

## ІНФОРМАТИКА, ОБЧИСЛЮВАЛЬНА ТЕХНІКА ТА АВТОМАТИЗАЦІЯ

УДК 004.932.2

DOI <https://doi.org/10.32838/2663-5941/2022.3/04>

*Арпендій С.П.*

Український науково-дослідний інститут спеціальної техніки та судових експертиз  
Служби безпеки України

### ТЕХНОЛОГІЯ ОПТИМАЛЬНОГО РОЗПОДІЛУ ШВИДКОСТЕЙ ТА ВІРТУАЛЬНИХ ОПОРНИХ КАДРІВ ДЛЯ ВИСОКОЕФЕКТИВНОГО КОДУВАННЯ ВІДЕО

Стандарт вискоефективного кодування відео (HEVC) зараз став найпопулярнішим рішенням кодування відео для відеоконференцій, трансляції та потокової передачі. Однак його продуктивність стиснення все ще є критичною проблемою для прийняття великої кількості нових відеопрограм з вищими просторовими та часовими роздільними можливостями. Щоб підвищити поточну продуктивність HEVC, в статті запропоноване ефективне рішення розподілу. Метою даної роботи є розробка простого, не стандартного інструменту кодування відео MPEG шляхом інтеграції рішення для створення віртуальної реальності в HEVC та уточнення існуючої схеми розподілу бітрейту. В роботі здійснено уникнення відомого підходу попереднього аналізу, який вимагає передбачити кілька кадрів або груп, що згодом створює значну складність і не підходить для ряду застосувань. Щоб досягти цієї мети пропонується підхід евристичного квантування. В загальному цей підхід зазвичай стикається з проблемою величезної обчислювальної складності кодування відео, різноманітних наборів даних, а також нелінійності процесу кодування. В запропонованій статті ці проблеми вирішені, використовуючи підхід динамічного програмування – розподіл ставок на основі решітки (TRA). Запропонований метод адаптивно розподіляє швидкість стиснення для кожного кодованого зображення в групі зображень за допомогою підходу динамічного програмування на основі решітки. Щоб досягти цього завдання в статті досліджені параметри квантування на основі решітки для кожного кадру в групі зображень з урахуванням положення тимчасового шару. Використане динамічне програмування на основі решітки для рекурсивного пошуку найкращого зміщення кожного тимчасового індексу (Tid). Для вивчення створення віртуальних опорних кадрів (VRF) у запропонованому HEVC, оцінено якість зображення та обчислювальна складність віртуальних опорних кадрів. Створення VRF, прийняте у статті, включає три основні кроки: ієрархічну оцінку руху (HME) для ініціалізації поля вектору руху; уточнення вектору руху (MVR) для зниження помилок квантування, які часто трапляються в декодованих посиланнях; компенсація руху для інтерполяції VRF. Одержані результати, отримані для широкого набору тестів, показали, що запропоновані рішення HEVC на основі TRA та VRF значно перевершують відповідні стандартні методи покращення HEVC.

**Ключові слова:** кодування відео, віртуальна реальність, евристичне квантування, динамічне програмування, віртуальні опорні кадри, тимчасовий індекс.

**Постановка проблеми.** Широке зростання мультимедійних додатків вимагає більш потужної передачі відео через Інтернет з високою продуктивністю стиснення та низькою складністю. Тому в даний час вже впроваджено вискоефективне кодування відео (в подальшому – HEVC – High Performance Video Encoding) зі зниженням бітрейту на 50 % при тій же якості сприйняття, що й попередній стандарт H.264/AVC [1]. Незва-

жаючи на це існує подальший попит на кодування за межами HEVC, щоб задовольняти все нові вимоги вищої роздільної здатності (наприклад, 4K, 8K) і різноманітного вмісту. У цьому контексті в новітньому стандарті стиснення відео H.266 / Versatile Video Coding (в подальшому – VVC), було запропоновано багато інструментів кодування, таких як розділення за допомогою quad-tree та двійкові адаптивні перетворення, нові

внутрішні режими, афінна оцінка руху, залежне квантування тощо [2]. Методи оцінки афінного руху, віртуальна система відліку та контроль тимчасової швидкості є активними темами дослідження для покращення продуктивності міжкодування. При цьому найвище зниження бітрейту забезпечують інструменти інтеркодування, а їхнє дослідження та застосування є актуальною проблемою.

**Аналіз останніх досліджень і публікацій.** Контроль швидкості є одним із найпотужніших інструментів кодування для підвищення швидкості спотворення (в подальшому – RD – rate distortion) будь-якого відеокодека. Загалом, при заданій якості, ставки розподіляються від грубої до тонкої (I) в групі зображень (в подальшому – GOP – group of pictures) або тимчасовому розподілі швидкості, потім (II) розподіляються між одиницями кодування, або (III) адаптивним розподілом [3].

Адаптивні підходи є складними і вимагають складності обчислень і пам'яті для попереднього або онлайн-аналізу. В результаті загальна умова тестування найновіших стандартів кодування відео, включаючи H.266/VVC, все ще сприяє простому ієрархічному розподілу швидкості на рівні кадру [4]. Однак із зростанням конкуренції з боку Media Video Coding (AV1) з адаптивним GOP, потрібне ефективне розподілення тимчасової швидкості для кодеків MPEG. HEVC приймає простий тимчасовий контроль швидкості через свою ієрархічну структуру кодування [5].

Однак простий метод у HEVC є не завжди ефективний. У роботі [6] запроваджено адаптивне розподілення бітрейту для кожного кадру за схемою, оптимізованою для швидкості. Однак ці методи вимагали модифікації декодера, який не відповідає стандарту HEVC.

Віртуальний опорний кадр (в подальшому – VRF – Virtual reference frame) є іншим підходом покращення міжкодування шляхом (I) генерування віртуальних кадрів і (II) використання віртуальних кадрів для керування реконструйованих кадрів [7]. У кількох роботах використовувалася оцінка руху або розширена структура глибокого навчання для інтерполяції із декодованих кадрів [8].

Підхід глибокого навчання показує більший приріст, але також вносить величезну складність як в кодері, так і в декодері [9, 10]. Для HEVC був запропонований алгоритм раннього контролю швидкості з використанням моделі уніфікованої швидкості квантування (R-Q – rate-quantization) по пікселям [11]. Ця модель R-Q майже така ж, як і звичайна квадратична модель R-D. Пізніше Лі та ін. у роботі [12], запропонували R-модель,

з множником Лагранжа на основі складності кадру. Цей алгоритм управління швидкістю розподіляє цільові біти для групи, кадру або STU. Подальше розширюючи цю роботу, автори в [13] запропонували R-модель на основі градієнта та керування частотою між кадрами для HEVC.

Щоб покращити деталізацію обличчя для відеоконференцій із кодуванням, група науковців у роботі [14] розробила алгоритм керування швидкістю на основі R-моделі, яка може виділити більше цільових бітів у відеокадрі. З огляду на продуктивність внутрішнього кодування, для оптимізації розподілу бітів на рівні STU був запропонований у [15] підхід теорії ігор на основі подібності структури (SSIM – structure similarity).

Аналогічно, розподіл бітів і алгоритми керування швидкістю були запропоновані в [16] для оптимізації розподілу бітів для ключових кадрів. В [17] автори представили паралельну апаратну архітектуру для оцінки швидкості у внутрішньому кодуванні HEVC. Хоча вищезгадані алгоритми контролю швидкості [18] показують багатообіцяючу продуктивність для HEVC, але їхня висока обчислювальна складність може не підходити для низки нових відеододатків.

**Формулювання цілей статті.** Хоча метод на основі віртуальних посилань забезпечив важливе покращення стиснення для HEVC, все ще є простір для подальшого покращення продуктивності HEVC. Попередні роботи досліджували віртуальну систему відліку або тимчасовий контроль швидкості окремо. По-перше, попереднє створення VR, в основному, покладалося на підхід на основі оцінки руху та інтерполяції, який може бути неефективним для вмісту швидкого руху та стиснутий з високими QR або низькою швидкістю. По-друге, попередні кадри VR використовуються для всіх V-зрізів, що завжди може бути ефективним з точки зору оптимізації швидкості викривлення (в подальшому – RDO – rate-distortion optimization), особливо для зображень у позиціях низького тимчасового шару. Нарешті, оскільки якість кадрів VR сильно залежить від якості існуючих декодованих посилань, адаптивні квантовані кадри на різних рівнях впливатимуть на продуктивність міжкодування. У цьому контексті в цій статті пропонується:

1. створення нового віртуального опорного кадру, в якому для генерації використовується метод оцінки руху з множинними гіпотезами;

2. ефективний алгоритм розподілу швидкості, в якому метод решіткового кодування використовується для вивчення тимчасового розподілу швидкості.

Тому метою даної роботи є розробити простий, не стандартний інструмент кодування відео MPEG шляхом інтеграції рішення для створення віртуальної реальності в HEVC та уточнення існуючої схеми розподілу бітрейту. В запропонованій роботі автор уникнув відомого підходу попереднього аналізу, який вимагає передбачити кілька кадрів або групових груп, що згодом створює значну складність і не підходить для ряду застосувань. Щоб досягти цієї мети, автор запропонував підхід евристичного квантування. В загальному підхід до навчання зазвичай стикається з проблемою величезної обчислювальної складності кодування відео, різноманітних наборів даних, а також нелінійності процесу кодування [19]. Автор вирішив ці проблеми, використовуючи підхід динамічного програмування – розподіл ставок на основі решітки (TRA).

**Виклад основного матеріалу дослідження. Інтеркодування HEVC**

У порівнянні з попередніми інструментами міжкодування стандарту H.264/AVC [20], HEVC вносить кілька покращень у інструменти кодування та структуру кодування, зокрема (I) розбиття блоків та (II) оцінку руху та керування опорним зображенням. HEVC прийняв інструменти міжкодування, щоб скористатися перевагами тимчасової кореляції між кадрами в GOP. У цьому інструменті представлено нове розподілення тимчасової швидкості та використання віртуального опорного кадру для міжкодування HEVC. Отже, опишемо базовий інструмент з інтеркодування HEVC, а далі проаналізуємо керування швидкістю HEVC. Міжкодування HEVC дозволяє стискати зображення з більшою кількістю форм блочних розділів, ніж внутрішнє кодування, зокрема, режими розділу PART 2N 2N, PART 2N N і PART

N 2N вказують на випадки, коли блок кодування не розділений та розділений на два блоки рівного розміру по горизонталі та вертикалі відповідно. Хоча PART N N відноситься до випадку, коли блок кодування розбивається на чотири блоки однакового розміру. Крім того, є чотири асиметричні розділи руху: PART 2N nU, PART 2N nD, PART nL 2N і PART nR 2N (рис. 1). Для кожної одиниці дерева кодування (CTU – coding tree unit) процес оптимізації швидкості спотворення RDO обчислюється, щоб знайти оптимальну одиницю кодування та прогнозування, зокрема від 64 x 64 до 8 x 8.

Оцінка руху та керування опорним зображенням: подібно до попереднього стандарту кодування відео, HEVC використовує оцінку руху, щоб знайти найкращий відповідний блок в опорних кадрах, щоб зменшити надмірність між послідовними кадрами. Для керування опорним зображенням HEVC зберігає попередньо декодовані зображення в буфері декодованого зображення (DPB – decoded picture buffer). Щоб ідентифікувати ці зображення, список ідентифікаторів кількості зображень (POC – picture order count) передається в кожному заголовку фрагмента, а набір зображень називається набором опорних зображень (RPS – reference picture set), як показано на рис. 2.

HEVC DPB також містить два списки, список 0 і список 1, на які посилаються через індекс опорного зображення. Для подвійного прогнозування, тобто в конфігурації кодування з довільним доступом, вибираються два зображення (по одному з кожного списку).

**Контроль швидкості HEVC**

Контроль швидкості був важливою темою дослідження в кодуванні відео протягом багатьох років. Часовий домен є ефективним способом зниження швидкості переходу. HEVC викорис-

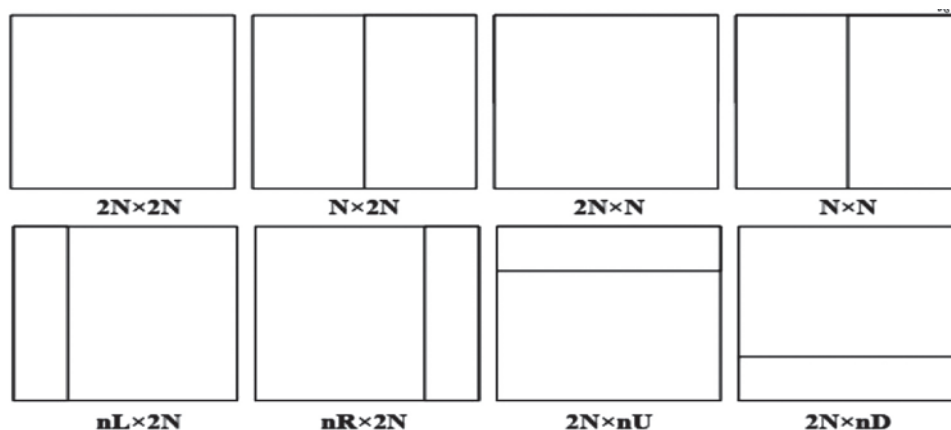


Рис. 1. Ілюстрація блочних розділів у міжкодуванні HEVC

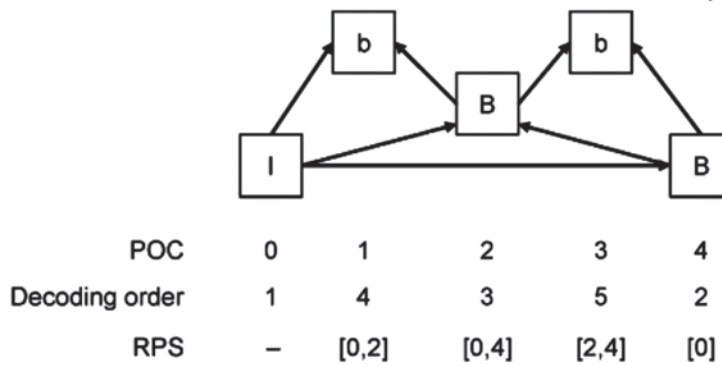


Рис. 2. Приклад структури в інтеркодуванні HEVC

товує фіксований ієрархічний розподіл бітової швидкості. Він керує бітами для кожного кадру в групі зображень GOP з тимчасовим індексом Tіd, пов'язаним з його ієрархічною опорною моделлю. За допомогою простого налаштування параметрів QP для кожного кадру, кадри з високим рівнем Tіd кодуються на вищий параметр квантування QP. Однак цей простий метод у HEVC не в змозі запобігти спотворенню. Тому для стандартів кодування відео були розроблені різні алгоритми контролю швидкості.

У цій роботі увага зосереджена на розробці алгоритму керування частотою кадрів. Для цього потрібні навчальні дані, щоб мати необроблений відеовміст різних розмірів і роздільної здатності. Таким чином, використання нещодавно вдосконалених платформ, таких як глибоке навчання [21], вимагало б функцій для кількох кадрів у кожній групі GOP, що вимагає величезної кількості обчислень не тільки під час навчання, але й під час розгортання. Крім того, структура навчання часто покладається на такі показники якості, як MSE, які погано відображають загальноприйнятую оцінку BDBR при розробці кодеків. В результаті, з обмеженими обчислювальними ресурсами, не можливо використовувати більш дешевий алгоритм навчання. Щоб вирішити цю проблему, запропоновано просте, але ефективне рішення розподілу тимчасової швидкості для HEVC, заснованого на використанні решітки.

Запропонований алгоритм TRA був інтегрований як в оригінальний HEVC, так і в HEVC з віртуальними системами відліку (VRF). Отже, у цьому розділі представлено новий HEVC з віртуальною системою відліку, за яким слідує запропоноване рішення TRA.

**Пропонована архітектура**

Запропонована архітектура кодування HEVC працює з виділеними модифікованими модулями.

Віртуальні опорні кадри генеруються з опорних декодованих кадрів на основі оцінки руху множин і потім передаються в DPB. Як раніше декодовані посилання, так і віртуальні посилання використовуються як опорні для оцінки руху поточного кадру.

Щоб досягти високоефективного кодера, автор створив новий віртуальний опорний кадр для кодування на основі ієрархічної оцінки руху (HME – hierarchical motion estimation) та методики компенсації. Щоб повністю використовувати статистичну інформацію з декодованих даних і кореляцію текстури між двома послідовними посиланнями, автор також пропонує вдосконалену тимчасову інтерполяцію з компенсацією руху (MCTI – motion-compensated temporal interpolation), засновану на створенні кадру VR, в якому поле вектору руху адаптивно генерується з використанням ієрархічної оцінки руху (ME – motion estimation) з блоком розміром 16 × 16 або 32 × 32. На цьому етапі була прийнята мінімізація середньої абсолютної різниці (RMAD – regularized mean absolute difference).

Оскільки великому розміру блоку важко охопити інформацію про рух малої активності або відео, знятого з віддаленої камери, можна вибрати більш тонкий розмір блоку. Після цього для уточнення поля руху треба прийняти уточнення вектору руху. Далі компенсація руху використовується для створення кадру VR. Створення кадру VR може бути виконано наступним чином. По перше, декодовані кадри, отримані з двох списків посилань (списку 0 і списку 1) фільтруються низькими частотами і використовуються як посилання в процесі оцінки руху. Запропоноване створення кадру VR використовує застосування прямого інтерпольованого кадру та зворотного інтерпольованого кадру. У цих модулях алгоритм узгодження блоків використовується для оцінки



руху між наступним і попереднім декодованими кадрами.

**Використання віртуальної опорної рамки**

Щоб отримати більш точну інформацію про вектор руху, треба виконати поле вектору руху (MVF – motion vector field). Щоб спростити проблему, необхідно підвищили вибірку опорних систем у 2 рази, як у горизонтальному, так і у вертикальному напрямках, як показано на рис. 3. Після підвищення вибірки MVF уточняється з урахуванням його 8 сусідніх MV.

Уточнення MV: для досягнення кращого поля руху використовується процес уточнення вектору руху (MVR – motion vector refinement). У MVR тимчасова двунправлена МЕ (BiME) і просторова зважена векторна медіана фільтрація (WVMF – weighted vector median filtering) вибираються для уточнення інформації про рух, отриманої з ієрархічного етапу МЕ. У BiME вектори руху кожного інтерпольованого блоку уточнюються в невеликій області пошуку з урахуванням припущення, що траєкторія руху між послідовними кадрами є лінійною. У той час як просторовий WVMF покращує просторово когерентне поле руху, шукаючи для кожного інтерпольованого блоку поле в сусідніх блоках, що може краще відображати траєкторію руху. Цей фільтр також регулюється керуючи його потужністю та залежно від спотворення блоку для кожного вектору руху. Оскільки якість декодованих послань і відео-контенту сильно впливає на кінцеву якість кадру VR, автором прийнято рішення для оптимізації параметрів на основі статистичного навчання для ініціалізації розміру блоку, діапазону пошуку та областей уточнення пошуку для запропонованого методу МСТІ. Компенсація руху: процес компенсації руху застосовується до декодованих кадрів MV для досягнення кадрів VR.

Щоб краще використовувати цей новий метод, автор виконав ряд експериментів, щоб визначити, яке положення кадру слід використовувати для створення кадрів VR. Створення кадру VR має переваги від низької складності оцінки руху та підходу до компенсації в поєднанні з оптимальною конфігурацією параметрів.

**Розподіл швидкості на основі решітки TRA**

Пошук решітки забезпечується алгоритмом динамічного програмування для пошуку оптимальної кодової послідовності. У кодуванні відео решітковий пошук використовується в оптимізованому квантуванні швидкості HEVC. Починаючи з початкового квантування (тобто  $x$ ), RDOQ рекурсивно вирішує оптимальний рівень квантування в кожному місці розташування коефіцієнта між  $x$  та  $x-1$  з метою мінімізації швидкості та вартості спотворення.

Автором змодельований розподіл бітрейту як оптимальний шлях пошуку коригування QP між тимчасовими ідентифікаторами Tid (див. рис. 4).

Далі виконується ітераційний пошук з максимальними фазами  $K$ . На кожній фазі шукаються оптимальні зміщення QP, починаючи від Tid 0 до Tid 3 щодо найкращого зменшення BDBR. Починаючи від найбільшого Tid 0 до найменшого Tid 3, уникаємо великих коливань в алгоритмі (тобто посилення BDBR має тенденцію зменшуватися після кожного кроку). Приклад структури Tid у HEVC показаний на рис. 4. Чим нижчий Tid, тим він важливіший, оскільки більше зображення має значення при нижчому Tid. У своїй практиці автор використовував тільки двофазний  $K = 2$ . Найкраще зміщення QP на ітерації  $k$ , і  $Tid_i$  (тобто позначений як  $DQP_k$ ) додається до оптимального списку  $DQP_k$  з повного списку  $\{0, \pm 1, \dots, \pm k\}$ , які вже видалили з попередньої ітерації (тобто

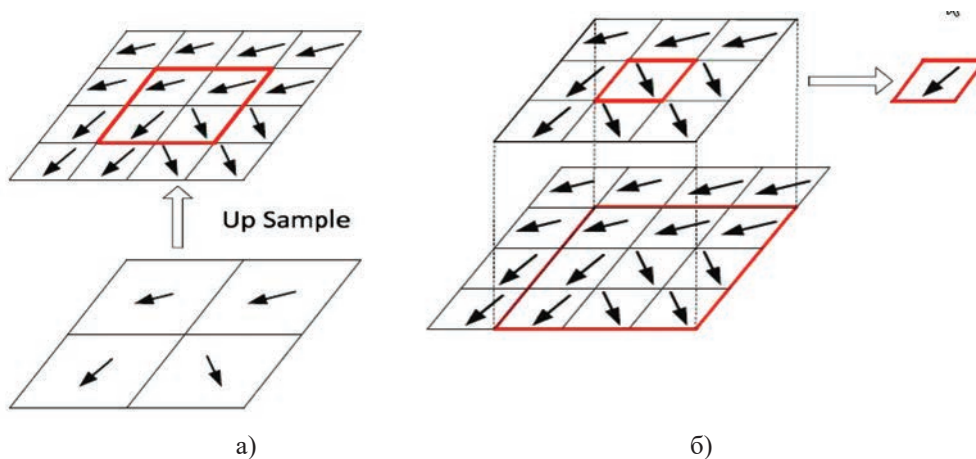


Рис. 3. Ієрархія МЕ: вибірка MVF (а); вибірка MV (б)

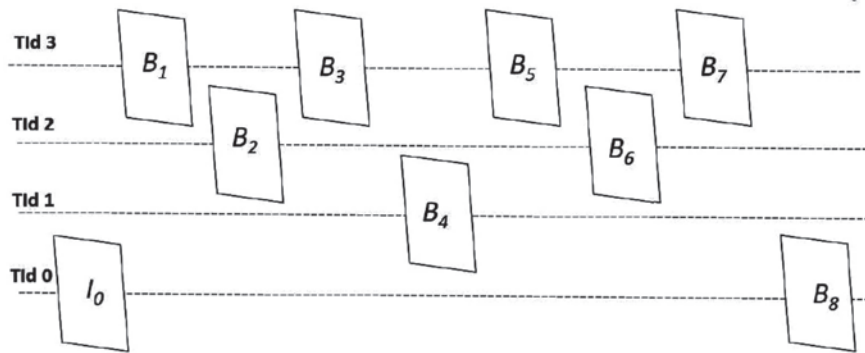


Рис. 4. Конфігурація довільного доступу HEVC

DQPk-1i). Для обмеження складності в цій роботі було встановлено максимальний діапазон зміщення  $k$ . На основі швидкості скорочення бітрейту (BDBR – Based on the bitrate reduction) даного набору відеоданих були обрані найкраще зміщення для сценаріїв високого та низького бітрейту. Враховуючи чотири Tids, у максимальному діапазоні 2 (тому, п'ять доступних варіантів), з шістьма QP для низької та високої швидкості, кількість тестів для кожної послідовності пошуку становить

$$nQP \times r^4 = 6 \times 5^4 = 3750. \quad (1)$$

Це все ще вимагає значної кількості обчислень. Тому автор використовував динамічне програмування на основі решітки для рекурсивного пошуку найкращого зміщення кожного Tld. Для одноетапного раунду пошуку (фази) він вимагає

$$nQP \times (5 + 4 + 4 + 4) = 6 \times 17 = 102, \quad (2)$$

для кожної послідовності, що значно зменшує кількість симуляцій. Незважаючи на те, що запропонований TRA може значно скоротити простір для пошуку, він все одно вимагає величезних обчислень, враховуючи численні тестові послідовності та кілька ітерацій. Тому автор пропонує спрощену версію TRA з двофазним ітеративним підходом. На першому етапі здійснено обмеження до  $\{0, \pm 1\}$ , що вимагає

$$nQP \times (3 + 2 + 2 + 2) = 6 \times 9 = 54, \quad (3)$$

проведення половини експериментів як у TRA. На другому етапі протестоване лише додаткову конфігурацію, яка становить  $\pm 1$  від фази 1. Якщо оптимальне зміщення дорівнює  $+1$ , тоді оцінюється тільки етапі 1. Отже, максимальна кількість додаткових експериментів за ітерацію дорівнює

$$nQP \times (1 + 1 + 1 + 1) = 6 \times 4 = 24. \quad (4)$$

Комбінація експериментів у (3) і (4) значно менша, ніж вихідна TRA в (2).

Для навчання TRA було прийнято рішення налаштувати рівень квантування або DQP на кожному Tld, пов'язаному з продуктивністю RD цілих послідовностей. Далі застосовано зміщення TRA до тестових послідовностей без тонкого налаштування, щоб уникнути додаткових витрат на складність. Таким чином, встановлюємо залежну групу квантування  $DQP_i, i = 0, 1, 2, 3$ .

#### Умови та оцінки тестування

Для оцінки ефективності кодування запропонованих методів Y-BDBR розраховується за конфігурацією довільного доступу (RA – Random Access) за загальних умов тестування. Автор використовував НМ 16.20 з GOP 8, внутрішнім періодом 32 і різними параметрами квантування для високої швидкості (QP: 22, 27, 32, 37), що відповідає загальній умові тесту в HEVC і низькій швидкості (QP: 32, 37, 42, 45). Для розподілу швидкості на основі решітки було обрано 16 послідовностей з різними роздільними можливостями та частотою кадрів, тоді як тестові послідовності є іншими поширеними послідовностями в HEVC [22]. Лише перші 64 кадри використовуються для скорочення часу навчання, але все одно забезпечують високу продуктивність.

Рис. 5 і 6 ілюструють перший кадр кожної відео-послідовності, а послідовності навчання та тестування можна побачити в табл. 1. Легко помітити, що TRA вже досягає хорошої продуктивності навіть на першій фазі, яка включає набір DQP  $\{\pm 1, 0\}$ . Більше того, однієї ітерації достатньо, щоб TRA зблизився в обох фазах. Для навчального набору 1080p на фазі 1 додатковий набір DQP  $\{\pm 2, \pm 1, 0\}$  у фазі 2 не показав покращення BDBR для TRA. Причина відсутності покращення TRA на фазі 2 пов'язана зі зміною DQP на рівні кадру, що істотно вплине на загальну продуктивність. Тому  $\pm 2$  може занадто значно відрізнятись від звичайної ставки. Однак у поєднанні з VRF DQP 2 показав значний

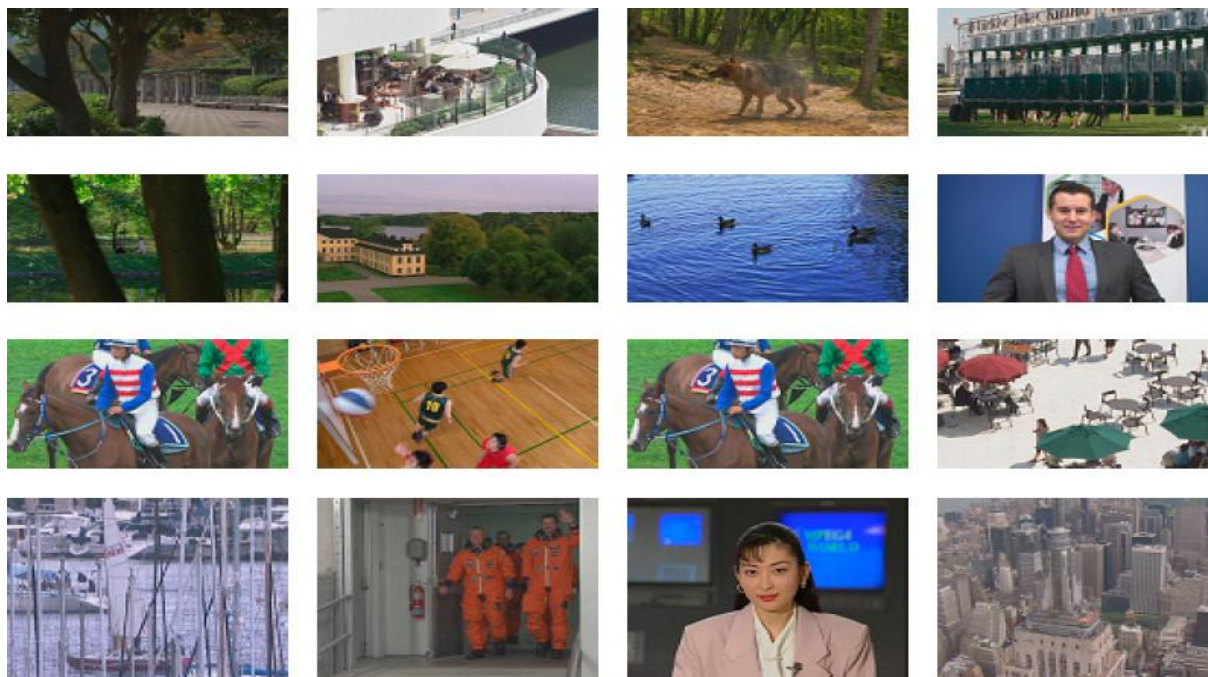


Рис. 5. Ілюстрація першого кадру навчального відео.

Зліва направо, зверху вниз: Park, Terrace, SnakeDry, ReadySteady, Parkjoy, InTree, DucksTake, Man, Horses, Basketball, Crew, Harbour, Race, Square, Woman, City.

приріст у середньому на 1 % BDBR у навчальному наборі.

Втрата якості на низькому рівні компенсується додатковими віртуальними кадрами. З'ясовано, що використання одного і того ж набору DQP для всіх навчальних даних демонструє незначне покращення якості з високою швидкістю. Це пов'язано з природою локального оптимуму динамічного програмування. При низькій швидкості додаткові 0,31 % Y-BDBR досягаються для низької швидкості, але майже ідентичні результати для високої. Тому автором прийнято простий адаптивний вибір DQP на основі роздільної здатності. Оскільки роздільна здатність доступна як вхідні дані, запропонований метод не вимагає додаткової складності порівняно з іншими фреймворками. Для складних адаптивних методів (наприклад, швидкий / повільний рух, проста / деталізована сцена тощо) це може бути цікавою темою для майбутніх робіт.

**Оцінка продуктивності стиснення.** Ефективність кодування тестових послідовностей оцінюється для BDBR і для порівняння BDPSNR. Автор порівняв запропонований метод TRA з VRF і без нього з HEVC з (TRA) і без VRF (TRA + VRF), найбільш релевантним методом адаптивного квантування, а також стандартним методом HEVC. Подібно до навчального набору, для порівняння використовуються низькошвидкісні та високошвидкісні BDBR.

Загалом, як TRA, так і VRF забезпечують послідовне покращення порівняно з базовим HEVC.

В основному це відбувається завдяки створенню додаткових посилок. Збільшення опорних кадрів також призводить до більшого часу кодування для оцінки руху. Цей компроміс між складністю та продуктивністю стиснення може запобігти широкому розгортанню VRF в деяких програмах кодування відео, де обчислювальна складність обмежена. Крім того, проблему складності в VRF можна вирішити шляхом пошуку методів створення VRF низької складності.

**Оцінка складності.** Хоча VRF продемонстрував високу продуктивність стиснення, коли він був прийнятий в архітектуру HEVC, він також вніс значну складність накладних витрат. Для подальшого вивчення створення VRF у запропонованому HEVC, у цьому пункті буде оцінено якість зображення та обчислювальна складність VRF.

Як показано у статті вище, створення VRF, прийняте в цій статті, включає три основні кроки: (I) ієрархічна оцінка руху HME для ініціалізації поля вектору руху, (II) уточнення вектору руху MVR для пом'якшення помилки квантування, які часто трапляються в декодованих посиленнях, і (III) компенсація руху MC для інтерполяції VRF. Щоб виявити внесок цих модулів, обчислимо та проілюструємо в таблиці 2 відсоток часу обчислень (%), вимірний для VRF щодо загального HEVC



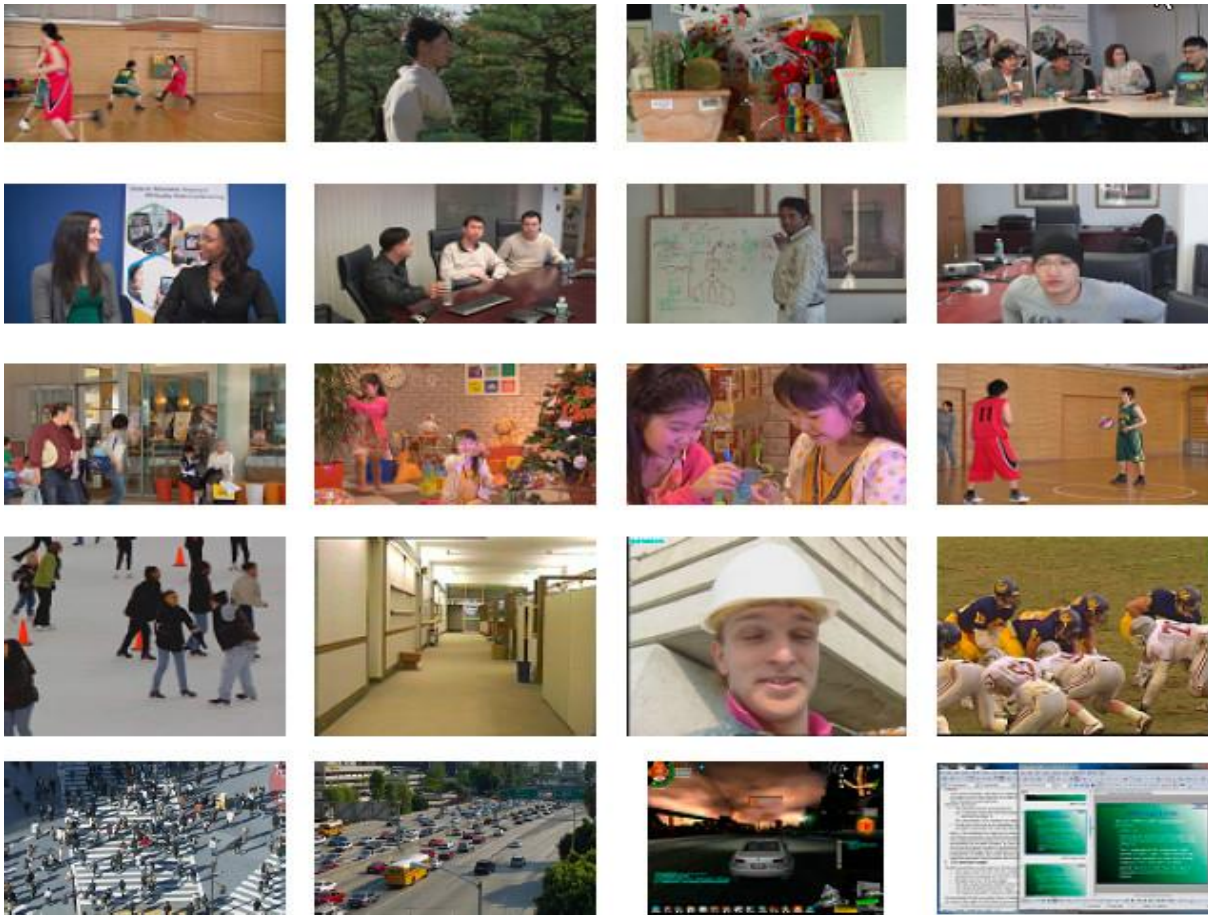


Рис. 6. Ілюстрація першого кадру тестового відео

Зліва направо, зверху вниз: BasketDrive, Kimono, Cactus, People, KristenSara, Vidio1, Vidio3, Vidio4, Mall, Party, Ice, BasketPass, Blowing, Hall, Foreman, Football, PeopleStreet, Traffic, Speed, SlideEditing

Таблиця 1

Послідовності навчання та тестування з різними роздільною здатністю та частотою кадрів

| Роздільна здатність | Навчальний відеоряд (рис. 6)         | Тестовий відеоряд (рис. 7)                               |
|---------------------|--------------------------------------|--|
| 1080p and Class A   | Park, Terrace, SnakeDry, ReadySteady | BasketDrive, Kimono, Cactus                              |
| 720p and Class      | Parkjoy, InTree, DucksTake, Man      | KristenSara, Vidio1, Vidio3, Vidio4, Speed, SlideEditing |
| 480p                | RaceHorses, BasketballDrill          | Mall, Party  |
| 4SIF                | Crew, Harbour                        | Ice  |
| 240p                | Race, Square                         | BasketPass, Blowing                                      |
| CIF                 | Woman, City                          | Hall   |

(VRFHEVC), а також для HME, MVR та Mcover VRF (HMEVRF, MVRRF, MCVRF).

$$VRF^{HEVC} = Time^{VRF} \times 100 / Time^{HEVC}, \quad (5)$$

$$HME^{VRF} = Time^{HME} \times 100 / Time^{VRF}, \quad (6)$$

$$MVR^{VRF} = Time^{MVR} \times 100 / Time^{VRF}, \quad (7)$$

$$MC^{VRF} = Time^{MC} \times 100 / Time^{VRF}. \quad (8)$$

Результати складності, отримані в таблиці 2 показують, що VRF споживає близько 11,70 %

обчислювальної складності в загальному кодуванні HEVC. У VRF HME потребував найбільшої обчислювальної складності, тобто 95,24 %. У цьому випадку скорочення часу HME має вирішальне значення.

**Висновки.** Ця стаття є піонерською роботою, яка розглядає вплив помилки квантування та відеоконтенту на якість кадрів.

У запропонованій роботі був представлений новий алгоритм розподілу бітрейту на основі



Оцінка складності VRF

| Video           | VRF <sub>HEVC</sub> | HME <sub>VRF</sub> | MVR <sub>VRF</sub> | MC <sub>VRF</sub> |
|-----------------|---------------------|--------------------|--------------------|-------------------|
| BasketballDrive | 12.71               | 96.06              | 2.37               | 1.56              |
| Kimono          | 12.41               | 96.09              | 2.38               | 1.53              |
| Cactus          | 14.12               | 96.03              | 2.37               | 1.59              |
| PeopleOnStreet  | 14.17               | 96.85              | 1.87               | 1.28              |
| Traffic         | 18.47               | 96.81              | 1.86               | 1.33              |
| FourPeople      | 13.70               | 95.21              | 2.85               | 1.94              |
| KristenAndSara  | 12.71               | 96.06              | 2.37               | 1.56              |
| Vidio1          | 13.72               | 95.22              | 2.82               | 1.96              |
| Vidio3          | 13.15               | 95.26              | 2.82               | 1.93              |
| Vidio4          | 13.24               | 95.24              | 2.85               | 1.91              |

решітки для міжкодування HEVC, де розглядається віртуальний опорний кадр. Запропонований метод VRF-HEVC спирається на декодовану інформацію, доступну як в кодері, так і в декодері; таким чином, не потрібно змінювати елементи синтаксису в стандартній специфікації, а також бітрейту.

Було виконано значну кількість експериментів із численними відеозаписами тестів, які продемонстрували збільшення швидкості BD приблизно на 25 % на основі останнього еталонного програмного забезпечення HEVC. Щоб оптимізувати якість

VRF, розроблено створення кадру VRF на основі статистичного навчання. Крім того, для досягнення вищої продуктивності стиснення HEVC, новий набір параметрів квантування введено для структури VRF-HEVC на основі підходу розподілу бітрейту шляхом динамічного програмування.

Одержані результати, отримані для широкого набору тестів показали, що запропоновані рішення HEVC на основі TRA та VRF значно перевершують відповідні стандартні методи покращення HEVC.

#### Список літератури:

1. Enhanced Motion-Compensated Video Coding With Deep Virtual Reference Frame Generation / L. Zhao et al. *IEEE Transactions on Image Processing*. 2019. Vol. 28, no. 10. P. 4832–4844. URL: <https://doi.org/10.1109/tip.2019.2913545> (date of access: 25.04.2022).
2. Neural Inter-Frame Compression for Video Coding / A. Djelouah et al. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, Korea (South), 27 October – 2 November 2019. 2019. URL: <https://doi.org/10.1109/iccv.2019.00652> (date of access: 25.04.2022).
3. QP refinement according to Lagrange multiplier for High Efficiency Video Coding / Bin Li et al. *2013 IEEE International Symposium on Circuits and Systems (ISCAS)*, Beijing, 19–23 May 2013. 2013. URL: <https://doi.org/10.1109/iscas.2013.6571884> (date of access: 25.04.2022).
4. A Novel Rate Control Algorithm Based on  $\rho$  Model for Multiview High Efficiency Video Coding / T. Yan et al. *Electronics*. 2020. Vol. 9, no. 1. P. 166. URL: <https://doi.org/10.3390/electronics9010166> (date of access: 25.04.2022).
5. HoangVan X. Adaptive Quantization Parameter Estimation for HEVC Based Surveillance Scalable Video Coding. *Electronics*. 2020. Vol. 9, no. 6. P. 915. URL: <https://doi.org/10.3390/electronics9060915> (date of access: 25.04.2022).
6. An adaptive QP offset determination method for HEVC / M. A. Papadopoulos et al. *2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA, 25–28 September 2016. 2016. URL: <https://doi.org/10.1109/icip.2016.7533155> (date of access: 25.04.2022).
7. Rate distortion modeling and adaptive rate control scheme for high efficiency video coding (HEVC) / L. Sun et al. *2014 IEEE International Symposium on Circuits and Systems (ISCAS)*, Melbourne VIC, Australia, 1–5 June 2014. 2014. URL: <https://doi.org/10.1109/iscas.2014.6865539> (date of access: 25.04.2022).
8. Xu M., Nguyen Canh T., Jeon B. Simplified Level Estimation for Rate-Distortion Optimized Quantization of HEVC. *IEEE Transactions on Broadcasting*. 2020. Vol. 66, no. 1. P. 88–99. URL: <https://doi.org/10.1109/tbc.2019.2941063> (date of access: 25.04.2022).
9. SSIM-Based Game Theory Approach for Rate-Distortion Optimized Intra Frame CTU-Level Bit Allocation / W. Gao et al. *IEEE Transactions on Multimedia*. 2016. Vol. 18, no. 6. P. 988–999. URL: <https://doi.org/10.1109/tmm.2016.2535254> (date of access: 25.04.2022).

10. Zhang Y., Lu C. A Highly Parallel Hardware Architecture of Table-Based CABAC Bit Rate Estimator in an HEVC Intra Encoder. *IEEE Transactions on Circuits and Systems for Video Technology*. 2019. Vol. 29, no. 5. P. 1544–1558. URL: <https://doi.org/10.1109/tcsvt.2018.2830126> (date of access: 25.04.2022).
11. Hoang Van X. Statistical search range adaptation solution for effective frame rate up-conversion. *IET Image Processing*. 2018. Vol. 12, no. 1. P. 113–120. URL: <https://doi.org/10.1049/iet-ipr.2016.0938> (date of access: 25.04.2022).
12. Bossen F. Common HM Test Conditions and Software Reference Configuration. In Proceedings of the 14th meeting of the Joint Collaborative Team on Video Coding (JCT-VC), Vienna, Austria, 25 July–2 August 2013. URL: [https://www.scirp.org/\(S\(i43dyn45teexjx455qlt3d2q\)\)/reference/ReferencesPapers.aspx?ReferenceID=2150073](https://www.scirp.org/(S(i43dyn45teexjx455qlt3d2q))/reference/ReferencesPapers.aspx?ReferenceID=2150073) (date of access: 25.04.2022).
13. Enhanced Ctu-Level Inter Prediction with Deep Frame Rate Up-Conversion for High Efficiency Video Coding / L. Zhao et al. *2018 25th IEEE International Conference on Image Processing (ICIP)*, Athens, 7–10 October 2018. 2018. URL: <https://doi.org/10.1109/icip.2018.8451465> (date of access: 25.04.2022).
14. Wang M., Ngan K. N., Li H. An Efficient Frame-Content Based Intra Frame Rate Control for High Efficiency Video Coding. *IEEE Signal Processing Letters*. 2015. Vol. 22, no. 7. P. 896–900. URL: <https://doi.org/10.1109/lsp.2014.2377032> (date of access: 25.04.2022).
15. Seong-Gyun Jeong, Chul Lee, Chang-Su Kim. Motion-Compensated Frame Interpolation Based on Multihypothesis Motion Estimation and Texture Optimization. *IEEE Transactions on Image Processing*. 2013. Vol. 22, no. 11. P. 4497–4509. URL: <https://doi.org/10.1109/tip.2013.2274731> (date of access: 25.04.2022).
16. An Overview of Core Coding Tools in the AV1 Video Codec / Y. Chen et al. *2018 Picture Coding Symposium (PCS)*, San Francisco, CA, 24–27 June 2018. 2018. URL: <https://doi.org/10.1109/pcs.2018.8456249> (date of access: 25.04.2022).
17. Overview of the High Efficiency Video Coding (HEVC) Standard / G. J. Sullivan et al. *IEEE Transactions on Circuits and Systems for Video Technology*. 2012. Vol. 22, no. 12. P. 1649–1668. URL: <https://doi.org/10.1109/tcsvt.2012.2221191> (date of access: 25.04.2022).
18. Xu M., Nguyen Canh T., Jeon B. Simplified Level Estimation for Rate-Distortion Optimized Quantization of HEVC. *IEEE Transactions on Broadcasting*. 2020. Vol. 66, no. 1. P. 88–99. URL: <https://doi.org/10.1109/tbc.2019.2941063> (date of access: 25.04.2022).
19. HoangVana X., Ascensob J., Pereiraa F. Improving predictive video coding performance with decoder side information. *2012 19th IEEE International Conference on Image Processing (ICIP 2012)*, Orlando, FL, USA, 30 September – 3 October 2012. 2012. URL: <https://doi.org/10.1109/icip.2012.6467505> (date of access: 25.04.2022).
20. Pixel-Wise Unified Rate-Quantization Model for Multi-Level Rate Control / H. Choi et al. *IEEE Journal of Selected Topics in Signal Processing*. 2013. Vol. 7, no. 6. P. 1112–1123. URL: <https://doi.org/10.1109/jstsp.2013.2272241> (date of access: 25.04.2022).
21. Parallelized Rate-Distortion Optimized Quantization Using Deep Learning / D. Kianfar et al. *2020 IEEE 22nd International Workshop on Multimedia Signal Processing (MMSP)*, Tampere, 21–24 September 2020. 2020. URL: <https://doi.org/10.1109/mmmsp48831.2020.9287135> (date of access: 25.04.2022).
22. Zhang Y., Zhang C., Fan R. Fast Motion Estimation in HEVC Inter Coding: An Overview of Recent Advances. *2018 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)*, Honolulu, HI, USA, 12–15 November 2018. 2018. URL: <https://doi.org/10.23919/apsipa.2018.8659650> (date of access: 25.04.2022).

## **Arpentii S.P. TECHNOLOGY OF OPTIMAL DISTRIBUTION OF SPEED AND VIRTUAL SUPPORT FRAMES FOR HIGHLY EFFICIENT VIDEO CODING**

*The High Performance Video Encoding (HEVC) standard has now become the most popular video encoding solution for video conferencing, broadcast and streaming. However, its compression performance is still a critical issue for the adoption of a large number of new video programs with higher spatial and temporal resolution. To increase the current performance of HEVC, the article proposes an effective distribution solution. The aim of this work is to develop a simple, non-standard MPEG video encoding tool by integrating a virtual reality solution in HEVC and refining the existing bitrate distribution scheme. The paper avoids the well-known approach of preliminary analysis, which requires the provision of several frames or group groups, which subsequently creates significant complexity and is not suitable for a number of applications. To achieve this goal, a heuristic quantization approach is proposed. In general, this approach usually faces the problem of the enormous computational complexity of video encoding, different data sets, and the nonlinearity of the encoding process. The proposed method adaptively distributes the compression rate for each encoded image in the image group using a dynamic lattice-based programming approach. To achieve this goal, the paper investigates lattice-based quantization parameters for each frame in a group of images, taking into account*

*the position of the time layer. Dynamic lattice-based programming is used to recursively search for the best offset of each time index (Tid). To study the creation of virtual reference frames (VRF) in the proposed HEVC, the image quality and computational complexity of virtual reference frames were evaluated. Creating a VRF, adopted in the article, includes three main steps: hierarchical motion estimation (HME) to initialize the motion vector field; motion vector refinement (MVR) to reduce quantization errors that often occur in decoded links; motion compensation for VRF interpolation. The results obtained for a wide range of tests showed that the proposed HEVC solutions based on TRA and VRF significantly exceed the corresponding standard methods of improving HEVC.*

**Key words:** *video encoding, virtual reality, heuristic quantization, dynamic programming, virtual reference frames, time index.*