

CHƯƠNG TRÌNH KC01
ĐỀ TÀI MÃ SỐ KC01-14
-----&-----

ĐỀ TÀI THUỘC CHƯƠNG TRÌNH KHCN CẤP NHÀ NƯỚC KC
01
MÃ SỐ KC 01.14

**NGHIÊN CỨU PHÁT TRIỂN ỨNG DỤNG
CÔNG NGHỆ ĐA PHƯƠNG TIỆN**

Chủ nhiệm đề tài: PGS.TS. Nguyễn Cát Hò

CẤP QUẢN LÝ: Nhà nước
CƠ QUAN CHỦ TRỊ: Viện công nghệ thông tin – Đại học Quốc
gia Hà nội

**BÁO CÁO CH CHUYÊN ĐỀ:
LÝ THUYẾT XỬ LÝ VIDEO**

CHỦ TRÌ CHUYÊN ĐỀ: PHAN THẾ HÙNG

6352-10
20/4/2007

HÀ NỘI, 4/2005

Mục lục

I. KHÁI NIỆM VỀ VIDEO	3
1. Khái niệm chung	3
2. Khái niệm Digital Video (Video số)	4
3. Đặc điểm Video số	4
II. NÉN DỮ LIỆU VIDEO	6
1. Sự cần thiết phải nén với hiệu suất (tỉ lệ nén) cao	6
2. Một số thuật toán nén dùng cho Video	7
III. CÁC ĐỊNH DẠNG VIDEO SỐ	11
1. Định dạng Video	11
IV. ÂM THANH	17
1. Âm thanh tự nhiên	17
2. Hình thức số hoá âm thanh	17
3. Âm thanh gốc	18
4. Phương thức lấy mẫu trong âm thanh	19
5. Một số chuẩn nén dữ liệu âm thanh	20
6. Tạo âm thanh	22
7. Âm thanh 3D thực	22
8. Định dạng âm thanh	23
V. CÁC THAM SỐ TRONG VIDEO VÀ AUDIO	24
1. Các tham số cho Video	24
2. Các tham số cho Audio	30

VI. CHUYỂN ĐỔI DỮ LIỆU TỪ VIDEO, BĂNG, ĐĨA CD THÀNH CÁC TỆP VIDEO-AUDIO CHO MÁY TÍNH VÀ NGƯỢC LẠI	33
1. Các cổng chuyển đổi tín hiệu	33
2. Thu tín hiệu từ các thiết bị phát Video-Audio vào máy tính	34
3. Chuyển đổi dữ liệu Video-Audio thành các định dạng khác nhau	36
<hr/>	
VII. MỘT SỐ KỸ THUẬT XỬ LÝ VIDEO-AUDIO TRÊN MÁY TÍNH	41
1. Kỹ thuật đánh dấu (Marker) và keyframe	41
2. Chuyển cảnh (Transition)	41
3. Kỹ thuật trộn (mix)	42
4. Kỹ thuật tạo độ trong suốt (Transparence)	44

I. KHÁI NIỆM VỀ VIDEO

1. Khái niệm chung

Video ra đời vào những năm đầu của thế kỷ XX nhưng nó phát triển khá chậm chạp và có nhiều người còn không tin vào khả năng của nó. Như Darryl Zanuck, giám đốc hãng phim Fox-TK20 phát biểu đầu năm 1946 “*Tivi sẽ không thể nào tiếp tục được trọng dụng quá 6 tháng. Người ta sẽ nhanh chóng chán ngay việc theo dõi một cái hộp gỗ mỗi tối*”. Video chỉ thực sự phát triển vào những năm cuối của thế kỷ XX. Với sự phát triển nhanh chóng của công nghệ, ngày nay Tivi-Video đã trở thành một thành phần không thể thiếu được trong đời sống xã hội.

Video là gì? Chúng ta có thể hiểu Video là một dạng dữ liệu bao gồm âm thanh và hình ảnh kết hợp với nhau và luôn có sự biến đổi về nội dung (khuôn hình) theo thời gian.

Các yêu cầu hệ thống của Video: Thông thường, nếu chúng ta xem một đoạn Video mà âm thanh và hình ảnh không khớp hay tốc độ quá chậm so với khả năng nhìn của chúng ta thì chắc chắn video không thể đăng tải được nội dung thực sự của nó. Nếu chúng ta xem các phim được sản xuất đầu thế kỷ XX thì chúng ta thấy các hình ảnh trên màn hình thường bị chậm hay bị giật so với hoạt động thực. Lý do là các máy quay đã không thu đủ 24 hình trên 1 giây. Do đó, các hệ thống Video hiện nay đều yêu cầu các thiết bị thu, phát, đường truyền video phải đảm bảo việc hiển thị hình ảnh và âm thanh trong thời gian thực. Hiện nay trên thế giới sử dụng 3 hệ Video chính: NTSC (National Television Standard Committee) theo chuẩn 29,97 hình/giây, PAL, SECAM theo chuẩn 25 hình/giây.

Truyền hình NTSC dùng mành 525 dòng và hiển thị đầy mành với tần số 30 mành mỗi giây, bằng phương pháp quét xen dòng 60 bán mành mỗi giây để phù hợp với tần số xoay chiều ở Mỹ là 60 Hz. Các ghép nối video NTSC sử dụng các đầu cắm và jack cắm chuẩn RCA. Các chuyên gia vô tuyến truyền hình thường nói đùa rằng NTSC là viết tắt của "Never Twice The Same Color" (cùng

một màu không bao giờ lặp lại hai lần) vì khả năng kiểm soát màu của chuẩn NTSC rất kém. Truyền hình NTSC được quảng bá ở Mỹ, Nhật và hầu hết các nước Trung và Nam Mỹ nhưng không dùng ở Châu Âu và Châu Á. Hầu hết các nước châu Âu và châu Á đều dùng chuẩn PAL dựa trên cơ sở tần số điện là 50 Hz.

2. Khái niệm Digital Video (Video số)

Cùng với sự ra đời và phát triển mạnh mẽ của máy tính và hệ thống viễn thông trong những thập kỷ cuối của thế kỷ XX, máy tính đã được sử dụng trong rất nhiều ngành công nghiệp. Do đó đặt ra yêu cầu cần phải có dữ liệu dạng âm thanh và hình ảnh cho máy tính (để máy tính có thể hiểu được). Chính vì vậy ra đời khái niệm Digital Video. *Digital Video là Video nhưng được ghi (mã hóa) dưới dạng số bằng các mã 0 và 1. Video thông thường được ghi dưới dạng tín hiệu tương tự (Analog)*

3. Đặc điểm Video số

- Video số sử dụng độ phân giải 72dpi.

Video số thường sử dụng độ phân giải 72 dpi (số điểm ảnh cho 1 inch). Đặc điểm này dựa trên giới hạn về khả năng nhìn của mắt người. Với màu huỳnh quang thì mắt người chỉ nhận biết được sự khác nhau của chất lượng hình ảnh chuyển động với độ phân giải dưới 72 dpi.

Chúng ta chỉ sử dụng độ phân giải trên 72 dpi cho một số trường hợp đặc biệt khi cần có Video hoặc ảnh chất lượng cao như các đoạn phim ảnh sử dụng cho việc phân tích khoa học cần phóng to lên nhiều lần hoặc các ảnh dùng trong ngành công nghiệp in ấn. Nếu muốn có một tấm ảnh in với chất lượng cao có thể người ta phải đặt độ phân giải lên trên 400 dpi.

- Kích cỡ tệp video là rất lớn.

Chúng ta có thể làm một phép tính như sau:

Nếu một đoạn Video thông thường hệ PAL (24 hình trên giây) có thời lượng là 1 phút có kích cỡ khung hình là 640x480 độ sâu màu 16 bit thì kích cỡ tệp này là: $16 \times 640 \times 480 \times 24 \times 60 = 7077888000\text{bit} = 843\text{MB}$

Như vậy chúng ta thấy rằng dữ liệu của video là rất lớn so với thiết bị lưu trữ thông tin hiện nay. Do vậy người ta luôn phải tìm mọi cách để giảm kích cỡ của Video. Có nhiều cách giảm kích cỡ tệp video ví dụ như giảm kích cỡ khuôn hình, giảm độ sâu màu... nhưng cách hiệu quả đó là sử dụng các thuật toán nén ảnh.

Các thuật toán nén ảnh đảm bảo cho việc sử dụng các tệp video trên máy tính cũng như truyền dữ liệu video trên mạng là hiệu quả, trong thời gian thực. Hiện nay, đã có rất nhiều thuật toán nén video khác nhau, nhưng nền tảng vẫn chủ yếu dựa trên các thuật toán nén ảnh như thuật toán RLE, LZW, Wavalet, DCT... Ngoài ra người ta còn có một số thuật toán giành riêng cho Video số như nén không gian màu, nén cấu trúc trong, nén dựa vào đối tượng cơ bản.

II. NÉN DỮ LIỆU VIDEO

1. Sự cần thiết phải nén với hiệu suất (tỷ lệ nén) cao

Như chúng ta đã biết dữ liệu Video số là rất lớn và yêu cầu hiển thị trong thời gian thực, do đó muốn sử dụng được video số một cách hiệu quả thì phải có thuật toán nén với hiệu suất cao.

Hiện nay, có nhiều thuật toán nén khác nhau nhưng có thể phân thành hai dạng chính: đó là thuật toán nén mất thông tin và nén không mất thông tin.

Nén không mất thông tin: đây là nhóm các thuật toán nén mà khi dữ liệu được phục hồi vẫn đảm bảo được chất lượng như dữ liệu gốc (chất lượng Video không hề thay đổi). Nhưng những thuật toán này có tỷ lệ nén rất thấp. Nó chỉ nén được khoảng 2 lần so với kích cỡ gốc. Trong thực tế chỉ sử dụng các thuật toán nén này để tạo các tệp video nguồn cho soạn thảo hoặc để di chuyển video từ hệ thống này sang hệ thống khác. Khi làm việc với các tệp video sử dụng thuật toán này chúng ta cần chú ý các tệp video là rất lớn đối với nhiều hệ thống máy tính đồng thời yêu cầu về tốc độ truyền dữ liệu cũng rất cao khi hiển thị (playback).

Nén mất thông tin: các thuật toán thuộc nhóm này thường có tỷ lệ nén rất cao có thể nén với tỷ lệ từ 10 đến 100 lần so với kích cỡ gốc. Ví dụ khi ta có 1 tệp Video kích cỡ 100 Mb, nếu áp dụng thuật toán nén này thì kích cỡ của tệp Video chỉ còn khoảng từ 1-10 Mb. Nhược điểm của các thuật toán này là chỉ đảm bảo chất lượng hình ảnh Video tương đối tốt nhưng không được như hình ảnh Video gốc. Tức là khi dùng các thuật toán nén này một nhóm thông tin của video đã được lượng tử hoá ví dụ một nhóm màu gần giống nhau gần nhau được chuyển thành một màu đặc trưng để giảm sự mã hoá màu như vậy sẽ làm giảm kích cỡ tệp video. Khi phục hồi các tệp video để hiển thị trên màn hình thì thông tin về màu sắc sẽ không được đầy đủ như tệp gốc nhưng kèm vào đó là các phương pháp xử lý màu giữa các vùng màu của thuật toán làm cho các cảnh video có chất lượng gần như ban đầu. Điểm hình là một số kiểu nén như JPEG, Planar RGB.

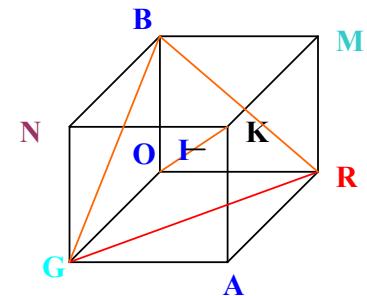
Trong thực tế thì người ta sử dụng nhiều các thuật toán nén này cho các tệp video sử dụng trên đĩa CD-ROM, trên Internet vì có thể thay đổi được chất lượng tệp video làm cho kích cỡ của tệp nhỏ đi, và tốc độ hiển thị (play back) nhanh hơn.

2. Một số thuật toán nén dùng cho Video

2.1 Nén không gian màu

Đây là thuật toán dựa trên nguyên lý làm giảm thông tin màu (trong không gian YUV) và sự kém nhạy cảm của mắt người với màu sắc đặc biệt với việc màu sắc liên tục thay đổi trong các chuyển động. Dựa trên 3 màu cơ bản RGB người ta có một không gian màu như sau:

Nếu lấy O làm gốc với ba trục là ba màu cơ bản đỏ, xanh và xanh lá cây (Red,Green,Blue), ta sẽ có không gian màu (ORGB). Trục KO là đường tổng hợp ánh sáng của 3 màu với giá trị bằng nhau do đó nó chính là đường thể hiện độ sáng của màu sắc. Tại gốc O sẽ là màu đen.

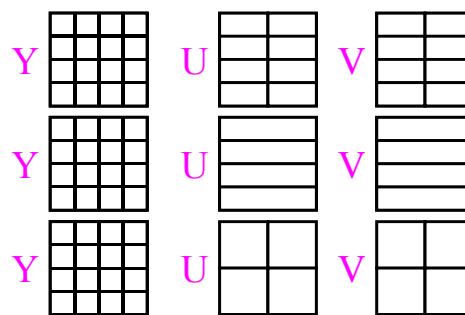


Từ không gian này ta xây dựng không gian YUV bằng cách: Dùng mặt phẳng GRB làm mặt phẳng màu. Mặt phẳng này có màu sắc được tổng hợp từ 3 màu cơ bản. Đặt tên mặt phẳng này là mặt phẳng (U,V) với hai đường thẳng U,V vuông góc với nhau và cắt nhau tại I.

Trục Y vuông góc với mặt phẳng (U,V) là đường thẳng KO thể hiện độ sáng (độ chói) của ánh sáng.

YUV (Luminance, 2 Color differences)

Y:U:V = 4:2:2



Y:U:V = 4:1:1

Y:U:V = 4:2:0

Trong thuật toán này người ta nén và làm giảm giá trị màu ở trên 2 trục U và V còn giá trị độ sáng Y được giữ nguyên vì giá trị này rất quan trọng(mắt người rất nhạy cảm đối với độ sáng). Người ta thường áp dụng rộng rãi thuật toán nén này trong các máy máy ghi Video, Tivi...

Nén không gian màu là cách mô phỏng ảo các vùng của khuôn hình với bản chất là việc tìm ra các mẫu và tạo lại các điểm ảnh. Ví dụ: trong một ảnh có vùng màu xanh da trời, thuật toán nén không gian sẽ nhận biết nhiều điểm màu xanh giống nhau trong không gian này. Để mô tả lại các điểm màu xanh này, thuật toán nén không gian ghi lại các mô tả ảnh một cách ngắn gọn nhất ví dụ như số điểm trong vùng màu xanh, mã màu trong khu vực đó...Như vậy nếu bạn tăng không gian nén (kích cỡ vùng màu) thì dữ liệu và kích cỡ tệp video sẽ giảm và ảnh sẽ bị mất độ nét. Vậy cấp độ nén có thể được điều khiển thông qua các chỉ số lựa chọn chất lượng và tốc độ truyền dữ liệu.

2.2 Nén cấu trúc bên trong

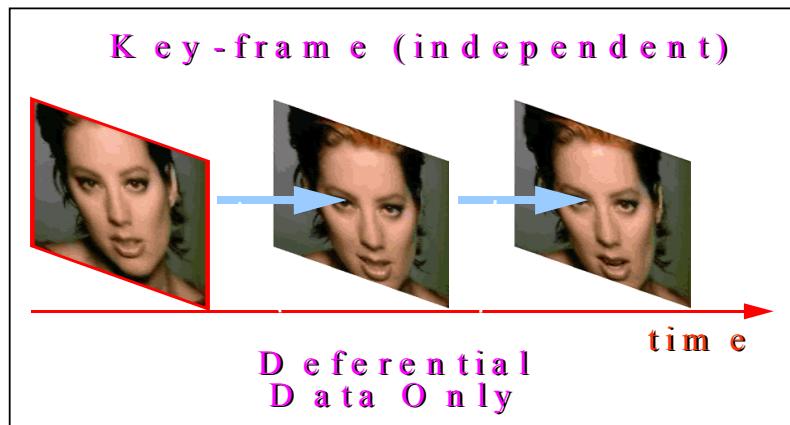
Thuật toán này dựa trên giải pháp nén theo cấu trúc và nén theo biến thời gian kết hợp với kỹ thuật bù chuyển động.

Đây là cách tìm kiếm các điểm ảnh thay đổi trong một khoảng thời gian hay một chuỗi các khuôn hình.

Ví dụ, trong một đoạn video có một nhân vật đang nói trên một nền tĩnh. Thuật toán nén theo thời gian sẽ nhận biết các pixel thay đổi từ khuôn hình này sang khuôn hình kia đó là hình khuôn mặt nhân vật đang nói. Còn tất cả các điểm khác không thay đổi. Để mô tả lại nhiều điểm ảnh và nhiều khuôn hình thì thuật toán này sẽ mô tả lại tất cả các điểm trong khuôn hình đầu tiên. Còn các khuôn hình tiếp theo thuật toán chỉ mô tả lại các điểm thay đổi. Phương pháp này gọi là phương pháp tính sai phân của khuôn hình.

Như vậy bằng cách phân tích trên thuật toán sẽ ghi lại khuôn hình đầu (keyframe) và khoảng thời gian nào đó cho chuyển động. Tiếp theo các nội dung của keyframe như vị trí, các vùng điểm màu.. Các vùng điểm ảnh thay đổi theo thời gian cũng sẽ được lưu lại. Khi hiển thị lại tệp video trên màn hình

chương trình dựa vào các chỉ số về keyframe, màu sắc, ánh sáng, các điểm ảnh chuyển động... để tái tạo lại chuyển động. Như vậy thuật toán này đã làm mất đi một số đáng kể các khuôn hình trong một khoảng thời gian và như vậy kích cỡ tệp video sẽ nhỏ đi.



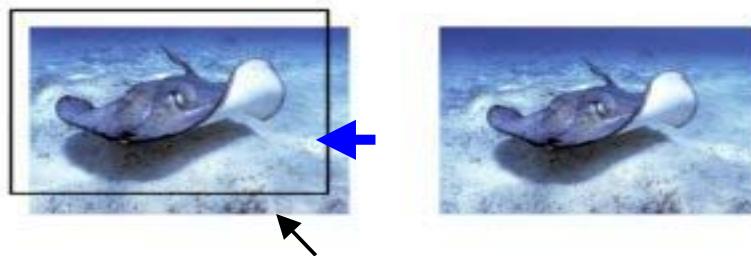
Theo thuật toán này nếu số keyframe càng nhiều (thời lượng giữa các keyframe càng ngắn) thì chất lượng video càng tốt. Vì vậy người ta dùng tham số keyframe và chất lượng khuôn hình (keyframe) để điều chỉnh cấp độ nén.

2.3 Nén dựa vào đối tượng cơ bản

Thuật toán này dựa trên kỹ thuật phân giải ảnh thành cấu trúc cây đối tượng sau đó véc-tơ hoá các đối tượng này (*Vector Quantization (VQ)*).

Ví dụ cảnh video như ở hình bên dưới: Hình đối tượng con cá sẽ được tách ra khỏi nền và được Vector hoá.

Như vậy việc ghi dữ liệu cho tệp Video với kiểu nén này chính là ghi các thông tin ảnh Vector và các thông tin chuyển động của ảnh véc-tơ trong một khoảng thời gian nào đó.



Khi giải nén chương trình dựa trên các thông tin về đường biên, màu sắc và hướng chuyển động của các đối tượng trong khuôn hình để xây dựng lại chuyển động. Tức là sẽ sinh ra một loạt các Frame để hình thành đoạn video.

Các thuật toán trên đều có chung một đặc điểm là chỉ lưu lại các thông tin cần thiết nhất của video và trên cơ sở đó xây dựng các frame cho video. Chính dựa trên đặc điểm này nên hầu hết các thuật toán này đều cho phép thay đổi chất lượng của video tùy vào mục đích sử dụng. Ví dụ như nếu người ta muốn sử dụng đoạn video cho Internet tốc độ chậm thì có thể điều chỉnh chất lượng kém đi và như vậy kích cỡ tệp video sẽ nhỏ đi nhiều. Trong trường hợp mục đích sử dụng cần có chất lượng cao mà không bị giới hạn về đường truyền như các tệp Video được ghi trên đĩa CD, DVD thì chúng ta có thể điều chỉnh chất lượng nén phù hợp để có được các tệp video thỏa mãn yêu cầu đặt ra.

III. CÁC ĐỊNH DẠNG VIDEO SỐ

1. Định dạng Video

1.1 Định dạng chuẩn cho hệ điều hành

1.1.1 Định dạng AVI

Đây là định dạng được thiết kế để dùng trong môi trường Windows. Định dạng này có thể sử dụng rất nhiều thuật toán nén video được phát triển từ trước đến nay.

1.1.2 Định dạng QuickTime

Đây là định dạng được thiết kế để dùng trong môi trường Macintosh. Định dạng Quicktime được xây dựng từ nhiều thuật toán nén ảnh và âm thanh trong môi trường Macintosh. Quick time cũng là định dạng được rất nhiều phần mềm ứng dụng hỗ trợ. Không những thế nó còn được cài đặt trong nhiều chíp xử lý của hệ thống.

Hiện nay cả hệ điều hành Windows và Macintosh đều cho phép sử dụng cả hai định dạng Video này.

1.2 Chuẩn quốc tế

MPEG-1/2/4- ISO (chuẩn quốc tế)

Định dạng MPEG-1/2/4 được phát triển bởi MPEG (Moving Picture Experts Group). Định dạng này được sử dụng để tạo các sản phẩm video trong ngành công nghiệp phát thanh truyền hình, Internet và các ứng dụng đồ họa .

1.2.1 MPEG-1

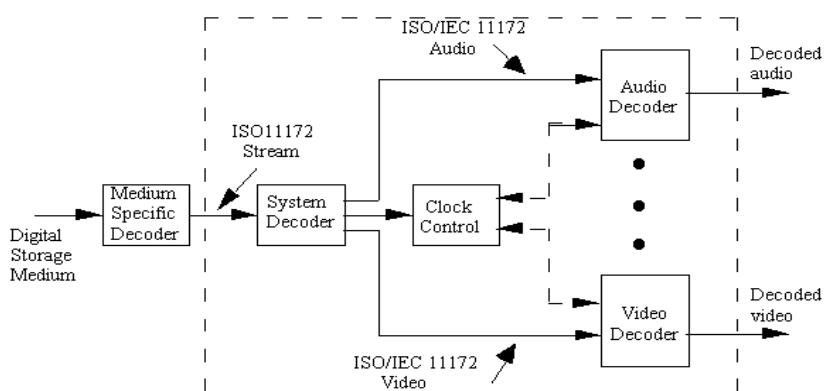
MPEG-1 được bắt đầu phát triển từ năm 1993 và được hoàn thiện vào năm 1998 với nhiều ưu điểm như: định dạng này có chất lượng cao tương đương với chất lượng hiển thị hình trên Tivi, có khả năng điều chỉnh chất lượng... Nó sử dụng thuật toán DCT(*Discrete Cosine Transformation*) với khuôn hình chuẩn

352x240 điểm với yêu cầu tốc độ đường truyền là 1.5Mbps (Mb trên một giây). Định dạng này được ứng dụng để xây dựng các sản phẩm Video trên đĩa CD-ROM.

MPEG-1 được thiết kế như sau :

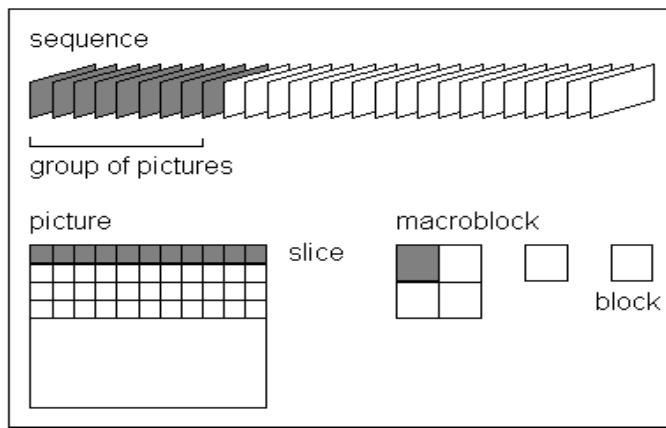
Theo chuẩn MPEG-1 nó phân định địa chỉ của các kênh dữ liệu âm thanh và hình ảnh kết hợp với thời gian. Đây là chức năng quan trọng vì từ dạng dữ liệu này sẽ được chuyển đổi thành các kênh dữ liệu phù hợp.

Sơ đồ nguyên mẫu chuẩn nén ISO/IEC 11172



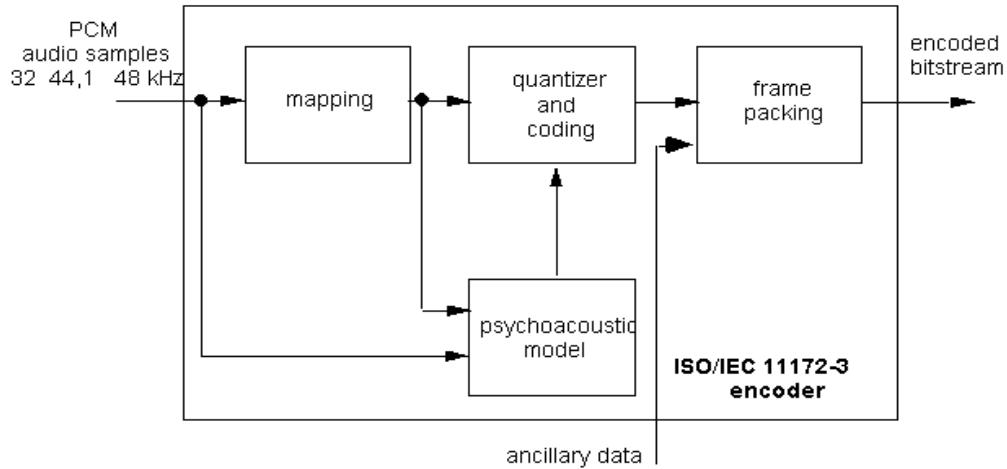
Theo sơ đồ này chúng ta thấy thông qua các địa chỉ được lưu trữ trong chuẩn MPEG-1, khi giải nén sẽ xác định được chính xác các kênh audio và video.

Thuật toán nén cho chuẩn MPEG-1 có khả năng nén cao. Đầu tiên người ta phải lựa chọn khung giờ phù hợp và giải pháp tín hiệu. Sau đó dùng thuật toán bù chuyển động và giảm thời gian dư thừa. Bù chuyển động được sử dụng trong việc tạo khung hình hiện tại dựa trên khung hình trước đó (chỉ cần một keyframe trước mà không cần dựa vào keyframe sau). Các tín hiệu khác, các lỗi được lượng hoá và nén bằng cách sử dụng thuật toán DCT (discrete cosine transform).



Hình 2: mô tả thời gian gian cấu trúc các khuôn hình trong định dạng MPEG-1

Hình 2 mô tả sự kết hợp giữa 3 loại thông tin về ảnh: Các pixel ảnh thay đổi, các chỉ số về vị trí , số lượng khuôn hình được sinh ra từ keyframe.

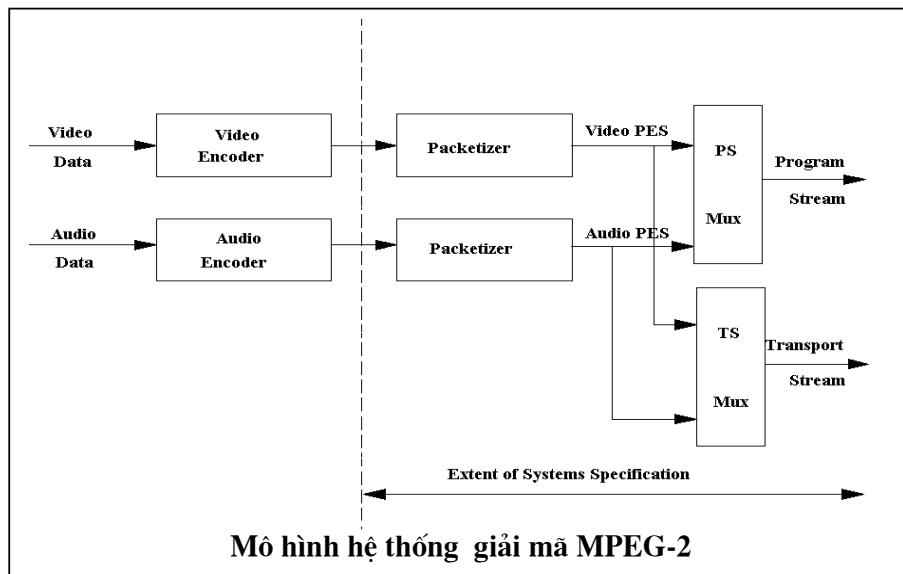


Hình 3 Cấu trúc giải nén cơ bản của Audio

Đầu vào là các tín hiệu mã với tần số 32, 44.1, 48 kHz. **Bản đồ** (mapping) sẽ lọc và lấy ra các mẫu đặc trưng. **A psychoacoustic model** là quá trình tập hợp dữ liệu và điều khiển việc lượng tử hoá và mã hoá để tạo ra khối các Frame. **Khối các frame** là các gói chuẩn (cơ sở) được mã hoá.

1.2.2 MPEG-2

Tháng 11/1994, MPEG-2 được phê chuẩn và bắt đầu được phát triển trên cơ sở các kỹ thuật nén tốt nhất của MPEG-1 nhưng phần mã hoá được mở rộng hơn. Các mã này được áp dụng cho các ảnh có độ phân giải 4:2:2 và cao hơn. Tuy nhiên MPEG-2 vẫn không được triển khai trong các ứng dụng video. Đến tháng 4/1997, MPEG-2 đưa thêm các mã phân định nhiều kênh audio. Một số thuật toán nén âm thanh được áp dụng trong phân này không còn bị lệ thuộc vào các thuật toán được áp dụng trong MPEG-1. Và chuẩn này đã được tổ chức ISO công nhận.



Theo mô hình này MPEG-2 đánh địa chỉ phối hợp một hoặc nhiều luồng dữ liệu của video và audio thành một luồng đơn thống nhất. Các dữ liệu trên luồng dữ liệu này được tổ chức phù hợp nhất cho việc lưu trữ và phát video. Việc tổ chức này dựa vào hai lớp chính: Program Stream và Transport Stream.

Program Stream(PS) là việc phối hợp một hoặc nhiều gói tin cơ bản PES (Packetised Elementary Streams) trong các luồng dữ liệu đơn thành một luồng dữ liệu đơn thống nhất. Gói tin Program Stream có độ lớn khác nhau và như vậy thời gian truyền gói tin là khác nhau. Trong quá trình truyền các gói tin, nếu phát hiện có một gói tin bị mất thì hệ thống sẽ yêu cầu truyền lại toàn bộ các gói tin.(vì Program Stream không xác định gói tin nào phải được truyền lại).

Program Stream được thiết kế cho việc sử dụng trong hệ thống đường truyền rất ít lỗi. Nó phù hợp với các ứng dụng có dùng phần mềm để xử lý.

Transport Stream(TS) là việc phối hợp một hoặc nhiều gói tin cơ bản PES (Packetised Elementary Streams) trong các luồng dữ liệu đơn thành một luồng dữ liệu đơn thống nhất. Gói tin Transport Stream có độ lớn bằng nhau là 188byte và như vậy thời gian truyền gói tin là như nhau. Trong quá trình truyền các gói tin, nếu phát hiện có một gói tin bị mất thì hệ thống không yêu cầu truyền lại toàn bộ các gói tin mà chỉ yêu cầu truyền lại gói tin bị mất.(vì Transport Stream đánh chỉ số cho các gói tin). Transport Stream được thiết kế cho việc sử dụng trong hệ thống đường truyền(môi trường) có nhiều lỗi.

Định dạng MPEG-2 có khuôn hình chuẩn là 720x480. Với yêu cầu đường truyền có tốc độ từ 5-20Mbps. Hiện nay MPEG-2 được ứng dụng cho việc xây dựng Video với chất lượng cao trên thiết bị DVD.

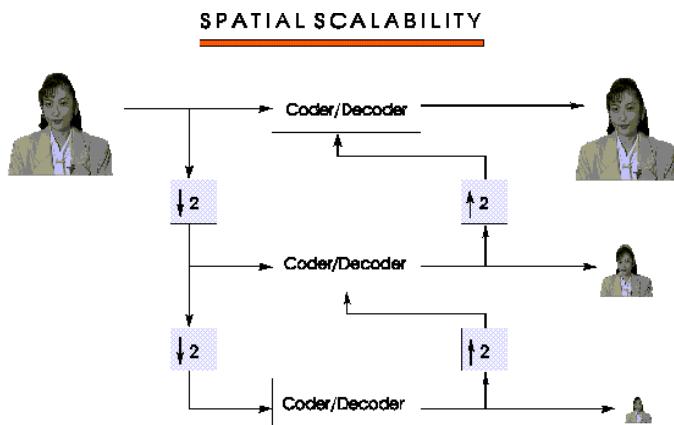
1.2.3 MPEG-4

MPEG-4 là chuẩn ISO/IEC được phát triển bởi MPEG (Moving Picture Experts Group). Uỷ ban này cũng đã phát triển chuẩn MPEG-1 và MPEG-2. Các chuẩn này cho phép phát hành video trên CD-ROM và truyền hình số. MPEG-4 là kết quả của hàng trăm nhà nghiên cứu và kỹ sư trên toàn thế giới. MPEG-4 được hoàn thành và tháng 10/1998 và trở thành chuẩn quốc tế tháng 1/1999. Cuối năm 1999 ra đời phiên bản 2 của MPEG-4.

MPEG-4 sử dụng thuật toán nén đối tượng cơ bản. Định dạng này yêu cầu tốc độ đường truyền thấp (64kbps) và không có kích cỡ khuôn hình chuẩn.

Nén hình ảnh trong Mpeg-4 sử dụng kỹ thuật phân lớp và lưu các thông tin đối tượng. Mỗi lớp lưu mã nén về nội dung của một chuỗi các ảnh(bao gồm: đường viền, quỹ đạo chuyển động, kết cấu bề mặt). Khi giải nén, thông qua nội dung được lưu trong các lớp để xây dựng lại từng phần của đoạn video.

Mô hình nén và giải nén theo từng lớp



Theo mô hình này, Video được nén trên 3 lớp với tỷ lệ giảm kích cỡ (không gian) là 2 lần trên từng lớp. Trên các lớp ngoài những thông tin về nội dung của một chuỗi hình ảnh còn lưu tỷ lệ giảm không gian của lớp so với lớp trên đó. Khi giải nén dựa vào tỷ lệ này cùng với các thông tin về đường viền, quỹ đạo chuyển động, kết cấu bề mặt của đối tượng trong video để phục hồi lại đoạn video gốc.

Như vậy dựa vào tỷ lệ giảm kích cỡ không gian video chúng ta có thể điều chỉnh được tốc độ phát hình cũng như kích cỡ của tệp video cho phù hợp với băng thông.

Với tính mềm dẻo của MPEG-4, nó đã được triển khai trong 3 lĩnh vực :

- ◆ Truyền hình số (Digital television)
- ◆ Tương tác tốt với các ứng dụng đồ họa (Interactive graphics applications)
- ◆ Tương tác với đa phương tiện (Interactive multimedia)

Chuẩn kỹ thuật của MPEG-4 được thống nhất cho sản xuất, phân phối các sản phẩm Video cho cả 3 lĩnh vực trên.

IV. ÂM THANH

Khi nói đến video bao giờ người ta cũng đề cập đến 2 vấn đề đó là hình ảnh và âm thanh. Có thể nói âm thanh là một phần không thể tách rời đối với video, vậy âm thanh là gì? Chúng ta sẽ xem xét các vấn đề về âm thanh ở các phần dưới đây.

1. Âm thanh tự nhiên

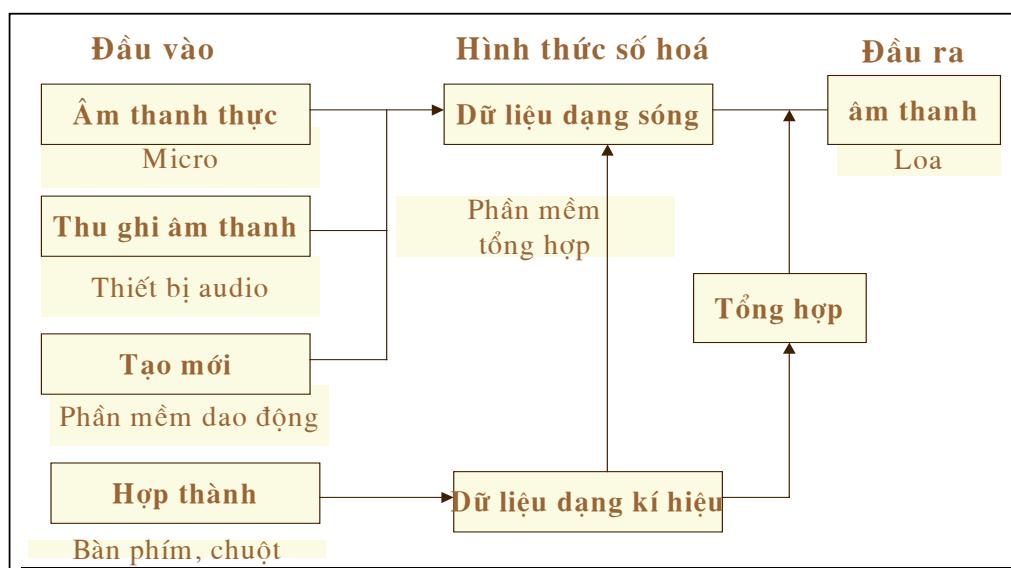
Có thể nói bản chất của âm thanh đó là sự dao động không khí. Khi một vật phát ra âm thanh chính là vật đó đã làm không khí xung quanh đó bị dao động. đo dao động của âm thanh bằng Hz và đơn vị đo độ ôn của âm thanh là dB.

$$\text{Độ ôn : } \text{dB} = 20 \cdot \log_{10} (\text{P}_1/\text{P}_2) \text{ với P là tần số âm thanh}$$

Đặc điểm: âm thanh giúp cho con người có thể hiểu nhanh, rõ ràng một vấn đề. Nó khác xa với các tín hiệu từ Text bởi vì sự phối hợp giữa âm thanh và hình ảnh giúp cho con người có thể hiểu rõ được mọi sự vật một cách nhanh chóng. Trong tự nhiên khả năng nghe của con người khoảng 40 Hz ~ 44KHz, nếu tần số âm thanh quá cao hoặc quá thấp thì người ta cũng không thể nghe được những âm thanh này. Ta có thể nghe thấy âm thanh có trong thực tế hoặc được con người sáng tạo ra.

2. Hình thức số hoá âm thanh

Người ta có thể số hoá video theo sơ đồ sau:

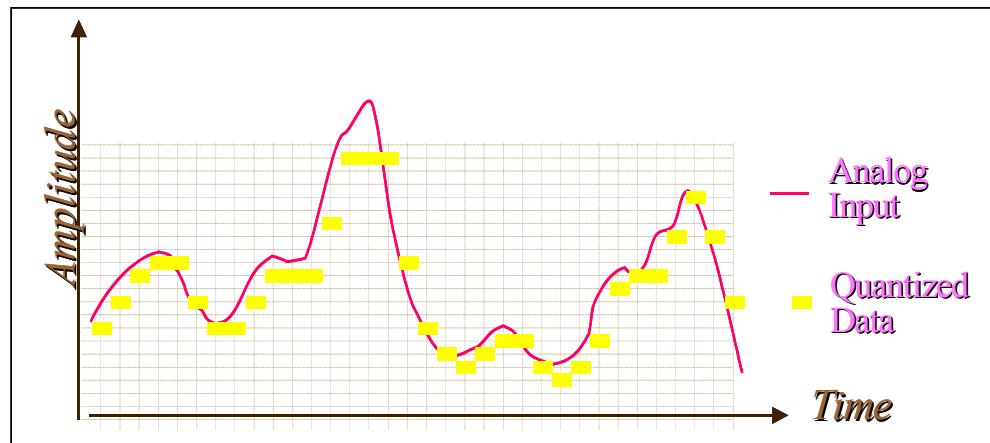


Theo sơ đồ trên chúng ta thấy đầu vào của âm thanh có từ rất nhiều nguồn khác nhau. Có thể là âm thanh thực có trong tự nhiên. Thông qua các thiết bị thu như micro chuyển hóa âm thanh thành dạng sóng điện từ và ghi vào băng đĩa. Chúng ta cũng có thể dùng các thiết bị sao chép âm thanh như đầu video, radio cassette,... để chuyển âm thanh từ băng, sang băng, từ băng sang đĩa,... Chúng ta cũng có thể tạo âm thanh bằng cách xây dựng các bộ dao động như các thiết bị âm nhạc... Với sự hỗ trợ của các phần mềm tổng hợp âm thanh chúng ta có thể tạo âm thanh từ các ký hiệu. Ví dụ: chúng ta có thể chơi nhạc bằng bàn phím, xây dựng một bản nhạc bằng cách soạn các nốt nhạc sau đó cho phát lại trên máy tính.

Từ các nguồn âm thanh khác nhau các âm thanh này đều được chuyển hóa thành sóng điện từ và được số hóa. Các dữ liệu sau khi được số hóa sẽ được máy tính xử lý. Sau đó các dữ liệu này sẽ được chuyển ngược thành âm thanh thực thông qua hệ thống loa.

3. Âm thanh gốc

Các tín hiệu âm thanh ở dạng nguyên thể có dạng hình sóng. Trước đây người ta thường thu tín hiệu âm thanh và ghi lại dưới dạng tương tự. Ngày nay, với sự phát triển của công nghệ số nên người ta đã số hóa âm thanh để có thể xử lý tốt hơn cho các ứng dụng thực tế.



Sơ đồ lượng tử hóa tín hiệu âm thanh

Âm thanh trong tự nhiên là sự dao động dạng sóng của khí. Khi được mã hoá dưới dạng sóng điện từ, âm thanh có có dạng đồ thị như trên. Để số hoá người ta sẽ lấy mẫu tại các điểm khác nhau dọc theo đồ thị của âm thanh. Số điểm lấy mẫu càng lớn chất lượng âm thanh số càng cao.

4. Phương thức lấy mẫu trong âm thanh

Vì âm thanh chính là sự dao động hình sóng quanh một trục nên người ta chỉ tính tần số âm thanh là phần trên của đồ thị thông qua trực đối xứng. Hay nói cách khác người ta chỉ lấy mẫu là một nửa chu kỳ dao động. Ví dụ nếu nói giọng nói có tần số là ~5.5 KHz thì tức là tần số thực khi nghe sẽ là 11KHz. Trong thực tế khi sản xuất đĩa CD nhạc thì người ta thường ghi với tần số nghe - >44.1KHz .

Khi xem xét vấn đề về tín hiệu chúng ta thấy rằng: nếu tín hiệu tần số vào lớn hơn khả năng nghe của con người thì sẽ gây ra các biến dạng âm thanh. Do đó cần phải có các phương pháp lọc bỏ các tần số không phù hợp.

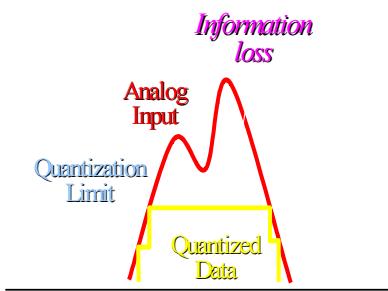
Khi số hoá người ta lấy mẫu trong từng khu vực và
ghi lại tần số đặc trưng trong khu vực đó.

Lượng tử hoá độ sâu

$$1 \text{ bit} = 20 \cdot \log_{10}(2) = 6.021 \text{ dB}$$

$$16 \text{ bit} = 6.021 * 16 = 96 \text{ dB}$$

Chuẩn DVD khoảng lượng tử hoá : 16/20/24
bit



Trong quá trình lượng tử hoá, người ta thường cắt tiếng ôn bằng cách đặt ra giới hạn khi lượng tử để không gây ra hiện tượng dữ liệu bị sai lệch. Nhưng cách tốt nhất vẫn là điều chỉnh mức thu âm thanh nguồn.

5. Một số chuẩn nén dữ liệu âm thanh

Như chúng ta đã biết dữ liệu của video là rất lớn. Trong đó không chỉ có dữ liệu hình ảnh phải nén mà dữ liệu âm thanh cũng phải nén vì kích cỡ của nó cũng rất lớn. Ví dụ một đoạn âm thanh 1 phút có kích cỡ khoảng 10MB. Hiện nay có nhiều phương pháp nén âm thanh khác nhau như có thể chia thành 2 loại:

Nén không mất thông tin

Thuật ngữ nén không mất thông tin ở đây được hiểu theo nghĩa là mọi file âm thanh nén đều được giải nén thành chính âm thanh gốc đã được nén trước đó.

- Nén dạng entropy với tỉ lệ: 1.5 ~3.0
- Nén kiểu LPAC với tỉ lệ: 1.5 ~4.0

Đây là chuẩn nén không mất thông tin dạng sóng 8 bit, 16 bit, 20 bit hoặc 24 bit (âm thanh đơn hoặc đa kênh) được hỗ trợ trong hầu hết các hệ điều hành: Windows, Linux và Solaris. Nó sử dụng thuật toán CRC đảm bảo quá trình xử lý, truyền phát không mất thông tin. Quá trình mã hóa nhanh trong thời gian thực (4x-12x trên máy 500 MHZ Pentium).

Nén không mất thông tin là ưu điểm chính của định dạng file LPAC so với các định dạng file âm thanh nén mất thông tin thông dụng hiện nay như MP3, WMA, RealAudio. Ngược lại, việc sử dụng thuật toán nén mất thông tin cho ta tỷ lệ nén âm thanh rất cao. MP3 với tốc độ 128 kbit/s có tỷ lệ nén là 11 trong khi LPAC chỉ đạt tỷ lệ nén từ 1,5 đến 4 và phụ thuộc hoàn toàn vào dữ liệu âm thanh. Ví dụ như LPAC có tỷ lệ nén là 2 cho âm thanh dạng nhạc pop và 2,5 cho loại âm nhạc cổ điển. Khi sử dụng định dạng nén này rất có thể chúng ta sẽ nhận được hoàn toàn âm thanh dạng bit đơn trong quá trình nén và giải nén file âm thanh. Hầu hết các định dạng nén nguyên thuỷ không mất thông tin khác như Zip, LZH, Gzip đều có tỷ lệ nén là 1 (hoàn toàn không nén được file âm thanh)

LPAC được sử dụng trong trường hợp file âm thanh cần đạt chất lượng tốt nhất trong quá trình phát mà định dạng MP3 không đáp ứng được. Các định

dạng file LPAC có đuôi là .PAC được xây dựng không mất thông tin và tương thích với mọi hệ điều hành cũng như bất kỳ quá trình xử lý âm thanh nào.

Nén mất thông tin:

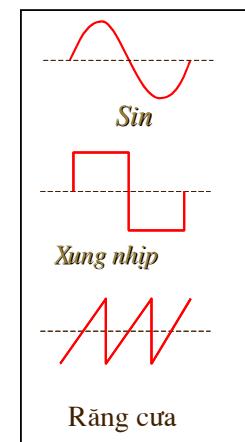
- Nén kiểu AAC (Advanced Audio Coding) có tỉ lệ nén: ~14lần được sử dụng trong MPEG-2/4

Người ta coi AAC là định dạng nén âm thanh có chất lượng tốt nhất trên Internet hay trên các đường truyền băng thông rộng. AAC được sử dụng rộng rãi trong các máy hát tự động và các thiết bị âm nhạc khác. Không những thế, AAC còn được coi là cơ sở hạ tầng trong việc truyền phát dữ liệu âm thanh trên Internet. Hãng Liquid Audio dự định phát triển một kỹ thuật tiên tiến nhất trong việc xử lý âm thanh để tích hợp vào AAC trong năm tới.

So sánh với MP3 người ta thấy rằng kỹ thuật AAC đã giảm tới 30% không gian lưu trữ dữ liệu. Kỹ thuật AAC đạt được điều này do đã loại trừ được tới 90% tín hiệu âm thanh gốc mà không hề ảnh hưởng tới chất lượng của âm thanh đó. AAC đã chính thức trở thành định dạng chuẩn quốc tế về âm thanh như các chuẩn kỹ thuật MPEG-2 hay MPEG-4.

AAC là kỹ thuật mã âm thanh dùng cho việc phát hành và phân phối các sản phẩm âm nhạc. Kỹ thuật AAC cho chất lượng nén cao. Các kiểm chứng độc lập nhau về hiệu quả của quá trình nén và giải nén cho thấy AAC hơn hẳn các định dạng âm thanh khác như MP3 hay bất kỳ mã nén âm thanh trực giác nào khác. AAC cung cấp 48 kênh âm thanh, và tốc độ lên tới 96kHz.

- Nén theo chuẩn MP3
- Nén kiểu WMA có tỉ lệ nén: ~15 lần được sử dụng làm Audio trong Windows
- TwinVQ có tỉ lệ nén: ~18 lần được sử dụng trong MPEG-4
- Nén không theo trực giác dùng phương pháp nén ADPCM (Adaptive Differential Pulse Code Modulation) có tỉ lệ nén: ~4.0



6. Tạo âm thanh

Người ta có thể tạo được các dạng âm thanh nhân tạo bằng cách xây dựng âm thanh dựa trên các đồ thị của các hàm toán học.

Tạo âm thanh dạng sóng hình Sin, hình xung nhịp, hình răng cưa... Việc tạo âm thanh nhân tạo này được ứng dụng trong rất nhiều trong các chương trình trò chơi giải trí.

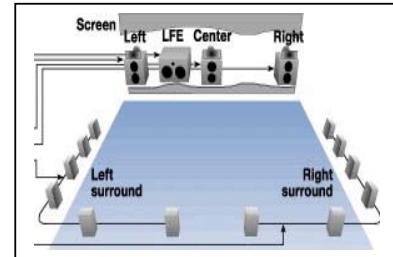
Âm thanh dạng ký hiệu cơ bản có nghĩa là ứng với một khoảng mức tần số âm thanh nào đó thì người ta mã hoá và chuyển thành một ký hiệu như các nốt nhạc (đô, rê, mi, fa, son, la, si ...) Nó có đặc điểm dữ liệu âm thanh không chính xác, chỉ mang tính giải thích logic chúng ta có thể thay đổi âm thanh bằng cách thay đổi cường độ, thời gian, vận tốc.. Chất lượng âm thanh phụ thuộc vào thiết bị đầu ra.

Đặc trưng của loại này là kích thước dữ liệu nhỏ ~1/1000 so với dữ liệu dạng sóng. Âm thanh ở dạng này thường có định dạng MIDI. Nó được ứng dụng trong các nhạc cụ điện tử.

7. Âm thanh 3D thực

Trước tiên muốn có được âm thanh 3D cần phải có một thiết bị ghi, thu đặc biệt. Thiết bị này sẽ thu âm thanh theo nhiều kênh khác nhau và ở các góc độ khác nhau.

Cách mô phỏng hay phát lại âm thanh người ta thường sử dụng hàm chuyển HRTF



Hiện nay người ta ứng dụng rất rộng rãi hàm HRTF để mô phỏng và tạo âm thanh 3D từ âm thanh 2D

Tạo âm thanh 3D được ứng dụng nhiều trong các trò chơi máy tính, hệ thống nhà hát nhỏ, họp từ xa...

8. Định dạng âm thanh

Đối với dữ liệu âm thanh dạng sóng chưa được xử lý

Người ta sử dụng định dạng WAV. Đây là định dạng dữ liệu dạng sóng được sử dụng trong môi trường Windows.

AIFF là định dạng Audio dùng trong các hệ máy Macintosh, Amiga, Silicon Graphics.

Đối với dữ liệu âm thanh dạng sóng đã được nén

Đặc trưng nhất của dữ liệu âm thanh này là định dạng MP3. MP3 có chất lượng cao, đồng thời tỷ lệ nén tốt. Nó được sử dụng nhiều trong các ấn phẩm phát hành trên Internet.

Dữ liệu dạng MIDI :

SMF là định dạng theo chuẩn MIDI đại diện cho sự mã hoá âm thanh bằng ký hiệu.

V. CÁC THAM SỐ TRONG VIDEO VÀ AUDIO

Trong thực tế, chúng ta phải sử dụng video và audio với các mục đích khác nhau trong các môi trường khác nhau do đó đặt các tham số cho video là rất cần thiết. Những tham số này sẽ xác định rõ chất lượng của sản phẩm. Ví dụ khi chúng ta muốn xây dựng một chương trình Video cho đĩa CD, hoặc DVD thì chúng ta phải có các lựa chọn nén khác so với các chương trình video xây dựng cho Web vì tốc độ truyền tín hiệu trong các thiết bị CD hoặc DVD lớn hơn rất nhiều so với tốc độ truyền tín hiệu trên Internet .

1. Các tham số cho Video

1.1 Compressor

Tham số này xác định các kiểu nén của video. Thông thường các kiểu này dựa vào các chuẩn nén khác nhau được viết cho Video.

1.1.1 Video cho Window người ta thường sử dụng các kiểu nén sau

Microsoft RLE : Kiểu này dùng để nén các frame có kích cỡ lớn và màu phẳng (các ảnh không có chiều sâu), ví dụ: để làm các phim hoạt hình. Kiểu nén này có mã độ dài 8 bit dùng thuật toán nén không mất thông tin RLE(Run –Length-Encoding). Chất lượng video cao.

- Microsoft Video1: Dùng cho nén video dạng tương tự (analog video). Kiểu mã nén này hỗ trợ các điểm có 8bit, 16 bit độ sâu.
- Indeo (R) video R3:2 : Sử dụng để nén video 24 bit dùng cho đĩa CD. Kiểu nén này có tỷ lệ nén tốt hơn, chất lượng tốt hơn, và tốc độ hiển thị (khi xem video) nhanh hơn so với kiểu nén Microsoft Video1. Cho kết quả tốt nhất nếu sử dụng mã nén Indeo Video trên dữ liệu video mà trước đó dữ liệu không bị nén với tỷ lệ cao. Khi sử dụng loại dữ liệu này để hiển thị lại thì chúng ta có thể so sánh các mã nén này với kiểu nén Cinepak.

- Cinepak code by Radius: Sử dụng để nén video 24 bit dùng cho CD-Rom hoặc Web. Đây là kiểu nén đạt được tỷ lệ nén cao hơn và tốc độ phát lại (giải nén) nhanh hơn so với kiểu nén video 1. Chúng ta có thể đặt chất lượng hình ảnh để có thể hiển thị lại video tốt với tốc độ 30KBps. Mã Cinepak được hiển thị lại rất nhanh nhưng khi nén mất rất nhiều thời gian. Nó không phù hợp cho việc soạn thảo video mà chỉ phù hợp cho việc chuyển một đoạn video thành kết quả cuối cùng.
- Intel Indeo 5.10 được sử dụng cho các định dạng Video phân tán trên mạng Internet cho các máy tính có bộ xử lý MMX or Pentium II. Đây là kiểu nén có đặc trưng: lựa chọn nén nhanh, mềm dẻo. Kiểu nén này cho phép người xử lý video có thể điều chỉnh việc hiển thị video đối với các băng thông khác nhau. Ví dụ có thể điều chỉnh để video có thể download với modem 56KB, 28,8KB hay đường cáp mạng... Mã nén này được thiết kế để làm việc phù hợp với mã Intel Audio Software.
- Intel Indeo Video Raw R1.1: được sử dụng tốt nhất cho việc thu các thông tin Video ở dạng dữ liệu không nén. Nó làm việc cùng với thiết bị Intel video-capture cards. Mã nén này cung cấp các hình ảnh chất lượng cao. Các tệp video dùng lựa chọn này có kích cỡ nhỏ hơn các tệp không dùng lựa chọn vì ở kiểu này màu sắc đã được chuyển từ model RGB thành model YUV .
- Intel Indeo Video Interactive: đây là kiểu nén tương tự như định dạng 5.10 nhưng nó có một số hỗ trợ cho các đặc tính trong suốt, nhiều đối tượng chuyển động.. trong video. Nó được hỗ trợ bởi các phần mềm tiện ích của Intel.

1.1.2 Video cho Macintosh thường sử dụng các kiểu nén.

- Component video: Được sử dụng cho thu video, lưu trữ video hay tạo các đoạn video trung gian (tạm thời). Kiểu nén này có tỷ lệ nén rất thấp do đó chiếm rất nhiều không gian đĩa

- Graphics: được sử dụng cho việc nén các ảnh chất lượng cao với độ sâu màu 8 bit. Mã nén (Graphics codec) này thường sử dụng cho các ảnh tĩnh nhưng đôi khi cũng sử dụng cho việc nén video bởi vì mã nén này không đạt được tỷ lệ nén cao. Nó thích hợp cho các tệp video được chạy (lưu trữ) trên đĩa cứng chứ không phù hợp với các tệp video đặt trên CD-ROM.
- Video: được sử dụng cho việc thu và nén các tín hiệu video có nguồn ở dạng Analog. Mã nén này cho kết quả cao khi xem lại các tệp video được lưu trữ trên ổ đĩa cứng. Cho chất lượng vừa phải nếu xem trên CD-ROM. Nó hỗ trợ cả hai loại nén: nén theo không gian và nén theo thời gian cho video 16 bit. Dữ liệu có thể nén lại hoặc dịch lại sau khi nén với tỷ lệ cao hơn mà chất lượng không suy giảm.
- Animation: Được sử dụng cho việc nén các khuôn hình có vùng màu có kích cỡ lớn. Ví dụ: các khuôn hình cho phim hoạt hình. Mã nén màu cho phép thay đổi tỷ lệ nén. Với tỷ lệ nén là 100%, video không bị nén. Nếu tỷ lệ dưới 100% tệp video bị nén ở dạng mất thông tin. Mã nén Animation dựa trên lý thuyết nén của Apple và thuật toán nén RLE.
- Motion JPEGA and Motion JPEG: được dùng cho mục đích chuyển mã video ví dụ như chuyển các tệp video trong máy tính, các đoạn video trên băng ra các thiết bị khác của máy tính như đĩa CD ... thông qua thiết bị thu video (video-capture card). Các mã nén này được hỗ trợ nhiều trong các chip có trên các thiết bị thu video như video-capture card do đó tốc độ xử lý rất nhanh.
- Photo-JPEG: kiểu nén này được dùng để nén các ảnh tĩnh có màu sắc biến đổi dần (các đường biên không rõ nét). Đây là kiểu nén mất thông tin nhưng có thể đặt được tham số nén để ảnh có chất lượng rất cao. Mã nén Photo-JPEG là kiểu nén đối xứng theo thời gian nhưng thời gian nén rất lâu. Các ảnh đã được nén theo kiểu này thì không nên dùng làm nguồn để soạn thảo vì nó đã bị mất thông tin. Tuy nhiên nó có tỷ lệ nén cao và chất lượng ảnh tốt nên có thể dùng để lưu trữ hoặc để di chuyển giữa các hệ thống máy tính.

- Chú ý: rất nhiều phần cứng nén (hardware compression) sử dụng định dạng JPEG. Với phần mềm QuickTime có thể chưa có trong danh sách các mã nén vì vậy không hiển thị được các tệp video. Chúng ta cần thêm mã nén của phần cứng đó vào danh sách mã nén cho QuickTime.
- H.263: Sử dụng tạo các video cho hội thảo. Mã nén này có tỷ lệ nén thấp. Không nên sử dụng chuẩn này cho soạn thảo video thông thường.
- DV - PAL and DV – NTSC : Sử dụng mã này tạo video số theo chuẩn PAL và NTSC. Mã nén này dùng để tạo các tệp video với định dạng PAL, NTSC phục vụ in ra băng theo các hệ trên hoặc ngược lại lấy từ băng vào máy tính thông qua digital-video capture card. Chuẩn nén này rất hữu dụng cho việc chuyển dữ liệu video từ hệ thống máy tính này sang hệ thống khác hoặc từ thiết bị này sang thiết bị khác.
- Cinepak : được sử dụng để nén video 24 bit. Các tệp sử dụng kiểu nén này để dùng cho CD-ROM và Web video. Mã nén này có tỷ lệ nén cao và tốc độ giải nén nhanh. Cinepak dung thuật toán nén không đối xứng các tệp video có kích cỡ nhỏ nhưng thời gian nén rất lâu. Cho kết quả tốt nhất nếu dùng mã nén này để tạo tệp video kết quả.
- Sorenson Video : được sử dụng để nén video 24 bit. Các tệp sử dụng kiểu nén này để dùng cho CD-ROM và Web video. Nó cũng giống như kiểu nén Cinepak nhưng đây là kiểu nén mới thiết kế để nén với chất lượng cao. Mã nén này cho hình ảnh tốt hơn, kích cỡ tệp video nhỏ hơn so với kiểu Cinepak vì vậy nó phù hợp cho việc tạo các tệp video cuối cùng chứ không phù hợp cho soạn thảo.
- Planar RGB: mã nén này được sử dụng hiệu quả cho các khuôn hình có vùng màu đặc như các tệp Animation. Nó sử dụng thuật toán nén RLE kết hợp với kỹ thuật tạo mã animation (Animation codec).

Đối với các thiết bị phần cứng hỗ trợ soạn thảo Video thường có các kiểu nén riêng được viết bởi nhà sản xuất thiết bị phần cứng.

1.2 Depth

Tham số này xác định độ sâu màu hay số màu của video khi hiển thị.

- ♦ Nếu độ sâu màu có giá trị là 8 bit tức là đoạn Video được thể hiện ở chế độ 256 màu
- ♦ Nếu độ sâu màu có giá trị là 24 bit tức là đoạn Video được thể hiện ở chế độ 16 triệu màu.
- ♦ Nếu độ sâu màu có giá trị trên 24 bit tức là đoạn Video được thể hiện ở chế độ trên 16 triệu màu.

1.3 Frame size

Tham số này xác định kích cỡ của khung hình tính bằng Pixel. Tham số này ảnh hưởng rất lớn đến kích cỡ của tệp video và thời gian xử lý video khi soạn thảo.

Khi đặt kích thước cho frame chúng ta phải chú ý đến tham số **Aspect**. Tham số này là hệ số co hay tỷ lệ co trong đồ họa máy tính. Giá trị của nó là trị số giữa kích thước theo chiều ngang và kích thước theo chiều dọc của một khung hình. Khi thay đổi cỡ hình phải duy trì tỷ lệ rộng/cao để tránh méo hình.

- ♦ Đối với chuẩn Video NTSC người ta thường sử dụng kích cỡ theo tỉ lệ 720x480
- ♦ Đối với chuẩn Video PAL người ta thường sử dụng kích cỡ theo tỉ lệ 4:3

1.4 Frame Rate

Tham số này xác định số khung hình trong một giây khi hiển thị Video. Tham số này ảnh hưởng lớn đến tốc độ xử lý của máy tính khi hiển thị. Chúng có một tệp Video được soạn thảo ở chế độ 24 hình trên giây khi được hiển thị ở tốc độ 12 hình trên giây thì chúng ta có cảm giác như các cảnh video bị chậm lại. Và ngược lại nếu chúng ta hiển thị tệp Video đó ở chế độ 48 hình trên giây thì cảnh video sẽ chuyển động nhanh gấp đôi.

1.5 Time base

Tham số này xác định số khung hình trong chế độ soạn thảo trong 1 giây.

- Khi soạn thảo Video ở chế độ NTSC thì chúng ta có thể đặt giá trị này bằng 29,97 hình (trên giây). Khi phát lại đoạn Video này theo chuẩn NTSC thì chúng ta sẽ được chất lượng tốt nhất .
- Khi soạn thảo Video ở chế độ PAL hay SECAM thì chúng ta có thể đặt giá trị này bằng 25 hình (trên giây). Khi phát lại đoạn Video này theo chuẩn PAL thì chúng ta sẽ được chất lượng tốt nhất .
- Khi soạn thảo Video cho các kiểu khác thì chúng ta có thể đặt giá trị này bằng 30 hình (trên giây).

1.6 Quality

Tham số này xác định chất lượng của các khung hình theo một chuẩn nén xác định. Giá trị của tham số này ảnh hưởng lớn đến chất lượng của khuôn hình cũng như kích cỡ của tệp Video. Nếu chất lượng thấp thì kích cỡ tệp video nhỏ. Nếu chất lượng cao thì kích cỡ của tệp video lớn. Một số trường hợp đặc biệt, trong một số chuẩn nén không đặt được tham số cho chất lượng.

1.7 Data Rate

Tham số này xác định giới hạn của tốc độ truyền dữ liệu của tệp Video ra màn hình hiển thị. Giá trị này được tính bằng Kilobit / giây (Kbps). Một số chuẩn nén cho phép thay đổi tham số này. Khi đặt tham số Data rate cần phải chú ý đến tốc độ truyền dữ liệu của hệ thống (không đặt tham số này cao hơn tốc độ truyền dữ liệu của hệ thống).

2. Các tham số cho Audio

2.1 Rate

Tham số này xác định dải tần số của Audio. Nó xác định chất lượng âm thanh khi phát lại. Thông thường tham số này cao thì chất lượng âm thanh sẽ tốt hơn nhưng kèm theo đó là kích cỡ của tệp Video-Audio cũng sẽ tăng lên đáng kể. Khi đặt tham số này trong chế độ soạn thảo Video-Audio chúng ta cũng cần phải chú ý đến tần số đầu vào của các tệp Video-Audio. Cách tốt nhất là đặt tham số này bằng với tần số đầu vào.

Đối với hệ thống Windows hoặc Macintosh các giá trị có thể đặt là : 5000Hz, 8000Hz, 11025 Hz, 32000Hz, 44100Hz, 48000 Hz

2.2 Format

Số bit để mã hoá âm thanh càng cao thì chất lượng âm thanh càng cao và kích cỡ tệp Video-Audio càng lớn. Có 4 định dạng chính:

- ◆ 8 bit – Mono (âm thanh đơn, 1 kênh âm thanh)
- ◆ 16 bit – Mono
- ◆ 8 bit – Stereo (âm thanh nổi, nhiều kênh âm thanh)
- ◆ 16 bit – Stereo

2.3 Compressor

Tham số này xác định các kiểu nén cho audio. Nó quyết định chất lượng cũng như độ lớn của tệp âm thanh.

2.3.1 Đối với hệ thống Window người ta thường sử dụng các kiểu nén âm thanh sau

- Intel Audio Software codec: Kiểu nén này sử dụng cho nhạc và lời nói trên Internet. Khả năng nén tối đa là 8:1. Mã nén này được thiết kế để làm việc với chương trình Intel Video Software .

- TrueSpeech : Kiểu nén này sử dụng cho nói chuyện trên mạng Internet với tốc độ truyền thấp.
- Microsoft GSM 6.10: Sử dụng cho telephone trên Internet ở châu Âu.
- MS-ADPCM (Microsoft implementation of Adaptive Differential Pulse Code Modulation) kiểu nén này sử dụng để tạo các tệp audio có chất lượng cao ghi trên đĩa CD-ROM. Kiểu nén này được sử dụng rộng rãi.
- Microsoft IMA ADPCM: kiểu nén này được sử dụng tạo các tệp audio cho các sản phẩm multimedia. Kiểu nén này dựa trên mã nén ADPCM được phát triển bởi IMA(Interactive Multimedia Association).
- Lucent Technologies SX8300P: kiểu nén này sử dụng cho giao tiếp trên Internet tốc độ thấp
- elemedia TM AX2400P: kiểu nén này được sử dụng tạo các tệp âm nhạc chất lượng cao trên Internet
- Voxware Audio Codecs: kiểu nén này sử dụng cho các tệp âm thanh dạng tiếng nói trên Internet tốc độ thấp

2.3.2 Đối với hệ thống Macintosh người ta thường sử dụng các kiểu nén âm thanh sau

- mLaw 2:1 : kiểu nén này được sử dụng cho việc chuyển các tệp audio để sử dụng tốt với các ứng dụng trên các máy trạm. Tại đó mLaw là định dạng audio chuẩn . mLaw được sử dụng cho telephone số ở Bắc Mỹ và Japan
- 16-bit BigEndian và 16-bit LittleEndian: kiểu nén này được sử dụng cho các phần cứng và phần mềm chuyên dụng nhưng thông thường nó không tốt cho việc soạn thảo.

- 24-bit Integer và 32-bit Integer: kiểu nén này được sử dụng cho audio 24-bit hoặc 32-bit nguyên. mã nén này được sử dụng tốt cho các phần cứng và phần mềm chuyên dụng nhưng thông thường nó không tốt cho việc soạn thảo.
- IMA 4:1: kiểu nén này được sử dụng cho các tệp audio tại các trạm. IMA 4:1 được phát triển bởi IMA sử dụng mã ADPCM.
- 32-bit Floating và 64-bit Floating: Các kiểu nén này được sử dụng trong các thiết bị phần cứng và phần mềm chuyên dụng nhưng thông thường không sử dụng cho soạn thảo Video.
- ALaw 2:1: giống mLaw nhưng được sử dụng cho telephone số ở Châu Âu
- QDesign Music Codec: Sử dụng để nén các tệp audio chất lượng cao sử dụng trên Internet. Chất lượng của nó tương đương các tệp âm thanh trên CD-ROM có định dạng 16-bit, 44.1 kHz. Các tệp âm thanh được nén theo kiểu này có thể nghe trực tiếp bằng đường Internet có tốc độ 28.8 Kbps.
- Qualcomm PureVoice: Đây là kiểu nén được thiết kế tạo các tệp âm thanh tốt nhất ở tần số 8 kHz. Cơ sở của kiểu nén này dựa trên thuật toán nén chuẩn của Telephone CDMA (Code Division Multiple Access)
- MACE 3:1 and MACE 6:1 :kiểu nén này để tạo các tệp audio thông thường cho hệ điều hành Macintosh. Nó dựa trên thuật toán nén âm thanh MACE (Macintosh Audio Compression and Expansion codec). Kiểu nén MACE 3:1 có tỷ lệ nén thấp hơn kiểu nén MACE 6:1 nhưng có chất lượng cao hơn

2.4 Data Rate

Tham số này xác định giới hạn của tốc độ truyền dữ liệu của tệp Audio khi phát lại. Giá trị này được tính bằng Kilobit/giây (Kbps).

Chuyển đổi dữ liệu từ video, băng, đĩa CD thành các tệp Video-Audio cho máy tính và ngược lại

3. Các cổng chuyển đổi tín hiệu

Chúng ta có thể chuyển dữ liệu từ nhiều nguồn video khác nhau, có thể là băng (tape), tập hợp các ảnh liên tiếp (tạo thành film) của một cảnh, từ ảnh động, hoạt hình, audio,... thành các tệp Video-Audio và được lưu trữ trên máy tính.

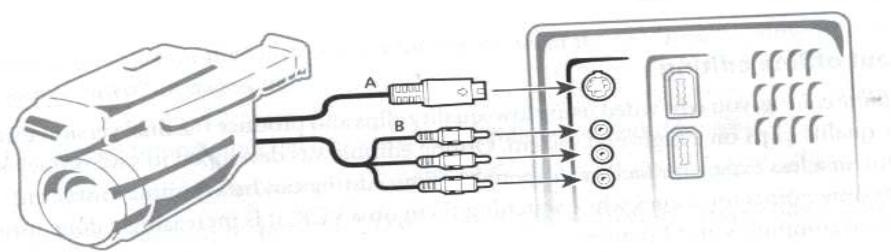
Digital media đây là các phương tiện số. Các phương tiện số có thể là các thiết bị thu, phát tín hiệu số như máy quay số, máy ảnh số.... Các thiết bị này lưu các tín hiệu Video-Audio ở dạng số và có định dạng theo các chuẩn mà máy tính có thể đọc và xử lý trực tiếp được.

Analog media: đây là các phương tiện lưu dữ liệu dưới dạng tín hiệu tương tự. Các thiết bị này thường là các thiết bị truyền thống ví dụ như máy quay thường, máy ảnh thường.... Như vậy muốn xử lý các dữ liệu này bằng máy tính thì chúng ta phải số hóa các dữ liệu này. Thông thường, người ta sử dụng các thiết bị phần cứng (video capture board) để thực hiện việc số hóa các tín hiệu Analog.

Các cổng kết nối từ các máy phát tín hiệu video vào máy tính:

Cổng vào ra tín hiệu Video-Audio dạng tương tự thông thường:

- ◆ Cổng S-video
- ◆ Cổng đa hợp (composite video).



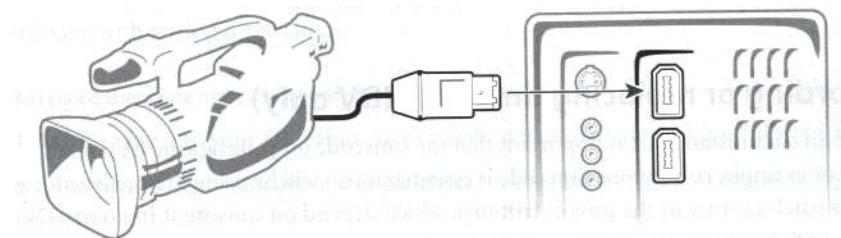
A là cổng S-video , B là cổng đa hợp

Nếu dùng cổng đa hợp người ta có thể điều chỉnh để thu: chỉ tín hiệu hình ảnh hoặc chỉ thu tín hiệu âm thanh hoặc cả hình ảnh và âm thanh với đầy đủ tín hiệu video (left, right).

Các thiết bị thu tín hiệu tương tự và chuyển thành dạng số thường sử dụng các cổng này để trao đổi thông tin ví dụ thiết bị DV500, DC2000, TV Capture board đều có sẵn các cổng này để trao đổi tín hiệu Video-Audio dạng tương tự

Cổng vào ra tín hiệu Video-Audio số thông thường:

Cổng IEEE 1394



Connecting to the IEEE 1394 (FireWire/i.Link) port on the computer

Thông thường để trao đổi thông tin giữa các máy phát tín hiệu Video-Audio và máy tính người ta sử dụng cổng IEEE 1394. Các thiết bị số như máy quay số, máy ảnh số, các thiết bị chuyển các dữ liệu Video-Audio số vào máy tính như Card VD500, DC2000... đều có sẵn cổng này. Các thiết bị hỗ trợ việc trao đổi tín hiệu Video-Audio số qua cổng này có tốc độ rất cao đảm bảo việc thu và hiển thị Video-Audio số trong thời gian thực.

4. Thu tín hiệu từ các thiết bị phát Video-Audio vào máy tính

4.1 Thu tín hiệu dạng tương tự

4.1.1 Yêu cầu phần cứng

Nhất thiết cần có thiết bị giao tiếp giữa máy tính và máy phát tín hiệu Video-Audio dạng tương tự. Thiết bị đó có thể là TV card, hay DV500... có

cổng giao tiếp là S-video hay Composite video. Thiết bị này được cài đặt trên máy tính và được hệ thống (Windows, Macintosh..) hỗ trợ.

Máy phát tín hiệu Video-Audio dạng tương tự. Nó có thể là Video-Cassette, máy quay video, thiết bị thu tín hiệu từ các đài phát hay vệ tinh như Angten...

Máy tính tốc độ cao, có bộ nhớ (RAM) lớn, không gian đĩa cứng (HDD) lớn và Card video có bộ nhớ lớn. Dây cáp nối các thiết bị Video-Audio với máy tính.

4.1.2 Yêu cầu phần mềm

Máy tính phải được cài đặt các phần mềm hỗ trợ thiết bị giao tiếp và có thể cả các chương trình soạn thảo Video-Audio. Ví dụ như chương trình Ulead, Adobe Premiere...

4.1.3 Các tham số cho công việc thu Video-Audio dạng tương tự

Kích cỡ của khuôn hình : Nếu các tín hiệu thu được được ghi với định dạng Quick time thì chúng ta đặt khuôn hình theo tỷ lệ 4:3. Nếu các tín hiệu thu được được ghi với định dạng Video cho Windows thì chúng ta có thể đặt kích thước khuôn hình bất kỳ.

Tốc độ của khuôn hình: chúng ta cần đặt tốc độ của khuôn hình phù hợp với hệ của Video khi phát. Nếu là hệ NTSC thì chúng ta đặt tốc độ là 29,97 khuôn hình trong 1 giây (fps). Nếu là hệ PAL hay SECAM thì chúng ta đặt tốc độ là 25 fps.

4.2 Thu tín hiệu dạng số

Để thu các tín hiệu Video-Audio số thì chúng ta cần có các yêu cầu sau:

4.2.1 Yêu cầu về phần cứng

Gần giống như các yêu cầu phần cứng để thu các tín hiệu dạng tương tự nhưng trên thiết bị giao tiếp (card DC2000, DV300, DV500...) phải có cổng giao tiếp IEEE 1394.

4.2.2 Yêu cầu phần mềm

Giống hoàn toàn như các yêu cầu để thu các tín hiệu dạng tương tự.

4.2.3 Đặt các tham số cho công việc thu Video-Audio dạng số

Kích cỡ của khuôn hình : đặt kích cỡ của khuôn phù hợp với tỷ lệ của khuôn hình trên các thiết bị phát. Ví dụ nếu khuôn hình trên máy phát digital camera là 720x480 dpi thì chúng ta có thể đặt kích cỡ khuôn hình khi thu là 720x480 hoặc 360x240... Nếu đặt như vậy thì tệp video thu được sẽ được đảm bảo về chất lượng.

Tốc độ của khuôn hình: chúng ta cần đặt tốc độ của khuôn hình phù hợp với hệ của Video khi phát.

Đặt dải tần số âm thanh: Chúng ta phải đặt dải tần số âm thanh phù hợp với hệ của Video. Thông thường DV audio thường sử dụng kiểu 16bit Stereo. Đối với hệ các máy phát video số (digital camera) sử dụng hệ NTSC hoặc PAL chúng ta đặt dải tần cho audio là 32kHz hoặc 48 kHz. Tuy nhiên trong một số trường hợp đặc biệt một số máy phát video số sử dụng dải tần Audio 44.1kHz vì vậy khi thu chúng ta phải đặt theo giá trị này.

5. Chuyển đổi dữ liệu Video-Audio thành các định dạng khác nhau

Sau khi chúng ta đã nạp được các dữ liệu Video-Audio vào máy tính chúng ta có thể sửa đổi dữ liệu đó phù hợp với mục đích và yêu cầu đặt ra. Sau khi đã có được các đoạn Video-Audio số phù hợp chúng ta phải chuyển chúng thành các tệp Video-Audio theo các định dạng khác nhau hoặc ghi ra băng, đĩa CD.

5.1 Chuyển các đoạn Video-Audio số đã được xử lý trên máy tính ra băng, đĩa

5.1.1 Yêu cầu về thiết bị

Trong trường hợp này các yêu cầu tương tự như yêu cầu về thiết bị khi thu tín hiệu Video-Audio vào máy tính. Đó là: thiết bị giao tiếp giữa máy tính và máy thu tín hiệu Video-Audio (Có thể thiết bị là card DV300, DV500, TV Capture board... .), máy tính có tốc độ cao, máy thu tín hiệu Video-Audio (video cassette, Video camera..).

5.1.2 Đặt các tham số cho đoạn Video-Audio khi in ra băng

Tỷ lệ khuôn hình: Khi in ra băng theo hệ nào thì chúng ta phải đặt tỷ số giữa độ rộng và chiều cao khuôn hình theo đúng hệ đó. Ví dụ nếu chúng ta định in ra băng theo hệ NTSC thì đặt khuôn hình theo tỷ lệ 720x480. Nếu muốn chất lượng hình ảnh trên băng cao thì chúng ta có thể ghi ở chế độ toàn màn hình.

Tốc độ khuôn hình (rate): Tốc độ khuôn hình phải đặt theo tốc độ khuôn hình của hệ thống thu. Đối với hệ PAL phải đặt tốc độ khuôn hình là 25 hình trên giây, hệ NTSC là 29,97 hình trên giây.

Đặt dải tần số cho audio : Giá trị này phải phù hợp với dải tần số âm thanh của hệ. Hầu hết các DV camera đều sử dụng dải tần 32KHz hoặc 48KHz.

Lựa chọn kiểu nén và đặt chất lượng Video-Audio: Nếu chúng ta in ra băng thì nên lựa chọn kiểu nén có chất lượng cao nhất. Thông thường người ta lựa chọn kiểu nén theo thiết bị phân cứng (các Card chuyển đổi tín hiệu).

5.1.3 Lựa chọn các định dạng tệp trung gian trước khi in ra băng

Trong trường hợp chúng ta không ghi trực tiếp được ra băng thì sau khi đã đặt các thông số cho tệp Video-Audio kết quả chúng ta sẽ ghi kết quả đó ra đĩa

cứng với định dạng AVI và MOV. Sau đó chúng ta phát lại tệp AVI và MOV trên máy tính để thực hiện việc thu tín hiệu.

5.2 Chuyển các đoạn Video-Audio số đã được xử lý trên máy tính để sử dụng cho Web hoặc CD-ROM

5.2.1 Yêu cầu về phần cứng

Trong trường hợp này yêu cầu về phần cứng không đòi hỏi phải có thiết bị giao tiếp giữa máy tính và máy thu phát tín hiệu Video-Audio. Nhưng yêu cầu một số phần mềm hỗ trợ hiển thị.

5.2.2 Các định dạng tệp Video-Audio

Lựa chọn định dạng cho tệp Video-Audio kết quả. Vì kết quả sử dụng cho các mục đích khác nhau do đó cần xác định chính xác định dạng tệp kết quả. Dưới đây là một số định dạng tiêu biểu:

Định dạng video theo chuẩn Quicktime :

Quick time lấy (down load) về từ mạng

Quick time cho phép xem trực tiếp trên mạng

Quick time CD-ROM

Định dạng audio theo chuẩn Real G2:

Real G2 dùng để lấy (down load) về từ mạng

Real G2 cho phép xem trực tiếp trên mạng

Định dạng Video-Audio theo chuẩn Windows

Windows media video

Windows media audio

AVI CD-ROM

Định dạng Video-Audio theo chuẩn MPEG

Cho Video dùng MPEG-1, MPEG-2

Cho Audio dùng MP3

Nếu kết quả dùng cho CD-ROM thì chúng ta phải lựa chọn định dạng tệp là MPEG-1, MPEG-2 có sử dụng thuật toán nén MPEG (Motion Picture Experts Group). Sử dụng định dạng MPEG-1 chúng ta có được sản phẩm Video-Audio có chất lượng tương đương VHS. Sử dụng định dạng MPEG-2 chúng ta có được sản phẩm Video-Audio có chất lượng tương đương SVHS. Tuy nhiên nếu chúng ta đặt tỷ lệ nén cao và chất lượng video thấp khi chất lượng video sẽ không được như mong muốn.

Nếu kết quả sử dụng để làm ảnh động cho Web thì chúng ta sử dụng định dạng GIF. Định dạng GIF sử dụng chuẩn nén không mất thông tin, sử dụng từ 2-256 màu để thể hiện hình ảnh và tốc độ khuôn hình thấp. Vì đặc tính trên nên định dạng GIF chỉ phù hợp cho việc tạo các logo, nút bấm động trên WEB. Chú ý khi sử dụng định dạng này không nên đặt quá nhiều khuôn hình vì như vậy kích cỡ của tệp quá lớn ảnh hưởng đến việc thể hiện lại trên WEB.

Nếu muốn có kết quả là tệp âm thanh thì chúng ta sử dụng định dạng MP3 (Motion Picture Experts Group 1 layer 3). Tỉ lệ nén của định dạng này là khá cao từ 5 đến 24 lần. Kích cỡ tệp nén nhỏ, chất lượng khá cao.

Nếu kết quả sử dụng trong môi trường Windows dùng định dạng AVI. Nếu kết quả sử dụng trong môi trường Macintosh thì dùng định dạng chuẩn Quick time.

Trong mỗi chuẩn định dạng dùng cho Web thì người ta phân ra làm nhiều mức chất lượng để phù hợp với tốc độ đường truyền. Ví dụ như 28,8k, 56k, ISDN, LAN...

Thông thường các định dạng trên đi kèm theo các chuẩn nén xác định.

5.3 Chuyển các đoạn Video-Audio số đã được xử lý trên máy tính thành chuỗi các ảnh

Không những chúng ta có thể chuyển kết quả Video đã được xử lý thành các tệp Video mà chúng ta còn có thể chuyển thành các khuôn hình trong

video thành các tệp ảnh để sử dụng trong các chương trình khác. Các định dạng ảnh kết quả thường là : BMP, GIF, TIFF, Targa

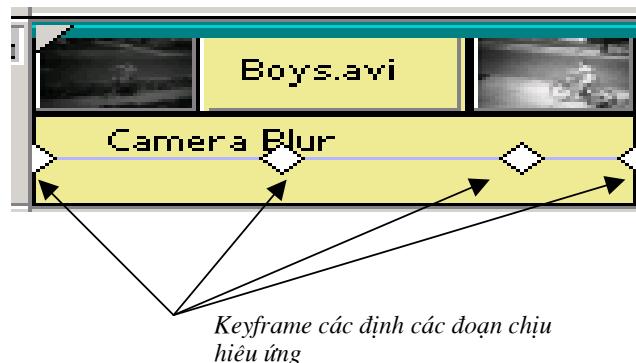
VI. MỘT SỐ KỸ THUẬT XỬ LÝ VIDEO-AUDIO TRÊN MÁY TÍNH

1. Kỹ thuật đánh dấu (Marker) và keyframe

Đánh dấu (marker) cho biết các điểm quan trọng trong một chuỗi các khuôn hình hay một số thời điểm nào đó trong một đoạn video. Đánh dấu giúp chúng ta biết được khu vực Video-Audio cần xử lý. Thường người ta dùng Marker để, chọn, xoá hoặc chuyển đến một đoạn Video một cách dễ dàng hơn.

Người ta chỉ dùng Marker trong chế độ soạn thảo và xử lý Video-Audio. Marker không tồn tại trong các tệp Video.

Trong các chương trình xử lý phim ảnh kỹ thuật keyframe dùng để đánh dấu khu vực Frame hoặc các Frame chịu tác động của một hiệu ứng nào đó. Ví dụ chúng ta có một đoạn video A (5 giây). Hai giây đầu của video A chúng ta muốn có hiệu ứng mở dần, hai giây tiếp theo có hiệu ứng rõ dần và giây cuối cùng có hiệu ứng sóng. Lúc này chúng ta phải sử dụng 4 keyframe để xác định 3 đoạn chịu các hiệu ứng trên.



2. Chuyển cảnh (Transition)

Chuyển cảnh (Transition) cho phép chúng ta tạo một hiệu ứng chuyển giữa cảnh A và cảnh B. Kỹ thuật này làm đoạn Video-Audio được mềm mại hơn.

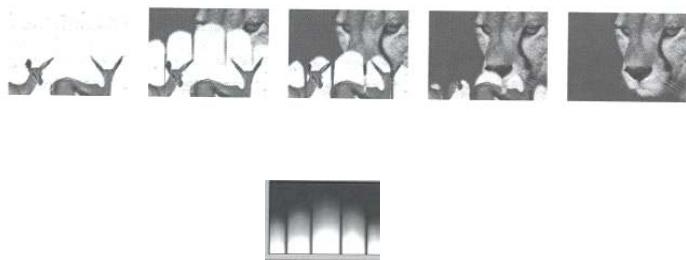
Ví dụ: Ta có cảnh cuối của đoạn video A là



Ta có cảnh đầu của đoạn video B là



Kết quả sau khi tạo chuyển cảnh



Lớp Mask giữa hai Video A và B:

(lớp Mask là một cảnh màu xám với dải màu 256 màu chuyển từ màu trắng sang đen và có độ trong suốt theo sắc độ sáng. Tại các điểm màu đen hơn thì cảnh ở phía dưới lớp Mask được hiển thị rõ hơn. Tại các điểm màu sáng hơn thì cảnh ở dưới lớp Mask hiển thị mờ hơn.

Trong chuyển cảnh trên người ta đã sử dụng một lớp Mask ở giữa hai đoạn video. Lớp này có tác dụng che dần cảnh A và hiện dần cảnh B. Hình dạng và độ trong suốt của lớp Mask có thể điều chỉnh được trong các chương trình soạn thảo Video-Audio.

3. Kỹ thuật trộn (mix)

Trộn hai kênh Video

Trong nhiều chương trình soạn thảo Video-Audio có hỗ trợ kỹ thuật này. Giả thiết ta có hai đoạn Video A và B được đặt chồng lên nhau. Thông thường nếu đặt như vậy thì ta chỉ có thể nhìn thấy được Video ở phía trên. Để nhìn thấy được cảnh ở của video phía dưới thì chúng ta cần phải thay đổi độ trong suốt của các điểm ảnh ở Video phía trên. Như vậy nếu ta thay đổi độ trong suốt của các điểm ảnh một cách đáng kể của video phía trên thì chúng ta có được một cảnh mà có cả hình ảnh Video A và cả hình ảnh video B.

Cảnh video A



Cảnh video B



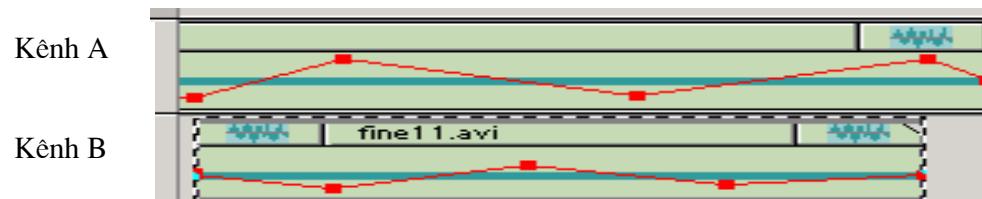
Đặt Video A chồng lên video B và thay đổi độ trong suốt của cảnh A ta thu được cảnh:



Sử dụng kỹ thuật này chúng ta có thể tạo được rất nhiều hiệu ứng cho Video từ các đoạn video đơn lẻ.

Trộn 2 kênh âm thanh

Nếu có hai kênh âm thanh A và B, quá trình trộn hai kênh là việc phát cùng một lúc hai kênh âm thanh với độ lớn khác nhau. Kết quả, chúng ta sẽ thu được một âm thanh hỗn tạp của hai kênh A và B.

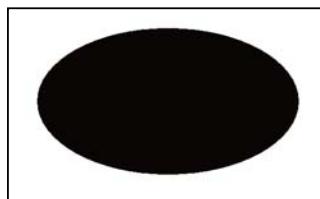


Đường gấp khúc thể hiện cường độ âm thanh của mỗi kênh tại các thời điểm khác nhau.

4. Kỹ thuật tạo độ trong suốt (Transparency)

Transparence là một kỹ thuật được áp dụng rất nhiều trong việc xử lý đồ họa. Đây là kỹ thuật làm mất một hoặc nhiều màu xác định cho một đoạn Video. Ví dụ chúng ta có hai đoạn Video-Audio A và B. Video A chồng lên video B. Video A có màu nền là màu đen. Chúng ta có thể dùng kỹ thuật này làm mất màu đen trên video A và như vậy tại các điểm có màu đen trên Video A sẽ không xuất hiện. Thay vào đó là các điểm trong suốt. Kết quả là tại các điểm có màu đen của video A ta có thể nhìn thấy các điểm ảnh của Video B. Còn tại các điểm khác chúng ta nhìn thấy video A.

Cảnh video A



Cảnh video B



Đặt Video A chồng lên video B và đặt Transparency cho màu đen trên video A ta thu được cảnh:



Trong các trường hợp làm mất một dải màu xám từ trắng đến đen (256 màu) thì chúng ta thu được kết quả :



Chú ý khi sử dụng kỹ thuật này màu bị mất không trùng với màu trong cảnh Video. Nếu trùng sẽ gây ra lỗi “thủng hình” tức là sẽ có một số vùng màu trong đối tượng chính mất đi không theo ý muốn

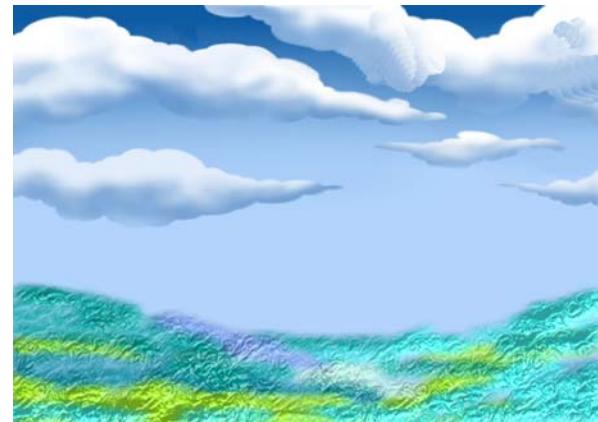
4.1 Kỹ thuật sử dụng kênh Alpha

Kỹ thuật sử dụng kênh Alpha là một kỹ thuật tiên tiến thường được sử dụng nhiều trong các chương trình soạn thảo phim và tạo các kỹ xảo cho phim ảnh. Phương pháp này dựa trên việc đánh dấu đường biên kín của các đối tượng trên một khuôn hình và lưu đường biên dưới dạng một kênh Alpha. Khi tiến hành ghép hai khuôn hình của hai đoạn phim với nhau người ta sẽ làm mất đi khu vực phía ngoài của tượng chứa kênh Alpha và như vậy nó sẽ đảm bảo chất của hình ảnh trong khu vực được đánh dấu.

Chúng ta có Video A là một nhân vật như cảnh bên:



Giả thiết chúng ta có Video B là một cảnh nền như cảnh bên:

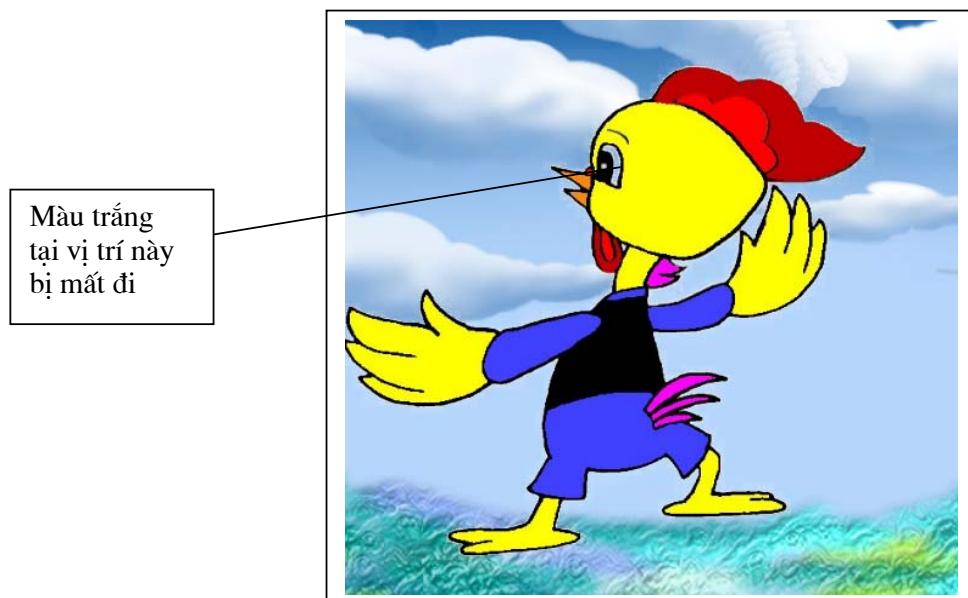


Đặt Video A chồng lên B .

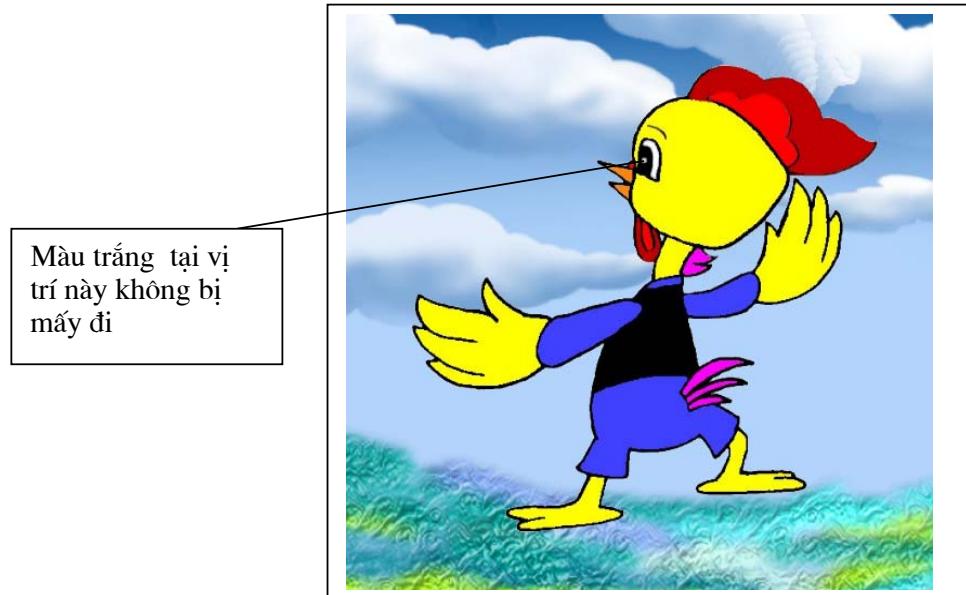
Nếu dùng kỹ thuật trộn hai kênh video A và B thì ta sẽ có kết quả là hai video ở chế độ không rõ nét và có nhiều phần hòa trộn với nhau.



Nếu dùng kỹ thuật Transparency làm mất màu trắng thì có hiện tượng phần ngoài của mắt gà cũng sẽ bị mất. Kết quả đạt chất lượng không cao



Trong trường hợp này chỉ có thể dùng kỹ thuật kênh Alpha. Lưu lại đường biên của đối tượng dưới dạng kênh Alpha và làm mất phần ngoài của kênh Alpha . Như vậy chất lượng đạt được là rất cao.



Kênh Alpha dùng cho khung hình trên video A có dạng sau :



Khi sử dụng kênh Alpha chúng ta cần lựa chọn định dạng ảnh hợp lý để lưu kênh Alpha. Hiện nay có định dạng ảnh TIFF đây là một định dạng theo chuẩn công nghiệp có chứa kênh Alpha. Khi sử dụng định dạng này chúng ta có thể tạo ra được những kỹ xảo chất lượng cao.

Tài liệu tham khảo

1. Adobe Primerie 6.0 uses guide
2. DV300 uses guide
3. Bài giảng của chuyên gia Yichi Kogure
4. <http://mpeg.telecomitalialab.com/standards/mpeg-1/mpeg-1.htm>
5. http://mpeg.telecomitalialab.com/mpeg_books.htm
6. <http://wwwam.hhi.de/mpeg-video/contact.htm>
7. http://www.ebu.ch/trev_dolby_frm.html
8. <http://www.msstate.edu/movies/>
9. http://wwwam.hhi.de/mpeg-video/#MPEG_Video_Group
10. <http://mp3.lycos.com/>