

УДК 004.627

**Б. М. Шевчук**, канд. техн. наук,**В. К. Задірака**, д-р фіз.-мат. наук,**В. І. Бовсунівський**, аспірант

Інститут кібернетики ім. В. М. Глушкова НАН України, м. Київ

## ДОСЛІДЖЕННЯ МЕТОДІВ І АЛГОРИТМІВ СТИСНЕННЯ ТА ВІДНОВЛЕННЯ ВІДЕОСИГНАЛІВ ДЛЯ ПОБУДОВИ ІНТЕЛЕКТУАЛЬНИХ МЕРЕЖЕВИХ ВІДЕОСЕНСОРІВ

На основі дослідження тестових кадрів первинних відео мавив даних досліджені спотворення форми обвідних відеосигналів з використанням стандартних методів стиснення відеоданих, проаналізовані типи спотворень відновлених після стиснення відеоданих, наведені результати моделювання алгоритмів відновлення обвідних відеосигналів та запропоновані методи і алгоритми оперативної фільтрації, стиснення та відновлення відеоданих, оптимізованих за точністю і швидкодією.

**Ключові слова:** *стиснення та відновлення відеосигналів, кодування інформації, моделювання алгоритмів стиснення.*

Відеоінформація характеризується підвищеною інформативністю даних, особливо рухоме відео (відеодані). Тому важливою проблемою є створення ефективних комп'ютерних систем збору, обробки та передачі відеоданих, інтелектуальних відеосенсорів, орієнтованих на використання в різноманітних галузях людської діяльності. Особливо актуальною проблемою є створення компактних, автономних та інтелектуальних відеосистем і відеосенсорів з засобами зв'язку для доставки оперативних відеоданих на центральний сервер спеціалізованих комп'ютерних мереж.

Основою побудови таких засобів є сигнальні процесори (ARM 9, Blackfin, DaVinci та ін.) провідних мікроелектронних та комп'ютерних фірм світу, доповнених модулями радіоприйомопередавачів відповідних технологій зв'язку. Методологічну та алгоритмічну основу побудови ефективних відео систем та відеосенсорів утворюють методи та алгоритми оперативної обробки і кодування відеоданих. При цьому велике значення при розробці програмно-апаратних засобів відеосистем мають алгоритми обробки і кодування відеоданих, від якості яких в значній степені залежать характеристики всієї відеосистеми [1; 2].

Метою статті є обґрунтування доцільності реалізації ефективних адаптивних методів і алгоритмів стиснення відеоданих для побудови інтелектуальних мережесенсорів на основі результатів дослідження та моделювання алгоритмів стиснення і відновлення відеосигналів.

Вихідні потоки сучасних відеосенсорів, наприклад *КМОН*-відеосенсорів (*КМОН*-комплементарний метало-оксидний напівпровідник), суттєво залежать від формату відеокадру, прийнятої схеми кольорового відеокодування, топології побудови світлофільтрів відеосенсора та формату вихідних даних.

У випадку використання відеосенсорів з *RGBG*-топологією розміщення піксельних світлофільтрів матриці зображення розміром  $M \times N$  ( $M$  — кількість пікселів в рядку,  $N$  — кількість пікселів в стовпчику (кількість рядків)), структура відліків відеосенсора має вигляд: перший рядок (непарний рядок) — *GRGRGR ... GRGR*, другий (парний рядок) — *BGBGBG ... BGBG*, де  $G$ ,  $R$  і  $B$  — коди зеленого, червоного і синього відліків відповідно. Вихідні відеодані в таких сенсорах передаються процесорним пристроям від кадру до кадру, при цьому для поточного кадру формуються службові керуючі сигнали кадру та відповідних рядків. Тому коди відповідних пікселів рядок за рядком по паралельній шині передаються на входи процесора.

Для якісного відображення відеоданих для кожного пікселя необхідно визначати три складові — *RGB*-відліки. Тому в залежності від вибраної схеми кодування відеоданих  $C_{КВД}$  процесорними засобами упорядковуються відліки  $R$ -,  $G$ - і  $B$ - сигналів або сигналу яскравості та кольорорізнцевих сигналів для виконання операцій фільтрації, стиску, захисту, накопичення і передачі даних. Сумарний потік первинних даних з відеосенсора визначається виразом  $V_{BC} = K \cdot M \cdot N \cdot q(C_{КВД})$ , де  $V_{BC}$  — об'єм даних в кбіт/с,  $K$  — кількість кадрів/с,  $M \cdot N$  — роздільна здатність відеосенсора,  $q(C_{КВД})$  — кількість біт, виділених для кодування кольорових сигналів в залежності від схеми кодування відеоданих.

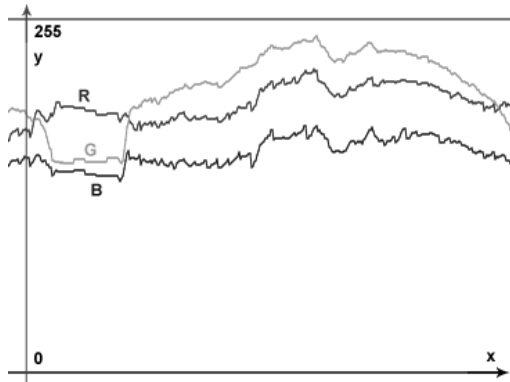
Серед останніх досягнень у сфері кодування відеоінформації на сьогодні можна виділити сім'ю відеокодеків H.264 [3]. Сам відеокодек було розроблено групою експертів по відео кодуванню (Video Coding Experts Group, VCEG) — робочою групою міжнародного союзу по телекомунікаціям (International Telecommunication Union, ITU-T). Групою VCEG було розроблено ряд візуальних телекомунікаційних стандартів. Кінцеві стадії розвитку стандарту H.264 були реалізовані об'єднанням (Join Video Team), яке складалося з експертів MPEG та VCEG, що зробили кінцеву версію стандарту в 2003 році. За результатами незалежних тестувань відеокодеків, опублікованих CS MSU Graphics&Media Lab в 2007 році, можна зробити висновок, що даний відеокодек випе-

реджає усі інші аналоги по швидкості обробки та якості передачі даних. Тому H.264 вважається лідером серед засобів стиску за рахунок високого коефіцієнту стиску і незначних втратах якості.

Ми вирішили розглянути ефективність роботи даного відеокордеку не в мультимедійних цілях. Відеосенсиори досить широко використовуються в науці та охоронних системах. В цих сферах основним критерієм якості є не ступінь стиску даних зі збереженням сприйнятності неозброєним оком людини, а висока чіткість та деталізація відео, водночас з можливістю оперативної передачі даних. Тому постає необхідність дослідити, наскільки якісно зберігаються найдрібніші деталі відеоданих, використовуючи різні алгоритми стиску. Взагалі усі методи кодування інформації за принципом реалізації можна поділити на дві основні групи. Це методи, що використовують сплайни, та евристичні методи. Методи, що базуються на використанні сплайнів, мають добре теоретичне підґрунтя, проте їх застосування рідко зустрічається на практиці. Ці методи ґрунтуються на інтерполяції локальними функціями. Евристичні методи, до яких саме і відноситься H.264, мають слабке теоретичне обґрунтування. Основна їх ідея базується на фізіологічній особливості людського сприйняття зображення. Алгоритми, що реалізують ці методи, дозволяють здійснювати стиснення зображення в десятки разів, завдяки чому отримали широке застосування на практиці. Проте такі методи можуть допускати незначні втрати даних, непомітні для ока людини [4].

В більшості випадків евристичні алгоритми розробляються для конкретних випадків, і для стиснення даних наперед визначеного типу. Спрямованість досліджень в окремих вузьких напрямках характеризується тим, що розроблені алгоритми стиснення, ефективні для одного виду даних, будуть зовсім не ефективні для іншого. Методи, що базуються на використанні сплайнів, менше залежать від структури зображення, до якого вони застосовується, тому такі методи є більш універсальними. Проте ситуація змінюється, якщо розглядати зображення як зміну певних величин в реальному часі або як сукупність слабкорельованих кривих, і застосовувати методи стиснення, ефективні саме для цього типу даних [5].

Процес кодування даних за допомогою методів, що використовують сплайни, передбачає обробку даних зображення чи відео, представлених у вигляді двохвимірної графіки. В такому графіку пікселі зображення розподіляються по осі абсцис, а значення кожного з елементів кольорів, в залежності від вибраної схеми представлення зображення, відображаються на осі ординат (рис. 1).



*Рис. 1. Представлення зображення у вигляді графіка*

В статті наведені результати дослідження того, наскільки якісно відображаються саме незначні деталі, сприйняті відеосенсором при невисоких коефіцієнтах стиску, а також змодельований алгоритм відновлення відеосигналів з використанням методів адаптивної фільтрації. Практично довільний відеосенсор зазвичай здійснює стиск зображення під час зйомки, при чому використовується стиск із втратами. Це пов'язано з метою економії носія, куди записується відеосигнал, якщо відеосенсор одразу передає сигнал для запису компактних відео-масивів даних на цифровий носій. При використанні стандартних інтерфейсів для передачі відеоданих, ні USB-, ні FireWire-інтерфейси не забезпечують такої швидкості передачі даних, щоб була передача повних кадрів без стиску інформації. Тому, зазвичай, такі відеосенсори, які не мають власного носія, а передають сигнал, самі здійснюють обробку відеосигналу.

Для отримання повного відео без стиску ми скористалися спеціалізованим відеосенсором, який дозволяє зберігати відеодані практично без втрат. В якості зразка відео була взята зйомка відеосенсора, встановленого на передній частині автомобіля. Відео містить запис руху автомобіля по проїжджій частині міста. Такі відеодані можуть бути корисні наприклад в аварійних ситуаціях. В таких випадках на відео будуть важливими найдрібніші деталі, втрата яких за рахунок стиску неприпустима.

На рис. 2 представлено кадр з відеосигналу контрольного зразку. Візьмемо цей кадр, як основу для порівняння з перекодованим відео. Здійснимо перекодування контрольного зразку відео в форматі H.264 з різним ступенем якості. В ролі показника якості візьмемо бітрейт відеоданих. Контрольний зразок має 221201 кбіт/с. При кодуванні ми отримали відео з низьким коефіцієнтом стиску  $K_{cm} \approx 27$  і швидкістю передачі даних 8153 кбіт/с (рис. 3), середнім коефіцієнтом стиску  $K_{cm} \approx 71$  і швидкістю передачі даних 3087 кбіт/с (рис. 4) та високим коефіцієнтом стиску  $K_{cm} \approx 142$  і швидкістю передачі даних

1548 кбіт/с (рис. 5), де коефіцієнт стиску  $K_{cm} = \frac{Ш_{КЗ}}{Ш_{ПЗ}}$  ( $Ш_{КЗ}$  — швидкість передачі даних контрольного зразку,  $Ш_{ПЗ}$  — швидкість передачі даних перекодованого зразку).

З усіх чотирьох відеоданих візьмемо конкретний кадр, і змодулюємо його подачу у вигляді дискретного *RGB* сигналу на спеціальне програмне забезпечення, призначене для аналізу таких даних. Ми отримаємо представлення кадру у вигляді графіку добре корельованих кривих червоного, зеленого і синього відліків.

На кадрі перекодованого відео з низьким коефіцієнтом стиску (рис. 3) білим кільцем виділено фрагмент, який ми порівняли з тією самою ділянкою контрольного зразку у вигляді накладки кривих червоного кольору (рис. 6). Як видно з порівняння двох кадрів, зображення практично не відрізняються одне від одного. Різницю можна лише помітити при безпосередній накладці кадрів одне на одне. Тоді будуть помічені спотворені ділянки відеосигналу. Як бачимо з порівняння кривих червоного кольору (рис. 6) відфільтрована крива в деяких місцях не повторює повністю контур контрольного сигналу. В деяких місцях при перегині крива урізається, таким чином описуючи значно меншу амплітуду стрибка, ніж на контрольному зразку. Ці спотворення не помітні на перший погляд неозброєним оком людини, проте в деяких випадках може бути втрачена корисна інформація.

На рис. 7 представлено порівняння кривої червоного кольору контрольного зразку з відео з середнім ступенем стиску (рис. 4). Неозброєним оком помітно відсутність незначних дефектів, які містилися на контрольному зразку. При порівнянні кривих виділеного фрагмента з кадром перекодованого відео з середнім ступенем стиску (рис. 4) з контрольним зразком, спостерігається значна відмінність кривої контрольного зразку від зразку перекодованого відео (рис. 7). Виділений фрагмент також значно відрізняється візуально. Основну тенденцію крива зберігає, проте вона стала дуже розгладженою, за рахунок чого проявляються видимі артефакти.

На рис. 8 представлено порівняння кривої червоного кольору контрольного зразку з відео з високим ступенем стиску на виділеному фрагменті (рис. 5). Вже при візуальному порівнянні контрольного зразку та перекодованого відео видно відсутність деяких дрібних предметів, що є неприпустимо на практиці в багатьох випадках, оскільки ці предмети можуть бути важливими елементами відео. На виділеному чорним кільцем фрагменті кадру перекодованого відео з високим ступенем стиску (рис. 5) знаходиться чисте небо, проте на цьому ж фрагменті контрольного зразку знаходиться електропровід на фоні неба. Як бачимо на рис. 8, крива, яка задавала червоний показник зображення електропроводника, практично повністю розгладжена, що свідчить про повну втрату інформації про цей предмет на даному фрагменті. На ін-

ших ділянках фрагмента, відфільтрована крива досить часто ігнорує поведінку кривої контрольного зразку, в деяких випадках взагалі відхиляється від основної тенденції контрольного зразку.

Отже, ми бачимо, що при збільшенні ступеню стиску відеоданих в форматі H.264, збільшується втрата дрібних поодиноких предметів. Деякі предмети стають розмитими. Водночас відеодані не набувають значної кількості артефактів, що і є одним з основних показників якості формату H.264. Проте «розмивання» дрібних деталей, яке дає цей формат при збільшенні ступеню стиску, в деяких випадках може бути неприпустимим.



*Рис. 2. Контрольний зразок (відео без стиску, 221201 кбіт/с)*



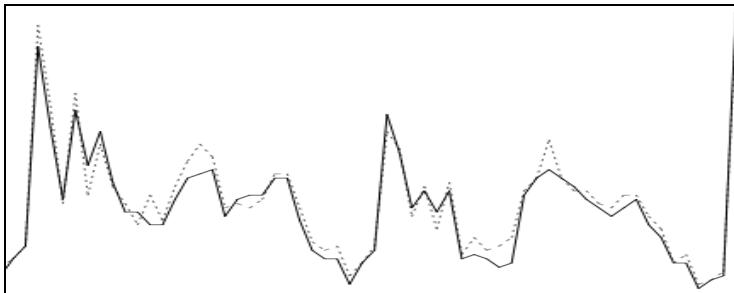
*Рис. 3. Кадр відео з малим коефіцієнтом стиску (8153 кбіт/с)*



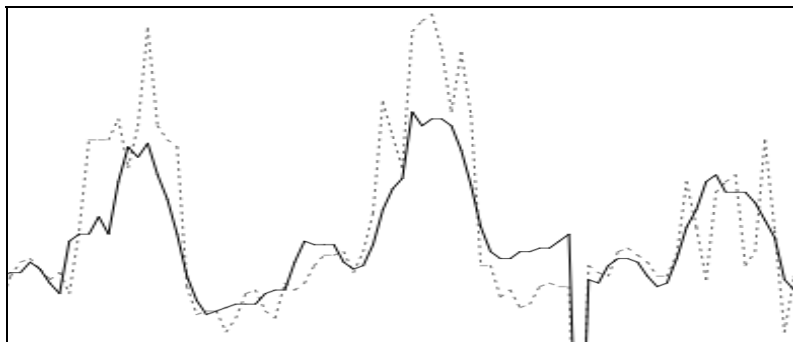
*Рис. 4. Кадр відео з середнім коефіцієнтом стиску (3087 кбіт/с)*



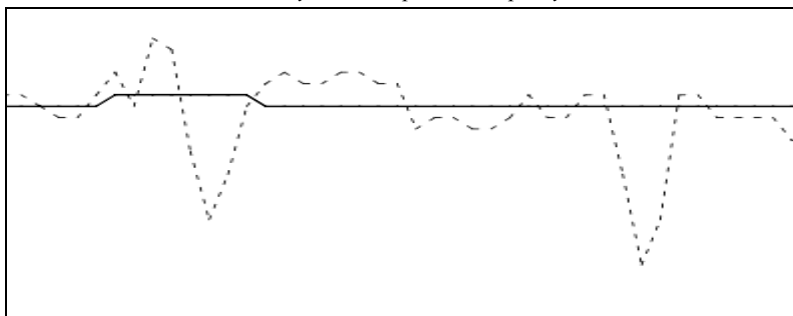
*Рис. 5. Кадр відео з високим коефіцієнтом стиску (1548 кбіт/с)*



*Рис. 6. Порівняння кривої червоного кольору на кадрі з низьким ступенем стиску та контрольної зразку*



*Рис. 7. Порівняння кривої червоного кольору на кадрі з середнім ступенем стиску та контрольного зразку*



*Рис. 8. Порівняння кривої червоного кольору на кадрі з високим ступенем стиску та контрольного зразку*

Представимо модель алгоритму адаптивної фільтрації та стиску даних, а також спосіб їх відновлення. Нехай ми маємо деяку систему, на виході якої ми отримуємо цифровий сигнал у вигляді дискретного потоку трійки байтів, які відповідають за певний відтінок і представляють один піксель. Нам не важливо, яким чином отримуються дані про зображення чи відео, зчитуванням файлу чи то конвертацією аналогового сигналу. Важливий лише сам факт отримання даних у дискретному вигляді. В процесі надходження такої інформації кожен з трьох компонент кольору описує певну криву (рис. 1). З метою проведення досліджень було розроблено програмне забезпечення для аналізу таких кривих, а також їх відображення у вигляді графіків.

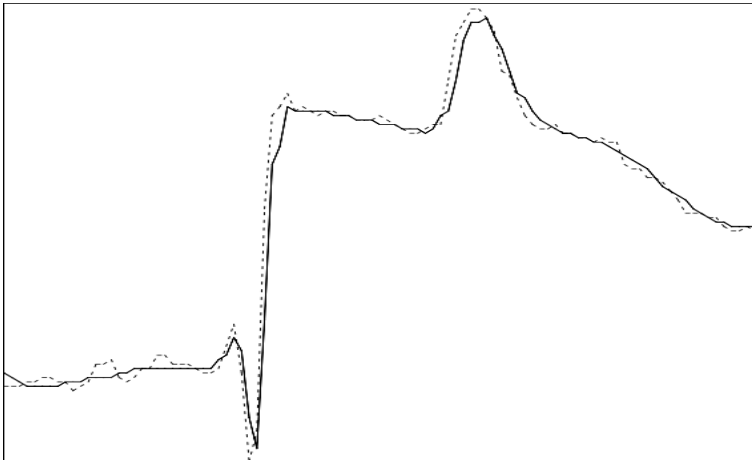
Проаналізувавши таку криву ми розділили її ділянки на 2 типи. Більш пологі ділянки, які містять незначні амплітудні відхилення ми віднесли до менш інформативних ділянок. Ділянки, які містять більші амплітуди кривої — до більш інформативних.



Використовуючи адаптивний метод фільтрації ми отримуємо гладку криву, позбавлену збиткових даних, корисні дані при цьому практично не змінюються (рис. 9). Основний принцип роботи цього методу фільтрації полягає в усередненні адаптивно підбраного вікна до певної ділянки даних, розмір якого обернено пропорційно залежить від загальної суми амплітудних величин сигналу у цьому вікні:

$$ШВ = F\left(\sum_{i=x_1}^{x_2} BB_i, KP\right), \text{ де } ШВ \text{ — ширина вікна для отримання ковзко-}$$

го середнього,  $KP$  — коефіцієнт регулювання залежності ширини вікна, який задається безпосередньо перед кодуванням,  $BB_i$  — величина відліку в точці  $i$ ,  $x_1$  та  $x_2$  — початок та кінець вікна відповідно.



**Рис. 9.** Результат адаптивної фільтрації кривої (пунктиром позначена оригінальна крива, суцільною — відфільтрована).

З метою компактного кодування  $R$ -,  $G$ - і  $B$ -сигналів доцільно на згладженій кривій відповідного сигналу оперативно визначати суттєві відліки ( $CB$ ), до яких відносяться екстремуми та точки перегину. Для прискореної обробки вихідних даних фільтрацію відліків сигналу необхідно виконувати на низько динамічних ділянках ковзким способом. За рахунок цього забезпечується мінімальне спотворення обвідної відеосигналу і при стисненні відліків сигналу амплітудно-часові характеристики екстремумів необхідно виявляти місцезнаходження змін тенденцій розвитку кривої сигналу шляхом аналізу змін знаку поточної різниці  $\Delta X_i^\phi = X_i^\phi - X_{i-1}^\phi$  між сусідніми відліками попереднього відфільтрованого сигналу, де  $X_i^\phi$  — поточний відлік відфільтро-

ваного сигналу (під зміною знаку величини  $\Delta X_i^\phi$  вважають зміну з «+» на «-», «-» на «+», з «+» на «0», з «-» на «0» і навпаки). Поточний відлік  $X_{i-1}^\phi$ , для якого величина  $\Delta X_i^\phi$  змінює свій знак, вважається екстремумом. Для оперативного визначення амплітудно-часових характеристик точок перегину доцільно визначати поточні різниці  $\Delta(\Delta X_i^\phi) = \Delta X_i^\phi - \Delta X_{i-1}^\phi$  між сусідніми приростами попередньо відфільтрованого сигналу та виявляти зміну знаку величини  $\Delta(\Delta X_i^\phi)$ : поточний відлік  $X_{i-1}^\phi$ , для якого величина  $\Delta(\Delta X_i^\phi)$  змінює свій знак, вважається точкою перегину.

Після аналізу та визначення амплітудно-часових характеристик  $CB$  на інтервалі даних вибірки сигналу, наприклад, поточного рядка або групи рядків матриці  $M \cdot N$ , здійснюється компактне кодування даних з урахуванням наявності загальної службової інформації, службової інформації і компактних даних суттєвих і несуттєвих відліків ( $NB$ ) відповідних рядків чи групи рядків поточного кадру.

Таким чином, компактну бітову послідовність поточних ібірок даних відеосигналу можна записати в такому вигляді:

$$\begin{aligned} & \{ZCI\} \left[ \left\{ CIP(GP)_{11} \right\} \left\{ KDP(GP)_{11} \right\} \left\{ CIP(GP)_{12} \right\} \left\{ KDP(GP)_{12} \right\} \dots \right. \\ & \quad \left. \dots \left\{ CIP(GP)_{1i} \right\} \left\{ KDP(GP)_{1i} \right\} \dots \left\{ CIP(GP)_{1M} \right\} \left\{ KDP(GP)_{1M} \right\} \right] \\ & \quad \left[ \left\{ CIP(GP)_{21} \right\} \left\{ KDP(GP)_{21} \right\} \dots \left\{ CIP(GP)_{2i} \right\} \left\{ KDP(GP)_{2i} \right\} \dots \right. \\ & \quad \left. \dots \left\{ CIP(GP)_{2M} \right\} \left\{ KDP(GP)_{2M} \right\} \right] \dots \left[ \left\{ CIP(GP)_{N1} \right\} \left\{ KDP(GP)_{N1} \right\} \dots \right. \\ & \quad \left. \dots \left\{ CIP(GP)_{NM} \right\} \left\{ KDP(GP)_{NM} \right\} \right], \end{aligned}$$

де  $ZCI$  — загальна службова інформація;  $CIP(GP)$  — службова інформація рядку (групи рядків);  $KDP(GP)$  — компактні дані рядку (групи рядків).

Для відновлення кривої переважно використовують сплайнами, які складаються з відрізків прямих та кубічних парабол. Значення параметрів можна однозначно визначити із системи лінійних рівнянь. Однак при використанні такої методики виникає ряд проблем. Сплайн визначається на обмеженому сегменті, і лише задавши всю множину вузлових точок, можна отримати значення сплайна в довільній точці.

Для побудови локальних сплайнів не потрібно розв'язувати систему лінійних рівнянь, оскільки в задоволенні умов стикування фраг-

ментів задіяна обмежена кількість найближчих вузлів. Відмінності локальних сплайнів полягають у різних способах знаходження неперервних похідних. Найбільш природнім є рівність першої похідної полінома на кінцях фрагменту першій похідній інтерпольованої функції. Такі сплайни називають ермітовими. Для ермітового кубічного сплайна з неперервною першою похідною значення в точці визначатимуть найближчі чотири вузли. Труднощі вибору крайових значень є менш актуальними, оскільки є декілька шляхів їх вирішення. Для розрахунку похідної можна задіяти лише існуючі вузли, або центральні різниці замінити розділеними. Інший підхід полягає в тому, щоб продовжити фрагменти ліворуч та праворуч. Для фрагментів однакової ширини продовжити аналогічними фрагментами. Для фрагментів довільної ширини продовжити в нескінченність. Для ермітових кубічних сплайнів такий спосіб дає хороші результати [4].

Кількість неперервних похідних для Ермітових сплайнів менша, ніж для глобальних. Вони також не мають оптимальних властивостей глобальних сплайнів. До локальних ермітових сплайнів відноситься ламана, яка є хорошою ілюстрацією їх основних переваг. Саме локальні ермітові сплайни є найбільш придатними до застосування в задачах обробки цифрових даних, їх перевага не лише в простоті розрахунків. В більшості випадків неперервність другої похідної та оптимальні властивості не є продиктованими природою даних. Значно важливіша відповідність першої похідної характеру даних, а розрив другої похідної є візуально непомітним.

Аналогія глобального кубічного сплайна з реальними фізичними об'єктами є досить умовною.

Практично форма глобального та локального сплайнів, побудованих на одних вузлах відрізняється мало. Локальність дозволяє будувати сплайн без обмежень на число фрагментів.

**Висновки.** З отриманих результатів випливає, що методи кодування H.264 не підходять для кодування спеціалізованих відеоданих, оскільки існує певний ризик втрати важливої інформації. Окрім того, таке складне кодування займає надто багато часу для ефективного його використання в оперативному режимі. Тому цей метод кодування є неефективними для оперативного стиску та передачі відеосигналу спеціального призначення, де вимагається висока чіткість і деталізація відеоданих.

З метою оперативного стиснення та відновлення відеоданих, оптимізованих за точністю і швидкістю, запропонований адаптивний метод фільтрації і стиску відеоданих, який ґрунтується на «сигнальному» підході. При цьому в процесі фільтрації і стиску відліків обвідної  $R$ - ,  $G$ - і  $B$ -сигналів здійснюється відбір та максимальне збереження

первинних амплітудно-часових характеристик найбільш інформативних відліків обвідних відеосигналів. За рахунок рейтингового відбору найбільш інформативних суттєвих відліків обвідних відеосигналів для заданих величин коефіцієнтів стиску відеоданих досягається точне відновлення параметрів найбільш інформативних СВ відеосигналів.

Для задачі відновлення сигналу найбільш придатними є ермітові сплайни, а також сплайни, утворені з локальних сплайнових базисів. При зчитуванні даних відновлюються точки локальних екстремумів кривої та точки перегину після чого решта даних відновлюється інтерполяцією на цих точках. Відновлений сигнал має досить хорошу наближеність до вихідних даних при відновленні за допомогою Ермітових сплайнів за рахунок простоти в розрахунках та відповідності першої похідної характеру даних.

#### Список використаних джерел:

1. Боюн В. П. Интеллектуальні відео системи реального часу / В. П. Боюн // Праці міжнар. конф. 50 років Інституту кібернетики ім. В. М. Глушкова НАН України. — К., 2008. — С. 177 — 181.
2. Боюн В. П. Интеллектуальные видеосистемы. Информационные основы и принципы построения / В. П. Боюн // Праці міжнар. конф. з індуктивного моделювання. Держ. НДІ Інформаційної Інфраструктури. — Львів, 2002. — С. 14 — 30.
3. Ричардсон Я. Видеокодирование. H.264 и MPEG-4 — стандарты нового поколения / Я. Ричардсон. — М., 2005. — 365 с.
4. Шелевицький І. В. Методи та засоби сплайн-технології обробки сигналів складної форми / І. В. Шелевицький. — Кривий Ріг, 2002. — 304 с.
5. Селомон Д. Сжатие данных, изображений и звука / Д. Селомон. — М., 2004. — 368 с.

Based on the research staff of primary test video data sets investigated distortion bypass video form using standard methods of compressing video, analyzed the types of distortion recovered after compressing video, the results of modeling algorithms reconstruction video bypass and proposed methods and algorithms for efficient filtering, compression and restoration of video optimized for accuracy and speed.

**Key words:** *compression, restore video, coding, modeling of compression algorithms.*

Отримано 16.09.2010