## Lưu trữ và truyền tải video:

Để hiểu về các định dạng lưu trữ video, ta cần tìm hiểu 02 khái niệm: Tập tin chứa dữ liệu dữ liệu đa phương tiện (Multimedia container formats) và Thuật toán mã hóa / giải mã (Codec).

### Tập tin chứa dữ liệu dữ liệu đa phương tiện (Multimedia container formats):

Ta tạm gọi các tập tin chứa nội dung đa phương tiện là các container. Các container chứa thông tin nhận diện về codec và dữ liệu xen kẽ giữa nhiều định dạng media khác nhau (tùy vào đặc tả của mỗi container). Các loại container đơn giản có thể chứa nhiều định dạng âm thanh khác nhau, trong khi các định dạng container nâng cao có thể chứa nhiều luồng (stream) âm thanh, video, phụ đề, thông tin các chương, và meta-data (tags) — kèm theo đó là thông tin để đồng bộ các luồng dữ liệu nhằm mục đích hiển thị, chơi cùng lúc các luồng đó. Các định dạng container được thiết kế riêng cho các mục đích khác nhau, các định dạng dành cho truyền dữ liệu qua internet sẽ khác với các định dạng chất lượng cao v.v…

Trong mỗi container, ta có các gói (packet) dữ liệu để chứa dữ liệu. Các định dạng khác nhau gọi tên các gói đó khác nhau, trong RIFF và PNG là "chunks, trong QuickTime/MP4 là "atoms", trong MPEG-TS là "packets" và là "segments" trong JPEG. Nội dung bên trong các gói dữ liệu đó được gọi là "data" hay "payload". Hầu hết các định dạng container lưu trữ các gói trong sequence, mỗi gói có header riêng, ngoại trừ TIFF lưu theo offsets. Cách lưu trữ này giúp chúng ta dễ dàng khôi phục trong các trường hợp lỗi file, mất hình (drop frame),…

Một vài container phổ biến đã được đăng ký riêng cho định dạng âm thanh:

* AIFF (định dạng IFF, sử dụng trên nền tảng Mac OS)
* WAV (định dạng RIFF, sử dụng trên nền tảng Windows)
* XMF (Định dạng âm thanh mở rộng)

Một và container khác thì được đăng ký riêng cho hình ảnh:

* FITS (Flexible Image Transport System) hình ảnh tĩnh, dữ liệu thô và metadata.
* TIFF (Tagged Image File Format) hình ảnh tĩnh và metadata.

Các container có thể chứa nhiều loại âm thanh, video và các loại dữ liệu đa phương tiện khác phổ biến hiện nay (Multimedia file container) như:

* 3GP (Sử dụng trên hầu hết các thiết bị di động; xây dựng dựa vào chuẩn ISO)
* ASF (được dùng để chứa 2 định dạng là WMA và WMV của Microsoft, tuy nhiên, ngày nay rất ít được dùng vì WMV đã có thể “tự chứa” được WMA)
* AVI (định dạng container chuẩn của Microsoft, xây dựng trên RIFF)
* DVR-MS ("Microsoft Digital Video Recording", được phát triển bới Microsoft, dựa trên ASF)
* Flash Video (FLV, F4V) (chứa video và audio của hãng Adobe Systems)
* IFF (định dạng container độc lập với nền tảng đầu tiên)
* Matroska (MKV) (Định dạng này là một chuẩn mở (open standard), mã nguồn mở, có thể chứa không giới hạn các định dạng media khác).
* MJ2 – Định dạng Motion JPEG 2000 xây dựng dựa vào chuẩn ISO về file media được đặc tả qua MPEG-4 Part 12 và JPEG 2000 Part 12
* QuickTime File Format (chuẩn video QuickTime của Apple Inc.)
* MPEG program stream (chuẩn container cho MPEG-1 và MPEG-2; thường sử dụng trên các DVD-Video hay VCD)
* MPEG-2 transport stream (còn gọi là MPEG-TS) (chuẩn container cho các định dạng chất lượng cao, có thể chứa nhiều luồng dữ liệu, chương trình điều khiển,… được sử dụng trên các đĩa Blue-ray hiện nay)
* MP4 (chuẩn container cho các định dạng đa phương tiện MPEG-4, đặc tả trong MPEG-4 Part 12 và JPEG 2000 Part 12).
* Ogg (chuẩn container âm thanh Vorbis của Xiph.org và định dạng video Theora)
* RM (RealMedia; chuẩn container dành cho RealVideo và RealAudio)
* WebM (Chuẩn container mở, xây dựng trên MKV, được hiện thực cho nền tảng HTML5)

Ngoài ra, còn rất nhiều định dạng container phổ biến khác như NUT, MXF, GXF, ratDVD, SVI, VOB và DivX,… mà trong giới hạn của đồ án, chúng tôi không thể giới thiệu hết.

### Thuật toán mã hóa (Codec = Compressor-Decompressor):

* 1. Codec là gì?

Tổng quát, Codec là một thiết bị phần cứng hoặc là một chương trình phần mềm cho phép mã hóa và giải mã các luồng dữ liệu số hoặc là tín hiệu. Codec là từ được kết hợptừ Compressor-Decompressor, hoặc là Coder-Decoder. Về mặt phần cứng, có thể lấy ví dụ chip xử lý âm thanh là một minh họa về codec. Codec phần cứng được thực hiện bằng các mạch điện tử được thiết kế chuyên biệt nên được thực hiện rất nhanh, tuy nhiên lại kém linh hoạt. Ta không thể sử dụng card âm thanh để xử lý…video được (!). Còn codec phần mềm được thực hiện bằng “công cụ vạn năng” là CPU, nên hết sức mềm dẻo và linh hoạt. Tuy nhiên, xử lý bằng phần mềm không thể nhanh bằng phần cứng, nhưng với tốc độ của các CPU hiện nay, hầu như không có loại Codec nào có thể “làm khó” máy tính của bạn.

* + 1. Tại sao phải sử dụng Codec?

Codec thực tế có nhiều ứng dụng, tuy nhiên, ở đây chúng ta chỉ xét trong lĩnh vực đa phương tiện – multimedia. Hẳn chúng ta ai cũng đã biết, các file multimedia nguyên gốc rất lớn. Lấy ví dụ về âm thanh. Một filewav (âm thanh chưa nén) chuẩn (lấy mẫu 16-bit ở tần số 44.1Khz) với độ dài khoảng 5 phút sẽ có dung lượng vào khoảng 50MB. Với video, con số này sẽ “khủng khiếp” hơn rất nhiều vì mỗi giây phải lưu giữ ít nhất 24 khung hình, mà mỗi khung hình, nếu theo chuẩn BITMAP chưa nén sẽ là 2,6MB cho kích thước HD-720p. Điều này gây ra nhiều khó khăn trong việc lưu trữ cũng như phát hành các file đa phương tiện qua mạng. Codec chính là giải pháp tuyệt vời cho trường hợp này. Với Codec, bạn có thể thu nhỏ các file multimedia xuống nhiều lần mà chất lượng vẫn giữ nguyên hoặc chỉ suy giảm chút ít. Chẳng hạn, với file âm thanh ở trên, nếu được mã hóa dưới dạng flac (một chuẩn âm thanh) thì dung lượng chỉ còn 20-25MB. Nếu chấp nhận bỏ bớt 1 chút chất lượng âm thanh, có thể dùng định dạng MP3 để thu nhỏ xuống còn 10 MB (với bitrate 320kbps) hay thậm chí hơn nữa… Để chơi được các file multimedia, player phải thực hiện 2 công việc: đọc được header của file để nhận diện xem đó file đó thuộc loại nào (video/audio), sử dụng chuẩn mã hóa nào (DivX/Xvid, x264 hay Quicktimes,…). Và sau đó là giải mã nội dung thật sự của file để đưa ra trình diễn.

### Các loại codec video:

Có khá nhiều loại codec video khác nhau. Trong quy mô đồ án, chúng tôi sẽ chỉ giới thiệu các codec phổ biến nhất, nổi bật nhất hiện nay. Cụ thể:

* 1. MPEG-4

MPEG-4 là một chuẩn nén video được phát triển bởi một nhóm gọi là “Moving Picture Experts Group” (hay gọi là MPEG). MPEG-4 được biết đến như là một “chuẩn nén của hình ảnh & âm thanh bit-rate thấp”. Có rất nhiều chuẩn được phát triển bởi nhóm này và đã được chấp nhận bởi tố chức tiêu chuẩn quốc tế (ISO – International Standards Organization) hay còn gọi là chuẩn ISO.

MPEG-4 là chuẩn-ISO/IEC #14496. Vì XviD đi theo chuẩn MPEG-2 part 2 nên nó còn được biết đến như là MPEG-4 chuẩn ISO #14496-2 (chuẩn mới nhất của MPEG-4 đó là MPEG-4 part 10/AVC (Advanced Video Coding) hay còn gọi là H.264).

Thực tế, MPEG-4 là một chuẩn nén video với rất nhiều các phần mở rộng mà đặc biệt được thiết kế để đạt tới chất lượng cao nhất của chuẩn nội dung video. Chuẩn nội dung ở đây là chuẩn các video trong thế giới thật, nó không được thiết kế để nén video dựng bởi các trình 3D, phim hoạt hình trên TV thông thường hay là anime (đó là lý do tại sao các video thuộc thể loại trên nén không tốt lắm khi sử dụng MPEG-4).

Các chuẩn codec thông thường thuộc họ MPEG-4 là DivX, 3ivx, Quicktime MPEG4 và XviD. Mặc dù chúng là họ hàng với nhau, tuy nhiên một phim được nén bởi một trong các codec trên sẽ không thể được giải mã bằng codec anh em của nó. Vì giữa chúng vẫn có những điểm khác nhau cơ bản không thể thay thế. Lấy ví dụ bằng việc xử lý tập tin video có nhiều hơn 1 tỷ frame – đó là việc mà DivX không thể thực hiện được, trong khi XviD lại có thể. Ví dụ khác sẽ là XviD có đến 3 điểm warppoint GMC (Global Motion Compensation) trong khi DivX chỉ có 1 điểm – điều này dẫn đến sự khác biệt trong cấu trúc phần cứng để có thể giải mã được đoạn video được nén bằng codec DivX hay XviD (phần cứng để giải mã DivX sẽ nhẹ nhàng hơn).

Có một số đầu đọc DVD phù hợp chuẩn MPEG-4 có thể đọc được video DivX. Để mở rộng khả năng đọc các chuẩn MPEG-4 khác sẽ không khó khăn mấy. Việc này chỉ cần được thực hiện bằng một số thay đổi nhỏ trong EEPROM liên kết với bộ xử lý video của đầu đọc – hay nói cách khác là flash lại bộ EEPROM của đầu đọc.

Kỹ hơn các lý thuyết cơ bản về MPEG-4, để có thể nén được 2h video chất lượng cao mà chỉ nhỏ để đút vừa một đĩa CD đòi hỏi cả một thuật toán phức tạp. MPEG-4 làm việc này bằng cách bỏ bớt những thông tin mà người xem không cảm nhận được bằng cách biến đổi dữ liệu về các điểm ảnh trên video thành công thức toán học gần giống với đoạn dữ liệu đó. Việc biến đổi này gần giống nhau đến mức con người bình thường sẽ rất khó có thể nhận ra sự khác biệt giữa video nguồn và video kết quả (tất nhiên, đó chỉ là lý thuyết).

Đầu tiên, mảng màu của khung hình ban đầu sẽ được biến đổi thành một mảng màu đặc biệt được gọi là YV12. Mắt người thông thường kém nhạy cảm với các màu sắc hơn là với các mức độ sáng tối (liếc nhìn thật nhanh qua bầu trời, bạn sẽ cảm nhận rõ nhất trời sáng hay trời tối hơn là bầu trời có màu gì). Chính vì vậy, MPEG-4 đánh giá độ sáng tối của hình ảnh quan trọng hơn là màu sắc. Điều này đem đến một kết quả là độ sáng tối được ghi lại trong mỗi pixel còn thông tin màu sắc sẽ chỉ được ghi lại sau mỗi 4 pixel.

Trong các codec phổ biến hiện nay, có thể kể đến DivX, XviD và H.264 Ngoài ra còn một số đối thủ cũng khá cạnh tranh khác (có điều cung cách phát triển hơi ích kỷ) là Real Media và Window Media Audio/Video.

* 1. DivX

Một codec nổi tiếng và rất rất phổ biến trong nhóm MPEG-4, có thể coi đã lên dạng lão làng khi có khá nhiều đầu đọc DVD đọc luôn cả DivX. So sánh giữa chuẩn MPEG-2 và DivX thì chất lượng có thể coi là như nhau có điều DivX trội hơn ở bit-rate thấp, điều này đồng nghĩa với việc chúng ta có thể có một đoạn phim tương tự DVD video nhưng nhẹ hơn chỉ cỡ bằng 1/2 của DVD video.

Tuy phổ biến như vậy nhưng so với các codec chị em của nó gần đây thì DivX bắt đầu tỏ ra yếu thế, đặc biệt là trước XviD và H264 và một số codec đang phát triển khác.

DivX homepage: http://www.divx.com

* 1. XviD

Xvid là một codec theo chuẩn MPEG-4. Nó có thể nén một file video thành những dữ liệu theo chuẩn của MPEG-4 và có thể được lưu trữ dưới dạng container .AVI .OGM .MP4 hay khác. Bản thân codec này không thể tự nén video mà nó phải nhờ một công cụ khác giúp nó, đó có thể là trình VirtualDub.

Thông thường một codec sẽ đi kèm cả Decoder và Encoder, nhưng cũng có trường hợp một bộ cài chỉ có Decoder mà thôi – với bộ cài đó bạn chỉ có để xem phim dùng codec đó mà không thể nén phim bằng codec đó. Bộ filter chỉ có decoder có thể kể đến là Nic’s decoder filter.

Xvid codec là một dự án mã nguồn mở GNU-GPL. Nghĩa là mã nguồn của nó có thể được download bởi bất kỳ ai và chỉnh sửa theo bất kỳ ý thích nào của họ. XviD là chuỗi kí tự viết ngược của DivX (phát triển bởi một công ty thương mại) – có lẽ nó giống như một sự thách thức với codec DivX. Sự thật cho thấy XviD vượt trội về chất lượng nén hơn hẳn DivX.

XviD codec homepage: http://www.koepi.info/xvid.html

* 1. RM/RMVB

Codec rất quen thuộc và phổ biến nhất Trung Quốc, nếu bạn bắt gặp 1 phim bộ tiếng Hoa bất kỳ được share trên mạng, thì chắc chắn rằng 99% sẽ có dạng \*.rmvb. Đây là codec riêng từ RealOne. Loại này cho ra ảnh phim rõ nét, sáng đẹp và dung lượng của file khá nhỏ, đặc biệt có chút nhỉnh hơn AVI nếu như là rip từ DVD. Tuy nhiên, codec này là đóng, và thông tin về nó không được RealOne cập nhật rõ ràng lắm.

Homepage: http://real.com

* 1. H.264

H264 là một codec thật sự ấn tượng.

Sức mạnh của H264 có thể thể hiện qua một đoạn anime dài 24 phút, phân giải 704×400 pixel với âm thanh MP3 128kbps mà chỉ có 76.2MB (video bit-rate là 301kbps) – chất lượng của nó gần bằng đoạn anime tương tự như thế nén với DivX nhưng nặng vào cỡ 233MB!

H264 có thể mang lại một đoạn video đẹp hơn cả video gốc ban đầu từ MPEG-1 (?!). Nguyên nhân của sự kỳ cục ấy là do H.264 sử dụng một thuật toán làm mờ đi các chi tiết mà nó thấy không cần thiết trong các khung hình. Các khung hình được de-cube - xoá nhoà các vết ô vuông của codec MPEG-1 tạo ra, mang lại cho người xem cảm giác hình ảnh mịn hơn, màu sắc đẹp hơn.

Mặc dù đây là phiên bản tính phí (chi phí sử dụng từ MPEG-LA khá cao), nhưng hiện có một project giúp H264 đến người sử dụng miễn phí đó là compressor / decompressor x264.

Chi tiết tại: http://www.videolan.org/developers/x264.html

* 1. VP8 (WebM)

Kể từ khi Google quan tâm tới VP8 bằng việc mua lại On2, VP8 bắt đầu trở thành đề tài nóng. Tháng 5-2010, một chuyện được xem là “động trời” với HTML5, khi Google mở mã VP8, miễn phí sử dụng, cùng với định dạng container mới WebM và dự án www.webmproject.org.

Với WebM, người dùng cuối không cần quan tâm tới các Plugin hay cài các codec khác, khi sử dụng chế độ mã hóa video theo chuẩn VP8 và audio theo Vorbis (mã mở). Điều quan trọng hơn, các trình duyệt thế hệ mới như Mozilla Firefox, Opera, Google Chrome và IE9 đều sẵn sàng hỗ trợ, trong khi Safari cũng hỗ trợ bất kì codec nào cài vào QuickTime và tương lai cũng sẽ cho phép WebM.

Epiphany đã bắt đầu hỗ trợ WebM thông qua nền tảng Gstreamer. Các trình phát đa phương tiện như VLC, Miro, Moovida, Winamp, Mplayer cũng đã sẵn sàng hỗ trợ. Nhóm phát triển Ffmpeg cũng nhanh chân hỗ trợ VP8. Theo thống kê, hiện có đến gần 50 đối tác, nền tảng đã sẵn sàng với WebM và VP8, trong đó có Android (công bố vào quý 4 – 2010), AMD, ARM, Broadcom, Intel, Qualcomm, NVIDIA, Sorenson Media, Logitech,…

Đây được xem là codec đáp ứng đủ 2 yếu tố tối ưu và miễn phí, được hỗ trợ trực tiếp với nền tảng HTML5 và hứa hẹn sẽ là codec thay thế các codec “già yếu” trước đây để trở thành đối thủ nặng ký của H264 và flash. Hiện nay codec này

Homepage: http://www.webmproject.org

### Các codec audio:

* 1. FLAC (Free Lossless Audio Codec)

FLAC là viết tắt của cho Codec Free Audio Lossless, định dạng âm thanh nén tương tự như MP3, nhưng không giảm chất lượng, có nghĩa là âm thanh được nén FLAC không có bất kỳ tổn thất nào về chất lượng. Điều này là tương tự như với cách nén Zip một file như thế nào, tuy nhiên với FLAC dữ liệu được nén tốt hơn nhiều bởi vì nó được thiết kế đặc biệt cho âm thanh, và ta có thể chơi lại tập tin nén FLAC trong máy nghe nhạc hoặc xe hơi hoặc dàn âm thanh nổi tại nhà giống như một tập tin MP3.

Là codec mã nguồn mở hỗ trợ nhạc Lossless (định dạng âm thanh chất lượng cao) nhanh và tốt nhất hiện nay với đầy đủ tài liệu, đặc tả API,…

Homepage: http://flac.sourceforge.net/

* 1. MP3:

MP3 là một dạng file đã được nén bằng cách nén lossy (có mất mát dữ liệu). Nó là một dạng âm thanh PCM (pulse-code modulation-encoded) nhỏ hơn rất nhiều so với dữ liệu ban đầu do bỏ đi những phần âm thanh được cho là không quan trọng trong khoảng nghe được của con người, tương tự như cách nén JPEG dành cho hình ảnh.

Tên của dạng này bắt nguồn từ "MPEG-1, lớp âm 3", còn được gọi chính thức hơn là ISO/IEC 11172-3 lớp 3. Những tập tin theo dạng này được lưu với phần mở rộng tên .mp3. Đôi khi những tập tin theo tiêu chuẩn MPEG-2, lớp âm 3 cũng sử dụng phần mở rộng này.

Có rất nhiều kỹ thuật đã được dùng trong chuẩn nén MP3 để xác định phần nào nên bỏ đi, trong đó có tâm thần âm học psychoacoustic. Dữ liệu MP3 có thể được tạo ra với nhiều bitrate khác nhau để có thể dễ dàng chọn lựa giữa chất lượng cao hay cỡ tập tin đầu ra nhỏ.

Cốt lõi của kỹ thuật nén MP3 là một dạng biến đổi phức để chuyến tín hiệu sóng ngang theo thời gian thành tín hiệu dựa theo tần số.

MP3 vòm là một định dạng MP3 hỗ trợ 5.1 kênh cho âm thanh vòm, được giới thiệu vào tháng 12 năm 2004. MP3 vòm có tính tương thích ngược với chuẩn MP3 trước đây, và kích cỡ file sau nén cũng tương tự.

Đây là định dạng âm thanh mã nguồn đóng phổ biến nhất thế giới hiện nay.

* 1. AAC (Advanced Audio Coding):

Advanced Audio Coding (AAC) - (ISO 14496-3) là một định dạng âm thanh đa năng nén kiểu lossy được định nghĩa theo tiêu chuẩn MPEG-2 và được phát triển bởi liên minh Fraunhofer, Dolby, Sony và AT&T. AAC được phát triển nhằm thay thế cho định dạng âm thanh đã quá nổi tiếng MP3 để tích hợp trong container MP4-một container của MPEG-4 tiêu chuẩn hỗ trợ đầy đủ các tính năng phụ (xem thêm phần MPEG-4).

Dạng định này được phát triển để xóa đi những chỗ yếu của MP3 và nâng cao phương pháp mã hóa đã có. Do vậy những tín hiệu thu của âm thanh hay tiếng động sẽ được nhận biết và mã hóa 1 cách hiệu quả hơn hoặc những vấn đề của Pre-Echo sẽ giảm thấp xuống nhiều.

AAC có thể tích hợp tới 48 kênh âm thanh (có sample rate tới 96KHz) cộng thêm 15 kênh âm thanh tần số thấp (Low Frequency Enhancement-LFE) giới hạn sample rate ở 120 Hz.

AAC là dạng định nén âm có tiêu hao về chất lượng, được sử dụng rộng rãi qua các kênh mua bán nhạc trực tuyến như iTunes Store, Real Music Store, LiquidAudio được gắn kèm với hệ thống chống sao chép DRM (ví dụ như FairPlay của Apple).

* 1. Vorbis:

Vorbis là một dự án miễn phí mã nguồn mở của Xiph.Org. Đây là một định dạng nén âm thanh theo kiểu lossy có chất lượng và mức độ nén gần tương đương với AAC, thường được chứa trong container OGG.

Thư viện của codec này đang được cập nhật liên tục với tài liệu cũng như API khá đầy đủ. Bộ thư viện mã nguồn mở theo chuẩn GNU.

Đây là codec được sử dụng trong container WebM để tạo ra file multimedia dành riêng cho HTML5. Đã được hỗ trợ bởi hầu hết các trình chơi âm thanh.

### Truyền tải video/audio thông qua WebRTC:

## Các công nghệ liên quan

### HTML5

### Hội thoại trực tuyến