності методів від вхідних дани та алфавітів. Для розв'язання цієї проблеми, далі у статті розроблено метод сегментації зображень на основі виявлення локальних областей когерентності, що характеризуються подібними структурно-візуальними властивостями. Запропоновано механізм апертурного поділу з урахуванням амплітудних варіацій пікселів та адаптивної апроксимації значень усередині апертур. Показано, що такий підхід дозволяє значно підвищити завадостійкість до помилок при передачі даних, локалізуючи їхній вплив у межах окремих фрагментів зображення. Проведено порівняльне експериментальне дослідження ефективності запропонованого методу у поєднанні з  $rANS$ -кодуванням. Отримані результати демонструють зниження обсягу закодованих даних без значущих втрат візуальної якості та суттєве зменшення масштабів спотворень при виникненні помилок у передаваному потоці. Завдяки врахуванню особливостей зображення, при внесенні каналних помилок, об'єкти на зображенні залишаються гарно видимими для людського ока. Надані експериментальні результати дослідження на прикладах, які підтверджують ефективність та запропонованого підходу у порівнянні до класичних методів сегментації, що підтверджує перспективність методів для використання в реальному часі в системах із підвищеними вимогами до надійності відеоінформації.*

**Ключові слова:** апертура, області когерентності, асиметричні системи числення, стиснення з втратами, шум, завадостійкість.

**Постановка проблеми.** Сучасні інформаційні системи, що працюють з великомасштабними зображеннями та відеопотоками, стикаються з викликами, пов'язаними із забезпеченням високої якості візуального сприйняття в умовах обмежених ресурсів – зокрема, пропускну здатності каналів зв'язку, обчислювальної потужності та енергоспоживання [1]. Ці виклики особливо актуальні для інтелектуальних сервісів реального часу,

таких як автоматизоване відеоспостереження, смарт-аналітика міського середовища, автономне керування транспортом, оперативне реагування у кризових ситуаціях. У подібних середовищах, де часто мають місце нестабільні або обмежені канали зв'язку (наприклад, у бездротових сенсорних мережах або мобільних пристроях), критичною стає потреба в ефективному стисненні зображень і відео [2]. Ключові вимоги до таких методів

включають зменшення обсягу переданих даних, зниження затримок і навантаження на мережу, швидке декодування та мінімізацію енергоспоживання. Водночас важливо зберігати не лише візуальну якість, але й семантичну інформативність даних для подальшої інтелектуальної обробки. Однак традиційні формати стиснення (JPEG, H.264/AVC, HEVC, AV1) демонструють обмеження в адаптації до складної структурної організації зображень [3, 4]. Особливо проблематичною є ситуація, коли через втрату навіть одного байта кодувального потоку відбувається порушення декодування всієї секції – що характерно, зокрема, для форматів JPEG XL та Jpegli [5]. Попри їхню високу ефективність і використання сучасних механізмів обробки (адаптивне квантування, модульна архітектура, ентропійне кодування на базі ANS) [6], ці формати залишаються вразливими до помилок, не маючи вбудованих засобів структурної локалізації або самовідновлення. Окремою проблемою є недостатня ефективність традиційного рівномірного поділу простору зображення (наприклад, блоки 8×8), який не враховує нерівномірну структурну насиченість. Унаслідок цього ресурси витрачаються нераціонально, а помилки в критично важливих зонах можуть спричинити каскадне погіршення якості [7].

**Аналіз останніх досліджень і публікацій.** Основною метрикою для областей когерентності є нормалізований коефіцієнт кореляції  $\underline{\mu}(x_{\xi,y}; x_{\xi,y+\tau})$ , який визначається наступною формулою:

$$\underline{\mu}(x_{\xi,y}; x_{\xi,y+\tau}) = \frac{\sum_{\xi=1}^M \sum_{y=1}^N (x_{\xi,y} - x_{\xi,y+\tau})(x_{\xi,y} - x_{\xi,y+\tau})}{R_{kov}},$$

де  $x_{\xi,y}$  та  $x_{\xi,y+\tau}$  відповідно  $\tau$ -ий елемент фрагмента зображення,  $R_{kov}$  – довжина фрагмента зображення, протягом якої спостерігається статистична залежність між елементами [8].

Водночас, нормалізований коефіцієнт кореляції неефективно враховує наявність психовізуальної надмірності в зображеннях. Тому для опису областей когерентності пропонується використовувати міру амплітуди незначних варіацій  $\delta$ . Значення  $\delta$  визначає послідовність (сегмент) зображення, для якої, з точки зору візуального сприйняття, прийнятно апроксимувати значення пікселів таким чином, що збільшує нормалізований коефіцієнт кореляції. В узагальненому вигляді це можна описати наступним чином. Якщо для послідовності  $X^\xi$  елементи  $x_{\xi,y}$  мають нерівні значення та виконується умова

$$x_{\xi,y+r} \in [a_\xi - \frac{\delta}{2}; a_\xi + \frac{\delta}{2}], r = 0, R_\xi - 1$$

тоді послідовність  $X^\xi$  буде областю когерентності (апертурою). Тут,  $R_\xi$  – довжина послідовності, для якої виконується умова,  $a_\xi$  – базовий параметр області когерентності, фізичний сенс якого полягає в тому, що при апроксимації опису елементів апертури  $X^\xi$  зі значенням  $a_\xi$ , це не призведе до втрат у якості візуального сприйняття зображення. Водночас, перший елемент, для якого умова перестане виконуватися, буде першим елементом наступної сусідньої апертури. Стиснення даних переважно базується на двох підходах до ентропійного кодування: кодуванні Хаффмана та арифметичному/діапазонному кодуванні. Перший підхід набагато швидший, але апроксимує ймовірності степенями 2, що зазвичай призводить до відносно низьких коефіцієнтів стиснення. Другий підхід використовує точні ймовірності, тим самим практично наближаючись до теоретичної межі ефективності стиснення (ентропія Шеннона), але ціною значно вищих обчислювальних витрат [9]. Асиметричні системи числення (ANS) представляють собою новий крок у точному ентропійному кодуванні, поєднуючи переваги обох підходів. Таким чином, для розміру алфавіту 256 швидкість кодування ANS приблизно на 50% вища, ніж кодування Хаффмана, наближаючись до ефективності стиснення арифметичного кодування [10]. Цього спрощення було досягнуто завдяки тому, що замість двох чисел для визначення діапазону використовується одне просте число як поточний стан. Окрім спрощення перенормування, це також дозволяє представити всю поведінку для заданого розподілу ймовірностей у відносно невеликій таблиці, що представляє автомат ентропійного кодування [11].

**Постановка завдання.** На підставі визначеного протиріччя постає актуальна науково-прикладна задача: підвищення цілісності інформації в умовах внесення каналних помилок, завдяки подальшому вдосконаленню методів сегментації та стиснення відеоданих. Отже, мета статті полягає у розробці методу стиснення відеозображень на основі асиметричного кодування та впровадженням механізмів локалізації впливу каналних помилок на цілісність відновлених відеоданих.

**Виклад основного матеріалу.** Як математичну модель було обрано діапазонно-орієнто-

ване ANS (rANS). Хоча воно має дещо нижчу ефективність стиснення, ніж табличний варіант (tANS), воно менш ресурсоємне, дозволяє працювати з більшими алфавітами та є прямою альтернативою класичному арифметичному кодуванню [12]. Стан на наступному кроці кодування обчислюється наступним чином:

$$blockId = \frac{x}{l_s}; slot = b_s + \text{mod}(s, l_s);$$

$$C(s, x) = blockId \cdot m + slot,$$

де  $x$  – попередній стан,  $s$  – позиція поточного символу, який потрібно закодувати, в алфавіті,  $l_s$  – частота появи символу в алфавіті,  $m = \sum_{i=1}^s l_i$ ,  $b_s$  – кумулятивна частота появи символу, розрахована за формулою  $b_s = \sum_{i=1}^{s-1} l_i$ , що враховує позицію  $s$  у поточному циклі. Крок декодування визначається наступним чином:  $blockId = \frac{x}{m}$ ;  $s = (\text{mod}(x, m)); D(x) = (s, l_s \cdot blockId + slot - b_s)$ .

Якщо ми виберемо  $m$  як ступень числа 2, множення та ділення на нього можна виконати за допомогою побітових зсувів, і  $\text{mod}(x, m)$  можна обчислити за допомогою побітового «І» з маскою. У цьому випадку для декодування потрібне лише одне множення на крок. Хоча rANS забезпечує відносно високий коефіцієнт стиснення та працює з високою швидкістю, були виявлені проблеми із завадостійкістю. У наступних експериментах у файл записувався лише результат кодування, тоді як вся службова інформація, така як алфавіт із частотами символів, зберігалася окремо. Навіть у цьому випадку внесення помилок у результат кодування суттєво вплинуло на вихідне зображення після декодування. Нижче наведено приклади зображень після декодування, де в декодовані дані було внесено різний ступінь помилок. Як показано на рис. 1–5, при рівномірній сегментації 8x8 без поділу апертури помилки, внесені в сегменти з високою нерівномірністю, викликають значне спотворення зображення. Це відбувається тому, що rANS має сувору залежність поточного стану кроку кодування від попереднього стану кроку. Отже, коли помилки вносяться в сегмент рівномірної сегментації, весь сегмент спотворюється на випадкові пікселі з алфавіту поточного сегмента. Щоб вирішити цю проблему, зображення пропонується розділити на апертури. В якості першого елемента апертури береться опорний елемент апертури  $a_\xi$ . Поріг для

незначних змін амплітуди  $\delta$  було встановлено на 40% значення, отже, умова поділу на апертури у вигляді  $x_{\xi, y+r} \in [0.8 \cdot a_\xi; 1.2 \cdot a_\xi]$ . В результаті, для зображення 1, представленого на рис. 1, з похибкою введення  $10e-1$ , було отримано наступний результат, який представлено на рис. 6.



Рис. 1. Вихідне зображення 1, рівномірна сегментація 8x8

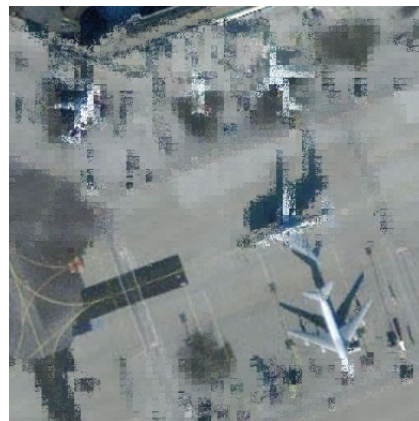


Рис. 2. Зображення 1 з введенням помилок 10e-4, рівномірна сегментація 8x8

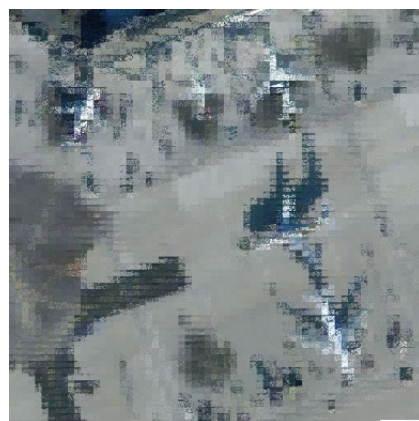


Рис. 3. Зображення 1 із введенням помилок 10e-3, рівномірна сегментація 8x8

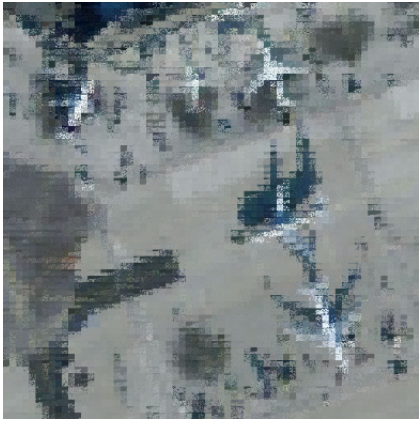


Рис. 4. Зображення 1 з введенням помилок  $10e-2$ , рівномірна сегментація  $8 \times 8$

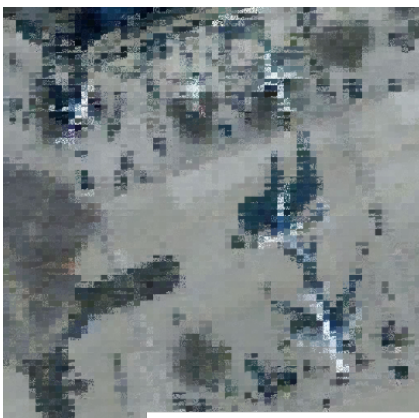


Рис. 5. Зображення 1 з введенням помилок  $10e-1$ , рівномірна сегментація  $8 \times 8$

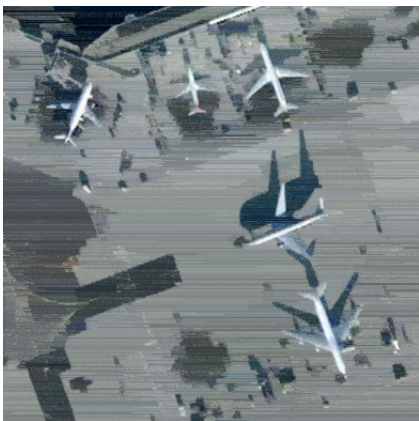


Рис. 6. Зображення 1 з введенням помилки  $10e-1$ , поділ апертури

Очевидно, що об'єкти на зображенні чітко розрізняються, що забезпечує практичне підтвердження ефективності використання апертурного поділу замість рівномірної сегментації. Цей ефект досягається тому, що алфавіт кожної апертури містить лише схожі пікселі. Навіть якщо в межах апертури виникає помилка та відновлення зображення після декодування гANS

може бути порушено, але значення пікселів залишаться в межах алфавітного діапазону, в результаті чого об'єкти на зображенні залишатимуться розрізними.

Таким чином, хоча гANS забезпечує високу ефективність стиснення даних, вона все ще залежить від правильного поділу зображення на фрагменти та побудови алфавітів з ймовірностями.

Розглянемо приклад зображення 1, представленого на рис. 1. Вихідний масив пікселів зображення, представлений у колірному просторі ARGB, має розмір 2 097 152 байти. Результат кодування за допомогою гANS з рівномірною сегментацією  $8 \times 8$  після запису у файл становить 406 181 байт. При простому поділі апертури з використанням розмір становить 395 595 байт, що суттєво не відрізняється від сегментації  $8 \times 8$ . Це пояснюється тим, що навіть пікселі з близькими значеннями все ще створюють свій власний символ в алфавіті, а різноманітність символів в алфавіті збільшує зростання стану на наступному кроці, особливо при великій кількості символів з низькими частотами. Однак, при виконанні апроксимації значень у межах кожної апертури:

$$x_{\xi,y} \in [0.9 \cdot a_{\xi}; 1.1 \cdot a_{\xi}]$$

ми отримуємо розмір 228 453 байти, при цьому візуальні характеристики об'єктів на зображенні залишаються незмінними, лише з незначним розмиттям в однорідних сегментах зображення. Цей результат вже перевершує результат для рівномірної сегментації  $8 \times 8$  і представлений на рис. 7.

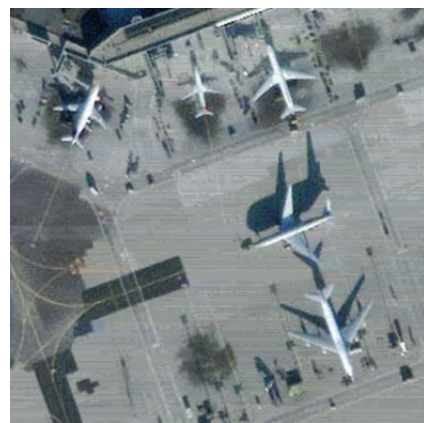


Рис. 7. Зображення 1 з поділом апертури та подальшою апроксимацією значення

Встановлення мінімальної довжини апертури також забезпечує значну перевагу в ефективності стиснення гANS, але також може призвести

до значного збільшення часу розрахунку, якщо мінімальна довжина апертури взята настільки великою. Те саме можна сказати і про великі коефіцієнти апертури. Крім того, в цьому випадку завадостійкість починає значно знижуватися.

**Висновки.** У цьому дослідженні запропоновано та практично підтверджено ефективність використання апертурного поділу замість рівно-

мірної сегментації для подальшого кодування rANS. Навіть проста реалізація апертурного поділу значно покращує завадостійкість кодovаних даних. Додавання апроксимацій у межах апертури підвищує ефективність подальшого кодування за допомогою ANS без шкоди для візуальної якості значних областей та об'єктів на зображенні.

#### Список літератури:

1. Метод групового кодування інфрачервоних зображень у спектрально-хвильовому просторі. Вісник НТУУ «КПІ». Серія Радіотехніка, Радіоапаратобудування. 2025. № 99. С. 24–34. DOI: 10.20535/RADAP.2025.99.24-34.
2. Ballé J., Laparra V., Simoncelli E.P. End-to-end Optimized Image Compression. 2016. DOI: 10.48550/arXiv.1611.01704.
3. Sun Y., Li L., Ding Y., Bai J., Xin X. Image Compression Algorithm Based On Variational Autoencoder. Journal of Physics: Conference Series. 2021. Vol. 2066. № 1. DOI: 10.1088/1742-6596/2066/1/012008.
4. Метод зберігання елементів службових складових криптокомпресійних кодограм відеозображень. Вісник НТУУ «КПІ». Серія Радіотехніка, Радіоапаратобудування. 2023. № 92. С. 28–40. DOI: 10.20535/RADAP.2023.92.28-40.
5. Rhatushnyak A., Wassenberg J., Sneyers J., Alakuijala J., Vandevenne L., Versari L., Obyrk R., Szabadka Z., Kliuchnikov E., Comsa I.-M., Potempa K., Bruse M., Firsching M., Khasanova R., van Asseldonk R., Boukourt S., Gomez S., Fischbacher T. Committee Draft of JPEG XL Image Coding System. DOI: 10.48550/arXiv.1908.03565.
6. Duda J. Asymmetric numeral systems: entropy coding combining speed of Huffman coding with compression rate of arithmetic coding. DOI: 10.48550/arXiv.1311.2540.
7. Barannik V., Onyshchenko R., Krasnorutsky A., Barannik D. The Methods of Intellectual Processing of Video Frames in Coding Systems in Progress Aeromonitor to Increase Efficiency and Semantic Integrity. 2022 IEEE 4th International Conference on Advanced Trends in Information Theory (ATIT). Kyiv, Ukraine. 2022. P. 53–56. DOI: 10.1109/ATIT58178.2022.10024208.
8. Krasnorutsky A., Barannik V., Shulgin S., Kolesnyk V., Kharchenko N., Revva K. Integration of Video Image Decryption Coding into a Remote Video Information Service. 2023 IEEE 5th International Conference on Advanced Information and Communication Technologies (AICT). Lviv, Ukraine. 2023. P. 1–6. DOI: 10.1109/AICT61584.2023.10452679.
9. Duda J., Niemiec M. Lightweight Compression with Encryption Based on Asymmetric Numeral Systems. Applied Mathematics and Computer Science. 2023. DOI: 10.34768/amcs-2023-0004.
10. Mamedov R. Analysis and Enhancement of Lossless Image Compression in JPEG-XL. 2024. DOI: 10.48550/arXiv.2404.19755.
11. Duda J. The Use of Asymmetric Numeral Systems Entropy Encoding in Video Compression. In: Distributed Computer and Communication Networks. 2019. Vol. 862. P. 125–139. DOI: 10.1007/978-3-030-36614-8\_10.
12. Технологія ковзного кодування нерівномірних діагональних послідовностей в двовимірному спектральному просторі трансформанти. Вісник НТУУ «КПІ». Серія Радіотехніка, Радіоапаратобудування. 2023. № 94. С. 13–23. DOI: 10.20535/RADAP.2023.94.13-23.

#### **Barannik D.V., Sidchenko Ye.S., Lykhodieiev O.S., Pomazuiev V.V., Fedorovskiy O.V. METHOD OF APPLYING APERTURES FOR ASYMMETRIC CODING VIDEO DATA SYSTEMS**

*The article considers the current problems of video image compression for information systems operating in conditions of limited or unstable communication channels. The shortcomings of classical and modern coding formats are analyzed, in particular, their low error resistance and non-adaptability to the structural complexity of images, in particular, the weakness of classical 8x8 segmentation, which does not take into account the features of the image, is shown on a practical example. The feasibility of using asymmetric number systems (ANS), in particular the rANS variant, as an effective entropy coding tool that combines high speed with approaching the theoretical compression limit is substantiated. The shortcomings of ANS are considered, namely, noise immunity and the dependence of the effectiveness of the methods on the input data and alphabets. To solve this problem, further in the article a method of image segmentation based on the detection of local coherence regions characterized by similar structural and visual properties is developed. An aperture division*

*mechanism is proposed, taking into account the amplitude variations of pixels and adaptive approximation of values inside the apertures. It is shown that such an approach allows to significantly increase the noise immunity to errors during data transmission, localizing their influence within individual image fragments. A comparative experimental study of the effectiveness of the proposed method in combination with rANS coding is carried out. The results obtained demonstrate a reduction in the volume of encoded data without significant loss of visual quality and a significant reduction in the scale of distortions when errors occur in the transmitted stream. Due to the consideration of the features of the image, when introducing channel errors, objects in the image remain clearly visible to the human eye. Experimental results of the study are provided on examples that confirm the effectiveness of the proposed approach in comparison with classical segmentation methods, which confirms the prospects of the methods for use in real time in systems with increased requirements for the reliability of video information.*

**Key words:** *aperture, coherence regions, asymmetric numeral systems, lossy compression, noise, noise immunity.*