

## АЛГОРИТМИ ШВИДКОГО ВИКОНАННЯ 4-ТОЧКОВИХ ЦІЛОЧИСЛОВИХ СИНУСНИХ ПЕРЕТВОРЕНЬ ТИПУ VII БЕЗ МНОЖЕННЯ І РОЗДІЛЬНІ НАПРАВЛЕНІ АДАПТИВНІ ПЕРЕТВОРЕННЯ ДЛЯ INTRA-ПРОГНОЗУВАННЯ В КОДУВАННІ ЗОБРАЖЕНЬ/ВІДЕО. I

**Анотація.** Запропоновано роздільні направлені адаптивні цілочислові косинусне та синусні типу VII перетворення з блоками  $4 \times 4$  для intra-прогнозування в кодуванні зображень і відео. Розроблено алгоритми швидкого виконання 2D 4-точкових роздільних направлених цілочислових косинусного та синусних перетворень. Вони порівняно з відомими алгоритмами потребують у сім разів менше операцій множення та забезпечують більш високий ступінь стиснення. Запропоновано два цілочислові синусні перетворення типу VII порядку 4 і розроблено алгоритми швидкого виконання 4-точкових цілочислових синусних перетворень типу VII, які потребують тільки цілочислових операцій. Вони порівняно з відомим алгоритмом дають змогу суттєво скоротити кількість операцій множення, або взагалі обійтися без них. Запропоновані цілочислові синусні перетворення є найближчими до дискретного синусного перетворення типу VII і мають більш високий ступінь стиснення, ніж відоме синусне перетворення, прийняте у стандарті H.265.

**Ключові слова:** дискретне косинусне перетворення, дискретне синусне перетворення, цілочислове косинусне перетворення, цілочислове синусне перетворення, модозалежне направлене перетворення, роздільне направлене адаптивне перетворення, факторизація, швидке виконання перетворення, intra-прогнозування, ефективність кодування, відеокодування, H.264, H.265.

### ВСТУП

Для усунення надлишковості в зображеннях і відеосигналах шляхом їхнього ефективного стиснення використовуються методи кодування на основі перетворення. Вибір найкращого перетворення для конкретного застосування залежить від величини допустимої похибки відновлення і від обчислювальних витрат. Більшість стандартів кодування зображень і відео, наприклад JPEG, H.264/AVC, VC-1, H.265/HEVC, використовують блочне кодування на основі перетворення як інструмент для ефективного стиснення вхідного зображення і відеосигналів. Стандарти стиснення відеоданих розширюють методи блочного кодування з перетворенням нерухомих зображень шляхом скорочення часової або міжкадрової надлишковості. Мета етапу перетворення під час оброблення відеокодеком — відображення цифрових даних, повних або залишкових кадрів в іншу координатну область (область перетворення). Дані піксельної області перетворюються на частотну область за допомогою перетворення блоку за блоком. Для типових зображень більша частина енергії зосереджена в низькочастотних коефіцієнтах перетворення. Після перетворення для більш високочастотних коефіцієнтів перетворення можна використати квантувач з більшим кроком квантування, який забезпечить більш ефективне і компактне упакування для досягнення кращого стиснення. Отже, потрібно розробити оптимальне перетворення для кожного блоку зображення, щоб отримати повністю декорельовані коефіцієнти перетворення. Відомо, що перетворення Карунена–Лоева (ПКЛ) має кілька привабливих властивостей, якими є, зокрема, квантування з високою роздільною здатністю гаусових сигналів і повна декореляція коефіцієнтів перетворення. Для блоку  $N \times N$  матриця ПКЛ

має розмір  $N^2 \times N^2$ . Тому практичне використання ПКЛ є обмеженим через високу обчислювальну складність і високі вимоги до пам'яті. В [1] показано, що дискретне косинусне перетворення (ДКП) є привабливою альтернативою ПКЛ з погляду ущільнення енергії та продуктивності, близьких до ПКЛ. Але з появою intra- (внутрішнього) прогнозування це вже не так, і оптимальне перетворення має бути адаптивним до inter- чи intra-прогнозувальної моди (режиму). Джайн [2, 3] показав, що ПКЛ для залишкових процесів буде дискретним синусним перетворенням (ДСП), коли граничні умови є доступними в обох напрямках. Мері та Юділевіч [4] спочатку кодували вміст блоку, потім окремо границі, а наприкінці використовували синусне перетворення типу «pinned sine transform». Для покращення кодування з перетворенням прогнозованих блоків у моді intra в [5] було введено модозалежне направлене перетворення («mode-dependent directional transform», MDDT), яке є ядром тестової моделі TMuC [6] стандарту відеокодування H.265/HEVC [7]. У роботі [8] для найкращої апроксимації нероздільного направленого перетворення [5] було використано тільки дві матриці перетворення: ДКП і отриманого ПКЛ. Було показано, що залишкові процеси, отримані в результаті intra-прогнозування під час кодування зображень/відео представляють сепарабельну кореляційну модель зображення з коефіцієнтами кореляції між сусідніми пікселями у вертикальній та горизонтальній проєкціях. Відмічено, що ПКЛ коваріаційної теплоцевої матриці, отриманої з рядка залишку після intra-прогнозування, наближено є ДКП. Оскільки коваріаційна матриця, отримана зі стовпця залишку, не є теплоцевою матрицею, то ДКП є підоптимальним (субоптимальним) перетворенням. Оптимальним перетворенням ПКЛ, яке приводить цю коваріаційну матрицю до діагонального вигляду, є трьохдіагональна матриця, власні вектори якої можуть бути обчислені [9] як дискретне синусне перетворення типу VII (ДСП-VII). Аналогічний результат в аналітичний спосіб отримано в [10] для гаусово-марковського процесу першого порядку, де показано, що ПКЛ для цього процесу є ДСП-VII для коефіцієнта кореляції  $\rho$ , що прямує до нуля, і було вперше опубліковано модозалежне ДКП/ДСП.

У пропозиції [8] до стандарту HEVC отримано цілочислові апроксимації матричного обчислення ДСП-VII порядку  $N = 4, 8, 16$ , а також наведено алгоритм 4-точкового швидкого виконання цілочислового синусного перетворення типу VII (ЦСП-VII), який не є досконалим за мультиплікативними операціями, оскільки потребує виконання восьми операцій множення.

У цій роботі представлено подальший розвиток апарату алгоритмічного забезпечення цілочислових роздільних направлених перетворень для intra-прогнозування блоків  $4 \times 4$  у кодуванні зображень і відео з метою суттєвого скорочення обчислювальних витрат та збільшення ступеню стиснення, що призводить до виграшу у бітрейті.

#### ЦІЛОЧИСЛОВЕ СИНУСНЕ ПЕРЕТВОРЕННЯ ТИПУ VII ПОРЯДКУ 4

Матрицю  $N$ -точкового ДСП типу VII (ДСП-VII) можна визначити, як [9, 11]

$$[S_N^{\text{VII}}]_{k,n} = \frac{2}{\sqrt{2N+1}} \sin \left[ \frac{(2k+1)(n+1)\pi}{2N+1} \right], \quad k, n = 0, 1, \dots, N-1.$$

У роботі [11] наведено три способи представлення матриці  $S_N^{\text{VII}}$  ДСП-VII через уявну частину  $2N+1$ -точкового дискретного перетворення Фур'є, а також представлено алгоритми швидкого обчислення ДСП-VII порядку 4 і 8. У роботах [8, 12] наведено цілочислові апроксимації матричного обчислення ДСП-VII, які для порядку  $N = 4$  збігаються, а для  $N = 8, 16$  є різними.

Розглянемо матрицю  $IST_4^{VII*}$  розміру  $4 \times 4$  цілочислового синусного перетворення типу VII (ЦСП-VII) з переставленими рядками і стовпцями, отриману з матриці  $IST_4^{VII}$  шляхом перестановки рядків на основі простих досконалих перестановок (ПДП) і перестановки стовпців на основі двійково-інверсних перестановок (ДІП) [13]

$$IST_4^{VII*} = \tilde{P}_4 IST_4^{VII} P_4, \quad (1)$$

де  $\tilde{P}_4, P_4$  — матриці  $4 \times 4$  ПДП і ДІП,  $\tilde{P}_4 = \text{antidiag}[I_2, I_2]$ ,  $P_4 = \text{diag}[1, \bar{I}_2, 1]$ ,  $I_2, \bar{I}_2$ , — одинична і антидіагональна одинична матриці  $2 \times 2$ ,

$$\tilde{P}_4 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, P_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \bar{I}_2 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Матрицю  $IST_4^{VII*}$  розміру  $4 \times 4$  ЦСП-VII з переставленими рядками і стовпцями можна представити матрицею ядра ЦСП-VII:

$$IST_4^{VII*} = B_4 S_4^{VII*}, \quad (2)$$

де  $S_4^{VII*}$  — матриця  $4 \times 4$  ядра ЦСП-VII з переставленими рядками і стовпцями,  $B_4$  — діагональна матриця  $4 \times 4$  коефіцієнтів нормування.

Матриця  $S_4^{VII*}$  містить цілі елементи  $\pm a, b, \pm c, \pm d$  і нуль та має такий вигляд:

$$S_4^{VII*} = \begin{bmatrix} d & -c & -a^* & b \\ b^* & c & -d & -a \\ a & c & b & d^* \\ c & 0 & c & -c \end{bmatrix}, \quad (3)$$

де  $a < b < c < d$ ,  $-a^* = b - d$ ,  $b^* = d - a$ ,  $d^* = a + b$ . Тут матрицю  $S_4^{VII*}$  можна представити прямокутною матрицею  $S_{3,4}^{VII*}$  розміру  $3 \times 4$  з вектор-рядком  $S_{1,4}^*$  розмірності 4. При цьому

$$S_{3,4}^{VII*} = \begin{bmatrix} d & -c & -a^* & b \\ b^* & c & -d & -a \\ a & c & b & d^* \end{bmatrix}, S_{1,4}^* = (c, 0, c, -c). \quad (4)$$

Матрицю  $S_{3,4}^{VII*}$  можна представити добутком двох фактор-матриць:

$$S_{3,4}^{VII*} = S_2 S_1, \quad (5)$$

де  $S_1, S_2$  — фактор-матриці розміру  $4 \times 4$  і  $3 \times 4$  алгоритму, запропонованого в [14] швидкого виконання 4-точкового прямого ЦСП-VII:  $S_1 = T_4$ ,  $S_2 = T_{3,4}$ . При цьому фактор-матриця  $4 \times 4$   $T_4$  містить ненульові елементи  $\pm 1$ .

Тут матрицю  $T_{3,4}$  можна представити як добуток двох матриць:

$$T_{3,4} = M_{3,4} R_4, \quad (6)$$

де  $R_4$  — діагональна матриця  $4 \times 4$  з елементами  $d, c, a, b$ ,  $M_{3,4}$  — фактор-матриця  $3 \times 4$ , яка містить в кожному рядку по три ненульових елементи  $\pm 1$ ,

$$R_4 = \text{diag}[d, c, a, b]. \quad (7)$$

#### АЛГОРИТМ ШВИДКОГО ВИКОНАННЯ 4-ТОЧКОВОГО ЦСП-VII

Алгоритм швидкого виконання 4-точкового ЦСП-VII обчислюється за три ітерації. На першій ітерації, що полягає у множенні фактор-матриці  $T_4$  (5), яка містить у трьох рядках по два ненульові елементи  $\pm 1$ , на вхідний вектор-стовпець  $x_4$  розмірності 4, потрібно виконати три операції додавання. На другій ітерації виконується множення отриманих результатів першої ітерації на цілі елементи  $d, c, a, b$  згідно з діагональною матрицею  $R_4$  (7), що потребує виконання чотирьох операцій множення. На третій ітерації, яка полягає у множенні прямокутної фактор-матриці  $M_{3,4}$ , що містить у трьох рядках по три ненульових елементи  $\pm 1$ , на вектор-стовпець  $X_4$  результатів другої ітерації, потрібно виконати шість операцій додавання. При цьому також виконується операція множення вектор-рядка  $S_{1,4}^*$ , який згідно з (4) містить елементи  $\pm c$  і нуль, на вхідний вектор-стовпець  $x_4$ , що потребує виконання двох операцій додавання і однієї операції множення на елемент  $c$ . Усього алгоритм потребує виконання п'яти операцій множення і 11 операцій додавання.

#### АЛГОРИТМ ШВИДКОГО ВИКОНАННЯ 4-ТОЧКОВОГО ОБЕРНЕНОГО ЦСП-VII

Матрицю  $(S_4^{\text{VII}})^{-1}$  оберненого ЦСП-VII порядку 4 можна отримати транспонуванням

$$(S_4^{\text{VII}})^{-1} = S_4^{\text{VII}*T}. \quad (8)$$

Матриця  $S_4^{\text{VII}*T}$  має вигляд

$$S_4^{\text{VII}*T} = \begin{bmatrix} d & b^* & a & c \\ -c & c & c & 0 \\ -a^* & -d & b & c \\ b & -a & d^* & -c \end{bmatrix}. \quad (9)$$

Матрицю  $S_4^{\text{VII}*T}$  можна представити прямокутною матрицею  $S_{3,4}^{\text{VII}*T}$  розміру  $3 \times 4$  і вектор-рядком  $S_{1,4}^{\text{VII}*T}$  розмірності 4. При цьому

$$S_{3,4}^{\text{VII}*T} = \begin{bmatrix} d & b^* & a & c \\ -a^* & -d & b & c \\ b & -a & d^* & -c \end{bmatrix}, \quad S_{1,4}^{\text{VII}*T} = (-c, c, c, 0). \quad (10)$$

Матрицю  $S_{3,4}^{\text{VII}*T}$  можна представити як добуток двох фактор-матриць:

$$S_{3,4}^{\text{VII}*T} = S_2^{-1} S_1^{-1}, \quad (11)$$

де  $S_1^{-1}, S_2^{-1}$  — обернені фактор-матриці алгоритму швидкого виконання 4-точкового оберненого ЦСП-VII:

$$S_1^{-1} = T_4^{-1}, \quad S_2^{-1} = T_{3,4}^{-1}. \quad (12)$$

При цьому обернена фактор-матриця  $T_4^{-1}$  розміру  $4 \times 4$  містить у трьох рядках по два ненульові елементи  $\pm 1$ , а обернена фактор-матриця  $T_{3,4}^{-1}$  розміру  $3 \times 4$  містить ненульові елементи  $a, b, \pm c, \pm d$  і може бути представлена як добуток двох матриць:

$$T_{3,4}^{-1} = M_{3,4}^{-1} R_4, \quad (13)$$

де  $R_4$  — діагональна матриця (7),  $M_{3,4}^{-1}$  — обернена фактор-матриця  $3 \times 4$ , яка містить в кожному рядку по три ненульових елементи  $\pm 1$ .

Отже, алгоритм швидкого виконання 4-точкового оберненого ЦСП-VII згідно з (10)–(13), як і алгоритм швидкого виконання прямого ЦСП-VII, потребує 5 операцій множення і 11 операцій додавання.

Матриця  $S_4^{\text{VII}}$  ядра ЦСП-VII порядку 4 має вигляд:

$$S_4^{\text{VII}} = \begin{bmatrix} a & b & c & d \\ d & c & 0 & -c \\ d & -a & -c & b \\ b & -d & c & -a \end{bmatrix}, \quad (14)$$

де  $a + b = d$ .

Елементи матриці  $S_4^{\text{VII}}$  однонормового масштабованого ЦСП-VII-1 із роботи [14] набувають таких значень:  $a = 14, b = 28, c = 37, d = 42$ . У стандарті Н.265 елементи матриці  $S_4^{\text{VII}}$  набувають таких значень:  $a = 29, b = 55, c = 74, d = 84$ . У [15] наведено цілочислову апроксимацію із заміною елементів матриці  $S_4^{\text{VII}}$ :  $a = 28, b = 56$  для зменшення операцій множення в алгоритмі швидкого виконання 4-точкового ЦСП-VII.

Матриця  $S_4^{\text{VII}-1}$  запропонованого однонормового масштабованого ЦСП-1 має вигляд

$$S_4^{\text{VII}-1} = \begin{bmatrix} 14 & 28 & 37 & 42 \\ 37 & 37 & 0 & -37 \\ 42 & -14 & -37 & 28 \\ 28 & -42 & 37 & -14 \end{bmatrix}. \quad (15)$$

Базисні вектори матриці  $S_4^{\text{VII}-1}$  масштабованого ЦСП-1 мають  $L_2$ -норму, яка наближається до числа степеня два:  $\|S_i\|^2 = 4096 + \Delta_i$ , де  $\|S_i\|^2$  —  $L_2$ -норма  $i$ -ї базисної функції матриці перетворення (яка прямує до одиниці для перетворення з ортонормованим базисом),  $\Delta_i$  (%) — відхилення  $L_2$ -норми  $i$ -ї базисної функції, яке в результаті проведеного аналізу становить  $\Delta_i = 0,42\%$ ,  $i = 0, 2, 3$ ,  $\Delta_1 = 0,27\%$ , при цьому неортогональність становить  $0,07\%$ . У роботі [14] запропоновано інше однонормове масштабоване ЦСП-2, елементи матриці якого представлено шістьма бітами, базисні вектори мають однакову  $L_2$ -норму, представлену трьохрозрядним числом, а неортогональність становить  $0,5\%$ .

Отже, запропоновані ЦСП-1,2 є найближчими до ДСП-VII і мають такі самі властивості, що й ЦСП-VII, прийняте у стандарті Н.265: коефіцієнти масштабованих перетворень 1 і 2 є достатньо близькими до ДСП-VII; неортогональність базисних векторів у результаті проведеного аналізу становить  $0,07\%$  і  $0,5\%$  (майже ортогональні); така сама схема квантування і деквантування, де множники залежать від значень параметру квантування QR [16].

Обчислювальну складність запропонованих алгоритмів ЦСП-1, ЦСП-2, відомого алгоритму ЦСП [8], прийнятого у стандарті Н.265, відомого алгоритму ЦСП [15] і відомого алгоритму ДСП [11] 1D 4-точкових перетворень наведено у табл. 1.

Алгоритм запропонованого ЦСП-1 потребує на  $37,5\%$  менше операцій множення порівняно з відомим алгоритмом [8], а алгоритм запропонованого ЦСП-2 потребує в чотири рази менше операцій множення і на  $60\%$  більше операцій додавання (або без операцій множення, але у два рази більше операцій додавання).

**Таблиця 1**

Операції	Оцінка обчислювальної складності 1D 4-точкових зображень				
	запропонованих		відомих		
	ЦСП-1	ЦСП-2	ЦСП (Н.265) [8]	ЦСП [15]	ДСП [11]
Множення	5	2 (0)*	8	5	5
Додавання + зсув	11	13+3 (15+6)*	10	11+3	11

\* Справедливо для оберненого перетворення.

**ЕФЕКТИВНІСТЬ КОДУВАННЯ НА ОСНОВІ ПЕРЕТВОРЕННЯ ДЛЯ 1D МОДЕЛЕЙ**

Важливим критерієм (мірою) оцінки характеристики стиснення є ефективність кодування  $G_{TC}$  на основі перетворення. У випадку оптимального розподілу бітового представлення коефіцієнтів перетворення у процесі квантування цих коефіцієнтів перетворення за характеристикою  $G_{TC}$  на основі перетворення, яка визначається співвідношенням середнього арифметичного і геометричного дисперсій коефіцієнтів перетворення  $\sigma_{xi}^2, 0 \leq i \leq N$ . Оцінка  $G_{TC}$  визначається як [17]

$$G_{TC} = 10 \log_{10} \frac{\frac{1}{N} \sum_{i=0}^{N-1} \sigma_{xi}^2}{\left( \prod_{i=0}^{N-1} \sigma_{xi}^2 \|f_i\|^2 \right)^{1/N}}, \text{ дБ,}$$

де  $\sigma_{xi}^2$  — дисперсія  $i$ -го коефіцієнта перетворення,  $\|f_i\|^2$  —  $L_2$ -норма  $i$ -ї базисної функції матриці перетворення (яка прямує до одиниці для перетворення з ортонормованим базисом).

**ЕФЕКТИВНІСТЬ КОДУВАННЯ НА ОСНОВІ ПЕРЕТВОРЕННЯ ДЛЯ 2D МОДЕЛЕЙ**

Для адекватного представлення блоку перетворення розміру  $N \times N$  застосовуються такі 2D кореляційні моделі [18]: проста сепарабельна двовимірна модель, проста ізотропна двовимірна модель, модифікована або узагальнена ізотропна несепарабельна 2D кореляційна модель [18, 19], несепарабельна кореляційна модель 2D ізотропного марковського процесу [17], сепарабельна 2D кореляційна модель [19] та оптимізована еліптична 2D модель [18] або узагальнена 2D кореляційна модель [19], яка вперше була застосована для ансамблю оцифрованих картин (зображень) [20].

Для оцінки 2D  $G_{2TC}$  (2D моделі), визначеної в [19], є інша міра, що визначається як [17]

$$2D G_{2TC} = 10 \log_{10} \frac{\frac{1}{N^2} \sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \sigma_R^2(i, j)}{\left( \prod_{i=0}^{N-1} \prod_{j=0}^{N-1} \sigma_R^2(i, j) \|f_i\|^2 \|f_j\|^2 \right)^{1/N^2}}, \text{ дБ,}$$

де  $\sigma_R^2(i, j)$  — дисперсія  $(i, j)$ -х елементів матриці  $V$ , які можна обчислити згідно з [19, 21],  $\|f_i\|^2, \|f_j\|^2$  —  $L_2$ -норми  $i$ -х,  $j$ -х базисних функцій матриці перетворення (що прямують до одиниці для ортонормованого базису). Величина  $G_{2TC}$  збільшується для більших розмірів блоків або для більших

Таблиця 2

Тестові зображення*	Для тестового зображення		Для блока 4 × 4	
	$\rho_v$	$\rho_h$	$\rho'_v$	$\rho'_h$
DSC00373 2048 × 2048	0,985	0,984	0,685	0,701
p9160541 1536 × 1536	0,952	0,954	0,619	0,668
p9160554 1024 × 1024	0,922	0,948	0,622	0,656
24259 800 × 800	0,928	0,939	0,336	0,427
Lev 256 × 256	0,845	0,814	0,566	0,472
Середнє значення			0,565	0,585
min	0,845	0,814	0,336	0,427
max	0,985	0,984	0,685	0,701

\* <http://people.csail.mit.edu/torr/alba/images/pictures.tar.gz>.

значень коефіцієнта кореляції  $\rho$  між сусідніми елементами. Загалом зі збільшенням розміру блоку збільшується як ступінь стиснення, так і обчислювальна складність. Найчастіше використовуються блоки розміру 8×8 і 16×16.

Коефіцієнти горизонтальної кореляції  $\rho_h$  і вертикальної кореляції  $\rho_v$  тестових зображень, середні значення  $\rho'_v$ ,  $\rho'_h$ , а також їхні мінімальні та максимальні значення для блоків 4×4 наведено у табл. 2.

Результати тестування за характеристикою 2D  $G_{2TC}$  для кореляційної моделі 2D ізотропного марковського процесу з кореляційною функцією  $R(m, n) = \rho^{\sqrt{m^2+n^2}}$  з нульовим математичним очікуванням і одиничною дисперсією для середніх значень коефіцієнтів кореляції  $\rho'_v$ ,  $\rho'_h = 0,4-0,75$  для відомого перетворення ЦСП (H.265) [8] і запропонованих ЦСП-1, ЦСП-2 для блоків 4×4 наведено у табл. 3.

Таблиця 3

Перетворення для блоків 4 × 4	Оцінка ефективності кодування за характеристикою 2D $G_{2TC}$ (дБ) для коефіцієнтів кореляції							
	0,4	0,45	0,5	0,55	0,6	0,65	0,7	0,75
ЦСП-VII (H.265) [8]	3,654	4,098	4,588	5,135	5,751	6,456	7,284	8,306
Запропоноване ЦСП-VII-1	3,811	4,274	4,786	5,355	6,000	6,715	7,551	8,556
Запропоноване ЦСП-VII-2	2,747	3,109	3,517	3,982	4,518	5,150	5,920	6,918

#### ДВОВИМІРНІ МОДОЗАЛЕЖНІ РОЗДІЛЬНІ НАПРАВЛЕНІ АДАПТИВНІ ПЕРЕТВОРЕННЯ ДЛЯ INTRA-ПРОГНОЗУВАННЯ В КОДУВАННІ ЗОБРАЖЕНЬ/ВІДЕО

Прогнозований блок P у моді intra формується на основі раніше закодованих і реконструйованих блоків. Він віднімається з поточного блоку перед кодуванням. Блок P будується для кожного блоку 4×4 і 8×8 або цілого макроблоку 16×16 семплів яскравості. У стандарті H.264/AVC [16, 22] усього є дев'ять прогнозованих мод для прогнозу блоків 4×4 і 8×8 і чотири прогнозовані моди для блоку 16×16. Зазвичай кодер вибирає моду прогнозу, намагаючись мінімізувати різницю між кодованим потоковим блоком і блоком P.

Для intra-прогнозування різниць (лишків) у [5] було вперше введено модозалежне направлене перетворення (MDDT). Це перетворення адаптивно змінюється залежно від моди, але виникає потреба у додатковій інформації, яка переважно є важливою для таких малих блоків, як 4×4 та 8×8. MDDT ґрунтується на ПКЛ, яке не є роздільним перетворенням і є надто вартісним з погляду потрібного обсягу пам'яті та обчислювальної складності.

Нехай  $X$  представляє матрицю  $N \times N$  блоку пікселів, тоді 2D роздільне направлене перетворення визначається як [5]

$$Y = C_i X R_i,$$

де  $C_i, R_i$  — матриці перетворення стовпців і рядків для intra-прогнозування моди  $i$  відповідно,  $Y$  — матриця  $N \times N$  коефіцієнтів 2D перетворення. Ще одна мода кодування intra, яка позначається I-PCM, дає змогу кодеру передавати безпосередньо значення семплів зображення (без прогнозу і перетворення). В деяких особливих випадках (наприклад, коли є аномальне зображення та/або задані малі параметри квантувача) ця мода може виявитися більш ефективною порівняно із «звичайним» процесом кодування intra, який складається із прогнозу, перетворення, квантування та ентропійного кодера. Але схема MDDT потребує зберігання 18 матриць перетворення. У [23] введено дві альтернативи. Перша має назву «діагональне модозалежне направлене перетворення» (OMDDT), в якому просто припускається, що  $C_i = R_i$ , і в такий спосіб обчислюється ПКЛ з використанням цього обмеження. У цьому випадку потрібно зберігати тільки дев'ять матриць перетворення замість 18. Друга альтернатива має назву «модозалежне ДКП» (MDDCT). Припустимо, що кожне ПКЛ є просто ДКП, тоді MDDCT отримують шляхом масштабування  $\pm 1$ :

$$Y = Q_i C X C^T Q_i,$$

де  $Q_i$  — діагональна матриця з елементами  $\pm 1$ ,  $C$  — матриця ДКП.

Автори [8] розширили подібні ідеї шляхом застосування комбінації ДСП і ДКП для інших прогнозованих мод в стандарті H.264/AVC і показали, що при цьому мають місце лише менші втрати виконання порівняно з MDDT. Модозалежне ДКП/ДСП було вперше описано в [10]. У [24–26] наведено схему ДКП/ДСП, яка була застосована тільки для блоків яскравості  $4 \times 4$  і  $8 \times 8$  та містить 34 прогнозовані моди. Цілочислову апроксимацію ДСП типу VII порядку 4 представлено у [26] для схеми ДКП/ДСП для блоків яскравості  $4 \times 4$ , прийнятої у стандарті H.265. Схему ДКП/ДСП для блоків кольоровості  $4 \times 4$  було вперше представлено у [27] і також було прийнято у стандарт H.265. У [28] представлено схему ДКП/ДСП, яка ґрунтується на традиційному ДКП чи ДСП типу VII, що застосовуються для відеокодування всіх intra-прогнозувальних мод: вертикальних, горизонтальних та косих. Експериментальні результати показали, що схема ДКП/ДСП забезпечує значне покращення швидкості передачі бітів (бітрейту) порівняно з традиційною базовою схемою ДКП для intra-прогнозування у відеопослідовностях. Для п'яти класів A, B, C, D і E відеопослідовностей середній вигреш бітрейту становить 0,7 %, 1,1 %, 1,0 % та 1,0 % для блоків  $4 \times 4$  з компонентами кольоровості для режимів intra високої ефективності, intra низької складності, випадкової вибірки високої ефективності та випадкової вибірки низької складності налаштувань.

Обчислювальна складність запропонованих роздільних направлених адаптивних перетворень цілочислового косинусного перетворення (ЦКП) [29]+ЦСП-1,2, як відомого [28], так і прийнятих у стандарті H.265 ЦКП [30]+ЦСП [8] 2D 4-точкових цілочислових обернених перетворень наведена у табл. 4.

Алгоритм запропонованих роздільних ЦКП+ЦСП-2 2D 4-точкових цілочислових обернених перетворень потребує у сім разів менше операцій множення і на 2,9 % більше операцій додавання, що в цілому зменшує його обчислювальну складність у 6,97 рази порівняно з відомим алгоритмом [28], або не потребує 56 операцій множення, але при цьому потребує на 10 % більше операцій додавання. Порівняно з роздільними перетвореннями зі стандарту H.265 запропоновані пе-



**Таблиця 4**

Операції	Оцінка обчислювальної складності 2D роздільних обернених перетворень для блоків $4 \times 4$				Порівняльний аналіз ЦКП+ЦСП-1 відносно	
	запропонованих ЦКП+ЦСП-1,2		відомих ЦКП+ЦСП		H.265 [30]+[8]	ЦКП+ЦСП [28]
	ЦСП-1	ЦСП-2	H.265[30]+[8]	із [28]		
Множення	20	8(0)	56	56	У 2,8 рази менше	У 2,8 рази менше
Додавання+зсув	100+24	108+36 (116+48)	72+16	108+32	На 40,9 % більше додавань	На 11,4 % менше додавань
Загальне зменшення					У 2,39 рази менше	У 2,91 рази менше

ретворення мають у 6,36 рази меншу обчислювальну складність. У випадку виконання запропонованих роздільних перетворень без операцій множення кількість операцій додавання збільшується на 13,9%.

#### ЕКСПЕРИМЕНТАЛЬНІ РЕЗУЛЬТАТИ

На рис. 1 представлено вихідні зображення класів В і С для тестування.

У табл. 5 представлено експериментальні результати ефективності кодування за характеристикою стандартної кількісної оцінки спотворення PSNR (дБ) для стиснення тестових зображень класу В (рис.1, *а*, *б*) з роздільною здатністю  $1920 \times 1072$  пікселів і класу С (рис. 1, *в*, *г*) з роздільною здатністю  $1280 \times 768$  пікселів для нормального (22–37) діапазону QR 2D роздільного направлено ЦКП/ЦСП для блоків  $4 \times 4$ . Ці результати представляють різницю на основі запропонованого роздільного направлено ЦКП/ЦСП-1 і на основі відомих роздільних направлених ЦКП (H.265) [30] /ЦСП [8].

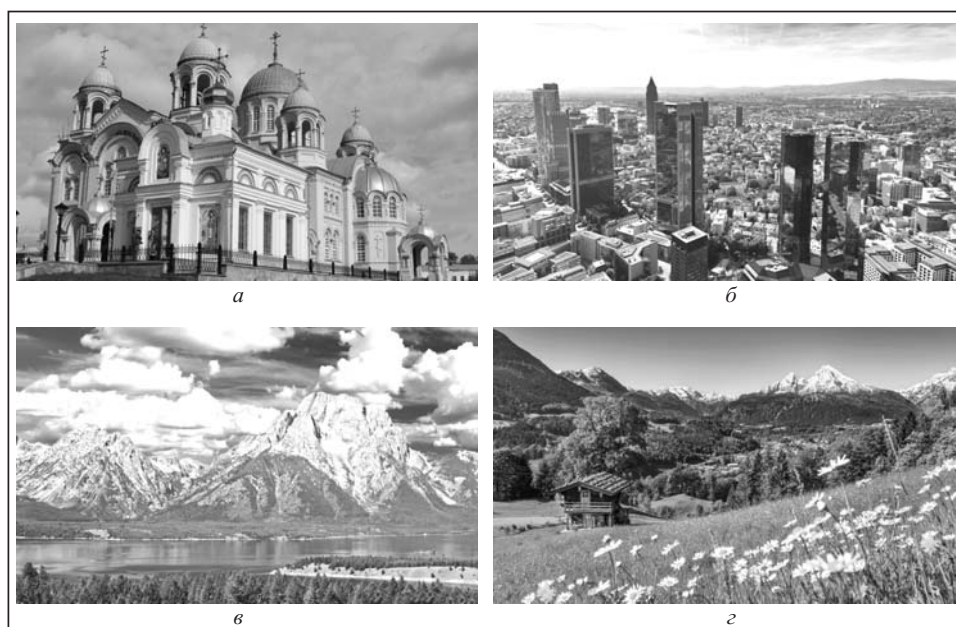


Рис. 1. Вихідні зображення класів В і С: храм, розмір  $1920 \times 1072$  (*а*); місто, розмір  $1920 \times 1072$  (*б*); гори, розмір  $1280 \times 768$  (*в*); пейзаж, розмір  $1280 \times 768$  (*г*)

Таблиця 5

Клас	Зображення з блоками 4 × 4	Результати ефективності кодування за характеристикою PSNR (дБ) для QP (діапазони)			
		22	27	32	37
В 1920 × 1072	Храм (СК)	0,10	-0,17	0,27	-0,12
	Місто (НК)	0,03	-0,02	0,04	-0,13
С 1280 × 768	Гори (СК)	0,06	0,06	0,08	-0,09
	Пейзаж (НК)	0,10	0,04	0,03	-0,01
Середнє значення		0,072	-0,020	0,103	-0,087

Таблиця 6

Клас	Зображення з блоками 4 × 4	Результати кодування за коефіцієнтом стиснення K : 1 для QP (діапазони)			
		22	27	32	37
В 1920 × 1072	Храм (СК)	2,37	2,76	3,16	3,63
	Місто (НК)	1,87	2,18	2,61	3,33
С 1280 × 768	Гори (СК)	1,84	2,15	2,58	3,29
	Пейзаж (НК)	1,48	1,73	2,13	2,91
Середнє значення		1,889	2,205	2,621	3,291

Таблиця 7

Клас	Зображення з блоками 4 × 4	Результати ефективності кодування за коефіцієнтом стиснення K (у %) для QP (діапазони)			
		22	27	32	37
В 1920 × 1072	Храм (СК)	-0,11	-0,12	-0,17	-0,26
	Місто (НК)	-0,57	-0,41	-0,31	-0,33
С 1280 × 768	Гори (СК)	-0,40	-0,25	-0,25	-0,35
	Пейзаж (НК)	-0,19	-0,32	-0,40	-0,44
Середнє значення		-0,318	-0,275	-0,282	-0,345

Експериментальні результати кодування за коефіцієнтом стиснення  $K : 1$  для запропонованого 2D роздільного направленої ЦКП/ЦСП-1 для блоків  $4 \times 4$  наведено у табл. 6. Результати ефективності кодування за коефіцієнтом стиснення  $K$  (у %), що представляють різницю на основі запропонованого роздільного направленої ЦКП/ЦСП-1 і на основі відомих роздільних направлених ЦКП (Н.265)/ЦСП, наведено у табл. 7.

Експериментальні результати ефективності кодування за характеристикою PSNR, які представляють різницю на основі запропонованого роздільного направленої ЦКП/ЦСП-2 і на основі відомих роздільних ЦКП (Н.265)/ЦСП, наведено у табл. 8.

Експериментальні результати кодування за коефіцієнтом стиснення  $K : 1$  для запропонованого 2D роздільного направленої ЦКП/ЦСП-2 для блоків  $4 \times 4$  наведено у табл. 9. Результати ефективності кодування за коефіцієнтом стиснення  $K$  (у %), що представляють різницю на основі запропонованого роздільного направленої ЦКП/ЦСП-2 і на основі відомих роздільних направлених ЦКП (Н.265)/ЦСП, наведено у табл. 10.

У таблицях наведено середні значення експериментальних результатів ефективності кодування за характеристикою PSNR і коефіцієнтом стиснення  $K$  для

Таблиця 8

Клас	Зображення з блоками 4 × 4	Результати ефективності кодування за характеристикою PSNR (дБ) для QP (діапазони)			
		22	27	32	37
В 1920 × 1072	Храм (СК)	-2,46	-0,97	-0,40	-0,03
	Місто (НК)	-1,49	-0,48	0,00	0,15
С 1280 × 768	Гори (СК)	-1,58	-0,46	-0,02	0,22
	Пейзаж (НК)	-1,27	-0,43	-0,14	0,06
Середнє значення		-1,700	-0,583	-0,139	0,102

Таблиця 9

Клас	Зображення з блоками 4 × 4	Результати кодування за коефіцієнтом стиснення $K$ ( $K:1$ ) для QP (діапазони)			
		22	27	32	37
В 1920 × 1072	Храм (СК)	2,69	3,25	3,84	4,56
	Місто (НК)	2,03	2,41	2,95	3,89
С 1280 × 768	Гори (СК)	1,99	2,38	2,96	3,98
	Пейзаж (НК)	1,54	1,83	2,33	3,35
Середнє значення		2,06	2,47	3,02	3,95

Таблиця 10

Клас	Зображення з блоками 4 × 4	Результати ефективності кодування за коефіцієнтом стиснення $K$ ( $\gamma$ %) для QP (діапазони)			
		22	27	32	37
В 1920 × 1072	Храм (СК)	4,94	5,32	5,46	5,32
	Місто (НК)	3,69	3,90	4,11	4,05
С 1280 × 768	Гори (СК)	3,83	4,24	4,72	4,89
	Пейзаж (НК)	2,39	2,95	3,63	4,09
Середнє значення		3,711	4,103	4,483	4,588

чотирьох тестових зображень класів В і С, де НК, СК — низько- та середньокорельовані зображення. Під час тестування використовувалося роздільне направле-не перетворення ЦКП/ЦСП без intra-прогнозування, при цьому ЦКП застосо-вувалося для перетворення стовпців, а ЦСП — для перетворення рядків з метою порівняння за двома характеристиками ефективності кодування (за якістю і сту-пенем стиснення) запропонованих синусних перетворень з відомими.

Запропоноване роздільне направле-не ЦКП/ЦСП-1 для блоку 4 × 4 порівняно з роздільними направленими перетвореннями стандарту H.265 за характе-ри-стикою спотворення PSNR для чотирьох тестових зображень класів В і С знижує се-реднє значення на 0,02–0,87 дБ, а для QP = 32 підвищує середнє значення на 0,103 дБ. При цьому середнє значення коефіцієнта стиснення  $K$  зменшується на 0,275–0,345 %. Запропоноване роздільне направле-не ЦКП/ЦСП-2 порівняно з

роздільними направленими перетвореннями стандарту H.265 за характеристикою спотворення PSNR для чотирьох тестових зображень класів В і С знижує середнє значення на 1,7–0,139 дБ для нормального QP = 22, 32, а для високого QP = 37 підвищує на 0,102 дБ. При цьому середнє значення коефіцієнта стиснення  $K$  збільшується на 3,71–4,59 %.

#### ВИСНОВКИ

Запропоновано алгоритм швидкого виконання 4-точкового ЦСП-VII, який потребує п'ять операцій множення, що на 37,5 % менше порівняно з відомим алгоритмом. Запропоновано два цілочислові синусні перетворення типу VII порядку 4 і розроблено алгоритми швидкого виконання цих перетворень, які потребують п'ять і дві операції множення, або без операцій множення, але в два рази більше операцій додавання. Розроблено алгоритми швидкого виконання 2D 4-точкових роздільних направлених цілочислових косинусного і синусних адаптивних перетворень для intra-прогнозування низької обчислювальної складності, яка у 2,91 рази і в 6,97 рази менша, ніж у відомому алгоритмі [28]. Ці адаптивні цілочислові перетворення найбільш близькі до ДКП/ДСП-VII і мають високі характеристики ефективності кодування за якістю та середні значення коефіцієнтів стиснення, які є більшими на 3,71–4,59%, ніж для перетворення зі стандарту H.265. Отже, обидва розроблені цілочислові синусні перетворення 1 і 2 типу VII порядку 4 можуть бути використані для покращення стандарту H.265 з метою збільшення швидкодії, ступеню стиснення (що призводить до покращення бітрейту) та зменшення обчислювальних та енергетичних витрат.

#### СПИСОК ЛІТЕРАТУРИ

1. Rao K.R., Yip P. Discrete cosine transform: algorithms, advantages, applications. Boston: Academic Press, 1990. 490 p.
2. Jain A.K. A fast Karhunen–Loeve transform for a class of stochastic process. *IEEE Trans. on Commun.* 1976. Vol. COM-24, N 9. P. 1023–1029.
3. Jain A.K. Image coding via nearest neighbors image model. *IEEE Trans. on Commun.* 1975. Vol. 23, N 3. P. 318–321.
4. Meiri A.Z., Yudilevich E.A. A pinned sine transform image coder. *IEEE Trans. on Commun.* 1981. Vol. 29, N 12, P. 1728–1735.
5. Ye Y., Karczewicz M. Improved Intra Coding. Doc. VCEG-AG11; ITU-T Q.6/SG11: Shenzhen, China, 2007.
6. McCann K., Bross B., Sekiguchi S., Han W.-J. HM4: High Efficiency Video Coding (HEVC) Test Model 4 Encoder Description. ITU-T doc. JCTVC-F802, Torino, IT, July, 2011.
7. 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.
8. Yeo C., Tan Y.H., Li Z. and Rahardja S. Mode-dependent fast separable KLT for block-based intra coding. ITU-T, doc. ICTVC-B024, Geneva, CH, July 2010.
9. Yueh W.C. Eigenvalues of several tridiagonal matrices. *Appl. Mathematics E-Notes*. 2005. Vol. 5. P. 66–74.
10. Han J., Saxena A., Rose K. Towards jointly optimal spatial prediction and adaptive transform in video/image coding. Proc. IEEE Int. Conf. Acoust., Speech, Signal Process (ICASSP). (March 14–19 2010, Dallas, TX, USA). Dallas, 2010. P. 726–729.
11. Chivukula R.K., Reznik Y.A. Fast computing of discrete cosine and sine transforms of types VI and VII. *Proc. SPIE 8135, Applications of Digital Image Processing XXXIV*, 813505 (19 September 2011). <https://doi.org/10.1117/12.903685>.

12. An J., Zhao X., Guo X., Lei S. Non-CE7: Boundary-dependent transform for inter-predicted residue. Resreport, Joint Collaborative Team on Video Coding (JCT-VC) of ITU-t SG16 WP3 and ISO IEC JTC1 SC29 WG11 JCTVC G281. Geneva, CH, Nov. 2011.
13. Шевчук Б.М., Задірака В.К., Гнатів Л.О., Фраєр С.В. Технологія багатофункціональної обробки і передачі інформації в моніторингових мережах. Київ: Наук. думка, 2010. 378 с.
14. Гнатів Л.О., Луц В.К. Метод побудови моде-залежного швидкого роздільного цілочисельного ПКЛ для адаптивного кодування зображень і відео. *Пр. міжн. наук. конф. «Питання оптимізації обчислень (ПОО-XL)»*. (вересень 2013, Україна, Крим, Велика Ялта, смт. Кацивелі). Кацивелі, 2013. С. 68–69.
15. Saxena A., Fernandes F. C. On fast implementation at 4-point DST type-7 with 5 multiplications. ITU-T, doc. JCTVC — F283, Torino, Italy, July 2011.
16. Ричардсон Я. В. Видеокодирование. H.264 і MPEG-4 — стандарти нового покоління. Москва: Техносфера, 2005. 368 с.
17. Гнатів Л.А. Целочисленные косинусные преобразования: методы построения новых быстрых преобразований порядка 8,16 и их применение. *Кибернетика и системный анализ*. 2014. Т. 50, № 6. С. 104–121.
18. Clarke R. J. Application of image covariance models to transform coding. *Int. J. Electronics*. 1984. Vol. 56, N 2. P. 245–260.
19. Akansu A.N., Haddad R.A. Factorization of the coefficient variance matrix in orthogonal transforms. *IEEE Trans. Signal Processing*. 1991. Vol. 39, N 3. P. 714–718.
20. Mauersberger W. Generalized correlation model for designing 2-dimensional image coders. *Electron. Lett.* 1979. Vol. 15, N 20. P. 664–665.
21. Гнатів Л.А., Шевчук Е.С. Методи синтезу ефективних ортогональних преобразований високої і низької кореляції і їх швидких алгоритмів для кодування і стиснення цифрових зображень. *Кибернетика и системный анализ*. 2002. Т. 38, № 6. С. 104–117.
22. ITU-T Rec. H.264|ISO/IEC 14496-10: 2009. Information technology — Coding of audio-visual objects-Part 10: Advanced Video Coding, 2009.
23. Budagavi M., Zhou M. Orthogonal MDDT and mode dependent DCT. ITU-T Q.6/SG16 VCEG, VCEG-AM20. Jan. 2010. Kyoto, Japan.
24. Saxena A., Fernandes F. Jointly optimal intra prediction and adaptive primary transform. ITU-T, doc. JCTVC-C108. October 2010. Guangzhou, China.
25. Saxena A., Fernandes F. CE 7: Mode-dependent DCT/DST for intra prediction in video coding. ITU-T, doc. ICTVC-D033. Jan. 2011. Daegu, Korea.
26. Saxena A., Fernandes F. CE7: Mode-dependent DCT/DST without  $4 \times 4$  full matrix multiplication for intra prediction. ITU-T, doc. JCTVC- E125. Mar. 2011. Geneva, CH.
27. Saxena A., Fernandes F., Alshina E., Chen J. Mode-dependent DCT/DST for  $4 \times 4$  chroma blocks. ITU-T, doc. JCTVC-F553. July 2011. Torino, Italy.
28. Saxena A., Fernandes F. C., Alshina E., Chen J. CE7: Mode-dependent DCT/DST for  $4 \times 4$  chroma blocks. ITU-T, doc. ICTVC-G107. Nov. 2011. Geneva, CH.
29. Гнатів Л.А. Целочисленные косинусные преобразования для высокоэффективного кодирования изображений и видео. *Кибернетика и системный анализ*. 2016. Т. 52, № 5. С. 161–176.
30. Fuldseth A., Bjøntegaard G., Sadafale M., Sze V. CE10: Core transform design for HEVC. ITU-T, doc. JCTVC-G495. Nov. 2011. Geneva, CH.

*Надійшла до редакції 01.04.2019*

**Л.А. Гнатив, В.К. Луц**

**АЛГОРИТМЫ БЫСТРОГО ВЫПОЛНЕНИЯ 4-ТОЧЕЧНЫХ ЦЕЛОЧИСЛЕННЫХ СИНУСНЫХ ПРЕОБРАЗОВАНИЙ ТИПА VII БЕЗ УМНОЖЕНИЯ И РАЗДЕЛЬНЫЕ НАПРАВЛЕННЫЕ АДАПТИВНЫЕ ПРЕОБРАЗОВАНИЯ ДЛЯ INTRA-ПРОГНОЗИРОВАНИЯ В КОДИРОВАНИИ ИЗОБРАЖЕНИЙ/ВИДЕО. I**

**Аннотация.** Предложены отдельные направленные адаптивные целочисленные косинусное и синусные типа VII преобразования с блоками  $4 \times 4$  для intra-прогнозирования в кодировании изображений и видео. Разработаны алгоритмы быстрого выполнения 2D 4-точечных отдельных направленных целочисленных косинусного и синусных преобразований. Они по сравнению с известными алгоритмами требуют в семь раз меньше операций умножения и обеспечивают более высокую степень сжатия. Предложены два целочисленных синусных преобразования типа VII порядка 4 и разработаны алгоритмы быстрого выполнения 4-точечных целочисленных синусных преобразований типа VII, которые требуют только целочисленных операций. Они по сравнению с известным алгоритмом позволяют существенно сократить число операций умножения или обойтись без них. Предложенные целочисленные синусные преобразования являются наиболее близкими к дискретному синусному преобразованию типа VII и имеют более высокую степень сжатия, чем известное синусное преобразование, принятое в стандарте H.265.

**Ключевые слова:** дискретное косинусное преобразование, дискретное синусное преобразование, целочисленное косинусное преобразование, целочисленное синусное преобразование, модозависимое направленное преобразование, отдельное направленное адаптивное преобразование, факторизация, быстрое выполнение преобразования, intra-прогнозирование, эффективность кодирования, видеокодирование, H.265.

**L.O. Hnativ, V.K. Luts**

**ALGORITHMS FOR FAST IMPLEMENTATION AT 4-POINT INTEGER SINE TRANSFORMS TYPE VII WITHOUT MULTIPLICATION AND SEPARABLE DIRECTIONAL ADAPTIVE TRANSFORMS FOR INTRA PREDICTION IN IMAGE/VIDEO CODING. I**

**Abstract.** The separable directional adaptive integer cosine and sine of type VII transforms with  $4 \times 4$  blocks are proposed for intra prediction in image and video coding. Algorithms for fast implementation of 2D 4-point separable directional integer cosine and sine transforms are developed. These algorithms requires 7 times less multiplication operations and provide higher compression ratio as compared with the well-known algorithms. Two order-4 integer sine type VII transforms are proposed and algorithms are developed for fast implementation of 4-point integer sine type VII transforms, which require integer operations only. As compared with the known algorithm, these algorithms can significantly reduce the number of multiplication operations or avoid them at all. These integer sine transforms are very close to discrete sine type VII transform and have higher compression ratio than the well-known sine transform accepted in standard H.265 has.

**Keywords:** discrete cosine transform, discrete sine transform, integer cosine transform, integer sine transform, mode-dependent directional transform, separate directional adaptive transform, factorization, fast transform, coding efficiency, intra predicted, video coding, H.264, H.265.

**Гнатів Лев Олексійович,**

кандидат техн. наук, старший науковий співробітник Інституту кібернетики ім. В.М. Глушкова НАН України, Київ, e-mail: levhnativ@gmail.com.

**Луц Василь Костянтинович,**

молодший науковий співробітник Інституту кібернетики ім. В.М. Глушкова НАН України, Київ, e-mail: vkluts@gmail.com.