

УДК 621.327

МАЗІАШВІЛІ А. Р., аспірант кафедри транспортного зв'язку,
КОРОЛЬОВА Н. А., к.т.н., доцент,
РОДІОНОВ С. В., к.т.н., доцент,
ІНДИК С. В., старший викладач (Український державний університет залізничного транспорту)

Розробка алгоритму статистичного кодування при компресії зображень на основі ієрархічної сіткової інтерполяції

У статті обґрунтовується значущість питання компресії зображень, а також розробки нових алгоритмів на основі ієрархічної сіткової інтерполяції. Показано системні проблеми зниження ефективності алгоритмів статистичного кодування, які пов'язані з тим, що раніше були поширені зйомки з літальних апаратів або космічного сегмента, тому зараз доводиться мати справу з даними, що містять великоформатні канали. У зв'язку з цим було запропоновано алгоритм, який може використовуватися в будь-яких методах компресії, в яких потрібно кодувати декорельований сигнал.

Ключові слова: стиск зображень, інтерполяція, квантування, кодування, алгоритм компресії, надмірність, метод стиснення.

Постановка проблеми і аналіз літератури

На сьогодні обсяги інформації, відповідні зображенням, продовжують збільшуватися. Прикладом може бути тривале зростання обсягів даних при зйомці будь-яких динамічних характеристик [1]. Якщо раніше поширені були тільки знімки з літаків і супутників, що містять одну або кілька компонент, то на сьогоднішній день доводиться мати справу з мультиспектральними і гіперспектральними даними (які призначені для військових цілей, для зв'язку спеціального призначення, а також для запобігання терористичній загрозі), які можуть містити сотні великоформатних каналів високої чіткості. Масове застосування безпілотних літальних апаратів і доступність результатів їх роботи ще більше посилює проблему обсягу даних, відповідних зображень.

Звичайно, зйомка земної поверхні, а також використання динамічних моделей для залізничного транспорту не є єдиним прикладом. Збільшення дозволу і кількості зображень існує в різних галузях знань, таких як гео-інформатика, медицина, поліграфія тощо [2].

Таким чином, виникає задача вибору (або побудови) алгоритму статистичного кодування для методу компресії зображень.

Мета статті: розглянути питання ефективності алгоритмів статистичного кодування при компресії зображень із втратами на основі математичного апарату методу ієрархічної інтерполяції; розробити новий двопотоковий алгоритм статистичного кодування.

Основна частина

При компресії на основі ієрархічної сіткової інтерполяції (ICI) пікселі зображення інтерполюються на основі проріджених версій того самого зображення, помилки інтерполяції (постінтерполяційні залишки) квантуються і кодуються.

В усіх розглянутих випадках у кінцевому підсумку кодуванню піддається якийсь квантований сигнал. Природно, етап декореляції в усіх методах компресії будується так, щоб квантований сигнал був якомога зручнішим для кодування («достискання»). Тому квантований сигнал має, як правило, специфічну (істотно нерівномірну) щільність розподілу ймовірностей. Щоб найкращим чином використовувати цю нерівномірність для підвищення коефіцієнта компресії, квантований сигнал зазвичай піддається статистичному (ентропійному) кодуванню, тобто кодуванню, що використовує нерівно-ймовірність відліків сигналу для скорочення обсягу даних.

Однак специфіка розподілу ймовірностей кодованого сигналу призводить до того, що пряме застосування алгоритму Хаффмена може спричинити істотну втрату ефективності кодування. Сигнал, що кодується (квантований), часто містить деяке значення (зазвичай нульове), яке трапляється набагато частіше за інших. Власне, причина втрати ефективності полягає в тому, що алгоритм Хаффмена не може призначити жодному символу (навіть який дуже часто трапляється) довжину коду, меншу за один біт.

Коефіцієнт стиснення, таким чином, не може перевищити розрядності кодованого сигналу, а обсяг стиснутих даних (у бітах на відлік) може істотно перевищити ентропію сигналу, що стискається, а це означає втрату ефективності кодування (помітне відставання від теоретичної межі).

У даній роботі розробляється двопотоковий статистичний алгоритм, що враховує специфіку розподілу ймовірностей кодованого сигналу при компресії зображень. Побудова пропонованого алгоритму проводиться для ієрархічного методу компресії, однак запропонована ідея може бути використана для широкого класу методів компресії, основаних на декореляції і подальшому квантуванні сигналу, що стискається.

Ієрархічне подання зображення

Метод ієрархічної сіткової інтерполяції (ICI) оснований на ієрархічному поданні зображення, що піддається компресії. Зображення має вигляд набору з L масштабних рівнів. Старший масштабний рівень, тобто рівень номер $(L-1)$, являє собою двовимірний масив відліків вихідного зображення, проріджених у $2L-1$ разів по кожній координаті. Наступний масштабний рівень, тобто масштабний рівень номер $(L-2)$, являє собою вихідні зображення, проріджені з кроком $2L-2$, з якого виключені відліки попереднього масштабного рівня номер $(L - 1)$, і так далі аж до рівня номер 0 (рис. 1).

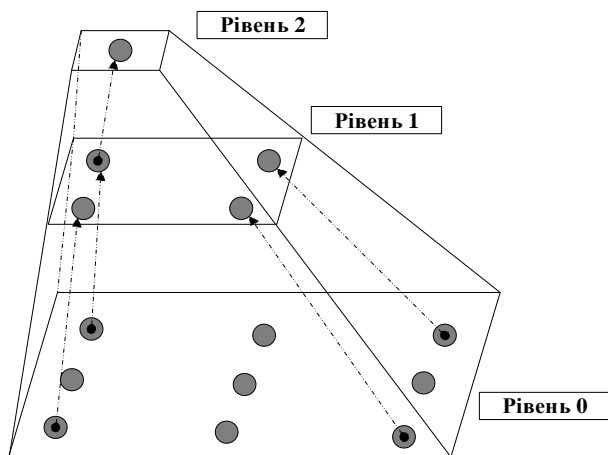


Рис. 1. Ієрархічне подання зображення у вигляді набору масштабних рівнів

Таке уявлення схоже на квадродререво (пірамідалне подання зображення), але має перевагу перед ним, тому що є ненадлишковим. У формалізованому вигляді ієрархічне подання зображення X у вигляді набору масштабних рівнів X_l може бути надане в такий спосіб:

$$X = \bigcup_{l=0}^{L-1} X_l, X_{L-1} = \{X_{L-1}(m, n)\}, \tag{1}$$

$$X_l = \{x_l(m, n)\} \setminus \{x_{l+1}(m, n)\}, 0 \leq l \leq L$$

де $x_l(m, n)$ – масив пікселів зображення, проріджений з кроком 2^l по обох координатах; L – кількість масштабних рівнів.

Описане ієрархічне представлення дає можливість компресувати масштабні рівні послідовно, починаючи зі старшого рівня X_{L-1} , причому пікселі більш проріджених (менш детальних) масштабних рівнів будуть використовуватися для інтерполяції пікселів менше проріджених (менш детальних) масштабних рівнів. Запишемо процедуру компресії формально.[3]

Ієрархічна компресія зображення

При компресії масштабні рівні зображення обробляються послідовно, від старшого до молодшого, тобто в порядку $X_{L-1}, X_{L-2}, \dots, X_1, X_0$. Вид процедури компресії старшого масштабного рівня X_{L-1} неважливий, оскільки його частка в загальному обсязі даних дуже мала (менше ніж піввідсотка для числа рівнів більше чотирьох). Тому запишемо процедуру компресії довільного нестаршого масштабного рівня $X_l (1 < L-1)$, яка складається з таких етапів (рис. 2).

Інтерполяція

До моменту інтерполяції масштабного рівня X_l вже оброблені всі більш проріджені масштабні рівні $X_k, l < k < L-1$. Їх об'єднання становить проріджені з кроком 2^{l+1} зображення $\{x_{l+1}(m, n)\}$, для всіх пікселів яких вже обчислені відновлені значення $\bar{x}_{l+1}(m, n)$, що збігаються зі значеннями, які будуть отримані при декомпресії. Позначимо \bar{X}_k безліч відновлених пікселів, відповідних масштабному рівню X_k . Таким чином, інтерполяція пікселів масштабного рівня X_l проводиться на основі відновлених пікселів, що відповідають вже обробленим більш прорідженим масштабним рівням:

$$\hat{x}_l(n, m) = P \left(\bigcup_{k=l+1}^{L-1} \bar{X}_k \right) = P(x_{l+1}(m, n)), \tag{2}$$

де P – функція, яка задає деякий інтерполятор.

Обчислення різницевого сигналу

Обчислюється різницевий сигнал, тобто масив різниць справжніх і передбачених значень відліків (постінтерполяційні залишки):

$$f_l = x_l(n, m) - \bar{x}_l(n, m). \quad (3)$$

Квантування різницевого сигналу

Виконується квантування наведеного вище різницевого сигналу $q_l(m, n)$. При цьому кожне різницеве значення $f_l(n, m)$ замінюється на квантування значення:

$$q_l(n, m) = Q_e(f_l(m, n)), \quad (4)$$

де Q_e – функція, яка задає деякий квантувач; $q_l(m, n)$ – квантований різницевий сигнал (або просто «квантований сигнал»).

Відновлення

На цьому етапі за квантованими значеннями $q_l(m, n)$ обчислюються відновлені значення $\bar{x}_l(m, n)$ пікселів масштабного рівня. Для цього необхідно спочатку обчислити відновлені значення різницевого сигналу:

$$f_l(m, n) = Q_d(q_l(m, n)), \quad (5)$$

де $Q_d(\dots)$ – функція деквантування, яка формує приблизне значення відліку різницевого сигналу з його квантованого значення;

а потім обчислити відновлені значення:

$$\bar{x}_l(m, n) = f_l(m, n) + x_l(m, n). \quad (6)$$

Ці відновлені значення пікселів ідентичні тим, що будуть отримані при декомпресії (процедура компресії ніби включає в себе процедуру декомпресії). Відновлені значення потрібні для інтерполяції більш проріджених ієрархічних рівнів ($X_k, 0 < k < l$). Тому на етапі компресії дана операція виконується для всіх масштабних рівнів (крім масштабного рівня X_0). Наявність описаного зворотного зв'язку (рис. 2), тобто використання для інтерполяції відновлених, а не вихідних пікселів, дає змогу гарантувати, що і при компресії, і при декомпресії інтерполятор буде працювати ідентичним чином (обчислювати одні і ті самі значення), що забезпечує можливість контролю похибки.

Виконується статистичне кодування квантованого сигналу $q_l(n, m)$. Оскільки розподіл ймовірностей цього сигналу є істотно нерівномірним, при кодуванні можна досягти значного скорочення обсягу даних, яке, власне, і є метою компресії зображення [4].

На цьому опис загальної структури алгоритму ієрархічної компресії завершено. Для конкретизації методу компресії необхідно вказати алгоритми інтерполяції, квантування і статистичного кодування.

У контексті цього дослідження, присвяченого статистичному кодуванню, вид алгоритму інтерполяції неважливий:

$$q_l(n, m) = \text{sign}(f_l(m, n)) \left\lceil \frac{|f_l(m, n) + \epsilon_{\max}|}{2^{\epsilon_{\max}} + 1} \right\rceil, \quad (7)$$

який гарантує, що декомпресоване зображення буде відрізнятися від вихідного не більше ніж на величину заданої максимальної похибки ϵ_{\max} :

$$|f_l(m, n)| = |x(m, n) - \bar{x}(m, n)| \leq \epsilon_{\max}. \quad (8)$$

Компресію ієрархічного рівня $X_l (1 < l < L)$ наведено на рис. 2.

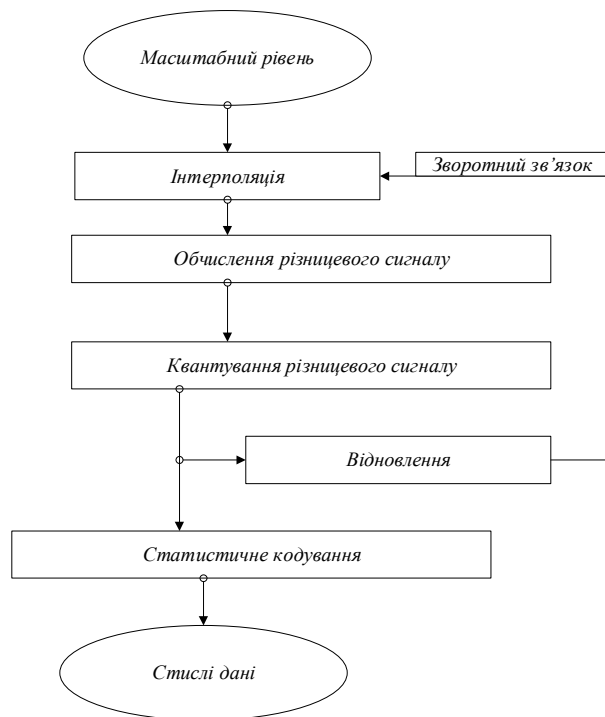


Рис. 2. Компресія ієрархічного рівня $X_l (1 < l < L)$

Статистичному кодуванню при ієрархічній компресії піддається квантований різницевий сигнал $q_i(n, m)$. Розподіл ймовірностей цього сигналу дуже схожий на розподіл ймовірностей «неквантованого» різницевого сигналу $f_i(m, n)$, тобто має максимум у нулі і швидко спадає по обидва боки від нуля симетричним чином. Відмінність розподілу ймовірності квантованого сигналу полягає в тому, що він є більш вузьким за рахунок об'єднання значень при квантуванні.

Для кодування сигналів такого виду, які характеризовані яскраво вираженою нерівномірністю розподілу ймовірностей, часто використовуються коди змінної довжини (символам, які часто трапляються, призначаються короткі коди, а які рідко трапляються, – довші коди). Ефективність кодера в цьому випадку визначається середньою довжиною кодового слова, яка в даному випадку збігається з обсягом стислих даних, який вимірюється в бітах на відлік:

$$B = \sum_{i=0}^{N-1} p_i b_i, \quad (9)$$

де N – кількість можливих значень кодового сигналу; p_i, b_i – ймовірності і довжини кодів кодованих значень.

Саме середню довжину кодового слова B і необхідно мінімізувати при побудові кодера. Однак відповідно до теореми Шеннона у середньої довжини кодового слова є теоретична межа: вона не може бути меншою за ентропію:

$$H = -\sum_{i=0}^{N-1} p_i \log_2 p_i. \quad (10)$$

З останніх двох виразів неважко побачити, що середня довжина кодового слова досягає теоретичної межі (тобто ентропії), якщо при побудові коду вдається забезпечити такі довжини кодів:

$$H = -\log_2 p_i, 0 \leq i \leq N. \quad (11)$$

Однак при ієрархічній компресії специфіка розподілу ймовірностей кодового (квантованого) сигналу часто не дає змоги це зробити, що призводить до втрати ефективності коду змінної довжини для цього сигналу. Цей ефект слід пояснити докладніше.

Чим точніше інтерполятор і чим більше значення має задана похибка квантування \mathcal{E}_{\max} , тим більше нульових значень у кодованому сигналі. Ймовірність

нуля p_0 в цьому випадку може істотно перевищувати значення 0,5. Кількість інформації в нульовому символі одна ($-\log_2 p_i (1/p_0)$) і може бути значно меншою одного біта при $p_0 > 0,5$, у той час як кодер змінної довжини привласнює цьому повідомленню код довжини один біт, за рахунок чого і відбувається втрата ефективності кодування, яка збільшується зі збільшенням похибки, що вноситься при компресії. Щоб уникнути зазначеної втрати ефективності, в даній роботі пропонується двопотоковий алгоритм статистичного кодування, що враховує частку нулів у кодованому сигналі (рис. 3).

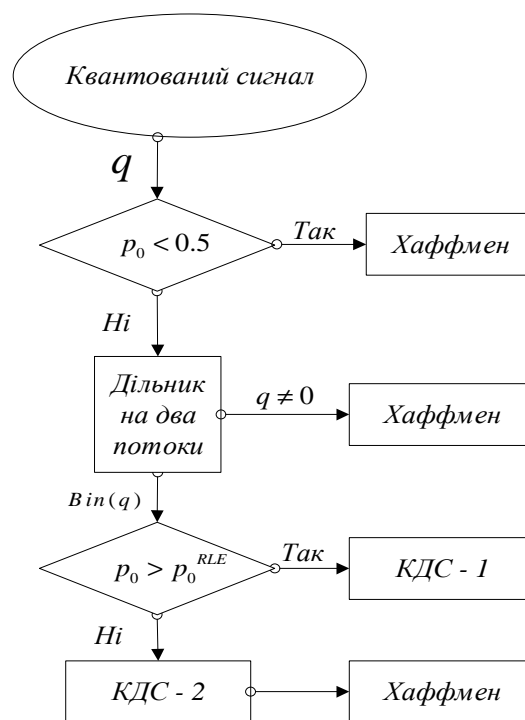


Рис. 3. Двопотоковий алгоритм статистичного кодування

Алгоритм працює таким чином: якщо вірогідність нуля $p_0 \leq 0,5$, то описаної втрати ефективності немає, тому для кодування використовується найкращий з кодів змінної довжини – кодер Хаффмена.

Якщо ймовірність нуля $p_0 > 0,5$, то сигнал, що кодується, розбивається на два потоки: в перший потік направляються значення всіх ненульових відліків, у другій потік прямує бінарна послідовність: на місце ненульових відліків вхідного потоку записуються одиниці, а на місце нульових – нулі. Далі потоки кодуються незалежно.

Другий (бінарний) потік кодується різними способами, залежно від частки нульових значень. Якщо ймовірність нуля не перевищує деякого

граничного значення $p_0 \leq p_0^{RLE}$ (тобто ймовірність нуля «велика, але не дуже»), то використовується алгоритм, оснований на алгоритмах кодування довжин серій (КДС) та Хаффмена. Спочатку здійснюється перехід від бінарної послідовності до послідовності, що складається з укрупнених символів КДС виду 1, 01, 001, ..., 00..01, 00..00. Відмінність від стандартного КДС полягає в тому, що кількість N_{RLE} різних укрупнених символів може не бути ступенем двійки.

Для завершення опису запропонованого алгоритму статистичного кодування слід зазначити, що він може використовуватися не тільки при ієрархічній компресії зображень, але і в складі інших методів компресії, що породжують декорельований сигнал з істотно нерівномірним розподілом ймовірностей.

Висновки

Розглянуто питання ефективності алгоритмів статистичного кодування при компресії зображень із втратами. На прикладі ієрархічного методу компресії розроблено двопотоковий алгоритм статистичного кодування, що входить до його складу. Запропонований алгоритм статистичного кодування може використовуватися в будь-яких методах компресії, в яких потрібно кодувати декорельований сигнал з нерівномірним розподілом ймовірностей, а також його доцільно використовувати в методах, які поєднані з математичним апаратом штучних нейронних мереж.

Список використаних джерел

1. Sayood, K. Introduction to data compression [Text] / K. Sayood. – 4th ed. – Waltham, MA: Morgan Kaufmann, 2012. – 768 p. – ISBN: 978-0-12-415796-5.
2. Претт, У. К. Цифровая обработка изображений [Текст] / У. К. Претт. – М.: Мир, 1982. – Кн. 1. – 312 с.
3. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео [Текст] / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. – М.: ДИАЛОГ – МИФИ, 2003. – 384 с.
4. Миано, Дж. Форматы и алгоритмы сжатия изображений в действии [Текст]: учеб. пособие / Дж. Миано; пер. с англ. – М.: Триумф, 2003. – 336 с.

Мазіашвілі А. Р., Королєва Н. А., Родионов С. В., Индык С. В. Розробка алгоритма статистичного кодування при компресії зображень на основі ієрархічної сеточної інтерполяції. В статті обосновується значимість питання компресії зображень, а також розробки нових алгоритмів на основі ієрархічної сеточної інтерполяції. Показані системні проблеми

зниження ефективності алгоритмів статистичного кодування, які пов'язані з тим, що раніше були розповсюджені знімки з летальних апаратів або космічного сегмента, тому зараз потрібно мати справу з даними, що містять великі формати каналів. В зв'язі з цим був запропонований алгоритм, який може використовуватися в різних методах компресії, в яких потрібно кодувати декорельований сигнал.

Ключевые слова: сжатие изображений, интерполяция, квантование, кодирование, алгоритм компресии, избыточность, метод сжатия.

Maziashvili A. R., Korolyova N. A., Rodionov S. V., Indyk S. V. Development of the algorithm of statistical coding for image compression based on hierarchical grid interpolation. At present, the volume of information corresponding to the image continues to increase. An example can be the continued growth of data volumes when shooting any dynamic characteristics. If previously only airborne and satellite images containing one or more components were common, then it is now necessary to deal with multi-spectral and hyper-spectral data (which are intended for military purposes, for special purpose communications, and also for the prevention of terrorist threat) that can contain hundreds of high-definition large-format channels. The massive use of unmanned aerial vehicles and the availability of the results of their work, further exacerbates the problem of the amount of data corresponding images.

Of course, shooting the earth's surface is not the only example. The increase in permission and number of images takes place in various areas of knowledge, such as geo-informatics, medicine, printing, forensic science, etc.

Key words: image compression, interpolation, quantization, encoding, compression algorithm, redundancy, compression method.

Надійшла 30.03.2018 р.

Мазіашвілі Артур Рамазійович аспірант кафедри транспортного зв'язку. Український державний університет залізничного транспорту, Харків, Україна. ORCID: <https://orcid.org/0000-0001-7778-5537>, e-mail: artmazo93@gmail.com

Корольова Наталія Анатоліївна к.т.н., доцент, доцент кафедри транспортного зв'язку. Український державний університет залізничного транспорту, Харків, Україна. ORCID: <https://orcid.org/0000-0001-7501-5301> E-mail: korolyova_na@ukr.net

Родіонов Сергій Вікторович к.т.н., доцент, доцент кафедри транспортного зв'язку. Український державний університет залізничного транспорту, Харків, Україна. E-mail: albano_rod@ukr.net, ORCID: <https://orcid.org/0000-0002-9511-3950>

Індик Сергій Володимирович старший викладач кафедри транспортного зв'язку. Український державний університет залізничного транспорту, Харків, Україна. E-mail: indyk.serhii@kart.edu.ua ORCID: <https://orcid.org/0000-0003-3124-8722>

Maziashvili A.R. postgraduate student, department of transport communication. Ukrainian State University of Railway Transport, Kharkiv, Ukraine, ORCID: <https://orcid.org/0000-0001-7778-5537>, e-mail: artmazo93@gmail.com

Korolyova N.A., candidate of technical sciences, associate professor, associate professor of the department of transport communications. Ukrainian State University of Railway Transport, Kharkiv, Ukraine. ORCID: <https://orcid.org/0000-0001-7501-5301> E-mail: korolyova_na@ukr.net

Rodionov S.V., candidate of technical sciences, associate professor, associate professor of the department of transport communications. Ukrainian State University of Railway Transport, Kharkiv, Ukraine., E-mail: albano_rod@ukr.net, ORCID: <https://orcid.org/0000-0002-9511-3950>

Indyk S.V., senior lecturer of the department of transport communications. Ukrainian State University of Railway Transport, Kharkiv, Ukraine. E-mail: indyk.serhii@kart.edu.ua, ORCID: <https://orcid.org/0000-0003-3124-8722>