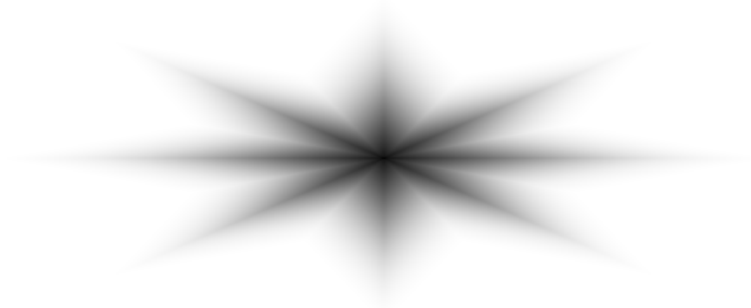
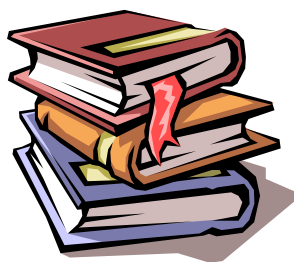


**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG.....**



Đồ án

CÔNG NGHỆ NÉN ẢNH H.264/MPEG - 4 AVC VÀ ỨNG DỤNG



Chương 1

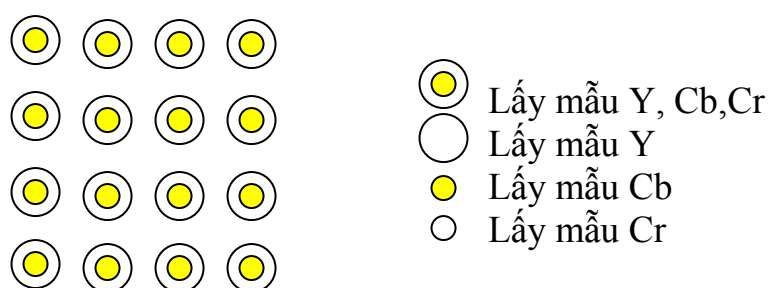
CÁC PHƯƠNG PHÁP NÉN ẢNH CƠ SỞ

1.1. TIÊU CHUẨN VIDEO SỐ THÀNH PHẦN

Trong kỹ thuật viễn thông, truyền hình số thường sử dụng tín hiệu video số thành phần cho cả hai tiêu chuẩn 625/50 và 525/60. Các tiêu chuẩn này khác nhau ở tỷ lệ giữa tần số lấy mẫu và phương pháp lấy mẫu giữa tín hiệu chói và tín hiệu màu (Y:Cb:Cr).

Các tiêu chuẩn đều dùng cấu trúc lấy mẫu loại trực giao với mã PCM lượng tử hoá đều, sử dụng 8 hoặc 16 bit/ mẫu cho tín hiệu chói và màu.

1.1.1. Tiêu chuẩn 4:4:4

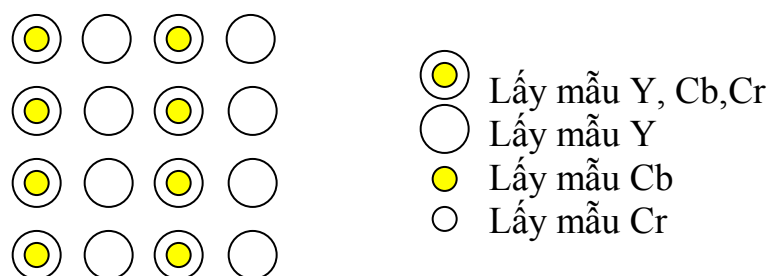


Hình 1.1. Tiêu chuẩn 4:4:4

- Tần số lấy mẫu: Y:13,5 MHz ; Cr/Cb: 13,5 MHz
- Phân bố lấy mẫu: Mật độ lấy mẫu của Y, Cr, Cb là như nhau.
- Tốc độ truyền (phụ thuộc hệ màu):
 - + Lấy mẫu 8 bit: $(720 + 720 + 720) \times 576 \times 8 \times 25 = 249 \text{ Mbit/s}$
 - + Lấy mẫu 10 bit: $(720 + 720 + 720) \times 576 \times 10 \times 25 = 311 \text{ Mbit/s}$

Nhận xét: Tốc độ dòng bit lớn nhất, chất lượng ảnh màu tốt nhất.

1.1.2. Tiêu chuẩn 4:2:2



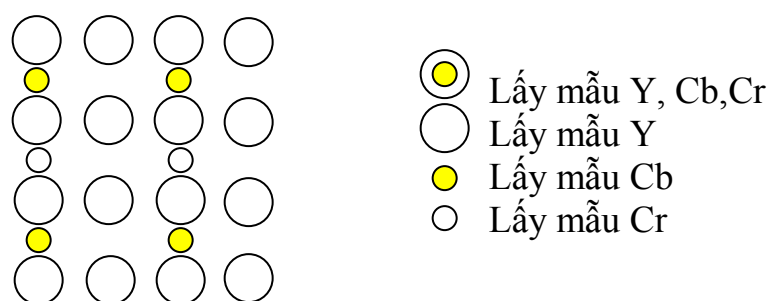
Hình 1.2. Tiêu chuẩn 4:2:2

- Tần số lấy mẫu: Y:13,5 MHz ; Cr/Cb: 6,75 MHz
- Phân bố lấy mẫu: Mật độ lấy mẫu của Y gấp đôi Cr, Cb. Khi giải mã màu điểm ảnh sau được suy từ điểm ảnh trước.
- Tốc độ truyền (phụ thuộc hệ màu):
 - + Lấy mẫu 8 bit: $(720 + 360 + 360) \times 576 \times 8 \times 25 = 166 \text{ Mbit/s}$
 - + Lấy mẫu 10 bit: $(720 + 360 + 360) \times 576 \times 10 \times 25 = 207 \text{ Mbit/s}$

Nhận xét:

Tốc độ truyền của tiêu chuẩn 4:2:2 nhỏ hơn tiêu chuẩn 4:4:4. Vì thế chất lượng ảnh màu kém hơn.

1.1.3. Tiêu chuẩn 4:2:0



Hình 1.3. Tiêu chuẩn 4:2:0

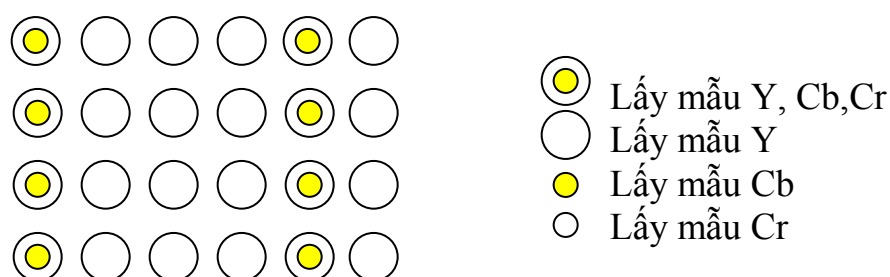
- Tần số lấy mẫu: Y:13,5 MHz ; Cr/Cb: 3,375 MHz
- Phân bố lấy mẫu: Mật độ lấy mẫu của Y gấp 4 lần Cr, Cb và được sắp xếp xen kẽ.

- Tốc độ truyền (phụ thuộc hệ màu):
 - + Lấy mẫu 8 bit: $(720 + 360) \times 576 \times 8 \times 25 = 124,4 \text{ Mbit/s}$
 - + Lấy mẫu 10 bit: $(720 + 360) \times 576 \times 10 \times 25 = 155,5 \text{ Mbit/s}$

Nhận xét:

Tốc độ truyền thấp nhất, chất lượng ảnh màu kém hơn tiêu chuẩn 4:2:2.

1.1.4. Tiêu chuẩn 4:1:1



Hình 1.4. Tiêu chuẩn 4:1:1

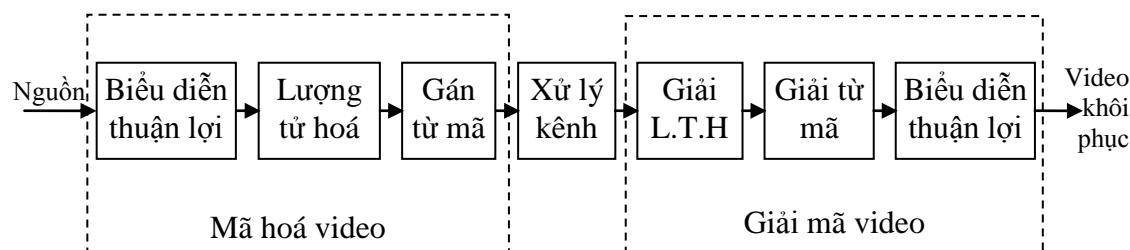
- Tần số lấy mẫu: Y: 13,5 MHz ; Cr/Cb: 3,375 MHz
- Phân bố lấy mẫu: Mật độ lấy mẫu của Y gấp 4 lần Cr, Cb
- Tốc độ truyền (phụ thuộc hệ màu):
 - + Lấy mẫu 8 bit: $(720 + 180 + 180) \times 576 \times 8 \times 25 = 124,4 \text{ Mbit/s}$
 - + Lấy mẫu 10 bit: $(720 + 180 + 180) \times 576 \times 10 \times 25 = 155,5 \text{ Mbit/s}$

Nhận xét:

Tốc độ truyền của tiêu chuẩn 4:1:1 bằng tốc độ truyền của tiêu chuẩn 4:2:0. Nhưng khi giải mã màu của 3 điểm ảnh sau phải suy từ điểm ảnh màu trước đó nên độ thật màu không bằng tiêu chuẩn 4:2:0.

1.2. CÁC KHÁI NIỆM

1.2.1. Mô hình nén ảnh



Hình 1.5. Mô hình hệ thống nén

– Mã hoá video:

+ Ban đầu, tín hiệu video được biểu diễn dưới dạng thuận tiện để nén có hiệu quả nhất. Sự biểu diễn có thể chứa nhiều mẫu thông tin để mô tả tín hiệu và các thông tin quan trọng chỉ tập trung cho một phần nhỏ của sự mô tả này. Trong cách biểu diễn tín hiệu có hiệu quả, chỉ có một phần nhỏ dữ liệu là cần thiết để truyền cho việc tái tạo lại tín hiệu video. Vì vậy điểm cốt yếu là phải xác định cái gì được mã hóa.

+ Lượng tử hoá là quá trình rời rạc hoá thông tin được biểu diễn thành một số hữu hạn các mức để truyền tín hiệu video qua một kênh số.

+ Gán các từ mã là việc biến các từ mã thành một chuỗi bit để biểu diễn các mức lượng tử hoá.

– Bộ giải mã video thì quá trình sẽ diễn ra ngược lại.

1.2.2. Dư thừa thông tin trong tín hiệu video

Nén số liệu là quá trình giảm lượng số liệu cần thiết để biểu diễn cùng một lượng thông tin cho trước. Giữa số liệu và thông tin có sự khác nhau, số liệu chỉ là phương tiện để truyền tải thông tin. Cùng một lượng thông tin cho trước có thể biểu diễn bằng các lượng số liệu khác nhau. Và điều này gây ra dư thừa số liệu.

Độ dư thừa số liệu là vấn đề trung tâm trong nén ảnh số. Để đánh giá độ dư thừa người ta đưa ra tỉ lệ nén (C_N).

Gọi N_1 và N_2 là lượng số liệu trong hai tập hợp số liệu cùng được biểu diễn một lượng thông tin cho trước thì độ dư thừa số liệu tương đối (R_D) của tập hợp số liệu thứ nhất so với tập hợp số liệu thứ hai được định nghĩa bởi hệ thức sau:

$$R_D = 1 - \frac{1}{C_N} \quad \text{Trong đó: } C_N = \frac{N_1}{N_2}$$

⇒ **Nhận xét:**

- Nếu $N_1 = N_2$ thì $C_N = 1 \Rightarrow R_D = 0$ Không có số liệu dư thừa.
- Nếu $\begin{cases} N_1 \neq N_2 \\ N_1 \leq N_2 \end{cases}$ thì $C_N = \infty \Rightarrow R_D \rightarrow 0$ Độ dư thừa số liệu tương đối của

tập số liệu thứ nhất là khá lớn so với tập dữ liệu thứ hai.

⇒ **Lưu ý:** Tỷ lệ nén càng cao sẽ làm giảm chất lượng hình ảnh và ngược lại. Trong đó chất lượng hình ảnh được tính bằng số bit cho một điểm ảnh trong ảnh nén, ký hiệu là N_b

$$N_b = \text{Số bit nén} / \text{Số điểm}$$

a) Dư thừa thống kê

Hầu như tất cả các ảnh đều chứa thông tin trùng lặp và tạo ra sự dư thừa thông tin. Sự dư thừa này không chỉ tồn tại trong phạm vi một bức ảnh (gọi là dư thừa trong không gian) mà còn trong các bức ảnh liên nhau trong chuỗi các bức ảnh tạo thành khung cảnh truyền hình (gọi là dư thừa theo thời gian). Tập hợp các dư thừa này gọi là dư thừa thống kê.

b) Dư thừa do cảm nhận sinh lý của mắt người

Mắt người chỉ phân biệt được có giới hạn tín hiệu chói và tín hiệu màu. Do khái niệm lưu ảnh của mắt cho nên có thể loại bỏ những thông tin vượt quá khả năng nhận biết của mắt người.

1.2.3. Sai lệch bình phương trung bình (RMS)

Sai lệch bình phương trung bình là hệ số cho phép đánh giá các giải thuật nén, chỉ ra sự khác nhau thống kê giữa ảnh nén và ảnh gốc. Ký hiệu là RMS (Root Mean Square) được tính bởi biểu thức:

$$RMS = \frac{1}{n} \sqrt{\sum_{i=0}^n (X_i - X_i')^2}$$

Trong đó: RMS: sai lệch bình phương trung bình

X_i : Giá trị điểm ảnh ban đầu

X_i' : Giá trị điểm ảnh sau khi giải nén

n : Tổng số điểm ảnh trong một ảnh

1.3. LÝ THUYẾT THÔNG TIN - ENTROPY

Lượng thông tin chứa đựng trong một chi tiết ảnh tỉ lệ nghịch với khả năng xuất hiện của nó. Lượng thông tin của một hình ảnh bằng tổng số lượng thông tin của từng phần tử ảnh. Khi đó ta xét đến:

– **Entropy** đo giá trị thông tin trung bình chứa đựng trong một bức ảnh và do đó entropy xác định lượng thông tin trung bình nhỏ nhất biểu diễn bởi mỗi giá trị nhị phân qua quá trình mã hoá để bảo toàn được khả năng khôi phục được ảnh gốc. Từ đó ta có nhận xét:

- Độ dài trung bình của từ mã qua một phương pháp nén không thể nhỏ hơn entropy của bức ảnh được mã hoá.
- Tốc độ bit sau khi nén nhỏ hơn giới hạn entropy của bức ảnh.

– **Lượng thông tin của từng phần tử ảnh:**

$$l(x_i) = \log_2 \frac{1}{P(x_i)} = -\log_2 P(x_i)$$

Trong đó: $l(x_i)$: lượng thông tin của phần tử ảnh x_i

$P(x_i)$: xác suất xuất hiện của phần tử ảnh x_i

Nếu một hình ảnh được biểu diễn bằng các phần tử $x_1, x_2, x_3 \dots$ thì xác suất hiện của các phần tử ảnh tương ứng là $P(x_1), P(x_2), P(x_3), \dots$

– **Lượng tin tức bình quân của hình ảnh (entropy của hình ảnh):**

$$H(x) = \sum_{i=0}^n P(x_i) l(x_i) = \sum_{i=0}^n P(x_i) \log_2 \frac{1}{P(x_i)}$$

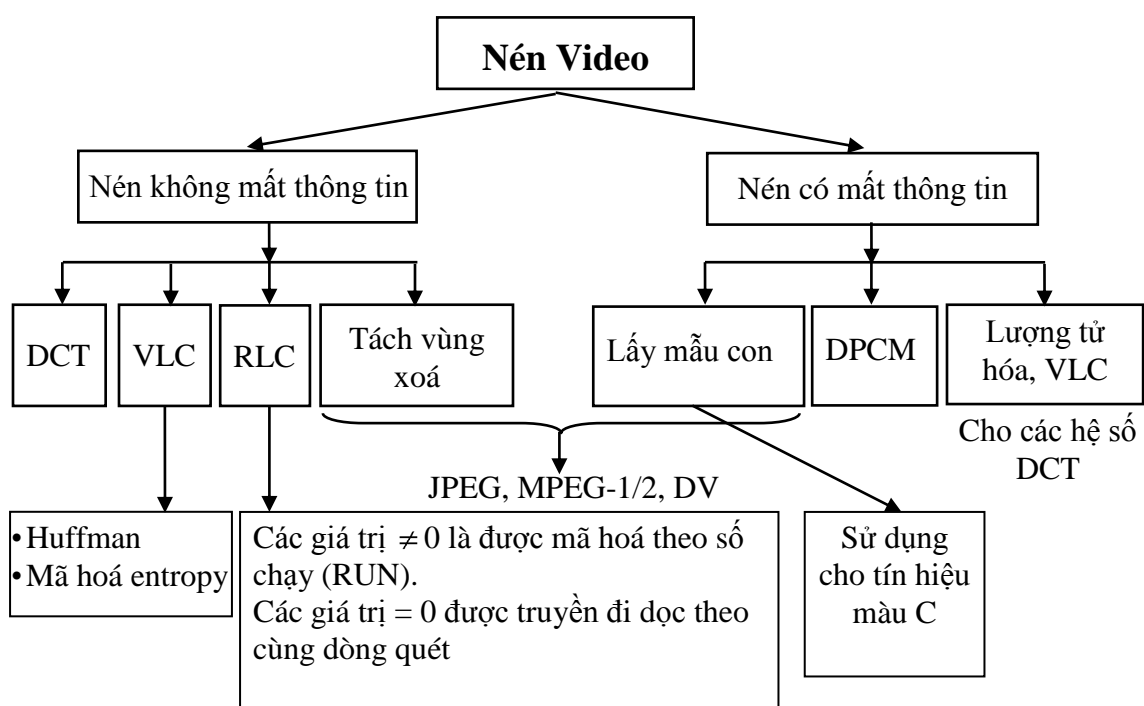
Entropy của hình ảnh xác định số lượng bit trung bình tối thiểu cần thiết để biểu diễn một phần tử ảnh. Trong công nghệ nén không tổn hao, entropy là

giới hạn dưới của tỉ số bit/pixel. Nếu tín hiệu video được nén với tỉ số bit/phần tử nhỏ hơn entropy, hình ảnh sẽ bị mất thông tin và quá trình nén sẽ có tổn hao.

1.4. CÁC PHƯƠNG PHÁP NÉN VIDEO

Các hệ thống nén là sự phối hợp của rất nhiều các kỹ thuật xử lý nhằm giảm tốc độ bit của tín hiệu số mà vẫn đảm bảo chất lượng ảnh phù hợp với một ứng dụng nhất định.

Hình 1.6 phân loại các kỹ thuật nén được sử dụng trong các chuẩn nén JPEG (Joint photographic Expert Group) và MPEG (Moving Picture Expert Group).



Hình 1.6. Các phương pháp nén và sự phối hợp kỹ thuật trong JPEG & MPEG

1.4.1. Nén không mất thông tin

Cho phép khôi phục lại đúng tín hiệu ban đầu sau khi giải nén. Hệ số nén nhỏ hơn 2:1.

a) Mã hóa với độ dài biến đổi (VLC)

Phương pháp này còn được gọi là mã hoá Huffman và mã hoá Entropy dựa trên khả năng xuất hiện của các biên độ trùng hợp trong một bức ảnh. Nó thiết lập một từ mã ngắn cho các giá trị có tần suất xuất hiện cao nhất và từ mã dài cho các giá trị còn lại.

b) Mã hoá với độ dài động (RLC)

RLC dựa trên sự lặp lại của cùng giá trị mẫu để tạo ra các từ mã đặc biệt biểu diễn sự bắt đầu và kết thúc của giá trị được lặp lại. Vì các mẫu có giá trị khác không mới được mã hoá, các mẫu có giá trị bằng không sẽ được truyền đi dọc theo cùng dòng quét.

c) Sử dụng khoảng xoá dòng và mảnh

Các thông tin xoá dòng và xoá mảnh sẽ không được ghi giữ và truyền đi mà được thay thế bằng các dữ liệu đồng bộ ngắn hơn tùy theo các ứng dụng.

d) Biến đổi cosin rời rạc (DCT)

DCT là phương pháp biến đổi tín hiệu rời rạc bằng hàm cosin. Trong đó mỗi một mảng 8 x 8 điểm ảnh sẽ được mã hóa bằng phương pháp DCT.

Quá trình DCT thuận và nghịch được coi là không mất thông tin nếu độ dài từ mã hệ số là 13 hoặc 14 bằng tần đối với dòng video số sử dụng 8 bit biểu diễn mẫu.

1.4.2. Nén có mất thông tin

Là sau khi nén một số thông tin sẽ bị mất và chất lượng ảnh bị suy hao do quá trình làm tròn và loại bỏ giá trị trong phạm vi khung hình hay giữa các khung hình. Hệ số nén cho phép từ 2:1 đến 100:1.

a) Lấy mẫu con (Subsampling)

Đây là phương pháp nén rất có hiệu quả nhưng độ phân giải của ảnh sau khi giải nén giảm so với hình ảnh ban đầu. Kỹ thuật này chỉ áp dụng cho lấy mẫu tín hiệu màu với tín hiệu video số thành phần, nhờ các cấu trúc lấy mẫu cho phép giảm tốc độ dữ liệu dòng bit (ví dụ cấu trúc 4:2:0; 4:1:1).

b) Điều xung mã visai (DPCM)

DPCM là phương pháp mã hóa dự đoán thay vì truyền đi cả một khung mẫu, kỹ thuật này chỉ mã hóa và truyền đi sự khác nhau giữa các giá trị mẫu. Giá trị sai lệch được cộng vào giá trị mẫu đã được giải mã trong quá trình giải nén để tạo lại giá trị mẫu cần thiết. Quá trình DPCM làm giảm lượng entropy của tín hiệu ban đầu.

Để hoàn thiện thêm thì kỹ thuật nén DPCM sử dụng các kỹ thuật dự đoán và lượng tử hoá thích nghi.

c) Lượng tử hoá và mã hoá VLC các hệ số DCT

Phối hợp 3 kỹ thuật này cho phép biểu diễn một khối các điểm ảnh bằng số ít các bit do đó tạo được hình ảnh nén cao.

1.5. MỘT SỐ MÃ DÙNG TRONG KỸ THUẬT NÉN

1.5.1. Mã hóa loạt dài RLC (Run Length Coding)

– Nguyên tắc:

Bước 1: Phát hiện loạt (loạt bit 0 giữa hai bit 1 hoặc ngược lại).

Bước 2: Ký hiệu lặp. Thay loạt bằng một chuỗi mới gồm chiều dài loạt (run length) và ký tự lặp. Ví dụ 12 giá trị 0 chỉ cần ghi ESC120 thay cho phải ghi 12 từ mã.

– Đặc điểm:

- + Chỉ có hiệu quả với chiều dài loạt lớn.
- + Tỷ lệ nén chưa cao → mã hóa loạt dài thích nghi hay biến đổi VLC (Mã hóa Huffman và mã hóa entropy).

1.5.2. Mã Huffman

– Nguyên tắc: Dựa vào mô hình thống kê của dữ liệu gốc, ký tự có xác suất càng cao thì mã hóa với từ mã càng ngắn.

– Thuật toán:

Bước 1: Sắp xếp xác suất của các ký hiệu theo thứ tự giảm dần.

Bước 2: Xét từ dưới lên trên, bắt đầu từ hai ký hiệu có xác suất bé nhất. Qui định mỗi nhánh là 0 (hoặc 1) hợp lại với nhau thành nút có xác suất bằng tổng hai xác suất hợp thành (nhánh trên có xác suất lớn hơn nhánh dưới).

Bước 3: Lặp lại cho đến hết (khi tập nút chỉ còn chứa 1 nút).

Xét ví dụ trên:

Ký hiệu	Xác suất		Mã	Tổng bit
A	15/39	1	1	15
B	7/39	0	000	21
C	6/39	1 13/39	001	18
E	6/39	0 24/39	010	18
D	5/39	1 11/39	011	15

Số bit trung bình: $87 / 39 = 2,23 (< 2,28) \rightarrow$ Hiệu quả hơn Shannon-Fano

↳ **Đặc điểm**: Độ dài trung bình của từ mã giới hạn trên bởi:

$$H + P + \log[2 \cdot \log e/e] = H + P + 0,086$$

Trong đó: H: entropy của nguồn

P: xác suất của tất cả các ký hiệu như nguồn.

1.5.3. Mã hóa DCT (Quá trình biến đổi cosin rời rạc)

Trong đó mỗi một mảng 8 x 8 điểm ảnh sẽ được mã hóa bằng phương pháp DCT thuận theo hệ thức sau:

$$F(u, v) = \frac{C(u)C(v)}{4} \sum_{i=0}^7 \sum_{j=0}^7 f(i, j) \cos \frac{(2i+1)u\pi}{16} \cos \frac{(2j+1)v\pi}{16}$$

Trong đó: f(i,j): Các mẫu ban đầu trong khối 8 x 8 điểm ảnh

F(u,v): Các hệ số biến đổi DCT khối 8 x 8

u: Tần số chuẩn hóa theo chiều ngang ($0 < u < 7$)

v: Tần số chuẩn hoá theo chiều đứng ($0 < v < 7$)

$$C(u), C(v) = \begin{cases} \frac{1}{\sqrt{2}} & \text{neu } u, v = 0 \\ 1 & \text{neu } u, v = 1, 2, 3, \dots, 7 \end{cases}$$

a) DCT một chiều

Hình 1.7 minh họa quá trình mã hoá DCT một chiều gồm 8 điểm ảnh biểu diễn tín hiệu chói theo chuẩn lấy mẫu 4:2:2.

b) DCT hai chiều, ba chiều

DCT hai chiều sẽ lần lượt thực hiện theo hàng và theo cột.

Thành phần DC là giá trị đầu tiên bên trái của khối 8 x 8 điểm ảnh. Nó biểu diễn mức trung bình một chiều của tín hiệu trong khối 8 x 8 điểm chói. Và được xác định bởi:

$$F(0,0) = \frac{1}{8} \sum_{i=0}^7 \sum_{j=0}^7 f(i, j) = 591$$

Thành phần AC là các giá trị còn lại, biểu diễn các thành phần tần số cao hơn trong tín hiệu ban đầu.

Hình 1.8 và 1.9 minh họa quá trình mã hóa DCT hai và ba chiều.

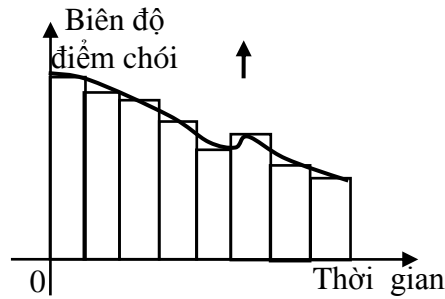
↳ **Nhận xét:**

+ DCT làm giảm độ tương quan không gian của thông tin trong khối. Vì vậy, việc biểu diễn DCT có độ dư thừa thông tin ít hơn.

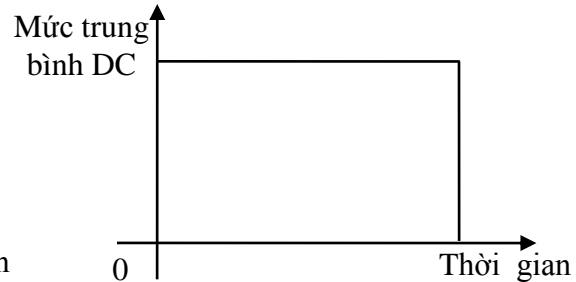
+ Đồng thời DCT chứa thông tin về nội dung tần số không gian của thông tin trong khối. Dựa vào đặc tính sinh lý thị giác, ta chỉ mã hóa những hệ số DCT quan trọng (gọi là nén).

8 điểm chói liên nhau trên một dòng

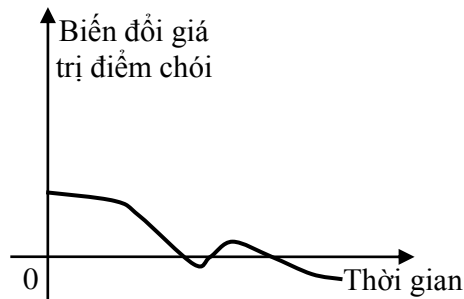
98	92	95	80	75	82	68	50
----	----	----	----	----	----	----	----



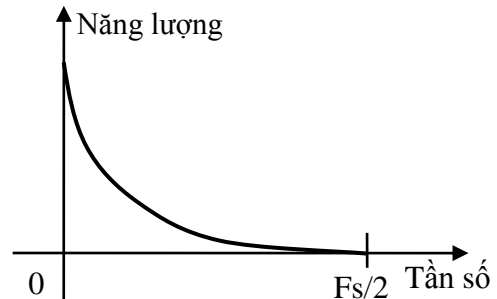
a) Biên độ chói của 8 điểm ảnh liên tiếp trên một dòng



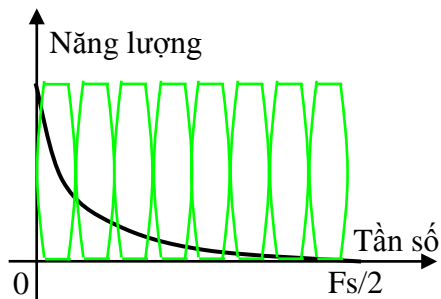
b) Mức trung bình DC của khối 8 điểm ảnh



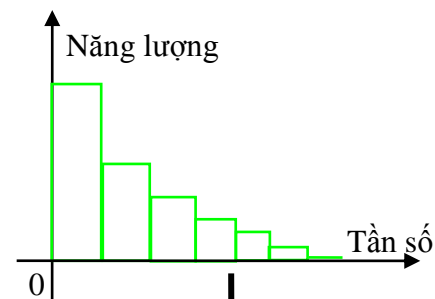
c) Biến đổi biên độ so với mức trung bình



d) Phổ khối 8 x 8 điểm chói



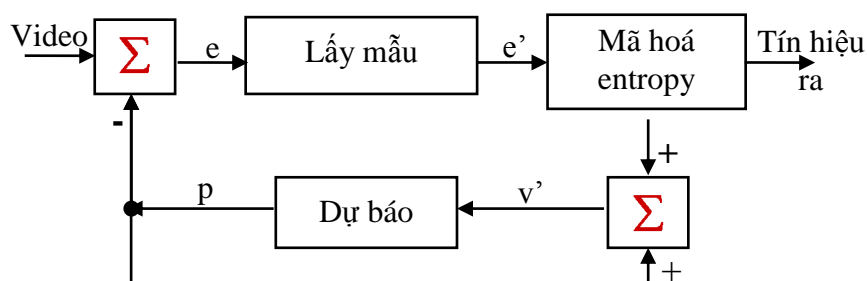
e) Phân chia các dải phổ



591	106	-18	28	-34	14	18	3
Hệ số DC		Các hệ số AC					

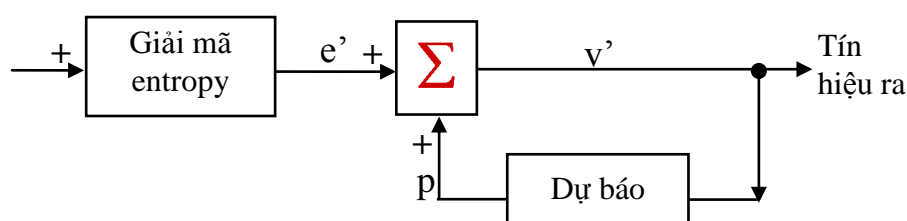
f) Giá trị hệ số DCT

Hình 1.7. Minh họa quá trình mã hoá DCT một chiều



Trong đó:

- e : sai số dự báo ($e = v - p$)
- e' : sai số lượng tử hoá
- v' : tín hiệu khôi phục ($v' = e' + p$)

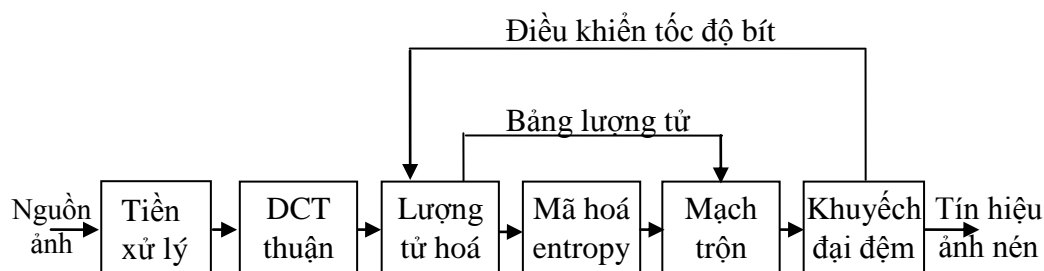


Hình 1.10. Mã hóa và giải mã DPCM

1.6. KỸ THUẬT NÉN TRONG ẢNH

1.6.1. Nguyên lý nén trong ảnh

Nén trong ảnh là loại nén làm 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. Quá trình này không sử dụng thông tin các ảnh trước và ảnh sau của ảnh đang xét. Trong đó kỹ thuật chính là phương pháp biến đổi DCT.

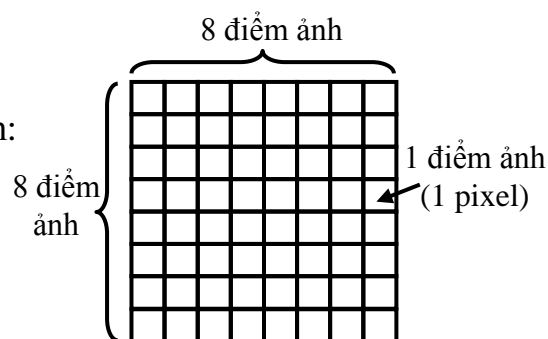


Hình 1.11. Nén trong ảnh (Intra Frame Compression)

– Miền thời gian cho phép biến đổi DCT được xác định là một mảng 8x8 điểm ảnh.

Trong đó, thông tin của 1 điểm ảnh bao gồm:

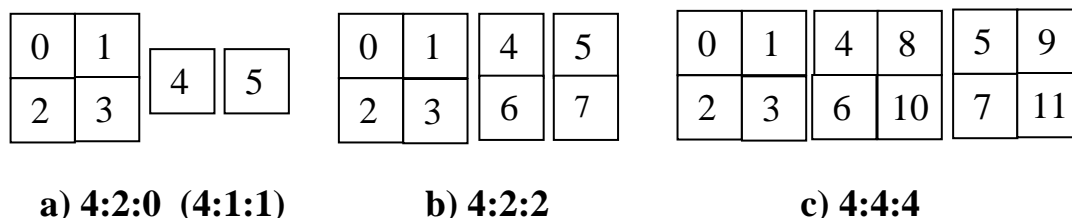
- Toạ độ của một điểm ảnh (x,y).
- Mức tín hiệu chói Y.
- Mức tín hiệu màu Cb, Cr.



1 mảng 8 x 8 điểm ảnh (1 block)

1.6.2. Tiền xử lý

Trước khi thực hiện biến đổi DCT thì ảnh được chia thành các khối riêng biệt không trùng nhau MB (Macro Block). Mỗi MB bao gồm 4 block các mẫu tín hiệu chói Uy và 2, 4 hoặc 8 block các mẫu tín hiệu màu Cb, Cr. Số các block các mẫu tín hiệu màu phụ thuộc vào tiêu chuẩn của tín hiệu video, hình 1.12



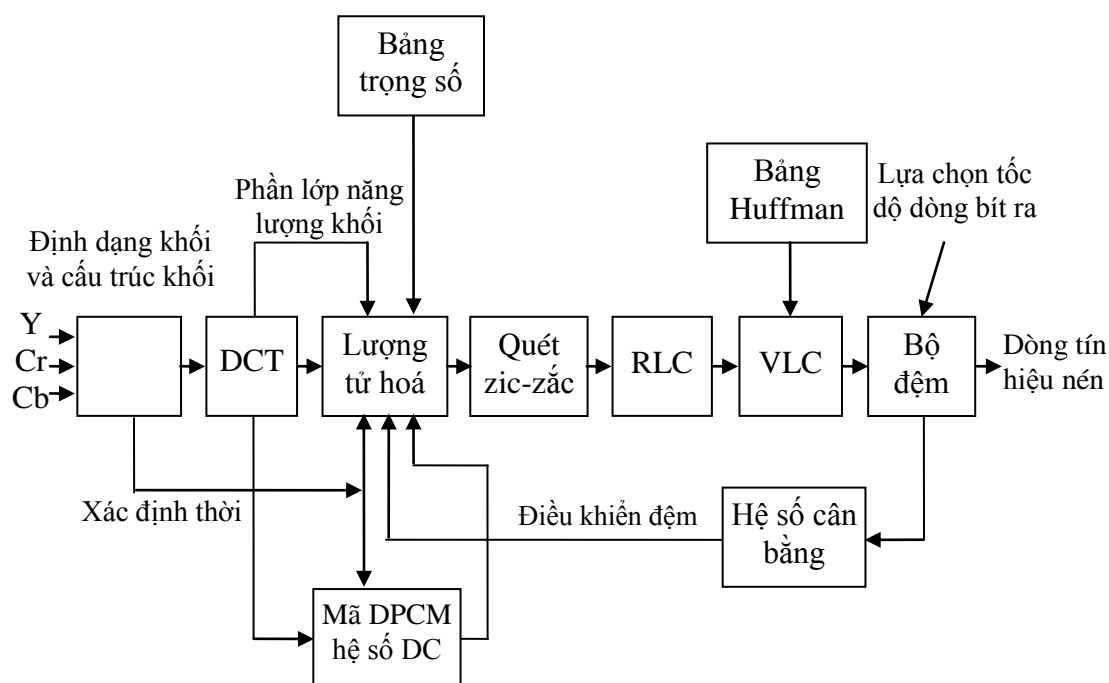
Hình 1.12. Cấu tạo của Macro Block

Tất cả các block có cùng kích thước và mỗi block là một ma trận điểm ảnh 8 x 8 được lấy từ một ảnh màn hình theo chiều từ trái qua phải, từ trên xuống dưới. Kích thước block được chọn dựa vào:

- Yêu cầu của phương pháp nén là loại bỏ các dư thừa về không gian, vì thế không cần quan tâm đến các khối pixel có kích thước lớn hơn 8 x 8.
- Sự tiện lợi cho việc tính toán và thiết kế mạch cứng.

1.6.3. Lượng tử hoá khối DCT

Tín hiệu Y, Cr, Cb sau tiền xử lý chia thành các MB (Macro block). Các mẫu $f(i,j)$ trong từng Block qua phép biến đổi DCT thuận tạo thành các hệ số biến đổi $F(u,v)$ sẽ được lượng tử hóa theo sơ đồ hình 1.13



Hình 1.13. Quá trình nén ảnh bằng phương pháp DCT

a) Nguyên tắc lượng tử

Chia các hệ số $F(u, v)$ cho các hệ số ở vị trí tương ứng trong bảng lượng tử $Q(u,v)$. Các hệ số tần thấp được chia cho các giá trị nhỏ, các hệ số ứng với tần cao được chia cho các giá trị lớn hơn. Sau đó các hệ số được làm tròn (bỏ đi các phần thập phân).

Kết quả nhận được bảng $F_q(u, v)$ mới. Trong đó phần lớn các hệ số tần cao sẽ bằng 0. Hệ số bảng lượng tử hóa thuận được xác định theo biểu thức:

$$F_q(u, v) = \left\lfloor \frac{F(u, v)}{Q(u, v)} \right\rfloor = \text{giá trị nguyên gần nhất} \left\lfloor \frac{F(u, v) + Q(u, v)/2}{Q(u, v)} \right\rfloor$$

Các giá trị $F_q(u, v)$ sẽ được mã hóa trong các công đoạn tiếp theo.

b) Bảng lượng tử (Bảng trọng số)

Bảng lượng tử được xây dựng theo nguyên tắc: mắt người ít cảm nhận được nội dung ở tần số cao và đặc biệt càng kém nhạy với nội dung ở tần số cao của kênh màu. Do đó:

- + Các hệ số tương ứng với thành phần DC và các thành phần tần số thấp có giá trị lớn nên phải được lượng tử chính xác.
- + Các hệ số tương ứng với thành phần tần số cao (AC) có giá trị nhỏ nên cho phép sai số, ta có thể chia với một số lớn hơn.

Bảng 1.1. Bảng lượng tử cho tín hiệu chói và màu:

16	11	10	16	24	40	51	61
12	12	14	19	26	58	60	55
14	14	16	24	40	57	69	56
14	17	22	29	51	57	69	56
18	12	37	56	68	109	103	77
24	35	55	64	81	104	113	92
49	64	78	87	103	121	120	101
72	92	95	98	112	100	103	99

Ma trận lượng tử kênh chói

17	18	24	47	99	99	99	99
18	21	26	66	99	99	99	99
24	26	56	99	99	99	99	99
47	66	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99
99	99	99	99	99	99	99	99

Ma trận lượng tử kênh màu

1.6.4. Mã hoá entropy

Mã hoá entropy gồm các quá trình sau:

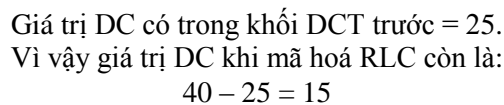
a) Quét zíc-zắc

Quét zíc-zắc nhằm biến đổi mảng hai chiều các hệ số $C'(u,v)$ thành chuỗi số một chiều. Với ma trận 8x8 sẽ thành vector 1x64 như ở hình 1.14.

b) Mã RLC và DPCM

– **Mã DPCM:** Do giá trị của thành phần DC là lớn, thay đổi nhưng gần với giá trị của block trước đó → Điều chế DPCM cho thành phần DC.

– **Mã RLC:** Thành phần AC sau quá trình quét zíc-zắc → Các hệ số có giá trị 0 giống nhau sẽ được thay bằng mã RLC. Dấu EOB (End of Block) được dùng để đánh dấu vị trí bắt đầu của chuỗi các số 0 liên tiếp.

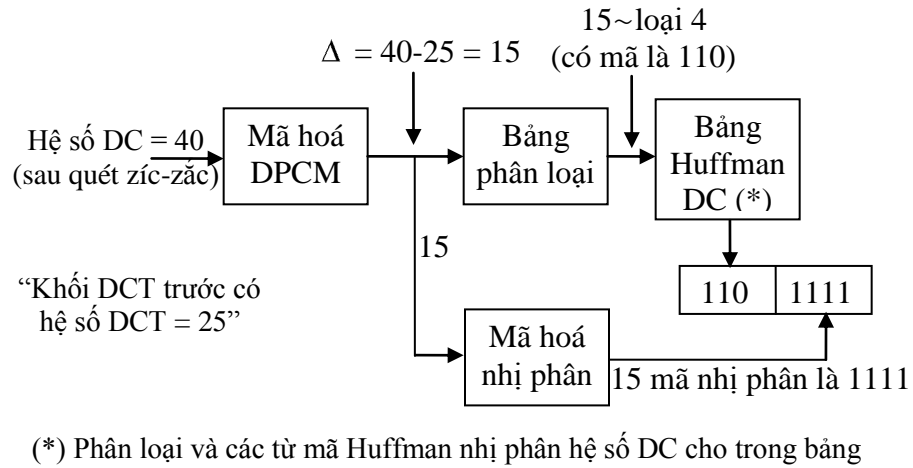


Hình 1.14. Ví dụ về mã RLC

c) Mã hoá VLC

Sau biến đổi RLC các từ mã có tần suất xuất hiện cao sẽ được mã hoá bằng các từ mã ngắn. Và các từ mã có tần suất xuất hiện thấp sẽ được mã hoá bằng các từ mã dài. Quá trình này được gọi là phương pháp mã hoá có độ dài từ mã thay đổi (VLC).

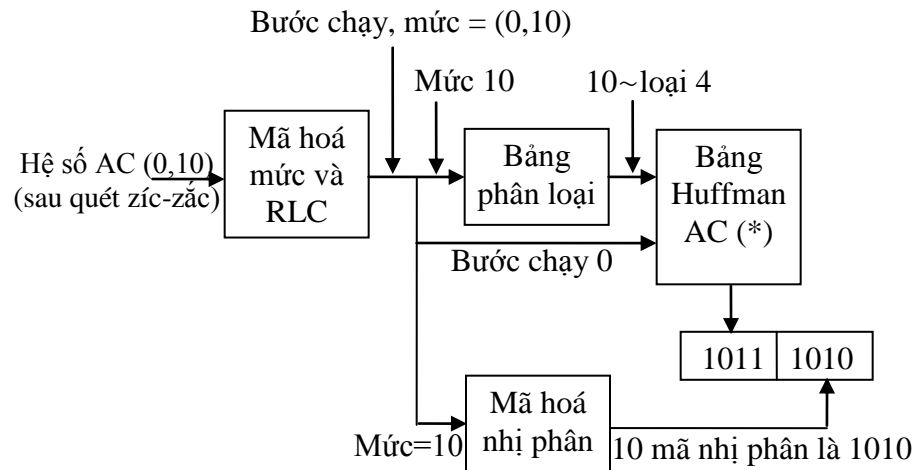
- Với thành phần DC: Giá trị sai lệch hệ số DC sẽ được mã hóa nhờ bảng phân loại và bảng Huffman (dựa vào đặc tính thống kê của tín hiệu) như ở hình 3.11.
- Với thành phần AC: Hệ số AC cũng được mã hóa nhờ bảng phân loại (giống như DC) và bảng Huffman (nhưng khác DC) như ở hình 1.15.



Hình 1.15. Mã hoá entropy thành phần DC

Bảng 1.2. Bảng phân loại và bảng Huffman cho thành phần DC chẵn

Các hệ số DC sai lệch	Phân loại	Từ mã (chẵn)
-255...-128; 128...255	8	1111110
-127...-64; 64...127	7	1111 10
-63...-32; 32...63	6	1111 0
-31...-16; 16...31	5	1110
-15...-8; 8...15	4	110
-7...-4; 4...7	3	101
-3; -2; 2; 3	2	01
-1; 1	1	00
0	0	100



(*) Phân loại, bước chạy và các từ mã Huffman nhị phân hệ số AC cho trong bảng

Hình 1.16. Mã hoá entropy thành phần AC

Bảng 1.3. Bảng Huffman các hệ số AC

Bước chạy	Phân loại	Độ dài mã	Từ mã
0	1	2	00
0	2	2	01
0	3	3	100
0	4	4	1011
1	1	4	1100
1	2	6	111001
2	1	5	11011
2	2	8	11111 000
3	1	6	111 010
4	1	6	111 011
5	1	7	1111 010
6	1	7	1111 011
EOB		4	1010

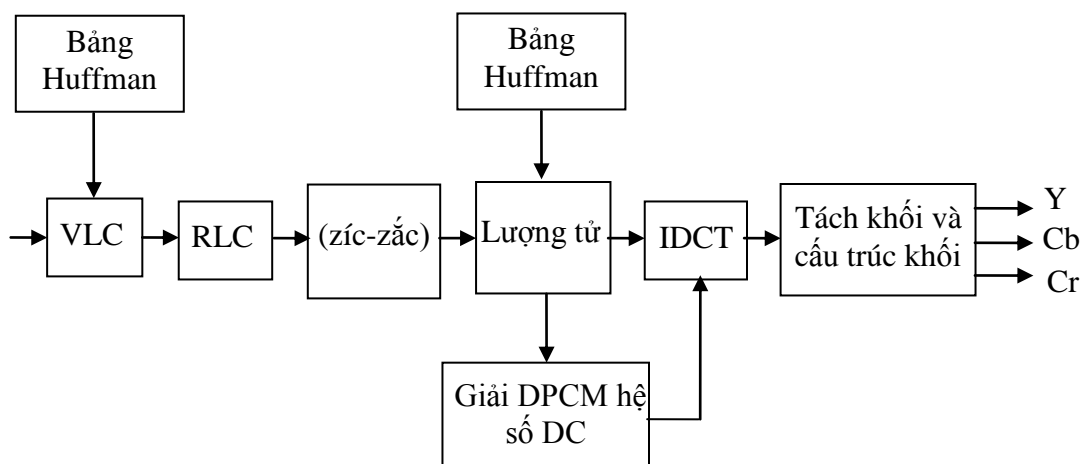
1.6.5. Bộ nhớ đệm

Các từ mã VLC tạo ra dòng số liệu biến đổi, phụ thuộc vào độ phức tạp của ảnh được mã hoá. Số liệu này sẽ được ghi vào bộ nhớ đệm. Các bit số liệu sẽ được đọc ra với tốc độ cố định theo bộ mã hoá.

Bộ đệm cần được điều khiển sao cho không xuất hiện hiện tượng tràn hay rỗng. Quá trình điều khiển được thực hiện bởi hệ số cân bằng cho bảng trọng số. Nếu bộ đệm đầy thì thực hiện lượng tử hoá thô nhằm tạo ra tốc độ dòng bit nhỏ lại (bằng cách tăng hệ số cân bằng của bộ lượng tử).

1.6.6. Giải mã DCT

❖ Sơ đồ khối của quá trình giải nén DCT:



Hình 1.17. Giải mã DCT

– Các bảng lượng tử và bảng Huffman xác định tại bộ mã hoá DCT được sử dụng để phục hồi các giá trị hệ số DCT của khối 8x8 điểm. Quá trình lượng tử hoá ngược $R_q(u,v)$ được tính theo biểu thức:

$$R_q(u,v) = F_q(u,v) Q(u,v)$$

– Các hệ số sẽ được biến đổi ngược DCT (IDCT) bằng quá trình $f^*(j,k)$ để tạo lại khối giá trị các điểm ban đầu theo biểu thức:

$$f^*(j,k) = \frac{1}{4} \sum_{u=0}^7 \sum_{v=0}^7 C(u)C(v) F(u,v) \cos \frac{(2j+1)u\pi}{16} \cos \frac{(2k+1)v\pi}{16}$$

Biểu thức biểu diễn quá trình biến đổi DCT và IDCT là tương tự nhau. Vì vậy, hệ thống nén và giải nén có thể dùng chung bằng một thiết bị phần cứng.

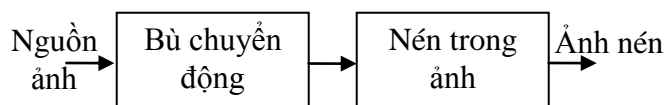
Sai lệch giữa khối ban đầu và các thiết bị tạo lại nguyên nhân do các lỗi xuất hiện trong quá trình nén. Lỗi được biểu diễn bởi :

$$e(j,k) = f(j,k) - f^*(j,k)$$

1.7. KỸ THUẬT NÉN NHÓM ẢNH

Tín hiệu video có chứa các thông tin dư thừa trong miền thời gian. Với một chuỗi ảnh liên tiếp, lượng thông tin chứa đựng trong mỗi ảnh thay đổi rất ít từ ảnh này sang ảnh khác. Vì vậy, nếu các ảnh liên tiếp mà giống nhau thì chỉ cần lấy thông tin của một ảnh và lưu nó lại, các ảnh giống nhau còn lại không cần phải lấy thông tin mà được tạo lại nhờ thông tin của một ảnh đã được lưu theo một quy luật mã hoá đã được lập trình từ trước.

1.7.1. Mô hình



Hình 1.18. Mô hình nén ảnh

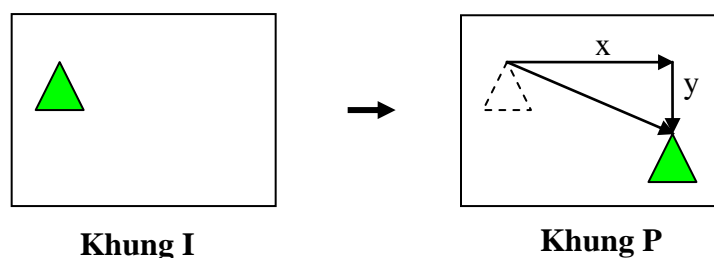
Trong đó:

- + Bù chuyển động loại bỏ phần dư thừa tạm thời giữa các frame kế nhau.
- + Kỹ thuật nén trong ảnh xử lý độ dư thừa trong không gian để tăng hiệu quả nén.

1.7.2. Kỹ thuật dự đoán bù chuyển động

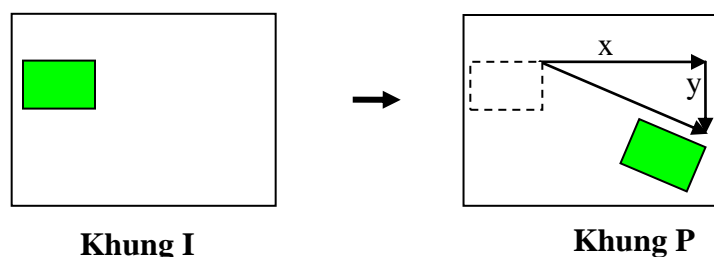
VD1: Xét ảnh là một hình tam giác xuất hiện lần lượt theo thời gian.

- + Thời điểm t_1 , hình tam giác xuất hiện ở góc trái (khung I)
- + Thời điểm t_2 , hình tam giác xuất hiện ở góc phải (khung P)



↳ **Nhận xét:** Hình tam giác ở khung P là hình tam giác ở khung I dịch chuyển đi một vectơ (x, y) . Ta chỉ cần lấy thông tin ở khung I cộng thêm một giá trị vectơ dịch chuyển (x, y) để suy ra thông tin của khung P. Việc làm này làm giảm bộ nhớ (nén ảnh). MPEG lấy ra các hình ảnh liên tiếp nhau để nghiên cứu các vectơ di chuyển. Việc tạo ra các vectơ di chuyển x, y gọi là “dự đoán”.

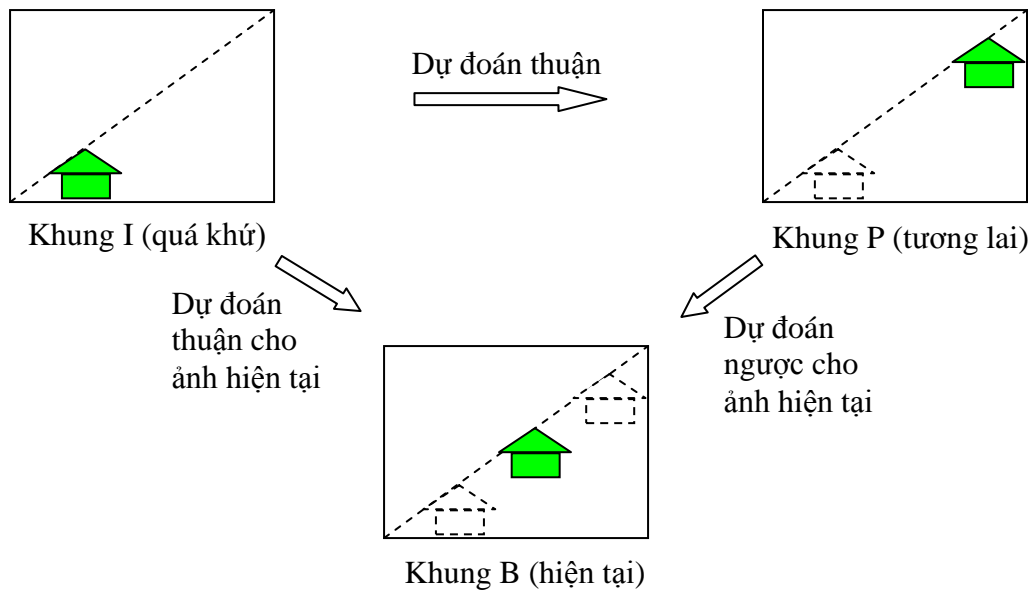
VD2: Xét ảnh là hình chữ nhật xuất hiện ở hai thời điểm



↳ **Nhận xét:** Ảnh ở khung P là ảnh của khung I dịch chuyển đi một vectơ (x, y) và xoay 5° . Vì thế một sự dịch chuyển đơn giản sẽ gây ra một lỗi (lỗi dự đoán). Do đó để suy ra ảnh của khung P phải có hai bước:

- Sử dụng ảnh của khung I và vectơ (x, y) .
- Hiệu chỉnh lỗi dự đoán (xác định hình chữ nhật xoay 5°).

VD3: Xét ảnh ở thời điểm khác nhau (t_1, t_2, t_3)



⇒ **Nhận xét:**

– Có thể dùng phương pháp biến đổi thuận ngược để suy ra ảnh hiện tại từ ảnh quá khứ hoặc ảnh tương lai. Kỹ thuật nén ảnh theo thời gian phải sử dụng 3 loại khung hình:

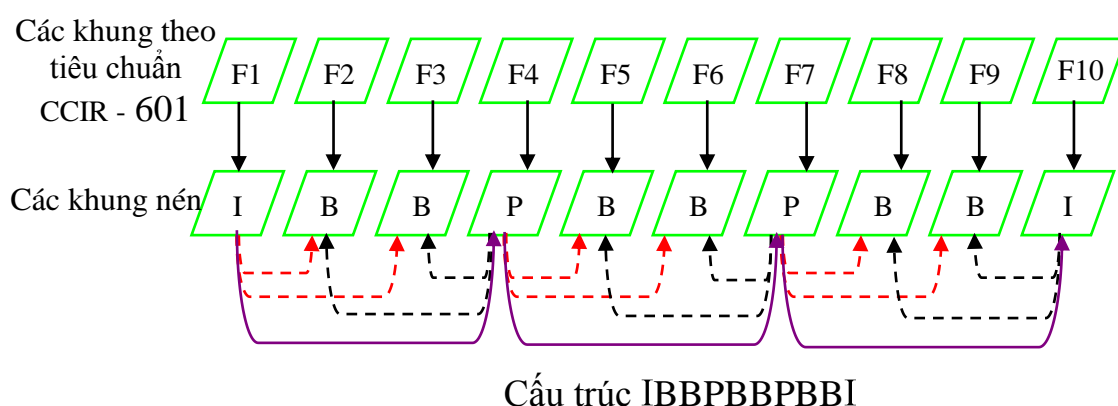
- **Khung I:** là khung được mã hoá cho ảnh thực (mã hoá nội bộ). Khung hình này được tạo ra không cần đến các khung hình khác.
- **Khung P:** là các khung được dự đoán thuận từ khung I hoặc từ khung B ngay trước nó. Khung P này không thể được tạo ra nếu không có các khung khác.
- **Khung B:** là các khung được dự đoán vừa thuận vừa ngược từ các khung I và P nằm lân cận trước và sau chúng.

– Về dung lượng: khung I là lớn nhất, khung B là nhỏ nhất. Do đó, thay vì sử dụng khung I, người ta sử dụng khung P và B để giảm dung lượng nhớ (nén hình theo thời gian).

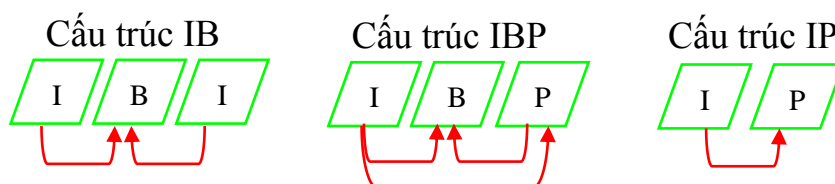
– Thứ tự mã hoá: đầu tiên là khung I, tiếp theo là khung P, sau đó là khung B. Thông thường cứ 15 khung thì chèn 1 khung với chu kỳ 0,5s.

– Tập hợp các khung I, P, B một cách hợp lý tạo thành một nhóm ảnh GOP (Group Of Picture). Nhóm ảnh này được tập hợp thành các chuỗi, thứ tự ảnh đầu ra sẽ khác với thứ tự ảnh khi đưa vào mã hoá. Chuỗi GOP có 2 thông số:

- m: số khung từ khung I đầu tiên cho đến khung cuối cùng B (P) - trước khung I tiếp theo.
- n: số khung B giữa hai khung P



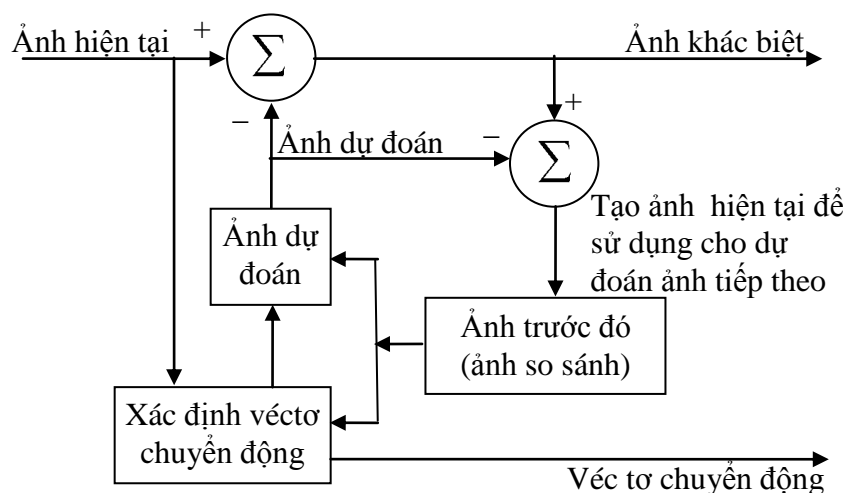
- Dự đoán thuận của khung P
 Dự đoán thuận của khung B
 Dự đoán ngược của khung B



Hình 1.19. Nhóm ảnh GOP trong hệ thống liên ảnh

1.7.3. Ảnh dự đoán trước (ảnh P)

Phương pháp mã hoá ảnh dự đoán trước sử dụng xác suất các ảnh liên tục trong chuỗi truyền hình. Ảnh dự đoán trước là các khung dự báo theo hướng thuận. Sử dụng với chuỗi ảnh tĩnh hoàn toàn. Hình 1.20 chỉ ra quá trình tạo ra ảnh dự báo trước:



Hình 1.20. Nén liên ảnh (Ảnh dự đoán trước)

➤ **Trong đó:**

+ Xác định véc tơ chuyển động có nhiệm vụ tính toán để tạo ra véc tơ chuyển động, sao cho ảnh hiện tại dự đoán một cách chính xác nhất.

+ Khởi trừ thực hiện cơ chế:

$$\text{Ảnh khác biệt} = \text{Ảnh hiện tại} - \text{Ảnh dự đoán}$$

(ảnh dự đoán được suy ra từ ảnh trước đó và véc tơ chuyển động).

+ Tạo ảnh trước đó (ảnh so sánh): là một vòng khép kín theo cơ chế:

$$\text{Ảnh trước đó} = \text{Ảnh khác biệt} - \text{Ảnh dự đoán}$$

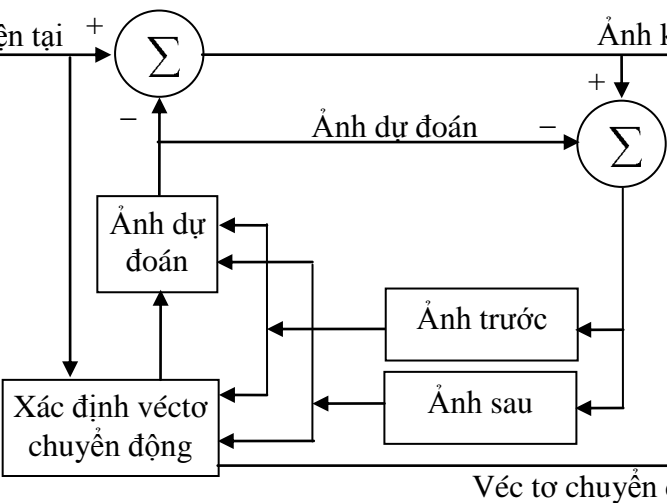
➤ **Nhận xét:**

– Dữ liệu truyền đi gồm 2 thành phần ảnh khác biệt và véc tơ chuyển động.

– Trong trường hợp véc tơ chuyển động xác định đúng thì ảnh khác biệt sẽ chứa rất ít thông tin và hiệu quả nén đạt cao nhất.

1.7.4. Ảnh dự đoán hai chiều (ảnh B)

Dự đoán hai chiều theo thời gian hay còn gọi là nội suy bù chuyển động. Sử dụng thông tin trong khung hình trước và một khung hình tham chiếu xuất hiện sau khung hình hiện tại để dự đoán khung hình hiện tại - gọi là khung hình B như hình 1.21



Hình 1.21. Nén liên ảnh (Ảnh dự đoán hai chiều)

↪ **Đặc điểm:**

- + Khả năng nén số liệu cao hơn so với dự đoán một chiều.
- + Hiệu ứng nhiễu trong khung hình hiện tại giảm bằng mức nhiễu trung bình của khung hình trước và sau nó.
- + Có khả năng dự đoán ngoài phạm vi tìm kiếm (nhờ các khung hình tiếp theo).

Sự khác biệt cơ bản của mạch tạo ảnh dự đoán hai chiều là ở bộ nhớ ảnh so sánh phải nhớ cả hai ảnh: ảnh trước và ảnh sau. Ảnh dự đoán hai chiều là kết quả nội suy giữa hai ảnh để xác định ảnh chuẩn của nó. Do đó, bộ mã hoá và giải mã phải đánh số các khung ảnh để xác định được ảnh trước và ảnh sau, phải dùng bộ nhớ lớn để lưu trữ các ảnh chuẩn.

Chương 2

CÔNG NGHỆ NÉN ẢNH H.264/ MPEG- 4AVC

2.1. GIỚI THIỆU CHUNG

Kể từ khi mới xuất hiện vào đầu những năm 90, chuẩn nén video MPEG-2 đã hoàn toàn thống lĩnh thế giới truyền thông. Cũng trong thập kỷ này, chuẩn nén MPEG-2 đã được cải tiến về nhiều mặt. Giờ đây nó có tốc độ bit thấp hơn và việc ứng dụng nó được mở rộng hơn nhờ có các kỹ thuật như đoán chuyển động, tiền xử lý, xử lý đối ngẫu và phân bổ tốc độ bit tùy theo tình huống thông qua ghép kênh thống kê.

Tuy nhiên, chuẩn nén MPEG-2 cũng không thể được phát triển một cách vô hạn định. Thực tế hiện nay cho thấy chuẩn nén này đã đạt đến hết giới hạn ứng dụng của mình trong lĩnh vực truyền truyền hình từ sản xuất tiền kỳ đến hậu kỳ và lưu trữ Video số. Bên cạnh đó, nhu cầu nén Video lại đang ngày một tăng cao kèm theo sự phát triển mạnh mẽ của mạng IP mà tiêu biểu là mạng Internet.

Chuẩn video MPEG-2 bị hạn chế bởi hai yếu tố trong định nghĩa ban đầu (original definition) của nó:

- Tốc độ bit nhắm tới của video được nén là khoảng 2–15 Mb/s (đối với main profile ở mail level). Tiêu chuẩn này không chứa giới hạn tốc độ bit thấp hơn bất kỳ vì điều này không được yêu cầu trong định nghĩa của bộ mã hóa tương thích. Hiển nhiên MPEG-2 cũng không hiệu quả với tốc độ bit thấp hơn.
- Silicon cho thực hiện MPEG-2 đã bị giới hạn đến trình độ công nghệ của những ngày đó. Điều này có nghĩa là trong năm 1994 ASIC (application Specific Integrated Circuit) được sử dụng trong thiết kế bộ giải mã với mật độ 120.000 gate/chip với kích thước gate 0.5 - 1 μ m. Trong khi đó công nghệ tiên tiến ngày nay đã đạt 25.000.000 gate/ASIC với kích thước gate nhỏ hơn 0.1 μ m.

Như vậy các kỹ thuật dựa trên MPEG-2 đã bị hạn chế trong việc thực hiện thực tế trong công nghệ ngày hôm nay.

Nhiều tiêu chuẩn mã hóa mới đã và đang nghiên cứu để khắc phục các nhược điểm của MPEG-2. Trong đó H.264/AVC, dự án tiêu chuẩn mã hóa video của nhóm chuyên gia mã hóa video của tổ chức ITU (ITU – T Video Coding Experts Group – VCEG) và nhóm chuyên gia ảnh động ISO/IEC (ISO/IEC Moving Picture Experts Group) là nổi bật hơn cả. Cái tên H.264/AVC chính là sự kết hợp tiêu chuẩn nghiên cứu riêng của từng nhóm: nhóm ITU là H.264, nhóm MPEG là AVC (Advanced Video Compression) tức MPEG-4 part 10. Các mục đích chính của việc nỗ lực tiêu chuẩn hóa này là phát triển một tiêu chuẩn mã hóa video đơn giản, với chỉ tiêu kỹ thuật nén tăng cường, và để có thể trình diễn video “thân thiện mạng” (network – friendly), nhắm tới các ứng dụng “hội thoại” (video telephone) và “không hội thoại” (lưu trữ, quảng bá hoặc streaming).

H.264/AVC đã thực hiện sự cải thiện đáng kể trong hiệu quả giảm tỷ lệ bit và méo ở tốc độ bit đó, cung cấp (một cách gần đúng) hệ số tiết kiệm tốc độ bit khoảng hai lần so với các tiêu chuẩn đang tồn tại như MPEG-2 video.

2.2. CÁC PHIÊN BẢN H.26X

Theo ITU-T, các tiêu chuẩn mã hoá Video được coi là các khuyến nghị gọi tắt là chuẩn H.26x (H.261, H.262, H.263 và H.264). Với tiêu chuẩn ISO/IEC, chúng được gọi là MPEG-x (như MPEG-1, MPEG-2 và MPEG-4).

2.2.1. H.261

2.2.1.1. Mục đích:

- H.261 được phát triển cho dịch vụ truyền hình hội nghị video phone qua ISDN ở tốc độ thấp 64kbps ($p=1..30$)
- Truyền hình hội nghị theo yêu cầu với chất lượng ảnh cao hơn, $p \geq 6$, tốc độ ≥ 384 kbps.

- Là cơ sở của chuẩn sau này như MPEG 1,2
- Các đặc tính:
 - + Trễ mã hóa < 150ms, → Truyền hình hội nghị song công, gây ấn tượng tốt cho khán giả.
 - + Thực hiện trên linh kiện VLSI (giá thành thấp) → Mở rộng thị trường ở các dịch vụ video phone, truyền hình hội nghị.

2.2.1.2. Các dạng ảnh đầu vào:

- Khả năng phối hợp giữa các chuẩn 625 và 525 dòng của TV, H.261 sử dụng dạng thức trung gian chung CIF (Common Intermediate Format) với các tốc độ bit thấp hơn. H.261 sử dụng dạng có tốc độ nhỏ hơn $\frac{1}{4}$ là QCIF (Quadrature).
- Với tốc độ 30 khung hình/s thì tốc độ dữ liệu của CIF là 37,3Mbps, QCIF là 9,35 Mbps. Tốc độ càng thấp thì càng giảm số khung hình/s.

2.2.2 H.263

2.2.2.1. Mục đích:

- Tiêu chuẩn cải tiến H.261 cho video tốc độ thấp, có thể truyền trên mạng điện thoại công cộng PSTN, được công nhận năm 1996.
- Giống như H.261, mã hóa DCT cho các MB trong I Frame và DCT sai biệt dự đoán trong P Frame.
- Tốc độ tối thiểu.

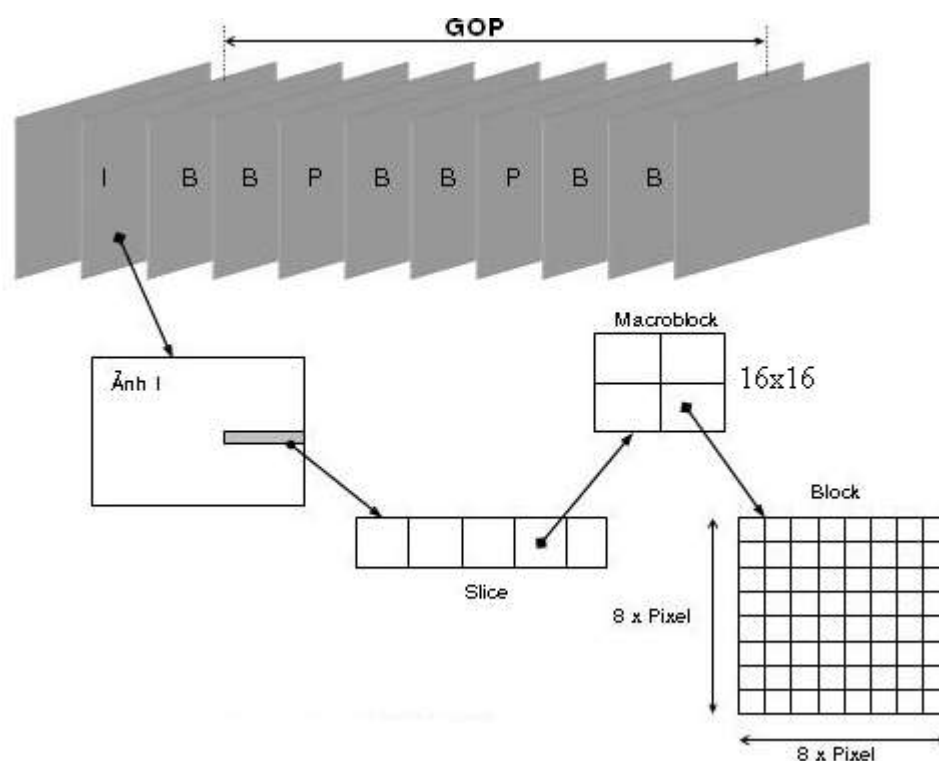
2.2.2.2. Ưu điểm:

- Chính xác sai biệt dự đoán tới $\frac{1}{2}$ pixels
- Không hạn chế vector chuyển động.
- Mã hóa số học theo cú pháp.
- Dự đoán thuận lợi với các khung P
- Ngoài CIF, QCIF, H.263 còn hỗ trợ SQCIF, 4 CIF và 16 CIF với độ phân giải tín hiệu chói tuần tự là 128x96, 704x576, 704x576, 1408x1152. Độ phân giải tín hiệu sắc bằng $\frac{1}{4}$ tín hiệu chói.

2.3. CẤU TRÚC DỮ LIỆU VÀ CÁC KIỂU NÉN

2.3.1. Chia ảnh thành các macroblock

- Mỗi ảnh video, frame hoặc field, được chia thành các macroblock (MB) có kích thước cố định bao trùm một diện tích ảnh hình chữ nhật gồm 16 x 16 mẫu thành phần luma và 8 x 8 mẫu cho mỗi một trong hai thành phần chroma.



Hình 2.1. Cấu trúc dòng video h.264

- Tất cả các mẫu macroblock luma hoặc chroma được dự đoán theo không gian hoặc thời gian, và dự đoán tại chỗ hợp thành được truyền đi nhờ dùng mã chuyển vị.

- Các macroblock được tổ chức thành các slice, biểu diễn các tập con của ảnh đã cho và có thể được giải mã độc lập. Thứ tự truyền các macroblock trong

dòng bit phụ thuộc vào bản đồ phân phối Macroblock (Macroblock Allocation Map) và không nhất thiết phải theo thứ tự quét.

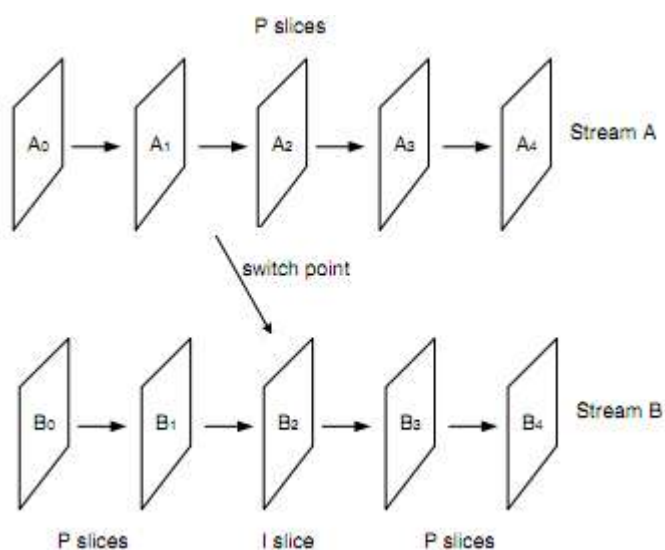
- H.264 / AVC hỗ trợ năm dạng mã hóa slice khác nhau:

- + I (Intra): Là slice đơn giản nhất, trong đó tất cả macroblock được mã hóa không có tham chiếu tới các ảnh khác trong dãy video.

- + P (Predicted): Các macroblock được mã hóa tham chiếu tới các ảnh trước nó

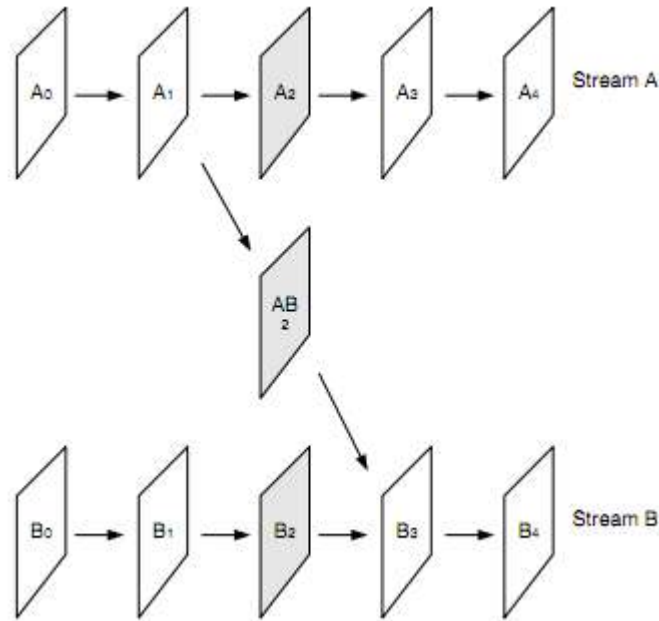
- + B (Bi- Predictive): Các macroblock được mã hóa tham chiếu tới cả ảnh trước lẫn ảnh sau.

- + SP (switching P) và SI (switching I), được xác định cho chuyển mạch hiệu quả giữa các dòng bit được mã hóa ở các tốc độ bit khác nhau.



Hình 2.2. Chuyển mạch cho slice P

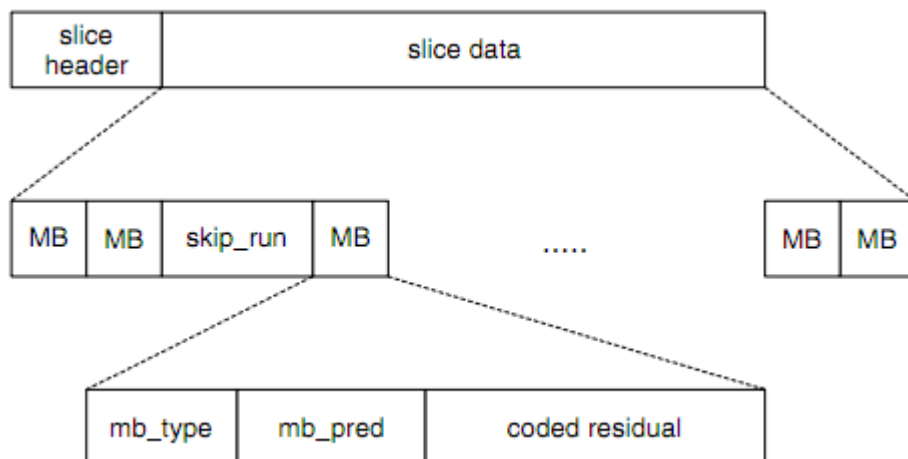
- Các tín hiệu dự đoán Inter của các dòng bit cho một frame SP được chọn được lượng tử hóa trong miền biến đổi, đưa chúng vào dải biên độ thô hơn để cho phép mã hóa tốc độ bit thấp tín hiệu khác nhau giữa các dòng bit.



Hình 2.3. Chuyển mạch cho slice P

- Các frame SI được xác định để thực hiện sự thích nghi hoàn thiện cho các frame SP trong trường hợp mà dự đoán Inter không thể được sử dụng do các lỗi truyền dẫn.

2.3.2. Cấu trúc các chuỗi bit:



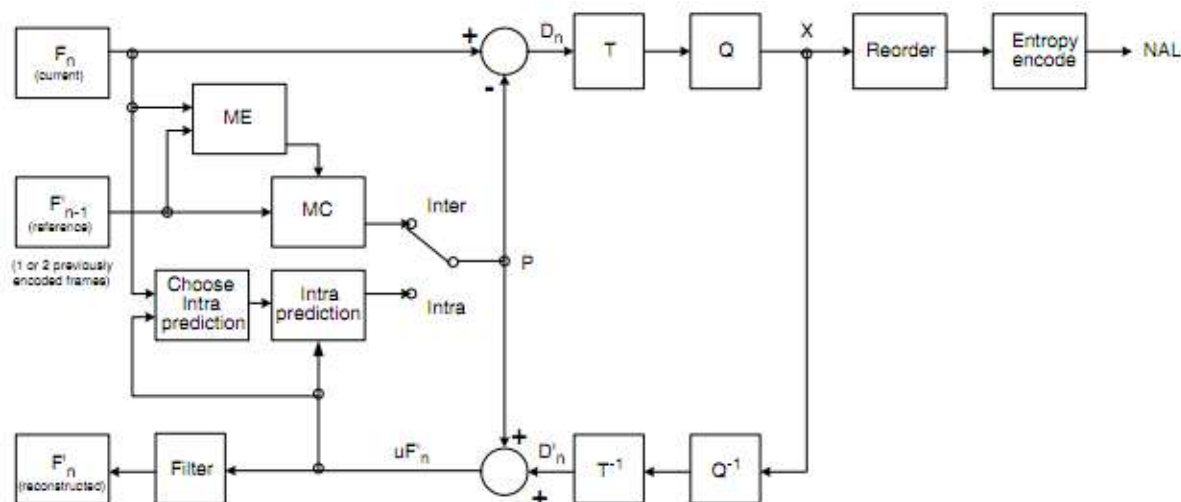
Hình 2.4. Cấu trúc dòng bit H.264

Trong đó:

- + **mb_type**: Xác định xem macroblock được mã hóa trong ảnh hay liên ảnh (P hoặc B); Xác định kích thước phân vùng macroblock
- + **mb_pred**: Xác định phương thức dự báo trong ảnh (macroblocks trong ảnh) và xác định các list 0 / hoặc list 1 để mã hóa phần khác biệt của ảnh chuyển động so với ảnh tham chiếu. Vector cho mỗi phân vùng macroblock (Các macroblocks liên ảnh, trừ MBs liên ảnh với kích thước 8x8 pixel).
- + **sub_mb_pred**: (Liên ảnh MBs với kích thước 8x8) Xác định kích thước các sub-macroblock cho mỗi phân vùng macroblock-sub, các list 0 hoặc list 1 của ảnh tham chiếu cho mỗi phân vùng macroblock; mã hóa sự khác biệt của vector chuyển động cho mỗi phân vùng sub-macroblock.
- + **coded_block_pattern**: Xác định được khối 8x8 (Luma và sắc) chứa vectors chuyển động được mã hóa cho mỗi phân vùng sub- macroblock.
- + **mb_qp_delta**: Thay đổi các tham số lượng tử.
- + **residual**: mã hóa chuyển đổi hệ số tương ứng với các mẫu ảnh còn sót lại sau khi dự đoán.

2.4. MÃ HÓA H.264

2.4.1. Sơ đồ khối mã hóa H.264:



Hình 2.5. Mã hóa H.264

Trong đó:

- F_n (current): Ảnh hiện tại
- F'_{n-1} (reference): Ảnh tham chiếu của 1 hoặc 2 khung hình được mã hóa trước đó.
- F'_n (reconstoncted): Ảnh khôi phục
- Intra prediction: Dự đoán trong ảnh
- Choose Intra Prediction: Chọn dự đoán trong ảnh
- Deblocking filter: Bộ lọc tách khối
- T , Q và T^{-1} , Q^{-1} : mã hóa separable integer transform-biến đổi nguyên tách biệt , lượng tử hóa và giải mã T , giải lượng tử hóa tương ứng.
- D_n , D'_n : Phần ảnh sai khác vào mã hóa và ra từ giải mã
- Reorder: Sắp xếp lại
- Entropy encode: Mã hóa entropy
- MC(Motion Compensation): Bù chuyển động.
- ME(Motion estimation): Đánh giá chuyển động

2.4.2. Định dạng mã hóa dữ liệu:

Thiết kế H. 264/ AVC hỗ trợ mã hóa video (trong định dạng 4:2:0) có chứa các frame liên tục hoặc xen kẽ, hoặc có thể được trộn với nhau trong cùng một dãy. Nhìn chung một frame video chứa hai field đan xen, được cách ly về mặt thời gian bởi một chu kỳ field (nửa thời gian của chu kỳ frame), có thể được mã hóa riêng biệt như hai ảnh field, hoặc cùng với nhau như một ảnh frame. Một frame liên tục phải luôn luôn được mã hóa như một ảnh frame đơn; tuy nhiên nó vẫn được xem như gồm hai field ở cùng một thời điểm.

2.4.2.1. Lớp trừu tượng mạng (NAL – Network Abstraction Layer)

Lớp trừu tượng mạng (NAL) được xác định để định dạng dữ liệu này và cung cấp thông tin header trong cách thích hợp cho việc chuyển chở bởi các lớp truyền tải hoặc môi trường lưu trữ. Tất cả dữ liệu được chứa trong các khối NAL, mỗi khối chứa một số nguyên byte. Một khối NAL xác định định dạng chung cho việc sử dụng trong cả hệ thống định hướng gói (packet – oriented) và hệ thống định hướng dòng bit (bitstream). Định dạng của các khối NAL là đồng nhất cho cả việc phân phối dòng truyền tải định hướng gói và định hướng dòng bit, ngoại trừ rằng mỗi khối NAL trong lớp truyền tải định hướng dòng bit có thể có một tiền tố mã hóa khởi hành ở trước.

2.4.2.2. Lớp mã hóa video (Video Coding Layer)

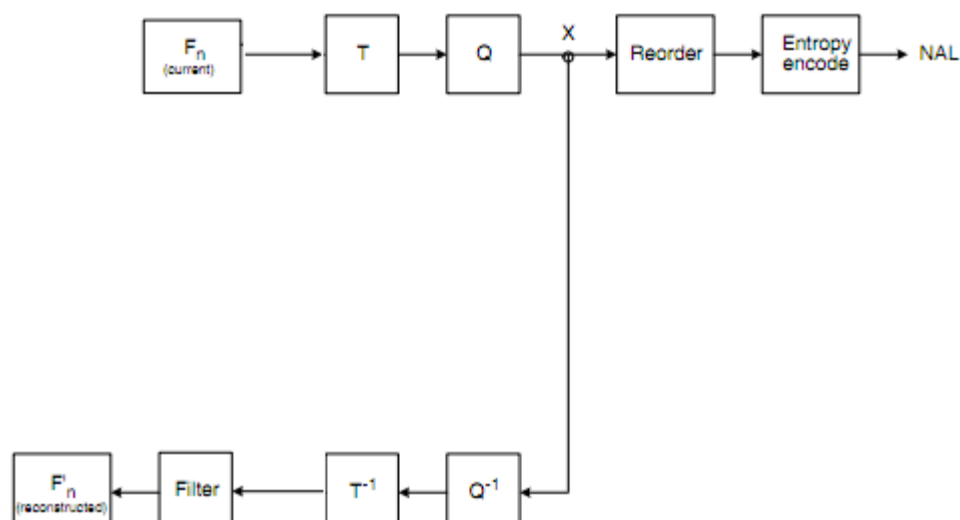
Lớp mã hóa video (VCL-Video Coding Layer), được định rõ để biểu diễn hiệu quả nội dung của dữ liệu video. Lớp mã hóa video của H.264/AVC thì tương tự với các tiêu chuẩn khác như MPEG-2 video. Nó là sự kết hợp dự đoán theo thời gian và theo không gian, và với mã chuyển vị.

Ảnh được tách thành các khối. Ảnh đầu tiên của dãy hoặc điểm truy nhập ngẫu nhiên thì được mã hóa (trong khối) “Intra”, có nghĩa là không dùng thông tin nào ngoài thông tin chứa trong bản thân ảnh. Mỗi mẫu của một khối trong một frame Intra được dự đoán nhờ dùng các mẫu không gian bên cạnh của các khối đã mã hóa trước đó. Đối với tất cả các ảnh còn lại của dãy hoặc giữa các điểm truy cập ngẫu nhiên, mã hóa “Inter” được sử dụng, dùng dự đoán bù chuyển động từ các ảnh được mã hóa trước.

2.5. KỸ THUẬT NÉN TRONG ẢNH

2.5.1. Sơ đồ mã hóa trong ảnh:

Intra Frame → Ảnh I thực hiện quá trình mã hóa trong ảnh cũng tương tự mã hóa MPEG-2, nhưng cũng có những điểm khác như trình bày ở các phần sau.



Hình 2.6. Mã hóa trong ảnh H.264

Trong đó:

- F'_n (reconstoncted): Ảnh khôi phục sẽ được lưu trong Memory để tìm vector chuyển động cho Frame P tiếp theo.

2.5.2. Chọn macroblock và thứ tự trong công đoạn nén:

Chuẩn nén MPEG-4 AVC có hai cải tiến mới trong lĩnh vực nén không gian.

- Trước hết, bộ lập mã này có thể tiến hành nén không gian tại các macroblock 16x16 điểm ảnh thay vì các block 8x8 như trước đây. Điều này giúp tăng cường đáng kể khả năng nén không gian đối với các hình ảnh có chứa nhiều khoảng lớn các điểm ảnh giống nhau.

- Thứ hai là thao tác nén được tiến hành trong miền không gian trước khi công đoạn DCT diễn ra. Chuẩn nén MPEG-4 AVC so sánh macroblock hiện thời với các macroblock kế bên trong cùng một khung, tính toán độ chênh lệch, và sau đó sẽ chỉ gửi đoạn chênh lệch tới DCT. Hoặc là nó có thể chia nhỏ macroblock 16x16 điểm ảnh thành các khối 4x4 nhỏ hơn và so sánh từng khối này với các khối kế bên trong cùng một macroblock. Điều này giúp cải thiện khả năng nén ảnh chi tiết.

2.5.3. Chuyển vị, co dẫn và lượng tử hóa:

Tương tự với các tiêu chuẩn mã hóa video trước đó, H.264/AVC cũng dùng mã chuyển vị cho dự đoán tiếp theo. Tuy nhiên trong H.264/AVC việc chuyển vị được áp dụng cho các khối 4×4 , và thay cho biến đổi cosin rời rạc (DCT) 4×4

a/ Biến đổi nguyên tách biệt (*separable integer transform*)

→ được sử dụng, với các tính chất giống như 4×4 DCT về cơ bản. Vì biến đổi ngược được xác định bởi các thuật toán nguyên chính xác nên tránh được sự không thích nghi của biến đổi ngược. Một biến đổi 2×2 bổ sung được áp dụng cho bốn hệ số DC của mỗi thành phần chroma. Nếu một macroblock được mã hóa trong mode Intra - 16×16 , thì một biến đổi 4×4 tương tự được thực hiện cho các hệ số DC 4×4 của tín hiệu luma.

Ví dụ: Đầu vào là khối luma 4×4 . Thực hiện biến đổi T

Đầu vào X

	$j = 0$	1	2	3
$i = 0$	5	11	8	10
1	9	8	4	12
2	1	10	11	4
3	19	6	15	7

Ra chuyển đổi W

	$j = 0$	1	2	3
$i = 0$	140	-1	-6	7
1	-19	-39	7	-92
2	22	17	8	31
3	-27	-32	-59	-21

b/ Lượng tử hóa các hệ số biến đổi

Để lượng tử hóa các hệ số biến đổi, H.264/AVC dùng lượng tử hóa vô hướng. Một trong số 52 bộ lượng tử hóa được chọn cho mỗi macroblock bởi tham số lượng tử hóa QP (Quantization Parameter).

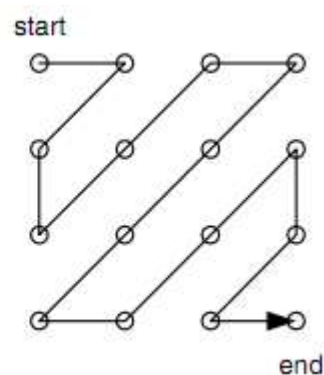
→ Lượng tử hóa thay đổi theo khoảng cách để đạt được hiệu quả nén cao theo biểu thức sau:

$$Z = W \text{ round } [PF / Qstep]$$

Sau lượng tử là Z

	$j = 0$	1	2	3
$i = 0$	17	0	-1	0
1	-1	-2	0	-5
2	3	1	1	2
3	-2	-1	-5	-1

Quét zic-zắc



Các bộ lượng tử hóa được sắp xếp sao cho có sự tăng khoảng 12.5% trong kích thước bước lượng tử hóa khi QP tăng một đơn vị.

c/ Quét zic-zắc

Quét zic-zac bắt đầu từ hệ số DC nhằm ánh xạ ma trận 4x4 thành vector 1x 16, đây vector là giá trị EOB (End of Block) → Nhóm các thành phần tần số thấp vào đỉnh vector.

2.5.4. Sắp xếp lại:

Sắp xếp lại khối 1x 16 sau quét zic- zac ở trên → Thành phần DC ở đỉnh vector 1x16, tiếp đến theo thứ tự lần lượt là 15 thành phần còn lại AC.

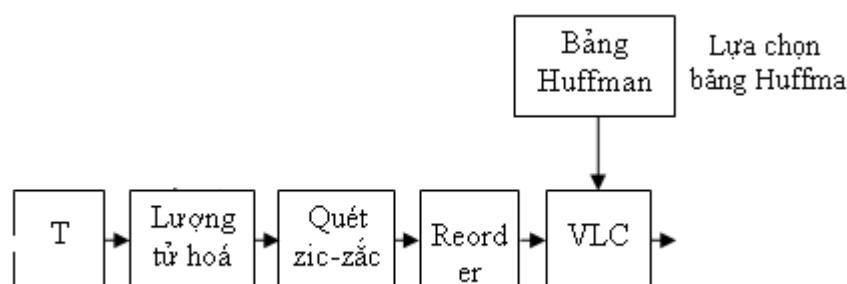
2.5.5.Mã hóa Entropy:

Trong H.264/AVC, hai phương pháp mã hóa entropy được hỗ trợ :

- Phương pháp mã hóa entropy mặc định (default) sử dụng một tập từ mã mở rộng vô hạn đơn cho tất cả phần tử cấu trúc, ngoại trừ các hệ số biến đổi được lượng tử hóa. Như vậy, thay cho việc một thiết kế bảng VCL khác nhau cho mỗi phần tử cấu trúc, chỉ có một ánh xạ cho một bảng từ mã đơn được thực

hiện theo thống kê dữ liệu. Bảng từ mã đơn được chọn là mã exp-Golomb với các tính chất giải mã rất đơn giản và hài hòa.

- Để truyền các hệ số biến đổi được lượng tử hóa, một phương pháp tinh tế hơn gọi là mã độ dài biến đổi thích nghi hoàn cảnh CAVLC (Context – Adaptive Variable Length Coding) được sử dụng.



Hình 2.7. Mã độ dài biến đổi thích nghi hoàn cảnh

Trong sơ đồ này, các bảng VLC được chuyển mạch tốt cho các phần tử cấu trúc khác nhau, phụ thuộc vào phần tử cấu trúc đã được truyền. Vì các bảng VLC được thiết kế tốt để thích nghi các thống kê có điều kiện tương ứng, chỉ tiêu kỹ thuật mã hóa entropy được cải thiện so với các sơ đồ dùng một bảng VLC đơn.

- Hiệu quả của mã hóa entropy có thể được cải thiện hơn nữa nếu mã hóa số học nhị phân thích nghi hoàn cảnh CABAC (Context – Adaptive Binary Arithmetic Coding) được sử dụng. Một mặt, việc dùng mã hóa số học cho phép ấn định một số không nguyên vẹn (non – integer) các bit cho mỗi symbol của bảng chữ cái và điều này là cực kỳ có ích đối với các xác suất symbol lớn hơn 0.5. Mặt khác, việc dùng mã thích nghi cho phép thích nghi với các thống kê symbol không tĩnh. Một tính chất quan trọng khác của CABAC là mô hình hóa hoàn cảnh của nó. Tính thống kê của các phần tử cấu trúc đã được mã hóa thì được sử dụng để đánh giá các xác suất điều kiện. Các xác suất điều kiện này được sử dụng để chuyển mạch hàng loạt model xác suất đã được đánh giá. Trong H.264/AVC, engine cốt lõi mã hóa số học và đánh giá xác suất kèm theo nó được xác định như các phương pháp ít phức tạp, không dùng phép nhân mà chỉ

dùng các phép dịch chuyển (shift) và các bảng look – up. So sánh với CAVLC, CABAC thường cho giảm tốc độ bit khoảng 10 – 15% khi mã hóa tín hiệu truyền hình có cùng chất lượng.

2.5.6. Bộ lọc tách khối trong vòng (In – loop deblocking filter)

Một đặc trưng riêng của mã hóa dựa trên cơ sở khối là có thể nhìn thấy các cấu trúc khối. Các mép khối được cấu trúc lại với độ chính xác kém hơn các pixel bên trong và nhìn chung “dạng khối” (blocking) được xem là một trong những nhiễu (artifact) dễ thấy nhất với các phương pháp nén hiện tại.



Hình 2.8. Tác dụng của bộ lọc tách khối đối với ảnh được nén nhiều

Trái là không có bộ lọc tách khối. Phải có bộ lọc tách khối

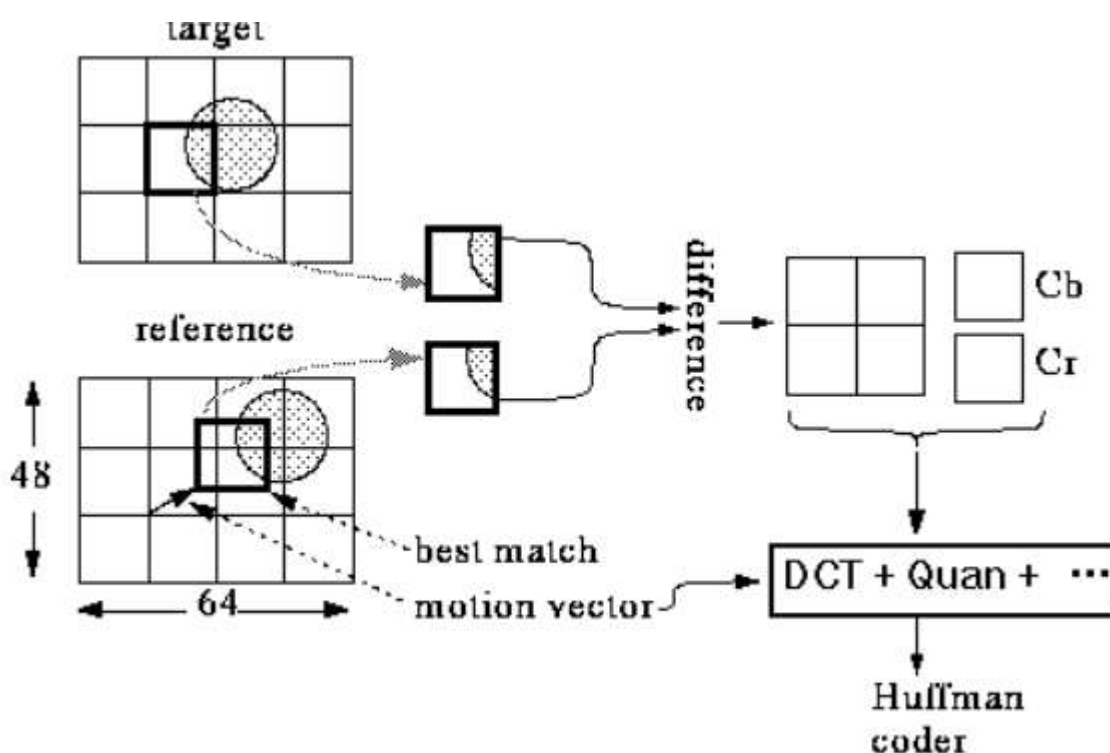
Do nguyên nhân này mà H.264/AVC xác định bộ lọc tách khối thích nghi trong vòng, ở đó cường độ lọc được điều khiển bởi giá trị của nhiễu phân tử cấu trúc.

Tính khối bị giảm mà không ảnh hưởng nhiều tới độ sắc nét của nội dung. Hậu quả là chất lượng chủ quan được cải thiện đáng kể. Đồng thời bộ lọc giảm tốc độ bit khoảng 10 – 15% khi tạo ra cùng chất lượng ảnh chủ quan với video không lọc.

2.6. KỸ THUẬT NÉN LIÊN ẢNH

2.6.1. Đánh giá chuyển động (ME-Motion estimation):

Hình ảnh được chia thành các khối (macroblock). Để xác định tọa độ mới từ ảnh n sang ảnh $n+1$ bằng cách so sánh tương quan giữa các mẫu trong macroblock. (hình 2.9)



Hình 2.9. Nguyên lý đánh giá chuyển động

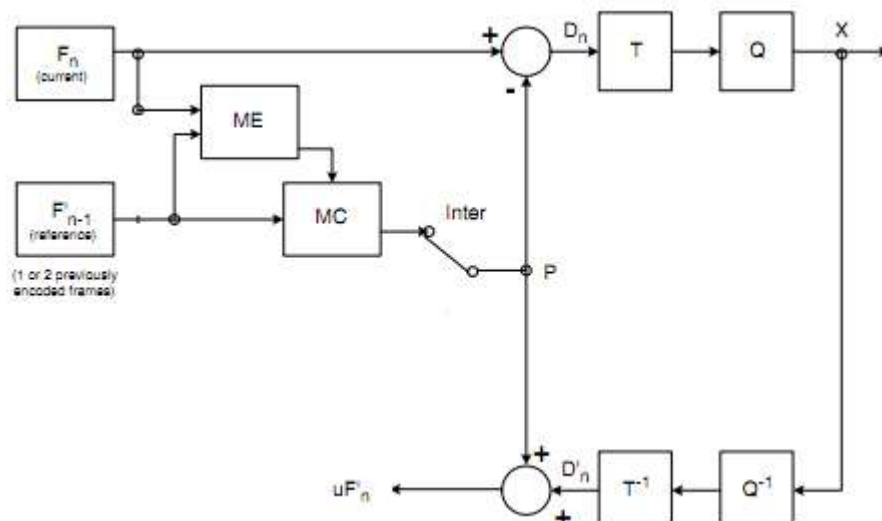
Giá trị so sánh thường là 1/2 hoặc lên tới 1/4 hay 1/8 pixel.

→ Kết quả cho ra một vector gọi là vector chuyển động

2.6.2. Nguyên lý bù chuyển động: (hình 2.10)

- Ảnh hiện tại F_n được đưa đến bộ trừ (Subtract) và ME.
- Ảnh tham chiếu F_{n-1} của 1 hoặc 2 khung hình được mã hóa trước đó (Lấy ra từ bộ nhớ ảnh) cũng đồng thời được đưa vào ME và MC (Bù chuyển động)

- F_n và F'_{n-1} qua ME sẽ xác định ra vector chuyển động để đưa vào MC
- Vector chuyển động cùng với F'_{n-1} qua MC sẽ xác định ra phần khác biệt của ảnh (hình 2.10)



Hình 2.10. Nguyên lý bù chuyển động

- Phần khác biệt từ MC và ảnh F_n vào bộ trừ sẽ đưa ra thông tin khác biệt giữa 2 ảnh D_n → D_n sẽ được thực hiện theo cơ chế nén trong ảnh.

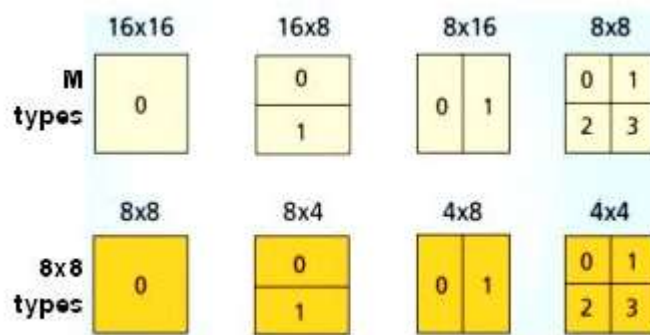
2.6.3. Bù chuyển động trong các slice P:

Macroblock P được mã hóa tham chiếu tới các ảnh trước nó. Các macroblock P được phân chia linh hoạt thành các khối thích hợp khi thực hiện cơ chế bù chuyển động.

a/ Phân chia các macroblock

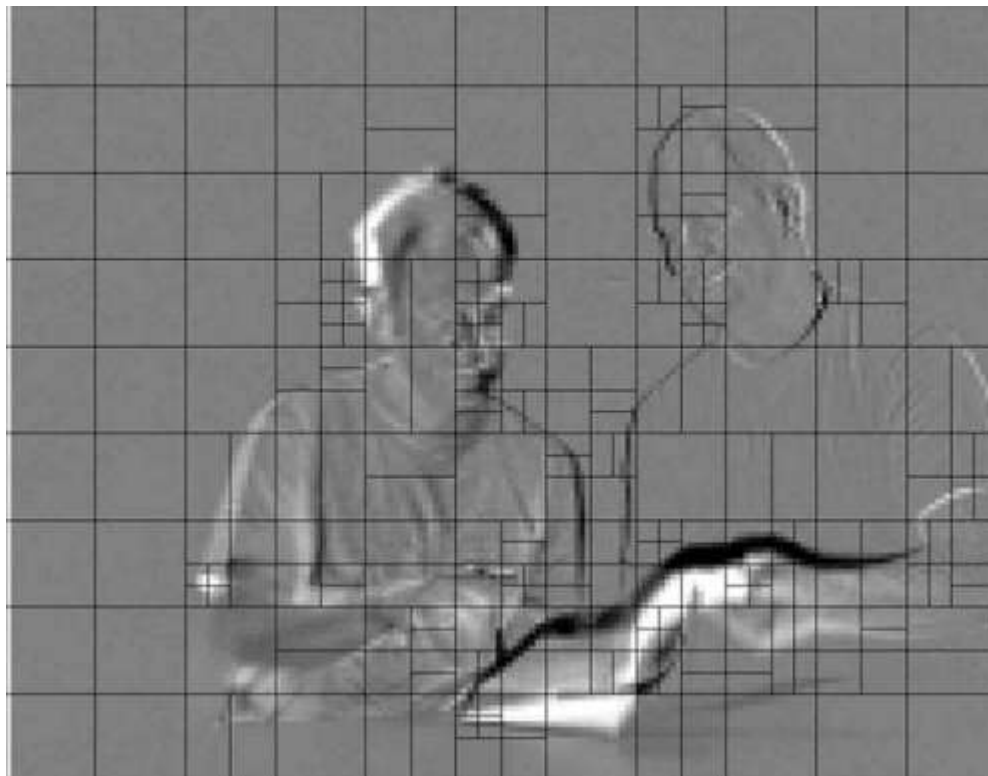
H.264 có thể chia thành phần chói của từng macroblock thành 4 cỡ:

- + Mẫu 16x16, 16x8, 8x16 hoặc 8x8.
- + Khi sử dụng khối 8x8, nó còn có thể chia tiếp 4 khối 8x8 này thành 4 cỡ nữa
là 8x8, 8x4, 4x8 hoặc 4x4.



Hình 2.11. Phân chia macroblock cho bù chuyển động

Việc phân chia các macroblock cho phép bộ lập mã xử lý được một vài loại chuyển động tùy theo độ phức tạp của chuyển động đó cũng như nguồn lực về tốc độ bit.



Hình 2.12. Sự lựa chọn các kích thước block khi bù chuyển động

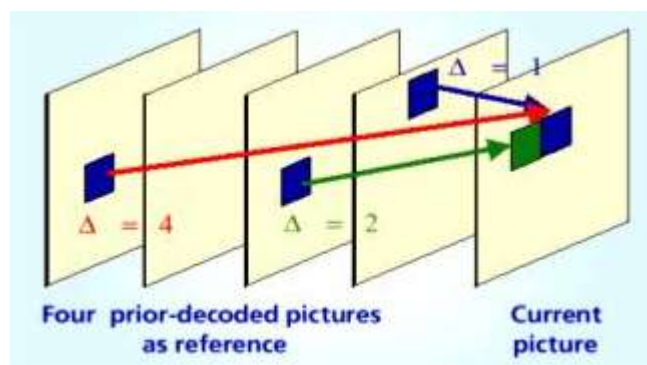
Nhìn chung, kích cỡ phân chia lớn phù hợp với việc xử lý chuyển động tại các khu vực giống nhau trong ảnh, trong khi đó kích cỡ phân chia nhỏ lại rất có

ích khi xử lý chuyển động tại các chỗ có nhiều chi tiết hơn. Kết quả là chất lượng hình ảnh cao hơn, ít bị vỡ khối hơn.

Các cuộc thử nghiệm đã chỉ ra rằng việc sắp xếp hợp lý các khung có thể tăng tỷ lệ nén thêm 15%.

b/ Chọn nhiều hình tham chiếu

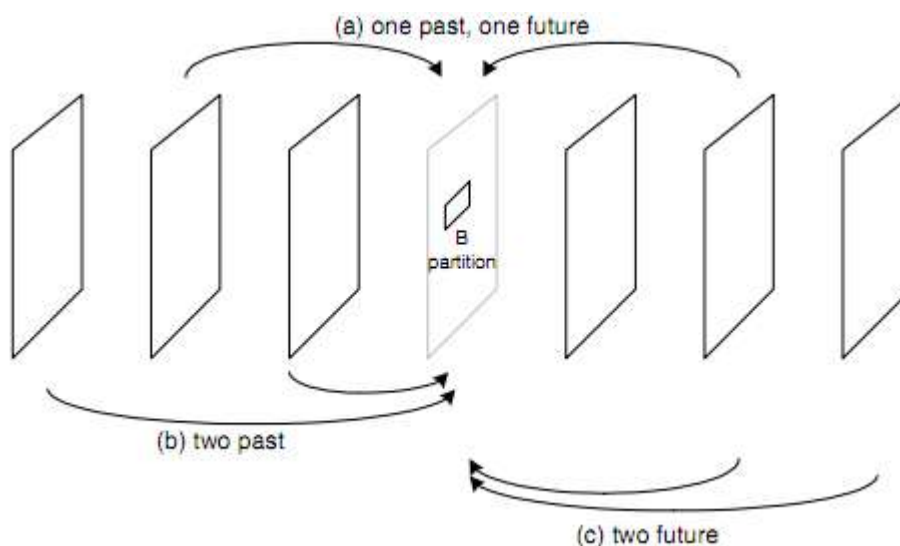
Chuẩn nén MPEG-2 chỉ dựa trên 2 khung tham chiếu để dự đoán các chuyển động mang tính chu kỳ. Tuy nhiên, khi camera thay đổi góc quay hay chuyển qua chuyển lại giữa các cảnh, việc chỉ sử dụng 2 khung tham chiếu không còn phù hợp để dự đoán chính xác chuyển động. Tương tự như vậy, để đoán trước các chuyển động phức tạp như sóng biển hay một vụ nổ, ta cần phải có nhiều hơn 2 khung tham chiếu. Vì thế, chuẩn MPEG-4 AVC cho phép có tới 5 khung tham chiếu phục vụ cho việc mã hoá giữa khung. Kết quả là chất lượng video tốt hơn và hiệu suất nén cao hơn.



Hình 2.13. Bù chuyển động nhiều Frame. Ngoài vector chuyển động, các tham số tham chiếu ảnh (Δ) cũng được truyền đi

Ngoài các mode macroblock bù chuyển động được mô tả ở trên, macroblock P – slice cũng có thể được mã hóa trong mode được gọi là SKIP. Đối với mode này cả tín hiệu lỗi dự đoán lượng tử hóa lẫn vector chuyển động hoặc tham số chỉ số tham chiếu đều không được truyền đi.

2.6.4. Bù chuyển động trong các slice B: (Hình 2.14)



Hình 2.14. Bù chuyển động trong slice B:

(a) trong quá khứ, tương lai, (b) quá khứ, (c) trong tương lai

Các slice B được mã hóa theo cách trong đó một số macroblock hoặc block có thể dùng trung bình trọng số của hai giá trị dự đoán bù chuyển động riêng biệt.

Trong các slice B, bốn dạng dự đoán ảnh inter khác nhau được hỗ trợ: dự đoán list 0, list 1, hai hướng (bi-predictive) và trực tiếp (direct).

+ Dự đoán list 0 → hiển thị tín hiệu dự đoán được tạo thành nhờ dùng bù chuyển động từ ảnh của bộ nhớ đệm ảnh tham chiếu đầu tiên

+ Dự đoán list 1 → ảnh tham chiếu thứ hai được sử dụng cho việc xây dựng tín hiệu dự đoán.

+ Trong mode dự đoán hai hướng, → tín hiệu dự đoán được tạo thành bởi trung bình trọng số của tín hiệu dự đoán bù chuyển động list 0 và list 1.

+ Mode dự đoán trực tiếp → được quy nạp từ các phần tử cấu trúc được truyền đi trước đó và có thể là dự đoán list 0 hoặc list 1, hoặc dự đoán hai hướng.

Các slice B dùng sự phân nhỏ macroblock tương tự với các slice P. Ngoài các mode Inter – 16 x 16, Inter – 16 x 8, Inter – 8 x 16, Inter – 8 x 8 và Intra, còn có dạng macroblock dùng dự đoán trực tiếp, có nghĩa là mode trực tiếp. Ngoài

ra, đối với mỗi sự phân chia 16×16 , 16×8 , 8×16 và 8×8 , phương pháp dự đoán (list 0, list 1, hai hướng) có thể được chọn lựa riêng biệt. Sự phân chia 8×8 của macroblock slice B cũng có thể được mã hóa trong mode trực tiếp. Nếu không có tín hiệu lỗi dự đoán được truyền đi cho mode macroblock trực tiếp, nó cũng được quy nạp đến mode SKIP slice B và có thể được mã hóa rất hiệu quả, tương tự mode SKIP trong các slice P. Việc mã hóa vector chuyển động là tương tự như đối với slice P với những biến đổi thích hợp vì các khối bên cạnh có thể được mã hóa khi dùng các mode dự đoán khác nhau.

2.7. SÁNH HIỆU QUẢ MÃ HÓA CỦA H.264/AVC VỚI CÁC TIÊU CHUẨN MÃ HÓA TRƯỚC ĐÓ.

2.7.1 Những ưu điểm kỹ thuật chính của H.264/AVC so với MPEG-2 Video:

Trong phần này chỉ so sánh những ưu điểm kỹ thuật chính của H.264/AVC so với MPEG-2 video, chuẩn nén đang được dùng phổ biến hiện nay. Từ phân tích đặc điểm kỹ thuật tiêu chuẩn H.264/AVC, ta thấy có một số công cụ được sử dụng, làm tăng hiệu quả nén của nó:

- Dự đoán hệ số AC/DC mà ở đó các hệ số macroblock được dự đoán từ các hệ số trong các macroblock bên cạnh và từ một số dữ liệu trong bản thân macroblock. Tất nhiên việc giải mã macroblock phải thay đổi khác với thuật toán dòng bit đơn giản trước đó để đưa vào một số tính toán cho mỗi hệ số.
- Các công cụ toán học mới như CAVLC và CABAC cho sự cải thiện đáng kể trong mã hóa entropy.
- Trong khi bù chuyển động trong MPEG-2 Part 2 được hạn chế đến nội suy hai chiều $\frac{1}{2}$ pixel thì H.264/AVC cho phép các vector chuyển động đến chính xác $\frac{1}{4}$ pixel và sau đó dùng nội suy nhiều chiều (bi-cubic). Nội suy nhiều chiều tạo ra sự thích hợp hơn cho macroblock, do vậy giảm năng lượng được lưu trong

ảnh lỗi, làm giảm số bit cần phải mã hóa. Tuy nhiên nội suy nhiều chiều đòi hỏi nhiều thuật toán hơn, sự phức tạp thực hiện bù chuyển động là cao hơn.

- Việc dùng bộ lọc tách khối ở cả phía bộ mã hóa và bộ giải mã làm giảm sự không liên tục ở các biên của khối do các hệ số chất lượng khác nhau được sử dụng cho các khối cạnh nhau gây ra. Công cụ này có ảnh hưởng đáng kể đến tính phức tạp của bộ mã hóa và bộ giải mã.

- Kích thước khối được thay đổi từ 16×16 xuống 4×4 . Kích thước khối giảm cung cấp độ lợi mã hóa mà không có sự tăng đáng kể tính phức tạp.

2.7.2. So sánh hiệu quả mã hóa của H.264/AVC đối với các tiêu chuẩn mã hóa trước đó

Từ các ưu việt kỹ thuật trên, các nhà nghiên cứu phát triển H.264/AVC cho rằng sự tiết kiệm tốc độ bit trung bình của H.264/AVC so với MPEG-2 là khoảng 65%. Điều này có nghĩa là tốc độ bit cho video truyền hình hoặc HD (ở mức chất lượng quảng bá và DVD) được giảm theo hệ số 2.25 hoặc 2.5 khi dùng mã hóa H.264/AVC so với MPEG-2. Tất nhiên mức tiết kiệm này còn phụ thuộc vào sản phẩm của từng hãng. Ví dụ hãng iVAST, California, USA đã hiển thị video chất lượng quảng bá ở tốc độ bit không đổi (CBR – Constant Bit Rate) 1.5 Mb/s so với MPEG-2 tốc độ bit thay đổi (VBR – Variable Bit Rate) 3.0 Mb/s. Trong khi đó vào tháng 09/2002, tại IBC ở Amsterdam, Hà Lan, công ty VideoLocus đã trình diễn codec H.264/AVC được tối ưu hóa cao của riêng họ, demo dòng video chất lượng DVD ở tốc độ 1 Mb/s và so sánh với dòng truyền video MPEG-2 tốc độ 5 Mb/s ở ngay bên cạnh.

Chương 3

ỨNG DỤNG CÔNG NGHỆ NÉN ẢNH H.264/MPEG-4 AVC

H.264/MPEG-4 AVC đánh dấu một bước ngoặt trong lĩnh vực nén video, áp dụng các kỹ thuật tiên tiến nhằm mục đích sử dụng băng thông hiệu quả hơn và đem lại chất lượng ảnh cao hơn. Điều này có nghĩa là MPEG-4 AVC là một ứng cử viên nặng ký có khả năng thay thế MPEG-2 trong những năm sắp tới.

3.1. CÁC PHẠM VI ỨNG DỤNG VÀ CÁC MODEL KINH DOANH MỚI.

Hiệu quả nén tăng của H.264/AVC tạo ra các phạm vi ứng dụng và các cơ hội kinh doanh mới. Ví dụ: bây giờ đã có thể truyền tín hiệu video chất lượng TV (ví dụ PAL) ở tốc độ 1 Mb/s, cho phép streaming qua các đường ADSL.

Một phạm vi kinh doanh đáng quan tâm khác là truyền dẫn truyền hình qua vệ tinh. Bằng việc chọn điều chế 8-PSK và mã hóa Turbo (hiện đang được nghiên cứu cho tiêu chuẩn DVB-S2) và dùng H.264/AVC, số chương trình qua vệ tinh có thể tăng gấp ba so với các hệ thống DVB-S hiện tại dùng MPEG-2. Nếu điều này xảy ra tất sẽ dẫn đến cuộc cách mạng trong việc thay đổi các máy thu DVB-S hiện có, cũng có nghĩa là đẩy mạnh công nghiệp sản xuất máy thu hình vệ tinh số.

Đối với DVB-T, H.264/AVC là một sự lựa chọn thú vị. Giả sử các tham số truyền dẫn như đã được chọn ở Đức (mode 8K, 16 – QAM, tỷ lệ mã 2/3, khoảng bảo vệ 1/4), ta có tốc độ bit 13.27 Mb/s cho mỗi kênh 8 MHz. Nếu dùng mã hóa MPEG-2 thì số chương trình truyền hình/kênh tần số bị hạn chế tới bốn. Trong khi đó nếu dùng H.264/AVC thì số chương trình có thể tăng lên đến mười hoặc hơn, vì không chỉ độ lợi mã hóa mà cả độ lợi ghép kênh thống kê cho các tốc độ bit thay đổi đều cao hơn do có nhiều hơn các chương trình khác nhau. Một tùy chọn khác là dùng QPSK, tỷ lệ mã 1/2 với H.264/AVC sẽ cho phép đạt được

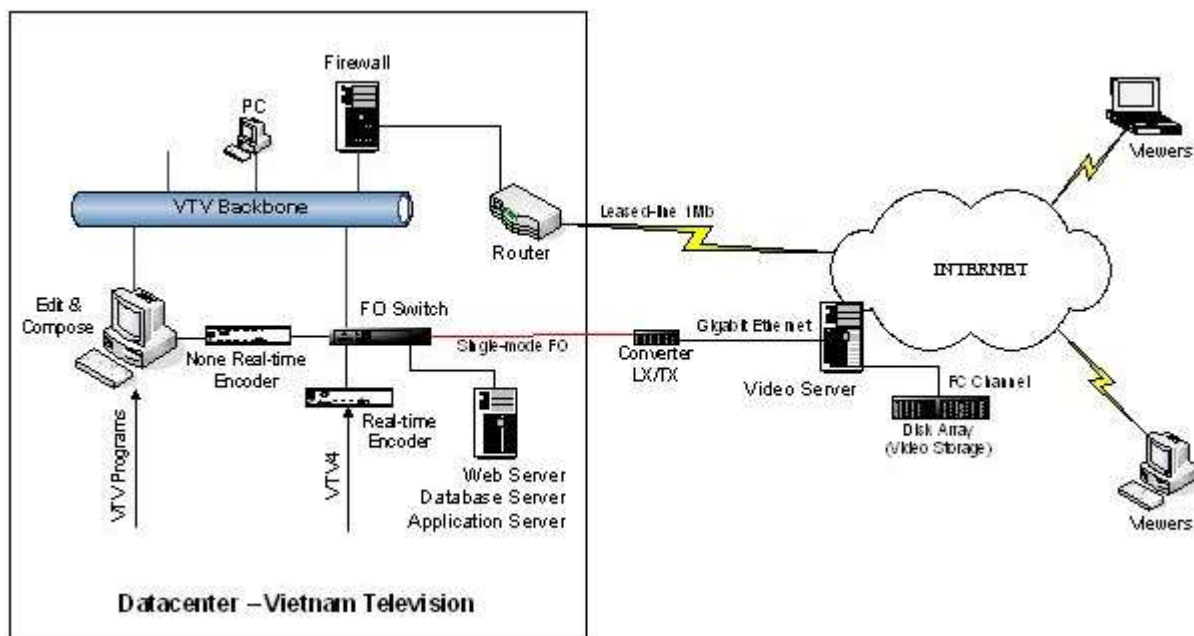
bốn chương trình / kênh, nhưng giảm công suất phát khoảng 15% so với mode truyền 16-QAM, 2/3 được nhắc tới ở trên.

Một phạm vi kinh doanh đáng quan tâm khác là truyền dẫn và lưu trữ truyền hình có độ phân giải cao (HD). Có thể mã hóa tín hiệu HD ở tốc độ 8 Mb/s, thích hợp với một DVD thông thường. Điều này sẽ làm tăng tốc thị trường home cinema. Cũng có thể truyền bốn chương trình HD qua một kênh vệ tinh hoặc cáp, làm cho dịch vụ này hấp dẫn hơn đối với các nhà quảng bá vì giá thành truyền dẫn thấp hơn nhiều so với MPEG-2.

Trong lĩnh vực thông tin di động, H.264/AVC có vai trò quan trọng vì hiệu quả nén tăng gấp đôi so với các sơ đồ mã hóa được xác định bởi hệ thống điện thoại di động 3GPP cho streaming, có nghĩa là H.263 Baseline, H.263+ và MPEG-4 Simple Profile. Điều này là cực kỳ quan trọng vì tốc độ dữ liệu có sẵn trong các hệ thống 3G là rất đắt. Với tốc độ bit thấp như vậy H.264/AVC sẽ cho phép tải nội dung multimedia cho các dụng cụ di động như điện thoại, PDA (Personal Digital Assistant) hoặc PC laptop qua các công nghệ mạng LAN không dây, DAB (Digital Audio Broadcasting), UMTS (Universal Mobile Telecommunications System) và cả DVB-T.

3.2. TRUYỀN HÌNH INTERNET-IPTV.

3.2.1. Sơ đồ hệ thống: được thể hiện trên hình 3.1



Hình 3.1. Hệ thống truyền hình Internet- VTV4

3.2.2. Mô tả hoạt động của hệ thống:

- Các chương trình truyền hình (được đưa vào dưới dạng tín hiệu tương tự hay băng từ ...) được đưa qua thiết bị dựng PC (nếu cần) và sau đó đưa vào thiết bị mã hóa Non Real-time Encoder để tạo chương trình dưới dạng file theo định dạng MPEG-4 AVC/H.264. Nhờ máy chủ Web Server (được cài đặt các phần mềm Apache Web Server, phần mềm hệ thống website VTV4 Online, Database server và Application Server), các file này được gửi đến thiết bị lưu trữ video Disk Array thông qua FO switch/cáp quang/converter LX/TX và lưu dưới dạng các thư mục (tùy theo tổ chức CSDL của hệ thống). Khi khách hàng có nhu cầu xem một chương trình truyền hình lưu trong hệ thống, máy chủ Video Server (có cài đặt phần mềm quản lý dòng video) sẽ tạo kết nối để khách hàng download chương trình về máy tính của mình. Do nhu cầu dựng không nhiều,

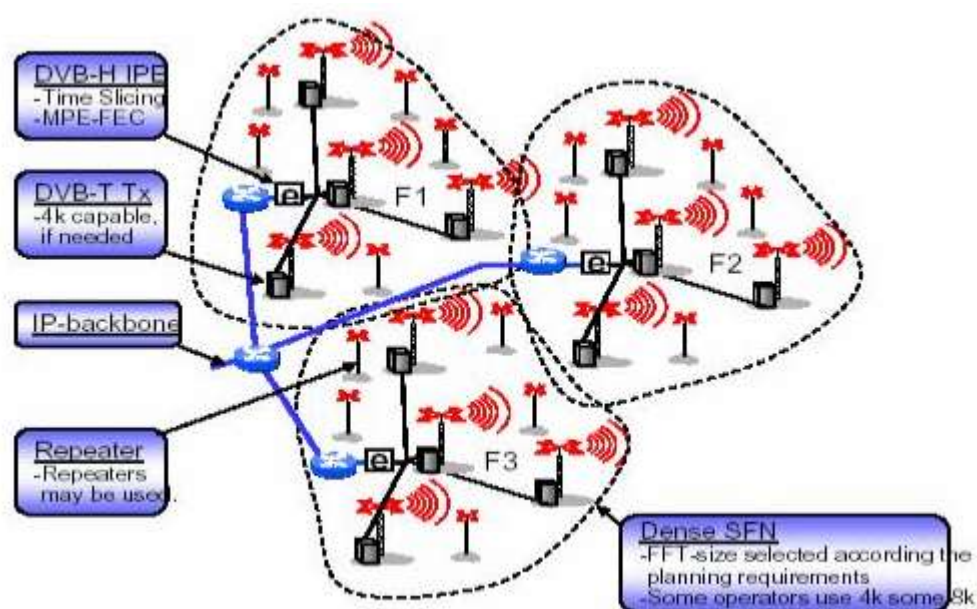
máy tính PC để dựng phi tuyến còn có thể dùng để update dữ liệu cho trang website VTV4 Online. Khi dịch vụ phát triển (số lượng người truy cập tăng cao, khối lượng video lưu trữ nhiều...), có thể tính đến việc tách các chức năng quản trị web và quản trị cơ sở dữ liệu trên hai máy chủ (cùng được đặt trong Trung tâm THDL của VTV).

- Vấn đề cung cấp điện cũng như an toàn, an ninh cho các thiết bị đặt tại Trung tâm Tích hợp dữ liệu (VTV) được bảo đảm bởi hệ thống an ninh mạng chung của Đài THVN. Với máy chủ Video Server đặt tại cổng Internet quốc tế, vấn đề an ninh hệ thống và chống xâm nhập sẽ được đảm nhận bởi hệ thống an ninh của nhà cung cấp dịch vụ ISP và phần mềm an ninh được cài đặt trực tiếp trên máy chủ, qui định và kiểm soát các cổng ra vào của các dòng dữ liệu.

3.3. TRUYỀN HÌNH DI ĐỘNG DVB- H.

DVB-H là hệ thống truyền dẫn quảng bá cho các datagram. Các datagram này có thể là các datagram IP hoặc các datagram khác, và có thể chứa các dữ liệu liên quan với các dịch vụ multimedia, các dịch vụ nạp file, hoặc các dịch vụ khác.

Mục đích của DVB-H là cung cấp phương thức hiệu quả để mang dữ liệu multimedia qua các mạng quảng bá mặt đất số đến các thiết bị handheld. Các đặc trưng chủ yếu liên quan với tính hiệu quả cần được xem xét là các hạn chế về nguồn cung cấp và các điều kiện truyền dòng truyền khác nhau do tính di động của thiết bị thu.



Hình 3.2. Mô hình mạng DVB-H riêng biệt

Để thực hiện mục đích đề ra, DVB-H đã bổ sung thêm một số giải pháp kỹ thuật vào tiêu chuẩn DVB-T, bao gồm: Time – slicing để tiết kiệm nguồn và hỗ trợ cho chuyển vùng trơn tru; MPE - FEC cho tính mạnh khỏe trong các môi trường truyền dẫn và bổ sung thêm tính di động, hỗ trợ cho tính di động; mode 4K tăng cường tính di động và tính mềm dẻo trong thiết kế mạng; báo hiệu DVB-H để giúp máy thu DVB-H để hoạt động đúng chức năng của nó.

DVB sử dụng mã hóa nguồn cho DVB-H sẽ là H.264/MPEG-4 AVC với các level và các thông số tương ứng cho ở bảng dưới. Dự kiến các dịch vụ DVB-H khởi đầu sẽ dùng máy thu mức B (điện thoại di động UMTS và thiết bị số cá nhân (Personal Digital Assistant – PDA)), sau đó sẽ mở rộng sang các máy thu dạng A và B. Ví dụ nếu ta dùng $\frac{1}{4}$ thông lượng kênh DVB-T như ở Đức hiện đang dùng (dung lượng tổng cộng 15 Mb/s) cho DVB-H thì sẽ có 3.7 Mb/s cho DVB-H, đủ cho truyền tải 6 DVB-H video stream hoặc 15 stereo audio signal với mã FEC mạnh bổ sung.

Bảng. Tốc độ dữ liệu cho video và audio theo H.264/MPEG-4 AVC.

Dạng máy thu DVB	H.264/AVC-level	Độ phân giải Video	Tốc độ bit cực đại	Ứng dụng mẫu
A	1	QCIF (180*144)	128 kb/s	UMTS Telephone
B	1.2	CIF (360*288)	384 kb/s	UMTS Telephone, PDA
C	2	CIF (360*288)	2 Mb/s	Máy thu bỏ túi (Pocket Receiver)
D	3	SDTV (720*576)	10 Mb/s	Máy thu hình
E	4	HDTV (1920*1080)	20 Mb/s	Máy thu hình

TÀI LIỆU THAM KHẢO

1. H.264 and MPEG-4 video compression. Lai E.G. Richardson
2. Giáo trình KT truyền hình, Đỗ Hoàng Tiến, Vũ Đức Lý. NXB KHKT-2001
3. <http://www.phatthanhtruyenhinh.vn/>