

DOI: 10.18372/2310-5461.66.20332

УДК 004.622: 517.927

В. В. Бараннік, д-р техн. наук, проф.
Харківський національний університет
імені В. Н. Каразіна
orcid.org/0000-0002-2848-4524
e-mail: vvbar.off@gmail.com;

Ф.О. Устименко
Харківський національний університет
імені В.Н. Каразіна
orcid.org/0009-0004-9291-5183
e-mail: fydor240402@gmail.com;

В. В. Бараннік
Харківський національний університет радіоелектроніки
orcid.org/0000-0003-3516-5553
e-mail: valera462000@gmail.com;

І. О. Мільцин
Харківського національного університету Повітряних Сил
orcid.org/0009-0004-2305-6207
e-mail: hhhoshimin@gmail.com;

О. К. Юдін, д-р техн. наук, проф.
Державний Науково-дослідний інститут технологій кібербезпеки
orcid.org/0000-0002-6417-0768
e-mail: yudin.ok8@gmail.com

МЕТОД СЕЛЕКТИВНОЇ ОБРОБКИ ВІДЕОСЕГМЕНТІВ ДЛЯ ПІДВИЩЕННЯ ЯКОСТІ ВІДЕОЗОБРАЖЕНЬ В СИСТЕМАХ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ

Вступ

Останніми роками активно розвиваються технології автоматизації, які змінюють підхід до обробки інформації у різних сферах – від промисловості до наукових досліджень. Одним із напрямів, що отримав особливу увагу, є інтелектуальні системи обробки зображень. Такі системи знаходять застосування в контролі якості, медичній діагностиці, транспорті, безпеці та інших галузях, де важлива швидка та точна обробка візуальної інформації [1, 2].

Попри зростання популярності, ефективність інтелектуальних систем обробки зображень усе ще обмежена низкою технічних чинників. Одним із головних є залежність якості роботи таких систем від характеристик вхідних зображень. Зниження роздільної здатності, цифровий шум чи спотворення даних можуть суттєво вплинути на результати аналізу, ускладнюючи або спотворюючи процес прийняття рішень [3, 4].

Ці обмеження стають особливо помітними під час стискання зображень, яке є невід'ємною частиною обробки та передачі відеоінформації. Хоча сучасні алгоритми демонструють певні успіхи у

зменшенні обсягів даних, вони досі не забезпечують необхідного рівня компромісу між коефіцієнтом стиснення та якістю відновленого зображення. Особливо це критично для інтелектуальних систем, де навіть незначна втрата візуальної інформації може знизити точність розпізнавання об'єктів, аналітики або автоматичного прийняття рішень.

З цього виходить актуальна **науково-прикладна задача** – підвищення якості відеозображень в процесі кодування та передачі з використанням інтелектуальних технологій обробки інформації.

Аналіз останніх досліджень та публікацій

Інтелектуальні системи обробки зображень відіграють ключову роль у сучасному світі. Завдяки інтеграції комп'ютерного зору, машинного навчання та штучного інтелекту ці системи здатні аналізувати, розпізнавати та інтерпретувати великі обсяги візуальної інформації.

Такі системи широко використовуються в різних сферах, забезпечуючи ефективні інструменти для автоматизації та вдосконалення численних процесів. Вони сприяють підвищенню продуктивності, зменшенню ризиків і досягненню більшої

точності результатів. З часом ці технології стають дедалі потужнішими, розширюючи свої можливості та змінюючи наше сприйняття й використання візуальної інформації [5, 6].

У зв'язку з важливістю цих систем у сучасному житті, вони стали предметом нашого дослідження. Проте інтелектуальні системи обробки зображень виявляють значну чутливість до якості вхідних даних. Рівень якості зображень безпосередньо впливає на їхню продуктивність і точність розпізнавання.

Традиційні підходи до покращення якості зображень, як-от фільтрація шуму, усунення артефактів або підвищення роздільної здатності, хоч і ефективні у постобробці, не вирішують проблеми втрати інформації під час первинного стиснення. Ця проблема стимулює розвиток селективних або адаптивних підходів до обробки відеозображень, що базуються на ідентифікації та збереженні найбільш інформативних ділянок зображення.

Серед сучасних підходів заслуговує на увагу робота «Агент навчання з підкріпленням на рівні блоків для контекстно-орієнтованого стиснення відео» [7], у якій було запропоновано метод RL-RC-DoT – агент глибокого навчання, що динамічно регулює параметри компресії на рівні блоків кадру, орієнтуючись на downstream-задачі, як-от розпізнавання об'єктів. Згідно з результатами, система демонструє покращення показників precision і recall без суттєвого зниження PSNR. Цей підхід підтримує ідею збереження високої якості саме інформативних сегментів, залишаючи фонові ділянки з більш агресивною компресією.

Схожі результати були отримані у дослідженні «Стиснення відео на основі глибокого навчання» [8], де глибокі нейронні мережі використовувалися для ідентифікації областей інтересу (ROI) на кадрах відео. Ці області потім стискалися з мінімальними втратами, що дозволяло досягати виграшу до 6 % BD-rate у порівнянні з VVC без втрати важливої візуальної інформації.

Інший підхід запропонує компресивне вимірювання з адаптацією до інформативних зон зображення для цілей відеоспостереження [9]. Метод дозволяє розділяти передній та фоновий план за допомогою глибокої оборотної нейромережі, оптимізуючи компресію з урахуванням змістовної насиченості сегментів. Застосування таких стратегій підтверджує ефективність ідеї селективної компресії, де важливі блоки зберігаються у вищій якості.

При цьому важливо враховувати і методи оцінки якості зображень. Як показано у великому огляді робіт за темою [10], традиційні метрики, такі як PSNR або SSIM, не завжди відображають

сприйману якість зображення. Гібридні підходи, засновані на машинному навчанні (наприклад, MOVIE, VMAF), краще відображають вплив селективної компресії на об'єктивну точність і суб'єктивне сприйняття відео.

Це підтверджується і класичною роботою Вінклера і Моандаса [11], де простежується еволюція метрик від PSNR до складних гібридних моделей, здатних враховувати просторову та часову інформативність відеосегментів. Саме такі метрики є найбільш релевантними при оцінці ефективності селективної обробки.

Зі свого боку, у опитуванні «Розуміння відео за допомогою великих мовних моделей» [12] демонструють, що навіть великі мовні моделі (LLM), натреновані на відеоаналіз, можуть ефективно виявляти семантично значущі сегменти кадру – підходи, які можна адаптувати для маркування областей інтересу перед компресією або інтерпретацією.

Таким чином, аналіз сучасної наукової літератури підтверджує доцільність розвитку методів селективної обробки зображень. Вони дозволяють не лише зберігати якість критично важливих блоків, а й оптимізувати загальну ефективність систем інтелектуального аналізу зображень в умовах обмеженого пропускнуго каналу або обчислювальних ресурсів.

Постановка проблеми

Сучасні технології дозволяють використовувати різноманітні методи підвищення якості зображень, спрямовані на усунення небажаних ефектів або покращення сприйняття отриманого візуального матеріалу. До таких методів відносяться, зокрема, фільтрація шуму, яка знижує шумові артефакти, що виникають через різні фактори, такі як низька якість сенсорів чи умови зйомки. Іншим поширеним підходом є усунення артефактів, що виникають через алгоритми стиснення, наприклад, JPEG, які можуть проявлятися у вигляді блокових структур або спотворень кольорів.

Крім того, велика увага приділяється методам підвищення роздільної здатності, що дозволяють покращити деталізацію зображення, особливо якщо воно спочатку має низьку роздільну здатність [13, 14]. Однак, хоча ці методи ефективні для постобробки зображень, вони не вирішують основної проблеми, яка виникає на початковому етапі – стисненні зображень. Стиснення зображень, особливо з втратами, необхідне для зменшення обсягу даних, але водночас воно неминуче призводить до втрати інформації, що знижує якість зображення. Такі втрати стають особливо критичними в тих випадках, коли потрібно

аналізувати або відновлювати деталі, важливі для подальшої обробки чи використання.

Тому, **метою** даної роботи є розробка селективного методу обробки відеозображень з метою підвищення їх якості в інтелектуальних технологій обробки інформації. Суть запропонованого підходу полягає у використанні інтелектуальних технологій обробки інформації, які дозволяють адаптивно реагувати на особливості кожного конкретного зображення чи його сегмента.

Обґрунтування необхідності розробки методу селективної обробки сегментів зображення

Для усунення недоліків існуючих методів пропонуємо розглянути новий метод селективної обробки зображень, який ґрунтується на відборі значущих сегментів зображення. Будь-який сегмент зображення можна позначити як $B(u, v)$ де (u, v) - координати сегменту у зображенні.

Цей сегмент може належати до однієї з трьох множин[15-17]: множини значущих сегментів

На рис. 1 наведено схематичне зображення розподілення кластерів інформативності на обраному тестовому зображенні.

зображення, далі позначених як $\{B_0, B_1, B_2, \dots, B_\psi\} = B_{inf}$,

де номер ψ останнього елементу у множині, а B_{inf} позначає усю множину;

– множини не-інформативних сегментів, яку далі ми будемо позначати як $\{B_0, B_1, B_2, \dots, B_k\} = B_{noinf}$,

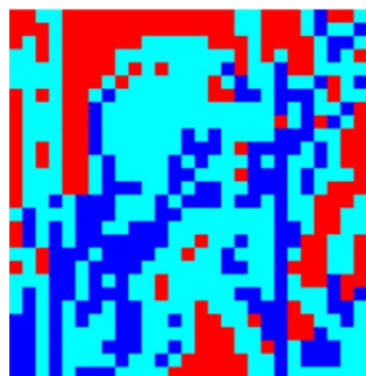
де k номер останнього елементу у множині, а B_{noinf} означає усю множину;

– множини сегментів змішаного значення, далі позначену як $\{B_0, B_1, B_2, \dots, B_f\} = B_{mix}$, де f номер останнього елементу у множині, а B_{mix} позначає усю множину.

Для демонстрації виділення кластерів сегментів зображення відповідно до рівня їх інформативності, а також для подальшого тестування існуючих та розробленого методів виділення інформативних кластерів зображення було обрано зображення відоме як «Лена» (Lenna) – один із найпопулярніших тестових образів у галузі обробки зображень[18].



a



b

Рис. 1. Схематичне позначення відношення між множинами $B_{(u,v)}$, B_{noinf} , B_{mix} , B_{inf} :
a – оригінальне зображення; b – схематичне виділення зображення інформативності сегментів:
синій – позначає найбільш інформативні сегменти; голубий – сегменти середньої інформативності;
червоний – не інформативні сегменти

Для коректного розподілу сегментів зображення на множини необхідно спочатку визначити значущі, тобто інформативні, сегменти. Це потребує виявлення тих областей зображення, які мають найбільшу структурну та семантичну вагу. На сьогодні існує кілька підходів до оцінки інформативності та семантичної ваги сегментів зображення.

Одним із таких підходів є порогування, що являє собою простий, але ефективний метод аналізу зображень. Він використовується для виділення об'єктів або областей інтересу з фону. Суть методу полягає у визначенні граничного значення або діапазону значень, які дозволяють відокремити передній план (об'єкти чи області

інтересу) від фону, що включає решту елементів зображення [19,20].

Математична формула для визначення порогу для зображення у градієнті сірого:

$$G(u, v) = 1 \text{ if } F(u, v) \geq \tau, 0 \text{ if } F(u, v) < \tau,$$

де $F(u, v)$ – інтенсивність пікселя в позиції (u, v) у вхідному зображенні, τ – порогове значення, а $G(u, v)$ – вихідне зображення.

Однією з основних труднощів при визначенні порогів є врахування шуму та нерівномірного освітлення. У таких умовах порогове значення може призводити до помилкових результатів: хибних спрацьовувань, коли пікселі переднього плану насправді належать до фону, або хибних

негативів, коли пікселі фону насправді є частиною переднього плану.

Метод виявлення контурів широко використовується для аналізу зображень, оскільки дозволяє визначати межі об'єктів або областей інтересу. Контури вказують на різкі зміни інтенсивності, такі як межі між об'єктами чи зміни текстури, і знаходять застосування в задачах розпізнавання об'єктів, сегментації та виділення ознак [21, 22]. Першим етапом цього процесу є обчислення градієнта зображення, що показує зміну інтенсивності пікселів. Для цього використовують різні оператори, наприклад, оператори Собеля, Превітта або Робертса.

Оператор Собеля є популярним вибором, оскільки обчислює горизонтальні та вертикальні градієнти окремо за допомогою двох 3×3 згорткових ядер. Отримані горизонтальні та вертикальні градієнти комбінуються для визначення

величини градієнта та напрямку краю, що допомагає ідентифікувати межі об'єктів із високою точністю. Математична формула для оператора Собеля:

$$D_u = [-101; -202; -101], \quad D_v = [-1 \ -2 \ -1; 000 \ 121 \ 1]:$$

$$D_{\text{mag}} = \sqrt{D_u^2 + D_v^2}; \quad D_{\text{dir}} = \arctan\left(\frac{D_v}{D_u}\right),$$

де D_u і D_v – горизонтальний і вертикальний градієнти, D_{mag} – величина градієнта, D_{dir} – напрямок краю.

На рис. 2 на прикладі тестового зображення наведено те, як працює градієнтний метод у контексті відбору інформативних сегментів зображення. Для цього прикладу, чорний колір використовується як маска для неінформативних сегментів зображення, а білий – як маска для інформативних сегментів зображення.

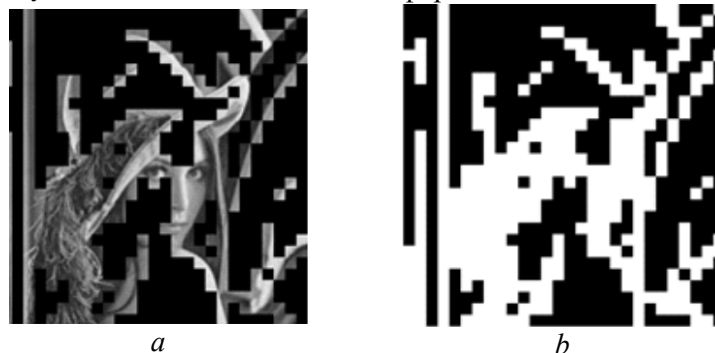


Рис. 2. Ілюстрація роботи методу градієнтного відбору інформативних сегментів зображення:

а – зображення із видаленими неінформативними кластерами;

б – маска інформативних сегментів зображення

Виявлення контурів є ефективним методом аналізу зображень, що дозволяє визначити межі об'єктів або цікавих областей. Однак цей підхід має свої обмеження, серед яких чутливість до шуму та необхідність точного вибору параметрів [23, 24].

Виділення ознак є іншим важливим методом, який використовується для вилучення значущої інформації або характеристик із зображення. Ці ознаки можуть слугувати основою для таких завдань, як розпізнавання об'єктів, класифікація або кластеризація. Ознаками можуть бути вимірювані властивості об'єкта або області, наприклад форма, текстура, колір або інтенсивність. Першим етапом у виділенні ознак є попередня обробка зображення, яка спрямована на покращення або нормалізацію властивостей, що становлять інтерес. Попередня обробка може включати фільтрацію, згладжування або нормалізацію даних.

Наступним етапом є виявлення цікавих елементів зображення, що може бути здійснене за допомогою різних методів, таких як виявлення

плям, кутів або країв. Виявлення плям використовується для ідентифікації областей із однорідною інтенсивністю або кольором. Цей метод може бути реалізований за допомогою таких інструментів, як фільтри Лапласа-Гаусса (L) або різниці Гаусса (D) [25–27]. Фільтр L є двовимірним ядром згортки, яке дозволяє виявляти плями на зображенні в різних масштабах. Математична формула для фільтра L :

$$L(u, v) = -\frac{1}{\pi\alpha^4} \left(1 - \frac{u^2 + v^2}{2\alpha^2}\right) \cdot \exp\left(-\frac{u^2 + v^2}{2\alpha^2}\right),$$

де $L(u, v)$ – відповідь фільтра L у позиції (u, v) , α – стандартне відхилення фільтра Гауса, а $\exp$ – експоненціальна функція.

На рис. 3 на прикладі тестового зображення наведено те, як працює метод виявлення контурів у контексті відбору інформативних сегментів зображення. Для цього прикладу, чорний колір використовується як маска для неінформативних сегментів зображення, а білий – як маска для інформативних сегментів зображення.

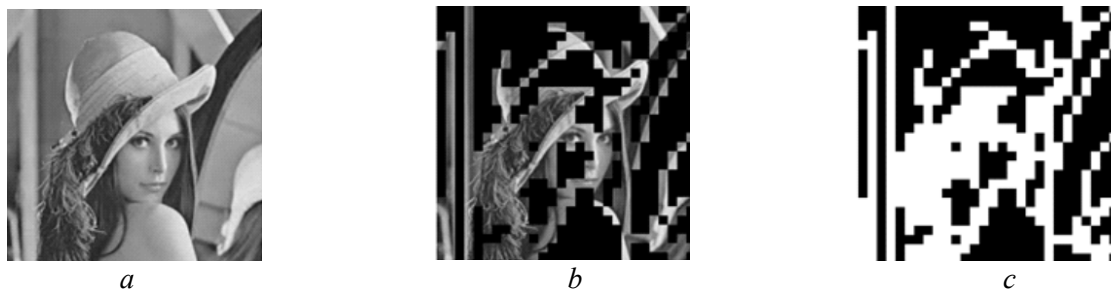


Рис. 3. ілюстрація роботи методу виявлення контурів для відбору інформативних сегментів зображення:
a – оригінальне зображення; *b* – зображення із видаленими неінформативними кластерами;
c – маска інформативних сегментів зображення

Виявлення кутів — це метод виявлення кутів або точок інтересу на зображенні. Кути можна виявити за допомогою таких методів, як детектор кутів Харріса або детектор кутів Ши-Томазі. Кутівий детектор Harris вимірює зміну інтенсивності або градієнта навколо точки та обчислює функцію кутового відгуку. Математична формула кутового детектора Харріса: $Z = D(J) - k * (\text{trace}(J))^2$, де J – матриця структурного тензора, $D(J)$ – визначник J , $\text{trace}(J)$ – слід J , а k – константа.

Після виявлення ознак наступним етапом є їх опис у формі, яка зберігає значущість і залишається незмінною до певних перетворень, таких як обертання, зміна масштабу чи освітлення. Для цього використовуються різні методи, зокрема гістограма орієнтованих градієнтів (HOG), масштабно-інваріантне перетворення ознак (SIFT) та прискорена надійна функція (SURF).

HOG застосовується для опису локальної форми або текстури зображення шляхом створення гістограм орієнтованих градієнтів у різних напрямках. Для обчислення цього дескриптора зображення розбивається на комірки, у кожній з яких визначаються орієнтації та величини градієнтів, після чого гістограми з усіх комірок об'єднуються. SIFT слугує для опису локальної текстури або зовнішнього вигляду зображення через масштабно-інваріантні ключові точки. Дескриптор створюється шляхом поділу області навколо кожної ключової точки на підобласті, розрахунку орієнтацій і величин градієнтів у цих підобластях і об'єднання отриманих гістограм. SURF використовується для опису текстури або зовнішнього вигляду зображення шляхом виділення ключових точок, які залишаються інваріантними до змін масштабу та обертання. Дескриптор SURF базується на обчисленні вейвлет-відповідей Хаара в різних масштабах і орієнтаціях, які потім об'єднуються у вектор для подальшого аналізу [28, 29].

Після опису ознак їх необхідно представити у відповідному форматі для подальшого аналізу. Це може включати перетворення ознак у вигляді векторів або гістограм, їх нормалізацію, а також застосування методів зменшення розмірності, таких як аналіз головних компонент (PCA) або лінійний дискримінантний аналіз (LDA). Після виділення та представлення ознак важливо оцінити їхню ефективність у контексті розглядуваної проблеми. Це може полягати у визначенні дискримінаційної спроможності, надійності чи обчислювальної ефективності функцій. Також може бути необхідним порівняння продуктивності різних наборів ознак або методів їх вилучення. Варто зазначити, що на основі оцінки продуктивності може знадобитися повторний вибір ознак з використанням більш обґрунтованих критеріїв. Наприклад, якщо деякі ознаки виявляються зайвими або нерелевантними, їх можна видалити з набору. До того ж, якщо нові ознаки виявляються більш дискримінаційними або інформативними, їх можна додати до набору функцій [30–32]. В цілому, виявлення ознак здатне ефективно ідентифікувати та визначати конкретні характеристики зображення, такі як текстура, форма та колір, проте це потребує ретельного підбору відповідних функцій для конкретного завдання та значних ресурсів. Усі ці методи, хоч і є досить ефективними, мають свої недоліки, які роблять їх використання у нашому випадку небажаним. Наприклад, у всіх зображеннях, оброблених за допомогою технологій, заснованих на цих методах, показник PSNR для інформативних сегментів зображення значно нижчий і залишається незмінним зі збільшенням коефіцієнта компресії (позначеного на графіку як K_{com}), у порівнянні з неінформативними та змішаними сегментами. На рис. 4 наведено гістограму залежності рівня PSNR від рівня інформативності сегменту зображення.

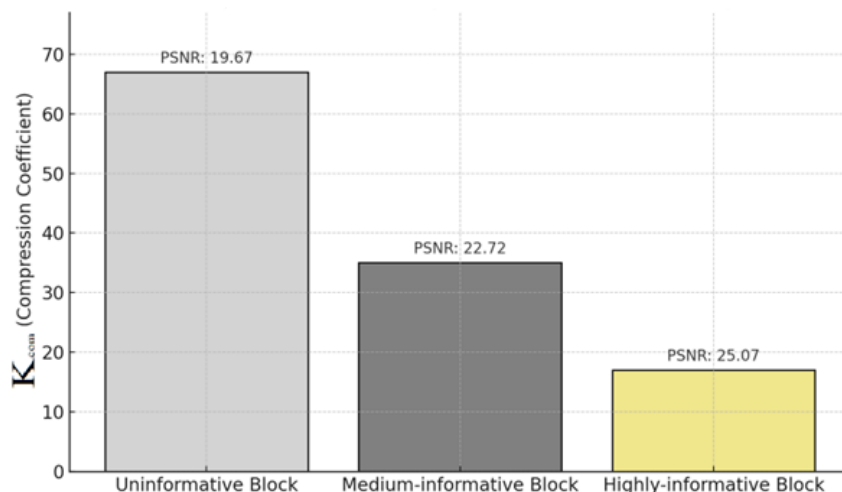


Рис. 4. Залежність зміни пікового відношення сигнал/шум від коефіцієнта компресії для зображення з урахуванням різного ступеня інформативності блоків з відповідним рівнем достовірності

На графіку зображено залежність між коефіцієнтом стиснення K_{com} та середнім значенням PSNR для блоків з різним рівнем інформативності. Найвищий коефіцієнт стиснення досягається для неінформативних блоків, однак при цьому спостерігається найнижче значення PSNR (19.67 дБ), що вказує на значну втрату якості. Навпаки, сильнонасичені блоки демонструють найвище значення PSNR (25.07 дБ), але стиснення є найменш ефективним (лише 17 за K_{com}). Це свідчить про ефективність адаптивного підходу до стиснення, при якому менш важливі області

зображення можуть бути стиснуті агресивніше, тоді як ключові сегменти потребують збереження високої точності. Такий підхід дозволяє досягти балансу між ефективністю стиснення та якістю відновленого зображення.

Даний показник також може впливати на функціонування систем штучного інтелекту. Так, наприклад, на рис. 5, наведена залежність показнику повноти (англ. recall), який показує, яку частку з усіх дійсно існуючих об'єктів модель знайшла в задачах класифікації, пошуку або детекції від показнику PSNR:

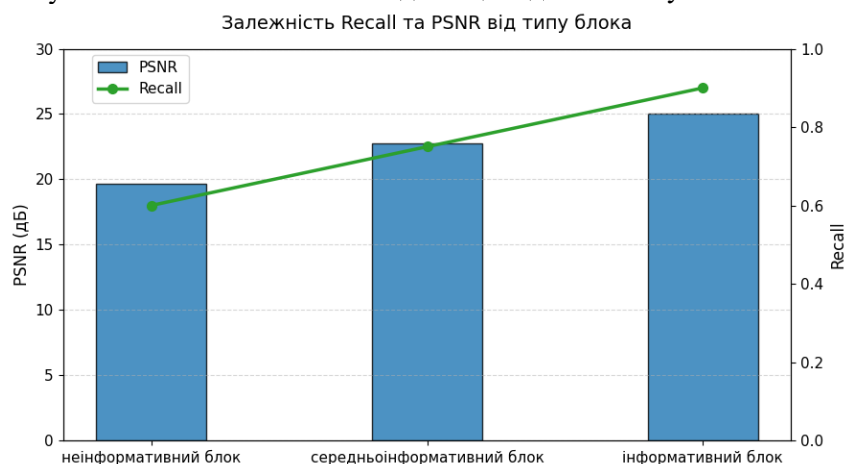


Рис. 5. Залежність показника повноти від пікового відношення сигнал/шум від коефіцієнта компресії для зображення з урахуванням різного ступеня інформативності блоків

На графіку показано залежність середнього значення PSNR та показника Recall від типу блоку зображення. Видно, що зі збільшенням інформативності блоків зростають як PSNR, так і Recall. Неінформативні блоки мають найнижчі значення обох показників, що свідчить про втрату точності та нижчу якість реконструкції внаслідок агресивнішого стиснення. І навпаки, інформативні блоки

демонструють найвищі результати, що вказує на збереження високої точності розпізнавання та відновлення. Така тенденція підтверджує доцільність селективного підходу до обробки блоків: важливіші ділянки повинні оброблятися із меншою втратою інформації. Це дозволяє забезпечити баланс між якістю реконструкції та ефективністю аналізу.

З метою вирішення перелічених нами проблем, у цій роботі буде розроблено інший метод визначення ступеню інформативності сегментів відеозображення, заснований на інтенсивності та діапазоні зміни кольорів у межах певного сегменту $B(u, v)$ [33,34]

Розробка методу селективної обробки сегментів зображення

Цей метод ґрунтується на припущенні, що частини зображення, які демонструють найвищу інтенсивність кольорів і широкий діапазон їх змін у межах певного сегмента, є найбільш інформативними.

Показники інтенсивності та діапазону зміни кольорів обрані з двох причин. По-перше, ці характеристики присутні в будь-якому сегменті зображення, незалежно від його розташування чи інших особливостей. По-друге, значення цих показників безпосередньо відображають інформативну та семантичну вагу сегмента, і навпаки – інформаційна та семантична вага сегмента безпосередньо впливає на ці характеристики.

Для спрощення порівняння сегментів зображення всі сегменти розглядаються у відтінках сірого [35–37]. Існує кілька способів переведення кольорового зображення у відтінки сірого, найпоширенішим із яких є обчислення середнього арифметичного значення кольорових компонентів. Формула для цього виглядає так: $Gr = (R + G + B)/3$, де R , G та B – значення червоного, зеленого та синього каналів відповідно, а Gr – відтінок сірого. Ця формула припускає, що всі канали мають однакову вагу.

Якщо ж необхідно врахувати різну вагу каналів, можна застосувати іншу формулу:

$$Gr = (R * 0.2989 + G * 0.5870 + B * 0.1140),$$

де 0.2989, 0.5870 та 0.1140 – це коефіцієнти вагомості каналів R , G та B відповідно.

Для класифікації інформативних сегментів зображення вводяться три нові поняття. Інформативні блоки — це сегменти зображення з великим діапазоном зміни кольорів у межах одного блоку. Змішані блоки представляють собою сегменти, у яких відмінності між пікселями незначні. Фонові блоки характеризуються або відсутністю змін кольорів, або діапазон цих змін не перевищує одного чи двох рівнів.

Однак проста класифікація блоків із використанням довільного методу визначення порогів не забезпечує достатньої надійності й стабільності. Для цього необхідно розробити метрику, яка дозволить порівнювати сегменти між собою та чітко розподіляти їх за різними класами.

Використовуючи дискретне косинусне перетворення (ДКП), можна більш наочно відобразити відмінності між сегментами зображення. ДКП дає

змогу перейти від просторово-часового представлення відеокадру до просторово-спектрального. Компоненти трансформанти ДКП виступають як інтегральні характеристики структурного змісту фрагмента зображення [38, 39]. Інтегральна залежність компонент трансформанти ДКП визначається так: у верхньому лівому куті трансформанти значення компонент пропорційні середній яскравості зображення. Вони відображають рівень насиченості блока низькочастотними перепадами, що можуть проявлятися у вигляді ступінчастих змін рівня яскравості або координат кольору. Компоненти, розташовані в середній частині трансформанти, відображають ступінь насиченості блоку зображення лінійними та рівномірними змінами рівня яскравості [40].

Виходячи з цього, можна зробити припущення, що чим більші значення у верхньому лівому куті трансформанти, тим інформативнішим загальом є кадр, оскільки в ньому наявні різкі перепади кольорів і яскравості. У подальшій роботі використовуватимуться такі позначення: $M(u, v)$ – трансформанта сегмента, $m(u, v)_{\sigma\phi}$ – компонент трансформанти з відповідним індексом.

Для визначення показника, який би слугував мірилом для кожної $Y(\alpha, \beta)$, пропонується розглядати трансформанту не як матрицю коефіцієнтів, а як матрицю чисел. Кожне з цих чисел може бути представлене у вигляді послідовності нулів і одиниць, тобто в тому форматі, в якому вони зберігаються у пам'яті комп'ютера. Наприклад, число 196 записується як 11000100, а число 10 – як 00001010. Загалом, чим більше число, тим більше одиниць у його бітовому представленні.

З цієї ідеєю трансформанта $M(u, v)$ перетворюється на бінарний куб, де кожен коефіцієнт буде представлений у бітовому форматі. Спочатку коефіцієнти $M(u, v)$ округлюються і беруться за модулем, що дозволяє отримати значення, придатні для переведення у байтовий формат, а також ігнорувати зайві нулі. Перетворена таким чином трансформанта позначається як $\{M(u, v)\} \in N$.

Після цього, перетворимо кожний $m(u, v)_{\sigma\phi}$ у байтовий вигляд, далі позначений як $[m(u, v)_{\sigma\phi}]_2$, відповідно, трансформанта, що складається з таких компонентів буде позначена як буде позначена як $[M(u, v)]_2$

Результатом цього стане куб, або, трьох вимірний масив, що складається з нулів та одиниць. Кожна площина цього кубу є байтовим представленням кожного з рядків коефіцієнтів трансформанти ДКП. Далі, у цій роботі кожна площина буде позначена як $M(u, v, \tau)_{\sigma\phi}$.

На рис. 6 наведено схематичне представлення трансформації довільного сегменту зображення із матриці коефіцієнтів ДКП до бінарного кубу.

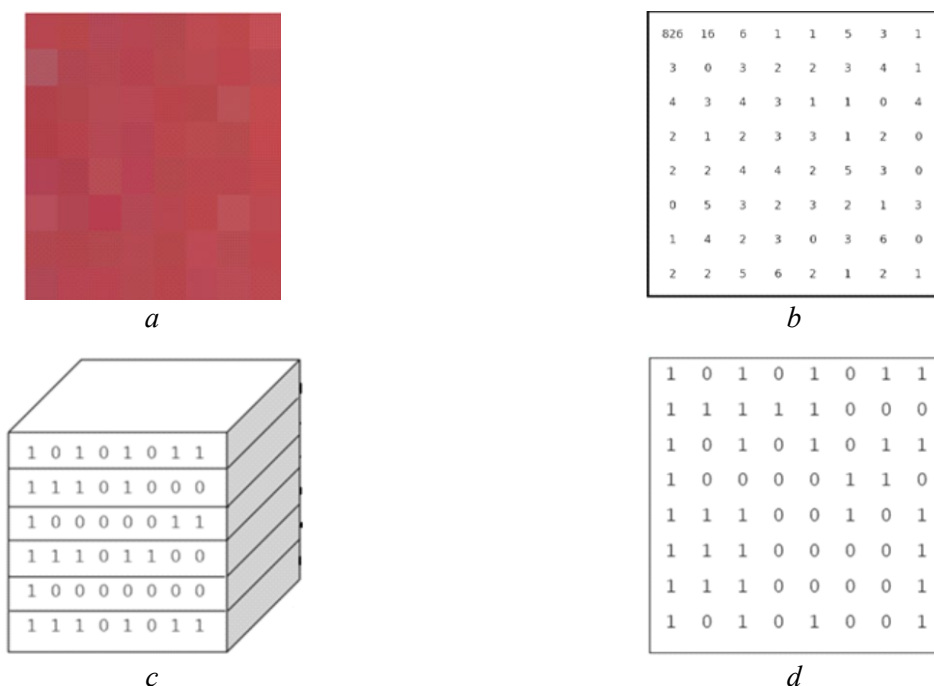


Рис. 6. Схема сегменту зображення для селективного відбору сегментів:

a – оригінальний сегмент тестового зображення, розміром 8×8 пікселів; *b* – ДКП трансформанта сегменту зображення; *c* – бітве представлення трансформанти ДКП у форматі куба; *d* – площина бітового кубу трансформанти ДКП

У цій роботі для позначення будь-якого елементу площини куба $M(u, v, \tau)_{\sigma\phi}$ було прийнято позначення $[m(u, v, \tau)_{(\sigma, \phi)}]_2$. На рис. 7 наведено наочна ілюстрація схеми відбору сегменті зображення для селективного відбору сегментів за їх інформативністю.

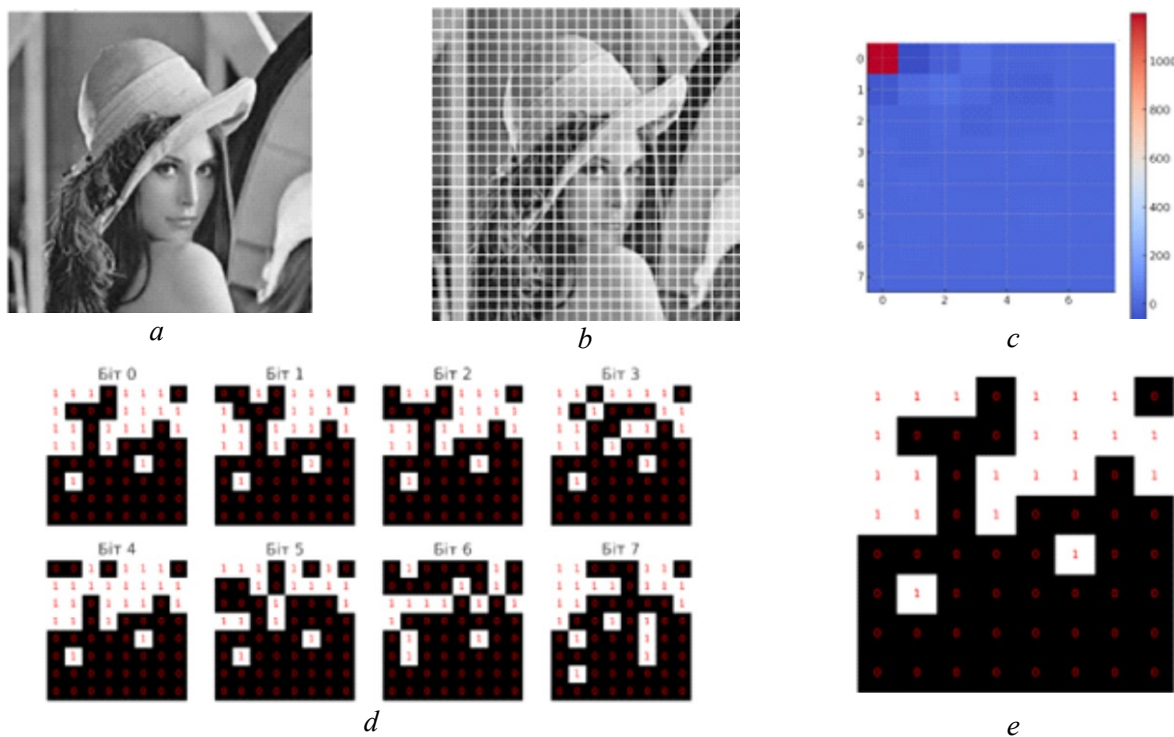


Рис. 7. Деталізація схеми відбору сегменті зображення для селективного відбору сегментів за їх інформативністю: *a* – оригінальне зображення; *b* – зображення поділені на блоки 8×8 перед застосуванням ДКП; *c* – матриця коефіцієнтів ДКП для довільного сегменту зображення; *d* – площини бітового кубу матриці коефіцієнтів ДКП; *e* – довільна площина бітового кубу

Тепер визначимо кількість серій одиниць для кожного $M(u, v, \tau)_{\sigma\phi}$. Як уже відомо, чим більше число, тим довшою буде нерозривна послідовність одиниць у його бінарному представленні. Цю послідовність будемо називати серією одиниць або позначати як ω .

Параметр ω визначає довжину найдовшої серії одиниць у бінарному представленні числа. Чим довшою є така серія, тим більше значення має цей параметр. Для кожного елемента матриці обчислюється ω , і отримані значення можуть

використовуватися для подальшого аналізу або класифікації елементів залежно від їхньої інформативності.

Визначення серій одиниць таким способом надає числову характеристику, що безпосередньо відображає інформативність елементів матриці. Довші серії одиниць зазвичай вказують на вищу інтенсивність змін або складнішу структуру в сегменті зображення. Це дозволяє використовувати параметр ω для поділу сегментів на інформативні та неінформативні.

На рис. 8 наведена блок-схема для розрахунку значення ω :

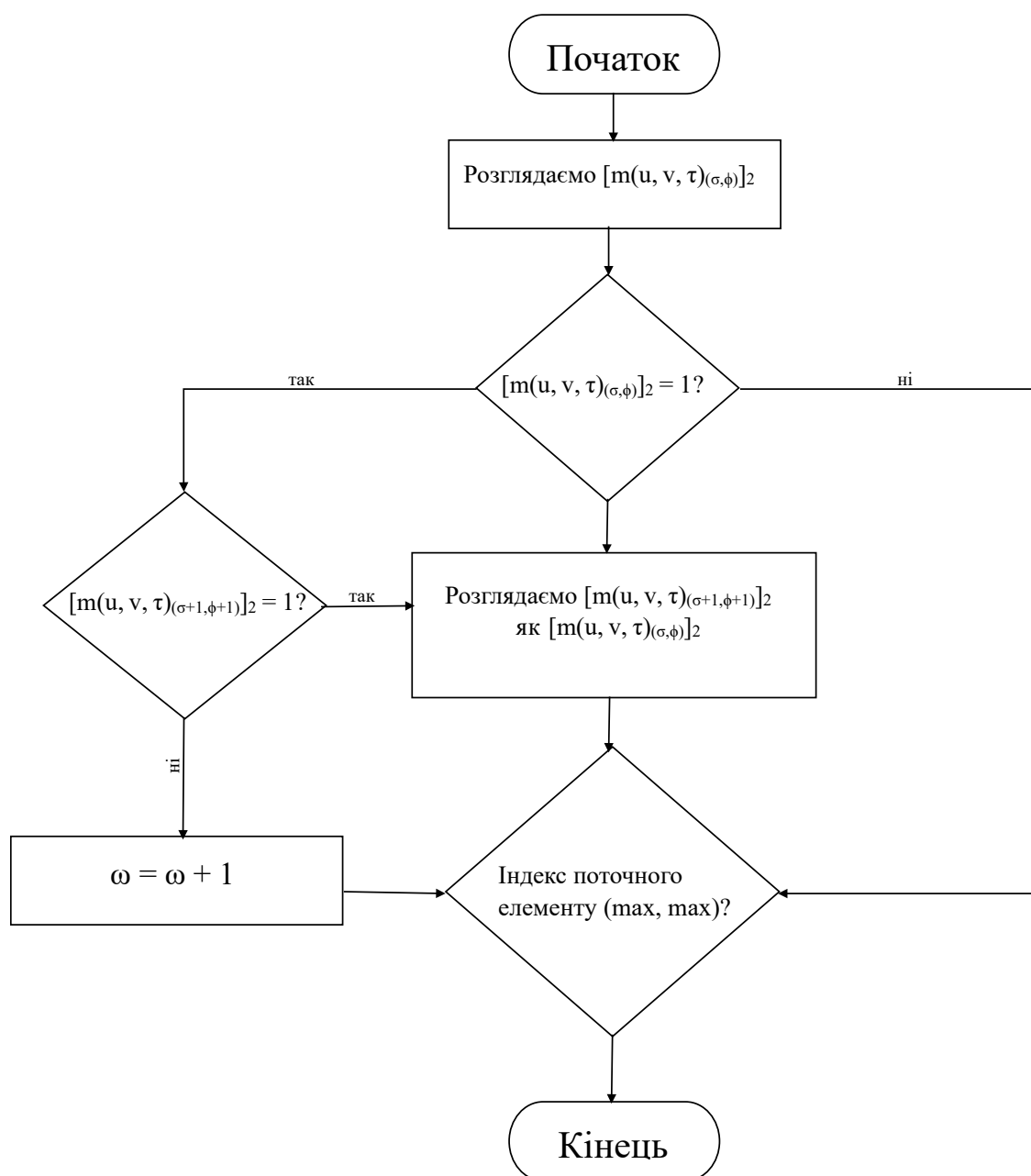


Рис. 8. Схема розрахунку нерозривної послідовності одиниць у байтовому представленні числа: $[m(u, v, \tau)_{(\sigma, \phi)}]_2$ – означає будь-якого елементу площини куба $M(u, v, \tau)_{\sigma\phi}$; ω – кількість серій одиниць для кожного $M(u, v, \tau)_{\sigma\phi}$ – байтовий куб, що є представленням довільного сегменту зображення

Таким чином, чим більше серій та чим вищі значення ω для кожного з $[m(u, v, \tau)(\sigma, \phi)]_2$, тим більші значення $m(u, v)\sigma\phi$ у відповідному рядку $M(u, v)$. У нашому попередньо розглянутому прикладі кількість серій одиниць для кожного рядка становить [14, 12, 8, 11, 7, 10, 7, 9].

Наступним кроком є визначення метрики, яка зможе описати ці показники для кожного сегменту. Бажано, щоб ця метрика виражалася одним числом і відображала загальний стан всієї матриці $M(u, v)$. У нашому дослідженні для цього обрано середнє арифметичне значення. Наприклад, для розглянутого вище прикладу середнє арифметичне значення ω дорівнює 9.75, яке ми позначимо як ω_{avg} .

Після цього необхідно встановити порогові значення для сортування сегментів зображення на інформативні та неінформативні. Всі сегменти зображення, для яких значення ω дорівнює або перевищує це порогове значення, будуть класифіковані як інформативні. Це порогове значення ми позначимо як $\Theta(\omega_{avg})_{inf}$.

Тобто, якщо для сегменту $B_{u,v}$: $(\omega_{avg} \geq \Theta(\omega_{avg})_{inf}) = \text{тоді}, B_{u,v} \in \{B_0, B_1, B_2, \dots, B_\Psi\} = B_{inf}$

Маючи ці порогові значення, ми можемо поділити сегменти зображення на інформативний та неінформативний кластери.

Висновки

1. Обґрунтовано підхід до підвищення якості відеозображень для інтелектуальних систем обробки, що базується на: селективному відборі сегментів відеозображення за рівнем їх інформативного та семантичного навантаження. Це дозволяє покращити якість важливих областей зображення, адаптуючись до специфічних вимог обробки; застосуванні спектрального аналізу та статистичних методів для оцінки і класифікації сегментів зображення. Це дозволяє виділяти найбільш значущі сегменти для подальшої обробки та зменшення надмірності.

2. Досягнуто покращення якості відеозображень шляхом: зниження впливу низькоінформативних і шумових областей на загальний результат обробки; забезпечення можливості точнішої класифікації сегментів зображення завдяки використанню адаптивних алгоритмів аналізу їхньої структури.

3. Розроблено новий метод селективної обробки сегментів відеозображень, який включає: впровадження метрики для оцінки інформативності сегментів на основі інтенсивності та діапазону кольорів у межах сегмента; програмне забезпечення для автоматизованого аналізу, класифікації

та відбору сегментів зображень, що дозволяє оптимізувати їх подальшу обробку.

Це дозволяє підвищити ефективність інтелектуальних систем обробки зображень шляхом адаптації до специфіки вхідних даних та збереження якості важливих сегментів при стисканні.

ЛІТЕРАТУРА

- [1] Бараннік В. В., Тарнополов Р. В., Власов А. В. Модель загроз безпеки відеоінформаційного ресурсу систем відеоконференцв'язку. *Наукоємні технології*. 2014. № 1 (21). С. 55–60.
- [2] Barannik V. V., Komolov Dm., Musienko A. P., Tarnopolov R. V. Methodological basis for determining the energy significance of the structural unit of a video frame based on the estimation of low-frequency components of the matrices of the DCT blocks of the luminance component. *Modern problems of radio engineering, telecommunications and computer science (TCSET'2016): materials of the XIV Internat. conf., (Lviv–Slavske, Ukraine, February, 22–26, 2016)*. Lviv–Slavske, 2016. P. 572–574.
- [3] Babenko Yurii, Barannik Dmitry, Balzhan Smailova, Hahanova Anna, Yroshenko Valerii, shmakov Vitalii, Shaikhanova Aigul, Veselska Olga, Karpiinski Mikolaj The Technology of Structural Classification of Video Frames in Intelligent Info-Communication Systems. “Development of technology analys for the content semantics,” in *Engineer of XXI Century - We Design the Future*, Bielsko-Biala, Poland: ATH, 2020. P. 122–132.
- [4] Бараннік В. В., Комолов Д. І. Методологія селективного захисту відеопотоку по базовим кадрам. *Інформаційно-управляючі системи на залізничному транспорті*. 2014. № 6. С. 69–77.
- [5] Бараннік В. В., Королева Н. А. Структурно-комбінаторне представлення даних в АСУ. Харків: ХУПС, 2009. 252 с.
- [6] Бараннік В. В., Тарнополов Р. В. Метод ефективного кодування розділених за комбінованою схемою аерофотознімків. *Радіoeлектроніка та інформатика*. 2016. № 2 (73). С. 33–37.
- [7] Uri Gadot, Assaf Shocher, Shie Mannor, Gal Chechik, Assaf Hallak RL-RC-DoT: A Block-level RL agent for Task-Aware Video Compression URL: https://www.researchgate.net/publication/388318023_RL-RC-DoT_A_Block-level_RL_agent_for_Task-Aware_Video_Compression
- [8] Aro Kim, Seung-taek Woo, Minho Park, Donghwi Kim, Hanshin Lim, Soon-heung Jung, Sangwoon Kwak & Sang-hyo Park Deep learning-guided video compression for machine vision tasks URL: <https://jivp-urasipjournals.springeropen.com/articles/10.1186/s13640-024-00649-w>

- [9] Lisha Gao, Zhoujun Ma, Shuo Han, Tiancheng Zhao, Qingcheng Liu, Zhangjie Fu ROI-aware video compressive sensing for surveillance URL: <https://www.mdpi.com/2078-2489/15/9/555>
- [10] Xiongkuo Min, Huiyu Duan, Wei Sun, Yucheng Zhu, Guangtao Zhai Perceptual Video Quality Assessment: A Survey URL: <https://arxiv.org/abs/2402.03413>
- [11] Stefan Winkler, Praveen Mohandas The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics URL: <https://www.researchgate.net/publication/3042136> The Evolution of Video Quality Measurement From PSNR to Hybrid Metrics
- [12] Yunlong Tang, Jing Bi, Siting Xu, Luchuan Song, Susan Liang, , Teng Wang, Daoan Zhang, Jie An, Jingyang Lin, Rongyi Zhu, Ali Vosoughi, , Chao Huang, Zeliang Zhang, Pinxin Liu, Mingqian Feng, Feng Zheng, , Jianguo Zhang, , Ping Luo, Jiebo Luo, and Chenliang Xu Video Understanding with Large Language Models: A Survey URL: <https://arxiv.org/html/2312.17432v5>
- [13] Бараннік В. В., Власов А. В. Модель загроз безпеки відеоінформаційного ресурсу систем відеоконференцв'язку. *Наукоємні технології*. 2014. № 3 (19). С. 299–304.
- [14] Бараннік В. В., Комолов Д. І. Метод суміщення кодової конструкції енергетично значимої структурної складової з вимогою методу блокового симетричного шифрування для закриття поточкових відеоданих на основі технології внутрікадрової селекції. *Наукоємні технології*. 2016. № 1. С. 39–47.
- [15] Бараннік В. В., Комолов Д. І., Селективний метод шифрування відеопотоку в телекомунікаційних системах на основі приховування базового І-кадру. *Наукоємні технології*. 2015. № 2. С. 69–77.
- [16] Бараннік В. В., Тарнополов Р. В., Власов А. В. Модель загроз безпеки відеоінформаційного ресурсу систем відеоконференцв'язку. *Наукоємні технології*. 2014. № 1 (21). С. 55–60.
- [17] Barannik V. V., Komolov Dm., Musienko A. P., Tarnopolov R. V. Methodological basis for determining the energy significance of the structural unit of a video frame based on the estimation of low-frequency components of the matrices of the DCT blocks of the luminance component. Modern problems of radio engineering, telecommunications and computer science (TCSET'2016): materials of the XIV Internat. conf., (Lviv–Slavske, Ukraine, February, 22–26, 2016). Lviv–Slavske, 2016. P. 572–574.
- [18] L. Torres, "The Lena image: The most widely used test image in image processing," *IEEE Signal Processing Magazine*, vol. 20, no. 5, pp. 105–107, 2003.
- [19] Bitstream-based JPEG Encryption in Real-time [Text] / S. Auer, A. Bliem, D. Engel, A. Uhl, A. Unterweger. *International Journal of Digital Crime and Forensics*. 2013. Vol. 5, Iss. 3. P. 1–14. DOI: 10.4018/jdcf.2013070101, 17.03.2023
- [20] Cryptographic and Information Security Approaches for Images and Videos [Text] / S. Ramakrishnan, et al. – CRC Press, 2018. 962 p. DOI: 10.1201/9780429435461.
- [21] Cartooning for Enhanced Privacy in Lifelogging and Streaming Videos [Text] / E. T. Hassan, R. Hasan, P. Shaffer, D. Crandall, A. Kapadia. Computer Vision and Pattern Recognition Workshops: proc. IEEE Conf. (CVPRW). 21–26 July 2017. Honolulu, USA, 2017. P. 1333–1342. DOI: 10.1109/CVPRW.2017.175.
- [22] Gonzalez R., Woods R. Digital Image Processing. Pearson Education. 2018. pp. 310–315.
- [23] ITU-T Recommendation H.264. International Telecommunication Union. 2017. pp. 45–48.
- [24] ITU-T Recommendation H.265. International Telecommunication Union. 2021. pp. 50–55.
- [25] AV1 Documentation. Alliance for Open Media. 2020. pp. 15–20.
- [26] Steve Göring, Alexander Raake Evaluation of Intra-Coding Based Image Compression; 2018. URL: <https://ieeexplore.ieee.org/document/8946162> (Дата звернення: 12.11.2024)
- [27] Miroslav Uhrina, Lukas Sevcik et al. Performance Comparison of VVC, AV1, HEVC, and AVC for High Resolutions URL: <https://www.mdpi.com/2079-9292/13/5/953>
- [28] Dan Grois , Tung Nguyen, Detlev Marpe Coding efficiency comparison of AV1/VP9, H.265/MPEG-HEVC, and H.264/MPEG-AVC encoders URL: <https://ieeexplore.ieee.org/document/7906321>
- [29] Dan Grois, Detlev Marpe, Amit Mulayoff, Benaya Itzhaky, Ofer Hadar Performance comparison of H.265/MPEG-HEVC, VP9, and H.264/MPEG-AVC encoders URL: <https://ieeexplore.ieee.org/abstract/document/6737766/authors#authors>
- [30] Shannon C. A Mathematical Theory of Communication. *Bell System Technical Journal*. 1948. Vol. 27. pp. 379–423.
- [31] Haralick R. M., Shanmugam K., Dinstein I. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*. 1973. Vol. 3. pp. 610–621.
- [32] Szeliski R. Computer Vision: Algorithms and Applications. Springer. 2020. pp. 225–230.
- [33] Daugman J. Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation. *Journal of the Optical Society of America*. 1985. Vol. 2. pp. 1160–1170.
- [34] Navidi W. Statistics for Engineers and Scientists. McGraw-Hill Education. 2019. pp. 145–150.

- [35] Poynton C. Digital Video and HD: Algorithms and Interfaces. Elsevier. 2012. pp. 130–135.
- [36] Bishop C. Pattern Recognition and Machine Learning. Springer. 2006. pp. 256–260.
- [37] Gonzalez R., Woods R. Digital Image Processing. Pearson Education. 2018. pp. 400–405.
- [38] ITU-T Recommendation JPEG. International Telecommunication Union. 1992. pp. 22–25.
- [39] Poynton C. Digital Video and HD: Algorithms and Interfaces. Elsevier. 2012. pp. 140–145.
- [40] Shannon C. A Mathematical Theory of Communication. *Bell System Technical Journal*. 1948. Vol. 27. pp. 423–425.

Бараннік В. В., Устименко Ф. О., Бараннік В. В., Мільцин І. О., Юдін О. К.
МЕТОД СЕЛЕКТИВНОЇ ОБРОБКИ ВІДЕО СЕГМЕНТІВ ДЛЯ ПІДВИЩЕННЯ ЯКОСТІ
ВІДЕОЗОБРАЖЕНЬ В СИСТЕМАХ ІНТЕЛЕКТУАЛЬНОГО АНАЛІЗУ

На даний момент світ переживає четверту світову революцію, пов'язану з появою і розвитком технологій автоматизації виробництва. Одна з груп технологій, що на даний момент все активніше використовуються у різних сферах індустрії та обслуговування це технології інтелектуальної обробки зображень. Однак, засоби, розроблені із використанням даних технологій, мають ряд обмежень, які лімітують їх ефективність. Одним з таких обмежень є існуюча залежність якості роботи систем інтелектуальної обробки зображень від якості зображень, що надходять у систему. Показано, що розвиток сучасних технологій обробки та стиснення відеозображень ставить нові вимоги до якості отриманих зображень для інтелектуальних систем обробки інформації. У роботі досліджено особливості селективної обробки відеосегментів із метою підвищення їх інформативності та адаптації до інтелектуальних технологій. Зроблено висновок, що традиційні методи обробки, такі як фільтрація шумів або усунення артефактів, не вирішують проблему втрати якості зображення на початкових етапах стиснення. Запропоновано новий підхід, який ґрунтується на аналізі інтенсивності та діапазону змін кольорів у сегментах зображення. Для досягнення цієї мети використано апарат статистичних методів, спектрального аналізу, а також комп'ютерне моделювання. Розроблено інноваційний метод селективного відбору значущих сегментів, який базується на їхньому інформативному та семантичному навантаженні. У роботі представлено метрику оцінки інформативності сегментів, яка дозволяє ефективно класифікувати сегменти на інформативні, змішані та неінформативні блоки. Це дозволяє покращити якість важливих ділянок зображення, знизити вплив шумів та забезпечити адаптивність обробки до специфіки вхідних даних. Методика селективної обробки підвищує ефективність роботи інтелектуальних систем шляхом збереження якості важливих областей при стисканні та адаптації до специфічних вимог обробки. Результати роботи можуть бути використані в системах відеоспостереження, автономному транспорті, медичній діагностиці та інших областях, де якість зображень має критичне значення.

Ключові слова: Відеозображення, селективні методи, класифікація сегментів, обробка зображень, інтелектуальні системи.

Barannik V., Ustymenko F., Barannik V., Miltcin I., Yudin O.
A METHOD OF SELECTIVE PROCESSING OF VIDEO SEGMENTS TO IMPROVE
THE QUALITY OF VIDEO IMAGES IN INTELLIGENT ANALYSIS SYSTEMS

The world is currently experiencing the fourth industrial revolution, marked by the emergence and development of production automation technologies. One of the technology groups that is increasingly being utilized in various industrial and service sectors is intelligent image processing technologies. However, tools developed using these technologies have a number of limitations that restrict their effectiveness. One such limitation is the existing dependence of the performance of intelligent image processing systems on the quality of the input images. It has been demonstrated that the advancement of modern technologies for processing and compressing video images imposes new requirements on the quality of the images received by intelligent information processing systems. This study investigates the features of selective video segment processing aimed at enhancing their informativeness and adaptation to intelligent technologies. It has been concluded that traditional processing methods, such as noise filtering or artifact removal, do not address the problem of image quality loss during the initial compression stages. A novel approach has been proposed, based on analyzing the intensity and range of color changes in image segments. To achieve this goal, a combination of statistical methods, spectral analysis, and computer modeling was employed. An innovative method for selectively identifying significant segments has been developed, relying on their informational and semantic load. The study presents a metric for evaluating segment informativeness, which enables effective classification of segments into informative, mixed, and non-informative blocks. This approach improves the quality of critical image areas, reduces noise impact, and ensures processing adaptability to the specificity of input data. The selective processing methodology enhances the efficiency of intelligent systems by preserving the quality of significant regions during compression and adapting to specific processing requirements. The results of this work can be applied in video surveillance systems, autonomous transportation, medical diagnostics, and other fields where image quality is of critical importance.

Keywords: Video images, selective methods, segment classification, image processing, intelligent systems.

Стаття надійшла до редакції 25.04.2025 р.
Прийнято до друку 11.06.2025 р.