

УДК 621.396.4

Новіков В.І., ст.викладач;
Наземнов В.В., магістр
НТУУ «КПІ»

МЕТОД ТА АЛГОРИТМ КОДУВАННЯ ВІДЕОІНФОРМАЦІЇ ПРИКЛАДНИХ ТЕЛЕВІЗІЙНИХ СИСТЕМАХ ІЗ МІНІМАЛЬНОЮ ПРОПУСКНОЮ ЗДАТНІСТЮ КАНАЛУ ТА МАКСИМАЛЬНИЙ ЯКОСТІ ПЕРЕДАЧІ ВІДЕОІНФОРМАЦІЇ

Запропонований метод та алгоритм кодування відеоінформації, що забезпечує в прикладних телевізійних системах, реалізованих як система на кристалі, мінімальну необхідну пропускну здатність каналу при максимумі якості відеоінформації. Даний алгоритм показав максимальну ефективність кодування, оцінювану згідно введеному узагальненому показнику ефективності кодера, на основі розробленої методики, що спирається на сформовану компактну представницьку вибірку сюжетів, що враховує запропоновану міру нестационарності сюжетів.

Предложенный метод и алгоритм кодирования видеoinформации, обеспечивающий в прикладных телевизионных системах, реализуемых как система на кристалле, минимальную требуемую пропускную способность канала при максимуме качества видеoinформации. Данный алгоритм показал максимальную эффективность кодирования, оцениваемую согласно введеному обобщенному показателю эффективности кодера, на основе разработанной методики, опирающейся на сформированную компактную представительную выборку сюжетов, учитывающую предложенную меру нестационарности сюжетов.

The proposed algorithm is a video encoding, providing applications in television systems, implemented as a system on a chip, the minimum required bandwidth at the maximum quality of the video. This algorithm showed the maximum coding efficiency, estimated according to the introduced generalized encoder performance indicators on the basis of the developed method, based on a representative sample of the generated compact plots, taking into account the measure proposed nonstationary scenes.

Вступ. Розробка методу та алгоритму адаптивного стиснення відеоінформації в системах на кристалі є актуальною, так як системи на кристалі є найсучаснішим напрямом розвитку мікроелектроніки. При цьому головним аспектом актуальності є економія пропускну здатності. Дефіцит пропускну спроможності - «вічна» проблема техніки зв'язку. Класичні розробки кодерів джерел відеоінформації в силу нестачі

апріорної інформації про джерело не дозволяють досягти граничної ефективності кодування. У зв'язку з націленістю даної роботи на проектування прикладних (в першу чергу бортових) телевізійних систем акцент в роботі робиться на синтезі кодера джерела без врахування обмежень на складність реалізації декодера на приймальному пункті.

Аналіз досліджень і публікацій. Істотний внесок у вирішення проблем передачі, характерних для змішаних систем зв'язку (безперервне джерело і цифровий канал зв'язку), внесли роботи: К. Шеннона, А. Н. Колмогорова, Б. Олівера, Р. Л. Добрушина, Б. С. Цибакова, Л. П. Ярославського, Л. І. Хромова, А. К. Цицуліна та ін. Дана робота є розвитком відомих результатів і додатком їх до кодування телевізійних сигналів, в якому застосовуються методи кодування з перетворенням (JPEG, MPEG і т.п.). У розвиток цих методів внесли вклад У. Претт, Р. Гонсалес, Р. Вудс, Д. Селомон, Я. Річардсон, Ю. Б. Зубарев, В. П. Дворкович, М. І. Кривошеєв, І. І. Цуккерман, А. А. Умбіталієв, Р. Є. Биков, А. І. Величкін та ін. Отже існуючи методи кодування та оцінки якості кодування відео потребують додаткового дослідження, удосконалення та систематизації в єдину методичку.

Постановка завдання. Мета роботи - визначення граничної ефективності кодування нестационарних телевізійних сигналів при обмеженні складності кодера. Для досягнення поставленої мети були поставлені та вирішені наступні завдання:

1) показана можливість досягнення інформаційного рівноваги, тобто рівності швидкості створення інформації джерелом і пропускної здатності каналу, в змішаній системі зв'язку, що складається з безперервного джерела і дискретного каналу, при певній комбінації апріорної інформації, критерію якості та обмеження;

2) розроблений адаптивний алгоритм кодування для прикладних бортових телевізійних систем з урахуванням обмеження складності кодування в умовах апріорної невизначеності та нестационарності телевізійних сигналів;

3) розроблена методика об'єктивної оцінки різних алгоритмів стиснення з урахуванням введених заходів нестационарності і узагальненого показника ефективності кодера;

4) сформовані компактні представницькі вибірки статичних зображень і відео для визначення ефективності кодування різними алгоритмами.

5) визначена гранична ефективність кодування при обмеженні складності кодера.

Розробка адаптивного алгоритму кодування джерела відеоінформації з перетворенням в частотну область.

Узагальнення концепції рівноважного погодження для змішаної системи зв'язку, на джерела відеоінформації, що володіють

нестационарністю, можливо за рахунок використання в алгоритмі кодування процедури адаптації до властивостей сюжету. Використання адаптації має на меті вирівнювання ймовірностей вторинних символів (спектральних коефіцієнтів перетворення), і при використанні як критерію якості максимуму модуля помилки, а як обмеження - сукупного посилення кодера, забезпечується рівність швидкості створення інформації джерелом і пропускної здатності каналу. В якості перетворення було вибрано тривимірне ДКП, так як воно володіє хорошими корелюючими властивостями і для нього існують швидкі алгоритми.

Обмеження на складність, типове для відеосистем на кристалі, диктує

необхідність розробки методу та алгоритму помірної складності. Розроблені метод та алгоритм кодування динамічно визначає оптимальні розміри відеосегментів, до якого застосовується тривимірне ДКП, на основі аналізу руху. Структурна схема кодера на основі тривимірного ДКП показана на рис. 1. Кодер включає наступні основні блоки: блок управління та аналізу руху; блок тривимірного ДКП; блок квантування і блок ентропійного кодування. Декодер містить ті ж блоки, становлення в зворотному порядку.

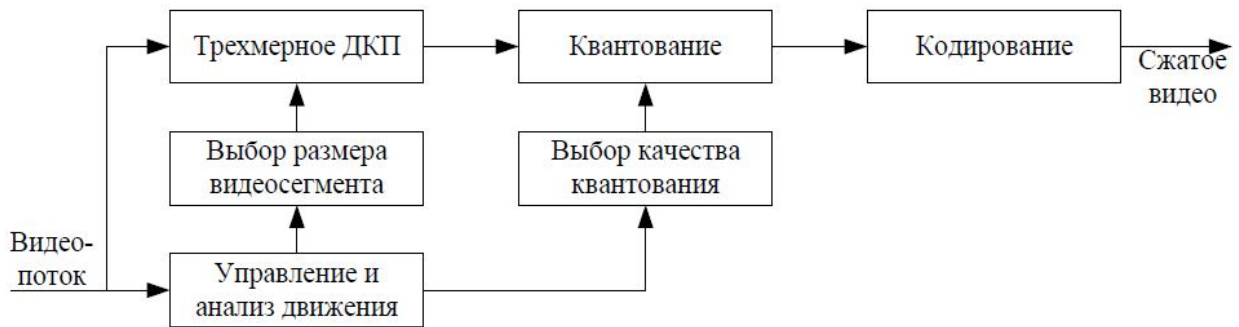


Рис. 1. Модель адаптивного кодування відеоінформації на основі тривимірного ДКП

Метод кодування включає 3 етапи.

Етап 1. Попередній аналіз послідовності кадрів, в результаті якого формуються три групи кадрів:

- групи без руху (ГБД), або статичні ;
- групи з малою швидкістю руху (ГМСД), або квазістатичні;
- групи з великою швидкістю руху (ГБСД), або динамічні.

При аналізі руху черговий кадр записується в буфер базового кадру (перший кадру формованої групи). Усі наступні кадри порівнюються з базовим шляхом вирахування для визначення залишкових кадрів (різниця між поточним і базовим кадрами). Ступінь тимчасової декореляції ε_t визначається за формулою:

$$\varepsilon_t = \frac{1}{NM} \sum_{i=0}^{N-1} \sum_{j=0}^{M-1} |Y_{ij} - Y'_{ij}|,$$

Де $Y_{ij}t$ та Y'_{ij} - яскравості пікселів поточного і базового кадрів відповідно; M і N - розміри кадру

Залежно від значення ε_t вибирається число кадрів в групі n_t :

- якщо $0 \leq \varepsilon_t \leq \varepsilon_{\text{доп1}}$ ($\varepsilon_{\text{доп1}}$ - поріг віднесення кадру до статичних), то n_t збільшується на одиницю і дана ітерація повторюється доти, поки відсутній рух. У результаті формуються ГБД, до яких застосовується двовимірне ДКП над базовим кадром. Просторові розміри сегмента в цьому випадку встановлюються рівними 16×16 пікселів.

- якщо $\varepsilon_{\text{доп1}} < \varepsilon_t \leq \varepsilon_{\text{доп2}}$ ($\varepsilon_{\text{доп2}}$ - поріг віднесення кадру до квазістатичним), то n_t збільшується на одиницю і дана ітерація повторюється доти, поки є повільний рух. У результаті формуються ГМСД, до яких застосовується тривимірне ДКП. Розміри відеосегментів при цьому складають $8 \times 8 \times n_t$.

- якщо $\varepsilon_t > \varepsilon_{\text{доп2}}$, то $n_t=4$ і є постійним. У результаті формуються ГБСД, до яких також застосовується тривимірне ДКП. Відеосегменти мають розміри $4 \times 4 \times 4$ пікселів.

Результати модельного експерименту над різними відеоданими, кодованими з яскравості здатністю 8 біт / піксель, показали, що доцільно встановити порогові $\varepsilon_{\text{доп1}}=4$ і $\varepsilon_{\text{доп2}}=24$.

Як і для алгоритму MPEG, особливою ситуацією є зміна сюжету, викликає короточасне порушення кореляції. Ця ситуація відстежується за величиною ε_t : якщо вона перевищує половину динамічного діапазону, то приймається рішення про зміну сюжету, і новий кадр стає опорним. Якщо зміна сталася при формуванні ГБСД, то що залишилися кадри групи кодуються як I- кадри MPEG.

У межах кожного з відеосегментів групи аналізується рух, і залежно від наявності або відсутності руху виконується тривимірне або двовимірне (по простору) ДКП.

Етап 2. Квантування і сканування.

Для квантування і z-сканування отриманий тривимірний набір просторово-часових частотних коефіцієнтів підрозділяється на слайси, відповідні одному значенню часової частоти. У кожному слайсі квантування і z-сканування проводяться незалежно від інших слайсів.

Просторово-частотні коефіцієнти слайса, отримані в результаті тривимірного ДКП, квантуються діленням на коефіцієнт квантування $QC_{ij} = 1 + (1 + i + j)R$, де i та j - координати відліку всередині слайса; R - параметр квантування, що задається користувачем.

Етап 3. Ентропійне кодування.

Квантовані коефіцієнти тривимірного ДКП зчитуються і кодуються методом Хаффмана.

Для оцінки розробленого алгоритму кодування в дисертації був введений узагальнений показник ефективності кодека, що враховує не тільки міру наближення до епсилон-ентропії джерела, але і те, якими інформаційними засобами досягнуто дане наближення. Слідуючи методиці векторного синтезу системи, сформований узагальнений показник ефективності кодека P , що включає вектор концепції кодека $\{c_i\}$ і сукупність приватних інформаційних показників якості кодека $\{P_i\}$: втрати корисної інформації ΔI , швидкості передачі R , складності кодера S_k і складності декодера S_d . Так як всі приватні інформаційні показники якості кодека пов'язані з помилкою передачі ε , то нижче вони позначені як $P_i(\varepsilon)$:

$$P = c_0 \Delta I(\varepsilon) + c_1 R(\varepsilon) + c_2 S_k(\varepsilon) + c_3 S_d(\varepsilon).$$

В окремому випадку, коли складність декодера не враховується, тобто $c_3 = 0$, що характерно для бортових систем прикладного телебачення, які акцентують увагу на складності кодера, узагальнений показник ефективності кодера має вигляд:

$$P = c_0 \Delta I(\varepsilon) + c_1 R(\varepsilon) + c_2 S_k(\varepsilon).$$

Очевидно, що в різних кодерах вектори концепції $\{c_i\}$ можуть істотно різнитися, що і призводить до існування безлічі різних оптимальних кодеків. Наприклад, навіть при рівності вагових коефіцієнтів при швидкості передачі і помилку передачі системи кодування в телевізійному мовленні і в аерокосмічних системах спостереження істотно розрізняються в співвідношенні вагових коефіцієнтів при складності кодера і декодера. У телевізійному мовленні істотну вагу в силу величезних накладів має складність декодера, в аерокосмічних системах зв'язку чинності жорстких масогабаритних обмежень має складність бортового кодера. Для забезпечення загального інформаційного підходу до вирішення задачі оптимізації кодека всі показники якості приведені до єдиного вигляду, що має розмірність інформації (біт). Зокрема, в силу того, що всі реальні безперервні сигнали доступні кодуванню тільки в суміші з шумом, точність передачі слід характеризувати не просто помилкою передачі ε , а втратою корисної інформації $\Delta I(\varepsilon)$.

Мінімум при варіації призначається помилки передачі ε може бути знайдений аналітично, якщо відомі статистичні властивості джерела і залежності складності і швидкості передачі від призначуваної помилки. Відомо, що узагальнений показник якості системи допускає безліч рішень. Ця різноманітність рішень в першу чергу пов'язано з зазначеним взаємообміном швидкості передачі і складності кодера. Узагальнений показник ефективності кодера дозволяє об'єктивно порівнювати довільні кодери, що відрізняються за всіма параметрами: помилку передачі, швидкості передачі і складності, якщо зафіксувати вибірку випробувальних сигналів.

Оскільки алгоритм адаптивного кодування, є спрощеним, то для нього необхідна методика оцінки його ефективності. Розробники алгоритмів кодування перевіряють ефективність їх роботи на деякій вибірці різних сюжетів, негласно вважаючи, що ця вибірка є представницькою. Оскільки йдеться про кодування нестационарних сигналів, то для порівняння ефективності кодування різних алгоритмів пропонується використовувати таку методику, засновану на обліку коефіцієнта нестационарності:

- 1) ввести міру нестационарності сигналів зображень;
- 2) розробити алгоритм обчислення введеної заходи ;
- 3) визначити межі зміни цього запобіжного для дуже великої вибірки сигналів (завідомо представницької);
- 4) зробити розбиття інтервалу зміни цього запобіжного на кілька (за можливості рівних) інтервалів, кількість яких має бути достатнім для впевненості в показності вибірки;
- 5) відібрати зображення, що належать до виділених інтервалах заходи нестационарності;
- 6) провести кодування сформованої вибірки кожним з досліджуваних алгоритмів;
- 7) для кожного алгоритму і для кожного елемента вибірки оцінити значення швидкості передачі інформації при заданій помилку передачі (іноді можлива інверсія - оцінити помилку передачі при заданій швидкості передачі);
- 8) при відсутності статистичних даних про ймовірності появи сюжетів з різною мірою нестационарності вважати їх рівноімовірними;
- 9) для кожного алгоритму усереднити по сформованої вибірці значення швидкості передачі при заданій помилку (або усереднити помилку передачі при заданій швидкості передачі);
- 10) порівняти результати усереднення і вибрати алгоритм, що забезпечує найменшу швидкість передачі при заданій помилку (найменшу помилку при заданій швидкості передачі).

Для реалізації пропонованої методики ключовими питаннями стають перші два пункти методики: вибір заходи нестационарності сигналів і розробка алгоритму обчислення цього заходу.

Для оцінки нестационарності сигналів зображень представляється доцільним скористатися методологією статистичної радіотехніки, в якій використовуються поняття стаціонарності в широкому і вузькому сенсі, тобто розрізняють процеси з постійною / змінною дисперсією, математичним очікуванням і автокорреляційною функцією. Це підказує, що узагальнена міра нестационарності повинна враховувати сукупна зміна різних статистик S_n сигналу зображення. Це можуть бути математичні очікування, дисперсії, інтервали (або коефіцієнти) кореляції яскравості і кольору та ін. Кожна з статистик обчислюється для деяких інтервалів: це

можуть бути оцінені за окремим адаптивним алгоритмам інтервали стаціонарності або апіорі компромісно обрані інтервали. Вони повинні бути не занадто малими для отримання прийнятних методологічних помилок обчислення статистик і не занадто великими для зменшення ймовірності зміни статистик усередині інтервалу. Можна вважати, що саме такі ж міркування лежали в основі вибору вікна (8 × 8) елементів при розробці стандартів MPEG.

Загальна міра нестаціонарності M повинна обчислюватися як сума (у загальному випадку зважена) заходів нестаціонарності окремих статистик:

$$M = \sum_{n=1}^K M_n$$

Для оцінки міри нестаціонарності кожної статистики сигналу пропонується використовувати нормовану статистику, яка обчислюється аналогічно середньоквадратической смузі частот з тією лише різницею, що в формулу повинні входити не самі сигнали, а їх статистики S_n , обчислювані окремо по кожному аргументу. Нижче для спрощення записаний один узагальнений дискретний аргумент $i \in [1, N]$. Формула для безперервних функцій неперервного аргументу включає ставлення інтегралів від квадрата похідної сигналу до інтеграла від квадрата самого сигналу. Для дискретних значень статистик можна використовувати не похідну, а першу кінцеву різницю:

$$M_n = \frac{\sum_{i=1}^{N-1} (S_{n(i+1)} - S_{ni})^2}{\sum_{i=1}^N S_{ni}^2}$$

Тут для обчислення заходи нестаціонарності розглядається сукупність статистичних характеристик (середнє значення m_i , дисперсія D_i і коефіцієнт кореляції r_i) вихідного зображення, обчислювані в межах вікна (8 x8):

$$M = M_1 + M_2 + M_3 = \frac{\sum_{i=1}^{N-1} (m_{i+1} - m_i)^2}{\sum_{i=1}^N m_i^2} + \frac{\sum_{i=1}^{N-1} (D_{i+1} - D_i)^2}{\sum_{i=1}^N D_i^2} + \frac{\sum_{i=1}^{N-1} (r_{i+1} - r_i)^2}{\sum_{i=1}^N r_i^2}$$

Важливою характеристикою ефективності кодування зображень різними алгоритмами є швидкість передачі інформації. У першому наближенні вона визначається смугою частот, займаною сигналом, тому очевидно, що характеризувати зображення тільки лише коефіцієнтом нестаціонарності можна. Тому, поряд з коефіцієнтом нестаціонарності, враховується також смуга частот, займана сигналом зображення. Для оцінки смуги частот нестаціонарного сигналу зображень використовується

формула, широко застосовувана в статистичній радіотехніці для оцінки середньоквадратичної смуги частот (заходи широкосмуговості):

$$W = \sqrt{\frac{\sum_{k=1}^{N/2} k^2 |\bar{X}(k)|^2}{\sum_{k=1}^{N/2} |\bar{X}(k)|^2}}$$

де $|\bar{X}(k)|$ - значення спектральної складової, отримане шляхом усереднення спектральних складових по всіх рядках зображення.

З урахуванням заходів нестационарності і заходи широкосмуговості була сформована компактна представницька вибірка статичних зображень. Поширюючи такий підхід на реальні відео, які характеризуються нестационарністю не тільки по простору, але і за часом, вдалося сформуванати компактну представницьку вибірку відео. На прикладі використання компактної представницької вибірки відео була проведена оцінка розробленого алгоритму стиснення на основі тривимірного ДКП і оцінені основні параметри, що визначають його ефективність. Ефективність запропонованого алгоритму показана на прикладі відеоданих трьох типів: з малої, середньої та великої швидкістю руху. На рис. 2 показані графіки залежності коефіцієнта стиснення K , швидкості передачі C [біт / піксел] і відносини сигнал / шум Ψ [дБ] від параметра квантування R , отримані в результаті моделювання обробки відеоданих з різними швидкостями руху адаптивним алгоритмом на основі тривимірного ДКП. На рис. 2 графіком 1 відповідає ГМСД, графіку 2 - ГБСД, а графіку 3 - ГБСД, закодована алгоритмом на основі тривимірного ДКП без адаптації з розміром відеосегментів $8 \times 8 \times 8$. Коефіцієнт стиснення визначався як відношення розміру вихідного (нестисненого) відеофайлу (в бітах) до розміру стисненого відеофайлу (в бітах). Відношення "сигнал / шум" визначалося за формулою $\Psi = 20 \lg(255/\epsilon)$, де

$$\epsilon = \sqrt{\frac{1}{MNT} \sum_{xyt} (Y_{и xyt} - Y_{в xyt})^2}$$

- середньоквадратична помилка кодування, причому T - кількість кадрів в відео потоці;

$Y_{и xyt}$ і $Y_{в xyt}$ - яскравості пікселів з координатами x, y, t в вихідному і відновленому після квантування кадрах відповідно.

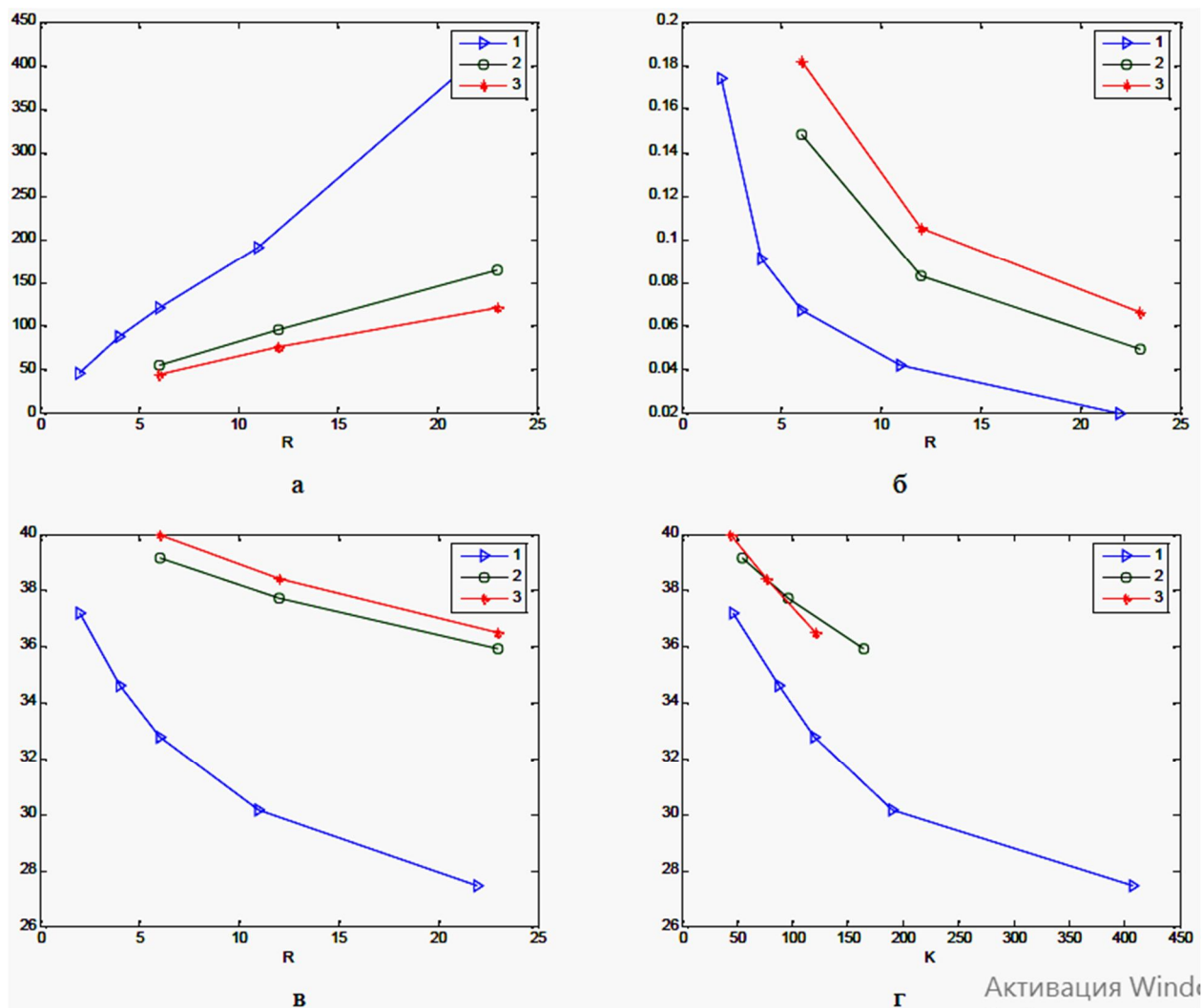


Рис. 2. Залежність коефіцієнта стиснення K (а), пропускної здатності C (б), відношення сигнал / шум Ψ (в) від параметра квантування R і залежність ставлення сигнал / шум Ψ (г) від коефіцієнта стиснення K

Зпорівняння результатів, отриманих при використанні алгоритму без адаптації розміром відеосегментів 8×8 пікселів з результатами застосування адаптивного алгоритму до ГБСД, сліднявність вирашу розробленого алгоритму при близькому якість відновлення відеоданих. Крім того, використання тривимірного ДКП при кодуванні відеоданих дозволяє на декілька порядків скоротити складність кодера джерела в порівнянні з кодером MPEG-4. Це обумовлено, по-перше, тим, що ДКП є асимптотично оптимальним (при великих розмірах фрагментів) перетворенням стаціонарного сигналу з експоненціальною автокорреляційною функцією, добре апроксимуючої реальні зображення. По-друге, ці два види кодерів оптимальні для різних характеристик зміни сюжету в часі: MPEG - 4 орієнтований на спостереження за об'єктами, а тривимірне ДКП виділяє будь-які зміни в сюжеті.

Висновки. Запропоновані метод та алгоритм кодування відеоінформації забезпечують в прикладних телевізійних системах, реалізованих як система на кристалі, мінімальну потребовану пропускну здатність каналу при максимумі якості відеоінформації. Метод і алгоритм показали максимальну ефективність кодування, оцінювану згідно введеному узагальненому показникові ефективності кодера, на основі розробленої методики, що спирається на сформовану компактну представницьку вибірку сюжетів, що враховує запропоновану міру нестационарності сюжетів.

Використані джерела інформації:

1. В.И. Воробьев, В.Г. Грибунин. Теория и практика вейлет-преобразования. СПб.: ВУС, 1999. 204 с.
2. Плахов А. Г., Чобану М. К Обзор элементной базы для построения систем сжатия видеосигнала. // Труды Междун. конгресса Академии информатизации ITS-2002 / МЭИ (ТУ), т. 3. М., 2002, с. 123-125.
3. Зубакин, И. А. Аппаратно-программное проектирование сложных функциональных блоков с использованием систем на кристалле [Текст] / Ш. С. Фахми, С. С. Шагаров // Научно-технический вестник Санкт-Петербургского государственного университета информационных технологий, механики и оптики. – 2010. – № 02 (66). – С. 90–98.
4. Математичні основи теорії телекомунікаційних систем / За загальною ред. В. В. Поповського. – Харків: ТОВ «Компанія СМІТ», 2006. – 564 с.