

Представлено аналіз співвідношення швидкості цифрового потоку та спотворень міжкадрового кодування. Описано різні варіанти застосування оптимізації. Проведено порівняльний аналіз описаних методів оптимізації

Ключові слова: співвідношення швидкості цифрового потоку/якість, оптимізація, кодування, MPEG-4/AVC, H.264

Представлен анализ соотношения скорости цифрового потока и искажений межкадрового кодирования. Описаны различные варианты применения оптимизации. Проведен сравнительный анализ описанных методов оптимизации

Ключевые слова: отношение скорость цифрового потока/качество, оптимизация, кодирование, MPEG-4/AVC, H.264

Analysis of rate-distortion optimization is given. Different variations of optimization are described and compared. Comparative analysis of described optimization method is given

Key words: rate-distortion ratio, optimization, coding, MPEG-4/AVC, H.264

АНАЛІЗ МЕТОДІВ ОПТИМІЗАЦІЇ ШВИДКОСТІ ЦИФРОВОГО ПОТОКУ ТА СПОТВОРЕНЬ МІЖКАДРОВОГО КОДУВАННЯ ВІДЕО- ПОСЛІДОВНОСТЕЙ

Є.О. Лопатін

Аспірант

Кафедра телебачення та радіомовлення
Одеська національна академія зв'язку ім. О.С. Попова
вул. Ковальська, 1, Одеса, Україна, 65026
Контактний тел.: 093-311-39-47
E-mail: lopatin_eo@list.ru

1. Вступ

Методи оптимізації співвідношення швидкості цифрового потоку та спотворень кодованого зображення отримали широке розповсюдження в системах кодування відеопослідовностей, що можна пояснити наявністю можливостей суттєвого збільшення ефективності стиснення відеоінформації. З метою досягнення більшого коефіцієнта стиснення в стандарті MPEG-4/AVC | H.264 [1] застосовано оптимізацію за гібридним принципом. В кодері для кожного блоку виконують обчислення всіх можливих комбінацій режимів передбачення, включаючи режими внутрішньо-кадрового та міжкадрового передбачення. В результаті зменшується розходження між вихідним та відтвореним для кодування зображеннями, проте значно збільшується об'єм обчислень і швидкість цифрового потоку. В Рекомендації H.264, як і в попередніх документах, визначено лише синтаксис цифрового потоку та процес декодування для забезпечення можливості створення різних реалізацій алгоритму кодування. На практиці реалізацію кодерів визначає велика кількість факторів, тому вирішення проблеми оптимізації співвідношення швидкості цифрового потоку та спотворень кодованого зображення має велике значення в аспекті впровадження технологій кодування відеоінформації в телевізійних та споріднених технологіях.

В якості вкладу в рішення цієї проблеми проведено відповідні дослідження і представлено кількісні оцін-

ки, які може бути використано в процесі прийняття технічних рішень під час вибору режимів кодування.

2. Аналіз і дослідження основних методів реалізації оптимізації

Попередні роботи було спрямовано на знаходження альтернативних варіантів обчислення множника, за допомогою котрого можна було б оптимізувати співвідношення швидкості цифрового потоку та спотворень кодованого зображення. Більшість робіт було присвячено впровадженню адаптивного множника Лагранжа, в якості такого регулюючого множника [2-5]. Через широку поширеність використання в програмних та апаратних реалізацій алгоритму кодування саме алгоритму оптимізації з використанням множника Лагранжа, виникла необхідність в проведенні порівняльного аналізу можливих наборів значень параметрів оптимізації та визначити можливі шляхи вдосконалення використання саме даного методу. Також доцільно визначити діапазон зміни значення множника Лагранжа в реальних застосуваннях оптимізації співвідношення швидкості цифрового потоку до спотворень кодованого зображення.

В випробувальній програмній еталонній моделі MPEG-4/AVC механізм оптимізації було реалізовано в додатковому алгоритмі вибору режиму передбачення «RDO (Rate-Distortion optimization)». Відповідно до

загально прийнятої методики, суть даного алгоритму полягає в виборі з K доступних режимів кодування одного з мінімальними внесеними спотвореннями. Наприклад, $X = (X_1, X_2, \dots, X_N)$ – група N макроблоків. Для вектора режимів кодування N макроблоків $M = (M_1, M_2, \dots, M_N)$ та обмеження швидкості R_T описаний алгоритм можна виразити наступним чином:

$$M^* = \arg \min_M D(X, M) \tag{1}$$

за умови, що $R(X, M) \leq R_T$, де M^* – вектор розташувань оптимального режиму, $D(X, M)$ – метрика оцінювання спотворень для групи кодованих і декодованих макроблоків X , $R(X, M)$ – швидкість кодування групи макроблоків X . Даний алгоритм також можна представити з використанням методу множника Лагранжа.

$$M^* = \arg \min_M \sum_{i=1}^N J(X_i) \tag{2}$$

де $J(X_i)$ – обчислювальні витрати алгоритму оптимізації i -го макроблоку, обчислюється наступним чином

$$J(X_i) = D(X_i, M) + \lambda R(X_i, M) \tag{3}$$

Множник Лагранжа обирається з урахуванням всіх N макроблоків та з урахуванням умови дотримання обмеження швидкості.

Зазвичай використання такого методу ускладнюється наявністю міжкадрової кореляції між макроблоками. Реалізацію алгоритму оптимізації можна спростити шляхом зазначення незалежності макроблоків шляхом введення додаткового рішення:

$$M^* = \sum_{i=1}^N \arg \min_{M_i} J(X_i) \tag{4}$$

В використовуваному алгоритмі обирається режим передбачення визначено, що макроблоки є незалежними, і тому для кожного макроблоку обирається окреме відповідне значення множника λ як функція від значення поточного коефіцієнта квантування (QP). Режим макроблоку визначається першим кодуванням макроблоку в кожному режимі, а потім обирається режим з мінімальним значенням витрат, що його отримано з використанням множника Лагранжа λ . В свою чергу множник Лагранжа визначається емпіричним шляхом з використанням визначеного параметру квантування (QP) для кожного макроблоку

$$\lambda = \text{LambdaWeight} * 2^{\frac{QP-12}{3}} \tag{5}$$

де множник LambdaWeight за замовчування дорівнює 0,85 для I-кадрів та 0,88 для P- та B-кадрів. Відповідно до такого методу, малі значення λ відповідають великим значенням PSNR та навпаки.

Для оцінювання якості зображень широко використовується метрика PSNR, що обчислюється за допомогою середньоквадратичного відхилення (СКВ) наступним чином:

$$\text{СКВ} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} (I(i, j) - K(i, j))^2 \tag{6}$$

$$\text{PSNR} = 20 \log_{10} \left(\frac{255}{\sqrt{\text{СКВ}}} \right) \tag{7}$$

Дана метрика базується на поелементному порівнянні вихідного та кодованого зображення, тому для

урахування спотворень структурних границь в зображенні було вирішено паралельно застосувати об'єктивну метрику SSIM [6] (Structural Similarity), яку можна описати наступним чином:

$$\text{SSIM}(x, y) = \frac{(2\mu_x \mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} \tag{8}$$

де x та y – вхідне та кодоване зображення; μ_x – середнє значення x ; μ_y – середнє значення y ; σ_x – дисперсія x ; σ_y – дисперсія y ; C_1 та C_2 – змінні нормуючі величини. Данні метрики мають чітку кореляцію з метрикою суб'єктивних оцінок MOS [7] (де оцінка «5» відповідає найвищій якості).

Таблиця 1

Порівняння метрик об'єктивного та суб'єктивного оцінювання якості

PSNR	MOS	SSIM
>37	5	>0.98
31-37	4	0.90 – 0.97
25-31	3	0.80 – 0.89
20 – 25	2	0.60 – 0.80
<20	1	<0.60

Відповідно до табл. 1, на рис. 1 та рис. 2 графічно відтворено відповідність значень об'єктивних метрик PSNR та SSIM значенням MOS. Дані рисунки надають можливість більш наглядно співвідносити результати досліджень, що їх представлено в табличній та графічній формі, зі зручними оцінками метрики MOS.

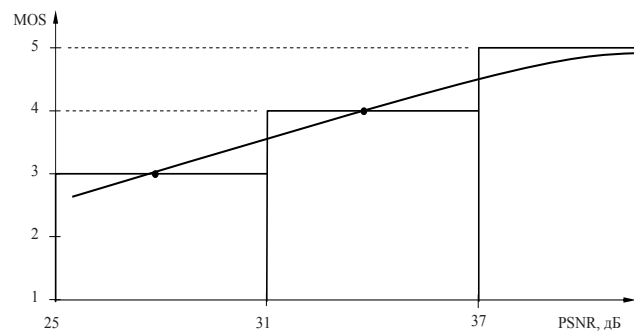


Рис. 1. Відповідність значень об'єктивної метрики PSNR значенням MOS

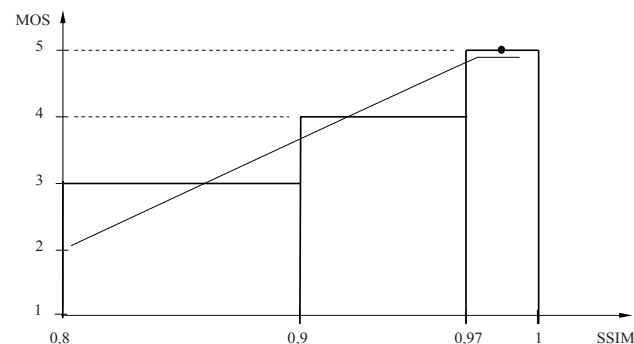


Рис. 2. Відповідність значень об'єктивної метрики SSIM значенням MOS

Дослідження було проведено з використанням випробувальних відеопослідовностей у відповідності до Рекомендації BT.1210 [8]. В роботі використано такі послідовності:

- Потяг і календар (mobile and calendar);
- Фрагмент футбольного матчу (Soccer action);
- Телевізійний рядок (TV trip);
- Сцена в японському магазині (Shinjuku);
- Людина в човні (Man in boat).

Кожна послідовність призначена для оцінювання певного типу спотворень і послідовності обрано таким чином, щоб охопити всі можливі типи спотворень, відповідно до Рекомендації BT.1210 [8]. Всі послідовності використані в форматі *.yuv, тобто дані представлені в колірному просторі YUV. Тому оцінка якості кодованого зображення обчислювалась для кожної компоненти простору в випадку метрики PSNR, та в випадку метрики SSIM лише для сигналів яскравості, що спричинене властивостями даної метрики.

Процес кодування за стандартом MPEG-4/AVC був відтворений на програмній еталонній моделі кодека H.264. Для порівняння результатів використання різних варіантів оптимізації якості зображення було вирішено використати адаптивне, стале та зважене (відповідно до формули 5) значення множника Лагранжа. Адаптивне значення обиралось автоматично кодеком. Зважене та стале значення задавалось вручну в діапазоні, границі котрого було визначено емпіричним шляхом. Також було використано різні значення параметра квантування (QP), який визначає коефіцієнти матриці квантування. Значення QP змінюється від 0 до 51. Значення 0 відповідає найменшим коефіцієнтам матриці квантування та відповідно мінімальним помилкам квантування та більшій якості зображення. Зі збільшенням значення QP збільшується крок квантування та знижується якість кодованого зображення. В дослідженні було використано середні значення з діапазону QP 28, 30 та 32.

На рис. 3 наведено результати дослідження параметрів кодування для I-кадру послідовності «mobile and calendar».

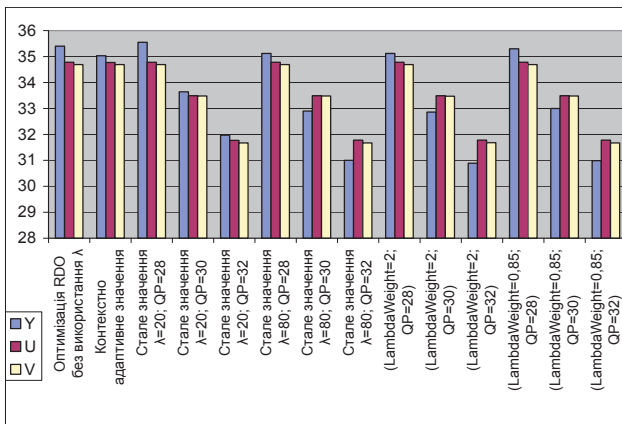


Рис. 3. Залежність якості кодованого зображення при різних режимах оптимізації співвідношення швидкості цифрового потоку до спотворень кодованого, I-кадр, послідовність «mobile and calendar»

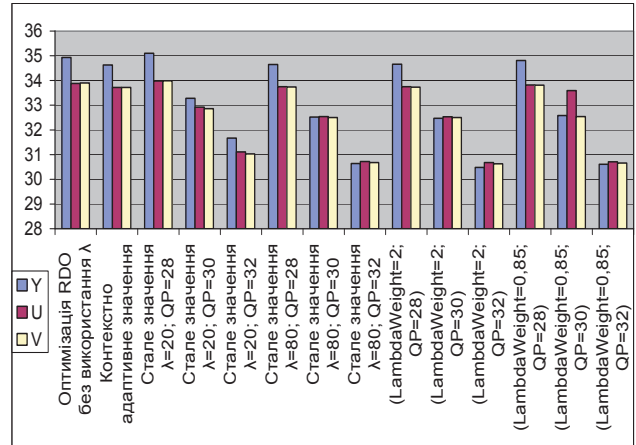


Рис. 4. Залежність якості кодованого зображення при різних режимах оптимізації співвідношення швидкості цифрового потоку до спотворень кодованого, P-кадр, послідовність «mobile and calendar»

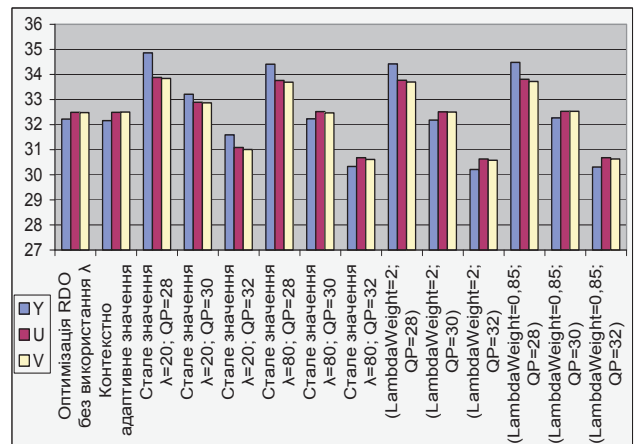


Рис. 5. Залежність якості кодованого зображення при різних режимах оптимізації співвідношення швидкості цифрового потоку до спотворень кодованого, B-кадр, послідовність «mobile and calendar»

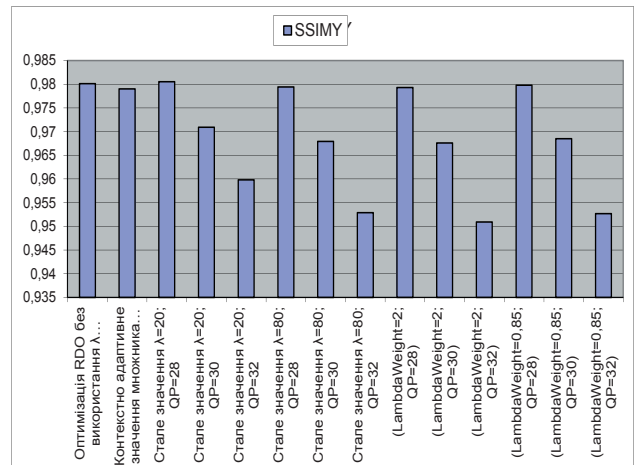


Рис. 6. Залежність якості кодованого зображення при різних режимах оптимізації швидкості цифрового потоку до спотворень кодованого, метрика SSIMY, P-кадр, послідовність «soccer action»

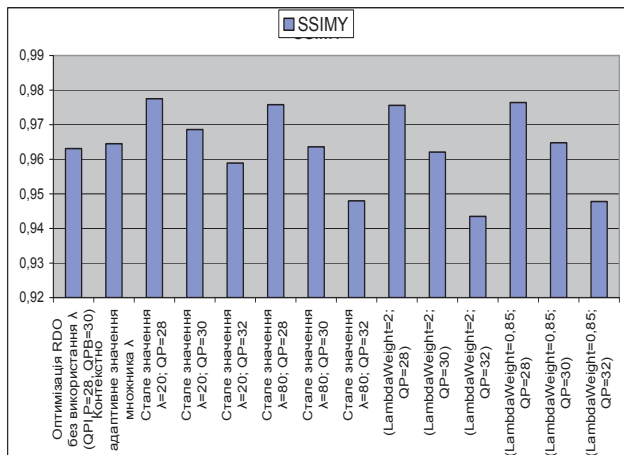


Рис. 7. Залежність якості кодованого зображення при різних режимах оптимізації співвідношення швидкості цифрового потоку до спотворень кодованого, метрика SSIMY, В-кадр, послідовність «soccer action»

3. Висновки

В ході випробувань було встановлено, що оптимальним мінімальним сталим значенням λ є 20. В випадку використання такого значення якість кодованого зображення є відмінною (значення PSNR сягає 100 дБ, а кількість бітів на зображення 5 млн.), та зменшення значення λ не призводить до суттєвих покращень, лише збільшується кількість бітів на зображення. Верхнє значення границі діапазону не обмежується та відрізняються для різних застосовань, для яких треба забезпечити різні рівні якості кодо-

ваного зображення та різні швидкості цифрового потоку. В результаті порівняння з використанням різних випробувальних відеопослідовностей було визначено оптимальне значення верхньої границі $\lambda=20$. Також було отримано такі значення зваженого множника, за яких рівні якості співпадають з такими, що відповідають використанню оптимального сталого значення λ . З рис. 4 та рис. 7 можна побачити, що найкращі оцінки якості кодованого зображення відповідають випадку використання сталого значення множника Лагранжа $\lambda = 20$. Особливий вплив дана оптимізація має на кодування В-кадрів, процес кодування в даному типі кадрів особливо чутливо залежить від зміни значення множника. Значення якості в такому діапазоні змінювання множника Лагранжа складає 5-6 дБ, що відповідає зміні рівня якості на 1 бал за шкалою MOS. Тобто, з використанням множника Лагранжа можна отримувати принципово різні рівні якості кодованого зображення, що може бути корисним для систем масштабованого кодування відеоінформації (SVC).

Таким чином, в результаті аналізу співвідношення швидкості цифрового потоку до спотворень міжкадрового кодування відеопослідовностей відповідно до стандарту MPEG-4/AVC | H.264 проведено порівняльний аналіз описаних методів реалізації та визначено діапазон оптимальних значень множника Лагранжа, що значно спрощує майбутнє використання даного методу оптимізації в системах кодування відеоінформації. Тобто, на основі використання випробувальних відеопослідовностей отримано чисельні значення залежності якості кодованого зображення для випадків використання окремих варіантів оптимізації та отримано значення границь діапазону доцільної зміни значень множника Лагранжа.

Література

1. Міжнародний стандарт 14496-10. Information technology – Coding of audio-visual objects – Part 10: Advanced Video Coding. – 2005. – 356 с.
2. G.J.Sullivan. Rate-Distortion Optimization for video compression / G.J.Sullivan, T.Wiegand // IEEE Signal Processing Magazine – 1998 – № 11. – С. 74-90.
3. Kwan-Jung Oh. Adaptive Rate-Distortion Optimization for H.264 / Kwan-Jung Oh, Yo-Sung Ho // PCM 2005, Part II, LNCS 3768 – 2005 – С. 617 – 628.
4. H. S. Malvar. Low-Complexity Transform and Quantization in H.264/AVC / H. S. Malvar, A. Hallapuro, M. Karzewicz // IEEE Trans. Circuits and System for Video Technology – 2003 – №13. – С. 598-603.
5. A. Ortega. Rate-Distortion Methods for Image and Video Compression / A. Ortega, K. Ramchandran // IEEE Signal Processing Magazine – 1998 – №5. – С. 23-50.
6. Zhou Wang. Image Quality Assessment: From Error Visibility to Structural Similarity / Zhou Wang, Alan Conrad Bovik, Hamid Rahim Sheikh // IEEE transactions on image processing. – 2004 – № 4. – С. 600-612.
7. Рекомендація ITU-R BT.500-11. Methodology for the subjective assessment of the quality of television pictures. – 2002.
8. Рекомендація ITU-R BT.1210-3. Test materials to be used in subjective assessment. – 2000-2004.