

УДК 621.397

О.В. Яковенко, кандидат технічних наук,
А.О. Красноруцький, кандидат технічних наук,
 старший науковий співробітник,
Д.В. Ісаєв

ОПИС АЛГОРИТМУ КОМПРЕСІЇ ВІДЕОДАНИХ СИСТЕМИ АУДІО-, ВІДЕОРЕЄСТРАЦІЇ НА БОРТОВИХ СИСТЕМАХ

Наведено аналіз алгоритму компресії відеоданих радіотехнічної системи аудіо-, відеореєстрації літального апарата, обґрунтовано шляхи підвищення ступеня компресії відеоданих з можливістю їх відновлення із заданою втратою якості шляхом заміни комбінації дискретного косинусного перетворення і компенсації руху на застосування ортогональних перетворень (ОП) на основі цілочисельних кусочно-постійних функцій і методів цілочисельного арифметичного кодування значень трансформант перетворення.

Ключові слова: дискретне косинусне перетворення, алгоритм компресії, ортогональні перетворення, базис Уолша, арифметичне кодування.

Предложен анализ алгоритма компрессии видеоданных радиотехнической системы аудио-, видеорегистрации летательного аппарата, обоснованы пути повышения степени компрессии видеоданных с возможностью их восстановления с заданной потерей качества путем замены комбинации дискретного косинусного преобразования и компенсации движения на применение ортогональных преобразований (ОП) на основе целочисленных кусочно-постоянных функций и методов целочисленного арифметического кодирования значений трансформант преобразования.

Ключевые слова: дискретное косинусное преобразование, алгоритм компрессии, ортогональные преобразования, базис Уолша, арифметическое кодирование.

The analysis of the algorithm of a compression of the video data of radioengineering system of audio-, videoregistration of a flying machine is offered, the ways of an increasing of degree of the videodata compression with the possibility of their restoration with the set loss of quality are proved.

Keywords: discrete cosine transform, algorithm of compression, orthogonal transformations, Walsh basis, arithmetic coding.

Під час проведення досліджень САВР-29 виникла низка певних проблем. До однієї з них можна віднести і те, що існуючий метод компресії даних не забезпечує такого ступеня скорочення обсягу інформації, який необхідний для оперативного доведення відеоінформації (передачу її в реальному масштабі часу). Це пов'язано з тим, що відтворення інформації в міжнародному файловому форматі AVI з використанням стандартних кодеків не дає нам необхідного скорочення обсягу і поліпшення чіткості зображення. Пояснюється це тим, що технологія стиску відео, яка заснована на використанні комбінації дискретного косинусного перетво-

рення (ДКП) і компенсації руху, не характерна для конкретного типу зображень, а саме для зображення місцевості з великою кількістю дрібних деталей тощо.

Такі недоліки, як: поява блокових артефактів при високому ступені декомпресії, випадкове розмиття і злам гострих граней, великі вимоги до обчислювальних потужностей роблять цю технологію стиску відео недоцільною для використання в САВР-29.

Тому *метою статті* є обґрунтування перспективного напрямку удосконалення технології стиску відео в форматі AVI шляхом заміни комбінації дискретного косинусного перетворення і компенсації руху на застосування ортогональних перетворень (ОП) на основі цілочисельних кусочно-постійних функцій і методів цілочисельного арифметичного кодування значень трансформант перетворення. Це зумовлено викладеними нижче чинниками.

1. Концентрацією значної частини енергії вихідного зображення в невеликому числі спектральних коефіцієнтів трансформант.

Можливість визначення більш і менш значимих коефіцієнтів дозволили вибрати дискретне ортогональне перетворення для створення ефективного методу стиску зображення.

2. Тим, що для сильно насичених зображень, які володіють певною анізотропією, застосування базису Уолша ефективніше від базису тригонометричних функцій, якщо критерієм вважати зменшення тимчасової складності перетворення й зниження помилки апроксимації.

3. Тим, що застосування арифметичного кодування трансформант перетворення Уолша дає більше наближення до оптимального кодування за рахунок більшого ступеня наближення частоти кодування до відносної частоти появи коефіцієнтів трансформант.

Основними вимогами до системи аудіо-, відеореєстрації (САВР-29) є:

– забезпечення номінального значення стиску 720x576 пікселів при 25 кадрів/с;

– формат кодування MPEG-4.

Основні етапи підготовки до процесу кодування в САВР-29 зображені на рис. 1.

Вхідні відеодані представляються в колірній моделі RGB, далі обирається технологія стиснення з втратами. Ця технологія включає в себе вибір стандартного кодеку, а саме MPEG-4, заснованого на принципі комбінації дискретного косинусного перетворення і компенсації руху.

Колірна модель RGB (від англ. Red, Green, Blue – червоний, зелений, голубий) є досить обґрунтованою. У цій моделі спектральна функція представляється як сума кривих чутливості для кожного типу колбочок з ненегативними ваговими

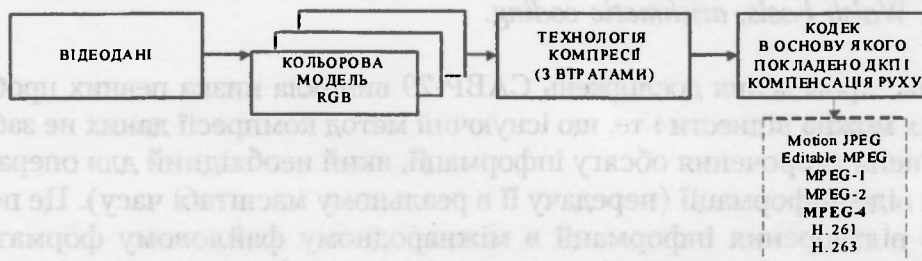


Рис. 1 Основні етапи підготовки до процесу кодування в САВР-29

коефіцієнтами (звичайно їх нормують від 0 до 1), які так і позначаються – R, G і B. Ця модель характеризується властивістю адитивності (ми складаємо кольори для одержання нових).

Процес стиснення відео в міжнародному файловому форматі AVI з використанням кодеку, в основу якого покладено алгоритм MPEG-4, зображено на рис. 2.



Рис. 2 Процес стиснення відео алгоритмом MPEG

Після того, як відео було представлено в кольоровій моделі RGB, відбувається процес кодування. Відео в MPEG ділиться на дві частини:

- 1) зменшення надмірності інформації в часовому вимірі (основною є невелика різниця між кадрами);
- 2) стиснення окремих зображень.

Для того, щоб задовольнити вимоги і збільшити гнучкість алгоритму, розглядається чотири типи кадрів:

- I-кадри – кадри, стиснені незалежно від інших кадрів (I-Intra pictures);
- P-кадри – стиснені з використанням посилання на одне зображення (P-Predicted);
- B-кадри – стиснені з використанням посилання на два зображення (B-Bidirection);
- DC-кадри – незалежно стиснені з великою втратою якості (використовуються тільки при швидкому пошуку).

I-кадри забезпечують можливість довільного доступу до будь-якого кадру, є вхідними точками в потік даних для декодера. P-кадри використовують при архівації посилання на один I- або P-кадр, підвищуючи тим самим ступінь стиснення фільму в цілому. B-кадри використовують посилання на два кадри, які знаходяться попереду і за ними, забезпечуючи найвищий ступінь стиснення, проте самі як посилання використовуватися не можуть. Після визначення кадрів формується їх послідовність у фільмі. Для прикладу вона може бути такою: IBVRVVRVVRVVRVVRV... Або, якщо ми не економимо на ступені стиснення, такою як на рис. 2

Частота I-кадрів вибирається залежно від вимог до часу довільного доступу і надійності потоку при передачі через канал з помилками. Співвідношення P- і B-кадрів підбирається з огляду на вимоги до величини компресії і обмежень декодера. Як правило, декодування B-кадрів потребує більше обчислювальних потужностей, але дозволяє підвищити ступінь стиснення. Саме зміна частоти кадрів різних типів забезпечує алгоритму необхідну гнучкість і можливість розширення. Зрозуміло, що для того, щоб розпакувати B-кадр, ми повинні розпакувати ті кадри, на які він посилается. Тому для послідовності IBVRVVRVVRVVRVVRV

кадри в фільмі будуть записані так: 0**312645..., де цифри – номери кадрів, а зірочкам відповідають В-кадри з номерами -1 і -2, якщо ми знаходимось в середині потоку, або порожні кадри, якщо ми на початку фільму.

Подібний формат володіє достатньо великою гнучкістю і здатен задовольнити найрізноманітніші набори вимог [8].

Одним з основних понять при стисненні декількох зображень є поняття макроблоку. При стисненні кадру після формування макроблоку 1 відбувається перетворення кадрів із кольорового простору RGB у кольоровий простір YUV. Кожна з площин стисненого зображення (Y, U, V) поділяється на блоки 8×8, з якими працює ДКП. Причому площини U і V відповідають компонентам кольору, беруться з роздільною здатністю в 2 рази меншою (по вертикалі і горизонталі) ніж вхідне зображення. Таким чином, ми одразу отримуємо стиснення в 2 рази, користуючись тим, що око людини гірше розрізняє колір окремої точки зображення, на відміну від її яскравості. Після ДКП відбувається квантування і кодування Хаффмана. Блоки 8×8 групуються в макроблоки. Макроблок – це група з чотирьох сусідніх блоків у площини компоненти яскравості Y (матриця пікселів 16×16) і два відповідних їм за розташуванням блоки із площини кольоровості U і V. Таким чином, кадр розбивається на незалежні одиниці, які несуть повну інформацію про частину зображення. При цьому розмір зображення має бути кратним 16 [3, 6, 8].

Окремі макроблоки стискаються незалежно, тобто в В-кадрах ми можемо стиснути макроблок конкретний як І-блок. Р-блок з посиланням на попередній кадр, Р-блок з посиланням на наступний кадр, і в кінці, як В-блок. Алгоритм компресії окремих кадрів в MPEG схожий на відповідний алгоритм для статичних зображень – JPEG. Якщо казати коротше, то ДКП вихідної матриці 8×8, стиснення вектору груповим кодуванням, квантування і стиснення по алгоритму Хаффмана. Перевагами цього алгоритму є підтримка достатньо прогресивних звукових стандартів, високий ступінь універсальності, підтримка новітніх технологій [8].

Недоліком є висока складність реалізації. Але в цій ситуації основною проблемою при обробці відео цим алгоритмом є великі обсяги дискового простору, необхідного для збереження навіть невеликих фрагментів, не кажучи вже про реалізацію передачі інформації в реальному масштабі часу. При цьому традиційні алгоритми стиснення зображень, орієнтовані на окремі кадри, не врятовують ситуацію, оскільки навіть при зменшенні потоку в 10 разів він залишає достатньо великі величини. Одним із перспективних напрямів рішення цієї проблеми є застосування замість комбінації ДКП дискретного перетворення Уолша, а замість квантування – арифметичного кодування. Процес стиснення алгоритмом зміниться і буде мати такий вигляд, як на рис. 3.



Рис. 3 Процес стиснення відео удосконаленим алгоритмом

Загальне використання дискретного перетворення Уолша (ДПУ) й арифметичного кодування дозволяє знизити час на стиск і відновлення зображень, а також скоротити додаткову статистичну надмірність. Ці результати зумовлені тим, що двовимірні ортогональні перетворення знижують ступінь статистичної залежності між елементами вихідних зображень як у рядках, так і у стовпчиках. При цьому отримані трансформанти характеризуються наявністю нерівномірності розподілу значень їхніх компонент. Звідси, організація ентропійного кодування трансформант ортогональних перетворень дозволяє виключити статистичну надмірність. У той же час ортогональні перетворення здійснюють поділ блоків вихідного зображення на низькочастотні і високочастотні складові. Це досягається за рахунок розкладання вихідних блоків за базисними функціями з різної частотою. Високочастотні складові несуть інформацію про дрібні об'єкти і тому їхня обробка з похибкою не викличе значного впливу при сприйнятті зоровою системою. Це дозволяє використовувати ортогональні перетворення для обліку психовізуальних закономірностей [7].

Варіант економного кодування трансформант перетворення Уолша полягає в арифметичному кодуванні. Арифметичний стиск дає краще наближення до нижньої межі, заданою ентропією джерела, за рахунок обробки кожного наступного символу з урахуванням статистичних характеристик попередніх елементів.

Метод арифметичного кодування полягає у виконанні наступних етапів:

1. Визначається вихідний робочий інтервал.

2. Вихідний інтервал розбивається на підінтервали з довжинами, рівними ймовірностям появи символів у трансформантах, які позначаються як діапазони відповідних компонентів. У результаті виконання цього етапу формується таблиця вихідних даних (рис. 4). Символи повідомлення в стовпці "символ повідомлення" вибудовуються в порядку убутання згідно з частотою появи символів у повідомленні. Кожному символу привласнюється ймовірність і діапазон символів повідомлення відповідно.

3. Процес кодування зводиться до зменшення робочого інтервалу діапазону. За вихідний робочий інтервал береться $[0,1)$, тобто

$$l_0 = 0, h_0 = 1,$$

де l_0 – початок інтервалу,

h_0 – кінець інтервалу.

Після цього інтервал розбивається на діапазони, що відповідають заданим частотам i , відповідно, ймовірностям появи символів повідомлення (згідно з таблицею).

Довжина першого робочого інтервалу задається діапазоном першого поточного символу, що кодується

символ повідомлення	частота появи символу повідомлення	вероятність появи символу повідомлення	діапазон символу повідомлення
y_i	f_i	p_i	$[a[y_i], b[y_i])$
y_i	f_i	p_i	$[a[y_i], b[y_i])$

Рис. 4. Таблиця вихідних даних для арифметичного кодування (за таблицею) і дорівнює ймовірності появи цього символу в повідомленні.

Тепер цей інтервал береться за вихідний для наступного символу, що кодується. Далі зчитується наступний символ, що кодується, і обчислюється його робочий інтервал і т.д.

Таким чином, за вихідний береться робочий інтервал, отриманий на попередньому етапі кодування символів. У підсумку, довжина робочого інтервалу зменшується пропорційно ймовірності появи поточного символу.

Робочий інтервал $[l_i, h_i)$ можна знайти з виразів:

$$l_i = l_{i-1} + a[y_i](h_{i-1} - l_{i-1}), \quad (2.68)$$

$$h_i = l_{i-1} + b[y_i](h_{i-1} - l_{i-1}), \quad (2.69)$$

де l_i – початок робочого інтервалу i -го символу, що кодується;

h_i – кінець робочого інтервалу i -го символу, що кодується;

l_{i-1} – початок робочого інтервалу попереднього символу, що кодується;

Узагальнюючи усе сказане вище, можна зробити такі висновки:

h_{i-1} – кінець робочого інтервалу попереднього символу, що кодується;

$a[y_i]$ – початок діапазону y_i -го символу, що кодується;

$b[y_i]$ – кінець діапазону y_i -го символу, що кодується.

Таким чином, кожний наступний інтервал символу, що кодується, в повідомленні вкладений у попередній інтервал. У підсумку, остаточно довжина інтервалу буде дорівнювати добутку ймовірностей всіх символів у повідомленні, а його початок залежить від порядку проходження символів у повідомленні.

Заключним етапом є знаходження кодового числа. Його знаходять шляхом визначення середнього числа між початком і кінцем робочого інтервалу останнього символу, що кодується, в повідомленні:

$$K_q = \frac{h_i + l_i}{2},$$

де K_q – кодове число повідомлення 1. Розроблений алгоритм дозволяє визначити граничні значення основних технічних характеристик САВР-29, для якої тимчасові витрати стиснення відеоданих за допомогою перетворення Уолша будуть нижче ніж тимчасові витрати на виконання комбінації дискретного косинусного перетворення

2. Мінімальний виграш за часом обробки для ДПУ відносно ДКП у середньому дорівнює 3,6 і 2,75 рази відповідно для розміру блоку і зумовлений тим, що, на відміну від ДКП, значення елементів матриці базису Уолша є цілими числами, що приймають тільки два значення 1 і -1. Тому виконання перетворення Уолша ґрунтується тільки на цілочисельних операціях додавання й віднімання.

3. Створена модель алгоритму дозволяє оцінити кількість інформації і статистичної надмірності в трансформантах Уолша в залежності від значень дисперсії елементів і кореляційної функції вихідного зображення.

4. Апроксимація реалістичних зображень, які не мають властивості когерентності (гладкості), на основі базисів Уолша формує менше значення показника середньоквадратичної похибки в порівнянні з величиною похибки, що одержується при апроксимації базисами ДКП.

5. За рахунок арифметичного кодування забезпечується виграш за кількістю статистичної надмірності, що скорочується, у трансформантах Уолша щодо кодування методом Хаффмана, що у середньому дорівнює 20 %.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Асташкин А.А. Космические системы, аппараты и приборы для решения задач природопользования и экологического контроля / А.А. Асташкин. – М. : ВИНТИ, 1991. – 142 с.
2. Виноградов Б.А. Аэрокосмический мониторинг экосистем / Б.А. Виноградов. – М. : Наука, 1984. – 319 с.
3. Гонсанс Р., Вудс Р. Цифровая обработка изображений / Р. Гонсанс, Р. Вудс ; пер. с англ. – М. : Техносфера, 2005. – 1072 с.
4. Кутовий О.П. Тенденції розвитку беспілотних літальних апаратів / О.П. Кутовий // Наука і оборона. – 2000. – № 4. – С. 39–47.
5. Матвеев С.И. Высокоточные системы РВиА: перспективы и основные направления работ по созданию разведывательно-ударных и разведывательно-огневых комплексов / С.И. Матвеев. – М. : Военная мысль. – 2005. – № 2. – С. 22–27.
6. Птачек М. Цифровое телевидение. Теория и техника / М. Птачек ; пер. с чешского под ред. Л.С. Виленчика. – М. : Радио и связь, 1990. – 528 с.
7. Красноруцкий А.А. Метод кодирования трансформант Уолша в системах мониторинга Земли / А.А. Красноруцкий, С.Я. Яценко // Системи обробки інформації (Харківський університет Повітряних Сил (ХВУ)). – 2007. – № 8(66). – С. 7–10.
8. Ватолин Д. Методы сжатия данных. Устройство архиваторов, сжатие изображений и видео / Д. Ватолин, А. Ратушняк, М. Смирнов, В. Юкин. – М. : ДИАЛОГ – МИФИ, 2003. – 338 с.

Отримано 11.05.2011