

A. Nén video theo chuẩn MPEG

1) *Khái quát về nén MPEG*

- MPEG (Moving Picture Expert Group) là nhóm chuyên gia về hình ảnh, được thành lập từ tháng 2 năm 1988 nhằm chuẩn hóa cho nén tín hiệu âm thanh và video. Nén tín hiệu video theo chuẩn MPEG là phương pháp nén ảnh động không những làm giảm dư thừa không gian (như JPEG) mà còn giảm dư thừa thời gian giữa các khung ảnh, đây chính là khác biệt so với IPEG.

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

- Nén MPEG là sự kết hợp hài hòa giữa bốn kỹ thuật đơn giản:

+ **Tiền xử lý (Preprocessing):** lọc ra các thông tin không cần thiết từ tín hiệu video và những thông tin khó mã hóa nhưng không quan trọng cho cảm thụ của mắt người.

+ **Đoán trước sự chuyển động của các frame ở bộ mã hóa (Temporal prediction):** 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 các frame đứng ngay phía trước và ngay sau nó. Các bộ mã hóa sẽ quét lần lượt từng phần nhỏ trong mỗi frame và chỉ phát những thay đổi giữa các khối trong frame hiện tại và các khối được bộ mã hóa tiên đoán tới phía thu. Nhờ vậy mà có thể loại bỏ dư thừa về thời gian.

+ **Bù chuyển động ở bộ giải mã (motion compensation):** tại phía thu có lưu trữ sẵn những thông tin mà không thay đổi của frame này tới frame khác trong bộ nhớ đệm và chúng được dùng để điền thêm một cách đều đặn vào các vị trí trống trong ảnh được khôi phục. Đây cũng chính là một trong những công cụ mạnh để làm giảm dư thừa về không gian giữa các bức ảnh.

+ **Mã lượng tử hóa (quantisation coding)**

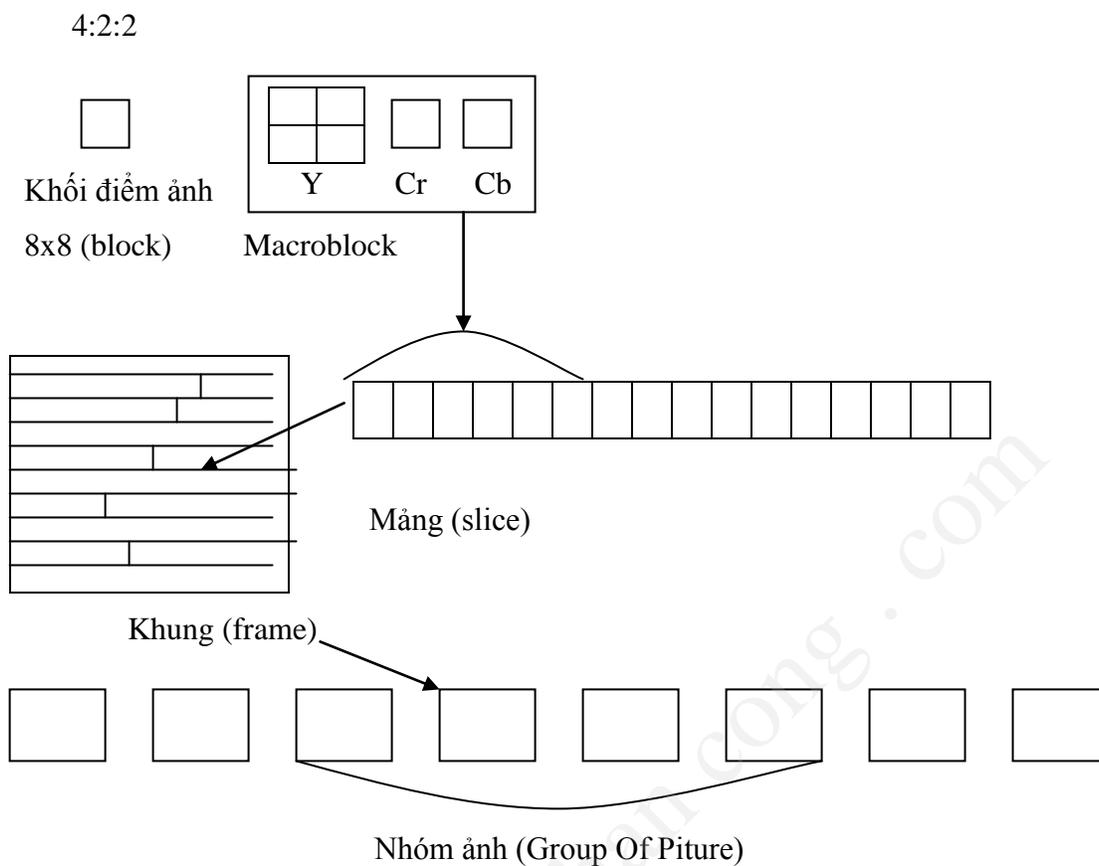
2) *Cấu trúc dòng bit*

- Cấu trúc dòng MPEG gồm 6 lớp:

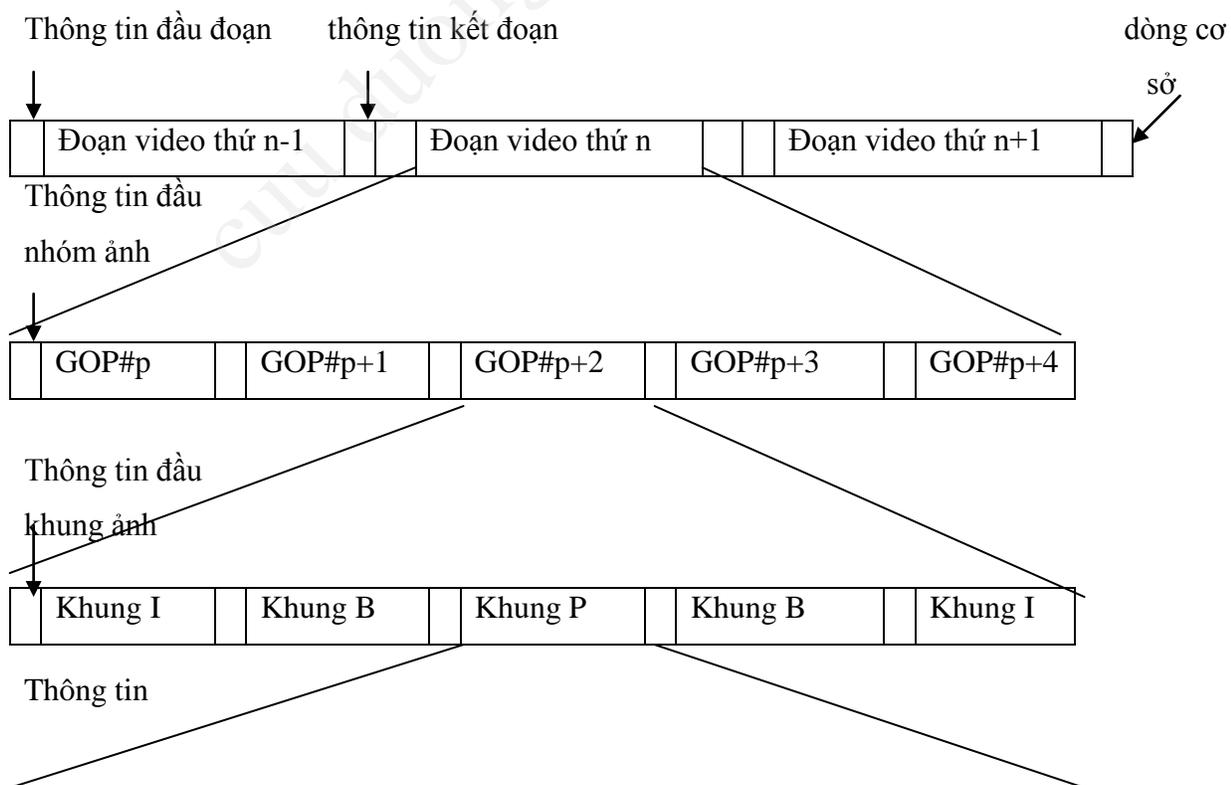
- Lớp khối (block) : gồm khối 8x8 các điểm ảnh tín hiệu chói và tín hiệu màu dùng cho phương pháp nén DCT.

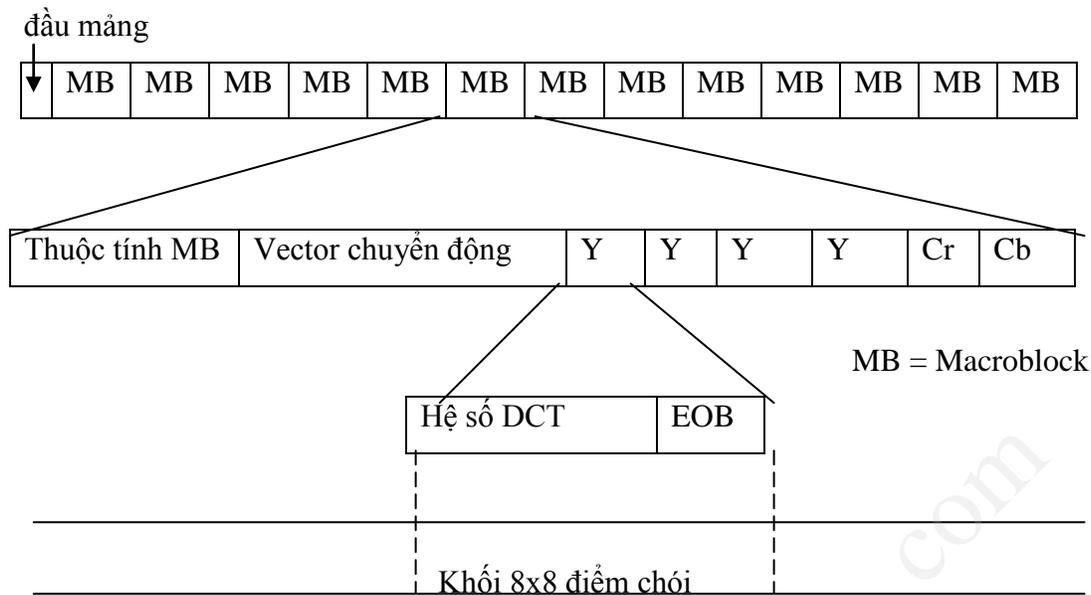
- Lớp macroblock: lớp tổ hợp cấu trúc khối là một nhóm các khối tương ứng với lượng thông tin chứa đựng trong kích thước 16x16 điểm trên bức ảnh. Cấu trúc khối này cũng xác định được lượng thông tin chứa trong đó sẽ thay đổi tùy theo cấu trúc mẫu được sử dụng. Thông tin đầu tiên trong cấu trúc khối mang dạng của nó và các vector bù chuyển động.

- Lớp cắt lát dòng bit (slice): Mảng bao gồm một và cấu trúc khối kề nhau. Kích thước lớn nhất của mảng có thể bao gồm toàn bộ bức ảnh và kích thước nhỏ nhất của mảng là một cấu trúc khối. Thông tin đầu tiên chứa đựng vị trí của mảng trong toàn bộ ảnh, và hệ số cân bằng lượng tử..
- Lớp ảnh (picture): cho phép bộ giải mã xác định loại ảnh được mã hóa. Thông tin đầu dùng để chỉ thứ tự truyền khung (để bộ mã hóa có thể sắp xếp các ảnh lại theo thứ tự đúng) và các thông tin về đồng bộ, độ phân giải và phạm vi vector chuyển động. Trong nén MPEG người ta sử dụng 3 loại ảnh sau:
 - Ảnh I (Intra Pictures): được mã hóa độc lập mà không cần tham khảo các ảnh khác. Hiệu quả nén tín hiệu đạt được do loại bỏ được dư thừa về không gian mà không có yếu tố thời gian tham gia vào quá trình. Ảnh I được dùng một cách tuần hoàn để tạo thành điểm tựa cho dòng dữ liệu trong quá trình giải mã.
 - Ảnh P (Predicted Pictures): được mã hóa từ ảnh I hoặc ảnh P ngay trước để bù chuyển động và chính nó cũng có thể dùng để tham khảo cho việc tiên đoán ảnh tiếp theo. Mỗi khối ảnh trong ảnh P có thể được mã hóa theo kiểu tiên đoán hoặc mã hóa một cách độc lập. Do sử dụng cả nén không gian và thời gian nên hiệu quả nén của ảnh P cao hơn đáng kể so với ảnh I.
 - Ảnh B (Bidirectionally Predictive Pictures) có thể sử dụng các ảnh I và P phía trước và phía sau nó để bù chuyển động vì vậy cho tỷ lệ nén cao nhất. Mỗi khối trong ảnh B có thể tiên đoán theo chiều ngược, xuôi, cả hai hướng hoặc được mã hóa độc lập. tuy nhiên để tiên đoán ngược từ một bức ảnh phía sau nó thì bộ mã hóa phải sắp xếp lại các bức ảnh. Do vậy sẽ tạo ra độ trễ do phải sắp xếp lại thông tin, độ trễ lớn hay nhỏ phụ thuộc vào số bức ảnh B liên tiếp nhau được truyền.
- Lớp nhóm ảnh (GOP): các ảnh I, P, B thường xuất hiện theo một thứ tự lặp đi lặp lại một cách tuần hoàn. Vì vậy xuất hiện nhóm ảnh GOP. Chất lượng ảnh không chỉ phụ thuộc vào tỷ lệ nén trong từng khuôn hình mà còn phụ thuộc vào độ dài của nhóm ảnh GOP. Chúng là đơn vị mang thông tin độc lập của MPEG và bắt buộc phải bắt đầu bằng một ảnh I. Tiếp sau đó là một loạt các ảnh P và B. Cấu trúc của một GOP được mô tả bởi 2 tham số: N(số ảnh trong GOP) và M(khoảng cách giữa các ảnh P)
- Lớp dãy ảnh (sequence): đoạn video bao gồm thông tin đầu, một số nhóm ảnh và thông tin kết đoạn.
 - Mỗi lớp này hỗ trợ một chức năng nhất định: một là chức năng xử lý tín hiệu (DCT và bù chuyển động). Hai là chức năng logic (tái đồng bộ, điểm truy xuất ngẫu nhiên).



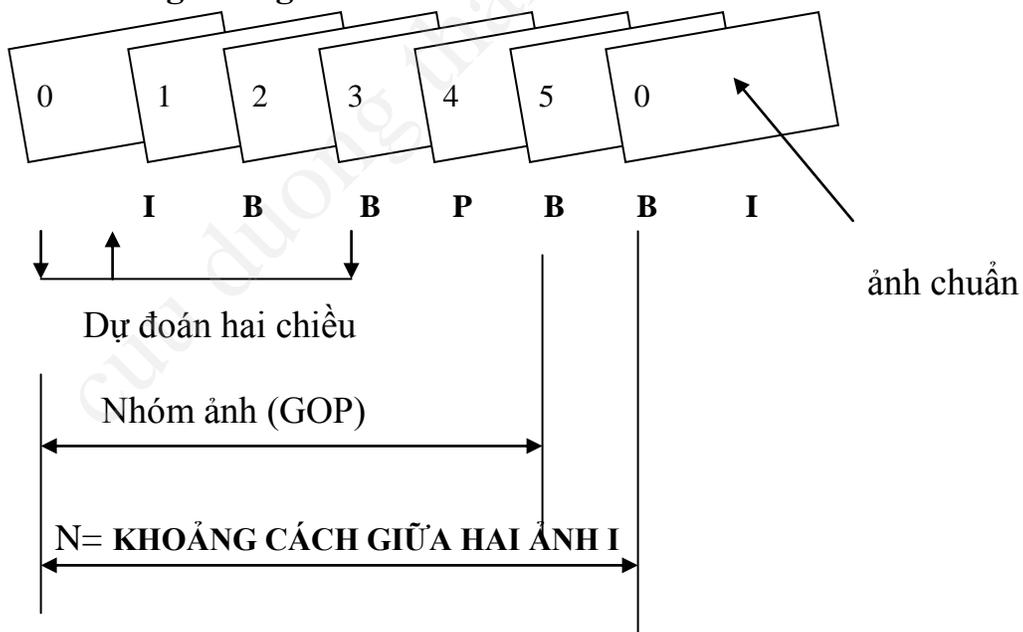
Hình 1: Cấu trúc dòng bit MPEG





Hình 2: Kiến trúc dòng dữ liệu MPEG

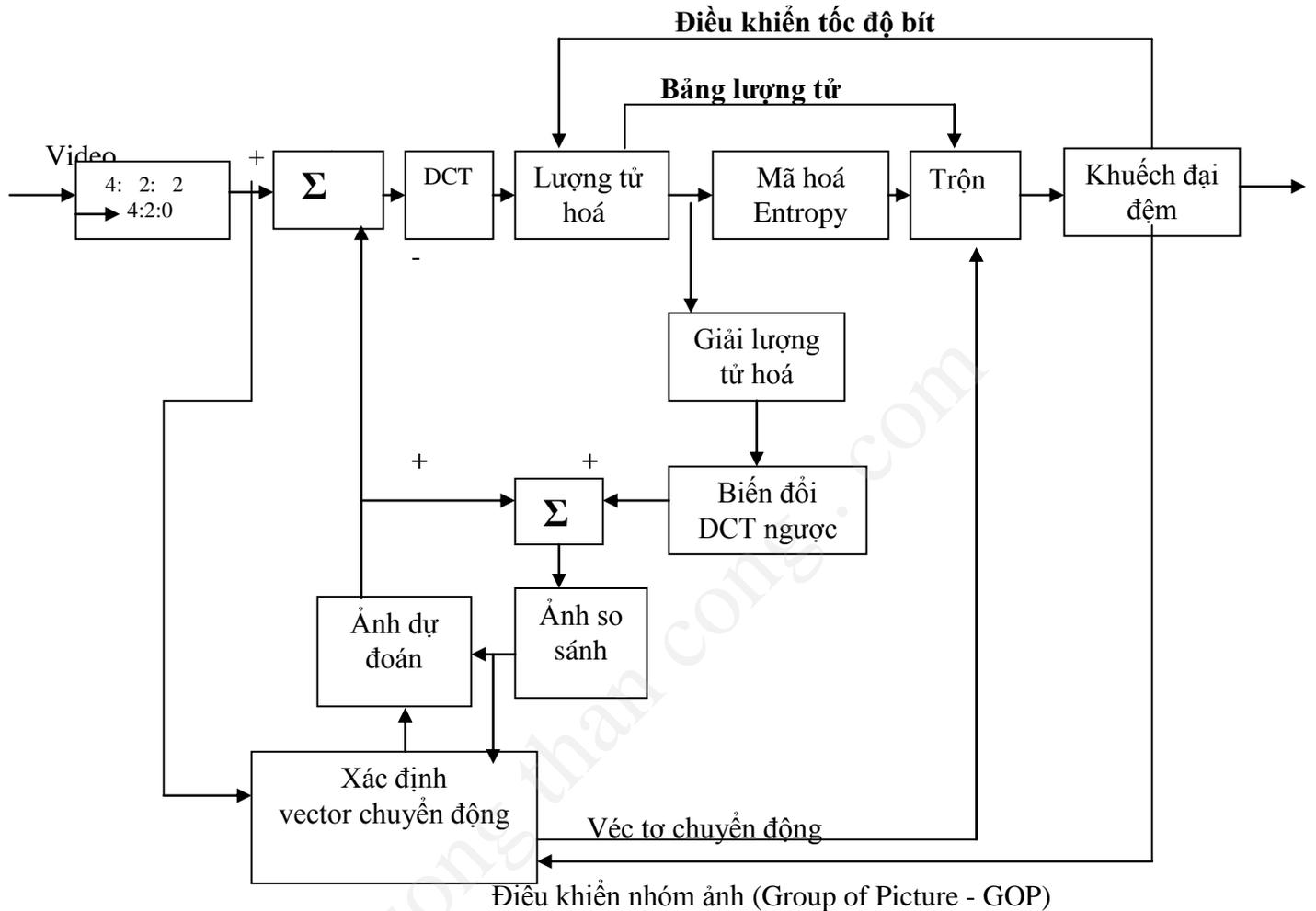
M = Khoảng cách giữa hai điểm so sánh



Hình 3: Cấu trúc ảnh MPEG

- Quá trình tạo dòng bit MPEG là ghép kênh: kết hợp dòng dữ liệu vào, dòng dữ liệu ra, điều chỉnh đồng bộ và quản lý bộ đệm. Cú pháp dòng MPEG bao gồm: lớp dòng bit (stream), lớp gói (back) và lớp gói tin (bucket).

3) Nguyên lý nén MPEG



Hình 4: Bộ mã hóa MPEG tiêu biểu

* Cơ sở của công nghệ nén video MPEG là sự kết hợp giữa nén trong ảnh (Intra -Frame Compression) và công nghệ nén liên ảnh (Inter-Frame Compression). Trong đó:

- **Nén trong ảnh (Intra -Frame Compression):** là loại nén nhằm giảm bớt thông tin dư thừa trong miền không gian. Nén trong ảnh sử dụng cả hai quá trình có tổn hao và không có tổn hao để giảm bớt dữ liệu trong ảnh. Quá trình này không sử dụng thông tin của các ảnh trước và sau ảnh đang xét.

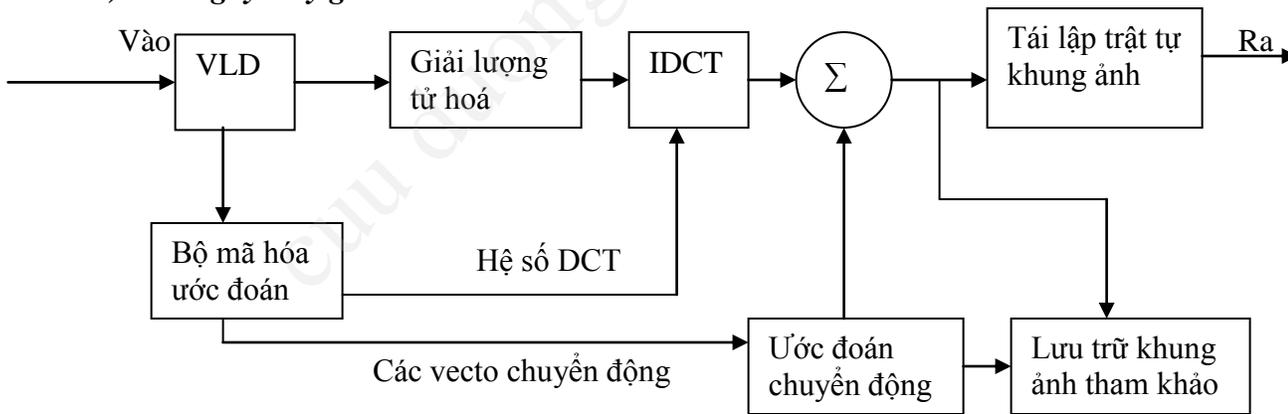
- **Nén liên ảnh (Intra -Frame Compression):** Trong tín hiệu video có chứa thông tin dư thừa trong miền thời gian. Nghĩa là với một chuỗi liên tục các ảnh, lượng thông tin chứa đựng trong mỗi ảnh thay đổi rất ít từ ảnh này sang ảnh khác. Tính toán sự dịch chuyển vị trí của nội dung ảnh là một phần rất

quan trọng trong kỹ thuật nén liên ảnh. Trong thuật nén MPEG, quá trình xác định Vector chuyển động được thực hiện bằng cách chia hình ảnh thành các Macro-Block, mỗi Macro-Block có 16 x 16 phần tử ảnh (tương đương với 4 Block, mỗi Block có 8 x 8 phần tử ảnh). Để xác định chiều chuyển động, người ta tìm kiếm vị trí của Macro-Block trong ảnh tiếp theo, kết quả của sự tìm kiếm sẽ cho ta Vector chuyển động của Macro-Block

*** Nguyên lý nén MPEG :**

- Tín hiệu đầu vào có dạng 4:2:2 hoặc 4:2:0 được nén liên ảnh nhằm tạo ra ảnh khác biệt ở đầu ra bộ cộng. Ảnh khác biệt này sau đó lại được nén trong ảnh qua các bước : biến đổi DCT, lượng tử hóa, mã hoá. Cuối cùng được trộn cùng với vector chuyển động đưa đến bộ khuếch đại đệm sẽ thu được ảnh đã nén.
- Tốc độ bit của tín hiệu video được nén không cố định, phụ thuộc vào nội dung ảnh đang xét (ví dụ một phần nén ít hơn hoặc nhiều hơn), nhưng tại đầu ra bộ mã hoá dòng bit phải cố định để xác định tốc độ cho dung lượng kênh truyền. Do đó tại đầu ra bộ mã hóa phải có bộ nhớ đệm đủ lớn. Bộ mã hóa phải kiểm tra trạng thái đầy của bộ nhớ đệm. Khi số liệu trong bộ nhớ đệm gần bằng dung lượng cực đại thì các hệ số biến đổi DCT ngược được lượng tử hóa ít chính xác hơn. Vì vậy bộ nhớ đệm chứa ít số liệu thì việc lượng tử hóa các hệ số sẽ tăng lên.

4) Nguyên lý giải nén MPEG



Hình 5: Bộ giải mã MPEG tiêu biểu

***Nguyên lý giải nén MPEG :**

- Chuỗi tín hiệu vào được giải mã Entropy tại VLD, sau đó tách dữ liệu ảnh (hệ số biến đổi DCT) ra khỏi các vector chuyển động. Số liệu ảnh sẽ được giải lượng tử hoá và biến đổi DCT ngược.

- Nếu ảnh là ảnh loại I bắt đầu ở mỗi nhóm ảnh trong chuỗi, ở đầu ra sẽ nhận được ảnh hoàn chỉnh bằng cách trên (vì ảnh loại I chỉ là nén trong ảnh, không có bù chuyển động, không dùng dữ liệu của ảnh khác). Nó được lưu trữ trong bộ nhớ ảnh và được và được dùng để giải mã các ảnh tiếp theo.
- Nếu ảnh là ảnh loại P thì cũng thực hiện giải lượng tử hóa và biến đổi DCT ngược kết hợp với việc sử dụng vector chuyển động và lưu vào bộ nhớ ảnh. Trên cơ sở đó xác định được dự đoán ảnh đang xét. Ta nhận được ảnh ra sau khi cộng dự đoán ảnh (ảnh dự đoán) và kết quả biến đổi DCT ngược. Ảnh này cũng được lưu vào bộ nhớ để có thể sử dụng như là chuẩn khi giải mã các ảnh tiếp theo.

B. CÁC CHUẨN NÉN MPEG

1) MPEG-1

- Tiêu chuẩn The Moving Picture Experts Group Phase 1 (MPEG-1) – Định dạng âm thanh, phim ảnh của nhóm chuyên gia ảnh động là một tiêu chuẩn cho định dạng âm thanh, phim ảnh của tổ chức ISO/IEC - International Organization for Standardization/International Electrotechnical Commission (Tổ chức tiêu chuẩn hóa quốc tế/Ủy ban kỹ thuật Điện quốc tế) phát triển nhằm hỗ trợ các thiết bị trình chiếu âm thanh, phim ảnh.
- Tiêu chuẩn MPEG-1 (hay còn gọi là ISO/IEC 11172 – Mã hóa ảnh động và âm thanh kết hợp cho phương tiện lưu trữ thông tin số 1,5Mbit/s) bao gồm 05 phần:

- ISO/IEC công bố 03 phần đầu tiên lần đầu vào tháng 08/1993 gồm :

+ **ISO/IEC-11172-1 :Hệ thống** (lưu trữ và đồng bộ âm thanh, phim ảnh và các dữ liệu khác)

+ **ISO/IEC-11172-2: Phim ảnh** (chứa nội dung của phim ảnh nén)

+ **ISO/IEC-11172-3 :Âm thanh** (chứa nội dung của âm thanh nén)

- Phần 4 của MPEG-1 **ISO/IEC-11172-4 : thử khả năng tương thích** (kiểm tra việc triển khai tiêu chuẩn) được công bố lần đầu vào tháng 03/1995

- Phần 5 của MPEG-1 **ISO/IEC-11172-5 Mô phỏng phần mềm** (phần mềm mẫu thực hiện việc mã hóa và giải mã như thế nào) được công bố lần đầu vào tháng 11/1998.

Nội dung các phần được trình bày tóm tắt như sau:

+ **Phần Hệ thống** quy định phương thức sử dụng, lưu trữ âm thanh, phim ảnh đã được mã hóa và các dữ liệu khác, đồng thời duy trì tính đồng bộ giữa các nội dung khác nhau. Định dạng tệp tin

MPEG được thiết kế để đáp ứng việc lưu trữ trên phương tiện (on media), truyền tải kênh dữ liệu một cách tin cậy;

+ **Phần Phim ảnh** chỉ ra phương pháp nén dữ liệu video để làm giảm tốc độ truyền tải dữ liệu trong một luồng thông tin video truyền tải qua mạng;

+ **Phần Âm thanh** chỉ ra phương pháp nén dữ liệu âm thanh để làm giảm tốc độ truyền tải dữ liệu âm thanh trong một luồng thông tin âm thanh truyền tải qua mạng;

+ **Phần Kiểm thử tương thích** cung cấp hướng dẫn và mô hình tham khảo cho việc kiểm thử tính tương thích của MPEG-1;

+ **Phần mềm tham chiếu** chỉ ra phương pháp mã hóa, giải mã như thế nào.

Nội dung của **Phần Phim ảnh** và **Phần Âm thanh** mô tả các định dạng được sử dụng phổ biến trong lưu trữ, trao đổi, xử lý âm thanh và phim ảnh (ví dụ: (.mpg), (.mpeg), (.mp1), (.mp2), (.mp3), (.m1v), (.m1a), (.m2a), (.mpa), (.mpv))

MPEG-1 phần Phim ảnh và phần Âm thanh

- MPEG-1 phần Phim ảnh (ISO/IEC-11172-2) sử dụng các phương pháp nén ảnh để làm giảm tốc độ dữ liệu theo yêu cầu của một luồng phim ảnh. Phương pháp nén ảnh sẽ giảm hoặc loại bỏ hoàn toàn thông tin/dữ liệu trong một dải tần số nhất định và các khu vực hình ảnh mà mắt người nhận biết một cách hạn chế. Các thuật toán nén phim ảnh và các bộ mã hóa/giải mã (codecs) đều kết hợp kỹ thuật nén hình ảnh theo không gian và kỹ thuật bù đắp chuyển động theo thời gian. Trong thực tế, hầu hết các bộ mã hóa/giải mã đều tiến hành kỹ thuật nén âm thanh và nén hình ảnh song song hay riêng biệt cho một tệp tin phim ảnh, tuy nhiên, các dòng dữ liệu được kết hợp lại và gửi đi như một gói dữ liệu. MPEG-1 phần Phim ảnh hỗ trợ độ phân giải lên đến 4095x4095 (12 bit) và tốc độ bit lên đến 100 Mbit/s. MPEG-1 phần Phim ảnh thường sử dụng độ phân giải SIF (Source Input Format): 352 x 240, 352 x 288 hoặc 320 x 240. Độ phân giải thấp kết hợp với một tỉ lệ tốc độ bit thấp hơn 1,5 Mbit/s tạo ra một dòng dữ liệu bit thấp với thông số hạn chế. Đây là thông số kỹ thuật tối thiểu mà bất kỳ bộ mã hóa/giải mã nào cũng xử lý được, đảm bảo sự cân bằng giữa chất lượng phim ảnh và hiệu quả sử dụng, cho phép triển khai trên các hệ thống phần cứng có chi phí không cao.

- MPEG-1 phần Âm thanh (ISO/IEC-11172-3) được chia làm 03 lớp: MPEG-1 phần Âm thanh Lớp I có tỉ lệ tốc độ bit từ 32 – 448 kbit/s (32, 64, 96... 416, 448 kbit/s), Lớp II có tỉ lệ tốc độ bit từ 32-384 kbit/s (32, 48, 56, 64, 80...256, 320, 384 kbit/s), Lớp III có tỉ lệ tốc độ bit từ 32-320 kbit/s (32, 40, 48, 56, 64, 80...224, 256, 320 kbit/s). Việc mã hóa/giải mã sẽ phức tạp hơn trong lớp cao hơn và sử dụng hiệu quả ở tỉ lệ tốc độ bit thấp hơn.

- MPEG-1 phần Âm thanh Lớp I có định dạng tệp tin là (.mp1) và định dạng tệp tin của Lớp II là (.mp2), chủ yếu được sử dụng trong phát sóng truyền thanh, truyền hình. MPEG-I phần Âm thanh Lớp III, thường được gọi là MP3, có định dạng tệp tin là (.mp3), là định dạng âm thanh phổ biến trong ngành công nghiệp âm nhạc, giải trí trên Internet do kích thước nhỏ gọn và chất lượng âm thanh của tệp tin .mp3 .

Cho đến nay, các phần được cập nhật nhiều lần. Lần cập nhật mới nhất cho các phần của MPEG-1 là ISO/IEC-11172-1 (1999); ISO/IEC-11172-2 (2006); ISO/IEC-11172-3 (1996), ISO/IEC-11172-4 (2007), ISO/IEC-11172-5 (2007).

MPEG-1 là tiêu chuẩn nén suy hao cho âm thanh, phim ảnh. MPEG-1 được thiết kế để nén âm thanh, phim ảnh xuống 1.5 Mbit/s (tỉ lệ nén tương ứng là 26:1 và 6:1) mà không giảm chất lượng âm thanh, phim ảnh quá mức.

Các đặc tính của MPEG-1

Cấu trúc dòng bit của MPEG -1 cũng tương tự như cấu trúc dòng bit của MPEG gồm:

- **Sequence (chuỗi ảnh)** : gồm nhiều nhóm ảnh GOP, có chức năng là dòng bit video.
- **GOP (Group of Picture)** : gồm từ 1- n ảnh bắt đầu bằng ảnh I, có chức năng là đơn vị truy xuất.
- **Picture I, P, B**: gồm nhiều Slice, chức năng là đơn vị mã hoá cơ bản.
- **Slice** : gồm nhiều các Macro Block, là đơn vị để tái đồng bộ phục hồi lỗi.
- **Macro-Block** : gồm 16 x 16 pixel, là đơn vị bù chuyển động.
- **Block** : gồm 8 x 8 pixel, là đơn vị tính DCT.

Bảng 2: Tham số theo tiêu chuẩn MPEG -1:

Tham số	Đặc điểm
Tín hiệu mã hoá	Tín hiệu mã hoá
Cấu trúc lấy mẫu	4:2:0
Kích thước ảnh tối đa(điểm ảnh x điểm ảnh)	4095 x 4095
Biểu diễn mẫu	8 bit
Độ chính xác của quá trình lượng tử hoá và biến đổi DCT	9 bit
Phương pháp lượng tử hoá hệ số DCT	DPCM tuyến tính
Cấu trúc khối trong quá trình lượng tử hoá thích nghi	16 x 16 bit
Độ chính xác cực đại của hệ số DC	8 bit
Biến đổi RLC	Mã Huffman

Bảng VLC	Không thể truyền tải
Hệ số cân bằng các khối	Có thể biến đổi
Bù chuyển động	Trong khung hình và giữa các khung hình
Quét	Tuần tự
Độ chính xác dự đoán chuyển động	½ điểm ảnh
Tốc độ khi nén	Tốc độ khi nén 1.85 Mb/s cho nén tham số
	100 Mb/s cho dòng đầy đủ tham số

MPEG-1 không thực sự định nghĩa thiết kế một bộ mã hóa, thay vào đó chuẩn này mô tả cú pháp mã hóa và bộ giải mã “tham khảo” giả thuyết.

Ứng dụng

MPEG-1 ứng dụng trong thiết kế CD chạy âm thanh, trình chiếu video, truyền hình cáp/vệ tinh kỹ thuật số... Ngoài ra, MPEG-1 thường được sử dụng là tiêu chuẩn cho đĩa quang VCD

2) MPEG-2

Tiêu chuẩn The Moving Picture Experts Group Phase 2 (MPEG-2) – Định dạng âm thanh (audio), phim ảnh (video) MPEG-2 là một tiêu chuẩn cho truyền hình kỹ thuật số thuộc họ tiêu chuẩn MPEG dùng để mã hóa luồng dữ liệu âm thanh, phim ảnh. MPEG-2 được công bố thành tiêu chuẩn ISO/IEC 13818, gồm 11 phần như bảng dưới đây:

STT	Tiêu chuẩn	Năm công bố lần thứ nhất	Năm cập nhật mới nhất	Chủ đề
Phần 1	<u>ISO/IEC 13818-1</u>	1996	2013	Hệ thống hỗ trợ phương thức mã hóa kết hợp phim ảnh và âm thanh
Phần 2	<u>ISO/IEC 13818-2</u>	1996	2013	Phương pháp biểu diễn dữ liệu phim ảnh đã mã hóa và quy trình giải mã theo yêu cầu khôi phục ảnh
Phần 3	<u>ISO/IEC 13818-3</u>	1995	1998	Phương pháp mã hóa dữ liệu âm thanh
Phần 4	<u>ISO/IEC 13818-4</u>	1998	2009	Thủ tục kiểm thử tính tương

				thích
Phần 5	<u>ISO/IEC TR 13818</u>	1997	2005	Mô phỏng phần mềm ngôn ngữ C của bộ mã hóa và giải mã
Phần 6	<u>ISO/IEC 13818-6</u>	1998	1998	Các mở rộng cho yêu cầu và điều khiển phương tiện lưu trữ số (Digital Storage Media Command and Control - DSM-CC)
Phần 7	<u>ISO/IEC 13818-7</u>	1997	2006	Tiêu chuẩn mã hóa âm thanh nâng cao
Phần 8 (đã bị thu hồi)				Mở rộng phim ảnh 10-bit, ứng dụng chính là phim ảnh trong studio (xưởng phim)
Phần 9	<u>ISO/IEC 13818-9</u>	1996	1996	Mở rộng giao diện thời gian thực cho các bộ giải mã hệ thống
Phần 10	<u>ISO/IEC 13818-10</u>	1999	1999	Mở rộng tương thích cho DSM-CC
Phần 11	<u>ISO/IEC 13818-11</u>	2004	2004	Quản lý sở hữu trí tuệ

Tiêu chuẩn MPEG-2 là mở rộng của MPEG-1 cơ sở với mục tiêu cung cấp nén có chất lượng cao cho việc truyền phim ảnh, cụ thể là để tạo ra hình ảnh chất lượng cao hơn MPEG-1 ở tốc độ bit (bit rate) cao hơn (từ 3 - 10 Mbps), tuy nhiên tại tốc độ bit thấp (<1,5 Mbps), MPEG-2 chưa chắc đã cho chất lượng ảnh tốt hơn MPEG-1. Một số định dạng tệp tin MPEG-2 như (.mpg), (.mpeg), (.mp2), (.mp3), (.m1v), (.m2v)...

Những điểm nâng cao được thêm vào MPEG-2 bao gồm:

Mã hóa hiệu quả video với chất lượng truyền hình: ứng dụng quan trọng nhất của MPEG-2 là truyền số liệu quảng bá. Các chức năng chính của MPEG-2 được đánh giá là có hiệu quả trong mã hóa độ phân giải truyền hình ở tốc độ bit trong khoảng từ 3-5Mbps.

Hỗ trợ mã hóa video kết hợp: MPEG-2 video có một vài đặc tính hỗ trợ mã hóa một cách linh hoạt video kết hợp. Hai trường tạo ra một khung kết hợp hoàn chỉnh có thể được mã hóa như những hình ảnh riêng biệt, mỗi trường sẽ được mã hóa thành một ảnh I, hoặc từ trường khác trong khung hiện tại.

Khả năng phân cấp: một luồng bit phân cấp bao gồm một lớp cơ sở và một hoặc nhiều lớp tăng cường. Lớp cơ sở có thể được giải mã để tạo ra chuỗi video có thể nhận diện mà có chất lượng nhìn bị hạn chế, và một chuỗi chất lượng cao hơn có thể tạo ra bởi việc giải mã lớp cơ sở cộng với những lớp tăng cường, mỗi lớp tăng cường thêm vào sẽ nâng cao chất lượng của chuỗi mã. MPEG-2 hỗ trợ 4 chế độ phân cấp:

+ Phân cấp theo không gian

- + Phân cấp theo thời gian
- + Phân cấp SNR
- + Phân vùng dữ liệu

a.Nén phim ảnh MPEG-2

Thuật toán nén phim ảnh MPEG-2 đạt tốc độ rất cao bằng cách khai thác các thông tin dư thừa trong phim ảnh. MPEG-2 loại bỏ cả sự dư thừa thời gian và dư thừa không gian xuất hiện trong phim ảnh chuyển động.

Dư thừa thời gian (Temporal redundancy) phát sinh khi khung hình kế tiếp của phim ảnh hiển thị hình ảnh (image) của cùng một cảnh. Nó có chung nội dung của một cảnh (scene) để giữ cố định hoặc để thay đổi một chút giữa khung hình kế tiếp.

Dư thừa không gian (Spatial redundancy) xảy ra do các vùng (pat) của một bức ảnh (picture) (được gọi là pels) được nhân bản (với thay đổi rất nhỏ) trong một khung hình phim ảnh duy nhất.

Rõ ràng, không phải lúc nào cũng có thể nén mọi khung hình của một đoạn phim ảnh với cùng mức độ - một số phần của một clip có thể có dư thừa không gian thấp (ví dụ như nội dung bức tranh phức tạp), trong khi một số phần khác có thể có dư thừa thời gian thấp (ví dụ như trình tự chuyển động nhanh). Do đó, các dòng phim ảnh đã nén sẽ có tốc độ bit biến đổi, phù hợp với tốc độ truyền. Yếu tố chính để kiểm soát tốc độ truyền dẫn là đặt dữ liệu đã nén trong một bộ nhớ đệm theo thứ tự giảm dần. Nén có thể được thực hiện bằng cách loại bỏ một số thông tin có chọn lọc. Chúng ta có thể đảm bảo tốc độ bit mà vẫn đạt được hình ảnh có chất lượng tốt (suy giảm tối thiểu chất lượng hình ảnh) bằng cách bỏ bớt một số thông tin quá chi tiết và giữ lại các nội dung chính.

b.Nén âm thanh MPEG-2

MPEG-2 định nghĩa một cơ chế mã hóa âm thanh theo nhiều định dạng khác nhau tại các tốc độ bit khác nhau

c.Dữ liệu MPEG-2

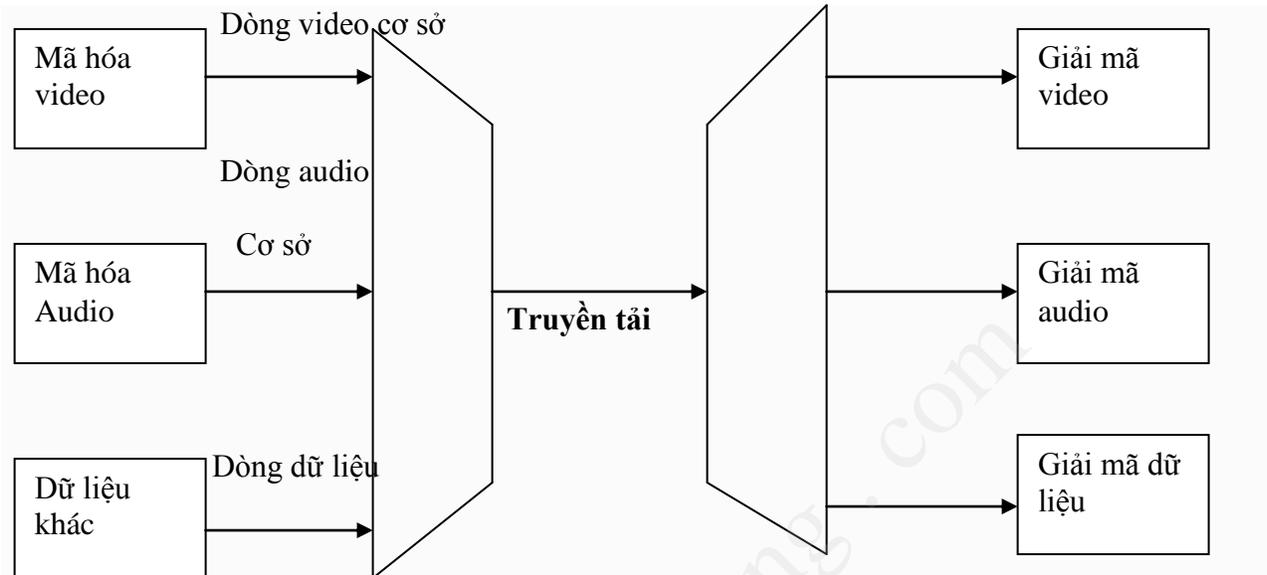
MPEG-2 hỗ trợ truyền 2 loại dữ liệu:

- Thông tin dịch vụ: Thông tin về phim ảnh, âm thanh và dòng dữ liệu được mang trên đường truyền MPEG-2
- Dữ liệu riêng tư: Thông tin cho một hoặc nhiều người dùng cụ thể (có thể là thiết bị thu).

- **Mã hoá MPEG-2**

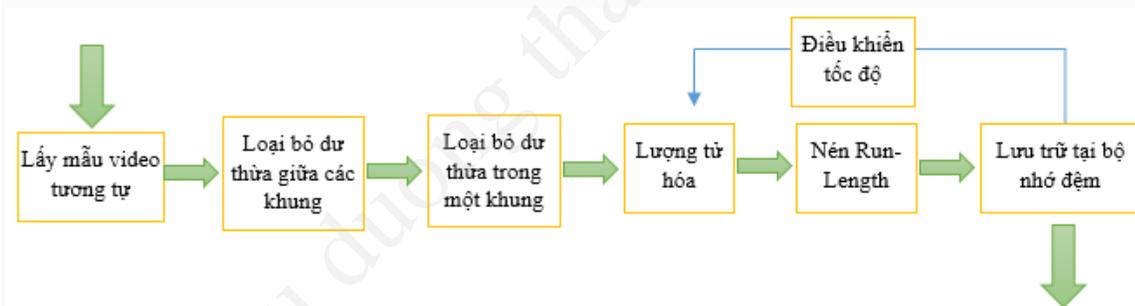
Tín hiệu Video và Audio được nén (theo như nguyên lý nén MPEG) và tạo thành các dòng dữ liệu cơ sở ES (Elementary Stream). Dòng ES được sử dụng để tạo nên dòng dữ liệu cơ sở được

đóng gói PES (Packetized Elementary Stream). Dòng PES lại được tiếp tục đóng gói tạo thành dòng truyền tải TS (Transport Stream)



Hình 6: Chuẩn nén MPEG-2

Các hoạt động cơ bản của bộ mã hóa được thể hiện như hình dưới đây:

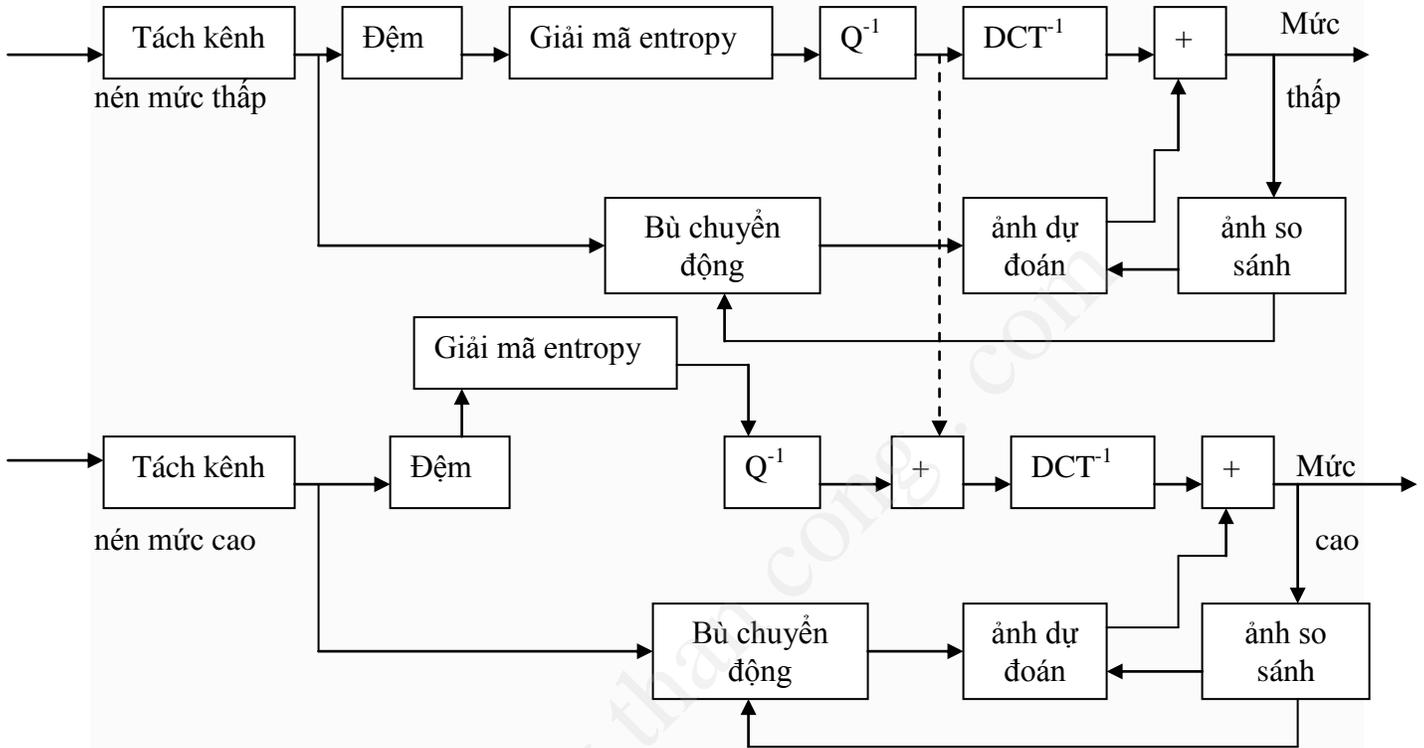


Hình 7: Hoạt động cơ bản của bộ mã hóa MPEG-2

MPEG-2 bao gồm một loạt các cơ chế nén. Do đó, một bộ mã hóa phải sử dụng cơ chế nén phù hợp nhất với một cảnh hoặc một chuỗi cảnh cụ thể. Nhìn chung, mức độ thích hợp của cơ chế nén được lựa chọn tỉ lệ thuận với chất lượng ảnh tại cùng một tốc độ bit truyền nhất định. Trong hình 1, nén Run-Length (nén loạt dài) là một hình thức nén dữ liệu rất đơn giản mà ở đó các chuỗi dữ liệu có giá trị giống nhau trong nhiều thành phần dữ liệu liên tiếp, được lưu trữ như một giá trị duy nhất thay vì nhiều giá trị. Trong thực tế, có nhiều bộ giải mã MPEG-2 khác nhau với nhiều tùy chọn kết nối, nhiều khả năng khác nhau (gồm cả khả năng xử lý phim ảnh chất lượng cao và khả năng xử lý khi gặp lỗi).

- **Giải mã MPEG -2**

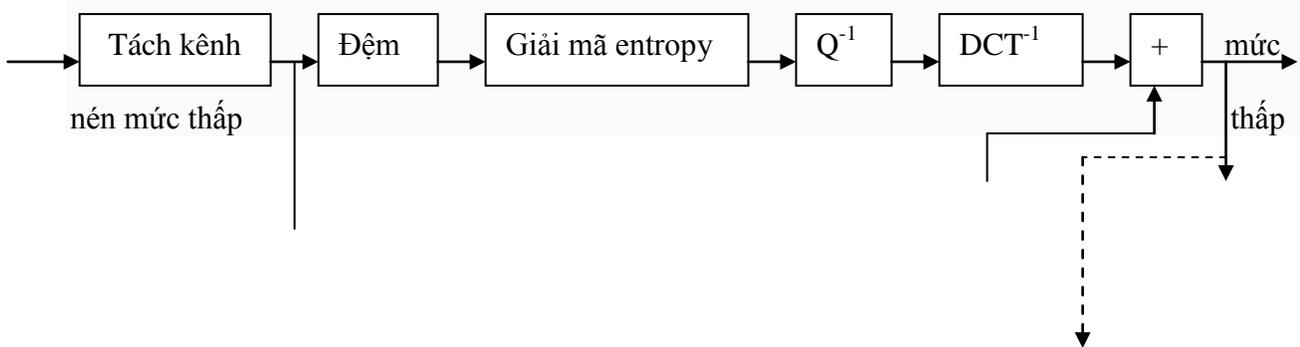
MPEG -2 Coder và Decoder không nhất thiết phải có cùng cấp chất lượng. Tính phân cấp cho phép các bộ giải mã MPEG đơn giản, rẻ tiền, có khả năng giải mã một phần của toàn bộ dòng bit và như vậy có khả năng tạo được hình ảnh tuy chất lượng có thấp hơn các bộ giải mã toàn bộ dòng bit.

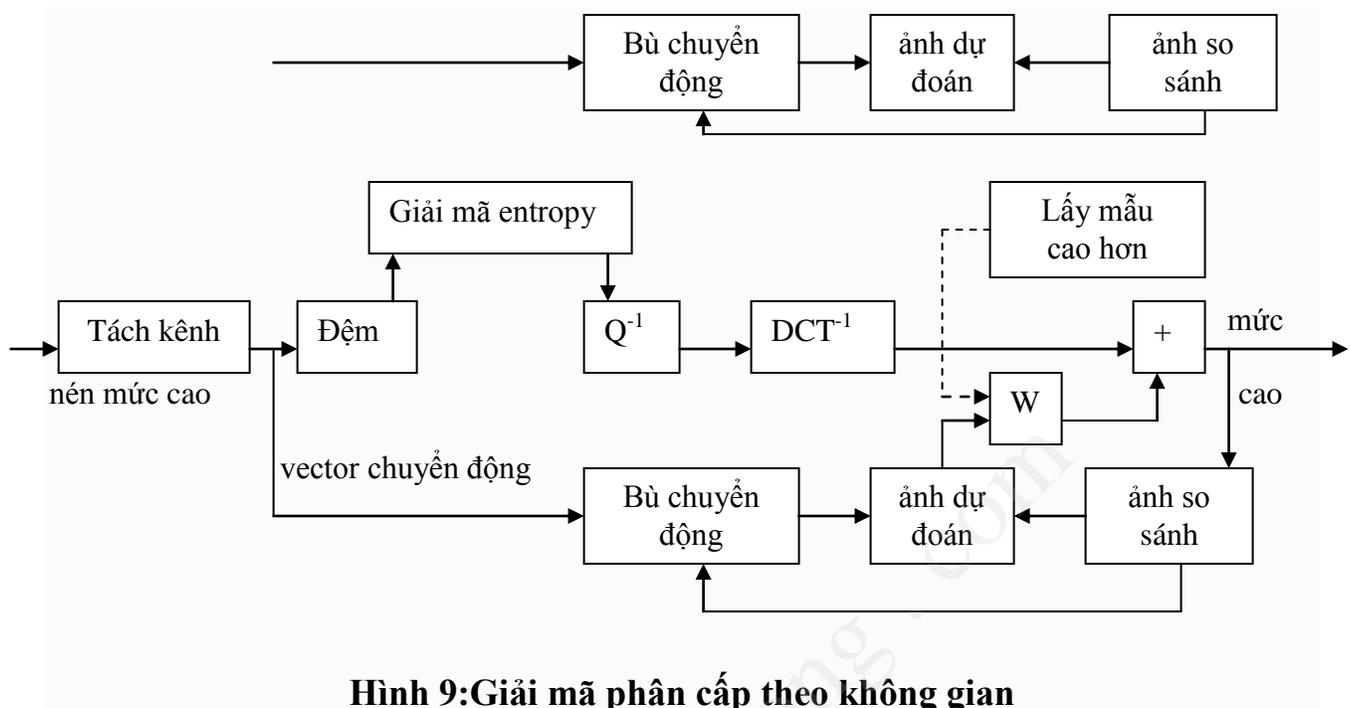


Hình 8: Giải mã phân cấp theo SNR

Tiêu chuẩn MPEG cho phép phân cấp theo tỷ số tín hiệu trên tạp âm (SNR) và theo độ phân giải. Trong đó :

- Tính phân cấp theo tỷ số tín hiệu trên tạp âm (SNR Scalability) có nghĩa là chất lượng hình ảnh và tỷ số tín hiệu trên tạp âm có tính thoả hiệp. Một bộ giải mã có tốc độ bit thấp, có thể có đầy đủ độ phân giải nhưng tỷ số tín hiệu trên tạp âm (SNR) thấp hơn so với bộ giải mã có tốc độ bit cao.
- Tính phân cấp theo không gian (Spatial Scalability) có nghĩa là có sự thoả hiệp đối độ phân giải. Một máy thu có tốc độ bit thấp cho hình ảnh có độ phân giải thấp hơn so với máy thu có khả năng giải mã toàn bộ dòng bit





Hình 9: Giải mã phân cấp theo không gian

- **Profiles và Levels**

+ Profiles: Là khái niệm cho ta biết cấp chất lượng bộ công cụ nén được sử dụng chuẩn nén này. Ở đây có sự thỏa hiệp giữa tỷ số nén và giá thành bộ giải nén.

+ Levels : Khái niệm Levels trong chuẩn MPEG-2 cho ta biết mức độ phân giải của ảnh, bao gồm từ định dạng trung gian cho nguồn tín hiệu SIF (Source Intermediate Format), định dạng cơ sở MPEG -1 (360 x 288 @ 25Hz hay 360 x 240 @ 30Hz), đến truyền hình số phân giải cao HDTV (hệ thống truyền hình với trên 1000 dòng quét)

+ Chuẩn MPEG -2 có 4 Levels (mức) và 5 Profiles (bộ công cụ).

- **profiles**

- Simple Profiles (Profiles đơn giản): Số bước nén thấp nhất, chỉ cho phép mã hoá các ảnh loại I và P. Do có tổn thất cao về tốc độ bit, nó không được sử dụng trong nén với kỹ thuật chuẩn đoán ảnh hai chiều (các ảnh B).

- Main Profiles (Profiles chính): Là sự thỏa hiệp tốt nhất giữa tổn hao tốc độ bit và chi phí, do nó sử dụng tất cả các ảnh I, P và B trong nén. Chất lượng tốt hơn Profiles đơn giản nhưng nó đòi hỏi phải sử dụng các thiết bị mã hoá và giải mã phức tạp hơn.

- SNR Profiles Scalable (Profiles phân cấp theo SNR): Tiêu chuẩn MPEG-2 cho phép phân cấp tỷ số tín hiệu trên tạp âm, có nghĩa là chất lượng hình ảnh và tỷ số tín hiệu trên tạp âm có tính thỏa hiệp. Chuỗi ảnh chia thành hai lớp phân biệt nhau về chất lượng. Các lớp thấp bao gồm ảnh có chất lượng cơ sở, ví dụ như chứa tín hiệu theo chuẩn 4:2:0. Các lớp cao bao gồm

lớp hoàn thiện hơn so với lớp thấp hơn, như với tín hiệu video trong chuẩn 4:2:2. Có thể mã hoá kênh khác nhau cho các lớp riêng.

- Spatially Scalable Profiles (phân cấp theo không gian): Tính phân cấp theo không gian có nghĩa là có sự thoả hiệp với độ phân giải. Chuỗi ảnh được chia ra thành hai lớp tương ứng với các độ phân giải khác nhau của ảnh. Lớp thấp hơn bao gồm ảnh có độ phân giải thấp như truyền hình tiêu chuẩn. Còn lớp cao hơn bao gồm ảnh có độ phân giải cao hơn như truyền hình độ phân giải cao (HDTV).

- High Profiles (Profiles cao): Cho phép cả hai loại thang mức được ứng dụng trong truyền hình HDTV với các định dạng 4:2:0 hay 4:2:2. Nó bao gồm toàn bộ các công cụ của Profiles trước cộng thêm khả năng mã hoá các tín hiệu màu khác nhau cùng một lúc. Như vậy, giữa các Profiles nói trên có sự tương thích cao dần, nghĩa là các bộ giải mã của một Profiles cao cấp hơn có khả năng giải mã tất cả các Profiles ở cấp đó và cấp thấp hơn.

▪ Levels

- Low Levels (mức thấp): phù hợp với độ phân giải SIF được sử dụng trong MPEG -1 (cho đến 360 x 288 pixel).

- Main Levels (mức chính): phù hợp với độ phân giải chuẩn 4:2:2 (tới 720 x 576 pixel).

- High Levels 1440 (mức cao 1440): nhằm vào truyền hình phân giải cao HDTV (độ phân giải tới 1440 x 1152 pixel).

- High Levels (mức cao): được tối ưu hoá đối với HDTV màn ảnh rộng (độ phân giải tới 1920 x 1152 pixel).

Trong thực tế ứng dụng, các nhà chế tạo đã chọn sẵn một số thoả hiệp giữa các mức Levels và các Profiles cho người sử dụng. Kết hợp 4 Levels và 5 Profiles ta được 20 tổ hợp khả năng và hiện nay đã có 11 khả năng được ứng dụng như trong Bảng 3. Trong các ô trong bảng 3 lần lượt từ trên xuống là: tỷ lệ lấy mẫu, điểm ảnh theo chiều ngang x điểm ảnh theo chiều dọc, vận tốc cao nhất của dòng dữ liệu sau khi nén, dòng cuối là các loại ảnh sử dụng để nén.

Profiles levels	Simple (Đơn giản)	Main (Chính)	SNR (Phân cấp theoSNR)	Spatially Theo không gian	High (Cao)
Low (thấp)		4:2:0 352 x288 4 Mb/s	4:2:0 352 x 288 4 Mb/s I, P, B		
Main (chính)	4:2:0 720 x 576 15 Mb/s I, P	4:2:0 720 x576 15 Mb/s I,P,Block	4:2:0 720 x 576 15 Mb/s I, P, B		4:2:0 720 x 576 20 Mb/s I, P, B

High 1440 (Cao 1440)		4:2:0 1440x1152 60Mb/s I, P, B		4:2:0 1440x1152 60Mb/s I, P, B	4:2:2 ;4:2:0 1440x1152 80Mb/s I, P, B
High (cao)		4:2:0 1920x1152 80Mb/s I, P, B			4:2:0;4:2:2 1920x1152 100Mb/s I, P, B

Ứng dụng

MPEG-2 là định dạng sử dụng phổ biến trong truyền hình số mặt đất, truyền hình cáp, truyền hình số vệ tinh DTH (Direct to home), định dạng đĩa quang kỹ thuật số DVD (Digital Versatile Disc) hoặc Digital Video Disc). Đối tượng sử dụng chính của MPEG-2 là các công ty phát sóng và cáp mạng yêu cầu chất lượng phát sóng phim ảnh và sử dụng bộ thu truyền hình vệ tinh và mạng cáp để truyền tải truyền hình cáp và truyền hình vệ tinh phát sóng trực tiếp.

MPEG -2 với phát sóng và sản xuất chương trình

Trong lĩnh vực phát sóng và sản xuất hậu kỳ, chuẩn nén MPEG -2 cuối cùng đã giải quyết được bi kịch đa dạng thức xưa nay. Để đạt được khả năng này, cần nắm vững yêu cầu và đặc tính của từng ứng dụng khác nhau.

- Sản xuất tin thời sự:

Chất lượng ảnh đòi hỏi cao khi đi quay ngoại cảnh, nhưng tín hiệu không phải in đi in lại nhiều lần. Thiết bị nhỏ, gọn, xách tay và dùng ắc quy. Do vậy các thiết bị này có thể sử dụng tốc độ bit thấp 18Mb/s (đỡ tốn năng lượng, giảm giá thành). Cấu trúc GOP đơn giản, thuận tiện cho công đoạn dựng hình.

- Lưu trữ:

Cần chất lượng cao hoặc ít nhất bằng ảnh gốc. Công nghệ lưu trữ còn cần sử dụng tốc độ bit lớn và cấu trúc IB-GOP. Tốc độ lưu trữ là 30 Mb/s.

- Sản xuất hậu kỳ:

Sản xuất hậu kỳ đòi hỏi cả chất lượng và mức độ cao về tính năng dựng hình. Tốc độ 50 Mb/s sản xuất và cấu trúc I-GOP là sự lựa chọn thích hợp.

- Phân phối chương trình:

Phân phối chương trình đòi hỏi khả năng lưu trữ về chất lượng phát sóng quảng bá. Tốc độ 20 Mb/s (hoặc 15 Mb/s để cung cấp cho máy phát NTSC/PAL thậm chí thấp hơn cho dịch vụ phát trực tiếp từ vệ tinh DBS