УДК 004.627, 004.932

В.К. ЛУЦ
Інститут кібернетики ім. В.М. Глушкова НАН України, Київ, Україна,
e-mail: *vkluts@gmail.com*.

НОВІ ПІДХОДИ І МЕТОДИ АДАПТИВНОГО КОДУВАННЯ ЗОБРАЖЕНЬ

Анотація. Розглянуто нові методи, підходи та ідеї подальшого розвитку кодування зображень для підвищення інформативності і зменшення обчислювальної складності. Розроблено архітектуру спрощеного швидкодійного кодека — набір адаптивних режимів кодування з мінімальними обчислювальними витратами на вибір і переключення між ними. Запропоновано прості універсальні алгоритми передбачення значень пікселів як на основі зовнішніх даних, так і з використанням значень сусідніх блоків. Запропоновано модифікацію mse (зсунуте mse), яка точніше відображає спотворення контурів. Запропоновано однонормове спрощене косинусне перетворення порядку 16 низької обчислювальної складності. Запропоновано новий алгоритм адаптивного вибору типу і порядку перетворення залежно від коефіцієнта стиснення.

Ключові слова: спрощений швидкодійний кодек, адаптивні режими кодування, алгоритми передбачення значень пікселів, алгоритми оцінювання похибок, інформативність перетворень, спрощені косинусні перетворення, швидкі алгоритми.

ВСТУП

Актуальність та важливість досліджень кодування зображень та відео зумовлені як сучасними задачами, так і перспективами узагальненої теорії кодування, перетворення і узгодження інформації різних типів. Подібність підходів до аналізу, оброблення інформації в різних предметних галузях дає змогу формулювати і досліджувати узагальнені принципи підвищення інформативності будь-яких даних за критеріями точності та структурованості. Одним із напрямів досліджень є розроблення спрощеного швидкодійного кодека зображень та відео для мінімізації витрат обчислювальних ресурсів для приладів інтернету речей та адаптивних режимів передавання інформації в умовах реального часу за додаткових обмежень.

Одним з прикладів зменшення обчислювальної складності є кодек низької складності (LCEVC) [1], в якому було використано новий алгоритм прогресивної передачі даних. Одночасно цей алгоритм є новим універсальним алгоритмом передбачення значень пікселів на основі середнього значення блока, яке є значенням пікселя зображення меншого масштабу, тобто вперше було використано алгоритм передбачення значень пікселів на основі зовнішніх даних. З іншого боку, кодек LCEVC є надбудовою над кодеками, оскільки зменшене зображення кодується за допомогою інших кодеків, а LCEVC кодує лише різниці між зображеннями різних масштабів. Тобто натепер фактично не існує жодного спрощеного швидкодійного кодека, який використовує сучасні результати і досягнення, але в спрощеній та адаптованій формі. Аналогічно у сучасних кодеків зображень та відео немає адаптивних режимів з мінімальними обчислювальними вимогами. Оскільки мінімальна швидкість оброблення відео складає 25 кадрів у секунду, мінімальні вимоги до обчислювальних ресурсів є достатньо високими. Додатковою перевагою адаптивних режимів з мінімізацією обчислень є енергоощадливість, що подовжує час роботи приладів інтернету речей від акумуляторів.

© В.К. Луц, 2024

ОСНОВНІ ІДЕЇ СПРОЩЕНОГО ШВИДКОДІЙНОГО КОДЕКА

Розглянемо базові ідеї загальної концепції спрощення і мінімізації обчислень, а саме:

— зменшення кількості алгоритмів;

— побудова альтернативних спрощених моделей;

— спрощення та редукція алгоритмів;

— зменшення обчислювальної складності алгоритмів.

Оскільки тривалий час обчислення, зокрема, стиснення відеоінформації, базувалися на концепціях збільшення кількості та складності алгоритмів або їхньої оптимізації, виникла нова складна і перспективна задача мінімізації обчислень, пошуку найбільш ефективних алгоритмів та їхніх комбінацій для зменшення витрат обчислювальних ресурсів.

Наприклад, у процесі комплексного порівняння ефективності різних алгоритмів кодування спостерігається ефект зниження сукупної ефективності, зокрема, використання алгоритмів передбачення значень пікселів зменшувало ефективність стиснення в алгоритмах блочних перетворень: $E_r < E_t = E_p + E_b$, де E_r — реальна ефективність комплексу алгоритмів передбачення та перетворення, E_t — теоретична оцінка ефективності одночасного застосування алгоритмів передбачення і перетворення, E_p — ефективність стиснення алгоритмами передбачення, E_b — ефективність блочних перетворень за умови незалежного використання. Наявність такого ефекту пояснюється тим, що ефективність стиснення блочними перетвореннями даних з високою кореляцією значно вища, ніж з низькою. Це слідує з теореми Шеннона: чим більш передбачувані дані, тим більше їх можна стиснути (при цьому існує межа, ліміт стиснення даних без втрат для кожного рівня кореляції, передбачуванності), тобто випадкову незалежну послідовність даних з однаковою ймовірністю неможливо стиснути без втрат. Оскільки алгоритми передбачення у процесі застосування значно зменшують кореляцію даних (взаємозв'язок, передбачуваність), ефективність стиснення у використовуваних після них алгоритмів блочного перетворення значно зменшується. Отже, потрібно оцінювати і порівнювати інтегральну ефективність стиснення комплексами алгоритмів, а не окремими алгоритмами. Таким чином, виникає задача підбору та оцінювання різних комбінацій алгоритмів для пошуку оптимального співвідношення між ефективністю та обчислювальною складністю [2].

Іншим напрямом досліджень є застосування комплексного підходу, тобто узгоджене одночасне використання декількох критеріїв. Наприклад, у разі узгодження ідеї прогресивної передачі даних з ідеєю універсального алгоритму передбачення значень пікселів можна отримати новий алгоритм, який реалізує одночасно обидві функції, використовуючи середнє значення блока пікселів як прогнозоване значення для всіх пікселів блока (яке одночасно є значенням пікселя зображення меншого порядку). Аналогічно у разі узгодження ідеї прогресивної передачі даних з ідеєю кодека, який використовує будь-які інші наявні кодеки, виникає новий алгоритм, під час застосування якого зображення менших масштабів кодується відомим кодеком, а різниці між зображеннями різних масштабів — новим кодеком [1]. Зауважимо, що ці два нові алгоритми повністю збігаються, хоча були отримані різними шляхами і в рамках різних концепцій.

Концепція мінімізації обчислень має свої переваги та особливості, зокрема, наявність мінімальної оптимізації та адаптації краща за їхню відсутність. Крім того, мінімізація обчислень відбувається з урахуванням усіх наявних алгоритмів та методів і полягає у пошуку спрощення і зменшення обчислювальної складності найкращих алгоритмів, а не лише у використанні найпростіших, які мають мінімальну складність за визначенням. Процедури, схеми спрощення ефектив-

них алгоритмів суттєво залежать від ситуативних умов і особливостей, що ускладнює їхнє узагальнення і формалізацію, тому теоретичні засади та методи теорії мінімізації обчислень лише на початковій стадії розвитку.

Важливо враховувати, що спрощення обчислень використовується не тільки окремо у формі спрощеного швидкодійного кодека, але й як частина складних кодеків як окремих режимів роботи у випадку малих обчислювальних ресурсів та обмеженнях каналу зв'язку. Іншим важливим застосуванням є обернена задача — ускладнення моделей, методів і алгоритмів, зокрема, з використанням комплексних методів та підходів.

АДАПТИВНІ РЕЖИМИ КОДУВАННЯ

Фіксовані алгоритми кодування були замінені адаптивними, суть яких полягає у тому, що залежно від додаткових умов вибирають той чи інший алгоритм, параметр кодування (або комплекс алгоритмів). Наприклад, адаптивно вибирають розмір блока передбачення, блока перетворення, параметр квантування (QP), тобто рівень стиснення [3–5]. Кількість адаптивних режимів та параметрів вибору буде збільшуватись, що дасть змогу оптимально підібрати узгоджений комплекс алгоритмів для конкретного типу зображення та додаткових вимог, задач і умов. Але емпіричний вибір між адаптивними параметрами кодування для зображень чи відеокадрів має велику обчислювальну складність (30–40 % обчислювальної складності кодування відео та 70 % — кодування зображень [3]). Тому перспективною є концепція адаптивних режимів з мінімальними обчислювальними витратами на вибір та переключення між ними, що дасть змогу в два-три рази зменшити складність кодування зображень чи на 30 % — складність кодування відео.

У рамках концепції мінімізації обчислень запропоновано три профілі нового швидкодійного кодека, кожен з яких має декілька нових адаптивних режимів кодування зображень, тобто комплексів алгоритмів, які є оптимальними для тих чи інших умов, задач та обмежень, на основі розробленої математичної моделі кодування зображень. Запропоновані профілі та адаптивні режими кодування вибираються на початку кодування і потребують мінімальних додаткових обчислень для вибору і переключення між ними, на відміну від поширеного підходу емпіричного підбору розміру блока передбачення, алгоритму передбачення та порядку блочного перетворення. Всі запропоновані адаптивні режими додатково використовують алгоритм масштабування, тобто універсальний алгоритм передбачення на основі зовнішніх даних (значень пікселів зменшених зображень), який також має низьку обчислювальну складність [1]. Різниця в обчислювальній складності між запропонованими режимами в декілька разів менша, ніж між різними профілями H.264 [3]. Причому навіть найскладніший запропонований режим потребує менше обчислень, ніж найбільш простий режим H.264, за рахунок використання лише одного простого універсального алгоритму передбачення замість декількох, а також блочних перетворень зі швидкими алгоритмами обчислень.

НОВИЙ АЛГОРИТМ АДАПТИВНОГО ВИБОРУ ТИПУ ПЕРЕТВОРЕННЯ

Як відомо, існують різні типи перетворень зі своїми особливостями, перевагами та недоліками. Відповідно в кодеках можуть адаптивно вибиратись ті чи інші перетворення залежно від додаткових параметрів та умов.

Для кращого збереження візуальної якості автор пропонує застосовувати в режимі низького та середнього стиснення косинусні перетворення [6], а в режимах значного стиснення — косинусно-ступінчасті [7] чи спрощені коси-

нусні [8–12], оскільки останні (особливо косинусно-ступінчасті) краще відображають перепади значень, межі та контури. Інакше кажучи, косинусні перетворення краще відображають гладкі та незначні зміни значень пікселів, що важливо для високої якості зображень, а ступінчасті та косинусно-ступінчасті перетворення краще відображають значні перепади.

Ефект зменшення різниці ефективності кодування між косинусними та іншими перетвореннями зі збільшенням QP (параметра квантування, від якого залежить рівень стиснення) неодноразово спостерігався під час експериментальних тестувань перетворень різних типів, зокрема в [7, 12], але алгоритм адаптивного використання різних типів перетворень залежно від QP пропонується вперше.

Крім покращення відображення контурів за значних коефіцієнтів стиснення, спрощені косинусні та косинусно-ступінчасті перетворення мають більшу швидкодію порівняно з косинусними як мінімум у два рази (згідно з оцінкою знизу, коли складність виконання операції множення і складність операцій зсуву та додавання вважаються однаковими). Таким чином, спрощені косинусні та косинусно-ступінчасті перетворення можна адаптивно застосовувати як у звичайних кодеках у режимах значного стиснення (за $QP > 32$), так і в кодеках низької складності.

НОВІ УНІВЕРСАЛЬНІ АЛГОРИТМИ ПЕРЕДБАЧЕННЯ ЗНАЧЕНЬ ПІКСЕЛІВ

Для адаптивного використання разом з універсальним алгоритмом передбачення значень пікселів за середнім значенням блока пікселів пропонується новий простий універсальний алгоритм, ефективний для тих блоків зображення, де є значні перепади значень, тобто контури та межі. Як відомо, кількість подібних блоків збільшується для зображень з меншою роздільною здатністю та для зображень з більш низькою кореляцією.

Розглянемо загальну схему роботи будь-якого алгоритму передбачення. Обчислюються прогнозовані значення O_p для поточного блока, знаходиться різниця між O_p та реальними значеннями пікселів блока, яка кодується блочним перетворенням з подальшим квантуванням та ентропійним кодуванням квантованих коефіцієнтів. Зазвичай для передбачення використовуються значення інших (найчастіше сусідніх) блоків початкового зображення, але можна скористатися і зовнішніми даними, наприклад, значенням пікселя зображення меншого масштабу.

Основною перевагою запропонованого алгоритму є незначне збільшення візуальної якості, тобто чіткіші контури та перепади значень. Для передбачення значень блока пікселів використовується одне значення, як і у випадку використання середнього значення блока, яке одночасно є значенням пікселя зображення меншого масштабу. В простішому випадку значення пікселя зображення меншого масштабу вибирається таким, що дорівнює значенню першого пікселя в блоці, в складнішому — знаходиться центральне значення в упорядкованій послідовності значень блока. Алгоритм знаходження медіанного значення добре відомий: значення пікселів блока переставляються від найменшого до найбільшого, знаходиться середнє значення між двома центральними елементами: $S = (P_{i/2-1} + P_{i/2}) / 2$, де i — кількість елементів блока. При цьому різниці прогнозування виявляються більшими, як і похибка відновлення, але контури і перепади відображаються точніше, ніж у разі використання алгоритму передбачення за середнім значенням блока. Отже, описано новий універсальний алгоритм передбачення, що може бути застосований для прогресивної передачі даних (адаптивно з відомим алгоритмом передбачення за середнім значенням), для підвищення візуальної якості [2].

Як простий універсальний алгоритм передбачення значень пікселів можна вибрати відомий DC-алгоритм, в якому прогнозоване значення отримують

внаслідок знаходження середнього значення для пікселів сусідніх блоків — стовпця лівого блока та рядка верхнього блока. Можна запропонувати незначну модифікацію цього алгоритму — використовувати значення не всього стовпця, а лише першого пікселя стовпця, що одночасно є сусіднім пікселем ліворуч першого пікселя поточного блока. Аналогічно замість рядка верхнього блока пропонується використати лише перший піксель рядка (що одночасно є сусіднім пікселем зверху першого пікселя поточного блока). Таким чином, як передбачення значень пікселів поточного блока буде середнє значення двох пікселів: сусідніх ліворуч та праворуч із першим пікселем поточного блока. Також можна використати значення пікселя лише одного сусіднього блока, наприклад, ліворуч. Таким чином, пропонуються два нові прості універсальні алгоритми передбачення для застосування в різних режимах кодування: прогресивної передачі даних з використанням медіанного значення блока пікселей і незалежного стиснення — модифікація DC-алгоритму.

ПОРІВНЯЛЬНІ ОЦІНКИ ОБЧИСЛЮВАЛЬНОЇ СКЛАДНОСТІ АЛГОРИТМІВ ПЕРЕДБАЧЕННЯ ЗНАЧЕНЬ ПІКСЕЛІВ

Порівняємо обчислювальну складність запропонованих універсальних алгоритмів передбачення з обчислювальною складністю адаптивного вибору одного алгоритму передбачення із відомих.

У відеокодеку H.263 не використовували алгоритмів передбачення взагалі, а у відеокодеку H.264 використовували дев'ять алгоритмів для блоків розміру 4×4 і додатково чотири алгоритми для блоків розміру 16×16 [3], адаптивно вибираючи один із них. Кодек зображень WebP також використовує чотири алгоритми передбачення для блоків порядку 16 і 10 алгоритмів для блоків порядку 4 [13]. Завдяки використанню алгоритмів передбачення на 30 % збільшується коефіцієнт стиснення порівняно з JPEG (в якому алгоритмів передбачення немає). Для відеокодека H.265 і кодека зображень HEIF (на основі H.265) вибирається один із 35 алгоритмів передбачення для блоків порядків від 4 до 64 [14, 15], а для H.266 — із 65 різних алгоритмів передбачення [16]. Обчислювальна складність схеми порівняння алгоритмів передбачення має широкий діапазон, що додає труднощів у її оцінюванні. Найпростіший критерій порівняння різних алгоритмів передбачення для знаходження оптимального є мінімізація суми абсолютних (обчислених за модулем) різниць між передбаченням та реальними значеннями пікселей блока. Зауважимо, що навіть у цьому випадку для обчислення кожного алгоритму передбачення потрібно виконати обернене блочне перетворення для вісьмох додаткових алгоритмів (оскільки буде використано один із дев'яти алгоритмів передбачення). Відповідно вже на першому етапі порівняння перевагою використання одного універсального алгоритму буде зменшення обчислювальної складності мінімум у вісім разів.

Алгоритм пошуку мінімальної суми модулів різниць дає достатньо наближені результати за мінімальної обчислювальної складності, але його можна використати для зменшення кількості алгоритмів, що будуть порівнюватись на наступних етапах. На другому етапі порівняння алгоритмів передбачення використовують пряме та обернене блочне перетворення до отриманих після передбачення різниць, обчислюють та порівнюють похибку відновлення для кожного алгоритму передбачення. Тобто порівняння й вибір одного з алгоритмів передбачення лише для одного розміру блока вже має значну обчислювальну складність, саме тому було запропоновано використовувати спочатку алгоритми передбачення для блоків більших розмірів. Наприклад, якщо алгоритм передбачення для блока порядку 16 дає достатню точність, то подальше розбиття на блоки розміром 4×4 вже не потрібне, що зменшує загальну обчислювальну складність.

З іншого боку, якщо подальше розбиття на менші блоки є необхідним, то обчислювальна складність зростає, особливо у випадку більшої кількості рівнів розбиття (HEIF використовує чотири рівні — від блока порядку 32 до блока порядку 4).

Таким чином, схема порівняння та вибору алгоритмів передбачення є обчислювально складною частиною кодування. Її обчислювальна складність видається надмірною для співвідношення ефективність/складність, про що свідчить той факт, що натепер найпоширенішим кодеком залишається H.264, який є оптимальнішим за швидкодією, ніж H.265, що був розроблений в 2013 р. (складність останнього на порядок більша, ніж у H.264, а збільшення стиснення становить 35–40 % для більшості даних). Використовуючи лише один алгоритм передбачення, можна додатково зменшити складність кодека H.264 ще на 30–40 % за незначного зменшення коефіцієнта стиснення (приблизно на 5 %) і отримати більш оптимальний відеокодек за швидкодією. Для кодека зображень на основі H.264 зменшення обчислювальних витрат буде становити приблизно 65–70 %, тобто приблизно в три рази.

ПОРІВНЯЛЬНИЙ АНАЛІЗ ЕФЕКТИВНОСТІ СТИСНЕННЯ УНІВЕРСАЛЬНИМ АЛГОРИТМОМ ПЕРЕДБАЧЕННЯ ЗНАЧЕНЬ ПІКСЕЛІВ

Порівняємо два варіанти кодування за допомогою косинусних перетворень різних порядків: без використання алгоритмів передбачення та з використанням одного алгоритму передбачення значень пікселів на основі значення одного пікселя сусіднього блока (лівого), тобто модифікованого DC-алгоритму. Зауважимо, що блок передбачення дорівнює блоку перетворення, відповідно для блока розміру 4×4 точність передбачення буде найвищою, а для блоків порядку 16 та 32 — найменшою.

Для тестування було вибрано тестове зображення «храм» [17] з середньою кореляцією між пікселями роздільної здатності 512×512 (з урахуванням алгоритму масштабування немає потреби тестувати зображення більшої роздільної здатності).

Експериментальні результати тестування перетворень різних порядків за характеристикою PSNR (дБ) для нормального діапазону QP (22, 27, 32) з одним алгоритмом передбачення наведено в табл. 1. Ефективність кодування за характеристикою коефіцієнта стиснення (у відсотках) на основі 2D цілочислових косинусних перетворень порядків 4, 8 та 16 запропонованого методу з використанням одного алгоритму передбачення порівняно з методом кодування без передбачення наведено у табл. 2. Якість кодування за коефіцієнтом стиснення (K:1) на основі 2D цілочислових косинусних перетворень порядків 4, 8 та 16 запропонованого методу з використанням одного алгоритму передбачення наведено у табл. 3.

Алгоритм обчислення коефіцієнта стиснення включає такі кроки: збереження отриманих квантованих і округлених до цілих чисел коефіцієнтів перетворення у файл та застосування програми архіватора (тобто алгоритмів ентропійного стиснення) до цього файлу. Неочікувано великими виявились розміри файлів з коефіцієнтами перетворень порядку 4 без передбачень. Це пояснюється тим, що найбільший DC-коефіцієнт для блока порядку 4 є один із 16 коефіцієнтів, а для блока порядку 8 — один із 64 коефіцієнтів (тобто частка DC-коефіцієнтів є більшою для меншого перетворення, а вони гірше стискаються, бо є найбільшими з коефіцієнтів).

З табл. 1 випливає, що 2D цілочислове косинусне перетворення порядку 8 забезпечує найкращу якість відновленого зображення за характеристикою PSNR. При цьому перетворення порядку 32 показує найгірші результати, тому в наступному тестуванні його не розглядають. Дані табл. 2 демонструють, що

Таблиця 1

Порядок 2D-перетворень	Результати кодування за характеристикою PSNR (дБ)		
	QP = 22	QP = 27	QP = 32
4	42.31	38.13	34.35
8	**42.80**	**38.78**	**34.46**
16	42.20	37.95	34.22
32	41.63	37.51	33.80

Таблиця 2

QP	Коефіцієнт стиснення для 2D-перетворень (%)		
	порядку 4	порядку 8	порядку 16
22	73.8	32	41.5
27	87.3	28.8	43.2
32	108.5	36.8	48.4

Таблиця 3

QP	Результати кодування за коефіцієнтом стиснення ($K:1$) для 2D-перетворень		
	порядку 4	порядку 8	порядку 16
22	3.73	4.24	4.36
27	4.89	5.26	5.82
32	6.56	7.25	8.08

застосування алгоритму передбачення дає кращий коефіцієнт стиснення, як і очікувалось, при цьому найбільший відносний приріст мають перетворення порядків 4 та 16 порівняно з аналогічними перетвореннями без використання алгоритму передбачення. Але з табл. 3 випливає, що абсолютний коефіцієнт стиснення найвищий для перетворень порядків 8 і 16, як теоретично і має бути. Зауважимо, що для блоків (і зображень) з високою кореляцією між пікселями перетворення порядку 16 буде мати ще вищий коефіцієнт стиснення, чому його і застосовують адаптивно разом з перетворенням порядку 8.

Таким чином, застосування лише одного простого універсального алгоритму передбачення на 29–36.8 % збільшує коефіцієнт стиснення для найкращого за якістю перетворення порядку 8 у разі мінімального збільшення обчислювальних ресурсів (за рахунок обчислення додаткового оберненого блочного перетворення для сусіднього блока). Зауважимо, що в H.263 використовувалось перетворення порядку 8, а в H.264 у базовому та основному (main) профілях застосували тільки перетворення порядку 4. Лише в більш потужних профілях використовували перетворення порядку 8 та його переваги (тобто на 10–12 % більший коефіцієнт стиснення (див. табл. 3)). У кодеку H.265 основними є перетворення порядків 8 та 16, як найбільш оптимальні, потім порядку 32, і найменше використовуються косинусні і синусні перетворення порядку 4 (як альтернатива відомому синусному перетворенню порядку 4 у [17] запропоновано синусне перетворення порядку 4).

У проведеному тестуванні перетворення порядку 16 показало високі результати за коефіцієнтом стиснення (від 41 до 48 %), тобто коефіцієнт стиснен-

ня на 12 % більший, ніж для перетворення порядку 8, незважаючи на те, що передбачення для блока порядку 16 є менш точним. Враховуючи, що обчислювальна складність алгоритму передбачення для блока порядку 16 дорівнює обчислювальній складності чотирьох алгоритмів передбачення для блока порядку 8, перетворення порядку 16 можна рекомендувати для режиму більшого стиснення за меншої обчислювальної складності алгоритмів передбачення.

НОВИЙ АЛГОРИТМ ОЦІНЮВАННЯ ПОХИБКИ

Запропоновано новий алгоритм оцінювання похибки, який враховує тільки відхилення більше за деяке X (вибране порогове значення), що дає змогу точніше визначати і оцінювати значні спотворення, зокрема, спотворення контурів, які важливі для візуальної якості та інформативності даних. Цей алгоритм також може бути корисним у будь-яких інших галузях науки, де досліджують та оцінюють похибки для випадків, коли важливі оцінки великих відхилень, більших за X, а не сума всіх відхилень. Також запропонований алгоритм нескладно узагальнити та застосувати до будь-яких оцінок похибок, наприклад, SSIM тощо.

Формулу обчислення зсунутої mse нескладно отримати з відомої

$$\text{mse} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \widehat{Y}_i)^2$$

додаванням умови, що значення $\widehat{Y}_i < X$, де X — вибране порогове значення похибки.

Цей алгоритм було запропоновано для оцінювання доцільності нового алгоритму передбачення значень пікселів на основі першого пікселя блока, оскільки традиційне обчислення похибки відновлення давало негативні результати, тобто похибка передбачення за середнім значенням була завжди менша, ніж за першим пікселем блока. І лише за оцінкою спотворень, більших за X, експерименти показали, що запропонований алгоритм має на 0.1–0.15 дБ меншу похибку, тобто краще відображає великі перепади значень, хоча й за рахунок збільшення похибки для малих перепадів значень пікселів. Оскільки візуально більш помітними є саме великі перепади, а не малі, дійшли висновку про доцільність застосування нового алгоритму.

Таблиця 4

Порядок перетворення	QP	Результати перетворення			
		Значення похибки		PSNR	
		X_1	X_2	X_1	X_2
8	8	24	12	29.89	34.66
	14	42	21	24.46	29.88
	24	69	29	20.54	26.79
16	8	30	15	27.8	33.02
	14	48	24	23.78	28.67
	24	90	45	18.47	23.56
32	8	48	24	23.87	28.73
	14	84	42	19.18	23.62
	24	117	58	16.3	21.23

Порівняємо перетворення різних порядків за новим критерієм модіфікованого mse для вибраного тестового зображення.

Як видно з табл. 4 (тут X_1 — максимальне відхилення, $X_2 = X_1 / 2$), існує суттєва різниця між перетвореннями порядків 8, 16 та 32. Найгірші показники має перетворення порядку 32, що підтверджує відому статистику частоти застосування різних перетворень (найчастіше використовуються перетворення порядків 8 і 16).

НОВИЙ АЛГОРИТМ ОЦІНЮВАННЯ СПОТВОРЕНЬ ДЛЯ ВИБОРУ ПАРАМЕТРА КВАНТУВАННЯ

Пропонується новий алгоритм оцінювання спотворень для вибору параметра квантування, що базується на обчисленні та оцінюванні максимально допустимих абсолютних відхилень для блока пікселів, замість відомих критеріїв вибору. Цей алгоритм призначений для точнішого відображення контурів для збереження та підвищення візуальної якості. У простішому випадку враховується лише найбільше значення похибки для блока, а в складнішому — сума найбільших похибок. Він може адаптивно застосовуватись у режимах високої якості та у випадках, коли точність є найбільш важливою характеристикою (супутникові, медичні знімки тощо).

ПОКРАЩЕННЯ ВІЗУАЛЬНОЇ ЯКОСТІ ТА ІНФОРМАТИВНОСТІ ЗОБРАЖЕНЬ ТА ВІДЕОКАДРІВ

На відміну від концепції найбільш точного збереження початкової інформації, нова концепція використовує різноманітні методи та підходи для збільшення інформативності та візуальної якості. Найбільш відомим є фільтрація зображень, яка, як відомо, не змінює значення похибок, але покращує візуальну якість. Подальшим розвитком може бути застосування різноманітних фільтрів за вимогою користувача, зокрема, мультиплікація зображень та відео (що також може призвести до збільшення стиснення).

Зараз розробляють підходи, що використовують штучний інтелект та статистичні методи для стиснення і відновлення зображень та відео. При цьому більш перспективним для мінімізації обчислювальних витрат є саме відновлення інформації (генерація подібних зображень та відеокадрів) на основі неповної інформації про початкові дані.

Одним із шляхів покращення візуальної якості та інформативності зображень і відеокадрів є зменшення розмитості фону. Для цього можна використовувати накладення, узгодження зображень та відеокадрів з різним фокусом, а у випадку відео — використовувати інформацію з інших кадрів.

МАТЕМАТИЧНА МОДЕЛЬ ВИКОРИСТАННЯ ПЕРЕТВОРЕНЬ

Математична модель, що відображає основні особливості застосування блочних перетворень різних порядків (зокрема, їхнього адаптивного вибору) для стиснення зображень та відеокадрів, базується на таких засадах.

Зображення M розкладається на суму частот φ_i, які можна розбити на групи високих, середніх та низьких частот. Під час розбиття зображення на блоки і застосування блочного перетворення отримуємо локальне представлення частот у базисі функцій перетворення певного порядку, при цьому кількість базисних функцій зростає з порядком перетворення, а відносна вага кожного коефіцієнта перетворення зменшується: $f(S) = kS$, $\omega(S) = l / S$, де $f(S)$ — кількість базисних функцій, S — розмір блока перетворення, $\omega(S)$ — відносна вага коефіцієнта перетворення, k та l — константи пропорційності.

Справедливі такі функціональні залежності:

$$h(S) = g(S) + z(1/S)\ ,$$

де $h(S)$ — розподіл всіх частот блока S, $g(S)$ — низькі частоти, вага яких прямо пропорційна розміру блока S, а $z(1/S)$ — високі частоти, вага яких обернено пропорційна розміру блока S. Отже, чим більший порядок перетворення, тим краще представлені низькі частоти і тим гірше — високі. При цьому коефіцієнти високих частот швидко зменшуються зі збільшенням коефіцієнта квантування QP: чим більший коефіцієнт стиснення даних C, тим більше втрачається інформація про високі частоти $z(C) = k/C$. Отже, можна дійти таких висновків: у випадку збільшення коефіцієнта стиснення краще використовувати перетворення більшого порядку, а у випадку зменшення — меншого. При цьому для блоків зображень, у яких переважають низькі частоти, краще використовувати перетворення великого порядку, де переважають середні частоти — перетворення середнього порядку, де переважають високі частоти — перетворення низького порядку. З урахуванням обох залежностей універсальнішими є перетворення середнього порядку для середніх коефіцієнтів стиснення.

Відповідно не існує найкращого перетворення, що є оптимальним за будь-яких умов, саме тому виконується емпіричний адаптивний вибір перетворень. Але алгоритм вибору порядку перетворення потребує великих обчислювальних ресурсів і не враховує сукупного, інтегрального коефіцієнта стиснення комплексом алгоритмів передбачення і перетворення.

На основі запропонованої математичної моделі розроблено адаптивні режими, які використовуються залежно від двох параметрів: коефіцієнта стиснення C і кореляції даних (КД). Адаптивні режими визначають для двох рівнів: зображення, відеокадру та фрагментів зображення. Кожний адаптивний режим формує комбінація зазначених параметрів, які набувають трьох значень. Отримуємо дев'ять можливих варіантів адаптивних режимів (табл. 5), для кожного пропонується своє перетворення певного порядку і типу, які об'єднуються в профіль main кодека.

Запропонований профіль basic має шість адаптивних режимів (табл. 6).

Зауважимо, що у табл. 5 і табл. 6 ICST — цілочислові косинусно-ступінчасті або спрощені косинусні перетворення, ICT — цілочислові косинусні перетворення.

Таблиця 5

Коефіцієнт стиснення даних, *C*	**Кореляція даних**		
	висока	середня	низька
високий	32 ICST	16 ICST	16 ICST
середній	32 ICST	16 ICT	8 ICT
низький	16 ICT	8 ICT	8 ICT

Таблиця 6

Коефіцієнт стиснення даних, *C*	**Кореляція даних**		
	висока	середня	низька
високий	16 ICST	16 ICST	16 ICST
середній	16 ICST	8 ICT	8 ICT

Профіль simplified використовує один універсальний режим і одне перетворення ICST порядку 16 для найбільшого стиснення за найменшої обчислювальної складності.

Ідею зменшення кількості перетворень для зменшення обчислювальної складності кодека було запропоновано і розглянуто в [12]. Оскільки за статистикою для кодека H.265 найчастіше використовують перетворення порядків 8 та 16, додаткове застосування перетворень порядків 4 та 32 збільшує коефіцієнт стиснення приблизно на 3–5 % і призводить до помітного збільшення обчислювальної складності. Подальше збільшення кількості перетворень і відповідно обчислювальної складності у кодеку H.266 збільшило коефіцієнт стиснення на 3–5 % [16], що може бути обґрунтовано і доцільно лише для окремих випадків.

Таким чином, запропоновані профілі та адаптивні режими побудовані на теоретичному узагальненні емпіричних даних і дають змогу замінити складні емпірічні алгоритми порівняння і вибору простими наборами схем з мінімальними обчислювальними витратами на підбір схеми для конкретних умов кодування, а також адаптивним переключенням між ними для різних фрагментів зображення.

НОВЕ СПРОЩЕНЕ КОСИНУСНЕ ПЕРЕТВОРЕННЯ ПОРЯДКУ 16

У [12] запропоновано нове спрощене косинусне перетворення порядку 16, яке має низьку обчислювальну складність та може використовуватись для адаптивного режиму високого рівня стиснення за низьких обчислювальних ресурсів або для передачі даних у режимі реального часу за обмежень. Для зменшення обчислювальної складності розроблено і протестовано покращений варіант цього перетворення, яке має одну норму для всіх базисних функцій, на відміну від початкового різнонормового. Тестування проводилось, на відміну від тестування в [12], за умови використання запропонованого універсального алгоритму передбачення на основі значення найближчого пікселя сусіднього блока і показало значно менші різниці між спрощеним косинусним та косинусним перетвореннями: замість попередніх 0.8–1.1 дБ різниці склали лише 0.1–0.2 дБ за характеристикою PSNR, що візуально не помітно.

Як було запропоновано раніше, в адаптивному режимі використання лише одного блочного перетворення для зменшення обчислювальної складності рекомендовано застосовувати перетворення порядку 8, яке має кращі показники для низької кореляції, ніж перетворення порядку 32, і менші похибки за великих коефіцієнтів компресії. Але у випадку необхідності значної компресії доцільно використовувати перетворення порядку 16, яке краще зберігає середні та низькі частоти, але демонструє меншу швидкість зростання похибки за великих коефіцієнтів компресії, ніж перетворення порядку 32, зокрема, за новим критерієм обчислення похибок, більших за X, тобто краще відображає контури і перепади. За умови жорстких вимог до обчислювальних ресурсів доцільно застосовувати спрощене косинусне перетворення порядку 16, оскільки воно вимагає як мінімум уп'ятеро меншу кількість обчислень, ніж косинусне перетворення порядку 16 (а також — у 2.5 раза менше за косинусне перетворення порядку 8).

ВИСНОВКИ

Розвинуто концепцію спрощеного швидкодійного кодека зображень для мінімізації обчислювальних ресурсів під час вибору та переключення адаптивних режимів кодування. Запропоновано три профілі кодека, для яких розглянуто 16 адаптивних режимів з вибором розміру блока передбачення і порядку перетворення для кожного режиму. Задачі порівняння і вибору адаптивних режимів будуть розглянуті в наступних публікаціях.

Запропонований комплекс адаптивних режимів кодування може застосовуватись як окремо, так і у складі складного сучасного кодека зображень та відео, значно розширюючи діапазон обчислювальних вимог кодування.

Розглянуто два способи зменшення обчислювальних витрат: зменшення кількості алгоритмів кодування з урахуванням їхнього вагового коефіцієнта впливу на кінцевий рівень стиснення та спрощення алгоритмів з урахуванням співвідношення складність/стиснення. Запропоновано і використано комплексний підхід до узгодження алгоритмів, пошуку найбільш ефективних і оптимальних комбінацій алгоритмів. Розглянуто ефект зменшення сумарного рівня стиснення комплексами алгоритмів, особливості окремих алгоритмів та їхній вплив на оброблені дані та кінцевий результат стиснення на прикладі зменшення кореляції зображення після застосування алгоритмів передбачення значень пікселів, що зменшує декореляційні можливості блочних перетворень та ефективність перетворень великого порядку.

Запропоновано новий простий алгоритм передбачення значень пікселів, що може адаптивно застосовуватись з відомим універсальним алгоритмом передбачення за середнім значенням блока. Новий алгоритм розрахований на використання для тих блоків зображення, де є значні перепади, тобто контури, переходи та межі. Кількість подібних блоків збільшується для зображень з меншою роздільною здатністю та для зображень з більш низькою міжпіксельною кореляцією.

Запропоновано новий алгоритм оцінювання похибки відновлення, який враховує тільки відхилення, більше за X, що дає змогу точніше визначати, оцінювати значні спотворення, зокрема, спотворення контурів та меж, які важливі для візуальної якості та інформативності даних.

Запропоновано нове однонормове спрощене косинусне перетворення порядку 16, яке має низьку обчислювальну складність та може використовуватись для адаптивного режиму високого рівня стиснення та кодування за наявності обмежень на обчислювальні ресурси.

СПИСОК ЛІТЕРАТУРИ

1. MPEG-5 part 2 LCEVC (Low Complexity Enhancement Video Coding). URL: https://www.lcevc.org/ (Accessed: 23.10.2023).
2. Luts Ya.V., Luts V.K. About the development of a high-speed simplified image codec. *Cybernetics and Computer Technologies.* 2021. N 1. P. 61–66. https://doi.org/10.34229/2707-451X.21.1.6.
3. Richardson I.E.G. H.264 and MPEG-4 video compression: Video coding for next-generation multimedia. John Wiley & Sons, Ltd, 2003. 368 p. https://doi.org/10.1002/0470869615.
4. Wien M. ABT coding for higher resolution video. Doc. JVT-B053, JVT of ISO/IEC MPEG and ITU-T VCEG, Geneva, Switzerland, Jan. 2002.
5. Cham W.K., Fong C.K., Dong J., Ngan K.N., Wong H.M., Wang L., Huo Y., Pun T. Adaptive block-size transform for AVS-X. 24th AVS Meeting, AVS-M2284. Lijiang: Audio Video Coding Standard Workgroup of China, March 2008.
6. Fuldseth A., Bjøntegaard G., Sadafale M., Budagavi M. Transform design for HEVC with 16 bit intermediate data representation. Doc. JCTVC-E243, Geneva, CH, Mar. 2011.
7. Hnativ L.O. A method to construct a large-size simple integer cosine transform for high-resolution image and video coding. *Cybernetics and Systems Analysis*. 2016. Vol. 52, N 6. P. 956–966. https://doi.org/10.1007/s10559-016-9898-y.
8. Fong C.K., Cham W.K. Simple order-16 integer transform for video coding. *IEEE International Conference on Image Processing.* Hong Kong, China, 2010. P. 161–164. https://doi.org/10.1109/ICIP.2010.5651737.

9. Dong J., Ngan K.N. A universal approach to developing fast algorithm for simplified order-16 ICT. *IEEE International Symposium on Circuits and Systems* (New Orleans, USA, 27–30 May, 2007). 2007. P. 281–284. https://doi.org/10.1109/ISCAS.2007.378390.
10. Lee B., Kim M., Park C. et al. A 16x16 transform kernel with quantization for (ultra) high definition video coding. ITU-T VCEG, Doc. VCEG-AK13, Yokohoma, Japan, April 2009.
11. Ma S., Kuo C. High-definition video coding with supermacroblocks. *Proc. SPIE Vis. Commun. Image Process.* 2007. Vol. 6508. P. 650816-1–650816-12. https://doi.org/10.1117/12.707582.
12. Луц В.К. Просте цілочисельне косинусне перетворення порядка 16 для кодування сигналів і зображень. *Штучний інтелект*. 2014. № 4. С. 31–43. URL: http://dspace.nbuv.gov.ua/handle/123456789/85309.
13. A new image format for the Web | WebP | Google Developers. URL: https://developers.google.com/speed/webp/ (Accessed: 25.10.2023).
14. HEIF, high efficiency image file format. URL: https://mpeg.chiariglione.org/standards/mpeg-h/image-file-format (Accessed: 25.10.2023).
15. ITU-T Rec. H.265|ISO/IEC 23008-2: 2013. Information technology — high efficiency coding and media delivery in heterogeneous environments. Part 2: High efficiency video coding, 2013.
16. Zhao X. et al., Transform coding in the VVC standard. *IEEE Transactions on Circuits and Systems for Video Technology.* 2021. Vol. 31, N 10. P. 3878–3890. https://doi.org/10.1109/TCSVT.2021.3087706.
17. Hnativ L.O., Luts V.K. Algorithms for fast implementation of 4-point integer sine type VII transforms without multiplication and separable directional adaptive transforms for intra prediction in image/video coding. *Cybernetics and Systems Analysis.* 2020. Vol. 56, N 1. P. 159–170. https://doi.org/10.1007/s10559-020-00231-y.

V.K. Luts

NEW APPROACHES AND METHODS OF IMAGE ENCODING

Abstract. The author analyzes new methods, approaches, and ideas for further development of image encoding to increase informativeness and reduce computational complexity. An architecture (a set of adaptive encoding modes with minimal computational expenses for selecting and switching between them) for a simplified high-speed codec is developed. Simple universal algorithms for predicting pixel values based on external data and using values of neighboring blocks are proposed. A modification of MSE (shifted MSE) is proposed, which more accurately reflects contour distortions. A one-norm simplified cosine transform of order 16 with low computational complexity is proposed. A new algorithm for adaptive use of transforms of different types and orders depending on the compression coefficient is given.

Keywords: simplified high-speed codec, adaptive coding modes, pixel value prediction algorithms, error estimation algorithms, transform informativeness, simplified cosine transforms, fast algorithms.

Надійшла до редакції 08.09.2023