

ĐẠI HỌC QUỐC GIA HÀ NỘI  
**TRƯỜNG ĐẠI HỌC CÔNG NGHỆ**

**CHU TRẦN HÙNG**

**NGHIÊN CỨU VÀ PHÂN TÍCH  
BỘ MÃ HÓA VIDEO DÙNG CHO MẠNG SENSOR**

Ngành : Công nghệ Kỹ thuật Điện tử, Truyền thông

Chuyên ngành : Kỹ thuật Điện tử

Mã số : 60520203

**LUẬN VĂN THẠC SĨ  
CÔNG NGHỆ KỸ THUẬT ĐIỆN TỬ TRUYỀN THÔNG**

NGƯỜI HƯỚNG DẪN KHOA HỌC: **PGS.TS TRỊNH ANH VŨ**

**TS. HOÀNG VĂN XIÊM**

**Hà Nội - 2016**

## LỜI CAM ĐOAN

Tôi xin cam đoan số liệu và kết quả nghiên cứu trong luận văn này là trung thực và chưa hề được sử dụng để bảo vệ học vị nào. Mọi sự giúp đỡ trong luận văn này đã được cảm ơn và các thông tin trích dẫn trong luận văn đã được chỉ rõ nguồn gốc rõ ràng và được phép công bố.

Hà Nội, ngày ... tháng ... năm 2016

Người thực hiện:

Chu Trần Hùng

# MỤC LỤC

## LỜI CAM ĐOAN

## MỤC LỤC

CÁC THUẬT NGỮ VIẾT TẮT

## DANH MỤC BẢNG

## DANH MỤC SƠ ĐỒ, HÌNH VẼ

LỜI NÓI ĐẦU .....	1
<b>CHƯƠNG 1: CƠ SỞ LÝ THUYẾT CỦA KỸ THUẬT MÃ HÓA VIDEO</b>	
<b>TRUYỀN THÔNG .....</b>	<b>2</b>
1.1. Yêu cầu mã hóa video.....	2
1.2. Các kỹ thuật mã hóa video then chốt.....	3
1.2.1. Mã hóa sai khác giữa ảnh thực tế và ảnh dự đoán .....	3
1.2.2. Mã hóa trong miền biến đổi.....	3
1.2.3. Phép lượng tử tuyến tính .....	4
1.2.4. Các phép dự đoán ảnh .....	5
1.2.5. Mã hóa Entropy.....	7
1.3. Các chuẩn mã hóa video phổ biến.....	9
1.3.1. Chuẩn H.264/AVC.....	9
1.3.2. Chuẩn H.265/HEVC .....	10
1.4. Nhược điểm của kỹ thuật mã hóa video truyền thống .....	12
1.4.1. Độ phức tạp cao tại phía mã hóa .....	12
1.4.2. Khả năng chống chịu nhiễu thấp .....	13
<b>CHƯƠNG 2: CƠ SỞ KỸ THUẬT CỦA MÃ VIDEO PHÂN TÁN .....</b>	<b>15</b>
2.1. Mã hóa dự đoán cổ điển.....	15
2.2. Định lý Slepian-Wolf.....	15
2.3. Định lý Winer-Ziv .....	17
2.4. Ví dụ minh họa .....	18
<b>CHƯƠNG 3: CÁC MÔ HÌNH THỰC NGHIỆM CỦA KỸ THUẬT MÃ VIDEO</b>	
<b>PHÂN TÁN .....</b>	<b>20</b>
3.1. Giải pháp STANFORD.....	20
3.1.1. Biến đổi và lượng tử hóa .....	21
3.1.2. Mã hóa kênh và bộ đệm .....	21
3.1.3. Tỷ lệ ước tính tối thiểu .....	21
3.1.4. Phân thông tin phụ được khai thác .....	21
3.1.5. Mô hình kênh ảo và tính toán đầu vào mềm .....	22
3.1.6. Giải mã Kênh và Kiểm tra CRC.....	22
3.1.7. Sự khôi phục và biến đổi ngược .....	22
3.2. Giải pháp PRISM .....	22
3.2.1. Quá trình huấn luyện.....	24
3.2.2. Quá trình mã hóa .....	26

3.2.3. <i>Quá trình giải mã</i> .....	32
<b>CHƯƠNG 4: ĐÁNH GIÁ ƯU NHƯỢC ĐIỂM BỘ MÃ HÓA DVC</b> .....	<b>33</b>
4.1. Điều kiện đánh giá.....	33
4.2. Đánh giá hiệu năng nén .....	35
4.3. Đánh giá độ phức tạp.....	38
<b>KẾT LUẬN</b> .....	<b>42</b>
<b>TÀI LIỆU THAM KHẢO</b> .....	<b>44</b>

## CÁC THUẬT NGỮ VIẾT TẮT

Từ viết tắt	Nghĩa tiếng Anh	Nghĩa tiếng Việt
AMVP	Advanced Motion Vector Predictor	Bộ dự đoán vector chuyển động nâng cao
AVC	Advanced Video Coding	Mã hóa video tiên tiến
BAC	Binary Arithmetic Coding	Mã hóa số học nhị phân
CABAC	Content Adaptive Binary Arithmetic Coding	Mã hóa số học nhị phân thích nghi theo thuộc tính
CAVLC	Content Adaptive Variable Length Coding	Mã chiều dài thay đổi thích nghi theo thuộc tính
CPU	Central Processing Unit	Đơn vị xử lý trung tâm
CTU	Coding Tree Unit	Đơn vị cây mã hóa
CU	Coding Unit	Đơn vị mã hóa
DCT	Discrete Cosine Transform	Biến đổi Cosine rời rạc
DPCM	Differential Pulse Code Modulation	Điều xung mã vi sai
DPB	Decoded Picture Buffer	Bộ đệm hình ảnh đã giải mã
DRAM	Dynamic Random Access Memory	Bộ nhớ truy cập ngẫu nhiên động
DVC	Distributed Video Coding	Mã hóa video phân tán
GOP	Group Of Picture	Nhóm ảnh
GMC	Global Motion Compensation	Bù chuyển động toàn phần
HD	High Definition	Độ phân giải cao
HEVC	High Efficiency Video Coding	Mã hóa video hiệu suất cao
HVS	Human Visual System	Hệ thống trực quan con người
IDCT	Inverse Discrete Cosine Transform	Biến đổi cosin rời rạc ngược
IEC	International Electrotechnical Commission	Ủy ban Kỹ thuật Điện Quốc tế
ISO	International Organization for Standardization	Tổ chức tiêu chuẩn hóa quốc tế

ITU-T	International Telecommunication Union Telecommunication Standardization Sector	Khu vực tiêu chuẩn hóa viễn thông thuộc Tổ chức viễn thông quốc tế Liên hợp quốc
JPEG	Joint Photographic Experts Group	Chuẩn nén ảnh của ủy ban quốc tế
MB	Macroblock	Khối lớn
MC	Motion Compensated	Bù chuyển động
MCP	Motion Compensated Prediction	Dự đoán bù chuyển động
MPEG	Moving Picture Experts Group	Nhóm các chuyên gia hình ảnh động
MPS	Most Probable Symbol	Biểu tượng có khả năng xuất hiện nhiều nhất
MVD	Motion Vector Difference	Sự khác biệt vector chuyển động
MVP	Motion Vector Predictor	Bộ dự đoán vector chuyển động
PPS	Picture Parameter Set	Tập hợp tham số hình ảnh
PSNR	Peak Signal to Noise Ratio	Tỉ lệ tín hiệu đỉnh/ tạp âm
PU	Prediction Unit	Đơn vị dự đoán
QG	Quantization Group	Nhóm lượng tử hóa
QP	Quantization Parameter	Tham số lượng tử hóa
RDO	Rate Distortion Optimization	Tối ưu hóa tốc độ/ méo
SAO	Sample Adaptive Offset	Bù thích nghi mẫu
SD	Standard Definition	Định dạng tiêu chuẩn
SPS	Sequence Parameter Set	Tập hợp tham số chuỗi
SRAM	Static Random Access Memory	Bộ nhớ truy cập ngẫu nhiên tĩnh
TU	Transform Unit	Đơn vị biến đổi
UHD	Ultra High Definition	Định dạng cực cao
VLC	Variable Length Code	Mã có chiều dài thay đổi
VLSI	Very Large Scale Integration	Tích hợp với quy mô rất rộng

## DANH MỤC BẢNG

Bảng 1.1. Giá trị lượng tử và xác suất tương ứng .....	7
Bảng 1.2. Từ mã Huffman và lượng bit cần mã hóa tương ứng .....	8
Bảng 4.1. Bảng mô tả tóm tắt các thông số sử dụng đánh giá .....	33
Bảng 4.2: Bảng lượng tử mã hóa DVC .....	34
Bảng 4.3: Giá trị lượng tử cho khung chính tại GOP=2, QCIF 15Hz .....	35

## DANH MỤC SƠ ĐỒ, HÌNH VẼ

Hình 1.1 Mô tả ví dụ về video với độ phân giải 2K và kích thước tương ứng của các loại video khác nhau.....	2
Hình 1.2: Minh họa sự sai khác giữa 2 khung liên tiếp trong dãy video .....	3
Hình 1.3: Giá trị của điểm ảnh trong miền pixel và miền DCT tương ứng .....	4
Hình 1.4. Mô hình phép lượng tử tuyến tính sử dụng trong mã hóa video .....	5
Hình 1.5. Mô tả ảnh gốc, ảnh dự đoán và ảnh dư thừa [20] .....	6
Hình 1.6. Tạo ảnh dự đoán trong khung.....	6
Hình 1.7. Tạo ảnh dự đoán liên khung .....	7
Hình 1.8. Sử dụng thuật toán cây tạo ra từ mã Huffman.....	8
Hình 1.9. Kiến trúc tổng quát của chuẩn H.264/AVC [3].....	9
Hình 1.10. Kiến trúc tổng quát của chuẩn H.254/HEVC [1].....	10
Hình 1.11: Phạm vi dò tìm vectơ dịch chuyển của block chứa bánh xe .....	12
Hình 1.12:Ước lượng véc tơ dịch chuyển.....	13
Hình 2.1: Mô hình mã hóa dự đoán cổ điển .....	15
Hình 2.2: Sơ đồ mã nguồn phụ thuộc thống kê.....	16
Hình 2.3: Biểu đồ vùng tỉ lệ tốc độ mã giữa 2 nguồn X,Y .....	16
Hình 2.4: Mô tả định lý Slepian-Wolf với thông tin phụ .....	17
Hình 2.5: Mã hóa mất mát thông tin với thông tin phụ ở phần giải mã.....	17
Hình 3.1. Cấu trúc khám phá thuật nén và giải nén [15].....	20
Hình 3.2: Sơ đồ mã hóa và giải mã PRISM [9] .....	24
Hình 3.3: Quét zig- zag.....	26
Hình 3.4: Mô hình mã Syndrome [8] .....	29
Hình 3.5: Mặt phẳng bit syndrome [9] .....	30
Hình 3.6: Chương trình quét xoắn ốc .....	32
Hình 4.1 Mô tả khung hình đầu tiên của 4 chuỗi video .....	33
Hình 4.2: Biểu đồ so sánh hiệu năng nén - Foreman .....	36
Hình 4.3: Biểu đồ so sánh hiệu năng nén - Hall monitor .....	36
Hình 4.4: Biểu đồ so sánh hiệu năng nén - Coast guard.....	37
Hình 4.5: Biểu đồ so sánh hiệu năng nén – Soccer.....	37
Hình 4.6: Biểu đồ so sánh thời gian mã hóa - Foreman .....	39
Hình 4.7: Biểu đồ so sánh thời gian mã hóa - Hall monitor .....	39
Hình 4.8: Biểu đồ so sánh thời gian mã hóa - Coastguard .....	40
Hình 4.9: Biểu đồ so sánh thời gian mã hóa - Soccer .....	40



## LỜI NÓI ĐẦU

Ngày nay, công nghệ mã hóa video đã và đang được sử dụng rất phổ biến, từ các ứng dụng truyền thống như truyền hình quảng bá, truyền hình hội nghị đến các ứng dụng mới xuất hiện như mạng cảm biến (sensor), mạng giám sát từ xa... Các công nghệ đều dựa trên mô hình mã hóa video dự đoán, với các chuẩn mã hóa phổ biến như MPEG-2/Video, H264/AVC hay H.265/HEVC. Mặc dù cho hiệu quả mã hóa cao, do khai thác có hiệu quả thông tin tương quan giữa các khung hình tại phía phát, công nghệ mã hóa này cũng luôn đi cùng với mức độ yêu cầu cao về độ phức tạp thuật toán tại phía phát, do đó cần những thiết bị hiện đại, đắt tiền ở bên phát. Mô hình mã hóa này phù hợp với hệ thống truyền hình quảng bá, ở đó đài truyền hình với nguồn tài chính mạnh đầu tư các thiết bị đắt tiền để phát sóng. Ở phía người dùng, bộ giải mã khá là đơn giản giúp cho người dùng có thể dễ dàng xem các chương trình với bộ đầu thu có giá thành rẻ.

Công nghệ ngày càng phát triển, các mô hình mạng video khác cũng được triển khai trong đời sống, chẳng hạn như các mạng video giám sát giao thông, camera an ninh lắp đặt trong các nhà hàng, bệnh viện... Những hệ thống này có đòi hỏi hoàn toàn khác so với hệ thống truyền hình quảng bá là yêu cầu phần phát đơn giản, gọn nhẹ, rẻ tiền trong khi phần phức tạp có thể chuyển về bộ xử lý trung tâm ở phía thu, nơi sẽ được đầu tư thiết bị tập trung và hiện đại hơn. Do đó các chuẩn mã hóa video truyền thống sẽ không còn phù hợp nữa. Yêu cầu đặt ra là phát triển mô hình mã hóa thế nào để đơn giản phần mã hóa mà vẫn không làm tổn thất đáng kể về hiệu suất nén so với các mô hình truyền thống.

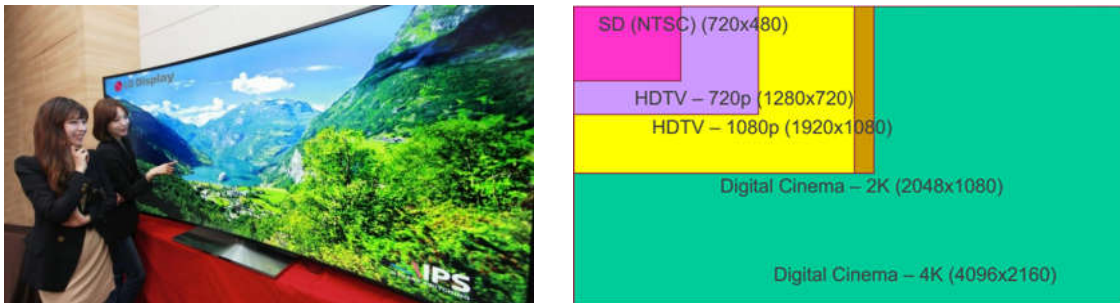
Trong nỗ lực đưa ra câu trả lời cho nhu cầu mã hóa video đáp ứng nhu cầu thực tiễn này, một mô hình mã hóa video mới được nghiên cứu, đã và đang được phát triển hiện nay đó là mô hình mã hóa video phân tán DVC (Distributed Video Coding). Mô hình này dựa trên kết quả của 2 định lý trong thuyết thông tin là các định lý của Slepian- Wolf và Wyner- Zip. Theo hướng nghiên cứu này luận văn tập trung trình bày về mô hình mã video phân tán DVC; bên cạnh đó, luận án cũng đưa ra mô hình mã hóa truyền thống để phân tích. Từ đó luận văn đưa ra những so sánh đánh giá ưu, nhược điểm của mô hình mới và các hướng phát triển tiếp theo cho mô hình này.

# CHƯƠNG 1: CƠ SỞ LÝ THUYẾT CỦA KỸ THUẬT MÃ HÓA VIDEO TRUYỀN THÔNG

Ra đời từ cách đây hơn bốn mươi năm, mã hóa video đã và đang đóng một vai trò vô cùng quan trọng trong sự phát triển cũng như trưởng thành của truyền thông số, truyền hình và Internet. Chương này sẽ tập trung vào điểm lại một số kỹ thuật then chốt được sử dụng trong các chuẩn mã hóa video truyền thống.

## 1.1. Yêu cầu mã hóa video

Một tín hiệu video số thường chứa một lượng lớn dữ liệu; do đó sẽ gặp rất nhiều khó khăn trong việc lưu trữ và truyền đi trên một kênh truyền có băng thông hạn chế. Ngoài ra, với sự phát triển của khoa học kỹ thuật, ngày nay đã sản xuất được bộ cảm biến màu có độ phân giải lên đến hàng chục triệu pixel và thực tế đã ứng dụng độ phân giải như  $1920 \times 1080$  pixel, hoặc lớn hơn là chuẩn 2K, 4K. Khi đó việc biểu diễn các thông tin video này càng tốn nhiều dữ liệu hơn.



(a)

(b)

*Video với độ phân giải 2K*

*Kích thước tương ứng của một số loại video*

*Hình 1.1 Mô tả ví dụ về video với độ phân giải 2K và kích thước tương ứng của các loại video khác nhau.*

Do đó để có thể tiết kiệm không gian lưu trữ và băng thông kênh truyền thì luôn cần phải mã hóa (nén) tín hiệu video.

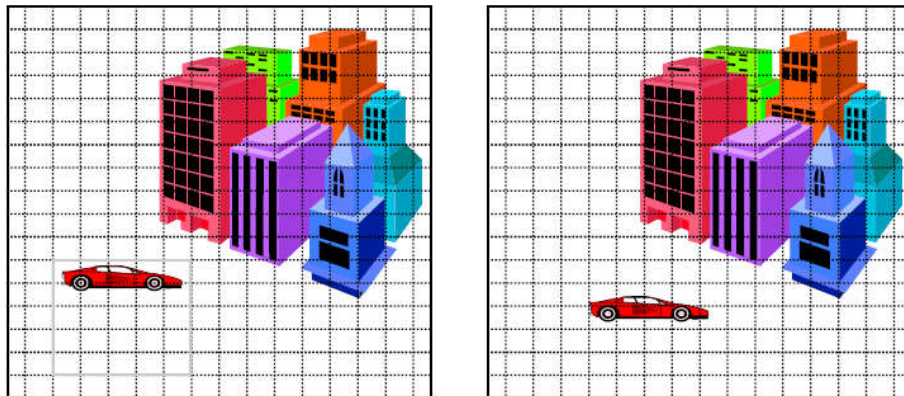
Quá trình nén ảnh thực hiện được là do thông tin trong bức ảnh có tổ chức, có trật tự, vì vậy nếu xem xét kỹ tính trật tự, cấu trúc của bức ảnh sẽ phát hiện và loại bỏ được các lượng thông tin dư thừa, chỉ giữ lại các thông tin quan trọng nhằm giảm số lượng bit khi lưu trữ cũng như khi truyền mà vẫn đảm bảo được thông tin hiển thị của bức ảnh. Tại đầu thu, bộ giải mã sẽ tổ chức, sắp xếp lại được bức ảnh xấp xỉ gần chính xác so với ảnh gốc nhưng vẫn đảm bảo thông tin cần thiết.

## 1.2. Các kỹ thuật mã hóa video then chốt

Mặc dù đã ra đời từ cách đây hơn bốn mươi năm, trải qua rất nhiều giai đoạn phát triển, từ các chuẩn MPEG 1 (1992) [1], MPEG 2 (1999) [2], H.264/AVC (2003) [3] đến chuẩn H.265/HEVC (2013)[4], các kỹ thuật căn bản sử dụng trong mã hóa video truyền thống như kỹ thuật mã hóa sự khác biệt, mã hóa trong miền biến đổi cô sin rời rạc, phép lượng tử tuyến tính, phép nội suy trong ảnh, liên ảnh, hay mã hóa entropy đều được giữ nguyên lại. Điều này phản ánh giá trị khoa học sâu sắc của các nghiên cứu kể trên.

### 1.2.1. Mã hóa sai khác giữa ảnh thực tế và ảnh dự đoán

Video là một chuỗi các bức ảnh liên tiếp cùng mô tả một nội dung và chứa đựng một thông tin, câu chuyện xuyên suốt nào đó. Do vậy, giữa các bức ảnh liên tiếp trong video luôn tồn tại các mối tương quan lớn như được mô tả ở hình 1.2.



Hình 1.2: Minh họa sự sai khác giữa 2 khung liên tiếp trong dãy video

Nhìn vào 2 khung hình trên (trong chuỗi khung liên tiếp của video), ta có thể thấy sự khác biệt căn bản nằm ở vị trí ô tô, do chuyển động, nên:

- Khi mã hóa hiệu 2 khung hình (theo pixel tương ứng) thì chỉ cần dùng một lượng bit ít hơn so với việc mã hóa toàn bộ thông tin trong ảnh thực tế.
- Nếu có cách *dự đoán* khung 2 từ khung 1 để có sự sai khác giữa khung 2 và khung dự đoán thì mã hóa còn có thể tốn ít bit mã hơn. Điều này là khả thi nếu có kỹ thuật dự đoán tốt. Tuy nhiên giá phải trả cho dự đoán tốt, như sẽ nói ở mục sau chính là sự gia tăng độ phức tạp tính toán tại phía mã hóa.

### 1.2.2. Mã hóa trong miền biến đổi

Trong miền điểm ảnh (pixel), các giá trị dư thừa thường không tập trung năng lượng một cách có hệ thống. Do vậy, để nâng cao tính hiệu quả của mã hóa thông tin dư thừa, biến đổi cô sin rời rạc (Discrete Cosine Transform - DCT) đã được thông qua

và sử dụng trong các chuẩn mã hóa video phổ biến như H.264/AVC [3], H.265/HEVC [4]. Sự tập trung năng lượng của dư thừa trong miền DCT được minh họa ở hình 1.4.

201	195	188	193	169	157	196	190
193	188	187	201	195	193	213	193
184	192	180	195	182	151	199	193
176	172	179	179	152	148	198	183
196	195	169	171	159	185	218	175
214	213	205	170	173	185	206	150
207	205	207	184	180	167	173	160
198	203	205	186	196	149	159	163

1480	49	33	-15	-14	33	-38	20
10	-52	11	-12	16	17	-13	-12
19	32	-22	-10	22	-20	9	8
16	10	17	27	-31	12	6	-5
-30	-6	13	-12	8	4	-3	-3
-25	16	6	-24	9	3	3	3
-2	17	4	-6	0	-4	-9	8
1	-2	6	0	7	-5	-8	-7

(a) Miền pixel

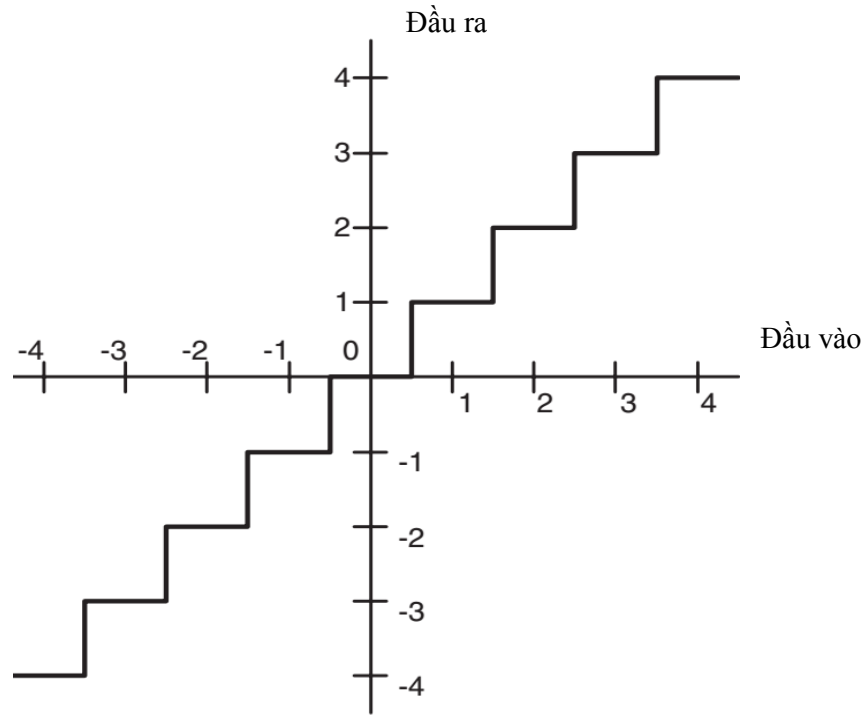
(b) Miền DCT

Hình 1.3: Giá trị của điểm ảnh trong miền pixel và miền DCT tương ứng

Như vậy, bằng cách tập trung năng lượng tại những thành phần tần số thấp; ví dụ: DC, AC1, AC2, AC3,..., các thông tin tại miền tần số cao sẽ có thể được loại bỏ (thông qua quá trình lượng tử mô tả ở phần tiếp theo); do vậy, thông tin mã hóa sẽ được giảm bớt. Cần lưu ý rằng, mắt người thường không nhạy cảm với các thành phần tần số cao. Do vậy, việc loại bỏ một vài thông tin tại tần số cao không hoặc rất ít làm ảnh hưởng tới cảm nhận về chất lượng của hình ảnh.

### 1.2.3. Phép lượng tử tuyến tính

Như đã mô tả ở trên, do đặc tính mắt người thường không nhạy cảm với thành phần tần số cao (ví dụ AC8, AC9,... AC15), việc loại bỏ các thành phần tần số này sẽ ít gây ảnh hưởng tới cảm nhận của mắt người đối với bức ảnh giải mã tại phía thu. Do vậy, một phép lượng tử tuyến tính thích hợp đã được sử dụng để loại bỏ các thành phần này như được minh họa ở hình 1.5.



Hình 1.4. Mô hình phép lượng tử tuyến tính sử dụng trong mã hóa video

Phép lượng tử mục đích chính là chuyển đổi các giá trị tín hiệu vào (ví dụ các hệ số DCT) trong một khoảng (gọi là khoảng lượng tử) tới 1 giá trị cụ thể (tín hiệu ra). Như vậy, bằng cách chuyển đổi này, nhiều hệ số DCT với giá trị gần nhau và trong cùng khoảng lượng tử có thể được chuyển đổi thành 1 giá trị lượng tử; giúp cho lượng thông tin cần phải mã hóa được giảm bớt. Tuy nhiên, cũng cần lưu ý rằng, phép lượng tử tuyến tính chính là thành phần chính gây ra tổn thất của chất lượng hình ảnh tái tạo tại phía thu.

#### 1.2.4. Các phép dự đoán ảnh

Ở kỹ thuật mã hóa sự sai khác giữa ảnh thực tế và ảnh dự đoán (mục 1.2.1) ta đã trao đổi về sự hiệu quả của việc mã hóa thông tin video thông qua việc mã hóa sự sai khác này. Sự sai khác giữa ảnh thực tế và ảnh dự đoán càng nhỏ, lượng thông tin cần mã hóa càng ít và hiệu quả nén sẽ càng cao. Do vậy, chất lượng của ảnh dự đoán sẽ là một trong những yếu tố then chốt, ảnh hưởng trực tiếp tới hiệu quả nén của bộ mã hóa video.

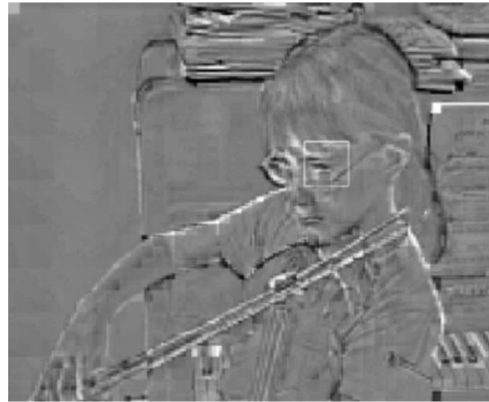
Có hai kỹ thuật tạo ảnh dự đoán cơ bản là kỹ thuật tạo ảnh dựa đoán trong khung (phổ biến với tên gọi Intra Prediction) và kỹ thuật tạo ảnh dự đoán liên khung (phổ biến với tên gọi Inter Prediction).



Ảnh gốc



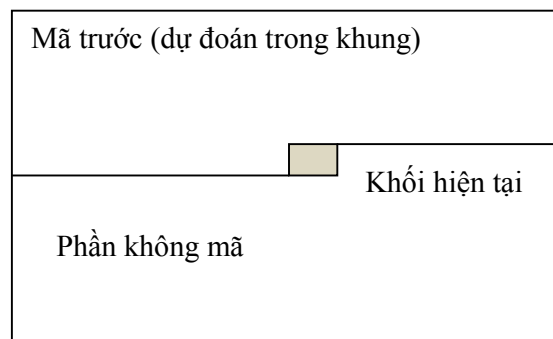
Ảnh dự đoán



Ảnh dư thừa (sai khác)

Hình 1.5. Mô tả ảnh gốc, ảnh dự đoán và ảnh dư thừa [5]

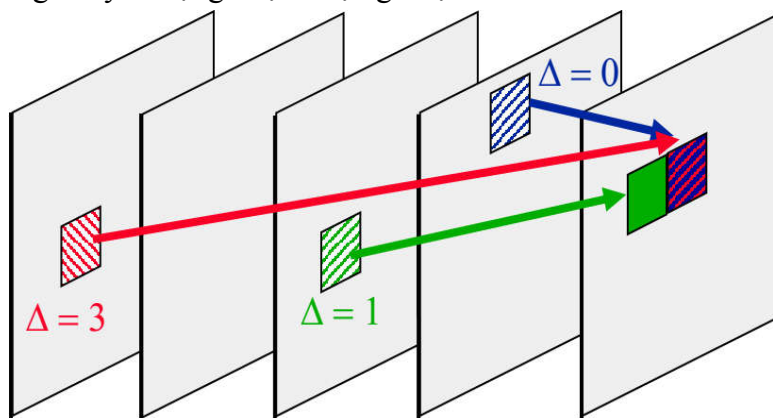
- Kỹ thuật dự đoán ảnh trong khung (intra prediction): Kỹ thuật này sử dụng các thông tin video đã được giải mã trong cùng một khung hình để tạo (ngoại suy) ra ảnh dự đoán. Hình 1.6 mô tả về kỹ thuật này.



Hình 1.6. Tạo ảnh dự đoán trong khung

Như vậy, có thể nói rằng kỹ thuật tạo ảnh dự đoán trong khung khai thác mối tương quan về mặt không gian giữa các khối ảnh trong cùng một khung hình.

- Kỹ thuật dự đoán ảnh liên khung (inter prediction): Với kỹ thuật này, thông tin tương quan giữa các khối ảnh trong các khung hình kế tiếp nhau được sử dụng. Như minh họa ở hình 1.7, các khung hình trước đó đã được giải mã tại phía phát sẽ được lưu trữ lại và sử dụng cho việc tạo ảnh dự đoán liên khung. Do khai thác tính tương quan về mặt thời gian, kỹ thuật dự đoán ảnh liên khung thường hiệu quả với những video có ít nội dung chuyển động hoặc được ghi lại bởi camera tĩnh.



4 khung tham chiếu trước khi mã hóa

Khung hiện tại

Hình 1.7. Tạo ảnh dự đoán liên khung

### 1.2.5. Mã hóa Entropy

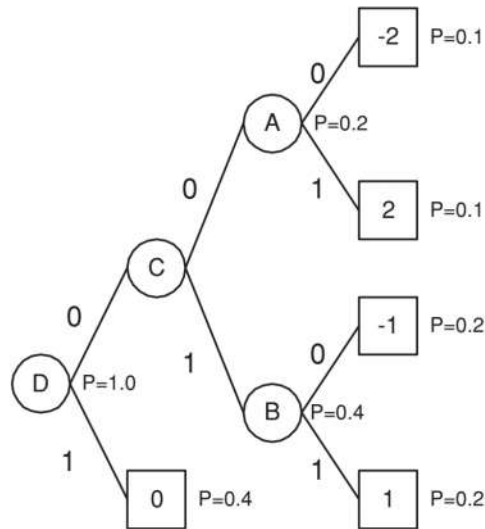
Giá trị lượng tử của sự khác biệt giữa thông tin gốc và thông tin dự đoán, tạo ra ở trên là một chuỗi các ký tự, thường là kiểu số nguyên. Do vậy, để truyền tải được, ta cần phải chuyển đổi các ký tự này sang chuỗi các số nhị phân 0, 1.

Cách đơn giản nhất để mã hóa các giá trị này là ta nhị phân hóa tất cả các giá trị lượng tử, sử dụng một số lượng bit cố định, ví dụ 8 bit. Tuy nhiên, cách làm này không hiệu quả do độ dài từ mã của mỗi giá trị lượng tử trong một khối ảnh có thể khác nhau. Do vậy, mã entropy với độ dài thay đổi thích ứng với từng khối ảnh, từng giá trị lượng tử của điểm ảnh đã được sử dụng trong các chuẩn mã hóa video truyền thống.

Trong kỹ thuật mã hóa Entropy, xác suất xuất hiện của từ mã được sử dụng để chỉ định độ dài cũng như giá trị của từ mã cho các giá trị lượng tử. Ví dụ như mã Huffman [6], với xác suất xuất hiện các ký tự được cho như ở bảng 1.1 sau:

Bảng 1.1. Giá trị lượng tử và xác suất tương ứng

Vector	xác suất p	$\log_2(1/p)$
-2	0.1	3.32
-1	0.2	2.32
0	0.4	1.32
1	0.2	2.32
2	0.1	3.32



Hình 1.8. Sử dụng thuật toán cây tạo ra từ mã Huffman

Khi đó, sử dụng giải thuật tạo mã Huffman (hình 1.8), ta có thể dễ dàng xác định được từ mã tương ứng và lượng bit cần mã hóa đối với mỗi giá trị lượng tử như sau:

Bảng 1.2. Từ mã Huffman và lượng bit cần mã hóa tương ứng

Vector	Từ mã	Bits thực tế	Bits lí tưởng
0	1	1	0.32
1	01	2	3.64
-1	001	3	3.84
2	0001	4	5.06
-2	0000	4	5.64

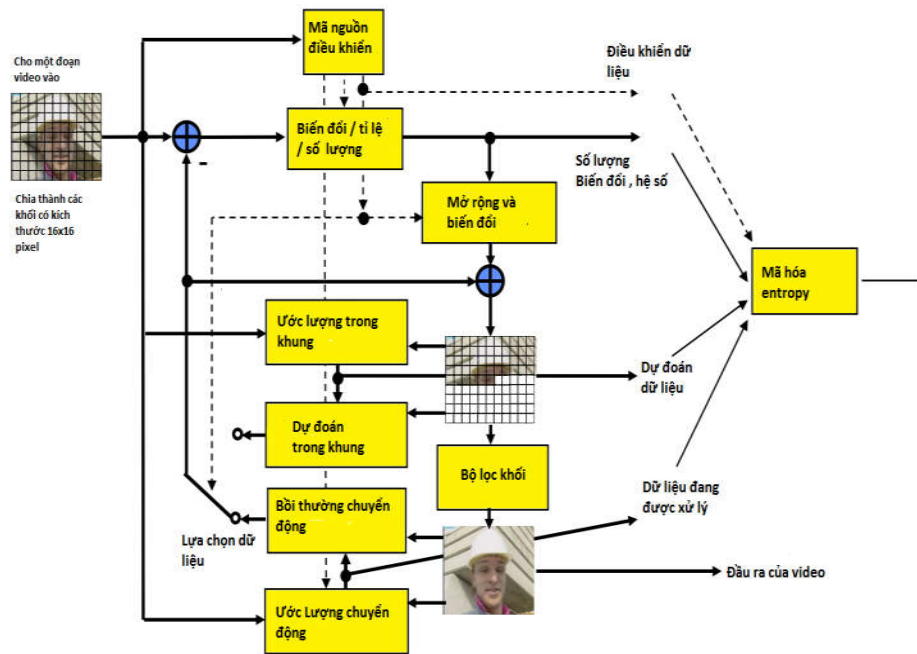
Trong thực tế, ngoài mã Huffman ra thì mã toán học (arithmetic coding) thường được sử dụng trong các chuẩn mã hóa video truyền thống như được mô tả ở các mục sau.



### 1.3. Các chuẩn mã hóa video phổ biến

#### 1.3.1. Chuẩn H.264/AVC

##### a) Kiến trúc tổng quát



Hình 1.9. Kiến trúc tổng quát của chuẩn H.264/AVC [3]

##### b) Các thành phần chính

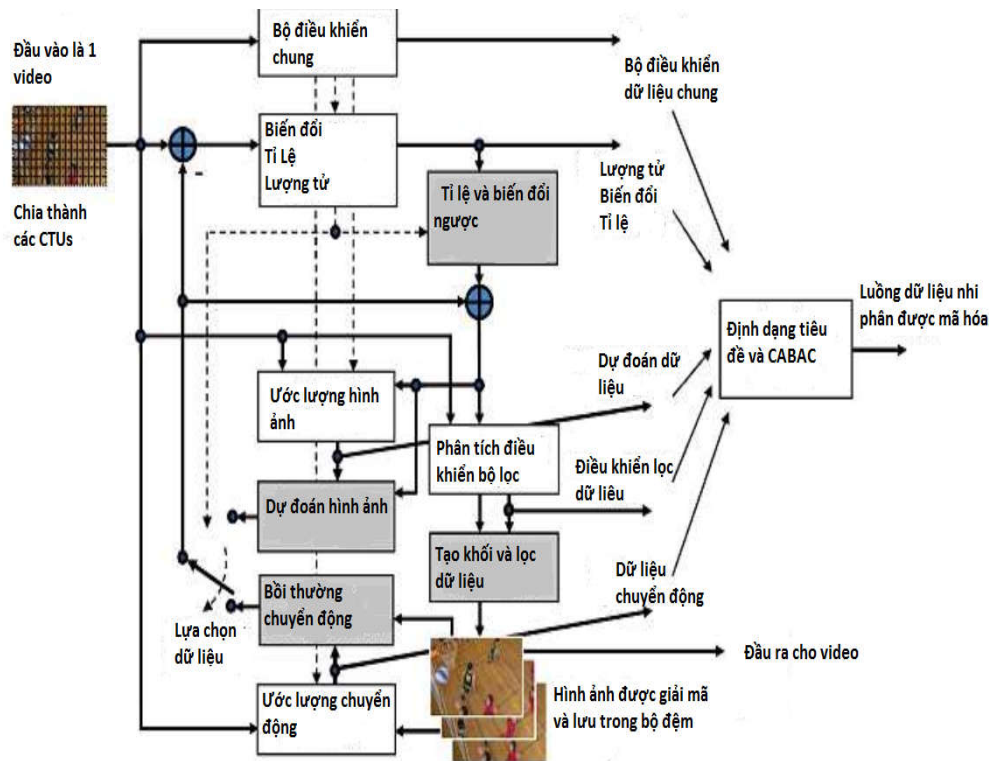
- **Dự đoán ảnh trong khung:** Chuẩn H.264/AVC cung cấp phép dự đoán ảnh trong khung với 9 khả năng lựa chọn và 2 loại kích cỡ khối (block) khác nhau (4×4 và 16×16).
- **Dự đoán ảnh liên khung:** Trong chuẩn H.264/AVC, dự đoán ảnh liên khung được cập nhật với một số các công cụ hiệu quả như dự đoán dựa trên nhiều khung hình tham khảo, dự đoán B-slides dạng liên cấp, vector dự đoán với độ chính xác tới 1/4 điểm ảnh, dự đoán với các khối ảnh có kích cỡ khác nhau, dự đoán kết hợp có trọng số và thêm hai lựa chọn dự đoán nữa là dự đoán bỏ qua và dự đoán trực tiếp.
- **Phép biến đổi cô sin rời rạc và lượng tử tuyến tính:** H.264/AVC sử dụng phép biến đổi cô sin rời rạc cho hai dạng kích cỡ khối là 4×4 và 8×8. Phép lượng tử tuyến tính tiếp tục được sử dụng trong chuẩn H.264/AVC. Cần nhấn mạnh rằng, H.264/AVC là chuẩn nén có tổn thất và thành phần chính tạo nên sự tổn thất này chính là phép lượng tử.

- **Mã hóa Entropy:** H.264/AVC cho phép người dùng lựa chọn một trong hai phép mã hóa entropy phổ biến và hiệu quả: i) context-adaptive binary arithmetic coding (CABAC) và ii) context-adaptive variable length coding (CAVLC). Việc lựa chọn phương pháp mã hóa entropy nào tùy thuộc vào mục đích và khả năng của người dùng. Trong đó, mã CABAC hiệu quả hơn CAVLC nhưng cũng đi kèm với sự phức tạp về mặt thuật toán cao hơn.
- **Bộ lọc giảm nhiễu khối:** Do đơn vị dự đoán và mã hóa nhỏ nhất của chuẩn nén video H.264/AVC là các khối (kích cỡ từ  $4 \times 4$  tới  $16 \times 16$ ), khi đó, việc khôi phục lại tín hiệu ban đầu sẽ luôn gặp phải vấn đề với nhiễu khối. Do vậy, chuẩn H.264/AVC sử dụng một bộ lọc với các thông số có thể được cấu hình bởi người dùng để làm giảm các nhiễu khối.

Tóm lại, với rất nhiều các công cụ mới, hiệu quả, chuẩn H.264/AVC [3] nâng cao được khả năng mã hóa video lên rất nhiều so với các chuẩn trước đó như chuẩn H.263 hay MPEG-2/Video [2].

### 1.3.2. Chuẩn H.265/HEVC

#### a) Kiến trúc tổng quát



Hình 1.10. Kiến trúc tổng quát của chuẩn H.265/HEVC [4]

## b) Các thành phần chính

- **Cấu trúc đơn vị mã hóa:** Một trong những đặc trưng khác biệt nhất của chuẩn H.265/HEVC so với các chuẩn trước đó như H.264/AVC hoặc MPEG-4/Video chính là cách cấu trúc lại các đơn vị mã hóa. Trong chuẩn H.264/AVC, đơn vị mã hóa cơ bản là các khối lớn (Macroblock) với kích cỡ tối đa là  $16 \times 16$ . Trong khi đó, chuẩn H.265/HEVC chia các khối mã hóa theo cấu trúc cây (coding tree block) với kích cỡ đa dạng hơn từ  $4 \times 4$  tới  $64 \times 64$ . Cấu trúc này đảm bảo hiệu quả nén tốt hơn so với chuẩn H.264/AVC khi các video với kích cỡ lớn như HD, 2K hay 4K được sử dụng.
- **Cấu trúc đơn vị dự đoán và đơn vị biến đổi:** Mỗi đơn vị khối mô tả ở phía trên được phân chia thành hai cấu trúc tương ứng với hai thành phần chính của chuẩn mã hóa H.265/HEVC; đó là cấu trúc đơn vị dự đoán và cấu trúc đơn vị biến đổi. Đơn vị dự đoán (coding unit) cho phép khối mã hóa có kích cỡ nhỏ nhất là  $8 \times 8$  đối với thành phần xám (luma) và  $4 \times 4$  với thành phần màu (chroma). Trong khi đó, đơn vị biến đổi có sin rời rạc có kích cỡ đa dạng hơn bao gồm  $4 \times 4$ ,  $8 \times 8$ ,  $16 \times 16$ ,  $32 \times 32$ .
- **Phương pháp dự đoán chuyển động nối vùng (merge mode):** Đây là một phương pháp dự đoán mới xuất hiện trong chuẩn H.265/HEVC. Phương pháp này kế thừa có hiệu quả phương pháp dự đoán bỏ qua và dự đoán trực tiếp trong chuẩn H.264/AVC mô tả ở trên. Trong đó có 2 sự khác biệt chính là việc gửi thông tin về véc tơ chuyển động được lựa chọn và thông tin về khung tham khảo được sử dụng.
- **Dự đoán ảnh trong khung:** H.265/HEVC hỗ trợ tới 35 mode dự đoán (so với 9 mode trong H.264/AVC). Ngoài ra, khối dự đoán trong khung có thể được tạo ra với hai loại khối,  $N \times N$  và  $2N \times 2N$ .
- **Dự đoán ảnh liên khung:** Dự đoán ảnh liên khung trong chuẩn H.265/HEVC cũng cho phép hỗ trợ dự đoán véc tơ chuyển động với độ chính xác lên tới  $\frac{1}{4}$  điểm ảnh (sub-pel). Các giá trị tại vị trí sub-pel được nội suy dựa trên hai bộ lọc với 7 hoặc 8 tham số cấu hình. Các khối hình sử dụng trong dự đoán ảnh liên khung cũng đa dạng, có thể đối xứng hoặc không đối xứng, ví dụ:  $2N \times 2N$ ,  $2N \times N$ ,  $N \times 2N$ ,  $N \times N$ ,  $2N \times nD$ ,  $nL \times 2N$ .
- **Mã hóa Entropy:** Khác với H.264/AVC, chuẩn H.265/HEVC chỉ cho phép người dùng sử dụng mã CABAC; mục đích là tăng tối đa hiệu quả nén chuẩn H.265/HEVC.
- **Bộ lọc giảm nhiễu khối:** Một chút thay đổi, nâng cấp được đề xuất để giảm nhiễu khối trong chuẩn H.265/HEVC.
- **Bộ lọc giảm nhiễu vòng:** Đây là thành phần mới trong chuẩn H.265/HEVC so với các chuẩn trước kia như H.264/AVC, MPEG 2,4,... Mục đích chính là giảm

nhiều vòng xuất hiện trong một số khung hình xuất hiện do việc loại bỏ thành phần tần số cao ở khối lượng tử tuyến tính.

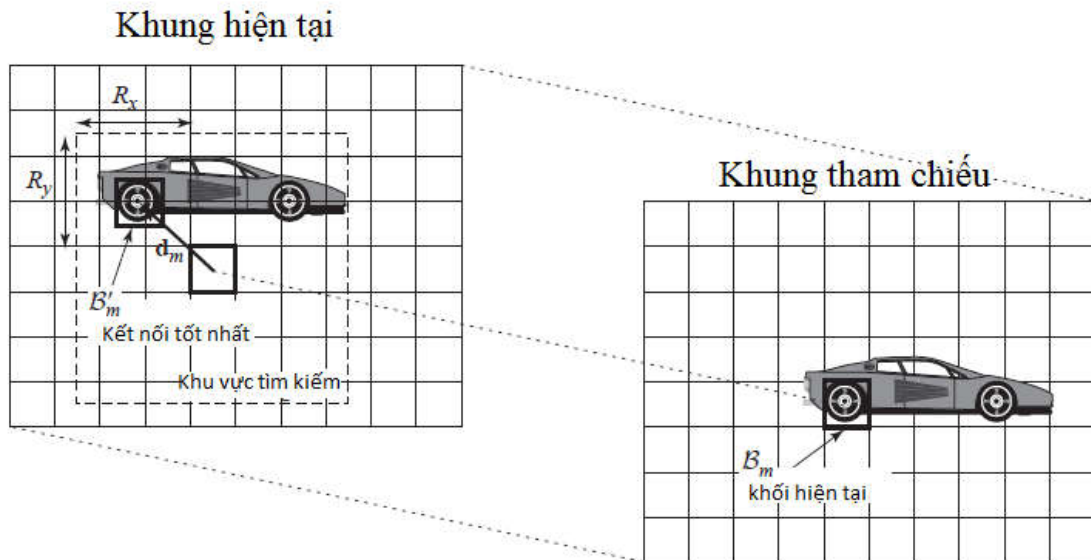
#### 1.4. Nhược điểm của kỹ thuật mã hóa video truyền thống

Mã hóa video truyền thống dựa vào phép dự đoán và biến đổi cô sin rời rạc tại phía thu đang được sử dụng phổ biến trong các ứng dụng truyền hình quảng bá, truyền hình hội nghị,... Tuy nhiên, các kỹ thuật này cũng có một số nhược điểm sau.

##### 1.4.1. Độ phức tạp cao tại phía mã hóa

Trong kỹ thuật truyền thống độ phức tạp tính toán nằm chính là ở phần dò tìm vecto dịch chuyển. Càng tìm được vecto dịch chuyển chính xác bao nhiêu, mã sai khác càng ít bit (nén video càng cao) và chất lượng khôi phục ảnh càng tốt. Để đạt được điều này thậm chí phải dò tìm chính xác đến mức  $\frac{1}{2}$  hoặc  $\frac{1}{4}$  giá trị điểm ảnh. Song như vậy số phép tính chuẩn bị cho dò tìm này càng lớn

Ta xét ví dụ minh họa sau trong hình sau:



Hình 1.11: Phạm vi dò tìm vecto dịch chuyển của block chứa bánh xe

Nhìn vào hình vẽ trên ta thấy mục tiêu là phải tìm vecto dịch chuyển  $d_m$  của block hiện tại  $B_m$  (ứng với bánh xe ô tô), nó là kết quả dịch từ block  $B'_m$  từ khung trước có sai khác ít nhất với block  $B'_m$ . Việc dò tìm thực hiện theo qui tắc:

- Dùng tổng lỗi tuyệt đối giữa các pixel tương ứng
- Vùng dò tìm: phụ thuộc dải chuyển động định trước
- Bước dò tìm: pixel hoặc nửa pixel

$$E_{DFD}(d_m) = \sum_{xBm} | \psi_2(x + d_m) - \psi_1(x) |^p \rightarrow \min$$

Dịch chuyển giữa block hiện tại và block khác ít nhất trong tham chiếu sẽ cho vecto dịch chuyển. Hay block hiện tại được dự đoán từ block tham chiếu và vecto dịch chuyển.

- Sự phức tạp của của thuật toán dò tìm chuyển động:

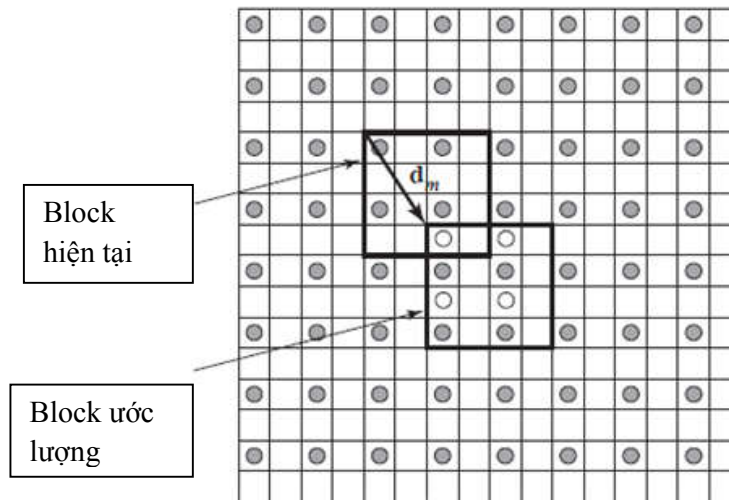
Để đánh giá hết sự phức tạp trong việc dò tìm vecto dịch chuyển phục vụ dự đoán khung hình theo thời gian, ta nêu ra tính toán như sau, giả sử:

- Kích thước khung là  $M \times M$
- Kích thước block là  $N \times N$
- Vùng dò tìm là  $(-R, R)$  cho mỗi chiều  $x, y$
- Bước dò tìm là: pixel
- Số phép toán trong block  $N^2$ , trong vùng dò tìm là  $(2R+1)^2 \cdot N^2$
- Trong toàn bộ khung sẽ là:  $(M/N)^2 \cdot (2R+1)^2 \cdot N^2$

Ví dụ:  $M=512$ ,  $N=16$ ,  $R=16$ , 30fps

Tổng số phép tính là  $2,85 \cdot 10^8$ /frame hay  $8,55 \cdot 10^9$ /giây

Nếu ước lượng dịch chuyển có độ chính xác nửa pixel thì số phép tính còn tăng lên nữa sau phép nội suy giá trị nửa pixel.



Hình 1.12: Ước lượng véc tơ dịch chuyển

#### 1.4.2. Khả năng chống chịu nhiễu thấp

Đối với các mô hình mã hóa video truyền thống, mã sai khác được sử dụng, các thành phần dự đoán (trong khung, liên khung) được thực hiện tại phía phát; các thông tin phụ như véc tơ chuyển động, thông tin dư thừa, mode lựa chọn,... được gửi tới phía thu để tạo ra thành phần dự đoán tương tự. Tuy nhiên, trong điều kiện không hoàn hảo của kênh truyền, có thể có nhiễu xuất hiện; đặc biệt là với phương pháp dự đoán liên

khung, nhiễu hoặc gói tin bị mất có thể ảnh hưởng trực tiếp tới chất lượng ảnh dự đoán của khung hình hiện tại và cả các khung hình kế tiếp. Do vậy, mã hóa video truyền thống rất nhạy cảm với các môi trường có nhiễu.

Trên cơ sở những nhược điểm kể trên, mã hóa video truyền thống chắc chắn không phù hợp với những ứng dụng truyền thông đa phương tiện mới như mạng cảm biến không dây, hệ thống điều khiển từ xa,... bởi các hệ thống này đòi hỏi một hệ nén có yêu cầu độ phức tạp thấp đồng thời có khả năng hoạt động tốt trong vùng có nhiễu. Do vậy, nghiên cứu này sẽ tập trung vào giới thiệu và phân tích một mô hình mã hóa video mới, tên gọi mã hóa video phân tán (Distributed Video Coding – DVC) như ở các chương kế tiếp.

## CHƯƠNG 2: CƠ SỞ KỸ THUẬT CỦA MÃ VIDEO PHÂN TÁN

Kỹ thuật mã hóa video phân tán (DVC) được xây dựng và phát triển dựa trên nền tảng lý thuyết thông tin kinh điển, và hai định lý của Slepian – Wolf [7] (cho mã hóa không tổn thất) và Wyner – Ziv [8] (cho mã hóa có tổn thất). Tuy nhiên, để có thể hiểu và so sánh với kỹ thuật mã hóa video truyền thống sử dụng trong các chuẩn. Chương này sẽ bắt đầu bằng phần trình bày về mô hình mã hóa video dự đoán và tiếp theo là mô hình mã hóa video phân tán với hai định lý Slepian-Wolf và Wyner-Ziv.

### 2.1. Mã hóa dự đoán cổ điển

Lý thuyết thông tin đã chứng minh được rằng:

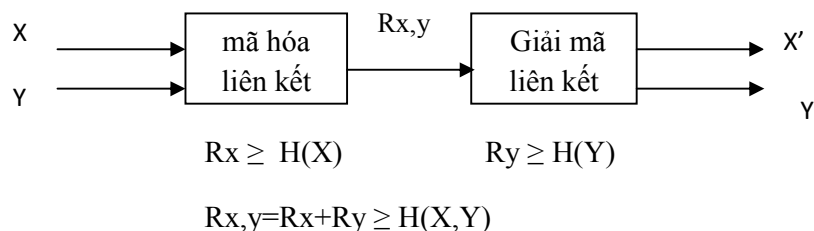
- Với các nguồn độc lập thì tốc độ tối thiểu để mã có giải mã không lỗi chính là entropy của nguồn, ví dụ 2 nguồn độc lập  $X$  và  $Y$  thì tốc độ tối thiểu mã là:  $H(X)$ ,  $H(Y)$
- Với 2 nguồn thống kê phụ thuộc, ví dụ  $X$  và  $Y$  có tương quan thì tốc độ tối thiểu để mã hóa mã và khôi phục không lỗi là  $H(X, Y)$

$$\text{Vì } H(X) + H(Y) \geq H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y) \quad (2.1)$$

Điều này có nghĩa là nếu khai thác được đặc tính phụ thuộc của 2 nguồn sẽ cho phép giảm tốc độ mã hóa đến entropy liên kết.

Hai nguồn thông tin phụ thuộc (có tương quan cao) này thực tế có thể là các khung liên tiếp của cùng 1 dãy video. Mã liên kết 2 khung hình video liên tiếp có thể là mã một khung ( $H(X)$ ) rồi mã tiếp dự đoán của  $Y$  dựa trên thông tin về  $X$  ( $H(Y|X)$ ) như công thức (2.1)

Tuy nhiên để có  $H(Y|X)$  nhỏ nhất, cần tìm vecto dự đoán có xác suất cao nhất điều này dẫn đến việc dò tìm phức tạp trong kiểu mã hóa liên kết của kỹ thuật mã hóa truyền thống.



Hình 2.1: Mô hình mã hóa dự đoán cổ điển

### 2.2. Định lý Slepian-Wolf

Định lý này không đòi hỏi phải trực tiếp mã *liên kết* theo  $H(X, Y)$  như nói ở trên mà có thể mã 2 nguồn phụ thuộc một cách *độc lập* mà vẫn khai thác đặc tính phụ

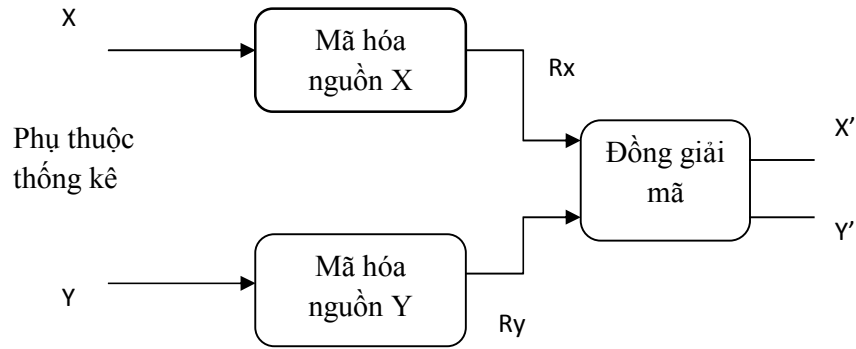
thuộc tức là không làm mất mát hiệu suất nén, song vẫn phải giải mã liên kết. Điều này cho xác suất khôi phục có lỗi tiến đến 0.

Tuy nhiên mã độc lập phải thỏa mã 3 điều kiện (bất đẳng thức) sau:

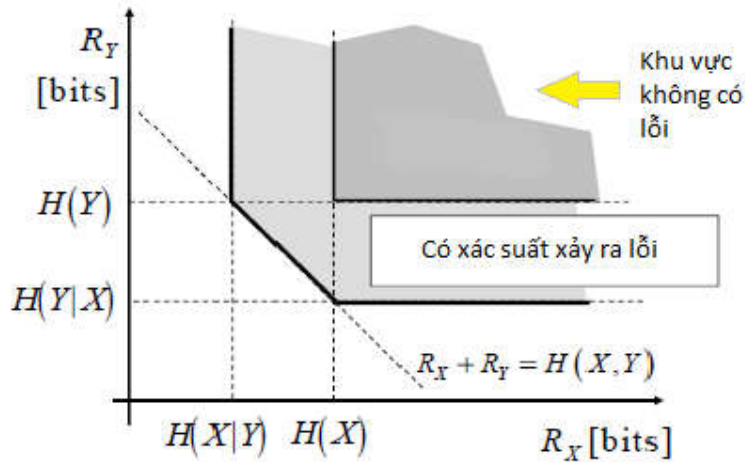
$$R_x + R_y \geq H(X,Y)$$

$$R_x \geq H(X/Y)$$

$$R_y \geq H(Y/X)$$



Hình 2.2: Sơ đồ mã nguồn phụ thuộc thống kê



Hình 2.3: Biểu đồ vùng tỉ lệ tốc độ mã giữa 2 nguồn X,Y

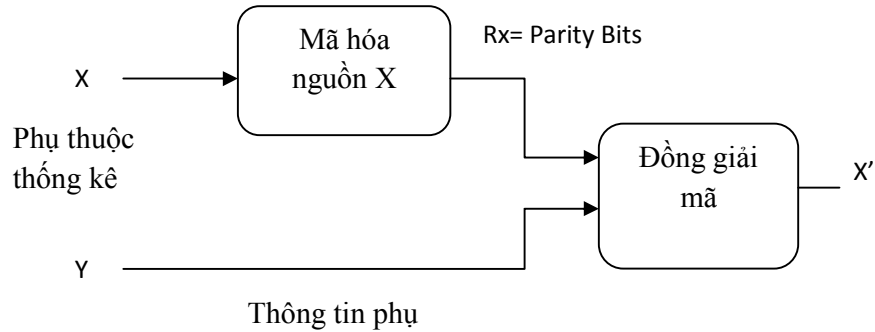
Hình 2.3 thể hiện rằng vùng tốc độ mã có thể khôi phục tốt đối với các tổ hợp  $R_x$  và  $R_y$  khác nhau song tuân theo hệ điều kiện trên.

Việc thực hiện định lý *Slepian-Wolf* có thể qua kỹ thuật mã kênh:

Coi X như một *phiên bản lỗi* của Y (như đi qua kênh truyền). Trong mã kênh để hiệu chỉnh lỗi giải pháp là chèn thêm các bit chẵn lẻ (Parity bits) nhằm tăng thêm khoảng cách giữa các từ mã lớn hơn 2 lần *khoảng cách lỗi*.



Khi đó nếu bên giải mã truy cập  $Y$  như thông tin lẻ (từ khung trước), bên mã hóa sẽ mã  $X$  bằng các bit chẵn lẻ (tạo nên các tập Coset mà  $X$  thuộc về một trong các tập Coset) và gửi các bit chẵn lẻ này đến bên thu, sau đó dựa vào  $Y$  có sẵn ở bên thu khôi phục chính xác được  $X$ . Ở kỹ thuật này càng ít bit chẵn lẻ, hệ số nén càng cao.

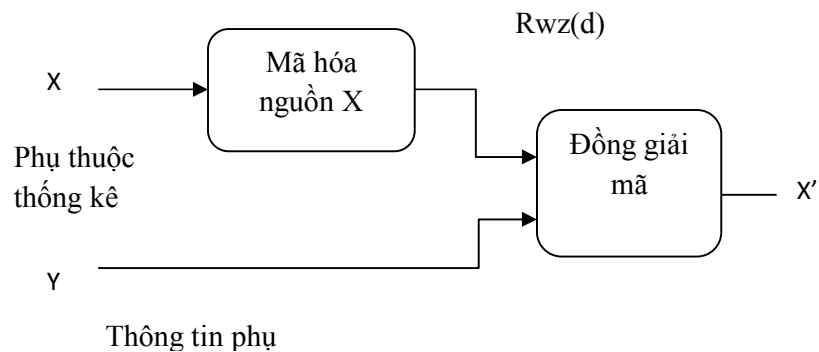


Hình 2.4: Mô tả định lý Slepian-Wolf với thông tin phụ

### 2.3. Định lý Winer-Ziv

Đây là sự mở rộng của định lý *Slepian-Wolf* với mã có mất mát cho các nguồn phụ thuộc có phân bố Gauss, định lý Wyner-Zip xem xét điểm tọa độ:  $R_x=(X/Y)$  và  $R_y=H(Y)$  với điều kiện  $Y$  được biết ở bên giải mã như thông tin lẻ, định lý phát biểu rằng: nếu 2 nguồn Gauss phụ thuộc không nhớ  $X, Y$ , ở đây  $Y$  là thông tin lẻ, được mã độc lập với độ méo  $d$  xác định (mã có mất mát), thì sẽ không mất hệ số nén so với mã cổ điển khi mã  $X$  và khôi phục  $X$  hình 2.5

Nếu thống kê nguồn không phải là Gauss sẽ có sự tăng thêm tốc độ mã  $WZ, R^{WZ}$  so với mã truyền thống. Sau này Zamir đã chứng minh sự tăng tốc độ giữa phương pháp  $WZ$  và phương pháp truyền thống là nhỏ hơn hay bằng 0.5 bit/mẫu.



Hình 2.5: Mã hóa mất mát thông tin với thông tin phụ ở phần giải mã

## 2.4. Ví dụ minh họa

Giả sử X và Y là các dữ liệu 3 bit có xác suất bằng nhau lấy trên 8 tổ hợp. Tuy nhiên X và Y lại tương quan với nhau sao cho khoảng cách hamming giữa chúng nhiều nhất là 1. Ví dụ nếu biết Y là 010 thì X chỉ có thể là 010, 000, 110, 011

Để mã nén X truyền đi có 2 kịch bản như sau:

Kịch bản 1: Y được coi là dự đoán của X từ khung trước (sau khi đã bù dịch chuyển). Thặng dư  $X \oplus Y$  chính là sai khác hay là lỗi dự đoán X từ Y được mã hóa và truyền đi chỉ bằng 2 bit truyền đi. Đây là tốc độ nhỏ nhất (có thể) để nén X truyền đi.

Ở đây có thể coi X là mã của block hiện tại trong 1 khung hình video. Y là dự đoán của X từ bộ nhớ khung trước (sau khi dự đoán vecto dịch chuyển rồi bù dịch chuyển này), thông tin này của Y cũng dễ dàng truy cập được ở bên thu (khi nhận được vecto dịch chuyển truyền từ bên phát sang). Tương quan X,Y là tương quan theo thời gian giữa X và dự đoán Y của nó nên mã này thuộc loại *mã dự đoán* (kiểu mã truyền thống nói trên).

Ở bên thu từ thông tin  $X \oplus Y$  và Y dễ dàng khôi phục lại được X.  
( $X \oplus Y \oplus Y = X$ )

Như ta thấy, để nén hiệu quả phải dự đoán tốt (làm cho sai khác tối thiểu và mã hóa sai khác ít nhất), nên phép dò tìm vecto dịch chuyển là *phức tạp ở bên mã hóa*.

Kịch bản 2:

Bộ mã hóa X không cần biết Y như một dự đoán (do đó không mất công cần dò tìm dịch chuyển) và do đó cũng không thể mã hóa hiệu giữa 2 giá trị này để truyền đi. Tuy nhiên nó cần biết X như một *cấu trúc tương quan* với khung trước đó là Y. Cấu trúc này biểu diễn N như tạp âm tương quan theo công thức:

$$X = Y + N$$

Dựa trên giá trị của tạp âm tương quan (được mô tả chi tiết ở chương sau) không gian từ mã của X được phân hoạch thành các tập con gọi là coset sao cho các từ mã trong coset có khoảng cách hamming lớn hơn 2 lần tạp âm tương quan nói trên.

Trong ví dụ trên bên mã hóa chia không gian từ mã của X được thành 4 tập con, mỗi tập 2 từ mã, chẳng hạn

Coset 1: 000,111

Coset 2: 001,110

Coset 3 010,101

Coset 4 100,011

Chú ý là khoảng cách hamming của 2 từ mã trong mỗi coset là 3 lớn hơn 2 lần khoảng cách Hamming (hay tạp âm gây nên) giữa X và Y là 1

Bộ mã sẽ mã chỉ số của Coset nào chứa X rồi truyền sang bên thu (cũng chỉ dùng 2 bit do có 4 tập coset, có độ nén tương đương kịch bản 1). Bên thu sau khi nhận chỉ số coset, sẽ dùng Y là giải mã của khung trước để xác định chính xác X là từ nào trong coset bằng cách tìm từ mã gần nhất với Y. Kiểu mã này gọi là mã mang thông tin lẻ.

Ví dụ: X là 110, ở bên thu Y của khung trước là 100 (với yêu cầu tương quan X và Y có khoảng cách hamming là 1). Khi đó:

Mã truyền đi là tập coset 2, trong tập này có 2 từ mã 001 và 110, lựa chọn cuối sẽ là X=110 vì gần Y hơn.

Ở kịch bản này rõ ràng *sự phức tạp chuyển sang bên thu* vì nếu trong coset có nhiều từ mã thì phải dò tìm từ mã nào có khoảng cách đến Y nhỏ nhất.

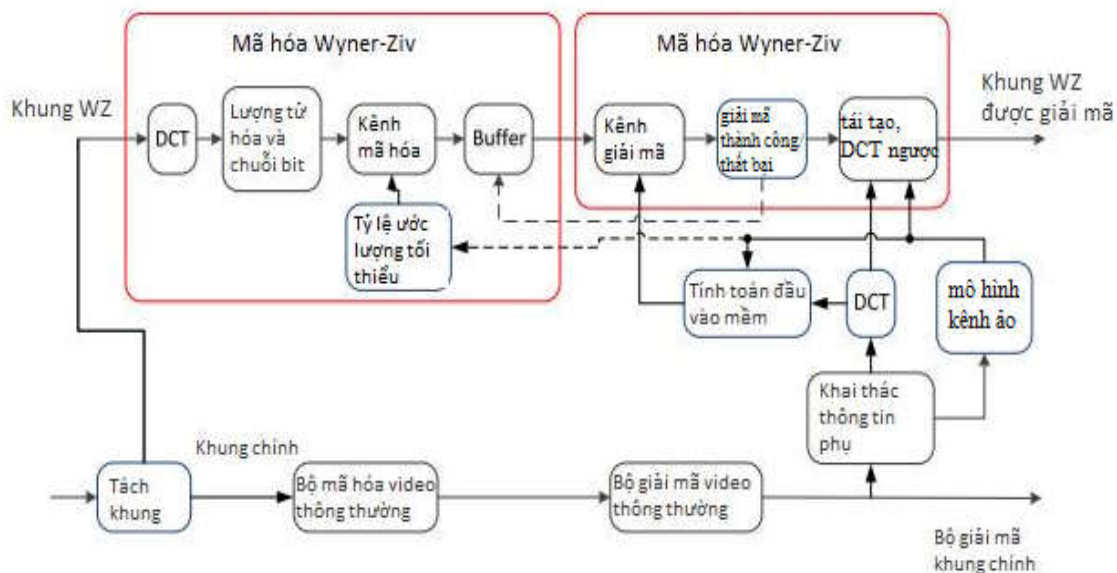
Chú ý là việc phân hoạch X dựa theo cấu trúc tương quan giữa X và Y nên cùng giá trị X trong Coset 3, Y bị 1 lỗi có thể nhận 1 trong các giá trị 010, 110, 000, 011 thì X vẫn khôi phục chính xác vì Y vẫn thỏa mã cấu trúc tương quan. Như vậy khác với mã dự đoán, không có sự phụ thuộc xác định giữa mã X và giá trị Y dự đoán mà chỉ cần có cấu trúc tương quan nhất định giữa X hiện tại và Y là giá trị khung trước đó. Kiểu mã này cung cấp khả năng chống nhiễu và mất dữ liệu cao hơn mã truyền thống.

## CHƯƠNG 3: CÁC MÔ HÌNH THỰC NGHIỆM CỦA KỸ THUẬT MÃ VIDEO PHÂN TÁN

Mặc dù các lý thuyết về mã hóa video phân tán đã xuất hiện từ những năm 70, tuy nhiên các kết quả thực nghiệm, các mô hình thực tế của phương pháp mã hóa video phân tán mới chỉ xuất hiện ở những năm đầu thế kỷ 20. Cụ thể là sự giới thiệu của hai giải pháp mã hóa video phân tán đề xuất bởi nhóm nghiên cứu tại đại học Stanford [9] và hướng thứ hai của nhóm nghiên cứu tại đại học Berkeley [10]. Chương này, do vậy, sẽ đi sâu vào mô tả hai hướng nghiên cứu này.

### 3.1. Giải pháp STANFORD

Mã hóa video phân tán theo phương pháp đề xuất tại trường Đại học Stanford bởi giáo sư Bernd Girod [9] là giải pháp mã hóa video theo mức độ khung hình, kết hợp giữa mã hóa nguồn và mã hóa kênh, có sử dụng kênh phản hồi (backward channel) từ phía thu gửi sang phía phát. Tiêu biểu cho hướng nghiên cứu này là bộ mã hóa DISCOVER, đề xuất và xây dựng bởi nhóm nghiên cứu tại viện IST, Bồ Đào Nha [11]. Trong giải pháp này, dữ liệu video được chia thành nhóm các khung hình (Group of pictures: GOPs). Khung đầu tiên và cuối cùng trong mỗi GOP, gọi là khung chính, được mã hóa theo kiểu mã hóa thông tin dự đoán truyền thống như H.264/AVC [3]. Các khung duy trì giữa các khung chính được gọi là khung *Wyner-Ziv* (WZ) và được mã hóa bằng việc sử dụng kỹ thuật DVC. So với giải pháp Stanford DVC ban đầu [8], bộ mã hóa DISCOVER đề xuất một thuật toán GOP thích nghi dựa trên nội dung phân tích video, kỹ thuật tạo thông tin phụ hiệu quả và sử dụng mã hóa kênh LDPC như mô tả ở hình 3.1.



Hình 3.1. Kiến trúc bộ mã hóa DISCOVER [11]

**Đối với khung WZ, cấu trúc khám phá thuật nén và giải nén bao gồm:**

**Tai bên mã hóa:**

***3.1.1. Biến đổi và lượng tử hóa***

Đầu tiên, biến đổi DCT được áp dụng trong các khối pixel 4 x 4 của mỗi khung WZ. Hệ số DCT của toàn bộ khung WZ sau đó được nhóm lại với nhau, theo vị trí của từng hệ số DCT trong một khối, tạo ra một hệ số DCT chuẩn. Mỗi chuẩn DCT sau đó được lượng tử hóa đồng nhất với một số mức độ tùy thuộc vào chất lượng đích. Đối với mỗi một chuẩn, các bits biểu diễn lượng tử sẽ được nhóm lại với nhau, tạo nên một “mặt phẳng” bit (bitplane), sau đó được mã hóa độc lập.

***3.1.2 Mã hóa kênh và bộ đệm***

Một tốc độ phù hợp mã tích lũy chẵn lẻ mật độ thấp (Low density parity codes accumulate: LDPCA) được áp dụng với mỗi chuẩn DCT bắt đầu từ bitplane lớn nhất (most significant bitplane: MSB). Thông tin chẵn lẻ từ mỗi bitplane được lưu trữ tại một vùng đệm ra và được gửi thành từng mảng khi có yêu cầu từ bộ giải mã hóa qua kênh phản hồi.

***3.1.3. Tỷ lệ ước tính tối thiểu***

DISCOVER codec đề xuất một modul mới gọi bộ ước lượng tốc độ tối thiểu được sử dụng để giảm số Syndrom tích trữ yêu cầu bởi bộ giải mã, vì vậy sẽ giảm sự phức tạp giải mã toàn thể. Tốc độ tối thiểu được tính toán dựa trên giới hạn Wyner-Ziv RD của hai nguồn phân bố Gauss tương quan [7].

**Tai bên giải mã:**

***3.1.4. Tạo và khai thác thông tin phụ***

DISCOVER codec sử dụng giải pháp nội suy bù trừ chuyển động (motion compensated interpolation: MCI) để tạo ra thông tin phụ (side information: SI) [12-13]. Tại MCI, một trường chuyển động gần chuyển động thực được ước lượng giữa trước (quá khứ) và sau (hiện tại) khung tham chiếu. Sau đó, sự bù trừ chuyển động giữa hai khung tham chiếu được thực hiện để có được các thông tin phụ. Thông tin phụ được hiểu như là ước lượng (phiên bản tap âm) của khung WZ gốc để mã hóa. Chất lượng thông tin phụ tốt hơn, lỗi ước tính nhỏ hơn để giải mã chính xác. Để khai thác dư thừa không gian, một phép biến đổi được áp, ở đây là phép biến đổi DCT được áp dụng cho thông tin phụ miền pixel.

### **3.1.5. Mô hình kênh ảo và tính toán đầu vào mềm**

DISCOVER codec sử dụng phân bố Laplacian để làm mô hình tạp âm tương quan (Phân bố lỗi giữa DCT chuẩn tương ứng của SI và khung WZ). Tuy nhiên, thay vì sử dụng một quy trình huấn luyện offline để tính toán tham số tạp âm tương quan [9], thì DISCOVER codec ước tính các tham số này online và tại bộ giải mã, theo một kỹ thuật được đề xuất trong [14]. Phương pháp ước lượng tham số Laplacian sử dụng khung dư thừa, nghĩa là sự sai khác giữa trước và sau khung tham chiếu đã bù chuyển động, như một sự ước tính thô của chất lượng SI. Mô hình phân phối mẫu Laplacian sau đó được sử dụng để chuyển các hệ số SI DCT thành thông tin đầu vào mềm cho Bộ giải mã LDPC.

### **3.1.6. Giải mã Kênh và Kiểm tra CRC**

Khi các hệ số SI DCT và thống kê tạp âm tương quan cho mỗi DCT chuẩn được biết, mỗi bitplane được giải mã LDPCA (bắt đầu từ MSB) để sửa lỗi bit trong SI. Trong trường hợp này, bộ giải mã LDPCA nhận từ bộ mã hóa các mảng bit chẵn lẻ nối tiếp theo yêu cầu thông qua kênh phản hồi. Để điều khiển số lượng bitplanes cần cho sự giải mã thành công của từng bitplane, bộ mã hóa sử dụng tiêu chuẩn dừng yêu cầu theo [15]. Vì một số lỗi dư thừa có thể còn lại sau việc dừng yêu cầu, có thể tạo nên ảnh hưởng tiêu cực đến chất lượng khung giải mã, Một kiểm tra tổng CRC được truyền để giúp bộ giải mã tìm ra và sửa các lỗi còn sót lại tại mỗi bitplane. Sau khi giải mã thành công bitplane MSB của một DCT chuẩn, bộ giải mã LDPCA sẽ xử lý với các bitplanes còn lại của cùng chuẩn DCT. Quá trình đó là sự ứng dụng liên tiếp cho các DCT còn lại.

### **3.1.7. Sự khôi phục và biến đổi ngược**

Sau giải mã kênh, tất cả các bitplanes liên quan đến mỗi DCT chuẩn sẽ được nhóm lại với nhau để tạo nên ký hiệu lượng tử gắn với mỗi chuẩn. Ký hiệu lượng tử được giải mã này sau đó được kết hợp với các thông tin phụ có sẵn tại bộ giải mã để khôi phục hệ số DCT. Giải pháp khôi phục thực hiện trong DISCOVER codec là lỗi bình phương tối thiểu (MMSE) được dựa trên giải pháp được đề xuất trong [16]. Sau khi tất cả các DCT chuẩn được khôi phục, biến đổi cosine ngược rời rạc (Inverse discrete cosine transform: IDCT) được áp dụng và nhận được khung WZ.

DISCOVER codec cung cấp hiệu quả RD cạnh tranh khi so sánh với những bộ codec khác có độ phức tạp mã hóa tương tự (thấp), cụ thể là giải pháp H.264/AVC intra.

## **3.2. Giải pháp PRISM**

- Giải pháp1: PRISM ban đầu, 2002 [10] tại ĐH Bekerley lãnh đạo bởi GS Kannan Ramchadran, sau đó được cải tiến đến phiên bản cuối 2007 [17], đây là loại

monoview, dựa trên block và không có phản hồi và không có tính chất cho kích thước thay đổi.

Trong giải pháp kiến trúc này có 2 điểm cho phép điều chỉnh là trễ và phân phối độ phức tạp giữa bên phát và bên thu tuy nhiên quan tâm hơn cả là dồn hết độ phức tạp cho bên thu.

- Giải pháp 2: Mã PRISM có thể co dãn kích thước, đây cũng là loại monoview, dựa trên khối và không có kênh phản hồi và là sự mở rộng của giải pháp 1. Mục đích của khả năng co dãn kích thước là các bộ giải mã có thể lựa chọn các tập con của dòng bit để cho ra chất lượng giải mã cũng như độ phân giải khác nhau. Điều này đem lại ứng dụng mềm dẻo trong mạng: Cùng 1 dòng bit gửi vào mạng có thể được giải với TV chất lượng cao, cũng có thể giải cho Mobile phone hay các thiết bị đơn giản khác. Đặc điểm này cũng mang lại khả năng tự thích nghi với điều kiện kênh truyền nếu như có thể bị rớt một số gói.

Về cơ bản sơ đồ xử lý cũng giống như giải pháp 1 song phải bổ sung thêm các khối cho phép:

Thay đổi kích thước SNR

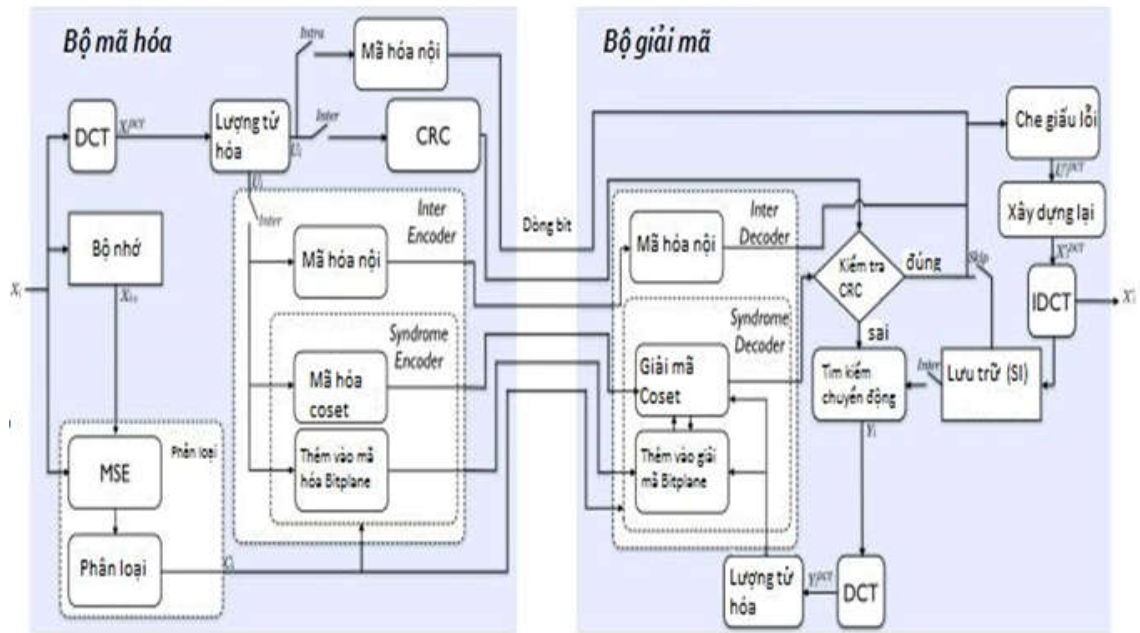
Thay đổi kích thước không gian

Thay đổi kích thước thời gian

- Giải pháp 3: Phát triển bởi C. Yeo and K. Ramchandran, Multiview, dựa trên block, không có kênh phản hồi. Dùng cho hệ nhiều camera với nhiều video view và có tương quan giữa các view, tương quan này được sử dụng trong mã hóa. Phương pháp này khai thác tương quan giữa các camera chồng lên nhau gọi là hệ nhiều camera phân tán đảm bảo chống lỗi cao trong môi trường vô tuyến và cho phép giảm tốc độ mã hóa hơn nữa, mặc dù các Camera này không có truyền tin gì cho nhau

Sự khác biệt cấu trúc trong giải pháp này chỉ tập trung ở bộ giải mã, ở đó thông tin từ nhiều Camera gửi về được tổng hợp đem lại độ tin cậy cao.

Dưới đây là sơ đồ khối cho mã và giải mã của kỹ thuật này được phát triển bởi M. Contente [18]



Hình 3.2: Sơ đồ mã hóa và giải mã PRISM [18]

Hệ thống này hoạt động qua 2 bước: Bước huấn luyện và bước mã hóa-giải mã. Sau đây là mô tả chi tiết về các bước này:

### 3.2.1. Quá trình huấn luyện

Quá trình này dùng một số dãy video đại diện nhằm đánh giá thống kê “tạp âm tương quan” phục vụ cho quá trình mã hóa. Quá trình này cho phép khối mã hóa Syndrome trong bộ mã hóa phân loại các block 8x8 về lượng tương quan thời gian và quyết định mặt phẳng bit cho các hệ số DCT được mã Intra (do tương quan thấp) và gửi đi các bit có trọng số thấp, còn các bit có trọng số lớn (do tương quan cao) được suy từ bộ giải mã. Mặt phẳng bit gửi đến bộ thu sẽ xác định tốc độ mã hóa.

Trong tình huống tối ưu, số bit phát cho bên thu tùy thuộc “tạp âm tương quan” giữa block hiện tại và block dự đoán nó từ khung trước, song như thế dễ phải dò tìm dịch chuyển ở bên mã hóa. Để tránh điều này ta xây dựng trước các cấp độ (lớp) tương quan từ pha huấn luyện. Cụ thể là 16 cấp độ tương quan được lựa chọn để gán cho mức tương quan của mỗi hệ số trong 64 hệ số DCT (block 8x8). Mỗi mức tương quan sẽ được phân một số bit mã hóa tương ứng. Tóm lại quá trình huấn luyện có mục tiêu là:

- Xác định hệ mức “tạp âm tương quan” của mỗi hệ số ứng với mỗi lớp “tạp âm tương quan”
- Sử dụng đặc trưng thống kê tương quan của hệ số, phân tốc độ cho mỗi hệ số tại bộ mã hóa.

Các khối chi tiết ở pha huấn luyện như sau:



- Khối phân loại tương quan: Mỗi khối sẽ so sánh với khối cùng vị trí ở khung trước qua đại lượng MSE. Giá trị nhận được sẽ ứng với 1 cấp độ (lớp) trong 16 cấp độ từ 0 đến 15

- Đếm lớp tương quan: biểu diễn bằng CC(c) tăng 1 sau mỗi bước

- Dò dịch chuyển: Được thực hiện bằng tính MSE giữa khối hiện tại và các khối trong 1 cửa sổ của khung trước để xác định khối có MSE nhỏ nhất. Cửa sổ dò chỉ là 16x16 (xung quanh block cùng vị trí) (chú ý là phép dò này có tính ước lượng thô, rất nhanh, không giống với phép dò trong mã truyền thống)

- DCT: áp dụng cho block hiện tại và block tốt nhất ở khung trước. Kết quả ta có DCT của cả 2 block

- Xác định phương sai và giá trị trung bình: bước này xác định  $E(x)$  và  $Var(x)$  giữa hiệu tuyệt đối của mỗi hệ số  $i$  trong mỗi cấp độ  $c$  giữa *tất cả các block* và các dự đoán tốt nhất của nó. Để tính giá trị trung bình, bộ đếm lớp sẽ chia tổng các sai số của mỗi hệ số trong mỗi lớp cho số lần tính,  $Var$  xác định (3.1):

$$Var(x) = E(x^2) - E(x)^2 \quad (3.1)$$

- Xác định tham số LRV: tập âm tương quan cho mỗi hệ số DCT có tính chất như biến ngẫu nhiên Laplace. Biến này được đặc trưng bởi hàm phân bố tích lũy như sau tùy thuộc  $E(x)$  và  $Var(x)$ :

$$F(x) = \frac{1}{2} \exp\left(-\frac{|x-\theta|}{\lambda}\right) \quad x \leq \theta$$

$$F(x) = 1 - \frac{1}{2} \exp\left(-\frac{|\theta-x|}{\lambda}\right) \quad x > \theta$$

$$\theta = E(x)$$

$$2\lambda^2 = Var(x) \quad (3.2)$$

Như thấy ở trên, mỗi LRV phụ thuộc 2 tham số  $\theta$  và  $\lambda$ , Các tham số này lại được xác định dựa trên  $E(x)$  và  $Var(x)$  của mỗi hệ số  $i$  trong lớp  $c$ . Điều này sẽ được ứng với  $E(i,c)$  và  $Var(i,c)$ . Ở đây tham số  $\theta(i,c)$  đơn giản là bằng trung bình lỗi tuyệt đối tạo ra trước đó  $E(i,c)$ . Để xác định  $\lambda(i,c)$  phải dùng  $Var(i,c)$ .

$$\begin{cases} 2(\lambda(i,c))^2 = Var(i,c) \\ \lambda(i,c) = \sqrt{\frac{Var(i,c)}{2}} \end{cases} \quad (3.3)$$

- Xác định giá trị tập âm tương quan ở mức hệ số DCT: việc xác định đúng CN sẽ xác định hiệu quả méo-tốc độ (RD) của bộ mã hóa. Mỗi hệ số của mỗi lớp sẽ gắn với giá trị LRV, nó cho phép xác định số phần trăm dưới giá trị CN là tập âm tương quan của 1 lớp tương quan xác định khi sử dụng đảo ngược

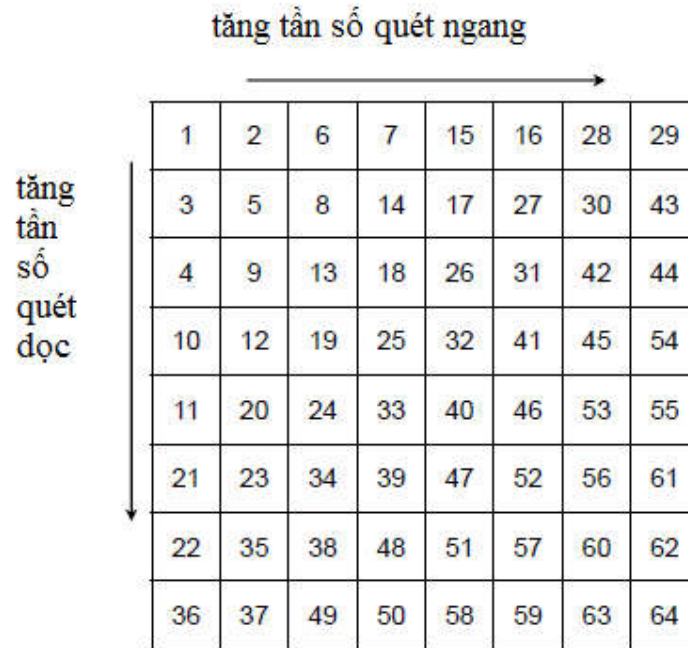
của hàm phân bố tích lũy. Ta lựa chọn CN thông qua giá trị  $p$ . Ví dụ nếu  $p=0.99$  thì CN là giá trị của trên 99% giá trị tập âm tương quan của hệ số này. Trên thực tế rất quan trọng quyết định giá trị  $p$  nào được dùng vì nó sẽ định nghĩa giá trị tập âm tương quan gắn với hệ số mỗi lớp:

$$\begin{cases} \text{CN}(i,c) = \theta(i,c) + \lambda(i,c) \ln(2p) & p \leq 0.5 \\ \text{CN}(i,c) = \theta(i,c) - \lambda(i,c) \ln(2(1-p)) & p > 0.5 \end{cases} \quad (3.4)$$

### 3.2.2. Quá trình mã hóa

#### 3.2.2.1. Biến đổi Cosin

DCT áp dụng cho 64 mẫu sáng (block 8x8, luminance) và nối nhau theo thứ tự zig-zag. Ô 1 ứng với hệ số tần số DC có giá trị tỉ lệ với trung bình tất cả các mẫu trong khối. Các hệ số khác là AC.



Hình 3.3: Quét zig-zag

$$\text{DCT}(u,v) = \frac{1}{\sqrt{2N}} C(u)C(v) \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} p(x,y) \cos\left[\frac{(2x+1)u\pi}{2N}\right] \cos\left[\frac{(2y+1)v\pi}{2N}\right]$$

$$C(g) = \begin{cases} \frac{1}{\sqrt{2}} & g = 0 \\ 1 & g > 0 \end{cases}$$

### 3.2.2.2. Lượng tử

Phân bước lượng tử khác nhau với các hệ số DCT khác nhau, vì mắt người có độ nhạy khác nhau. Chọn bước lượng tử thích hợp cho phép tăng hệ số nén và trong khi giữ nguyên chất lượng ảnh, vì mắt người ít nhạy với tần số cao nên những hệ số này gán bước lượng tử lớn.

Mã syndromer dựa trên khoảng cách giữa các từ mã lượng tử liên tiếp thể hiện biên độ lượng tử các hệ số DCT để chọn mặt phẳng bit mã (kích thước coset). Nếu từ mã chưa đủ xa với nhau, bộ mã syndromer phải tăng kích thước coset cho đảm bảo khoảng cách.

Lượng tử được giả thiết có thể điều chỉnh được chất lượng, tức là cho phép trả giá giữa độ nén và chất lượng video giải mã. Chất lượng được đánh số từ 1-99 với 1 là chất lượng kém nhất hay nén cao nhất. Mỗi chỉ số chất lượng gán với 1 ma trận lượng tử ví dụ ma trận có chất lượng chỉ số 50 là:

$$QS_{50} = \begin{bmatrix} 16 & 11 & 10 & 16 & 24 & 40 & 51 & 61 \\ 12 & 12 & 14 & 19 & 26 & 58 & 60 & 55 \\ 14 & 13 & 16 & 24 & 40 & 57 & 69 & 56 \\ 14 & 17 & 22 & 29 & 51 & 87 & 80 & 62 \\ 18 & 22 & 37 & 56 & 68 & 109 & 103 & 77 \\ 24 & 35 & 55 & 64 & 81 & 104 & 113 & 92 \\ 49 & 64 & 78 & 87 & 103 & 121 & 120 & 101 \\ 72 & 92 & 95 & 98 & 112 & 100 & 103 & 99 \end{bmatrix}$$

Ma trận lượng tử tạo ra từ lúc bắt đầu mã hóa. Chất lượng tính theo phương trình:

$$\begin{cases} QS_k(u,v) = QS_{50}(u,v) \frac{100-k}{50} & k \geq 50 \\ QS_k(u,v) = QS_{50}(u,v) \frac{50}{k} & k < 50 \end{cases}$$

Sau đó là xác định ma trận hệ số được lượng tử QDCT(u,v) từ ma trận độ lớn DCT (u,v), quá trình tính bước lượng tử khác nhau với DC và AC:

$$QDCT_{DC} = \left[ \frac{DCT}{QSk} \right]$$

$$QDCT(u,v) = \text{round} \left( \frac{DCT(u,v)}{QSk(u,v)} \right)$$

### 3.2.2.3. Phân loại block

Có mục đích phân 1 trong 16 lớp tap âm tương quan cho mỗi block 8x8 trong khung được mã. Thực hiện như sau:

a) *Khởi phát*: đây là khung đầu tiên có thông tin nội khung (Intra).

b) *Xác định MSE*: từ khung thứ 2 MSE được xác định cho mỗi block trong khung so với khung trước đó MSE(xb, yb) ở đó xb, yb là chỉ số hàng và cột của block trong khung

Công thức tính MSE như sau:

$$MSE = \frac{1}{m.n} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [ I(i,j) - K(i,j) ]^2$$

c) *Cập nhật khung đã qua*: Khi các khối trong khung đã được phân loại, nó được lưu trữ và quá trình lại quay lại bước b).

Liên quan đến 3 mode mã là:

i) *Mode block* bỏ qua (skip) không cần bất kỳ mã bổ sung nào, chỉ tạo tốc độ phân loại khối

ii) *Block intra* : được mã như H263

iii) *Block inter*: được mã như mô tả sau

### 3.2.2.4. Mã syndrome

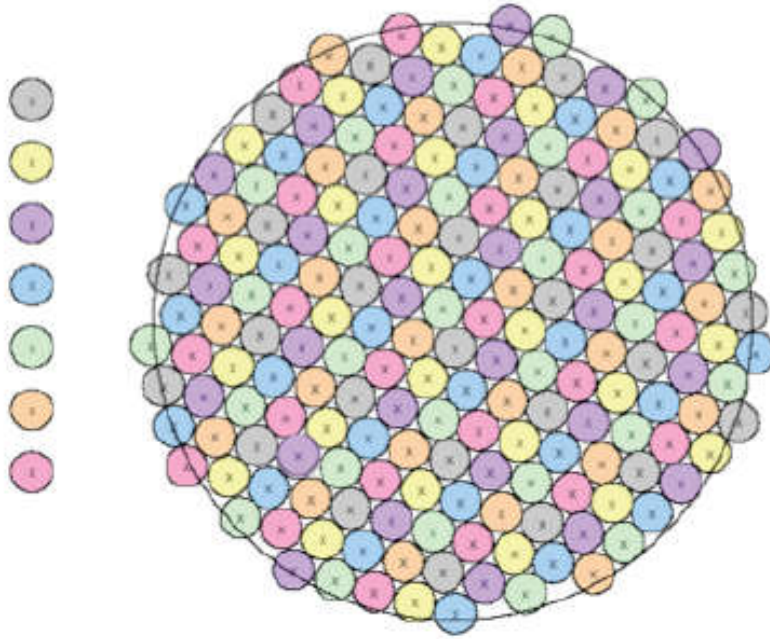
Dựa trên 2 phương pháp: (liên kết hoặc tách biệt), chúng là Mã coset syndrome và mã syndrome có mặt phẳng bit bổ sung.

Mục tiêu của mã syndrome nhằm đạt độ nén với độ phức tạp thấp. Lý thuyết của nó (định lý WZ) là giả thiết bộ giải mã truy cập vào thông tin lẻ, bổ sung chúng tương quan với thông tin block hiện tại. Trong trường hợp xác định thông tin lẻ là khối đã bù chuyển động có được từ khung giải mã trước đó. Mỗi khối đã bù chuyển động là *ứng cử viên* của thông tin lẻ. Nếu cấu trúc lỗi hay phân bố thống kê giữa từ mã lượng tử hiện tại X và thông tin lẻ Y được biết (tại mức hệ số DCT) thì có thể giảm lượng thông tin truyền đến bộ giải mã để có thể giải mã chính xác X dựa trên Y.

Hiệu X-Y gọi là tap âm tương quan N, chúng không thể biết chính xác ở bộ mã hóa cũng như ở bộ giải mã song có thể được ước lượng (Đó chính là giá trị CN tại pha huấn luyện). Bộ mã hóa ước lượng tap âm tương quan dùng khung gốc ban đầu mà không cần ước lượng dịch chuyển (dò tìm phức tạp) mà dựa trên hiệu giữa các khối cùng vị trí gắn với pha huấn luyện để xác định dải tap âm tương quan.

Chiến lược mã hóa là chia không gian từ mã X thành tập nhiều từ. các tập này (chứa các từ mã có khoảng cách bằng nhau) gọi là coset xác định theo chỉ số hay syndrome chúng, Chỉ số này cần ít thông tin hơn để biểu diễn. Nếu khoảng cách giữa các từ mã trong coset đủ lớn hơn 2 lần CN được ước lượng thì có thể khôi phục từ mã X dùng Y và tập coset được phát.

Coset là tập con chọn từ tập gốc. Tất cả các coset phải không chồng lấn lên nhau và lấp đầy tập gốc. Trên hình minh họa vòng tròn lớn như là tập không gian từ mã X chứa các tập từ mã như các vòng tròn nhỏ có bán kính CN. Các vòng tròn cùng màu gắn với 1 coset



Hình 3.4: Mô hình mã Syndrome [18]

Trên thực tế mỗi hệ số được lượng tử khác 0 được mã hóa độc lập. Chỉ số coset được định nghĩa bởi  $Csize()$  là các bit trong số nhẹ, nó chia không gian từ mã thành  $2^{Csize()}$  tập coset khác nhau.

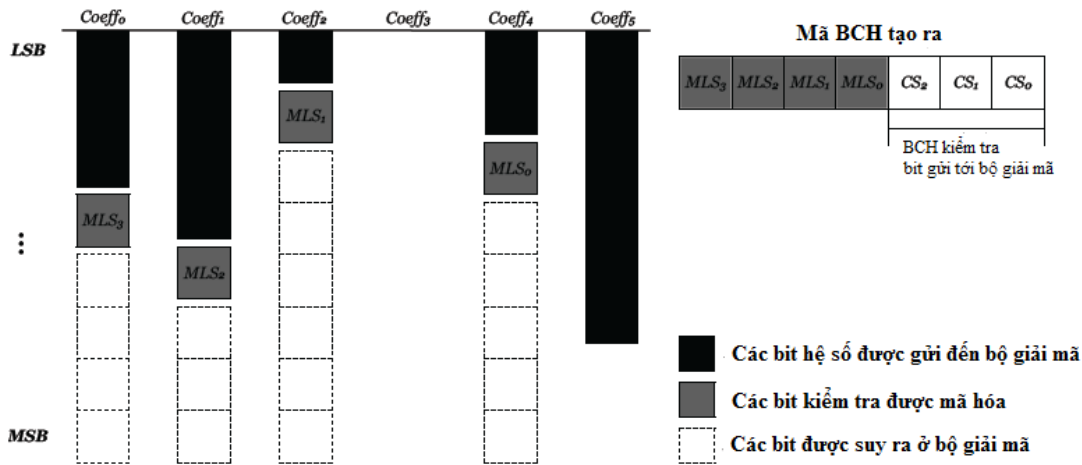
Khoảng cách Hamming giữa các từ mã trong coset là  $2^{Csize()}$ . Để xác định khoảng cách các giá trị DCT phải nhân khoảng cách Hamming với bước lượng tử của hệ số xác định  $i$ , hay là  $QS(i)$ . Để giải mã thành công, khoảng cách các giá trị DCT trong coset phải lớn hơn 2 lần tạp âm tương quan định nghĩa cho các hệ số DCT tương ứng.

$$QS(i) \times 2^{Csize(c,i)} > 2 \times CN(c,i)$$

Trong DVC, không phải tất cả các hệ số trong block đều được mã syndrome. 15 hệ số zig-zag đầu tiên được mã như intra. Điều này là do lượng tương quan của mỗi hệ số ở đây không đủ lớn. Cần chú ý là những giá trị này gắn với số bit được chọn cho CRC.

### 3.2.2.5. Mã hóa coset syndrome

Quá trình này nhằm xác định kích thước và mã hóa chỉ số coset cho mỗi hệ số DCT khác 0 được mã hóa syndrome của mỗi block mã liên khung. Trên thực tế chỉ số coset được mã chính là các bit có trọng số nhỏ nhất trong mỗi hệ số DCT tương ứng với bit đen.



Hình 3.5: Mặt phẳng bit syndrome [18]

Để nhận được ưu điểm xác suất cao của các hệ số zero, lõi ra của bộ mã có 4 nhánh.

### Quá trình mã hóa coset như sau:

- Khởi phát: Nhằm xác định khoảng cách tối thiểu giữa các từ mã lượng tử liên tiếp trong một coset cho mỗi hệ số DCT  $i$  trong mỗi lớp tương quan  $c$ . Giá trị tap âm tương quan  $CN(c,i)$  nhận được tại cuối pha huấn luyện. Bây giờ đối với mỗi hệ số của mỗi lớp trung gian, kích thước chỉ số coset hay số bit trong số nhẹ nhất  $Clsize(c,i)$  được dùng để mã coset được xác định. Điều này đạt được bằng cách tăng  $Clsize(c,i)$  bắt đầu từ 0 đến khi (3.13) được thỏa mãn.

- Tạo mã (last, run, level): Một khối lượng từ quét zig-zag nhận được và mã dùng 3 ký hiệu (last, run, level) cho mỗi hệ số. Last set=1 khi hệ số khác 0 cuối cùng theo thứ tự zig-zag kết thúc. Run xác định số hệ số zero trước hiện tại và level chỉ thị giá trị hệ số DCT hiện tại khác zero.

- Xác định coset: mỗi hệ số DCT khác zero được xác định bởi 1 coset. Bước đầu tiên là cộng 127 với level hệ số AC để nhận được số dương. DC không được làm như vậy vì nó vốn là số nguyên dương. Khi chỉ số coset được định nghĩa bởi các bit cuối  $Clsize(c, i)$  của mỗi từ mã lượng tử, được xác định trong bước bắt đầu, một phép cộng nhị phân được cấp giữa giá trị được tạo ra và mặt nạ với độ dài  $Clsize(c,i)$ . Ví dụ nếu  $Clsize(c, i)$ , cũng được gọi là độ sâu, là 3 và level là -20. Giá trị tạo ra là  $-20+127=107$  (nhị phân là 0110 1011) với độ sâu này mặt nạ là: 0000 0111, coset được xác định sẽ là: 0110 1011 x 0000 0111 = 011 (in decimal 3).

- Tạo mã (last, run, depth, path): mục tiêu bước này là mã syndrome với chữ cái thích hợp. Điều này thực hiện bằng chuyển mã được tạo ra trước đó sang 4 chữ cái, nó thích hợp hơn với tính chất mã syndrome. Các trường last, run có ý nghĩa như trước, xác định dương các hệ số DCT trong block. Trường level được thay bằng depth và path chịu trách

nhiệm biểu diễn coset. Depth xác định số bit của coset, path lưu trữ các bit nhận diện coset. Trong thực hành Depth là Clsize và Path bằng kết quả trong bước trước.

- Mã entropy: nhằm có ưu điểm thống kê của chữ cái 4 nhánh trong bước trước. Mỗi từ mã (*Last, Run, Depth, Level*) sẽ được mã dùng mã độ dài thay đổi.

### 3.2.2.6. Mã hóa mặt phẳng bit bổ sung syndrome

Mô dul này nhằm nén và truyền mặt phẳng bit cosey bổ sung. Giá trị tạp âm tương quan ước lượng tại bộ mã hóa không thật chính xác so với thực tế. Đối với đa số block, giá trị tạp âm tương quan ước lượng lớn hơn giá trị thật nên giải điều chế thành công. Tuy nhiên có 1 tỷ lệ nhỏ block, giá trị ước lượng này nhỏ hơn giá trị thật, nên khoảng cách từ mã trong coset không đủ nên sẽ chọn nhầm ứng cử viên ở bên giải mã. Và nếu 1 số hệ số giải mã sai dẫn đến CRC được truyền để phát hiện cũng không thành công. Đây là giá phải trả cho việc ước lượng tạp âm tương quan. Để giảm điều này kích thước các chỉ số coset phải tăng thêm một lượng để số phần trăm giải mã không chính xác có thể chấp nhận được. Tuy nhiên như vậy lại không đạt được độ nén mong muốn.

Do nguyên nhân này một chiến lược lai ghép được triển khai: Thay cho việc truyền thêm các bit cho mỗi coset, các bit bổ sung được truyền phối hợp với giải mã để dùng tổng kiểm tra ngắn hơn. Các từ được thành các tổng kiểm tra gọi là từ mã mặt phẳng bit chỉ số bổ sung. Bên phía giải mã, tổng kiểm tra có khả năng hiệu chỉnh một lượng giới hạn chỉ số coset mà bị lựa chọn sai do khoảng cách từ mã không đủ trong coset.

Khi giải mã syndrome kết thúc, bộ giải mã mặt phẳng bit bổ sung truy cập từ mã lượng từ đã giải cho phép tạo lại từ mã mặt phẳng bit chỉ số bổ sung. Đây chính là một phiên bản lỗi của từ mã mặt phẳng bit bên mã hóa, từ đó tổng kiểm tra sẽ hiệu chỉnh nó. Sau khi hiệu chỉnh chỉ số coset, giải mã sẽ dùng chỉ số tăng thêm này tiến hành giải mã lại với cùng những ứng cử viên trước đó.

Trong bộ codec IST-PRISM, các bit thêm (bit xám trong hình 3.5) được mã phối hợp dùng mã BCH. Phân mã hệ thống trong BCH chưa bit bổ sung bị loại, chỉ có tổng kiểm tra được truyền đi đến bộ giải mã. Mã BCH được lựa chọn vì chúng khá đơn giản và thích hợp với các khối nhỏ.

### 3.2.2.7. Mã CRC

Mã CRC được thiết kế để chọn từ mã đúng trong các ứng cử viên, vì có nhiều ứng viên trong cửa sổ dò tìm bên phải giải mã. Trong PRISM mã lựa chọn là CRC-16 có đa thức tạo mã là:

$$x^{16} + x^{12} + x^5 + 1$$

Cuối cùng của bộ mã hóa là việc tạo ra các dòng bit có cấu trúc như sau:

CT 2 bits	Syndrome Variable length	Intra Coefficients Variable length	CRC 16 bits
--------------	-----------------------------	---------------------------------------	----------------



### 3.2.3. Quá trình giải mã

#### 3.2.3.1. Giải mã syndrome

Mục tiêu là tìm từ mã trong coset nhận được.

Trước tiên là xác định coset, sau đó là hiệu chỉnh coset dùng tổng kiểm tra BCH và từ mã mặt phẳng bit chỉ số bổ sung.

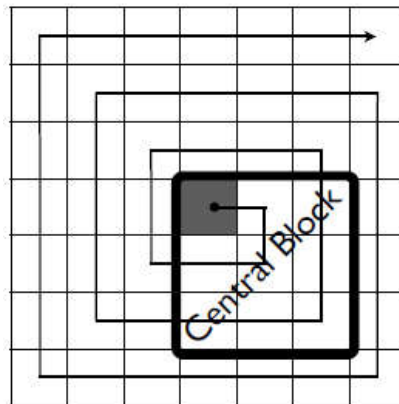
Do có nhiều ứng cử viên trong coset cùng có tương quan với từ mã cần tìm nên cần dò tìm và thử để tìm được từ mã đúng. Phép thử thành công khi kết hợp với mã kiểm tra CRC. Chi tiết các khối trong bộ giải mã như sau:

Modul này có chức năng chọn từ mã trong coset sử dụng thông tin lẻ. Thông tin lẻ được dùng như một block DCT được lượng tử và quét zig-zag.

- i) Giải mã entropy
- ii) Giải mã chỉ số coset
- iii) Tạo lại ký hiệu
- iv) Giải mã mặt phẳng bit bổ sung syndrome

#### 3.2.3.2. Dò chuyển động

Module này tạo ra các ứng cử kết hợp với syndrome nhận được để giải mã đúng. DVC thích hợp cho dãy có nội dung chuyển động giới hạn như video surveillance and videotelephony. Với dãy có nội dung chuyển động thấp các block bù chuyển động sẽ nằm gần nhất block đang giải mã là có nhiều tương quan với nó. Đó là nguyên nhân tác giả dùng cách quét theo đường xoắn ốc để tìm các ứng cử.



Hình 3.6: Chương trình quét xoắn ốc



## CHƯƠNG 4: ĐÁNH GIÁ ƯU NHƯỢC ĐIỂM BỘ MÃ HÓA DVC

### 4.1. Điều kiện đánh giá

Chương này sẽ đưa ra các kịch bản so sánh đánh giá mã DVC và mã truyền thống, tập trung vào đánh giá hiệu năng nén và độ phức tạp cho 4 chuỗi video với các đặc tính và nội dung khác nhau. Hình 4.1. hiển thị khung hình đầu tiên của 4 chuỗi video này.



Hình 4.1 Mô tả khung hình đầu tiên của 4 chuỗi video

Do các thành phần màu (Chroma) không ảnh hưởng đáng kể tới hiệu năng nén của bộ mã hóa nên trong thực nghiệm, chỉ những thành phần độ chói (Luma) được sử dụng để mã hóa và đánh giá so sánh. Các thông số khác của chuỗi video được mô tả tóm tắt như ở bảng 4.1.

Bảng 4.1. Bảng mô tả tóm tắt các thông số sử dụng đánh giá

Tên chuỗi video	Foreman	Hall Monitor	Coastguard	Soccer
Số khung hình	299	329	299	299
Kích cỡ	QCIF (176 × 144)			
Tốc độ khung hình	15 Hz			
Kích cỡ nhóm ảnh	2 (Key – WZ – Key)			
Tham số lượng tử	QM1, QM3, QM5, QM7			

- **Bảng lượng tử sử dụng cho mã hóa DVC**

Bảng 4.2: Bảng lượng tử mã hóa DVC

QM1	QM3	QM5	QM7
$\begin{bmatrix} 16 & 8 & 0 & 0 \\ 8 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 32 & 8 & 4 & 0 \\ 8 & 4 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 32 & 16 & 8 & 4 \\ 16 & 8 & 4 & 4 \\ 8 & 4 & 4 & 0 \\ 4 & 4 & 0 & 0 \end{bmatrix}$	$\begin{bmatrix} 64 & 32 & 16 & 8 \\ 32 & 16 & 8 & 4 \\ 16 & 8 & 4 & 4 \\ 8 & 4 & 4 & 0 \end{bmatrix}$

- **Các bộ mã hóa tham chiếu:**

- **H.264/Intra**: Đây là bộ mã hóa phổ biến, cấu hình từ chuẩn H.264/AVC và được tải về từ [19]; trong đó, tất cả các khung hình chỉ sử dụng phương pháp mã hóa dự đoán trong khung. Do vậy, phương pháp mã hóa này đòi hỏi thời gian mã hóa thấp, rất phù hợp với các ứng dụng như mạng sensor,... Cần lưu ý rằng, H.264/Intra chính là đối tượng so sánh chính của bộ mã hóa video phân tán DVC.
- **H.264/No ME**: Đây cũng là bộ mã hóa video phổ biến, cấu hình từ chuẩn H.264/AVC và được tải về từ [19]. Tuy nhiên, không chỉ phương pháp dự đoán trong khung mà cả liên khung với giả thiết vector dự đoán bằng không được sử dụng. Trong bộ mã hóa này, phương pháp ước lượng vector chuyển động không được sử dụng. Do vậy, yêu cầu thời gian mã hóa cũng rất thấp. Tuy nhiên, do có khai thác tính tương quan giữa các khung tại phía mã hóa, bộ mã hóa này rất nhạy cảm với các nhiễu liên khung. .
- **DVC-UET**: Đây là bộ mã hóa được trình bày và phân tích trong luận văn này. Bộ mã hóa được xây dựng tại Visual Processing and Communications Group, VNU-UET [20] và dựa trên mô hình của nhóm nghiên cứu tại đại học Stanford. Trong đó, các khung hình (key frames) được mã hóa với H.264/AVC Intra bởi bộ mã hóa H.264/AVC Intra mang lại hiệu năng nén cao đồng thời yêu cầu độ phức tạp thấp nhất do không khai thác bộ ước lượng chuyển động tại phía phát. Các giá trị lượng tử cho khung chính được lựa chọn với mục tiêu là có chất lượng video được giải mã gần như liên tục cho toàn bộ các khung hình và được mô tả như ở bảng 4.3.

Bảng 4.3: Giá trị lượng tử cho khung chính tại GOP=2, QCIF 15Hz

<i>Sequences</i>	<i>QM1</i>	<i>QM3</i>	<i>QM5</i>	<i>QM7</i>
Foreman	40	38	34	29
Hall Monitor	37	36	33	29
Coastguard	38	37	30	30
Soccer	44	41	36	31

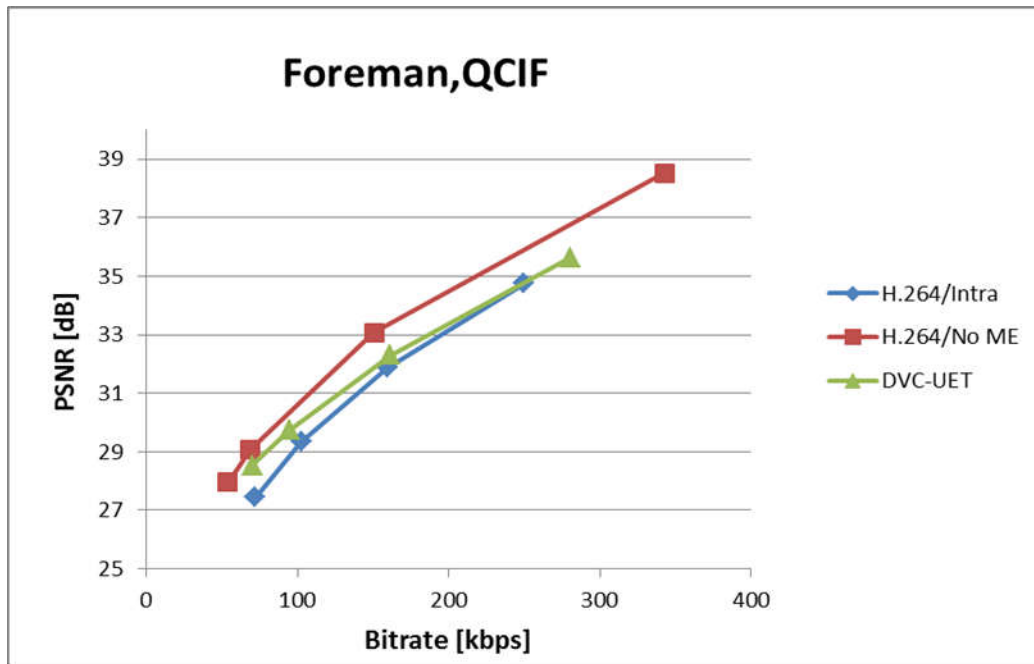
#### 4.2. Đánh giá hiệu năng nén

Hiệu năng nén là một trong những yêu cầu đầu tiên và quan trọng nhất đối với các phương pháp mã hóa video; một phương pháp được đánh giá là tốt nếu hiệu năng nén của nó cao hơn so với các phương pháp trước đó. Trong việc đánh giá hiệu năng nén của mã hóa video, mối tương quan giữa chất lượng hình ảnh sau giải mã, đo bằng giá trị đỉnh của tín hiệu trên tạp âm (Peak Signal to Noise Ration – PSNR) và lượng bit cần sử dụng để mã hóa (bitrate – bit per second – bps). Trong đó, PSNR được xác định như sau:

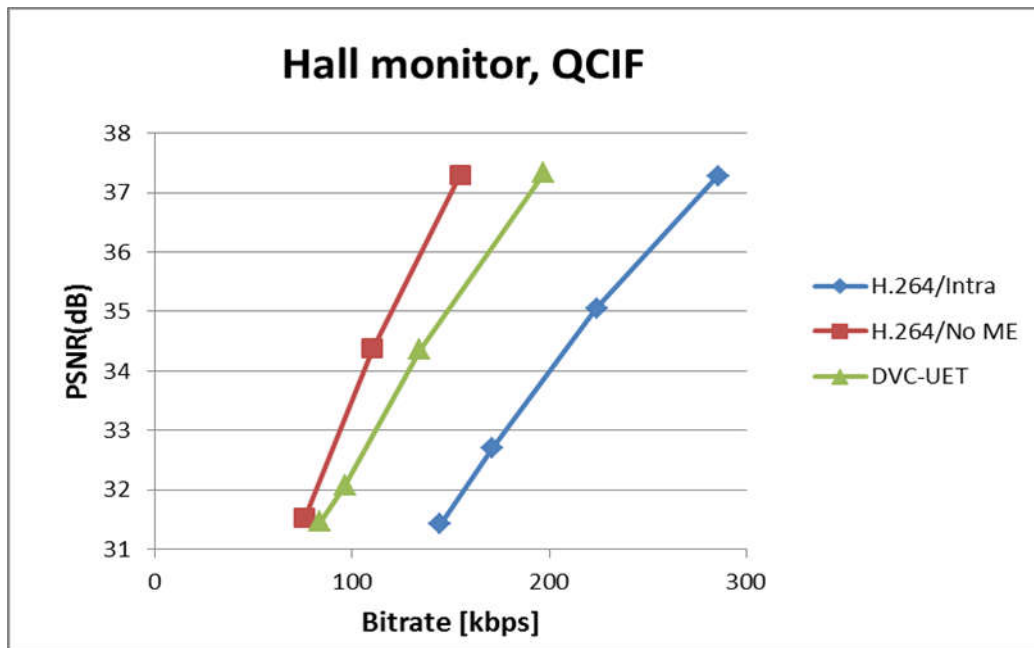
$$PSNR = 10 \log_{10} \frac{255^2}{MSE}$$

MSE: Mean Squared Error là giá trị trung bình bình phương lỗi (sự khác biệt) giữa tín hiệu trước và sau giải mã.

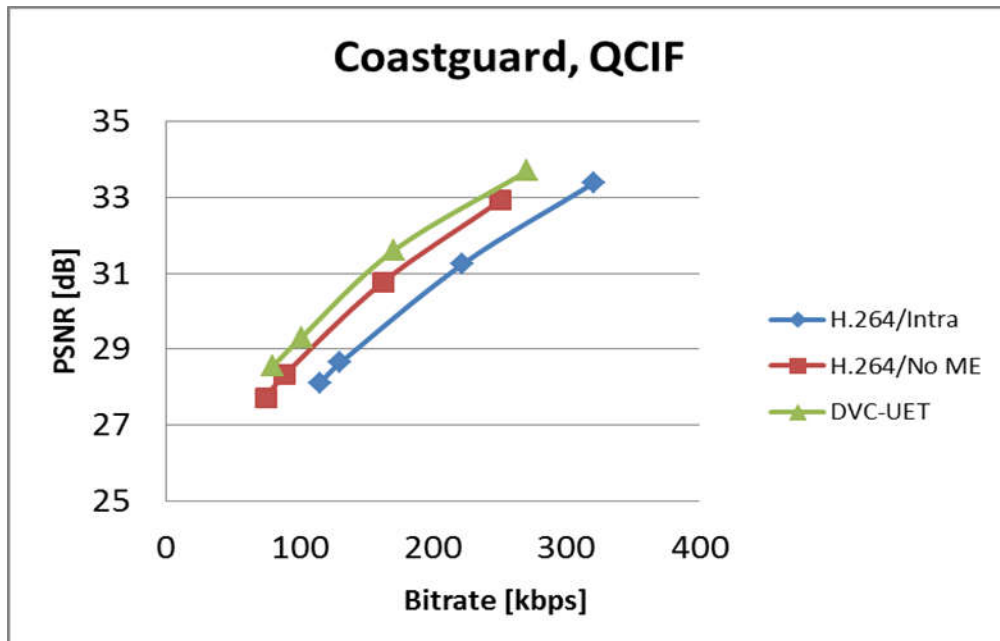
Trên cơ sở đó, việc đánh giá hiệu năng nén và so sánh với chuẩn H.264/AVC được xác định cho 4 trường hợp lượng tử khác nhau như mô tả ở bảng 4.2 và 4.3. Hình 4.2 đến 4.5 thể hiện biểu đồ so sánh hiệu năng nén cho các chuỗi video được đánh giá.



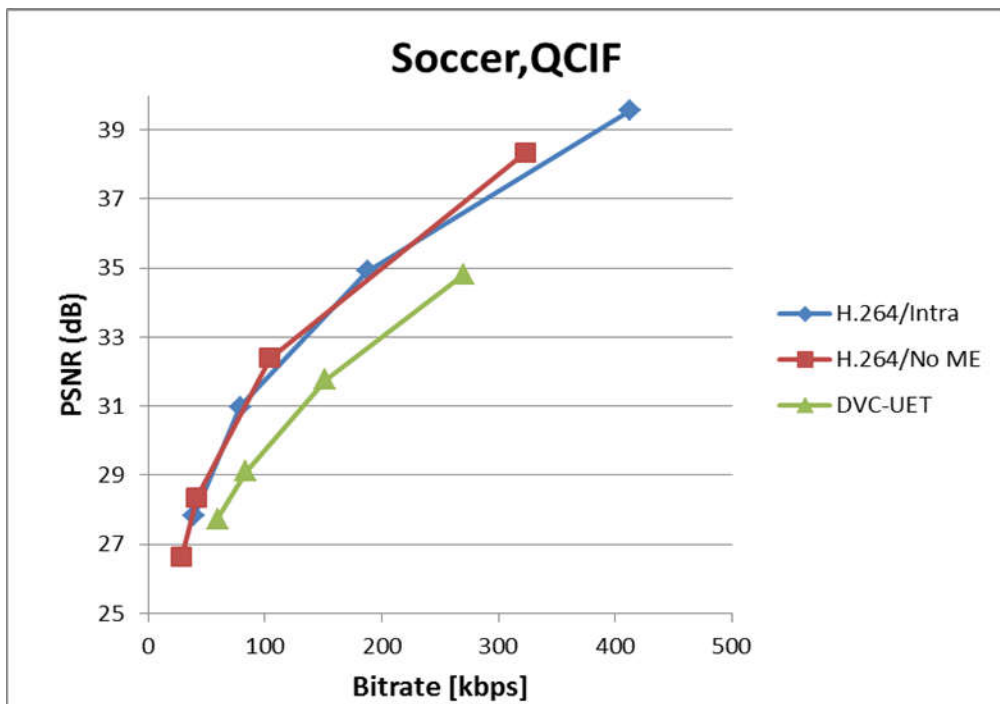
Hình 4.2: Biểu đồ so sánh hiệu năng nén - Foreman



Hình 4.3: Biểu đồ so sánh hiệu năng nén - Hall monitor



Hình 4.4: Biểu đồ so sánh hiệu năng nén - Coast guard



Hình 4.5: Biểu đồ so sánh hiệu năng nén – Soccer

Trên cơ sở các kết quả thu được như mô tả ở hình 4.2 – 4.5, một số kết luận có thể rút ra như sau;

- **So sánh với H.264/Intra**

- Trong hầu hết các trường hợp so sánh thì mô hình mã hóa DVC đều có hiệu suất nén cao hơn so với mô hình mã hóa H264/Intra
- Ở trong 3 trường hợp Coastguard, Hall Monitor và Foreman thì vật thể chuyển động với tốc độ không cao nên sử dụng mã DVC đạt hiệu quả cao. Chỉ trong trường hợp Soccer vật thể chuyển động với tốc độ cao thì mô hình truyền thống lại đạt được hiệu quả nén tốt hơn. Điều này là do việc ước lượng chuyển động tại phía mã hóa của chuẩn nén video truyền thống

- **So sánh với H.264/ No ME**

Trong các trường hợp so sánh thì mô hình mã hóa DVC có hiệu suất nén đạt xấp xỉ so với mô hình H.264/ No ME. Trong trường hợp chuỗi Coastguard thậm chí còn đạt hiệu suất nén cao hơn tuy nhiên đây không phải là mục tiêu so sánh chính trong luận văn này

### 4.3. Đánh giá độ phức tạp

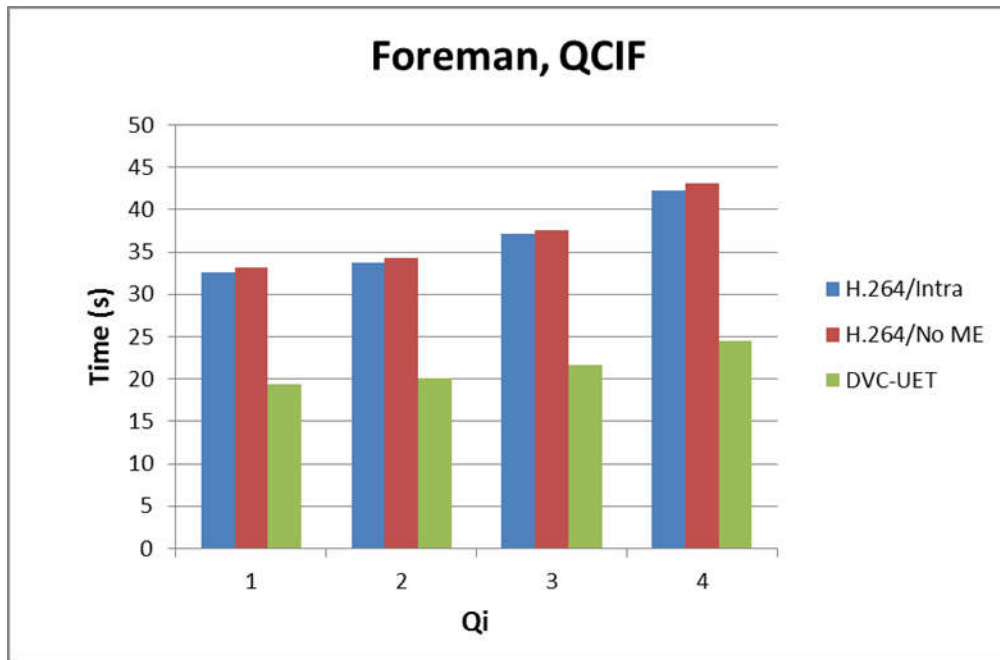
Độ phức tạp trong việc mã hóa, cùng với hiệu năng nén và khả năng chống chịu lỗi là ba yêu cầu cơ bản của một chuẩn nén video hiệu quả. Do vậy, nghiên cứu sẽ không thể trọn vẹn nếu không đánh giá độ phức tạp của phương pháp mã hóa video phân tán DVC, nhất là so sánh với chuẩn H.264/AVC.

Có nhiều cách đánh giá *độ phức tạp* của một chuẩn mã hóa, như xác định số lượng các phép tính trong các thuật toán hay một cách phổ biến và đơn giản nhất là đo thời gian mã hóa của các phương pháp mã hóa trong cùng một điều kiện kiểm thử chuẩn.

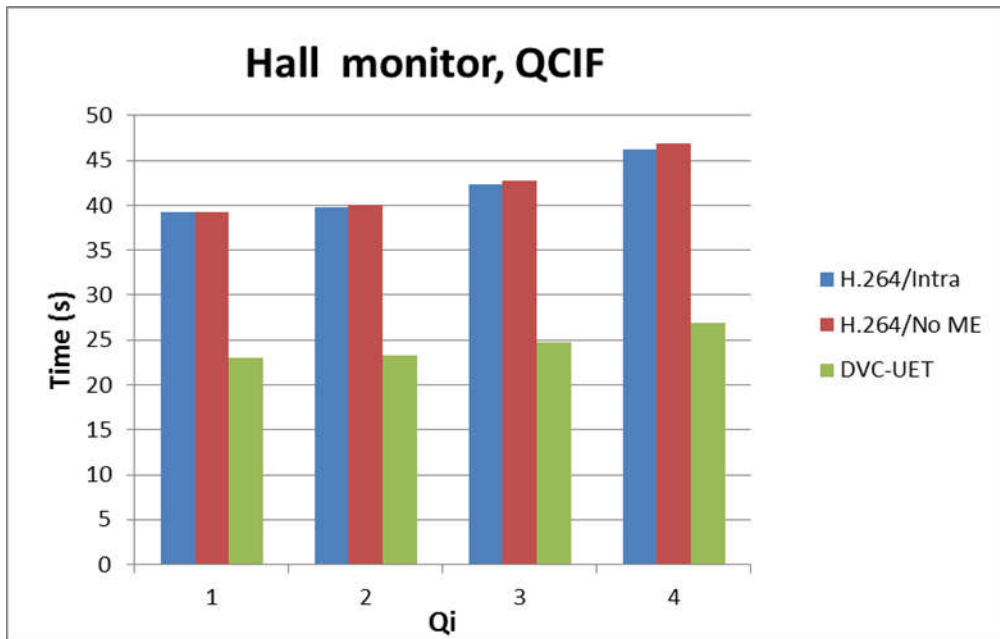
Về số lượng phép tính có thể thấy từ chương 1 khi dò tìm vecto dịch chuyển trung bình sẽ tốn rất nhiều phép tính (trong ví dụ đưa ra là  $8,55 \times 10^9$  phép tính trên giây ) trong khi ở DVC không có phép dò tìm vecto dịch chuyển ở bên mã hóa, số phép tính chủ yếu cho các khung reference theo mã hóa intra và tạo ra thông tin phụ.

Trong nghiên cứu này, phương pháp thứ hai, đo lường thời gian mã hóa của các phương pháp mã hóa như DVC-UET, H.264/Intra và H.264/No ME được sử dụng. cùng trên máy tính với cấu hình bao gồm 1 chip dual core Pentium D 3.4 GHz, 2GB RAM, ngôn ngữ lập trình C++ và Microsoft Visual Studio được sử dụng.

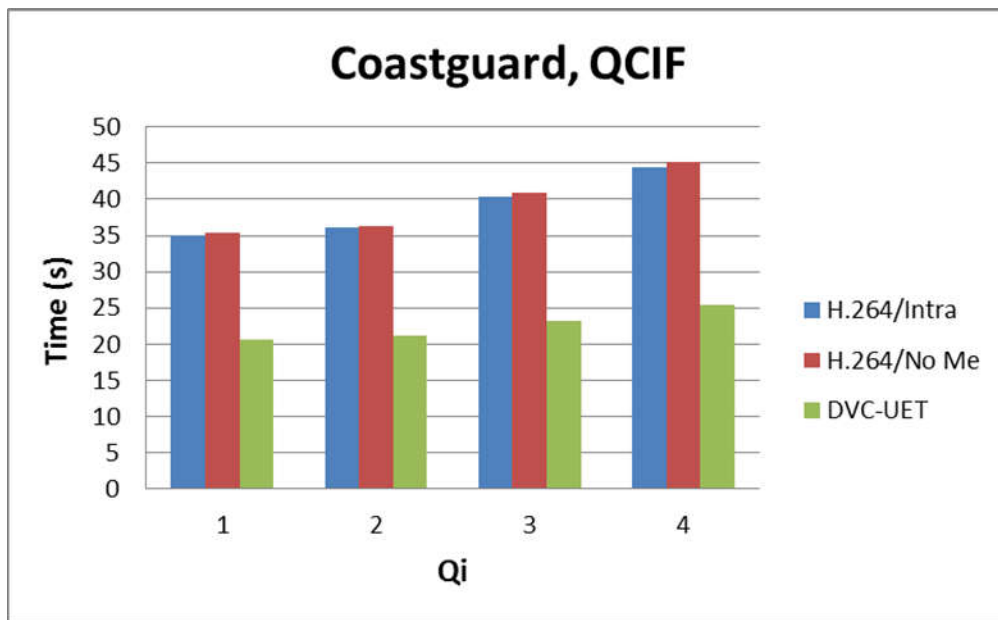
Hình 4.6 đến 4.9 mô tả và so sánh thời gian mã hóa đối với các chuẩn H.264/AVC và phương pháp DVC.



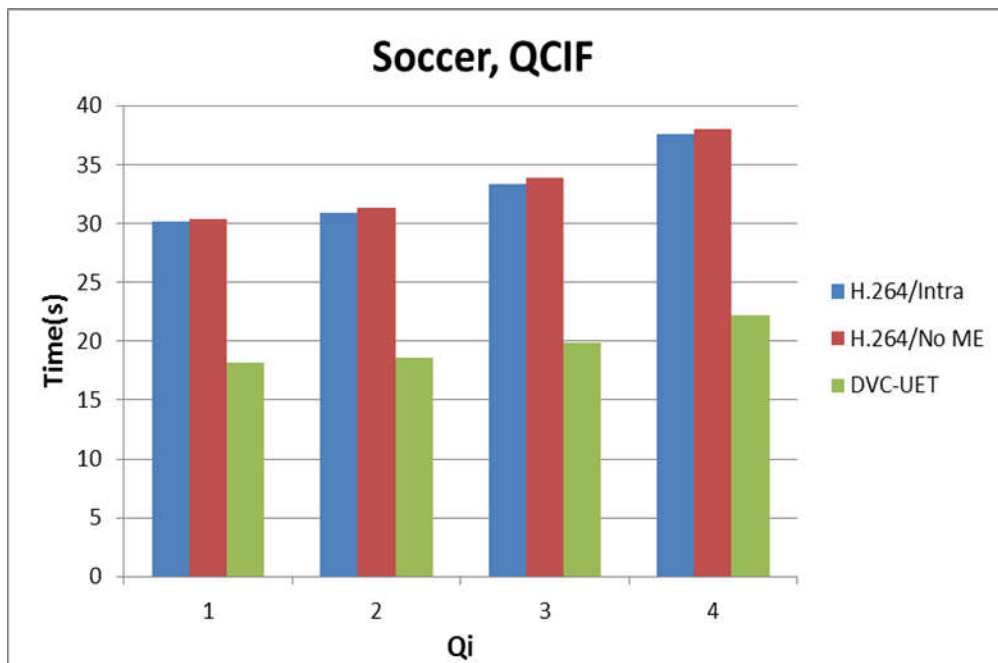
Hình 4.6: Biểu đồ so sánh thời gian mã hóa - Foreman



Hình 4.7: Biểu đồ so sánh thời gian mã hóa - Hall monitor



Hình 4.8: Biểu đồ so sánh thời gian mã hóa - Coastguard



Hình 4.9: Biểu đồ so sánh thời gian mã hóa - Soccer



### **Nhận xét:**

- Các biểu đồ so sánh thời gian mã hóa đã cho thấy mô hình mã hóa DVC có thời gian mã hóa thấp hơn hẳn so với các mô hình còn lại trong tất cả các trường hợp nên có thể khẳng định rằng mô hình mã hóa DVC có độ phức tạp thấp nhất ở phần mã hóa, phù hợp với nhu cầu phát triển các mạng video giám sát. Điều này là bởi so với bộ mã hóa H.264/Intra, bộ mã hóa DVC-UET không sử dụng cách tạo dự đoán năng nề (với 9 mode dự đoán) và bộ tối ưu hệ thống (Rate-Distortion Optimization RDO) rất phức tạp tại phía phát.

## KẾT LUẬN

Trước nhu cầu ngày càng cao về một bộ mã hóa video có hiệu quả nén tốt nhưng đòi hỏi thời gian mã hóa thấp; đặc biệt là đối với các ứng dụng trong mạng sensor, luận văn này đã tập trung vào việc giới thiệu, phân tích và so sánh phương pháp mã hóa video phân tán – DVC. Mã hóa video phân tán DVC là một trong những hướng nghiên cứu đang được quan tâm hiện nay do những đặc điểm như hiệu năng mã hóa cao, độ phức tạp mã hóa thấp và khả năng chống chịu nhiễu hiệu quả.

Luận văn được chia làm năm chương. Chương 1 giới thiệu về các kỹ thuật được sử dụng trong bộ mã hóa video dự đoán truyền thống. Chương 2 tóm tắt qua phần cơ sở lý thuyết của bộ mã hóa video phân tán, cụ thể là hai định lý Slepian-Wolf và Wyner-Ziv. Tiếp theo, chương 3 trình bày và phân tích hai hướng xây dựng bộ mã hóa video phân tán trong thực tế, đề xuất bởi hai nhóm nghiên cứu tại trường đại học Stanford và đại học Berkeley. Chương 4 trình bày kết quả so sánh hiệu quả nén và độ phức tạp mã hóa của mô hình mã hóa video phân tán DVC.

Như đã trình bày ở chương 2 và chương 3, kỹ thuật mã hóa video phân tán có khả năng mang lại hiệu quả nén cao để thích ứng với các kênh truyền vô tuyến hạn hẹp, tuy nhiên có điểm khác biệt căn bản so với kỹ thuật mã hóa truyền thống (trình bày ở chương 1) là độ phức tạp về mặt thuật toán. Cụ thể, mã hóa video truyền thống rất nặng ở phần mã hóa còn kỹ thuật DVC phần mã hóa lại rất nhẹ. Do vậy, mã hóa video phân tán rất phù hợp với nhu cầu phát triển các mạng giám sát hiện nay.

Như trình bày ở chương 4, sự khác biệt này là ở mã hóa hóa truyền thống thực hiện việc dò tìm chính xác vecto dịch chuyển giữa khung reference và khung inter, trong khi ở mã DVC chỉ là dựa trên ước lượng tap âm tương quan.

Ở phương pháp DVC (tiêu biểu là PRISM và DISCOVER), sự phức tạp được dịch chuyển về phía giải mã trong đó việc dò tìm các từ mã các hệ số DCT đúng trong tập phân hoạch.

Thực hiện so sánh chất lượng (theo PSNR) và độ phức tạp (theo thời gian) cho thấy:

DVC có kết quả đánh giá chất lượng tốt hơn khi các chuỗi video là các hình ảnh với vật thể chuyển động chậm hơn. Về thời gian mã hóa thì trong mọi trường hợp DVC tốt hơn hẳn so với mô hình truyền thống.

Nhận thấy tiềm năng lớn của hướng nghiên cứu này, trong tương lai một hệ thống DVC toàn diện, tương thích ngược với chuẩn H.265/HEVC chính là một trong những hướng nghiên cứu tiềm năng.

## LỜI CẢM ƠN

Được sự phân công của Khoa Điện tử viễn thông, trường Đại học Công nghệ, dưới sự hướng dẫn của PGS.TS Trịnh Anh Vũ và TS Hoàng Văn Xiêm tôi đã thực hiện luận văn: “ Nghiên cứu và phân tích bộ mã hóa video dùng cho mạng sensor ”.

Để hoàn thành luận văn này tôi xin cảm ơn các Thầy cô trong Khoa Điện tử viễn thông, trường Đại học Công nghệ đã nhiệt tình dạy dỗ, cung cấp phương pháp, kiến thức trong suốt thời gian học tập. Tôi xin chân thành cảm ơn tới hai thầy giáo PGS.TS Trịnh Anh Vũ và TS Hoàng Văn Xiêm đã tận tình hướng dẫn và chỉ bảo tôi rất nhiều cả về kiến thức cũng như phương pháp làm việc trong quá trình thực hiện và hoàn thành đề án này. Tôi cũng xin gửi lời cảm ơn tới nhóm nghiên cứu “Visual Processing and Communication” trường Đại học Công nghệ đã hỗ trợ trong việc tìm hiểu và phân tích bộ mã hóa DVC.

Mặc dù đã cố gắng và tích cực trong việc thực hiện luận văn nhưng không thể tránh được những sai sót nhất định, tôi rất mong nhận được sự đóng góp của các Thầy cô và các bạn đồng nghiệp để hoàn chỉnh luận văn và có hướng phát triển hơn nữa.

Tôi xin chân thành cảm ơn!

## TÀI LIỆU THAM KHẢO

- [1] D. Le. Gall, "MPEG: a video compression standard for multimedia applications", *Communications of the ACM*, Vol. 34, No. 4, pp. 46-56, Apr. 1991.
- [2] B. G. Haskell, A. Puri, and A. N. Netravali, "Digital video: an introduction to MPEG-2", *Spingher US Publisher*, 2002.
- [3] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC Video Coding Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 560-576, Jul. 2003.
- [4] G. J. Sullivan, J. R. Ohm, W. J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 12, pp. 1649-1668, Dec. 2012.
- [5] Iain E. G. Richardson, "H.264 And MPEG-4 Video Compression", *The Robert Gordon University, Aberdeen, UK*.
- [6] F. MacWilliams and N. Sloane, "The Theory of Error Correcting Codes" The Netherlands: Elsevier, Amsterdam, Netherlands, 1977.
- [7] D. Slepian and J. K. Wolf, "Noiseless coding of correlated information sources," *IEEE Transactions on Information Technology*, vol. 19, pp. 471-480, Jul. 1973.
- [8] A. D. Wyner and J. Ziv, "The rate-distortion function for source coding with side information at the decoder," *IEEE Transactions on Information Technology*, vol. 22, no. 1, pp. 1-10, Jan. 1976.
- [9] B. Girod, et al., "Distributed video coding," *Proceeding of the IEEE*, vol. 93, no. 1, pp. 71-83, Jan. 2005.
- [10] R. Puri and K. Ramchandran, "PRISM: A new robust video coding architecture based on distributed compression principles," in *40th Allerton Conf. Communication, Control and Computing*, Allerton, IL, USA, 2002.
- [11] X. Artiga, J. Ascenso, M. Dalai, S. Klomp, D. Kubasov, anh M. Ouaret, "The DISCOVER codec: architecture, techniques and evaluation" in *Picture Coding Symposium*, Lisbon, Portugal, December 2007.
- [12] L. Nataro, C. Brites, J. Ascenso, F. Pereira, "Side information extrapolation for low-delay pixel-domain distributed video coding",

in *International Workshop on Very Low Bitrate Video*, Sardinia, Italy, September 2005

- [13] J. Ascenso, C. Brites, and F. Pereira, "Improving Frame Interpolation with Spatial Motion Smoothing for Pixel Domain Distributed Video Coding," *5th EURASIP Conf. on Speech, Image Processing, Multimedia Communications and Services*, Jul. 2005.
- [14] C. Brites, J. Ascenso and F. Pereira, "Studying temporal correlation noise modeling for pixel based Wyner-Ziv video coding" in *IEEE International Conference on Image Processing*, Atlanta GA, USA, October 2006.
- [15] J. Ascenso, C. Brites and F. Pereira, "Content adaptive Wyner-Ziv video coding driven by motion activity" in *IEEE International Conference on Image Processing*, Atlanta, USA, October 2006
- [16] D. Kubasov et al., "Optimal reconstruction in Wyner-Ziv video coding with multiple side information," *IEEE Workshop on Multimedia Signal Processing*, pp. 183-186, Crete, Greece, Oct. 2007
- [17] R. Puri, A. Majumdar and K. Ramchandran, "PRISM: A Video Coding Paradigm With Motion Estimation at the Decoder", *IEEE Transactions on Image Processing*, vol. 16, no. 10, pp. 2436-2448, October 2007.
- [18] M. Contente, "Low complexity video coding for sensor networks", *Master Thesis, IST-Lisbon University*, Oct. 2009
- [19] <http://iphome.hhi.de/suehring/tml/>
- [20] X. HoangVan and B. Jeon, "Flexible complexity control solution for transform domain Wyner-Ziv video", *IEEE Transactions on Broadcasting*, Vol. 58, No 2, pp. 209-220, Jun. 2012.