

**BỘ NÔNG NGHIỆP VÀ PHÁT TRIỂN NÔNG THÔN
TRƯỜNG CAO ĐẲNG NGHỀ CƠ ĐIỆN XÂY DỰNG VIỆT XÔ
KHOA: CÔNG NGHỆ THÔNG TIN VÀ NGOẠI NGỮ**

GIÁO TRÌNH
MÔ ĐUN: CÔNG NGHỆ ĐA PHƯƠNG
TIỆN
NGHỀ: TIN HỌC VĂN PHÒNG
TRÌNH ĐỘ: TRUNG CẤP

TUYÊN BỐ BẢN QUYỀN

Tài liệu này thuộc loại sách giáo trình nên các nguồn thông tin có thể được phép dùng nguyên bản hoặc trích dùng cho các mục đích về đào tạo và tham khảo.

Mọi mục đích khác mang tính lệch lạc hoặc sử dụng với mục đích kinh doanh thiếu lành mạnh sẽ bị nghiêm cấm.

LỜI GIỚI THIỆU

Giáo trình Công nghệ Multimedia được biên soạn nhằm đáp ứng yêu cầu học tập của sinh viên bước đầu làm quen với lý thuyết của Công nghệ Đa phương, đồng thời giúp cho sinh viên có một tài liệu học tập, rèn luyện tốt khả năng thiết kế, tạo nền tảng vững chắc cho các môn học tiếp theo.

Giáo trình không chỉ phù hợp cho người mới bắt đầu mà còn phù hợp cho những người cần tham khảo.

Khi biên soạn, chúng tôi đã tham khảo các giáo trình và tài liệu giảng dạy môn học này của một số trường Cao đẳng, Đại học để giáo trình vừa đạt yêu cầu về nội dung vừa thích hợp với đối tượng là sinh viên của các trường Cao đẳng Nghề.

Chúng tôi hy vọng sớm nhận được những ý kiến đóng góp, phê bình của bạn đọc về nội dung, chất lượng và hình thức trình bày để giáo trình này ngày một hoàn thiện hơn.

Ninh Bình, ngày..... tháng.... năm 2018

Tham gia biên soạn

1. Chủ biên: Nguyễn Thị Thanh

MỤC LỤC

TUYÊN BỐ BẢN QUYỀN	2
LỜI GIỚI THIỆU	2
MÔ ĐUN CÔNG NGHỆ MULTIMEDIA	6
Bài 1: Các kiến thức cơ bản	6
1. Dữ liệu Multimedia	6
1.1. Ảnh	8
1.1.1. Ảnh và ứng dụng.....	9
1.1.2. Thu ảnh	9
1.1.3. Kỹ thuật nén.....	10
1.2. Âm thanh.....	14
1.2.1. Các ứng dụng âm thanh	14
1.2.2. Ghi âm thanh.....	14
1.2.3. Kỹ thuật nén.....	15
1.3. Video	15
1.3. 1. Các ứng dụng video	15
1.3.2. Nén video	16
2. Công nghệ Multiedia.....	21
3. Các dịch vụ Multimedia.....	22
3.1. Hệ thống hội thảo truyền hình (Videoconference Systems).....	23
3.1.1. Yêu cầu thiết bị	23
3.1.2. Yêu cầu đường truyền.....	23
3.2. Hệ thống truyền hình theo yêu cầu (Video-on-demand System)(VOD)	25
THỰC HÀNH ĐA PHƯƠNG TIỆN	27
CÂU HỎI BÀI TẬP	39
HƯỚNG DẪN TRẢ LỜI.....	40
BÀI 2: CÁC CHUẨN CÔNG NGHỆ MULTIMEDIA	44
1. Chuẩn MPEG trong Audio.....	44
1.1. Giới thiệu chung về MPEG	44
1.2. Kỹ thuật Audio	47
1.2.1. Khái niệm.....	47
1.2.2. Ứng dụng.....	48
1.2. Nguyên lý nén dữ liệu	51

1.2.1. Tầm quan trọng của nén dữ liệu	51
1.2.2. Các nguyên tắc cơ bản của nén dữ liệu.....	51
2. Chuẩn MPEG trong Video	52
2.1. Các tiêu chuẩn nén ảnh	52
2.2. Kỹ thuật nén ảnh JPEG	53
2.3. Ứng dụng MPEG-2 trong nén tín hiệu video	56
2.4. Tiêu chuẩn MPEG-4	57
2.5. Chuẩn H.261	58
3. Multimedia trên Internet	58
3.1. Internet là gì?	58
3.2. Kết nối Internet bằng cách nào?.....	58
3.3. Các máy tính trong Internet giao tiếp với nhau bằng cách nào?	59
3.4. Một số ứng dụng của Internet đối với công nghệ Multimedia	59
a. Hệ thống hội thảo truyền hình (Videoconference Systems)	59
b. Hệ thống truyền hình theo yêu cầu (Video-on-demand System)	61
CHƯƠNG 3: TÍCH HỢP DỮ LIỆU ĐA PHƯƠNG TIỆN	63
1. Sản phẩm audio	63
1.1. Quản trị sản phẩm	63
1.2. Quản trị sản phẩm audio.....	63
1.3. Phòng thu	63
1.4. Số hoá	64
1.5. Một số vấn đề liên quan đến âm thanh	64
1.6. Kết luận	64
2. Sản phẩm video	64
2.1. Quản lý sản phẩm video	64
2.2. Nguyên tắc cơ bản của video	65
2.3. Các chuẩn truyền hình video	66
3. Sản phẩm hoạt hình	66
4. Tích hợp dữ liệu multimedia	68
4.1. Các công cụ lập trình	68
4.2. Các giai đoạn	68
5. Bản quyền.....	69
6. Tạo hồ sơ sản phẩm.....	70
6.1. Cần thiết của tư liệu.....	70

6.2. Kết thúc đề án.....	70
6.3. Kết luận.....	70
YÊU CẦU VỀ ĐÁNH GIÁ KẾT QUẢ HỌC TẬP	71
1. Nội dung.....	71
2. Phương pháp.....	72
TÀI LIỆU THAM KHẢO.....	73

TaiLieu.vn

MÔ ĐUN CÔNG NGHỆ ĐA PHƯƠNG TIỆN

Mã mô đun: MĐ 22

Vị trí, tính chất, ý nghĩa và vai trò của mô đun:

- Vị trí: Môn học được bố trí sau khi người học học xong các môn học chung và trước các môn học cơ sở về đồ họa.
- Tính chất : Là môn học tự chọn rèn luyện kỹ năng đồ họa sử dụng trong công nghệ multimedia.
- Ý nghĩa và vai trò: Đây là mô đun chuyên môn nghề cung cấp cho sinh viên các kỹ năng cơ bản nhất về công nghệ Multimedia, xây dựng các sản phẩm đa phương tiện.

Mục tiêu của mô đun:

- Trang bị cho người học các kiến thức quản lý dự án đa phương tiện; về các loại dữ liệu đa phương tiện; về tích hợp dữ liệu đa phương tiện;
- Rèn luyện kỹ năng về tích hợp dữ liệu đa phương tiện; về thu thập dữ liệu ảnh, hình động, video, âm thanh.

Nội dung của mô đun:

Bài 1: Các kiến thức cơ bản

Mã mô đun: MĐ 22-01

Giới thiệu:

Bài giảng giới thiệu về dữ liệu multimedia, công nghệ và dịch vụ multimedia

Mục tiêu:

- Trang bị các kiến thức cơ sở về đa phương tiện. Cung cấp kỹ năng về thu thập dữ liệu đa phương tiện.
- Rèn luyện tư duy, sáng tạo, tác phong công nghiệp

Nội dung chính:

1. Dữ liệu Multimedia

Mục tiêu:

- Trình bày được khái niệm dữ liệu multimedia, kỹ thuật nén ảnh, video, âm thanh.
- Phân biệt được các kỹ thuật nén ảnh, video, âm thanh

Trước tiên người ta có thể hỏi đa phương tiện là gì ? Đa phương tiện là tích hợp của văn bản, âm thanh, hình ảnh của tất cả các loại và phần mềm có điều khiển trong một môi trường thông tin số. Định nghĩa về đa phương tiện sẽ đề cập sau. Dữ liệu đa phương tiện gồm dữ liệu về :

- Văn bản
- Hình ảnh
- Âm thanh
- Hình động

Khái niệm về multimedia

Con người có nhu cầu diễn tả các trạng thái của mình; và họ có nhiều loại hình thể hiện. Con người có nhu cầu truyền thông, do đó cách thể hiện trên đường truyền rất quan trọng. Trên Internet thông dụng với mọi người, cái đẹp của trang Web phải được thể hiện cả ở nội dung và hình thức.

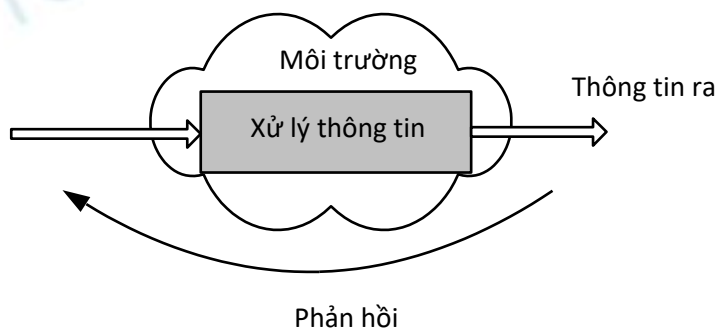
Đa phương tiện có nhiều loại, những phương tiện công cộng về đa phương tiện: Radio, vô tuyến, quảng cáo, phim, ảnh...

Dữ liệu đa phương tiện được chia thành hai lớp là các dữ liệu liên tục và các dữ liệu không liên tục. Các dữ liệu liên tục bao gồm các dữ liệu âm thanh, video thay đổi theo thời gian. Các dữ liệu không liên tục là các dữ liệu không phụ thuộc vào thời gian, các loại dữ liệu đặc trưng cho dạng này là các dữ liệu văn bản (có hoặc không có định dạng), hình ảnh tĩnh và các đối tượng đồ họa.

Các kiểu dữ liệu thông thường của một CSDL multimedia bao gồm:

- Dữ liệu văn bản (có hoặc không có định dạng).
- Đồ họa: là các bản vẽ, minh họa được mã hóa như các tệp postscript.
- Hình ảnh: là các hình ảnh được mã hóa sử dụng các dạng thức chuẩn như là JPEG hoặc MPEG.
- Các hoạt hình.
- Âm thanh. Hay video

Nhu cầu về tương tác người-máy luôn đặt ra trong hệ thống thông tin. Vấn đề chính về tương tác người-máy không là quan hệ giữa con người với máy tính mà là con người với con người. Con người có vai trò quan trọng trong hệ thống thông tin.



Hình 1.1. Hệ thống thông tin

Định nghĩa đa phương tiện (theo nghĩa rộng) là bao gồm các phương tiện: văn bản, hình vẽ tĩnh (vẽ, chụp), hoạt hình (hình ảnh động), âm thanh

Cuối cùng người ta có thể định nghĩa đa phương tiện: *đa phương tiện là kỹ thuật mô phỏng và sử dụng đồng thời nhiều dạng phương tiện chuyển hoá thông tin và các tác phẩm từ các kỹ thuật đó*

Liên quan đến định nghĩa đa phương tiện, người ta cần lưu ý những khía cạnh sau:

- Thông tin cần phải được số hoá, phù hợp với xu thế và rẻ;
- Phải dùng mạng máy tính, để đảm bảo truyền bá, hay truyền tải tốt;
- Sử dụng phần mềm có tương tác, cho phép người dùng trao đổi với phần mềm và thay đổi theo ý người dùng;
- Phải thiết kế giao diện người máy phù hợp với phát triển của đa phương tiện, tức giao

Kế hoạch về thu thập dữ liệu đa phương tiện. Thu thập dữ liệu: Một cách thể hiện đa phương tiện là liệt kê các dữ liệu thu thập được. Một số thành phần cơ bản trong một thể hiện đa phương tiện đã được liệt kê gồm

- (i) văn bản;
- (ii) hình ảnh;
- (iii) âm thanh;
- (iv) hình động và phim (movie);

Lên kế hoạch về dữ liệu ảnh tĩnh : ảnh tĩnh là hình ảnh được số hoá hay ảnh chụp, nhờ máy quét hay máy ảnh số. Muốn lên kế hoạch về ảnh tĩnh đòi hỏi tổ chức nêu lên các nhu cầu về ảnh... Các chi tiết trong danh sách ảnh tĩnh là

- (i) chỉ số của ảnh;
- (ii) tên file ảnh;
- (iii) chỉ số cảnh;
- (iv) kích thước file dữ liệu ảnh;
- (v) màu sắc;
- (vi) ghi chú đi kèm ảnh;

Lên kế hoạch về dữ liệu ảnh động, dữ liệu video. Cũng như ảnh tĩnh, nhưng kế hoạch về dữ liệu ảnh động cần có thêm chi tiết dạng nén hay dạng mã hoá các khung hình trong đoạn ảnh động;

Lên kế hoạch về âm thanh: âm thanh có thể là tiếng người, âm thanh nhạc cụ hay âm tổng hợp. Dữ liệu âm thanh được số hoá trên đĩa quang từ, xử lý trên phần mềm âm thanh, sử dụng micro, máy tính đa phương tiện, bìa âm thanh.

Chuẩn bị dữ liệu

Chuẩn bị dữ liệu đa phương tiện nhằm có đủ dữ liệu, thông tin để tích hợp thành sản phẩm đa phương tiện. Do người ta chia ra 4 loại dữ liệu chính, nên việc chuẩn bị dữ liệu đa phương tiện theo 4 loại dữ liệu là:

- (i) văn bản;
- (ii) hình ảnh;
- (iii) hình động;
- (iv) âm thanh.

• Tạo và số hoá ảnh động

Hình động đòi hỏi nhiều công sức của người sản xuất. Người ta có thể dựng nhiều hình tĩnh để tạo ra hình động. Việc dùng các phần mềm tạo hình động là một xu hướng; tuy nhiên một số hình động được dựng trực tiếp từ các đoạn video.

• Quay và số hoá dữ liệu video

Việc tạo dữ liệu động sẽ được trình bày trong phần sau. Trong phần đầu, cần lưu ý đến tính phức tạp của hình động và yêu cầu về kỹ thuật trong việc xử lý dữ liệu hình động.

Để có đoạn video trên máy, cần :

- Quay tại hiện trường;
- Đưa đoạn video vào máy tính; các dữ liệu tương tự được chuyển sang dạng số;
- Chọn chuẩn nén cho file dữ liệu hình động, phù hợp với yêu cầu tích hợp sau này.

MỘT SỐ DỮ LIỆU ĐA PHƯƠNG TIỆN

1.1. Ảnh

1.1.1. Ảnh và ứng dụng

Hiện tại người ta đòi hỏi các ứng dụng máy tính xử lý nhiều loại ảnh khác nhau trong nhiều ứng dụng khác nhau. Nhu cầu của họ thay đổi tùy theo loại ảnh cần hỗ trợ. Ảnh bitonal (trắng và đen) bao gồm văn bản trong các tài liệu kinh doanh như thư từ hay các giấy khổ A4. Thông thường những ảnh này được quét và lưu trữ trong file folder để sử dụng trong các ứng dụng. Công nghệ lưu trữ và quét quang học cũng đang thay thế microform trong hệ quản lý hồ sơ, nơi lưu trữ các tài liệu như bằng sáng chế, báo cáo y khoa, mẫu đơn thức và báo cáo ngân hàng. Những đề mục nhỏ như biên lai, séc và thẻ tiện dụng được xử lý trong hệ thống xử lý giao dịch khối lượng lớn.

Một loại ảnh bitonal thứ 2, được gọi là line art, bao gồm các đồ họa kỹ thuật trong ứng dụng thiết kế được máy tính hỗ trợ (CAD), biểu đồ trong sổ tay kỹ thuật dành cho lĩnh vực quốc phòng và hàng không, lược đồ, lưu đồ, sơ đồ mạch, bản đồ và hoạt hình. Một số tài liệu kinh doanh như đơn từ, là tổng hợp nhiều dòng, văn bản in và menuscrypt, để xử lý những ảnh như vậy cần phải sử dụng hỗn hợp công nghệ nhận dạng và quét.

Ảnh chụp, ảnh nửa tông hoặc khung đơn là các loại ảnh tông liên tục có thang độ xám hoặc màu. Ảnh thang độ xám chứa đựng nhiều bóng xám. Chúng được sử dụng trong các ứng dụng như dàn trang và các thư viện cho việc biên soạn và phát hành các bài báo hay các ứng dụng về khoa học kỹ thuật như không ảnh, thông tin vệ tinh và dữ liệu về động đất. Thông thường các ứng dụng này yêu cầu ảnh phải có chất lượng cao hơn ảnh hệ thống xử lý tài liệu đã được đề cập trước đó. Chẳng hạn, nhờ vào các ảnh y khoa chụp từ máy quét ảnh cộng hưởng từ MRI và máy quét chụp cắt lớp bằng tia X dưới sự hỗ trợ của máy tính, các bác sĩ có thể chẩn đoán bệnh từ xa thông qua tia phóng xạ.

Các ứng dụng chuyển biết được thiết lập riêng cho loại ảnh màu (đa quang phổ) chẳng hạn như sách cũ và bản thảo như ở thư viện hoặc ảnh hội họa chất lượng cao cả các đề mục trưng bày trong viện nghệ thuật và viện bảo tàng. Nhu cầu về ảnh chụp có màu trong hệ thống truyền thông đa phương tiện thường ngày như các loại ứng dụng cũng tăng lên. Điển hình là hiện thời, người tiêu dùng và các chuyên gia có thể xử lý và lưu trữ ảnh màu trên đĩa compact ảnh để sau đó hiển thị chúng trên màn hình máy tính hoặc truyền hình. Trong các buổi trình bày trong kinh doanh, các doanh nghiệp có thể sử dụng bộ sưu tập ảnh trên đĩa mềm hoặc CD-ROM.

Ảnh có thể được ứng dụng vào nhiều lĩnh vực hiện đại hơn nữa, tuy nhiên nếu kết hợp giữa ảnh và các công nghệ khác, chẳng hạn như hệ cơ sở tri thức và thuật toán so khớp mẫu - con người sẽ bước vào kỷ nguyên dần triển vọng hơn và sự kết hợp đó phục vụ cho quá trình điều tra và phát triển, chẳng hạn như dấu tay và ảnh chụp có mục đích nhận diện trong an ninh.

1.1.2. Thu ảnh

Thông thường hầu hết các loại ảnh đề cập như trên đều được thu giữ bằng máy chụp hay máy quét quang học có công dụng chuyển đổi ảnh vào mảng điểm hình chữ nhật gọi là các phần tử ảnh (pixels). Hệ quét quang học bao gồm một nguồn sáng, một giá đỡ tài liệu và một bộ dò ánh sáng. Sau mỗi lần chạy, ánh

sáng phản xạ được chuyển đổi thành tín hiệu điện, và sau đó sẽ được chuyển đổi dạng số để xử lý và lưu trữ thành mảng phần tử ảnh, kích thước của mảng này phụ thuộc vào loại ảnh được thu:

- Ảnh bitonal chỉ có giá trị cường độ và do đó có lúc lưu giữ một bit một phần tử ảnh với giá trị là 1 hoặc 0.
- Ảnh thang độ xám có nhiều mức xám. Ảnh được lưu giữ trong n bit một phần tử ảnh, nơi mà tổng số độ xám là $2^n - 1$ (ví dụ, 1 ảnh có 15 độ xám + trắng cần được lưu giữ trong 4 bit một phần tử ảnh).
- Cường độ của ba màu chính và màu xám định rõ đặc điểm của ảnh màu. Số lượng màu hiện có trong n bit là $2^n - 1$ (ví dụ, cần 8bit một phần tử ảnh để lưu trữ một ảnh chứa 25màu + trắng).

Kích thước của mảng cũng phụ thuộc vào mật độ, đó là số lượng phần tử ảnh có trong 2.54cm theo một hướng. Thuật ngữ mật độ cũng được dùng để mô tả độ phân giải của máy quét linh theo số lượng điểm trong 2.54cm (dpi). Khi lựa chọn độ phân giải, cần phải xét đến độ phân giải của thiết bị xuất bởi vì chúng có mối quan hệ lẫn nhau. Chẳng hạn, độ phân giải của màn hình hiển thị máy tính nằm giữa 70 và 200 dpi. của máy in laser thông thường là 300 dpi, nhưng của máy in offset lên đến 1000 dpi.

Tốc độ thu giữ ảnh cũng thay đổi từ 3 trang A4 trong một phút (đối với loại máy quét để bàn dùng cho máy tính cá nhân) đến 30 trang A4 một phút (đối với loại máy quét tốc độ cao). Loại máy quét như thế thu ảnh của cả những đề mục nhỏ như biên lai, tín dụng hay chi phiếu séc phục vụ cho quá trình xử lý nghiệp vụ. Để thu ảnh, người ta cũng sử dụng loại máy quay số có độ phân giải cao - chẳng hạn như dùng 2 máy camera thu đồng thời nửa phần dưới của tài liệu để đạt tốc độ yêu cầu. Các mảng của thiết bị nạp phát (CCDS) được lắp đặt trong kiểu máy camera như thế.

Đối với loại ảnh có độ phân giải cao (tới 2200x1700) trong lĩnh vực nghệ thuật màu, người ta sử dụng máy quay ảnh hiện có thu giữ, cũng có thu được khung tĩnh từ chuỗi video động bằng bộ số hoá video hay bằng bộ chụp khung. Cần phải lắp đặt các thiết bị đặc biệt đa số hoá ảnh quét MRI và CT, giúp cho các bác sĩ có thể chẩn đoán thông qua ảnh quét được hiển thị trên màn hình có độ phân giải cao (2500 x 2000 phần tử có 256 độ xám). Thiết bị ra ảnh có thể là máy in đen trắng, máy in màu hay máy vẽ (ploter). Nhìn chung, các hệ thống thu nhận ảnh thực hiện hai quá trình:

- Cảm biến: biến đổi năng lượng quang học (ánh sáng) thành năng lượng điện
- Tổng hợp năng lượng điện thành ảnh

1.1.3. Kỹ thuật nén

– Tại sao phải nén

Cần rất nhiều byte để hiển thị một ảnh chưa nén. Lấy một mặt giấy A4 làm ví dụ. Như mô tả ở trên, máy quét có thể thu thông tin trên giấy theo thang độ xám hay bitonal. Sau đó dữ liệu qua thường được lưu giữ tạm thời trên đĩa từ, Bảng 1.1 cho thấy số lượng không gian lưu trữ mà tài liệu này chiếm khi nó được quét với mức độ 200, 300, và 400 dpi

Bảng 1.1 - Yêu cầu lưu trữ của khổ giấy A4 chưa nén

Độ phân giải dpi	Bitonal(Mb)	Thang độ xám (Mb)	Màu sắc (Mb)
200	0,48	1,9 – 7,7	15-61
300	1,09	4,4 – 17,4	35-140
400	1,93	7,7 – 30,9	62-247

Trong đó: 1 tờ giấy A4 có kích thước 210 x 297mm hoặc 8,27 x 11,69mm

Ảnh bitonal cần 1 bit / một phần tử

Ảnh thang độ xám cần 4 - 6 bit / phần tử ảnh

Ảnh màu cần 32 - 128 bit / 1 phần tử ảnh

Để giảm bớt khoảng không lưu trữ tài liệu, ảnh phải được chuyển đổi sang dạng khác và nhỏ hơn bằng cách loại bỏ những thông tin dư thừa. Nói 1 cách khác, ảnh cần phải được nén lại để giảm không gian lưu trữ. Một số phương pháp nén ảnh sẽ được trình bày chi tiết trong các phần tiếp theo.

– **Nén ảnh, JPEG**

Công nghệ nén ảnh JPEG (Joint Photographic Experts Group) là một trong những công nghệ nén ảnh hiệu quả, cho phép làm việc với các ảnh có nhiều màu và kích cỡ lớn, tỷ lệ nén ảnh đạt mức so sánh tới vài chục lần (chứ không phải phần trăm). Tuy nhiên được cái này bạn phải mất cái khác, đó là quy luật cộng trừ tự nhiên.

Thông thường các ảnh màu hiện nay dùng 8 bit (1 byte) hay 256 màu thay cho từng mức cường độ của các màu đỏ, xanh lá cây và xanh da trời. Như thế mỗi điểm của ảnh cần 3 bít để lưu mã màu, và lượng byte một ảnh màu này chiếm gấp 24 lần ảnh trắng đen cùng cỡ. Với những ảnh này các phương pháp nén ảnh như IFF (Image File Format) theo phương pháp RLE (Run Length Encoding) không mang lại hiệu quả vì hệ số nén chỉ đạt tới 2:1 hay 3:1 (tất nhiên là kết quả nén theo phương pháp RLE phụ thuộc vào cụ thể từng loại ảnh, ví dụ như kết quả rất tốt và các loại ảnh ít đổi màu). Ưu điểm cao của phương pháp nào là ảnh đã nén sau khi bung sẽ trùng khớp với ảnh ban đầu. Một số phương pháp nén khác không để mất thông tin như của Lempel - Ziv - Welch (LZW) có thể cho hệ số nén tới 6:1. Nhưng như thế cũng chưa thật đáp ứng yêu cầu đòi hỏi thực tế.

Phương pháp nén ảnh theo chuẩn JPEG có thể cho hệ số nén tới 80:1 hay lớn hơn, nhưng bạn phải chịu mất thông tin (ảnh sau khi bung nén khác với ảnh ban đầu), lượng thông tin mất mát tăng dần theo hệ số nén. Tuy nhiên sự mất mát thông tin này không bị làm một cách cầu JPEG tiến hành sửa đổi thông tin ảnh khi nén sao cho ảnh mới gần giống như ảnh cũ, khiến phần đông mọi người không nhận thấy sự khác biệt. Và bạn hoàn toàn có thể quản lý sự mất mát này bằng cách hạn chế hệ số nén. Như thế người dùng có thể cân nhắc giữa cái lợi của việc tiết kiệm bộ nhớ và mức độ mất thông tin của ảnh, để chọn phương án thích hợp.

Phương pháp nén ảnh JPEG dựa trên nguyên lý sau: ảnh màu trong không gian của 3 màu RGB (red Green Blue) được biến đổi về hệ YUV (hay YCBCr) (điều này không phải là nhất thiết, nhưng nếu thực hiện thì cho kết quả nén cao hơn). Hệ YUV là kết quả nghiên cứu của các nhà sản xuất vô tuyến truyền hình

hệ Pal, Secam và NTSC, nhận thấy tín hiệu video có thể phân ra 3 thành phần Y, U, V (cũng như phân theo màu chuẩn đỏ, xanh lá cây và xanh da trời). Và một điều thú vị là hệ nhân thị của con người rất nhạy cảm với thành phần Y và kém nhạy cảm với hai loại U và V, phương pháp JPEG đã nắm bắt phát hiện này để tách những thông tin thừa của ảnh, hệ thống nén thành phần Y của ảnh với mức độ ít hơn so với U, V, bởi người ta ít nhận thấy sự thay đổi của U và V so với Y.

Giai đoạn tiếp theo là biến đổi những vùng thể hiện dùng biến đổi cosin rời rạc (thông thường là những vùng 8×8 pixel). Khi đó thông tin về 64 pixel ban đầu sẽ biến đổi thành ma trận có 64 hệ số thể hiện "thực trạng" các pixel. Điều quan trọng là ở đây hệ số đầu tiên có khả năng thể hiện "thực trạng" cao nhất, khả năng đó giảm rất nhanh với các hệ số khác. Nói cách khác thì lượng thông tin của 64 pixel tập trung chủ yếu ở một số hệ số ma trận theo biến đổi trên. Trong giai đoạn này có sự mất mát thông tin, bởi không có biến đổi ngược chính xác. Nhưng lượng thông tin bị mất này chưa đáng kể so với giai đoạn tiếp theo. Ma trận nhận được sau biến đổi cosin rời rạc được lược bớt sự khác nhau giữa các hệ số. Đây chính là lúc mất nhiều thông tin vì người ta sẽ vứt bỏ những thay đổi nhỏ của các hệ số. Như thế khi bung ảnh đã nén bạn sẽ có được những tham số khác của các pixel. Các biến đổi trên áp dụng cho thành phần U và V của ảnh với mức độ cao hơn so với Y (mất nhiều thông tin của U và V hơn). Sau đó thì áp dụng phương pháp mã hóa của Hoffman: phân tích dãy số, các phần tử lặp lại nhiều được mã hóa bằng ký hiệu ngắn (marker). Khi bung ảnh người ta chỉ việc làm lại các bước trên theo quá trình ngược lại cùng với các biến đổi ngược.

Vì phương pháp này thực hiện với các vùng ảnh (thông thường là 8×8 pixel) nên hay xuất hiện sự mất mát thông tin trên vùng biên của các vùng (block) này. Hiện nay người ta đã giải quyết vấn đề này bằng cách làm trơn ảnh sau khi bung nén để che lấp sự khác biệt của biên giới giữa các block. Một hệ nén ảnh theo chuẩn JPEG cùng algorithm làm trơn ảnh đã được công ty ASDG đưa ra trong hệ Art Department Professional.

– Nén Fractal

Tất cả các phương pháp nén ảnh đều dựa trên một nguyên lý đơn giản: trong dữ liệu có nhiều phần tử thừa và nén ảnh dựa trên cơ sở tìm ra những phần tử đó và mã hóa chúng. Ví dụ, như số 9999997777 có thể mã hóa thành 6947. Các hình ảnh trên màn hình máy vi tính đặc trưng bởi số điểm (pixel) là số bit dành cho mã màu của mỗi điểm (bit/pixel).

Phần lớn các hình ảnh (nhất là có độ phân giải cao) không có quy luật giữa các điểm gần nhau, do đó các phương pháp thông dụng hiện nay như biến đổi cosin rời rạc. Wavelet Image Compression (WIC) (theo chuẩn JPEG và MPEG) phải dùng đến biến đổi toán học và xấp xỉ các mối tương quan giữa các pixel. Với các phương pháp nén ảnh fractal bạn có thể nén ảnh tới tỷ lệ 20:1 - 30:1. Nhưng những ảnh này (vì bị mất thông tin) chỉ là những ảnh gần đúng với ảnh ban đầu, ngoài ra còn có thể xuất hiện biến dạng hình ảnh như đối với phương pháp biến đổi cosin rời rạc.

a. Hình học Fractal và biến đổi Fractal

Một cuộc cách mạng trong vấn đề xử lý ảnh ghế giới thực" đã xảy ra cùng với sự ra đời cuốn sách " Hình học Fractal của tự nhiên (the Fractal Geometry of Nature) của tác giả Mandelbrot. Theo tác giả, khái niệm Fractal là cấu trúc thể hiện sự gần giống nhau về hình dạng của các hình thể kích cỡ khác nhau. Nếu bạn nghiền một củ khoai tây rán giòn bạn sẽ có vô số những mảnh vỏ lớn nhỏ, các mảnh này có thể gọi là Fractal, Mandelbrot chỉ ra rằng. Có thể tìm ra các cấu trúc và qui luật để tạo các hình dạng Fractal, do đó có thể coi Fractal như là các hình cơ bản của hình học phẳng Ô-cơ-lit cùng với đường thẳng, hình chữ nhật, hình tròn. Fractal không phụ thuộc vào độ phân giải của hình, đó là những hình ảnh nhỏ, có thể vẽ được bằng một bộ hữu hạn thuật toán như quay hình, co giãn, biến đổi từ một hình nào đó. Các phép toán trên thực hiện với các hệ số được gọi là hệ số affin. Một bức tranh có thể được Fractal hóa và ta có thể khôi phục nó ngòi các hệ số affin. Trên thực tế đối với các hình rất ngẫu nhiên thì các hệ số gian tìm được rất khó. Trước kia làm bằng tay, người ta phải mất hàng ngày, hàng tuần. Hiện nay công việc đó có thể làm trong 5 phút. Quá trình Fractal hóa đã được hãng Integrated Systems nghiên cứu và giữ bản quyền. Sau đây là một số bước của quá trình đó.

b. Nén hình ảnh

Chia ảnh thành những vùng không phủ nhau còn gọi là domen (chẳng hạn bằng các đường thẳng ngang và đứng). Các vùng này phải phủ kín hình ảnh.

Lấy bộ các vùng cơ sở, các vùng này không nhất thiết phủ kín bề mặt bức tranh. Thực hiện biến đổi Fractal. Với mỗi vùng domen ta tìm vùng cơ sở mà sau biến đổi affin xấp xỉ nhất với domen.

Lưu các hệ số nhìn vào file, File này gồm 2 phần: đầu thể chứa thông tin về vị trí các domen và vùng cơ sở sau đó là bảng các thông số affin cho từng domen.

c. Vẽ lại hình ảnh

Tạo hai hình ảnh cùng cỡ A và B. Cỡ các ảnh này có thể khác với ảnh ban đầu. Các ảnh này có thể là trắng hay đen. Biến đổi các điểm của A vào B. Để làm điều đó trước hết chia D thành các domen như quá trình nén ảnh trên, với mỗi domen của B ta thực hiện biến đổi affin áp dụng với vùng cơ sở A (Các hệ số affin lấy từ file). Kết quả có được ta ghi vào B.

Biến đổi giá trị của B vào A giống như lần trước, chỉ có điều đổi vị trí chúng.

Thực hiện biến đổi trên nhiều lần cho đến khi A và B không khác gì nhau.

Quá trình này dẫn đến việc là ta khôi phục được bức tranh ban đầu mà độ chính xác phụ thuộc vào độ chính xác của các biến đổi affin.

Thuật toán quá trình nén và giải nén ảnh được công ty Integrated Systems đưa ra sử dụng số học nguyên cùng các phương pháp làm giảm sự tăng dần của sai số trong các phép toán làm tròn. Các thuật toán đã được tối ưu về mặt thời gian thực hiện. Tuy thế quá trình nén ảnh đó phải thực hiện một khối lượng tính toán lớn nên đòi hỏi khá nhiều thời gian so với việc giải nén ảnh. Với máy 386, tốc độ 33MHz và màn hình VGA các trình thí nghiệm đã thử phim video màu với tốc độ 20 ảnh loại này trong một giây.

d. Những ưu điểm của phương pháp Fractal

Trong quá trình Fractal hóa, bạn sẽ nhận được bộ các chữ số rất nhỏ thể hiện hình ảnh. Do đó hệ số nén của phương pháp là rất lớn, tuy thế chất lượng ảnh sau khi nén được bảo đảm khá chính xác. Phương pháp rất hiệu quả với những ảnh có độ phân giải cao. Phương pháp này đã được áp dụng không những trong nén dữ liệu mà còn để thể hiện các mối quan hệ giữa các phần tử của các ảnh xạ.

1.2. Âm thanh

1.2.1. Các ứng dụng âm thanh

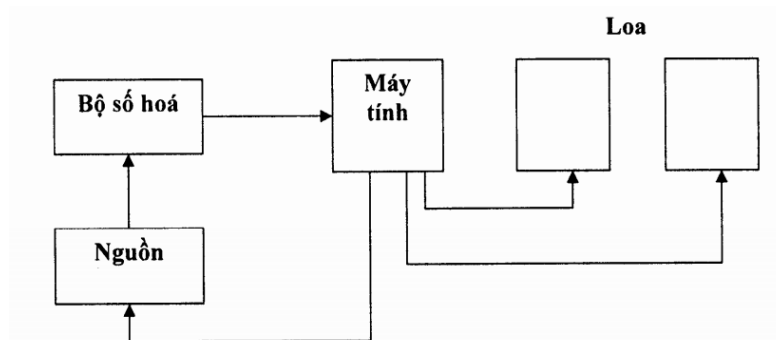
Âm thanh đóng vai trò quan trọng trong các ứng dụng truyền thông đa phương tiện. Các hiệu ứng đặc biệt của như âm nhạc và tiếng nói có thể được đưa vào các ứng dụng, đặc biệt là các ứng dụng trong hệ thống đào tạo và bán hàng tự động hoặc hệ thống điểm thông tin. Một lời chú thích bằng tiếng nói có thể được dùng để diễn tả những gì đang diễn ra trên màn hình hoặc để làm nổi bật và nhấn mạnh những khái niệm then chốt. Âm thanh có thể được sử dụng kết hợp với hình ảnh tĩnh hoặc động để giải thích cho người sử dụng một ý tưởng hay một quy trình hiệu quả hơn theo cách giải thích chỉ đơn giản bằng văn bản hay đồ họa, âm nhạc có thể được sử dụng để thu hút sự chú ý của khách hàng hoặc để tạo ra được một phong cách riêng biệt.

Trong một số lĩnh vực chuyên dụng tự âm thanh có thể hình thành nên được cái lõi của một ứng dụng truyền thông đa phương tiện, chẳng hạn như các hệ thống giúp cho người tàn tật nhìn thân được. Một dự án mới đây đã đưa đến việc chuyển tải nhật báo đến một thiết bị máy tính đặt tại nhà người sử dụng. Chỉ cần ngồi ở nhà, người sử dụng có thể chọn nghe hệ thống xử lý tiếng nói đọc lớn các bài báo đã được chọn lọc hoặc cho các bài báo đó hiển thị trên màn hình với kiểu chữ lớn. Một khi chi phí giảm và công nghệ được cải tiến thì mối quan tâm của người dùng đến việc sử dụng hệ thống xử lý và nhận dạng tiếng nói trong các ứng dụng kinh doanh nói chung sẽ tăng lên.

1.2.2. Ghi âm thanh

Bộ số hoá âm thanh được sử dụng để ghi và số hoá âm thanh tương tự từ băng âm thanh, đĩa ghi, CD-ROM và phiên bản đĩa compact audio gốc hay CD-DA (compact âm thanh kỹ thuật số). Có thể dùng micro để thu lại nhạc gốc hoặc sử dụng các nhạc cụ được cài sẵn trong máy tính để soạn nhạc thông qua giao diện MIDI.

Sau khi âm thanh được thu, âm thanh sẽ được lưu trữ ở đĩa cứng hoặc ở môi trường quang học. Nó có thể được hiệu chỉnh và phát lại qua bộ loa nối với máy tính hoặc qua tai nghe. Hiện tại máy tính có bộ xử lý âm thanh và loa cài sẵn trong máy. Tuy nhiên, hệ thống loa riêng biệt bên ngoài sẽ phát ra âm thanh hay hơn. Nếu nguồn âm thanh là ổ đĩa compact riêng cần phải kết nối với loa (hình 1.1)



Hình 1.1. Cấu hình hệ thống âm thanh

1.2.3. Kỹ thuật nén

Do con người dễ nhạy cảm với những thay đổi về chất lượng âm thanh hơn là chất lượng hình ảnh nên đòi hỏi cần phải có các hệ thống truyền thông đa phương tiện để hỗ trợ các chuẩn âm thanh cao. Hiện nay các kỹ thuật dùng để mã hoá thông tin đã được phát triển rất tốt.

Âm thanh được tạo ra từ các vi sai áp suất trong không khí. Micro tiếp nhận những vi sai này và phát ra thông qua bộ khuếch đại. Đầu tiên, tín hiệu tương tự này được số hoá bằng bộ chuyển mã tương tự sang tín hiệu số (ADC). Sau đó máy tính lấy mẫu dạng sóng nhập vào theo những khoảng cách đều đặn, rồi sử dụng phương pháp điều biến mã xung (PCM) để chuyển đổi biên độ sang mã nhị phân.

Đối với tiếng nói, tốc độ lấy mẫu là 8kHz và 8 bit - đại diện cho 256 giá trị biên độ khác nhau - được dùng để mã hoá mỗi mẫu. Kỹ thuật giới hạn số lượng như thế này được gọi là phép lượng tử hoá. Phương pháp mã hoá này sẽ tạo ra một dòng dữ liệu âm thanh liên tục 64.000 bit trong một giây (64kbit/s), sau đó được xếp thành từng bó tin để truyền qua mạng. Đối với loại nhạc có chất lượng phát từ đĩa compact, tốc độ lấy mẫu của tín hiệu là 44,1 kHz và dùng 16 bit để giải mã mỗi mẫu. Trong hệ âm thanh nổi, phương pháp mã hoá này sẽ tạo dòng dữ liệu âm thanh liên tục 1,4 triệu bit trong 1 giây.

Việc loại bỏ yên lặng hay dùng các phương pháp mã hoá tốt hơn thì có thể đạt được độ nén cao hơn nữa:

Phương pháp PCM phi tuyến ấn định các điểm giá trị biên độ phi tuyến.

PCM sai phân mã hoá vi sai của tín hiệu thay chỉ vì mã hoá tín hiệu này. Dãy vi sai thường nhỏ hơn dãy biên độ.

Phương pháp điều biến mã xung không tương ứng (ADPCM) điều chỉnh động dãy giá trị biên độ để tương ứng với dãy biên độ có trong dòng dữ liệu nhập.

1.3. Video

1.3.1. Các ứng dụng video

Các ứng dụng video được chia thành 2 nhóm:

Nhóm phát lại chất liệu nghe nhìn:

Nhóm truyền thông nghe nhìn ở thời gian thực

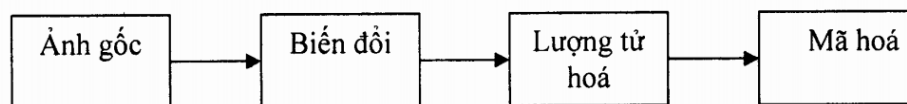
1.3.2. Nén video

Kỹ thuật nén ảnh số đang đóng một vai trò cực kỳ quan trọng trong các hệ thống viễn thông và multimedia để giải quyết vấn đề băng thông của đường truyền. Các kỹ thuật nén video đều cố gắng làm giảm lượng thông tin cần thiết cho một chuỗi các bức ảnh mà không làm giảm chất lượng của nó đối với người xem. Nói chung, tín hiệu video thường chứa đựng một lượng lớn các thông tin thừa, chúng thường được chia thành hai loại: thừa tĩnh bên trong từng frame (statistical) và thừa động giữa các frame (subjective). Mục đích của nén video là nhằm làm giảm số bit khi lưu trữ và khi truyền bằng cách phát hiện để loại bỏ các lượng thông tin dư thừa này và dùng các kỹ thuật Entropy mã hoá để tối thiểu hoá lượng tin quan trọng cần giữ lại.

Nén dữ liệu được chia thành hai dạng cơ bản: Nén không mất dữ liệu (Lossless) và nén có mất dữ liệu (Lossy). Đối với dạng nén không mất dữ liệu, ảnh được khôi phục hoàn toàn giống ảnh gốc, tuy nhiên điều này đòi hỏi phải có thiết bị lưu trữ và đường truyền lớn hơn. Các thuật toán của nén không mất dữ liệu thường dựa vào việc thay thế một nhóm các ký tự trùng lặp bởi một nhóm các ký tự đặc biệt khác ngắn hơn mà không quan tâm tới ý nghĩa của dòng bộ dữ liệu. Các ví dụ của dạng nén không mất dữ liệu là Run-length Encoding (RLE), Huffman Coding, Arithmetic coding, Shannon-Fano Coding, tz78, LZH, LZW....

Đối với dạng nén có mất dữ liệu, ảnh được khôi phục không giống hoàn toàn với ảnh gốc, dạng nén này thích hợp cho việc lưu trữ và truyền ảnh, video qua một mạng có băng thông hạn chế. Các dạng nén này thường cho hệ số nén cao hơn, nó liên quan tới việc dùng các phép biến đổi tín hiệu từ miền này sang miền khác. Các ví dụ của biến đổi có mất dữ liệu gồm: Differential Encoding, Discrete Cosine Transform(DCT), Vector Quantization, JPEG (Joint Photographic Experts Group) và MPEG (Motion Picture Experts Group).

Các phương pháp nén ảnh có mất tín hiệu gồm có 4 bước như hình 1.2.



Hình 1.2. Sơ đồ cơ bản của bộ mã hoá

Ảnh gốc được biến đổi theo nhiều cách khác nhau. Vào những năm 1980, việc nén và giải nén tín hiệu video dựa trên kỹ thuật DPCM (differential pulse code modulation) đã được CCITT chuẩn hoá theo tiêu chuẩn H.120. Các phương pháp nén dùng DPCM dựa trên nguyên tắc phát hiện sự giống nhau và khác nhau giữa các điểm ảnh (pixels) gần nhau để tìm cách loại bỏ các thông tin thừa. Tuy nhiên, chất lượng ảnh động không đạt được các yêu cầu cần thiết. Để cải thiện chất lượng ảnh động mà không làm tăng số lượng bộ so' với yêu cầu, kỹ thuật mã hoá chuyển sang dùng các phép biến đổi mà chúng có thể xử lý đồng thời một nhóm các pixels và ta có khái niệm về các bộ mã hoá trên các khối (block-based codecs). Đối với các bộ mã hoá trên các khối ảnh, mỗi điểm ảnh (pixel) sẽ cần ít hơn 1 bit để mã hoá.

Các bộ mã hoá khối có thể dựa trên hai nguyên tắc biến đổi cơ bản: Discrete Cosine Transform (DCT) và Vector Quantization (VQ). DCT được dùng để biến đổi các khối ảnh hai chiều có kích thước 8X8 từ miền không gian sang miền tần số.

Biến đổi DCT là tương tự như biến đổi DFT (Discrete Fourier Transform). Các hệ số DCT nhận được sẽ được lượng tử hoá (Quantization) và mã hoá (Encode).

Các hệ số DCT nhận được sẽ được lượng tử hoá (Quantisation coding) thành tập các hệ số đơn giản hơn nữa. Mục đích của nó là làm giảm hơn nữa số bit đặc trưng cho một hệ số. Tại bộ mã hoá sẽ có một bảng mã (code book) và bảng các chỉ số nội bộ, từ đó có thể chọn được các từ mã (code word) tương ứng một cách tốt nhất cho tập các hệ số được tạo ra. Quá trình lượng tử hoá cũng đồng thời làm tròn giá trị của các hệ số ở mức nhỏ hơn, đây chính là nguyên nhân gây ra mất tín hiệu, tuy vậy ảnh được khôi phục đạt chất lượng ở mức độ có thể chấp nhận được đối với người xem.

Trong phương pháp VQ, bức ảnh được chia thành các khối có kích thước cố định, một bảng mã (code book) được xây dựng với các chỉ số tương ứng với các khối ảnh này. Như vậy, thay cho việc phải truyền lần lượt các khối của bức ảnh, ta chỉ cần truyền các chỉ số tương ứng của các khối ảnh hoặc chỉ số tương ứng gần đúng nhất so với các khối ảnh cần truyền. Hai phương pháp này cho kết quả không khác nhau nhiều về chất lượng nén ảnh động, tuy nhiên ngày nay biến đổi DCT tỏ ra được ứng dụng rộng rãi hơn trong các sơ đồ nén và giải nén các bức ảnh (theo tiêu chuẩn JPEG) và xử lý ảnh động (theo tiêu chuẩn của MPEG).

– Nén tín hiệu ảnh dùng MPEG

MPEG (Moving Picture Expert Group) được ra đời vào năm 1988 nhằm mục đích chuẩn hoá cho nén tín hiệu âm thanh và video. MPEG - 1 có thể nén tín hiệu video tới 1.5Mbit/s với chất lượng VHS và âm thanh lập thể (stereo audio) với tốc độ 102 bit/s. Nó được dùng để lưu trữ video và âm thanh trên CD-ROM.

Vào những năm 1990, MPEG-2 đã ra đời nhằm đáp ứng các tiêu chuẩn nén video cho truyền hình. MPEG-2 có khả năng mã hoá tín hiệu truyền hình ở tốc độ 3-15Mbit/s và truyền hình độ nét cao ở tốc độ tới 15-30Mbit/s. MPEG-2 cho phép mã hoá tín hiệu video với nhiều mức độ phân giải khác nhau, chúng có khả năng đáp ứng cho nhiều ứng dụng khác nhau. Nhiều thuật toán tương ứng với nhiều các ứng dụng khác nhau đã phát triển và được tập hợp lại thành một bộ tiêu chuẩn đầy đủ của MPEG. Việc áp dụng toàn bộ các đặc điểm của chuẩn MPEG-2 trong tất cả các bộ mã hoá và giải mã là không cần thiết do sự phức tạp của thiết bị cũng như sự tốn kém về dải thông của đường truyền. Vì vậy trong hầu hết các trường hợp ta chỉ sử dụng một phần nhất định trong toàn bộ các đặc điểm của chuẩn MPEG-2, chúng thường được gọi là profiles và levels. Một profile sẽ xác định một thuật toán (điều chỉnh bitstream và độ phân giải màu) và một level sẽ xác định một số tiêu chí bắt buộc cho các tham số của bức ảnh (ví dụ như kích thước ảnh và số lượng bit).

MPEG-4 trở thành một tiêu chuẩn cho nén ảnh kỹ thuật truyền hình số, các ứng dụng về đồ họa và video tương tác hai chiều (games, videoconferencing) và các ứng dụng multimedia tương tác hai chiều (World Wide Web hoặc các ứng dụng nhằm phân phát dữ liệu video như truyền hình cáp, Internet video...) vào năm 1999. Ngày nay, MPEG-4 đã trở thành một tiêu chuẩn công nghệ trong quá trình sản xuất, phân phối và truy cập vào các hệ thống video. Nó đã góp phần giải quyết vấn đề về dung lượng cho các thiết bị lưu trữ, giải quyết vấn đề về băng thông của đường truyền tín hiệu video hoặc kết hợp cả hai vấn đề trên.

MPEG không phải là một công cụ nén đơn lẻ mà ưu điểm của nén ảnh dùng MPEG chính là ở chỗ MPEG có một tập hợp các công cụ mã hoá chuẩn, chúng có thể được kết hợp với nhau một cách linh động để phục vụ cho một loạt các ứng dụng khác nhau.

Nén MPEG là sự kết hợp hài hoà của bốn kỹ thuật cơ bản: Tiền xử lý (Preprocessing), đoán trước sự chuyển động của các frame ở bộ mã hoá (temporal prediction), bù chuyển động ở bộ giải mã (motion compensation) và mã lượng tử hoá (quantisation coding). Các bộ lọc tiền xử lý sẽ lọc ra những thông tin không cần thiết từ tín hiệu video và những thông tin khó mã hoá nhưng không quan trọng cho sự cảm thụ của mắt người. Kỹ thuật đoán chuyển động dựa trên nguyên tắc là các ảnh trong chuỗi video dường như có liên quan mật thiết với nhau theo thời gian: Mỗi frame tại một thời điểm nhất định sẽ có nhiều khả năng giống với các frame đứng ngay phía trước và ngay phía sau nó. Các bộ mã hoá sẽ tiến hành quét lần lượt từng phần nhỏ trong mỗi frame gọi là macro blocks, sau đó nó sẽ phát hiện macro block nào không thay đổi từ frame này tới frame khác. Bộ mã hoá sẽ tiên đoán trước sự xuất hiện của các macro blocks khi biết vị trí và hướng chuyển động của nó. Do đó chỉ những sự thay đổi giữa các khối trong frame hiện tại (motion compensated residual) và các khối được tiên đoán mới được truyền tới bên phía thu. Phía bên thu tức bộ giải mã đã lưu trữ sẵn những thông tin mà không thay đổi từ frame này tới frame khác trong bộ nhớ đệm của nó và chúng được dùng để điền thêm một cách đều đặn vào các vị trí trống trong ảnh được khôi phục.

Như chúng ta đều biết, nén tín hiệu video được thực hiện nhờ việc loại bỏ cả sự dư thừa về không gian (spatial coding) và thời gian (temporal coding). Trong MPEG, việc loại bỏ dư thừa về thời gian (nén liên ảnh) được thực hiện trước hết nhờ sử dụng các tính chất giống nhau giữa các ảnh liên tiếp (Inter-frame techniques). Chúng ta có thể sử dụng tính chất này để tạo ra các bức ảnh mới nhờ vào những thông tin từ những ảnh đã gửi trước nó ("predicted"). Do vậy ở phía bộ mã hoá, ta chỉ cần gửi những bức ảnh có thay đổi so với những ảnh trước, sau đó ta lại dùng phương pháp nén về không gian để loại bỏ sự dư thừa về không gian trong chính bức ảnh sai khác này. Nén về không gian dựa trên nguyên tắc là phát hiện sự giống nhau của các điểm ảnh (pixels) lân cận nhau (Intra-frame coding techniques). JPEG chỉ áp dụng phương pháp nén theo không gian vì nó được thiết kế để xử lý và truyền các ảnh tĩnh. Tuy nhiên nén tín hiệu theo phương pháp của JPEG cũng có thể được dùng để nén các bức ảnh một cách độc lập trong dãy tín hiệu video. ứng dụng này thường được gọi là JPEG

động (Motion JPEG). Trong một chu kỳ gửi một dãy các bức ảnh theo kiểu JPEG động, ảnh đầu tiên được nén nhờ sự loại bỏ độ dư thừa về không gian, sau đó các ảnh tiếp theo được nén nhờ sự loại bỏ độ dư thừa về thời gian (nén liên ảnh). Quá trình được lặp đi lặp lại cho một dãy các bức ảnh trong tín hiệu video.

Thuật toán nén MPEG cũng dựa trên phép biến đổi DCT cho các khối ảnh 8×8 pixels để tìm ra sự thừa về không gian một cách có hiệu quả giữa các điểm ảnh trong cùng một bức ảnh. Tuy nhiên, trong trường hợp có mối tương quan chặt chẽ giữa các điểm ảnh trong các bức ảnh kế tiếp nhau tức là trong trường hợp hai bức ảnh liên tiếp có nội dung trùng nhau, kỹ thuật Inter-frame coding techniques sẽ được dùng cùng với việc tiên đoán sự dư thừa về không gian để tạo thành kỹ thuật tiên đoán bù chuyển động giữa các bức ảnh (Motion compensated prediction between frames). Trong nhiều sơ đồ nén MPEG, người ta thường kết hợp cả việc tiên đoán bù chuyển động theo thời gian và phép biến đổi thông tin theo không gian để đạt hiệu quả nén cao (Hybrid DPCM/DCT coding of video).

Hầu hết các sơ đồ nén MPEG đều dùng kỹ thuật lấy mẫu bỏ xung (Subsampling) và lượng tử hoá (Quantization) trước khi mã hoá. Lấy mẫu bỏ xung nhằm mục đích để làm giảm kích thước bức ảnh đầu vào theo cả theo chiều ngang và chiều dọc, như vậy sẽ giảm số lượng các điểm ảnh trước mã hoá. Cũng nên nhớ rằng trong một số trường hợp người ta còn lấy mẫu bỏ xung theo thời gian để làm giảm số lượng các bức ảnh trong dãy ảnh trước khi mã hoá. Đây được xem như là một kỹ thuật rất cơ bản nhằm loại bỏ sự dư thừa dựa vào khả năng lưu ảnh của mắt người cảm thụ. Thường thường, chúng ta có thể phân biệt sự thay đổi về độ sáng của ảnh (changes in Brightness) tốt hơn so với sự thay đổi về màu (Chromaticity changes). Do đó trước hết các sơ đồ nén MPEG sẽ tiến hành chia bức ảnh thành các thành phần Y (Luminance hay brightness plane) và UY (Chrominance hay color planes) tức là một thành phần về độ sáng và hai thành phần về độ màu. Các tín hiệu video thành phần này sẽ được lấy mẫu (samples) và số hoá (digitised) để tạo nên các điểm ảnh rời rạc theo tỷ lệ 4 : 2 : 2 và 4 : 2 : 0.

Kỹ thuật tiên đoán bù chuyển động được sử dụng như là một trong những công cụ mạnh để làm giảm sự dư thừa về không gian giữa các bức ảnh. Khái niệm về bù chuyển động là dựa trên sự phán đoán hướng chuyển động của các bức ảnh tức là các ảnh thành phần trong dãy video sẽ được thay thế gần đúng. Kỹ thuật tiên đoán bù chuyển động giữa các bức ảnh được xem như là biện pháp để hạn chế bớt các thông số của chuyển động bởi việc dùng các vector chuyển động để mô tả sự dịch chuyển của các điểm ảnh. Kết quả tiên đoán tốt nhất của một điểm ảnh là dựa trên sự tiên đoán bù chuyển động từ một bức ảnh đã mã hoá được truyền phía trước của nó. Cả hai thông số, sai số chuyển động (biên độ) và các vectors chuyển động (hướng chuyển động) đều được truyền tới phía bên nhận. Tuy nhiên do có mối quan hệ tương quan chặt chẽ giữa các điểm ảnh về không gian (trùng về không gian), một vector chuyển động có thể được dùng cho một khối các điểm ảnh gồm các pixels lân cận nhau (MPEG -1 và MPEG -2 dùng các khối 16×16 pixels).

Trong MPEG-2, có nhiều phương pháp để tiên đoán sự chuyển động. Ví dụ một khối ảnh có thể được tiên đoán xuôi từ những ảnh đã được truyền trước nó (Forward Predicted) có thể đoán ngược từ những ảnh truyền sau nó (Backward Predicted) hoặc theo cả hai chiều (Bidirectionally Predicted). Các phương pháp dùng để tiên đoán các khối trong cùng một ảnh cũng có thể không giống nhau, chúng có thể thay đổi từ khối nọ sang khối kia. Hơn nữa, hai trường (fields) trong cùng một khối cũng có thể được tiên đoán theo hai cách khác nhau dùng các vector độc lập nhau hoặc chúng có thể dùng chung một vector. Đối với mỗi khối ảnh, bộ mã hoá sẽ chọn các phương pháp tiên đoán thích hợp, cố gắng đảm bảo chất lượng ảnh tốt nhất khi được giải mã trong điều kiện yêu cầu khắt khe về số bit. Các thông số liên quan tới chọn phương pháp tiên đoán cũng được truyền tới bộ giải mã cùng với dự đoán sai số nhằm khôi phục gần chính xác ảnh gốc.

Trong MPEG, có 3 kiểu ảnh khác nhau được dùng để mã hoá cho các khối ảnh. Kiểu ảnh 'Intra' (I-pictures) là ảnh được mã hoá một cách độc lập mà không cần tham khảo tới các ảnh khác. Hiệu quả nén tín hiệu đạt được do loại bỏ sự thừa về không gian mà không có yếu tố thời gian tham gia vào quá trình. I-pictures được dùng một cách tuần hoàn để tạo thành các điểm tựa cho dòng dữ liệu trong quá trình giải mã.

Ảnh 'Predictive' (P-pictures) có thể sử dụng các ảnh I hoặc P ngay sát phía trước nó để bù chuyển động và chính nó cũng có thể được dùng để tham khảo cho việc tiên đoán các ảnh khác tiếp theo. Mỗi khối ảnh trong P-picture có thể hoặc được mã theo kiểu tiên đoán (predicted) hoặc được mã một cách độc lập (intra-coded). Do sử dụng cả nén theo không gian và thời gian, hiệu quả nén của P-pictures được tăng lên một cách đáng kể so với I-pictures.

Ảnh 'Bidirectionally-predictive' pictures hay B- Pictures có thể sử dụng các ảnh I hoặc P phía trước hoặc phía sau nó cho việc bù chuyển động và do vậy cho kết quả nén cao nhất. Mỗi khối trong B-pictures có thể được tiên đoán theo chiều ngược, xuôi, cả hai hướng hoặc được mã một cách độc lập. Để có thể tiên đoán ngược từ một bức ảnh phía sau nó, bộ mã hoá sẽ tiến hành sắp xếp lại các bức ảnh từ thứ tự xuất hiện một cách tự nhiên sang một thứ tự khác của các ảnh trên đường truyền. Do vậy từ đầu ra của bộ mã hoá, B-pictures được truyền sau các ảnh dùng để tham khảo ở phía trước và phía sau của nó. Điều này sẽ tạo ra độ trễ do phải sắp xếp lại thông tin(độ trễ này lớn hay nhỏ là tùy thuộc vào số các bức ảnh B-pictures liên tiếp nhau được truyền. Các ảnh I P. B-pictures thường xuất hiện theo một thứ tự lặp đi lặp lại một cách tuần hoàn, do đó ta có khái niệm về nhóm các bức ảnh GOP (Group of Pictures). Một ví dụ của GOP ở dạng ảnh tự nhiên xuất hiện theo thứ tự như sau:

B 1 B2B I3 B4 B5 B7 B8 P9 B10 B11 P12 Thứ tự xuất hiện của chúng trên đường truyền bị thay đổi do sự sắp xếp lại của bộ mã hoá như sau:

I3 B1 B2 P6 B4 B5 p9 B7 B8 P12 B10 B11

Cấu trúc của một GOP có thể được mô tả bởi hai tham số: N là số các ảnh trong GOP và M là khoảng cách giữa các ảnh P-pictures. Nhóm GOP này được miêu tả như $N = 12$ và $M = 3$.