MỤC LỤC

Chương 8. Video số 10
8.1 Vị trí của chương trong cuốn sách 12
8.2 Ảnh là một hàm nhiều biến 15
8.3 Lấy mẫu không gian và độ phân giải 16
8.4 Bao nhiêu điểm ảnh là đủ 20
8.5 Lượng tử hoá và độ sâu bit 21
8.6 Ảnh trên phim nhựa — một đối chiếu 25
8.7 Bài thực hành số 1 — lấy mẫu, độ phân giải và độ sâu bit 26
8.8 Từ RGB sang Y′CbCr trong thực tế 27
8.9 Lấy mẫu con màu 29
8.10 Quét xen kẽ và quét liên tục 36
8.11 Tần số hình 40
8.12 Đồng bộ, khoảng xoá và thời mã 43
8.13 Di sản analog 45
8.14 Vì sao phải nén 48
8.15 Nén không tổn hao và nén có tổn hao 53
8.16 Nén trong khung 56
8.17 Nén liên khung 62
8.18 Điều khiển tốc độ bit 72
8.19 Bài thực hành số 2 — nhìn thấy nén 77
8.20 Bản đồ các họ chuẩn 78
8.21 H.264 / AVC 79
8.22 H.265 / HEVC 82
8.23 AV1 và H.266 / VVC 84
8.24 Codec để dựng và codec để phát hành 85
8.25 Định dạng chứa và cấu trúc tệp 87
8.26 Bài thực hành số 3 — codec, bộ chứa và chất lượng 91
8.27 Giao diện truyền dẫn cơ sở 92
8.28 Video trên hạ tầng IP 97
8.29 Phát sóng số mặt đất ở Việt Nam 99
8.30 Phát trực tuyến và OTT 103
8.31 Chuẩn giao nộp và kiểm tra chất lượng 105
8.32 Quy trình sản xuất trên tệp 108
8.33 Câu hỏi và bài tập ôn tập 112

DANH MỤC HÌNH VẼ

Hình 8.1 Vị trí của Chương 8 trong chuỗi kỹ thuật, và năm phần của chương.....12
Hình 8.2 Ba phép rời rạc hoá của việc số hoá hình ảnh, minh hoạ trên cùng một cảnh.16
Hình 8.3 Cùng một cảnh, lấy mẫu với mật độ giảm dần.....17
Hình 8.4 Hiện tượng răng cưa và vân moiré, và tác dụng của bộ lọc thông thấp.....17
Hình 8.5 Sáu họ định dạng ảnh, vẽ đúng tỉ lệ, kèm số điểm ảnh và tỉ số so với SD.....18
Hình 8.6 Vì sao 720 × 576 mẫu lại hiển thị đúng khuôn hình 4:3.....19
Hình 8.7 Khoảng cách ngồi xem theo độ phân giải, và góc nhìn tương ứng.....20
Hình 8.8 Sai số lượng tử trên một dòng ảnh, và quan hệ giữa số bit với tỉ số tín hiệu trên nhiễu.....22
Hình 8.9 Hiện tượng dải màu và tác dụng của dither.....23
Hình 8.10 Dải hẹp và dải rộng, và hậu quả khi gán sai.....24
Hình 8.11 Hạt phim so với điểm ảnh, và đường đặc tính của hai loại vật mang.....25
Hình 8.12 Một cảnh tách thành ba thành phần Y′, Cb và Cr.....27
Hình 8.13 Hai bộ hệ số độ chói, và hậu quả khi giải mã bằng nhầm ma trận.....28
Hình 8.14 Thí nghiệm hai chiều ngược nhau: làm nhoè màu mà giữ độ chói (b) hầu như không thấy, làm nhoè độ chói mà giữ màu (c) thấy ngay.....29
Hình 8.15 Cách đọc ký hiệu J:a:b trên khối tham chiếu 4 × 2 điểm ảnh: chấm Y′ là mẫu độ chói, hình thoi là cặp mẫu màu (Cb, Cr), vùng tô là các điểm ảnh dùng chung một mẫu màu.....31
Hình 8.16 Kích thước các mặt phẳng Y′, Cb và Cr của khung 1920 × 1080 ở bốn cấu trúc, vẽ đúng tỉ lệ, kèm số byte mỗi khung ở 8 bit.....32
Hình 8.17 Cạnh màu trên một dòng điểm ảnh sau khi lọc xuống và nội suy lên: mẫu màu càng thưa thì vùng nhoè càng rộng.....34
Hình 8.18 Cùng một mẫu thử màu qua bốn cấu trúc lấy mẫu con, phóng to từng điểm ảnh: vùng xám không đổi, hỏng ở vạch mảnh, đường xiên và ô cờ.....35
Hình 8.19 Hai quy ước đặt vị trí mẫu màu trong cấu trúc 4:2:0.....35
Hình 8.20 Nguyên lý quét xen kẽ, và so sánh với quét liên tục ở cùng số lần quét mỗi giây.....37
Hình 8.21 Cách đọc một ký hiệu định dạng, và các ký hiệu thường gặp.....38
Hình 8.22 Sequence Settings (trái) và Interpret Footage (phải) trong Adobe Premiere Pro: khai báo thứ tự trường và chế độ quét (tệp mẫu tổng hợp, quét liên tục) 38
Hình 8.23 Răng lược ở vật chuyển động và kết quả sau khi khử xen kẽ.....39
Hình 8.24 Bốn tần số hình thường gặp và lý do lịch sử của từng con số.....40
Hình 8.25 Pulldown 2:3 cho hệ 60 trường, và ba cách đưa 24 lên 25 hình/giây.....42
Hình 8.26 Cấu trúc một khung hình HD 1080: vùng tích cực và khoảng xoá.....43
Hình 8.27 Cấu trúc thời mã SMPTE, và phân biệt bỏ khung với không bỏ khung.....44
Hình 8.28 Trống đầu từ quay của máy ghi băng từ video.....45
Hình 8.29 Phổ tín hiệu tổng hợp, và phân biệt ba bộ ký hiệu hiệu số màu.....46
Hình 8.30 Tốc độ bit thô và dung lượng một giờ của bốn định dạng.....48
Hình 8.31 Bốn loại dư thừa mà kỹ thuật nén khai thác.....49
Hình 8.32 Vì sao nén được: biểu đồ giá trị điểm ảnh trải rộng, còn biểu đồ sai số sau dự đoán co về một đỉnh nhọn quanh 0, entropy giảm rõ rệt.....51
Hình 8.33 Đường tốc độ–méo đo trên một bộ mã hoá ảnh thu nhỏ: công cụ tốt hơn dịch đường sang trái, và bộ mã hoá chọn điểm làm J = D + λR nhỏ nhất.....52
Hình 8.34 Đường chất lượng – tốc độ bit của bốn thế hệ codec, và ba chỉ số đo chất lượng.....55
Hình 8.35 64 hàm cơ sở của biến đổi cosin rời rạc 8 × 8.....57

Hình 8.36 Bốn bước của nén trong khung trên một khối 8 × 8 thật....58
Hình 8.37 Bảng lượng tử hoá độ chói của JPEG, và cùng một khối ở hai mức chất lượng....59
Hình 8.38 Đường quét zig-zag và dãy hệ số thật của khối ở hình trước....60
Hình 8.39 Hai hình dạng đặc trưng của lỗi nén trong khung....60
Hình 8.40 Dự đoán trong khung trên khối 4 × 4: ba hướng dự đoán, phần dư của từng hướng, và hướng có tổng sai số nhỏ nhất được chọn....62
Hình 8.41 Hai khung liên tiếp của một cảnh mô phỏng: ảnh hiệu gần như đen và biểu đồ |hiệu| dồn về 0....63
Hình 8.42 Lấy hiệu thẳng giữa hai khung: tốt với cảnh tĩnh, hỏng khi có chuyển động; bù chuyển động theo khối đưa phần dư về mức nhiễu....64
Hình 8.43 Dò chuyển động theo khối: cửa sổ tìm kiếm, bản đồ SAD với điểm cực tiểu, khối dự đoán và khối phần dư....65
Hình 8.44 Phần dư đi qua cùng đường ống DCT, lượng tử hoá và mã hoá entropy như ảnh tĩnh: nén liên khung cần ít hệ số và ít bit hơn nén trong khung....66
Hình 8.45 Vòng phản hồi của bộ mã hoá lai (a) và hậu quả khi dự đoán từ khung gốc thay vì khung tái tạo (b)....67
Hình 8.46 Khung B tham chiếu cả quá khứ lẫn tương lai: thứ tự hiển thị khác thứ tự mã hoá, và vùng nền vừa lộ ra chỉ có ở khung tương lai....68
Hình 8.47 Cấu trúc một nhóm ảnh 12 khung với quan hệ tham chiếu....69
Hình 8.48 Kích thước từng khung trong một nhóm ảnh, và ba cách nhìn cùng một nhóm....70
Hình 8.49 Số bit giảm dần qua từng tầng kỹ thuật, đo trên một đoạn phim thật ở cùng PSNR (trục lôgarit)....71
Hình 8.50 Điều khiển theo chất lượng không đổi (RF) và các tuỳ chọn bộ mã hoá x264 trong HandBrake....72
Hình 8.51 Hai chế độ điều khiển tốc độ bit và mức bộ đệm tương ứng....73
Hình 8.52 Export Settings của Adobe Media Encoder: định dạng HEVC (H.265), điều khiển tốc độ bit CBR, tốc độ bit đích 5,81 Mbit/s (nguồn là tệp thanh màu mẫu tổng hợp)....74
Hình 8.53 Thiết lập mã hoá H.264 VBR hai lượt (mục tiêu 20 Mbit/s, tối đa 24 Mbit/s) trong màn hình Export của Adobe Premiere Pro....75
Hình 8.54 Các tổ chức làm ra chuẩn nén và bảng đối chiếu tên gọi....78
Hình 8.55 Dòng thời gian các chuẩn nén video từ 1993 tới nay....79
Hình 8.56 Dự đoán trong khung: số hướng của H.264 so với HEVC....80
Hình 8.57 Profile quy định dùng công cụ nào, Level quy định được đến cỡ nào....81
Hình 8.58 Ô cố định của H.264 so với cây tứ phân của HEVC....83
Hình 8.59 Bảng so sánh bốn thế hệ codec theo bảy tiêu chí....85
Hình 8.60 Vị trí của các codec trên hai trục: tốc độ bit và công giải mã....86
Hình 8.61 Thông tin codec của từng luồng trong một tệp MP4 hiển thị bởi VLC....88
Hình 8.62 Một tệp video gồm những gì, và các bộ chứa thường gặp....88
Hình 8.63 Thông số bộ chứa và codec của một tệp MP4 (H.264, 1920 × 1080, 4:2:0, 8 bit) trong MediaInfo....90
Hình 8.64 Ba mức đóng gói: luồng cơ sở, gói hoá và luồng truyền tải....91
Hình 8.65 Họ giao diện SDI theo tốc độ đường truyền, và so sánh với HDMI....93
Hình 8.66 Đầu nối BNC đực trên cáp đồng trục — giao diện vật lý của SDI....94
Hình 8.67 Đầu nối HDMI loại A đực....95
Hình 8.68 Trường quay theo lối SDI so với trường quay trên hạ tầng IP....97
Hình 8.69 Từ ST 2022-6 tới ST 2110, và vai trò của đồng bộ thời gian PTP....97
Hình 8.70 Dây chuyền phát sóng số mặt đất ở Việt Nam, và ba con số cần nhớ....99

Hình 8.71 Ăng-ten Yagi–Uda thu truyền hình mặt đất.....99
Hình 8.72 Chòm sao điều chế và bốn tham số chính của DVB-T2.....100
Hình 8.73 Mặt sau đầu thu truyền hình số mặt đất DVB-T2: cổng ăng-ten vào, HDMI và S/PDIF ra.....101
Hình 8.74 Thang tốc độ bit của một dịch vụ phát trực tuyến, và cách đầu phát chuyển bậc.....103
Hình 8.75 Danh mục kiểm trước khi giao nộp một bản hoàn chỉnh.....105
Hình 8.76 Quy trình sản xuất trên tệp từ thẻ nhớ tới lưu trữ, với hai điểm kiểm bắt buộc 108

DANH MỤC BẢNG BIỂU

Bảng 8.1 Năm phần của Chương 8 11
Bảng 8.2 Cùng một khái niệm, hai miền tín hiệu 13
Bảng 8.3 Bộ thuật ngữ dùng trong Chương 8 13
Bảng 8.4 Các định dạng ảnh thường gặp 19
Bảng 8.5 Tỉ lệ điểm ảnh của một số định dạng 19
Bảng 8.6 Khoảng cách ngồi xem và góc nhìn ngang theo độ phân giải (khuôn hình 16:9, tính từ (8.7)) 21
Bảng 8.7 Độ sâu bit và số mức 22
Bảng 8.8 Phạm vi mức theo độ sâu bit 24
Bảng 8.9 Phim nhựa và cảm biến số — bốn khác biệt cốt lõi 25
Bảng 8.10 Bốn phần của bài thực hành số 1 26
Bảng 8.11 Hai bộ hệ số độ chói thường gặp 28
Bảng 8.12 Bốn cấu trúc lấy mẫu con màu 31
Bảng 8.13 Kích thước mặt phẳng và dung lượng một khung 1920 × 1080 (8 bit mỗi mẫu) 33
Bảng 8.14 Một dòng 8 điểm ảnh có cạnh màu qua 4:2:2 và 4:1:1 (nhân bản khi nội suy) 33
Bảng 8.15 Hai quy ước vị trí mẫu màu 35
Bảng 8.16 Chọn cấu trúc lấy mẫu con theo công việc 36
Bảng 8.17 Bốn chỗ cần đính chính về quét xen kẽ 37
Bảng 8.18 Thứ tự trường của các định dạng thường gặp 39
Bảng 8.19 Ba mức kỹ thuật khử xen kẽ 40
Bảng 8.20 Nguồn gốc các tần số hình 41
Bảng 8.21 Ba cách đưa phim 24 hình/giây lên 25 hình/giây 42
Bảng 8.22 Số dòng và số mẫu của các định dạng chính 43
Bảng 8.23 Các dạng thời mã 44
Bảng 8.24 Ba hệ truyền hình màu tương tự 46
Bảng 8.25 Phân biệt ba bộ ký hiệu hiệu số màu 47
Bảng 8.26 Ba mức chất lượng của giao diện tương tự 47
Bảng 8.27 Tốc độ bit thô, tính cho 10 bit và cấu trúc 4:2:2 48
Bảng 8.28 Bốn loại dư thừa và cách khai thác 50
Bảng 8.29 Cùng bộ khung bốn việc, ba thế hệ làm tinh hơn dần 53
Bảng 8.30 Ba chỉ số đo chất lượng bằng máy 54
Bảng 8.31 Ba trường hợp dùng nén không tổn hao 56
Bảng 8.32 Ba lớp mã hoá entropy 60
Bảng 8.33 Ba lỗi nén trong khung và cách nhận ra 61
Bảng 8.34 SAD của ba vị trí ứng viên trong ví dụ khối 4 × 4 64
Bảng 8.35 Ba loại khung 68
Bảng 8.36 Chọn độ dài nhóm ảnh theo mục đích 69
Bảng 8.37 Ba mức tinh vi của bù chuyển động 70
Bảng 8.38 Đánh đổi giữa hiệu quả nén và độ trễ 71
Bảng 8.39 Ba chế độ điều khiển tốc độ bit 72
Bảng 8.40 Tốc độ bit khởi điểm theo định dạng và codec 76
Bảng 8.41 Bốn bước kiểm tốc độ bit đã chọn 76
Bảng 8.42 Bốn phần của bài thực hành số 2 77
Bảng 8.43 Các thế hệ chuẩn nén và bối cảnh ra đời 79
Bảng 8.44 Sáu công cụ làm nên hiệu quả của H.264 80

Bảng 8.45 Các profile chính của H.264 82
Bảng 8.46 Bốn công cụ đặc trưng khác của HEVC 83
Bảng 8.47 Khuyến nghị chọn codec cho hoàn cảnh Việt Nam hiện nay 85
Bảng 8.48 Hai lớp codec 86
Bảng 8.49 Ba họ codec dựng và các bậc của chúng 86
Bảng 8.50 Quy tắc chọn codec cho từng khâu 87
Bảng 8.51 Bốn bộ chứa thường gặp 89
Bảng 8.52 Ba mức đóng gói 91
Bảng 8.53 Bốn phần của bài thực hành số 3 92
Bảng 8.54 Họ giao diện SDI 93
Bảng 8.55 SDI so với HDMI 95
Bảng 8.56 Hai cách đưa video lên hạ tầng IP 98
Bảng 8.57 Các phần chính của bộ ST 2110 98
Bảng 8.58 Ba mức truyền video trên mạng 98
Bảng 8.59 Bốn tham số chính của DVB-T2 theo ETSI EN 302 755 101
Bảng 8.60 Các yêu cầu chính của QCVN 63:2020/BTTTT 102
Bảng 8.61 Hai chuẩn truyền theo HTTP thích ứng 103
Bảng 8.62 Một thang tốc độ bit tham khảo cho nội dung Full HD 104
Bảng 8.63 Độ trễ của các cách phát 104
Bảng 8.64 Một hồ sơ giao nộp điển hình cho kênh HD của đài Việt Nam 105
Bảng 8.65 Bốn lỗi thường gặp và cách phòng 106
Bảng 8.66 Hai kiểu hồ sơ giao nộp 107
Bảng 8.67 Ba vấn đề của lưu trữ dài hạn 109
Bảng 8.68 Bốn phần của bài thực hành số 4 110
Bảng 8.69 Mười bài tập tính toán 115

DANH MỤC KÝ HIỆU VÀ CHỮ VIẾT TẮT

Ký hiệu viết tắt	Thuật ngữ tiếng Anh đầy đủ	Ý nghĩa chuyên môn trong kỹ thuật Audio - Video
ADSR	Attack - Decay - Sustain - Release	Mô hình đường bao biên độ 4 giai đoạn theo thời gian trong thiết kế tiếng động
AES	Audio Engineering Society	Hội Kỹ sư Âm thanh Quốc tế — tổ chức ban hành các tiêu chuẩn âm thanh chuyên nghiệp
DAW	Digital Audio Workstation	Trạm làm việc âm thanh số (Pro Tools, DaVinci Resolve Fairlight, Nuendo, Reaper)
dB	Decibel	Đơn vị logarit biểu diễn tỷ số giữa hai đại lượng biên độ hoặc công suất
dB SPL	Decibel Sound Pressure Level	Mức áp suất âm thanh trong không khí quy chiếu theo ngưỡng nghe 20 µPa
dBFS	Decibels Relative to Full Scale	Thang đo decibel số toàn dải; 0 dBFS là trần tuyệt đối của hệ thống kỹ thuật số
dBTP	Decibels True Peak	Mức đỉnh thực của tín hiệu số có tính toán hiện tượng liên mẫu (inter-sample peaks)
dBA / dBC	A-weighted / C-weighted Decibels	Mức đo decibel qua mạch lọc bù trừ đặc tuyến độ nhạy của tai người
EBU	European Broadcasting Union	Hiệp hội Phát thanh Truyền hình Châu Âu (tác giả tiêu chuẩn EBU R128)
HRTF	Head-Related Transfer Function	Hàm truyền liên quan đến đầu người — cơ sở định vị không gian 3 chiều và âm thanh nhị nhĩ
ILD	Interaural Level Difference	Độ chênh lệch mức áp suất âm thanh giữa hai tai (chủ đạo ở tần số cao > 1,5 kHz)
ITD	Interaural Time Difference	Độ chênh lệch thời gian âm thanh truyền đến hai tai (chủ đạo ở tần số thấp < 1,5 kHz)
ITU-R	ITU Radiocommunication Sector	Phân ban Vô tuyến của Liên minh Viễn thông Quốc tế (tác giả chuẩn BS.1770)
LKFS / LUFS	Loudness K-weighted / Units Full Scale	Đơn vị đo lường độ to cảm nhận chuẩn hóa theo ITU-R BS.1770
LRA	Loudness Range	Dải độ to — chỉ số mô tả độ động giữa đoạn to và đoạn nhỏ của chương trình (đơn vị LU)
NC / NR	Noise Criteria / Noise Rating	Đường cong tiêu chí đánh giá mức độ ồn nền của phòng thu

Ký hiệu viết tắt	Thuật ngữ tiếng Anh đầy đủ	Ý nghĩa chuyên môn trong kỹ thuật Audio - Video
NIOSH	National Institute for Occupational Safety and Health	Viện Quốc gia về An toàn và Sức khỏe Nghề nghiệp Hoa Kỳ
RMS	Root Mean Square	Giá trị căn bậc hai trung bình bình phương (năng lượng hiệu dụng của tín hiệu)
RT60	Reverberation Time 60 dB	Thời gian để năng lượng âm vang suy giảm 60 dB sau khi ngắt nguồn âm
SMPTE	Society of Motion Picture and Television Engineers	Hiệp hội Kỹ sư Điện ảnh và Truyền hình Quốc tế
SNR	Signal-to-Noise Ratio	Tỷ số tín hiệu trên tạp âm (đo bằng dB)

CHƯƠNG 8
VIDEO SỐ

Cấu trúc tín hiệu · Nén · Phân phối

Bảy chương trước của cuốn sách bao quát hai chặng. Sáu chương đầu trình bày toàn bộ miền âm thanh, từ sóng áp suất trong không khí đến các thiết bị xử lý nằm giữa hai bộ chuyển đổi, và Chương 7 mở đầu miền hình ảnh, đi từ ánh sáng thật đến con số qua vật lý ánh sáng, sinh lý thị giác, hệ toạ độ màu CIE, rồi gamma, đường log và HDR.

Chương 8 tiếp nối quãng đường còn lại, bắt đầu từ thời điểm mỗi điểm ảnh đã trở thành ba con số và kết thúc khi khán giả nhìn thấy hình trên màn hình. Giữa hai điểm ấy là năm khối kiến thức: cách số hoá ảnh, cách sắp xếp tín hiệu, nguyên lý nén, các chuẩn nén cùng định dạng tệp, và cuối cùng là truyền dẫn, phát sóng cùng quy trình sản xuất trên tệp.

Mục tiêu của chương

Sau khi đọc xong chương này, người đọc có thể:

1. Giải thích ba phép rời rạc hoá trong số hoá hình ảnh (không gian, thời gian, biên độ) và đọc đúng các thông số của một tệp video, gồm độ phân giải, tần số hình và độ sâu bit.
2. Áp dụng định lý lấy mẫu cho hai chiều không gian, nhận biết hiện tượng răng cưa và vân moiré, đồng thời nêu được cách xử lý.
3. Giải thích tỉ lệ điểm ảnh và tính được khoảng cách xem tối ưu theo độ phân giải.
4. Trình bày độ sâu bit, hiện tượng dải màu và dither, đồng thời phân biệt dải hẹp 16 – 235 với dải rộng 0 – 255 cùng hậu quả khi gán sai.
5. Giải thích lý do tách tín hiệu thành Y′CbCr và tính được phép chuyển đổi giữa RGB và Y′CbCr bằng hai bộ hệ số BT.601 và BT.709.
6. Đọc và giải thích ký hiệu lấy mẫu con màu J:a:b (4:4:4, 4:2:2, 4:2:0, 4:1:1) trên khối tham chiếu, tính được số mẫu mỗi điểm ảnh và dung lượng khung hình, và chọn được cấu trúc phù hợp với từng công việc.
7. Phân biệt trường với khung, quét xen kẽ với quét liên tục, đọc được ký hiệu định dạng, và giải thích tần số hình cùng pulldown.
8. Trình bày cấu trúc khoảng xoá và thời mã, kể cả thời mã bỏ khung.
9. Giải thích từ nguyên lý (dự đoán và mã hoá phần bất ngờ) lý do nén được, bốn loại dư thừa, và lý do codec thế hệ mới nén kỹ hơn nhưng đòi hỏi nhiều tính toán hơn.

10. Mô tả chuỗi nén trong khung (dự đoán, biến đổi cosin, lượng tử hoá, quét zig-zag, mã hoá entropy) trên một khối thật và nhận biết các lỗi nén bằng mắt.
11. Trình bày từng bước chuyển từ nén ảnh sang nén video: dò chuyển động theo khối, mã hoá phần dư, vòng phản hồi, khung I, P, B và nhóm ảnh; đồng thời giải thích điều khiển tốc độ bit và bộ đệm.
12. So sánh H.264, HEVC, AV1 và VVC, phân biệt codec với bộ chứa, và phân biệt codec để dựng với codec để phát hành.
13. Mô tả giao diện SDI, video trên hạ tầng IP (ST 2110, NDI, SRT), phát sóng DVB-T2 ở Việt Nam theo QCVN 63:2020/BTTTT và phát trực tuyến thích ứng.
14. Lập và kiểm hồ sơ giao nộp, áp dụng quy tắc ba – hai – một và kiểm tổng kiểm trong quy trình sản xuất trên tệp.

Cấu trúc của chương

Bảng 8.1 Năm phần của Chương 8

Phần	Tên phần	Các mục	Nội dung chính
A	Từ ảnh tới con số	8.1 – 8.7	Ba phép rời rạc hoá; lấy mẫu không gian và răng cưa; tỉ lệ điểm ảnh; khoảng cách ngồi xem; lượng tử hoá, độ sâu bit và phạm vi mức; đối chiếu với phim nhựa
B	Cấu trúc tín hiệu video	8.8 – 8.13	Y′CbCr trong thực tế; lấy mẫu con màu; quét xen kẽ và khử xen kẽ; tần số hình và pulldown; khoảng xoá và thời mã; di sản analog
C	Nguyên lý nén	8.14 – 8.19	Bốn loại dư thừa; nén trong khung với DCT và lượng tử hoá; nén liên khung với GOP và bù chuyển động; điều khiển tốc độ bit
D	Chuẩn nén và định dạng tệp	8.20 – 8.26	Bản đồ các họ chuẩn; H.264, HEVC, AV1 và VVC; profile và level; codec dựng; bộ chứa và luồng truyền tải
E	Truyền dẫn, phát sóng và quy trình	8.27 – 8.32	SDI và HDMI; ST 2110, NDI, SRT; DVB-T2 và QCVN 63:2020/BTTTT; phát trực tuyến; chuẩn giao nộp; quy trình sản xuất trên tệp

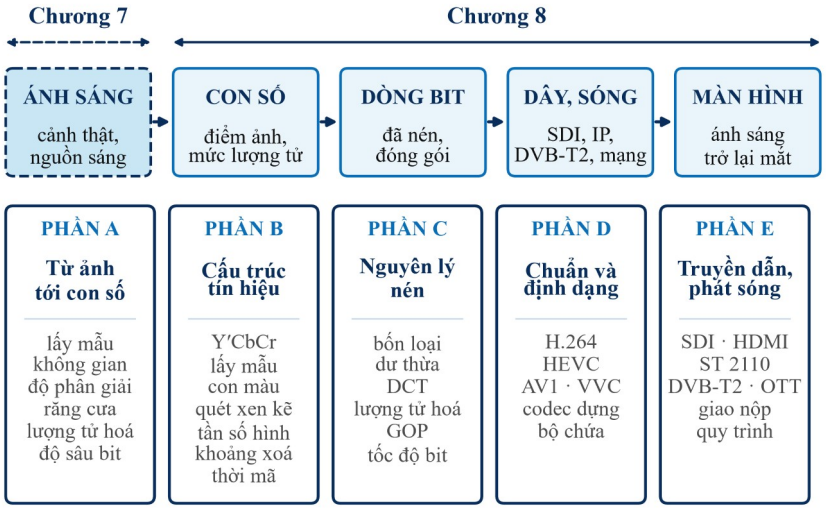
Chương gồm 32 mục, 65 hình vẽ, 34 công thức và bốn bài thực hành đặt ở cuối các phần có nội dung đo đạc hoặc thao tác máy.

8.1 Vị trí của chương trong cuốn sách

8.1.1 Chương 7 dừng ở đâu

Chương 7 đưa người đọc đi trọn quãng đường từ ánh sáng thật đến con số: bắt đầu từ bản chất vật lý của ánh sáng, qua sinh lý thị giác, đến việc dựng hệ toạ độ màu CIE, và kết thúc bằng lời giải thích vì sao tín hiệu hình được mã hoá phi tuyến và vì sao cần có gamma, đường log, HDR. Đến cuối Chương 7, mỗi điểm ảnh của cảnh quay đã trở thành ba con số.

Tuy nhiên, ba con số ấy vẫn còn nằm nguyên trong máy quay và chưa thể đến được với người xem. Từ máy quay đến lúc khán giả nhìn thấy hình trên màn hình còn cả một quãng đường dài, và toàn bộ quãng đường ấy là nội dung của Chương 8.



Hình 8.1 Vị trí của Chương 8 trong chuỗi kỹ thuật, và năm phần của chương

8.1.2 Bốn câu hỏi mà chương phải trả lời

Nội dung của chương có thể tóm tắt bằng bốn câu hỏi; người đọc trả lời được cả bốn câu bằng ngôn ngữ của mình thì có thể xem là đã nắm vững chương.

- **Một cảnh quay trở thành con số bằng cách nào, và cần bao nhiêu con số?** Câu hỏi này thuộc Phần A: lấy mẫu theo không gian, theo thời gian và theo biên độ.
- **Những con số ấy được sắp xếp ra sao trước khi nén?** Câu hỏi này thuộc Phần B: cấu trúc Y′CbCr, lấy mẫu con màu, quét xen kẽ, tần số hình và thời mã.
- **Vì sao nén được, và nén bằng cách nào?** Câu hỏi này thuộc Phần C và Phần D: bốn loại dư thừa, biến đổi cosin, bù chuyển động, rồi đến các chuẩn nén cụ thể và các định dạng tệp.
- **Tín hiệu đến với người xem bằng đường nào, và bản giao nộp phải đáp ứng những yêu cầu gì?** Câu hỏi này thuộc Phần E: SDI, hạ tầng IP, phát sóng DVB-T2 ở Việt Nam, phát trực tuyến, chuẩn giao nộp và quy trình sản xuất trên tệp.

8.1.3 Quan hệ với Chương 3 — cùng một định lý

Cần lưu ý ngay từ đầu rằng phần lớn cơ sở lý thuyết của chương này đã được trình bày ở Chương 3, chương về tín hiệu số của phần âm thanh: định lý lấy mẫu, hiện tượng răng cưa, sai số lượng tử và quan hệ giữa số bit với tỉ số tín hiệu trên nhiễu đều được áp dụng nguyên vẹn. Khác biệt duy nhất là ở hình ảnh, việc lấy mẫu diễn ra trên hai chiều không gian chứ không phải trên một trục thời gian.

Bảng 8.2 Cùng một khái niệm, hai miền tín hiệu

Khái niệm	Trong âm thanh (Chương 3)	Trong hình ảnh (chương này)
Lấy mẫu	Theo thời gian: 48 000 mẫu mỗi giây	Theo hai chiều không gian (điểm ảnh) và theo thời gian (khung hình)
Tần số Nyquist	Một nửa tần số lấy mẫu, tính bằng hertz	Một nửa tần số lấy mẫu không gian, tính bằng chu kỳ trên một điểm ảnh
Răng cưa	Tần số cao gấp ngược lại thành tần số thấp — nghe thành tiếng rít lạ	Chi tiết mịn gấp ngược lại thành hoa văn giả — nhìn thành vân moiré
Bộ lọc chống răng cưa	Bộ lọc thông thấp điện tử trước bộ chuyển đổi	Bộ lọc quang thông thấp đặt ngay trước cảm biến
Lượng tử hoá	16 hoặc 24 bit; sai số nghe thành nhiễu nền	8, 10 hoặc 12 bit; sai số NHÌN thành đường viền giả
Tỉ số tín hiệu trên nhiễu	Xấp xỉ 6,02n + 1,76 dB	Cùng công thức, cùng con số
Dither	Thêm nhiễu rất nhỏ để xoá méo lượng tử	Thêm nhiễu rất nhỏ để xoá đường viền giả

GHI NHỚ 8.1 — MỘT LỜI KHUYÊN HỌC TẬP

Mỗi khi gặp một khái niệm mới trong chương này, người đọc nên tự hỏi khái niệm ấy ứng với đại lượng nào bên âm thanh. Hầu như luôn có câu trả lời, và sự đối chiếu đó rút ngắn đáng kể thời gian tiếp thu.

Phép so sánh chỉ không còn áp dụng được ở một điểm: hình ảnh có thêm chiều màu, và màu được lấy mẫu thưa hơn độ chói. Đặc điểm này không có tương ứng bên âm thanh và là nội dung riêng của Phần B.

8.1.4 Thuật ngữ dùng trong chương

Thuật ngữ tiếng Việt trong kỹ thuật video chưa hoàn toàn thống nhất, trong khi tài liệu kỹ thuật hầu như luôn được viết bằng tiếng Anh. Vì vậy, bảng dưới đây ghi bộ thuật ngữ dùng xuyên suốt chương, kèm từ tiếng Anh tương ứng để người đọc tra cứu tài liệu gốc.

Bảng 8.3 Bộ thuật ngữ dùng trong Chương 8

Thuật ngữ dùng trong sách	Tiếng Anh	Ghi chú
Điểm ảnh	pixel	Một MẪU, không phải một ô vuông có màu

Thuật ngữ dùng trong sách	Tiếng Anh	Ghi chú
Độ phân giải	resolution	Số mẫu theo hai chiều; phân biệt với độ nét thật đo được
Tỉ lệ điểm ảnh	pixel aspect ratio (PAR)	Tỉ số rộng trên cao của MỘT mẫu khi hiển thị
Khuôn hình	display aspect ratio (DAR)	Tỉ số của cả khung hình
Độ sâu bit	bit depth	Số bit cho mỗi thành phần màu
Dải hẹp · dải rộng	narrow (video, studio) range · full range	16 – 235 và 0 – 255 ở tám bit
Dải màu (hiện tượng)	banding	Đường viền giả trong vùng chuyển sắc êm
Lấy mẫu con màu	chroma subsampling	Cấu trúc 4:2:2, 4:2:0…
Trường · mành · bán ảnh	field	Tập các dòng lẻ hoặc chẵn; hai trường thành một khung
Khung hình	frame	Một ảnh hoàn chỉnh
Quét xen kẽ · quét liên tục	interlaced · progressive	
Khử xen kẽ	deinterlacing	
Răng lược	combing	Lỗi hình khi ghép hai trường lệch thời gian
Thời mã	timecode	
Khoảng xoá	blanking interval	
Nhóm ảnh	group of pictures (GOP)	
Bù chuyển động	motion compensation	
Vỡ khối · nhiễu muỗi	blocking · mosquito noise	Hai lỗi đặc trưng của nén trong khung
Bộ chứa · định dạng chứa	container format	MP4, MOV, MXF, TS — khác với codec
Luồng truyền tải	transport stream (TS)	Gói đều 188 byte
Thang tốc độ bit	bitrate ladder	Bộ nhiều bậc chất lượng của một dịch vụ phát trực tuyến
Tổng kiểm	checksum	Dùng để kiểm bản sao đúng tới từng bit

Về cách viết số, theo thông lệ Việt Nam, dấu phẩy được dùng làm dấu thập phân và dấu cách được dùng để tách hàng nghìn. Do đó, trong sách này "1 037 Mbit/s" là một nghìn không trăm ba mươi bảy, còn "29,97 hình/giây" là hai mươi

chín phẩy chín bảy. Tài liệu tiếng Anh dùng quy ước ngược lại, nên cần chú ý khi tra cứu.

8.2 Ảnh là một hàm nhiều biến

8.2.1 Viết cảnh quay thành một hàm

Muốn số hoá một đối tượng, trước hết phải biểu diễn nó thành một hàm. Với cảnh quay, hàm ấy xác định một giá trị độ chói tại mỗi vị trí trong không gian, tại mỗi thời điểm và với mỗi bước sóng.

$$L = f(x,\ y,\ t,\ \lambda) \tag{8.1}$$

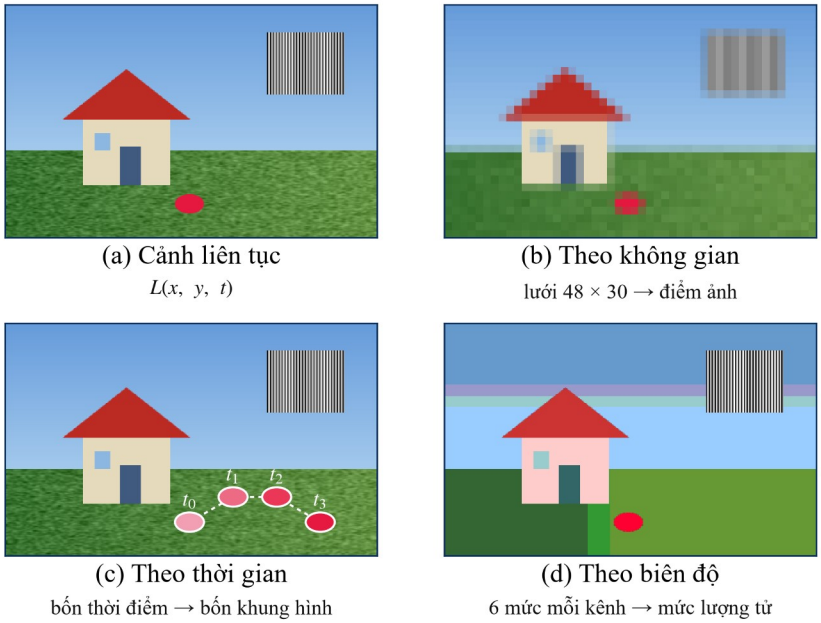
Bốn biến của hàm này tương ứng với bốn đại lượng mà máy quay phải xử lý: hai biến đầu là vị trí trên mặt phẳng ảnh, biến thứ ba là thời gian, và biến thứ tư là bước sóng, tức là màu. Phần còn lại của chương chỉ ra phép xử lý áp dụng cho từng biến.

8.2.2 Ba phép rời rạc hoá

Máy quay không lưu được một hàm liên tục mà phải biến hàm ấy thành một mảng số hữu hạn, thông qua ba phép rời rạc hoá được thực hiện đồng thời.

$$L_{\mathrm{d}}[i,\ j,\ k] = Q\{\ L(i\,\Delta x,\ \ j\,\Delta y,\ \ k\,\Delta t)\ \} \tag{8.2}$$

- **Rời rạc hoá theo không gian.** Mặt phẳng ảnh được chia thành lưới ô vuông, mỗi ô cho một giá trị duy nhất, và kết quả là *điểm ảnh*. Số điểm ảnh theo hai chiều chính là độ phân giải.
- **Rời rạc hoá theo thời gian.** Thời gian được chia thành các khoảng đều nhau, mỗi khoảng cho một ảnh, và kết quả là *khung hình*. Số khung mỗi giây chính là tần số hình.
- **Rời rạc hoá theo biên độ.** Mỗi giá trị được làm tròn về mức gần nhất trong một thang hữu hạn, và kết quả là *mức lượng tử*. Số mức được quyết định bởi độ sâu bit.



Hình 8.2 Ba phép rời rạc hoá của việc số hoá hình ảnh, minh hoạ trên cùng một cảnh

Biến thứ tư là bước sóng được xử lý theo cách khác hẳn: thay vì lấy mẫu dày đặc phổ, người ta lấy đúng ba mẫu ứng với ba loại tế bào nón của mắt người. Nội dung này đã được trình bày đầy đủ ở Chương 7 và không nhắc lại ở đây.

> **GHI NHỚ 8.2 — ĐIỀU CẦN NHỚ**
>
> Mọi dòng thông số kỹ thuật của một tệp video, chẳng hạn "1920 × 1080, 25 hình/giây, 10 bit, 4:2:2", đều là kết quả của ba phép rời rạc hoá trên, cộng với cách xử lý màu. Khi nắm được điều này, việc đọc thông số tệp không còn là học thuộc lòng.

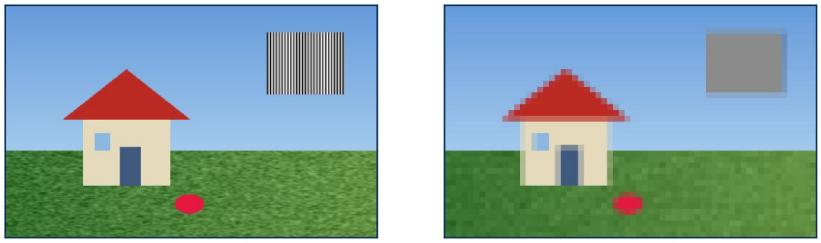
8.3 Lấy mẫu không gian và độ phân giải

8.3.1 Điểm ảnh và số mẫu

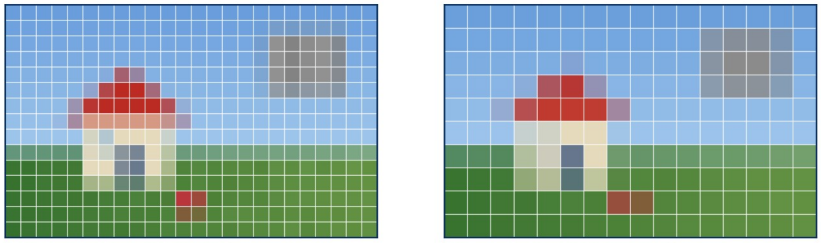
Điểm ảnh là mẫu nhỏ nhất của ảnh. Cần lưu ý rằng điểm ảnh không phải là một ô vuông nhỏ có màu, mà là một giá trị đo được lấy tại một vị trí; việc màn hình hiển thị giá trị ấy dưới dạng ô vuông chỉ là cách trình bày, không phải bản chất của điểm ảnh.

$$N_{\text{px}} = W \times H \quad ; \quad f_s = W \times H \times f_{\text{fr}} \tag{8.3}$$

Từ công thức trên suy ra một con số quan trọng là tốc độ lấy mẫu của tín hiệu hình. Với định dạng Full HD ở 25 hình mỗi giây, tốc độ này là 1920 × 1080 × 25 ≈ 51,8 triệu mẫu mỗi giây cho riêng thành phần độ chói, gấp hơn một nghìn lần tốc độ lấy mẫu của âm thanh chất lượng cao. Con số này giải thích vì sao hình ảnh cần băng thông lớn, và vì sao nén là yêu cầu bắt buộc chứ không phải tuỳ chọn.



(a) 384 × 240 điểm ảnh (b) 64 × 40 điểm ảnh



(c) 24 × 15 điểm ảnh (d) 16 × 10 điểm ảnh

Hình 8.3 Cùng một cảnh, lấy mẫu với mật độ giảm dần

Hình trên cũng cho thấy một quy luật đáng ghi nhớ: khi giảm số mẫu, chi tiết không mất đều. Vùng trời phẳng gần như không đổi, trong khi mảng hoa văn và cạnh xiên của mái nhà suy giảm trước tiên. Theo ngôn ngữ kỹ thuật, vùng nào có *tần số không gian* cao thì vùng đó chịu suy giảm trước.

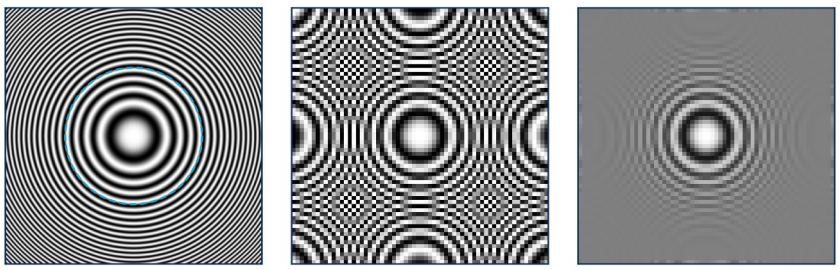
8.3.2 Tần số không gian và định lý lấy mẫu

Tần số không gian đo mức độ dày của hoa văn trong ảnh, tính bằng số chu kỳ trên một đơn vị chiều dài hoặc trên một điểm ảnh. Một bức tường trơn có tần số không gian gần bằng không, còn một tấm vải kẻ sọc mảnh có tần số không gian rất cao.

Định lý lấy mẫu được áp dụng nguyên vẹn, với điểm khác biệt là phải áp dụng cho cả hai chiều.

$$f_{s,x} \geq 2f_{\max,x} \quad ; \quad f_{s,y} \geq 2f_{\max,y} \tag{8.4}$$

Khi điều kiện này bị vi phạm, tức là ảnh có chi tiết mịn hơn mức mà lưới điểm ảnh phân biệt được, chi tiết ấy không biến mất mà bị *gấp ngược* thành hoa văn giả có tần số thấp. Trong ảnh, hiện tượng này được gọi là răng cưa nếu xảy ra trên một cạnh, và vân moiré nếu xảy ra trên một mảng hoa văn tuần hoàn.



(a) Ảnh gốc
tần số tăng dần ra phía rìa

(b) Lấy mẫu không lọc
vân moiré ở vùng rìa

(c) Lọc rồi lấy mẫu
mờ dần thành xám, vân giả yếu hẳn

Hình 8.4 Hiện tượng răng cưa và vân moiré, và tác dụng của bộ lọc thông thấp

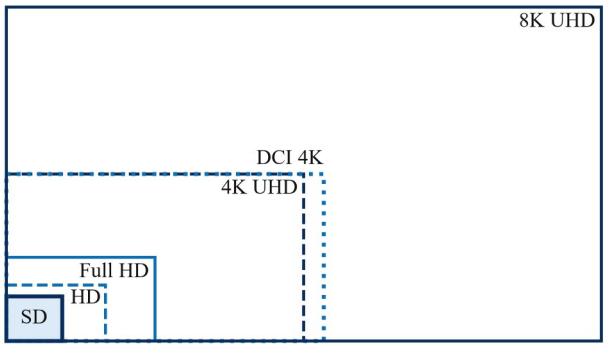
8.3.3 Bộ lọc quang thông thấp

Cách xử lý cũng tương tự như bên âm thanh: lọc bỏ những tần số vượt quá giới hạn trước khi lấy mẫu. Với hình ảnh, việc này được thực hiện bằng quang học thay vì điện tử: một tấm lọc quang thông thấp, thường làm bằng tinh thể lưỡng chiết, được đặt ngay trước cảm biến và có tác dụng làm nhoè ảnh ở mức cần thiết.

> **GHI NHỚ 8.3 — MỘT ĐÁNH ĐỔI CÓ THẬT TRONG NGHỀ**
>
> Bộ lọc quang thông thấp loại bỏ được vân moiré nhưng cũng làm ảnh kém sắc nét hơn. Từ khoảng năm 2012 trở đi, nhiều máy quay và máy ảnh cao cấp bỏ hẳn bộ lọc này để đạt độ nét tối đa, đồng thời chấp nhận rủi ro xuất hiện vân moiré.
>
> Trong thực tế quay phim, khi quay áo kẻ sọc mảnh, rèm cửa, mái ngói ở xa hoặc màn hình máy tính bằng máy không có bộ lọc, vân màu lạ rất dễ xuất hiện. Tại hiện trường, có thể xử lý bằng cách đổi góc máy, đổi khoảng cách, hoặc mở khẩu lớn hơn để độ sâu trường ảnh nông đi và chi tiết nền tự nhoè bớt.

8.3.4 Các họ định dạng ảnh

Trong sản xuất hiện nay có sáu họ định dạng thường gặp. Khi so sánh các họ này, yếu tố cần chú ý nhất là số điểm ảnh, đại lượng tăng theo bình phương, chứ không phải con số về chiều dài.



Định dạng	Số mẫu	Số điểm ảnh (Mpx)	Gấp SD
8K UHD	7680 × 4320	33,18	80,0 ×
DCI 4K	4096 × 2160	8,85	21,3 ×
4K UHD	3840 × 2160	8,29	20,0 ×
Full HD	1920 × 1080	2,07	5,0 ×
HD	1280 × 720	0,92	2,2 ×
SD PAL	720 × 576	0,41	1,0 ×

Hình 8.5 Sáu họ định dạng ảnh, vẽ đúng tỉ lệ, kèm số điểm ảnh và tỉ số so với SD

Bảng 8.4 Các định dạng ảnh thường gặp

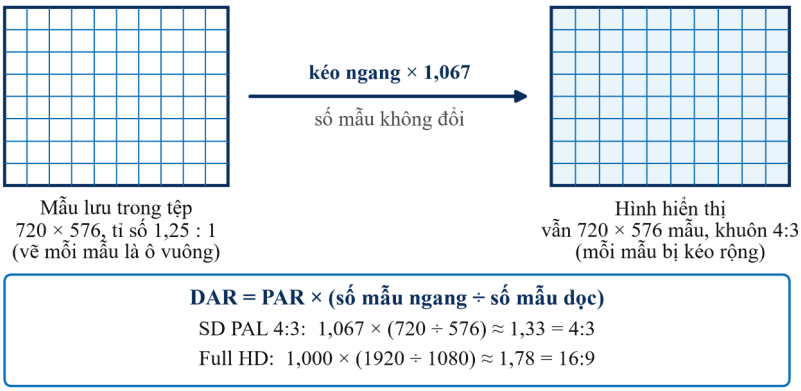
Tên gọi	Số mẫu	Khuôn hình	Số điểm ảnh	Ghi chú
SD PAL	720 × 576	4:3 hoặc 16:9	0,41 Mpx	Điểm ảnh KHÔNG vuông; di sản của hệ PAL 625 dòng
HD	1280 × 720	16:9	0,92 Mpx	Chỉ có bản quét liên tục; hay dùng cho phát sóng thể thao
Full HD	1920 × 1080	16:9	2,07 Mpx	Định dạng phát sóng phổ biến nhất ở Việt Nam hiện nay
4K UHD	3840 × 2160	16:9	8,29 Mpx	Đúng gấp đôi Full HD theo mỗi chiều, gấp bốn về số điểm ảnh
DCI 4K	4096 × 2160	256:135	8,85 Mpx	Chuẩn của điện ảnh số; rộng hơn 4K UHD một chút
8K UHD	7680 × 4320	16:9	33,18 Mpx	Gấp mười sáu lần Full HD; hiện chủ yếu dùng ở màn hình LED lớn

8.3.5 Tỉ lệ điểm ảnh — một chỗ hay bị bỏ sót

Một khái niệm thường gây nhiều khó khăn trong thực tế là tỉ lệ điểm ảnh. Câu hỏi cụ thể là vì sao định dạng SD PAL có 720 × 576 mẫu, tức tỉ số 1,25 : 1, nhưng lại hiển thị đúng khuôn hình 4:3, tức 1,33 : 1.

Nguyên nhân là điểm ảnh của SD không vuông: mỗi mẫu, khi hiển thị, được kéo rộng ra theo chiều ngang. Quan hệ giữa ba đại lượng được biểu diễn như sau.

$$\mathrm{DAR} = \mathrm{PAR} \times \frac{W}{H} \tag{8.5}$$



Hình 8.6 Vì sao 720 × 576 mẫu lại hiển thị đúng khuôn hình 4:3

Bảng 8.5 Tỉ lệ điểm ảnh của một số định dạng

Định dạng	Số mẫu	Tỉ lệ điểm ảnh	Khuôn hiển thị
SD PAL 4:3	720 × 576	≈ 1,067	4:3 (1,33)
SD PAL 16:9	720 × 576	≈ 1,422	16:9 (1,78)
HD, Full HD, UHD	1280 × 720 và lớn hơn	1,000 (vuông)	16:9 (1,78)
DCI 4K	4096 × 2160	1,000 (vuông)	1,90

CẢNH BÁO 8.1 — LỖI THƯỜNG GẶP

Khi mở một tệp SD bằng phần mềm không nhận biết tỉ lệ điểm ảnh, hình người sẽ bị co hẹp theo chiều ngang hoặc bị giãn rộng, tuỳ chiều khai báo sai.

Hiện tượng này là lỗi khai báo chứ không phải lỗi của tệp. Video HD và UHD dùng điểm ảnh vuông nên không gặp tình huống này, nhưng tư liệu SD cũ trong kho của các đài thì có, và người thực hiện phóng sự tư liệu sẽ thường xuyên gặp phải.

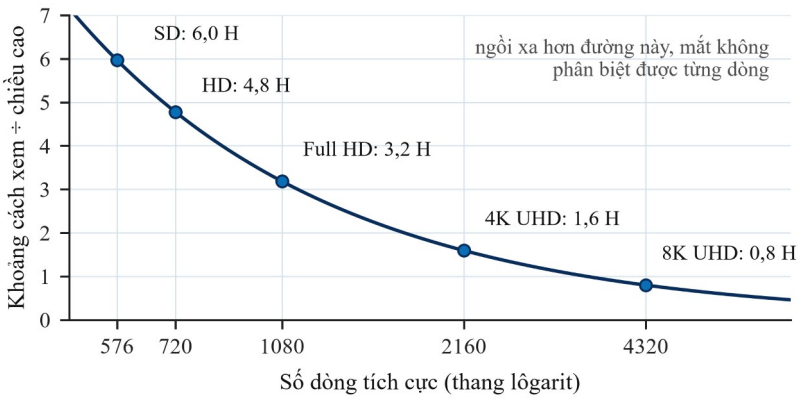
8.4 Bao nhiêu điểm ảnh là đủ

8.4.1 Từ góc phân giải của mắt tới số dòng quét

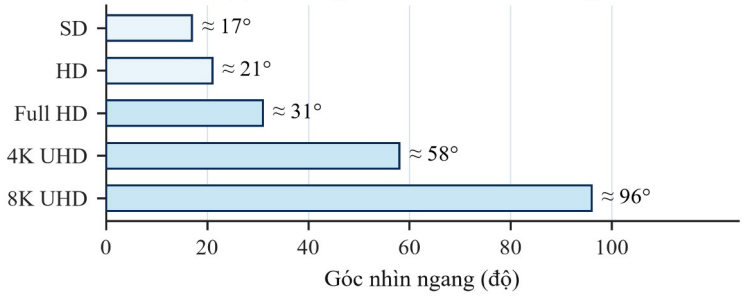
Số dòng quét cần thiết được quyết định bởi góc phân giải của mắt chứ không phải bởi ý muốn của nhà sản xuất. Lập luận này đúng, nhưng các con số quen thuộc về số dòng quét vốn được tính cho truyền hình độ nét tiêu chuẩn và nay đã lạc hậu [4].

Xuất phát điểm là một dữ kiện sinh lý đã học ở Chương 7: mắt người phân biệt được hai điểm cách nhau khoảng một phút cung, tức một phần sáu mươi của một độ. Từ đó suy ra khoảng cách xem mà tại đó mắt vừa đủ phân biệt được từng dòng của màn hình.

$$\frac{d}{H} = \frac{1}{N_{\text{row}} \cdot \theta} \quad , \quad \theta = 1' = 2{,}909 \times 10^{-4} \text{ rad} \tag{8.6}$$



(a) Khoảng cách xem theo số dòng



(b) Góc nhìn ngang tương ứng

Hình 8.7 Khoảng cách ngồi xem theo độ phân giải, và góc nhìn tương ứng

8.4.2 Đọc kết quả

Bảng 8.6 Khoảng cách ngồi xem và góc nhìn ngang theo độ phân giải (khuôn hình 16:9, tính từ (8.7))

Định dạng	Số dòng tích cực	Khoảng cách ÷ chiều cao màn hình	Góc nhìn ngang	Khuyến nghị
SD	576	6,0 lần	≈ 17°	Khuyến nghị cũ cho truyền hình độ nét tiêu chuẩn
HD	720	4,8 lần	≈ 21°	
Full HD	1080	3,2 lần	≈ 31°	Khuyến nghị của SMPTE cho phòng xem tại nhà
4K UHD	2160	1,6 lần	≈ 58°	Khuyến nghị của ITU-R BT.2100 cho UHDTV
8K UHD	4320	0,8 lần	≈ 96°	Chỉ đạt được ở rạp hoặc màn hình LED sân khấu

$$\alpha = 2\arctan\left(\frac{W_{\text{scr}}}{2d}\right) \tag{8.7}$$

Hai kết luận thực tế rút ra từ bảng này. Thứ nhất, một chiếc tivi 4K đặt cách người xem ba lần chiều cao màn hình thì phần lớn độ phân giải bị lãng phí vì mắt không phân biệt được. Thứ hai, độ phân giải 8K chỉ có ý nghĩa ở nơi màn hình đủ lớn hoặc người xem đủ gần, chẳng hạn rạp chiếu, màn hình LED sân khấu, hoặc màn hình đặt sát mắt như kính thực tế ảo.

> **QUY TẮC 8.1 — ĐIỀU NÊN GIẢI THÍCH CHO KHÁCH HÀNG**
>
> Khi khách hàng hỏi vì sao phải quay 4K trong khi tivi của họ đặt ở xa, lý do thật sự không nằm ở độ nét khi xem mà ở khâu hậu kỳ: quay 4K rồi phát hành Full HD cho phép cắt cúp lại khung hình, ổn định hình rung và giữ được độ nét sau khi thu nhỏ.
>
> Nghĩa là độ phân giải dư là phần dự trữ cho hậu kỳ chứ không phải cho mắt người xem.

8.5 Lượng tử hoá và độ sâu bit

8.5.1 Số mức và bước lượng tử

Sau khi đã rời rạc hoá theo không gian và thời gian, mỗi mẫu vẫn là một số thực, và bước cuối cùng là làm tròn mẫu đó về một mức trong thang hữu hạn.

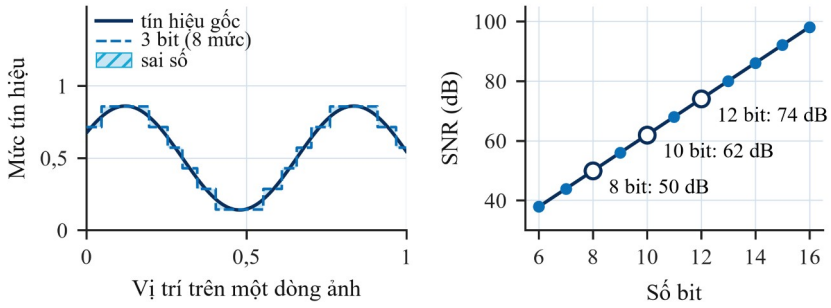
$$L = 2^{n} \quad ; \quad \Delta = \frac{V_{\max} - V_{\min}}{2^{n} - 1} \tag{8.8}$$

Bảng 8.7 Độ sâu bit và số mức

Độ sâu bit	Số mức mỗi thành phần	Số màu tổ hợp được	Dùng ở đâu
8 bit	256	khoảng 16,7 triệu	Phát sóng SD và HD, tệp phát hành thông thường
10 bit	1 024	khoảng 1,07 tỉ	Sản xuất chuyên nghiệp, ảnh log, HDR, chuẩn giao nộp của đài
12 bit	4 096	khoảng 68,7 tỉ	Tệp RAW của máy quay điện ảnh, bản gốc lưu trữ
16 bit	65 536	rất lớn	Bản trung gian trong phần mềm chỉnh màu, không dùng để giao nộp

Sai số do làm tròn được gọi là nhiễu lượng tử, và quan hệ giữa nhiễu này với số bit đúng như quan hệ đã trình bày ở Chương 3.

$$\mathrm{SNR}_q \approx 6{,}02\,n + 1{,}76 \quad \text{(dB)} \tag{8.9}$$



(a) Sai số lượng tử (b) SNR theo số bit

Hình 8.8 Sai số lượng tử trên một dòng ảnh, và quan hệ giữa số bit với tỉ số tín hiệu trên nhiễu

8.5.2 Hiện tượng dải màu

Ở bước này có một khác biệt quan trọng giữa âm thanh và hình ảnh. Với âm thanh, sai số lượng tử được nghe như nhiễu nền, khó chịu nhưng dễ bỏ qua; với hình ảnh, sai số lượng tử không hiện ra như nhiễu mà như những đường viền giả chạy ngang qua các vùng lẽ ra phải chuyển sắc êm, chẳng hạn bầu trời, phông vải hoặc vệt sáng đèn.

Nguyên nhân là mắt người rất nhạy trong việc phát hiện biên. Một bước nhảy chỉ một mức trên hai trăm năm mươi sáu mức, nếu chạy thành một đường liền trên nền phẳng, vẫn bị mắt nhận ra, và hiện tượng này được gọi là dải màu.



(a) 8 bit: 256 mức, chuyển sắc liền mạch



(b) 5 bit: 32 mức, thấy rõ đường viền giả



(c) 5 bit cộng dither: cùng 32 mức, viền tan thành nhiễu mịn

Hình 8.9 Hiện tượng dải màu và tác dụng của dither

Phương pháp khắc phục tương tự như trong âm thanh là thêm một lượng nhiễu ngẫu nhiên rất nhỏ trước khi làm tròn. Dither không bổ sung thông tin mới; kỹ thuật này chỉ chuyển một lỗi có cấu trúc thành nhiễu ngẫu nhiên, và thị giác con người chịu đựng nhiễu ngẫu nhiên tốt hơn nhiều.

GHI NHỚ 8.4 — VÌ SAO MƯỜI BIT LÀ MỨC TỐI THIỂU CHO ẢNH LOG

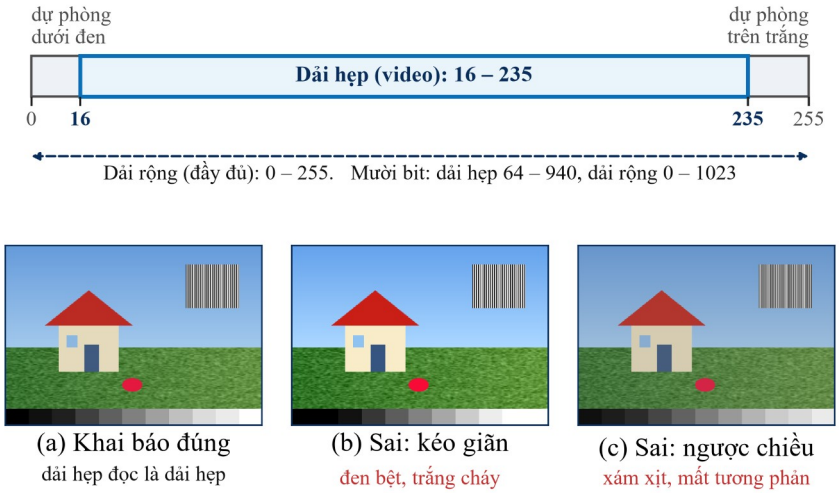
Như Chương 7 đã trình bày, đường log dồn nhiều stop dải động vào cùng một khoảng mã, do đó mỗi mức mã phải biểu diễn một khoảng độ chói lớn hơn so với đường gamma thông thường.

Với 8 bit, sau khi đưa ảnh log về không gian hiển thị, các bước nhảy giữa hai mức kề nhau đủ lớn để mắt nhận ra, và ảnh xuất hiện hiện tượng dải màu ngay ở vùng bầu trời. Với 10 bit, số mức tăng gấp bốn nên bước nhảy nhỏ đi bốn lần và hiện tượng này biến mất. Đây là lý do kỹ thuật khiến mọi máy quay có chế độ log đều ghi ở 10 bit trở lên.

8.5.3 Dải hẹp và dải rộng

Hiện tượng này là nguyên nhân của lỗi phổ biến nhất trong phòng dựng.

Khi tín hiệu hình được lượng tử hoá, thang mã không được sử dụng toàn bộ: ở tám bit, mức đen được đặt tại 16 và mức trắng tại 235, còn hai đầu thang được chừa lại làm khoảng dự phòng. Cách này gọi là *dải hẹp*, còn được gọi là dải video hoặc dải studio, trong khi cách dùng trọn thang từ 0 tới 255 gọi là *dải rộng* hay dải đầy đủ.



Hình 8.10 Dải hẹp và dải rộng, và hậu quả khi gán sai

Bảng 8.8 Phạm vi mức theo độ sâu bit

Độ sâu bit	Thành phần	Dải hẹp	Dải rộng
8 bit	Độ chói Y′	16 – 235	0 – 255
8 bit	Màu Cb, Cr	16 – 240 (tâm ở 128)	0 – 255
10 bit	Độ chói Y′	64 – 940	0 – 1023
10 bit	Màu Cb, Cr	64 – 960 (tâm ở 512)	0 – 1023

Việc chừa khoảng trống xuất phát từ hai lý do lịch sử vẫn còn giá trị. Thứ nhất, các bộ lọc trong dây chuyền xử lý có thể sinh ra các giá trị vượt quá nhất thời ở hai đầu, và nếu không có khoảng dự phòng thì tín hiệu sẽ bị cắt ngọn. Thứ hai, khoảng trống này còn được dùng để mang tín hiệu đồng bộ và dữ liệu phụ trong hệ thống tương tự.

$$D_{\text{full}} = (D_{\text{narrow}} - 16) \times \frac{255}{219} \quad \text{(8 bit)} \tag{8.10}$$

CẢNH BÁO 8.2 — LỖI SỐ MỘT TRONG PHÒNG DỰNG

Nếu phần mềm đọc một tệp dải hẹp nhưng coi đó là dải rộng, thang mã bị kéo giãn: mức 16 trở thành 0 và mức 235 trở thành 255. Kết quả là vùng tối bị bẹt đen và vùng sáng bị cháy trắng, làm mất toàn bộ chi tiết ở hai đầu.

Trong trường hợp sai theo chiều ngược lại, hình trở nên xám nhạt, đen không đạt mức đen và trắng không đạt mức trắng.

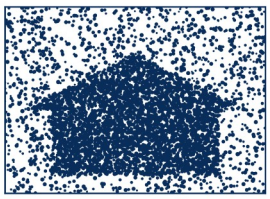
Cả hai lỗi đều không phải do hỏng tệp và đều có thể sửa bằng một khai báo, nhưng khó khăn nằm ở việc nhận diện: người chưa nắm nội dung này thường chỉ nhận xét rằng hình có vẻ không đúng mà không xác định được vấn đề.

8.6 Ảnh trên phim nhựa — một đối chiếu

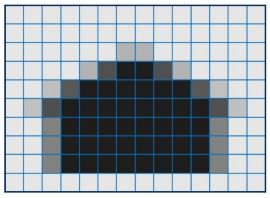
8.6.1 Hạt bạc và điểm ảnh

Ảnh trên phim nhựa được trình bày ở đây như một đối chiếu giúp hiểu ảnh số rõ hơn, không phải như một nội dung độc lập.

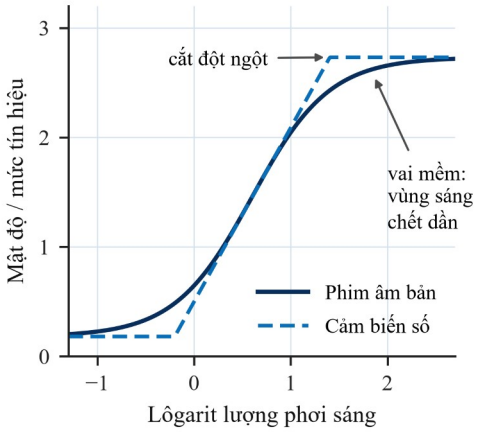
Lớp nhũ tương của phim chứa các tinh thể bạc halogenua. Khi photon tác động, tinh thể bị kích hoạt, và trong khâu tráng, phản ứng khử biến ion bạc thành bạc kim loại màu đen: $Ag^{+} + e^{-} \rightarrow Ag$. Vùng nhận nhiều ánh sáng sẽ có mật độ đen lớn, và đó là nguyên lý của phim âm bản.



(a) Phim: hạt ngẫu nhiên



(b) Cảm biến: lưới đều



(c) Đường đặc tính

Hình 8.11 Hạt phim so với điểm ảnh, và đường đặc tính của hai loại vật mang

Bảng 8.9 Phim nhựa và cảm biến số — bốn khác biệt cốt lõi

Vấn đề	Phim nhựa	Cảm biến số
Cách sắp xếp mẫu	Hạt bạc phân bố NGẪU NHIÊN, kích thước không đều	Điểm ảnh xếp theo LƯỚI ĐỀU, cùng kích thước, cố định vị trí
Hệ quả với nhiễu	Nhiễu hạt phân bố ngẫu nhiên, mắt cảm thấy tự nhiên; không có vân moiré vì không có lưới	Nhiễu có cấu trúc theo lưới, dễ lộ; có thể sinh vân moiré khi gặp hoa văn tuần hoàn
Đường đặc tính	Hình chữ S: chân uốn ở vùng tối, VAI UỐN MỀM ở vùng sáng nên vùng sáng chết dần và còn giữ chi tiết	Tuyến tính rồi CẮT ĐỘT NGỘT: quá điểm bão hoà thì mọi chi tiết thành trắng như nhau
Cách khắc phục	Không cần — đặc tính vốn đã mềm	Sinh ra các đường log và HDR ở Chương 7 chính là để mô phỏng lại cái vai mềm ấy

8.6.2 Vì sao vẫn cần hiểu phim

Có ba lý do rất thực tế, không phải hoài cổ.

– Kho tư liệu của các đài truyền hình và của Viện Phim Việt Nam còn lưu giữ rất nhiều phim nhựa, và việc số hoá đúng đòi hỏi phải hiểu đường đặc tính của bản gốc, nếu không bản quét sẽ bị mất chi tiết ở vùng sáng hoặc vùng tối.

– Toàn bộ ngôn ngữ nghề nghiệp về phơi sáng, dải động và "chất phim" đều bắt nguồn từ phim nhựa; khi người quay phim nói "giữ được vai sáng", họ đang nói về đường đặc tính của phim, kể cả khi sử dụng máy quay số.
– Các đường log và các bảng chuyển đổi màu ở Chương 7 phần lớn được thiết kế để mô phỏng đặc tính của phim, vì vậy việc hiểu hình dạng của chúng đòi hỏi phải hiểu đặc tính của bản gốc.

Xét về độ phân giải tương đương, phim 35 mm quét ở 4K thường được coi là đủ để thu lại toàn bộ thông tin có ích, mặc dù con số chính xác phụ thuộc loại phim, ống kính và điều kiện chụp. Quan trọng hơn con số ấy, chất lượng của một hình không nằm ở số điểm ảnh mà ở hình dạng của đường đặc tính, đúng như bảng trên đã chỉ ra.

8.7 Bài thực hành số 1 — lấy mẫu, độ phân giải và độ sâu bit

Mục tiêu

Sau bài này, người đọc đo được độ phân giải thực của một máy quay, tạo ra và nhận biết được vân moiré, quan sát được hiện tượng dải màu và phân biệt được hậu quả của việc gán sai phạm vi mức.

Thiết bị cần có

– Một máy quay cho phép chỉnh tay và thay đổi độ phân giải ghi.
– Một bảng đo độ phân giải in trên giấy (biểu đồ hình nêm) hoặc hiển thị trên màn hình máy tính.
– Một vật có hoa văn mảnh, chẳng hạn áo kẻ sọc, rèm cửa hoặc chính màn hình máy tính quan sát từ xa.
– Phần mềm dựng có waveform và histogram, trong đó bản miễn phí của DaVinci Resolve là đủ.

Nội dung thực hành

Bảng 8.10 Bốn phần của bài thực hành số 1

Phần	Việc phải làm	Kết quả cần rút ra
1. Đo độ phân giải thật	Quay bảng đo ở khoảng cách sao cho bảng chiếm trọn chiều cao khung hình. Phóng to hình trên máy tính, tìm chỗ các vạch của biểu đồ hình nêm bắt đầu nhập vào nhau. Đọc con số ở chỗ đó và so với độ phân giải ghi trên nhãn máy.	Độ phân giải THẬT gần như luôn thấp hơn số điểm ảnh ghi trên nhãn. Giải thích được vì sao: ống kính, bộ lọc quang thông thấp, và cách khử ghép màu của ma trận Bayer đều làm giảm độ nét.
2. Tạo và khử vân moiré	Quay vật có hoa văn mảnh, lùi xa dần cho tới khi thấy vân màu lạ. Ghi lại khoảng cách. Sau đó thử ba cách xử lý: đổi góc máy, đổi khoảng cách, mở khẩu lớn hơn.	Vân moiré xuất hiện khi tần số không gian của vật vượt quá giới hạn của lưới điểm ảnh. Ba cách xử lý đều nhằm hạ tần số ấy xuống, không phải "sửa" ở khâu hậu kỳ.

Phần	Việc phải làm	Kết quả cần rút ra
3. Nhìn thấy dải màu	Dựng một dải chuyển sắc êm (bầu trời hoặc phông vải trơn). Xuất hai bản: một bản 8 bit, một bản 10 bit. Mở cả hai, đẩy mạnh tương phản, so sánh.	Đường viền giả xuất hiện ở bản 8 bit và biến mất ở bản 10 bit. Thử thêm dither cho bản 8 bit và quan sát viền tan thành nhiễu mịn.
4. Gán sai phạm vi mức	Lấy một cảnh quay bất kỳ, trong phần mềm dựng đổi khai báo phạm vi mức từ dải hẹp sang dải rộng và ngược lại. Quan sát waveform trong từng trường hợp.	Nhận ra hai kiểu hỏng: bệt đen và cháy trắng ở một chiều, xám xịt mất tương phản ở chiều kia. Quan trọng hơn: học cách ĐỌC waveform để phát hiện lỗi này chứ không đoán bằng mắt.

Kết quả cần có

Một bản ghi chép ngắn gồm: con số độ phân giải thật đo được và so sánh với nhãn máy; ảnh chụp màn hình cho thấy vân moiré trước và sau khi xử lý; hai ảnh so sánh dải màu ở 8 bit và 10 bit; và bốn ảnh chụp waveform ứng với bốn cách khai báo phạm vi mức, mỗi ảnh kèm một câu giải thích bằng ngôn ngữ kỹ thuật của chương.

8.8 Từ RGB sang Y′CbCr trong thực tế

8.8.1 Nhắc lại kết quả của Chương 7

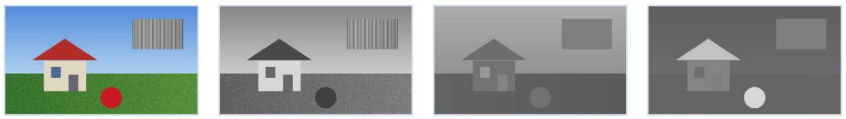
Mục 7.17 của Chương 7 đã xây dựng phép biến đổi từ ba thành phần màu sang một thành phần độ chói và hai thành phần hiệu số màu. Phần này không nhắc lại lý thuyết mà đi vào những hệ quả kỹ thuật mà người làm nghề gặp hằng ngày [1], [2].

Tóm lược lại, thành phần độ chói là tổ hợp có trọng số của ba thành phần màu đã mã hoá gamma.

$$Y' = K_R R' + K_G G' + K_B B' \quad , \quad K_R + K_G + K_B = 1 \tag{8.11}$$

Hai thành phần còn lại không mang thông tin sáng tối mà chỉ mang thông tin về mức lệch màu so với xám, nên chúng được xây dựng từ hai hiệu số rồi được tỉ lệ hoá về khoảng giá trị thuận tiện.

$$C_b = \frac{B' - Y'}{2(1 - K_B)} \quad ; \quad C_r = \frac{R' - Y'}{2(1 - K_R)} \tag{8.12}$$



(a) Ảnh gốc R′, G′, B′ — (b) Y′ độ chói, giữ chi tiết — (c) Cb B′ − Y′, phẳng và êm — (d) Cr R′ − Y′, phẳng và êm

Hình 8.12 Một cảnh tách thành ba thành phần Y′, Cb và Cr

Hình trên minh hoạ điều quan trọng nhất của cả Phần B: gần như toàn bộ chi tiết của cảnh nằm trong ảnh Y′, trong khi hai ảnh màu tương đối phẳng và biến

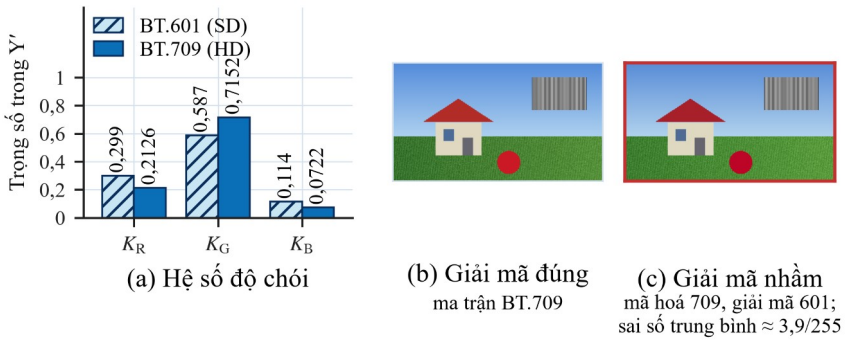
thiên chậm. Đây chính là chỗ có thể tiết kiệm dữ liệu, và cách thực hiện là nội dung của mục sau.

8.8.2 Hai bộ hệ số, hai chuẩn

Trọng số của ba thành phần màu trong tổng độ chói không phải là các con số tuỳ ý mà phản ánh độ nhạy của mắt với từng vùng phổ. Do các chuẩn khác nhau định nghĩa ba màu cơ bản ở những toạ độ khác nhau, trọng số cũng khác nhau.

Bảng 8.11 Hai bộ hệ số độ chói thường gặp

Chuẩn	K_R (đỏ)	K_G (lục)	K_B (lam)	Dùng cho
ITU-R BT.601	0,299	0,587	0,114	Tư liệu SD: 720 × 576 và 720 × 480
ITU-R BT.709	0,2126	0,7152	0,0722	HD và Full HD; cũng là hệ số của sRGB
ITU-R BT.2020	0,2627	0,6780	0,0593	UHD gam màu rộng, HDR



Hình 8.13 Hai bộ hệ số độ chói, và hậu quả khi giải mã bằng nhầm ma trận

CẢNH BÁO 8.3 — LỖI TRỘN TƯ LIỆU SD VỚI TƯ LIỆU HD

Nếu một tệp được mã hoá theo BT.709 nhưng phần mềm giải mã bằng ma trận BT.601, toàn bộ cảnh sẽ lệch sắc theo cùng một hướng. Trên một khung tĩnh, sai lệch có vẻ không lớn; tuy nhiên, khi ghép với cảnh khác được giải mã đúng, hai cảnh không khớp màu và sản phẩm không đạt yêu cầu nghiệm thu.

Tình huống thường gặp nhất là dựng phóng sự có xen tư liệu SD cũ của đài. Biện pháp phòng ngừa là khai báo rõ không gian màu cho từng đoạn tư liệu ngay khi nhập vào dự án, thay vì để phần mềm tự suy đoán.

8.8.3 Từ số thực sang mã số

Ba thành phần sau khi biến đổi vẫn là số thực, và bước cuối cùng là đưa chúng về thang mã nguyên, nơi quy ước dải hẹp đã học ở mục 8.5 được áp dụng.

$$D_Y = \text{round}(219\,Y' + 16) \quad ; \quad D_C = \text{round}(224\,C + 128) \tag{8.13}$$

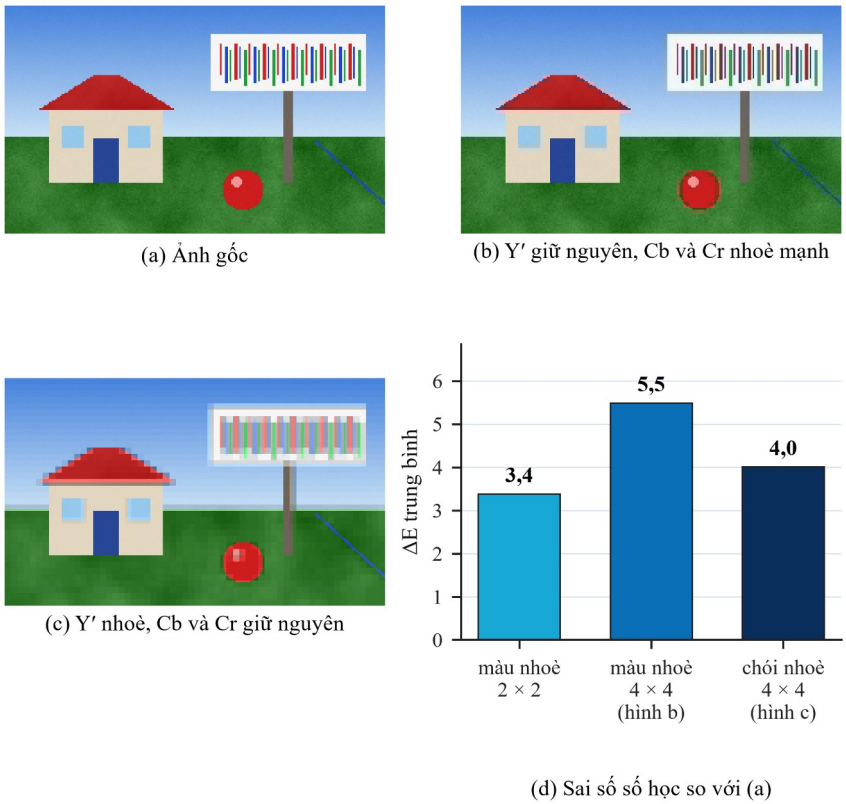
Cần lưu ý hai điểm. Thứ nhất, hệ số của độ chói là 219 chứ không phải 255, đúng bằng khoảng cách từ 16 tới 235. Thứ hai, hai thành phần màu được cộng thêm 128 để đưa giá trị âm về vùng dương, và đó là lý do mức không màu trong tín hiệu số nằm ở 128 chứ không phải ở 0.

8.9 Lấy mẫu con màu

8.9.1 Vì sao có thể bớt mẫu màu

Chương 7 (mục 7.10) đã cho thấy mắt người phân biệt chi tiết mịn của độ chói tốt hơn nhiều so với chi tiết mịn của màu. Mục 8.8 đã tách tín hiệu hình thành ba thành phần: độ chói Y′ mang gần hết chi tiết nhìn thấy, còn hai hiệu số màu Cb và Cr biến đổi chậm và êm. Từ hai đặc điểm này, ý tưởng nén tự nhiên là giữ độ chói ở độ phân giải đầy đủ và lấy mẫu thưa hơn đối với màu, và kỹ thuật đó được gọi là lấy mẫu con màu (chroma subsampling) [32].

Ý tưởng này có thể được kiểm chứng bằng thực nghiệm theo hai chiều ngược nhau trên cùng một cảnh. Hình 8.14 làm nhoè mạnh hai thành phần màu nhưng giữ nguyên độ chói (ảnh b), rồi làm điều ngược lại: nhoè độ chói và giữ nguyên màu (ảnh c). Với ảnh (b), mắt gần như không thấy khác biệt so với ảnh gốc, còn với ảnh (c), ảnh mờ được nhận ra ngay. Mức làm nhoè như nhau nhưng hậu quả nhìn thấy hoàn toàn khác nhau. Đáng chú ý là về mặt số học, sai số màu của ảnh (b) (ΔE trung bình 5,5) còn lớn hơn sai số của ảnh (c) (4,0), nhưng mắt chỉ khó chịu với ảnh (c). Con số đo khoảng cách màu không phản ánh độ nhạy của mắt với chi tiết, nên thiết kế nén phải dựa vào cảm nhận chứ không chỉ vào sai số số học.



Hình 8.14 Thí nghiệm hai chiều ngược nhau: làm nhoè màu mà giữ độ chói (b) hầu như không thấy, làm nhoè độ chói mà giữ màu (c) thấy ngay

Đó là toàn bộ cơ sở của kỹ thuật này: loại bỏ thông tin ở đúng nơi mắt kém nhạy. Về phân loại, lấy mẫu con màu thuộc loại dư thừa thứ tư của mục 8.14, tức dư thừa cảm nhận, nên là phép nén có tổn hao. Mất mát chỉ xảy ra ở thành phần màu và độ chói không bị tác động. Phép này diễn ra ngay ở đầu dây chuyền, trước

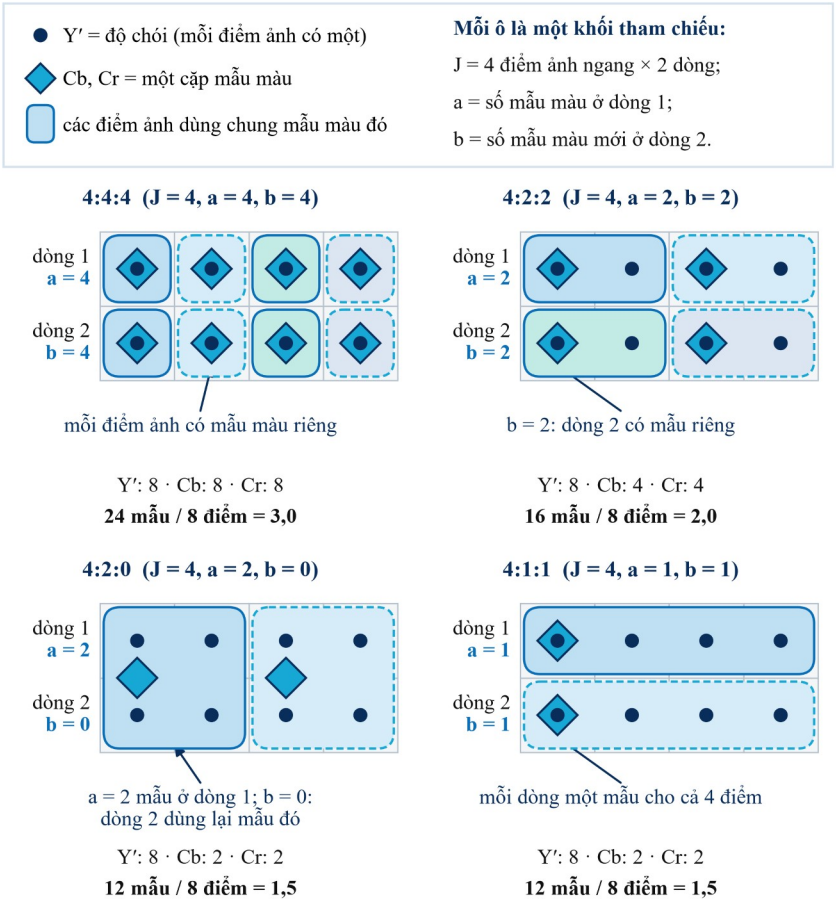
cả bước biến đổi cosin của mục 8.16, nên dữ liệu đưa vào bộ nén đã giảm đi một phần ba hoặc một nửa.

> **GHI NHỚ 8.5 — CƠ SỞ SINH LÝ CỦA LẤY MẪU CON MÀU**
>
> Việc giảm số mẫu màu dựa trên đặc điểm của thị giác: mắt nhạy với chi tiết độ chói hơn nhiều lần so với chi tiết màu, như mục 7.10 của Chương 7 đã trình bày.
>
> Vì vậy kỹ thuật này giữ nguyên độ chói và chỉ lấy thưa các thành phần màu; nếu mắt nhạy với màu hơn với độ chói thì phép nén này sẽ không có ý nghĩa.

8.9.2 Ký hiệu J:a:b — đọc từng bước

Ký hiệu ba số như 4:2:0 thường bị hiểu là một tỉ lệ, nhưng thực chất nó mô tả một *khối tham chiếu*. Cách đọc chính xác nhất là vẽ khối ra rồi đếm, theo bốn bước sau (Hình 8.15):

1. **J** — chiều rộng của khối tham chiếu, tính bằng số điểm ảnh. Theo quy ước, J luôn bằng 4 và khối cao 2 dòng, nên mỗi khối gồm 4 × 2 = 8 điểm ảnh.
2. **Đếm độ chói.** Mỗi điểm ảnh luôn có mẫu Y′ riêng, nên mọi khối đều có đúng 8 mẫu Y′ bất kể cấu trúc nào, và độ chói không bao giờ bị lấy thưa.
3. **a** — số mẫu màu lấy được trên dòng thứ nhất của khối, trong đó một mẫu màu là một cặp (Cb, Cr) đi cùng nhau.
4. **b** — số mẫu màu được thêm ở dòng thứ hai; giá trị 0 nghĩa là dòng thứ hai không có mẫu mới và dùng chung mẫu của dòng trên.



Hình 8.15 Cách đọc ký hiệu J:a:b trên khối tham chiếu 4 × 2 điểm ảnh: chấm Y' là mẫu độ chói, hình thoi là cặp mẫu màu (Cb, Cr), vùng tô là các điểm ảnh dùng chung một mẫu màu

Từ cách đọc ấy có thể suy ra số mẫu trung bình cho mỗi điểm ảnh. Trong một khối, độ chói có 2J mẫu (J mẫu mỗi dòng, hai dòng), còn mỗi thành phần Cb và Cr có a + b mẫu. Chia tổng số mẫu cho số điểm ảnh 2J:

$$s = \frac{2J + 2(a+b)}{2J} = \frac{J + a + b}{J} \tag{8.14}$$

Áp dụng vào bốn cấu trúc: 4:4:4 cho (4 + 4 + 4)/4 = 3,0; 4:2:2 cho (4 + 2 + 2)/4 = 2,0; 4:2:0 cho (4 + 2 + 0)/4 = 1,5; 4:1:1 cho (4 + 1 + 1)/4 = 1,5.

Bảng 8.12 Bốn cấu trúc lấy mẫu con màu

Cấu trúc	Mẫu màu theo chiều ngang	Mẫu màu theo chiều dọc	Mẫu mỗi điểm ảnh	Tiết kiệm
4:4:4	đủ	đủ	3,0	0 %
4:2:2	một nửa	đủ	2,0	33 %
4:2:0	một nửa	một nửa	1,5	50 %
4:1:1	một phần tư	đủ	1,5	50 %

CẢNH BÁO 8.4 — Ý NGHĨA CỦA CHỮ SỐ 0 TRONG 4:2:0

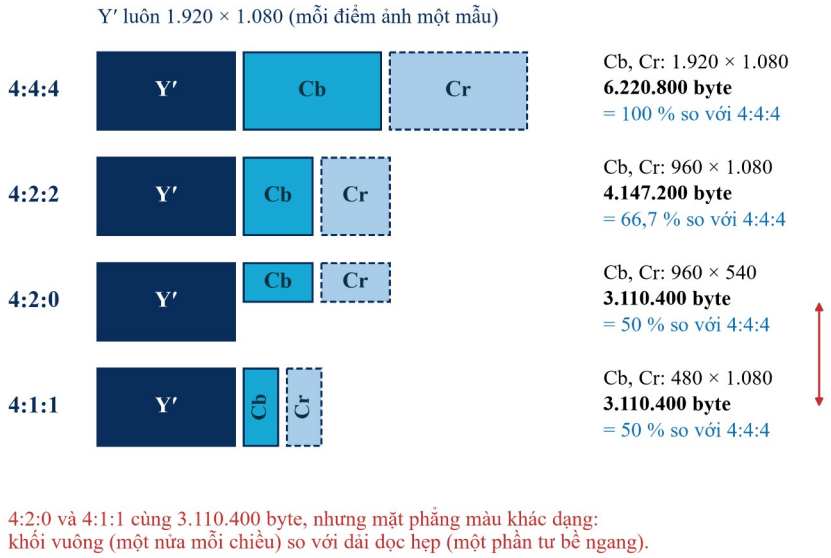
Chữ số cuối chỉ cho biết dòng thứ hai của khối không lấy thêm mẫu màu mới (b = 0). Mọi điểm ảnh vẫn có màu, nhưng bốn điểm ảnh của một khối 2 × 2 dùng chung một cặp (Cb, Cr).

Cấu trúc 4:2:0 lấy mẫu cả hai thành phần Cb và Cr với tần số bằng một nửa theo cả chiều ngang lẫn chiều dọc.

Bảng trên cho thấy 4:2:0 và 4:1:1 có cùng số mẫu, 1,5 mẫu mỗi điểm ảnh, nhưng hệ quả của hai cấu trúc này khác nhau do hình dạng vùng dùng chung. Cấu trúc 4:1:1 gộp bốn điểm ảnh nằm ngang thành một mẫu màu, nên màu bị nhoè mạnh theo chiều ngang và giữ nguyên theo chiều dọc, và mắt nhận ra rõ ở các cạnh đứng. Cấu trúc 4:2:0 gộp một khối vuông 2 × 2, nên màu nhoè đều theo hai chiều và mắt khó nhận ra hơn. Vì vậy 4:2:0 chiếm ưu thế trong mọi chuẩn phát hành, còn 4:1:1 chỉ còn gặp ở một số định dạng băng từ cũ. Mục 8.9.4 sẽ minh hoạ trực quan điều này.

8.9.3 Từ khối 4 × 2 lên cả khung hình — mặt phẳng màu nhỏ đi bao nhiêu

Một khung hình không phải là ảnh đơn mà gồm ba *mặt phẳng* mẫu: mặt phẳng Y′ và hai mặt phẳng màu Cb, Cr. Mặt phẳng Y′ luôn có đủ số điểm ảnh của khung, còn cấu trúc lấy mẫu con màu chỉ quyết định kích thước của hai mặt phẳng màu. Hình 8.16 vẽ đúng tỉ lệ ba mặt phẳng ấy cho khung 1920 × 1080 ở cả bốn cấu trúc.



Hình 8.16 Kích thước các mặt phẳng Y′, Cb và Cr của khung 1920 × 1080 ở bốn cấu trúc, vẽ đúng tỉ lệ, kèm số byte mỗi khung ở 8 bit

Bảng 8.13 Kích thước mặt phẳng và dung lượng một khung 1920 × 1080 (8 bit mỗi mẫu)

Cấu trúc	Mặt phẳng Y′	Mỗi mặt phẳng Cb, Cr	Số byte mỗi khung	So với 4:4:4
4:4:4	1920 × 1080	1920 × 1080	6 220 800	100 %
4:2:2	1920 × 1080	960 × 1080	4 147 200	66,7 %
4:2:0	1920 × 1080	960 × 540	3 110 400	50 %
4:1:1	1920 × 1080	480 × 1080	3 110 400	50 %

Theo bảng, ở 4:2:2, mỗi dòng màu chỉ có một nửa số cột nhưng vẫn đủ 1080 dòng. Ở 4:2:0, mặt phẳng màu chỉ cao 540 dòng, tức một dòng màu phục vụ hai dòng độ chói liền kề. Ở 4:1:1, mặt phẳng màu giữ đủ 1080 dòng nhưng chỉ còn 480 cột. Hai cấu trúc 4:2:0 và 4:1:1 cùng chiếm 3 110 400 byte, song mặt phẳng màu của chúng có hình dạng khác hẳn nhau: một bên giảm một nửa theo mỗi chiều, bên kia giảm một phần tư theo chiều ngang. Nhân số mẫu với độ sâu bit sẽ được tốc độ bit thô ở mục 8.14.

8.9.4 Mất mát xảy ra ở đâu — lọc xuống rồi nội suy lên

Lấy mẫu con màu gồm hai nửa nằm ở hai đầu của dây chuyền. Ở bộ mã hoá có hai bước: (1) lọc, tức lấy trung bình các điểm ảnh sẽ dùng chung một mẫu màu (nếu bỏ lọc mà chỉ chọn một điểm ảnh thì sinh răng cưa màu, đúng như hiện tượng ở mục 8.3); (2) giảm mẫu, tức giữ lại một mẫu cho mỗi nhóm. Ở bộ giải mã có bước (3) nội suy, nhằm gán cho mỗi điểm ảnh một giá trị màu; cách đơn giản nhất là nhân bản mẫu cho cả nhóm, còn cách tốt hơn là nội suy tuyến tính giữa các mẫu.

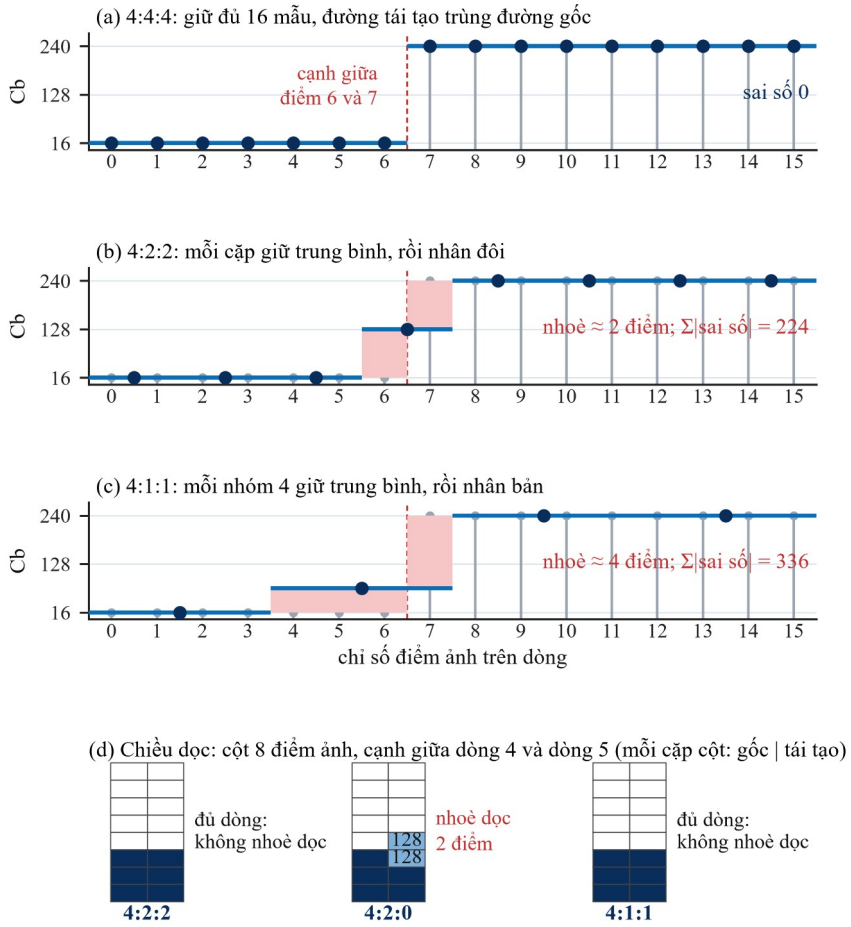
Hiện tượng được quan sát rõ nhất trên một dòng điểm ảnh có một cạnh màu. Giả sử tám điểm ảnh liên tiếp có giá trị Cb như bảng dưới đây, với cạnh nằm giữa điểm thứ ba và thứ tư.

Bảng 8.14 Một dòng 8 điểm ảnh có cạnh màu qua 4:2:2 và 4:1:1 (nhân bản khi nội suy)

Dòng	Giá trị Cb theo thứ tự điểm ảnh 1 – 8	Tổng sai số tuyệt đối
Gốc (4:4:4)	16 · 16 · 16 · 240 · 240 · 240 · 240 · 240	0
Giữ lại ở 4:2:2 (trung bình cặp)	16 · 128 · 240 · 240	
Tái tạo từ 4:2:2	16 · 16 · 128 · 128 · 240 · 240 · 240 · 240	224
Giữ lại ở 4:1:1 (trung bình nhóm 4)	72 · 240	
Tái tạo từ 4:1:1	72 · 72 · 72 · 72 · 240 · 240 · 240 · 240	336

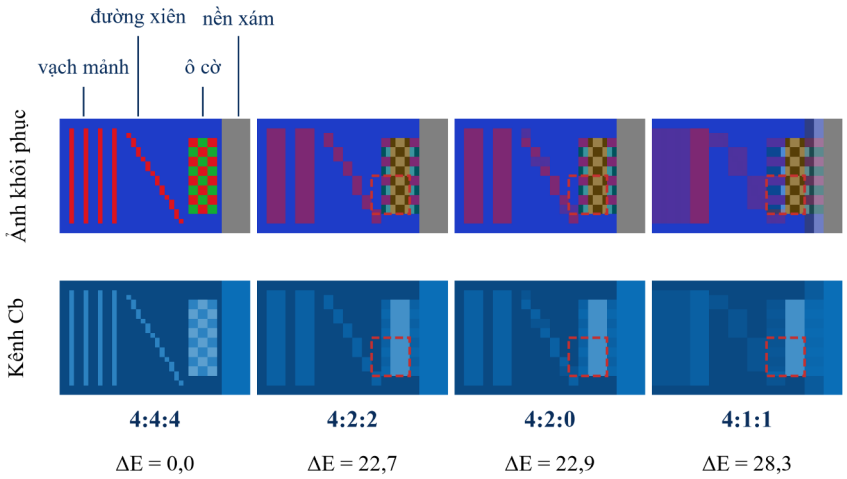
Cặp thứ hai (điểm 3 và 4) nằm vắt qua cạnh nên mẫu giữ lại là trung bình 128 của hai màu, và cả hai điểm ảnh nhận giá trị 128 khi tái tạo, do đó cạnh sắc biến thành một bậc trung gian rộng hai điểm ảnh. Ở 4:1:1, nhóm bốn điểm đầu

cũng vắt qua cạnh, mẫu giữ lại là 72 và bậc nhoè rộng bốn điểm ảnh với sai số lớn hơn. Từ đó rút ra hai nhận xét: mất mát chỉ xảy ra ở chỗ màu đổi nhanh, và mẫu càng thưa thì vùng nhoè càng rộng. Nếu cạnh trùng đúng ranh giới nhóm (giữa điểm 4 và 5) thì cả hai cấu trúc tái tạo trọn vẹn, không có sai số, nghĩa là mức độ ảnh hưởng còn phụ thuộc vào vị trí cạnh so với ranh giới nhóm. Hình 8.17 vẽ hiện tượng này trên một dòng dài hơn, kèm chiều dọc: 4:2:0 nhoè cả hai chiều, còn 4:1:1 chỉ nhoè theo chiều ngang nhưng rộng gấp đôi 4:2:2.



Hình 8.17 Cạnh màu trên một dòng điểm ảnh sau khi lọc xuống và nội suy lên: mẫu màu càng thưa thì vùng nhoè càng rộng

Thao tác tương tự được thực hiện trên ảnh hai chiều. Hình 8.18 cho cùng một mẫu thử màu đi qua cả bốn cấu trúc (RGB sang Y′CbCr, giảm mẫu, nội suy, rồi về RGB), phóng to từng điểm ảnh để quan sát rõ.

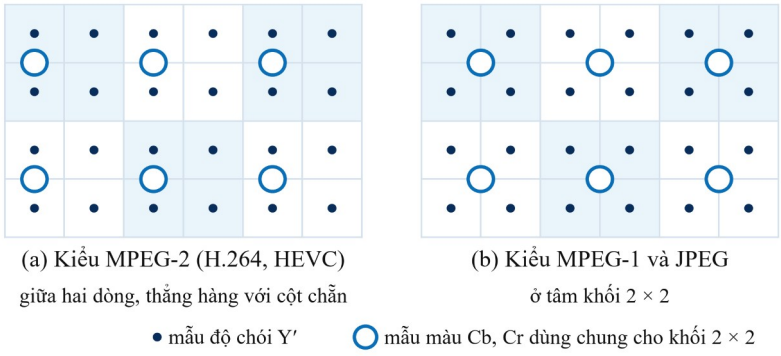


Hình 8.18 Cùng một mẫu thử màu qua bốn cấu trúc lấy mẫu con, phóng to từng điểm ảnh: vùng xám không đổi, hỏng ở vạch mảnh, đường xiên và ô cờ

Hình cho thấy những đặc điểm mà bảng số liệu không thể hiện được. Vùng xám gần như không đổi qua cả bốn cấu trúc vì độ chói được giữ nguyên, trong khi hư hỏng tập trung ở các vạch đỏ mảnh, đường xiên và ô cờ, tức là những vị trí màu thay đổi nhanh. Sai số màu trung bình so với bản 4:4:4 là ΔE = 22,7 ở 4:2:2, 22,9 ở 4:2:0 và 28,3 ở 4:1:1. Các giá trị này lớn vì mẫu thử được chọn có chủ ý với toàn màu bão hoà và tương phản mạnh, nên không dùng để so sánh với ảnh thật; điều đáng quan sát là dạng hư hỏng, trong đó 4:1:1 làm vạch mảnh và đường xiên trải rộng nhất theo chiều ngang, còn 4:2:0 làm nhoè đều theo cả hai chiều.

8.9.5 Vị trí đặt mẫu màu

Một chi tiết nhỏ nhưng gây lỗi thực sự là vị trí được quy ước của cặp mẫu màu khi bốn điểm ảnh dùng chung cặp mẫu ấy.



Hình 8.19 Hai quy ước đặt vị trí mẫu màu trong cấu trúc 4:2:0

Bảng 8.15 Hai quy ước vị trí mẫu màu

Quy ước	Vị trí	Dùng trong
Kiểu MPEG-2	Đồng vị trí với cột chẵn theo chiều ngang; nằm GIỮA hai dòng theo chiều dọc	MPEG-2, H.264, HEVC — tức gần như mọi chuẩn phát sóng và phát hành
Kiểu MPEG-1 / JPEG	Nằm ở TÂM của khối 2 × 2	MPEG-1, ảnh JPEG, và một số bộ giải mã đời cũ

Hai quy ước về vị trí này lệch nhau nửa điểm ảnh, vì vậy chuyển tư liệu từ quy ước này sang quy ước kia mà không bù sẽ làm màu dịch đi nửa điểm ảnh so với độ chói. Trên hình tĩnh, lỗi này khó nhận thấy; tuy nhiên, ở mép các vật có màu tương phản mạnh sẽ xuất hiện một viền màu mảnh chạy dọc theo cạnh, và người có kinh nghiệm nhận ra lỗi ấy ngay.

8.9.6 Hệ quả thực tế

Kết quả ở mục 8.9.4 cho thấy mất mát tập trung ở những vị trí màu thay đổi nhanh. Bảng dưới đây chuyển nhận xét ấy thành lựa chọn cấu trúc theo từng loại công việc.

Bảng 8.16 Chọn cấu trúc lấy mẫu con theo công việc

Công việc	Cấu trúc tối thiểu	Lý do
Xem và phát hành cho khán giả	4:2:0	Mắt không nhận ra khác biệt ở điều kiện xem bình thường; tiết kiệm một nửa dữ liệu màu
Dựng thông thường, cắt ghép	4:2:0 vẫn chấp nhận được	Không xử lý màu cục bộ thì không lộ nhược điểm
Chỉnh màu sâu, đổi màu cục bộ	4:2:2	Kéo màu mạnh làm lộ ra biên nhoè của mẫu màu thưa
Khoá màu (phông xanh, phông lục)	4:2:2, tốt nhất là 4:4:4	Ranh giới giữa người và phông là ranh giới MÀU; mẫu màu thưa làm mép tóc răng cưa, không tách sạch được
Đồ hoạ, chữ màu trên nền màu	4:4:4	Chữ mảnh có màu là tần số màu cao nhất có thể; mọi cấu trúc thưa hơn đều làm chữ nhoè viền

GHI NHỚ 8.6 — KINH NGHIỆM HIỆN TRƯỜNG

Khi biết trước sẽ phải khoá màu, nên ghi ra thiết bị ghi ngoài ở 4:2:2 10 bit qua cổng SDI hoặc HDMI thay vì dùng tệp 4:2:0 8 bit ghi trong thẻ của máy quay, vì chênh lệch về công sức ở hiện trường là nhỏ trong khi chênh lệch về chất lượng ở khâu hậu kỳ là rất lớn.

Ngược lại, với phóng sự thời sự thông thường, 4:2:0 8 bit là đủ; việc cố ghi ở 4:2:2 chỉ làm tăng tải cho thiết bị và tốn dung lượng thẻ mà không đem lại lợi ích tương xứng.

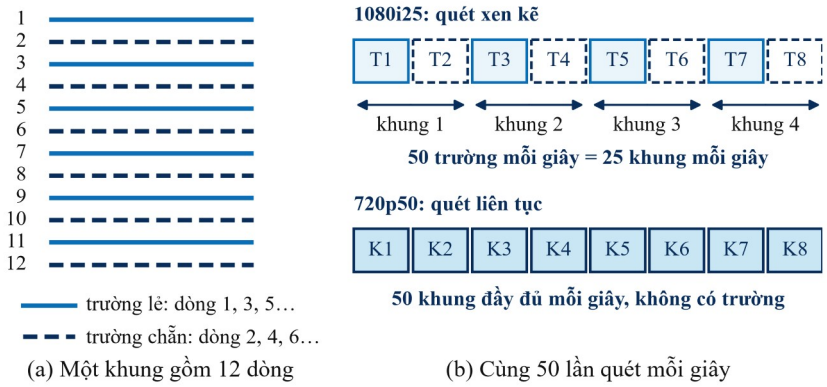
8.10 Quét xen kẽ và quét liên tục

8.10.1 Nguồn gốc: một cuộc đánh đổi

Quét xen kẽ ra đời từ những năm 1930 nhằm giải quyết một bài toán cụ thể: để mắt không thấy hiện tượng chớp, màn hình phải được làm mới ít nhất khoảng 50 lần mỗi giây, nhưng truyền đủ 50 ảnh hoàn chỉnh mỗi giây thì đòi hỏi băng thông gấp đôi, trong khi băng thông thời kỳ đó là tài nguyên đắt đỏ nhất.

Giải pháp là mỗi lần chỉ truyền một nửa số dòng, xen kẽ dòng lẻ và dòng chẵn. Mắt vẫn nhận đủ 50 lần cập nhật mỗi giây nên không thấy chớp, còn lượng dữ liệu chỉ tương đương 25 ảnh hoàn chỉnh.

$$f_{\text{fld}} = 2f_{\text{fr}} \quad ; \quad N_{\text{row/fld}} = \frac{N_{\text{row/fr}}}{2} \tag{8.15}$$



(a) Một khung gồm 12 dòng (b) Cùng 50 lần quét mỗi giây

Hình 8.20 Nguyên lý quét xen kẽ, và so sánh với quét liên tục ở cùng số lần quét mỗi giây

8.10.2 Trường và khung — sửa bốn chỗ nhầm

Phần lớn nhầm lẫn trong mục này bắt nguồn từ việc lẫn giữa *trường* (còn gọi là mành hay bán ảnh) và *khung* (còn gọi là ảnh hoàn chỉnh).

Bảng 8.17 Bốn chỗ cần đính chính về quét xen kẽ

#	Cách nói thường gặp (sai)	Thực tế
1	"50 bán ảnh (PAL) hoặc 60 bán ảnh (NTSC) tương đương với một ảnh hoàn chỉnh"	HAI bán ảnh mới thành một ảnh hoàn chỉnh. Hệ PAL: 50 trường mỗi giây bằng 25 khung mỗi giây
2	"Quét liên tục đòi hỏi phải có tần số gấp đôi bình thường tức là 1 giây ghi được 50 hoặc 60 hình"	Ở CÙNG tần số khung, quét liên tục không cần tần số gấp đôi: 1080i25 và 1080p25 có cùng số dòng mỗi giây. Chỉ khi so 1080i25 với 1080p50 mới gấp đôi
3	Ký hiệu "720/25p"	Định dạng này không tồn tại trong phát sóng. Chuẩn 720p sinh ra để chạy ở 50 hoặc 60 khung mỗi giây; nếu chỉ cần 25 khung thì người ta dùng 1080
4	Mô tả "quét mành" chỉ như sự lặp lại của quét dòng	Mành (trường) là TẬP các dòng lẻ hoặc tập các dòng chẵn của một khung, không phải một lần lặp của quét dòng

8.10.3 Đọc ký hiệu định dạng cho đúng



Ký hiệu	Kiểu quét	Mỗi giây	Hay dùng cho
1080i25	xen kẽ	25 khung = 50 trường	PAL, DVB-T2 Việt Nam
1080/50i	xen kẽ	50 trường	cùng nghĩa với 1080i25
1080p25	liên tục	25 khung	phim, phóng sự
1080p50	liên tục	50 khung	thể thao, gấp đôi dữ liệu
720p50	liên tục	50 khung	phát sóng châu Âu
2160p25	liên tục	25 khung	UHD, không có bản xen kẽ

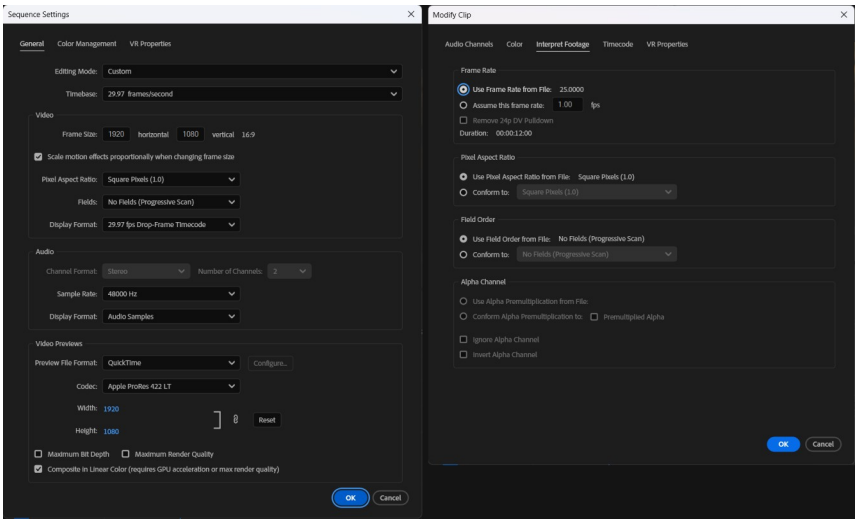
Hình 8.21 Cách đọc một ký hiệu định dạng, và các ký hiệu thường gặp

Hai cách viết cùng tồn tại và đều hợp lệ, chỉ khác nhau ở chỗ đếm trường hay đếm khung: ký hiệu 1080i25 đếm khung, còn ký hiệu 1080/50i đếm trường, và hai ký hiệu này có cùng ý nghĩa. Khi đọc tài liệu nước ngoài, cần xác định vị trí của chữ i hoặc p để biết con số bên cạnh đang đếm đại lượng nào.

8.10.4 Thứ tự trường

Trong một khung xen kẽ, thứ tự ghi của hai trường phụ thuộc vào định dạng, và việc ghi sai thứ tự sẽ làm chuyển động bị giật lùi từng nhịp nhỏ.

Hình 8.22 đặt cạnh nhau hai nơi khai báo thứ tự trường trong Premiere Pro: mục Fields của Sequence Settings và mục Field Order của hộp Interpret Footage, cả hai đều ở trạng thái quét liên tục (No Fields).



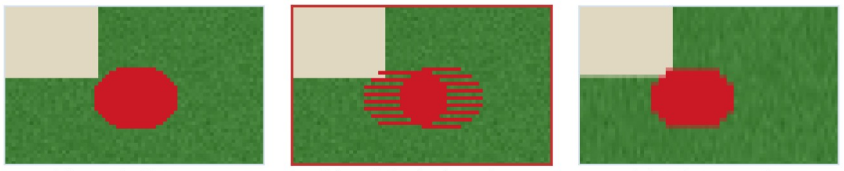
Hình 8.22 Sequence Settings (trái) và Interpret Footage (phải) trong Adobe Premiere Pro: khai báo thứ tự trường và chế độ quét (tệp mẫu tổng hợp, quét liên tục)

Nguồn: ảnh chụp màn hình do người soạn thực hiện (Adobe Premiere Pro 2026), 07/11/2026.

Bảng 8.18 Thứ tự trường của các định dạng thường gặp

Định dạng	Thứ tự trường	Ghi chú
SD PAL 576i (phát sóng)	Trường trên trước	Riêng tư liệu DV PAL ghi trường dưới trước; là nguồn lỗi hay gặp khi dựng tư liệu SD cũ
SD NTSC 480i	Trường dưới trước	Ngược với trực giác; đúng cho cả tư liệu phát sóng lẫn DV
HD 1080i	Trường trên trước	Khác với SD NTSC: chuyển tư liệu giữa hai hệ phải đổi thứ tự
Mọi định dạng quét liên tục	Không có khái niệm này	

8.10.5 Các lỗi hình do quét xen kẽ



(a) Quét liên tục
cả khung cùng một thời điểm

(b) Ghép hai trường
răng lược ở vật chuyển động

(c) Khử xen kẽ
hết răng lược, mất một ít độ nét

Hình 8.23 Răng lược ở vật chuyển động và kết quả sau khi khử xen kẽ

Ba lỗi dưới đây đều là hệ quả tất yếu của việc ghép hai thời điểm vào một khung chứ không phải do hỏng thiết bị.

- **Răng lược.** Vật chuyển động ngang xuất hiện ở hai vị trí khác nhau trên hai trường, nên khi ghép lại thành một khung, mép vật có dạng răng lược; vật chuyển động càng nhanh thì hiện tượng càng rõ.
- **Nhấp nháy đường ngang mảnh.** Một đường ngang chỉ dày một dòng chỉ xuất hiện ở một trong hai trường nên nhấp nháy 25 lần mỗi giây; vì lẽ đó, đồ hoạ truyền hình có quy tắc không dùng nét mảnh hơn hai điểm ảnh và không dùng hoa văn kẻ sọc ngang mau.
- **Hỏng khi phóng to hoặc xoay.** Mọi phép biến hình đều giả định các điểm ảnh thuộc cùng một thời điểm, nhưng với tư liệu xen kẽ giả định này không đúng, nên phóng to hay xoay đều làm lộ răng lược và không thể xoá được.

8.10.6 Khử xen kẽ

Bảng 8.19 Ba mức kỹ thuật khử xen kẽ

Phương pháp	Cách làm	Ưu và nhược
Nhân đôi dòng	Bỏ một trường, nhân đôi các dòng của trường còn lại	Nhanh, không bao giờ sai; nhưng mất một nửa độ phân giải dọc và hình trông thô
Ghép trường	Ghép nguyên hai trường thành một khung	Giữ đủ độ phân giải ở vùng ĐỨNG YÊN; nhưng để nguyên răng lược ở vùng chuyển động
Thích nghi chuyển động	Xét từng vùng ảnh: vùng đứng yên thì ghép trường, vùng chuyển động thì nội suy	Cho kết quả tốt nhất và là phương pháp mọi phần mềm dựng nghiêm túc dùng; đổi lại tốn nhiều công máy

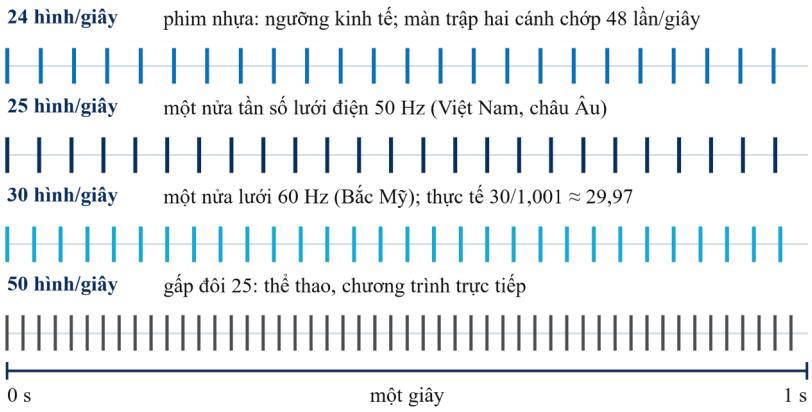
GHI NHỚ 8.7 — VÌ SAO VẪN PHẢI BIẾT PHẦN NÀY

Sản xuất mới hầu hết đã dùng quét liên tục nên có ý kiến cho rằng quét xen kẽ thuộc về quá khứ, song nhận định này không phù hợp với hoàn cảnh Việt Nam.

Thứ nhất, kho tư liệu của mọi đài đều ở dạng 576i và 1080i. Thứ hai, nhiều kênh vẫn phát ở 1080i để tiết kiệm băng thông trong bộ ghép kênh. Thứ ba, người mới vào làm việc tại đài phải xử lý tư liệu xen kẽ ngay trong tuần đầu tiên.

8.11 Tần số hình

8.11.1 Không con số nào là ngẫu nhiên



Hình 8.24 Bốn tần số hình thường gặp và lý do lịch sử của từng con số

Bảng 8.20 Nguồn gốc các tần số hình

Tần số	Nguồn gốc	Còn dùng ở đâu
24 hình/giây	Ngưỡng kinh tế của phim nhựa những năm 1920: chậm hơn thì thấy giật, nhanh hơn thì tốn phim. Máy chiếu dùng màn trập HAI CÁNH nên chớp 48 lần mỗi giây, đủ để mắt không thấy nhấp nháy	Điện ảnh trên toàn thế giới
25 hình/giây	Bằng một nửa tần số lưới điện 50 Hz của châu Âu và Việt Nam. Chia hết cho tần số nhấp nháy của đèn lưới nên không sinh vằn khi quay	Truyền hình Việt Nam, châu Âu, phần lớn châu Á và châu Phi
30 hình/giây	Bằng một nửa tần số lưới 60 Hz của Bắc Mỹ	Truyền hình Bắc Mỹ, Nhật Bản, một phần Nam Mỹ
50 và 60 hình/giây	Gấp đôi hai con số trên, dùng khi cần chuyển động thật mượt	Thể thao, chương trình trực tiếp, video trên mạng
100, 120 và cao hơn	Không phải tần số phát hành mà là tần số GHI, để phát chậm lại	Quay chậm; khi phát ở 25 hình/giây, tư liệu 100 hình/giây cho hình chậm bốn lần

8.11.2 Các tần số phân số

Trong tài liệu kỹ thuật thường gặp những con số khác thường: 29,97 thay vì 30; 23,976 thay vì 24; 59,94 thay vì 60. Đây không phải sai số đo mà là các giá trị chính xác, và chúng có nguồn gốc lịch sử cụ thể.

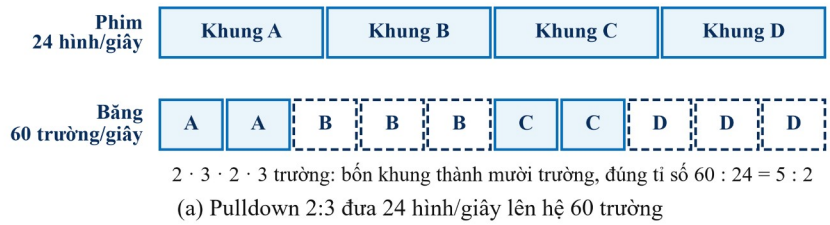
$$f_{\text{NTSC}} = \frac{30}{1{,}001} \approx 29{,}97 \quad ; \quad f_{\text{film}} = \frac{24}{1{,}001} \approx 23{,}976 \tag{8.16}$$

Khi hệ NTSC bổ sung màu vào năm 1953, sóng mang màu phải được đặt vào giữa phổ của tín hiệu độ chói mà không được gây nhiễu lên sóng mang tiếng, và cách duy nhất là hạ tần số khung xuống đúng một phần nghìn không phẩy một. Hệ PAL của châu Âu bổ sung màu theo cách khác nên giữ nguyên được 25 hình/giây, và Việt Nam thừa hưởng lợi thế này.

> **GHI NHỚ 8.8 — HỆ QUẢ THỰC TẾ VỚI NGƯỜI VIỆT NAM**
>
> Khi làm việc ở Việt Nam, tần số hình luôn là số nguyên, 25 hoặc 50, nên không phải xét đến 29,97 hay thời mã bỏ khung.
>
> Tuy nhiên, khi nhận tư liệu từ nước ngoài, nhất là tư liệu Mỹ hoặc tư liệu tải từ mạng, rất dễ gặp 29,97 hoặc 23,976. Việc đưa trực tiếp các tư liệu này vào dự án 25 hình/giây mà không chuyển đổi sẽ làm hình giật nhẹ theo chu kỳ; lỗi này khó thấy trên máy tính nhưng lộ rõ khi phát trên tivi.

8.11.3 Đưa 24 hình/giây lên tần số truyền hình

Phim điện ảnh được quay ở 24 hình/giây, còn truyền hình phát ở 25 hoặc 30 hình/giây; hai hệ thống này phải được kết nối với nhau, và cách kết nối khác nhau giữa hai hệ truyền hình.



2 · 3 · 2 · 3 trường: bốn khung thành mười trường, đúng tỉ số 60 : 24 = 5 : 2

(a) Pulldown 2:3 đưa 24 hình/giây lên hệ 60 trường

Tăng tốc 4 %	24 khung chiếu trong 0,96 giây; phim 100 phút còn 96 phút; cao độ tiếng nhích lên khoảng 0,7 nửa cung nếu không bù
Lặp một khung	lặp một khung sau mỗi 24 khung; giữ đúng độ dài và cao độ; chuyển động giật một nhịp nhỏ mỗi giây
Nội suy khung	phần mềm tạo khung mới; giữ độ dài, chuyển động mượt; tốn công máy, dễ biến dạng ở mép vật chuyển động nhanh

(b) Ba cách đưa 24 lên 25 hình/giây

Hình 8.25 Pulldown 2:3 cho hệ 60 trường, và ba cách đưa 24 lên 25 hình/giây

Ở hệ 60 trường mỗi giây, tỉ số giữa hai tần số là một phân số đơn giản nên có cách chuyển không làm thay đổi độ dài phim.

$$\frac{N_{\text{fld}}}{N_{\text{fr}}} = \frac{60}{24} = \frac{5}{2} \quad \Rightarrow \quad \text{2:3 pulldown} \tag{8.17}$$

Cách làm này gọi là pulldown 2:3: khung thứ nhất được ghi thành 2 trường, khung thứ hai thành 3 trường, và cứ thế xen kẽ, nhờ đó bốn khung phim tạo thành mười trường băng, đúng tỉ số 60 chia 24.

Ở hệ 50 trường của châu Âu và Việt Nam, tỉ số 50 chia 24 không phải là phân số đơn giản nên phải chọn một trong hai cách chuyển kém hoàn hảo. Cách phổ biến là tăng tốc 4 phần trăm, tức là chiếu 24 khung trong 0,96 giây thay vì trong 1 giây.

$$k = \frac{25}{24} \approx 1{,}0417 \quad ; \quad \Delta p = 12\log_2 k \approx 0{,}71 \tag{8.18}$$

Bảng 8.21 Ba cách đưa phim 24 hình/giây lên 25 hình/giây

Cách làm	Ưu điểm	Nhược điểm
Tăng tốc 4 %	Chuyển động hoàn toàn mượt, không lặp khung nào	Phim ngắn đi 4 %: bản 100 phút còn 96 phút. Cao độ tiếng nhích lên khoảng 0,7 nửa cung nếu không xử lý bù
Lặp một khung sau mỗi 24 khung	Giữ đúng độ dài và đúng cao độ tiếng	Chuyển động giật một nhịp rất nhỏ mỗi giây, người tinh mắt nhận ra
Nội suy khung bằng phần mềm	Giữ độ dài, chuyển động mượt	Đòi nhiều công máy; dễ sinh lỗi biến dạng ở mép vật chuyển động nhanh

CẢNH BÁO 8.5 — ĐỪNG QUÊN BÙ CAO ĐỘ TIẾNG

Tăng tốc 4 % làm cao độ tiếng cao lên khoảng 0,7 nửa cung. Với lời thoại, thay đổi này hầu như không được nhận ra; với nhạc, người có thính giác tốt nhận ra ngay vì bản nhạc quen thuộc trở nên "cao hơn" và nhanh hơn.

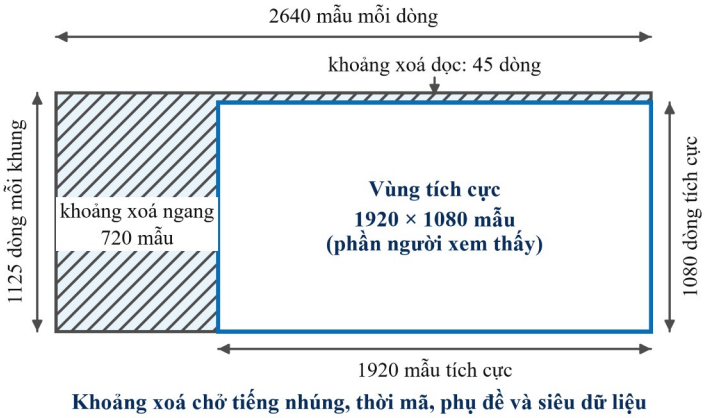
Mọi phần mềm dựng đều có chức năng bù cao độ khi đổi tốc độ, và nên bật chức năng này vì đó là một tiêu chí phân biệt bản dựng chuyên nghiệp với bản dựng nghiệp dư.

8.12 Đồng bộ, khoảng xoá và thời mã

8.12.1 Vì sao tín hiệu số vẫn giữ khoảng xoá

Trong hệ tương tự, tia điện tử quét xong một dòng phải quay về đầu dòng sau, và quét xong một khung phải quay về đầu khung sau. Trong hai quãng quay về ấy, tia bị tắt để không vẽ vệt sáng lên màn hình; các quãng này gọi là *khoảng xoá* ngang và khoảng xoá dọc.

Tín hiệu số không có tia điện tử nhưng vẫn giữ nguyên cấu trúc đó vì hai lý do: bảo đảm tương thích với thiết bị cũ trong dây chuyền, và tận dụng khoảng trống ấy để chở dữ liệu phụ.



Hình 8.26 Cấu trúc một khung hình HD 1080: vùng tích cực và khoảng xoá

Bảng 8.22 Số dòng và số mẫu của các định dạng chính

Định dạng	Dòng TỔNG	Dòng TÍCH CỰC	Mẫu mỗi dòng	Mẫu tích cực
SD PAL 576i25	625	576	864	720
SD NTSC 480i29,97	525	480	858	720
HD 1080i25	1125	1080	2640	1920
HD 1080p25	1125	1080	2640	1920
HD 720p50	750	720	1980	1280

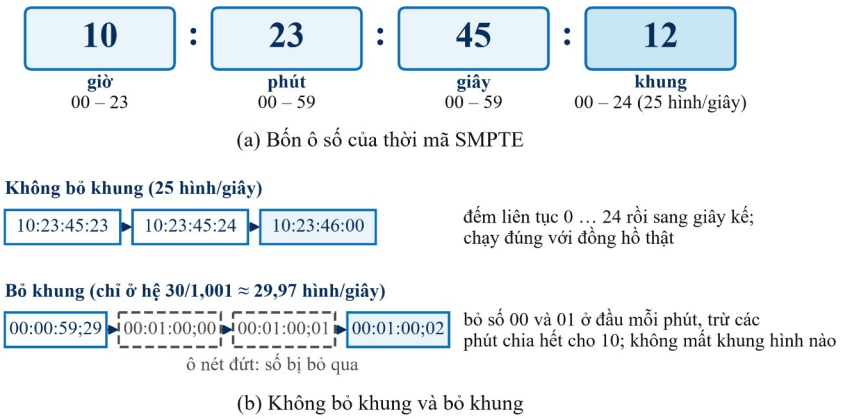
GHI NHỚ 8.9 — SỐ DÒNG CỦA ĐỊNH DẠNG HD 1080

Định dạng HD 1080 có tổng cộng 1125 dòng mỗi khung, trong đó 1080 dòng là dòng tích cực; chênh lệch 45 dòng chính là khoảng xoá dọc.

Khoảng xoá của tín hiệu số ngày nay chở khá nhiều thành phần: tới 16 kênh tiếng nhúng, thời mã, phụ đề, dữ liệu điều khiển máy quay và siêu dữ liệu cảnh quay. Nhờ đó, một sợi cáp SDI duy nhất mang được toàn bộ tín hiệu của một máy quay.

8.12.2 Thời mã

Thời mã là hệ thống đánh số từng khung hình để mọi thiết bị và mọi người trong dây chuyền cùng xác định thống nhất vị trí trong tư liệu [14].



(a) Bốn ô số của thời mã SMPTE

(b) Không bỏ khung và bỏ khung

Hình 8.27 Cấu trúc thời mã SMPTE, và phân biệt bỏ khung với không bỏ khung

Điều đầu tiên cần ghi nhớ là thời mã không phải một con số mà gồm bốn ô số đếm độc lập, trong đó ô cuối cùng đếm khung chứ không đếm phần trăm giây. Ở 25 hình/giây, ô khung chạy từ 00 tới 24 rồi mới sang giây kế tiếp.

Bảng 8.23 Các dạng thời mã

Dạng	Ghi ở đâu	Đặc điểm
LTC — thời mã dọc	Trên một rãnh tiếng riêng hoặc một dây riêng	Đọc được cả khi băng chạy nhanh hoặc chạy lùi; nhưng không đọc được khi đứng hình
VITC — thời mã trong khoảng xoá dọc	Nhét trong khoảng xoá dọc của chính tín hiệu hình	Đọc được khi đứng hình; nhưng không đọc được khi chạy nhanh
Thời mã trong tệp	Một rãnh riêng trong bộ chứa MXF, MOV, MP4	Dạng thường gặp nhất trong sản xuất trên tệp hiện nay

CẢNH BÁO 8.6 — BỎ KHUNG KHÔNG PHẢI LÀ MẤT HÌNH

Đây là nhầm lẫn phổ biến nhất về thời mã. Ở hệ 29,97 hình/giây, nếu đếm liên tục từ 0 tới 29 thì sau một giờ thời mã chạy nhanh hơn đồng hồ thật khoảng 3,6 giây; để khớp lại, người ta bỏ qua một số con số trong dãy đếm, cụ thể là bỏ 00 và 01 ở đầu mỗi phút, trừ các phút chia hết cho 10.

Không có khung hình nào bị mất, chỉ có con số bị bỏ qua, tương tự như toà nhà không có tầng 13 vẫn đủ số tầng.

Ở Việt Nam, khi làm việc ở 25 hình/giây thì thời mã luôn khớp với đồng hồ thật nên không cần dùng bỏ khung.

8.12.3 Quy ước thời mã bắt đầu

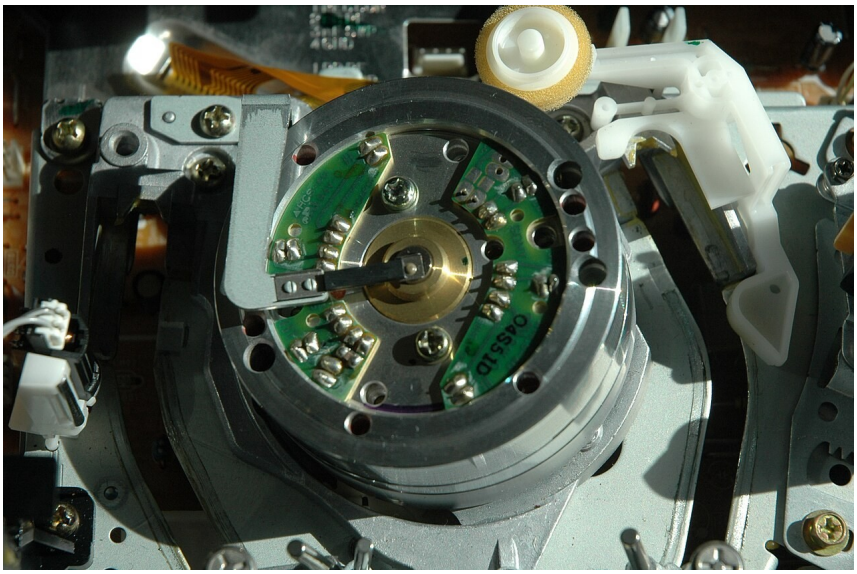
Hầu hết đài truyền hình yêu cầu bản giao nộp bắt đầu ở thời mã 10:00:00:00, với thanh màu chuẩn và tín hiệu âm thanh chuẩn đặt trước đó, nhằm chừa chỗ cho phần dẫn đầu mà không phải dùng tới thời mã âm. Đây là một trong bốn lỗi bị trả bài nhiều nhất và sẽ được trình bày kỹ ở mục 8.31.

8.13 Di sản analog

8.13.1 Vì sao vẫn cần một mục về analog

Truyền hình tương tự mặt đất ở Việt Nam đã tắt sóng hoàn toàn trong năm 2020. Cuốn sách vẫn dành một mục cho hệ này vì ba lý do cụ thể sau.

Hình 8.28 cho thấy trống đầu từ quay trong một máy ghi băng từ, bộ phận ghi tín hiệu hình tương tự lên băng bằng các rãnh quét xiên.



Hình 8.28 Trống đầu từ quay của máy ghi băng từ video

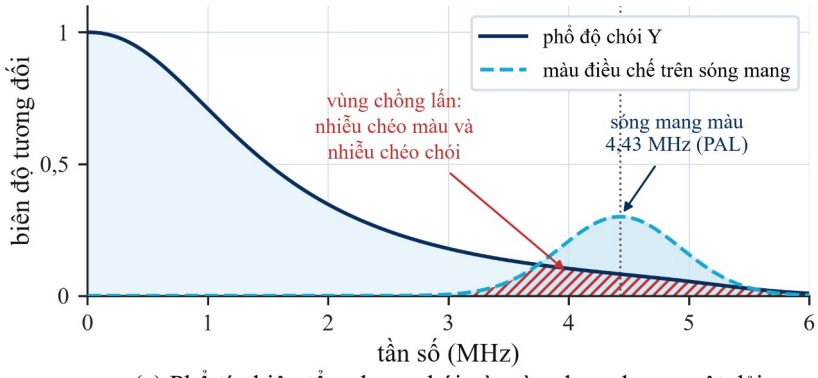
Nguồn: Erikfog, CC BY-SA 4.0, Wikimedia Commons (truy cập 05/11/2026).

– Toàn bộ những con số được dùng hiện nay, gồm 25 hình/giây, 720 × 576 mẫu, cấu trúc 4:2:2 và khoảng xoá 45 dòng, đều là di sản trực tiếp của hệ tương tự, và nếu không hiểu nguồn gốc thì người đọc chỉ có thể học thuộc.

– Thiết bị cũ trong các cơ sở đào tạo và các đài địa phương vẫn còn nhiều cổng tương tự, nên người làm nghề vẫn có lúc phải kết nối với các cổng này.
– Các bộ ký hiệu Y/U/V, Y/R−Y/B−Y và Y′/Cb/Cr xuất phát từ ba giai đoạn khác nhau của lịch sử này, nên thường bị nhầm lẫn với nhau.

8.13.2 Tín hiệu tổng hợp

Trong tín hiệu tổng hợp, độ chói và màu dùng chung một dải tần. Thông tin màu được điều chế lên một sóng mang đặt ở phần trên của dải, với tần số 4,43 MHz ở hệ PAL và 3,58 MHz ở hệ NTSC.



(a) Phổ tín hiệu tổng hợp: chói và màu chen chung một dải

Bộ ký hiệu	Thuộc về	Hệ số tỉ lệ và mức dịch
Y, R−Y, B−Y	thành phần tương tự	hiệu số màu thô, chưa tỉ lệ hoá; đi trên ba dây riêng
Y, U, V	PAL tương tự	đã tỉ lệ hoá để điều chế lên sóng mang màu mà không vượt biên độ cho phép
Y′, Cb, Cr	video số	tỉ lệ hoá theo BT.601 hoặc BT.709 và dịch mức để lượng tử hoá, tâm ở 128 (8 bit)

(b) Ba bộ ký hiệu hiệu số màu không đồng nghĩa

Hình 8.29 Phổ tín hiệu tổng hợp, và phân biệt ba bộ ký hiệu hiệu số màu

Cách ghép này tiết kiệm băng thông nhưng có một nhược điểm không thể khắc phục: hai loại thông tin chồng lấn nhau trong cùng một dải. Hệ quả là hai hiện tượng được người làm nghề gọi bằng tên riêng: *nhiễu chéo màu*, khi chi tiết mảnh của ảnh bị bộ giải mã hiểu nhầm thành màu và sinh ra vệt cầu vồng trên áo kẻ sọc; và *nhiễu chéo chói*, khi sóng mang màu lọt sang kênh độ chói và sinh ra các điểm lấm tấm trên vùng màu bão hoà.

Bảng 8.24 Ba hệ truyền hình màu tương tự

Hệ	Số dòng · tần số hình	Cách mã hoá màu	Vùng dùng
NTSC	525 dòng · 29,97 hình/giây	Điều chế cầu phương; nhạy với sai pha nên màu dễ lệch	Bắc Mỹ, Nhật Bản
PAL	625 dòng · 25 hình/giây	Như NTSC nhưng ĐẢO PHA thành phần V ở mỗi dòng, nhờ đó sai pha tự triệt tiêu	Châu Âu, Việt Nam, phần lớn châu Á
SECAM	625 dòng · 25 hình/giây	Truyền hai hiệu số màu LUÂN PHIÊN theo dòng, điều tần	Pháp, Nga và một số nước Đông Âu

GHI NHỚ 8.10 — SO SÁNH HAI HỆ PAL VÀ NTSC

Khác biệt thực sự giữa hai hệ là PAL có nhiều dòng hơn (625 so với 525) nên độ phân giải dọc cao hơn, còn NTSC có tần số hình cao hơn (29,97 so với 25) nên chuyển động ít giật hơn và ít nhấp nháy hơn. Đây là hai sự đánh đổi khác nhau chứ không phải hệ này vượt trội hệ kia.

8.13.3 Ba bộ ký hiệu KHÔNG đồng nghĩa

Nội dung này là điểm quan trọng nhất của mục và liên hệ trực tiếp với mục 7.17.4 của Chương 7.

Bảng 8.25 Phân biệt ba bộ ký hiệu hiệu số màu

Bộ ký hiệu	Thuộc về	Định nghĩa
Y, R−Y, B−Y	Tín hiệu THÀNH PHẦN tương tự	Hai hiệu số màu THÔ, chưa tỉ lệ hoá. Đi trên ba dây riêng, không chen chung dải tần
Y, U, V	Hệ PAL tương tự	Hai hiệu số màu ĐÃ TỈ LỆ HOÁ với hệ số riêng để điều chế lên sóng mang màu mà không làm tín hiệu tổng hợp vượt biên độ cho phép
Y′, Cb, Cr	Video SỐ	Hai hiệu số màu đã tỉ lệ hoá theo hệ số của BT.601 hoặc BT.709 VÀ dịch mức để lượng tử hoá, tâm ở 128 (8 bit) hoặc 512 (10 bit)

GHI NHỚ 8.11 — BA BỘ KÝ HIỆU KHÔNG ĐỒNG NGHĨA

Các bộ ký hiệu Y, Cr, Cb và Y, U, V không phải là ba tên gọi của cùng một tín hiệu thành phần.

Chúng khác nhau ở hệ số tỉ lệ và ở mức dịch; dùng lẫn các bộ ký hiệu này sẽ dẫn đến sai màu, đúng là loại lỗi mà hình ở mục 8.8 đã minh hoạ.

8.13.4 Các giao diện tương tự còn gặp

Bảng 8.26 Ba mức chất lượng của giao diện tương tự

Giao diện	Số dây tín hiệu	Cách truyền	Chất lượng
Tổng hợp (composite)	1	Chói và màu chen chung một dải	Thấp nhất; có nhiễu chéo màu và nhiễu chéo chói
S-Video (Y/C)	2	Tách riêng chói và màu thành hai dây	Khá hơn hẳn; hết nhiễu chéo nhưng màu vẫn bị nén băng thông
Thành phần (component)	3	Ba tín hiệu Y, R−Y, B−Y đi riêng	Tốt nhất trong các giao diện tương tự; là tiền thân trực tiếp của SDI

Ba mức này minh hoạ một quy luật xuyên suốt của kỹ thuật video: các thành phần càng được tách riêng thì chất lượng càng cao và số dây càng nhiều. Quy luật

ấy còn xuất hiện lại ở Phần E, khi so sánh cách bọc nguyên tín hiệu SDI vào gói IP với cách tách riêng ba luồng của chuẩn ST 2110.

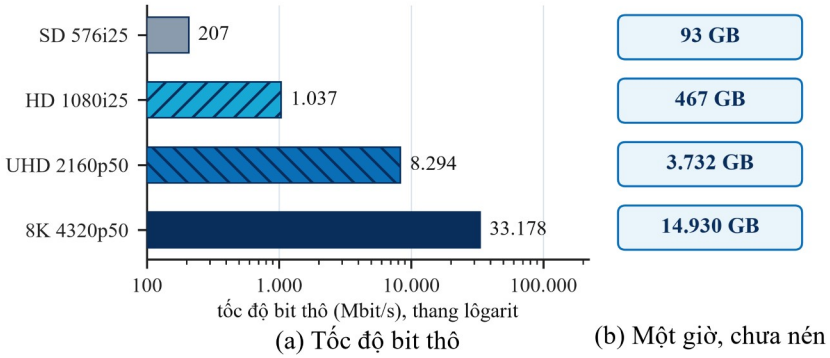
8.14 Vì sao phải nén

8.14.1 Con số tự nó trả lời

Câu hỏi vì sao phải nén có thể được trả lời ngắn gọn bằng cách tính tốc độ bit của tín hiệu chưa nén.

$$R_{\text{raw}} = W \times H \times f_{\text{fr}} \times n \times s \tag{8.19}$$

Trong công thức trên, W và H là số mẫu theo hai chiều, f là tần số hình, n là độ sâu bit và s là số mẫu trung bình mỗi điểm ảnh theo cấu trúc lấy mẫu con màu, đúng với đại lượng đã tính ở mục 8.9.



Hình 8.30 Tốc độ bit thô và dung lượng một giờ của bốn định dạng

Bảng 8.27 Tốc độ bit thô, tính cho 10 bit và cấu trúc 4:2:2

Định dạng	Phép tính	Tốc độ bit	Một giờ nặng
SD 576i25	720 × 576 × 25 × 10 × 2	207 Mbit/s	93 GB
HD 1080i25	1920 × 1080 × 25 × 10 × 2	1 037 Mbit/s	467 GB
UHD 2160p50	3840 × 2160 × 50 × 10 × 2	8 294 Mbit/s	3 732 GB
8K 4320p50	7680 × 4320 × 50 × 10 × 2	33 178 Mbit/s	14 930 GB

Giá trị ở hàng thứ hai đáng chú ý nhất: hơn một gigabit mỗi giây cho một tín hiệu HD thông thường, trong khi một kênh phát sóng số mặt đất ở Việt Nam chỉ có vài megabit mỗi giây cho mỗi chương trình. Chênh lệch này lên tới hàng trăm lần, và toàn bộ Phần C, Phần D trình bày các phương pháp bù đắp chênh lệch ấy.

8.14.2 Một phép tính hay bị nhân sai thừa số

GHI NHỚ 8.12 — TÍNH TỐC ĐỘ BIT CỦA TÍN HIỆU THÀNH PHẦN SỐ

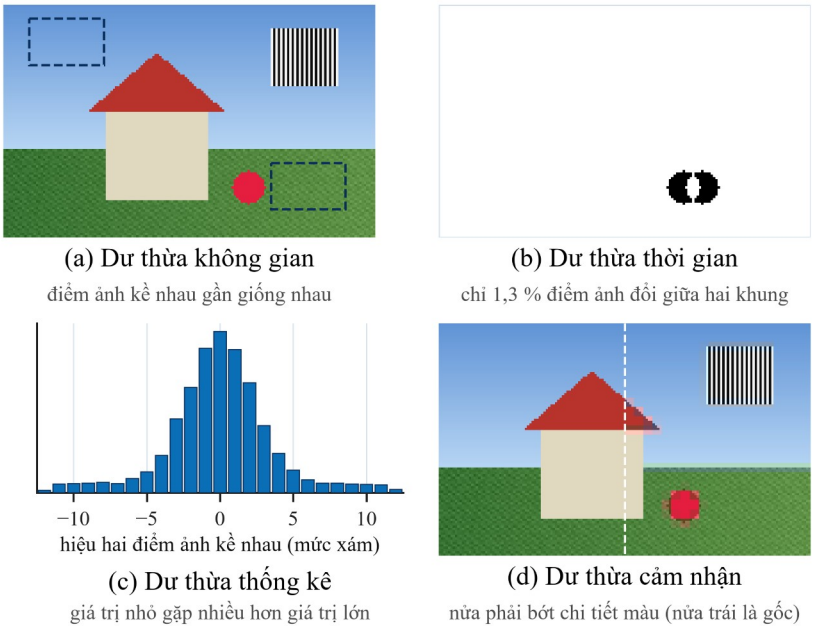
Tín hiệu thành phần số tiêu chuẩn SD có 864 × 625 điểm ảnh mỗi khung (kể cả khoảng xoá), 25 khung mỗi giây và 10 bit mỗi mẫu.

Tín hiệu 4:2:2 có trung bình hai mẫu mỗi điểm ảnh, gồm một mẫu độ chói cộng nửa mẫu Cb và nửa mẫu Cr, còn chỉ tín hiệu 4:4:4 mới có ba mẫu mỗi điểm ảnh.

Do đó, tốc độ lấy mẫu là 864 × 625 × 25 = 13,5 MHz, nhân với 10 bit và với 2 cho ra 270 Mbit/s, đúng bằng tốc độ của giao diện SD-SDI; với tám bit, kết quả tương ứng là 216 Mbit/s.

8.14.3 Bốn loại dư thừa

Tín hiệu hình có thể nén được vì chứa rất nhiều thông tin thừa. Có bốn loại dư thừa, và việc phân biệt được chúng là cơ sở để hiểu mọi codec.



Hình 8.31 Bốn loại dư thừa mà kỹ thuật nén khai thác

Bảng 8.28 Bốn loại dư thừa và cách khai thác

Loại	Bản chất	Khai thác bằng	Có tổn hao?
① Dư thừa KHÔNG GIAN	Các điểm ảnh cạnh nhau trong cùng một khung gần giống nhau	Biến đổi cosin rời rạc, dự đoán trong khung	Không, nếu chỉ biến đổi
② Dư thừa THỜI GIAN	Hai khung liên tiếp gần như trùng nhau; thường dưới 5 % điểm ảnh thật sự đổi	Bù chuyển động, khung P và khung B	Không, về nguyên tắc
③ Dư thừa THỐNG KÊ	Các giá trị nhỏ xuất hiện nhiều hơn giá trị lớn	Mã chạy dài, mã độ dài thay đổi, mã số học	KHÔNG — hoàn toàn khôi phục được
④ Dư thừa CẢM NHẬN	Mắt người không nhìn thấy một phần thông tin có trong tín hiệu	Lấy mẫu con màu, bảng lượng tử hoá có trọng số thị giác	CÓ — đây là chỗ mất mát thật sự

GHI NHỚ 8.13 — RANH GIỚI QUAN TRỌNG NHẤT CỦA CẢ PHẦN C

Ba loại đầu là dư thừa thực sự, vì sau khi loại bỏ vẫn có thể khôi phục nguyên vẹn từng bit; đó là nén không tổn hao.

Loại thứ tư không có tính chất này, vì nó đánh đổi thông tin lấy dung lượng dựa trên độ nhạy kém của mắt người; đó là nén có tổn hao.

Mọi codec video thực dụng đều khai thác cả bốn loại dư thừa, và vấn đề cần xác định không phải là có tổn hao hay không mà là mức tổn hao bao nhiêu thì còn chấp nhận được.

8.14.4 Nguyên lý đầu tiên: nén là bỏ phần mà bộ giải mã đoán được

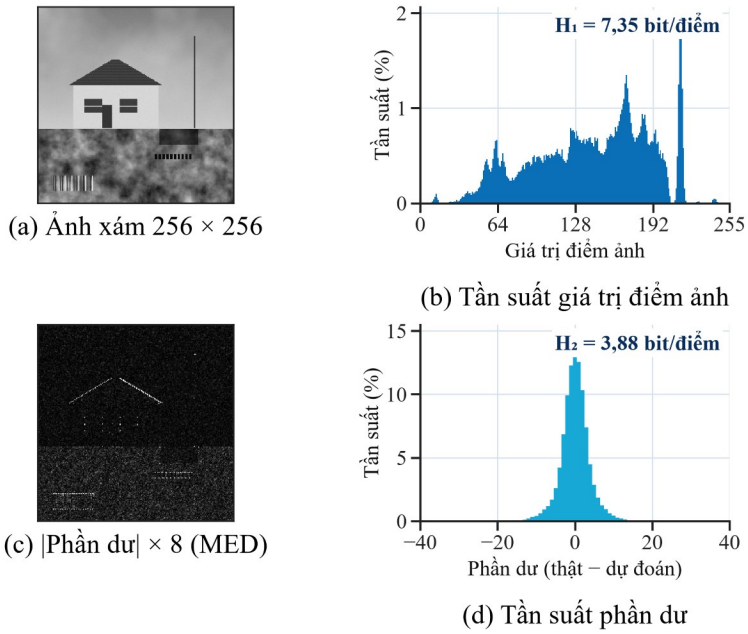
Bốn loại dư thừa nêu trên có thể quy về một nguyên lý chung, và khi hiểu nguyên lý ấy thì các kỹ thuật ở những mục sau trở nên dễ tiếp nhận. Một tín hiệu hình chứa hai thành phần lẫn vào nhau: phần mà bên nhận có thể *dự đoán* được từ những gì đã có, và phần *bất ngờ* không thể dự đoán. Chỉ phần bất ngờ mới thật sự mang thông tin, còn phần dự đoán được mà vẫn gửi đi là lãng phí. Về bản chất, nén gồm hai công việc: (1) dự đoán, tức là dựng một bản đoán từ dữ liệu mà bộ giải mã cũng đã có; (2) mã hoá sai số, tức là chỉ gửi hiệu giữa bản thật và bản đoán, đồng thời gửi hiệu ấy ở dạng gọn nhất.

Hiệu giữa bản thật và bản đoán gọn hơn bản thật vì khi dự đoán tốt, hiệu hầu hết gần bằng 0, mà giá trị nào xuất hiện thường xuyên thì được ghi bằng ít bit. Mức gọn này đo được bằng *entropy*: nếu một mẫu nhận các giá trị với xác suất p_i thì entropy là số bit trung bình tối thiểu để ghi mỗi mẫu bằng một bộ mã không tổn hao, theo Shannon [27]:

$$H = -\sum_i p_i \log_2 p_i \quad \text{(bit/m¤u)} \tag{8.20}$$

Một giá trị hiếm (p nhỏ) là một bất ngờ lớn và tốn nhiều bit, còn một giá trị hay gặp (p gần 1) hầu như không tốn bit nào. Hình 8.32 đo điều đó trên một ảnh mô phỏng: biểu đồ giá trị điểm ảnh trải rộng khắp thang 0 – 255 với entropy 7,35

bit mỗi điểm; nhưng khi dự đoán mỗi điểm ảnh từ các láng giềng đã biết rồi chỉ giữ sai số, biểu đồ co lại thành một đỉnh nhọn quanh 0 và entropy chỉ còn 3,88 bit mỗi điểm, bằng 53 % lúc đầu. Quá trình này không làm mất thông tin, vì bộ giải mã dựng lại đúng bản đoán rồi cộng sai số vào là thu được nguyên điểm ảnh ban đầu.



(a) Ảnh xám 256 × 256

(b) Tần suất giá trị điểm ảnh

(c) |Phần dư| × 8 (MED)

(d) Tần suất phần dư

Hình 8.32 Vì sao nén được: biểu đồ giá trị điểm ảnh trải rộng, còn biểu đồ sai số sau dự đoán co về một đỉnh nhọn quanh 0, entropy giảm rõ rệt

Xét lại bốn loại dư thừa theo cách tiếp cận này: dư thừa không gian là phần điểm ảnh có thể suy ra từ các điểm ảnh lân cận; dư thừa thời gian là phần khung sau có thể suy ra từ khung trước; dư thừa thống kê là các giá trị thường gặp được ghi bằng mã ngắn; còn dư thừa cảm nhận là phần thông tin thực sự tồn tại nhưng mắt người không cần đến nên được loại bỏ, và đây là khâu duy nhất gây mất mát.

GHI NHỚ 8.14 — BỐN VIỆC CỦA MỌI BỘ NÉN VIDEO

Dù là JPEG, H.264, HEVC hay AV1, mọi bộ nén đều thực hiện bốn việc theo cùng một thứ tự: dự đoán từ những gì đã có; biến đổi phần dư để dồn năng lượng về một vài hệ số; lượng tử hoá để loại bỏ phần mắt không cần (khâu duy nhất có tổn hao); mã hoá entropy để ghi gọn phần còn lại.

Các mục 8.16 và 8.17 lần lượt trình bày việc dự đoán trong một khung và giữa các khung. Các thế hệ codec khác nhau không ở thứ tự của các việc ấy mà ở mức độ tinh vi khi thực hiện từng việc, nội dung của mục tiếp theo.

8.14.5 Vì sao nén kỹ hơn thì tốn nhiều tính toán hơn

Ở cùng một mức chất lượng, bộ mã hoá tốt hơn dùng ít bit hơn. Mỗi thế hệ codec giữ nguyên bộ khung bốn việc ở trên nhưng thực hiện từng việc tinh hơn, với ý nghĩa cụ thể như sau:

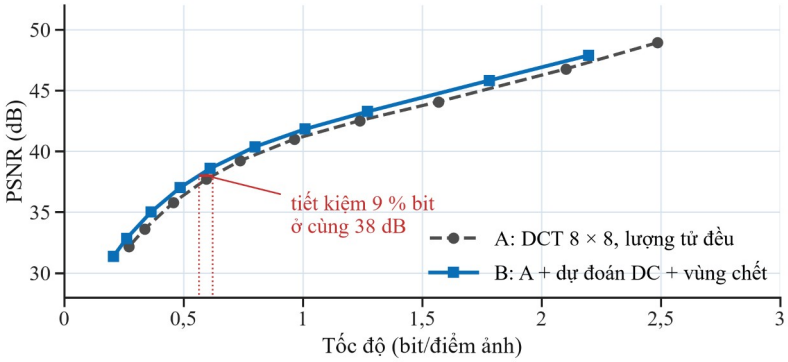
– **Dự đoán tinh hơn.** Bộ mã hoá thử nhiều hướng dự đoán trong khung, nhiều cỡ khối, nhiều vectơ chuyển động và nhiều khung tham chiếu hơn, nên bản dự đoán sát bản thật hơn và phần dư nhỏ hơn.

– **Biến đổi linh hoạt hơn.** Biến đổi cho phép nhiều cỡ khối, trong đó khối lớn dùng cho vùng phẳng và khối nhỏ dùng cho vùng có chi tiết, thay vì một cỡ cố định.
– **Mã hoá entropy thích nghi hơn.** Xác suất của từng ký hiệu được cập nhật theo ngữ cảnh đang mã hoá, nhờ đó tiến gần giới hạn entropy ở công thức (8.20).
– **Lọc và phân bổ bit hiệu quả hơn.** Bộ lọc trong vòng lặp loại bỏ lỗi khối trước khi khung được dùng làm tham chiếu, và bit được dồn vào những vùng mà mắt dễ nhận ra lỗi.

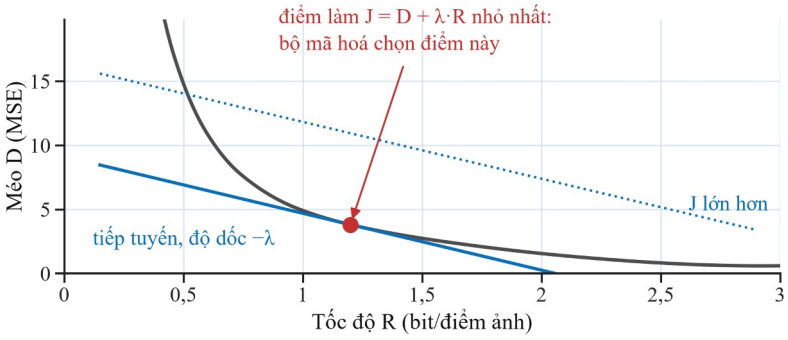
Cái giá phải trả là bộ mã hoá phải thử nhiều phương án hơn: để chọn cách tốt nhất cho mỗi khối, bộ mã hoá thử từng phương án, đo số bit R và độ méo D của phương án đó, rồi giữ phương án cho hàm giá sau đây nhỏ nhất:

$$J = D + \lambda R \tag{8.21}$$

Ở đây D là độ méo (thường là tổng bình phương sai khác so với bản gốc), R là số bit và λ là hệ số quy đổi bit sang độ méo: λ càng lớn thì bộ mã hoá càng ưu tiên tiết kiệm bit hơn là giữ chất lượng. Hình 8.33 đo đường tốc độ–méo của hai cấu hình bộ mã hoá ảnh thu nhỏ: công cụ tốt hơn dịch cả đường sang trái, nghĩa là cùng chất lượng thì cần ít bit hơn (ở mức 38 dB, tiết kiệm khoảng 9 % trong thí nghiệm này), và điểm mà bộ mã hoá chọn là điểm tiếp xúc với đường tiếp tuyến có độ dốc $-\lambda$.



(a) Hai bộ mã hoá, cùng ảnh, đo thật



(b) Đường D–R của A và tiếp tuyến

Hình 8.33 Đường tốc độ–méo đo trên một bộ mã hoá ảnh thu nhỏ: công cụ tốt hơn dịch đường sang trái, và bộ mã hoá chọn điểm làm $J = D + \lambda R$ nhỏ nhất

Số phương án phải thử tăng rất nhanh: chỉ riêng cách chia một khối 64 × 64 thành các khối nhỏ bằng cây tứ phân xuống tới 8 × 8 đã có 83 522 cách chia khác nhau. Gọi f(k) là số cách chia một khối còn chia được k lần thì f(0) = 1 và f(k) = 1 + f(k − 1)⁴ (hoặc không chia, hoặc chia làm bốn khối con và mỗi khối con tự chia), suy ra f(1) = 2, f(2) = 17, f(3) = 83 522. Mỗi cách chia lại phải thử nhiều hướng dự đoán, nhiều vectơ và nhiều cỡ biến đổi cho từng khối con. Vì vậy bộ mã hoá cho mức nén cao nhất chạy chậm hơn nhiều lần so với bộ mã hoá của thế hệ trước, trong khi bộ giải mã chỉ thực hiện theo các quyết định đã ghi trong luồng bit nên chỉ nặng hơn ít.

Bảng 8.29 Cùng bộ khung bốn việc, ba thế hệ làm tinh hơn dần

Công cụ	H.264 / AVC (2003)	H.265 / HEVC (2013)	H.266 / VVC (2020)
Khối mã hoá lớn nhất	16 × 16 (macroblock)	tới 64 × 64 (CTU, cây tứ phân)	tới 128 × 128 (CTU, cây đa dạng)
Số hướng dự đoán trong khung	9 hướng cho khối 4 × 4 và 8 × 8	35 chế độ	67 chế độ
Cỡ biến đổi	4 × 4 và 8 × 8	từ 4 × 4 tới 32 × 32	tới 64 × 64
Mã hoá entropy	CAVLC hoặc CABAC	CABAC	CABAC cải tiến
Bộ lọc trong vòng lặp	lọc khối	lọc khối và lọc thích nghi mẫu (SAO)	thêm lọc vòng thích nghi (ALF)
Tiết kiệm bit so với thế hệ trước	—	khoảng 40 – 50 % [29]	khoảng 40 – 50 % [30]
Chi phí mã hoá so với thế hệ trước	—	gấp nhiều lần	gấp nhiều lần nữa

Hai dòng cuối của bảng là các con số gần đúng, đo bằng phần mềm tham chiếu ở một số cấu hình thử và thay đổi theo nội dung cũng như mức chất lượng. Cần ghi nhận hai điều về xu hướng và quy luật: mỗi thế hệ mới tiết kiệm cỡ một nửa số bit và đổi lại phải tính toán nhiều hơn rất nhiều, đồng thời mức tiết kiệm thu được ở mỗi bước ngày càng khó đạt hơn. Chi phí không chỉ nằm ở tính toán: thiết bị cũ không giải mã được, và các thế hệ mới còn kéo theo vấn đề bản quyền, sẽ bàn ở mục 8.24. Các mục 8.21 – 8.23 mô tả cụ thể từng công cụ trong bảng.

8.15 Nén không tổn hao và nén có tổn hao

8.15.1 Hai định nghĩa chặt

- **Nén không tổn hao.** Sau khi giải nén, dữ liệu giống bản gốc từng bit. Tỉ số nén đạt được với ảnh thật thường chỉ từ 1,5 tới 2 lần, và đáng chú ý hơn là không thể cố định tốc độ bit vì tỉ số phụ thuộc nội dung; một cảnh nhiều chi tiết có thể gần như không nén được.
- **Nén có tổn hao.** Sau khi giải nén, dữ liệu khác bản gốc nhưng ở những chỗ mắt ít nhận ra, đổi lại tỉ số nén đạt hàng chục tới hàng trăm lần và tốc độ bit có thể đặt theo yêu cầu.

$$\mathrm{CR} = \frac{R_{\text{raw}}}{R_{\text{enc}}} \quad ; \quad S = \frac{R_{\text{enc}} \times t}{8} \tag{8.22}$$

GHI NHỚ 8.15 — DIGITAL BETACAM LÀ NÉN CÓ TỔN HAO NHẸ

Digital Betacam ghi ở tốc độ 128 Mbit/s, tỉ số nén khoảng 2:1 so với luồng 270 Mbit/s của ITU-R BT.601 và khoảng 1,6:1 so với 207 Mbit/s của phần hình tích cực ở 10 bit.

Hệ này dùng biến đổi cosin rời rạc có tổn hao ở mức rất nhẹ, được xếp vào loại "gần như không nhìn thấy khác biệt" chứ không thuộc nhóm nén không tổn hao.

Có một cách kiểm tra đơn giản: nén không tổn hao không bao giờ cố định được tốc độ bit. Do đó một định dạng ghi băng cam kết tốc độ bit không đổi chắc chắn phải có tổn hao, vì nếu không thì cảnh quá phức tạp sẽ không ghi vừa băng.

8.15.2 Đo chất lượng như thế nào

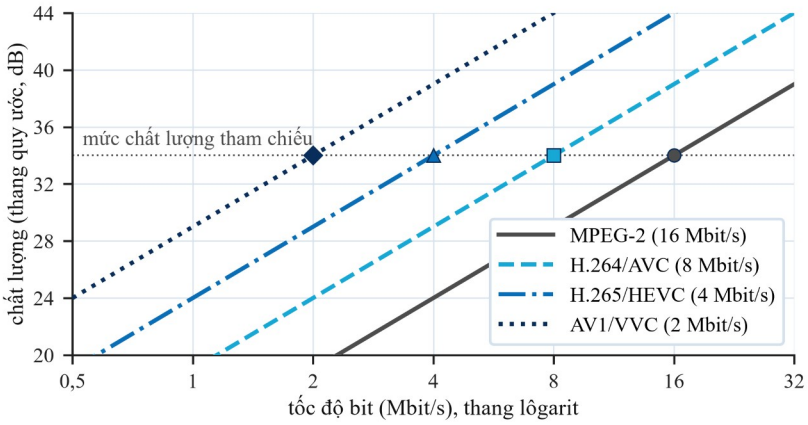
Khi nén có tổn hao thì tổn hao ấy cần được đo, và có hai cách: đo bằng máy và đo bằng người [5].

$$\mathrm{MSE} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[I(i,j) - \hat{I}(i,j)\right]^2 \tag{8.23}$$

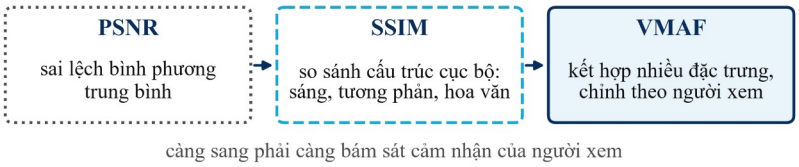
$$\mathrm{PSNR} = 10\log_{10}\frac{(2^n - 1)^2}{\mathrm{MSE}} \quad (\mathrm{dB}) \tag{8.24}$$

Bảng 8.30 Ba chỉ số đo chất lượng bằng máy

Chỉ số	Cách đo	Ưu điểm	Nhược điểm
PSNR	Từ sai lệch bình phương trung bình so với ảnh gốc	Rất dễ tính, có mặt trong mọi công cụ	KHÔNG khớp cảm nhận: ảnh mờ đều thường được điểm cao hơn ảnh sắc nét có ít nhiễu
SSIM	So sánh cấu trúc cục bộ: độ sáng, tương phản, hoa văn	Gần với mắt người hơn PSNR đáng kể	Vẫn không tính tới chuyển động và tới điều kiện xem
VMAF	Kết hợp nhiều đặc trưng, hiệu chỉnh theo điểm chấm của người xem thật	Bám sát cảm nhận nhất trong ba chỉ số; là chỉ số các nền tảng phát trực tuyến dùng nhiều nhất	Tốn công tính; kết quả phụ thuộc mô hình đã huấn luyện



(a) Mỗi thế hệ đạt cùng chất lượng ở khoảng một nửa tốc độ bit



(b) Ba chỉ số đo chất lượng bằng máy

Hình 8.34 Đường chất lượng – tốc độ bit của bốn thế hệ codec, và ba chỉ số đo chất lượng

Cách đo bằng người được quy định trong ITU-R BT.500: một nhóm người xem không biết trước nội dung chấm điểm theo thang năm bậc, trong phòng có điều kiện chiếu sáng và khoảng cách xem được chuẩn hoá. Phương pháp này vẫn là căn cứ cuối cùng khi hai chỉ số đo bằng máy cho kết quả trái ngược nhau.

8.15.3 Tổn hao thế hệ

Một hiện tượng thường bị người mới vào nghề bỏ qua là mã hoá lại một tệp đã nén có tổn hao sẽ làm chất lượng giảm thêm, dù tốc độ bit ở lần sau bằng hoặc cao hơn lần trước, bởi bộ mã hoá lần hai phải mã hoá cả những lỗi mà lần một đã sinh ra.

> **QUY TẮC 8.2 — QUY TẮC BA LẦN**
>
> Trong một dây chuyền sản xuất thông thường, tư liệu bị mã hoá lại ít nhất ba lần: khi máy quay ghi vào thẻ, khi xuất bản dựng và khi nền tảng phát hành mã hoá lại lần cuối, và mỗi lần đều làm mất đi một phần chất lượng.
>
> **Kết luận thực hành:** không nên dựng trực tiếp trên tệp đã nén sâu rồi xuất ra tệp nén sâu; nên dựng trên codec chỉ nén trong khung (mục 8.24) và chỉ nén sâu ở khâu cuối cùng.

8.15.4 Nén không tổn hao dùng ở đâu

Nén không tổn hao ít được dùng cho video, song không phải là không dùng, và có ba trường hợp mà nó thực sự cần thiết.

Bảng 8.31 Ba trường hợp dùng nén không tổn hao

Trường hợp	Vì sao cần không tổn hao	Định dạng thường dùng
Lưu trữ tư liệu gốc của viện phim và thư viện	Bản lưu trữ phải khôi phục được nguyên vẹn; mọi tổn hao đều tích luỹ qua các lần chuyển đổi về sau	FFV1 trong bộ chứa Matroska; là lựa chọn được nhiều thư viện phim trên thế giới dùng vì có tài liệu công khai và không ràng buộc bản quyền
Ảnh tĩnh và đồ hoạ trong dây chuyền sản xuất	Chữ và đường nét sắc là tần số cao nhất có thể; nén có tổn hao làm nhoè viền ngay ở mức nhẹ	PNG cho ảnh bitmap; TIFF không nén hoặc nén LZW cho bản trung gian
Bản trung gian giữa hai công đoạn hậu kỳ	Tránh cộng dồn tổn hao thế hệ khi tư liệu qua nhiều phần mềm	ProRes 4444 XQ hoặc DNxHR 444 — về danh nghĩa vẫn có tổn hao nhưng ở mức không đo được trong thực tế

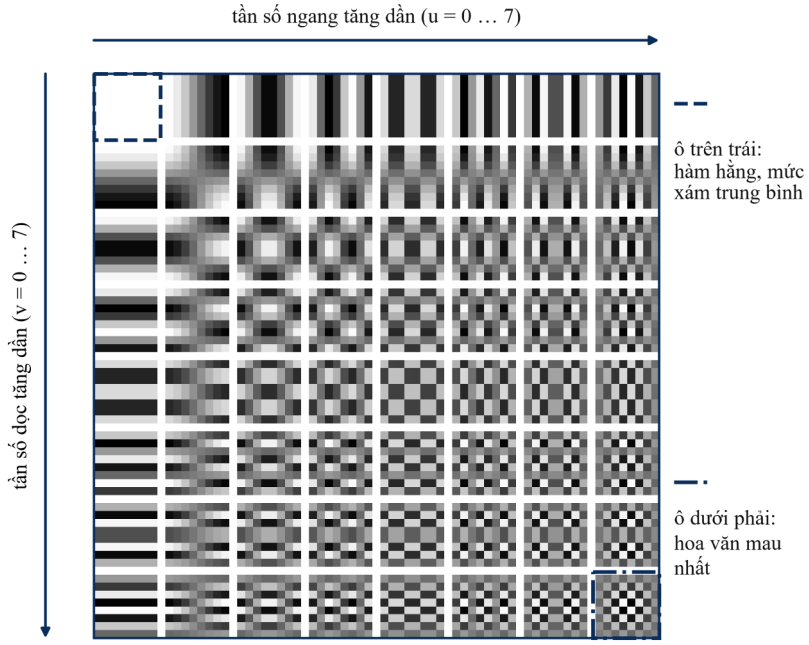
Điểm chung của cả ba trường hợp là đều thuộc khâu lưu trữ và trung gian chứ không phải khâu phát hành. Khi đến tay khán giả, nén có tổn hao luôn có ưu thế vì tỉ số nén hàng chục lần không thể đạt được bằng cách nào khác.

8.16 Nén trong khung

8.16.1 Vì sao đi qua miền tần số

Nén trong khung là nén từng ảnh một cách độc lập, không dùng đến các ảnh lân cận, và bước đầu tiên cũng là bước quyết định là chuyển ảnh sang miền tần số bằng biến đổi cosin rời rạc.

Lý do là trong miền điểm ảnh mọi giá trị đều quan trọng như nhau nên không có căn cứ để loại bỏ giá trị nào. Sau khi biến đổi, thông tin được phân loại: phần quan trọng dồn về vài hệ số ở góc trên trái, còn phần mắt ít nhạy tập trung ở góc dưới phải, nhờ đó có thể loại bỏ có chọn lọc.



(a) 64 hàm cơ sở của biến đổi DCT 8 × 8

Hình 8.35 64 hàm cơ sở của biến đổi cosin rời rạc 8 × 8

$$F(u,v)=\frac{1}{4}C_uC_v\sum_{x=0}^{7}\sum_{y=0}^{7}f(x,y)\cos\frac{(2x+1)u\pi}{16}\cos\frac{(2y+1)v\pi}{16} \tag{8.25}$$

$$C_w=\frac{1}{\sqrt{2}}\ (w=0)\quad;\quad C_w=1\ (w>0) \tag{8.26}$$

Phép biến đổi ngược đưa các hệ số trở lại thành giá trị điểm ảnh và, nếu không có bước lượng tử hoá xen giữa, khôi phục chính xác từng giá trị.

$$f(x,y)=\frac{1}{4}\sum_{u=0}^{7}\sum_{v=0}^{7}C_uC_vF(u,v)\cos\frac{(2x+1)u\pi}{16}\cos\frac{(2y+1)v\pi}{16} \tag{8.27}$$

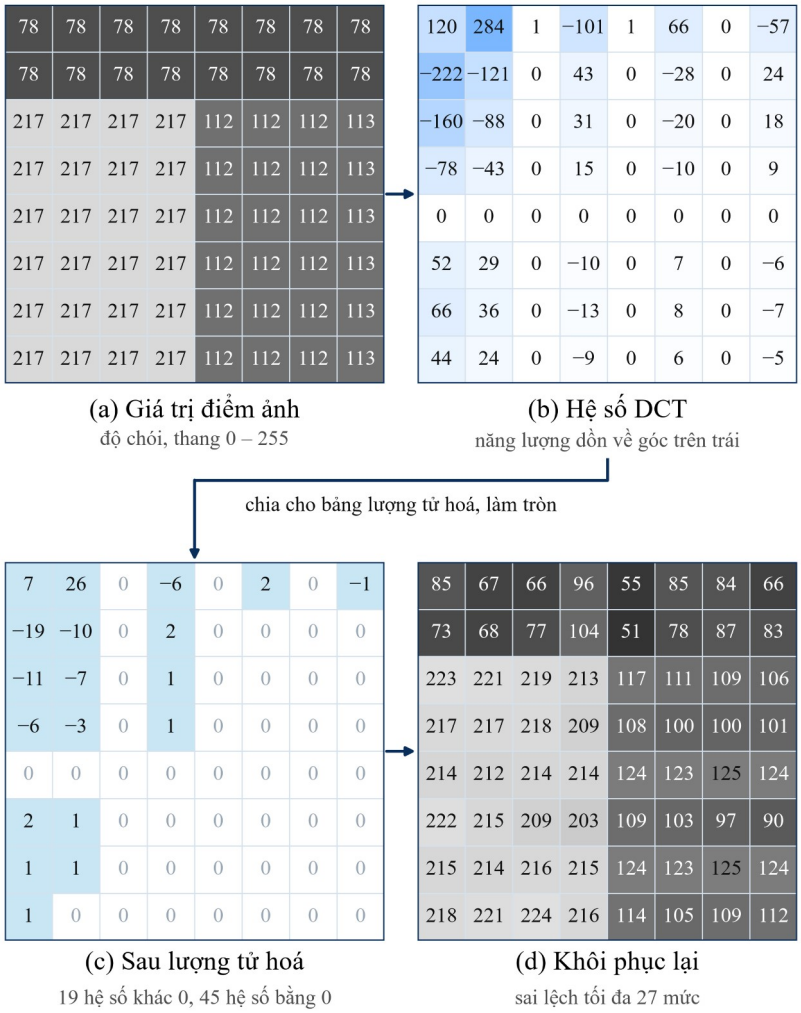
GHI NHỚ 8.16 — ĐIỀU CẦN NHỚ VỀ DCT

Bản thân biến đổi cosin rời rạc không làm mất thông tin nào, vì nó chỉ viết lại cùng một khối 64 số dưới dạng khác.

Toàn bộ mất mát của nén trong khung xảy ra ở bước sau đó, tức là lượng tử hoá các hệ số; việc nắm được điều này phân biệt người hiểu bản chất với người chỉ ghi nhớ chuỗi bước.

8.16.2 Chuỗi bước trên một khối thật

Để hiểu rõ, cần theo dõi một khối 8 × 8 thực tế đi qua toàn bộ chuỗi. Khối dưới đây được cắt từ chính bức ảnh mẫu dùng trong chương, tại vị trí có ranh giới giữa tường nhà và mặt đất.



Hình 8.36 Bốn bước của nén trong khung trên một khối 8 × 8 thật

$$F_q(u,v) = \text{round}\left(\frac{F(u,v)}{q \cdot Q(u,v)}\right) \tag{8.28}$$

Kết quả đáng chú ý là sau bước lượng tử hoá, 45 trong 64 hệ số trở thành 0, nghĩa là chỉ cần ghi 19 giá trị thay vì 64, trong khi sai lệch tối đa của khối khôi phục chỉ là vài chục mức trên thang 256 và tập trung ở đúng vị trí có ranh giới sắc nét, nơi mắt ít chú ý đến sai lệch nhất.

8.16.3 Bảng lượng tử hoá

Bảng lượng tử hoá là nơi kiến thức về thị giác ở Chương 7 được đưa trực tiếp vào kỹ thuật. Bảng này không đều: các ô ứng với tần số thấp chia cho số nhỏ nên giữ gần nguyên, còn các ô ứng với tần số cao chia cho số lớn nên bị loại bỏ gần hết.

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

(a) Bảng lượng tử hoá JPEG

15	52	0	−13	0	3	0	−2
−37	−20	0	5	0	−1	0	1
−23	−13	0	3	0	−1	0	1
−11	−5	0	1	0	0	0	0
0	0	0	0	0	0	0	0
4	2	0	0	0	0	0	0
3	1	0	0	0	0	0	0
1	1	0	0	0	0	0	0

(b) Bảng × 0,5: chất lượng cao

2	6	0	−2	0	0	0	0
−5	−3	0	1	0	0	0	0
−3	−2	0	0	0	0	0	0
−1	−1	0	0	0	0	0	0
0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

(c) Bảng × 4: chất lượng thấp

bảng nhân	số hệ số khác 0	sai lệch tối đa
× 0,5	24	11 mức
× 1	19	27 mức
× 4	11	55 mức

hàng × 1 chính là khối ở Hình 8.36

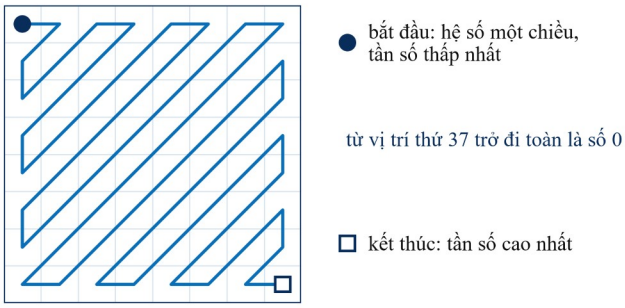
(d) Tóm tắt ba mức

Hình 8.37 Bảng lượng tử hoá độ chói của JPEG, và cùng một khối ở hai mức chất lượng

Việc nhân cả bảng với một hệ số chính là thao tác mà các phần mềm gọi là "thanh chất lượng": hệ số nhỏ giữ lại nhiều hệ số, cho tệp lớn và chất lượng cao, còn hệ số lớn cho kết quả ngược lại. Cơ chế này chỉ là một phép nhân.

8.16.4 Zig-zag và mã hoá entropy

Sau lượng tử hoá, khối chứa rất nhiều số 0 nhưng chúng nằm rải rác nếu đọc theo hàng. Phép quét zig-zag sắp xếp lại 64 hệ số thành một dãy đi từ tần số thấp tới tần số cao, nhờ đó các số 0 dồn hết về cuối.



(a) Đường quét zig-zag

7	26	−19	−11	−10	0	−6	0	−7	−6	0	−3	0	2	0	2
0	1	0	0	2	1	1	0	1	0	0	0	−1	0	0	0
0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

ô nền xanh nhạt: 28 số 0 liên tiếp cuối dãy, thay bằng một ký hiệu "hết"

(b) Dãy 64 hệ số của khối ở Hình 8.36, theo thứ tự quét

Hình 8.38 Đường quét zig-zag và dãy hệ số thật của khối ở hình trước

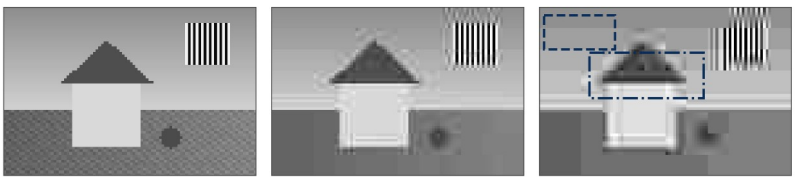
Bảng 8.32 Ba lớp mã hoá entropy

Kỹ thuật	Nguyên lý	Ví dụ
Mã chạy dài (RLC)	Thay một chuỗi giá trị giống nhau bằng cặp (giá trị, số lần lặp)	Hai mươi tám số 0 liên tiếp ở cuối dãy chỉ cần một ký hiệu "hết"
Mã độ dài thay đổi (VLC)	Giá trị hay gặp được gán mã NGẮN, giá trị hiếm gặp được gán mã DÀI	Bảng Huffman trong JPEG và MPEG-2; CAVLC trong H.264
Mã số học	Mã hoá cả dãy thành một số thực duy nhất, đạt gần giới hạn lý thuyết của entropy	CABAC trong H.264 và HEVC; hiệu quả hơn Huffman khoảng 10 – 15 %

Cả ba kỹ thuật này đều không tổn hao: chúng khai thác loại dư thừa thứ ba trong bảng ở mục 8.14 và có thể khôi phục chính xác từng hệ số đã đưa vào.

8.16.5 Nhận ra lỗi nén bằng mắt

Lỗi của nén trong khung có hình dạng đặc trưng nên người có kinh nghiệm nhận ra và gọi được tên ngay; đây là kỹ năng thực hành quan trọng, vì nhận ra các lỗi này cho biết tốc độ bit đang đặt quá thấp.



(a) Gốc (b) Nén vừa (c) Nén nặng

- - - vỡ khối: ô vuông 8 × 8 hiện lên ở vùng lẽ ra chuyển sắc êm (bầu trời)

- · - nhiễu muỗi: chấm lấm tấm bám quanh mép sắc nét (mép mái nhà)

Hình 8.39 Hai hình dạng đặc trưng của lỗi nén trong khung

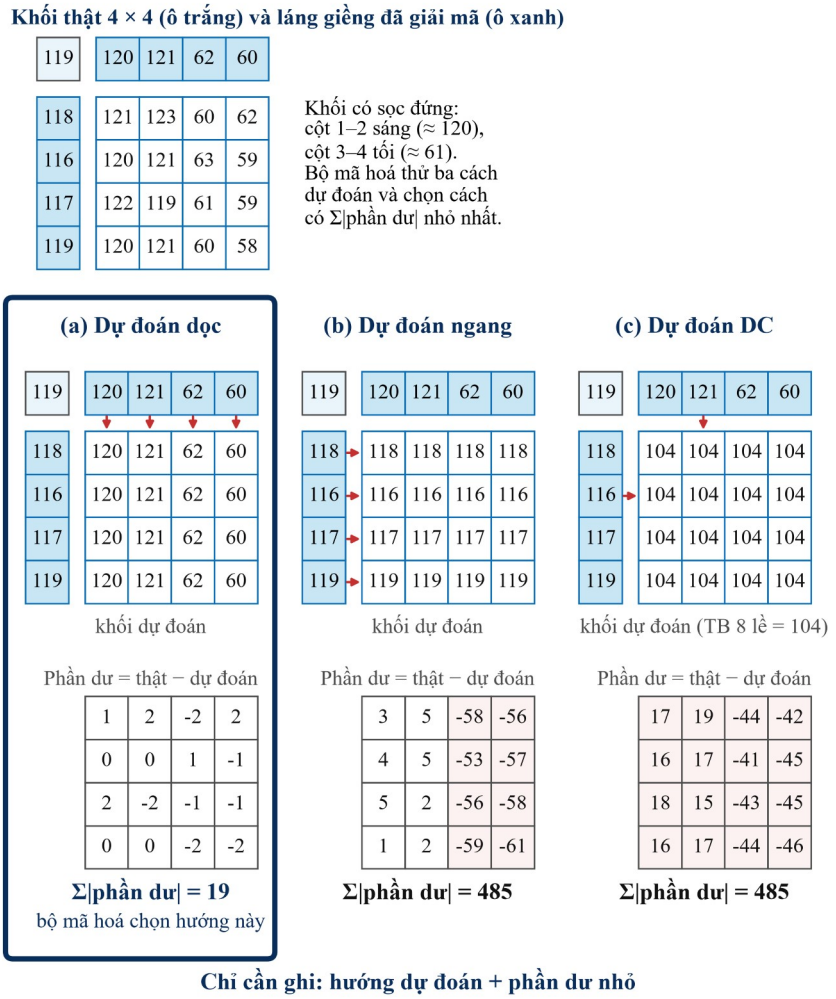
Bảng 8.33 Ba lỗi nén trong khung và cách nhận ra

Tên lỗi	Hình dạng	Nguyên nhân
Vỡ khối	Các ô vuông 8 × 8 hiện lên rõ ràng trong vùng lẽ ra phải chuyển sắc êm — bầu trời, phông vải, làn da	Mỗi khối được lượng tử hoá độc lập nên mức trung bình của các khối cạnh nhau lệch nhau một bậc
Nhiễu muỗi	Những chấm lấm tấm bám quanh mép sắc nét, nhất là quanh chữ và quanh cành lá trên nền trời	Mép sắc nét cần rất nhiều hệ số tần số cao; xoá bớt chúng sinh ra dao động dư quanh mép
Bệt chi tiết	Vùng có hoa văn mịn — cỏ, tóc, vải — biến thành mảng phẳng	Toàn bộ hệ số tần số cao của khối bị xoá về 0

8.16.6 Dự đoán trong khung — nén ảnh của các codec hiện đại

JPEG nén từng khối 8 × 8 độc lập bằng biến đổi cosin trực tiếp trên điểm ảnh. Từ H.264 trở đi, các bộ nén video thực hiện thêm một bước trước khi biến đổi, đúng theo nguyên lý ở mục 8.14.4: khối đang mã hoá được dự đoán từ các điểm ảnh lân cận đã giải mã (hàng phía trên và cột bên trái), rồi chỉ phần dư được biến đổi và lượng tử hoá. Cách làm này khai thác dư thừa không gian ngay trong miền điểm ảnh, bổ sung cho việc dồn năng lượng của biến đổi cosin.

Hình 8.40 minh hoạ với một khối 4 × 4 có sọc đứng và ba cách dự đoán đơn giản nhất: dự đoán dọc (chép điểm ảnh phía trên xuống), dự đoán ngang (chép điểm ảnh bên trái sang) và dự đoán DC (lấy trung bình các điểm ảnh lân cận). Với mỗi cách, bộ mã hoá tính khối dự đoán, trừ đi để có phần dư rồi cộng các giá trị tuyệt đối của phần dư, và cách cho tổng nhỏ nhất được chọn: ở khối này dự đoán dọc cho tổng 19, ngang 485 và DC 485 (hai tổng sau bằng nhau là do số liệu của khối này). Luồng bit chỉ cần ghi hướng đã chọn và phần dư nhỏ ấy.



Hình 8.40 Dự đoán trong khung trên khối 4 × 4: ba hướng dự đoán, phần dư của từng hướng, và hướng có tổng sai số nhỏ nhất được chọn

Một chi tiết quan trọng là các điểm ảnh lân cận dùng để dự đoán phải là điểm ảnh đã giải mã (tái tạo) chứ không phải điểm ảnh gốc, vì bộ giải mã chỉ có bản tái tạo; chi tiết này lặp lại ở quy mô lớn hơn trong nén liên khung (mục 8.17.5). Số hướng dự đoán tăng theo từng thế hệ codec: 9 hướng ở H.264, 35 ở HEVC và 67 ở VVC, như Hình 8.56 ở mục 8.21 sẽ cho thấy; càng nhiều hướng thì bản dự đoán càng sát và phần dư càng nhỏ, nhưng bộ mã hoá phải thử nhiều hơn.

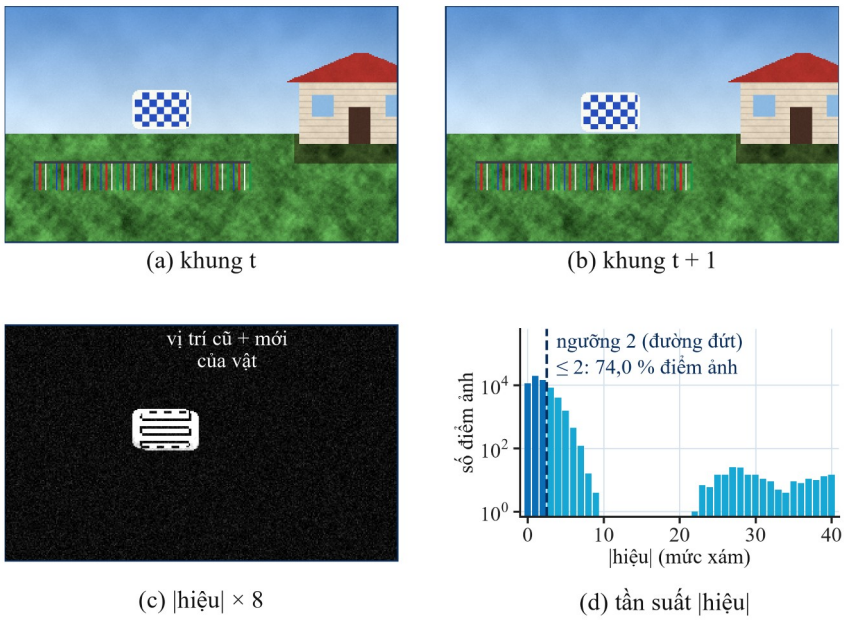
8.17 Nén liên khung

8.17.1 Từ ảnh sang video: hai khung liên tiếp giống nhau tới mức nào

Nén trong khung mới chỉ khai thác dư thừa không gian, trong khi một đoạn phim gồm hàng chục ảnh mỗi giây và các ảnh liền nhau gần như giống nhau. Vì vậy cần xác định bằng số mức độ giống nhau đó.

Hình 8.41 so sánh hai khung liên tiếp của một cảnh mô phỏng có một vật nhỏ chuyển động và nhiễu cảm biến. Ảnh hiệu |khung t + 1 − khung t| gần như đen, với 74,0 % điểm ảnh gần như không đổi (|hiệu| ≤ 2 trên thang 255). Phần còn lại chủ yếu là nhiễu cảm biến (hiệu của hai khung có nhiễu độc lập có độ lệch chuẩn gấp √2 lần nhiễu một khung, nên con số này do nhiễu quyết định nhiều hơn

do chuyển động) và hai vệt ở nơi vật chuyển động. Entropy của chính khung là 7,46 bit mỗi điểm, còn của ảnh hiệu chỉ 3,37 bit mỗi điểm. Đó chính là hiện tượng của mục 8.14.4 nhưng theo chiều thời gian, trong đó khung trước đóng vai trò bản dự đoán của khung sau.

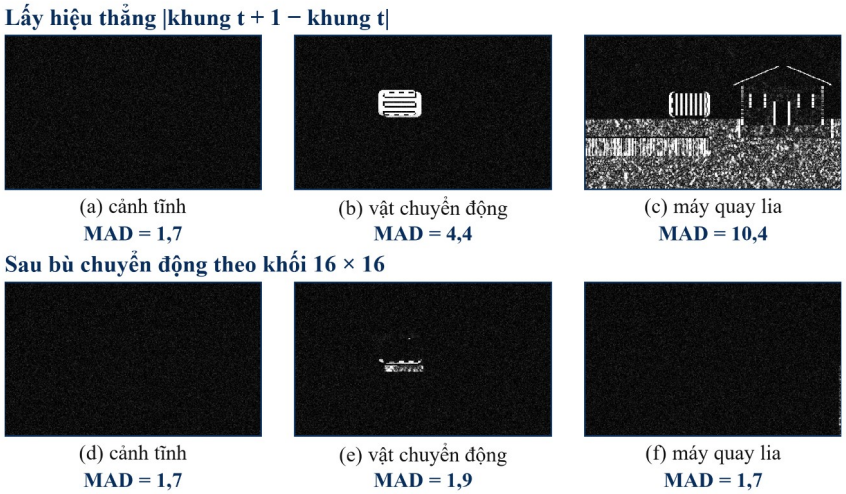


Hình 8.41 Hai khung liên tiếp của một cảnh mô phỏng: ảnh hiệu gần như đen và biểu đồ |hiệu| dồn về 0

Mức độ giống nhau phụ thuộc vào cảnh: phóng sự có máy quay đặt yên giống nhau nhiều hơn một cảnh hành động quay cầm tay. Tuy nhiên, ở phần lớn nội dung thực, phần ảnh thật sự đổi nội dung giữa hai khung liên tiếp rất nhỏ so với toàn khung, và toàn bộ nén liên khung được xây dựng từ nhận xét ấy qua ba cấp, từ thô đến tinh.

8.17.2 Cấp một — lấy hiệu thẳng giữa hai khung

Cách đơn giản nhất là dự đoán khung t bằng chính khung t − 1 và chỉ mã hoá hiệu. Hình 8.42 (hàng trên) thử cách này ở ba tình huống. Với cảnh tĩnh, ảnh hiệu chỉ còn nhiễu, độ lệch tuyệt đối trung bình (MAD) là 1,7, đạt kết quả tốt. Với vật chuyển động, ảnh hiệu để lại hai vệt, một ở vị trí cũ và một ở vị trí mới, và MAD tăng lên 4,4. Với máy quay lia, cả nền dịch đi vài điểm ảnh nên mọi cạnh trong cảnh đều sáng lên, MAD là 10,4. Như vậy lấy hiệu trực tiếp chỉ tốt khi không có gì chuyển động; ở cảnh có chuyển động, phần dư lớn ngang một ảnh mới, do đó cần một cách dự đoán tốt hơn: xác định vật đã dịch chuyển đến đâu.



Hình 8.42 Lấy hiệu thẳng giữa hai khung: tốt với cảnh tĩnh, hỏng khi có chuyển động; bù chuyển động theo khối đưa phần dư về mức nhiễu

8.17.3 Cấp hai — dò chuyển động theo khối

Ý tưởng là nếu vật chỉ dịch chỗ thì phần ảnh của nó ở khung t gần giống một phần ảnh ở khung t − 1, chỉ khác ở vị trí. Thay vì lấy hiệu tại cùng vị trí, bộ mã hoá tìm phần giống đó, theo bốn bước [31]:

1. **Chia khối.** Khung đang mã hoá được chia thành các khối, ví dụ 16 × 16 điểm ảnh; dùng khối thay vì từng điểm ảnh vì một vectơ cho mỗi điểm ảnh sẽ tốn nhiều bit hơn mức lợi thu được.
2. **Tìm trong cửa sổ.** Với mỗi khối, xét mọi vị trí (dx, dy) trong một cửa sổ tìm kiếm ± R điểm ảnh quanh vị trí cũ của khối ở khung tham chiếu.
3. **Đo độ giống bằng SAD.** Với mỗi vị trí ứng viên, tính tổng giá trị tuyệt đối của hiệu từng cặp điểm ảnh giữa khối hiện tại B_t và khối ứng viên của khung tham chiếu B_{t−1}:

$$\mathrm{SAD}(d_x, d_y) = \sum_{i,j} \left| B_t(i,j) - B_{t-1}(i + d_x,\ j + d_y) \right| \tag{8.29}$$

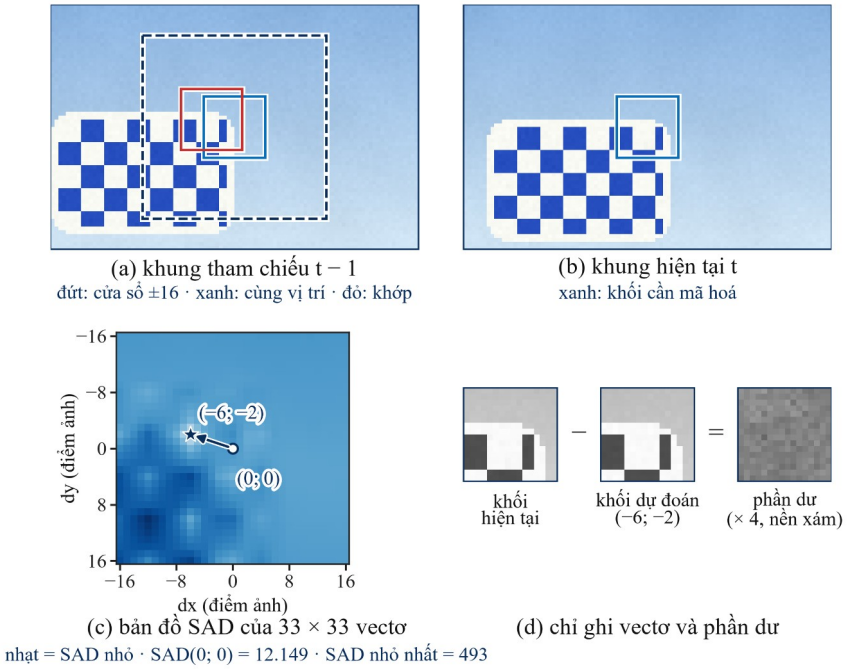
4. **Chọn vectơ.** Vị trí có SAD nhỏ nhất cho vectơ chuyển động (dx, dy); khối ở vị trí đó của khung tham chiếu là khối dự đoán, và hiệu giữa khối hiện tại và khối dự đoán là phần dư.

Một ví dụ nhỏ giúp minh hoạ từng con số. Một khối hiện tại 4 × 4 có cạnh đứng, trong đó mỗi dòng là 10, 10, 80, 80; dòng tương ứng của khung tham chiếu, mở rộng sang hai bên, là 10, 10, 10, 80, 80, 80, 80 (vật đã dịch một điểm ảnh). Xét ba vị trí theo phương ngang:

Bảng 8.34 SAD của ba vị trí ứng viên trong ví dụ khối 4 × 4

Vị trí ứng viên (dx)	Một dòng của khối tham chiếu	Hiệu tuyệt đối trên một dòng	SAD cả khối (4 dòng)
0	10 · 10 · 10 · 80	0 · 0 · 70 · 0	280
+1	10 · 10 · 80 · 80	0 · 0 · 0 · 0	0
+2	10 · 80 · 80 · 80	0 · 70 · 0 · 0	280

SAD nhỏ nhất bằng 0 ở dx = +1, do đó vectơ chuyển động là (+1, 0) và phần dư bằng 0, nghĩa là toàn bộ khối được mã hoá chỉ bằng một vectơ. Hình 8.43 thực hiện đúng phép tính đó cho một khối 16 × 16 thật của cảnh mô phỏng, với bản đồ SAD tính cho mọi vectơ trong cửa sổ ±16: điểm có SAD nhỏ nhất là vectơ (-6, -2) với SAD 493, so với 12149 nếu không bù (vectơ (0, 0)). Vectơ trỏ từ khối hiện tại về vị trí khối giống nó ở khung tham chiếu nên có chiều ngược với chiều dời của vật.



Hình 8.43 Dò chuyển động theo khối: cửa sổ tìm kiếm, bản đồ SAD với điểm cực tiểu, khối dự đoán và khối phần dư

Tìm kiếm vét cạn có chi phí tính toán rất lớn. Với cửa sổ ±16 có 33 × 33 = 1 089 vị trí ứng viên, mỗi vị trí cần 256 phép tính hiệu cho khối 16 × 16, tức khoảng 279 nghìn phép mỗi khối. Một khung 1920 × 1080 có khoảng 8 100 khối nên mỗi khung cần hơn hai tỉ phép, và ở 25 hình/giây con số này thành khoảng 56 tỉ phép mỗi giây chỉ riêng cho việc dò chuyển động. Vì vậy các bộ mã hoá thực tế dùng thuật toán tìm kiếm nhanh (thu hẹp dần, bắt đầu từ vectơ của khối láng giềng), và đây là một nguồn gốc của chi phí tính toán được nêu ở mục 8.14.5.

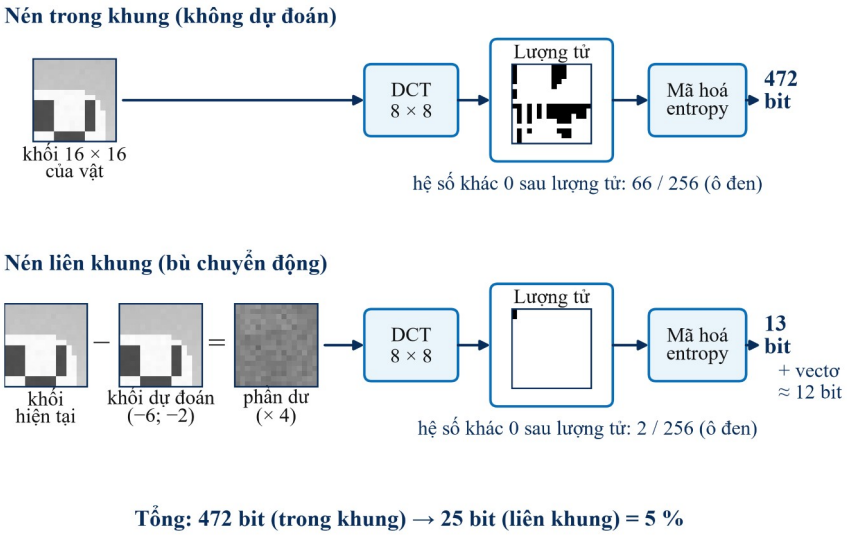
8.17.4 Cấp ba — phần dư đi qua chính bộ máy nén ảnh

Phần dư của một khối cũng là một khối điểm ảnh, nên nó đi qua đúng đường ống đã học ở mục 8.16: biến đổi cosin, lượng tử hoá, quét zig-zag và mã hoá entropy. Điểm khác biệt là phần dư đi vào đường ống ở dạng nhỏ và phẳng, do đó sau lượng tử hoá hầu hết hệ số về 0. Gọi P là khối dự đoán, B là khối thật, E là phần dư đã đi qua đường ống và được giải mã lại (Ê), bộ giải mã dựng lại khối như sau:

$$E = B - P \quad ; \quad \hat{B} = P + \hat{E} \tag{8.30}$$

Hình 8.44 so sánh hai cách mã hoá cùng một khối 16 × 16 của vật chuyển động. Nén trong khung (không dự đoán) cần 472 bit với 66 hệ số khác 0, còn nén

liên khung, trong đó khối được trừ đi khối dự đoán đã bù chuyển động rồi mới đi qua đường ống, chỉ cần 25 bit, gồm 12 bit cho vectơ chuyển động và 13 bit cho phần dư với 2 hệ số khác 0, tức khoảng 5,3 % số bit của nén trong khung. (Các số này tính trên bộ mã hoá thu nhỏ của hình, chỉ dùng để so sánh tương đối giữa hai cách mã hoá và không phải số bit của một chuẩn cụ thể.)



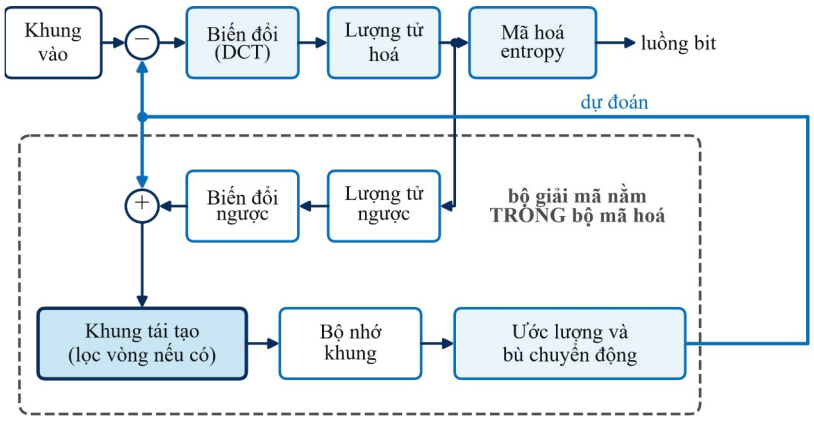
Hình 8.44 Phần dư đi qua cùng đường ống DCT, lượng tử hoá và mã hoá entropy như ảnh tĩnh: nén liên khung cần ít hệ số và ít bit hơn nén trong khung

Điều này cho thấy quan hệ giữa nén ảnh và nén video: nén video không thay thế nén ảnh mà đặt thêm một tầng dự đoán phía trước. Chuỗi bốn việc ở mục 8.14.4 vẫn được giữ nguyên, chỉ khác ở chỗ bản đoán nay được lấy từ khung đã giải mã trước đó, không chỉ từ láng giềng trong cùng khung.

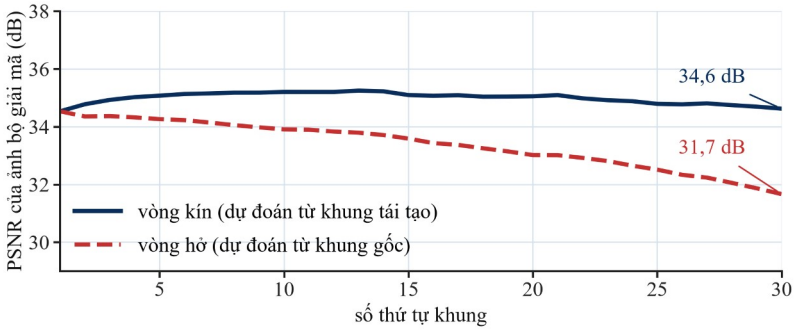
8.17.5 Vòng phản hồi — vì sao bộ mã hoá phải chứa một bộ giải mã

Một chi tiết tinh tế quyết định thiết kế của mọi bộ mã hoá video là bộ giải mã không có khung gốc mà chỉ có khung đã giải mã, vốn đã bị lượng tử hoá nên khác khung gốc đôi chút. Nếu bộ mã hoá dự đoán từ khung gốc trong khi bộ giải mã chỉ có thể dự đoán từ khung đã giải mã thì hai bên dùng hai bản đoán khác nhau. Sai khác nhỏ ấy cộng dồn qua từng khung và làm hình ảnh trôi dần khỏi bản gốc.

Cách khắc phục là buộc bộ mã hoá chỉ dùng những gì bộ giải mã có: bên trong bộ mã hoá đặt một bộ giải mã thu nhỏ (lượng tử ngược, biến đổi ngược, cộng với bản đoán), và bản đoán của khung kế tiếp được dựng từ khung tái tạo đó. Hình 8.45 (a) vẽ sơ đồ khối, còn hình (b) mô phỏng 30 khung để so sánh hai cách: với dự đoán từ khung gốc ("vòng hở"), PSNR của ảnh giải mã giảm xuống 31,7 dB ở khung cuối và tiếp tục xấu đi theo từng khung, trong khi với dự đoán từ khung tái tạo ("vòng kín"), PSNR ổn định ở 34,6 dB. Chênh lệch trong mô phỏng ngắn này chỉ cỡ 3 dB, nhưng sai lệch cộng dồn nên chuỗi càng dài thì hình càng trôi xa bản gốc.



(a) sơ đồ khối của bộ mã hoá lai



(b) mô phỏng 30 khung: lỗi cộng dồn nếu bộ mã hoá không giải mã

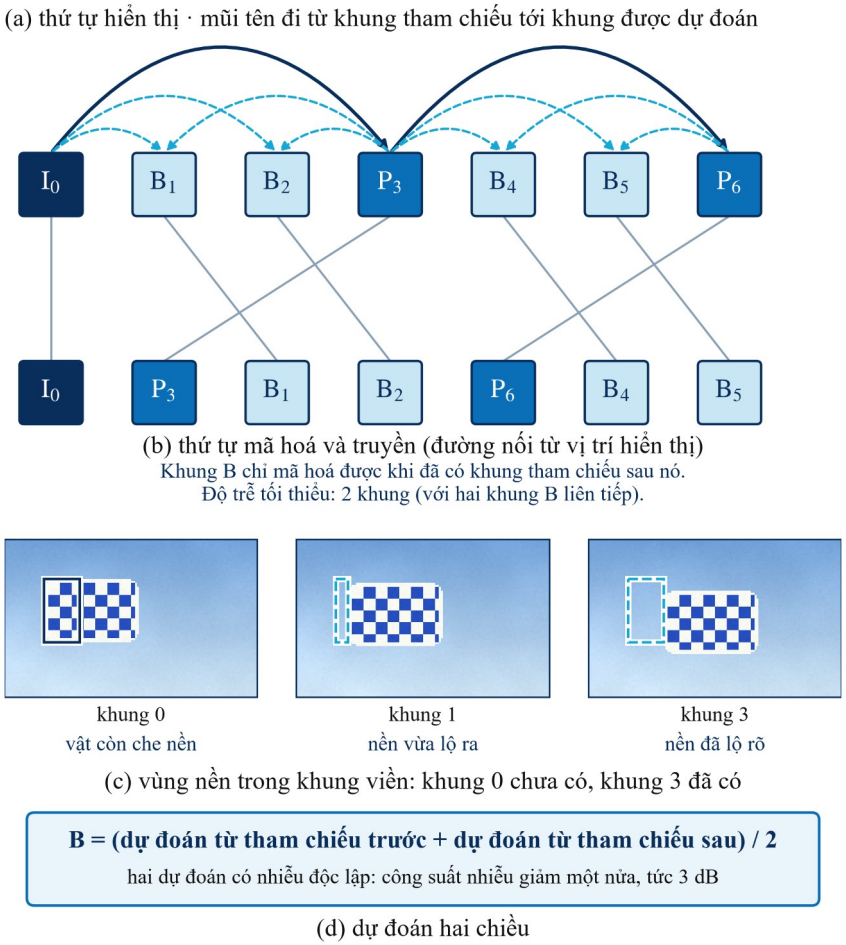
Hình 8.45 Vòng phản hồi của bộ mã hoá lai (a) và hậu quả khi dự đoán từ khung gốc thay vì khung tái tạo (b)

Từ đó rút ra hai hệ quả thực tế. Thứ nhất, chuẩn nén chỉ quy định bộ giải mã, gồm cú pháp của luồng bit và cách dựng lại hình, mà không quy định bộ mã hoá; bộ giải mã của mọi nhà sản xuất phải dựng ra cùng một kết quả đến từng bit nên mới dùng chung được một tệp, còn bộ mã hoá được tự do cải tiến, và đó là lý do cùng chuẩn H.264 vẫn có bộ mã hoá tốt và bộ mã hoá kém. Thứ hai, bộ giải mã phải cho kết quả chính xác từng bit chứ không chỉ gần đúng, vì sai một bit trong khung tham chiếu sẽ lan sang các khung sau.

8.17.6 Khung B — nhìn cả tương lai

Dự đoán chỉ từ quá khứ có một hạn chế: khi vật chuyển động che hoặc để lộ nền, phần nền vừa lộ ra không có trong khung trước nên không có cơ sở để dự đoán, nhưng phần này thường có trong một khung ở tương lai. Khung B (Bi-directional) giải quyết đúng vấn đề đó vì được phép tham chiếu cả khung quá khứ lẫn khung tương lai.

Hình 8.46 cho thấy hai điều. Điều thứ nhất là vùng nền "vừa lộ ra" mà khung quá khứ không có nhưng khung tương lai có. Điều thứ hai là hệ quả về thứ tự: muốn mã hoá khung B thì phải có khung tương lai trước, nên thứ tự mã hoá và truyền khác thứ tự hiển thị. Chẳng hạn, các khung hiển thị theo thứ tự $I_0\ B_1\ B_2\ P_3$ được mã hoá theo thứ tự $I_0\ P_3\ B_1\ B_2$. Bộ giải mã nhận khung theo thứ tự truyền, giữ lại khung tham chiếu rồi sắp xếp về thứ tự hiển thị, vì vậy khung B gây ra độ trễ (mục 8.17.11).



Hình 8.46 Khung B tham chiếu cả quá khứ lẫn tương lai: thứ tự hiển thị khác thứ tự mã hoá, và vùng nền vừa lộ ra chỉ có ở khung tương lai

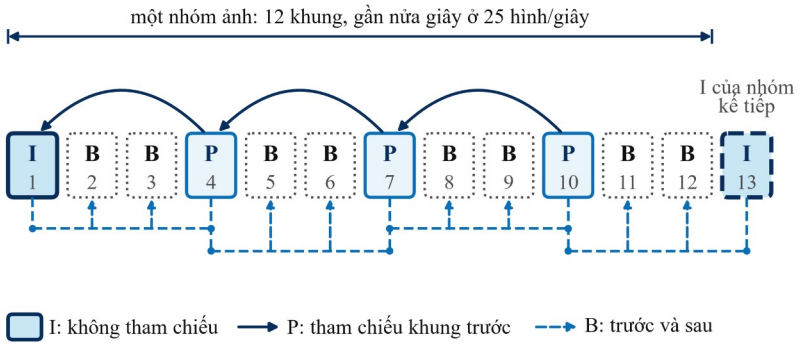
Khung B còn có một lợi thế thứ hai: việc lấy trung bình hai bản đoán (một từ quá khứ, một từ tương lai) làm giảm nhiễu. Nếu nhiễu trong hai bản đoán độc lập với nhau thì trung bình của chúng có công suất nhiễu bằng một nửa, tức tốt hơn 3 dB. Đó là một lý do khiến khung B vừa nhỏ nhất vừa chính xác nhất trong ba loại khung.

8.17.7 Ba loại khung

Nén trong khung mới chỉ khai thác dư thừa không gian, trong khi dư thừa thời gian vốn lớn hơn nhiều, nên muốn khai thác loại dư thừa này thì phải cho các khung tham chiếu lẫn nhau.

Bảng 8.35 Ba loại khung

Loại	Tham chiếu	Kích thước tương đối	Đặc điểm
Khung I (Intra)	Không tham chiếu khung nào	100	Tự giải mã được một mình. Là điểm duy nhất có thể CẮT và TUA tới
Khung P (Predicted)	Tham chiếu khung đã giải mã TRƯỚC nó	≈ 34	Chỉ ghi phần khác so với khung tham chiếu
Khung B (Bi-directional)	Tham chiếu CẢ khung trước lẫn khung sau	≈ 13	Nhỏ nhất, nhưng gây độ trễ vì phải chờ khung sau mới giải mã được



Hình 8.47 Cấu trúc một nhóm ảnh 12 khung với quan hệ tham chiếu

8.17.8 Nhóm ảnh

Chuỗi khung lặp lại theo chu kỳ và mỗi chu kỳ bắt đầu bằng một khung I; chu kỳ ấy được gọi là nhóm ảnh. Nhóm ảnh được mô tả bằng hai tham số là độ dài nhóm và khoảng cách giữa hai khung tham chiếu liên tiếp.

Bảng 8.36 Chọn độ dài nhóm ảnh theo mục đích

Độ dài nhóm	Ưu điểm	Nhược điểm	Dùng khi
1 khung — chỉ khung I	Cắt được ở mọi khung; dựng rất nhẹ máy	Tệp lớn nhất	Codec dựng: ProRes, DNxHR, XAVC-I
12 – 15 khung	Cân bằng giữa hiệu quả nén và khả năng cắt	Chỉ cắt được mỗi nửa giây	Phát sóng DVB-T2, đĩa DVD
48 – 250 khung	Nén hiệu quả nhất, tệp nhỏ nhất	Tua và dò kênh chậm; cắt phải mã hoá lại	Phát trực tuyến, lưu trữ nội dung đã hoàn thiện

GHI NHỚ 8.17 — NHÓM ẢNH ĐÓNG VÀ NHÓM ẢNH MỞ

Nhóm đóng: không khung nào trong nhóm tham chiếu ra ngoài nhóm, nhờ vậy có thể cắt và ghép nối ở ranh giới nhóm mà không cần mã hoá lại. Vì lý do này, chuẩn giao nộp của các đài thường bắt buộc nhóm đóng.

Nhóm mở: khung B ở đầu nhóm được phép tham chiếu khung I của nhóm trước, nên nén hiệu quả hơn đôi chút nhưng không thể cắt ở ranh giới nhóm.

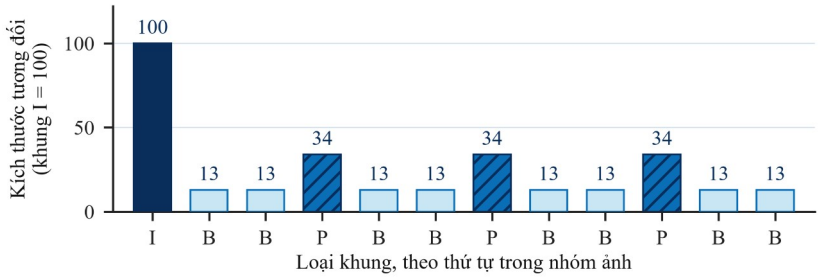
8.17.9 Các mức tinh vi của bù chuyển động

Phần trên mô tả nguyên lý với khối 16 × 16 và vectơ chính xác tới điểm ảnh, còn các chuẩn khác nhau ở mức độ tinh vi trong việc thực hiện chính ý tưởng ấy.

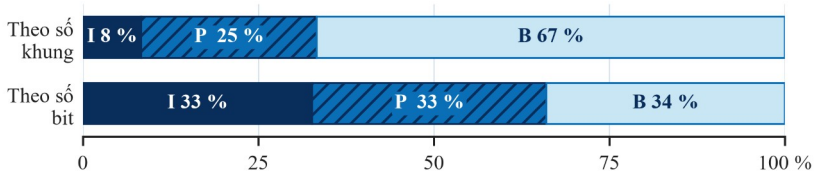
Bảng 8.37 Ba mức tinh vi của bù chuyển động

Kỹ thuật	Nội dung	Có trong chuẩn nào
Khối cố định, chính xác tới điểm ảnh	Khối 16 × 16, vectơ là số nguyên	MPEG-1, MPEG-2
Khối kích thước thay đổi, chính xác tới một phần tư điểm ảnh	Khối từ 16 × 16 xuống 4 × 4; vectơ có phần lẻ, nội suy để lấy giá trị giữa hai điểm ảnh	H.264 trở đi
Nhiều ảnh tham chiếu, khối tuỳ ý	Mỗi khối được chọn tham chiếu từ nhiều khung khác nhau, kể cả khung cách xa; hình dạng khối không nhất thiết vuông	HEVC, AV1, VVC

8.17.10 Bit phân bố ra sao giữa các loại khung



(a) bit trung bình của mỗi khung



(b) tỉ lệ I, P, B theo số khung và theo số bit

Hình 8.48 Kích thước từng khung trong một nhóm ảnh, và ba cách nhìn cùng một nhóm

GHI NHỚ 8.18 — ĐỌC ĐÚNG TỈ LỆ PHẦN TRĂM CỦA KHUNG I, P, B

Tỉ lệ phần trăm của khung I, P, B chỉ có ý nghĩa khi nêu rõ đang đếm theo số khung hay theo số bit, vì hai cách đếm cho hai bức tranh rất khác nhau.

Theo số khung, trong nhóm 12 khung kiểu I B B P B B P B B P B B, khung I chiếm 8 %, khung P chiếm 25 % và khung B chiếm 67 %.

Theo số bit thì bức tranh khác hẳn: khung I tuy chỉ chiếm 8 % số khung nhưng chiếm khoảng một phần ba tổng số bit, vì mỗi khung I nặng gần gấp ba khung P và gấp gần tám lần khung B.

Điều này giải thích vì sao nhóm ảnh dài giúp giảm tốc độ bit: khung I càng thưa thì càng ít khung có chi phí bit cao.

8.17.11 Độ trễ mã hoá

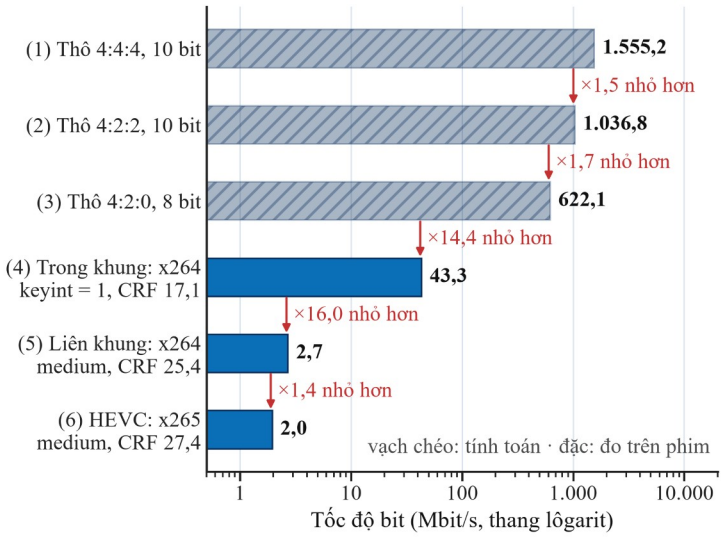
Khung B tham chiếu cả khung sau nên bộ mã hoá phải chờ khung sau tới rồi mới mã hoá được khung B. Việc chờ này sinh ra độ trễ, và độ trễ không thể chấp nhận trong truyền hình trực tiếp.

Bảng 8.38 Đánh đổi giữa hiệu quả nén và độ trễ

Cấu hình	Độ trễ điển hình	Dùng cho
Chỉ khung I	Gần bằng không	Sản xuất trong trường quay, dựng
I và P, không có B	Một khung	Truyền hình trực tiếp, hội nghị truyền hình, truyền hình lưu động
I, P và B, nhóm 12 khung	Vài khung	Phát sóng thông thường
Nhóm dài, nhiều tầng khung B	Hàng chục khung	Phát trực tuyến nội dung đã hoàn thiện, lưu trữ

8.17.12 Cộng dồn: mỗi tầng kỹ thuật giữ lại bao nhiêu bit

Để tổng hợp, Hình 8.49 ghi tốc độ bit sau từng tầng, tính cho tư liệu 1920 × 1080 ở 25 hình/giây. Ba cột đầu là kết quả tính thuần tuý theo công thức ở mục 8.14: 4:4:4 ở 10 bit cần 1 555,2 Mbit/s; chuyển sang 4:2:2 còn 1 036,8 Mbit/s (nhỏ đi 1,5 lần); chuyển sang 4:2:0 ở 8 bit còn 622,1 Mbit/s (nhỏ đi thêm 1,7 lần). Ba cột sau là số đo thật trên một đoạn phim quay ngoài trời 1080p25, chuyển động vừa phải, ở cùng chất lượng (PSNR của độ chói khoảng 40 dB): nén trong khung (x264, chỉ khung I) còn 43,3 Mbit/s, nén liên khung bằng x264 còn 2,7 Mbit/s và bằng x265 (HEVC) còn 2,0 Mbit/s.



Hình 8.49 Số bit giảm dần qua từng tầng kỹ thuật, đo trên một đoạn phim thật ở cùng PSNR (trục lôgarit)

Xét theo từng bước, dự đoán và biến đổi trong khung nén được khoảng 14,4 lần so với mẫu thô 4:2:0, tầng dự đoán liên khung nén thêm khoảng 16,0 lần nữa,

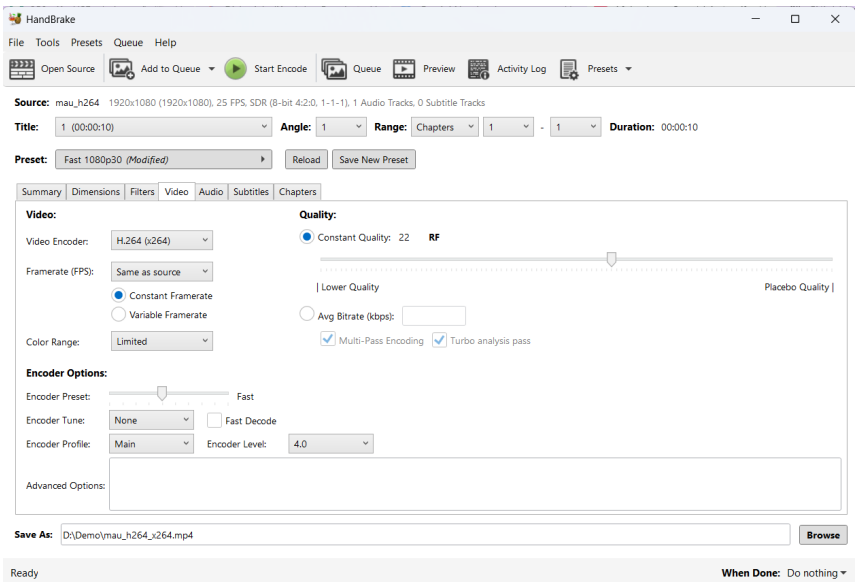
và thế hệ codec mới nén thêm khoảng 1,4 lần so với H.264. Mỗi tầng ứng với một nguyên lý đã học: lấy mẫu con màu (dư thừa cảm nhận), nén trong khung (dư thừa không gian), nén liên khung (dư thừa thời gian) và các công cụ tinh vi hơn của thế hệ mới (mục 8.14.5). Đoạn phim đo vốn đã qua nén trước khi đo nên các số tuyệt đối thấp hơn so với tư liệu thô từ máy quay. Cần đọc các con số một cách thận trọng, vì tỉ số phụ thuộc mạnh vào nội dung: cảnh ít chuyển động nén được nhiều hơn rất nhiều so với cảnh nhiều chuyển động hay nhiều nhiễu, nên các số trên chỉ minh hoạ bậc độ lớn của từng tầng và không phải hằng số.

8.18 Điều khiển tốc độ bit

8.18.1 Ba chế độ

Bộ mã hoá phải quyết định dùng bao nhiêu bit cho mỗi khung, và việc này có ba cách tiếp cận.

Hình 8.50 cho thấy thẻ Video của HandBrake, nơi chế độ chất lượng không đổi (Constant Quality, RF) được đặt cùng với bộ mã hoá x264, mức tinh chỉnh (Encoder Preset), profile và level.

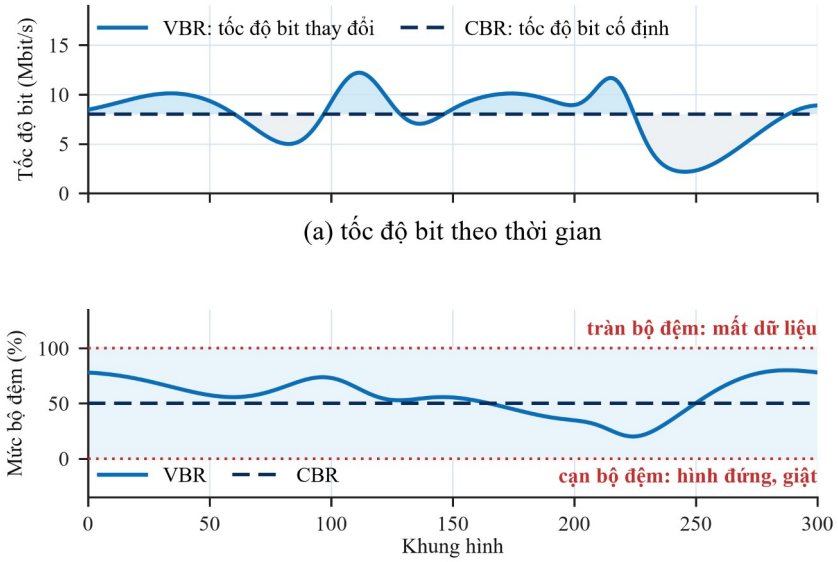


Hình 8.50 Điều khiển theo chất lượng không đổi (RF) và các tuỳ chọn bộ mã hoá x264 trong HandBrake

Nguồn: ảnh chụp màn hình do người soạn thực hiện (HandBrake 1.11), 05/11/2026.

Bảng 8.39 Ba chế độ điều khiển tốc độ bit

Chế độ	Cách làm	Ưu	Nhược
Tốc độ bit cố định (CBR)	Giữ tốc độ bit gần như không đổi theo thời gian	Dễ tính băng thông; bắt buộc với phát sóng và ghép kênh	Cảnh phức tạp bị thiếu bit nên vỡ hình; cảnh đơn giản lại thừa bit
Tốc độ bit thay đổi (VBR)	Cho phép tốc độ bit lên xuống theo độ phức tạp của cảnh	Chất lượng đồng đều hơn ở cùng dung lượng trung bình	Không đoán trước được đỉnh; khó ghép kênh
Chất lượng cố định	Đặt một mức chất lượng, để tốc độ bit tự do	Cho chất lượng ổn định nhất; hợp với lưu trữ và với bản trung gian	Hoàn toàn không biết trước dung lượng tệp

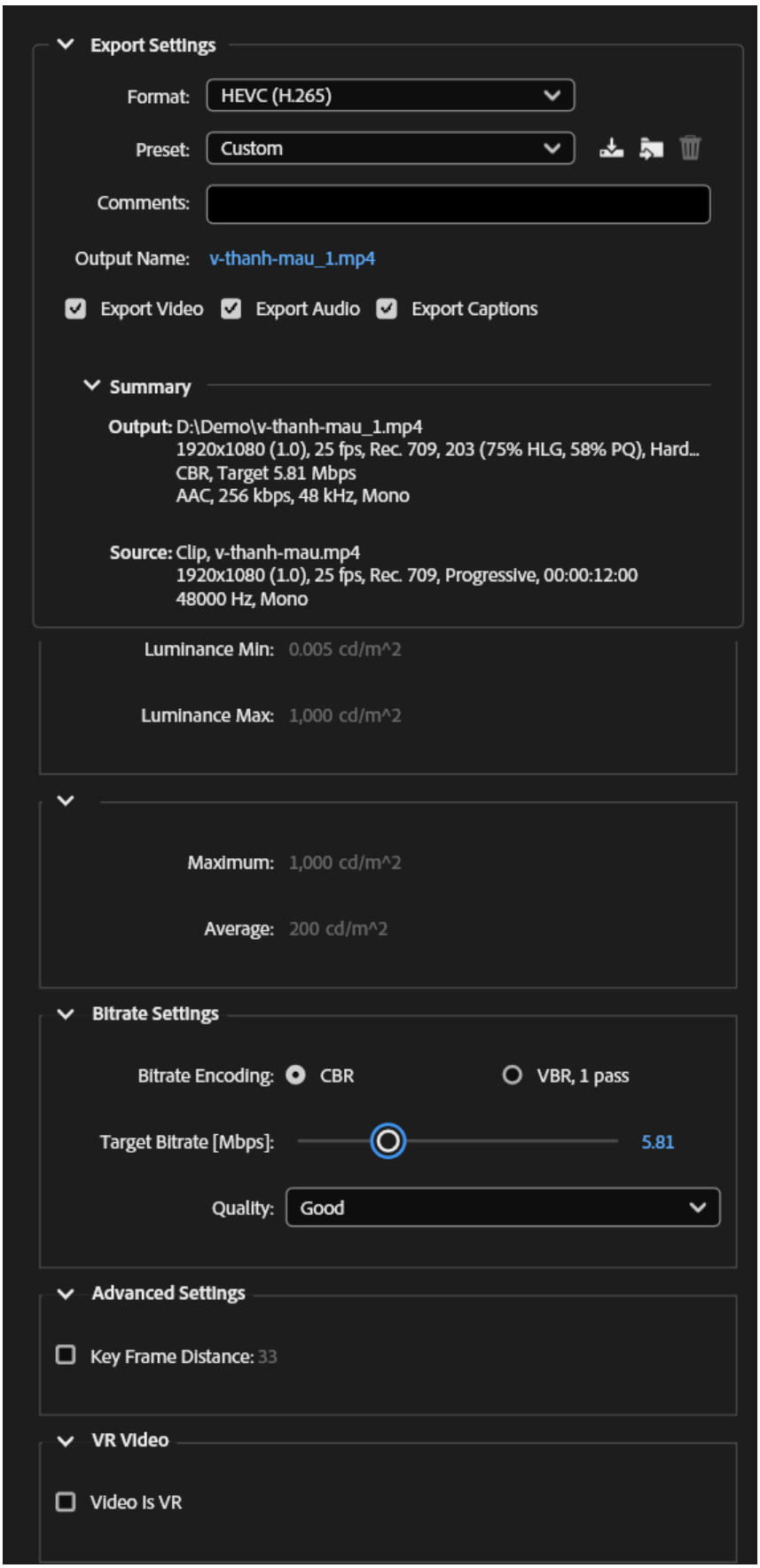


Hình 8.51 Hai chế độ điều khiển tốc độ bit và mức bộ đệm tương ứng

8.18.2 Bộ đệm — chỗ dung hoà

Bộ đệm cho phép tốc độ bit tức thời dao động quanh giá trị đặt, miễn là mức bộ đệm không bị tràn và không bị cạn.

Hình 8.52 là hộp thoại Export Settings của Adobe Media Encoder khi chọn bộ mã hoá HEVC (H.265) và chế độ tốc độ bit không đổi (CBR), cho thấy tóm tắt đầu ra và mục Bitrate Settings với tốc độ bit đích.



Hình 8.52 Export Settings của Adobe Media Encoder: định dạng HEVC (H.265), điều khiển tốc độ bit CBR, tốc độ bit đích 5,81 Mbit/s (nguồn là tệp thanh màu mẫu tổng hợp)

Nguồn: ảnh chụp màn hình do người soạn thực hiện (Adobe Media Encoder 2026), 07/11/2026.

$$B(t) = B(0) + \int_0^t [R_{\text{ch}} - R_{\text{enc}}(\tau)]d\tau \quad ; \quad 0 \le B(t) \le B_{\max} \tag{8.31}$$

CẢNH BÁO 8.7 — HAI SỰ CỐ CỦA NGƯỜI XEM, MỘT NGUYÊN NHÂN

Bộ đệm tràn: bộ mã hoá đẩy dữ liệu vào nhanh hơn đường truyền chuyển đi, dữ liệu bị loại bỏ và người xem thấy hình vỡ hoặc mất khối.

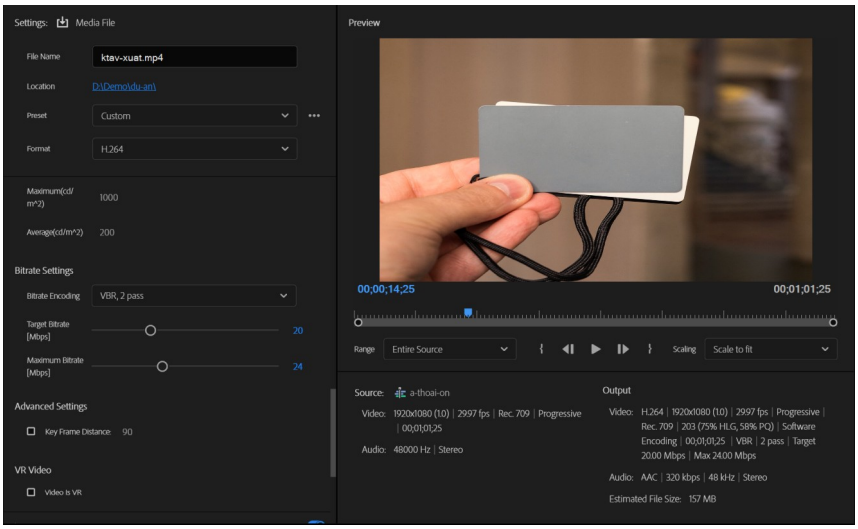
Bộ đệm cạn: bộ giải mã lấy dữ liệu ra nhanh hơn tốc độ nhận, không còn dữ liệu để giải mã và người xem thấy hình đứng lại rồi nhảy.

Mọi chuẩn phát sóng đều quy định kích thước bộ đệm tối đa. Một tệp vượt quá quy định vẫn mở bình thường trên máy tính vì máy tính có bộ nhớ dư dả, nhưng sẽ bị giật trên đầu thu của khán giả. Loại lỗi này chỉ phát hiện được bằng công cụ kiểm tra, không phát hiện được bằng cách xem thử.

8.18.3 Mã hoá một lượt và hai lượt

Mã hoá một lượt: bộ mã hoá đi thẳng từ đầu đến cuối và quyết định phân bổ bit dựa trên những gì đã xử lý. Cách này bắt buộc phải dùng khi phát trực tiếp.

Hình 8.53 là màn hình Export của Premiere Pro với định dạng H.264: mục Bitrate Settings chọn VBR hai lượt với tốc độ bit mục tiêu 20 Mbit/s và tối đa 24 Mbit/s, kèm khoảng cách khung khoá và tóm tắt đầu ra ở bên phải.



Hình 8.53 Thiết lập mã hoá H.264 VBR hai lượt (mục tiêu 20 Mbit/s, tối đa 24 Mbit/s) trong màn hình Export của Adobe Premiere Pro

Nguồn: ảnh chụp màn hình do người soạn thực hiện (Adobe Premiere Pro 2026), 07/11/2026.

Mã hoá hai lượt: lượt thứ nhất chỉ phân tích toàn bộ tư liệu để xác định đoạn phức tạp và đoạn đơn giản, còn lượt thứ hai mới mã hoá thật và phân bổ bit theo bản đồ đã có. Kết quả tốt hơn rõ rệt ở cùng dung lượng, nhưng thời gian xử lý gấp đôi và chỉ áp dụng được với tệp đã hoàn chỉnh.

8.18.4 Chọn tốc độ bit bao nhiêu

Không có con số nào dùng được cho mọi trường hợp, vì tốc độ bit cần thiết phụ thuộc độ phân giải, tần số hình, mức chuyển động của nội dung và thế hệ

codec. Bảng dưới đây là điểm khởi đầu thực dụng cho nội dung có mức chuyển động trung bình.

Bảng 8.40 Tốc độ bit khởi điểm theo định dạng và codec

Định dạng	MPEG-2	H.264	HEVC	Ghi chú
SD 576i25	4 – 6 Mbit/s	2 – 3 Mbit/s	1 – 1,5 Mbit/s	Kênh SD trong bộ ghép kênh DVB-T2
HD 1080i25	15 – 18 Mbit/s	6 – 10 Mbit/s	3 – 5 Mbit/s	Kênh HD phát sóng
Full HD 1080p25	không dùng	8 – 12 Mbit/s	4 – 6 Mbit/s	Phát trực tuyến chất lượng cao
UHD 2160p50	không dùng	35 – 50 Mbit/s	15 – 25 Mbit/s	UHD phát trực tuyến

Cách dùng bảng là lấy con số làm điểm xuất phát, mã hoá thử một đoạn khó nhất của chương trình, xem trên màn hình đúng cỡ mà khán giả sẽ xem rồi điều chỉnh. Nội dung nhiều chuyển động (thể thao, ca nhạc có ánh sáng nhấp nháy) cần tốc độ bit cao hơn đáng kể, còn nội dung tĩnh (bản tin, phỏng vấn) cần ít hơn.

8.18.5 Kiểm lại tốc độ bit đã chọn

Việc chọn xong tốc độ bit chưa phải là hoàn tất công việc, mà cần thêm một quy trình kiểm tra ngắn, chỉ mất chừng mười lăm phút.

Bảng 8.41 Bốn bước kiểm tốc độ bit đã chọn

Bước	Việc làm	Điều cần thấy
1. Chọn đoạn KHÓ NHẤT	Tìm trong chương trình đoạn có nhiều chuyển động nhất, hoặc nhiều chi tiết mịn nhất — thường là cảnh lia máy qua đám đông, cảnh mưa, cảnh có khói hoặc pháo hoa.	Nếu đoạn khó nhất đạt thì cả chương trình đạt. Kiểm bằng đoạn dễ là tự lừa mình.
2. Mã hoá thử 30 giây	Mã hoá riêng đoạn ấy đúng cấu hình dự định dùng cho cả chương trình.	Không cần mã hoá cả chương trình chỉ để kiểm.
3. Xem trên màn hình ĐÚNG CỠ	Xem trên loại màn hình mà khán giả sẽ dùng, ở khoảng cách thật. Xem trên màn hình máy tính đặt gần mắt sẽ thấy lỗi nặng hơn thực tế; xem trên điện thoại sẽ thấy nhẹ hơn thực tế.	Áp dụng đúng lập luận về khoảng cách ngồi xem ở mục 8.4.
4. Kiểm bằng số	Nếu có công cụ, tính thêm chỉ số VMAF hoặc SSIM so với bản gốc; ghi lại con số làm mốc cho các chương trình sau.	Sau vài lần, sẽ có một bảng mốc riêng cho từng loại nội dung mà đơn vị mình hay làm — quý hơn mọi bảng số in sẵn.

8.19 Bài thực hành số 2 — nhìn thấy nén

Mục tiêu

Sau bài này, người đọc nhận ra được vỡ khối và nhiễu muỗi bằng mắt, đo được ảnh hưởng của độ dài nhóm ảnh tới việc tua và cắt, so sánh được hai thế hệ codec ở cùng tốc độ bit, và nhận biết được tổn hao thế hệ.

Thiết bị và phần mềm

- Máy tính có cài FFmpeg (miễn phí, chạy bằng dòng lệnh) hoặc một phần mềm chuyển mã có giao diện.
- Một phần mềm dựng để tua và cắt thử.
- Hai đoạn tư liệu tự quay, mỗi đoạn khoảng 30 giây: một đoạn tĩnh (phỏng vấn, máy đặt trên chân máy) và một đoạn chuyển động mạnh (lia máy nhanh qua cây lá, hoặc người chạy).

Nội dung thực hành

Bảng 8.42 Bốn phần của bài thực hành số 2

Phần	Việc phải làm	Kết quả cần rút ra
1. Tìm ngưỡng vỡ hình	Mã hoá mỗi đoạn tư liệu bằng H.264 ở các tốc độ bit giảm dần: 8, 4, 2, 1, 0,5 rồi 0,25 Mbit/s. Xem từng bản trên màn hình lớn.	Ghi lại tốc độ bit mà tại đó mỗi đoạn bắt đầu lộ lỗi. Giải thích vì sao đoạn chuyển động mạnh vỡ sớm hơn đoạn tĩnh rất nhiều.
2. Gọi tên lỗi	Ở bản tốc độ bit thấp nhất, phóng to và chụp màn hình ba vùng: một vùng bầu trời hoặc phông trơn, một vùng có chữ hoặc cành lá trên nền sáng, một vùng có hoa văn mịn.	Chỉ ra và gọi đúng tên ba lỗi: vỡ khối, nhiễu muỗi, bệt chi tiết. Giải thích nguyên nhân của từng lỗi theo mục 8.16.5.
3. So sánh hai thế hệ codec	Mã hoá cùng một đoạn bằng H.264 và bằng HEVC ở CÙNG tốc độ bit, chọn mức tốc độ bit đủ thấp để thấy khác biệt. So sánh trực quan; nếu công cụ cho phép, tính thêm PSNR của cả hai.	HEVC cho chất lượng cao hơn rõ rệt ở cùng tốc độ bit. Đo thử thời gian mã hoá của hai bản và nhận xét về cái giá phải trả.
4. Nhóm ảnh và tổn hao thế hệ	Mã hoá cùng một đoạn với nhóm ảnh 1 khung, 12 khung và 250 khung; so dung lượng ba tệp và thử tua, cắt trong phần mềm dựng. Sau đó, lấy bản 2 Mbit/s mã hoá LẠI năm lần liên tiếp ở cùng tốc độ bit.	Nhóm dài cho tệp nhỏ nhất nhưng tua nặng và cắt kém linh hoạt. Sau năm lần mã hoá lại, chất lượng giảm rõ dù tốc độ bit không đổi — đó là tổn hao thế hệ.

Kết quả cần có

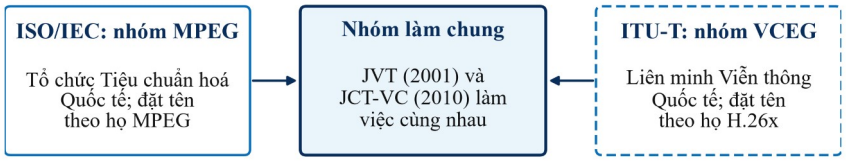
Kết quả cần có gồm một bảng ghi tốc độ bit ngưỡng của hai đoạn tư liệu; ba ảnh chụp màn hình có chú thích gọi tên lỗi; hai ảnh so sánh H.264 với HEVC kèm thời gian mã hoá; một bảng so sánh ba độ dài nhóm ảnh; và một ảnh so sánh

bản gốc với bản đã mã hoá lại năm lần. Kèm theo là một đoạn nhận xét ngắn về codec và tốc độ bit nên chọn cho loại nội dung mà mình thường thực hiện.

8.20 Bản đồ các họ chuẩn

8.20.1 Ai làm ra chuẩn nén

Trước khi đi vào từng chuẩn, cần làm rõ vì sao cùng một kỹ thuật lại có hai tên gọi và vì sao các tên gọi ấy dễ gây nhầm lẫn.



Tên họ MPEG	Tên họ ITU-T	Số hiệu ISO/IEC	Năm	Dùng chủ yếu cho
MPEG-1 Video	—	ISO/IEC 11172-2	1993	đĩa VCD
MPEG-2 Video	**H.262**	ISO/IEC 13818-2	1995	DVD; truyền hình số thế hệ đầu
MPEG-4 Part 2	—	ISO/IEC 14496-2	1999	video trên mạng thời kỳ đầu
MPEG-4 Part 10	**H.264 / AVC**	ISO/IEC 14496-10	2003	chuẩn phổ biến nhất hiện nay
MPEG-H Part 2	**H.265 / HEVC**	ISO/IEC 23008-2	2013	UHD và HDR
MPEG-I Part 3	**H.266 / VVC**	ISO/IEC 23090-3	2020	8K, video 360 độ

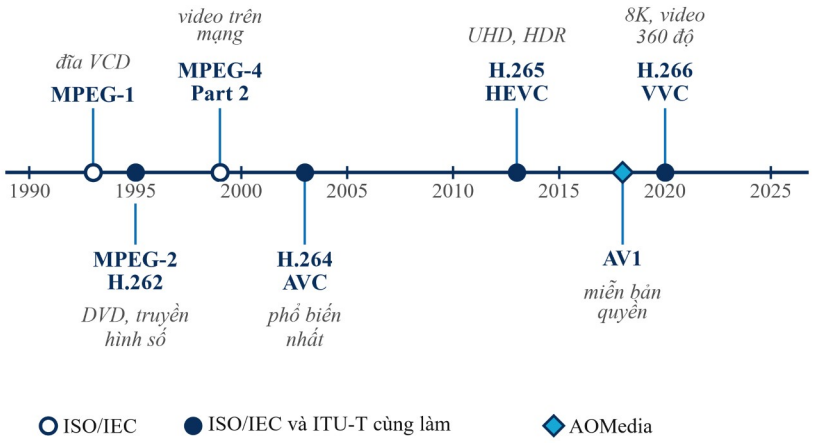
Hình 8.54 Các tổ chức làm ra chuẩn nén và bảng đối chiếu tên gọi

Hai tổ chức lớn cùng thực hiện công việc này. Tổ chức Tiêu chuẩn hoá Quốc tế cùng Uỷ ban Kỹ thuật điện Quốc tế lập ra nhóm chuyên gia hình động, quen gọi là nhóm MPEG, và đặt tên sản phẩm theo họ MPEG; Liên minh Viễn thông Quốc tế lập ra nhóm chuyên gia mã hoá hình, quen gọi là VCEG, và đặt tên theo họ H.26x.

Từ năm 2001, hai bên bắt đầu làm chung: nhóm JVT cho H.264 và nhóm JCT-VC cho HEVC. Sản phẩm chung có hai tên gọi đều hợp lệ, cùng chỉ một văn bản kỹ thuật duy nhất, nên "H.264" và "MPEG-4 Part 10" không phải hai chuẩn khác nhau.

> **GHI NHỚ 8.19 — MPEG-7 VÀ MPEG-21 KHÔNG PHẢI CHUẨN NÉN**
>
> MPEG-7 không nén hình mà là chuẩn mô tả nội dung đa phương tiện bằng siêu dữ liệu, phục vụ việc tìm kiếm và phân loại; MPEG-21 cũng không phải chuẩn nén mà là khung quản lý quyền và phân phối nội dung.
>
> Vì vậy MPEG-7 không thuộc cùng nhóm với MPEG-1, MPEG-2 và MPEG-4, và thế hệ codec kế tiếp của MPEG-4 là H.264.

8.20.2 Dòng thời gian



Hình 8.55 Dòng thời gian các chuẩn nén video từ 1993 tới nay

Bảng 8.43 Các thế hệ chuẩn nén và bối cảnh ra đời

Chuẩn	Năm	Sinh ra để giải bài toán gì	Còn gặp ở đâu
MPEG-1	1993	Nhét được một bộ phim vào đĩa CD ở khoảng 1,5 Mbit/s	Đĩa VCD; tư liệu số hoá thời kỳ đầu
MPEG-2 / H.262	1995	Chất lượng phát sóng ở vài megabit mỗi giây; hỗ trợ quét xen kẽ	Đĩa DVD; truyền hình số thế hệ đầu; một số kênh SD
MPEG-4 Part 2	1999	Video trên đường truyền hẹp thời Internet quay số	Hầu như không còn dùng
H.264 / AVC	2003	Giảm một nửa tốc độ bit so với MPEG-2, ở mọi độ phân giải	Chuẩn phổ biến nhất thế giới; DVB-T2 ở Việt Nam; đĩa Blu-ray
H.265 / HEVC	2013	Giảm tiếp một nửa để chở nổi UHD và HDR	UHD phát trực tuyến; lưu trữ; máy quay ghi thẻ
AV1	2018	Tiết kiệm thêm khoảng 20 – 30 % so với HEVC, miễn bản quyền	Các nền tảng phát trực tuyến lớn
H.266 / VVC	2020	Giảm tiếp một nửa so với HEVC; nhắm 8K và video 360 độ	Còn đang phổ cập

Điểm đáng chú ý trong bảng là khoảng cách giữa các thế hệ ngày càng ngắn nhưng thời gian phổ cập thì không ngắn lại. H.264 ra đời năm 2003 và đến nay vẫn là chuẩn được hỗ trợ rộng nhất, bởi gần như mọi thiết bị đều giải mã được chuẩn này bằng phần cứng. Vì vậy, gọi H.264 là "chuẩn nén mới nhất" là đã lạc hậu hai thế hệ, do sau nó còn có HEVC (2013), AV1 (2018) và VVC (2020) [6].

8.21 H.264 / AVC

8.21.1 Vì sao vẫn phải học kỹ chuẩn của năm 2003

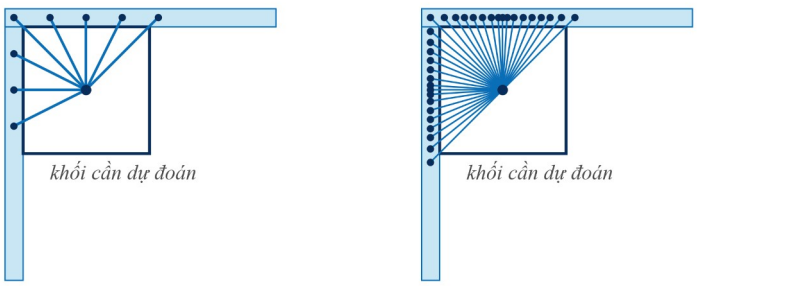
H.264 xứng đáng được trình bày kỹ nhất trong bốn thế hệ vì ba lý do. Thứ nhất, đây là chuẩn mà QCVN 63:2020/BTTTT quy định cho đầu thu truyền hình

số mặt đất ở Việt Nam, nên mọi bản giao nộp cho đài phải dùng chuẩn này. Thứ hai, đây là chuẩn có mặt trên nhiều thiết bị nhất thế giới. Thứ ba, mọi chuẩn sau nó đều xây dựng trên cùng bộ khái niệm, nên khi đã hiểu H.264 thì có thể nắm bắt HEVC và AV1 nhanh chóng [7], [28].

8.21.2 Các công cụ chính

Bảng 8.44 Sáu công cụ làm nên hiệu quả của H.264

Công cụ	Nội dung	Đóng góp
Macroblock 16 × 16 chia nhỏ được	Mỗi macroblock chia được xuống tới khối 4 × 4 để bù chuyển động	Bám sát biên vật chuyển động, giảm phần dư
Dự đoán trong khung theo hướng	Khối được đoán từ các điểm ảnh đã giải mã ở hàng trên và cột trái, theo tám hướng cộng chế độ trung bình	Giảm mạnh kích thước khung I so với MPEG-2
Nhiều ảnh tham chiếu	Một khối được phép tham chiếu tới nhiều khung khác nhau, không chỉ khung liền trước	Xử lý tốt cảnh có vật che khuất rồi hiện lại
Biến đổi số nguyên 4 × 4	Thay biến đổi cosin dấu phẩy động bằng phép biến đổi số nguyên xấp xỉ	Mọi bộ giải mã cho kết quả GIỐNG HỆT nhau, không sai lệch tích luỹ
Bộ lọc khử khối trong vòng lặp	Làm mềm ranh giới giữa các khối ngay trong quá trình mã hoá	Giảm hẳn hiện tượng vỡ khối ở tốc độ bit thấp
CABAC	Mã số học thích nghi theo ngữ cảnh, thay cho mã Huffman	Tiết kiệm thêm khoảng 10 – 15 % so với CAVLC



Hình 8.56 Dự đoán trong khung: số hướng của H.264 so với HEVC

8.21.3 Lát cắt và đơn vị NAL

Một khung được chia thành các *lát cắt*, mỗi lát cắt được giải mã độc lập với các lát cắt khác trong cùng khung. Nhờ vậy, việc mất một gói dữ liệu chỉ làm hỏng một lát chứ không làm hỏng cả khung, đặc điểm này rất quan trọng khi truyền trên đường truyền có lỗi.

Dòng bit H.264 được đóng thành các đơn vị NAL, mỗi đơn vị là một gói tự mô tả và thuộc một trong hai loại: đơn vị mang dữ liệu ảnh và đơn vị mang *thông số*, gồm bộ thông số dãy và bộ thông số ảnh. Hai bộ thông số này chứa toàn bộ

khai báo về độ phân giải, tần số hình, không gian màu và phạm vi mức; khi chúng bị thiếu hoặc sai, bộ giải mã phải suy đoán, và việc suy đoán sai chính là nguồn gốc của các lỗi màu đã nêu ở mục 8.8.

8.21.4 Profile và Level

Profile và level là hai khái niệm dễ gây nhầm lẫn nhất khi đọc thông số kỹ thuật, đồng thời cũng là dạng yêu cầu mà QCVN 63:2020/BTTTT đặt ra.

Profile *được dùng những công cụ nào*

Baseline	không có khung B, không có CABAC *hội nghị truyền hình, thiết bị yếu*
Main	thêm khung B và CABAC *truyền hình độ nét tiêu chuẩn*
High	thêm biến đổi 8 × 8 và ma trận tỉ lệ *HD, Blu-ray, DVB-T2; thông dụng nhất*
High 10	thêm độ sâu 10 bit *lưu trữ, hậu kỳ, bản trung gian*
High 4:2:2	thêm cấu trúc 4:2:2 và 10 bit *codec dựng chuyên nghiệp, ghi ngoài*
High 4:4:4 Predictive	thêm 4:4:4, 14 bit, không tổn hao *bản trung gian đặc biệt; ít thiết bị hỗ trợ*

(a) profile: công cụ cho phép, càng xuống càng nhiều

Level *được phép đến cỡ nào*

Level	Cỡ ảnh · tần số hình	Tốc độ bit tối đa
3.0	720 × 576 · 25 hình/s	10 Mbit/s
4.0	1.920 × 1.080 · 30 hình/s	20 Mbit/s
4.2	1.920 × 1.080 · 60 hình/s	50 Mbit/s
5.1	3.840 × 2.160 · 30 hình/s	240 Mbit/s
6.2	8.192 × 4.320 · 120 hình/s	800 Mbit/s

(b) level: giới hạn cỡ ảnh, tần số hình và tốc độ bit

Hình 8.57 Profile quy định dùng công cụ nào, Level quy định được đến cỡ nào

- **Profile** xác định những công cụ nén nào được sử dụng: profile càng cao thì dùng càng nhiều công cụ và nén càng tốt, nhưng đòi hỏi bộ giải mã mạnh hơn.
- **Level** xác định giới hạn được phép: level quy định cỡ ảnh, tần số hình, tốc độ bit và kích thước bộ đệm tối đa.

Bảng 8.45 Các profile chính của H.264

Profile	Công cụ được dùng thêm	Dùng cho
Baseline	Không có khung B, không có CABAC	Hội nghị truyền hình, thiết bị yếu
Main	Thêm khung B và CABAC	Truyền hình độ nét tiêu chuẩn
High	Thêm biến đổi 8 × 8 và ma trận tỉ lệ	HD, đĩa Blu-ray, phát sóng DVB-T2 — profile thông dụng nhất
High 10	Thêm độ sâu 10 bit	Lưu trữ, hậu kỳ, bản trung gian
High 4:2:2	Thêm cấu trúc 4:2:2 và 10 bit	Codec dựng chuyên nghiệp, ghi ngoài trên hiện trường
High 4:4:4 Predictive	Thêm 4:4:4, 14 bit và chế độ không tổn hao	Bản trung gian đặc biệt; rất ít thiết bị hỗ trợ

GHI NHỚ 8.20 — ĐỌC MỘT DÒNG THÔNG SỐ

Dòng mô tả "H.264 High Profile Level 4.0, 4:2:0, 8 bit, 1920 × 1080i25, 8 Mbit/s CBR" cung cấp đầy đủ các thông tin cần thiết:

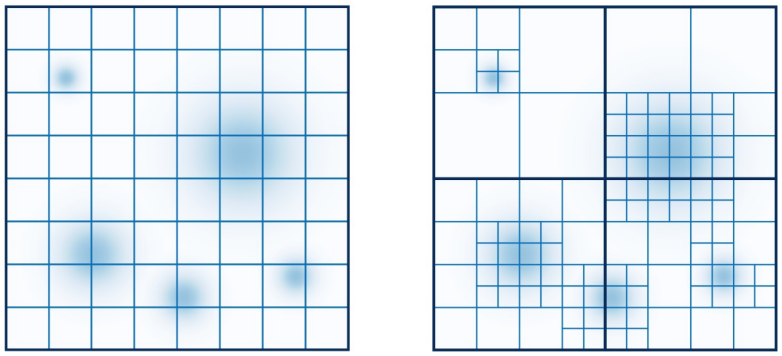
H.264 High Profile cho phép dùng biến đổi 8 × 8 và ma trận tỉ lệ; Level 4.0 quy định không vượt quá Full HD ở 30 hình/giây và tốc độ bit tối đa 25 Mbit/s với High Profile (20 Mbit/s với Main Profile); 4:2:0 8 bit cho biết cấu trúc lấy mẫu con màu và độ sâu bit; phần còn lại thể hiện định dạng ảnh và chế độ tốc độ bit.

QCVN 63:2020/BTTTT quy định đầu thu ở Việt Nam phải giải mã được H.264 Main Profile Level 3 cho độ nét tiêu chuẩn và High Profile Level 4 cho độ nét cao; các mức này chính là giới hạn kỹ thuật của bản giao nộp cho truyền hình mặt đất.

8.22 H.265 / HEVC

8.22.1 Khác biệt cốt lõi: cây tứ phân

Điểm cốt lõi của HEVC là việc thay ô mã hoá cố định bằng một *cây tứ phân* [8], [29].



64 macroblock 16 × 16, bằng nhau

(a) H.264: ô cố định

4 CTU 64 × 64, tự chia tới 8 × 8

(b) HEVC: cây tứ phân

Hình 8.58 Ô cố định của H.264 so với cây tứ phân của HEVC

Trong H.264, ảnh được chia thành các macroblock 16 × 16 bằng nhau, nên vùng trời phẳng và vùng lá cây dày đặc đều tiêu tốn số ô như nhau. Trong HEVC, đơn vị mã hoá lớn nhất là 64 × 64 và được chia tứ phân xuống tới 8 × 8 tuỳ theo độ phức tạp của từng vùng, do đó một mảng trời có thể chỉ dùng một ô 64 × 64, còn việc chia nhỏ được dành cho những nơi thật sự cần.

Phần lớn mức tiết kiệm khoảng 50 phần trăm của HEVC so với H.264 đến từ cơ chế này, cùng với việc số hướng dự đoán trong khung tăng từ tám lên ba mươi ba.

8.22.2 Các công cụ khác

Bảng 8.46 Bốn công cụ đặc trưng khác của HEVC

Công cụ	Nội dung	Tác dụng
Đơn vị dự đoán và đơn vị biến đổi tách rời	Kích thước khối dùng cho bù chuyển động không nhất thiết bằng kích thước khối dùng cho biến đổi	Linh hoạt hơn hẳn; bám sát cấu trúc ảnh
Bộ lọc bù dịch mẫu (SAO)	Sau bộ lọc khử khối, thêm một bước hiệu chỉnh mức theo từng vùng	Giảm hiện tượng vòng dư quanh mép sắc nét
Tile và wavefront	Chia ảnh thành các mảng hoặc các dòng có thể mã hoá SONG SONG	Cho phép dùng nhiều lõi xử lý; quan trọng với UHD
Profile Main 10	Hỗ trợ 10 bit ngay trong profile cơ bản, không cần profile riêng	Là nền tảng kỹ thuật của HDR — nối thẳng với mục 7.22 của Chương 7

8.22.3 Cái giá phải trả

HEVC thay thế H.264 chậm hơn so với dự đoán năm 2013, với hai nguyên nhân cụ thể.

– **Chi phí mã hoá.** Cây tứ phân và ba mươi ba hướng dự đoán buộc bộ mã hoá phải thử nhiều phương án hơn hẳn, nên thời gian mã hoá tăng khoảng mười

lần so với H.264 ở cùng cấu hình chất lượng; với phát trực tiếp, đây là rào cản thực sự.

– **Bản quyền sáng chế.** H.264 chỉ có một nhóm quản lý bản quyền với biểu giá rõ ràng, trong khi HEVC có ba nhóm khác nhau, cộng thêm một số chủ sáng chế đứng ngoài và điều kiện chưa rõ ràng; doanh nghiệp không tính được chi phí thì không thể triển khai.

GHI NHỚ 8.21 — HỆ QUẢ LỊCH SỬ

Tình trạng bản quyền phức tạp của HEVC đã dẫn tới việc một nhóm các công ty công nghệ lớn lập ra liên minh AOMedia năm 2015 và công bố AV1 năm 2018, một chuẩn có hiệu quả tương đương nhưng miễn phí bản quyền.

Trường hợp này cho thấy một chuẩn được chấp nhận không chỉ nhờ ưu thế kỹ thuật mà còn nhờ điều kiện sử dụng.

8.23 AV1 và H.266 / VVC

8.23.1 AV1

AV1 do liên minh AOMedia công bố năm 2018. Về kỹ thuật, chuẩn này đi theo hướng của HEVC nhưng mở rộng hơn: siêu khối lớn nhất là 128 × 128, có thêm nhiều chế độ dự đoán và nhiều bộ lọc mới, trong đó có bộ lọc tự thích nghi theo hướng của cạnh [10].

Về hiệu quả, AV1 tiết kiệm thêm khoảng 20 tới 30 phần trăm so với HEVC ở cùng chất lượng. Về chi phí mã hoá, các bản mã hoá đầu tiên chậm hơn HEVC hàng chục lần, nhưng các bản sau đã được cải thiện đáng kể. Ưu thế quyết định của AV1 nằm ở điều kiện sử dụng hơn là kỹ thuật: chuẩn này miễn phí bản quyền nên được các nền tảng phát trực tuyến lớn áp dụng nhanh chóng.

8.23.2 VVC

H.266, còn gọi là VVC, được công bố tháng 7 năm 2020 và cũng do nhóm chung của ISO/IEC và ITU-T xây dựng. Mục tiêu của chuẩn là giảm thêm khoảng một nửa tốc độ bit so với HEVC, và hướng tới ba loại nội dung mà HEVC xử lý chưa tốt: độ phân giải 8K, video 360 độ và nội dung màn hình máy tính có nhiều chữ và đồ hoạ [9], [30].

8.23.3 Bảng so sánh bốn thế hệ

	MPEG-2	H.264 / AVC	H.265 / HEVC	AV1 · H.266 / VVC
Năm công bố	1995	2003	2013	2018 · 2020
Ô mã hoá lớn nhất	16 × 16	16 × 16	64 × 64	128 × 128
Tốc độ bit cho cùng chất lượng	100 %	≈ 50 %	≈ 25 %	≈ 12 – 20 %
Chi phí mã hoá (thời gian máy)	1 ×	≈ 3 ×	≈ 10 ×	≈ 30 – 100 ×
Giải mã bằng phần cứng	mọi thiết bị	mọi thiết bị	thiết bị từ 2015	thiết bị từ 2020
Bản quyền sáng chế	đã hết hạn	một nhóm, giá rõ ràng	ba nhóm, giá không rõ	AV1 miễn phí; VVC còn đang bàn
Ở Việt Nam dùng cho	tư liệu cũ	DVB-T2, giao nộp cho đài	UHD, lưu trữ	phát trực tuyến

Hình 8.59 Bảng so sánh bốn thế hệ codec theo bảy tiêu chí

Bảng trong hình trên là công cụ hỗ trợ lựa chọn, không phải nội dung cần học thuộc. Bảng được đọc từ dưới lên, bắt đầu bằng câu hỏi "bản này sẽ được xem ở đâu và trên thiết bị nào", rồi mới chọn codec.

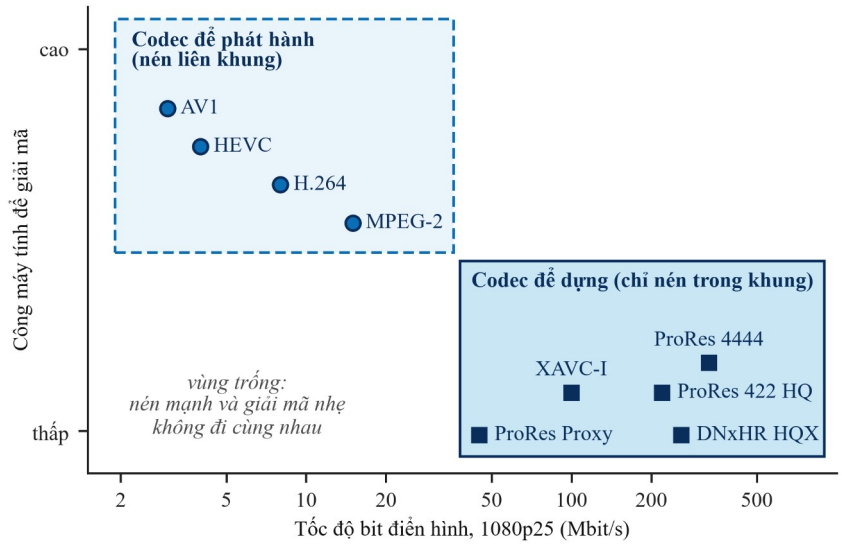
Bảng 8.47 Khuyến nghị chọn codec cho hoàn cảnh Việt Nam hiện nay

Mục đích	Nên dùng	Lý do
Giao nộp cho đài truyền hình	H.264 High Profile	QCVN 63:2020/BTTTT quy định đầu thu phải giải mã H.264; quy chuẩn KHÔNG nêu HEVC nên bản HEVC không bảo đảm thu được
Phát hành trên mạng xã hội và nền tảng video	H.264 cho bản tải lên; để nền tảng tự mã hoá lại	Nền tảng nào cũng nhận H.264; tải lên bằng codec lạ chỉ thêm một lần mã hoá lại
Lưu trữ bản hoàn chỉnh dài hạn	HEVC 10 bit, hoặc codec dựng nếu có đủ dung lượng	Tiết kiệm một nửa dung lượng so với H.264 ở cùng chất lượng
Quay và ghi trên hiện trường	Codec dựng: ProRes, DNxHR, XAVC-I	Xem mục 8.24 — không nên ghi thẳng bằng codec nén liên khung nếu có lựa chọn
Dịch vụ phát trực tuyến tự xây dựng	H.264 làm bậc nền, thêm AV1 nếu có nguồn lực mã hoá	H.264 bảo đảm mọi thiết bị xem được; AV1 giảm chi phí băng thông cho các bậc cao

8.24 Codec để dựng và codec để phát hành

8.24.1 Hai lớp codec, hai mục đích

Sự phân biệt này quyết định trực tiếp việc phòng dựng vận hành ổn định hay có hiệu năng thấp.



Hình 8.60 Vị trí của các codec trên hai trục: tốc độ bit và công giải mã

Bảng 8.48 Hai lớp codec

Tiêu chí	Codec để DỰNG	Codec để PHÁT HÀNH
Kiểu nén	Chỉ nén TRONG KHUNG; mọi khung đều là khung I	Nén LIÊN KHUNG với nhóm ảnh dài
Tốc độ bit	Cao: 100 – 900 Mbit/s tuỳ bậc	Thấp: 2 – 50 Mbit/s
Công giải mã	Rất nhẹ — giải mã một khung không cần khung nào khác	Nặng — phải giải mã cả nhóm ảnh để lấy được một khung
Tua và cắt	Tức thì, cắt được ở mọi khung	Chậm; cắt ngoài khung I phải mã hoá lại
Chịu được mã hoá lại	Tốt — tổn hao thế hệ rất nhỏ	Kém — mỗi lần mã hoá lại đều lấy đi chất lượng
Ví dụ	ProRes, DNxHR, XAVC-I, Cineform	H.264, HEVC, AV1, MPEG-2

8.24.2 Các họ codec dựng thường gặp

Bảng 8.49 Ba họ codec dựng và các bậc của chúng

Họ	Các bậc	Đặc điểm
ProRes	422 Proxy · 422 LT · 422 · 422 HQ · 4444 · 4444 XQ	Của Apple; phổ biến nhất trong hậu kỳ điện ảnh và quảng cáo. Bậc 4444 mang thêm kênh trong suốt
DNxHR	LB · SQ · HQ · HQX · 444	Của Avid; định dạng chuẩn trong dây chuyền hậu kỳ dùng Avid. Bậc HQX là 12 bit
XAVC-I	Nhiều mức tốc độ bit theo độ phân giải	Của Sony; ghi thẳng trong máy quay, dùng H.264 nhưng CHỈ khung I

XAVC-I dùng chính chuẩn H.264 nhưng cấu hình nhóm ảnh chỉ gồm một khung, điều này cho thấy ranh giới giữa hai lớp codec không nằm ở chuẩn nén mà nằm ở cách cấu hình chuẩn đó.

8.24.3 RAW không phải video đã mã hoá

> **GHI NHỚ 8.22 — MỘT PHÂN BIỆT CẦN NÓI RÕ**
>
> Tệp RAW của máy quay không phải là video đã mã hoá mà là dữ liệu đọc thẳng từ cảm biến, trước khi khử ghép màu Bayer, trước khi cân bằng trắng và trước khi áp đường truyền.
>
> Do đó, việc mở tệp RAW đòi hỏi phần mềm thực hiện lại toàn bộ các bước ấy, và đổi lại người dựng có quyền quyết định lại cân bằng trắng và phơi sáng như khi đang ở hiện trường.
>
> Chi tiết về cảm biến và ma trận Bayer thuộc Chương 9; ở đây chỉ cần ghi nhận rằng RAW nằm ngoài hệ thống phân loại codec của chương này.

8.24.4 Proxy và quy tắc chọn codec cho từng khâu

Proxy là bản sao có độ phân giải thấp và tốc độ bit thấp, dùng để dựng trên máy có cấu hình yếu. Sau khi dựng xong, phần mềm tự thay proxy bằng bản gốc theo danh sách quyết định dựng. Đây là kỹ thuật đã có từ lâu nhưng vẫn rất hữu ích với tư liệu UHD.

Bảng 8.50 Quy tắc chọn codec cho từng khâu

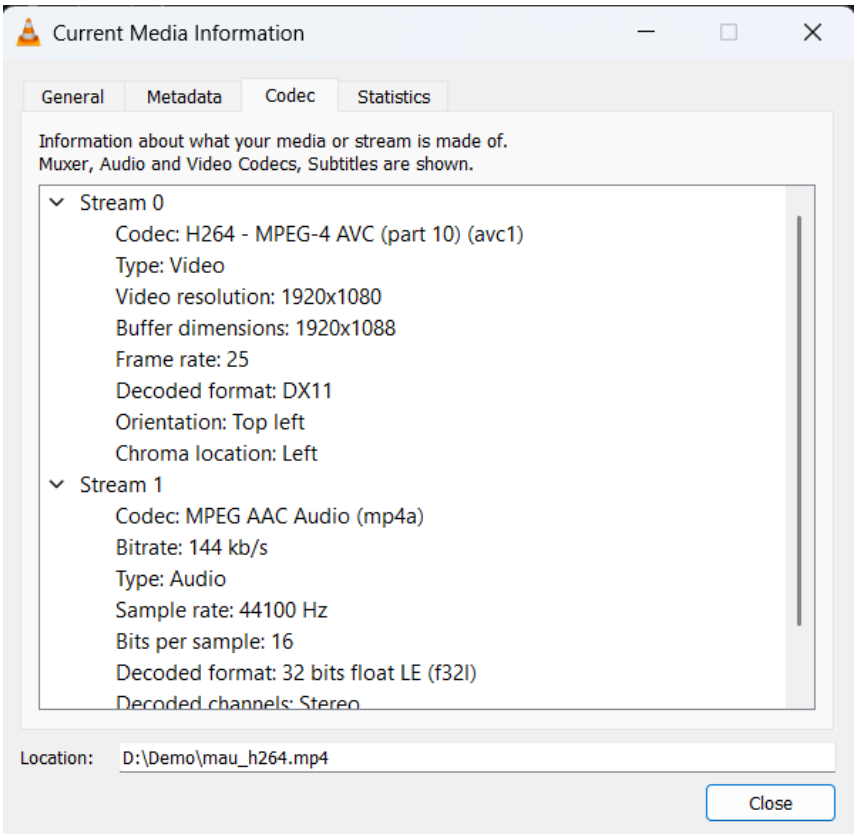
Khâu	Nên dùng	Tránh
Quay và ghi	Codec dựng nếu thiết bị cho phép; nếu không thì chấp nhận codec liên khung của máy	Đặt tốc độ bit thấp nhất chỉ để tiết kiệm thẻ
Sao lưu bản gốc	Giữ NGUYÊN tệp gốc, không chuyển mã	Chuyển mã "cho gọn" ngay khi đổ thẻ — mất bản gốc là mất vĩnh viễn
Dựng	Codec dựng, hoặc proxy nếu máy yếu	Dựng trực tiếp trên tệp H.264 hoặc HEVC tốc độ bit thấp
Chỉnh màu	Bản gốc hoặc bản trung gian 10 bit trở lên	Bản 8 bit đã nén sâu — kéo màu sẽ lộ dải màu
Giao nộp	Đúng codec và profile mà hồ sơ yêu cầu	Tự ý đổi sang codec "tốt hơn"
Lưu trữ dài hạn	Codec phổ biến, có tài liệu công khai; HEVC hoặc codec dựng	Codec độc quyền của một hãng thiết bị — mười năm nữa có thể không còn phần mềm mở được

8.25 Định dạng chứa và cấu trúc tệp

8.25.1 Bộ chứa không phải codec

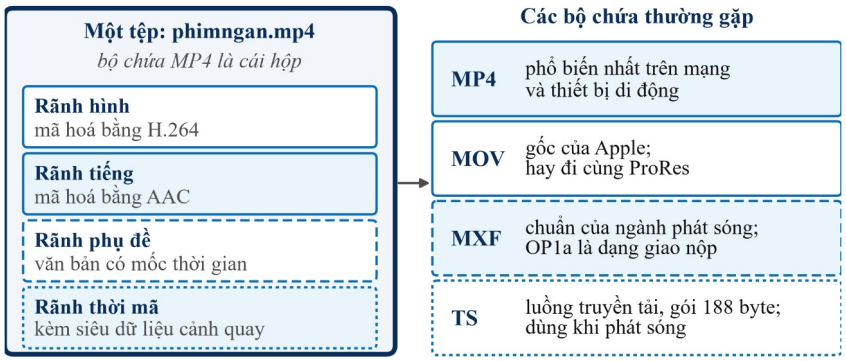
Đây là nguyên nhân của câu hỏi thường gặp trong phòng máy: "vì sao tệp MP4 mở được trên máy này mà không mở được trên máy kia" [11], [12].

Hình 8.61 là hộp thoại thông tin phương tiện của VLC, liệt kê riêng từng luồng của một tệp: luồng hình mã hoá H.264 và luồng tiếng mã hoá AAC, nằm trong cùng một bộ chứa.



Hình 8.61 Thông tin codec của từng luồng trong một tệp MP4 hiển thị bởi VLC

Nguồn: ảnh chụp màn hình do người soạn thực hiện (VLC 3.0.24), 05/11/2026.



Hình 8.62 Một tệp video gồm những gì, và các bộ chứa thường gặp

Có thể hình dung đơn giản rằng bộ chứa là chiếc hộp, còn codec là cách đóng gói món hàng bên trong. Đuôi tệp chỉ cho biết loại hộp mà không cho biết cách đóng gói: một hộp MP4 có thể chứa hình mã hoá bằng H.264, HEVC hoặc AV1, và để biết chắc thì phải mở thông số của tệp ra kiểm tra.

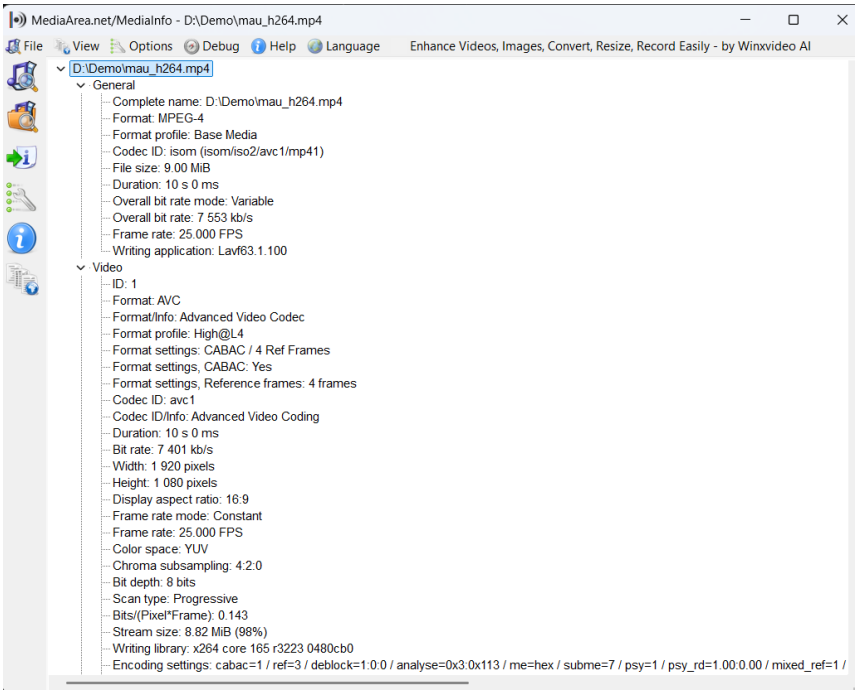
Bảng 8.51 Bốn bộ chứa thường gặp

Bộ chứa	Chuẩn	Đặc điểm	Dùng cho
MP4	ISO/IEC 14496-14, dựa trên ISO base media file format	Nhẹ, phổ biến nhất; hỗ trợ phát dần khi tải	Video trên mạng, thiết bị di động
MOV	Định dạng của Apple, cùng gốc với MP4	Linh hoạt hơn MP4 về số rãnh và loại dữ liệu	Hậu kỳ, hay đi cùng ProRes
MXF	SMPTE ST 377-1; dạng OP1a là ST 378	Mang được rất nhiều siêu dữ liệu chuẩn hoá; thiết kế cho phát sóng	Giao nộp cho đài, lưu trữ chuyên nghiệp
TS	ISO/IEC 13818-1	Gói đều 188 byte, chịu lỗi tốt, ghép được nhiều chương trình	Phát sóng DVB-T2, truyền qua đường có lỗi

8.25.2 Bên trong một tệp có gì

- Một hoặc nhiều *rãnh hình*, mỗi rãnh được mã hoá bằng một codec.
- Một hoặc nhiều *rãnh tiếng*, có thể dùng codec khác với rãnh hình.
- *Rãnh phụ đề*, gồm văn bản kèm mốc thời gian hoặc ảnh bitmap kèm mốc thời gian.
- *Rãnh thời mã* và siêu dữ liệu, gồm ngày quay, máy quay, ống kính, thông số ghi, không gian màu và phạm vi mức.
- *Bảng chỉ mục*, là bảng tra vị trí từng khung trong tệp để phục vụ việc tua nhanh; vì vậy một tệp bị ngắt giữa chừng khi ghi thường không mở được do bảng chỉ mục chưa kịp được viết.

Hình 8.63 minh hoạ cách phần mềm MediaInfo hiển thị cấu trúc bên trong một tệp MP4: nhánh General mô tả bộ chứa, nhánh Video nêu codec, profile và level, tốc độ bit, kiểu lấy mẫu con màu và độ sâu bit.



Hình 8.63 Thông số bộ chứa và codec của một tệp MP4 (H.264, 1920 × 1080, 4:2:0, 8 bit) trong MediaInfo

Nguồn: ảnh chụp màn hình do người soạn thực hiện (MediaInfo 26.10), 05/11/2026.

GHI NHỚ 8.23 — BA HỆ QUẢ THỰC TẾ

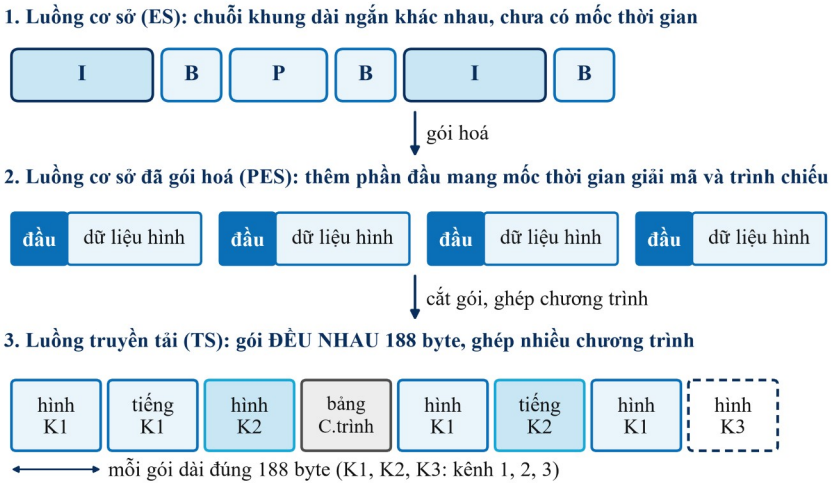
① Việc đổi đuôi tệp không làm đổi codec bên trong; chẳng hạn đổi .mkv thành .mp4 chỉ khiến hệ điều hành nhận diện sai, còn máy vẫn không mở được tệp.

② Việc tách và ghép rãnh giữa các bộ chứa không cần mã hoá lại, nên không làm mất chất lượng và thực hiện rất nhanh; đây là thao tác thường xuyên khi chuẩn bị bản giao nộp.

③ Để biết một tệp thực sự chứa gì, cần dùng công cụ phân tích, trong đó MediaInfo là công cụ miễn phí phổ biến nhất, thay vì suy đoán qua đuôi tệp hay biểu tượng.

8.25.3 Từ luồng cơ sở tới luồng truyền tải

Có ba mức đóng gói, và việc phân biệt được ba mức này là điều kiện để hiểu phần phát sóng ở Phần E [13].



Hình 8.64 Ba mức đóng gói: luồng cơ sở, gói hoá và luồng truyền tải

Bảng 8.52 Ba mức đóng gói

Mức	Tên	Nội dung
①	Luồng cơ sở (ES)	Kết quả thô của bộ mã hoá: một chuỗi khung dài ngắn khác nhau, chưa có mốc thời gian
②	Luồng cơ sở đã gói hoá (PES)	Thêm phần đầu mang MỐC THỜI GIAN giải mã và mốc thời gian trình chiếu — nhờ hai mốc này mà hình và tiếng khớp nhau
③a	Luồng chương trình (PS)	Gói dài tuỳ ý, dùng cho môi trường ÍT LỖI như đĩa DVD
③b	Luồng truyền tải (TS)	Gói ĐỀU NHAU 188 byte, ghép được nhiều chương trình; dùng cho môi trường NHIỀU LỖI như phát sóng

$$R_{\text{TS}} = \frac{188}{204} R_{\text{ch}} \quad ; \quad N_{\text{pkt}} = \frac{R_{\text{TS}}}{188 \times 8} \tag{8.32}$$

Con số 188 byte không ngẫu nhiên: gói đủ nhỏ để việc mất một gói không gây hỏng nhiều, đồng thời đủ lớn để phần đầu gói không chiếm tỉ lệ quá cao. Khi thêm mã sửa lỗi ngoài, mỗi gói 188 byte trở thành 204 byte.

8.26 Bài thực hành số 3 — codec, bộ chứa và chất lượng

Mục tiêu

Sau bài này, người đọc đọc được thông số đầy đủ của một tệp video bằng công cụ phân tích, phân biệt được bộ chứa với codec bằng thực nghiệm, và cảm nhận được khác biệt giữa codec dựng với codec phát hành khi dựng thực tế.

Thiết bị và phần mềm

- MediaInfo hoặc công cụ phân tích tệp tương đương (miễn phí).
- FFmpeg để tách, ghép và chuyển mã.
- Một phần mềm dựng.

– Năm tệp video có nguồn gốc khác nhau: một tệp quay bằng điện thoại, một tệp quay bằng máy quay, một tệp tải từ mạng, một tệp tư liệu cũ của đài nếu xin được, và một tệp do chính mình xuất ra.

Nội dung thực hành

Bảng 8.53 Bốn phần của bài thực hành số 3

Phần	Việc phải làm	Kết quả cần rút ra
1. Đọc thông số năm tệp	Dùng công cụ phân tích, lập một bảng đối chiếu năm tệp theo mười cột: bộ chứa, codec hình, profile, level, độ phân giải, tỉ lệ điểm ảnh, tần số hình, kiểu quét, phạm vi mức, cấu trúc lấy mẫu con màu.	Nhận ra rằng đuôi tệp giống nhau không có nghĩa là bên trong giống nhau. Chỉ ra ít nhất hai tệp có cùng đuôi nhưng khác codec hoặc khác profile.
2. Tách và ghép rãnh	Tách rãnh hình và rãnh tiếng khỏi một tệp MP4, ghép chúng lại vào một bộ chứa MOV mà KHÔNG mã hoá lại. So sánh dung lượng và thông số trước sau.	Codec không đổi, chất lượng không đổi, chỉ cái hộp đổi. Đây là bằng chứng thực nghiệm cho phân biệt ở mục 8.25.
3. Thử đổi đuôi tệp	Đổi đuôi một tệp MKV thành MP4 rồi thử mở bằng ba phần mềm khác nhau. Ghi lại phần mềm nào mở được, phần mềm nào không.	Giải thích vì sao có phần mềm mở được: chúng đọc phần đầu tệp để nhận dạng bộ chứa thật, chứ không tin vào đuôi.
4. Cảm nhận codec dựng	Chuyển một đoạn tư liệu 2 phút sang ProRes 422 hoặc DNxHR SQ. Đưa cả bản gốc và bản chuyển vào phần mềm dựng, thử tua nhanh, cắt, đặt hiệu ứng. Đo thời gian phản hồi và so dung lượng hai tệp.	Bản codec dựng nặng gấp nhiều lần nhưng thao tác mượt hơn hẳn. Nhận ra đánh đổi: dung lượng đổi lấy công máy.

Kết quả cần có

Bảng đối chiếu mười cột của năm tệp; ảnh chụp thông số trước và sau khi ghép lại bộ chứa; bảng ghi kết quả thử đổi đuôi tệp trên ba phần mềm; và bảng so sánh dung lượng cùng thời gian phản hồi giữa bản gốc và bản codec dựng. Kèm một đoạn kết luận: với cấu hình máy đang có, nên chọn quy trình dựng nào.

8.27 Giao diện truyền dẫn cơ sở

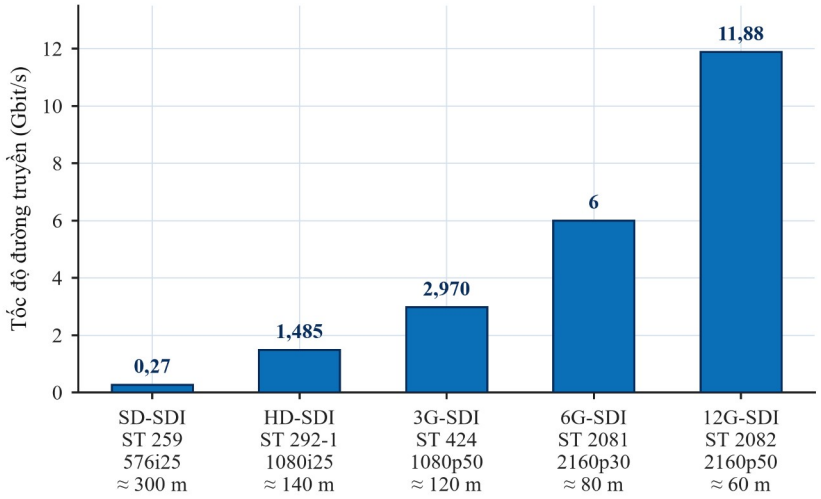
8.27.1 Vì sao trong trường quay lại KHÔNG nén

Trong Phần C và Phần D, cả chương tập trung vào nén; đến đây cần nêu một điều tưởng như mâu thuẫn: trong trường quay, người ta cố gắng không nén tín hiệu.

Lý do gồm ba điểm. Thứ nhất, nén sinh ra *độ trễ* vì bộ mã hoá phải gom đủ dữ liệu mới xử lý được, trong khi ở truyền hình trực tiếp, độ trễ giữa hình và tiếng hoặc giữa các máy quay là không chấp nhận được. Thứ hai, tín hiệu trong trường quay đi qua rất nhiều thiết bị như bàn trộn, bộ tạo chữ và bộ chỉnh màu, nên nếu

mỗi lần đều nén rồi giải nén thì tổn hao thế hệ tích luỹ rất nhanh. Thứ ba, khoảng cách trong một trường quay chỉ vài chục mét, đủ ngắn để truyền tín hiệu chưa nén.

8.27.2 Họ giao diện SDI



(a) Tốc độ, định dạng chở được và cự ly thực dụng của họ SDI

Tiêu chí	SDI	HDMI
Nơi dùng	Trường quay, thiết bị chuyên dụng	Thiết bị dân dụng, màn hình, máy chiếu
Đầu cắm, cự ly	BNC khoá chặt; hàng chục tới hàng trăm mét	Cắm ma sát; dưới 10 m với cáp đồng
Băng thông tối đa	11,88 Gbit/s (12G-SDI)	48 Gbit/s (HDMI 2.1)
Bảo vệ nội dung	Không có	HDCP: thiết bị ghi có thể bị từ chối tín hiệu
Tiếng nhúng	Tới 16 kênh	Tới 8 kênh, tuỳ phiên bản

(b) SDI so với HDMI

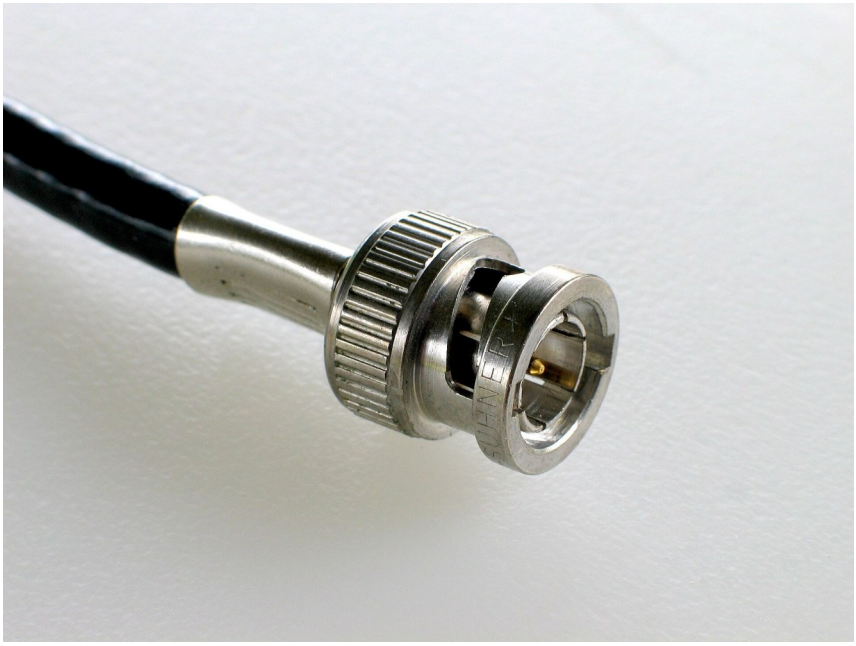
Hình 8.65 Họ giao diện SDI theo tốc độ đường truyền, và so sánh với HDMI

Bảng 8.54 Họ giao diện SDI

Tên	Chuẩn SMPTE	Tốc độ	Chở được	Cự ly thực dụng
SD-SDI	ST 259	270 Mbit/s	SD 576i25 hoặc 480i	≈ 300 m
HD-SDI	ST 292-1	1,485 Gbit/s	HD 1080i25, 720p50	≈ 140 m
3G-SDI	ST 424	2,970 Gbit/s	Full HD 1080p50	≈ 120 m
6G-SDI	ST 2081	6 Gbit/s	UHD 2160p30	≈ 80 m
12G-SDI	ST 2082	11,88 Gbit/s	UHD 2160p50 và 2160p60	≈ 60 m

Cự ly trong bảng là giá trị thực dụng với cáp đồng trục chất lượng tốt, không phải giới hạn tuyệt đối. Quy luật cần ghi nhớ là tốc độ càng cao thì cáp càng ngắn; vì vậy một trường quay UHD không thể chỉ thay tín hiệu trên hệ thống dây cũ mà phải tính lại toàn bộ sơ đồ đi dây, hoặc chuyển sang truyền trên hạ tầng IP [15].

Hình 8.66 là đầu nối BNC đực gắn trên cáp đồng trục, loại đầu nối khoá xoay 75 Ω thường dùng cho tín hiệu SDI trong trường quay.



Hình 8.66 Đầu nối BNC đực trên cáp đồng trục — giao diện vật lý của SDI

Nguồn: Jonas Bergsten, Public domain, Wikimedia Commons (truy cập 05/11/2026).

8.27.3 SDI mang được những gì

Một sợi SDI không chỉ chở hình mà trong khoảng xoá đã nêu ở mục 8.12, còn nhúng được các thành phần sau:

- Tới 16 kênh tiếng số, đủ cho một chương trình có nhiều ngôn ngữ và nhiều bản trộn.
- Thời mã, đồng bộ với từng khung hình.
- Phụ đề và dữ liệu điều khiển máy quay.
- Siêu dữ liệu cảnh quay, gồm tên máy, thông số ống kính và cân bằng trắng.

Khả năng nhúng này tạo nên ưu thế thực dụng của SDI: một sợi cáp với một đầu cắm khoá chặt mang trọn tín hiệu của một máy quay. Trong một trường quay có sáu máy, hệ thống chỉ cần sáu sợi cáp thay vì hàng chục.

8.27.4 HDMI và DisplayPort

Bảng 8.55 SDI so với HDMI

Tiêu chí	SDI	HDMI
Nơi dùng	Thiết bị chuyên dụng, trường quay	Thiết bị dân dụng, màn hình, máy chiếu
Đầu cắm	BNC, xoay khoá chặt	Cắm ma sát, dễ tuột
Cự ly thực dụng	Hàng chục tới hàng trăm mét	Dưới 10 mét với cáp đồng thường
Băng thông tối đa	11,88 Gbit/s với 12G-SDI	48 Gbit/s với HDMI 2.1
Bảo vệ nội dung	Không có	HDCP — thiết bị ghi hình chuyên dụng có thể bị từ chối tín hiệu
Tiếng nhúng	Tới 16 kênh	Tới 8 kênh ở HDMI 1.x và tới 32 kênh từ HDMI 2.0 trở lên

Hình 8.67 cho thấy đầu nối HDMI loại A đực với mười chín chân tiếp xúc, khác hẳn đầu nối BNC của SDI về kích thước và cơ chế khoá.



Hình 8.67 Đầu nối HDMI loại A đực

Nguồn: Evan-Amos, Public domain, Wikimedia Commons (truy cập 05/11/2026).

CẢNH BÁO 8.8 — MỘT SỰ CỐ HAY GẶP KHI THỰC HÀNH

Khi cắm đầu ra HDMI của một đầu phát đĩa hoặc một hộp thu truyền hình vào thiết bị ghi hình mà màn hình đen hoặc báo lỗi, nguyên nhân gần như luôn là HDCP chứ không phải hỏng cáp.

HDCP là cơ chế bảo vệ bản quyền, không phải lỗi kỹ thuật, và không nên tìm cách vượt qua. Nếu cần ghi lại nội dung có bản quyền cho mục đích giảng dạy thì phải xin nguồn tư liệu hợp pháp.

8.27.5 Đồng bộ chuẩn và bộ chuyển đổi

Hai thiết bị nữa luôn có mặt trong một trường quay SDI nhưng ít được nhắc tới trong sách vở, mặc dù thiếu chúng thì hệ thống không thể vận hành.

- **Bộ phát tín hiệu đồng bộ chuẩn.** Mọi máy quay và mọi thiết bị trong trường quay phải quét hình cùng một nhịp, nếu không bàn trộn sẽ bị giật một khung khi chuyển cảnh. Một máy phát tín hiệu chuẩn gửi nhịp ấy tới từng thiết bị, và thao tác đưa máy quay vào nhịp chung được gọi là khoá đồng bộ; trong hệ IP, vai trò này chuyển sang đồng hồ PTP nêu ở mục sau.
- **Bộ chuyển đổi và bộ khung hình.** Khi đưa vào hệ thống một nguồn không khoá đồng bộ được, như máy tính đồ hoạ, tín hiệu từ hiện trường về qua vệ tinh hoặc tệp phát lại, nguồn đó phải qua một bộ khung hình để đưa về đúng nhịp. Bộ chuyển đổi đảm nhận việc đổi định dạng, chẳng hạn từ HDMI sang SDI, từ 1080i sang 1080p, hoặc từ 59,94 sang 50 hình mỗi giây.

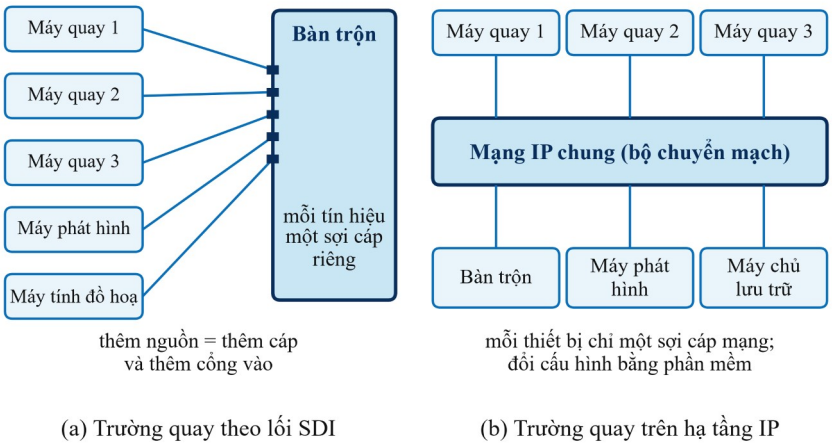
GHI NHỚ 8.24 — MỘT CHI TIẾT DỄ BỎ SÓT KHI LẬP DỰ TOÁN

Bộ chuyển đổi tần số hình, chẳng hạn từ 59,94 sang 50 hình mỗi giây, là thiết bị có giá thành cao nếu yêu cầu chất lượng cao, vì nó phải nội suy khung chứ không chỉ lặp khung.

Vì vậy các đài luôn cố gắng chuẩn hoá toàn bộ dây chuyền về một tần số hình duy nhất, và ở Việt Nam tần số ấy là 25 hình mỗi giây.

8.28 Video trên hạ tầng IP

8.28.1 Vì sao trường quay chuyển sang IP

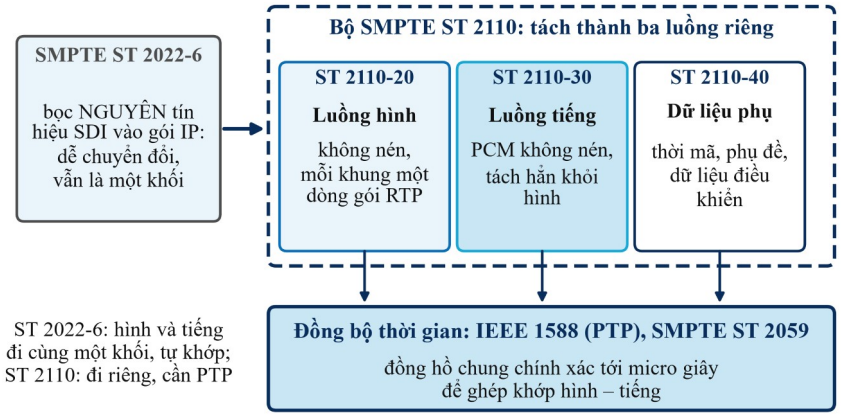


Hình 8.68 Trường quay theo lối SDI so với trường quay trên hạ tầng IP

Trong sơ đồ SDI, mỗi tín hiệu cần một sợi cáp riêng và mỗi đường đi là một chiều cố định từ nguồn tới đích, nên thêm một máy quay nghĩa là kéo thêm cáp và mở thêm cổng vào bàn trộn. Với trường quay lớn, số cáp lên tới hàng trăm và việc thay đổi cấu hình đòi hỏi nhiều công sức.

Trên hạ tầng IP, mỗi thiết bị chỉ cần một sợi cáp mạng nối vào một bộ chuyển mạch chung, và bất kỳ thiết bị nào cũng nhận được tín hiệu của bất kỳ thiết bị nào khác. Việc đổi cấu hình được thực hiện bằng phần mềm thay vì thay đổi dây.

8.28.2 Hai thế hệ chuẩn



Hình 8.69 Từ ST 2022-6 tới ST 2110, và vai trò của đồng bộ thời gian PTP

Bảng 8.56 Hai cách đưa video lên hạ tầng IP

Chuẩn	Cách làm	Ưu	Nhược
SMPTE ST 2022-6	Bọc NGUYÊN tín hiệu SDI vào gói IP, không tách gì cả	Chuyển đổi rất dễ; thiết bị SDI cũ dùng lại được qua bộ chuyển đổi	Vẫn là một khối; muốn lấy riêng tiếng thì phải nhận cả hình
Bộ SMPTE ST 2110	TÁCH hình, tiếng và dữ liệu phụ thành ba luồng riêng biệt	Lấy riêng từng luồng; băng thông chỉ dùng cho thứ thật sự cần	Hình và tiếng không còn tự khớp; bắt buộc phải có đồng hồ chung

Bảng 8.57 Các phần chính của bộ ST 2110

Số hiệu	Nội dung
ST 2110-10	Khung tổng thể và mô hình đồng bộ thời gian
ST 2110-20	Luồng HÌNH không nén, đóng gói theo RTP
ST 2110-30	Luồng TIẾNG PCM không nén
ST 2110-40	Luồng DỮ LIỆU PHỤ: thời mã, phụ đề, dữ liệu điều khiển

8.28.3 Đồng bộ thời gian

Khi ba luồng đi riêng trên một mạng gói, chúng đến đích ở những thời điểm hơi khác nhau. Để ghép lại đúng khớp hình và tiếng, mọi thiết bị phải dùng chung một đồng hồ có độ chính xác tới mức micro giây, và đó là nhiệm vụ của giao thức thời gian chính xác theo IEEE 1588, được ngành phát sóng cụ thể hoá trong SMPTE ST 2059 [16] [17].

> **GHI NHỚ 8.25 — ĐIỀU CẦN NHỚ VỀ TRƯỜNG QUAY IP**
>
> Trong hệ SDI, đồng bộ là điều đương nhiên vì hình và tiếng đi cùng một sợi nên không thể lệch nhau.
>
> Trong hệ IP, đồng bộ trở thành một hệ thống riêng cần được thiết kế, mua sắm thiết bị và bảo trì; đây là chi phí ẩn mà các dự án chuyển đổi thường bỏ sót khi lập dự toán.

8.28.4 NDI và SRT — hai giải pháp thực dụng

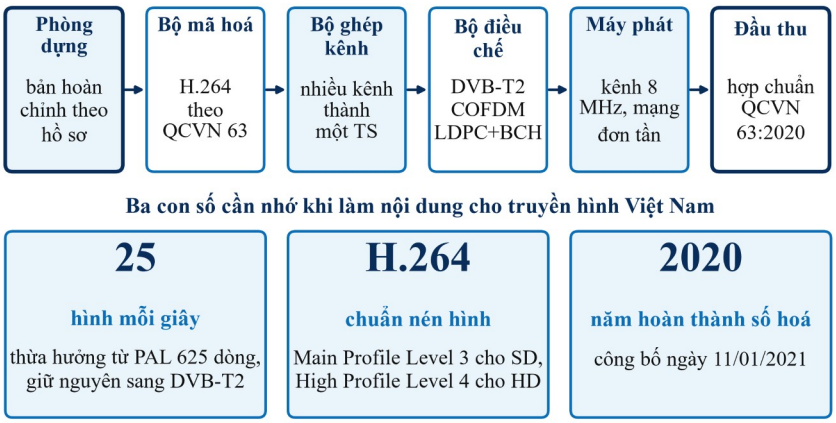
Bảng 8.58 Ba mức truyền video trên mạng

Giải pháp	Mức nén	Yêu cầu mạng	Hợp với
ST 2110	Không nén	Mạng chuyên dụng 10 Gbit/s trở lên, có PTP	Đài truyền hình lớn, trường quay xây mới
NDI	Nén nhẹ, độ trễ rất thấp	Mạng gigabit thông thường	Sản xuất nhỏ, phòng thu của trường, sự kiện trực tuyến
SRT	Dùng tín hiệu đã nén sâu; tự sửa lỗi mất gói	Internet công cộng	Truyền hình lưu động, đưa tín hiệu từ hiện trường về đài

Với điều kiện của một trường quay nhỏ, NDI là lựa chọn thực tế nhất: giao thức này chạy được trên mạng máy tính sẵn có, phần mềm phần lớn miễn phí, và đủ để dựng một dây chuyền sản xuất nhiều máy quay.

8.29 Phát sóng số mặt đất ở Việt Nam

8.29.1 Từ phòng dựng tới nhà khán giả



Hình 8.70 Dây chuyền phát sóng số mặt đất ở Việt Nam, và ba con số cần nhớ

Bản hoàn chỉnh rời phòng dựng đi qua năm khâu trước khi tới màn hình khán giả: mã hoá, ghép kênh, điều chế, phát sóng và giải mã ở đầu thu. Mỗi khâu đặt ra một ràng buộc đối với người làm nội dung, và mục này trình bày từng ràng buộc đó.

Hình 8.71 là ăng-ten thu truyền hình mặt đất kiểu Yagi–Uda lắp ngoài trời, loại ăng-ten định hướng thường dùng để thu tín hiệu truyền hình mặt đất.



Hình 8.71 Ăng-ten Yagi–Uda thu truyền hình mặt đất

Nguồn: Sajad-HasanAhmadi, CC BY-SA 4.0, Wikimedia Commons (truy cập 05/11/2026).

8.29.2 Ghép kênh

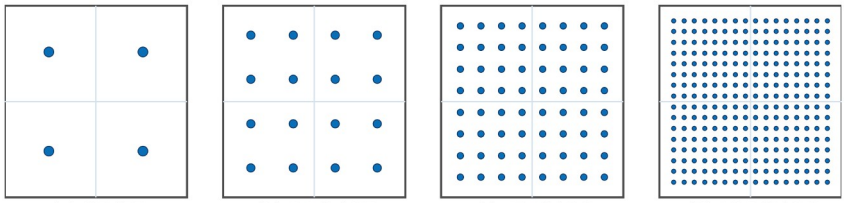
Một kênh tần số 8 MHz không chỉ chở một chương trình mà chở nhiều chương trình cùng lúc. Bộ ghép kênh nhận các luồng truyền tải của từng chương trình, trộn thành một luồng duy nhất và chèn thêm các bảng thông tin dịch vụ để đầu thu biết trong luồng có những thành phần nào.

$$n_{\text{prog}} = \left\lfloor \frac{R_{\text{mux}} - R_{\text{ovh}}}{R_{\text{prog}}} \right\rfloor \tag{8.33}$$

Trong (8.33), n_{prog} là số chương trình ghép được, R_{mux} là dung lượng ích của kênh, R_{ovh} là tốc độ dành cho các bảng thông tin dịch vụ và phần phụ trợ, còn R_{prog} là tốc độ bit của một chương trình.

Vì tổng dung lượng là cố định, mức cấp cho mỗi chương trình là kết quả của một quyết định kinh doanh chứ không thuần tuý là quyết định kỹ thuật. Thêm một kênh HD nghĩa là phải giảm tốc độ bit của các kênh còn lại, và đây là lý do chất lượng hình của một số kênh trên sóng mặt đất thấp hơn mong đợi dù bản gốc rất tốt.

8.29.3 Điều chế DVB-T2



(a) QPSK — 2 bit mỗi ký hiệu; (b) 16-QAM — 4 bit mỗi ký hiệu; (c) 64-QAM — 6 bit mỗi ký hiệu; (d) 256-QAM — 8 bit mỗi ký hiệu

Tham số	Các giá trị	Đánh đổi
Chòm sao	QPSK, 16-QAM, 64-QAM, 256-QAM	dày hơn thì chở nhiều bit hơn nhưng cần tỉ số tín hiệu trên nhiễu cao, vùng phủ hẹp
Mã sửa lỗi	LDPC + BCH, tỉ lệ 1/2 tới 5/6	tỉ lệ cao thì chở nhiều dữ liệu hơn nhưng kém chịu nhiễu
FFT	1k, 2k, 4k, 8k, 16k, 32k sóng mang con	lớn thì chịu được phản xạ xa hơn nhưng nhạy với dịch tần do chuyển động
Khoảng bảo vệ	1/128 tới 1/4 độ dài ký hiệu	dài thì máy phát trong mạng đơn tần đặt xa nhau hơn nhưng tốn dung lượng

Hình 8.72 Chòm sao điều chế và bốn tham số chính của DVB-T2

DVB-T2 dùng phương pháp ghép kênh phân chia theo tần số trực giao có mã hoá, quen gọi tắt là COFDM. Thay vì một sóng mang duy nhất, tín hiệu được rải lên hàng nghìn sóng mang con, mỗi sóng mang chở một phần dữ liệu ở tốc độ rất thấp, nhờ đó hệ thống chịu được phản xạ nhiều đường, vốn là điều kiện thường gặp ở đô thị Việt Nam [18].

Hình 8.73 là mặt sau của một đầu thu DVB-T2, trong đó có cổng vào ăng-ten cùng các cổng ra HDMI, âm thanh số và mạng.



Hình 8.73 Mặt sau đầu thu truyền hình số mặt đất DVB-T2: cổng ăng-ten vào, HDMI và S/PDIF ra

Nguồn: Raimond Spekking, CC BY-SA 4.0, Wikimedia Commons (truy cập 05/11/2026).

$$R_{\text{ch}} = R_{\text{sym}} \cdot \log_2 M \cdot r_c \cdot \frac{1}{1+\Delta_g} \tag{8.34}$$

Trong (8.34), R_{ch} là dung lượng ích của kênh, R_{sym} là tốc độ ký hiệu, M là số điểm của chòm sao điều chế ($\log_2 M$ là số bit trên một ký hiệu), r_c là tỉ lệ mã sửa lỗi và Δ_g là độ dài khoảng bảo vệ tính theo tỉ lệ của độ dài ký hiệu.

Bảng 8.59 Bốn tham số chính của DVB-T2 theo ETSI EN 302 755

Tham số	Các giá trị	Đánh đổi
Chòm sao điều chế	QPSK, 16-QAM, 64-QAM, 256-QAM	Chòm sao dày chở được nhiều bit hơn nhưng cần tỉ số tín hiệu trên nhiễu cao hơn, nên vùng phủ hẹp lại
Tỉ lệ mã sửa lỗi	Từ 1/2 tới 5/6, dùng LDPC ghép với BCH	Tỉ lệ cao thì ít bit dành cho sửa lỗi nên chở được nhiều dữ liệu hơn, nhưng kém chịu nhiễu
Kích thước FFT	1k, 2k, 4k, 8k, 16k, 32k sóng mang con	FFT lớn thì ký hiệu dài, chịu được phản xạ xa hơn, nhưng nhạy hơn với dịch tần do chuyển động
Khoảng bảo vệ	Từ 1/128 tới 1/4 độ dài ký hiệu	Khoảng bảo vệ dài cho phép các máy phát trong mạng đơn tần đặt xa nhau hơn, nhưng tốn dung lượng

Một đặc điểm quan trọng của DVB-T2 là *mạng đơn tần*, trong đó nhiều máy phát cùng phát trên một tần số. Ở hệ tương tự, hai máy phát cùng tần số sẽ triệt tiêu nhau, còn ở COFDM có khoảng bảo vệ, tín hiệu từ máy phát thứ hai được coi như một tia phản xạ và được cộng vào thay vì gây nhiễu lẫn nhau. Do đó, việc phủ sóng một vùng rộng chỉ cần một kênh tần số.

8.29.4 QCVN 63:2020/BTTTT

Quy chuẩn này là quy chuẩn kỹ thuật quốc gia về thiết bị thu truyền hình số mặt đất DVB-T2, ban hành kèm Thông tư 17/2020/TT-BTTTT. Quy chuẩn áp

dụng cho thiết bị thu tín hiệu truyền hình số mặt đất không khoá mã, bao gồm cả đầu thu rời lẫn tivi tích hợp [24].

Bảng 8.60 Các yêu cầu chính của QCVN 63:2020/BTTTT

Nhóm yêu cầu	Nội dung
Giải mã hình MPEG-2	Các độ phân giải 720 × 576, 544 × 576, 480 × 576, 352 × 576 ở 25 hình/giây, khuôn 4:3 và 16:9
Giải mã hình H.264 — độ nét tiêu chuẩn	Main Profile Level 3 theo ISO/IEC 14496-10, cùng các độ phân giải và khuôn hình như trên
Giải mã hình H.264 — độ nét cao	High Profile Level 4 theo ISO/IEC 14496-10, phù hợp ETSI TS 101 154
Tốc độ bit hình tối thiểu	600 kbit/s ở độ phân giải 720 × 576
Giải mã tiếng	MPEG-1 Layer II hai kênh; HE-AAC Level 2 cho một và hai kênh, Level 4 cho tới 5.1 kênh, ở 48 kHz
Tham số cao tần	FFT từ 1k tới 32k; QPSK tới 256-QAM; LDPC ghép BCH với tỉ lệ mã 1/2 tới 5/6; khoảng bảo vệ 1/128 tới 1/4
Chuẩn viện dẫn	ETSI EN 302 755 (DVB-T2), EN 300 468 (thông tin dịch vụ), TS 101 154 (mã hoá hình và tiếng), EN 300 743 (phụ đề)

GHI NHỚ 8.26 — RÀNG BUỘC RẤT CỤ THỂ VỚI NGƯỜI LÀM NỘI DUNG

Quy chuẩn không nêu HEVC, nghĩa là một đầu thu hợp chuẩn ở Việt Nam chỉ bảo đảm giải mã được MPEG-2 và H.264.

Hệ quả trực tiếp là bản giao nộp cho truyền hình mặt đất phải được mã hoá bằng H.264, thường là High Profile Level 4 cho kênh HD. Việc nộp bản HEVC với lý do nén tốt hơn sẽ tạo ra một bản mà nhiều khán giả không xem được.

Đây là ví dụ điển hình cho một nguyên tắc xuyên suốt Phần E: kỹ thuật được chọn phải theo quy chuẩn của nơi phát hành, không theo phương án tốt nhất về mặt lý thuyết.

8.29.5 Lộ trình số hoá truyền hình Việt Nam

Đề án số hoá truyền dẫn, phát sóng truyền hình mặt đất đến năm 2020 được phê duyệt theo Quyết định 2451/QĐ-TTg ngày 27 tháng 12 năm 2011. Đề án triển khai trong chín năm, chia thành nhiều giai đoạn theo nhóm tỉnh thành, và hoàn thành trong năm 2020; việc hoàn thành được công bố chính thức ngày 11 tháng 1 năm 2021 [26].

- Việt Nam đi thẳng vào công nghệ DVB-T2 và bỏ qua thế hệ DVB-T mà nhiều nước đã triển khai trước đó, nhờ vậy tiết kiệm được một vòng đầu tư thiết bị và đạt dung lượng cao hơn trên cùng một kênh tần số.
- Vùng phủ sóng truyền hình số mặt đất tăng từ khoảng một nửa dân số năm 2011 lên khoảng 80 phần trăm dân cư vào năm 2020.

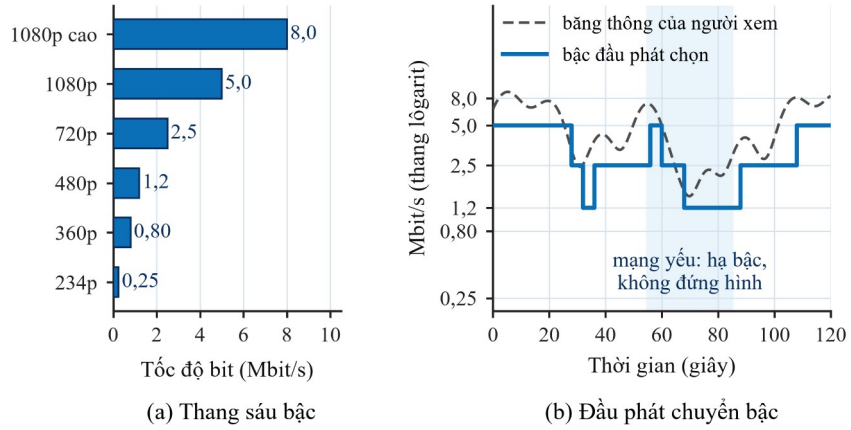
– Tần số hình vẫn giữ 25 hình mỗi giây, thừa hưởng từ hệ PAL 625 dòng; vì vậy mọi bản giao nộp trong nước đều ở 25 hoặc 50 hình mỗi giây, và người làm nghề ở Việt Nam không phải xử lý vấn đề thời mã bỏ khung.

8.30 Phát trực tuyến và OTT

8.30.1 Vì sao không dùng một tốc độ bit duy nhất

Phát sóng và phát trực tuyến khác nhau ở một điểm căn bản. Phát sóng là truyền một chiều, trong đó đài phát đi và mọi máy thu nhận cùng một tín hiệu, nên máy thu ở nơi sóng yếu phải chấp nhận chất lượng kém. Phát trực tuyến là truyền hai chiều, trong đó mỗi người xem có một đường truyền riêng với băng thông riêng, và băng thông ấy thay đổi từng phút.

Nếu chỉ chọn một tốc độ bit duy nhất thì người dùng mạng yếu không xem được, hoặc người dùng mạng mạnh phải xem ở chất lượng thấp. Giải pháp là mã hoá sẵn nhiều bản ở nhiều tốc độ bit khác nhau, rồi để đầu phát tự chọn bản phù hợp.



(a) Thang sáu bậc (b) Đầu phát chuyển bậc

Hình 8.74 Thang tốc độ bit của một dịch vụ phát trực tuyến, và cách đầu phát chuyển bậc

8.30.2 Cơ chế truyền theo HTTP thích ứng

– Chương trình được cắt thành các *đoạn* ngắn, thường từ 2 tới 10 giây, và mỗi đoạn là một tệp riêng.
– Mỗi đoạn được mã hoá thành nhiều *bậc* chất lượng khác nhau.
– Một *bản kê* liệt kê tất cả các bậc cùng địa chỉ của từng đoạn.
– Đầu phát tải bản kê, đo tốc độ tải của đoạn vừa nhận, rồi chọn bậc cho đoạn tiếp theo, do đó bậc có thể thay đổi ở ranh giới của mỗi đoạn.

Bảng 8.61 Hai chuẩn truyền theo HTTP thích ứng

Chuẩn	Nguồn gốc	Đặc điểm
HLS	Của Apple; nay là RFC 8216	Bản kê dạng văn bản đơn giản; được hỗ trợ rộng nhất, nhất là trên thiết bị di động
MPEG-DASH	ISO/IEC 23009-1	Bản kê dạng XML, linh hoạt hơn; độc lập với codec nên dùng được cả AV1

Cơ chế này giải thích vì sao video trên mạng mờ đi rồi nét lại thay vì đứng hình. Từ góc nhìn của người xem, hình mờ đi trong chốc lát dễ chấp nhận hơn nhiều so với đứng hình để chờ tải, và cơ chế trên được thiết kế đúng theo cảm nhận ấy [21], [22].

8.30.3 Dựng một thang tốc độ bit

Bảng 8.62 Một thang tốc độ bit tham khảo cho nội dung Full HD

Bậc	Độ phân giải	Tốc độ bit hình	Dành cho
1	416 × 234	0,25 Mbit/s	Mạng di động rất yếu
2	640 × 360	0,8 Mbit/s	Mạng di động thông thường
3	854 × 480	1,2 Mbit/s	Mạng cố định chậm
4	1280 × 720	2,5 Mbit/s	Mạng cố định thông thường
5	1920 × 1080	5,0 Mbit/s	Mạng cáp quang
6	1920 × 1080	8,0 Mbit/s	Bậc chất lượng cao

Nguyên tắc dựng thang là các bậc liền kề cách nhau khoảng một phẩy năm tới hai lần về tốc độ bit (trừ bậc thấp nhất, vốn phải đủ nhẹ cho mạng rất yếu), để việc chuyển bậc không quá đột ngột nhưng cũng không quá dày. Bậc thấp nhất phải đủ nhẹ để chạy được trên mạng di động kém nhất mà dịch vụ nhắm tới.

8.30.4 Mạng phân phối nội dung và độ trễ

Nếu mọi người xem đều lấy dữ liệu từ một máy chủ duy nhất thì máy chủ ấy không đáp ứng nổi và người ở xa sẽ tải rất chậm. Mạng phân phối nội dung giải quyết vấn đề này bằng cách đặt bản sao của từng đoạn ở nhiều điểm gần người xem.

Bảng 8.63 Độ trễ của các cách phát

Cách phát	Độ trễ điển hình	Nguyên nhân
Truyền hình mặt đất DVB-T2	3 – 6 giây	Mã hoá, ghép kênh, điều chế, và bộ đệm ở đầu thu
Phát trực tuyến thông thường	20 – 45 giây	Đoạn dài cộng nhiều tầng bộ đệm ở máy chủ và ở đầu phát
Phát trực tuyến độ trễ thấp	3 – 8 giây	Đoạn ngắn, truyền theo từng phần nhỏ ngay khi mã hoá xong
Hội nghị truyền hình	Dưới 0,5 giây	Không dùng khung B, đoạn rất ngắn, chấp nhận chất lượng thấp hơn

GHI NHỚ 8.27 — MỘT CON SỐ CÓ Ý NGHĨA KINH TẾ

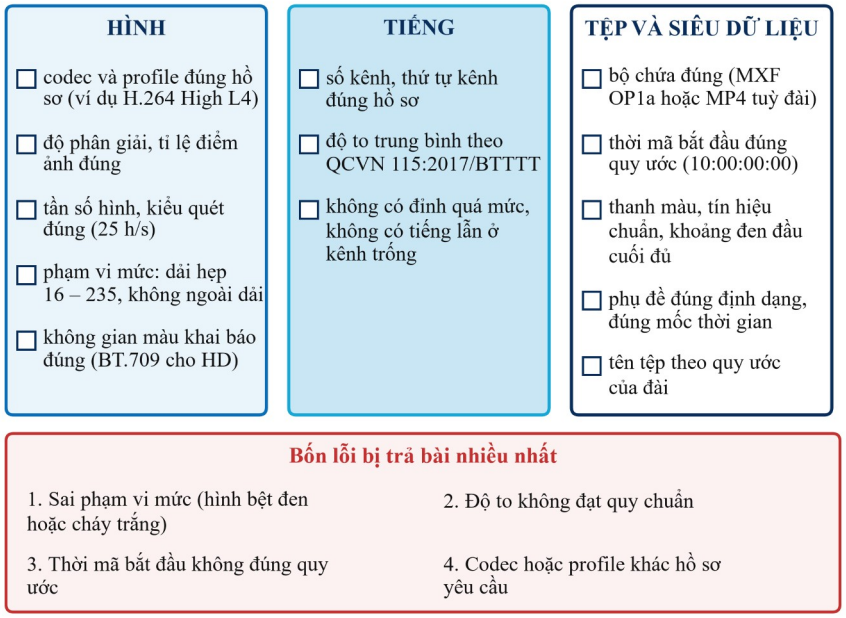
Với một dịch vụ có hàng triệu lượt xem, việc giảm được 30 phần trăm tốc độ bit ở cùng chất lượng đồng nghĩa với giảm 30 phần trăm chi phí băng thông. Vì vậy các nền tảng lớn đầu tư mạnh vào AV1 và VVC, dù chi phí mã hoá cao gấp hàng chục lần.

Điều này cũng trả lời cho câu hỏi về mục đích của codec mới khi mạng đã nhanh: codec mới nhằm giảm chi phí băng thông hơn là làm hình xem nét hơn.

8.31 Chuẩn giao nộp và kiểm tra chất lượng

8.31.1 Hồ sơ giao nộp là gì

Mỗi đài và mỗi nền tảng đều có một văn bản quy định bản hoàn chỉnh phải đáp ứng những yêu cầu nào. Văn bản ấy gọi là hồ sơ giao nộp, và nó là điều kiện nghiệm thu chứ không chỉ là gợi ý.



Hình 8.75 Danh mục kiểm trước khi giao nộp một bản hoàn chỉnh

Bảng 8.64 Một hồ sơ giao nộp điển hình cho kênh HD của đài Việt Nam

Hạng mục	Yêu cầu điển hình
Bộ chứa	MXF OP1a, hoặc MP4 tuỳ quy định từng đài
Codec hình	H.264 High Profile Level 4
Độ phân giải và khuôn hình	1920 × 1080, điểm ảnh vuông, khuôn 16:9
Tần số hình và kiểu quét	25 hình/giây, quét liên tục hoặc xen kẽ theo quy định kênh
Cấu trúc lấy mẫu con màu và độ sâu bit	4:2:0 8 bit, hoặc 4:2:2 10 bit với chương trình chất lượng cao
Phạm vi mức và không gian màu	Dải hẹp 16 – 235; BT.709

Hạng mục	Yêu cầu điển hình
Tốc độ bit	8 – 12 Mbit/s, thường yêu cầu chế độ tốc độ bit thay đổi có giới hạn đỉnh
Nhóm ảnh	Nhóm ĐÓNG, độ dài 12 hoặc 15 khung
Tiếng	Hai kênh hoặc 5.1; độ to trung bình theo QCVN 115:2017/BTTTT
Thời mã bắt đầu	10:00:00:00
Phần dẫn đầu	Thanh màu và tín hiệu âm thanh chuẩn, tiếp theo là bảng thông tin chương trình, rồi đếm ngược, rồi khoảng đen
Phụ đề	Tệp riêng đúng định dạng quy định, mốc thời gian khớp thời mã của hình
Tên tệp	Theo quy ước đặt tên của đài

8.31.2 Nối với Chương 6: độ to

Phần tiếng của hồ sơ giao nộp không phải là nội dung mới. QCVN 115:2017/BTTTT về mức độ to của tín hiệu âm thanh trong truyền hình đã được trình bày ở Chương 6, cùng với cách đo theo ITU-R BS.1770 và cách chuẩn hoá theo EBU R 128 [25].

Nội dung của hai chương gặp nhau ở đúng điểm này: một bản hoàn chỉnh có hình đạt chuẩn nhưng tiếng không đạt độ to quy định vẫn bị trả về, và ngược lại. Do đó, người dựng phải kiểm tra cả hai.

8.31.3 Bốn lỗi bị trả bài nhiều nhất

Bảng 8.65 Bốn lỗi thường gặp và cách phòng

Lỗi	Biểu hiện	Cách phòng
Sai phạm vi mức	Hình bệt đen và cháy trắng, hoặc xám xịt mất tương phản	Kiểm bằng waveform, không kiểm bằng mắt; khai báo rõ dải hẹp khi xuất
Độ to không đạt quy chuẩn	Tiếng to hoặc nhỏ hơn các chương trình khác trên cùng kênh	Đo bằng công cụ đo độ to trước khi xuất; xem lại Chương 6
Thời mã bắt đầu sai	Bản nộp bắt đầu ở 00:00:00:00 thay vì 10:00:00:00	Đặt thời mã bắt đầu ngay khi tạo dòng thời gian, không sửa ở cuối
Codec hoặc profile khác hồ sơ	Đài không giải mã được, hoặc giải mã được nhưng không phát được	Đọc kỹ hồ sơ TRƯỚC khi bắt đầu dựng; xuất thử một đoạn 30 giây và gửi kiểm trước khi xuất bản đầy đủ

QUY TẮC 8.3 — KIỂM BẰNG CÔNG CỤ, KHÔNG KIỂM BẰNG MẮT

Mắt không nhận ra được profile, tỉ lệ điểm ảnh hay thời mã bắt đầu, và tai cũng không nhận ra được độ to trung bình tính theo LUFS.

Mọi bản giao nộp đều phải chạy qua một công cụ phân tích tệp và một công cụ đo độ to. Nhiều đài lớn còn dùng hệ thống kiểm tự động để trả lại tệp không đạt mà không cần người xem.

Cách có chi phí thấp nhất để tránh bị trả bài là đối chéo bài giữa các nhóm và kiểm cho nhau theo danh mục.

8.31.4 Giao nộp cho đài khác giao nộp cho nền tảng

Hai nơi phát hành có triết lý khác nhau, nên hồ sơ giao nộp cũng khác nhau về bản chất, không chỉ về các con số.

Bảng 8.66 Hai kiểu hồ sơ giao nộp

Tiêu chí	Đài truyền hình	Nền tảng phát trực tuyến
Triết lý	Bản nộp đi THẲNG vào dây chuyền phát sóng, nên phải đúng tuyệt đối mọi thông số	Bản nộp là bản GỐC để nền tảng tự mã hoá lại thành nhiều bậc, nên ưu tiên chất lượng cao
Codec yêu cầu	Đúng codec và profile quy định — thường H.264 High Profile ở Việt Nam	Càng ít nén càng tốt; nhiều nền tảng nhận cả ProRes và DNxHR
Tốc độ bit	Bị khống chế cả trần lẫn sàn để vừa khe trong bộ ghép kênh	Càng cao càng tốt trong giới hạn dung lượng tải lên
Phần dẫn đầu	Bắt buộc: thanh màu, tín hiệu âm thanh chuẩn, đếm ngược, thời mã bắt đầu 10:00:00:00	Thường KHÔNG có; nội dung bắt đầu ngay từ khung đầu tiên
Độ to	Theo QCVN 115:2017/BTTTT	Theo quy định riêng của nền tảng, thường quanh −14 tới −16 LUFS, khác với truyền hình
Phụ đề	Tệp riêng đúng định dạng của đài	Tệp riêng, thường theo định dạng phổ thông; có thể nộp nhiều thứ tiếng

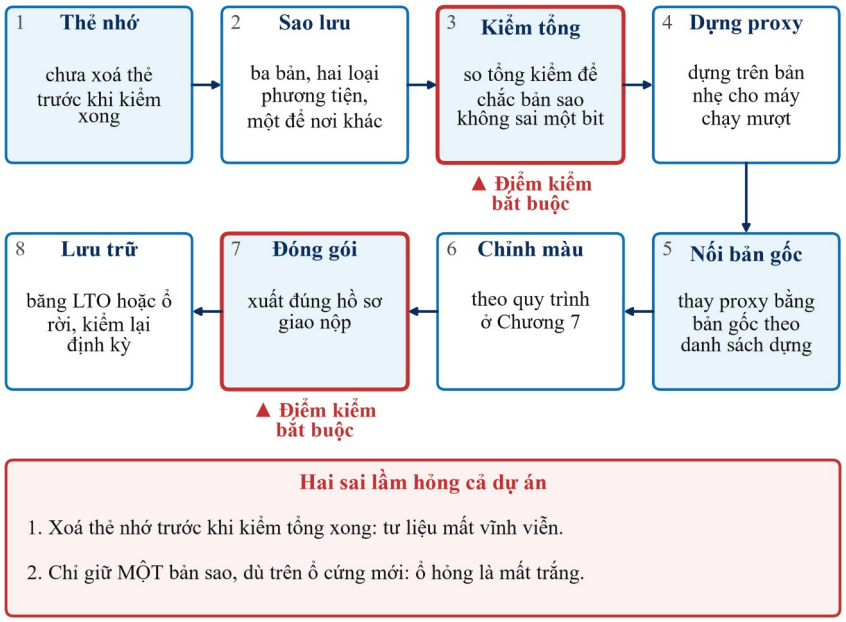
CẢNH BÁO 8.9 — MỘT SAI LẦM RẤT PHỔ BIẾN

Một lỗi thường gặp là lấy nguyên bản đã nén sẵn cho đài rồi tải lên nền tảng trực tuyến. Bản ấy đã qua một lần nén sâu, nền tảng lại nén thêm lần nữa, nên tổn hao thế hệ cộng dồn và kết quả kém hơn rõ rệt.

Cách đúng là giữ một bản trung gian chất lượng cao, rồi từ bản ấy xuất riêng cho từng nơi phát hành theo hồ sơ của nơi đó.

8.32 Quy trình sản xuất trên tệp

8.32.1 Toàn cảnh



Hình 8.76 Quy trình sản xuất trên tệp từ thẻ nhớ tới lưu trữ, với hai điểm kiểm bắt buộc

Phần này không đòi hỏi kỹ thuật phức tạp, nhưng lại là khâu mà người mới vào nghề dễ mất tư liệu nhất. Khi một dự án hỏng vì mất tư liệu thì mọi kiến thức về codec và bộ chứa đều không còn giá trị.

8.32.2 Đổ thẻ và sao lưu

QUY TẮC 8.4 — QUY TẮC BA – HAI – MỘT

BA bản sao của mọi tư liệu gốc.

Tư liệu được lưu trên hai loại phương tiện khác nhau, ví dụ ổ cứng và băng LTO, hoặc ổ cứng và dịch vụ lưu trữ đám mây.

Trong số các bản đó, một bản được để ở địa điểm khác, phòng khi cháy, ngập hoặc mất trộm.

Quy tắc này áp dụng cho tư liệu gốc, không áp dụng cho mọi tệp trung gian.

Sau khi sao chép, bắt buộc phải kiểm tổng kiểm. Tổng kiểm là phép tính một chuỗi ký tự đại diện cho toàn bộ nội dung tệp; nếu bản sao và bản gốc cho cùng một chuỗi thì bản sao chính xác tới từng bit. Sao chép mà không kiểm tra thì kết quả chỉ dựa trên sự phỏng đoán.

CẢNH BÁO 8.10 — HAI SAI LẦM HỎNG CẢ DỰ ÁN

① Xoá thẻ nhớ trước khi kiểm tổng kiểm xong. Nếu bản sao lỗi mà thẻ đã bị xoá thì tư liệu mất vĩnh viễn và không có cách khôi phục.

② Chỉ giữ một bản sao, kể cả khi đó là ổ cứng mới mua. Ổ cứng có thể hỏng mà không báo trước, và xác suất hỏng trong năm đầu không nhỏ.

Cả hai sai lầm đều không tốn chi phí để tránh, chỉ đòi hỏi kỷ luật làm việc.

8.32.3 Dựng bằng proxy và nối lại bản gốc

Với tư liệu UHD hoặc tư liệu quay nhiều máy, dựng trực tiếp trên bản gốc làm máy hoạt động chậm. Cách làm chuẩn là tạo bản proxy nhẹ, dựng trên proxy, rồi để phần mềm tự thay lại bản gốc trước khi chỉnh màu và xuất.

Danh sách quyết định dựng là tệp mô tả mọi lát cắt, gồm cảnh nào được dùng và từ thời mã nào tới thời mã nào. Nhờ tệp này, bản dựng chuyển được giữa các phần mềm khác nhau và việc thay proxy bằng bản gốc diễn ra tự động.

8.32.4 Lưu trữ dài hạn

Bảng 8.67 Ba vấn đề của lưu trữ dài hạn

Vấn đề	Nội dung	Cách xử lý
Phương tiện xuống cấp	Ổ cứng cơ hỏng sau vài năm; ổ thể rắn mất dữ liệu nếu để lâu không cấp điện; băng LTO bền hơn nhưng cần điều kiện bảo quản	Kiểm lại định kỳ hằng năm; sao chép sang phương tiện mới sau mỗi năm tới bảy năm
Định dạng lỗi thời	Codec độc quyền của một hãng có thể không còn phần mềm mở được sau mười năm	Chọn codec phổ biến, có tài liệu công khai; tránh codec riêng của một dòng máy
Mất siêu dữ liệu	Tệp còn nhưng không ai nhớ nó là cảnh gì, quay ngày nào, thuộc dự án nào	Đặt tên tệp và tổ chức thư mục theo quy ước ngay từ đầu; ghi kèm một tệp mô tả bằng văn bản thuần

8.32.5 Bài thực hành số 4 — giao nộp một bản hoàn chỉnh

Mục tiêu của bài thực hành là người đọc tự xuất được một bản hoàn chỉnh đúng hồ sơ giao nộp, tự kiểm theo danh mục, và kiểm chéo được bài của người khác.

Bảng 8.68 Bốn phần của bài thực hành số 4

Phần	Việc phải làm	Kết quả cần rút ra
1. Tập đổ thẻ và sao lưu	Đổ một thẻ nhớ sang ổ cứng theo quy tắc ba – hai – một; tính và so tổng kiểm của bản gốc với bản sao. Thử làm hỏng cố ý một bản sao rồi kiểm lại.	Thấy tận mắt rằng tổng kiểm phát hiện được cả sai lệch một bit, trong khi mắt và dung lượng tệp thì không.
2. Xuất theo hồ sơ	Tự đặt ra một hồ sơ giao nộp giả định, dựa theo hồ sơ điển hình ở mục 8.31. Xuất một phim ngắn của mình đúng theo hồ sơ đó, kể cả thanh màu, đếm ngược, thời mã bắt đầu và khoảng đen cuối.	Nắm được quy trình đặt thời mã bắt đầu và dựng phần dẫn đầu — hai việc mà phần mềm không tự làm.
3. Tự kiểm theo danh mục	Chạy tệp vừa xuất qua công cụ phân tích và công cụ đo độ to. Điền đầy đủ danh mục kiểm ở mục 8.31, đánh dấu từng mục đạt hoặc không đạt.	Phát hiện được ít nhất một mục không đạt ngay lần xuất đầu tiên — điều này gần như luôn xảy ra và chính là bài học.
4. Kiểm chéo	Đổi bài với một người khác cùng làm. Kiểm bài của họ theo cùng danh mục, lập biên bản ghi rõ mục nào không đạt và số đo thực tế.	Học cách viết một biên bản kỹ thuật: nêu số đo, nêu yêu cầu, nêu chênh lệch — không nhận xét cảm tính.

Kết quả cần có gồm tệp bản hoàn chỉnh, ảnh chụp kết quả tổng kiểm, danh mục kiểm đã điền đầy đủ của chính mình, và biên bản kiểm chéo bài của người khác.

TÓM TẮT CHƯƠNG 8

Phần A — từ ảnh tới con số

- Số hoá hình ảnh gồm ba phép rời rạc hoá: theo không gian thành điểm ảnh, theo thời gian thành khung hình và theo biên độ thành mức lượng tử, và mọi dòng thông số của một tệp video chỉ phản ánh kết quả của ba phép ấy.
- Định lý lấy mẫu áp dụng nguyên vẹn cho hai chiều không gian; việc vi phạm định lý sinh ra răng cưa và vân moiré, và biện pháp khắc phục là bộ lọc quang thông thấp đặt trước cảm biến.
- Tỉ lệ điểm ảnh giải thích vì sao 720 × 576 mẫu hiển thị đúng khuôn 4:3, trong khi định dạng HD và UHD dùng điểm ảnh vuông.
- Độ phân giải chỉ có ý nghĩa khi xét trong quan hệ với khoảng cách xem: bốn lần chiều cao màn hình là ngưỡng của Full HD, còn 4K yêu cầu ngồi gần gấp đôi.
- Dải hẹp 16 – 235 và dải rộng 0 – 255 là hai quy ước khác nhau, và việc gán sai gây ra hiện tượng đen bị bệt, trắng bị cháy hoặc hình bị xám. Mười bit là mức tối thiểu cho ảnh log và HDR.

Phần B — cấu trúc tín hiệu

- Việc tách Y′CbCr là bước quyết định cho mọi khâu phía sau, vì chi tiết tập trung ở độ chói còn hai thành phần màu tương đối phẳng nên có thể lấy mẫu thưa.
- Mắt nhạy với chi tiết độ chói hơn chi tiết màu, và đây là cơ sở của lấy mẫu con màu.
- Bốn cấu trúc 4:4:4, 4:2:2, 4:2:0 và 4:1:1 khác nhau ở hình dạng vùng dùng chung, không chỉ ở số mẫu. Công đoạn hậu kỳ có khoá màu đòi hỏi 4:2:2 trở lên.
- Trường không phải khung: hai trường mới thành một khung. Ký hiệu 1080i25 và 1080/50i là một, còn 720/25p không tồn tại.
- Ba bộ ký hiệu Y/U/V, Y/R−Y/B−Y và Y′/Cb/Cr không đồng nghĩa với nhau, vì chúng khác nhau ở hệ số tỉ lệ và mức dịch.
- Ký hiệu J:a:b mô tả một khối tham chiếu 4 × 2 điểm ảnh, trong đó J = 4, a là số mẫu màu ở dòng một và b là số mẫu màu có thêm ở dòng hai; số mẫu mỗi điểm ảnh là s = (J + a + b)/J. Chữ số 0 trong 4:2:0 không có nghĩa là không có màu.

Phần C — nguyên lý nén

- Có bốn loại dư thừa gồm không gian, thời gian, thống kê và cảm nhận. Ba loại đầu có thể khai thác mà không mất thông tin, còn loại thứ tư chính là ranh giới của nén có tổn hao.
- Biến đổi cosin rời rạc tự nó không làm mất thông tin; mất mát xảy ra ở bước lượng tử hoá hệ số bằng một bảng có trọng số thị giác.
- Vỡ khối và nhiễu muỗi có hình dạng riêng, nhận ra được bằng mắt, và đều là dấu hiệu của tốc độ bit đặt quá thấp.
- Khung I chiếm 8 phần trăm số khung nhưng khoảng một phần ba số bit. Nhóm ảnh dài nén hiệu quả hơn, song kém linh hoạt khi cắt.
- Bộ đệm là thành phần dung hoà giữa tốc độ bit cố định và tốc độ bit thay đổi; nếu vượt quy định về bộ đệm thì đầu thu của khán giả sẽ bị giật.
- Nén là dự đoán rồi mã hoá phần bất ngờ, vì entropy của phần dư sau dự đoán nhỏ hơn rất nhiều so với entropy của chính điểm ảnh. Mọi bộ nén đều thực hiện bốn việc: dự đoán, biến đổi, lượng tử hoá và mã hoá entropy.
- Codec thế hệ mới nén kỹ hơn vì mỗi bước được thực hiện tinh vi hơn; bộ mã hoá phải thử rất nhiều phương án để làm J = D + λR nhỏ nhất nên tốn nhiều tính toán, còn bộ giải mã chỉ làm theo quyết định ghi trong luồng bit.
- Nén video là nén ảnh cộng thêm tầng dự đoán liên khung: chuyển động được dò theo khối bằng SAD, phần dư được mã hoá bằng cùng đường ống DCT, và bộ mã hoá phải chứa một bộ giải mã để luôn dự đoán từ khung tái tạo.

Phần D — chuẩn nén và định dạng

- Mỗi thế hệ codec đạt cùng chất lượng ở khoảng một nửa tốc độ bit của thế hệ trước, đổi lại chi phí mã hoá tăng nhiều lần.

- Profile quy định các công cụ được phép dùng, còn Level quy định cỡ tối đa được phép; QCVN 63:2020/BTTTT cũng có dạng yêu cầu này.
- Codec để dựng chỉ nén trong khung nên tệp lớn nhưng máy xử lý nhẹ, còn codec để phát hành nén liên khung nên tệp nhỏ nhưng máy xử lý nặng khi tua.
- Bộ chứa khác với codec. Việc đổi đuôi tệp không làm đổi codec, và việc tách ghép rãnh giữa các bộ chứa không cần mã hoá lại.

Phần E — truyền dẫn, phát sóng và quy trình

- Trong trường quay, tín hiệu thường được giữ ở dạng không nén vì hai lý do là độ trễ và tổn hao thế hệ. SDI mang trọn hình, tiếng, thời mã và siêu dữ liệu trong một sợi.
- ST 2110 tách hình, tiếng và dữ liệu thành ba luồng riêng, đổi lại hệ thống phải có đồng bộ thời gian PTP như một hệ thống riêng.
- Truyền hình mặt đất Việt Nam dùng DVB-T2 và hoàn thành số hoá năm 2020. QCVN 63:2020/BTTTT quy định H.264 Main Profile Level 3 cho SD và High Profile Level 4 cho HD; quy chuẩn không nêu HEVC.
- Phát trực tuyến dùng thang tốc độ bit và truyền theo HTTP thích ứng, nên hình mờ đi rồi nét lại thay vì đứng hình.
- Quy tắc ba – hai – một và kiểm tổng kiểm là hai biện pháp có chi phí thấp nhất nhưng bảo vệ được nhiều dự án nhất.

8.33 Câu hỏi và bài tập ôn tập

A. Câu hỏi lý thuyết

Phần A

1. Ba phép rời rạc hoá của việc số hoá hình ảnh là gì, và mỗi phép cho ra đại lượng nào trong dòng thông số của một tệp video?
2. Vì sao chi tiết mịn bị hỏng trước khi giảm số điểm ảnh, còn vùng phẳng thì gần như không đổi?
3. Vân moiré khác răng cưa ở điểm nào, và có thể xử lý bằng ba cách nào tại hiện trường?
4. Bộ lọc quang thông thấp có tác dụng gì và cái giá phải trả là gì?
5. Vì sao định dạng SD PAL có 720 × 576 mẫu nhưng vẫn hiển thị đúng khuôn 4:3?
6. Một màn hình cao 60 cm hiển thị Full HD thì người xem nên ngồi cách bao xa để tận dụng hết độ phân giải?
7. Vì sao khi tăng độ phân giải từ Full HD lên 4K, số điểm ảnh tăng bốn lần chứ không phải hai lần?
8. Dải hẹp và dải rộng khác nhau ra sao, và gán sai gây ra hai kiểu hỏng nào?
9. Vì sao ảnh log bắt buộc phải ghi ở 10 bit trở lên?
10. Dither có thêm thông tin vào ảnh hay không, và nếu không thì nó có tác dụng gì?

11. Hạt phim và điểm ảnh có ba khác biệt cốt lõi nào?
12. Vì sao người quay phim vẫn nói tới "vai sáng" dù đang dùng máy quay số?

Phần B

1. Vì sao việc tách tín hiệu thành Y′CbCr là bước quyết định cho việc nén?
2. Hai bộ hệ số BT.601 và BT.709 khác nhau ở đâu, và điều gì xảy ra khi giải mã bằng nhầm ma trận?
3. Vì sao mức "không màu" trong tín hiệu số 8 bit nằm ở 128 chứ không phải ở 0?
4. Ký hiệu 4:2:0 nghĩa là gì? Giải thích từng con số.
5. Cấu trúc 4:2:0 và 4:1:1 có cùng số mẫu, vậy vì sao 4:2:0 lại chiếm ưu thế?
6. Vì sao khoá màu đòi tư liệu 4:2:2 trở lên?
7. Hai quy ước về vị trí mẫu màu khác nhau ở điểm nào và gây ra lỗi gì?
8. Quét xen kẽ ra đời nhằm giải quyết bài toán nào, và cái giá phải trả là gì?
9. Ký hiệu 1080i25 và 1080/50i có cùng nghĩa không? Giải thích.
10. Vì sao "720/25p" không phải là một định dạng phát sóng có thật?
11. Quét xen kẽ gây ra ba lỗi hình nào, và nguyên nhân của từng lỗi là gì?
12. So sánh ba phương pháp khử xen kẽ về chất lượng và chi phí tính toán.
13. Vì sao truyền hình Việt Nam dùng 25 hình/giây trong khi Bắc Mỹ dùng 30?
14. Tần số 29,97 hình/giây từ đâu ra?
15. So sánh ba cách đưa phim 24 hình/giây lên 25 hình/giây.
16. Vì sao tín hiệu số vẫn giữ khoảng xoá dù không còn tia điện tử?
17. Định dạng HD 1080 có bao nhiêu dòng tổng và bao nhiêu dòng tích cực?
18. Thời mã bỏ khung có làm mất hình không? Giải thích.
19. Vì sao ba bộ ký hiệu Y/U/V, Y/R−Y/B−Y và Y′/Cb/Cr không đồng nghĩa với nhau?
20. Vẽ khối tham chiếu 4 × 2 và chỉ ra vị trí các mẫu Y′ cùng các mẫu màu của 4:4:4, 4:2:2, 4:2:0 và 4:1:1, từ đó suy ra số mẫu trên mỗi điểm ảnh bằng công thức $s = (J + a + b)/J$.
21. Khi một cạnh màu rơi vào giữa một nhóm điểm ảnh dùng chung mẫu màu, hiện tượng nào xảy ra sau khi lọc xuống và nội suy lên, và vì sao mất mát chỉ xuất hiện ở những vùng màu thay đổi nhanh?

Phần C và D

1. Nêu bốn loại dư thừa mà kỹ thuật nén khai thác và cho biết loại nào gây tổn hao.
2. Vì sao nén không tổn hao không thể cố định tốc độ bit?
3. Biến đổi cosin rời rạc có làm mất thông tin hay không, và mất mát xảy ra ở bước nào?
4. Vì sao bảng lượng tử hoá không đồng đều, và những ô nào được chia cho số lớn vì lý do gì?
5. Quét zig-zag phục vụ bước nào tiếp theo trong chuỗi nén?

6. Phân biệt vỡ khối, nhiễu muỗi và hiện tượng làm phẳng chi tiết theo hình dạng và nguyên nhân.
7. Ba loại khung I, P và B khác nhau như thế nào về tham chiếu và kích thước?
8. Vì sao nhóm ảnh đóng được chuẩn giao nộp ưa dùng?
9. Bù chuyển động làm giảm phần dư bằng cách nào, và đối với khối đứng yên thì kết quả ra sao?
10. Trong một nhóm ảnh 12 khung, khung I chiếm bao nhiêu phần trăm số khung và bao nhiêu phần trăm số bit?
11. Vì sao truyền hình trực tiếp hạn chế dùng khung B?
12. So sánh CBR, VBR và chế độ chất lượng cố định về ưu điểm, nhược điểm và phạm vi sử dụng.
13. Bộ đệm tràn và bộ đệm cạn gây ra hiện tượng gì đối với người xem?
14. Vì sao MPEG-7 không phải một chuẩn nén hình?
15. Khác biệt cốt lõi giữa macroblock của H.264 và CTU của HEVC là gì?
16. Vì sao HEVC không thay thế H.264 nhanh như dự đoán?
17. AV1 ra đời trong bối cảnh nào, và điểm mạnh mang tính quyết định của nó là gì?
18. Profile và Level trả lời hai câu hỏi khác nhau nào?
19. Vì sao không nên dựng trực tiếp trên tệp H.264 có tốc độ bit thấp?
20. Bộ chứa khác codec ở những điểm nào, và có ba hệ quả thực tế nào?
21. Vì sao luồng truyền tải dùng gói 188 byte có kích thước cố định, còn luồng chương trình thì không?
22. Giải thích bằng entropy vì sao phần dư sau dự đoán nén được nhiều hơn chính ảnh gốc, và bốn loại dư thừa tương ứng với phần "đoán được" hay phần "bất ngờ" nào của tín hiệu?
23. Vì sao codec thế hệ mới nén hiệu quả hơn nhưng bộ mã hoá lại chạy chậm hơn nhiều lần, và hàm giá $J = D + \lambda R$ có vai trò gì trong đó?
24. Trình bày ba cấp của nén liên khung gồm lấy hiệu trực tiếp, dò chuyển động theo khối và mã hoá phần dư, đồng thời chỉ ra mỗi cấp khắc phục nhược điểm nào của cấp trước.
25. Vì sao bộ mã hoá video phải chứa một bộ giải mã, và điều gì xảy ra nếu dự đoán từ khung gốc?
26. Vì sao khung B làm cho thứ tự mã hoá khác thứ tự hiển thị và gây ra độ trễ?

Phần E

1. Vì sao trong trường quay người ta cố gắng không nén tín hiệu?
2. Vì sao tốc độ SDI càng cao thì cự ly cáp càng ngắn?
3. Một sợi SDI mang được những gì ngoài hình?
4. HDCP là gì, và cơ chế này gây ra sự cố nào trong thực hành?
5. ST 2022-6 và ST 2110 khác nhau ở điểm nào về nguyên lý?
6. Vì sao trường quay IP bắt buộc phải có đồng hồ PTP, còn trường quay SDI thì không?

7. So sánh ST 2110, NDI và SRT về mức nén và yêu cầu đối với mạng.
8. Vì sao việc thêm một kênh HD vào bộ ghép kênh làm giảm chất lượng của các kênh còn lại?
9. Nêu bốn tham số chính của DVB-T2 và sự đánh đổi gắn với từng tham số.
10. Mạng đơn tần hoạt động được nhờ đặc điểm nào của COFDM?
11. QCVN 63:2020/BTTTT quy định đầu thu phải giải mã được những chuẩn nào, quy chuẩn có nêu HEVC hay không, và hệ quả đối với người làm nội dung là gì?
12. Việt Nam hoàn thành số hoá truyền hình mặt đất vào năm nào và lựa chọn chuẩn nào?
13. Vì sao phát trực tuyến phải mã hoá sẵn nhiều bậc chất lượng?
14. Vì sao video trên mạng mờ đi rồi nét trở lại thay vì đứng hình?
15. Nêu nguyên tắc dựng một thang tốc độ bit.
16. Vì sao một codec tiết kiệm băng thông hơn lại mang giá trị kinh tế trực tiếp đối với nhà cung cấp dịch vụ?
17. Nêu đầy đủ danh mục kiểm tra một bản giao nộp cho kênh HD.
18. Bốn lỗi bị trả bài nhiều nhất là gì, và cách phòng ngừa từng lỗi?
19. Quy tắc ba – hai – một quy định điều gì, và vì sao phải kiểm tra tổng kiểm?
20. Nêu ba vấn đề của lưu trữ dài hạn và cách xử lý từng vấn đề.

B. Bài tập tính toán

Bảng 8.69 Mười bài tập tính toán

#	Đề bài	Gợi ý
1	Tính tốc độ bit thô của tín hiệu Full HD 1080p50, 10 bit, cấu trúc 4:2:2. Một giờ tư liệu như vậy nặng bao nhiêu gigabyte?	Dùng công thức (8.15) rồi nhân với 3600 và chia cho 8
2	Một kênh DVB-T2 có dung lượng ích 22 Mbit/s. Nếu mỗi kênh HD cần 6 Mbit/s và mỗi kênh SD cần 2 Mbit/s, có thể ghép tối đa bao nhiêu kênh HD nếu vẫn phải giữ 4 kênh SD? Dành 1,5 Mbit/s cho các bảng thông tin dịch vụ.	Dùng công thức (8.34)
3	Một màn hình 55 inch theo đường chéo, khuôn 16:9. Tính chiều cao màn hình, rồi tính khoảng cách ngồi xem tối ưu cho Full HD và cho 4K.	Chiều cao bằng đường chéo nhân 0,49; dùng công thức (8.6)
4	Một tệp SD PAL có 720 × 576 mẫu và tỉ lệ điểm ảnh 1,422. Khuôn hiển thị của nó là bao nhiêu?	Dùng công thức (8.5)
5	Tín hiệu 8 bit dải hẹp có mã độ chói bằng 100. Nếu bị gán nhầm thành dải rộng, giá trị hiển thị tương ứng là bao nhiêu, và lệch bao nhiêu phần trăm?	Dùng công thức (8.10)

#	Đề bài	Gợi ý
6	Một cảnh quay 4K 2160p25, 10 bit, 4:2:2 được ghi bằng codec dựng ở tỉ số nén 5:1. Tính tốc độ bit và dung lượng của 20 phút tư liệu.	Tính tốc độ thô bằng (8.15), chia cho tỉ số nén, rồi dùng (8.22)
7	Một phim 100 phút quay ở 24 hình/giây được phát ở Việt Nam bằng cách tăng tốc. Phim dài bao nhiêu phút sau khi phát? Cao độ tiếng thay đổi bao nhiêu nửa cung?	Dùng công thức (8.19)
8	Một khối 8 × 8 sau lượng tử hoá có 12 hệ số khác 0. Nếu mỗi hệ số khác 0 cần trung bình 6 bit và ký hiệu kết thúc cần 4 bit, khối này chiếm bao nhiêu bit? So với 64 mẫu 8 bit chưa nén, tỉ số nén là bao nhiêu?	Đếm bit rồi dùng (8.22)
9	Một nhóm ảnh 12 khung có kích thước tương đối I = 100, P = 34, B = 13, theo cấu trúc I B B P B B P B B P B B. Tính tỉ lệ số bit của mỗi loại khung. So sánh với tỉ lệ số khung.	Cộng tổng rồi chia; đối chiếu với mục 8.17.4
10	Tính tỉ số dung lượng giữa một giờ tư liệu 8K 4320p50 chưa nén và một giờ tư liệu SD 576i25 chưa nén, cùng 10 bit 4:2:2. Kết quả nói lên điều gì về nhu cầu lưu trữ?	Dùng (8.15) cho cả hai rồi lập tỉ số
11	Tính số byte mỗi khung 8 bit của khung UHD 3840 × 2160 ở bốn cấu trúc 4:4:4, 4:2:2, 4:2:0, 4:1:1 và so với khung Full HD 1920 × 1080 cùng cấu trúc.	Mặt phẳng Y′ có 8 294 400 mẫu; dùng s = (J + a + b)/J. Đáp số 24 883 200; 16 588 800; 12 441 600; 12 441 600 byte.
12	Một dòng Cb gồm tám điểm ảnh: 30, 30, 30, 30, 30, 200, 200, 200. Lập dòng sau khi qua 4:2:2 (trung bình từng cặp, nhân bản khi nội suy) và qua 4:1:1 (trung bình từng nhóm 4), rồi tính tổng sai số tuyệt đối.	4:2:2 giữ 30, 30, 115, 200 và cho tổng sai số 170; 4:1:1 giữ 30 và 158 (làm tròn) và cho tổng sai số 254.
13	Một khối 64 × 64 được chia bằng cây tứ phân, mỗi khối con tối đa chia tiếp tới 16 × 16. Có bao nhiêu cách chia? Tính với $f(k) = 1 + f(k-1)^4$, $f(0) = 1$.	Có hai lần chia: f(1) = 2, f(2) = 17. Đáp số 17 cách.

C. Bài thảo luận

Bài 1 — Chọn codec cho một đài địa phương. Một đài truyền hình tỉnh muốn chuyển toàn bộ kho tư liệu sang lưu trữ số và đồng thời nâng cấp dây chuyền sản xuất trong điều kiện ngân sách có hạn. Người đọc lập luận để chọn codec lưu trữ, codec sản xuất và codec giao nộp phát sóng, đồng thời nêu rõ những ràng buộc đến từ QCVN 63:2020/BTTTT và những ràng buộc đến từ yếu tố kinh tế.

Bài 2 — SDI hay IP cho một trường quay nhỏ. Một cơ sở đào tạo dự định xây dựng một trường quay ba máy để thực hành. Người đọc so sánh

phương án SDI truyền thống với phương án dùng NDI trên mạng máy tính sẵn có về chi phí thiết bị, độ phức tạp vận hành, khả năng mở rộng và giá trị đào tạo, sau đó đề xuất một phương án và bảo vệ lựa chọn của mình.

Bài 3 — Vì sao 8K chưa phổ cập. Dựa trên các con số của mục 8.4 về khoảng cách ngồi xem, mục 8.14 về tốc độ bit thô và mục 8.23 về chi phí mã hoá, người đọc lập luận về những trường hợp độ phân giải 8K có ý nghĩa và những trường hợp không có ý nghĩa, đồng thời nêu ít nhất hai ứng dụng mà 8K thật sự cần thiết.

TÀI LIỆU THAM KHẢO CHƯƠNG 8

[1] ITU-R. Recommendation ITU-R BT.601: Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios. Geneva: International Telecommunication Union.

[2] ITU-R. Recommendation ITU-R BT.709: Parameter values for the HDTV standards for production and international programme exchange. Geneva: International Telecommunication Union.

[3] ITU-R. Recommendation ITU-R BT.2020: Parameter values for ultra-high definition television systems for production and international programme exchange. Geneva: International Telecommunication Union.

[4] ITU-R. Recommendation ITU-R BT.2100: Image parameter values for high dynamic range television for use in production and international programme exchange. Geneva: International Telecommunication Union.

[5] ITU-R. Recommendation ITU-R BT.500: Methodologies for the subjective assessment of the quality of television images. Geneva: International Telecommunication Union.

[6] ISO/IEC 13818-2 = ITU-T H.262. Information technology — Generic coding of moving pictures and associated audio information — Video (MPEG-2 Video).

[7] ISO/IEC 14496-10 = ITU-T H.264. Advanced Video Coding for generic audiovisual services.

[8] ISO/IEC 23008-2 = ITU-T H.265. High Efficiency Video Coding.

[9] ISO/IEC 23090-3 = ITU-T H.266. Versatile Video Coding.

[10] Alliance for Open Media. AV1 Bitstream and Decoding Process Specification.

[11] ISO/IEC 13818-1. MPEG-2 Systems: luồng chương trình và luồng truyền tải.

[12] ISO/IEC 14496-12 và 14496-14. ISO base media file format và định dạng MP4.

[13] SMPTE ST 377-1 và ST 378. Material Exchange Format (MXF) và dạng OP1a.

[14] SMPTE ST 12-1. Time and Control Code.

[15] SMPTE ST 259, ST 292-1, ST 424, ST 2081, ST 2082. Họ giao diện SDI.

[16] SMPTE ST 2022-6 và bộ SMPTE ST 2110 (-10, -20, -30, -40). Video chuyên nghiệp trên hạ tầng IP.

[17] IEEE 1588 và SMPTE ST 2059. Đồng bộ thời gian chính xác.

[18] ETSI EN 302 755. DVB-T2: Frame structure, channel coding and modulation for a second generation digital terrestrial television broadcasting system. Sophia Antipolis: European Telecommunications Standards Institute.

[19] ETSI TS 101 154. Specification for the use of video and audio coding in broadcasting applications based on the MPEG-2 transport stream.

[20] ETSI EN 300 468. Specification for Service Information in DVB systems.

[21] ISO/IEC 23009-1. Dynamic adaptive streaming over HTTP (MPEG-DASH).

[22] Pantos, R., & May, W. (2017). HTTP Live Streaming. RFC 8216. Internet Engineering Task Force.

[23] ETSI TS 101 154 (đã nêu ở [19]) và ETSI EN 300 468 (đã nêu ở [20]) được dùng trong thực tế cùng EN 302 755.

[24] Bộ Thông tin và Truyền thông (2020). QCVN 63:2020/BTTTT: Quy chuẩn kỹ thuật quốc gia về thiết bị thu truyền hình số mặt đất DVB-T2, ban hành kèm Thông tư số 17/2020/TT-BTTTT, Hà Nội.

[25] Bộ Thông tin và Truyền thông (2017). QCVN 115:2017/BTTTT: Quy chuẩn kỹ thuật quốc gia về mức âm thanh của nội dung truyền hình, Hà Nội.

[26] Thủ tướng Chính phủ (2011). Quyết định 2451/QĐ-TTg ngày 27 tháng 12 năm 2011 phê duyệt Đề án số hoá truyền dẫn, phát sóng truyền hình mặt đất đến năm 2020, Hà Nội.

[27] Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27(3), 379–423.

[28] Wiegand, T., Sullivan, G. J., Bjøntegaard, G., & Luthra, A. (2003). Overview of the H.264/AVC video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 13(7), 560–576.

[29] Sullivan, G. J., Ohm, J.-R., Han, W.-J., & Wiegand, T. (2012). Overview of the High Efficiency Video Coding (HEVC) standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 22(12), 1649–1668.

[30] Bross, B., Wang, Y.-K., Ye, Y., Liu, S., Chen, J., Sullivan, G. J., & Ohm, J.-R. (2021). Overview of the Versatile Video Coding (VVC) standard and its applications. *IEEE Transactions on Circuits and Systems for Video Technology*, 31(10), 3736–3764.

[31] Richardson, I. E. G. (2003). *H.264 and MPEG-4 Video Compression: Video Coding for Next-generation Multimedia*. Chichester: Wiley.

[32] Poynton, C. (2012). *Digital Video and HD: Algorithms and Interfaces* (2nd ed.). Waltham, MA: Morgan Kaufmann.