

**ХАРАКТЕРИСТИКИ СИСТЕМИ КОДУВАННЯ ЗОБРАЖЕНЬ  
H.264 | MPEG-4 AVC**

**Лопатін Є.О.**

**CHARACTERISTICS OF VIDEO CODING SYSTEM H.264 | MPEG-4 AVC**

**Lopatin E.O.**

Одеська національна академія зв'язку (ОНАЗ) ім. О.С. Попова

На даному етапі впровадження цифрового телебачення на території України виникла гостра необхідність у визначенні різних підходів до стиснення відеоінформації, в оцінюванні можливого ефекту від використання методів кодування в телевізійних і споріднених їм застосуваннях. Для поступового та всебічного освоєння даної проблеми необхідно провести аналіз підходів до стиснення відеоінформації і оцінити можливі результати при використанні таких методів в телевізійних і споріднених їм застосуваннях.

На сьогодні стандарт H.264 | MPEG-4 AVC являє собою найбільш досконалим стандартом стиснення відеоінформації. В ньому застосовано велику кількість нових прийомів, що відсутні в попередніх стандартах MPEG-2 Video, MPEG-4 Visual та Рекомендації

ITU-T H.263, на основі яких сумісними зусиллями ITU та ISO і був створений стандарт кодування MPEG-4 AVC. Основною метою при створенні стандарту відео кодування H.264 | MPEG-4 AVC було поліпшення алгоритму стиснення та створення “дружньої” середі передавання не мовленнєвих та мовленнєвих застосуваннях.

Новий стандарт стиснення H.264 | MPEG-4 AVC створений з метою вирішення технічних рішень у всіх можливих відео застосуваннях. В ньому поліпшені або повністю перероблені алгоритми стиснення відеоінформації, що призводить до кращих результатів в порівнянні з попередніми версіями стандартів MPEG:

- компенсація руху з використанням змінних розмірів блоку, включаючи малі розміри блоку. Стандарт підтримує досить велику гнучкість у виборі розмірів і форм блоків при компенсації руху в порівнянні з усіма попередніми стандартами, мінімальний розмір блоку становить  $4 \times 4$  елементи зображення;
- компенсація руху з точністю до чверті елемента зображення. Більшість попередніх стандартів підтримували компенсацію руху з точністю не вище половини елемента зображення. Новий стандарт підвищує точність компенсації руху до чверті елемента зображення, як і в розширеному профілі MPEG-4 Visual, але в новому стандарті зменшена складність інтерполяції відліків у половину й чверть елемента зображення. Інтерполяція відліків у половину елемента зображення здійснюється округленням сусідніх цілочисельних відліків з використанням 6 фільтрів з ваговими коефіцієнтами  $(1/32, -5/32, 5/8, 5/8, -5/32, 1/32)$ . Інтерполяція відліків у чверть елемента зображення здійснюється лінійно по парі сусідніх відліків у половину елемента зображення;
- вектори руху, що виходять за межі зображення. В стандарті MPEG-2 Video і попередніх йому стандартах вектори руху могли вказувати тільки на елементи зображення, що перебувають у границях декодованого опорного зображення. Методика екстраполяції за межі зображення, що з'явилася як необов'язкова в Рекомендації ITU-T H.263 (Додаток O) [4], включена в новий стандарт. Сама методика екстраполяції проста й ґрунтується на “заміні” елементів зображення поза областю кодованого зображення, на які спрямований вектор руху, граничними елементами;

- незалежність порядку відтворення зображень і порядку опорних зображень. У попередніх стандартах встановлювалася жорстка залежність між порядком проходження зображень для використання при компенсації руху й порядком проходження зображень при відтворенні. У новому стандарті ці обмеження значною мірою усунуті, що дозволяє кодера обирати порядок зображень для компенсації руху й для відтворення з високим ступенем гнучкості, що обмежена тільки об'ємом пам'яті, що гарантує можливість декодування. Усунення обмеження також дозволяє в ряді випадків виключити додаткову затримку, раніше пов'язану із двоспрямованим передбаченням;
- поліпшена обробка “пропущених” (skipped) блоків і блоків з “прямим” (direct) передбаченням руху. У попередніх стандартах “загублені” області зображення, кодованого із передбаченням, не могли містити руху, що негативно позначалося на ефективності кодування, коли кодована відеопослідовність містила рух сцени в цілому. Новий стандарт припускає можливу наявність руху в “загублених” областях. Для областей із двоспрямованим передбаченням новий стандарт також передбачає розширений метод компенсації руху, відомий за назвою “пряма компенсація руху”, що дозволяє поліпшити “пряме” передбачення, уведене в додаток до H.263 [4] і MPEG-4 [1];
- направлене просторове передбачення для внутрикадрового кодування. Нова методика екстраполяції країв раніше декодованих частин поточного зображення застосовується в областях зображень, кодованих за допомогою внутрикадрової методики (тобто без посилань на зміст інших зображень). Ця методика підвищує якість сигналу, що його використовують для передбачення, а також дозволяє використовувати для передбачення сусідніх областей, які були закодовані не за допомогою внутрикадрової методики (це не допускалося в стандартах H.263 і MPEG-4). Можливі 9 режимів передбачення блоку яскравості 4×4;
- деблокуюче фільтрування в циклі кодування. Кодування відео, засноване на обробці блоків, приводить до шуму, що має назву “блокінг-ефект” (блоковим шумом). Причиною може служити як передбачення, так і кодування залишків передбачення. Застосування адаптивного деблокуючого фільтра – добре відомий метод підвищення якості відновленого відео. При правильному проектуванні використання такого фільтра може підвищити як суб'єктивне, так і об'єктивна якість. Побудований на основі концепції додатка до стандарту H.263, деблокувальний фільтр у новому стандарті внесений у петлю зворотного зв'язка пророкування й компенсації руху після зворотного перетворювача в декодері перед буфером зберігання макроблоків для подальшого передбачення. Таким чином, підвищення якості зображення може бути використане в міжкадровому передбаченні, що, у свою чергу, поліпшує можливість передбачення руху інших зображень.

На додаток до поліпшених методів передбачення з метою підвищення ефективності кодування були також розширені інші частини стандарту:

- компенсація руху з використанням деревоподібного алгоритму. Відліки яскравості в кожному блоці 16×16 відліків можуть бути поділені чотирма різними шляхами, таким чином рух може бути компенсовано поділенням макроблоку 16×16 відліків на дві частини розміром 16×8 (або 8×16) відліків або чотири частини розміром 8×8 відліків. При використанні поділення макроблоку на чотири частини розміром 8×8 відліків можливе подальше поділення на частини розміром 4×8 (або 8×4) відліків або 4×4 відліків. Такі поділи надають можливість роботи з великою кількістю можливих варіантів компенсації руху в кожному макроблоці. Для кожного поділення необхідно використовувати окремий вектор руху. Кожний вектор руху має бути кодованим та переданим разом з кодованим варіантом поділення в стисненому потоці даних. Обирання варіанту з великим розміром частин (16×16, 16×8, 8×16) визначає меншу кількість біт, необхідних для передавання інформації щодо вектору руху та типу поділення, але такий

метод може призвести до значної надлишковості в областях з високим рівнем деталізації. Обирання малого розміру частин ( $8 \times 4$  або  $4 \times 4$ ) призводить до зменшення надлишковості після використання компенсації руху але потребує великої кількості біт для передавання вектору руху та інформації щодо методу поділу макроблоку. Великі частини обиранються для представлення однотонних частин зображення, малі частини служать для компенсації руху в областях з високим рівнем деталізації;

- перетворення з використанням 16-розрядної арифметики. Всі попередні стандарти потребували від кодерів і декодерів більше складних обчислень під час перетворення, як правило, з використанням 32-розрядної арифметики. Новий стандарт використовує тільки 16-розрядну точність обчислень;
- точне зворотне перетворення. У попередніх стандартах відеокодування перетворення, що його використовували для представлення відеоінформації (дискретне косінусне перетворення – ДКП), обмовлялося тільки в рамках припустимих помилок через практичну нереалізованість точного зворотного перетворення. У підсумку, кожна реалізація декодера могла злегка розрізнитись у кінцевому декодованому відеозображенні, що спричиняло розбіжність подання декодованого відеозображення на кодері й декодері, що приводило до зниження ефективної якості відеозображень. Новий стандарт першим дозволив досягти точного збігу декодованого відеозображення на всіх декодерах;
- арифметичне ентропійне кодування. У стандарт включений передовий метод ентропійного кодування, відомий як арифметичне кодування. Хоча арифметичне кодування включалося як необов'язкова можливість в Рекомендації ІТУ-Т Н.263, новий стандарт використовує більш ефективну методику, яка має назву САВАС (контекстно-адаптивне двійкове арифметичне кодування);
- контекстно-адаптивне ентропійне кодування. У новому стандарті застосовуються два методи ентропійного кодування - САVLC (контекстно-адаптивне кодування кодами змінної довжини) і САВАС (контекстно-адаптивне двійкове арифметичне кодування); ці методи використовують адаптивність, основу на контексті, що підвищує продуктивність кодування в порівнянні з методиками в попередніх стандартах.

Стійкість до помилок, втрат даних і гнучкість роботи в різних мережах у новому стандарті визначаються цілим рядом нових методик, у тому числі таких як:

- структура набору параметрів. Набір параметрів забезпечує стійку й ефективну передачу заголовної інформації. При використанні попередніх стандартів втрата декількох ключових біт інформації (таких як заголовки послідовності або заголовки зображення) може внести сильний негативний вплив на процес декодування. Для запобігання такого впливу новий стандарт передбачає можливість відділення ключової інформації від інших даних і обробки її більш гнучким спеціалізованим способом;
- синтаксична структура блоку рівня мережної абстракції (NAL unit). Кожна синтаксична структура в новому стандарті розташована в логічному пакеті даних, що має назву "блок NAL". Синтаксична структура блоків NAL передбачає досить велику гнучкість для передачі відеоконтенту способом, сумісним з кожним специфічним видом мереж;
- гнучкий розмір зрізу. На відміну від сталої та незмінної структури зрізу в стандарті MPEG-2, що зменшує ефективність кодування, збільшуючи обсяг заголовної інформації й зменшуючи ефективність передбачення, розміри зрізу в новому стандарті задаються дуже гнучко (як це було раніше в стандарті MPEG-1). Зріз має розміри  $(16 \text{ рядків}) \times (n \times 16 \text{ елементів зображення})$ ;

- гнучке упорядкування макроблоків. В H.264/AVC була розроблена нова можливість розподілу зображення на області, під назвою «група зрізів», причому кожний зріз становить незалежну підмножиною групи зрізів, яку декодують. При ефективному використанні гнучке упорядкування макроблоків може істотно підвищити стійкість до втрати даних шляхом обробки просторових залежностей між областями, які кодуються в кожному зрізі;

- довільне упорядкування зрізів. Новий стандарт дозволяє посилати й одержувати зрізи в довільному порядку відносно один одного, тому що кожний зріз кодованого зображення може бути декодованим незалежно від інших зрізів (при певних обмеженнях кодування). Ця можливість, що вперше з'явилася в необов'язковому додатку до стандарту H.263, може знизити затримку в додатках реального часу, особливо при використанні на мережах, що мають режим роботи “доставка позачергово” (наприклад, в IP-мережах);

- надлишкові зображення. H.264 | MPEG-4 AVC містить можливість посилки кодером надлишкового подання областей зображень, дозволяючи відтворити області зображень (звичайно з деякою втратою якості), для подання яких дані були загублені в процесі передачі, що збільшує стійкість до втрати даних;

- поділ даних. Деякі види кодової інформації для подання кожної області (наприклад, вектори руху й інша інформація пророкування) мають більшу значимість для подання відеоконтенту, ніж інші види інформації. Новий стандарт дозволяє розділити синтаксис кожного зрізу на частини (до трьох частин) для передачі залежно від категорії синтаксичних елементів. Ця частина стандарту побудована на основі стандарту MPEG-4 і необов'язкового додатку до H.263. У новому стандарті методика спрощена завдяки використанню єдиного синтаксису.

Стандарт H.264 | MPEG-4 AVC є значним кроком на шляху технічного прогресу методів стиснення відеоінформації. Слід чекати найближчим часом нових кроків на цьому шляху. А це означає, що як в мовленнєвих, так і в немовленнєвих застосуваннях спеціалістам всіх напрямків прийдеться займатися проблемами сумісності “старого” і “нового” і віддавати перевагу гнучким рішенням при побудові систем.

### Література

- 1 F. Pereira and T. Ebrahimi, A. Tesher F. Jones. The MPEG-4 BOOK
- 2 Ian E.G. Richardson. H.264 and MPEG-4 VIDEO compression. Video Coding for Next-generation Multimedia.
- 3 Recommendation ITU-T H.264. Advanced Video Coding for generic audiovisual services
- 4 Recommendation ITU-T H.263. Video Coding for low bit rate communication
- 5 Recommendation BT.1692 09-2004 Optimization of the quality of colour reproduction in television
- 6 Safdar Ali Syed Abedi “Exploring discrete cosine transform for multi-resolution analysis”
- 7 Ahumada, Albert J., Jr. Peterson, Heidi A.: “A visual detection model for DCT coefficient quantization”, AIAA Computing in Aerospace Conference, 9th, San Diego, CA, Oct. 19-21, 1993, Technical Papers. Pt. 1 (A94-11401 01-62), Washington, American Institute of Aeronautics and Astronautics, 1993, p. 314-318