

МЕТОД СКОРОЧЕННЯ ЗАТРИМОК ПРИ КОДУВАННІ ВІДЕОДАНИХ В СИСТЕМІ КОДУВАННЯ СТАНДАРТУ H.265/HEVC

Бараннік В.В., Додух О.М., Ковалевський С.М., Харченко Н.А.

Харківський національний університет Повітряних Сил ім. Івана Кожедуба, Харків, Україна,

E-mail: aleksandr.dodukh@gmail.com

Анотація – Проводиться аналіз процесу кодування відеоданих при використанні стандарту H.265/HEVC. Розглядаються етапи прийняття рішень, для забезпечення оптимального варіанту кодування відео. Пропонуються можливості, які забезпечують високу ступінь стиснення зображень при обробці стандартом H.265/HEVC. Обґрунтована необхідність у розробці швидких методів вибору режимів кодування. Пропонується для прискорення процесу стиснення відео, проводити розбиття кадру на блоки кодування CU і блоки передбачення PU.

Ключові слова – стандарт кодування, ступінь стиснення, піксель, кодування-декодування, передача даних.

I. Вступ

На сьогоднішній день постійно зростають вимоги до високоякісного цифрового відеоматеріалу, який передається за допомогою супутникової, кабельної та наземної системи мовлення. Однак існуюча інфраструктура не в змозі задовільнити вимоги для передачі споживачам відеоматеріалу з надвисокою роздільною здатністю. Тому стандарт високо-ефективного відеокодування (High Efficiency Video Coding HEVC) призначений для заміни стандартів H.264/MPEG-4 Advanced Video Encoding (AVC) та підвищення ступені стиснення за рахунок адаптивного розбиття відеозображення на блоки кодування у порівнянні зі стандартом попереднього покоління.

Таке розбиття підлаштовується в процесі кодування під характер зображення, забезпечуючи максимально високу ступінь стиснення. З іншого боку, в новому стандарті не визначені критерії вибору варіанта розбивки зображення. Вибір такого критерію залишається за розробниками конкретних кодуючих систем.

Існують декілька варіантів розбиття зображення. Вибирається варіант, який забезпечує найкращу ступінь стиснення відеоданих при найменшому внесенні спотворень, які вносяться при здійсненні кодування зображення. Такий підхід призводить до збільшення обсягу обчислень і в результаті цього, відбувається зниження його ефективності. За таких умов, кодуюча система здійснює кодування одного відеокадру з високою роздільною здатністю за 2-3 хв., і це за умови використання сучасних високопродуктивних комп'ютерних платформ. Як результат, час, який витрачається на кодування одного зображення призводить до зниження можливостей практичного використання кодуючої системи, яка базується на використанні нового стандарту.

В останній час, проводяться активні дослідження та пошук швидких методів і алгоритмів, які б надали змогу забезпечити значне скорочення обчислювальних затрат на кодування. Основна увага таких досліджень спрямовується на процедуру вибору варіанта розбиття зображення на блоки кодування [1, 2]. Тому розробка методу скорочення затримок при кодуванні відеоданих в системі кодування стандарту H.265/HEVC є актуальною.

II. Основна частина. Вибір оптимального варіанту кодування відеоданих

У стандарті H.265 High Efficiency Video Coding (HEVC) застосовується блоковий гібридний підхід кодування відеоданих. Такий підхід полягає в усуненні часової статистичної надмірності шляхом застосування міжкадрового (inter) пе-

редбачення, і усунення просторової надмірності шляхом внутрішньокадрового (intra) передбачення, а так само кодування з перетворенням сигналів передбачення для подальшого усунення статистичної надмірності.

Для того, щоб провести порівняння варіантів кодування визначимо критерій. Як правило, в таких випадках використовується метрика RDC (від англ. Rate-Distortion Cost) [3, 4]. Величина RDC складається з двох складових:

$$RDC = D + \lambda \times R, \quad (1)$$

де $D = \sum_{x,y} |P_{org}(x,y) - P_{rec}(x,y)|$, $x, y = 0, 1, \dots, N-1$; $P_{org}(x,y)$, –

значення відліків кодованого блоку; $P_{rec}(x,y)$, – значення

декодуючих відліків; N – розмір кодованого блоку в пікселях; R – кількість біт, які представляють кодований блок в бітовому потоці на виході ентропійного кодера; λ – множник Лагранжа.

Величина RDC , визначається, з одного боку, величиною спотворень D , що вносяться в зображення при кодуванні на етапі квантування спектральних коефіцієнтів, з іншого боку – ступенем стиснення кодованого блоку R .

Необхідно зазначити, що для обчислення величини R для деякого варіанту кодування, необхідно провести процес кодування цього варіанту. Для того щоб отримати величину D необхідні значення $P_{rec}(x,y)$, які отримуються при проведенні декодування. Отже, для розрахунку величини RDC , яка відповідає одному варіанту кодування, необхідно виконати повний цикл кодування і декодування ділянки зображення. Отже можна стверджувати, що обчислювальна складність оцінки вартості одного варіанта кодування є занадто великою.

Ухвалення рішення про вибір того чи іншого варіанту кодування, будь то розбиття блоків кодування, режим внутрішньокадрового або міжкадрового передбачення, здійснюється так, щоб мінімізувати величину RDC . Такий підхід забезпечує максимальну ступінь стиснення відеозображення при мінімізації спотворень, що вносяться в процесі кодування зображення.

При виборі структури розбиття, блоки CU (блоки кодування), PU (блоки передбачення) і TU (блоки перетворення) в H.265/HEVC можуть мати різний розмір. Розмір, який вибраний при кодуванні може залежати від різних факторів.

Якщо проводиться внутрішньо кадровий прогноз, то чим складніша текстура в даній області зображення, тим, імовірно, менший розмір блоків слід використовувати. Для між кадрового кодування, таке твердження виявляється не вірним.

При вдалому пошуку вектора руху і виборі опорного кадру прогнозування може давати малу кількість спотворень D для великих блоків PU , отже, немає необхідності кодувати безліч додаткових структур, а значить і величина R буде менше.

Для пошуку оптимального рішення про розбивку CTU (блоки дерева кодування) на підблоки необхідно обчислити RDC для всіх можливих комбінацій розбиття кожного блоку, що і призводить до вкрай високих обчислювальних ви-

трат. Так, наприклад, для кожного блоку зображення розміром 64×64 пікселя, величина RDC обчислюється 1 раз для всього блоку цілком, 4 рази - для чотирьох підблоків розміром 32×32 , 16 разів - для підблоків розміром 16×16 , 64 рази - для підблоків розміром 8×8 і 256 разів - для підблоків розміром 4×4 .

У процесі кодування кадру розбиття зображення на блоки є початковою точкою для всіх подальших обчислень.

Таким чином, при розробці методів і алгоритмів, які забезпечують найбільш близьке до оптимального розбиття кадру на блоки кодування і блоки передбачення, досягається збільшення швидкості кодування.

Як показує аналіз, для вибору режиму передбачення необхідно здійснити велику кількість перевіряємих розбиттів, при цьому виконуються складні в обчислювальному сенсі розрахунки значення RDC . Кількість обчислень величини RDC не обмежується числом розбиття. Для кожного PU необхідно вибирати режим передбачення. Цей вибір так само відбувається на підставі оцінки RDC . Число режимів прогнозів залежить від типу передбачення - внутрішньокадрове (просторове) або міжкадрове.

При просторовому прогнозі для формування передбаченого сигналу використовуються відновлені раніше значення пікселів, що належать сусіднім блокам PU , що межує з поточним блоком (рис. 1).

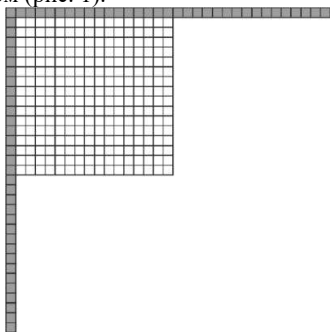


Рис. 1. Пікселі, які використовуються для просторового передбачення (сірі) значення пікселів поточного PU (білі)

Значення цих пікселів використовуються для передбачення відповідає одному з 35 доступних алгоритмів, які називаються режимами прогнозувань. Кожен режим має свій певний порядковий номер - індекс. Всі режими можна умовно розділити на три групи:

- Planar (режим 0) обчислює передбачені значення пікселів блоку передбачення як середнє арифметичне отриманих в результаті лінійної інтерполяції в горизонтальному і вертикальному напрямках величин.

- DC (режим 1) застосовується для ділянок зображення з монотонною текстурою. Блок PU , передбачений в режимі DC привласнює всім пікселям однакове значення, рівне середньому арифметичному значень пікселів сусідів.

- Решта 33 режиму передбачення (2 - 34) називаються кутовими режимами. При кутовому прогнозі значення пікселів сусідніх елементів поширюються в напрямку, відповідному обраному режиму.

Таким чином, для пошуку оптимального режиму просторового передбачення необхідно виконати 35 оцінок значення RDC . На практиці цей пошук вимагає проведення повного циклу кодування-декодування для кожного режиму алгоритму RMD (від англ. Rough Mode Decision - груба оцінка режимів) і є менш точним [5]. Крім того необхідно ввести додаткову метрику RDC_{SATD} (від англ. Sum of Absolute Hadamard Transformed Differences):

$$RDC_{SATD} = SATD + \lambda \times R_{pred}, \quad (2)$$

$$RDC_{SATD} = \left(\sum_{i,j} DiffT(i,j) \right) / 2, \quad (3)$$

$$DiffT = had(P_{org} - P_{pred}), \quad (4)$$

де P_{org} - вихідне зображення; P_{pred} - передбачене зображення; had - перетворення Адамара; P_{pred} - кількість біт, необхідне для кодування інформації про прогнозування. Як видно, для розрахунку цієї метрики немає необхідності виконувати повне кодування блоку, так що оцінка вартості кодування блоку за даним критерієм значно швидше оцінки за метрикою RDC .

При використанні метрики RDC_{SATD} вибирається фіксоване число кращих кандидатів, для яких виконується повна перевірка за критерієм RDC .

Таким чином, для вибору режиму просторового передбачення існує можливість для зменшення трудовитрат на перевірку всіх 35 варіантів, при цьому необхідно попередньо відсіяти більшу частину режимів на підставі спрощеної оцінки RDC_{SATD} .

III. Висновки

Отже можна стверджувати, що на кожному етапі кодування відеоданих можна виділити етапи обробки, які потребують значних обчислювальних затрат. Тому розробка та використання алгоритмів і методів, що дозволяють швидко визначити найбільш близькі до оптимального розбиття кадру на блоки кодування і блоки передбачення, є значущою і дозволяє стверджувати, що при визначенні оптимального варіанту розбиття CTU на CU необхідно вибрати оптимальні режими для великої кількості блоків різних розмірів. Внаслідок цього, забезпечується швидке прийняття рішення про перевірку або не перевірку будь-якого варіанту розбиття блоку CU , що може привести до великої економії часу в процесі кодування відео.

IV. Список літератури

- [1] P. B. Johns, "A symmetrical condensed node for the TLM method," *IEEE Trans. Microwave Theory Tech.*, vol. MTT-35, pp.370-377, Apr. 1997.
- [2] D Grois, "Performance Comparison of H.265/MPEG-HEVC, VP9, and H.264/MPEG-AVC Encoders / D. Grois, M. Marpe, A. Mulyoff, B. Itzhaky, O. Nadar // 30th Picture Coding Symposium 2013 (PCS 2013). —San Jose, CA, USA. —8-11 Dec 2013. —P.1-4.
- [3] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, no. 12, pp. 1648-1667, Dec. 2012.
- [4] Vanne J. Efficient mode decision schemes for HEVC inter prediction / J. Vanne, M. Viitanen, T.D. Hämmäläinen // *IEEE Transactions on Circuits and Systems for Video Technology*. – Sept. 2014. – Vol. 24, № 9. – P. 1579–1593.
- [5] Kim, I.-K. High Efficiency Video Coding (HEVC) Test Model 15 (HM15) Encoder Description / I.-K. Kim, K. McCann, K. Sugimoto, B. Bross, W.-J. Han, G. Sullivan // 17th Meeting: Valencia, ES. – 27 Mar. – 4 Apr. 2014.