

**BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI**

---

**Hoàng Văn Tú**

**NGHIÊN CỨU CHUẨN NÉN H.264  
VÀ CÁC ỨNG DỤNG**

**LUẬN VĂN THẠC SĨ KỸ THUẬT  
CHUYÊN NGÀNH ĐIỆN TỬ VIỄN THÔNG**

**HOÀNG VĂN TÚ**

**ĐIỆN TỬ VIỄN THÔNG**

**2009 - 2011**

**Hà Nội – 2011**

**BỘ GIÁO DỤC VÀ ĐÀO TẠO  
TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI**

---

**Hoàng Văn Tú**

**NGHIÊN CỨU CHUẨN NÉN H.264  
VÀ CÁC ỨNG DỤNG**

Chuyên ngành : Điện tử viễn thông

**LUẬN VĂN THẠC SĨ KỸ THUẬT  
CHUYÊN NGÀNH ĐIỆN TỬ VIỄN THÔNG**

**NGƯỜI HƯỚNG DẪN  
TS. Nguyễn Hoàng Hải**

Hà Nội – 2011

## LỜI CAM ĐOAN

Tôi xin cam đoan nội dung trình bày trong luận văn này không phải là sự sao chép của luận văn, tài liệu, sách báo mà là sự nghiên cứu và tổng hợp xây dựng của tôi. Các nội dung trong luận văn có sự tham khảo các sách báo, tài liệu, tạp chí trong và ngoài nước. Tôi xin chịu trách nhiệm về các thông tin đề cập trong luận văn này.

*Hà Nội, ngày 21 tháng 9 năm 2011*  
*Người cam đoan*

**Hoàng Văn Tú**

## MỤC LỤC

<b>Danh mục viết tắt.....</b>	<b>1</b>
<b>Danh mục hình vẽ.....</b>	<b>4</b>
<b>Danh mục bảng biểu.....</b>	<b>6</b>
<b>Lời nói đầu .....</b>	<b>7</b>
<b>CHƯƠNG 1: CƠ BẢN VỀ NÉN VIDEO SỐ.....</b>	<b>10</b>
1.1 Kỹ thuật nén Video số.....	10
1.1.1 Khái niệm.....	10
1.1.2 Mô hình nén ảnh .....	10
1.2 Các đặc điểm của nén tín hiệu .....	11
1.2.1 Xác định hiệu quả của quá trình nén tín hiệu số.....	11
1.2.2 Độ dư thừa số liệu.....	12
1.2.3 Sai lệch bình phương trung bình .....	12
1.3 Các phương pháp nén .....	13
1.3.1 Nén không tổn hao .....	14
1.3.2 Nén có tổn hao.....	15
1.4 Tiêu chuẩn nén MPEG.....	15
1.4.1 Cấu trúc ảnh .....	15
1.4.2 Nhóm ảnh (GOP: Group Of Picture) .....	17
1.4.3 Cấu trúc dòng bit MPEG Video .....	18
1.5 Tiêu chuẩn nén MPEG-2 .....	20
1.5.1 Đặc tính và định mức (profile and level) .....	20
1.5.2 MPEG-2 4:2:2P@ML .....	23
<b>CHƯƠNG 2: CHUẨN NÉN VIDEO H.264 .....</b>	<b>25</b>
2.1 Giới thiệu chung về bộ CODEC H.264 .....	25
2.1.1 Bộ mã hoá (Encoder) .....	25
2.1.2 Bộ giải mã (Decoder) .....	27
2.2 Cấu trúc .....	28
2.2.1 Định dạng video (Video Format).....	28
2.2.2 Định dạng dữ liệu được mã hóa.....	28
2.2.3 Slice .....	29

2.2.4	Macroblock .....	31
2.2.5	Ảnh tham chiếu(Reference Picture).....	32
2.3	Profile.....	33
2.4	Một số kỹ thuật trong H.264 .....	34
2.4.1	Dự đoán nội khung (Intra Prediction).....	34
2.4.2	Dự đoán liên khung (Inter Prediction).....	37
2.4.3	Chuyển đổi (transform) và lượng tử .....	42
2.4.4	Mã hóa entropy .....	45
2.4.5	Bộ lọc Deblocking .....	45
2.4.6	Điều khiển tốc độ trong H.264 (Rate Control) .....	46
<b>CHƯƠNG 3: CÁC ỨNG DỤNG CỦA CHUẨN NÉN H.264 .....</b>		<b>54</b>
3.1	Ưu điểm nổi bật và các ứng dụng của H.264 .....	54
3.1.1	Ưu điểm của nén không gian .....	54
3.1.2	Ưu điểm của nén thời gian .....	54
3.1.3	Kích cỡ khối .....	54
3.1.4	Ưu điểm về lượng tử hoá và biến đổi.....	56
3.1.5	Ưu điểm đối với mã hoá Entropy .....	57
3.1.6	Các ứng dụng của chuẩn nén H.264 .....	57
3.2	Ứng dụng của H.264 trong mô hình IPTV .....	58
3.2.1	Mô hình truyền thông của IPTV .....	58
3.2.2	Lớp mã hoá video .....	59
3.2.3	Lớp đóng gói video .....	60
3.2.4	Lớp cấu trúc dòng truyền tải .....	63
3.2.5	Lớp giao thức truyền tải thời gian thực( tùy chọn) .....	69
3.2.6	Lớp truyền tải .....	74
3.2.7	Lớp IP .....	82
3.2.8	Lớp liên kết dữ liệu .....	91
3.2.9	Lớp vật lí .....	94
<b>CHƯƠNG 4: KẾT LUẬN VÀ KIẾN NGHỊ.....</b>		<b>96</b>
<b>Tóm tắt luận văn.....</b>		<b>97</b>
<b>Tài liệu tham khảo.....</b>		<b>99</b>

## Danh mục viết tắt

<b>Thuật ngữ</b>	<b>Tiếng anh</b>	<b>Ý nghĩa tiếng Việt</b>
AVC	Advanced Video Coding	Phương pháp mã hóa video tiên tiến
B Picture	Bidirectionally Picture	Ảnh dự đoán hai chiều
B-VOP	Bidirectionally predicted Inter-coded	Mặt phẳng đối tượng video loại B
CABAC	Context-adaptive Arithmetic Coding	
CAVLC	Context-adaptive Variable Length Coding	
CODEC	Coder and Decoder	Bộ mã hóa và giải mã
DCT	Discrete Cosine Transform	Biến đổi Cosin rời rạc
DPCM	Pulse Code Modulation	Điều chế xung mã vi phân
DPCM	Differential Pulse Code Modulation	Điều chế xung mã vi sai
GOB	Group of Blocks	Nhóm block
GOP	Group Of Picture	Nhóm ảnh
H.264		Là một chuẩn nén tín hiệu hiện đại hay còn gọi là AVC/ MPEG-4 part 10
HDTV	High Definition Television	Truyền hình độ phân giải cao
I_PCM		

IPTV	Internet Protocol Television	Truyền hình theo giao thức IP
IPTVCD	Internet Protocol Television Consumer Device	Thiết bị để sử dụng dịch vụ IPTV
IPTVCM		
I Picture	Intra-Code Picture	Ảnh I
JPEG	Joint Photographic Experts Group	Chuẩn nén ảnh JPEG
ME	Motion Estimation	Ước lượng chuyển động
MP@HL	Main Profile High Level	
MP@ML	Main Profile Main Level	
MB	Macro block	Khối macro
MPEG	Moving Picture Experts Group	Chuẩn nén ảnh động MPEG
MV	Motion Vector	Vector chuyển động
NAL	Network Abstraction Layer	Lớp mạng trừu tượng
NRI	nal_ref_idc	
PAT		
PMT		
P Picture	Predictive Code Picture	Ảnh P
PPS	picture_parameter_set	
PSNR	Peak Signal To Noise Ratio	Tỉ số đỉnh tín hiệu trên tạp âm (thông số đánh giá chất lượng)

PES	Packetized Element Stream	Dòng đóng gói cơ bản
QoS	Quality of Service	Chất lượng dịch vụ
QP	Quantization Parameter	Tham số lượng tử hóa
SDTV	Standart Digital Television	Truyền hình số độ phân giải chuẩn
RBSP		
SNR (S/N)	Signal Noise Ratio	Tỉ số tín hiệu trên tạp âm
RD		
RMS	Root Mean square	
RLC	Run Length Codes	
RTP	Real Time Protocol	Giao thức vận chuyển theo thời gian thực
RVLC	Reversible Variable Length Codes	Bộ mã bước chạy theo chiều ngược được
SP	Switching P	
SPS	Sequence Parameter Set	
TCP	Transmission Control Protocol	Giao thức điều khiển truyền
TS	Transport Stream	Dòng vận tải
UDP	User Datagram Protocol	Giao thức truyền tin không tin cậy lớp 4 trong mô hình OSI
VCL	Video Coding Layer	Lớp mã hóa video
VLC	Variable length codes	Bộ mã bước chạy

## Danh mục hình vẽ

- Hình 1.1: Mô hình hệ thống nén Video
- Hình 1.2: Sự phối hợp các kỹ thuật trong JPEG và MPEG
- Hình 1.3: Dự đoán bù chuyển động một chiều và hai chiều
- Hình 1.4: Kiến trúc dòng Video MPEG 11
- Hình 1.5: Cấu trúc số liệu nén ảnh MPEG 12
- Hình 2.1: Sơ đồ bộ mã hoá
- Hình 2.2: Sơ đồ bộ mã hoá cụ thể
- Hình 2.3: Sơ đồ bộ giải mã
- Hình 2.4: Cấu trúc mã hoá dữ liệu của H.264
- Hình 2.5: Chuỗi đơn vị NAL
- Hình 2.6: Slices
- Hình 2.7: Cấu trúc của slice
- Hình 2.8: MacroBlock
- Hình 2.9: Ảnh tham chiếu và bù chuyển động
- Hình 2.10: Các profile
- Hình 2.11: Dự đoán nội suy
- Hình 2.12: Dự đoán intra 4x4
- Hình 2.13: Các chế độ dự đoán intra\_4x4
- Hình 2.14: Các chế độ dự đoán intra\_16x16
- Hình 2.15: Phân mảnh macroblock
- Hình 2.16: Phân mảnh submacroblock
- Hình 2.17: Tính nội suy cho các chuyển động có khoảng cách không nguyên
- Hình 2.18: Minh họa dự đoán bù chuyển động đa khung
- Hình 2.19: Dự đoán liên khung trong slice B
- Hình 2.20: Ví dụ về một khung hình sử dụng bộ lọc Deblocking
- Hình 3.1: Mô hình truyền thông IPTV
- Hình 3.2: Cấu trúc của khối NAL
- Hình 3.3: Định dạng gói MPEG PES

Hình 3.4: Định dạng gói MPEG TS

Hình 3.5: Ánh xạ gói truy cập AVC sang gói MPEG PES

Hình 3.6: Ứng dụng nhãn thời gian với các gói MPEG PES

Hình 3.7: Mối liên hệ giữa PMT và PAT

Hình 3.8: Định dạng RTP header

Hình 3.9: Các gói MPEG TS

Hình 3.10: Ánh xạ nội dung H264/AVC (từng khối NAL riêng biệt) sang RTP payload

Hình 3.11: Ánh xạ nội dung H264/AVC (nhiều khối NAL riêng biệt) sang 1 RTP payload

Hình 3.12: Ánh xạ nội dung một H264/AVC NAL sang nhiều RTP payload

Hình 3.13: Cơ chế điều khiển luồng của TCP

Hình 3.14: Quá trình truyền thông trong mạng IPTV

Hình 3.15: Định dạng datagram dựa trên UDP

Hình 3.16: Định dạng gói video IPv4

Hình 3.17: Các lớp địa chỉ IP

Hình 3.18: Cấu trúc header của IPv6

## Danh mục bảng

- Bảng 1.1: Bảng thông số chính profile và level của tín hiệu chuẩn MPEG –2
- Bảng 2.1: Các loại slice
- Bảng 2.2: Các thành phần cấu trúc của Macroblock
- Bảng 3.1: Cấu trúc của một gói MPEG PES
- Bảng 3.2: Cấu trúc gói MPEG TS
- Bảng 3.3: Cấu trúc của gói IPTV dựa trên RTP
- Bảng 3.4: Định dạng của TCP segment
- Bảng 3.5: Cấu trúc datagram IPTV dựa trên UDP
- Bảng 3.6: Cấu trúc gói video IPv4
- Bảng 3.7: Các lớp địa chỉ IPv4
- Bảng 3.8: Mô tả trường của IPv6
- Bảng 3.9: Cấu trúc của Ethernet header
- Bảng 3.10: Cấu trúc khung Ethernet được dùng để mang nội dung MPEG-2
- Bảng 3.11 Tổng kết các lớp trong mô hình IPTV

## LỜI NÓI ĐẦU

### 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 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. Khối lượng nội dung mà các công ty truyền thông cũng như các nhà cung cấp dịch vụ thông tin có thể mang lại ngày càng lớn, ngoài ra họ còn có thể cung cấp nhiều dịch vụ theo yêu cầu thông qua hệ thống cáp, vệ tinh và các hạ tầng viễn thông đặt biệt là mạng Internet.

Các tiêu chuẩn mã hoá Video ra đời và phát triển với mục tiêu cung cấp các phương tiện cần thiết để tạo ra sự thống nhất giữa các hệ thống được thiết kế bởi những nhà sản xuất khác nhau đối với mọi loại ứng dụng Video; Nhờ vậy thị trường Video có điều kiện tăng trưởng mạnh. Chính vì lý do này nên những người sử dụng bộ giải mã cần có một chuẩn nén mới để đi tiếp chặng đường mà MPEG-2 đã bỏ dở.

Hiệp hội viễn thông quốc tế (ITU) và tổ chức tiêu chuẩn quốc tế/ Ủy ban kỹ thuật điện tử quốc tế (ISO/IEC) là hai tổ chức phát triển các tiêu chuẩn mã hoá Video. 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).

Những khuyến nghị của ITU được thiết kế dành cho các ứng dụng truyền thông Video thời gian thực như Video Conferencing hay điện thoại truyền hình. Mặt khác, những tiêu chuẩn MPEG được thiết kế hướng tới mục tiêu lưu trữ Video chẳng hạn như trên đĩa quang DVD, quảng bá Video số trên mạng cáp, đường truyền số DSL, truyền hình vệ tinh hay những ứng dụng truyền dòng Video trên mạng Internet hoặc thông qua mạng không dây (wireless).

Với đối tượng để truyền dẫn Video là mạng Internet thì ứng cử viên hàng đầu là chuẩn nén MPEG-4 AVC, còn được gọi là H.264, MPEG-4 part 10, H.26L hoặc JVT.

Mục tiêu chính của chuẩn nén H.264 đang phát triển nhằm cung cấp Video có chất lượng tốt hơn nhiều so với những chuẩn nén Video trước đây. Điều này có thể đạt được nhờ sự kế thừa các lợi điểm của các chuẩn nén Video trước đây. Không chỉ thế, chuẩn nén H.264 còn kế thừa phần lớn lợi điểm của các tiêu chuẩn trước đó là H.263 và MPEG-4 bao gồm 4 đặc điểm chính như sau:

Phân chia mỗi hình ảnh thành các Block (bao gồm nhiều điểm ảnh), do vậy quá trình xử lý từng ảnh có thể được tiếp cận tới mức Block.

Khai thác triệt để sự dư thừa về mặt không gian tồn tại giữa các hình ảnh liên tiếp bởi một vài mã của những Block gốc thông qua dự đoán về không gian, phép biến đổi, quá trình lượng tử và mã hoá Entropy (hay mã có độ dài thay đổi VLC).

Khai thác sự phụ thuộc tạm thời của các Block của các hình ảnh liên tiếp bởi vậy chỉ cần mã hoá những chi tiết thay đổi giữa các ảnh liên tiếp. Việc này được thực hiện thông qua dự đoán và bù chuyển động. Với bất kỳ Block nào cũng có thể được thực hiện từ một hoặc vài ảnh mã hoá trước đó hay ảnh được mã hoá sau đó để quyết định véc tơ chuyển động, các véc tơ này được sử dụng trong bộ mã hoá và giải mã để dự đoán các loại Block.

Khai thác tất cả sự dư thừa về không gian còn lại trong ảnh bằng việc mã các block dư thừa. Ví dụ như sự khác biệt giữa block gốc và Block dự đoán sẽ được mã hoá thông qua quá trình biến đổi, lượng tử hoá và mã hoá Entropy. Chính từ những

ưu điểm vượt trội trên tôi đã chọn việc nghiên cứu chuẩn H.264 làm luận văn tốt nghiệp với đề tài “ *nghiên cứu chuẩn nén H.264 và các ứng dụng* ”

*Nội dung của luận văn gồm 4 chương:*

### **Chương 1: Cơ bản về nén Video số**

Chương 1 trình bày những kiến thức cơ bản về nén Video số như khái niệm, đặc điểm, phương pháp nén và giới thiệu một chuẩn nén rất điển hình là MPEG.

### **Chương 2: Chuẩn mã hóa Video tiên tiến H.264**

Chương này đi vào chi tiết chuẩn mã hóa H.264 như: cấu trúc bộ Codec H.264 (bộ mã hóa và giải mã hóa), cấu trúc dữ liệu trong H.264, các profile của H.264

### **Chương 3: Các ứng dụng của chuẩn nén H.264**

H.264 có tỉ lệ nén rất cao nên nó được ứng dụng trong rất nhiều lĩnh vực. Chương này thể hiện các ứng dụng của chuẩn H.264 trong thực tế và nghiên đi sâu nghiên cứu ứng dụng của nó trong mô hình IPTV và cho chúng ta biết về các lớp trong mô hình IPTV.

### **Chương 4: Kết luận và kiến nghị**

Qua quá trình nghiên cứu chuẩn H.264 chương này đưa ra những kiến nghị, kết luận và hướng phát triển trong tương lai.

Tôi xin gửi lời cảm ơn chân thành tới thầy giáo TS. Nguyễn Hoàng Hải người đã hướng dẫn và giúp đỡ tôi hoàn thành được luận văn này. Tôi cũng xin gửi lời cảm ơn tới những thầy cô giáo người đã trang bị những kiến thức để tôi làm luận văn này. Mặc dù đã cố gắng nhưng chắc chắn luận văn vẫn còn rất nhiều thiếu sót, tôi mong tiếp thu được các ý kiến đóng góp của thầy cô và đồng nghiệp để luận văn của tôi được hoàn thiện hơn

Hà Nội, tháng 9/2011

**Hoàng Văn Tú**

## CHƯƠNG I:

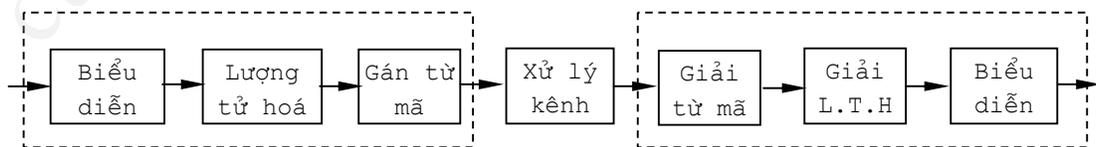
### CƠ BẢN VỀ NÉN VIDEO SỐ

#### 1.1 Kỹ thuật nén Video số

##### 1.1.1 Khái niệm

Nén về cơ bản là một quá trình trong đó lượng số liệu (data) biểu diễn lượng thông tin của một ảnh hoặc nhiều ảnh được giảm bớt bằng cách loại bỏ những số liệu dư thừa trong tín hiệu Video. Các chuỗi ảnh truyền hình có nhiều phần ảnh giống nhau, vậy tín hiệu truyền hình có chứa nhiều dữ liệu dư thừa, ta có thể bỏ qua mà không làm mất thông tin ảnh. Đó là các phần xoá dòng, xoá màn hình, vùng ảnh tĩnh hoặc chuyển động rất chậm, vùng ảnh nền giống nhau, mà ở đó các phần tử liên tiếp giống nhau hoặc khác nhau rất ít. Thường thì chuyển động trong ảnh truyền hình có thể được dự báo, do đó chỉ cần truyền các thông tin về chuyển động. Các phần tử lân cận trong ảnh thường giống nhau, nên chỉ cần truyền các thông tin biến đổi. Các hệ thống nén sử dụng đặc tính này của tín hiệu Video và các đặc trưng của mắt người là kém nhạy với sai số trong hình ảnh có nhiều chi tiết, các phần tử chuyển động. Quá trình sau nén ảnh là dẫn (giải nén) ảnh để tạo lại ảnh gốc hoặc xấp xỉ ảnh gốc.

##### 1.1.2 Mô hình nén ảnh



*Hình 1.1: Mô hình hệ thống nén Video.*

Ở tầng đầu tiên của bộ mã hoá video, tín hiệu video được trình bày dưới dạng thuận tiện để nén có hiệu quả nhất. Điểm cốt yếu là chỉ xác định cái gì được mã hoá. Sự biểu diễn có thể chứa nhiều mẫu thông tin để mô tả tín hiệu hơn là chính tín hiệu, nhưng hầu hết thông tin quan trọng chỉ tập trung trong một phần nhỏ của sự

mô tả này. Trong cách biểu diễn có hiệu quả, chỉ có phần nhỏ dữ liệu cần thiết để truyền cho việc tái tạo tín hiệu video.

-Hoạt động thứ hai của bộ mã hoá là lượng tử hoá, giúp rời rạc hoá thông tin được biểu diễn. Để truyền tín hiệu video qua một kênh số, những thông tin biểu diễn được lượng tử hoá thành một số hữu hạn các mức.

-Hoạt động thứ 3 là gán các từ mã. Các từ mã này là một chuỗi bit dùng để biểu diễn các mức lượng tử hoá.

Các quá trình sẽ ngược lại trong bộ giải mã video.

Mỗi hoạt động cố gắng loại bỏ phần dư thừa trong tín hiệu video và tận dụng sự giới hạn của khả năng nhìn của mắt người. Nhờ bỏ đi các phần dư thừa, các thông tin giống nhau hoặc có liên quan đến nhau sẽ không được truyền đi. Những thông tin bỏ đi mà không ảnh hưởng đến việc nhìn cũng không được truyền đi.

## 1.2 Các đặc điểm của nén tín hiệu số

### 1.2.1 Xác định hiệu quả của quá trình nén tín hiệu số

Hiệu quả nén được xác định bằng tỉ lệ nén, nghĩa là tỷ số giữa số lượng dữ liệu của ảnh gốc trên số lượng dữ liệu của ảnh nén.

Độ phức tạp của thuật toán nén được xác định bằng số bước tính toán trong cả hai quá trình mã hoá và giải mã. Thông thường thì thuật toán nén càng phức tạp bao nhiêu thì hiệu quả nén càng cao nhưng ngược lại giá thành và thời gian thực hiện lại tăng. Đối với thuật toán nén có tổn thất thì độ sai lệch được xác định bằng số thông tin bị mất đi khi tái tạo lại hình ảnh từ dữ liệu nén. Với nén không tổn thất thì chúng ta có thể có những thuật toán mã hoá càng gần với Entropy của thông tin nguồn, bởi vì lượng entropy của nguồn chính là tốc độ nhỏ nhất mà bất cứ một thuật toán nén không tổn thất nào cũng có thể đạt được.

Ngược lại, trong các nén có tổn thất thì mối quan hệ giữa tỷ lệ nén và độ sai lệch thông tin được Shannon nghiên cứu và biểu diễn dưới dạng hàm  $R_D$  (hàm về độ sai lệch thông tin). Lý thuyết của ông cũng chỉ ra rằng với thuật toán nén có tổn thất

thì chúng ta sẽ có hiệu quả cao nhất, nhưng ngược lại ta bị mất thông tin trong quá trình tái tạo lại nó từ dữ liệu nén. Trong khi đó nén không tổn thất, mặc dù đạt hiệu quả thấp nhưng ta lại không bị mất thông tin trong quá trình tái tạo lại nó. Vì vậy, ta phải tìm ra một biện pháp nhằm trung hoà giữa hai thuật toán nén này để tìm ra một thuật toán nén tối ưu sao cho hiệu quả nén cao mà lại không bị mất mát thông tin.

### 1.2.2 Độ dư thừa số liệu

Nén số liệu là quá trình giảm lượng số liệu cần thiết để biểu diễn cùng một lượng thông tin cho trước. Số liệu và thông tin không đồng nghĩa với nhau, số liệu chỉ là phương tiện dùng để truyền tải thông tin. Cùng một lượng thông tin cho trước có thể biểu diễn bằng các lượng số liệu khác nhau.

Độ dư thừa số liệu là vấn đề trung tâm trong nén ảnh số. Đánh giá cho quá trình thực hiện giải thuật nén là tỷ lệ nén ( $C_N$ ) được xác định như sau: Nếu  $N_1$  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 với tập hợp số liệu thứ hai có thể được định nghĩa như sau:

$$R_D = 1 - 1/C_N$$

$$\text{Trong đó: } C_N = N_1/N_2$$

Trong trường hợp  $N_1 = N_2$  thì  $C_N = 1$  và  $R_D = 0$ , có nghĩa là so với tập số liệu thứ hai thì tập số liệu thứ nhất không chứa số liệu dư thừa. Khi  $N_2 \ll N_1$  thì  $C_N$  tiến tới vô cùng và  $R_D$  tiến tới 1, có nghĩa là độ dư thừa số liệu tương đối của tập số liệu thứ nhất là khá lớn hay tập số liệu thứ hai đã được nén khá nhỏ.

### 1.2.3 Sai lệch bình phương trung bình

Một đánh giá thống kê khác có thể đánh giá cho nhiều giải thuật nén là sai lệch bình phương trung bình so với ảnh gốc RMS (Root Mean Square) được tính bởi biểu thức:

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

Trong đó:

RMS – sai lệch bình phương trung bình

$X_i$  – Giá trị điểm ảnh ban đầu

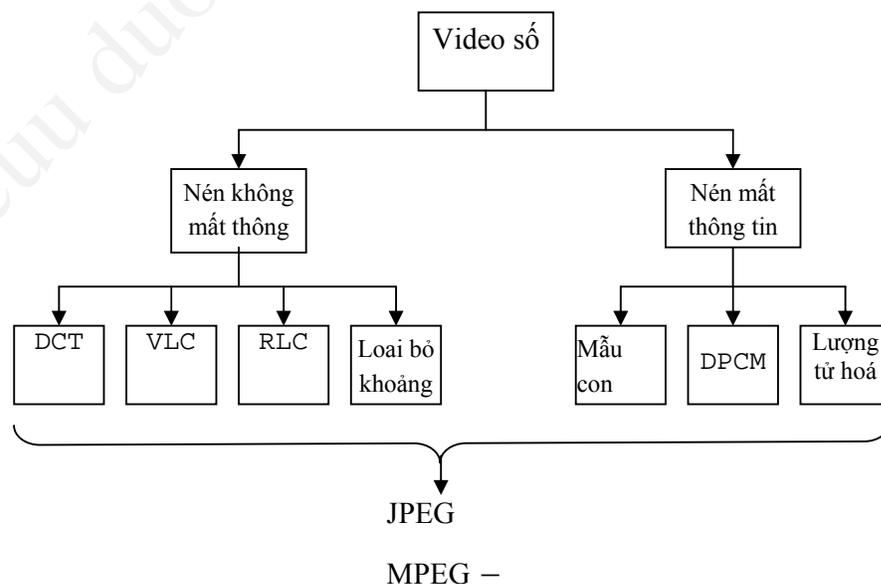
$X_i'$  – Giá trị điểm ảnh sau khi nén

$n$  – Tổng số điểm ảnh trong một ảnh

RMS chỉ ra sự khác nhau thống kê giữa ảnh ban đầu và ảnh sau khi nén. Đa số trường hợp khi nén chất lượng của ảnh nén là tốt với RMS thấp. Tuy nhiên, trong một số trường hợp có thể xảy ra là chất lượng ảnh nén với RMS cao tốt hơn ảnh với RMS thấp hơn.

### 1.3 Các phương pháp nén

Các hệ thống nén số liệu 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 hình ảnh phù hợp ứng với một ứng dụng nhất định. Nhiều kỹ thuật nén mất và không mất thông tin (loss/lossless data reduction techniques) đã được phát triển trong nhiều năm qua. Chỉ có một số ít trong chúng có thể áp dụng cho nén video số.



Hình 1.2: Sự phối hợp các kỹ thuật trong JPEG và MPEG

Hình 1.2 minh họa kỹ thuật nén được sử dụng để tạo thành các tín hiệu nén JPEG (Joint Photographic Expert Group) và MPEG (Moving Picture Expert Group). Sử dụng các kỹ thuật này một cách riêng rẽ thực tế không đưa lại một kết quả nào về giảm tốc độ dòng tín hiệu.

Tuy nhiên, phối hợp một số các kỹ thuật này sẽ đem lại những hệ thống nén vô cùng hiệu quả như hệ thống nén JPEG, MPEG-1, MPEG-2.

### 1.3.1 Nén không tổn hao

Nén không mất thông tin cho phép phục hồi lại đúng tín hiệu ban đầu sau khi giải nén. Đây là một quá trình mã hoá có tính thuận nghịch. Hệ số nén phụ thuộc vào chi tiết ảnh được nén. Hệ số nén của phương pháp nén không mất thông tin nhỏ hơn 2:1. Các kỹ thuật nén không mất thông tin bao gồm:

#### a. Mã hoá với độ dài thay đổi (VLC)

Phương pháp này còn được gọi là mã hoá Huffman và mã hoá Entropy, dựa trên khả năng xuất hiện của các giá trị biên độ trùng hợp trong một bức ảnh và 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. Khi thực hiện giải nén, các thiết lập mã trùng hợp sẽ được sử dụng để tái tạo lại giá trị tín hiệu ban đầu.

#### b. Mã hoá với độ dài động (RLC)

Phương pháp này dựa trên sự lặp lại của cùng giá trị mẫu để tạo ra các mã đặc biệt biểu diễn sự bắt đầu và kết thúc của giá trị được lặp lại.

Chỉ các mẫu có giá trị khác không mới được mã hoá. Số mẫu có giá trị bằng không sẽ được truyền đi dọc theo cùng dòng quét.

#### c. Sử dụng khoảng xoá dòng, xoá mảnh

Vùng thông tin xoá được loại bỏ khỏi dòng tín hiệu để truyền đi vùng thông tin tích cực của ảnh. Theo phương pháp đó, thông tin xoá dòng và xoá mảnh sẽ không được ghi giữ và truyền đi. Chúng được thay bằng các dữ liệu đồng bộ ngắn hơn tùy theo ứng dụng.

#### d. Biến đổi cosin rời rạc (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. Nếu độ dài từ mã hệ số của phép biến đổi DCT nhỏ hơn, quá trình này trở nên có mất thông tin.

Trong truyền hình, phương pháp nén không tổn hao được kết hợp trong các phương pháp nén có tổn hao sẽ cho tỷ lệ nén tốt mà không gây mất mát về độ phân giải.

### 1.3.2 Nén có tổn hao

Nén có tổn hao chấp nhận mất mát một ít thông tin để gia tăng hiệu quả nén, rất thích hợp với nguồn thông tin là hình ảnh và âm thanh. Như vậy, nén có tổn hao mới thật sự có ý nghĩa đối với truyền hình. Nó có thể cho tỷ lệ nén ảnh cao để truyền dẫn, phát sóng đồng thời cho một tỷ lệ nén thích hợp cho xử lý và lưu trữ ảnh trong studio.

Nén tổn hao thường thực hiện theo 3 bước liên tục:

-*Bước 1:* Biến đổi tín hiệu từ miền thời gian (không gian) sang miền tần số bằng cách sử dụng các thuật toán chuyển vị như biến đổi cosin rời rạc DCT. Bước này thực hiện việc giảm độ dư thừa của pixel trong ảnh, tuy nhiên quá trình này không gây tổn hao.

-*Bước 2:* Thực hiện lượng tử hoá các hệ số DCT, số liệu được “làm tròn” bằng cách làm tròn. Việc mất mát số liệu xảy ra ở giai đoạn làm tròn này.

*Bước 3:* Nén số liệu đã biến đổi và làm tròn bằng cách mã hoá Entropy, ở đây sử dụng các mã không tổn hao như mã Huffman, RLC,...

## 1.4 Tiêu chuẩn nén MPEG

### 1.4.1 Cấu trúc ảnh

MPEG định nghĩa các loại ảnh khác nhau cho phép sự linh hoạt để cân nhắc giữa hiệu quả mã hoá và truy cập ngẫu nhiên. Các loại ảnh đó như sau:

*a. ẢNH I: (Intra- Code Picture)*

Các ảnh I được mã hoá theo mode intra để có thể giải mã mà không cần sử dụng dữ liệu từ bất cứ một ảnh nào khác. Đặc điểm của phương pháp mã hoá này như sau:

- Chi loại bỏ được sự dư thừa không gian.
- Dùng các điểm trong cùng một khung để dự báo.
- Không có bù chuyển động.
- Các thông tin được mã hoá rõ ràng, minh bạch nên số lượng bit yêu cầu lớn.

Do được mã hoá Intra, ảnh I bao giờ cũng là ảnh đầu tiên trong một nhóm ảnh hay một chuỗi ảnh. Nó cung cấp thông tin khởi động các ảnh tiếp theo trong nhóm.

*b. ẢNH P (Predictive Code Picture)*

Ảnh P được mã hoá liên ảnh một chiều (Interframe một chiều):

- Dự báo Inter một chiều.
- Ảnh dự báo được tạo ảnh tham chiếu trước đó (dự báo nhân quả). Ảnh tham chiếu này có thể là ảnh I hoặc ảnh P gần nhất.
- Có sử dụng bù chuyển động. Thông tin ước lượng chuyển động của các khối nằm trong vector chuyển động (motion vector). Vector này xác định Macroblock nào được sử dụng từ ảnh trước.

Do vậy ảnh P bao gồm cả những MB mã hoá (I-MB) là những Macroblock chứa thông tin lấy từ ảnh tham chiếu và những MB mã hoá Intra là những MB chứa thông tin không thể mượn từ ảnh trước.

Ảnh P có thể sử dụng làm ảnh tham chiếu tạo dự báo cho ảnh sau.

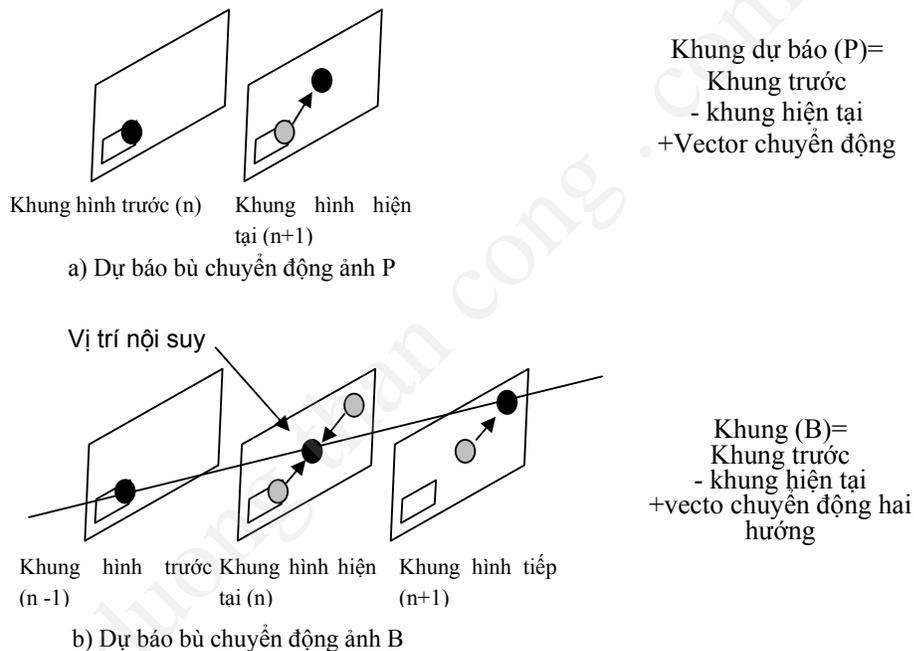
*c. Ảnh B (Bidirectionally Picture)*

Ảnh B là ảnh mã hoá liên ảnh hai chiều. Tức là:

- Có sử dụng bù chuyển động.
- Dự báo không nhân quả, ảnh dự báo gồm các MB của cả khung hình trước đó và sau đó.

Việc sử dụng thông tin lấy từ ảnh trong tương lai hoàn toàn có thể thực hiện được vì tại thời điểm mã hoá thì bộ mã hoá đã sẵn sàng truy cập tới ảnh phía sau.

Ảnh B không được sử dụng làm ảnh tham chiếu tạo dự báo cho các ảnh sau.



Hình 1.3: Dự đoán bù chuyển động một chiều và hai chiều

#### 1.4.2 Nhóm ảnh (GOP- Group Of Picture)

Nhóm ảnh là một tập các ảnh mà đầu tiên phải là ảnh hoàn chỉnh I, tiếp sau đó là một loạt các ảnh P, B. Nhóm ảnh có hai loại:

- *Cấu trúc mở*: Luôn bắt đầu từ một ảnh I và kết thúc bằng một ảnh I tiếp theo, nghĩa là ảnh cuối cùng của GOP dùng ảnh đầu tiên của GOP tiếp theo làm chuẩn.

- *Cấu trúc khép kín*: Việc dự đoán ảnh không sử dụng thông tin của GOP khác. Ảnh cuối cùng của một GOP bao giờ cũng là ảnh P.

### 1.4.3 Cấu trúc dòng bit MPEG Video

Cấu trúc số liệu Video MPEG-1 và MPEG-2 bao gồm 6 lớp như sau

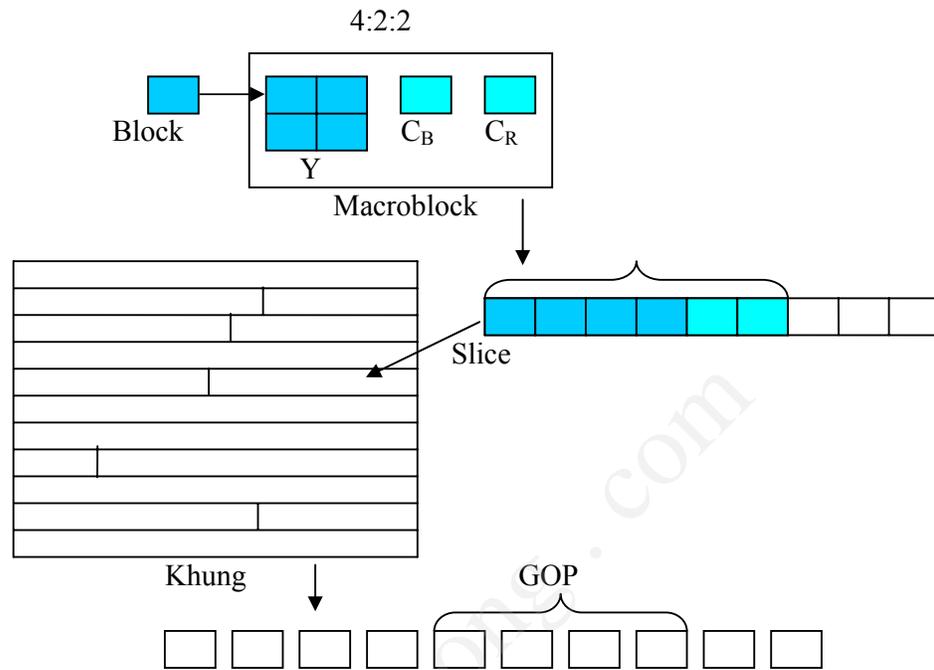
- *Khối (Block)*: Là đơn vị cơ bản cho chuyển đổi DCT. Bao gồm 8x8 điểm ảnh tín hiệu chói hoặc tín hiệu màu.

- *Khối Macro Block*: Là nhóm các khối DCT tương ứng với thông tin của một cửa sổ 16x16 điểm ảnh gốc. Có nhiều dạng Macro Block khác nhau phụ thuộc vào cấu trúc lấy mẫu được sử dụng.

Phần đầu đề (Header) của Macroblock chứa thông tin phân loại (Y hay CB, CR) và vector bù chuyển động tương ứng.

- *Lát (slice)*: Được cấu thành từ một hay một số MB liên tiếp nhau. Phần header của slice chứa thông tin về vị trí của nó trong ảnh và tham số quét lượng tử (quantized scaling factor). Kích cỡ của slice quyết định bởi mức bảo vệ lỗi cần có trong ứng dụng vì bộ giải mã sẽ bỏ qua slice bị lỗi. Hệ số một chiều DC được định vị tại điểm bắt đầu mỗi slice.

- *Ảnh* : Lớp ảnh cho bên thu biết về loại mã hoá khung I, P, B. Phần Header mang thứ tự truyền tải của khung để bên thu hiển thị khung theo đúng thứ tự, ngoài ra còn có một số thông tin bổ sung như thông tin đồng bộ, độ phân giải và vector chuyển động.

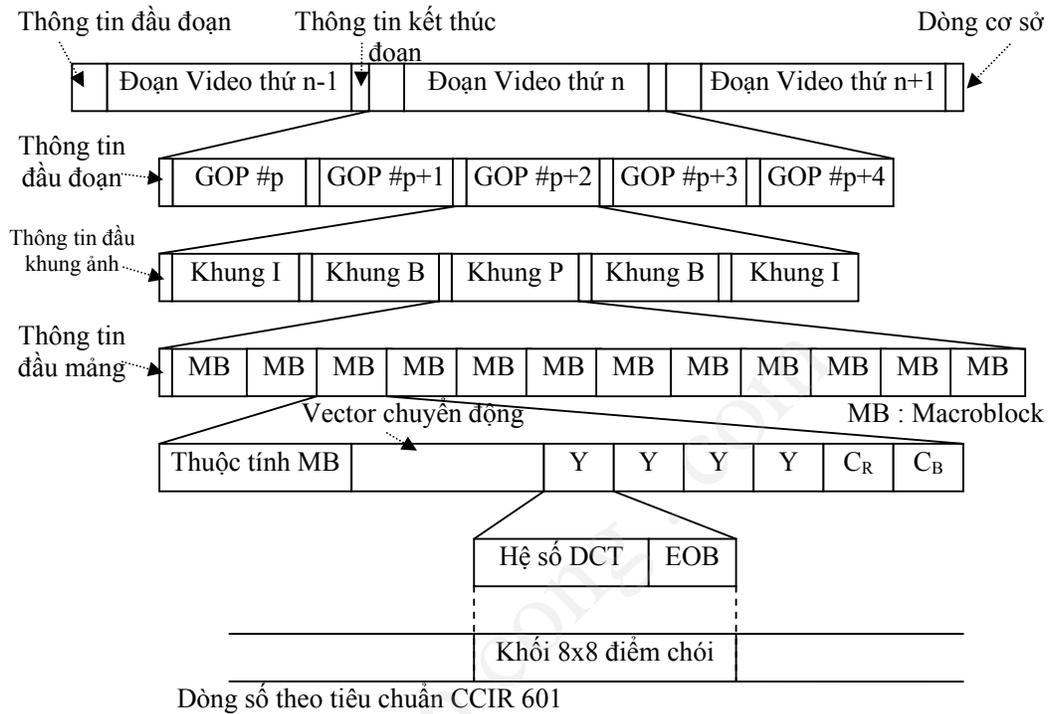


Hình 1.4: Kiến trúc dòng Video MPEG

- Nhóm ảnh (Group of Picture): Gồm cấu trúc các ảnh I, B và P. Mỗi nhóm bắt đầu bằng ảnh I cung cấp điểm vào ra và tìm kiếm.
- Chuỗi video ( Video Sequence): Lớp chuỗi bao gồm phần Header, một hoặc một số nhóm ảnh (Picture Group) và phần kết thúc chuỗi (Sequence End Code).

Thông tin quan trọng nhất của phần Header là kích thước (dọc, ngang) của mỗi ảnh, tốc độ bit, tốc độ ảnh và dung lượng đi hỏi bộ đệm dữ liệu bên thu.

Thông tin chuỗi ảnh và phần Header của chuỗi là dòng bit đã mã hoá, còn gọi là dòng video cơ bản.



Hình 1.5: Cấu trúc số liệu nén ảnh MPEG

## 1.5 Tiêu chuẩn nén MPEG-2

### 1.5.1 Đặc tính và định mức (profile and level)

Nén MPEG –2 có một chuỗi các mức (level) và đặc tính (profile) được dùng cho nhiều ứng dụng khác nhau. Cấu trúc tín hiệu số trong chuẩn MPEG –2 rất phức tạp. Việc sử dụng tiêu chuẩn MPEG –2 không phải lúc nào cũng cần thiết hoặc có ý nghĩa. Vì thế dẫn đến việc phân chia cấu trúc thành phần các tập con gọi là các profiles. Trong phạm vi mỗi profile chỉ cho phép sử dụng các phần của các phần tử vừa phải trong cấu trúc tín hiệu đầy đủ. Có 5 định nghĩa về profile:

- *Simple profile* (profile đơn giản): Số bước nén thấp, chỉ cho phép mã hoá các ảnh loại I hoặc P. Việc tách các ảnh loại B sẽ làm giảm bộ nhớ cho giải mã chuỗi.

□ *Main profile*: Cho phép sử dụng tất cả các loại ảnh, nhưng không tạo các mức bất kỳ. Chất lượng tốt hơn simple profile nhưng tốc độ bit không thay đổi.

□ *SNR profile scalable* (profile phân cấp theo SNR): Tiêu chuẩn MPEG –2 cho phép phân cấp theo tỷ số nén tín hiệu trên tạp âm (S/N). Tính phân cấp theo S/N có nghĩa là chất lượng hình ảnh và tỷ số S/N có tính thoả hiệp.

□ *Spatially Scalable profile* (phân cấp theo không gian): Tính phân cấp theo không gian có nghĩa là có sự thoả hợp 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 ví dụ như truyền hình tiêu chuẩn, lớp cao hơn bao gồm ảnh có độ phân giải cao hơn ví dụ như truyền hình độ phân giải cao (HDTV).

□ *High profile* (profile cao): Cho phép đối với cả hai loại thang mức và chuẩn 4:2:2 của tín hiệu video. Nó bao gồm toàn bộ công cụ của profile trước cộng thêm khả năng mã hoá các tín hiệu khác nhau cùng một lúc. Hay nói đúng hơn là “High profile” là một hệ thống hoàn hảo được thiết kế cho toàn bộ ứng dụng mà không bị giới hạn bởi tốc độ bit cao.

Vấn đề hạn chế các mức có liên quan đến độ phân giải cực đại của ảnh. Có 4 mức hạn chế sau:

- *Low level* (Mức thấp): Ứng với độ phân giải của MPEG –1, có nghĩa là bằng độ phân giải truyền hình tiêu chuẩn.

- *Main level* (mức chính): Độ phân giải của truyền hình tiêu chuẩn.

- *High 1440 level* (mức cao 1440): Độ phân giải của HDTV với 1440 mẫu/dòng.

- *High level* (mức cao): Độ phân giải HDTV với 1920 mẫu/dòng.

**Bảng 1.1: Bảng thông số chính profile và level của tín hiệu chuẩn  
MPEG –2**

Profile Level	Đơn giản (Simple)	Chính (Main)	Phân cấp theo SNR	Phân cấp theo không gian	Cao (High)
Thấp (Low)		4:2:0 352x288 4 Mbit/s	4:2:0 352x 288 4 Mbit/s I,B,P		
Chính (Main)	4:2:0 720x576 15 Mbit/s I,P	4:2:0 720x576 15 Mbit/s I,B,P	4:2:0 720x 576 15 Mbit/s I,B,P		4:2:0 720x576 20 Mbit/s I,B,P
Cao 1440 (High 1440)		4:2:0 1440x1152 60 Mbit/s I,B,P		4:2:0 1440x1152 60 Mbit/s I,B,P	4:2:0; 4:2:2 1440x11 52 80 Mbit/s I,B,P

Cao (High)		4:2:0			4:2:0
		1920×1152			;4:2:2
		80 Mbit/s			1920×1152
		I,B,P			52
					100
					Mbit/s
					I,P,B

Kết hợp 4 level và 5 profile ta được tổ hợp 20 khả năng và hiện nay đã có 11 khả năng được ứng dụng như bảng ( theo tài liệu của Techtronic).

Với MPEG –2 MP@ML có thể nén tín hiệu truyền hình xuống còn (3Π5) Mbit/s, rất phù hợp và đáp ứng được tính kinh tế cho phát quảng bá các chương trình truyền hình tiêu chuẩn (SDTV). Còn đối với HDTV thì sử dụng MPEG –2 P@HL và MPEG –2 4:2:2 MP@HL

### 1.5.2 MPEG-2 4:2:2P@ML

Trong bảng 1.1: các tiêu chuẩn đều lấy mẫu theo tiêu chuẩn 4:2:2 và cho tốc độ bit thấp rất phù hợp cho công đoạn truyền dẫn, phát sóng. Tuy nhiên nó không thoả mãn yêu cầu chất lượng cho công đoạn sản xuất hậu kỳ. Chuẩn 4:2:0 không thể cho một hình ảnh chất lượng studio sau một vài thế hệ gia công tín hiệu bởi phép nội suy tín hiệu mẫu. Sử dụng tốc độ bit 15 Mbit/s với GOP nhỏ chất lượng hình ảnh sẽ kém, GOP lớn sẽ gây khó khăn cho tất cả các thiết bị có chuyển đổi tín hiệu trong thời gian xoá màn hình. Từ năm 1994 nhiều nhà sản xuất và sử dụng thấy cần phải có tiêu chuẩn MPEG –2 4:2:2 P@ML (Profile Main Level) với tốc độ bit đạt 50 Mbit/s có thể đáp ứng được nhu cầu chất lượng trong các ứng dụng chuyên nghiệp.

Tháng 1/1996, MPEG –2 4:2:2P@ML trở thành tiêu chuẩn Quốc tế. Nó hơn hẳn MPEG –2 MP@ML trên nhiều khía cạnh: tốc độ bit bằng 50 Mbit/s và có thể đáp ứng được cả hai chuẩn Video 4:2:2 và 4:2:0. Hệ thống này có đặc điểm chính sau đây:

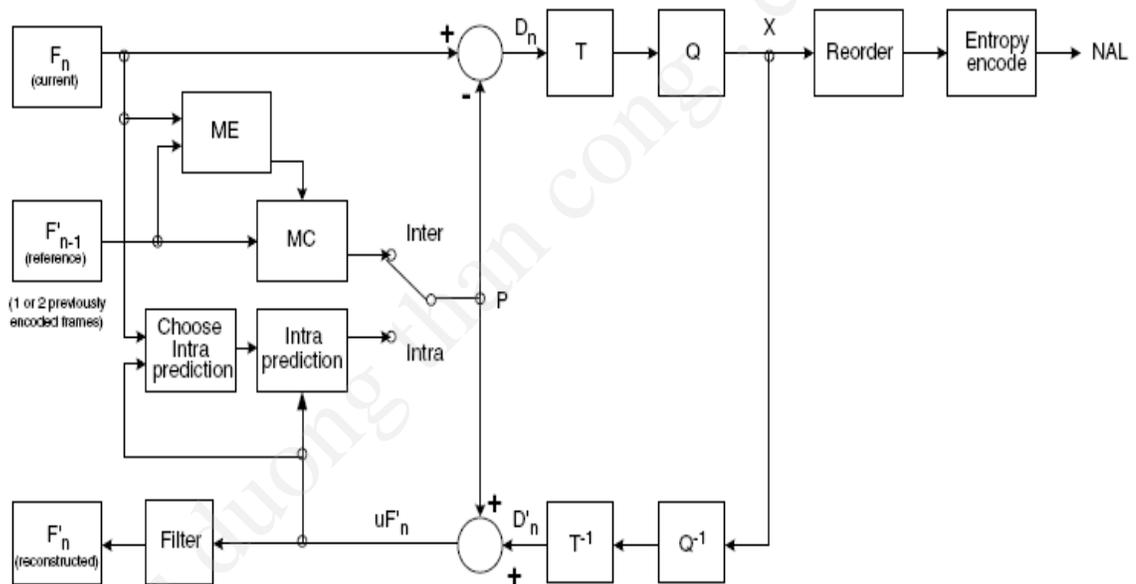
- Có độ mềm dẻo cao và tính khai thác hỗn hợp. Có khả năng giải mã trong phạm vi (15→50)Mbit/s với bất kỳ loại phối hợp nào giữa các ảnh I, P và B.
- Chất lượng cao hơn MP@ML.
- Độ phân giải màu tốt hơn MP@ML.
- Xử lý hậu kỳ sau khi nén và giải nén.
- Nén và giải nén nhiều lần.
- Nhóm ảnh nhỏ, thuận tiện cho công nghệ dựng hình.
- Có khả năng biểu thị tất cả các dòng tích cực của tín hiệu Video.

Có khả năng biểu thị thông tin trong khoảng thời gian xoá màn.

## CHƯƠNG 2: CHUẨN NÉN H.264

### 2.1 Giới thiệu chung về bộ CODEC H.264

#### 2.1.1 Bộ mã hoá (Encoder)



Hình 2. 2: Sơ đồ bộ mã hoá

Điểm khác biệt của H.264 so với các bộ mã hóa khác là có sự lựa chọn chế độ mã hóa liên ảnh (Inter) hoặc trong ảnh (Intra). Chế độ mã hóa trong ảnh cho phép một MB có thể được nội suy từ giá trị các điểm ảnh ở MB lân cận trong cùng ảnh và nhờ đó làm tăng hiệu quả nén trong miền không gian.

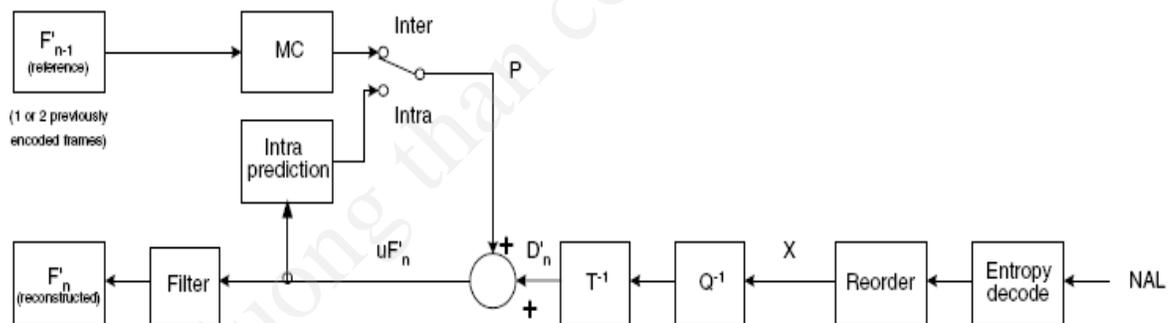


tin của véc tơ chuyển động... được nén thành bitstream, qua NAL (Network Abstraction Layer) để truyền đi hay lưu trữ.

➤ **Quá trình khôi phục** : quá trình mã hoá thực hiện từ phải qua trái

- Cũng giống như quá trình mã hoá và truyền đi từng block trong 1 Macroblock, quá trình mã hoá sẽ giải mã block nhằm mục đích tạo ra block tham chiếu cho lần dự đoán kế tiếp.
- Block hiệu D'n cộng với P khôi phục lại block. Bộ lọc có nhiệm vụ giảm nhiễu trong từng block. Nhiều block được khôi phục sẽ tạo thành bức tranh tham chiếu F'n.

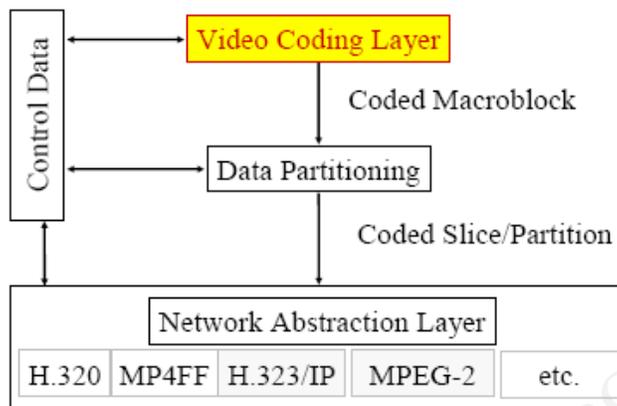
### 2.1.2 Bộ giải mã (Decoder)



Hình 2.3: Sơ đồ bộ giải mã

Bộ giải mã hóa nhận được 1 luồng dữ liệu nén từ NAL và giải mã entropy nhưng thành phần cơ bản của dữ liệu để tạo ra tập các hệ số được lượng tử hóa X. Những hệ số này được "scale" và chuyển đổi ngược thành  $D_n$ . Sử dụng thông tin tiêu đề được giải mã từ lượng bit, bộ giải mã tạo ra khối dự đoán PRED, phân biệt với khối PRED được tạo ở bộ mã hóa.

## 2.2 Cấu trúc



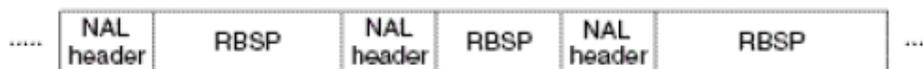
Hình 2.4: cấu trúc mã hoá dữ liệu của H.264

### 2.2.1 Định dạng video (Video Format)

H.264 hỗ trợ mã hóa và giải mã video 4:2:0 quét liên tục hoặc xen kẽ. Khung quét xen kẽ bao gồm 2 trường (trên và dưới) tách biệt theo thời gian với định dạng mặc định.

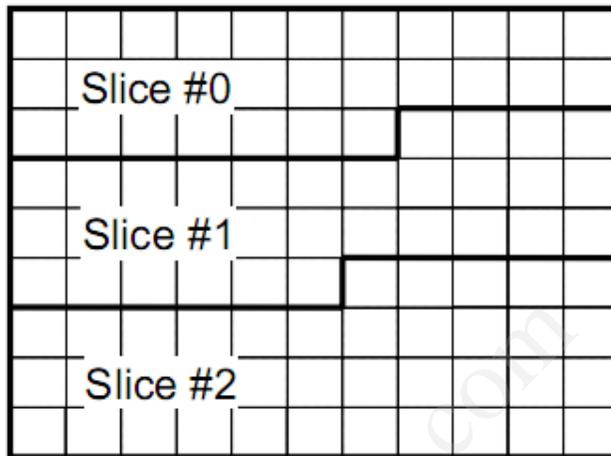
### 2.2.2 Định dạng dữ liệu được mã hóa

H.264 phân biệt lớp mã hóa video (Video Coding layer VCL) và lớp mạng trừu tượng (Network Abstraction Layer – NAL). Đầu ra của quá trình mã hóa là dữ liệu lớp mã hóa video VCL (chuỗi bit biểu diễn dữ liệu video đã được mã hóa) sẽ được ánh xạ vào các đơn vị của lớp mạng trừu tượng- NAL trước khi truyền dẫn hay lưu trữ. Mỗi đơn vị NAL bao gồm chuỗi byte thô về thứ tự tải, và một tập các thông tin ứng với dữ liệu video hay còn gọi là thông tin header. Một chuỗi video được mã hóa được biểu diễn bởi chuỗi các đơn vị NAL mà có thể được truyền dẫn trên các mạng gói hay luồng bit trên đường truyền hay lưu ra file. Mục đích của việc phân chia các lớp VCL và NAL là để phân biệt đặc tính mã hóa tại lớp VCL và đặc tính truyền dẫn tại lớp NAL.



Hình 2. 5: Chuỗi đơn vị NAL

### 2.2.3 Slice

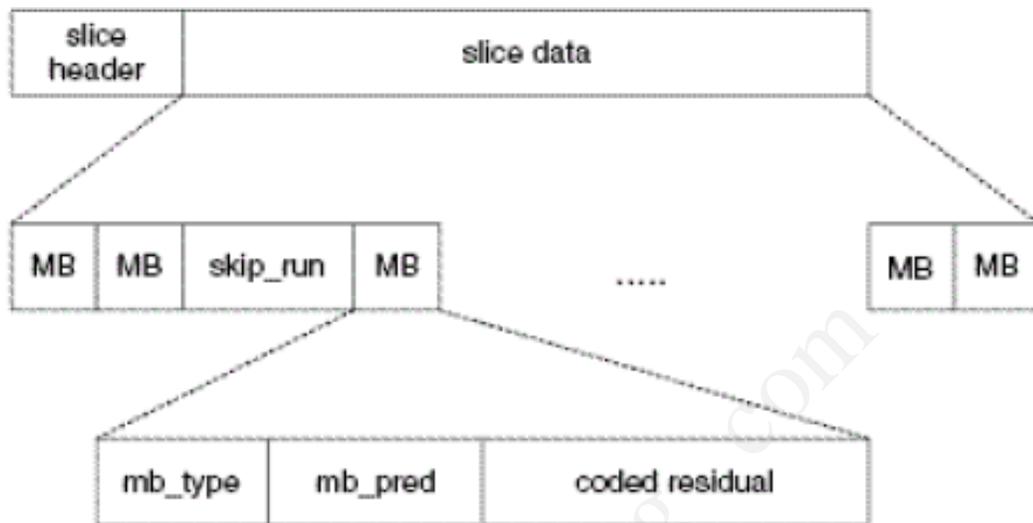


Hình 2. 6: Slices

Một ảnh video được mã hóa gồm một hay nhiều slice. Mỗi slice bao gồm số nguyên các khối macro. Số lượng khối macro trong một bức ảnh có thể không cố định. Có sự phụ thuộc lẫn nhau tối thiểu giữa các slice đã được mã hóa để giúp giảm sự lan truyền lỗi. Chuẩn mã hoá H.264 có 5 loại slice được mã hóa và một ảnh được mã hóa có thể bao gồm nhiều loại slice khác nhau, ví dụ ảnh được mã hóa trong profile cơ bản có thể bao gồm các slice I và P và ảnh được mã hóa ở profile chính hay mở rộng có thể gồm I,P,B slice. Header slice (phần thông tin mào đầu) định nghĩa loại slice và ảnh mã hóa mà slice đó thuộc về và có thể kèm theo các thông tin hướng dẫn liên quan đến quản lý ảnh tham chiếu. Phần dữ liệu của slice bao gồm chuỗi các khối macro được mã hóa và chỉ thị bỏ qua (không được mã hóa) khối macro. Mỗi khối macro bao gồm chuỗi các thành phần header và dữ liệu dư thừa được mã hóa.

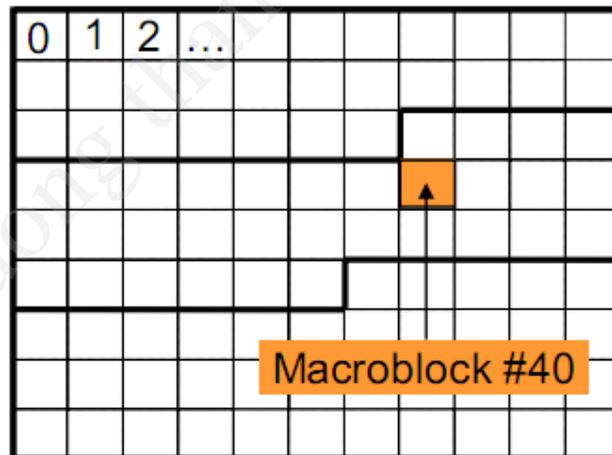
**Bảng 2.1: Các loại slice**

<b>Loại Slice</b>	<b>Mô tả</b>	<b>Profile</b>
I (Intra)	Chỉ bao gồm khối macro I (mỗi khối hoặc khối macro được dự đoán từ dữ liệu được mã hóa trước đó trong cùng một slice)	Tất cả
P (Predicted)	Bao gồm khối macro P (mỗi khối macro hoặc vùng macro được dự đoán từ danh sách các ảnh trong list 0 hoặc là khối macro I	Tất cả
B (Dự đoán hai chiều)	Bao gồm các khối macro B (mỗi khối hay một vùng khối macro được dự đoán từ danh sách ảnh list 0 hoặc list 1) hoặc là khối macro I	Mở rộng hoặc chính
SP (Switching P)	Tạo điều kiện thuận lợi cho việc chuyển đổi giữa các luồng dữ liệu đã được mã hoá, chứa macroblock loại P hoặc I.	Mở rộng
SI (Switching I)	Tạo điều kiện thuận lợi cho việc chuyển đổi giữa các luồng dữ liệu đã được mã hoá, chứa macroblock loại SI (một dạng đặc biệt của macroblock mã hoá kiểu intra)	Mở rộng



Hình 2.7: Cấu trúc của slice

#### 2.2.4 Macroblock



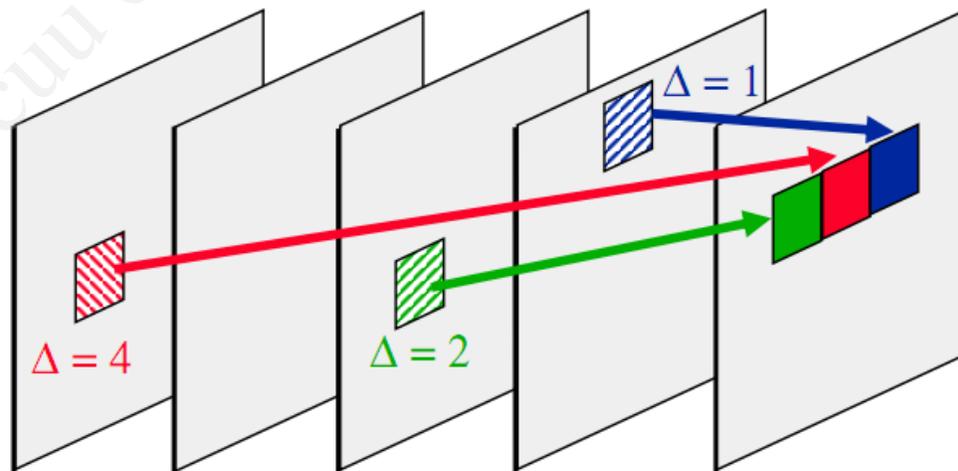
Hình 2.8: MacroBlock

Một khối macro bao gồm các dữ liệu được mã hóa ứng với vùng 16x16 mẫu của khung video. (16x16 mẫu độ chói, 8x8 Cb và 8x8 Cr) và bao gồm các thành phần cú pháp theo bảng ở . Khối macro được đánh số theo thứ tự quét trong khung.

**Bảng 2.2: Các thành phần cấu trúc của Macroblock**

Mb_type	Xác định liệu khối macro là loại I hay P, xác định kích cỡ một vùng trong khối macro
Mb_pred	Xác định chế độ mã hóa trong ( khối macro intra), danh sách tham chiếu List 0 hay List 1 và mã hóa các vector chuyển động khác nhau cho mỗi phần của khối macro
Sub_mp_pred	Xác định kích cỡ khối macro con. Danh sách tham chiếu List 0 hay 1 cho mỗi vùng khối macro và mã hóa các vector chuyển động khác nhau cho mỗi vùng con
Coded_block_pattern	Xác định khối 8x8 nào ( độ chói hay sắc) sẽ mang hệ số biến đổi
Mb_qp_delta	Thay đổi tham số lượng tử hóa
Residual	Hệ số biến đổi đã được mã hóa ứng với mẫu dư thừa sau khi dự đoán

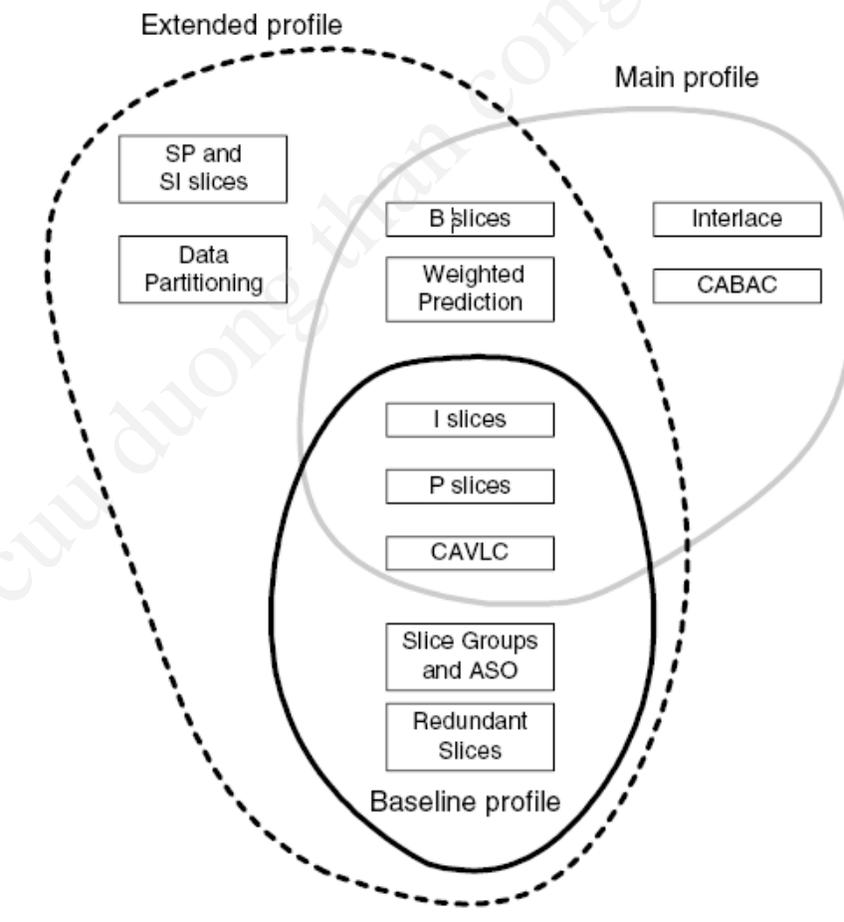
### 2.2.5 Ảnh tham chiếu(Reference Picture)



Hình 2. 9: Ảnh tham chiếu và bù chuyển động

Bộ mã hóa H.264 có thể sử dụng hai hoặc nhiều các ảnh được mã hóa trước đó để làm tham chiếu cho dự đoán bù chuyển động cho mã hóa ngoài các khối macro hoặc phân tách khối macro. Điều này cho phép bộ mã hóa tìm kiếm khối macro giống nhất với khối macro được tách ra từ bức ảnh vừa được mã hóa. Bộ mã hóa và giải mã luôn giữ một hoặc hai danh sách các ảnh tham chiếu, bao gồm ảnh đã vừa được mã hóa hay giải mã (xuất hiện trước hoặc sau ảnh hiện tại). Mã hóa ngoài các khối macro hay vùng của khối macro trong slice P được dự đoán từ một danh sách các ảnh –list 0. Mã hóa ngoài khối macro và vùng các khối macro trong slide B có thể được dự đoán từ hai danh sách list 0 và list 1.

### 2.3 Profile



Hình 2.10: Các profile

H.264 định nghĩa 3 profile trong đó mỗi profile hỗ trợ 1 tập cụ thể các hàm mã hóa và chỉ ra những gì được yêu cầu của bộ mã hóa/giải mã phù hợp với từng profile. Base profile hỗ trợ mã hóa trong và liên khung (sử dụng slice I và slice P) và phương pháp mã hóa entropy CAVLC. Main profile bao gồm video quét xen kẽ, mã hóa liên khung sử dụng slice B, mã hóa liên khung dùng dự đoán có trọng số và phương pháp mã hóa entropy CABAC. Extended profile không hỗ trợ video quét xen kẽ và phương pháp mã hóa entropy CABAC nhưng có thêm chế độ cho phép việc chuyển đổi giữa các luồng bit được mã hóa.

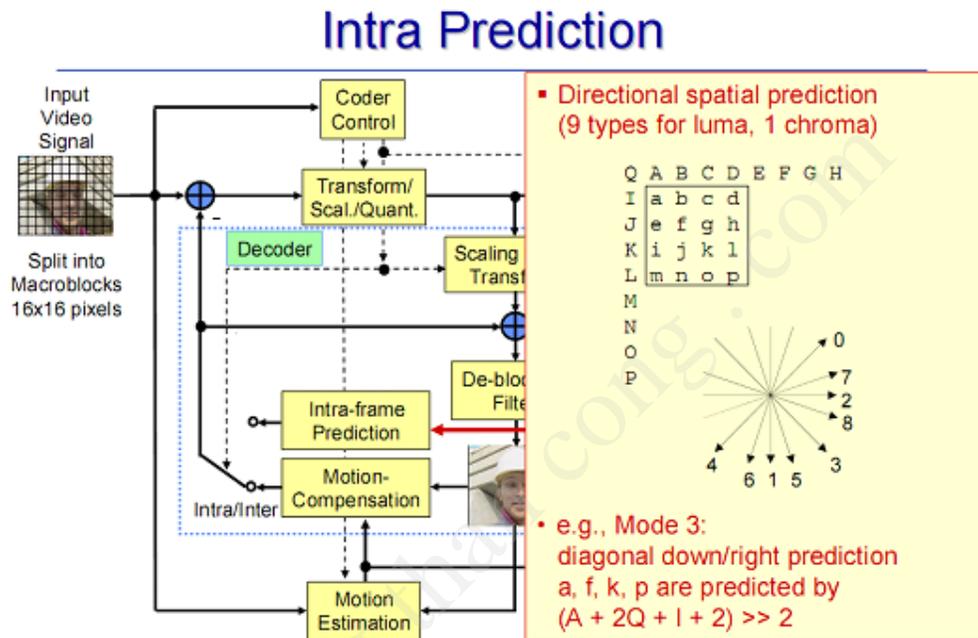
Ứng dụng tiềm năng của profile Baseline bao gồm thoại video, hội thảo truyền hình, và truyền thông không dây. Ứng dụng tiềm năng của Main profile là truyền hình quảng bá và lưu trữ dữ liệu. Profile mở rộng có thể hữu ích trong ứng dụng streaming. Tuy nhiên mỗi profile có sự mềm dẻo đủ để hỗ trợ một loại ứng dụng khác nhau.

## 2.4 Một số kỹ thuật trong H.264

### 2.4.1 Dự đoán nội khung (Intra Prediction)

Trong dự đoán nội khung, các block được dự đoán từ các block trong cùng slice của khung hiện tại. Các kiểu dự đoán intra trong H.264 bao gồm: 9 chế độ dự đoán block intra\_4x4 (kích cỡ 4x4) cho độ chói, 4 chế độ dự đoán block intra\_16x16 (kích cỡ 16x16) cho độ chói, 4 chế độ dự đoán cho 2 đại diện của thành phần màu, cùng với chế độ dự đoán I\_PCM. Trong đó các chế độ dự đoán intra\_4x4 phù hợp với các vùng ảnh có độ phức tạp về các chi tiết cao, các chế độ dự đoán intra\_16x16 phù hợp với các vùng đơn giản của ảnh. Chế độ I\_PCM cho phép bộ mã hóa bỏ qua các tiến trình dự đoán và chuyển đổi thay vào đó bộ mã hóa sẽ gửi trực tiếp các giá trị của các mẫu được mã hóa. I\_PCM được sử dụng cho các trường hợp như: để bộ mã hóa biểu diễn chính xác giá trị của các mẫu có các tham số lượng tử rất nhỏ, cung cấp một cách biểu diễn chính xác các nội dung bất thường của ảnh ... Với các trường hợp đó thì chế độ này hiệu quả hơn là các chế độ intra khác.

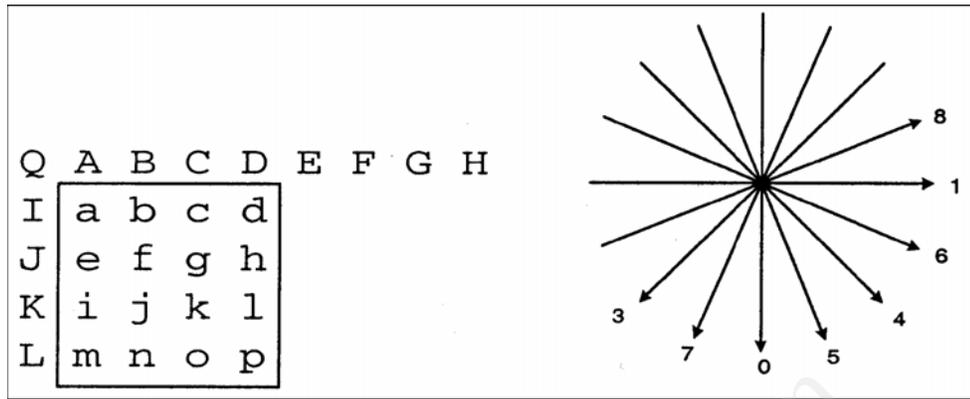
Các chuẩn trước thì dự đoán intra được điều khiển trong miền không gian chuyển đổi, trái lại dự đoán Intra trong H.264 luôn luôn được điều khiển trong miền không gian bằng việc tham chiếu các mẫu bên cạnh đã được mã hóa trước đó (ở bên trái, hoặc phía trên của block được dự đoán).



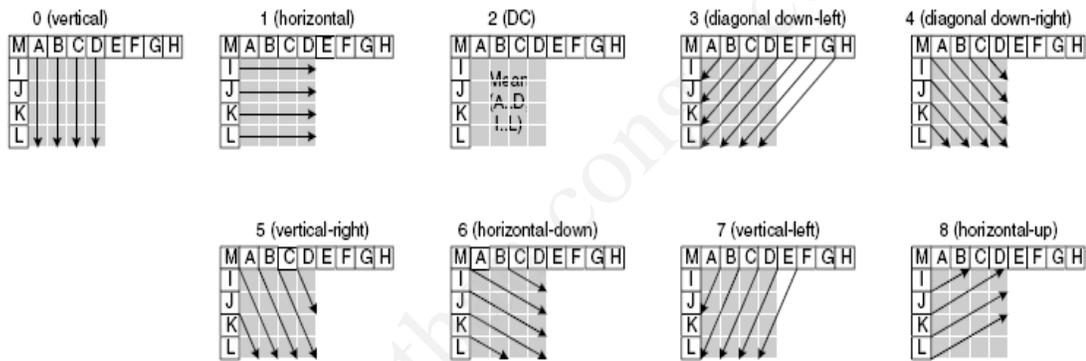
Hình 2.11: Dự đoán nội suy

#### ➤ Chế độ dự đoán intra 4x4

Khi sử dụng chế độ dự đoán intra\_4x4 thì mỗi block 4x4 được dự đoán từ các mẫu bên cạnh. 16 mẫu của block 4x4 được đánh nhãn từ a tới p được dự đoán sử dụng các mẫu của block kề cận được đánh nhãn A-Q (như hình 2.19). Với mỗi block 4x4 thì một trong 9 chế độ dự đoán được sử dụng, các chế độ đó được minh họa ở các hình dưới đây.



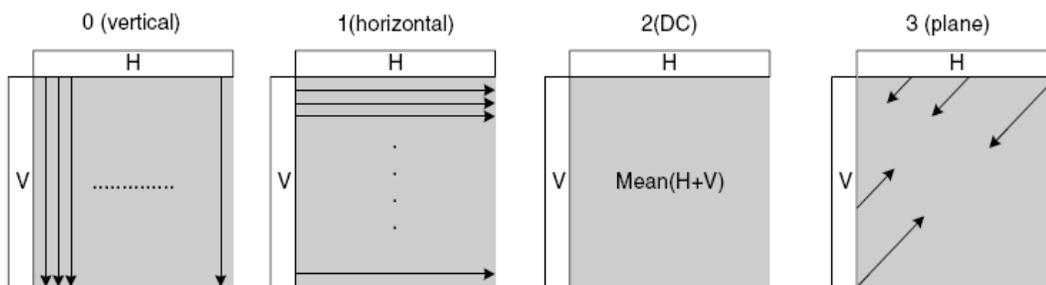
Hình 2.12: Dự đoán intra 4x4



Hình 2.13: Các chế độ dự đoán intra\_4x4

➤ **Chế độ dự đoán Intra\_16x16:**

Khi sử dụng chế Intra\_16x16, tất cả các thành phần độ chói của một MB được dự đoán. Có 4 chế độ dự đoán intra16x16 là dự đoán theo chiều dọc, dự đoán theo chiều ngang, dự đoán DC, dự đoán phẳng như được thể hiện trong hình dưới đây:



Hình 2.14: Các chế độ dự đoán intra\_16x16

➤ **Chế độ dự đoán intra cho các thành phần màu:**

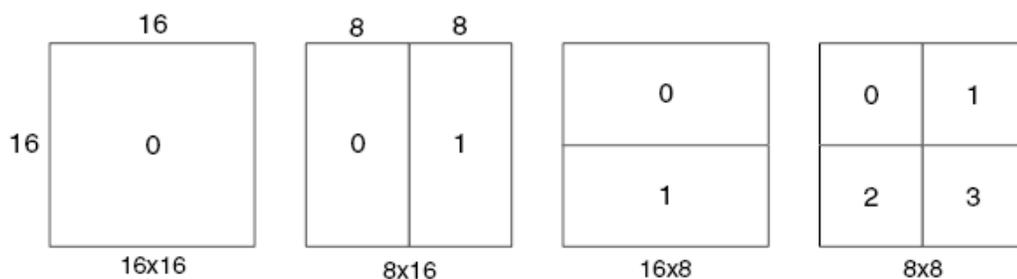
Mỗi thành phần màu 8x8 của một MB được dự đoán từ các mẫu màu được mã hóa trước đó, các thành phần ở trên và bên trái luôn luôn được sử dụng cùng 1 chế độ dự đoán. Dự đoán các thành phần màu có 4 chế độ như dự đoán intra\_16x16 của độ chói.

**2.4.2 Dự đoán liên khung (Inter Prediction)**

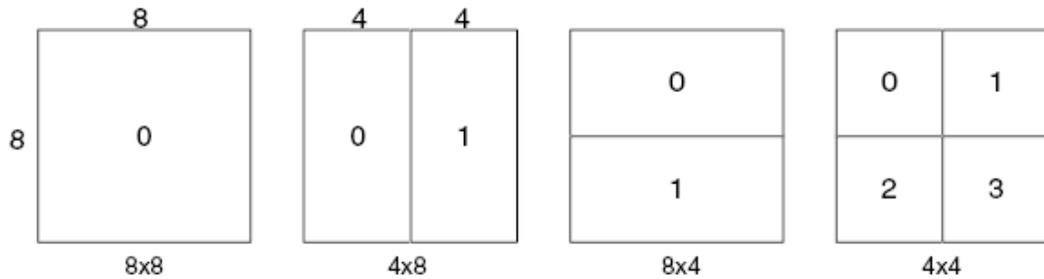
Các dự đoán liên khung trong H.264 cho các block có thể tham chiếu tới nhiều khung đã được mã hóa trước đó với cả 2 chiều (theo thời gian của video nguồn). Sự khác biệt quan trọng nhất với các chuẩn trước đó là H.264 hỗ trợ một dải các kích cỡ block (từ 16x16 cho tới 4x4) và điều chỉnh vector chuyển động chính xác tới độ phân giải  $\frac{1}{4}$  mẫu đối với các thành phần độ chói. Các dự đoán liên khung có thể phân thành dự đoán liên khung trong slice P, dự đoán liên khung trong slice B và dự đoán có trọng số.

➤ **Dự đoán liên khung trong các P slice:**

Mỗi Macroblock P được phân mảnh thành các nhiều dạng block được sử dụng cho dự đoán bù chuyển động. Các phân mảnh đó có các kích cỡ block là 16x16, 16x8, 8x8. Trong các trường hợp phân mảnh với các mẫu 8x8 được truyền phần tử cú pháp chỉ ra liệu phân mảnh 8x8 tương ứng có được phân mảnh nhỏ thành các mảnh nhỏ hơn 8x4, 4x8, 4x4 hay không (với thành phần độ chói và thành phần màu).

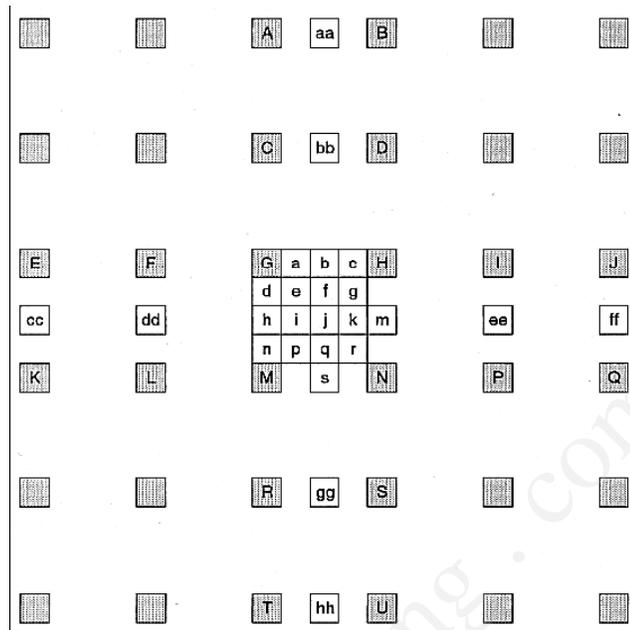


Hình 2.15: Phân mảnh macroblock



Hình 2.16: Phân mảnh submacroblock

Tính hiệu dự đoán cho mỗi block  $M \times N$  của độ chói có được bằng cách thể một vùng của ảnh tham chiếu tương ứng, vùng này được chỉ ra bởi một vecto chuyển động tịnh tiến và chỉ số của ảnh tham chiếu. Ví thế nếu MB này được mã hóa sử dụng 4 phân mảnh  $8 \times 8$  và mỗi phân mảnh  $8 \times 8$  được chia nhỏ hơn thành 4 phân mảnh  $4 \times 4$ , tối đa 16 vecto chuyển động được truyền cho một MB kiểu P. Độ chính xác của vecto chuyển động có thể chỉ tới khoảng cách  $\frac{1}{4}$  của khoảng cách các mẫu. Trong trường hợp vecto chuyển động chỉ tới một vùng mẫu nguyên thì tín hiệu dự đoán bao gồm các mẫu tương ứng của ảnh tham chiếu ngược lại thì mẫu tương ứng được tính bằng cách sử dụng phương pháp nội suy để tính các vị trí vecto chuyển động chỉ tới vị trí không nguyên. Các giá trị dự đoán tại các vị trí  $\frac{1}{2}$  của khoảng cách mẫu có được bằng cách sử dụng bộ lọc FIR một chiều 6-tap theo chiều ngang và chiều dọc. Các giá trị dự đoán tại các vị trí  $\frac{1}{4}$  khoảng cách mẫu mẫu được tạo ra bằng các mẫu trung bình tại các mẫu nguyên và các mẫu có khoảng cách  $\frac{1}{2}$  vừa được tính, như được biểu diễn trong hình 2.24 dưới đây.



Hình 2.17: Tính nội suy cho các chuyển động có khoảng cách không nguyên

Hình 2.24 minh họa nội suy các mẫu phân số cho các mẫu từ a-k và từ n-r. Các mẫu này tại các vị trí  $\frac{1}{2}$  mẫu được đánh nhãn b và h và được tính từ các giá trị trung gian sau:

$$b_1 = E - 5F + 20G + 20H - 5I + J$$

$$h_1 = A - 5C + 20G + 20M - 5R + T$$

Giá trị dự đoán cho các vị trí b và h có được như sau:

$$b = b_1 + 16$$

$$h = h_1 + 16$$

Các mẫu tại các vị trí mẫu  $\frac{1}{2}$  được đánh nhãn là j và được tính thông qua  $j_1$

$$j_1 = cc + 5dd + 20h_1 + 20m_1 - 5ee + ff$$

Trong đó các giá trị trung gian cc, dd, ee,  $m_1$ , ff được tính tương tự như  $h_1$  và giá trị dự đoán j được tính là  $j = j_1 + 512$  và được lấy giá trị trong khoảng 0-255, nếu lớn hơn 255 thì lấy giá trị 255 và nếu nhỏ hơn 0 thì lấy giá trị 0.

Các mẫu tại các vị trí  $\frac{1}{4}$  được đánh nhãn a, c, d, n, f, l, k, và q được tạo ra bằng cách lấy trung bình làm tròn lên của 2 giá trị mẫu gần nhất tại các vị trí mẫu nguyên và mẫu  $\frac{1}{2}$  như sau:

$$a = G + b + 1$$

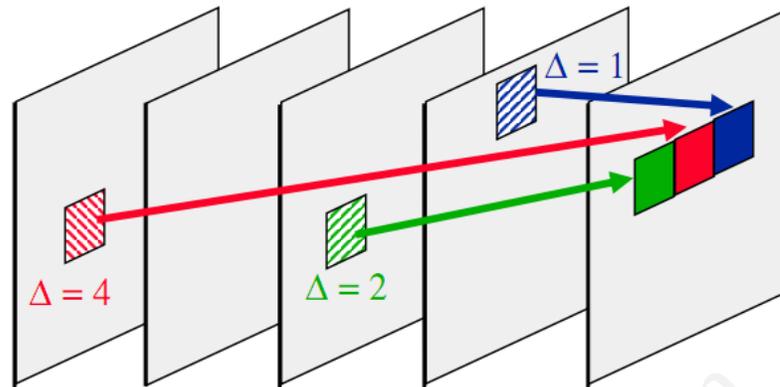
Các mẫu tịa vị trí  $\frac{1}{4}$  được đánh nhãn là e, g, p và r được tạo ra bằng cách lấy trung bình làm tròn lên của 2 mẫu gần nhất của các vị trí mẫu  $\frac{1}{2}$  trong hướng chéo như sau:

$$e = b + h + 1$$

Các giá trị dự đoán cho thành phần màu được tính bằng nội suy song tuyến tính. Do việc lưới lấy mẫu của thành phần màu có độ phân giải thấp hơn lưới lấy mẫu của thành phần độ chói nên các chuyển dịch được sử dụng cho thành phần màu có độ chính xác  $\frac{1}{8}$  mẫu. Dự đoán chuyển động chính xác hơn sử dụng các dự đoán đầy đủ mẫu,  $\frac{1}{2}$  mẫu hoặc  $\frac{1}{4}$  mẫu tiêu biểu cho một trong các cái tiến chính của mã H.264 so với các chuẩn trước đó.

Cú pháp trên (phương pháp trên) cho phép vecto chuyển động chỉ tới tận biên của ảnh, trong trường hợp này khung tham chiếu được ngoại suy ngoài các biên ảnh bằng cách lặp lại các mẫu biên trước khi nội suy. Các thành phần vecto chuyển động được mã hóa vi sai hoặc là sử dụng số trung bình hoặc là sử dụng dự đoán định hướng từ các block bên cạnh. Dự đoán thành phần vecto không chuyển động (hoặc dạng khác của dự đoán) diễn ra qua các biên slice.

Cú pháp này hỗ trợ dự đoán bù chuyển động đa khung như hình dưới đây. Dự đoán bù chuyển động đa khung yêu cầu cả bộ mã hóa và bộ giải mã phải lưu trữ các ảnh tham chiếu được sử dụng cho dự đoán liên khung trong bộ đệm đa ảnh. Bộ giải mã tái tạo bộ đệm đa ảnh của bộ mã hóa theo các hoạt động điều khiển quản lý bộ nhớ được chỉ ra trong dòng bit. Trừ khi kích cỡ của bộ đệm đa ảnh được thiết lập là 1 ảnh thì chỉ số tại đó ảnh tham chiếu được đặt bên trong bộ đệm đa ảnh phải được báo hiệu. Tham số chỉ mục tham chiếu được truyền cho mỗi block bù chuyển động  $16 \times 16$ ,  $16 \times 8$ ,  $8 \times 16$ ,  $8 \times 8$ . Bù chuyển động cho các vùng nhỏ hơn  $8 \times 8$  sử dụng cùng loại chỉ mục tham chiếu cho dự đoán của tất cả các block trong vùng  $8 \times 8$ .



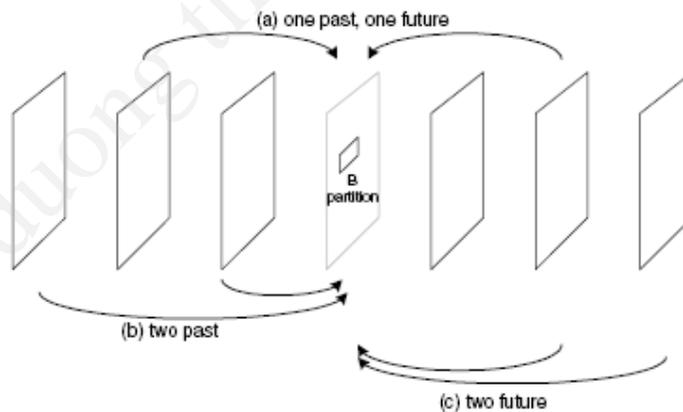
Hình 2.18: Minh họa dự đoán bù chuyển động đa khung

Thêm vào các chế độ bù chuyển động MB được miêu tả ở trên một MB P cũng có thể được mã hóa theo kiểu được gọi là P\_skip. Với kiểu mã hóa này cả báo hiệu lỗi dự đoán lượng tử và vecto chuyển động hoặc tham số chỉ mục tham chiếu đều không được truyền. Tín hiệu khôi phục có được tương tự như tín hiệu dự đoán của kiểu MB P\_16x16 mà tham chiếu ảnh được đặt index 0 trong bộ đệm đa ảnh. Vector chuyển động được sử dụng cho việc khôi phục MB P\_Skip tương tự như bộ dự đoán vecto chuyển động cho block 16x16. Hiệu quả hữu ích của định nghĩa này của kiểu mã hóa P\_skip là các vùng lớn đó không thay đổi hoặc chuyển động hằng số giống như quét chậm có thể được biểu diễn với rất ít bit.

#### ➤ Dự đoán liên khung trong slice B:

Sự khác nhau chủ yếu giữa các slice B và P là slice B được mã hóa theo cách mà trong đó một số MB hoặc các block có thể sử dụng trung bình trọng số của 2 giá trị dự đoán bù chuyển động riêng biệt cho việc xây dựng tín hiệu dự đoán. Các slice B sử dụng các ảnh tham chiếu trong 2 danh sách riêng, 2 danh sách này được tham chiếu như là danh sách ảnh tham chiếu thứ nhất (list 0) và danh sách thứ 2 (list 1) tương ứng. Các ảnh này được đặt thực sự trong mỗi danh sách ảnh tham chiếu là một vấn đề điều khiển bộ đệm đa ảnh và một hoạt động rất giống với các ảnh B MPEG-2 có thể được cho phép nếu muốn bỏ bộ mã hóa. Trong slice B, 4 kiểu khác nhau của dự đoán liên khung được hỗ trợ: list 0, list 1, dự đoán 2 hướng và dự đoán trực tiếp. Với chế độ dự đoán 2 hướng, tín hiệu dự

đoán được thực hiện bằng một trung bình trong số của các tín hiệu dự đoán chuyển động list 0 và list 1. Chế độ dự đoán trực tiếp được tham chiếu từ các phân tử cú pháp được truyền trước đó và có thể hoặc là dự đoán list 0 hoặc là dự đoán list 1 hoặc là dự đoán 2 hướng. Các slice B dùng một phân mảnh MB tương tự như slice P. Bên cạnh các dự đoán  $P_{16 \times 16}$ ,  $P_{16 \times 8}$ ,  $P_{8 \times 16}$ ,  $P_{8 \times 8}$ , các kiểu mã hóa intra, dự đoán 2 chiều một kiểu dự đoán khác được sử dụng là dự đoán trực tiếp. Với mỗi phân mảnh  $16 \times 16$ ,  $16 \times 8$ ,  $8 \times 16$  và  $8 \times 8$  phương pháp dự đoán này (list 0, list 1, 2 hướng) có thể được chọn một cách riêng biệt. Một phân mảnh  $8 \times 8$  của MB B cũng có thể được mã hóa trong chế độ trực tiếp. Nếu không có tín hiệu lỗi dự đoán được truyền cho chế độ MB trực tiếp thì nó cũng có thể xem như chế độ  $B_{\text{skip}}$  và có thể được mã hóa rất hiệu quả như chế độ  $P_{\text{skip}}$  trong slice P. Mã hóa vectơ chuyển động giống như slice P với sự thay đổi thích hợp bởi vì các block bên cạnh có thể được mã hóa sử dụng các chế độ dự đoán khác.



Hình 2.19: Dự đoán liên khung trong slice B

### 2.4.3 Chuyển đổi (transform) và lượng tử.

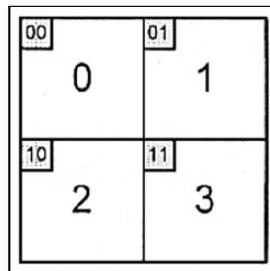
Giống như các chuẩn mã hóa trước, H.264 dùng mã hóa chuyển đổi phần dư dự đoán. Tuy nhiên trong H.264 chuyển đổi này được áp dụng cho các block  $4 \times 4$  và thay vì sử dụng DCT (discrete cosine transform)  $4 \times 4$  một chuyển đổi số tự nhiên có

thể phân tách được với các thuộc tính như DCT 4x4 được sử dụng. Ma trận chuyển đổi đó như sau:

$$H = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 2 & 1 & -1 & -2 \\ 1 & -1 & -1 & 1 \\ 1 & -2 & 2 & -1 \end{bmatrix}.$$

Do chuyển đổi ngược được định nghĩa bởi toán tử tự nhiên chính xác do đó tránh được các sai lệch chuyển đổi ngược. Tiến trình mã hóa chuyển đổi cơ bản rất giống với các chuẩn trước. Tại bộ mã hóa, tiến trình này bao gồm chuyển đổi, quét zig-zag, lượng hóa và làm tròn khi tiến trình lượng tử hóa theo sau bởi mã hóa entropy. Tại bộ giải mã tiến trình ngược của tiến trình mã hóa được thực hiện ngoại trừ làm tròn.

Điều đã được đề cập đến là các chế độ dự đoán Intra\_16x16 và các chế độ dự đoán màu được dành cho việc mã hóa các vùng ảnh mịn, chi tiết đơn giản. Vì lý do này, các hệ số DC bị chuyển đổi lần thứ 2 với kết quả này chúng ta có các hệ số chuyển đổi bao phủ toàn bộ MB. Một chuyển đổi 2x2 bổ xung được áp dụng cho các hệ số DC của các block 4x4 của mỗi thành phần màu. Thủ tục cho một block màu được minh họa trong hình dưới. Các block nhỏ bên trong các block lớn biểu diễn các hệ số DC cho 4 block màu 4x4 được đánh số 0, 1, 2, 3. 2 chỉ số tương ứng với các chỉ số của chuyển đổi Hamadar 2x2 ngược.



Để giải thích ý tưởng đằng sau các chuyển đổi được lặp lại này, phải chỉ ra thuộc tính chung của chuyển đổi 2 hướng của nội dung rất mịn (nơi có tương

quan mẫu xấp xỉ 1). Trong tình huống này, độ chính xác phục hồi tỉ lệ nghịch với kích cỡ một hướng của chuyển đổi. Do đó với vùng rất mịn thì lỗi phục hồi với một chuyển đổi bao phủ block 8x8 bị giảm đi  $\frac{1}{2}$  so với việc chỉ sử dụng chuyển đổi 4x4. Một nhân tố căn bản tương tự có thể được sử dụng cho chuyển đổi thứ 2 được kết nối tới chế độ intra\_16x16. Có nhiều lý do cho việc sử dụng một chuyển đổi kích cỡ nhỏ hơn:

- Cải thiện tiến trình dự đoán cho cả inter và intra. Kết quả này cho tín hiệu phân dư có tương qua ít hơn. Điều này có nghĩa là chuyển đổi ít phải đề ra việc giải tương quan. Điều này cũng có nghĩa rằng một chuyển đổi 4x4 về cơ bản là hiệu quả trong loại bỏ tương quan thống kê như là chuyển đổi lớn.
- Giống như khả năng nén đối tượng, chuyển đổi nhỏ 4x4 có lợi trực quan, tạo ra ít nhiễu xung quanh biên.
- Chuyển đổi nhỏ hơn yêu cầu tính toán ít và việc xử lý độ dài từ nhỏ hơn. Do tiến trình chuyển đổi cho H.264 chỉ bao hàm các phép cộng và phép dịch nên nó cũng chỉ rõ nhiều sai lệch giữa bộ mã hóa và bộ giải mã được tránh (đây là một vấn đề với các chuẩn DCT 8x8 trước).

Một tham số lượng tử được sử dụng cho việc xác định lượng tử của các hệ số chuyển đổi trong H264. Tham số này có 52 giá trị. Các giá trị này được sắp xếp để khi tăng 1 giá trị có nghĩa là tăng kích cỡ lượng tử xấp xỉ 12% (nếu tăng lên 6 đơn vị thì bước lượng tử tăng lên 2 lần). Có thể chú ý rằng sự thay đổi bước lượng tử xấp xỉ 12% có nghĩa nói chung là giảm tốc độ bit xấp xỉ 12%. Các hệ số chuyển đổi lượng tử của một block được quét zig-zag và được truyền đi sử dụng phương pháp mã hóa entropy. Các hệ số DC 2x2 của thành phần màu được quét theo trật tự raster. Tất cả các phép toán chuyển đổi ngược trong H264 có thể được thực hiện chỉ sử dụng phép toán cộng và dịch các giá trị 16bit. Tương tự, các truy cập bộ nhớ 16bit là cần thiết cho việc thực hiện tốt tiến trình chuyển đổi và lượng tử trong bộ mã hóa.

#### 2.4.4 Mã hóa entropy

Trong H264 có 2 phương pháp cho việc mã hóa entropy được hỗ trợ. Phương pháp mã hóa entropy đơn giản hơn sử dụng một bảng từ mã có kích thước không xác định cho tất cả các phần tử cú pháp ngoại trừ các hệ số chuyển đổi lượng tử. Vì thế thay vì thiết kế một bảng VLC cho mỗi phần tử cú pháp, chỉ một ánh xạ tới bảng từ mã đơn được tùy biến theo các thống kê dữ liệu. Bảng từ mã đơn được chọn là một mã exp-Golomb với các thuộc tính toán giải mã rất đơn giản rõ ràng. Để truyền các hệ số lượng tử phương pháp hiệu quả hơn được gọi là CAVLC (context-adaptive variable length coding) được sử dụng. Trong mô hình H264, bảng VLC cho nhiều phần tử cú pháp được chuyển mạch phụ thuộc và phần tử cú pháp đã được truyền. Do các bảng VLC được thiết kế để phù hợp với các thống kê có điều kiện tương ứng, thực thi mã hóa entropy được cải thiện so với các mô hình sử dụng bảng VLC đơn.

Trong phương pháp mã hóa entropy CAVLC, số lượng hệ số lượng tử 0 và kích cỡ thực và vị trí của các hệ số được mã hóa một cách riêng rẽ. Sau khi quét zig-zag các hệ số lượng tử, phân bố thống kê của chúng chỉ ra các giá trị lớn cho phần tần số thấp giảm các giá trị nhỏ sau đó trong việc quét phần tần số cao.

#### 2.4.5 Bộ lọc Deblocking

Cấu trúc khối cơ bản của H.264 là 4x4 cho biến đổi và bù chuyển động. Do vậy, ảnh sẽ hình thành các đường biên giữa các khối. Bộ lọc sẽ triệt tiêu các biên này để hình ảnh được tự nhiên hơn.



Hình 2.20: Ví dụ về một khung hình sử dụng bộ lọc Deblocking

Trong hình 2.27, ảnh bên trái là sử dụng bộ lọc Deblocking nhìn ảnh không bị có cấu trúc ô vuông như ảnh bên phải không sử dụng Deblocking.

Một bộ lọc được sử dụng cho mỗi macroblock được mã hóa nhằm làm giảm méo. Bộ lọc deblocking được sử dụng sau biến đổi ngược ở bộ mã hóa (trước khi tái tạo lại và lưu trữ) và ngược lại đối với bộ giải mã. Bộ lọc giúp cải thiện chất lượng hình ảnh. Hình ảnh được lọc sẽ sử dụng để làm dự đoán bù chuyển động cho các hình ảnh trong tương lai và điều này có thể giúp cải thiện chất lượng nén bởi vì những hình ảnh đã được lọc chân thực hơn so với các ảnh không được lọc.

- **In-loop Deblocking**: Dựa trên kết quả của bù chuyển động và biến đổi mã hóa nó giúp loại bỏ dư thừa về mặt không gian.



## 2.4.6 Điều khiển tốc độ trong H.264 (Rate Control)

### 2.4.6.1 Giới thiệu về điều khiển tốc độ trong H.264

Một bộ mã hoá thực thi việc điều khiển tốc độ như là một công cụ để thay đổi đặc tính tốc độ bit của luồng bit đã được mã hoá nhằm mục đích tạo ra các khung hình mã hoá chất lượng cao ở tốc độ bit mong muốn. Chuẩn mã hoá H.264 được sử dụng chủ yếu cho việc truyền video tại những khu vực mà băng thông hoặc khả năng lưu trữ hạn chế. Một thuật toán điều khiển tốc độ thay đổi các tham số mã hoá, chủ yếu là hệ số lượng tử hoá QP để đạt được tốc độ bit mong muốn. Nó cấp phát một lượng bit cho một nhóm ảnh, một ảnh đơn hoặc một sub-picture (ảnh con)

trong một chuỗi video. Điều khiển tốc độ cho phép việc chọn lựa các tham số mã hoá để đạt được chất lượng video tốt nhất trong điều kiện ràng buộc về tốc độ bit và bộ đệm bên giải mã video.

Điều khiển tốc độ trong H.264 có thể chia làm 3 mức chính:

- mức nhóm ảnh.
- mức ảnh.
- mức macro-block.

Ở tại mỗi mức, thuật toán điều khiển tốc độ sẽ chọn giá trị lượng tử hoá QP- giá trị này quyết định quá trình lượng tử hoá của các hệ số chuyển đổi. Điều khiển tốc độ trong H.264 bao gồm 5 mode khác nhau, 1 mode two-pass (tạm dịch mã hoá 2 lần), 4 mode 1-pass (mã hoá 1 lần).

#### 2.4.6.2 Chế độ ratecontrol trong x264

##### a) Constant QP

X264 khi chạy ở chế độ này thì dùng 1 giá trị QP không đổi trong suốt quá trình mã hóa ảnh. Tùy thuộc loại khung ảnh (I,P,B) mà có giá trị QP tương ứng. Giá trị mặc định là 26, khi tăng giá trị QP thì tỉ số nén sẽ tăng lên đồng thời làm cho tốc độ bit đầu ra giảm đi nhưng đồng thời chất lượng ảnh sẽ giảm đi do dữ liệu bị mất đi so với ảnh gốc tăng lên. Nếu bạn mong muốn 1 tốc độ bit cụ thể thì không dùng được chế độ này do không thể dự đoán được kích thước cuối cùng của đầu ra

##### b) Constant rate factor (CRF)

X264 khi chạy ở chế độ này thì dùng 1 giá trị QP trung bình trong suốt quá trình mã hóa ảnh. X264 sẽ tính toán sự khác biệt giữa ảnh chuyển động nhanh và ảnh chuyển động chậm. Ảnh chuyển động sẽ được mã hóa với lượng bit nhiều hơn còn ảnh chuyển động chậm thì ngược lại. Như vậy chế độ CRF cho chất lượng hình ảnh bằng với chế độ CQP nhưng có tỉ số nén cao hơn. Khoảng giá trị của crf từ 18- 26 cho chất lượng tốt với kích thước file không quá lớn

## Các đặc điểm nổi bật trong thiết kế của H.264

Các thiết kế làm cho H.264 tạo ra sự thân thiện với các mạng:

- Cấu trúc dữ liệu sau mã hóa thành các gói dữ liệu logic được gọi là NALU (Network abstract layer unit). Cấu trúc cú pháp NALU cho phép tùy biến lớn hơn cho các mạng khác nhau.
- Kích cỡ slice mềm dẻo: Không giống như cấu trúc slice cứng nhắc trong MPEG-2 các kích cỡ slice trong H.264 rất mềm dẻo để đáp ứng cho các yêu cầu về phân vùng ảnh, trễ mã hoá, dự đoán chuyển động trong các ứng dụng khác nhau.
- Sắp thứ tự các Macroblock mềm dẻo (FMO): là một khả năng mới trong H.264 để phân mảnh hình thành các vùng được gọi là các slice group. Mỗi slice sẽ trở thành một tập con của một slice group, các slice có thể giải mã độc lập được. Khi sử dụng hiệu quả FMO, có thể làm tăng độ vững chắc một cách đáng kể cho các tổn thất dữ liệu bằng cách quản lý mối quan hệ không gian giữa các vùng được mã hoá trong mỗi slice.
- Sắp thứ tự các slice tùy ý (ASO): Do mỗi slice của một ảnh có thể được giải mã một cách độc lập với các slice khác của ảnh, thiết kế H.264 cho phép gửi và nhận các slice này của ảnh theo bất cứ thứ tự nào tương ứng với chúng. Khả năng này được đề cập đầu tiên trong phần bổ sung của H.263+ để cải thiện trễ end to end trong các ứng dụng thời gian thực đặc biệt khi được sử dụng trên các mạng có trật tự phân phát dữ liệu tự do như các mạng gói, mạng internet.

Các thiết kế làm cho H.264 có độ nén cao hơn mà vẫn đảm bảo chất lượng tương đương với các chuẩn khác trước nó. Không có phần tử nào trong VLC là thành phần chủ yếu cho việc cải thiện đáng kể hiệu quả nén so với các chuẩn trước mà hiệu quả nén được tăng lên đáng kể là do nhiều cải thiện nhỏ trong bộ mã hóa H.264 tạo nên:

➤ Kích cỡ block bù chuyển động có thể thay đổi nhỏ hơn với các kích cỡ block nhỏ: H.264 hỗ trợ linh hoạt hơn trong việc lựa chọn các kích cỡ bù chuyển động hơn các chuẩn trước đó với kích cỡ block bù chuyển động tối thiểu cho độ chói là  $4 \times 4$ .

➤ Bù chuyển động chính xác tới  $\frac{1}{4}$  mẫu: Hầu hết các chuẩn mã hoá trước cho phép vecto chuyển động chính xác tới  $\frac{1}{2}$  mẫu. H.264 cải thiện độ chính xác của vecto chuyển động tới  $\frac{1}{4}$  mẫu, tính năng này kế thừa từ tính năng nâng cao của MPEG-4 Visual (part 2) nhưng H.264 đã giảm độ phức tạp trong tiến trình nội suy hơn so với MPEG-4 Visual (part 2).

➤ Vecto chuyển động có thể vượt quá các biên của ảnh: Trong khi vecto chuyển động trong MPEG-2 các chuẩn trước đó yêu cầu chỉ ra vùng trong hình than chiếu cho việc dự đoán chuyển động, H.264 kế thừa kỹ thuật ngoại suy biên hình trong H.263 để có thể tăng hiệu quả dự đoán chuyển động.

➤ Bù chuyển động với nhiều ảnh tham chiếu: Các ảnh dự đoán mã hoá trong MPEG-2 và các chuẩn trước chỉ sử dụng 1 ảnh trước đó để dự đoán các giá trị trong ảnh đầu vào. Thiết kế mới này mở rộng từ kỹ thuật lựa chọn ảnh tham chiếu trong tính năng nâng cao của H.263++ có thể mã hoá hiệu quả hơn bằng cách cho phép bộ mã hoá sử dụng nhiều ảnh tham chiếu để dự đoán chuyển động. Song song với khả năng sử dụng nhiều ảnh tham chiếu là việc áp dụng dự đoán chuyển động 2 hướng (sử dụng cho các khung B)

➤ Tách biệt mối quan hệ giữa thứ tự tham chiếu khỏi thứ tự hiển thị: trong các chuẩn trước có một sự phụ thuộc rất chặt chẽ giữa trật tự ảnh cho các mục đích tham chiếu bù chuyển động và trật tự của ảnh cho mục đích hiển thị. Trong H.264, các giới hạn này bị loại bỏ cho phép bộ mã hoá chọn trật tự các ảnh cho việc tham chiếu và cho mục đích hiển thị ở mức độ mềm dẻo cao chỉ ràng buộc bởi giới hạn về tổng dung lượng bộ nhớ được sử dụng để đảm bảo khả năng giải mã. Việc loại bỏ giới hạn này cũng cho phép loại bỏ trễ phụ được kết hợp với mã hoá hai chiều trước đó.

➤ Tách biệt phương pháp biểu diễn ảnh với khả năng tham chiếu ảnh: Trong các chuẩn trước đó các bức ảnh được tham chiếu 2 chiều (ảnh B) không được sử dụng cho việc dự đoán các ảnh khác trong chuỗi video. H.264 loại bỏ sự ràng buộc này làm cho việc mã hóa mềm dẻo hơn.

➤ Dự đoán có trọng số: làm một cải tiến mới trong H.264 cho phép tín hiệu dự đoán bù chuyển động được đặt trọng số và bù bằng những lượng nhất định được chỉ ra bởi bộ mã hoá. Điều này có thể cải thiện hiệu quả mã hoá rất lớn cho các cảnh mờ và có thể được sử dụng mềm dẻo cho các mục đích khác nhau.

➤ Cải thiện suy luận chuyển động trong vùng mã hóa trực tiếp và loại bỏ (skipped và direct): Trong các chuẩn mã hóa trước các vùng sử dụng mã hóa skipped và direct thì không chuyển động trong nội dung cảnh. Điều này tạo ra ảnh hưởng xấu khi video mã hóa có chứa chuyển động chung. H.264 được thiết kế có thể suy luận chuyển động trong các vùng skipped và direct này (bù chuyển động direct kế thừa từ MPEG-4).

➤ Dự đoán không gian trực tiếp cho mã hóa intra: Một kỹ thuật mới của việc ngoại suy các biên của những phần được mã hoá trước của ảnh hiện tại được áp dụng trong các vùng ảnh được mã hoá intra. Điều này cải thiện chất lượng của tín hiệu dự đoán và cũng cho phép dự đoán từ các vùng lân cận mà chưa được mã hoá sử dụng mã hoá intra (Điều này là không thể khi sử dụng phương pháp dự đoán miền transform trong MPEG-4 Visual).

➤ Chuyển đổi miền không gian kích cỡ block nhỏ: Các chuẩn mã hoá trước sử dụng kích cỡ block chuyển đổi là 8x8 trong khi thiết kế mới của H.264 cơ bản dựa trên chuyển đổi 4x4. Điều này cho phép bộ mã hoá biểu diễn các tín hiệu trong một hình dạng có sự tương thích cục bộ cao hơn làm giảm các thành phần lạ. Các block nhỏ cũng chỉ được điều chỉnh một cách chừng mực bằng việc nâng cao khả năng dự đoán nội dung của video sử dụng kỹ thuật này tốt hơn.

➤ Chuyển đổi không gian block phân cấp: Trong khi hầu hết các trường hợp sử dụng kích cỡ block chuyển đổi nhỏ 4x4 là có cảm giác là có ích thì có

một số vùng tín hiệu có đủ tương quan để có thể sử dụng mã hóa block lớn hơn. H.264 cho phép 2 cách: 1) bằng việc sử dụng chuyển đổi phân cấp để mở rộng kích cỡ block hiệu quả sử dụng cho thông tin màu tần số thấp với ma trận  $8 \times 8$ . 2) bằng cách cho phép bộ mã hoá lựa chọn một kiểu mã hoá đặc biệt cho mã hoá intra, cho phép mở rộng độ dài của chuyển đổi độ chói cho thông tin tần số thấp tới kích cỡ block  $16 \times 16$  và tương tự với thành phần tín hiệu đại diện cho tín hiệu màu cũng vậy.

- Mã hóa thuật toán entropy: phương pháp mã hóa entropy nâng cao là mã hóa thuật toán được đưa vào bộ mã hóa H.264 để tạo ra khả năng mã hóa entropy mạnh hơn đó là mã hóa CABAC trong khi trong H.263 chỉ là một tính năng mở rộng.

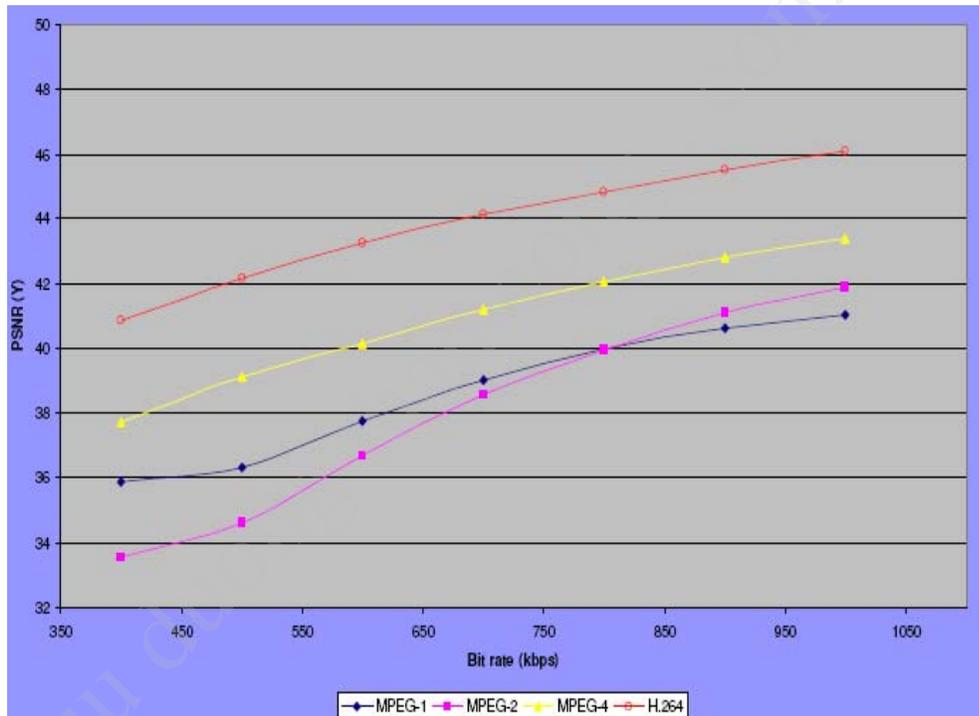
- Context-adaptive entropy coding: 2 phương pháp mã hóa entropy được áp dụng trong H.264 là CAVLC (context-adaptive variable-length coding) và CABAC. Cả 2 phương pháp này đều sử dụng tương thích nội dung để cải thiện hiệu quả thực hiện mã hóa so với các chuẩn trước.

- Data Partitioning: Do một số thông tin biểu diễn cho các vùng (vecto chuyển động và các thông tin dự đoán khác) quan trọng hơn nhiều thông tin khác biểu diễn nội dung video do đó H.264 cho phép cú pháp của mỗi slice được tách làm 3 phần khác nhau cho việc truyền, phụ thuộc vào sự phân nhóm các phần tử cú pháp. Tính năng này kế thừa từ MPEG-4 và phân mở rộng của H.263++.

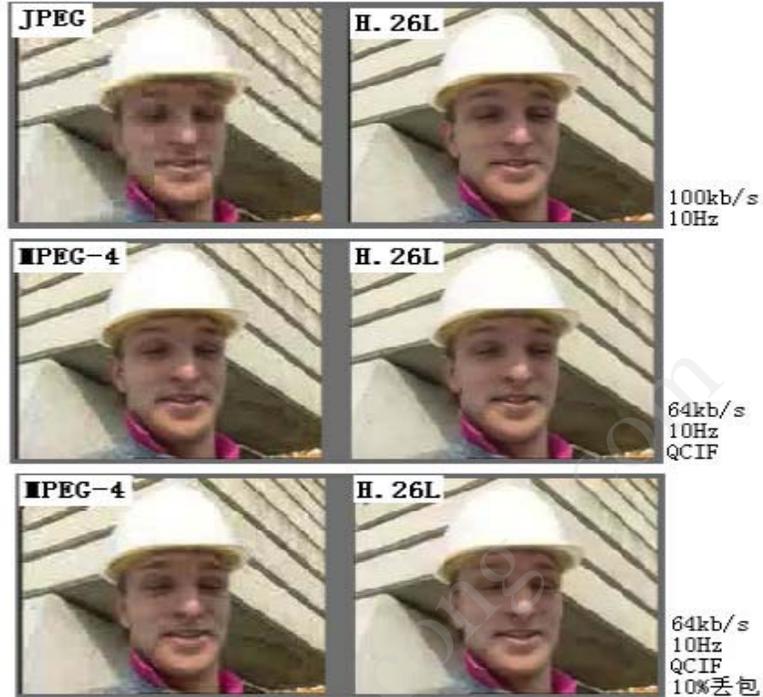
- Bù chuyển động có kích cỡ block thay đổi: H.264 hỗ trợ mềm dẻo hơn trong việc lựa chọn kích cỡ bù chuyển động bù chuyển với kích cỡ nhỏ nhất là các block có kích cỡ  $4 \times 4$  điểm ảnh.

- Bù chuyển động chính xác tới  $\frac{1}{4}$  pixel: hầu hết các chuẩn trước chỉ cho phép vecto chuyển động chính xác tới  $\frac{1}{2}$  pixel thì H.264 kế thừa một tính năng mở rộng của MPEG-4 với độ chính xác của vecto chuyển động lên tới  $\frac{1}{4}$  pixel nhưng giảm sự phức tạp của tiến trình nội suy so với chuẩn trước đó.

➤ Vecto chuyển động có thể tới tận các biên của ảnh. Trong khi vecto chuyển động của MPEG-2 và các chuẩn mã hóa trước đó yêu cầu chỉ ra vùng trong khung ảnh tham chiếu được mã hóa trước cho dự đoán chuyển động thì H.264 kế thừa kỹ thuật ngoại suy biên ảnh của H.263 (đây là một tính năng mở rộng của H.263).



*PSNR của H.264 so chuẩn MPEG*



H.264 so với một số chuẩn nén khác.

## **CHƯƠNG 3:**

### **CÁC ỨNG DỤNG CỦA CHUẨN NÉN H.264**

#### **3.1 Ưu điểm nổi bật và các ứng dụng của chuẩn nén H.264**

##### **3.1.1. Ưu điểm của nén không gian**

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.

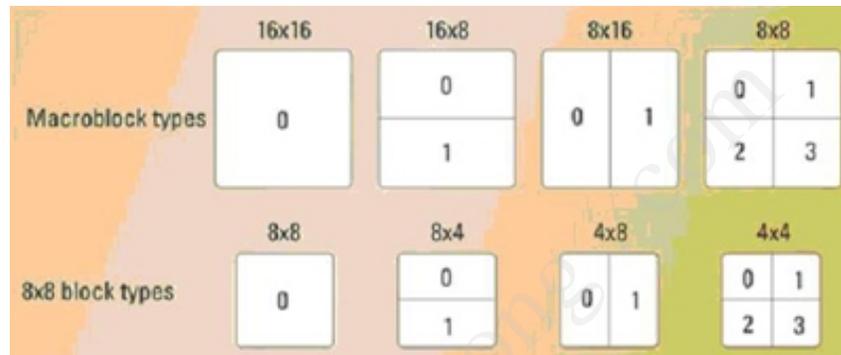
##### **3.1.2. Ưu điểm của nén thời gian**

Điểm cải tiến lớn nhất ở MPEG-4 AVC là chế độ mã hoá giữa. Những phương pháp tiên tiến ở chế độ này khiến cho nén thời gian đạt đến một cấp độ cao hơn nhiều, cùng với chất lượng chuyển động tốt hơn so với các chuẩn MPEG trước đây.

##### **3.1.3. Kích cỡ khối**

Ở chế độ giữa khối, MPEG-2 chỉ hỗ trợ các macroblock 16x16 điểm ảnh, không đủ độ phân giải để mã hoá chính xác các chuyển động phức tạp hoặc phi tuyến tính, ví dụ như phóng to thu nhỏ. Ngược lại, MPEG-4 AVC lại tăng cường hiệu chỉnh chuyển động bằng cách cho phép bộ lập mã biến đổi kích cỡ thành phần

chói của mỗi macroblock. (Bộ lập mã sử dụng thành phần chói như vậy là do mắt người nhạy cảm với chuyển động chói hơn nhiều so với chuyển động màu.) Như có thể thấy trong Hình 2, MPEG-4 AVC có thể chia thành phần chói của từng macroblock thành 4 cỡ: 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 24. MPEG-4 AVC có thể phân chia thành phần chói của từng MacroBlock theo nhiều cách để tối ưu hoá việc 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. 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%. MPEG-4 AVC lấy phần chói của ảnh gốc và sử dụng các macroblock đã được chia nhỏ tại các khu vực có nhiều chi tiết nhằm tăng cường khả năng hiệu chỉnh chuyển động.

**\* Độ chính xác trong hiệu chỉnh chuyển động:**

Trong đa số trường hợp, chuyển động tại rìa mỗi macroblock hay khối thường diễn ra với độ phân giải nhỏ hơn một điểm ảnh. Do vậy, chuẩn nén MPEG-4 AVC có thể đảm bảo độ chính xác trong hiệu chỉnh chuyển động lên tới 1/4 hoặc 1/8 điểm ảnh, trong khi các chuẩn MPEG trước đây chỉ dừng lại ở mức 1/2 điểm ảnh. Khả năng đạt mức chính xác 1/8 ảnh điểm của MPEG-4 AVC giúp tăng hiệu suất mã hoá tại tốc độ bit cao và độ phân giải video cao. Các thử nghiệm cho thấy độ chính xác đến 1/4 điểm ảnh có thể làm giảm tốc độ bit xuống hơn 15% so với độ chính xác 1 điểm ảnh.

**\* 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, giống như trong trò kéo quân. 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.

**\* Giải khối tích hợp:**

Video số sau khi nén thường tạo ra một hiệu ứng gọi là "kết khối", có thể thấy rõ tại điểm giao nhau giữa các khối, đặc biệt là khi có tốc độ bit thấp. Hiệu ứng này là do công đoạn xử lý sử dụng nhiều loại chuyển động và bộ lượng tử khác nhau. Đối với MPEG-2, cách duy nhất để ngăn chặn hiệu ứng này là sử dụng các cơ chế hậu xử lý phù hợp, tuy nhiên các cơ chế này lại không tương thích được với tất cả các máy thu. Chuẩn nén MPEG-4 AVC đưa vào sử dụng một bộ lọc giải khối hoạt động ở hai cấp độ: macroblock 16x16 và khối 4x4. Việc giải khối thường tạo ra một tỉ số tín hiệu trên nhiễu (PSNR) cực điểm thấp hơn, tuy nhiên nhìn một cách chủ quan thì nó tạo ra hình ảnh chất lượng tốt hơn.

### 3.1.4. Ưu điểm về lượng tử hoá và biến đổi.

Chấm di động 8x8 DCT cùng với dung sai của lỗi làm tròn chính là phần cốt lõi của các chuẩn MPEG trước đây. MPEG-4 AVC độc đáo hơn ở chỗ nó sử dụng biến đổi không gian nguyên (gần giống như DCT) đối với các khối 4x4 điểm ảnh. Kích cỡ nhỏ giúp giảm bớt hiện tượng "kết khối", trong khi thông số nguyên tuyệt đối giúp loại bỏ nguy cơ không thích ứng giữa bộ lập mã và giải mã trong phép biến đổi ngược. Thêm vào đó, dãy hệ số xích lượng tử lớn hơn khiến cho cơ chế kiểm soát tốc độ dữ liệu ở bộ lập mã hoạt động một cách linh hoạt hơn dựa trên một tỉ lệ phức hợp vào khoảng 12,5% thay cho một mức tăng lượng gia không đổi.

### 3.1.5. Ưu điểm đối với mã hoá entropy.

Sau khi tiến hành hiệu chỉnh, biến đổi và lượng tử hoá chuyển động, các bộ lập mã MPEG trước đây sẽ vạch ra các symbol biểu diễn vectơ chuyển động và hệ số đã lượng tử hoá thành các bit thực sự. Ví dụ như chuẩn nén MPEG-2 sử dụng phương pháp mã có chiều dài biến thiên tĩnh (VLC) không thể tối ưu hoá trong môi trường video thời gian thực (trong đó nội dung và các cảnh biến đổi theo thời gian).

MPEG-4 AVC sử dụng mã hoá thuật toán nhị phân theo tình huống CABAC (Context-Adaptive Binary Arithmetic Coding). Hiệu suất mã hoá của CABAC cao hơn hẳn nhờ khả năng thích nghi với các thay đổi có thể xảy ra trong phân bố symbol. Ví dụ, nó có thể khai thác sự tương quan giữa các symbol và từ đó sử dụng sự tương quan bit và thuật toán mã hoá. Cơ chế này có thể giúp tiết kiệm thêm một lượng bit vào khoảng hơn 5%.

### 3.1.6. Các ứng dụng của chuẩn nén H.264

Với những ưu điểm về mặt kĩ thuật như trình bày ở trên, H.264 được ứng dụng rộng rãi trong các ứng dụng sau :

1) Về mặt chất lượng hình ảnh : H.264 là chuẩn nén sử dụng công nghệ âm thanh, hình ảnh mới khả năng nén tốt hơn so với các chuẩn nén trước đó. Do đó, chuẩn nén này cho phép cung cấp dịch vụ phân phát hình ảnh chất lượng cao qua mạng băng thông giới hạn.

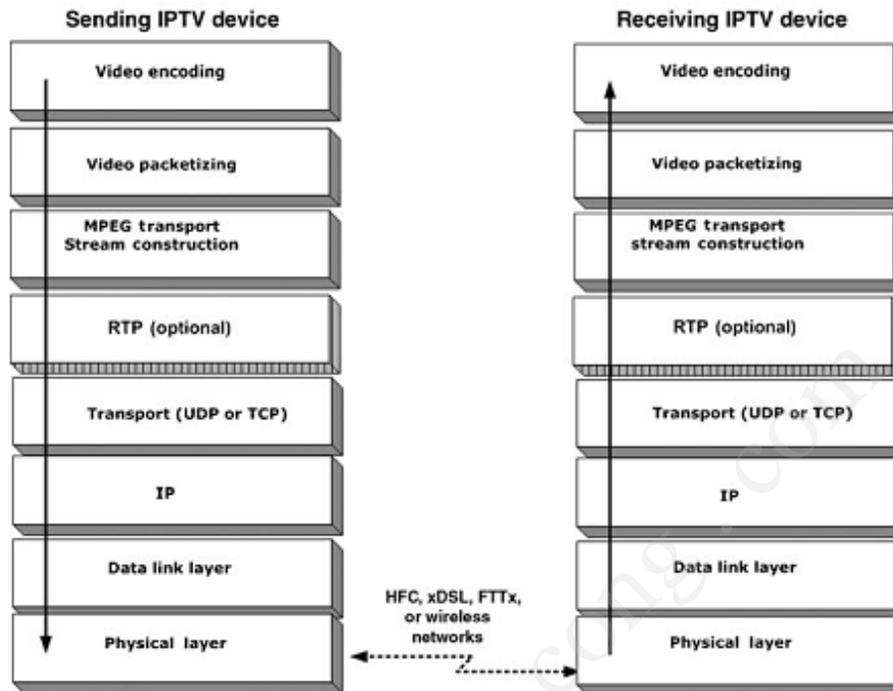
- 2) Yêu cầu băng thông thấp: Chất lượng hình ảnh của H.264 gần giống với MPEG-2 nhưng H.264 cần ít băng thông để truyền tải tín hiệu với cùng chất lượng. Đặc điểm này rất phù hợp để sử dụng trong hệ thống IPTV.
- 3) Có khả năng kết hợp với các thiết bị xử lý video có sẵn như MPEG-2 và hạ tầng mạng dựa trên IP đã có sẵn .
- 4) Hỗ trợ truyền hình độ phân giải cao: Khi sử dụng tối ưu chuẩn nén có thể làm tăng khả năng truyền dữ liệu của mạng. Do đó các nhà cung cấp dịch vụ truyền thông có thể sử dụng chuẩn nén này để cung cấp chương trình video độ phân giải cao qua mạng sẵn có.
- 5) Hỗ trợ nhiều ứng dụng : Chuẩn nén H.264 được sử dụng trong nhiều ứng dụng, với nền khác nhau thì có những yêu cầu riêng. Ví dụ, ứng dụng truyền đa điểm trong IPTV yêu cầu phải hiển thị hình ảnh ở dạng chuẩn truyền hình, trong khi, đối với các ứng dụng giải trí di động, hình ảnh phải hiển thị được trên các thiết bị di động. Để phù hợp với mọi ứng dụng, chuẩn nén H.264 có rất nhiều profile và level. Đặc điểm của profile và level là tốc độ bit và kích thước ảnh.
- 6) Có thể truyền độc lập: Chuẩn nén H.264 có thể truyền qua nhiều giao thức như ATM, RTP, UDP, TCP và các dòng MPEG-2.
- 7) Dễ dàng thích nghi với các mạng chất lượng kém nhờ cơ chế sửa lỗi .

Như vậy, từ những đặc điểm và ưu điểm của H.264 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. Với các kỹ thuật này, H.264 có thể giảm tốc độ bit xuống hơn 50% so với chuẩn MPEG-2. Do đó, H.264 có thể được lựa chọn để ứng dụng trong IPTV.

## **3.2 Ứng dụng của H.264 trong mô hình IPTV**

### **3.2.1 Mô hình truyền thông của IPTV**

Hình 3.1 cho ta thấy nội dung video đã được nén qua các lớp như thế nào khi chuyển từ lớp trên xuống lớp dưới trong mô hình IPTV.



Hình 3.1: Mô hình truyền thông IPTV

### 3.2.2 Lớp mã hóa video

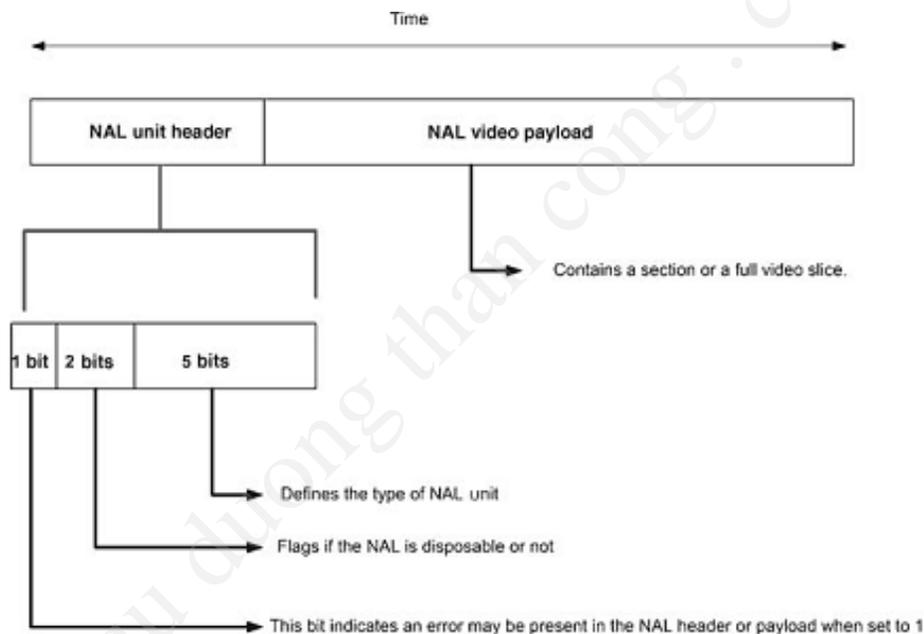
Quá trình truyền thông bắt đầu ở lớp mã hóa (video coding), các tín hiệu tương tự hoặc số được nén. Tín hiệu đầu ra của bộ nén là các dòng MPEG cơ bản. Các dòng MPEG cơ bản được định nghĩa là các tín hiệu số liên tục thời gian thực. Có nhiều loại dòng cơ bản. Ví dụ âm thanh được mã hóa sử dụng MPEG được gọi là “dòng cơ bản âm thanh”. Một dòng cơ bản thực ra chỉ là tín hiệu ra thô từ bộ mã hóa. Các dòng dữ liệu được tổ chức thành các khung tại lớp này. Các thông tin chứa trong một dòng cơ bản có thể bao gồm:

- Loại khung và tốc độ
- Vị trí của những block dữ liệu trên màn hình
- Tỷ số cạnh

Điều quan trọng phải chú ý là lớp này được chia thành 2 lớp phụ theo đặc tính của chuẩn H.264/AVC: lớp mã hóa video(VCL) và lớp trừu tượng(NAL). Lớp phụ

VCL quan tâm tới việc nén các nội dung video. Tín hiệu đầu ra của lớp này là chuỗi các slice ảnh. Dòng bit ở lớp NAL được tổ chức thành các gói rời rạc được gọi là các khối NAL. Định dạng của các khối NAL được mô tả trong hình 3.2

Các khối trong hình 3.2 mô tả 1 khối NAL với phần payload nội dung video. Nó có thể chứa các loại payload khác trong phần này như thông tin điều khiển. Những khối như thế được xếp vào loại không phải khối VCL (non-VCL unit). Các khối NAL được kết hợp với nhau thành chuỗi, định dạng nên khối truy cập. chú ý rằng khối NAL được tạo nên từ chuẩn H.264/AVC có thể hỗ trợ cho cả cấu trúc mạng dựa trên giao thức IP và các mạng không dựa trên giao thức IP.

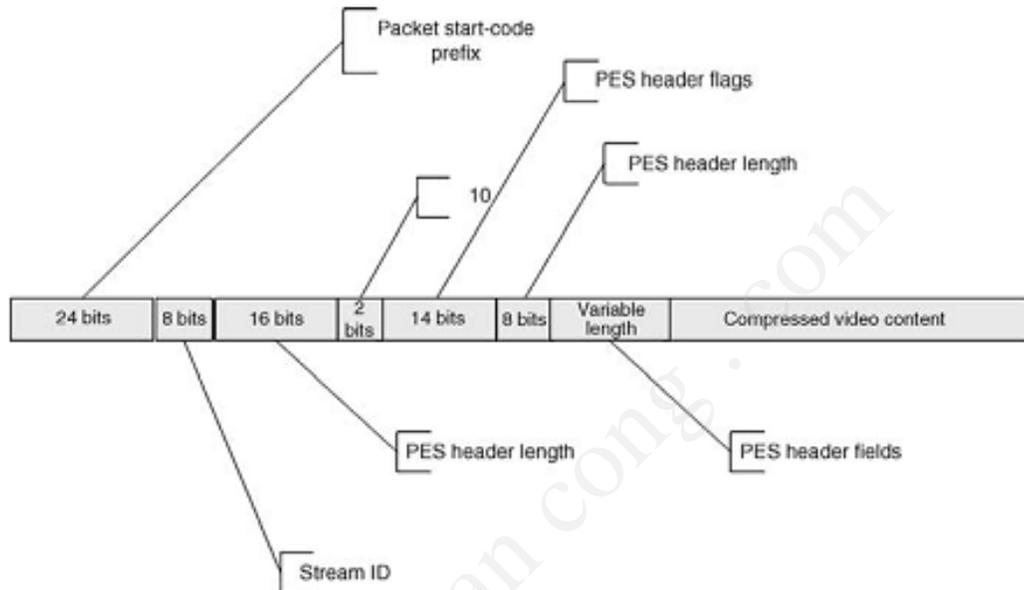


Hình 3.2: Cấu trúc của khối NAL

### 3.2.3 Lớp đóng gói Video

Để truyền các dòng cơ bản âm thanh, dữ liệu và hình ảnh qua mạng số, mỗi dòng cơ bản này phải được chuyển đổi sang một dòng được chèn của gói PES đã được đánh dấu thời gian (PES- Parketized Element Stream ). Một dòng PES chỉ bao gồm 1 loại dữ liệu từ 1 nguồn. Một gói PES có thể có kích thước khối cố định hoặc

thay đổi, có thể lên tới 65536 byte/gói. Bao gồm 6 byte header, và số byte còn lại chứa nội dung chương trình. Định dạng của 1 PES header được minh họa trong hình 3.3 và giải thích trong bảng 3.1 dưới:



Hình 3.3: Định dạng gói MPEG PES

**Bảng 3.1: Cấu trúc của một gói MPEG PES**

Tên trường	Chức năng
Tiền tố mã bắt đầu gói	Gói PES bắt đầu với tiền tố 0x000001
Nhận dạng dòng (1 byte)	Trường này nhận dạng loại payload trong gói. Một mẫu bit 111x xxxx cho biết đó là gói audio, còn mẫu bit 1110 xxxx cho biết rằng đó là gói video. Giá trị "X" được sử dụng để biểu thị các số của các dòng MPEG
Độ dài gói PES	Trường dài 2 byte để chỉ thị độ dài gói
Mã đồng bộ	Trường được dùng để đồng bộ nội dung video và audio
Cờ header PES	Trường 14 bit chứa các cờ chỉ thị PES khác nhau hay các cờ, cung cấp phần cứng hay phần mềm bộ giải mã của set - top box IP với thông tin thêm vào. Các loại cờ gồm: Điều khiển tranh chấp PES: Cờ này báo cho bộ giải mã gói có được bảo đảm hay không thông qua xử lý tranh chấp. Độ ưu tiên PES: Cờ này cung cấp cho bộ giải mã thông tin về mức độ ưu tiên của gói PES. Bộ chỉ thị sắp xếp dữ liệu: Bộ chỉ thị này quyết định Payload của PES bắt đầu với bit video hay audio Thông tin bản quyền: Khi bit này được thiết đặt, nội dung video được bảo vệ bởi bản quyền. Bản gốc hay bản sao: Cờ này chỉ ra nội dung này là bản gốc hay bản sao.
Độ dài dữ liệu header của PES	Trường này nhận chỉ ra tổng số byte bị chiếm bởi các trường header khác nhau.
Trường header của PES	Trường này chứa một số các bit tùy chọn.
Payload của PES	Payload của PES gồm các dòng audio hay dữ liệu video.

Do bản chất của mạng, thứ tự hay chuỗi các khung video từ lối ra của trung tâm dữ liệu IPTV có thể khác thứ tự các khung do các thiết bị của người dùng nhận được. Do đó, để giúp đỡ quá trình đồng bộ, các hệ thống dựa trên MPEG thường dán nhãn các gói PES khác nhau trong chuỗi video.

Có 2 loại nhãn thời gian được sử dụng đối với mỗi gói PES: nhãn thời gian trình diễn(PTS), và nhãn thời gian giải mã(DTS):

., PTS---- nhãn thời gian trình diễn có giá trị thời gian 33 bit, được đặt trong trường PES header. Mục đích của việc sử dụng PTS cho mỗi gói là để xác định xem khi nào và theo trật tự nào thì gói đó được xem (bởi người xem).

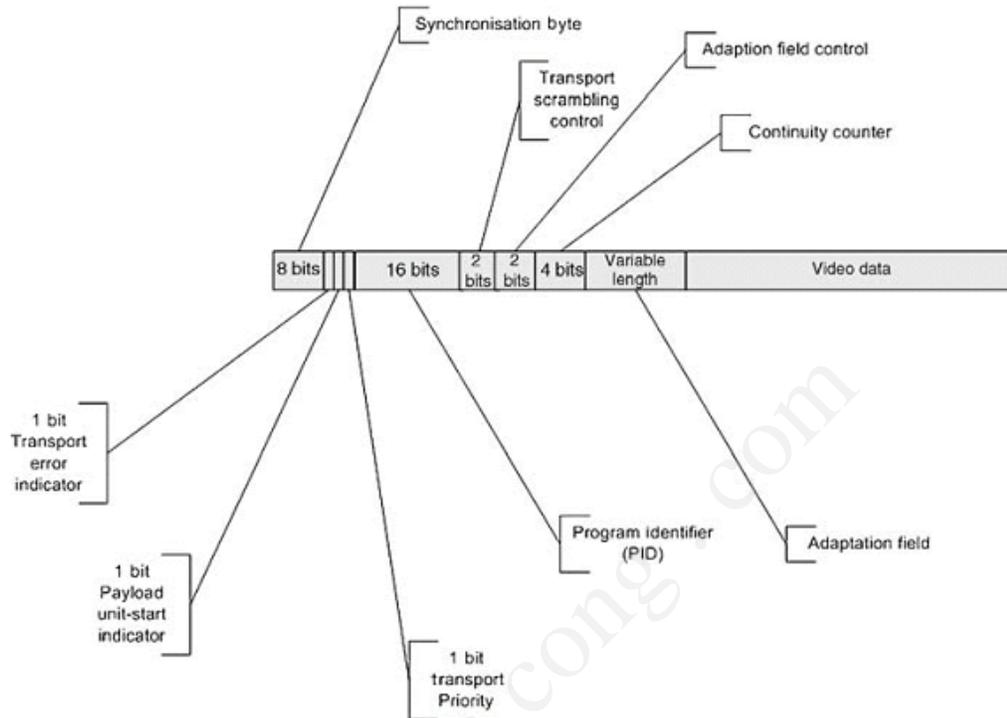
., DTS---nhãn giải mã để sử dụng để giúp bộ giải mã ở thiết bị của người sử dụng biết khi nào xử lý gói đó.

Khái niệm ứng dụng những nhãn thời gian khác nhau đối với mỗi gói PES trong dòng mã hóa MPEG được minh họa trong hình 3.4.

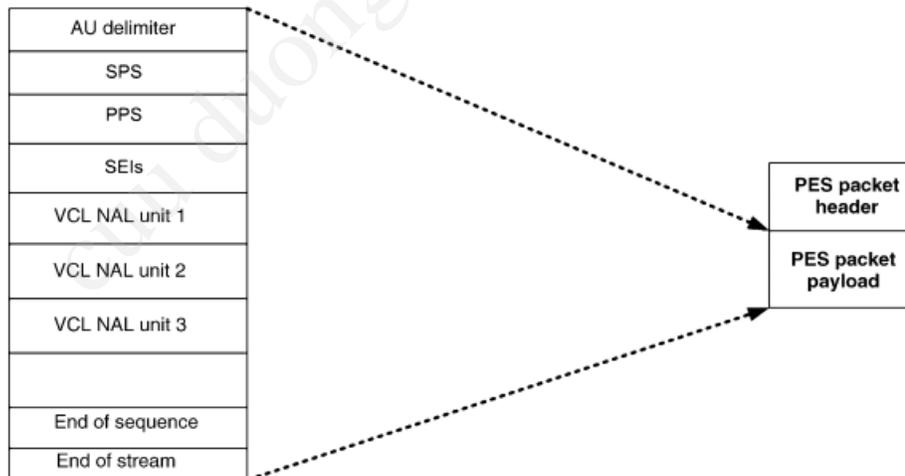
Như đã chỉ ra ở trên, thứ tự các gói được truyền đi qua mạng khác với thứ tự các gói nhận được ở thiết bị của người sử dụng. Thiết bị người sử dụng IPTV sẽ dùng các nhãn PTS và DTS để tái tạo lại nội dung video gốc. Bên cạnh việc gửi đi các nội dung nén MPEG-2, PES còn có khả năng truyền tải các khối H.264/AVC qua mạng IPTV. Các chi tiết quá trình ánh xạ được mô tả ở hình 3.5.

### **3.2.4 Lớp cấu trúc dòng truyền tải**

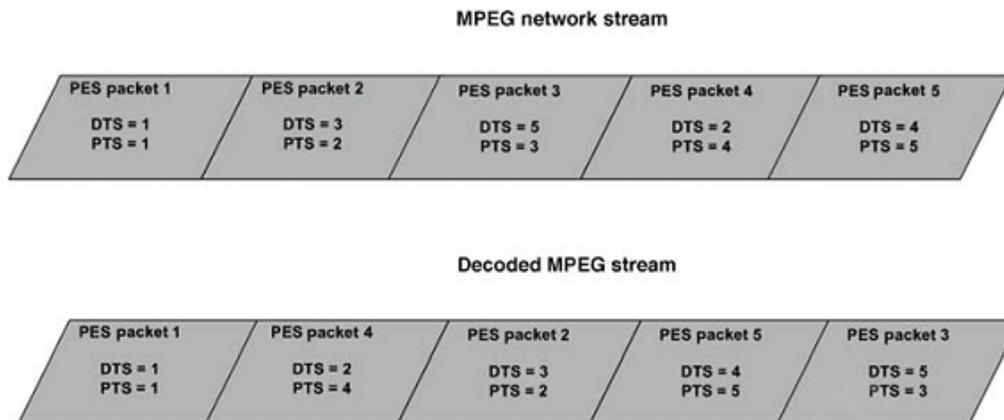
Lớp tiếp theo trong mô hình truyền thông IPTV làm nhiệm vụ tạo nên dòng truyền tải, bao gồm 1 dòng liên tiếp các gói. Những gói này thường được gọi là các gói TS, được tạo ra bằng cách ngắt các gói PES thành các gói TS có kích thước cố định là 188 byte độc lập với thời gian. Sử dụng thời gian độc lập này làm giảm khả năng mất gói tin trong quá trình truyền và giảm ồn. Mỗi gói TS bao gồm 1 trong 3 định dạng truyền thông: dữ liệu, âm thanh, hình ảnh. Do đó, các gói TS mang cố định 1 loại hình truyền thông. Mỗi gói TS bao gồm 184 byte payload và 4 byte header. Các thành phần của TS header được mô tả trong hình 3.6 và giải thích trong bảng sau.



Hình 3.4: Định dạng gói MPEG TS



Hình 3.5: Ảnh xạ gói truy cập AVC sang gói MPEG PES



Hình 3.6: Ứng dụng nhãn thời gian với các gói MPEG PES

**Bảng 3.2: Cấu trúc gói MPEG TS**

Tên trường	Chức năng
Trường đồng bộ	Phần header thường bắt đầu bằng các bit đồng bộ (8 bit), thường là các bit 0. Trường này dùng để xác định điểm bắt đầu của 1 gói IPTV.
Trường chỉ thị lỗi	Bit cờ này sẽ chỉ ra 1 lỗi (nếu có) liên quan đến dòng truyền tải
Trường chỉ thị điểm bắt đầu khối truyền tải	Bit cờ này sẽ chỉ ra 1 điểm bắt đầu của khối truyền tải
Mức ưu tiên truyền tải	Khi đặt cờ này sẽ chỉ ra mức ưu tiên khối payload
ID chương trình	Trường quan trọng nhất trong phần header là 13 bit để xác định ID chương trình. Nó sẽ chỉ ra gói nào thuộc dòng nào. Các gói thuộc dòng nào thì sẽ có cùng ID chương trình. Bộ phân kênh trong bị của người sử dụng dùng thông tin để

	phân biệt các loại gói khác nhau. Chú ý rằng các gói null có ID chương trình = 8191. Các gói không có ID chương trình thì sẽ bị thiết bị nhận IPTV loại bỏ.
Điều khiển việc tranh chấp các truyền tải	2 bit của trường này sẽ cho biết trạng thái mã hóa của phần payload gói truyền tải.
Trường điều khiển thích nghi	Trường 2 bit này sẽ cho biết liệu header của gói dữ liệu có liên quan tới dòng truyền tải có bao gồm trường thích nghi và payload không
Bộ đếm tiến	Bộ đếm tiến sẽ đếm tăng lên 1 khi 1 gói dòng truyền tải với cùng một ID chương trình. Nhờ đó có thể xác định được nếu có mất hoặc bị lặp gói. Điều này có thể ảnh hưởng hình ảnh.
Trường thích nghi	Trường này có thể có hoặc không có trong phần header. Trường thích nghi này bao gồm nhiều thông tin khác nhau được sử dụng để định thời và điều khiển, bao gồm cả PCR. PCR được sử dụng để đồng bộ đồng hồ IPTVCD với đồng hồ bộ mã hóa. Giá trị PCR có độ dài 42 bit và được tăng theo tốc độ đồng hồ chuẩn, 27MHz. Ngay sau khi đồng bộ, việc giải mã MPEG - 2 IPTV được tiến hành
Trường thích nghi	Trường này có thể có hoặc không có trong phần header. Trường thích nghi này bao gồm nhiều thông tin khác nhau được sử dụng để định thời và điều khiển, bao gồm cả PCR. PCR được sử dụng để đồng bộ đồng hồ IPTVCD với đồng hồ bộ mã hóa. Giá trị PCR có độ dài 42 bit và được tăng Theo tốc độ đồng hồ chuẩn, 27MHz. Ngay sau khi đồng bộ, việc giải mã MPEG - 2 IPTV được tiến hành.

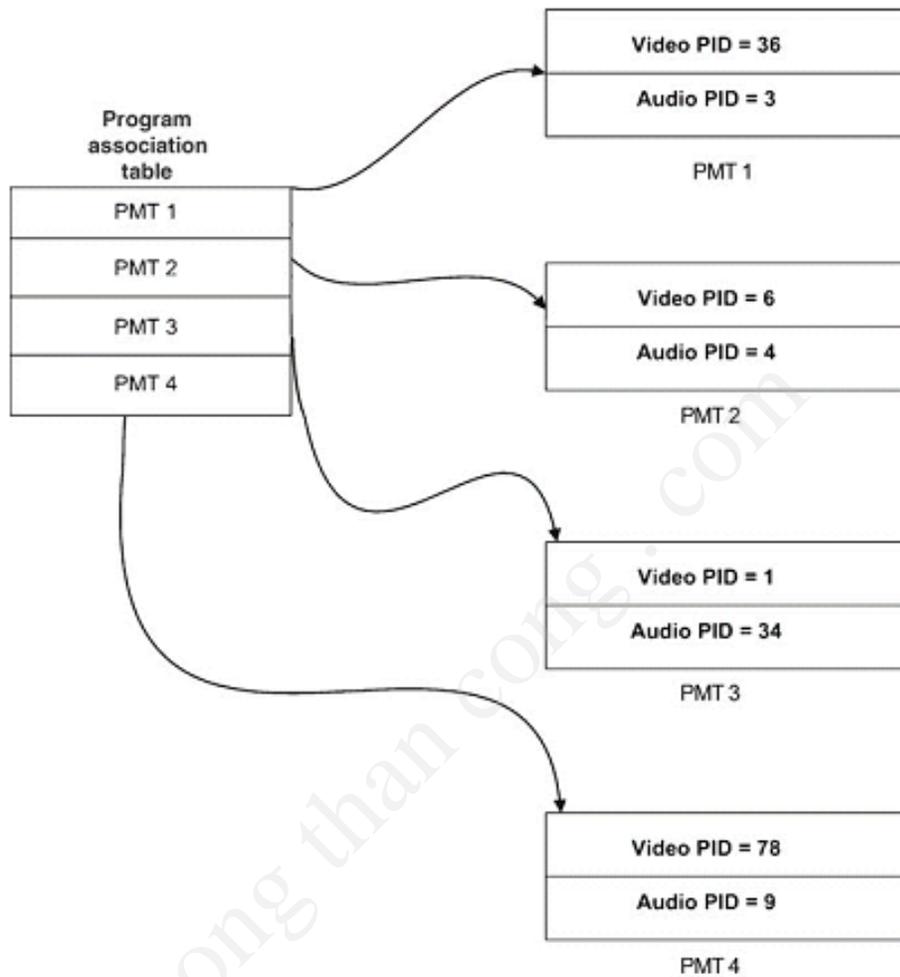
Lớp này cũng cung cấp chức năng để tạo ra các dòng chương trình. Một dòng chương trình là một gói PES chứa 1 vài dòng cơ bản được mã hóa sử dụng cùng đồng hồ chủ, hoặc đồng hồ hệ thống. Các kiểu dòng này được phát triển cho những ứng dụng như lưu trữ nội dung video trên các đĩa quang hoặc đĩa cứng.

Bên cạnh các nội dung hình ảnh và âm thanh đã được nén, dòng truyền tải bao gồm nhiều thông tin đặc trưng của chương trình hoặc metadata mô tả các dòng bit. Thông tin này được chứa trong 4 bảng PSI.

1) Bảng chương trình kết hợp (PAT)- Việc truyền các bảng PAT là bắt buộc và là điểm vào các bảng PSI. Bảng chương trình kết hợp luôn có ID chương trình là 0. Bảng này đưa ra các liên kết giữa chỉ số chương trình và ID của chương trình

2) Bảng ánh xạ chương trình: Bảng ánh xạ chương trình cũng là bắt buộc và mang các thông tin về một chương trình cụ thể. Bảng ánh xạ chương trình liệt kê các ID chương trình cho các gói mang các thành phần của một chương trình cụ thể (âm thanh, hình ảnh, dữ liệu, và các thông tin PCR). Hình 3.7 mô tả một ví dụ về mối quan hệ giữa bảng chương trình kết hợp (PAT) và bảng chương trình ánh xạ (PMT).

Vì thế khi một thiết bị người dùng (IPTVCD) yêu cầu một chương trình, bảng chương trình kết hợp sẽ được kiểm tra, sau đó sẽ kiểm tra bảng chương trình ánh xạ để định nghĩa các IP chương trình về gói âm thanh, hình ảnh và dữ liệu liên kết với chương trình đó. Trong ví dụ này, thuê bao lựa chọn chương trình 1 và thiết bị IPTV của người sử dụng định vị toàn bộ các gói truyền tải với ID chương trình là 36 đối với phần hình ảnh của chương trình và các gói với ID chương trình là 3 với các phần âm thanh của chương trình. Nếu dữ liệu là quảng bá với chương trình, thì bảng ánh xạ chương trình sẽ bao gồm các chi tiết trên đó xác định các gói dữ liệu truyền tải.



Hình 3.7: Mối liên hệ giữa PMT và PAT

3) Bảng điều kiện truy cập (CAT)- bảng điều kiện truy cập là một bảng tùy chọn PSI bao gồm các ID chương trình của EMMs (các tin quản lý quyền truy cập). Tin quản lý quyền truy cập bao gồm các thông tin về mức cho phép đối với hệ thống truy cập.

4) Bảng thông tin mạng (NIT)- Bảng thông tin mạng NIT là một bảng tùy chọn lưu trữ các thông tin như tần số kênh và số dòng truyền tải. Set top box sẽ sử dụng thông tin này để chỉnh sóng tới các chương trình cụ thể.

Khi TS được cấu trúc và định dạng, nó sẽ được chuyển xuống lớp truyền tải trực tiếp hoặc tới lớp sử dụng giao thức truyền tải thời gian thực (RTP).

### 3.2.5 Lớp giao thức truyền tải thời gian thực( tùy chọn)

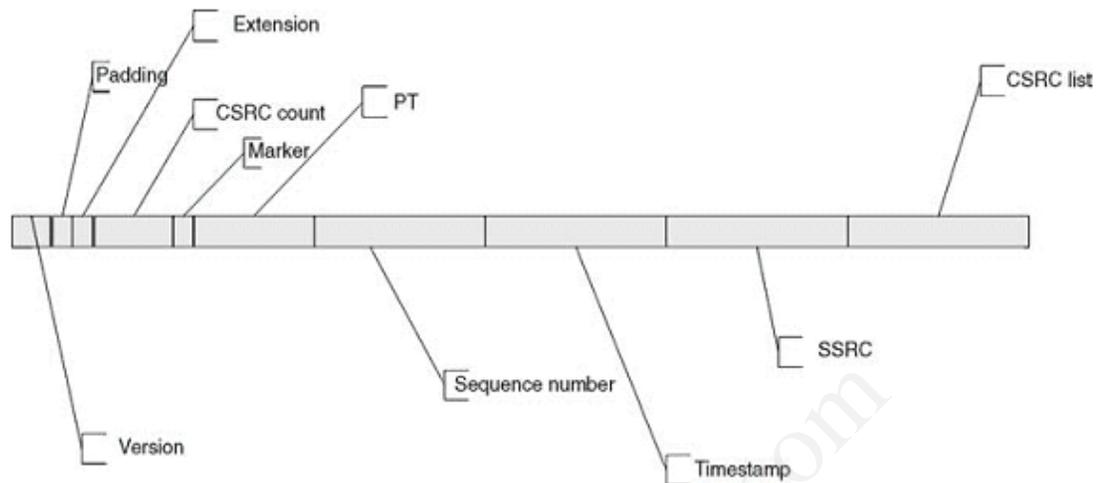
Lớp tùy chọn này được sử dụng rộng rãi trong nhiều ứng dụng khác nhau. Lớp này hoạt động như 1 lớp trung gian giữa các nội dung được nén MPEG-2, H.264/AVC ở lớp cao hơn và cả lớp thấp hơn trong mô hình IPTV. Giao thức RTP chính là lõi của lớp này và thường là block cơ sở hỗ trợ truyền dòng nội dung theo thời gian thực qua mạng IP.

Giao thức truyền tải thời gian thực phân phát các dòng âm thanh và hình ảnh bằng cách đóng gói các nội dung này trong một định dạng riêng biệt được gọi gói. Mỗi gói bao gồm phần header và payload(dữ liệu IPTV). Để sử dụng hiệu quả băng thông, phần payload thường bao gồm nhiều hơn một gói MPEG-TS.

Phần header bao gồm các chức năng cốt yếu để các thẻ truyền thành công các dữ liệu thời gian thực qua mạng. Header của RTP có thể nhận biết với header của UDP có giá trị là 5004, và bao gồm rất nhiều trường. Chi tiết về các trường khác nhau được minh họa trong hình 3.8 và bảng 3.3.

Có một điều đáng chú ý, đó là giao thức thời gian thực không có trường dài trong phần header bởi vì nó phụ thuộc vào giao thức truyền tải cơ bản để cung cấp loại thông tin này. Như đã miêu tả trong bảng dưới, lợi ích chính của việc chèn các nội dung video đã được nén và trong các gói RTP là:

- 1)Thêm số chuỗi vào gói để giúp cả bộ giải mã ở phía nhà cung cấp và thiết bị người dùng có thể sắp xếp lại các gói nhận được từ mạng IP.
- 2)Trường nhãn thời gian giúp khắc phục các vấn đề như jitter và mất đồng bộ giữa nguồn và đích.



Hình 3.8: Định dạng RTP header

Khi header truyền tải thời gian thực được thêm vào payload video, gói truyền tải thời gian thực được gửi tới giao thức TCP hoặc UDP để tiếp tục xử lý.

Định dạng phần payload của RTP cho việc đóng gói dòng bit nén MPEG-2 : thay vì sử dụng UDP để mang các gói TS MPEG-2 thì một vài hệ thống IPTV sử dụng lớp RTP thêm vào lớp UDP để truyền các gói. Việc ánh xạ các gói MPEG-TS sang các gói RTP là khá đơn giản. Cấu trúc bao gồm phần header và payload của gói MPEG-2 TS. Mỗi gói có độ dài 188 byte. Hình 3.9 mô tả cấu trúc truyền tải nội dung DVB dựa trên MPEG-2 qua mạng IP.

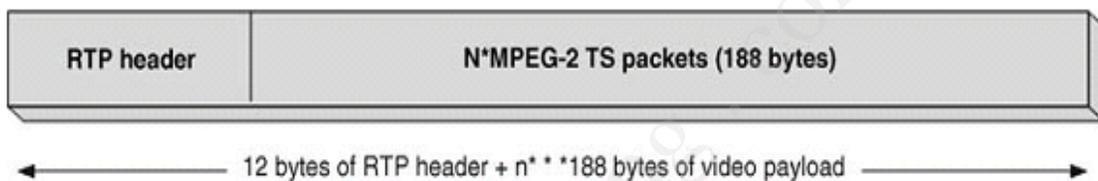
Định dạng payload RTP cho việc đóng gói dòng bit được nén dùng chuẩn H.264/AVC: RFC 3984 cung cấp khuyến nghị về giải pháp truyền các nội dung H264/AVC và định nghĩa 3 cơ chế để chèn các khối NAL vào RTP payload:

- 1) Một gói NAL riêng biệt: kỹ thuật này định nghĩa sự ánh xạ gói NAL sang từng payload RTP.
- 2) Gói NAL tập hợp: Kỹ thuật này định nghĩa sự ánh xạ nhiều gói NAL sang một gói RTP.

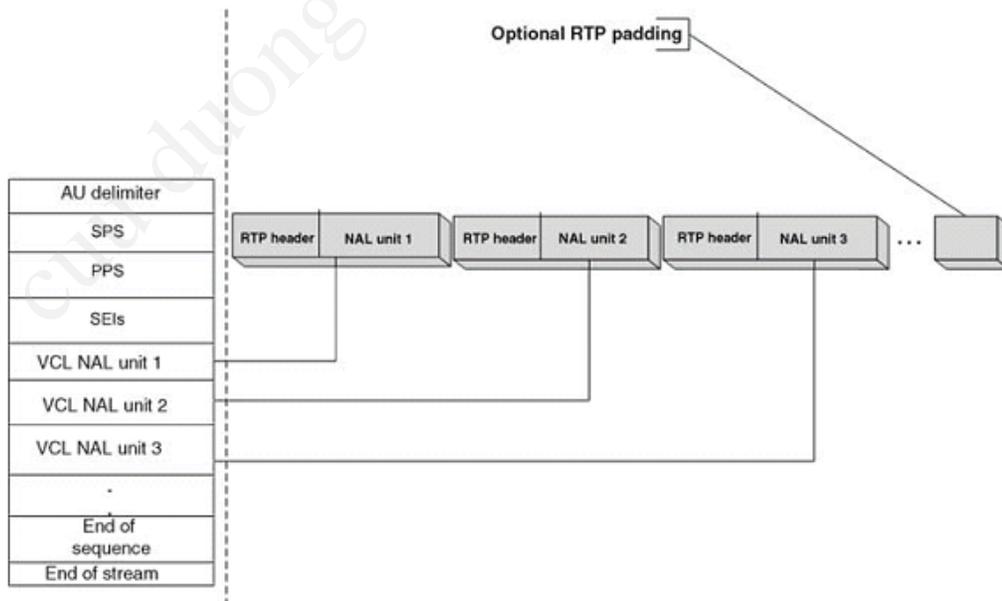
**Bảng 3.3: Cấu trúc của gói IPTV dựa trên RTP**

<b>Tên trường</b>	<b>Chức năng</b>
Phiên bản (V)	Trường này xác định phiên bản RTP được dùng trong gói IPTV.
Phần đệm (P)	Trường này xác định có byte đệm trong gói RTP hay không.
Phần mở rộng (X)	Nếu bit này được đặt bằng 1 thì phần mở rộng theo ngay sau tiêu đề cố định.
Tổng số nguồn góp (CSRC)	Trường này chứa thông tin số bộ nhận diện CSRC có trong gói.
Bit dấu	Chức năng của nó được xác định bởi mô tả RTP. Thường được sử dụng để xác định ranh giới khung.
Loại Payload (PT)	Trường này chứa thông tin về định dạng payload của IPTV. Ví dụ, giá trị 34 chỉ ra nội dung video được mã hóa sử dụng H.263
Số thứ tự gói	Trường này giúp tìm ra được những gói bị mất, lỗi. Giúp cho IPTVCD sắp xếp lại các gói được gửi tới không theo thứ tự, xác định đúng kính thước gói không đúng và chỉ ra gói bị lặp. Giá trị trong trường được tăng lên một mỗi lần một gói RTP được gửi qua mạng. Khi dòng IPTV bắt đầu, một giá trị bất kì được gán cho trường này để giảm rủi ro bị hacker tấn công.
Dấu thời gian	Trường này giữ dấu thời gian của gói, được khởi tạo từ một đồng hồ đáng tin cậy. Trường này được sd để thêm vào trong các gói âm thanh và hình ảnh đúng theo thứ tự thời gian của dòng IPTV.

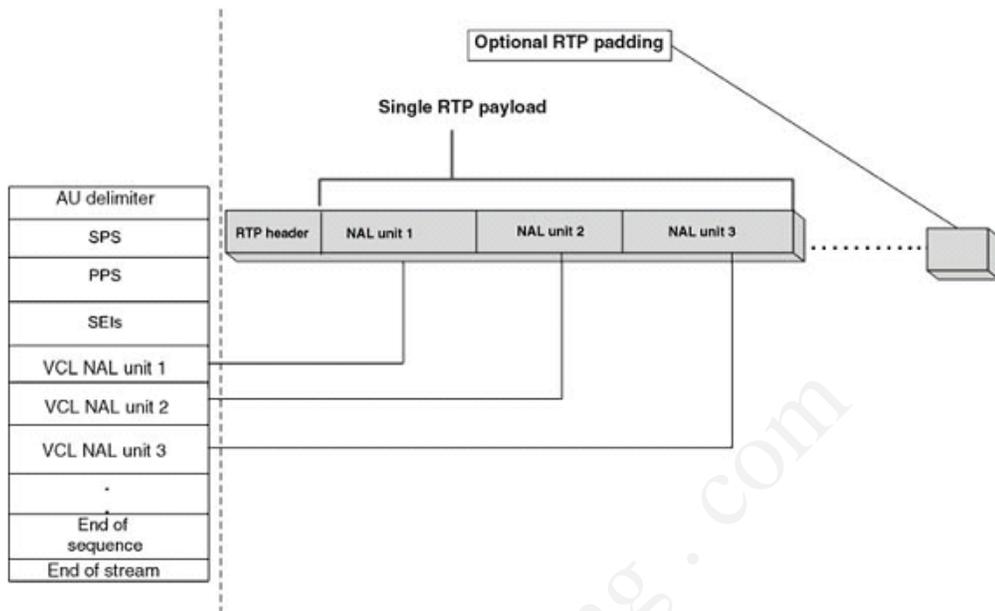
<p>Nguồn đồng bộ (SSRC)</p>	<p>Mục đích của trường này để chỉ ra nguồn đồng bộ trong mạng IPTV. Trường này thường được sử dụng kết hợp với trường số thứ tự gói để sửa những vấn đề xảy ra trong chuỗi IPTV.</p>
<p>Danh sách CSRC</p>	<p>Mục đích của trường 32 bit này để chỉ ra những nguồn video và audio góp vào payload IPTV.</p>



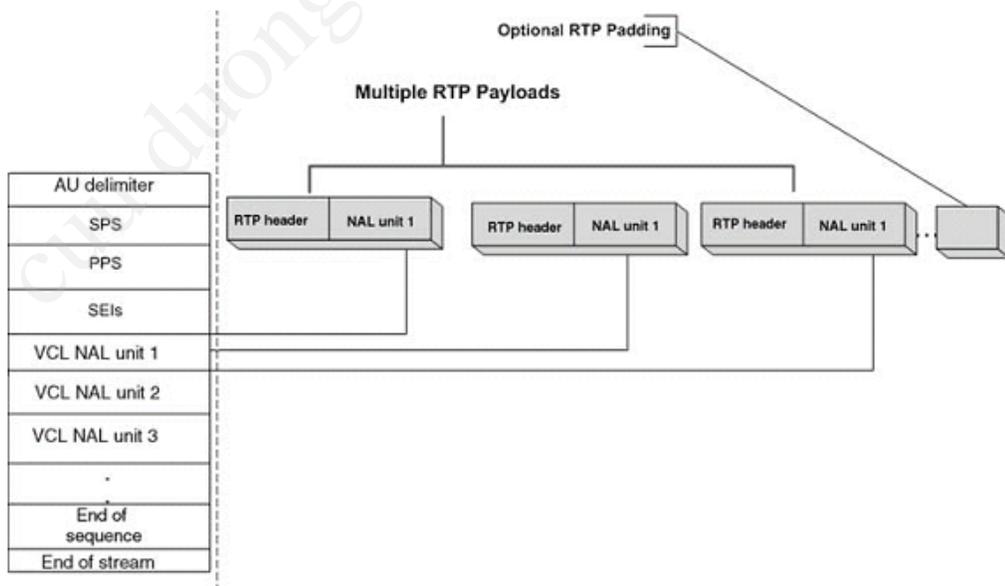
Hình 3.9: Các gói MPEG TS



Hình 3.10: Ảnh xạ nội dung H264/AVC ( từng khối NAL riêng biệt ) sang RTP payload



Hình 3.11: Ảnh xạ nội dung H264/AVC (nhiều khối NAL riêng biệt) sang một RTP payload



Hình 3.12: Ảnh xạ nội dung một H264/AVC NAL sang nhiều RTP payload

Khối NAL tập hợp được định nghĩa để xác định dung lượng gói lớn nhất đối với mỗi mạng. VD, với mạng Ethernet kích thước gói lớn nhất là 1500 byte, còn với mạng ATM kích thước gói lớn nhất là 54 byte. Dùng các gói NAL tập hợp để ánh xạ nhiều gói NAL sang một phần payload RTP không cần phải chuyển mã và thêm nhiều tiêu đề gói khi triển khai IPTV trên các nền.

3) Gói NAL phân tách: Đây là kỹ thuật để ánh xạ 1 khối NAL riêng rẽ ra khỏi nhiều phần payload RTP.

Điểm đáng chú ý là gói phân tách NAL phải được gửi qua mạng theo một trật tự liên tiếp. Điều này là có thể khi sử dụng các số tăng dần trong header của RTP. Kỹ thuật này đem lại hai lợi ích cho nhà cung cấp dịch vụ. Thứ nhất, điều này giúp truyền lượng lớn chương trình có độ phân giải cao dựa trên IP. Thứ hai là giúp tăng khả năng sửa lỗi.

Chú ý rằng, kỹ thuật RTP thường được triển khai trong các mạng không đảm bảo chất lượng dịch vụ QoS để truyền các dịch vụ IPTV. Mặc dù RTP giúp làm tăng khả năng các dòng tới đích trong trật tự đúng, nhưng không được thiết kế để đảm bảo các mức chất lượng dịch vụ. Do đó, trách nhiệm của nhà cung cấp dịch vụ là đảm bảo video luôn được ưu tiên khi chúng được truyền đi trong hạ tầng mạng.

### 3.2.6 Lớp truyền tải

Thông thường các gói RTP là dạng đầu vào của lớp truyền tải. Điều đáng chú ý là có thể ánh xạ trực tiếp các gói MPEG-TS sang payload giao thức của lớp truyền tải.

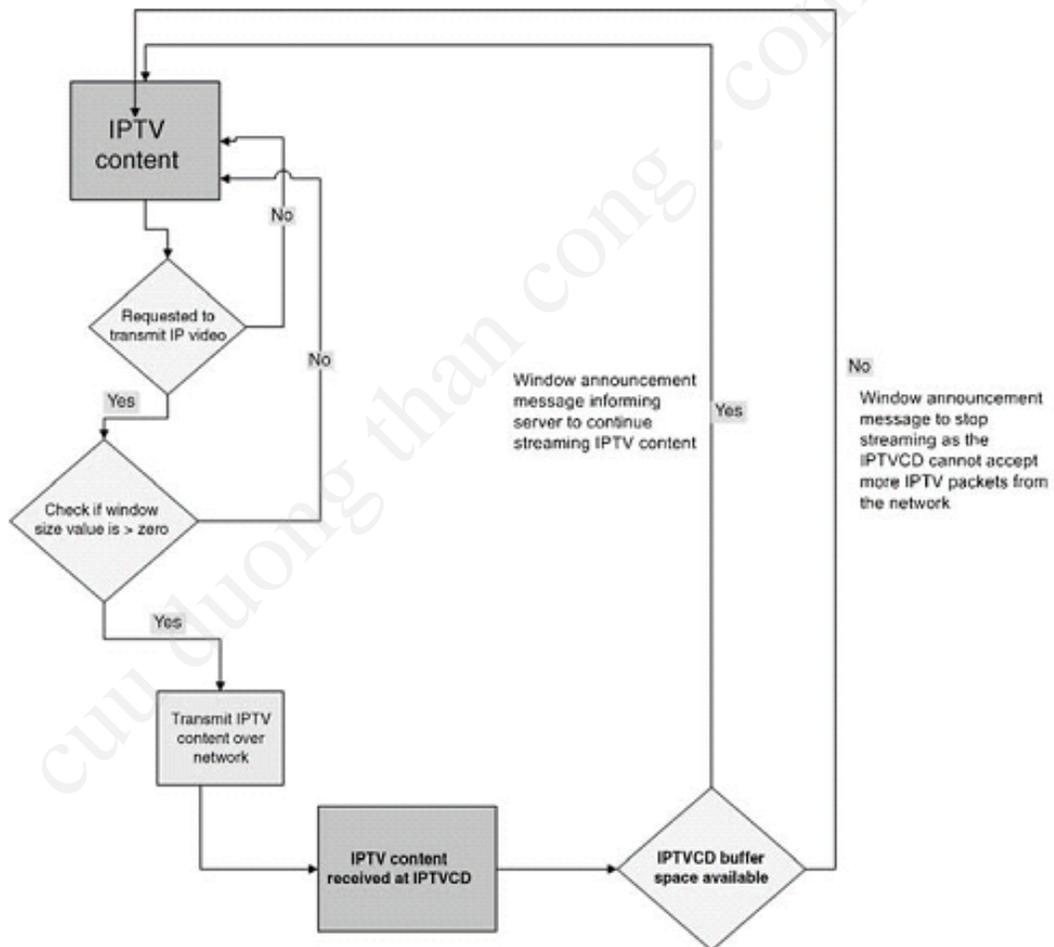
Lớp truyền tải IPTV được thiết kế để đảm bảo các kết nối đầu cuối là tin cậy. Nếu dữ liệu tới thiết bị người nhận đúng, lớp truyền tải sẽ truyền lại. Lớp truyền tải thông báo với lớp trên để có các thông tin chính xác hơn.

TCP và UDP là 2 giao thức quan trọng nhất được sử dụng ở lớp này.

Sử dụng TCP để định tuyến các gói IPTV :

TCP là giao thức cốt lõi của bộ giao thức internet và được xếp vào loại định hướng kết nối. Điều này cơ bản có nghĩa là kết nối được thiết lập giữa đầu cuối nhà cung cấp và thiết bị IPTV của người sử dụng để truyền các chương trình qua mạng.

TCP có khả năng điều khiển lỗi xảy ra trong quá trình truyền các chương trình qua mạng. Các lỗi như mất gói, mất trật tự gói, hoặc lặp gói thường gặp trong môi trường truyền IPTV. Để xử lý các tình huống này, TCP sử dụng hệ thống các số liên tục để cho phép thiết bị gửi có thể gửi lại các dữ liệu hình ảnh bị mất hoặc hỏng. Hệ thống số liên tục này là trường có độ dài 32 bit trong cấu trúc gói. Trường đầu tiên chứa chuỗi số bắt đầu của dữ liệu trong gói và trường thứ hai chứa giá trị của chuỗi số tiếp theo mà video server đang đợi (mong) nhận trở lại từ IPTVCD



Hình 3.13: Cơ chế điều khiển luồng của TCP

Bên cạnh việc sửa các lỗi có thể xảy ra trong quá trình truyền nội dung video qua mạng IP băng rộng, TCP còn có điều khiển luồng dữ liệu. Điều này có thể đạt được bằng cách sử dụng trường kích thước cửa sổ, với thuật toán được gọi là cửa sổ

trượt. Giá trị trong trường này xác định số các byte có thể truyền đi qua mạng trước khi nhận được xác nhận từ phía thiết bị nhận.

Trong môi trường IPTV, giá trị trường kích thước cửa sổ chính là kích thước vùng đệm trong IPTVCD trừ đi lượng nội dung đã có trong vùng đệm tại một thời điểm. Dữ liệu này sẽ được giữ cho tới khi bản tin thông báo đã nhận được gửi về từ IPTVCD.

Khi giá trị của trường này bằng 0, IPTVCD ở phía đầu thu sẽ không đủ khả năng xử lý các dữ liệu IPTV ở tốc độ đủ lớn. Khi đó, TCP sẽ chỉ thị cho video server dừng hoặc làm chậm lại tốc độ gửi các gói dữ liệu tới IPTVCD. Điều này sẽ đảm bảo rằng IPTVCD sẽ không bị tràn các gói dữ liệu tới. Khi IPTVCD đã xử lý xong các các gói dữ liệu trong vùng đệm và video server đã biết được điều đó thì giá trị tại vùng đệm sẽ tăng lên, và video server sẽ bắt đầu truyền tiếp các nội dung. Trong môi trường IPTV lí tưởng, số của cửa sổ được báo về từ IPTVCD sẽ báo cho server biết không gian vùng đệm còn trống chính là tốc độ mà tại đó các nội dung video được gửi đi từ video server.

Các cổng TCP và Socket: Mỗi điểm cuối của 1 liên kết IPTV thì có 1 địa chỉ IP và 1 giá trị cổng liên quan. Vì thế mỗi liên kết có 4 thành phần khác nhau:

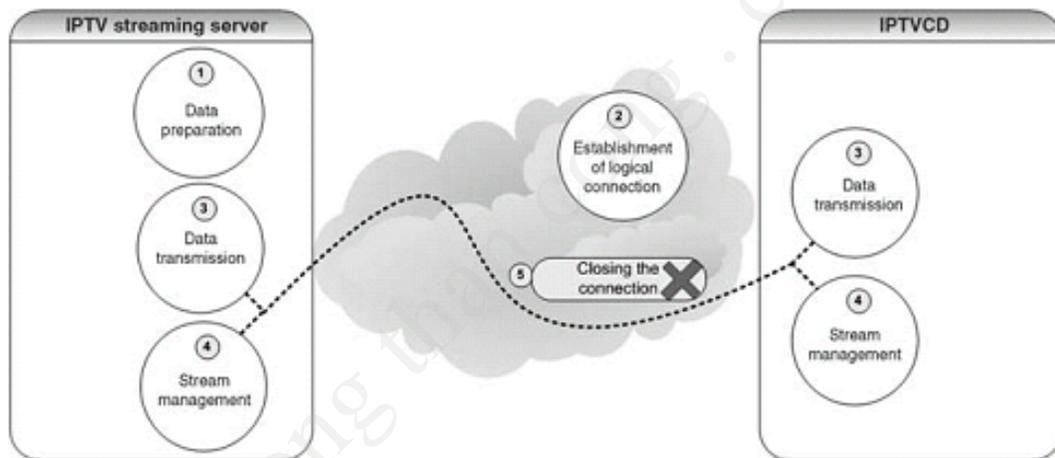
- (1) Địa chỉ IP của video server
- (2) Số của cổng của video server
- (3) Địa chỉ IP của IPTVCD
- (4) Số của cổng của IPTVCD

Việc kết hợp địa chỉ IP và số của cổng cho phép một tiến trình trên IPTVCD có thể liên lạc trực tiếp với tiến trình đang chạy trên một trong các máy server được đặt ở trung tâm dữ liệu IPTV. Một cổng gồm 16 bit để định nghĩa hướng để truyền các thông báo giữa các lớp mạng. Có 2 loại cổng :

- 1) Cổng well known có giá trị từ 1 đến 1023. Loại cổng này thường được các server sử dụng và được quản lí bởi IANA.

2) Cổng Ephemeral được thiết lập bởi IPTVCD ở trạng thái tạm thời khi liên lạc với IPTV server. Các cổng thường được nhớ trong ngăn xếp phân mềm IP. Các giá trị này thường lớn hơn 1024 và nhỏ hơn 65535. Cổng này không chịu sự quản lý của IANA.

Socket cũng là một thành phần quan trọng khác trong mô hình truyền thông IP. Một socket về cơ bản là một giao diện ứng dụng chương trình (API), được sử dụng để làm cho việc liên lạc giữa các tiến trình đang chạy trên 1 thiết bị IP. Một socket được thiết lập bằng cách kết hợp địa chỉ IP với số của cổng.



Hình 3.14: Quá trình truyền thông trong mạng IPTV

Để hiểu hơn về mối liên hệ giữa địa chỉ IP và socket, xét các bước để thiết lập một kênh truyền thông giữa 1 tiến trình chạy trên IPTVCD và một tiến trình chạy trên trung tâm cung cấp dữ liệu IPTV. Bước được mô tả như sau:

1) Chuẩn bị dữ liệu: Tiến trình gửi chạy trên hệ thống server IPTV chuẩn bị nội dung và gọi module truyền thông TCP/IP để truyền các dữ liệu tới 1 tiến trình đang chạy trên một IPTVCD. Các tiến trình truyền thông bắt đầu và thông tin header được thêm vào nội dung khi truyền qua các lớp trong IPTVCM.

2) Thiết lập kết nối logic TCP: Cả 2 đầu kết nối đều được định nghĩa bởi 1 địa chỉ IP và 1 số cổng. Kết hợp giữa địa chỉ IP và số cổng gọi là socket. Hệ thống địa chỉ đối với liên kết truyền thông bao gồm các thành phần sau:

Giao thức :

Địa chỉ IP của máy chủ IPTV

ID của tiến trình chạy trên máy chủ IPTV

Địa chỉ IP của IPTVCD

ID của tiến trình chạy trên IPTVCD

- Truyền dữ liệu: Truyền thông bắt đầu thông qua socket giữa 2 tiến trình từ phía IPTV server đến IPTVCD.
- Quản lí các dòng nội dung IPTV: giao thức TCP quản lí các dòng IPTV trong khi kết nối được thiết lập.
- Hủy bỏ kết nối: Khi hoàn thành việc truyền các nội dung IPTV, IPTVCD hoặc trung tâm dữ liệu sẽ hủy bỏ socket và kết nối mạng.

Header: Thông tin này giúp cho segment được truyền đi từ nguồn đến đích. Header mang thông tin chính là số cổng của nguồn và đích, số chuỗi của segment và kiểm tra tổng. Các số tổng đảm bảo rằng dữ liệu có thể tới và trở về từ đúng các tiến trình đang chạy trên mỗi thiết bị IP. Số chuỗi giúp TCP có thể hiểu được bằng cách nào để đưa dữ liệu về dạng trước khi bị ngắt thành các segment.

**Bảng 3.4: Định dạng của TCP segment**

Tên trường	Chức năng
Cửa nguồn	Trường 16 bit. Chỉ ra số của cổng nguồn.
Cổng đích	Trường 16 bit. Chỉ ra số của cổng đích.
Chỉ số dãy	Chỉ ra chỉ số dãy của đoạn TCP. Điều này giúp TCP theo được vết của mỗi gói IPTV chuyển qua mạng.
Chỉ số chấp nhận	Chứa chỉ số dãy tiếp theo mà thiết bị gửi mong muốn nhận được. Trường này chỉ chứa giá trị này nếu bit điều khiển ACK được thiết lập.
Độ dời dữ liệu	Trường ngắn này chỉ ra vị trí trong đoạn mà dữ liệu video bắt đầu.
Được đặt trước	Trường này được đặt trước để sử dụng sau này và có giá trị bằng 0
Các bit điều khiển	Trường này gồm có 6 bit điều khiển: URG: Urgent pointer ACK: Acknowledgment PSH: Push RST: Reset (đặt lại liên kết) SYN: Synchronize (đồng bộ chỉ số dãy) FIN: không có thêm dữ liệu từ bên gửi
Cửa sổ	Chỉ kích thước cửa sổ cho dữ liệu phân phối giữa những lần nhận
Checksum	Kiểm tra lỗi để bảo vệ dữ liệu được gửi qua mạng
con trỏ khẩn cấp	Chỉ được dịch khi trường con trỏ khẩn cấp được cho phép
Các tùy chọn và đệm	Các tùy chọn là các bội số 8 bit có sẵn và có thể chứa các kiểu biến khác nhau. Các tùy chọn được lựa chọn sẽ quyết định độ dài của đoạn TCP. Đệm là phần cuối cùng của tiêu đề và tạo thành bởi các số 0

TCP ánh xạ segment sang giao thức IP sau khi chèn các thông tin cần thiết vào trong phần header. Như đã mô tả ở trên, TCP cung cấp toàn bộ các chức năng để truyền tải các dữ liệu qua mạng IP.

### Sử dụng UDP để định hướng các gói IPTV:

UDP là giao thức thuộc về bộ giao thức Internet. UDP cho phép máy chủ kết nối với mạng băng rộng để gửi tới các IPTVCD dịch vụ truyền hình quảng bá có chất lượng hài lòng người dùng. UDP giống với TCP nhưng là phiên bản sơ lược hơn, đưa ra cho số lượng tối thiểu các dịch vụ truyền tải. UDP là giao thức không liên kết, điều đó có nghĩa là kết nối giữa video server và IPTVCD ko cần phải thiết lập trước khi dữ liệu được truyền đi. Video server đơn giản chỉ thêm vào địa chỉ IP đích và số cổng vào datagram và gửi tới cơ sở mạng để phân phát tới địa chỉ IP đích. Khi trên mạng, UDP sử dụng cách tốt nhất để cố gắng thu được dữ liệu về điểm đích của nó. Chú ý rằng UDP sử dụng các khối dữ liệu được gọi là các datagram để truyền nội dung qua mạng.

### **Ưu điểm và nhược điểm của UDP:**

Ưu điểm của UDP:

- Không có ngắt trong quá trình truyền nội dung video: không có trễ trong quá trình phân phối ngay cả khi trong mạng có các gói bị trễ hoặc bị hỏng. Ngược lại, khi sử dụng TCP, có thể xảy ra sự ngắt quãng khi phải chờ các gói bị trễ và các khung hình tới hoặc phải chờ các gói bị hỏng được thay thế.
- Dung lượng thấp: Kích thước header của UDP chỉ bao gồm có 8 byte trong khi TCP header chiếm tới 20 byte.
- Tốc độ thiết lập kết nối: thời gian thiết lập và hủy bỏ kết nối giữa IPTVCD và các thiết bị ở trung tâm dữ liệu IPTV ngắn. Do đó, việc phân phối các gói sử dụng giao thức UDP thương nhanh hơn so với sử dụng giao thức TCP.
- Hỗ trợ truyền 1 chiều: UDP không yêu cầu đường về, do đó cho phép các công ty sử dụng vệ tinh có thể truyền nội dung IPTV truyền đa điểm tới khách hàng của mình.

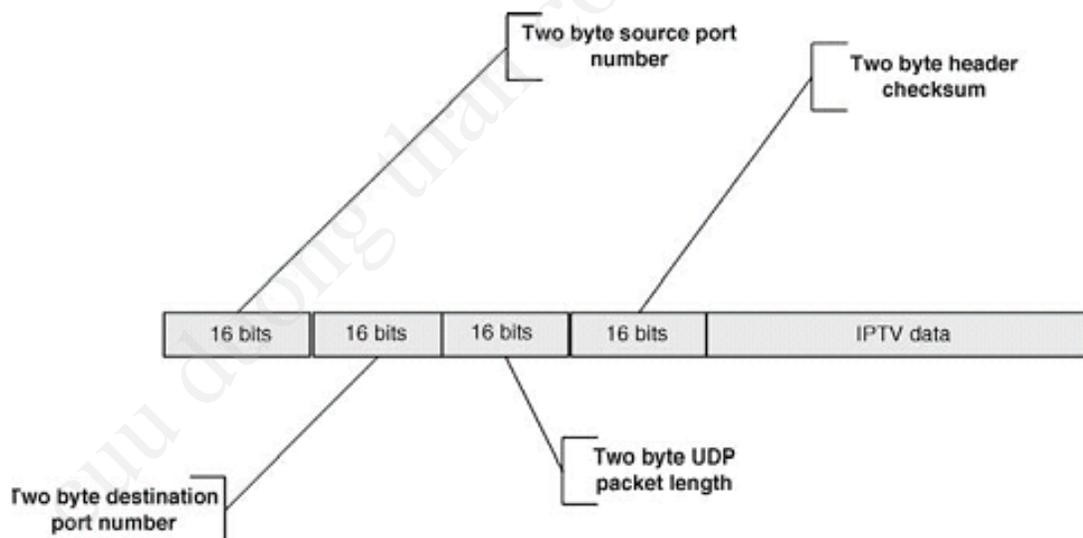
Nhược điểm của UDP:

Mặc dù UDP là nhanh chóng và hiệu quả đối với các ứng dụng cần thời

gian, và sẽ là không hiệu quả trong trường hợp:

□ Tính toàn vẹn của dữ liệu: Tính toàn vẹn của dữ liệu khi sử dụng UDP là không được bảo đảm khi UDP chỉ cung cấp 1 dịch vụ duy nhất là kiểm tra tổng và multiplexing thông qua số cổng. bất kì vấn đề nào cũng có thể xảy ra trong quá trình truyền thông ở tại đầu cuối nào cần được điều khiển độc lập với các ứng dụng. Các vấn đề thường gặp như là phát lại, đóng gói và lắp ráp lại, truyền lại các gói bị mất, sự tắc nghẽn, và điều khiển luồng nằm ngoài khả năng sửa lỗi của UDP.

□ Khó khăn trong việc vượt qua các tường lửa: Nhiều loại tường lửa trên mạng chặn các thông tin UDP gây ra các lỗi trong quá trình truyền thông. Đây không phải là vấn đề lớn đối với các nhà cung cấp dịch vụ IPTV, tuy nhiên nó cũng ảnh hưởng tới các công ty cung cấp dịch vụ Internet TV.



Hình 3.15: Định dạng datagram dựa trên UDP

**Bảng 3.5: Cấu trúc datagram IPTV dựa trên UDP**

<b>Tên trường</b>	<b>Chức năng</b>
Cửa nguồn	Chỉ ra chỉ số cửa của quá trình gửi datagram. Nó là cửa tùy chọn và nếu không được dùng thì sẽ được điền đầy bằng các số 0.
Cửa đích	Chỉ ra chỉ số cửa của quá trình đích đang chạy trên IPTVCD.
Độ dài	Giúp cho IPTVCD xác định độ dài và kích thước của datagram UDP đang đến. Trường độ dài bao gồm một giá trị hệ 8, bao gồm cả tiêu đề và dữ liệu video thực sự.
Checksum	Trường này dài 2 byte chứa số được định trước, cho phép một IPTVCD kiểm lại tính nguyên vẹn của UDP đang đến dựa trên datagram IPTV.
Dữ liệu video	Phần này của datagram chứa dữ liệu video. Trong trường hợp môi trường IPTV, dữ liệu là một phần của datagram UDP được định dạng bởi giao thức dòng video và audio mà được sử dụng tại đầu cuối IPTV.

Đối với IPTV, UDP tỏ ra hữu ích khi trung tâm dữ liệu cần gửi các nội dung video IP tới nhiều IPTVCD và là giao thức mức truyền tải phổ biến nhất mà các nhà cung cấp dịch vụ IPTV.

### 3.2.7 Lớp IP

Sau lớp truyền tải là lớp IP (còn được gọi là lớp liên mạng). Nhiệm vụ chính của lớp này là đưa các dữ liệu tới các vị trí mạng riêng biệt thông qua nhiều mạng

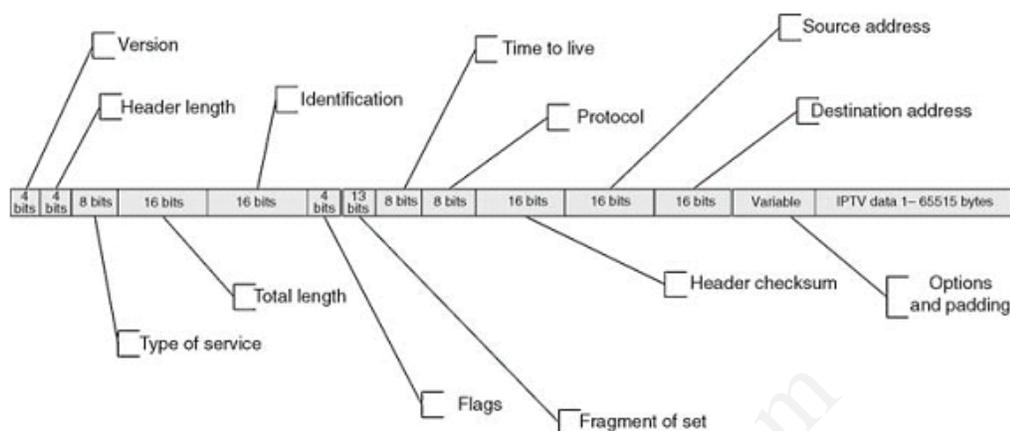
độc lập được liên kết với nhau được gọi là liên mạng. Lớp này được sử dụng để gửi các dữ liệu thông qua các đường khác nhau tới đích. IP là giao thức tốt nhất được sử dụng trong lớp liên mạng. Giao thức này cung cấp dịch vụ phân phát gói cơ bản cho tất cả các dịch vụ IPTV. Các loại dịch vụ này với hệ thống truyền đơn điểm, nơi các gói được truyền từ nguồn tới một IPTVCD đích, khác với hệ thống truyền đa điểm nơi mà các gói được truyền từ máy chủ tới nhiều IPTVCD.

IPv4 là giao thức phổ biến nhất được sử dụng trong mạng IPTV ngày nay. Nhiệm vụ chính của IP là phân phát các bit dữ liệu trong các gói từ nguồn tới đích. IP sử dụng kỹ thuật có hiệu quả cao nhất để phân phát dữ liệu. Nói cách khác không có tiến trình nào đảm bảo quá trình phân phát thông tin qua mạng. Các khối cơ sở của giao thức IP là các đoạn bit dữ liệu được đặt trong các gói và được định địa chỉ.

Gói IP là đơn vị dữ liệu bao gồm dữ liệu video thực và các thông tin của việc nhận video từ trung tâm cung cấp dữ liệu IPTV tới đích IPTVCD.

Cách đánh địa chỉ IP: trong môi trường IPTV, địa chỉ IPv4 thường được dùng để định nghĩa IPTVCD và trung tâm cung cấp dữ liệu. Địa chỉ IPv4 là chuỗi 4 số được ngăn cách với nhau bằng các dấu chấm để định nghĩa một cách chính xác vị trí vật lý của một thiết bị, ví dụ như set-top box, trong mạng. Địa chỉ IPv4 gồm 32 bit trong hệ nhị phân. Các số nhị phân này được chia thành 4 octet, mỗi octet 8 bit, mỗi octet được đại diện bởi một số hệ thập phân nằm trong khoảng từ 0 đến 255. Mỗi octet được ngăn cách bởi một dấu chấm trong hệ thập phân. Địa chỉ IP được tổ chức thành hai phần:

- 1) Địa chỉ mạng dùng để định nghĩa mạng băng rộng mà IPTVCD kết nối tới.
- 2) Địa chỉ host dùng để định nghĩa các thiết bị IPTV.



Hình 3.16: Định dạng gói video IPv4

**Bảng 3.6: Cấu trúc gói video IPv4**

Tên trường	Chức năng
Phiên bản	Chỉ ra phiên bản của IP được sử dụng trên mạng _ IPv4 hay IPv6.
Độ dài Header	Mô tả kích thước của header, giúp cho IPTVCD nhận ra payload của dữ liệu video bắt đầu ở đâu.
Loại dịch vụ	Còn được biết là điểm mã phân biệt dịch vụ (DSCP_Differentiated services code point). Trường này khá quan trọng cho việc phân phối nội dung IPTV bởi vì nó cho nhà cung cấp dịch vụ thiết đặt loại nội dung được mang trong gói dữ liệu. Thông tin này sau đó được xử lý bởi bộ định tuyến IP trong hệ thống mạng. Điều này cho phép các bộ định tuyến chấp nhận và làm cho mức chất lượng dịch vụ thích hợp với những loại lưu lượng khác nhau.
Tổng chiều dài	Báo cho IPTVCD biết toàn bộ chiều dài của gói IPTV. Giá trị này có độ dài 16bit, có nghĩa là một gói có thể có kích thước tối đa là 65535 byte
Thẻ nhận dạng	Được bộ định tuyến sử dụng để phân một gói lớn thành nhiều mảnh nhỏ. Khi được phân mảnh, bộ định tuyến sử dụng trường này để phân biệt các mảnh khác nhau của gói ban đầu

Cờ	Các cờ xác định các kiểu phân mảnh khác nhau, hoặc gói là một mảnh, hoặc nó được cho phép phân mảnh, hoặc gói đó là mảnh cuối cùng hay có thêm mảnh khác.
Độ dài của mảnh	Khi một gói IPTV được phân mảnh và chuyển đi qua mạng thì chức năng của IPTVCD là tổng hợp các mảnh theo đúng thứ tự. Trường này đánh số mỗi mảnh, cho phép IPTVCD tổng hợp lại theo đúng thứ tự.
Thời gian sống	Vì một gói IP chuyển qua mạng, trường thời gian sống sẽ được mỗi bộ định tuyến dọc theo đường truyền kiểm tra và giá trị bên trong trường này sẽ được giảm dần. Quá trình này tiếp tục cho đến khi giá trị trường này giảm về 0. Khi điều này xảy ra có nghĩa là gói này hết hạn và sẽ bị loại bỏ. Chức năng chính của trường này là để loại bỏ các gói trên mạng mà không có khả năng truyền đến đích. Điều này sẽ làm giảm tắc nghẽn trên mạng. Giá trị của trường này trong khoảng từ 30 đến 32.
Giao thức	Trường dài 8 bit, chỉ ra loại giao thức đóng gói bên trong datagram IP. Nếu giá trị là một thì sau đó ICMP được dùng, là hai tương ứng với IGMP, là 6 được sử dụng cho lưu lượng TCP, và 17 được sử dụng cho ứng dụng UDP.
Kiểm tra tổng header	Cho phép IPTVCD tách datagram khỏi header lỗi. Gói bị lỗi thường bị IPTVCD loại bỏ hay hủy
Địa chỉ nguồn	Chứa địa chỉ IP của thiết bị gửi gói IPTV. Trong môi trường IPTV, thiết bị gửi thường là máy chủ VoD hoặc là bộ mã hóa.
Địa chỉ đích	Chứa địa chỉ của thiết bị nhận được chỉ định. Trong môi trường IPTV, địa chỉ này thường là địa chỉ của IPTVCD
Tùy chọn và đệm	Được nhà cung cấp IPTV sử dụng để cung cấp thêm các đặc điểm khác
Dữ liệu	Là nội dung video

Một điểm đáng chú ý là một vài bit đầu tiên của địa chỉ sẽ định nghĩa các bit còn lại của trường địa chỉ sẽ được phân chia thế nào cho host và mạng. Để thuận lợi cho việc sử dụng và quản lí, địa chỉ IP được chia thành các lớp khác nhau. Chi tiết địa chỉ các lớp được mô tả trong bảng dưới đây 3.7.

Bên cạnh việc chia thành các lớp, một số địa chỉ IP được dành riêng cho các mạng tư nhân. Các địa chỉ này nằm trong dải:

10.0.0.0 to 10.255.255.255

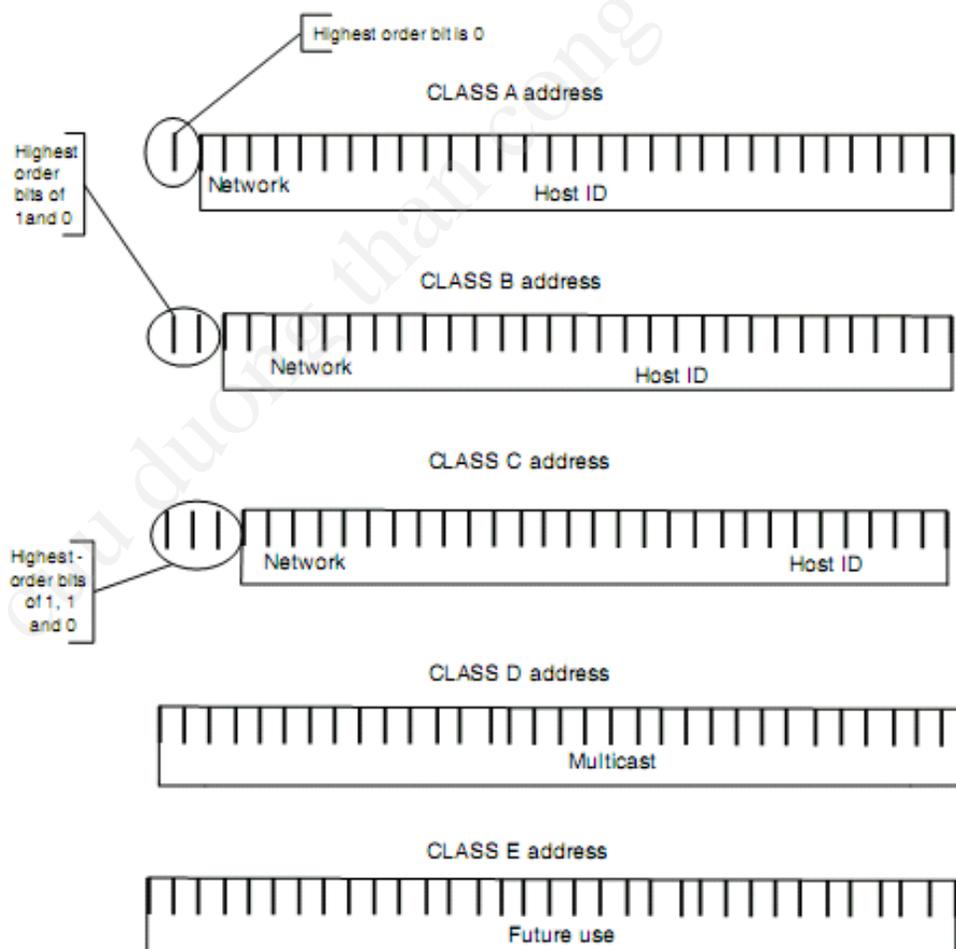
172.16.0.0 to 172.31.255.255

192.168.0.0 to 192.168.255.255

**Bảng 3.7: Các lớp địa chỉ IPv4**

Lớp địa chỉ IP	Mô tả
A	<p>Mô tả một mạng sở hữu số đầu tiên trong địa chỉ IP, có giá trị từ 0 đến 128. Ba số còn lại được dùng để xác định một IPTVCD, máy chủ, hay thiết bị mạng khác. Do đó, một địa chỉ lớp A có địa chỉ mạng 7 bit và địa chỉ host 24 bit. Bit có thứ tự cao nhất được thiết lập bằng 0</p> <p>Có 126 địa chỉ mạng lớp A trên thế giới và mỗi mạng trong số đó có đủ số địa chỉ IP để hỗ trợ hơn 16 triệu thiết bị mạng. Tất cả các địa chỉ IP lớp A được cấp phép từ InterNIC từ nhiều năm trước.</p>
B	<p>Một mạng lớp B có địa chỉ gồm số đầu tiên có giá trị từ 128 đến 191. Giá trị này tương đương với địa chỉ mạng 14 bit và một địa chỉ cục bộ 16 bit. Giá trị của 1 và 0 gán cho 2 bit có thứ tự cao nhất. Có khoảng 16000 mạng lớp B trên Internet, mỗi mạng có khả năng hỗ trợ 64000 thiết bị mạng. Những tổ chức lớn hơn và các nhà cung cấp dịch vụ Internet đã cấp phép cho hầu hết hoặc gần hết các địa chỉ này. Ví dụ về địa chỉ lớp B: 132.6.2.24, trong đó, 132.6 chỉ ra mạng, 2.24 chỉ ra host.</p>

C	Mạng lớp C có địa chỉ có số đầu tiên có giá trị từ 192 đến 223. Số này tương đương với một địa chỉ mạng 21 và địa chỉ cục bộ 8 bit. Giá trị 1,1 và 0 được gán cho 3 bit có thứ tự cao nhất. Có gần 2 triệu địa chỉ mạng lớp C, mỗi mạng có khả năng hỗ trợ đánh địa chỉ cho 254 thiết bị mạng.
D	Là phần đầu tiên của địa chỉ có giá trị từ 224 đến 239. Những địa chỉ IP này được sử dụng cho mục đích truyền đa điểm.
E	Địa chỉ mạng lớp E có giá trị từ 240 đến 247 và được đặt trước để sử dụng sau này.



Hình 3.17: Các lớp địa chỉ IP

### **Mạng con IPTV:**

Trong các mạng lớn IPTV với hàng ngàn IPTVCD trải rộng trên 1 khu vực địa lí rộng, mạng dựa trên IP này cần được chia thành các mạng nhỏ hơn gọi là mạng con. Mạng con của mạng IPTV cho phép nhà cung cấp dịch vụ định nghĩa và giám sát các phần riêng biệt trong mạng mà không cần địa chỉ IPv4 mới. Người điều hành mạng cũng sử dụng địa chỉ mạng con để giấu đi cấu trúc mạng nội bộ để tránh bị tấn công từ mạng Internet công cộng. Người quản trị mạng IPTV sử dụng các số riêng biệt được gọi là địa chỉ subnet mask để tạo các mạng con trong môi trường IPTV. Subnet mask là địa chỉ IP 32 bit. Giá trị mặc định của subnet mask của lớp A, B, C là:

Class A----- 255.0.0.0

Class B----- 255.255.0.0

Class C-----255.255.255.0

Các tiến trình chạy trên mạng con thường được quản lí bởi thiết bị phần cứng gọi là router. Một router có thể nối với nhiều mạng và quyết định thông tin sẽ được gửi đến đâu trong mạng.

### Tương lai của địa chỉ IP:

Khi cấu trúc địa chỉ IP của mạng Internet mới được phát triển trong những năm đầu của thập kỉ 80, người ta đã cho rằng nó đáp ứng được nhu cầu của người dùng hiện tại và trong tương lai. Địa chỉ IP 32 bit trong vesion IPv4 có thể đánh địa chỉ cho hơn 4 tỉ máy trạm trong khoảng 16,7 triệu mạng khác nhau. Địa chỉ IPv4 không đáp ứng được nhu cầu tốc độ phát triển của mạng Internet như hiện nay. Tốc độ phát triển của mạng Internet đã nằm ngoài dự đoán của những người phát triển giao thức IP và số lượng mạng kết nối với mạng Internet tăng lên từng tháng.

Để tìm giải pháp cho hạn chế của địa chỉ IPv4, Nhóm Kỹ Sư mạng Internet đã đưa ra version mới IPv6 với 128 bit địa chỉ. IPv6 có thể cung cấp lượng địa chỉ gấp hàng tỉ lần số địa chỉ IPv4. IPv6 có cả các khả năng cung cấp hỗ trợ việc xác thực,

tính toàn vẹn QoS, mã hóa và bí mật. Việc sử dụng IPv6 trên mạng IPTV là thích hợp bởi vì cơ chế QoS bên trong nó và có khả năng hỗ trợ số lượng không giới hạn các IPTVCD. Trong 5 tới 10 năm tới, IPv6 sẽ dần thay thế IPv4.

Tại sao lại sử dụng IPv6 khi triển khai IPTV ? IPv6 có tính năng vượt trội hơn so với IPv4. Các nguyên nhân chính khiến các nhà cung cấp dịch vụ xem xét việc sử dụng IPv6:

Quy mô được tăng lên: IPv4 dùng địa chỉ 32 bit trong khi IPv6 dùng địa chỉ 128 bit, lớn hơn gấp 4 lần. Điều này cho phép nhà cung cấp dịch vụ IPTV mở rộng số lượng thiết bị có thể quản lý .

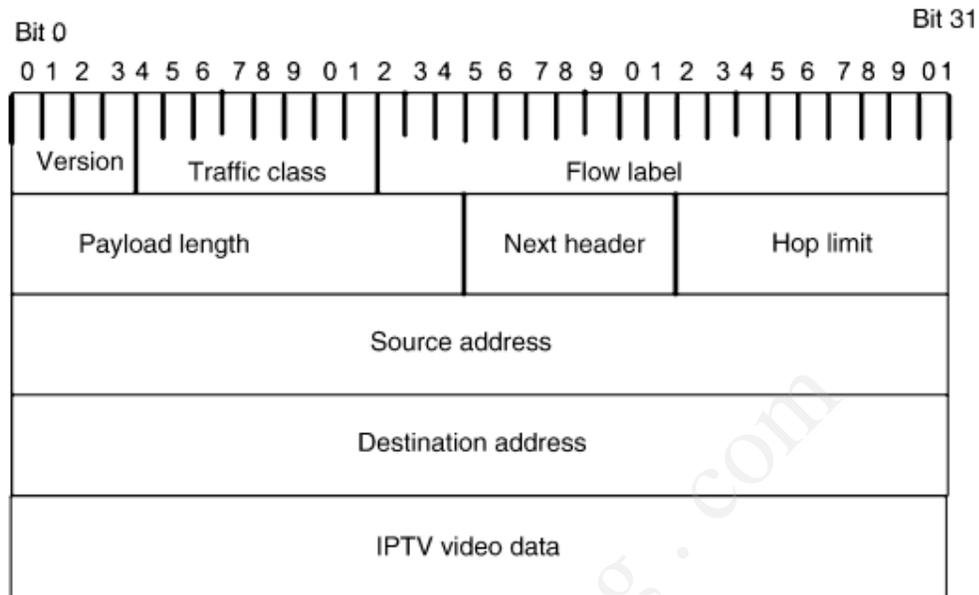
Cấu trúc header đơn giản: IPv6 giảm kích thước header xuống còn 40 byte cố định và đơn giản cấu trúc của trường header. Các thành phần cơ bản của IPv6 header được mô tả trong hình 42 và bảng dưới.

Tăng mức độ bảo vệ: IPv6 có 2 đặc điểm giúp tăng mức độ bảo vệ:

- 1) Bao gồm header xác thực-- bao gồm các bản tin xác nhận và kiểm tra người gửi gói.
- 2) Payload bảo mật được đóng gói-- đặc điểm này đảm bảo tính toàn vẹn của dữ liệu IPTV và bảo mật giữa các máy chủ trung tâm dữ liệu IPTV và các IPTVCD khác.

Lưu lượng thời gian thực tốt hơn: khả năng dán nhãn luồng của IPv6 cho phép nhà cung cấp dịch vụ đánh dấu các gói riêng, phụ thuộc vào từng loại dịch vụ. Trong môi trường triple-play, các router có thể coi các gói IP được dán nhãn với một nhận dạng video là các gói IP được dán nhãn khi mang nội dung web.

Tự động cấu hình: Khả năng plug and play của IPv6 giúp giảm bớt độ phức tạp khi cài đặt dịch vụ IPTV tại nhà của khách hàng.



Hình 3.18: Cấu trúc header của IPv6

Bảng 3.8: Mô tả trường của IPv6

Tên trường	Chức năng
Mã bắt đầu của gói	Các gói PES bắt đầu bằng các bit 0
Định dạng dòng	Trường này dùng để định nghĩa loại payload chứa trong gói. Dạng 111xxxxx là chỉ gói âm thanh, và 1110 xxxx chỉ các gói hình ảnh. “x” là số dòng MPEG
Độ dài gói PES	2 byte của trường này dùng để định nghĩa độ dài gói
Mã đồng bộ	Trường này được sử dụng để đồng bộ âm thanh và nội dung video cung cấp phần cứng và phần mềm trong bộ mã hóa đặt trong set-top box IP với các thông tin thêm. Trong trường này còn bao gồm các Cờ: Cờ điều khiển trộn âm PES: báo cho bộ mã hóa biết liệu gói có được đảm bảo không trong quá trình xử lý trộn âm. Cờ ưu tiên PES: Cung cấp thông tin về mức ưu tiên của gói PES cho bộ giải mã. Data_alignment_indicator : Cờ chỉ thị này sẽ xác định xem phần

	payload của PES bắt đầu với bit âm thanh hay hình ảnh. Thông tin bản quyền: khi bit này được thiết lập, chương trình đó là chương trình có bản quyền. Cờ nguyên bản hay bản sao: Cờ này chỉ thị liệu chương trình đó có phải là bản gốc, hay là bản sao.
Độ dài dữ liệu header	Trường này dùng để định nghĩa tổng số byte được sử dụng bởi các trường header khác.
Các trường header trong PES	Trường này chứa các số các bit tùy chọn.
Trường payload và dữ liệu video	Trường payload của PES chứa các dòng cơ bản âm thanh.

Vì những ưu điểm kể trên IPv6 được xem như giải pháp lâu dài để hỗ trợ triển khai các thiết bị số ở quy mô lớn, có thể sử dụng nhiều loại ứng dụng dựa trên IP.

Nhược điểm chính khi sử dụng giao thức IP là không có gì đảm bảo rằng khi nào các gói tới đúng đích hay gói có đến đúng lúc không. Ngay cả thứ tự các gói được chuyển đến cũng không được xác định. Do đó, lớp IP làm việc cùng với giao thức lớp truyền tải để đảm bảo rằng các gói đến IPTVCD đúng lúc và theo trật tự đúng. IP cũng làm cho quá trình phân phát nội video bị trễ.

### 3.2.8 Lớp liên kết dữ liệu

Lớp liên kết dữ liệu lấy các dữ liệu thô từ lớp IP và định dạng chúng thành các gói phù hợp để truyền qua mạng vật lí. Chú ý, lớp liên kết dữ liệu khác với các giao thức mạng. Kỹ thuật Ethernet là một trong những kỹ thuật phổ biến hơn được sử dụng trong hệ thống IPTV. Lớp liên kết dữ liệu bao gồm các chức năng dành cho các mạng dựa trên Ethernet:

Encapsulation---- Lớp này thêm vào các gói IPTV 1 header. Ethernet header là loại Encapsulation phổ biến nhất dùng trong lớp liên kết dữ liệu của IPTVCD. Các thành phần cơ bản của Ethernet header được giải thích trong bảng 3.9

Định địa chỉ---- Lớp liên kết dữ liệu xử lý các địa chỉ vật lý của mạng người sử dụng và các thiết bị chủ. Hệ thống địa chỉ khác nhau với các topo mạng. Ví dụ, địa chỉ MAC được sử dụng trong mạng Ethernet. Mỗi thiết bị kết nối với mạng IPTV thì có 1 địa chỉ MAC. Độ dài của địa chỉ MAC là 48 bit và thường được biểu diễn bằng 12 số trong hệ 16. Trong 12 số hệ 16 này, 6 số đầu tiên để dành cho nhà sản xuất thiết bị IPTV và các số còn lại được dùng để định nghĩa giao diện mạng ảo.

Kiểm tra lỗi----- chức năng kiểm tra lỗi được dùng trong vài lớp của mô hình IPTV, bao gồm cả lớp liên kết dữ liệu. Các gói bị ngắt là lỗi thường gặp trong quá trình truyền các nội dung video qua mạng dựa trên IP. Phương pháp sửa lỗi thường dùng là kiểm tra dư thừa vòng (CRC) trong IPTV để tìm và loại bỏ các gói bị ngắt. Sử dụng kỹ thuật CRC thiết bị gửi IPTV thực hiện việc tính toán trên các gói và lưu trữ kết quả trong gói. Các phép tính toán tương tự cũng được thực hiện trên thiết bị nhận khi nhận được các gói. Nếu kết quả tính toán là như nhau, thì các gói được xử lý bình thường. Tuy nhiên, nếu kết quả này là khác nhau, thì gói bị lỗi sẽ bị loại bỏ. Thiết bị gửi sẽ tạo một gói mới và gửi lại nó. Thông báo với lớp trên trong mô hình IPTV khi có lỗi xảy ra là nhiệm vụ chính của lớp liên kết dữ liệu trong kỹ thuật kiểm tra lỗi mà các hệ thống IPTV end to end.

Điều khiển luồng---- Điều khiển luồng là một trong chức năng của lớp truyền tải. Trong mạng IPTV, điều khiển luồng cho thiết bị IPTV của người sử dụng không bị tràn bởi các nội dung. Lớp liên kết dữ liệu cùng với lớp truyền tải thực hiện bất kỳ yêu cầu điều khiển luồng nào.

**Bảng 3.9: Cấu trúc của Ethernet header**

<b>Tên trường trong Header</b>	<b>Kích thước (bit)</b>	<b>Chức năng</b>
Địa chỉ đích mạng Ethernet	48	Chỉ ra địa chỉ của giao diện đích
Địa chỉ nguồn mạng Ethernet	48	Chỉ ra địa chỉ của giao diện nguồn
Loại mã	16	Chỉ ra giao thức được sử dụng trong việc định dạng gói. Ví dụ kiểu gói "TCP/IP" chứa giá trị kiểu hex "0 x 80 0 x 00"

**Bảng 3.10: Cấu trúc khung Ethernet được dùng để mang nội dung MPEG-2**

<b>Mô tả</b>	<b>Kích thước</b>
Các dòng gói truyền tải MPEG bảy	Mỗi gói có kích thước 188 byte (184 byte chứa nội dung video, cộng với 4 byte được dùng làm thông tin header). Gói MPEG - TS bảy này chiếm 1316 byte (10528 bit) của khung
Tiêu đề RTP	Tiêu đề này chiếm 12 byte của khung Ethernet
Tiêu đề UDP	Tiêu đề này chiếm 8 byte của khung Ethernet
Tiêu đề IP	Tiêu đề này chiếm 20 byte của khung Ethernet
Tiêu đề Ethernet	Tiêu đề này chiếm 14 byte của khung Ethernet

### 3.2.9 Lớp vật lí

Lớp vật lí quy định luật lệ truyền các bit số qua mạng. Nó đề cập đến việc đưa các dữ liệu qua các mạng vật lí riêng biệt như x DSL, và không dây. Lớp này định nghĩa cấu hình mạng vật lí, thông số kĩ thuật, điện trong môi trường truyền.

Khi dòng bit được truyền qua mạng, các gói được chuyển từ lớp thấp đến lớp cao trong mô hình truyền thông IPTV. Ví dụ lớp liên kết dữ liệu sẽ kiểm tra các gói và loại bỏ đi phần header Ethernet và trường sửa lỗi CRC. Tiếp đó sẽ kiểm tra trường kiểu mã của Ethernet header và xác định gói cần được xử lí bởi giao thức IP. Do đó gói dữ liệu được chuyển lên lớp mạng. Lớp mạng kiểm tra và loại bỏ đi phần IP header và chuyển gói đó lên lớp truyền tải. Phương pháp bỏ đi phần header khi qua các lớp khác nhau gọi là bóc gói. Quá trình này tiếp tục được thực hiện cho tới khi gói dữ liệu lên đến tầng trên cùng trong mô hình. Hình ảnh gốc được thể hiện trên màn hình TV của người xem.

**Bảng 3.11 Tổng kết các lớp trong mô hình IPTV**

Số thứ tự lớp	Tên lớp	Tổng quan
1	Vật lí	Xác định các thuộc tính của các phương tiện truyền thông trong mạng chịu trách nhiệm truyền các bit dữ liệu IPTV.
2	Liên kết dữ liệu	Quản lí các cơ chế được sử dụng để truy cập thiết bị truyền thông trong mạng. Kiểm soát lỗi, đồng bộ và điều khiển luồng là những chức năng khác mà lớp này cung cấp.
3	IP	Chức năng của lớp này có quan hệ với việc định tuyến các gói IP truyền trên mạng. Các cơ chế như là đánh địa chỉ và điều khiển tắc nghẽn thường được lớp IPTVCM này sử dụng.
4	Giao vận	Chức năng chính của lớp này là để đảm bảo rằng các gói IPTV đến được đích. TCP và UDP đều vận hành tại lớp này.

5	RTP (tùy chọn)	Mặc dù UDP là giao thức được lớp giao vận ưu tiên sử dụng hơn cho việc phân phối nội dung IPTV trên mạng băng rộng, nhưng nó cũng không phải là giao thức đáng tin cậy và nó không hỗ trợ việc sửa lỗi hay giải quyết vấn đề các gói đến IPTVCD không theo đúng thứ tự. Do đó, một số các nhà cung cấp dịch vụ sử dụng RTP để giải quyết những thiếu hụt cố hữu trong giao thức UDP.
6	Xây dựng dòng truyền tải	Lớp này đóng gói các dòng bit video và audio. Các gói có độ dài 188 byte.
7	Lớp đóng gói Video	Lớp này tạo một dòng các gói PES đã được đánh dấu thời gian.
8	Mã hóa Video	Các dòng audio và video cơ bản làm thành nền tảng cho lớp IPTVCM. Định dạng của dòng truyền được sử dụng ở lớp này phụ thuộc vào thuật toán nén sử dụng bởi bộ mã hóa.

Dịch vụ IPTV với những ưu điểm nổi trội so với các chuẩn truyền hình truyền thống đã và đang được phát triển mạnh mẽ trên thế giới cũng như ở Việt Nam. Đi kèm với IPTV là rất nhiều dịch vụ giá trị gia tăng khác như: truyền hình theo yêu cầu, mạng giáo dục từ xa, hội thảo từ xa ...

Trong phạm vi đề án này chỉ nêu phương pháp đóng gói dữ liệu video qua từng lớp. Chuẩn nén H.264 này không chỉ được sử dụng trong IPTV mà còn được ứng dụng trong hầu hết các công nghệ truyền hình số hiện đại như: truyền hình số mặt đất, truyền hình số vệ tinh, truyền hình cáp, truyền hình di động đến truyền hình phân giải cao HDTV. Đóng gói dữ liệu video qua từng lớp cho thấy cấu trúc dữ liệu trong từng lớp và các cơ chế điều khiển luồng, cơ chế sửa lỗi, đảm bảo chất lượng dịch vụ làm hài lòng người xem

## CHƯƠNG 4:

### KẾT LUẬN VÀ KIẾN NGHỊ

#### Kết luận

Qua quá trình nghiên cứu chuẩn H.264 cho thấy những yêu điểm nổi trội của nó so với các chuẩn khác. Nó cho tỉ lệ nén rất cao nhờ đó có thể truyền Video qua những đường truyền tốc độ thấp ( x.64kbps).

Vì H.264 là nén tổn hao nên khi tỉ lệ nén cao thì chất lượng giảm. Vậy nên phải tính toán để cân bằng giữa chất lượng và dung lượng.

H.264 có những yêu điểm nhưng ngược lại nó sử dụng những thuật toán phức tạp nên khi xử lý real time đòi hỏi tốc độ xử lý phải cao.

Nhờ có H.264 mà IPTV có tính khả thi cao.

Chuẩn nén H.264 này không chỉ được sử dụng trong IPTV mà còn được ứng dụng trong hầu hết các công nghệ truyền hình số hiện đại như: truyền hình số mặt đất, truyền hình số vệ tinh, truyền hình cáp, truyền hình di động đến truyền hình phân giải cao HDTV. Nhưng luận văn mới đi sâu tìm hiểu được ứng dụng của chuẩn nén H.264 trong IPTV. Đóng gói dữ liệu video qua từng lớp cho thấy cấu trúc dữ liệu trong từng lớp và các cơ chế điều khiển luồng, cơ chế sửa lỗi, đảm bảo chất lượng dịch vụ làm hài lòng người xem

#### Kiến nghị:

Luận văn mới dừng lại ở việc nghiên cứu lý thuyết chưa can thiệp được vào code để thực thi việc mã hóa và giải mã H.264 cũng như làm những bộ codec bằng phần cứng.

Hướng phát triển tương lai của luận văn có thể là can thiệp vào code để làm cho bộ codec tối ưu hơn như điều khiển tốc độ (rate control), xây dựng một mô hình

IPTV hoàn chỉnh trong đó dữ liệu video được mã hóa H.264 rồi truyền trong mạng IPTV và được giải mã khi nhận được tín hiệu.

*CuuDuongThanCong.com*

**TÀI LIỆU THAM KHẢO**

- [1] Iain E. G. Richardson, (2003), *H.264 and MPEG-4 Video Compression Video Coding for Next-generation Multimedia*, John Willey & Son
- [2] Thomas Wiegand, Gary J. Sullivan, Gisle Bjontegaard, and Ajay Luthra, (6/2003), “Overview of the H.264 / AVC Video Coding Standard”, IEEE Transactions on circuits and systems for video technology.
- [3] Đỗ Hoàng Tiến và Dương Thanh Phương, (2004), *Giáo trình Kỹ thuật truyền hình*, Nhà xuất bản KHKT, Hà Nội
- [4] Nguyễn Thanh Bình và Võ Nguyễn Quốc Bảo, (2007), Xử lý âm thanh, hình ảnh, Học viện Công nghệ bưu chính viễn thông.
- [5] Wes Simpson & Howard Greenfield, (2005), *IPTV and Internet Video: New Markets in Television Broadcasting*, Focal Press, United States of America.
- [6] *RFC 3626- request for comments*
- [7] Jirka Klaue, Berthold Rathke, and Adam Wolisz, (2003), EvalVid-A Framework for Video Transmission and Quality Evaluation, Telecommunication Network Group (TKN).