

6. Мултимедия и поточни данни.

Васил Георгиев



`is.fmi.uni-sofia.bg/t3/`



`v.georgiev@fmi.uni-sofia.bg`

Съдържание

- Компресия и декомпресия на статични и динамични изображения и на звук
- Пренос на поточни данни
- Системи и приложения за интерпретация на мултимедия

Мултимедия и компресия

- ✦ Мултимедия – комбинирано съхранение, пренос и интерпретация на текст, графика, образи, анимация, кино, звук и други феномени, които служат като посредник между абстрактната информация и човешките възприятия
- ✦ компресията и декомпресията е представяне на данните в компактна форма с цел пренос или съхранение и последващото им възстановяване с уместна точност
- ✦ мултимедийните данни се разглеждат като последователност знаци от крайна дискретна азбука (т.е. азбука, за чиито елементи съществува целочислово представяне)
- ✦ пример: данните за статично изображение са набор от масиви (по един за всеки съставен цвят), чиито стойности са интензивността на цвета в точка, съответстваща на пространственото положение на фрагмент от изображението – “пиксел”

Първични ММ данни (изображение)

- ✦ статично изображение

- ✦ форматиран текст

- ✦ страница A4 стандартен обем $80 \text{ символа} * 64 \text{ реда} * 16 \text{ b} = 84 \text{ Kb}$
 $= 5.7 \text{ S}$ (14.4 Kb/S модем)

- ✦ графика – набор човеко- или машинно-генерирани обекти

- ✦ подлежат на манипулация – избор, местене, въртене, изтриване...

- ✦ образ – RGB кадър, 8b/цвят, 24b/пиксел

- ✦ B/W страница A4 стандартен печат (600 dpi) $> 100 \text{ MB}$

- ✦ $1024 * 768 * 24 \text{ b/p} = 18.8 \text{ Mb} = 21 \text{ min}$ (14.4 Kb/S)

- ✦ динамично изображение

- ✦ анимация (подвижна графика)

- ✦ цифрово кино (подвижни образи) $4\text{K} * 2\text{K} * 3 \text{ (цвята)} * 12 \text{ b}$
 $= 48 \text{ MB/frame}$, 1.15 GB/S, 69GB/min

Първични ММ данни (звук)

✦ звукови формати

формат	сканиране (квантуване), kHz	честотна лента, kHz	честотен обхват на звука, Hz	побитов темп, Kbps
телефон	8.0	3.0	200 - 3200	64
телекон- ференция	16.0	7.0	50 - 7000	256
CD	44.1	20.0	20 - 20000	1410

Предпоставки за компресията

- статистически излишък от периодически и аperiodично повторение или подобие на данните
 - пространствена корелация
 - локална – повтаряемост на съседни пиксели
 - глобална – повтаряемост на части от изображения
 - спектрална корелация – между интензивностите на базовите цветове
 - времева корелация – между последователни кадри
- допуск (tolerance)
 - излишък при възприемане
 - ограничения в производителността

Имплементации на компресията

- апаратна имплементация със специализирани логически схеми – ASIC (Application Specific Integrated Circuit)
 - сложно и скъпо проектиране, тясно специализирано приложение
- програмна имплементация на общ процесор
 - ниска или негарантирана скорост на обработка
- апаратно-програмна имплементация на DSP- или VSP-чип
 - хибридни свойства
- подпроцеси:
 - сканиране → квантуване → компресия → пренос/запис → декомпресия → интерпретация

Имплементации на компресията

- апаратна имплементация със специализирани логически схеми – ASIC (Application Specific Integrated Circuit)
 - сложно и скъпо проектиране, тясно специализирано приложение
- програмна имплементация на общ процесор
 - ниска или негарантирана скорост на обработка
- апаратно-програмна имплементация на DSP- или VSP-чип
 - хибридни свойства
- подпроцеси:
 - сканиране → квантуване → компресия → пренос/запис → декомпресия → интерпретация

Сканиране и квантуване

- ✦ честота (темп) на сканиране – sampling
- ✦ квантуване – ADC/DAC (codec)
- ✦ ADC: [периодично – от sampling] цифрово (т.е. крайно-точно) отчитане на параметъра на [електрически] сигнал
 - ✦ за ADC при MM – медийният феномен (образ/звук) предварително се преобразува до [ел.] сигнал
- ✦ DAC е необходим само при възпроизвеждане (не при съхранение и пренос)
- ✦ освен темпа осн. х-ка е точността (напр. резолюция, точност на кодиране на пиксел) – определя заетата памет, необходимата скорост на пренос и скоростта на конверсия (codec-устройство) – табл. 6.9
- ✦ теорема на Nyquist: при периодичните сигнали темпа трябва да е поне 2x честотата на сигнала – напр. HiFi звук (20Hz – 20kHz) → темп 40kHz

Точна компресия на изображения

- ✦ базира се на информационен излишък (контекстнозависим)
- ✦ точна компресия – без загуби след декомпресия
 - ✦ предпочита се при компютърно- или човеко-генерирани изображения (напр. чертежи) или при отговорни данни (архивиране, томография)
 - ✦ методи
 - ✦ RLE (Run-length encoding) – коефициенти за повторение на байтове или думи; подходящ за палитрови изображения (неголям брой цветове) – напр. икони; прилага се във форматите/стандартите BMP, TIFF, JPEG (вариант със загуби)
 - ✦ ентропно кодиране – променлива дължина на кодовете на отделните символи от входната азбука, като най-повтаряните символи се означават с най-къси кодове (напр. 1-2 бита), а за редките символи се ползват дълги кодове
 - ✦ LZW (Lempel-Ziv-Welch) – базира се на динамична (за всяка входна последователност) транслираща таблица за двойките <думи (низове), кодове с фиксирана дължина (напр. 12 бита)>; думите обикновено са от два символа; при следващо възникване на същата дума, вместо нея се записва кода от таблицата. Ако се изчерпят всички комбинации от кодове с тази дължина, тогава дължината се увеличава с нов старши бит; прилага се в GIF

Наподобяваща компресия на изображения

- наподобяващата компресия е със загуби след декомпресия
- предпочита се при дигитализирани фотографии като се разчита на ограничените възможности на окото
- методи:
 - цветова редукция – съкращаване на палитрата на дискретните цветове със запазване на най-масовите (в главата на файла), към които останалите са апроксимират; често се комбинира с размазване на ръбовете при декодиране (при което се връща по-сложна палитра); скъсява кодирането на всяка точка
 - цветово подтискане (chroma subsampling) – по-пълно кодиране на яркостта на точките, отколкото на цвета, видеоформати и JPEG
 - трансформации – използва се фуриерова трансформация – напр. DCT (Discrete Cosign Transform) за да филтрира част от входните данни (експериментално доказано по-несъществени за възприемането на образа), след което се прилага ентропно кодиране (JPEG)

Компресия на видео

- видеоинформацията е 3D масив от “кадри” – двумерни изображения във времева дименсия
- при видеокомпресията освен пространствено-цветови излишък се използва и времеви излишък
- не се прилагат методи за точна компресия
- кадрова компресия (не използва времевата дименсия)
- междукадрови методи (основен метод при видео)
 - непроменените зони от последователни кадри се копират или се подлагат на елементарни линейни операции
 - срещу загуба на базов кадър (която проваля декомпресията на останалите кадри) периодично се прилага кадрова компресия без копиране от предишни кадри
- при компресията се прилага предимно DCT

Компресия на звук

- със (50-60% редукция) или без загуби (5-20% редукция) както при изображенията
- звуковите поточни данни са по-хаотични (неповторими) и динамични от визуалните
- компресията се базира на определяне на звуковия спектър (коллекция честоти, които се възприемат от ухото при даден контекст) – както определянето на базовата палитра
- ползва се модифицирана DCT, като се определя и отделен тегловен шаблон за времевата скала и друг – за честотната скала
- звуковите поточни данни се представят като последователност (във времето) от вектори със стойности на интензивността за базовите честоти от спектъра
- речевото кодиране се различава от звуковото кодиране и води до значително по-голяма компресия
- звуковото кодиране се оценява по възприеманото качество, коеф. на редукция, скорост на операцията и възможност за РВ-пренос (при интерактивна мултимедия)

Стандарти за наподобяващо компресиране на изображения - JPEG и GIF

- стандартизиран се процедурите в двете посоки (codec) и файловете формати
- JPEG (Joint Photographic Experts Group) – компресия на изображения с настройваемо ниво на загуби
 - може да се вгражда в други формати (напр. markup, а също в други графични формати като TIFF)
- GIF (Graphics Interchange Format) – редуцира 24-битова палитра на изображението до 8-битова: 256 цвята на кадър, върху който се прилага LZW; подходяща предимно за малки или регулярни изображения
 - поддържа кратка анимация (серия от свързани изображения)
 - поддържа компресиране на набор от изображения в един файл, при което чрез управляваща информация се линии от изображенията се записват и възпроизвеждат по специален ред, който позволява частично възстановяване преди пълното зареждане/декодиране на файла
 - модификация на GIF е true color GIF – изображението се разделя на малки области, за които се прилага GIF, но за цялото кодиране точността е близка до 24-битовата

Стандарти за точно компресиране на изображения - BMP, PNG и TIFF

- ✦ BMP – пикселно кодиране на Microsoft, не се извършва компресия, а само дискретизация до фиксирана палитра
 - ✦ палитри: 2 цвята (1 bpp), 16 цв. (4 bpp), 256 цв. или сива палитра (8 bpp), 65536 цв. (16 bpp) или 16,7m цв. (24 bpp)
 - ✦ на практика не е подходящ за пренос на изображение
- ✦ PNG (Portable Network Graphics) – 24-битово (true color) кодиране на пикселите (RGB или сива палитра) за обмен в Интернет
 - ✦ файлът се състои от глава и управляващи полета (chunks), с които се задава състава на палитрата, едно поле или няколко полета за данни (за постепенно зареждане на образа)
 - ✦ освен цветовете запазва контраста на преходите и е подходящ за редактиране на изображението (поддържа се от графичните редактори)
 - ✦ $800 \times 600 \times 24b \rightarrow 1.4 \text{ MB}$ в BMP и типично между 80 и 250 KB в PNG.
- ✦ TIFF (Tagged Image File Format) – настройваем фолмат, подходящ за принтерни файлове (PS) и за графични редактори
 - ✦ в главата на файла се съдържат марките на включените файлове, вк. и използваната компресия за всеки от тях – напр. JPEG
 - ✦ подходящ също за архивиране на документи

Стандартно компресиране на видео - MPEG-2

- MPEG-2 (Motion Pictures Expert Group) – масов стандарт за съвместно кодиране на видео и звук при цифров запис и предаване в мрежа или ОТА (наземно или спътниково)
 - инкорпорира два формата – транспортен поток (за AV данните) и програмен поток (за управляваща информация когато се предава от надеждна и бърза медия като оптичните носители – за интерактивни настройки напр. FF)
 - оптимизиран за скорост 3Mb/S и неприложим при <1Mb/S
 - прилага видеокомпресия (пространствена и времева – т.е. междукадрови методи), оптимизирана за скорост 50 - 60 fps (TV стандарти в Европа и САЩ)
 - един от методите за компресия на видео тук е последователно предаване или запис само на четните и само на нечетните редове от кадъра и смесването им при възпроизвеждане (поради това при няколко презаписа качеството чувствително пада)
 - звуковото кодиране се базира на разделяне на честотния спектър на слоеве и на сканиране на отделните слоеве с различна (намаляваща) честота

Стандартно компресиране на звук - MP3

- MP3 (търговско название на MPEG-1 Audio Layer 3) – специализиран стандарт за наподобяваща но реалистична звукова компресия
 - честотата на сканиране (sampling) на звука съответства на HiFi: 32, 44.1 и 48 kHz
 - скорост на обмен: 32, 40, ... 128, 160, ... 256 и 320 kbit/s
 - и двата параметъра са настройваеми
 - при запис се отчита характера на звуковата картина
 - при пренос и интерпретиране се отчитата скоростта на обмен в канала
 - информацията е структурирана на независими кадри – глава и данни
 - кадрите с данни съдържат настройваемо 384, 576 или 1152 последователни моментни стойности на звуковия спектър (samples)
 - отпадането на даден кадър данни не блокира декодирането
 - стерео-звук се синхронизира “кадър към кадър” от двата потока

ММ на оптичен носител - DVD

- DVD (Digital Versatile Disc) е формат за съхранение на големи по обем данни на оптичен носител
 - поради структурирането на данните е възможен запис на AV, само на аудио или само на числови данни, както и комбинации
 - при AV данни използва редуциран MPEG-2
 - задава следните режими
 - формат на кадрите - 720 × 480, 704 × 480, 352 × 480, 352 × 240 пиксела
 - пропорция - 4:3 и 16:9
 - темп 29.97 fps (NTSC – САЩ) и 25 fps (PAL - Европа)
 - прилага защитен алгоритъм CPPM (Content Protection for Prerecorded Media – IBM, Intel, Matsushita, Toshiba)
- устройствата за запис на оптичен носител поддържат и други стандарти като DVD+/-R/RW, CD-R/RW, MP3, JPEG, PNG, MPEG-4 (DivX/Xvid)

Пренос на поточни данни

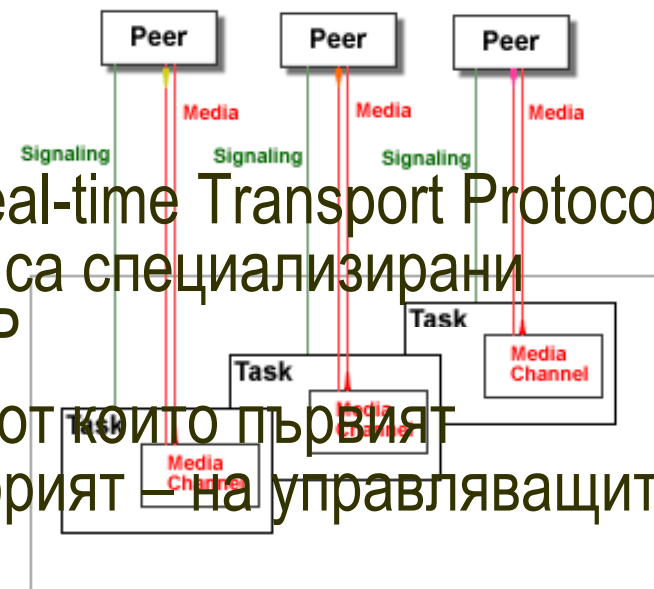
- Освен с компресия (за редуциране на обема) мултимедията се асоциира с поточни данни –
 - комбинирано съхранение и
 - синхронна интерпретация и пренос на няколко типа медия
- преноса на качествена мултимедия изисква RV предаване в изосинхронен режим – дефинирани са както максимално така и минимално закъснение на последователните пакети данни
- от протоколите за общи комуникации поточни данни се предават с основно с UDP, възможностите на TCP са ограничени

UDP и TCP

- ✦ UDP – данните се предават като последователност от къси пакети
 - ✦ ефикасност в надеждни мрежи поради малкия комуникационен свръхтовар
 - ✦ в мрежи със загуби недоставените пакети се изоставят (не се препредават)
 - ✦ при загуба на пакети е възможно прилагането на методи за корекция на грешки
 - ✦ TCP – гарантирано предаване на всеки бит от поточните данни, но ограничено приложение в MM поради високия свръхтовар и противоречието с PB
 - ✦ система от срокове и препредаване на грешните данни
 - ✦ при закъснения TCP изисква преустановяване на потока MM данни и препредаване на загубените или грешни пакети
 - ✦ някои MM системи ползват TCP но в режим на предварително буфериране на данните
- ✎ N.B.: защитните стени обикновено се конфигурират да блокират UDP-базиран обмен, но не и TCP-обмен

РВ предаване на ММ

- RTSP (Real-time Streaming Protocol), RTP (Real-time Transport Protocol) и RTCP (Real-time Transport Control Protocol) са специализирани протоколи за ММ поточни данни – върху UDP
- RTP и RTCP са комплементарни протоколи, от които първият обслужва предаването на поточна ММ, а вторият – на управляващите обмена метаданни – фиг. 6.21
- RTCP периодично настройва сървера като изпраща контролни пакети, а междувременно сканира поточните данни относно QoS:
 - брой приети байтове и пакети, брой изоставени пакети, десинхронизиране (jitter), закъснение на предаването (roundtrip)
- RTP е фокусиран само на обмена на ММ поточни данни
 - стандартизира формата на видео- и аудио-пакетите
 - определя типа на данните, съответно допуските и загубите
 - номерирането на пакетите и формата на времевите марки
 - единично или групово предаване (unicast/multicast)
 - поддържа и upload – “push-to-talk”, което го прави удобен за интерфейс на телеконференентни приложения и за някои функции на RTSP



RTSP

- RTSP позволява отдалечен контрол на сървера, предаващ поточни ММ данни
 - ползва интерфейсите примитиви на RTP
 - възможно е конфигуриране и за други протоколни стекове
 - обмена е в режим на локално съхранена сесия (с ID на сесията) – няма нужда от TCP, но е statefull протокол
 - имплементира командите DESCRIBE (заявка за описание на поддържаните потоци от сървера – примерен отговор един видео и два аудио, по специфицирани стандарти), SETUP (задава URL и порт на приемащия данните процес) PLAY, PAUSE, RECORD (заявка за предаване на поточни данни към сървера) и TEARDOWN (прекратяване на потока)
- пример: поддържани стекови конфигурации на QuickTime поточният сървер на Apple:
 - RTSP върху RTP (RTP върху UDP)
 - RTSP/RTP вложен (tunneling) в HTTP
 - RTSP върху TCP
 - RTP върху UDP
 - RTP върху Reliable UDP (Apple)

Windows Media Player

- ▶ самостоятелно приложение е библиотека на MS за представяне на AV и изображения
- ▶ режимите на обмен поддържат локален и отдалечен източник и прогресивно зареждане (буфериране)
- ▶ изглаждане на изображението чрез вмъкване на интерполирани кадри
- ▶ изисква стандартна MS-OC/Mac/Solaris
- ▶ има версии за мобилни персонални устройства Pocket PC / Windows Mobile edition
 - ▶ 640*480 пиксела на кадър и стерео звук