

УДК 621.397

О.Г. Гуйда, А.М. Негода
О.М. Кушнір, М.Г. Гордєєв**ПЕРСПЕКТИВИ ВИКОРИСТАННЯ СПЛАЙН-ФУНКЦІЙ В ТЕХНОЛОГІЯХ
СТИСНЕННЯ ЦИФРОВОГО ВІДЕО**Національний авіаційний університет,
e-mail: guydasg@ukr.net, negoda@nevo.com.ua, okushnir@bigmir.net, skyfox46@ukr.net.

Розглянуто технології стиснення цифрового відео з втратами. Запропоновано замість дискретного-косинусного перетворення використовувати сплайн-функції. Представлено базисні функції ермітового кубічного сплайна.

Ключові слова: стиснення цифрового відео, стандарти стиснення відео, дискретне-косинусне перетворення, сплайн-функції, ермітів кубічний сплайн.

Вступ

Стиснення або компресія відео інформації являє собою процес її ущільнення, при якому цифрові дані приводяться до форми, яка потребує меншого об'єму пам'яті для їх зберігання. Стиснення даних досягається за рахунок усунення надлишковості компонентів даних, без яких можна обійтися для правильного відображення вихідної інформації.

Багато типів даних мають в собі статичну надлишковість. Такі дані можна ефективно стискувати використовуючи компресію без втрат. Нажаль, стиснення без втрат при застосуванні його до цифрового відео дає відносно невеликий вигравш. Найкращий результат, який можна досягнути використанням стандартного алгоритму JPEG-LS, дає коефіцієнт стиснення приблизно в 3-4 рази по відношенню до вихідного об'єму даних [3]. Для досягнення більшої ефективності стиснення відео даних приходиться застосовувати стиснення з втратами. Методи стиснення відео з втратами ґрунтуються на усуненні суб'єктивної надлишковості тих елементів зображення, які можна усунути без помітного впливу на зорове сприйняття відео.

Починаючи з появи першого стандарту компресії відео MPEG-1 (MPEG - Moving Picture Experts Group) відбувається неперервний розвиток і удосконалення методів компресії і декомпресії цифрового відео. На даний час, дане питання представляє собою надзвичайно актуальну сферу наукових досліджень.

Аналіз останніх досліджень і публікацій

Успіхи цифрової відеоіндустрії (перш за все широкоформатного цифрового відео і DVD-відео) базувалися на міжнародному стандарті ISO/IEC 13818, широко відомому під абревіатурою MPEG-2 (названий так по імені робочої групи експертів по рухомих зображеннях, - Moving Picture Experts Group). Необхідність кращого стиснення відеоданих спонукало розробку подальших стандартів відеостиснення, відомих під назвами ISO/IEC 14496 Part 2 (MPEG - 4 Visual) і рекомендації організації ITU, - H.264/ISO/IEC 144496 Part 10 (скорочено H.264). Стандарти MPEG - 4 Visual і H.264 мають загальне походження і багато загальних рис. Вони були розроблені на основі більше ранніх стандартів стиснення. Проте вони розвивають старі стандарти в істотно різних напрямках. Стандарт MPEG-4 Visual віддаляється від прямокутного відеокадру, пропонує гнучкий і відкритий погляд на візуальні комунікації і використовує високоефективне відеостискування і об'єктно-орієнтовану обробку даних. Стандарт H.264 має прагматичніший погляд. Він прагне виконувати ті ж дії, що і попередні стандарти (забезпечуючи механізм стискування для прямокутних кадрів), але з більшою ефективністю і стійкістю, забезпечуючи сумісність зі всіма широко поширеними типами додатків, такими як широкоекранне телебачення, зберігання візуальної інформації і передача потокового відео [1,3,4].

Постановка задачі

Підвищення ефективності програмних засобів стиснення відеоданих та створення нового методу стиснення зображень і відео, який має вищий коефіцієнт стиснення та кращу швидкодію декомпресії, ніж ті математичні методи, які використовуються в відомих кодексах MPEG - 4, H.264, DivX є перспективним і актуальним напрямком досліджень. З огляду на це, перспективним є використання сплайн-функцій у задачах обробки цифрового відео.

Виклад основного матеріалу

Для стиснення цифрового відео необхідно мати дві системи, які доповнюють одна іншу: компресор (кодер) та декомпресор (декодер). Кодер перетворює відеодані в стиснену форму для подальшої передачі або зберігання відео, а декодер робить зворотне перетворення, повертаючи стиснені відеодані до вигляду придатного для показу глядачу. Пару кодер/декодер прийнято називати "кодеком."

Більшість методів кодування відео використовують часову та просторову надлишковість. В часовій області існує значна подібність між відеокадрами які знаходяться в близькі моменти часу. Суміжні по часу кадри (які ідуть один за одним) мають високий ступінь кореляції, особливо при великій частоті кадрів. В просторовій області також спостерігається висока залежність величин пікселів (семплів), які знаходяться поруч.

Кодек перетворює початковий відеоряд за допомогою певної моделі. Модель кодування - це ефективно кодоване представлення відеоданих, за допомогою якого можна реконструювати ці дані з певним ступенем точності. У ідеалі модель повинна представляти послідовність з найменшим числом бітів і найбільшою можливою точністю. Ці дві мети (висока якість і ефективність стискування) зазвичай суперечать одна одній, оскільки високий ступінь стискування відеоданих припускає істотне зниження якості на виході декодера.

Відеокодек складається з наступних основних функціональних блоків[3]:

- блоку перетворення кольорного простору;
- блоку усунення часової надлишковості;
- блоку усунення просторової надлишковості;
- блоку ентропійного кодера.

Більшість алгоритмів використовують перетворення з кольорового простору RGB в кольоровий простір YUV (YCbCr), де Y - компоненту, що відповідає за яскравість, а U(Cb) і V(Cr) – компоненти, що відповідають за колір, для відділення яскравості від кольору. Людський зір чутливіший до яскравості, чим до кольору. Користуючись цим, можна ще до апроксимації зменшити кількість інформації про колір в 2 або 4 рази[2]. Спрощено переклад з кольорного простору RGB в колірний простір YUV можна представити за допомогою матриці переходу:

$$\begin{pmatrix} Y \\ Cb \\ Cr \end{pmatrix} = \begin{pmatrix} 0,299 & 0,587 & 0,114 \\ 0,5 & -0,4187 & -0,0813 \\ 0,1687 & -0,3313 & 0,5 \end{pmatrix} * \begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} 0 \\ 128 \\ 128 \end{pmatrix}$$

Перетворення $[Y Cb Cr]$ у $[R G B]$ (зворотно до попереднього перетворення). RGB-колір може бути обчислений безпосередньо з Y, Cb, Cr у такий спосіб:

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1,402 \\ 1 & -0,34414 & -0,71414 \\ 1 & 1,772 & 0 \end{pmatrix} * \begin{pmatrix} Y \\ Cb - 128 \\ Cr - 128 \end{pmatrix}$$

Очевидно, що при зменшенні інформації про колір знижується якість зображення.

Загальна схема усунення просторової надлишковості для стандартів MPEG і H.263/264 визначає процес стискування окремого кадру відеопотоку і збігається з процедурою стискування статичного повнокольорового зображення за стандартом JPEG (рис.1) [1].



Рис. 1 Основні етапи процедури стиснення по стандарту JPEG

На початку зображення переводиться з кольорного простору RGB в колірний простір YCrCb.

Потім зображення розбивається на блоки розміром 8x8. Такі блоки створюються для кожної колірної складової. Для компонент Cr і Cb формування блоку відбувається через строчку і через стовпець, тобто з початкового блоку 16x16 виходить один робочий блок 8x8. Надалі до

кожного з блоків застосовується дискретне-косинусне перетворення, яке описується виразом:

$$F(u,v) = \frac{c(u,v)}{4} \sum_{x=0}^7 \sum_{y=0}^7 f(x,y) \cos\left(\frac{2x+1}{16} u \pi\right) \cos\left(\frac{2y+1}{16} v \pi\right),$$

де - $u, v = 0, 1 \dots 7$; $c(u,v) = 1/2$, коли $u = v = 0$; $c(u,v) = 1$ – в інших випадках.

Для зменшення інформації, що зберігається, проводиться квантування отриманої матриці частотних коефіцієнтів. Квантуванням є ділення матриці частотних коефіцієнтів на матрицю квантування.

Для кожної з колірних складових зазвичай використовується своя матриця квантування $q[u,v]$:

$$Yq[u,v] = \text{IntegerRound}\left(\frac{Y[u,v]}{q[u,v]}\right).$$

Іноді для формування матриці квантування може використовуватися спеціальна вагова функція, що дозволяє сформувати коефіцієнти квантування, які перетворюють в 0 значну кількість високо і середньо частотних коефіцієнтів.

Етап квантування – це етап, на якому відбуваються найбільші втрати в якості зображення і етап, на якому здійснюється управління ступенем стискування зображення. При сильному квантуванні низькочастотні коефіцієнти можуть настільки змінитися, що зображення розпадеться на візуально помітні квадрати. Втрати у високочастотних коефіцієнтах приводять до появи «німба» навколо контурів з різким переходом кольору.

Квантування відображає числовий сигнал з областю значень X в квантований сигнал області Y із зменшеним числом значень. Це дає можливість представити квантовані величини з меншим числом бітів в порівнянні з початковими неквантованими величинами.

Проквантована матриця потім перетвориться у вектор за допомогою ZigZag сканування. Після ZigZag сканування низькочастотні елементи опиняються в «голіві» вектора, а високочастотні коефіцієнти в «хвості» вектора.

Далі виконується так зване вторинне стискування з використанням ентропійного кодера. Ентропійний кодер перетворює елементи послідовності символів, в стиснутий потік бітів, який можна зберегти у файлі або передати по мережах зв'язку.

Перспективним є використання на третьому кроці конвеєра сплайна-функцій, які використовуються для відновлення дискретної інформації. Сплайни у ряді ситуацій мають кращі апроксимаційні властивості, які забезпечують мінімально можливу при даній розмірності похибку. При їх застосуванні істотно зменшується обсяг обчислень.

Базисні функції ермітова кубічного сплайна X будуть визначатися за формулами [5]:

$$\bar{x}_{i0} = {}^1X_{i2} I_x(\Delta\tilde{x}_2) + {}^2X_{i1} I_x(\Delta\tilde{x}_1), \quad i = \overline{1, m_2},$$

$$\bar{x}_{i1} = {}^1X_{i3} I_x(\Delta\tilde{x}_3) + {}^2X_{i2} I_x(\Delta\tilde{x}_2) + {}^3X_{i1} I_x(\Delta\tilde{x}_1), \quad i = \overline{1, m_3},$$

$$\bar{x}_{ij} = {}^1X_{ij+2} I_x(\Delta\tilde{x}_{j+2}) + {}^2X_{ij+1} I_x(\Delta\tilde{x}_{j+1}) + {}^3X_{ij} I_x(\Delta\tilde{x}_j) + {}^4X_{ij-1} I_x(\Delta\tilde{x}_{j-1}), \quad j = \overline{2, r-2}, \quad i = \overline{1 + m_1, m_{r-2}},$$

$$\bar{x}_{ir-1} = {}^2X_{i r} I_x(\Delta\tilde{x}_r) + {}^3X_{i r-1} I_x(\Delta\tilde{x}_{r-1}) + {}^4X_{i r-2}, \quad i = \overline{1 + m_{r-3}, m_r},$$

$$\bar{x}_{ir} = {}^3X_{i r} I_x(\Delta\tilde{x}_r) + {}^4X_{i r-1} I_x(\Delta\tilde{x}_{r-1}), \quad i = \overline{1 + m_{r-2}, m_r}$$

де
$$I_x(\Delta\tilde{x}_j) = \begin{cases} 1 & \text{при } x \in \Delta\tilde{x}_j = \tilde{x}_j - \tilde{x}_{j-1} \\ 0 & \text{при } x \notin \Delta\tilde{x}_j, j = \overline{1, r}; \end{cases}$$

$${}^1X_{ij} = -\frac{h_j^2 x_{ij} (1 - x_{ij})^2}{h_{j-1} (h_{j-1} + h_j)}, \quad j = \overline{2, r}, \quad i = \overline{1 + m_{j-1}, m_j}; \quad {}^2X_{i1} = 1 - x_{i1} - \frac{h_1 x_{i1}^2 (1 - x_{i1})}{(h_1 + h_2)}, \quad i = \overline{1, m_1};$$

$${}^2X_{ij} = 1 - x_{ij} - \frac{h_j x_{ij}^2 (1 - x_{ij})}{(h_j + h_{j+1})} + \frac{h_j x_{ij} (1 - x_{ij})^2}{h_{j-1}}, \quad j = \overline{2, r-1}, \quad i = \overline{1 + m_{j-1}, m_j};$$

$${}^2X_{ir} = 1 - x_{ir} - \frac{h_r x_{ir} (1 - x_{ir})^2}{h_{r-1}}, \quad i = \overline{1 + m_{r-1}, m_r}; \quad {}^3X_{i1} = x_{i1} - \frac{h_1 x_{i1}^2 (1 - x_{i1})}{h_2}, \quad i = \overline{1, m_1};$$

$${}^3X_{ij} = x_{ij} - \frac{h_j x_{ij}^2 (1 - x_{ij})}{h_{j+1}} - \frac{h_j x_{ij} (1 - x_{ij})^2}{h_{j-1} + h_j}, \quad j = \overline{2, r-1}, \quad i = \overline{1 + m_{j-1}, m_j};$$

$${}^3X_{ir} = x_{ir} - \frac{h_r x_{ir} (1 - x_{ir})^2}{h_{r-1} + h_r}, \quad i = \overline{1 + m_{r-1}, m_r};$$

$${}^4X_{ij} = -\frac{h_j^2 x_{ij}^2 (1 - x_{ij})}{h_{j+1} (h_j + h_{j+1})}, \quad j = \overline{1, r-1}, \quad i = \overline{1 + m_{j-1}, m_j};$$

$$x_{ij} = \frac{x_i - \tilde{x}_{j-1}}{h_j}; \quad h_j = \tilde{x}_j - \tilde{x}_{j-1}; \quad j = \overline{1, r}; \quad x_i \in [\tilde{x}_{j-1}, \tilde{x}_j], \quad j = \overline{1, r-1}; \quad x_i \in [\tilde{x}_{r-1}, \tilde{x}_r];$$

$$m_j = \sum_{u=1}^j K_u, \quad j = \overline{1, r}; \quad m_{-1} = m_0 = 0; \quad m_r = N;$$

де K_u - кількість відліків на u -му відрізку.

Значення ермітова кубічного сплайна в довільній точці $x \in [\tilde{x}_{j-1}, \tilde{x}_j]$, $j = \overline{1, r}$ визначається за формулою:

$$S_3(x_i) = \hat{a}_{j-2} {}^1X_{ij} + \hat{a}_{j-1} {}^2X_{ij} + \hat{a}_j {}^3X_{ij} + \hat{a}_{j+1} {}^4X_{ij} \\ i = \overline{1 + m_{j-1}, m_j}, \quad j = \overline{1, r}, \quad a_{-1} = a_{r+1} = 0.$$

Висновки

Актуальність та перспективність використання сплайнів у задачах обробки цифрового відео безпосередньо впливає з їх хороших наближувачих та обчислювальних властивостей. Так як сплайни – кусочно-поліноміальні функції, то вони можуть легко бути використані при обчисленнях.

Перевагою розглянутого підходу є його універсальність і відносна простота. Отримані сплайнові бази дозволяють ефективно розв'язувати інтерполяційні задачі, що є основою методів обробки цифрових сигналів. Альтернатива широкого вибору сплайнових базисів одночасно ставить серйозну задачу вибору найкращого з них для вирішення конкретної задачі.

Найбільш доцільно користуватися кубічним сплайном, оскільки він має неперервну першу похідну і ефективний у розрахунках. Сплайни вищих степенів використовувати в загальному випадку недоцільно у зв'язку зі зростанням обсягів розрахунків при відносно малих виграшах якості наближення.

Список літературних джерел

1. Ватолин Д., Ратушняк А., Смирнов М. Юкин В. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео. // -М. – Диалог-МИФИ. -2003. - 384 с.
2. Гонсалес Р., Вудс Р. Цифровая обработка изображений. // Пер. с англ.- Москва.- Техносфера. – 2006. -1072 с.
3. Ричардсон Я. Видеокодирование H.264 и MPEG-4 – стандарты нового поколения. // – М. – Техносфера. - 2004. – 368 с.
4. Сэломон Д. Сжатие данных, изображений и звука. // – М. – Техносфера. - 2004. – 368 с.
5. Шутко Н.А. Методы и средства повышения достоверности обработки измерительной информации и контроля параметров радиоэлектронных систем управления воздушным движением: дис. докт. техн. наук: 05.22.14 // - К. - КИИГА, - 1991. - 245 с.