Підготували студенти
групи І-34:
Шпортак Уляна
Постумент Марія
Віятик Христина
Мандебура Андрій

Необхідність стиску із втратами зображень, звуку, відео

jpeg2000

Стандарт JPEG2000 розроблений тією ж групою експертів в області фотографії, що і JPEG. Формування JPEG як міжнародного стандарту було закінчено в 1992 році. В 1997 стало ясно, що необхідний новий, більш гнучкий і потужний стандарт, який і був доопрацьований до зими 2000 року.

jpeg2000

Основні відмінності алгоритму JPEG2000 від JPEG полягають в наступному:

- Краща якість зображення при сильному ступені стиснення.
- Підтримка кодування окремих областей з кращою якістю. Тут передбачається, що мається можливість людині «на око» визначити: які галузі можна стиснути з меншим якістю, а для яких залишити колишнє. В результаті при однаковому суб'єктивному якості зображення можуть бути досягнуті більш високі ступені стиснення.
- Алгоритм заснований на вейвлет-перетворенні. З'явилася можливість поступового прояви зображення при його завантаженні по мережах зв'язку.
- Біт-орієнтоване арифметичне кодування. Арифметичне кодування передбачалося використовувати ще в стандарті JPEG, але тоді воно було захищено патентами. Зараз термін дії основного патенту закінчився і стало можливим його застосування в JPEG2000.
- Підтримка стиснення без втрат. Завдяки цій опції з'явилася можливість застосування JPEG2000 при стисканні медичних зображень, поліграфії, для задач розпізнавання тексту і т.п., критичних до втрат інформації.
- Підтримка стиснення однобітних (2-кольорових) зображень.

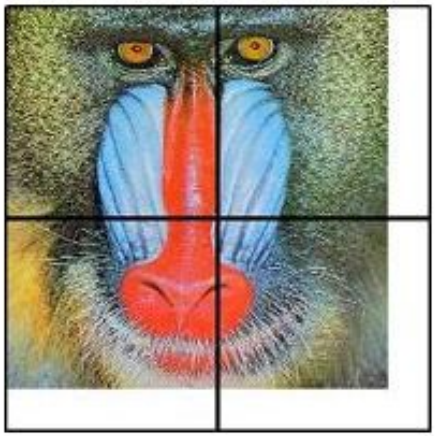
jpeg2000

В якості основного механізму компресії в JPEG2000, на відміну від JPEG, використовується хвильове (wavelet) перетворення - система фільтрів, що застосовуються до всього зображення. Відзначимо основні моменти цього принципу стиснення:

1. Спочатку точно так же, як і для JPEG, відбувається конвертація зображення з RGB в систему YCrCb.

2. Первинне видалення надлишкової інформації (шляхом вже відомого об'єднання сусідніх пікселів в блоки 2x2).

3. Потім все зображення ділиться на частини однакового розміру (tile), над кожною з яких незалежно від інших і будуть відбуватися подальші перетворення (це знижує вимоги до обсягу пам'яті і обчислювальних ресурсів).

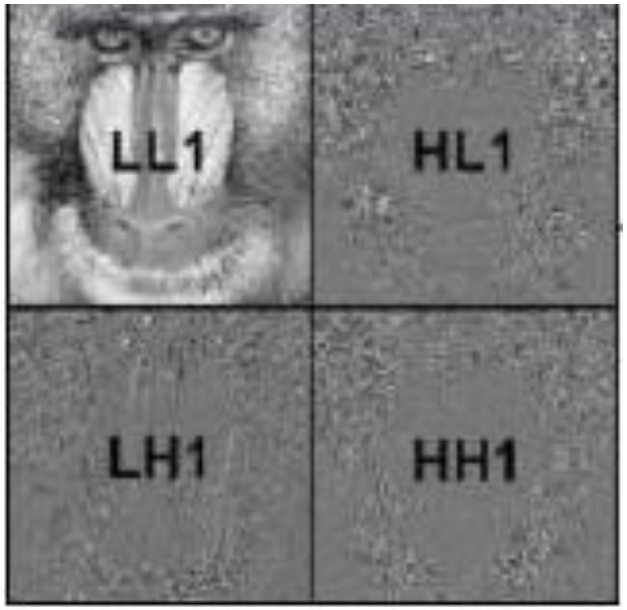


Пример разбиения изображения на тайлы.

jpeg2000

4. Далі кожен канал проходить фільтрацію низькочастотним і високочастотним фільтрами окремо по рядках і по стовпцях, в результаті чого після першого проходу в кожній частині формуються чотири дрібніших зображення (subband). Всі вони несуть інформацію про вихідне зображенні, але їх інформативність сильно відрізняється.

LL - низькі частоти по рядках і стовпчиках
HL - високі частоти по рядках і низькі за стовпцями
LH - низькі частоти по рядках і високі за стовпцями
HH - високі частоти по рядках і стовпчиках



jpeg2000

Зображення, отримане після низькочастотної фільтрації по рядках і стовпцях (угорі ліворуч), несе найбільшу кількість інформації, а отримане після високочастотної - мінімальне. Інформативність у зображень, отриманих після НЧ-фільтрації рядків і ВЧ для стовпців (і навпаки), середня.

5.Найбільш інформативне зображення знову піддається фільтрації, а отримані складові, як і при jpeg-компресії, квантуються.

Так відбувається кілька разів: для стиснення без втрат цикл зазвичай повторюється 3 рази, з втратами - розумним компромісом між розміром, якістю і швидкістю декомпресії вважається 10 ітерацій. В результаті виходить одне маленьке зображення і набір картинок з дрібними деталями, послідовно і з певною точністю відновлюють його до нормального розміру. Очевидно, що найбільша ступінь компресії виходить на великих зображеннях, оскільки можна встановити більшу кількість циклів.

Приклад

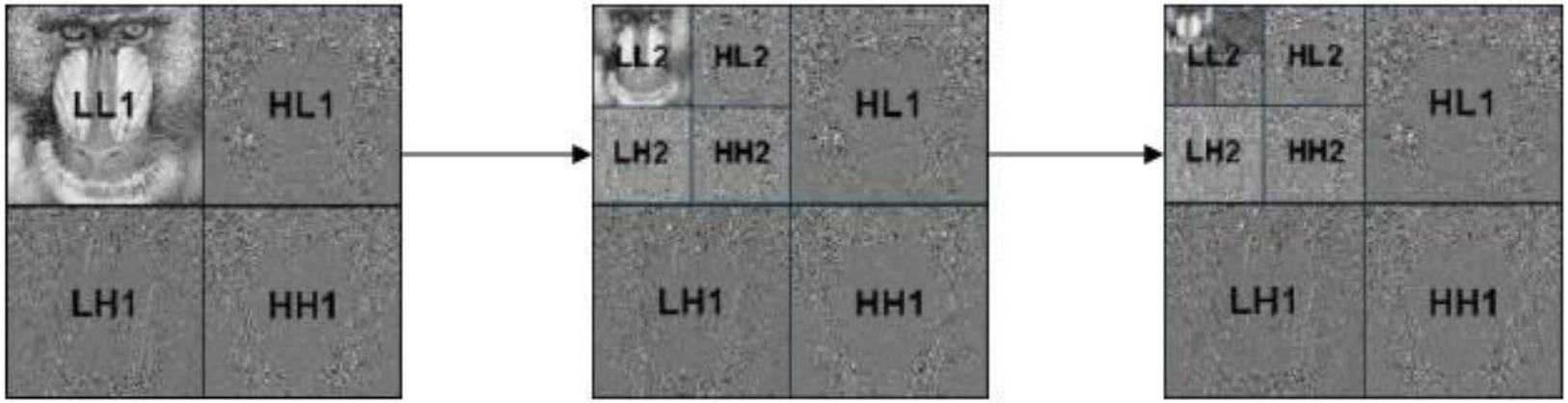
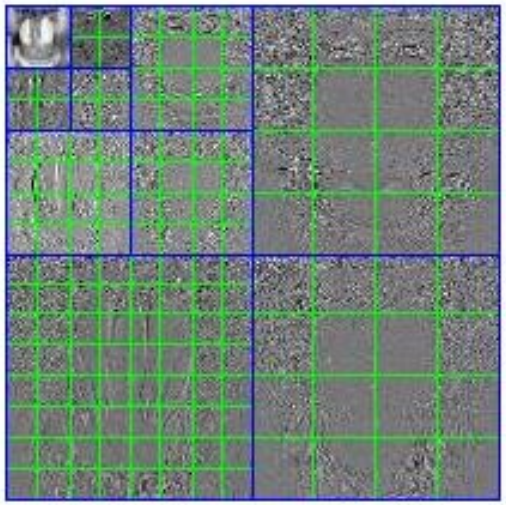


Рис. 11. Последовательное применение ДВП к изображению.

jpeg2000

Кодування отриманих округлених коефіцієнтів виконується поблочно. За стандартом JPEG2000 безпосередньо перед кодуванням фрагменти розбиваються на досить малі блоки (наприклад, розміром 32x32 або 64x64) так, щоб всі блоки одного фрагмента були однакового розміру. Розбиття на блоки виконується для того, щоб здійснити більш гнучку організацію стислій інформації для підвищення завадостійкості і так далі.



. Пример разбиения фрагментов изображения на блоки.

mp3

MP3 — ліцензований формат файлу для зберігання аудіо-інформації.

Розроблений наприкінці 1980-тих років Карлхайнц Бранденбургом, аспірантом Університету Ерлангена—Нюрнберга, Німеччина. Широко використовується в файлообмінних мережах для передачі музичних творів. Базується на теоремі Котельникова-Шеннона. Є форматом стиснення з втратами, тобто частина звукової інформації, яку вухо людини майже не сприймає, безповоротно видаляється з запису. Розробка формату була пов'язана з тим, що формат Audio-CD — дуже великий за розміром файл (1 хвилина ~ 10-15 Мб інформації), і для пересилання в інтернеті був дуже незручним. Тому треба було зменшити розмір файлу та зберегти якість звучання.

mp3

Історія

MP3 розроблений робочою групою інституту Фраунгофера MPEG (Motion Pictures Expert Group — група експертів в галузі кінематографу), на чолі з Карлгайнцом Бранденбургом. Основою розробки MP3 став експериментальний кодек ASPEC (Adaptive Spectral Perceptual Entropy Coding). Першим кодувальником у формат MP3 стала програма L3Enc, випущена влітку 1994 року. Через один рік з'явився перший програмний MP3-плеєр — Winplay3.

При розробці алгоритму тести проводилися на цілком конкретних популярних композиціях. Основною піддослідною стала пісня Сюзани Веги — Tom's Dinner. Звідси виник жарт, що «MP3 був створений виключно заради комфортного прослуховування улюбленої пісні Бранденбурга».

Опис формату

У цьому форматі звуки кодуються частотним чином (без дискретних партій); є підтримка стерео. MP3 є форматом стиснення з втратами, тобто частина звукової інформації, яку (згідно з психоакустичною моделлю) вухо людини сприйняти не може або сприймається не всіма людьми, із запису видаляється безповоротно.

Ступінь стиснення можна варіювати, зокрема в межах одного файлу. Інтервал можливих значень швидкості становить 8 — 320 кбіт/с. Для порівняння, потік даних із звичайного компакт-диска формату AUDIO-CD дорівнює 1411,2 кбіт/с при частоті дискретизації 44100 Гц.

MP3 і «якість AUDIO-CD»

Поширена думка, що запис з швидкістю 128 кбіт/с підходить для музичних творів, призначених для прослуховування більшістю людей, забезпечуючи якість звучання AUDIO-CD. Насправді все набагато складніше. По-перше, якість отриманого MP3 залежить не тільки від швидкості, але і від кодуючої програми. По-друге, крім режиму CBR (у якому, простіше кажучи, кожна секунда аудіо кодується однаковим числом бітів) існують режими ABR і VBR (у яких швидкість варіюється, забезпечуючи вищу якість звучання). По-третє, межа 128кбіт/с є умовною, оскільки її було «винайдено» в епоху становлення формату, коли на неякісних аудіокартах і колонках було практично неможливо відрізнити MP3 від оригіналу.

У наш час вважається, що звучання, яке годі відрізнити від оригінального (при правильно вибраному і налаштованому кодеку) як правило можна досягти при швидкості від 160кбіт/с і вище — залежно від початкового аудіофайлу, слухача і його аудіосистеми. Деякі аудіофіли вважають за найкраще стискати музику з «максимальною якістю» — 320кбіт/с . Проте існують семпли (фрагменти аудіозапису), які не піддаються якісному стисненню з втратами: на всіх можливих швидкостях не складає особливих труднощів відрізнити стиснене аудіо від оригіналу.

Режими кодування і опції

Існує три версії MP3 кодека для різних потреб: MPEG-1, MPEG-2 і MPEG-2.5. Відрізняються вони можливими діапазонами швидкості і частоти дискретизації:

- 32 — 320 кбіт/с при частотах дискретизації 32000 Гц, 44100 Гц і 48000 Гц для MPEG-1 Layer 3;
- 8 — 160 кбіт/с при частотах дискретизації 16000 Гц, 22050 Гц і 24000 Гц для MPEG-2 Layer 3;
- 8 — 160 кбіт/с при частотах дискретизації 8000 Гц і 11025 Гц для MPEG-2.5 Layer 3.

Режими управління кодуванням звукових каналів

Оскільки формат MP3 підтримує двохканальне кодування (стерео), існує 3 режими:

Стерео — двохканальне кодування, при якому канали кодуються незалежно один від одного. Таким чином, заданий бітрейт ділиться на два канали. Наприклад, якщо заданий бітрейт 192кбіт/с , то для кожного каналу він буде рівний тільки 96кбіт/с .

Моно — одноканальне кодування. Якщо закодувати двохканальний матеріал в цей спосіб, відмінності між каналами будуть повністю стерті, оскільки два канали змішуються в один, він кодується і він же відтворюється в обох каналах стереосистеми. Об'єднане стерео (Joint Stereo) — оптимальний спосіб двохканального кодування, при якому лівий і правий канали перетворяться в їх суму і різницю. Для більшості звукових файлів канал з різницею виходить набагато тихіше за канал з сумою, тому на суму відводиться більша частина бітрейта. Таким чином, якість вихідного файлу разюче відрізняється в кращу сторону від режиму Стерео при однаковому бітрейті, особливо при низькому.

Теги

Теги — мітки у межах Mp3-файла (на початку і в кінці). У них може бути записана інформація про авторство, альбом, році випуску і інша інформація про трек. У пізніших версіях тегів можливе зберігання обкладинок альбомів і тексти пісні. Існують різні версії тегів.

ID3 — формат метаданих. Найчастіше використовується разом з аудіоформатом MP3. Містить дані про назву трека, альбому, ім'я виконавця тощо. Ця інформація може використовуватись, наприклад, програвачами мультимедіа для відображення відомостей про трек чи автоматичного сортування. Назва «ID3» — це скорочення «Identify a MP3». Існує дві несумісних версії ID3: ID3v1 та ID3v2. Хоч ID3 розроблявся для MP3, розробникам немає перешкод вбудовувати його також і в інші формати.

ID3v1 – після створення формату MP3 виникла проблема зберігання метаданих. MP3 цього не реалізовував. У 1996 році Еріку Кемпу прийшла ідея як вирішити цю проблему: додати маленький шматочок даних до файлу. Стандарт, тепер відомий як ID3v1 швидко став стандартом де-факто зберігання метаданих у MP3-файлі.

ID3v2 – у відповідь на критику стандарту ID3v1 у 1998 році був розроблений стандарт ID3v2. Він мало схожий на першу версію.

Переваги і недоліки

MP3 є лідером за поширеністю, але при цьому не є найкращим за технічними параметрами. Наприклад, існують формати, що дозволяють отримати порівнянну якість (суб'єктивно) при меншій щільності. Також у форматі MP3 відсутній режим кодування без втрат придатний для професіоналів.

MP3 непридатний для професійного використання музикантами вже через те, що дані стискаються з втратами, і при кожному редагуванні файлу якість погіршується. При цьому формат цілком підходить (з професійної точки зору) для розповсюдження демонстраційних композицій або інших способів «роздачі» своєї музики завдяки повсюдній поширеності програвачів.

Формат розвивається і зараз вже є формат MP3 Pro — більш вдосконалений, який підтримує технологію Dolby Surround.

Sorenson Media

Sorenson Media була створена в грудні 1995 року через ринковий попит на онлайн медіа-контенті, розвиваючи інноваційні, економічно ефективної технології кодування відео, що значно скоротити вимоги до пропускної здатності при збереженні високої якості відео. Оригінальна назва Sorenson Vision. Компанія вперше представила свою “codec" (інструмент для стиснення і декомпресії) в попередньому перегляді розробника на MacWorld Expo в січні 1997 року.

Sorenson Media грає важливу роль в залученні Інтернет відео в QuickTime і пов'язаних додатків на платформах Windows, і Macintosh та в зв'язку із їх ліцензійною угодою з Apple, були створені додатки для покращення досвіду онлайн відео-творців контенту, менеджерів і споживачів.

З моменту свого випуску, відео технологія кодування Sorenson Media була використана у сайтах компанії Apple Computer і фільмів для студій, таких як Disney, Lucas Film, MGM і Paramount і музичних кліпів від iTunes до переходу на стандартний формат H.264.

Продукти

- Sorenson Video - Apple, в комплекті QuickTime 3 із компонентами для декодування і основний кодування;
- Sorenson Video Pro - Apple, поширений, повнофункціональний, Двопрохідний VBR QuickTime компонент кодування для розробників;
- Sorenson MPEG-1/2;
- Sorenson MPEG-4 Part 24;
- Sorenson H.264 ;
- Sorenson Squeeze - для Windows і Macintosh повнофункціональний додаток кодування;
- Sorenson 360 - онлайн відео платформа (OVP);
- Squeeze Server - сервер на базі живого кодування;
- Sorenson Server 2;
- Sorenson Spark.

Технології кодування

Sorenson codec може відноситися до будь-якого з трьох фірмових codecs:

- Sorenson Video
- Sorenson Video 3
- Sorenson Spark

Для обробки відео-та аудіо, а саме стиснення і розпакування, будь-яке програмне забезпечення (і не тільки Flash) використовує спеціальний модуль, званий кодеком (codec). Кодек Flash (Spark video codec) складається фактично з двох підмодулів: кодера і декодера. Кодер (або компресор) забезпечує стиснення даних, а декодер (або декомпресор) виконує їх розпакування. Flash Player містить тільки декодер.

Sorenson Spark

Sorenson Spark - це включений в Flash MX кодер / декодер відео, який дає можливість додавати в документи Flash відеовміст. Цей високоякісний кодер / декодер дозволяє стиснути відеодані досить сильно для того, щоб фільм Flash можна було без особливих проблем передавати по повільних каналах зв'язку. І, в той же час, Sorenson забезпечує досить високу якість зображення.

Включення в пакет Flash алгоритму Sorenson Spark значно розширює і без того багаті можливості програми. Адже в попередніх версіях Flash можна було моделювати відео, тільки використовуючи послідовність растрових зображень.

Зараз доступні дві версії Flash Sorenson: Sorenson Spark Standard Edition, увімкнений в Flash MX, і Flash Player 6. Особливо гарна якість Spark Standard edition забезпечує для відео з незначними рухами, типу зображення переконливої людини.

Про стиск

Є два типи стиснення цифрових відеоданих: spatial (просторовий) і temporal (тимчасовий).

- Temporal-стиск ідентифікує відмінності між кадрами і запам'ятовує їх так, щоб кадри описувалися на підставі їх відмінності від попереднього кадру. Незмінні області просто повторюються з попереднього кадру. Temporal-стиск часто називають interframe.
- Spatial-стиск застосовується до одного кадру, незалежно від оточуючих кадрів. Spatial-стиск може бути lossless (при цьому дані з зображення не відкидаються) або lossy (коли дані вибірково відкидаються). Стиснутий таким чином кадр часто називають intraframe.

Sorenson Flash - по суті interframe кодер / декодер. Ефективний interframe-стиск Sorenson Spark вигідно відрізняє його від інших технологій стиснення, так як вимагає набагато більш низьку швидкість передачі даних, ніж у більшості інших кодер / декодерів і забезпечує при цьому гарну якість відео. Багато інших кодер / декодерів використовують intraframe-стиск; наприклад, JPEG - intraframe кодер / декодер.

Поради для створення Flash-відео з стисненням Flash Sorenson

Те як стискається відео, в значній мірі залежить від його змісту. Особливості сучасних алгоритмів стиснення відео (в тому числі, і Sorenson Spark) такі, що малорухомі сцени стискаються зовсім не так, як динамічні, швидко мінливі сюжети. Тепер кілька порад:

- Уникайте сцен зі складними рухами та змінами в кадрах, вони дуже погано стискаються. Зазвичай найкращим рішенням є або швидка зміна сюжетів або плавні переходи.
- Коли відео призначене для Інтернету, файли повинні бути розраховані на більш низькі швидкості передачі даних, ніж для внутрішніх мереж.
- Якщо є кліп з динамічними сюжетами, більш низька швидкість кадрів може поліпшити відтворення на кінцевих комп'ютерах.
- Коли для відео вибирається розмір кадру, слід одночасно враховувати швидкість кадрів, початковий матеріал і персональні переваги.
- Стандартні розміри кліпу:
  - Модеми: 160 x 120
  - ISDN: 192 x 144
  - Швидкодіючі канали: T1 / DSL / Cable: 320 x 240

Поради для створення Flash-відео з стисненням Flash Sorenson

- Під час завантаження відеокліпу непогано було б використовувати інше зображення, яке з'являється і "маскує" завантаження. У розрахунках для короткого кліпу використовуйте наступну формулу: Пауза = час завантаження - час програвання + 10% часу програвання. Наприклад, якщо кліп - 30 секунд і завантаження займає одну хвилину, потрібно дати кліпу 33-секундний буфер: 60 секунд - 30 секунд + 3 секунди = 33 секунди.
- Чим вище якість оригіналу, тим краще кінцевий фільм.
- Щоб досягти хорошого звукового стиснення, потрібно починати з чистого звуку. Якщо кодується матеріал з CD, спробуйте записувати файл, використовуючи пряму цифрову передачу замість аналогового вводу звукової плати.

:)

Дякуємо за увагу.