



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0052369

(51)^{2020.01} H04B 1/66

(13) B

(21) 1-2021-07232

(22) 15/02/2021

(86) PCT/US2021/018099 15/02/2021

(87) WO2021/202000 A1 07/10/2021

(30) 63/003,112 31/03/2020 US; 17/087,865 03/11/2020 US

(45) 27/10/2025 451

(43) 27/06/2022 411A

(73) TENCENT AMERICA LLC (US)

2747 Park Boulevard, Palo Alto, California 94306, United States of America

(72) CHOI, Byeongdoo (KR); LIU, Shan (US); WENGER, Stephan (DE).

(74) Công ty TNHH một thành viên Sở hữu trí tuệ VCCI (VCCI-IP CO.,LTD)

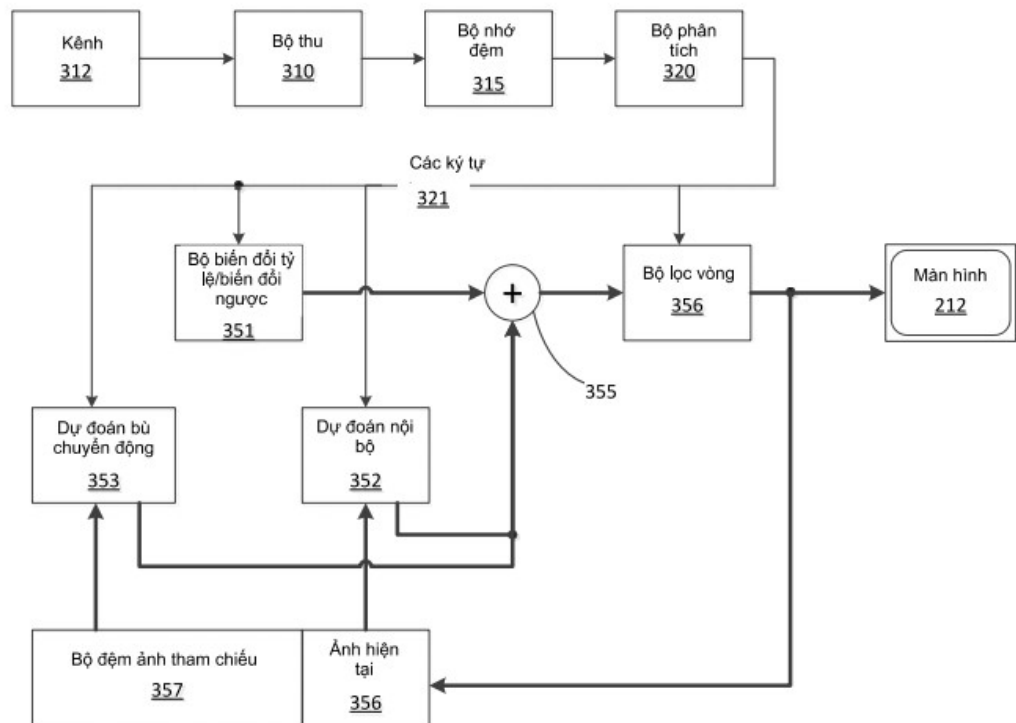
(54) PHƯƠNG PHÁP VÀ HỆ THỐNG MÁY TÍNH MÃ HÓA DỮ LIỆU VIDEO VÀ
PHƯƠNG TIỆN BẤT BIẾN ĐỌC ĐƯỢC BẰNG MÁY TÍNH

(21) 1-2021-07232

(57) Sáng chế đề cập đến phương pháp và hệ thống máy tính mã hóa dữ liệu video và phương tiện bất biến đọc được bằng máy tính. Dữ liệu video bao gồm ảnh hiện tại và một hoặc nhiều ảnh khác được thu. Cờ thứ nhất tương ứng với việc ảnh hiện tại có được tham chiếu bởi một hoặc nhiều ảnh khác theo thứ tự giải mã hay không được kiểm tra. Cờ thứ hai tương ứng với việc ảnh hiện tại có được xuất ra không; được kiểm tra. Dữ liệu video được giải mã dựa trên các giá trị tương ứng với cờ thứ nhất và cờ thứ hai.

FIG. 3

Bộ giải mã 210



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến lĩnh vực xử lý dữ liệu, và cụ thể hơn là đến mã hóa và giải mã video.

Tình trạng kỹ thuật của sáng chế

Việc mã hóa và giải mã video nhờ sử dụng dự đoán liên ảnh với việc bù chuyển động đã được biết đến trong nhiều thập kỷ. Video số không được nén có thể bao gồm một loạt ảnh, mỗi ảnh mà có kích thước không gian, ví dụ, là 1920 x 1080 mẫu độ chói và các mẫu sắc độ được kết hợp. Một loạt ảnh có thể có tốc độ ảnh cố định hoặc biến thiên (cũng thường được gọi là tốc độ khung), ví dụ, bằng 60 ảnh trên giây hoặc 60 Hz. Video không nén có các yêu cầu tốc độ bit quan trọng. Ví dụ, video 1080p60 4:2:0 tại 8 bit trên mẫu (độ phân giải mẫu độ chói 1920x1080 tại tốc độ khung 60 Hz) yêu cầu gần 1,5 Gbit/s băng thông. Mỗi giờ của video này yêu cầu nhiều hơn 600 Gbyte của không gian lưu trữ.

Một mục đích của việc mã hóa và giải mã video có thể là giảm dư lượng trong tín hiệu video đầu vào, thông qua việc nén. Việc nén có thể giúp làm giảm các yêu cầu băng thông hoặc không gian lưu trữ nêu trên, trong một vài trường hợp xuống bởi hai bậc độ lớn hoặc nhiều hơn. Cả việc nén tổn hao và nén không tổn hao, cũng như kết hợp của chúng có thể được sử dụng. Việc nén không tổn hao liên quan đến các kỹ thuật trong đó bản sao chính xác của tín hiệu gốc có thể được khôi phục từ tín hiệu gốc được nén. Khi sử dụng việc nén tổn hao, tín hiệu được khôi phục có thể không đồng nhất với tín hiệu gốc, nhưng độ méo giữa các tín hiệu gốc và tín hiệu được khôi phục là đủ nhỏ để tạo ra tín hiệu được khôi phục hữu ích cho việc áp dụng dự định. Trong trường hợp của video, việc nén tổn hao được sử dụng rộng rãi. Lượng biến dạng chấp nhận được phụ thuộc vào ứng dụng; ví dụ, những người dùng của các ứng dụng phát trực tuyến tiêu dùng nhất định có thể chịu được sự biến dạng cao hơn so với những người dùng của các ứng dụng

bài báo tivi. Việc đạt được tỷ lệ nén có thể phản ánh rằng: sự biến dạng cho phép/chấp nhận được cao hơn có thể mang lại các tỷ lệ nén cao hơn.

Bộ mã hóa video và bộ giải mã có thể sử dụng các kỹ thuật từ một vài hạng mục chung, bao gồm, ví dụ, bù chuyển động, biến đổi, lượng tử hóa, và mã hóa entropi, một vài trong số chúng sẽ được giới thiệu dưới đây.

Về mặt lịch sử, các bộ mã hóa và giải mã video có xu hướng hoạt động trên kích cỡ ảnh định trước mà, trong hầu hết các trường hợp, được xác định và cố định đối với chuỗi video được mã hóa (CVS-coded video sequence), Nhóm của các ảnh (GOP-Group of Pictures), hoặc khung thời gian liên ảnh. Ví dụ, trong MPEG-2, các thiết kế hệ thống đã được biết đến để thay đổi độ phân giải chiều ngang (và, liên quan đến, kích cỡ ảnh) phụ thuộc vào các yếu tố như hoạt động của các cảnh, nhưng chỉ tại các ảnh I, một cách điển hình đối với GOP. Việc lấy mẫu lại của các ảnh tham chiếu cho việc sử dụng của các độ phân giải khác nhau trong CVS được biết đến, ví dụ, từ ITU-T Rec.H.263 Phụ lục P. Tuy nhiên, ở đây kích cỡ ảnh không thay đổi, chỉ các ảnh tham chiếu được lấy mẫu lại, dẫn đến có khả năng chỉ trong các phần của ảnh chỉ trong các phần của vài ảnh được sử dụng (trong trường hợp lấy mẫu xuống), hoặc chỉ các phần của cảnh được chụp (trong trường hợp lấy mẫu lên). Ngoài ra, H.263 phụ lục Q cho phép việc tái lấy mẫu của khối macro riêng biệt bởi hệ số hai (trong mỗi chiều), hướng lên hoặc hướng xuống. Một lần nữa, kích cỡ ảnh vẫn giữ nguyên. Kích cỡ của khối macro được cố định trong H.263, và do đó không cần được báo hiệu.

Sự thay đổi của kích cỡ ảnh trong các ảnh được dự đoán trở thành đối tượng chính hơn so với trong mã hóa video hiện đại. Ví dụ, VP9 cho phép việc tái lấy mẫu ảnh tham chiếu và thay đổi của độ phân giải đối với toàn bộ ảnh. Một cách tương tự, các đề xuất cụ thể được thực hiện cho VVC (bao gồm, ví dụ, Hendry, et.al, “Sự thay đổi trên độ phân giải thích ứng (ARC) dùng cho VVC”, Tài liệu nhóm video chung JVET-M0135-v1, Jan 9-19, 2019, được kết hợp ở đây trong toàn bộ nội dung của bản mô tả) cho phép tái lấy mẫu toàn bộ các hình ảnh tham chiếu để phân biệt các độ phân giải cao hơn hoặc thấp hơn. Trong tài liệu

này, các độ phân giải ứng viên khác nhau được đề xuất để được mã hóa trong tập hợp tham số chuỗi và được viện dẫn bởi các phân tử cú pháp theo ảnh trong tập hợp tham số ảnh.

Bản chất kỹ thuật của sáng chế

Các phương án đề cập đến phương pháp, hệ thống, và phương tiện đọc được bằng máy tính để mã hóa dữ liệu video. Theo một khía cạnh, phương pháp mã hóa dữ liệu video được đề xuất. Phương pháp này có thể bao gồm thu dữ liệu video bao gồm ảnh hiện tại và một hoặc nhiều ảnh khác. Cờ thứ nhất tương ứng với việc ảnh hiện tại được tham chiếu bởi một hoặc nhiều ảnh khác theo thứ tự giải mã có được kiểm tra không. Cờ thứ hai tương ứng với việc ảnh hiện tại có được xuất ra không; được kiểm tra. Dữ liệu video được giải mã dựa trên các giá trị tương ứng với cờ thứ nhất và cờ thứ hai.

Theo một khía cạnh khác, hệ thống máy tính mã hóa dữ liệu video được đề xuất. Hệ thống máy tính có thể bao gồm một hoặc nhiều bộ xử lý, một hoặc nhiều bộ nhớ đọc được bằng máy tính, một hoặc nhiều thiết bị lưu trữ hữu hình đọc được bằng máy tính, và các lệnh chương trình được lưu trữ trên ít nhất một trong một hoặc nhiều thiết bị lưu trữ dùng để thực thi bởi ít nhất một trong một hoặc nhiều bộ xử lý thông qua ít nhất một trong một hoặc nhiều bộ nhớ, nhờ đó hệ thống máy tính có khả năng thực hiện phương pháp. Phương pháp có thể bao gồm thu dữ liệu video bao gồm ảnh hiện tại và một hoặc nhiều ảnh khác. Cờ thứ nhất tương ứng với việc ảnh hiện tại được tham chiếu bởi một hoặc nhiều ảnh khác theo thứ tự giải mã có được kiểm tra không. Cờ thứ hai tương ứng với việc ảnh hiện tại có được xuất ra không; được kiểm tra. Dữ liệu video được giải mã dựa trên các giá trị tương ứng với cờ thứ nhất và cờ thứ hai.

Theo một khía cạnh khác nữa, phương tiện đọc được bằng máy tính mã hóa dữ liệu video được đề xuất. Phương tiện đọc được bằng máy tính có thể bao gồm một hoặc nhiều thiết bị lưu trữ đọc được bằng máy tính và các lệnh chương trình được lưu trữ trên ít nhất một trong một hoặc nhiều thiết bị lưu trữ hữu hình, các lệnh chương trình thực thi được bằng bộ xử lý. Các lệnh chương trình thực thi

được bằng bộ xử lý để thực hiện phương pháp mà có thể theo đó bao gồm bước thu dữ liệu video mà bao gồm ảnh hiện tại và một hoặc nhiều ảnh khác. Cờ thứ nhất tương ứng với việc ảnh hiện tại có được tham chiếu bởi một hoặc nhiều ảnh khác theo thứ tự giải mã hay không được kiểm tra. Cờ thứ hai tương ứng với việc ảnh hiện tại có được xuất ra không; được kiểm tra. Dữ liệu video được giải mã dựa trên các giá trị tương ứng với cờ thứ nhất và cờ thứ hai.

Mô tả vắn tắt các hình vẽ

Các đối tượng này và các đối tượng khác, các đặc điểm và ưu điểm sẽ trở nên rõ ràng với phần mô tả chi tiết dưới đây của các phương án minh họa, mà được đọc cùng với các hình vẽ kèm theo. Các đặc điểm khác nhau của các hình vẽ không biến đổi tỷ lệ như các hình minh họa để làm rõ để thuận tiện cho việc hiểu về sáng chế của người có hiểu biết trung bình trong lĩnh vực kỹ thuật kết hợp với phần mô tả chi tiết. Trên các hình vẽ:

FIG.1 là hình minh họa giản lược của sơ đồ khối đơn giản của hệ thống truyền thông theo phương án.

FIG.2 là hình minh họa giản lược của sơ đồ khối đơn giản của hệ thống truyền thông theo phương án.

FIG.3 là minh họa giản lược của sơ đồ khối đơn giản hóa của bộ giải mã theo phương án của sáng chế.

FIG.4 là minh họa giản lược của sơ đồ khối đơn giản hóa của bộ mã hóa theo phương án của sáng chế.

FIG.5 là hình minh họa giản lược của các tùy chọn để báo hiệu các tham số ARC theo kỹ thuật đã biết hoặc phương án, như được chỉ báo.

FIG.6 là ví dụ của bảng cú pháp theo phương án.

FIG.7 là hình minh họa giản lược của hệ thống máy tính theo phương án.

FIG.8 là ví dụ về cấu trúc dự đoán dùng cho khả năng lấy tỷ lệ với sự thay đổi độ phân giải thích ứng.

FIG.9 là ví dụ của bảng cú pháp theo phương án.

FIG.10 là hình minh họa giản lược của sơ đồ khối đơn giản của việc phân tích và giải mã vòng đếm thứ tự ảnh (poc) trên mỗi đơn vị truy nhập và giá trị đếm đơn vị truy nhập.

FIG.11 là hình minh họa giản lược của cấu trúc dòng bit video bao gồm các ảnh con nhiều lớp.

FIG.12 là hình minh họa giản lược của màn hiển thị của ảnh con được lựa chọn với độ phân giải được nâng cao.

FIG.13 là sơ đồ khối của quá trình giải mã và hiển thị cho dòng bit video bao gồm các ảnh con được tạo đa lớp.

FIG.14 là hình minh họa giản lược của hiển thị video 360 với lớp nâng cao của ảnh con.

FIG.15 là ví dụ về thông tin bản thiết kế của các ảnh con và lớp tương ứng của nó và cấu trúc dự đoán hình ảnh.

FIG.16 là ví dụ về thông tin bản thiết kế của các ảnh con và lớp tương ứng của nó và cấu trúc dự đoán hình ảnh với phương thức lấy tỷ lệ theo không gian.

FIG.17 là ví dụ của bảng cú pháp dùng cho thông tin bản thiết kế ảnh con.

FIG.18 là ví dụ của bảng cú pháp của bản tin SEI dùng cho thông tin bản thiết kế ảnh con.

FIG.19 là ví dụ về bảng cú pháp để chỉ báo các lớp đầu ra và thông tin tệp hồ sơ/tầng/cấp độ cho mỗi tập hợp lớp đầu ra.

FIG.20 là ví dụ về bảng cú pháp để chỉ báo chế độ lớp đầu ra trên cho mỗi tập hợp lớp đầu ra.

FIG.21 là ví dụ về bảng cú pháp để chỉ báo ảnh con hiện tại của mỗi lớp cho mỗi tập hợp lớp đầu ra.

FIG.22 là ví dụ về bảng cú pháp của tập tham số video RBSP.

FIG.23 là ví dụ về bảng cú pháp để chỉ báo tập lớp đầu ra với chế độ tập lớp đầu ra.

FIG.24 là ví dụ về bảng cú pháp của tiêu đề ảnh để chỉ báo thông tin đầu ra ảnh.

Mô tả chi tiết sáng chế

Các phương án chi tiết của các cấu trúc và phương pháp được bảo hộ được đề xuất ở đây; tuy nhiên, có thể hiểu rằng các phương án được bộc lộ chỉ là minh họa của các cấu trúc và phương pháp được bảo hộ mà có thể được thực hiện theo các dạng khác nhau. Các cấu trúc và phương pháp có thể, tuy nhiên, được thực hiện trong các dạng khác nhau và không nên được hiểu là bị giới hạn ở tập phương án mẫu được đặt ra ở đây. Hơn nữa, các phương án mẫu này được đề xuất để phần bộc lộ sẽ xuyên suốt và hoàn chỉnh và sẽ truyền tải đầy đủ phạm vi cho những người có hiểu biết trung bình trong lĩnh vực kỹ thuật. Trong phần mô tả, các chi tiết của các đặc điểm và kỹ thuật đã biết có thể được bỏ qua để tránh tạo ra rào cản không cần thiết đối với các phương án được thể hiện.

Như được mô tả trước đây, các bộ mã hóa và giải mã video có xu hướng hoạt động trên kích cỡ ảnh định trước mà, trong hầu hết các trường hợp, được xác định và cố định đối với chuỗi video được mã hóa (CVS-coded video sequence), Nhóm của các ảnh (GOP-Group of Pictures), hoặc khung thời gian liên ảnh. Tuy nhiên, ảnh có thể hoặc có thể không được tham chiếu bởi ảnh kế tiếp cho bù chuyển động hoặc dự đoán tham số khác. Ảnh có thể hoặc có thể không được xuất ra. Có thể có ưu điểm, do đó, để báo hiệu thông tin tham chiếu và thông đầu ra ảnh trong một hoặc nhiều tập tham số.

Các khía cạnh được mô tả ở đây viển dẫn tới các minh họa bằng lưu đồ và/hoặc các sơ đồ khối của các phương pháp, thiết bị (các hệ thống), và phương tiện đọc được bằng máy tính theo các phương án khác nhau. Sẽ được hiểu rằng mỗi khối của các minh họa bằng lưu đồ và/hoặc các sơ đồ khối, và các kết hợp của các khối trong các minh họa bằng lưu đồ và/hoặc các sơ đồ khối, có thể được thực

hiện bởi các lệnh chương trình đọc được bằng máy tính.

FIG.1 minh họa sơ đồ khối đơn giản của hệ thống truyền thông (100) theo phương án của sáng chế. Hệ thống (100) có thể bao gồm ít nhất hai thiết bị đầu cuối (110-120) được kết nối với nhau qua mạng (150). Đối với việc truyền dữ liệu đơn hướng, thiết bị đầu cuối thứ nhất (110) có thể mã hóa dữ liệu video tại vị trí cục bộ cho việc truyền tới thiết bị đầu cuối khác (120) thông qua mạng (150). Thiết bị đầu cuối thứ hai (120) có thể thu dữ liệu video được mã hóa của thiết bị đầu cuối khác từ mạng (150), giải mã dữ liệu được mã hóa và hiển thị dữ liệu video được khôi phục. Việc truyền dữ liệu đơn hướng có thể dùng chung trong các ứng dụng phục vụ nội dung đa phương tiện và loại tương tự.

FIG.1 minh họa cặp thiết bị đầu cuối thứ hai (130, 140) được bố trí để hỗ trợ việc truyền hai chiều của video được mã hóa mà có thể diễn ra, ví dụ, trong hội nghị video. Đối với việc truyền dữ liệu hai chiều, mỗi thiết bị đầu cuối (130, 140) có thể mã hóa dữ liệu video được thu tại vị trí cục bộ cho việc truyền tới thiết bị đầu cuối khác thông qua mạng (150). Mỗi thiết bị đầu cuối 130, 140 cũng có thể thu dữ liệu video được mã hóa được truyền bởi thiết bị đầu cuối khác, có thể giải mã dữ liệu được mã hóa và có thể hiển thị dữ liệu video được khôi phục tại thiết bị hiển thị cục bộ.

Trên FIG.1, các thiết bị đầu cuối (110-140) có thể được minh họa như là các máy chủ, máy tính cá nhân và điện thoại thông minh tuy nhiên các nguyên tắc của sáng chế có thể không bị giới hạn ở đó. Các phương án của sáng chế tìm kiếm ứng dụng bằng các máy tính xách tay, máy tính bảng, các máy chơi nội dung đa phương tiện và/hoặc thiết bị hội nghị video chuyên dụng. Mạng (150) biểu diễn số lượng mạng bất kỳ mà truyền dữ liệu video được mã hóa giữa các thiết bị đầu cuối (110-140), bao gồm ví dụ các mạng truyền thông đường dây và/hoặc không dây. Mạng truyền thông (150) có thể trao đổi dữ liệu trong các kênh được chuyển đổi tròn và/hoặc các kênh được chuyển đổi theo gói tin. Các mạng đại diện bao gồm các mạng viễn thông, các mạng vùng cục bộ, các mạng diện rộng và/hoặc Internet. Nhằm mục đích thảo luận hiện tại, cấu trúc và cấu trúc liên kết của mạng (150) có

thể không được xem xét đối với hoạt động của sáng chế trừ khi được giải thích dưới đây.

FIG.2 minh họa, như là ví dụ đối với ứng dụng cho đối tượng được bộc lộ, việc bố trí của bộ mã hóa video và bộ giải mã video trong môi trường tạo dòng. Đối tượng được bộc lộ có thể được áp dụng đồng đều tới các ứng dụng cho phép video khác, bao gồm, ví dụ, hội nghị video, TV số, lưu trữ video được nén trên nội dung đa phương tiện số bao gồm CD, DVD, thẻ nhớ và loại tương tự, và v.v.

Hệ thống tạo dòng có thể bao gồm hệ thống thu phụ (213), mà có thể bao gồm nguồn video (201), ví dụ camera số, tạo ra, ví dụ, dòng mẫu video không được nén (202). Dòng mẫu (202) này, được thể hiện như là đường đậm để nhấn mạnh khối lượng dữ liệu cao khi được so sánh với các dòng bit video được mã hóa, có thể được xử lý bởi bộ mã hóa (203) được ghép nối tới camera (201). Bộ mã hóa (203) có thể bao gồm phần cứng, phần mềm, hoặc kết hợp của chúng để cho phép hoặc thực hiện các khía cạnh của đối tượng được bộc lộ được mô tả chi tiết dưới đây. Dòng bit video được mã hóa (204), được thể hiện như là đường mỏng để nhấn mạnh khối lượng dữ liệu thấp hơn khi được so sánh với dòng mẫu, có thể được lưu trữ trên máy chủ tạo dòng (205) cho việc sử dụng trong tương lai. Một hoặc nhiều các máy khách tạo dòng (206, 208) có thể truy nhập máy chủ tạo dòng (205) để phục hồi các bản sao (207, 209) của dòng bit video được mã hóa (204). Máy khách (206) có thể bao gồm bộ giải mã video (210) mà giải mã bản sao đi vào của dòng bit video được mã hóa (207) và tạo ra video dòng mẫu đi ra (211) mà có thể được kết xuất đồ họa trên màn hình (212) hoặc thiết bị kết xuất đồ họa khác (không được thể hiện). Trong các hệ thống tạo dòng khác, các dòng bit video được mã hóa (204, 207, 209) có thể được mã hóa theo các tiêu chuẩn nén/mã hóa video định trước. Các ví dụ của các tiêu chuẩn này bao gồm khuyến nghị ITU-T H.265. Đang phát triển là tiêu chuẩn mã hóa video được biết đến không chính thức là Mã hóa video đa năng hoặc VVC. Đối tượng được bộc lộ có thể được sử dụng trong ngữ cảnh của VVC.

FIG.3 có thể là sơ đồ khối chức năng của bộ giải mã video 210 theo

phương án của sáng chế.

Bộ thu (310) có thể thu một hoặc nhiều chuỗi video được mã hóa-giải mã cần được giải mã bởi bộ giải mã (210); trong cùng phương án hoặc phương án khác, một chuỗi video được mã hóa tại thời điểm, trong đó việc giải mã của mỗi chuỗi video được mã hóa độc lập với các chuỗi video được mã hóa khác. Chuỗi video được mã hóa có thể thu được từ kênh (312), mà có thể là liên kết phần cứng/phần mềm tới thiết bị lưu trữ mà lưu trữ dữ liệu video được mã hóa. Bộ thu (310) có thể thu dữ liệu video được mã hóa với dữ liệu khác, ví dụ, dữ liệu audio được mã hóa và/hoặc các dòng dữ liệu hỗ trợ, mà có thể được chuyển tiếp tới các phần tương ứng sử dụng các thực thể (không được thể hiện). Bộ thu (310) có thể tách biệt chuỗi video được mã hóa từ dữ liệu khác. Để đối phó với sự biến động mạng, bộ nhớ đệm (315) có thể được ghép nối giữa bộ thu (310) và bộ giải mã entropy / bộ phân tích (320) (sau đây gọi là "bộ phân tích"). Khi bộ thu (310) đang thu dữ liệu từ thiết bị lưu trữ/chuyển tiếp có băng thông và khả năng điều khiển đủ, hoặc từ mạng đồng bộ, bộ đệm (315) có thể không cần thiết, hoặc có thể là nhỏ. Để sử dụng trên các mạng gói tin nổi lực cao nhất như là mạng Internet, bộ đệm (315) có thể được yêu cầu, có thể là tương đối lớn và có thể có lợi về kích cỡ thích ứng.

Bộ giải mã video (210) có thể bao gồm bộ phân tích (320) để khôi phục các ký tự (321) từ chuỗi video được mã hóa entropy. Các mục của các ký tự này bao gồm thông tin được sử dụng để quản lý hoạt động của bộ giải mã (210), và thông tin có khả năng để điều khiển thiết bị kết xuất đồ họa như màn hình (212) mà không phải phần liên khối của bộ giải mã mà có thể được ghép nối tới đó, như được thể hiện trong Fig. 2. Thông tin điều khiển dùng cho các thiết bị kết xuất đồ họa có thể nằm trong dạng của thông tin nâng cao bổ sung (các bản tin SEI -Supplementary Enhancement Information) hoặc các phân đoạn tập hợp tham số thông tin có khả năng sử dụng video (VUI-Video Usability Information) (không được thể hiện). Bộ phân tích (320) có thể phân tích/giải mã entropy chuỗi video được mã hóa thu được. Việc mã hóa chuỗi video được mã hóa có thể tuân theo kỹ thuật hoặc tiêu chuẩn mã hóa video, và có thể tuân theo các nguyên tắc đã được

biết đến rộng rãi đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật, bao gồm mã hóa độ dài biến thiên, mã hóa Huffman, mã hóa số học có hoặc không có cảm nhận ngữ cảnh, và v.v. Bộ phân tích (320) có thể tách từ chuỗi video được mã hóa, tập hợp của các tham số nhóm con đối với ít nhất một trong số các nhóm con của các điểm ảnh trong bộ giải mã video, dựa trên ít nhất một tham số tương ứng với nhóm. Các nhóm con có thể bao gồm các Nhóm của các ảnh (GOPs), các ảnh, các ô, các lát, các khối macrô, các đơn vị mã hóa (CU-Coding Unit), các khối, các đơn vị biến đổi (TU-Transform Unit), các đơn vị dự đoán (PU-Prediction Unit) và v.v. Bộ giải mã entropi / bộ phân tích có thể cũng trích xuất từ thông tin chuỗi video được mã hóa như là các hệ số biến đổi, các giá trị tham số lượng tử hóa, các vector chuyển động, và v.v..

Bộ phân tích (320) có thể thực hiện hoạt động giải mã entropi/phân tích trên chuỗi video thu được từ bộ đệm (315), để tạo ra các ký tự (321).

Việc khôi phục của các ký tự (321) có thể bao hàm các đơn vị khác nhau phụ thuộc vào kiểu của hình ảnh video được mã hóa hoặc các phần của nó (như là: hình ảnh liên đới và hình ảnh nội bộ, khối liên đới và khối nội bộ), và các hệ số khác. Các đơn vị mà được bao hàm, và cách thức, có thể được điều khiển bởi thông tin điều khiển nhóm con mà được phân tích từ chuỗi video được mã hóa bởi bộ phân tích (320). Tiến trình của thông tin điều khiển nhóm con này giữa bộ phân tích (320) và nhiều đơn vị dưới đây không được mô tả một cách rõ ràng.

Ngoài các khối chức năng đã được đề cập, bộ giải mã 210 có thể được phân chia về mặt khái niệm thành các bộ phận chức năng như được mô tả dưới đây. Trong vận hành thực hiện thực tế theo các điều kiện thương mại, nhiều bộ phận này tương tác khăng khít với nhau và có thể, ít nhất một phần, được tích hợp vào nhau. Tuy nhiên, nhằm mục đích mô tả các đối tượng được bộc lộ, việc chia nhỏ khái niệm thành các đơn vị chức năng dưới đây là phù hợp.

Bộ phận thứ nhất là bộ biến đổi tỷ lệ/bộ biến đổi ngược (351). Bộ biến đổi tỷ lệ/bộ biến đổi ngược (351) thu hệ số biến đổi được lượng tử hóa cũng như thông tin điều khiển, bao gồm biến đổi nào được sử dụng, kích cỡ khối, hệ số

lượng tử hóa, các ma trận biến đổi tỷ lệ lượng tử hóa, v.v như các ký tự (321) từ bộ phân tích (320). Bộ biến đổi tỷ lệ/bộ biến đổi ngược (351) có thể xuất ra các khối bao gồm các giá trị mẫu, mà có thể được đưa vào bộ kết hợp (355).

Trong một số trường hợp, các mẫu đầu ra của bộ định tỷ lệ / biến đổi ngược (351) có thể thuộc về khối được mã hóa nội bộ; tức là: khối mà không sử dụng thông tin mang tính dự đoán từ các hình ảnh được tái cấu trúc trước đó, nhưng có thể sử dụng thông tin mang tính dự đoán từ các phần được tái cấu trúc trước đó của hình ảnh hiện thời. Thông tin dự đoán này có thể được cấp bởi bộ dự đoán nội ảnh (352). Trong một vài trường hợp, bộ dự đoán nội ảnh (352) tạo ra khối có cùng dạng và kích cỡ của khối đang được khôi phục, nhờ sử dụng thông tin đã được khôi phục xung quanh được truy xuất từ ảnh hiện tại (được khôi phục một phần) (356). Bộ tổng hợp (355), trong một số trường hợp, thêm, trên cơ sở mỗi mẫu, thông tin dự đoán của bộ phận dự đoán nội bộ (352) được tạo ra cho thông tin mẫu đầu ra như được đưa ra bởi bộ định tỷ lệ / bộ phận biến đổi ngược (351).

Trong các trường hợp khác, các mẫu đầu ra của bộ định tỷ lệ / bộ phận biến đổi ngược (351) có thể thuộc về khối được mã hóa liên đới và có khả năng được bù chuyển động. Trong trường hợp này, bộ phận dự đoán bù chuyển động (353) có thể truy nhập bộ nhớ hình ảnh tham chiếu (357) để lấy các mẫu được sử dụng cho việc dự đoán. Sau khi bù chuyển động các mẫu được truy xuất theo các ký tự (321) thuộc về khối, các mẫu này có thể được bổ sung bởi bộ kết hợp (355) tới đầu ra của bộ biến đổi tỷ lệ/bộ biến đổi ngược (trong trường hợp này gọi là các mẫu dư hoặc tín hiệu dư) để tạo ra thông tin mẫu đầu ra. Các địa chỉ trong dạng bộ nhớ ảnh tham chiếu trong đó bộ bù chuyển động truy xuất các mẫu dự đoán có thể được điều khiển bởi các vectơ chuyển động, khả dụng đối với bộ bù chuyển động trong dạng của các ký tự (321) mà có thể có, ví dụ X, Y, và các thành phần ảnh tham chiếu. Việc bù chuyển động cũng có thể bao gồm việc nội suy các giá trị mẫu như được truy xuất từ bộ nhớ ảnh tham chiếu khi các vectơ chuyển động chính xác mẫu con được sử dụng, các cơ chế dự đoán vectơ chuyển động, và v.v.

Các mẫu đầu ra của bộ kết hợp (355) có thể được đưa vào các kỹ thuật lọc vòng khác nhau trong bộ lọc vòng (356). Các kỹ thuật nén video có thể bao gồm các kỹ thuật lọc vòng trong mà được điều khiển bởi các tham số được chứa trong dòng bit được mã hóa và khả dụng đối với bộ lọc vòng (356) như các ký tự (321) từ bộ phân tích (320), nhưng có thể cũng phản hồi lại siêu thông tin thu được trong khi giải mã các phần trước đó (trong thứ tự giải mã) của ảnh được mã hóa hoặc chuỗi video được mã hóa, cũng như phản hồi lại các giá trị mẫu được lọc vòng và được khôi phục trước đó.

Đầu ra của bộ lọc vòng (356) có thể là dòng mẫu mà có thể được xuất ra tới thiết bị kết xuất đồ họa (212) cũng như được lưu trữ trong bộ nhớ ảnh tham chiếu (356) để sử dụng trong việc dự đoán liên ảnh tương lai.

Các ảnh được mã hóa định trước, được khôi phục đầy đủ một lần, có thể được sử dụng như là các ảnh tham chiếu cho việc dự đoán tương lai. Một khi ảnh được mã hóa được khôi phục đầy đủ và ảnh được mã hóa đã được nhận dạng là ảnh tham chiếu (ví dụ, bởi bộ phân tích (320)), ảnh tham chiếu hiện tại (356) có thể trở thành một phần của bộ đệm ảnh tham chiếu (357), và bộ nhớ ảnh hiện tại làm mới có thể được cấp phát lại trước khi bắt đầu khôi phục ảnh được mã hóa tiếp theo..

Bộ giải mã video 320 có thể thực hiện các hoạt động giải mã theo kỹ thuật nén video định trước mà được mô tả trong tiêu chuẩn, như khuyến nghị ITU-T H.265. Chuỗi video được mã hóa có thể tuân theo cú pháp được chỉ rõ bởi kỹ thuật nén video hoặc tiêu chuẩn đang được sử dụng, trong trường hợp mà tuân theo cú pháp của kỹ thuật hoặc tiêu chuẩn nén video, như được chỉ rõ trong tiêu chuẩn hoặc tài liệu kỹ thuật nén video và một cách cụ thể trong tài liệu mẫu. Điều cần tuân theo có thể là việc độ phức tạp của chuỗi video được mã hóa nằm trong giới hạn đã được xác định bởi cấp độ của kỹ thuật hoặc tiêu chuẩn nén video. Trong một số trường hợp, các cấp độ giới hạn kích cỡ ảnh tối đa, tốc độ khung tối đa, tỷ lệ mẫu tái cấu trúc tối đa (được đo theo, ví dụ các siêu mẫu trên giây), kích cỡ ảnh tham chiếu tối đa, và v.v.. Các giới hạn được thiết lập bởi các cấp độ có thể,

trong một số trường hợp, còn được giới hạn thông qua các mô tả kỹ thuật về Bộ giải mã tham chiếu giả thuyết (HRD) và siêu dữ liệu dùng cho việc quản lý bộ đệm HRD được báo hiệu trong chuỗi video được mã hóa.

Trong phương án của sáng chế, bộ thu (310) có thể thu dữ liệu bổ sung (dư thừa) với video được mã hóa. Dữ liệu bổ sung có thể được bao gồm như là một phần của (các) chuỗi video được mã hóa. Dữ liệu bổ sung có thể được sử dụng bởi bộ giải mã video (320) để giải mã một cách thích hợp dữ liệu và/hoặc khôi phục chính xác hơn dữ liệu video gốc. Dữ liệu bổ sung có thể nằm trong dạng của, ví dụ, các lớp nâng cao theo thời gian, không gian, các lát dư thừa, các ảnh dư thừa, các mã sửa lỗi tiến, và loại tương tự.

FIG.4 có thể là sơ đồ khối chức năng của bộ mã hóa video (203) theo phương án của sáng chế.

Bộ mã hóa (203) có thể thu các mẫu video từ nguồn video (201) (mà không phải một phần của bộ mã hóa) mà có thể chụp các ảnh video cần được mã hóa bởi bộ mã hóa (203).

Nguồn video (201) có thể cấp chuỗi video nguồn cần được mã hóa bởi bộ mã hóa (203) trong dạng của dòng mẫu video số mà có thể có độ sâu bit thích hợp bất kỳ (ví dụ: 8 bit, 10 bit, 12 bit, ...), bất kỳ không gian màu (ví dụ, BT.601 Y CrCb, RGB, ...), và bất kỳ cấu trúc lấy mẫu thích hợp (ví dụ Y CrCb 4:2:0, Y CrCb 4:4:4). Trong hệ thống phục vụ nội dung đa phương tiện, nguồn video (201) có thể là thiết bị lưu trữ mà lưu trữ video được xử lý trước. Trong hệ thống hội thảo video, nguồn video (203) có thể là camera mà chụp thông tin ảnh cục bộ như chuỗi video. Dữ liệu video có thể được cung cấp dưới dạng các hình ảnh chuyên biệt mà thể hiện chuyển động khi được nhìn theo chuỗi. Bản thân các ảnh có thể được sắp xếp như là mảng không gian của các điểm ảnh, trong đó mỗi điểm ảnh có thể bao gồm một hoặc nhiều mẫu phụ thuộc vào cấu trúc lấy mẫu, không gian màu, v.v được sử dụng. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật có thể dễ hiểu mối liên hệ giữa các điểm ảnh và các mẫu. Phần mô tả dưới đây tập trung vào các mẫu.

Theo phương án của sáng chế, bộ mã hóa (203) có thể mã hóa và nén các ảnh của chuỗi video nguồn thành chuỗi video được mã hóa (443) trong thời gian thực hoặc theo bất kỳ điều kiện thời gian khác như được yêu cầu bởi ứng dụng. Cường ép tốc độ mã hóa thích hợp là một chức năng của bộ điều khiển (450). Bộ điều khiển sẽ điều khiển các bộ phận chức năng khác như được mô tả dưới đây và được ghép nối chức năng tới các bộ phận này. Việc ghép nối không được mô tả rõ ràng. Các tham số được thiết lập bởi bộ điều khiển có thể bao gồm các tham số dành cho điều khiển tốc độ (bỏ qua ảnh, lượng tử hóa, giá trị lamda của các kỹ thuật tối ưu hóa biến dạng tỷ lệ, ...), kích cỡ ảnh, nhóm bản thiết kế ảnh (GOP), phạm vi tìm kiếm vectơ chuyển động tối đa, và v.v.. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật có thể dễ dàng nhận ra các chức năng khác của bộ điều khiển (450) do chúng có thể thuộc về bộ mã hóa video (203) được tối ưu hóa cho thiết kế hệ thống định trước.

Một vài bộ mã hóa video hoạt động theo cách mà người có hiểu biết trung bình trong lĩnh vực kỹ thuật dễ dàng nhận biết là “vòng mã hóa”. Đối với phần mô tả đơn giản hóa, vòng mã hóa có thể bao gồm phần mã hóa của bộ mã hóa (430) (sau đây gọi là “bộ mã hóa nguồn”) (chịu trách nhiệm tạo ra các ký tự dựa trên ảnh đầu vào cần được mã hóa, và ảnh tham chiếu(s)), và bộ giải mã (cục bộ) (433) được gắn trong bộ mã hóa (203) mà khôi phục các ký tự để tạo ra dữ liệu mẫu, bộ giải mã (tù xa) cũng sẽ tạo ra (khi việc nén bất kỳ giữa các ký tự và dòng bit video được mã hóa là không tổn hao trong các kỹ thuật nén video được xem xét trong đối tượng được bộc lộ). Dòng mẫu được khôi phục này được đưa vào bộ nhớ ảnh tham chiếu (434). Khi việc giải mã của dòng ký tự dẫn tới các kết quả chính xác bit độc lập với vị trí bộ giải mã (cục bộ hoặc tù xa), nội dung bộ đệm ảnh tham chiếu cũng là chính xác bit giữa bộ mã hóa cục bộ và bộ mã hóa tù xa. Nói cách khác, phần dự đoán của bộ mã hóa “thấy được” như là các mẫu ảnh tham chiếu một cách chính xác tương tự các giá trị mẫu khi bộ giải mã “thấy được” khi sử dụng việc dự đoán trong khi giải mã. Nguyên tắc cơ bản này của đồng bộ ảnh tham chiếu (và độ trượt hệ quả, nếu tính đồng bộ có thể không được duy trì, ví dụ do các lỗi kênh) cũng được biết đến với người có hiểu biết trung bình trong lĩnh

vực kỹ thuật.

Hoạt động bộ giải mã “cục bộ” (433) có thể giống với của bộ giải mã “từ xa” (210), mà đã được mô tả chi tiết nêu trên dựa trên FIG.3. Tuy nhiên, viện dẫn ngắn gọn tới FIG.3, khi các ký tự là khả dụng và việc mã hóa/giải mã các ký tự đối với chuỗi video được mã hóa bởi bộ mã hóa entropi (445) và bộ phân tích (320) có thể là không tốn hao, các phần giải mã entropi của bộ giải mã (210), bao gồm kênh (312), bộ thu (310), bộ đệm (315), và bộ phân tích (320) có thể không được thực hiện đầy đủ trong bộ giải mã cục bộ (433).

Việc quan sát có thể được thực hiện tại điểm này đó là bất kỳ kỹ thuật bộ giải mã ngoại trừ việc phân tích/giải mã entropi mà được hiện diện trong bộ giải mã cũng cần được hiện diện, về cơ bản trong dạng chức năng đồng nhất, trong bộ mã hóa tương ứng. Vì lý do này, đối tượng được bộc lộ tập trung vào hoạt động bộ giải mã. Phần mô tả của các kỹ thuật của bộ mã hóa có thể được vắng tắt khi chúng ngược với các kỹ thuật của bộ giải mã được mô tả hoàn chỉnh. Chỉ trong các phần nhất định thì việc mô tả chi tiết hơn được yêu cầu và được cung cấp dưới đây.

Đối với một phần của hoạt động, bộ mã hóa nguồn (430) có thể thực hiện mã hóa dự đoán bù chuyển động, mã mã hóa khung đầu vào theo cách dự đoán có viện dẫn tới một hoặc nhiều khung được mã hóa trước đó từ chuỗi video mà được chỉ định là “các khung tham chiếu.” Theo cách thức này, cơ cấu mã hóa (432) mã hóa các độ chênh lệch giữa các khối điểm ảnh của khung đầu vào và các khối điểm ảnh của các khung tham chiếu mà có thể được lựa chọn như là các tham chiếu dự đoán đối với khung đầu vào.

Bộ giải mã video cục bộ (433) có thể giải mã dữ liệu video được mã hóa của các khung mà có thể được chỉ định là các khung tham chiếu, dựa trên các ký tự đã được tạo ra bởi bộ mã hóa nguồn (430). Các thao tác của cơ cấu mã hóa (432) có thể là các quá trình tổn hao. Khi dữ liệu video được mã hóa có thể được giải mã ở bộ giải mã video (không được thể hiện trên FIG.4), chuỗi video được tái cấu trúc thường có thể là bản sao của chuỗi video nguồn với một số sai số. Bộ giải mã video cục bộ (433) sao chép các xử lý giải mã mà có thể được thực hiện bởi bộ giải

mã video trên các khung tham chiếu và có thể làm cho các khung tham chiếu được khôi phục được lưu trữ trong bộ lưu trữ đệm ảnh tham chiếu (434). Theo cách thức này, bộ mã hóa (203) có thể lưu trữ các bản sao của các khung tham chiếu được khôi phục cục bộ mà có nội dung chung như các khung tham chiếu được khôi phục mà sẽ thu được bởi bộ giải mã video phía xa (không có các lỗi truyền).

Bộ dự đoán (435) có thể thực hiện dự đoán tìm kiếm cho cơ cấu mã hóa (432). Tức là, đối với khung mới cần được mã hóa, bộ dự đoán (435) có thể tìm kiếm bộ nhớ ảnh tham chiếu (434) đối với dữ liệu mẫu (như các khối điểm ảnh tham chiếu ứng viên) hoặc siêu dữ liệu định trước như các vectơ chuyển động ảnh tham chiếu, các dạng khối, và v.v, mà có thể đóng vai trò là tham chiếu dự đoán thích hợp đối với các ảnh mới. Bộ dự đoán (435) có thể thao tác trên cơ sở khối với điểm ảnh khối mẫu để tìm ra các tham chiếu dự đoán thích hợp. Trong một số trường hợp, như được xác định bằng cách tìm kiếm các kết quả được thu nhận bởi bộ dự đoán (435), hình ảnh đầu vào có thể có các tham chiếu dự đoán được rút ra từ các hình ảnh tham chiếu được lưu trữ trong bộ nhớ hình ảnh tham chiếu (434).

Bộ điều khiển (450) có thể quản lý các thao tác mã hóa của bộ mã hóa video (430), bao gồm, ví dụ, việc thiết lập của các tham số và các tham số nhóm con được sử dụng để mã hóa dữ liệu video.

Đầu ra của tất cả các bộ phận chức năng nêu trên có thể được đưa vào mã hóa entropy trong bộ mã hóa entropy (445). Bộ mã hóa entropy dịch các ký tự như được tạo ra bởi các bộ phận chức năng khác nhau thành chuỗi video được mã hóa, bằng cách nén không tổn hao các ký tự theo các kỹ thuật đã biết đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật như, ví dụ mã hóa Huffman, mã hóa độ dài biến thiên, mã hóa số học, và v.v.

Bộ truyền (440) có thể lưu trữ đệm (các) chuỗi video được mã hóa như được tạo ra bởi bộ mã hóa entropy (445) để chuẩn bị cho việc truyền thông qua kênh truyền thông (460), mà có thể là liên kết phần cứng/phần mềm tối thiểu bị lưu trữ mà sẽ lưu trữ dữ liệu video được mã hóa. Bộ truyền (440) có thể hợp nhất dữ liệu video được mã hóa từ bộ mã hóa video (430) với dữ liệu khác cần được

truyền, ví dụ, dữ liệu audio được mã hóa và/hoặc các dòng dữ liệu hỗ trợ (các nguồn không được thể hiện).

Bộ điều khiển (450) có thể quản lý hoạt động của bộ mã hóa (203). Trong khi mã hóa, bộ điều khiển (450) có thể gán tới mỗi ảnh được mã hóa loại ảnh được mã hóa định trước, mà có thể ảnh hưởng tới các kỹ thuật mã hóa mà có thể được áp dụng tới ảnh tương ứng. Ví dụ, các ảnh thường có thể được gán như là một trong số các loại khung sau đây:

Ảnh nội dự đoán (ảnh I) có thể là ảnh mà có thể được mã hóa và giải mã mà không sử dụng bất kỳ khung khác trong chuỗi như là nguồn dự đoán. Một vài bộ mã hóa-giải mã video cho phép các loại ảnh nội dự đoán khác nhau, bao gồm, ví dụ các ảnh làm mới bộ giải mã độc lập. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật sẽ nhận ra các biến thể của các ảnh I và các ứng dụng và các đặc tính tương ứng của chúng.

Ảnh dự đoán (ảnh P) có thể là ảnh mà có thể được mã hóa và giải mã nhờ sử dụng việc nội dự đoán hoặc việc dự đoán liên đới nhờ sử dụng nhiều nhất một vector chuyển động và chỉ số tham chiếu để dự đoán các giá trị mẫu của mỗi khối.

Ảnh dự đoán hai chiều (ảnh B) có thể là ảnh mà có thể được mã hóa và được giải mã nhờ sử dụng việc nội dự đoán hoặc việc dự đoán liên đới nhờ sử dụng nhiều nhất hai vector chuyển động và các chỉ số tham chiếu để dự đoán các giá trị mẫu của mỗi khối. Tương tự, các ảnh đa dự đoán có thể sử dụng nhiều hơn hai ảnh tham chiếu và siêu dữ liệu được kết hợp cho việc khôi phục khối đơn.

Các ảnh nguồn thường có thể được phân chia về mặt không gian thành nhiều khối mẫu (ví dụ, các khối của các mẫu 4×4 , 8×8 , 4×8 , hoặc 16×16) và được mã hóa trên cơ sở lần lượt từng khối. Các khối có thể được mã hóa dự đoán có viện dẫn tới các khối khác (đã được mã hóa) như được xác định bởi việc gán mã hóa được áp dụng tới các ảnh tương ứng của các khối. Ví dụ, các khối của các ảnh I có thể được mã hóa không dự đoán hoặc chúng có thể được mã hóa dự đoán có viện dẫn tới các khối đã được mã hóa của cùng ảnh (dự đoán không gian hoặc nội

dự đoán). Các khối điểm ảnh của các ảnh P có thể được mã hóa không dự đoán, thông qua dự đoán không gian hoặc thông qua dự đoán theo thời gian có viện dẫn tới một ảnh tham chiếu được mã hóa trước đó. Các khối của các ảnh B có thể được mã hóa không dự đoán, thông qua dự đoán theo không gian hoặc thông qua dự đoán theo thời gian có viện dẫn tới một hoặc hai ảnh tham chiếu được mã hóa trước đó.

Bộ mã hóa video (203) có thể thực hiện các hoạt động mã hóa theo kỹ thuật hoặc tiêu chuẩn mã hóa video định trước, như Khuyến nghị ITU-T H.265. Trong hoạt động của bộ mã hóa video (203), bộ mã hóa video (603) có thể thực hiện các hoạt động nén khác nhau, bao gồm các hoạt động mã hóa dự đoán mà khai thác các dư thừa về mặt không gian và thời gian trong chuỗi video đầu vào. Do đó, dữ liệu video được mã hóa, có thể tuân theo cú pháp được chỉ rõ bởi kỹ thuật hoặc tiêu chuẩn mã hóa video đang được sử dụng.

Trong phương án của sáng chế, bộ truyền (440) có thể truyền dữ liệu bổ sung với video được mã hóa. Bộ mã hóa video (430) có thể bao gồm dữ liệu này như là một phần của chuỗi video được mã hóa. Dữ liệu bổ sung có thể bao gồm các lớp nâng cao SNR/theo thời gian/không gian, các dạng khác của dữ liệu dư thừa như các ảnh và các lát dư thừa, các bản tin thông tin tăng cường bổ sung (SEI-Supplementary Enhancement Information), các phân đoạn tập hợp tham số thông tin khả năng sử dụng trực quan (VUI-Visual Usability Information), và v.v.

Trước khi mô tả các khía cạnh nhất định của đối tượng được bộc lộ một cách chi tiết hơn, số ít thuật ngữ cần được đưa ra mà sẽ liên quan đến phần còn lại của phần mô tả này.

Ảnh con sau đây gọi là, trong một vài trường hợp, sắp xếp hình chữ nhật của các mẫu, các khối, các khối macrô, các đơn vị mã hóa, hoặc các thực thể tương tự mà được nhóm lại về mặt ngữ nghĩa học, và có thể được mã hóa độc lập trong độ phân giải được thay đổi. Một hoặc nhiều ảnh con có thể dùng cho ảnh.

Một hoặc nhiều ảnh con được mã hóa có thể tạo thành ảnh được mã hóa. Một hoặc nhiều ảnh con có thể được ghép thành hình ảnh, và một hoặc nhiều ảnh con có thể được trích xuất từ hình ảnh. Trong các môi trường nhất định, một hoặc nhiều ảnh con được mã hóa có thể được ghép trong miền được nén mà không chuyển mã tới mức mẫu thành ảnh được mã hóa, và trong các trường hợp khác hoặc tương tự, một hoặc nhiều ảnh con được mã hóa có thể được tách từ ảnh được mã hóa trong miền được nén.

Việc Thay đổi độ phân giải thích nghi (ARC-Adaptive Resolution Change) sau đây liên quan đến cơ chế mà cho phép sự thay đổi độ phân giải của ảnh hoặc ảnh con trong chuỗi video được mã hóa, ví dụ, bằng cách tái lấy mẫu ảnh tham chiếu. Các tham số ARC sau đây liên quan đến thông tin điều khiển được yêu cầu để thực hiện việc thay đổi độ phân giải thích nghi, mà có thể bao gồm, ví dụ, các tham số lọc, các hệ số biến đổi tỷ lệ, các độ phân giải đầu ra và/hoặc các ảnh tham chiếu, các cờ điều khiển khác nhau, và v.v.

Phần mô tả nêu trên tập trung vào việc mã hóa và giải mã ảnh video đơn được mã hóa độc lập về mặt ngữ nghĩa. Trước khi mô tả ý nghĩa của việc mã hóa/giải mã của các ảnh con với các tham số ARC độc lập và tính phức tạp bổ sung của việc này, các tùy chọn để báo hiệu các tham số ARC có thể được mô tả.

Viện dẫn tới FIG.5, một vài tùy chọn mới để báo hiệu các tham số ARC được thể hiện. Đối với mỗi tùy chọn, chúng có ưu điểm và nhược điểm nhất định từ khía cạnh hiệu quả mã hóa, độ phức tạp, và cấu trúc. Tiêu chuẩn hoặc kỹ thuật mã hóa video có thể lựa chọn một hoặc nhiều trong số các tùy chọn này, hoặc các tùy chọn đã biết từ kỹ thuật trước đó, để báo hiệu các tham số ARC. Các tùy chọn có thể không loại trừ lẫn nhau, và có thể được hiểu là có thể hoán đổi dựa trên các nhu cầu ứng dụng, kỹ thuật tiêu chuẩn liên quan, hoặc sự lựa chọn của bộ mã hóa.

Các lớp của các tham số ARC có thể bao gồm:

-các hệ số lấy mẫu lên/xuống, riêng biệt hoặc được kết hợp trong chiều X và Y,

-các hệ số lấy mẫu lên/xuống, với việc bổ sung chiều thời gian, mà chỉ báo việc thu/phòng tốc độ cố định đối với số lượng định trước,

-Một trong hai lớp nêu trên có thể bao gồm việc mã hóa của một hoặc nhiều phần tử cú pháp ngắn có thể giả định mà có thể trở vào Bảng mà chứa các hệ số này,

-độ phân giải, trong chiều X hoặc Y, trong các đơn vị của các mẫu, các khối, các khối macrô, các CU, hoặc mức chi tiết phù hợp bất kỳ khác, của hình ảnh đầu vào, hình ảnh đầu ra, hình ảnh tham chiếu, hình ảnh được mã hóa, được kết hợp hoặc được phân tách. Nếu có nhiều hơn một phân giải (như là, ví dụ, một cho hình ảnh đầu vào, một cho hình ảnh tham chiếu) sau đó, trong một số trường hợp nhất định, một tập các giá trị có thể được suy ra từ một tập giá trị khác. Phần trên được giám sát, ví dụ, bằng cách sử dụng các cờ. Đối với ví dụ được mô tả chi tiết hơn, xem phần dưới đây.

—các tọa độ “cong” gần giống với các tọa độ này trong Phụ lục P H.263, có cùng mức chi tiết phù hợp như được mô tả ở trên. Phụ lục P H.263 xác định một cách hiệu quả để mã hóa các tọa độ cong này, nhưng mặt khác, cũng có thể nghĩ ra các cách hiệu quả hơn có tiềm năng hơn. Ví dụ, độ dài biến có thể đảo ngược, mã hóa kiểu “Huffman” của các tọa độ cong theo Phụ lục P có thể được thay thế bởi mã hóa nhị phân độ dài thích hợp, trong đó độ dài của từ mã nhị phân có thể, ví dụ, thu được từ kích cỡ ảnh lớn nhất, có thể được nhân bởi hệ số và dịch vị định trước với giá định trước, để cho phép “làm cong” phía ngoài các biên của kích cỡ ảnh lớn nhất.

-các tham số bộ lọc lấy mẫu lên hoặc lấy mẫu xuống. Trong các trường hợp đơn giản nhất, có thể chỉ là bộ lọc đơn nhất cho việc lấy mẫu lên và/hoặc lấy mẫu xuống. Tuy nhiên, trong các trường hợp nhất định, có thể có ưu điểm để cho phép tính linh hoạt hơn trong thiết kế bộ lọc, và có yêu cầu báo hiệu của các tham số lọc. Các tham số này có thể được lựa chọn thông qua chỉ số trong danh sách của các phương án bộ lọc khả dụng, bộ lọc có thể được chỉ rõ đầy đủ (ví dụ thông qua danh sách của các hệ số lọc, nhờ sử dụng các kỹ thuật mã hóa entropy thích hợp),

bộ lọc có thể được lựa chọn ẩn thông qua các tỷ lệ lấy mẫu lên/xuống mà được báo hiệu theo bất kỳ trong số các cơ chế được đề cập nêu trên, và v.v..

Sau đây, phần mô tả giả định việc mã hóa của tập hợp hữu hạn các hệ số lấy mẫu lên/xuống (hệ số giống nhau cần được sử dụng trong cả chiều X và Y), được chỉ báo thông qua từ mã. Từ mã này có thể một cách hiệu quả được mã hóa độ dài biến thiên, ví dụ nhờ sử dụng mã Golomb số mũ dùng chung cho các phần tử cú pháp nhất định trong các tiêu chuẩn kỹ thuật mã hóa video như H.264 và H.265. Một việc ánh xạ thích hợp của các giá trị tới các hệ số lấy mẫu lên/xuống có thể, ví dụ, tuân theo theo bảng sau đây:

Từ mã	Mã số mũ Golomb	Độ phân giải gốc / đích
0	1	1 / 1
1	010	1 / 1,5 (biến đổi tỷ lệ lên 50%)
2	011	1,5 / 1 (biến đổi tỷ lệ xuống 50%)
3	00100	1 / 2 (biến đổi tỷ lệ lên 100%)
4	00101	2 / 1 (biến đổi tỷ lệ xuống 100%)

Nhiều ánh xạ tương tự có thể được nghĩ ra theo nhu cầu về ứng dụng và khả năng của các cơ chế biến đổi tỷ lệ lên và xuống khả dụng trong kỹ thuật hoặc tiêu chuẩn nén video. Bảng có thể được mở rộng thêm nhiều giá trị. Các giá trị có thể cũng được thể hiện bởi các cơ chế mã hóa entropi ngoài các mã Ext-Golomb, ví dụ bằng cách sử dụng mã hóa nhị phân. Điều này có thể có các ưu điểm nhất định khi bản thân các hệ số lấy mẫu được quan tâm nằm ngoài cơ cấu xử lý video

(bộ mã hóa và bộ giải mã đầu tiên), ví dụ bởi MANE. Lưu ý rằng, đối với trường hợp chung nhất (có thể giả định) trong đó việc thay đổi độ phân giải không được yêu cầu, mã Golomb số mũ có thể được lựa chọn là ngắn; trong Bảng nêu trên, chỉ là một bit. Có thể có ưu điểm về hiệu quả mã hóa qua việc sử dụng các mã nhị phân cho trường hợp chung nhất.

Số lượng mục trong Bảng, cũng như các ngữ nghĩa học của chúng có thể được cấu hình đầy đủ hoặc một phần. Ví dụ, tổng quát cơ bản của Bảng có thể được truyền trong tập hợp tham số “cao” như chuỗi hoặc tập hợp tham số bộ giải mã. Ngoài ra, một hoặc nhiều bảng này có thể được xác định trong kỹ thuật hoặc tiêu chuẩn mã hóa video, và có thể được lựa chọn thông qua, ví dụ, tập hợp tham số chuỗi hoặc bộ giải mã.

Sau đây, sáng chế mô tả làm thế nào hệ số lấy mẫu lên/xuống (thông tin ARC), được mã hóa như được mô tả nêu trên, có thể được chứa trong cú pháp kỹ thuật hoặc tiêu chuẩn mã hóa video. Các xem xét tương tự có thể áp dụng tới một, hoặc một ít, từ mã mà điều khiển các bộ lọc lấy mẫu lên/xuống. Xem phần bên dưới đề cập khi lượng tương đối lớn dữ liệu được yêu cầu cho bộ lọc hoặc các cấu trúc dữ liệu khác.

H.263 Phụ lục P bao gồm thông tin ARC 502 có dạng bốn tọa độ cong vênh trong tiêu đề hình ảnh 501, cụ thể trong phần mở rộng của tiêu đề H.263 PLUSPTYPE (503). Điều này có thể là sự lựa chọn thiết kế hợp lý khi a) có tiêu đề hình ảnh khả dụng, và b) các thay đổi thường xuyên của thông tin ARC được mong đợi. Tuy nhiên, thông tin tiêu đề khi bằng cách sử dụng báo hiệu kiểu H.263 có thể là cực cao, và các hệ số biến đổi tỷ lệ có thể là không liên quan trong số các biên ảnh như tiêu đề hình ảnh có thể là của tính chất nhất thời.

JVCET-M135-v1, được viện dẫn ở trên, bao gồm thông tin tham chiếu ARC 505 (chỉ số) được đặt trong tập tham số hình ảnh 504, lập chỉ số Bảng 506 bao gồm các độ phân giải mục tiêu mà đến lượt được đặt trong tập tham số chuỗi (507). Vị trí của độ phân giải khả thi trong Bảng 506 trong tập tham số chuỗi 507 có thể, theo tuyên bố bằng lời của các tác giả, được điều chỉnh bằng cách sử dụng

SPS làm điểm đàm phán khả năng tương tác trong khi trao đổi khả năng. Độ phân giải có thể thay đổi, nằm trong tập hợp giới hạn bởi các giá trị trong bảng 506 từ ảnh đến ảnh bằng cách tham chiếu tập tham số hình ảnh thích hợp 504.

Vấn viện dẫn tới FIG.5, các tùy chọn bổ sung dưới đây có thể tồn tại để truyền tải thông tin ARC trong dòng bit video. Mỗi trong số các tùy chọn đó có các ưu điểm nhất định ngoài kỹ thuật đã biết như được mô tả ở trên. Các tùy chọn có thể được đồng thời thể hiện trong cùng một kỹ thuật hoặc tiêu chuẩn mã hóa video.

Trong phương án, thông tin ARC 509 như là hệ số (thu phóng) tái lấy mẫu có thể được thể hiện trong tiêu đề lát, tiêu đề GOB, tiêu đề ô, hoặc tiêu đề nhóm ô (tile group header) 508. Điều có thể là đủ của thông tin ARC là nhỏ, như là độ dài biến đơn nhất $ue(v)$ hoặc độ dài được cố định từ mã của số ít bit, ví dụ như được thể hiện ở trên. Việc có thông tin ARC trong tiêu đề nhóm ô trực tiếp có ưu điểm bổ sung của thông tin ARC có thể là áp dụng được cho ảnh con được thể hiện bởi, ví dụ, nhóm ô, chứ không phải toàn bộ hình ảnh. Xem phần bên dưới. Ngoài ra, ngay cả khi kỹ thuật hoặc tiêu chuẩn nén video thấy được chỉ các thay đổi về độ phân giải thích ứng của toàn bộ hình ảnh (trái ngược với, ví dụ, các thay đổi về độ phân giải thích ứng dựa trên nhóm ô), việc đưa thông tin ARC vào tiêu đề nhóm ô trái ngược với việc đưa nó vào trong tiêu đề hình ảnh kiểu H.263 có các ưu điểm nhất định khía cạnh phục hồi lỗi.

Trong cùng một phương án hoặc một phương án khác, thông tin ARC 512 tự nó có thể được thể hiện trong tập tham số thích hợp 511 như là, ví dụ, tập tham số hình ảnh, tập tham số tiêu đề, tập tham số ô, tập tham số thích ứng, và v.v. (Tập tham số thích ứng được mô tả). Phạm vi của tập tham số có thể không lớn hơn hình ảnh, ví dụ nhóm ô. Việc sử dụng của thông tin ARC là sử dụng ngầm thông qua việc kích hoạt của tập tham số tương đương. Ví dụ, khi kỹ thuật hoặc tiêu chuẩn mã hóa video chỉ thiết kế ARC dựa trên hình ảnh, sau đó tập tham số hình ảnh hoặc tương đương có thể là thích hợp.

trong cùng một phương án hoặc một phương án khác, thông tin tham chiếu ARC (513) có thể thể hiện trong tiêu đề nhóm ô 514 hoặc cấu trúc dữ liệu tương tự. Thông tin tham chiếu 513 có thể viện dẫn tới tập con của thông tin ARC 515 khả dụng trong tập tham số 516 với phạm vi vượt xa hình ảnh đơn nhất, ví dụ tập tham số chuỗi, hoặc tập tham số bộ giải mã.

Cấp độ bổ sung của việc kích hoạt ngầm không trực tiếp của PPS từ tiêu đề nhóm ô, PPS, SPS, như được sử dụng trong JVET-M0135-v1 dường như là không cần thiết, như các tập tham số hình ảnh, giống như các tập tham số chuỗi, có thể (và có các tiêu chuẩn nhất định như là RFC3984) được sử dụng cho việc đàm phán khả năng hoặc các thông báo. Tuy nhiên, nếu thông tin ARC cần áp dụng được cho ảnh con được đại diện, ví dụ, cũng bởi các nhóm ô, tập tham số với phạm vi kích hoạt được giới hạn đến nhóm ô, như là tập tham số thích ứng hoặc tập tham số tiêu đề có thể là lựa chọn tốt hơn. Cũng vậy, nếu thông tin ARC lớn hơn kích cỡ không đáng kể—ví dụ chứa thông tin điều khiển bộ lọc như là các hệ số bộ lọc số—sau đó tham số có thể là lựa chọn tốt hơn so với bằng cách sử dụng tiêu đề 508 trực tiếp với hiệu quả mã hóa, như các thiết lập đó có thể là sử dụng lại được bởi các hình ảnh tương lai hoặc các ảnh con bằng cách tham chiếu cùng tập tham số.

Khi bằng cách sử dụng tập tham số chuỗi hoặc một tập tham số cao hơn với các hình ảnh mở rộng phạm vi, các nội dung cân nhắc có thể áp dụng:

1. Tập tham số lưu trữ bảng thông tin ARC 516 có thể là, trong một số trường hợp, tập tham số chuỗi, nhưng trong các trường hợp khác tập tham số bộ giải mã một cách thuận lợi. Tập tham số bộ giải mã có thể có phạm vi kích hoạt của các CVS, cụ thể là dòng video được mã hóa, tức là tất cả các bit video được mã hóa từ khi bắt đầu phiên cho đến khi kết thúc phiên. Phạm vi này có thể là thích hợp hơn vì các hệ số ARC khả thi có thể là đặc điểm của bộ giải mã, có thể được thực hiện trong phần cứng, và các đặc điểm phần cứng có xu hướng không thay đổi với CVS bất kỳ (mà trong ít nhất một số hệ thống giải trí là Nhóm ảnh, độ dài một giây hoặc ít hơn). Điều đó cho thấy rằng, việc đưa bảng vào tập tham số

chuỗi rõ ràng là nằm trong các tùy chọn vị trí được mô tả ở đây, cụ thể là kết hợp với điểm 2 dưới đây.

2. Thông tin tham chiếu ARC 513 có thể được đặt trực tiếp trong ảnh/ô lát/GOB/tiêu đề nhóm ô (tile group header) 514 hơn là trong tập tham số hình ảnh được thiết lập như trong JVCET-M0135-v1, với lý do như sau: khi bộ mã hóa muốn thay đổi giá trị đơn nhất trong tập tham số hình ảnh, ví dụ như thông tin tham chiếu ARC, sau đó phải tạo PPS mới và tham chiếu PPS mới. Giả sử rằng chỉ thông tin tham chiếu ARC thay đổi, nhưng thông tin khác như là, ví dụ, thông tin ma trận lượng tử trong PPS giữ nguyên. Thông tin này có thể là kích cỡ cơ sở, và cần được truyền lại để hoàn thành PPS mới. Đối với thông tin tham chiếu ARC có thể là từ mã đơn nhất, như là chỉ số trong bảng 513 và chỉ cần là giá trị mà thay đổi, việc truyền lại toàn bộ sẽ rườm rà và lãng phí, ví dụ, truyền lại thông tin ma trận lượng tử. Trong chừng mực, có thể được xét đến một cách tốt hơn từ khía cạnh hiệu quả mã hóa để tránh việc không trực tiếp đi qua PPS, như được đề xuất trong JVET-M0135-v1. Một cách tương tự, việc đưa thông tin tham chiếu ARC vào PPS có thêm nhược điểm là thông tin ARC được tham chiếu bởi thông tin tham chiếu ARC (513) cần áp dụng cho toàn bộ ảnh và không dùng với ảnh con, theo phạm vi của việc kích hoạt tập tham số hình ảnh là hình ảnh.

Trong cùng một phương án hoặc một phương án khác, báo hiệu của các tham số ARC có thể tuân theo ví dụ chi tiết như đã nêu trong các FIG.6. FIG.6 mô tả các sơ đồ cú pháp trong sự thể hiện như được sử dụng trong các tiêu chuẩn mã hóa video sớm nhất từ năm 1993. Ký hiệu của các sơ đồ cú pháp này thường tuân theo lập trình loại C. Các đường in đậm chỉ báo các phần tử cú pháp thể hiện trong dòng bit, các đường nhạt thường chỉ báo dòng điều khiển hoặc thiết lập của các biến.

Tiêu đề nhóm ô (601) như cấu trúc cú pháp mẫu của tiêu đề ứng dụng được cho phần (có thể là chữ nhật) của hình ảnh có thể chứa có điều kiện, độ dài biến, phần tử cú pháp được mã hóa Exp-Golomb dec_pic_size_idx (602) (được mô tả in đậm). Sự thể hiện của phần tử cú pháp này trong tiêu đề nhóm ô có thể

được giám sát khi sử dụng độ phân giải thích ứng (603) ở đây, giá trị của cờ không được mô tả in đậm, mà có nghĩa rằng cờ được thể hiện trong dòng bit ở điểm mà xuất hiện trong sơ đồ cú pháp. Có hay không việc độ phân giải thích ứng khi sử dụng hình ảnh hoặc các phần này nó có thể được báo hiệu trong cấu trúc cú pháp bậc cao bất kỳ bên trong hoặc bên ngoài dòng bit. Trong ví dụ được thể hiện, được báo hiệu trong tập tham số chuỗi như được đưa ra dưới đây.

Cũng viện dẫn tới FIG.6, cũng là trích xuất của tập tham số chuỗi (610). Phần tử cú pháp thứ nhất được thể hiện là `adaptive_pic_resolution_change_flag` (611). Khi báo đúng, cờ có thể chỉ báo việc sử dụng của độ phân giải thích ứng mà, đến lượt có thể yêu cầu thông tin điều khiển cụ thể. Trong ví dụ, thông tin điều khiển mà được thể hiện có điều kiện dựa trên giá trị của cờ dựa trên câu lệnh `if()` trong tập tham số (612) và tiêu đề nhóm ô (601).

Khi độ phân giải thích ứng đang sử dụng, trong ví dụ này, được mã hóa là độ phân giải đầu ra trong các đơn vị của các mẫu (613). Số 613 liên quan đến cả `output_pic_width_in_luma_samples` và `output_pic_height_in_luma_samples`, mà có thể cùng nhau xác định độ phân giải của ảnh đầu ra. Ở những nơi khác trong kỹ thuật hoặc tiêu chuẩn mã hóa video, các giới hạn cụ thể cho một trong hai giá trị có thể được xác định. Ví dụ, định nghĩa mức có thể giới hạn số lượng tổng mẫu đầu ra, mà có thể là tích của giá trị của hai phần tử cú pháp này. Ngoài ra, các tiêu chuẩn hoặc kỹ thuật mã hóa video định trước, hoặc các kỹ thuật hoặc tiêu chuẩn mở rộng như, ví dụ, các tiêu chuẩn hệ thống, có thể giới hạn phạm vi đánh số (ví dụ, một hoặc cả hai chiều phải có thể chia được cho lũy thừa 2), hoặc tỷ lệ co (ví dụ, độ cao và độ rộng phải nằm trong hệ thức như 4:3 hoặc 16:9). Các giới hạn này có thể được đưa ra để thuận tiện cho các cách thực hiện phần cứng hoặc cho các lý do khác, và đã biết trong kỹ thuật.

Trong các ứng dụng nhất định, có thể được đề xuất rằng bộ mã hóa chỉ dẫn bộ giải mã để sử dụng kích cỡ ảnh tham chiếu định trước mà không giả định ảnh rằng kích cỡ là kích cỡ ảnh đầu ra. Trong ví dụ này, phần tử cú pháp `reference_pic_size_present_flag` (614) tạo công sự có mặt có điều kiện của các

chiều ảnh tham chiếu (615) (một lần nữa, số này liên quan đến các độ rộng và độ cao).

Cuối cùng, được thể hiện là Bảng độ rộng và các độ cao của hình ảnh có thể giải mã. Bảng này có thể được đưa ra, ví dụ, bởi chỉ báo Bảng (`num_dec_pic_size_in_luma_samples_minus1`) (616). “minus1” có thể đề cập đến sự giải thích của giá trị của phần tử cú pháp đó. Ví dụ, nếu giá trị được mã hóa bằng không, một mục nhập bảng được thể hiện. Nếu giá trị bằng năm, sáu mục nhập bảng được thể hiện. Đối với mỗi “đường” trong Bảng, độ rộng và độ cao hình ảnh được mã hóa sau đó nằm trong cú pháp (617).

Các mục Bảng được thể hiện (617) có thể được lấy chỉ số bằng cách sử dụng phần tử cú pháp `dec_pic_size_idx` (602) trong tiêu đề nhóm ô, nhờ đó cho phép các kích cỡ được giải mã khác nhau—có hiệu lực, các hệ số thu phóng—trên mỗi nhóm ô.

Các tiêu chuẩn hoặc kỹ thuật mã hóa video định trước, ví dụ VP9, hỗ trợ khả năng biến đổi tỷ lệ không gian bằng cách thực hiện các dạng định trước của việc tái lấy mẫu ảnh tham chiếu (được báo hiệu khác với đối tượng được bọc lộ) kết hợp với khả năng biến đổi tỷ lệ thời gian, để cho phép khả năng biến đổi không gian. Cụ thể, các ảnh tham chiếu định trước có thể được lấy mẫu lên nhờ sử dụng các kỹ thuật kiểu ARC tới độ phân giải cao hơn để tạo thành cơ sở của lớp tăng cường không gian. Các hình ảnh được lấy mẫu lên này có thể được xác định, bằng cách sử dụng các cơ chế dự đoán thông thường với độ phân giải cao, để thêm chi tiết.

Đối tượng được bọc lộ có thể được sử dụng trong môi trường này. Trong các trường hợp nhất định, trong cùng phương án hoặc phương án khác, giá trị trong thông tin tiêu đề đơn vị NAL, ví dụ trường ID thời gian, có thể được sử dụng để chỉ báo không chỉ thời gian mà còn lớp không gian. Việc này có các ưu điểm nhất định đối với các thiết kế hệ thống định trước; ví dụ, các đơn vị chuyển tiếp được lựa chọn (SFU-Selected Forwarding Unit) hiện tại được tạo ra và được tối ưu hóa đối với lớp thời gian được lựa chọn để chuyển tiếp dựa trên giá trị ID thời

gian thông tin tiêu đề đơn vị NAL, có thể được sử dụng mà không cần cải biến, đối với các môi trường có khả năng biến đổi tỷ lệ. Để cho phép điều này, có thể có yêu cầu đối với việc ánh xạ giữa kích cỡ ảnh được mã hóa và lớp thời gian được chỉ báo bởi trường ID thời gian trong thông tin tiêu đề đơn vị NAL.

Trong một số kỹ thuật mã hóa video, Đơn vị truy nhập (AU) có thể viện dẫn tới (các) hình ảnh được mã hóa, (các) lát, (các) ô, (các) đơn vị NAL, và v.v., mà được chụp và được xử lý bên trong ảnh/lát/ô/dòng bit đơn vị NAL tương ứng ở thời điểm cụ thể. Thời điểm cụ thể này có thể là thời điểm xử lý.

Trong HEVC, và các kỹ thuật mã hóa video cụ thể khác, giá trị số đếm thứ tự hình ảnh (POC) có thể được sử dụng để chỉ báo hình ảnh tham chiếu được lựa chọn trong số nhiều hình ảnh tham chiếu được lưu trữ trong bộ đệm hình ảnh được giải mã (DPB). Khi đơn vị truy nhập (AU) bao gồm một hoặc nhiều ảnh, lát, hoặc ô, mỗi ảnh, lát, hoặc ô thuộc về cùng một AU có thể mang cùng một giá trị POC, mà có thể được lấy từ đó mà chúng được tạo ra từ nội dung của cùng một thời điểm xử lý. Nói cách khác, trong kịch bản trong đó hai ảnh/lát/ô mang cùng một giá trị POC đã cho, mà có thể chỉ báo hai ảnh/lát/ô thuộc về cùng một AU và có cùng một thời điểm xử lý. Ngược lại, hai ảnh/ô/lát mà có các giá trị POC khác nhau có thể chỉ báo các ảnh/lát/ô thuộc về các AU khác nhau và có các thời điểm xử lý khác nhau.

Theo phương án của đối tượng được bộc lộ, quan hệ cứng nói trên có thể làm mềm trong đó đơn vị truy nhập có thể bao gồm các ảnh, các lát hoặc các ô với các giá trị POC khác nhau. Bằng cách cho phép các giá trị POC khác nhau trong AU, có thể sử dụng được giá trị POC để nhận dạng các ảnh/lát/ô có khả năng giải mã được với thời gian thể hiện như nhau. Việc đó theo lượt có thể cho phép hỗ trợ nhiều lớp có thể lấy tỷ lệ được mà không cần thay đổi báo hiệu lựa chọn hình ảnh tham chiếu (ví dụ như báo hiệu tập hình ảnh tham chiếu hoặc báo hiệu danh sách hình ảnh tham chiếu), như được mô tả chi tiết hơn dưới đây.

Tuy nhiên, vẫn mong muốn có thể nhận dạng AU mà hình ảnh/lát/ô thuộc về, đối với hình ảnh/các lát/các ô khác mà có các giá trị POC khác nhau, chỉ từ giá trị POC. Việc này có thể đạt được, như được mô tả dưới đây.

Trong các phương án giống hoặc khác nhau, số đếm đơn vị truy nhập (AUC) có thể được báo hiệu trong cấu trúc cú pháp bậc cao, như là tiêu đề đơn vị NAL, tiêu đề lát, tiêu đề nhóm ô, bản tin SEI, tập tham số hoặc dấu phân tách AU. Giá trị của AUC có thể được sử dụng để nhận dạng các đơn vị NAL, các ảnh, các lát hoặc các ô mà thuộc về AU đã cho. Giá trị của AUC có thể tương ứng với khoảng thời điểm xử lý riêng biệt. Giá trị AUC có thể bằng nhiều giá trị POC. Bằng cách chia giá trị POC cho giá trị nguyên, giá trị AUC có thể được tính toán. Trong các trường hợp cụ thể, các thao tác phân chia có thể đặt một gánh nặng lên các cách thực hiện của bộ giải mã. Trong các trường hợp đó, các giới hạn nhỏ trong khoảng đếm số của các giá trị AUC có thể cho phép thay thế thao tác phân chia thành các thao tác dịch. Ví dụ, giá trị AUC có thể bằng giá trị Bit quan trọng nhất (MSB) của khoảng giá trị POC.

Trong các phương án giống nhau, giá trị của POC trên mỗi AU (`poc_cycle-au`) có thể được báo hiệu trong cấu trúc cú pháp bậc cao, như là tiêu đề đơn vị NAL, tiêu đề lát, tiêu đề nhóm ô, bản tin SEI, tập tham số hoặc dấu phân tách AU. `poc_cycle-au` có thể chỉ báo có bao nhiêu giá trị POC khác nhau và liên tiếp có thể được kết hợp với cùng một AU. Ví dụ, nếu giá trị của `poc_cycle-au` bằng 4, các ảnh, các lát hoặc các ô có giá trị POC bằng 0 – 3, bao gồm cả hai giá trị đầu và cuối, được kết hợp với AU với giá trị AUC bằng 0, và các ảnh, các lát hoặc các ô với giá trị POC bằng 4 – 7, bao gồm cả hai giá trị đầu và cuối, được kết hợp với AU với giá trị AUC bằng 1. Kể từ đây, giá trị của AUC có thể được suy ra bằng cách chia giá trị POC cho giá trị của `poc_cycle-au`.

Trong các phương án giống nhau hoặc một phương án khác, giá trị của `poc_cyle-au` có thể được lấy từ thông tin, được đặt ví dụ trong tập tham số video (VPS), mà nhận dạng số lượng các lớp không gian hoặc SNR trong chuỗi video được mã hóa. Quan hệ khả thi này được mô tả ngắn gọn dưới đây. Trong khi việc

thu nhận như được mô tả ở trên có thể tiết kiệm số ít bit trong VPS và do đó có thể cải thiện hiệu quả mã hóa, có thể là ưu điểm khi mã hóa rõ poc_cycle_au trong cấu trúc cú pháp bậc cao thích hợp theo phân cấp bên dưới tập tham số video, để có thể tối thiểu hóa poc_cycle_au cho phần nhỏ đã cho của dòng bit như là hình ảnh. Việc tối ưu hóa này có thể tiết kiệm nhiều bit hơn mà có thể được tiết kiệm thông qua quá trình thu nhận ở trên vì các giá trị POC (và/hoặc các giá trị của các phần tử cú pháp viện dẫn không trực tiếp đến POC) có thể được mã hóa trong các cấu trúc cú pháp cấp độ thấp.

Các kỹ thuật để báo hiệu các tham số độ phân giải thích nghi được mô tả nêu trên, có thể được thực hiện như là phần mềm máy tính nhờ sử dụng các lệnh có thể đọc được bởi máy tính và được lưu trữ vật lý trong một hoặc nhiều phương tiện đọc được bởi máy tính. Ví dụ, FIG.7 thể hiện hệ thống máy tính 700 thích hợp để thực hiện các phương án của đối tượng được bộc lộ.

Phần mềm máy tính có thể được mã hóa nhờ sử dụng bất kỳ mã máy hoặc ngôn ngữ máy tính thích hợp, mà có thể được tập hợp, biên dịch, liên kết, hoặc cơ chế tương tự để tạo ra mã bao gồm các lệnh mà có thể được thực thi trực tiếp, hoặc thông qua diễn dịch, thực thi vi mã, và loại tương tự, bởi các bộ xử lý trung tâm (CPU) máy tính, các bộ xử lý đồ họa (GPU), và loại tương tự.

Các lệnh có thể được thực thi trên các loại máy tính khác nhau hoặc các thành phần của chúng, bao gồm, ví dụ, các máy tính cá nhân, máy tính bảng, máy chủ, điện thoại thông minh, thiết bị chơi trò chơi, thiết bị Internet vạn vật, và loại tương tự.

Các thành phần được thể hiện trên FIG.7 đối với hệ thống máy tính 700 là ví dụ về mặt bản chất và không dự định đề xuất bất kỳ giới hạn nào đối với phạm vi sử dụng hoặc chức năng của phần mềm máy tính mà thực hiện các phương án của sáng chế. Cấu trúc của các bộ phận không được hiểu là có bất kỳ sự phụ thuộc hoặc yêu cầu liên quan đến bất kỳ một hoặc kết hợp của các bộ phận được minh họa trong phương án ví dụ của hệ thống máy tính 700.

Hệ thống máy tính 700 có thể bao gồm các thiết bị đầu vào giao diện người dùng định trước. Thiết bị đầu vào giao diện người dùng này có thể phản hồi lại việc nhập vào bởi một hoặc nhiều người dùng thông qua, ví dụ, việc nhập xúc giác (như: nhấn phím, trượt, di chuyển bao dữ liệu), nhập audio (như: thoại, vỗ tay), nhập trực quan (như: các cử chỉ), nhập khứu giác (không được thể hiện). Các thiết bị giao diện người dùng có thể cũng được sử dụng để quay nội dung đa phương tiện không cần thiết liên quan trực tiếp đến việc nhập nhận thức bởi người dùng, như audio (như: âm nhạc xung quanh, âm nhạc thoại), các ảnh (như: các ảnh được quét, các ảnh chụp thu nhận từ camera ảnh tĩnh), video (như video hai chiều, video ba chiều bao gồm video lập thể).

Các thiết bị giao diện người dùng đầu vào có thể bao gồm một hoặc nhiều trong số (chỉ một trong số các thiết bị được thể hiện): bàn phím (701), chuột (702), bàn di chuột (703), màn chạm (710), găng tay dữ liệu 704, cần điều khiển (705), microphôn (706), máy quét (707), camera (708).

Hệ thống máy tính 700 có thể cũng bao gồm các thiết bị đầu ra giao diện người dùng định trước. Các thiết bị đầu ra giao diện người dùng này có thể là sự mô phỏng các giác quan của một hoặc nhiều người dùng thông qua, ví dụ, xúc giác, âm thanh, ánh sáng, và mùi/vị giác. Các thiết bị đầu ra giao diện người dùng này có thể bao gồm các thiết bị đầu ra xúc giác (ví dụ xúc giác phản hồi bởi màn chạm (710), găng tay dữ liệu 704, hoặc cần điều khiển (705), nhưng có thể cũng có các thiết bị phản hồi xúc giác mà không đóng vai trò là các thiết bị đầu vào), các thiết bị xuất audio (như: các loa (709), các tai nghe (không được thể hiện)), các thiết bị đầu ra trực quan (như các màn hình (710) bao gồm các màn hình CRT, màn hình LCD, màn hình plasma, màn hình OLED, mỗi trong số chúng có hoặc không có khả năng nhập màn chạm, có hoặc không có khả năng phản hồi xúc giác—một vài trong số chúng có thể có khả năng xuất ra đầu ra trực quan hai chiều hoặc đầu ra nhiều hơn ba chiều thông qua phương tiện như đầu ra lập thể; các kính thực tế ảo (không được thể hiện), các màn hình toàn ký và bề khối (không được thể hiện)), và các máy in (không được thể hiện).

Hệ thống máy tính 700 có thể cũng bao gồm các thiết bị lưu trữ có thể truy nhập được bởi người dùng và phương tiện được kết hợp của chúng như phương tiện quang bao gồm CD/DVD ROM/RW 720 với CD/DVD hoặc phương tiện tương tự 721, ổ chóp 722, ổ cứng tháo rời hoặc ổ bán dẫn 723, phương tiện từ truyền thống như băng và đĩa mềm (không được thể hiện), các thiết bị dựa trên ROM/ASIC/PLD chuyên dụng như các khóa điện tử bảo mật (không được thể hiện), và loại tương tự.

Người có hiểu biết trung bình trong lĩnh vực kỹ thuật cũng sẽ hiểu rằng thuật ngữ “phương tiện đọc được bởi máy tính” như được sử dụng kết hợp với đối tượng được bộc lộ không bao hàm môi trường truyền, các sóng mang, hoặc các tín hiệu chuyển tiếp khác.

Hệ thống máy tính (700) có thể cũng bao gồm giao diện tới một hoặc nhiều mạng truyền thông. Các mạng có thể ví dụ là mạng không dây, nối dây, quang học. Các mạng có thể là mạng cục bộ, mạng diện rộng, mạng đô thị, mạng phương tiện giao thông và công nghiệp, thời gian thực, dung sai trễ, và v.v. Các ví dụ của các mạng bao gồm các mạng cục bộ như Ethernet, LAN không dây, mạng tế bào bao gồm GSM, 3G, 4G, 5G, LTE và loại tương tự, TV các mạng số diện rộng có dây hoặc không dây bao gồm TV cáp, TV vệ tinh, và TV quảng bá mặt đất, phương tiện giao thông và công nghiệp bao gồm CANBus, và v.v. Các mạng nhất định thường yêu cầu các bộ thích ứng giao diện mạng phía ngoài mà được gắn vào cổng dữ liệu mục đích chung định trước hoặc các kênh truyền ngoại vi (749) (như, ví dụ các cổng USB của hệ thống máy tính (700)); các phần khác thường được tích hợp vào lõi của hệ thống máy tính (700) bằng cách gắn tới kênh truyền hệ thống như được mô tả dưới đây (ví dụ giao diện Ethernet vào hệ thống máy tính PC hoặc giao diện mạng tế bào vào hệ thống máy tính điện thoại thông minh). Nhờ sử dụng bất kỳ các mạng này, hệ thống máy tính 700 có thể truyền thông với các thực thể khác. Việc truyền thông này có thể là chỉ thu đơn hướng (ví dụ, TV quảng bá), chỉ gửi đơn hướng (ví dụ các thiết bị CANbus tới CANbus định trước), hoặc hai chiều, ví dụ tới các hệ thống máy tính khác nhờ sử dụng các mạng

số diện rộng hoặc cục bộ. Các giao thức định trước và các ngăn giao thức có thể được sử dụng trên mỗi mạng và giao diện mạng này như được mô tả nêu trên.

Các thiết bị giao diện người dùng, các thiết bị lưu trữ có thể truy nhập được bởi người dùng, và các giao diện mạng nêu trên có thể được gắn vào lõi 740 của hệ thống máy tính 700.

Lõi (740) có thể bao gồm một hoặc nhiều bộ xử lý trung tâm (CPU) (741), các bộ xử lý đồ họa (GPU) (742), các bộ xử lý khả trình chuyên dụng dưới dạng của mảng cổng khả trình dạng trường (FPGA) (743), các bộ tăng tốc phần cứng đối với các tác vụ định trước (744), và v.v. Các thiết bị này, cùng với bộ nhớ chỉ đọc (ROM) (745), bộ nhớ truy nhập ngẫu nhiên (746), bộ lưu trữ khối bên trong như các ổ cứng không thể truy nhập bởi người dùng bên trong, SSD, và loại tương tự (747), có thể được kết nối thông qua kênh truyền hệ thống (748). Trong một vài hệ thống máy tính, kênh truyền hệ thống 748 có thể truy nhập được trong dạng của một hoặc nhiều đầu cắm vật lý để cho phép các phần mở rộng bởi các CPU, GPU bổ sung, và loại tương tự. Các thiết bị ngoại vi có thể được gắn trực tiếp tới kênh truyền hệ thống của lõi 748, hoặc thông qua kênh truyền ngoại vi 749. Các cấu trúc của các kênh truyền ngoại vi bao gồm PCI, USB, và loại tương tự.

Các CPU 741, GPU 742, FPGA 743, và các bộ tăng tốc 744 có thể thực thi các lệnh định trước mà, một cách kết hợp, có thể thực thi mã máy tính nêu trên. Mã máy tính có thể được lưu trữ trong ROM 745 hoặc RAM 746. Dữ liệu chuyển tiếp có thể cũng được lưu trữ trong RAM 746, trong khi dữ liệu cố định có thể được lưu trữ ví dụ, trong bộ nhớ khối bên trong 747. Việc lưu trữ nhanh và truy xuất tới bất kỳ thiết bị bộ nhớ có thể được cho phép thông qua việc sử dụng của bộ nhớ lưu trữ đệm, mà có thể được kết hợp mật thiết với một hoặc nhiều CPU 741, GPU 742, bộ lưu trữ cỡ lớn 747, ROM 745, RAM 746, và loại tương tự.

Phương tiện đọc được bởi máy tính có thể có mã máy tính trên đó để thực hiện các hoạt động được thực hiện bởi máy tính khác nhau. Phương tiện và mã máy tính có thể được thiết kế và xây dựng đặc biệt cho mục đích của sáng chế,

hoặc chúng có thể là loại được biết đến rộng rãi và khả dụng đối với người có hiểu biết trung bình trong lĩnh vực phần mềm máy tính.

Theo ví dụ của sáng chế và không bị giới hạn, hệ thống máy tính có cấu trúc 700, và cụ thể, lõi 740 có thể cung cấp chức năng như là kết quả của các bộ xử lý (bao gồm các CPU, GPU, FPGA, các bộ tăng tốc, và loại tương tự) mà thực thi phần mềm trong một hoặc nhiều phương tiện hữu hình đọc được bởi máy tính. Các phương tiện đọc được bởi máy tính này có thể là phương tiện được kết hợp với bộ lưu trữ khối có thể truy nhập bởi người dùng như được giới thiệu nêu trên, cũng như bộ lưu trữ định trước của lõi 740 mà có bản chất bất biến, như bộ lưu trữ khối bên trong của lõi 747 hoặc ROM 745. Phần mềm mà thực hiện các phương án của sáng chế có thể được lưu trữ trong các thiết bị này và được thực thi bởi lõi 740. Phương tiện đọc được bởi máy tính có thể bao gồm một hoặc nhiều thiết bị bộ nhớ hoặc chip, theo nhu cầu cụ thể. Phần mềm có thể làm cho lõi 740 và một cách cụ thể, các bộ xử lý trong đó (bao gồm CPU, GPU, FPGA, và loại tương tự) thực hiện các xử lý cụ thể hoặc các phần cụ thể của các xử lý cụ thể được mô tả ở đây, bao gồm xác định các cấu trúc dữ liệu được lưu trữ trong RAM 746 và cải biến các cấu trúc dữ liệu này theo các xử lý được xác định bởi phần mềm. Ngoài ra hoặc như là tùy chọn khác, hệ thống máy tính có thể cung cấp chức năng như là kết quả của mạng cứng logic hoặc được lắp trong mạch (ví dụ: bộ tăng tốc 744), mà có thể hoạt động trong vị trí của hoặc cùng với phần mềm để thực hiện các xử lý cụ thể hoặc các phần cụ thể của các xử lý cụ thể được mô tả ở đây. Việc tham chiếu tới phần mềm có thể bao hàm logic, và ngược lại, nếu cần. Việc dẫn tới phương tiện đọc được bằng máy tính có thể bao trùm mạch (như là mạch tích hợp (IC)) mà lưu trữ phần mềm cho việc thực thi, mạch thể hiện logic cho việc thực thi, hoặc cả hai, mà là phù hợp. Sáng chế bao trùm kết hợp phù hợp bất kỳ nào của phần cứng và phần mềm.

FIG.8 thể hiện ví dụ của cấu trúc chuỗi video với sự kết hợp của temporal_id, layer_id, các giá trị POC và AUC có sự thay đổi độ phân giải thích ứng Trong ví dụ này, hình ảnh, lát hoặc ô trong AU thứ nhất với AUC = 0 có thể có temporal_id = 0 và layer_id = 0 hoặc 1, trong khi đó hình ảnh, lát hoặc ô trong

AU thứ hai với $AUC = 1$ có thể có $temporal_id = 1$ và $layer_id = 0$ hoặc 1, theo cách tương ứng. Giá trị của POC tăng lên 1 trên mỗi ảnh bất kể các giá trị của $temporal_id$ và $layer_id$. Trong ví dụ này, giá trị của poc_cycle_au có thể bằng 2. Tốt nhất là, giá trị của poc_cycle_au có thể được thiết lập bằng số lượng các lớp (có khả năng lấy tỷ lệ theo không gian). Trong ví dụ này, do đó, giá trị của POC được tăng thêm 2, trong khi đó giá trị của AUC tăng thêm 1.

Trong các phương án nêu trên, tất cả hoặc tập con của cấu trúc dự đoán ảnh liên đới hoặc cấu trúc dự đoán lớp liên đới và sự chỉ báo hình ảnh tham chiếu có thể được hỗ trợ bằng cách sử dụng báo hiệu tập hình ảnh tham chiếu (RPS) đã biết trong HEVC hoặc báo hiệu danh sách hình ảnh tham chiếu (RPL). Trong RPS hoặc RPL, hình ảnh tham chiếu được chọn được chỉ báo bởi báo hiệu giá trị của POC hoặc giá trị delta của POC giữa hình ảnh hiện thời và hình ảnh tham chiếu được lựa chọn. Đối với đối tượng được bộc lộ, RPS và RPL có thể được sử dụng để chỉ báo cấu trúc dự đoán ảnh liên đới hoặc dự đoán lớp liên đới mà không cần thay đổi báo hiệu, nhưng có các giới hạn sau đây. Nếu giá trị của $temporal_id$ của hình ảnh tham chiếu lớn hơn giá trị của $temporal_id$ của hình ảnh hiện tại, hình ảnh hiện tại có thể không sử dụng hình ảnh tham chiếu cho việc bù chuyển động hoặc các dự đoán khác. Nếu giá trị của $layer_id$ của hình ảnh tham chiếu lớn hơn giá trị $layer_id$ của hình ảnh hiện tại, hình ảnh hiện tại có thể không sử dụng hình ảnh tham chiếu cho việc bù chuyển động hoặc các dự đoán khác.

Trong các phương án giống nhau và các phương án khác, vector chuyển động biến đổi tỷ lệ dựa trên độ chênh lệch POC đối với dự đoán vector chuyển động theo thời gian có thể được vô hiệu qua các ảnh trong đơn vị truy nhập. Kể từ đây, mặc dù mỗi hình ảnh có thể có giá trị POC khác nhau trong đơn vị truy nhập, vector chuyển động không được lấy tỷ lệ và được sử dụng cho dự đoán vector chuyển động theo thời gian nằm trong đơn vị truy nhập. Đó là do hình ảnh tham chiếu với POC khác nhau trong cùng một AU được coi là hình ảnh tham chiếu mà có cùng đối tượng thời gian. Do đó, trong các phương án nêu trên, chức năng lấy tỷ lệ vector chuyển động có thể trở về 1, khi hình ảnh tham chiếu thuộc về AU được kết hợp với hình ảnh hiện tại.

Trong các phương án giống nhau và các phương án khác, việc lấy tỷ lệ vector chuyển động dựa trên độ chênh lệch POC cho dự đoán vector chuyển động theo thời gian có thể một cách tùy chọn được vô hiệu qua nhiều ảnh, khi độ phân giải theo không gian của hình ảnh tham chiếu khác với độ phân giải theo không gian của hình ảnh hiện tại. Khi việc lấy tỷ lệ vector chuyển động được cho phép, vector chuyển động được lấy tỷ lệ dựa trên cả độ chênh lệch POC và tỷ lệ độ phân giải theo không gian giữa hình ảnh hiện tại và hình ảnh tham chiếu.

Trong cùng một phương án hoặc một phương án khác, vector chuyển động có thể được lấy tỷ lệ dựa trên độ chênh lệch AUC thay vì độ chênh lệch POC, cho dự đoán vector chuyển động theo thời gian, đặc biệt khi poc_cycle_au có giá trị không đồng nhất (khi $vps_contant_poc_cycle_per_au == 0$). Ngược lại (khi $vps_contant_poc_cycle_per_au == 1$), việc lấy tỷ lệ vector chuyển động dựa trên độ chênh lệch AUC có thể giống với việc lấy tỷ lệ vector chuyển động dựa trên độ chênh lệch POC.

Trong cùng một phương án hoặc một phương án khác, khi vector chuyển động được lấy tỷ lệ dựa trên độ chênh lệch AUC, vector chuyển động tham chiếu trong cùng một AU (với cùng một giá trị AUC) với hình ảnh hiện tại không được lấy tỷ lệ dựa trên độ chênh lệch AUC và được sử dụng cho dự đoán vector chuyển động mà không cần biến đổi tỷ lệ hoặc có biến đổi tỷ lệ dựa trên tỷ lệ độ phân giải theo không gian giữa hình ảnh hiện tại và hình ảnh tham chiếu.

Trong các phương án giống nhau và các phương án khác, giá trị AUC được sử dụng để nhận dạng biên của AU và được sử dụng cho thao tác bộ giải mã tham chiếu giả định (HRD), mà cần sự định thời đầu vào và đầu ra với mức chi tiết AU. Trong hầu hết các trường hợp, hình ảnh được giải mã với lớp cao nhất trong AU có thể được xuất ra để hiển thị. Giá trị AUC và giá trị layer_id có thể được sử dụng để nhận dạng hình ảnh đầu ra.

Trong phương án nêu trên, hình ảnh có thể gồm có một hoặc nhiều ảnh con. Mỗi ảnh con này có thể phủ vùng cục bộ toàn bộ vùng của hình ảnh. Vùng được hỗ trợ bởi ảnh con có thể hoặc có thể không bị chồng lấn với vùng được hỗ

trợ bởi một ảnh con khác. Vùng được xử lý bởi một hoặc nhiều ảnh con có thể hoặc có thể không phủ toàn bộ vùng của hình ảnh. Nếu hình ảnh gồm ảnh con, vùng được hỗ trợ bởi ảnh con giống với vùng được hỗ trợ bởi hình ảnh.

Trong cùng một phương án, ảnh con có thể được mã hóa bằng phương pháp mã hóa giống với phương pháp mã hóa được sử dụng cho hình ảnh được mã hóa. Ảnh con có thể được mã hóa độc lập hoặc có thể được mã hóa phụ thuộc vào một ảnh con khác hoặc ảnh được mã hóa. Ảnh con có thể hoặc có thể không có phần dư thừa phân tích bất kỳ từ một ảnh con khác hoặc ảnh được mã hóa.

Trong cùng một phương án, ảnh con được mã hóa có thể được chứa trong một hoặc nhiều lớp. Ảnh con được mã hóa trong lớp có thể có độ phân giải theo không gian khác nhau. Ảnh con gốc có thể được tái lấy mẫu theo không gian (lấy mẫu lên hoặc lấy mẫu xuống), được mã hóa với các tham số độ phân giải không gian khác nhau, và được chứa trong dòng bit tương ứng với lớp.

Trong các phương án giống nhau hoặc một phương án khác, ảnh con với (W, H) , trong đó W chỉ báo độ rộng của ảnh con và H chỉ báo độ cao của ảnh con, theo cách tương ứng, có thể được mã hóa và được chứa trong dòng bit được mã hóa tương ứng với lớp 0, trong khi đó ảnh con được lấy mẫu lên (hoặc được lấy mẫu xuống) từ ảnh con với độ phân giải theo không gian gốc, với $(W^*S_{w,k}, H^*S_{h,k})$, có thể được mã hóa và được chứa trong dòng bit được mã hóa tương ứng với lớp k , trong đó $S_{w,k}, S_{h,k}$ chỉ báo các tỷ lệ tái lấy mẫu, theo chiều ngang và theo chiều dọc. Nếu các giá trị của $S_{w,k}, S_{h,k}$ lớn hơn 1, việc tái lấy mẫu bằng lấy mẫu lên. Trong khi đó, nếu các giá trị của $S_{w,k}, S_{h,k}$ nhỏ hơn 1, việc tái lấy mẫu bằng lấy mẫu xuống.

Trong các phương án giống nhau hoặc một phương án khác, ảnh con được mã hóa trong lớp có thể có chất lượng quan sát khác nhau từ chất lượng quan sát của ảnh con được mã hóa trong một lớp khác trong cùng ảnh con hoặc các ảnh con khác nhau. Ví dụ, ảnh con i trong lớp, n , được mã hóa với tham số lượng tử hóa, $Q_{i,n}$, trong khi đó ảnh con j trong lớp, m , được mã hóa với tham số lượng tử hóa, $Q_{j,m}$.

Trong các phương án giống nhau hoặc một phương án khác, ảnh con được mã hóa trong lớp có thể giải mã được một cách độc lập, mà không cần bất kỳ phân tích hoặc giải mã phần dư nào từ ảnh con được mã hóa trong một lớp khác của cùng vùng cục bộ. Lớp ảnh con, mà có thể giải mã được một cách độc lập mà không cần tham chiếu đến một lớp ảnh con khác của cùng vùng cục bộ, là lớp ảnh con độc lập. Ảnh con được mã hóa trong lớp ảnh con độc lập có thể hoặc có thể không có giải mã hoặc phân tích phần dư từ ảnh con được mã hóa trước đó trong cùng lớp ảnh con, nhưng ảnh con được mã hóa có thể không có bất kỳ phần dư nào từ ảnh được mã hóa trong một lớp ảnh con khác.

Trong các phương án giống nhau hoặc một phương án khác, ảnh con được mã hóa trong lớp có thể giải mã được một cách phụ thuộc, với phân tích hoặc giải mã phần dư nào từ ảnh con được mã hóa trong một lớp khác của cùng vùng cục bộ. Lớp ảnh con, mà có thể giải mã được một cách phụ thuộc với tham chiếu đến một lớp ảnh con khác của cùng vùng cục bộ, là lớp ảnh con phụ thuộc. Ảnh con được mã hóa trong ảnh con phụ thuộc có thể tham chiếu đến ảnh con được mã hóa thuộc về cùng một ảnh con, ảnh con được mã hóa trước trong cùng một lớp ảnh con, hoặc cả hai ảnh con tham chiếu.

Trong các phương án giống nhau hoặc một phương án khác, ảnh con được mã hóa gồm có một hoặc nhiều lớp ảnh con độc lập và một hoặc nhiều lớp ảnh con phụ thuộc. Tuy nhiên, ít nhất một lớp ảnh con độc lập có thể được thể hiện cho ảnh con được mã hóa. Lớp ảnh con độc lập có thể có giá trị của ký hiệu nhận dạng lớp (`layer_id`), mà có thể được thể hiện trong tiêu đề đơn vị NAL hoặc một cấu trúc cú pháp bậc cao khác, bằng 0. Lớp ảnh con với `layer_id` bằng 0 là lớp ảnh con gốc.

Trong các phương án giống nhau hoặc một phương án khác, hình ảnh có thể gồm có một hoặc nhiều ảnh con phía trước và một ảnh con phía sau. Vùng được hỗ trợ bởi ảnh con phía sau có thể bằng vùng của hình ảnh. Vùng được hỗ trợ bởi ảnh con phía trước có thể chồng lấn với vùng được hỗ trợ bởi ảnh con phía sau. Ảnh con phía sau có thể là lớp ảnh con gốc, trong khi đó ảnh con phía trước có

thể là lớp ảnh con không gốc (nâng cao). Một hoặc nhiều lớp ảnh con không gốc có thể tham chiếu cùng lớp gốc để giải mã. Mỗi lớp ảnh con không gốc với `layer_id` bằng a có thể tham chiếu lớp ảnh con không gốc với `layer_id` bằng b , trong đó a lớn hơn b .

Trong cùng phương án hoặc một phương án khác, hình ảnh có thể gồm có một hoặc nhiều ảnh con phía trước với hoặc không cần một ảnh con phía sau. Mỗi ảnh con có thể có lớp ảnh con gốc của nó và một hoặc nhiều lớp không gốc (nâng cao). Mỗi lớp ảnh con gốc có thể được tham chiếu bởi một hoặc nhiều lớp ảnh con không gốc. Mỗi lớp ảnh con không gốc với `layer_id` bằng a có thể tham chiếu lớp ảnh con không gốc với `layer_id` bằng b , trong đó a lớn hơn b .

Trong cùng phương án hoặc một phương án khác, hình ảnh có thể gồm có một hoặc nhiều ảnh con phía trước với hoặc không cần một ảnh con phía sau. Mỗi ảnh con được mã hóa trong lớp ảnh con (gốc hoặc không gốc) có thể được tham chiếu bởi một hoặc nhiều lớp ảnh con không gốc thuộc về cùng một ảnh con và một hoặc nhiều lớp ảnh con không gốc, mà không thuộc về cùng một ảnh con.

Trong cùng phương án hoặc một phương án khác, hình ảnh có thể gồm có một hoặc nhiều ảnh con phía trước với hoặc không cần một ảnh con phía sau. Ảnh con trong lớp a có thể còn được phân vùng thành nhiều ảnh con trong cùng một lớp. Một hoặc nhiều ảnh con được mã hóa trong lớp b có thể tham chiếu ảnh con được phân vùng trong lớp a .

Trong các phương án giống nhau hoặc một phương án khác, chuỗi video được mã hóa (CVS) có thể là nhóm của các ảnh được mã hóa. CVS có thể gồm có một hoặc nhiều chuỗi ảnh con được mã hóa (CSPS), trong đó CSPS có thể là nhóm ảnh con được mã hóa mà phủ cùng vùng cục bộ của hình ảnh. CSPS có thể có độ phân giải theo thời gian giống hoặc khác nhau so với của chuỗi video được mã hóa.

Trong các phương án giống nhau hoặc một phương án khác, CSPS có thể được mã hóa và được chứa trong một hoặc nhiều lớp. CSPS có thể gồm có một

hoặc nhiều lớp CSPA. Giải mã một hoặc nhiều lớp CSPA tương ứng với CSPA có thể tái cấu trúc chuỗi của ảnh con tương ứng với cùng vùng cục bộ.

Trong các phương án giống nhau hoặc một phương án khác, số lượng lớp CSPA tương ứng với CSPA có thể giống với hoặc khác với số lượng lớp CSPA tương ứng với một CSPA khác.

Trong cùng một phương án hoặc một phương án khác, lớp CSPA có thể có độ phân giải theo thời gian khác nhau (ví dụ như tốc độ khung) từ một lớp CSPA khác. Chuỗi ảnh con (không nén) gốc có thể được tái lấy mẫu theo thời gian (lấy mẫu lên hoặc lấy mẫu xuống), được mã hóa với các tham số độ phân giải theo thời gian khác nhau, và được chứa trong dòng bit tương ứng với lớp.

Trong cùng một phương án hoặc một phương án khác, chuỗi ảnh con với tốc độ khung, F , có thể được mã hóa và được chứa trong dòng bit được mã hóa tương ứng với lớp 0, trong khi đó chuỗi ảnh con được lấy mẫu lên (hoặc lấy mẫu xuống) theo thời gian từ chuỗi ảnh con gốc, với $F \cdot S_{t,k}$, có thể được mã hóa và được chứa trong dòng bit được mã hóa tương ứng với lớp k , trong đó $S_{t,k}$ chỉ báo tỷ lệ lấy mẫu theo thời gian cho lớp k . Nếu giá trị của $S_{t,k}$ lớn hơn 1, quá trình tái lấy mẫu theo thời gian bằng sự chuyển đổi lên tốc độ khung. Trong đó, nếu giá trị của $S_{t,k}$ nhỏ hơn 1, quá trình tái lấy mẫu theo thời gian bằng sự chuyển đổi xuống tốc độ khung.

Trong các phương án giống nhau hoặc một phương án khác, khi ảnh con với lớp CSPA a là tham chiếu bởi ảnh con với lớp CSPA b cho việc bù chuyển động hoặc dự đoán lớp liên đới bất kỳ, nếu độ phân giải theo không gian của lớp CSPA a khác với độ phân giải theo không gian của lớp CSPA b , các điểm ảnh được giải mã trong lớp CSPA a được tái lấy mẫu và được sử dụng cho việc tham chiếu. Quá trình tái lấy mẫu có thể cần lọc lấy mẫu lên hoặc lọc lấy mẫu xuống.

Trong cùng một phương án hoặc một phương án khác, FIG.9 thể hiện ví dụ của các bảng cú pháp của các bảng cú pháp để báo hiệu phần tử cú pháp của `vps_poc_cycle_au` trong VPS (hoặc SPS), mà chỉ báo `poc_cycle_au` được sử dụng cho toàn bộ ảnh/các lát trong chuỗi video được mã hóa, và phần tử cú pháp của

slice_poc_cycle_au, mà chỉ báo poc_cycle_au của lát hiện thời, trong tiêu đề lát. Nếu giá trị POC tăng lên một cách thống nhất trên mỗi AU, vps_contant_poc_cycle_per_au trong VPS được thiết đặt bằng 1 và vps_poc_cycle_au được báo hiệu trong VPS. Trong trường hợp này, slice_poc_cycle_au không được báo hiệu rõ ràng, và giá trị của AUC cho mỗi AU được tính toán bằng cách chia giá trị của POC bởi vps_poc_cycle_au. Nếu giá trị POC không làm tăng một cách thống nhất trên mỗi AU, vps_contant_poc_cycle_per_au trong VPS được thiết đặt bằng 0. Trong trường hợp này, vps_access_unit_cnt không được báo hiệu, trong khi đó slice_access_unit_cnt được báo hiệu trong tiêu đề lát cho mỗi lát hoặc ảnh. Mỗi lát hoặc ảnh có thể có giá trị khác nhau của slice_access_unit_cnt. Giá trị của AUC cho mỗi AU được tính toán bằng cách chia giá trị của POC bởi slice_poc_cycle_au. FIG.10 thể hiện sơ đồ khối minh họa tiến trình làm việc tương đương.

Trong các phương án giống nhau hoặc các phương án khác, ngay cả khi giá trị của POC của hình ảnh, lát, hoặc ô có thể là khác nhau, hình ảnh, lát, hoặc ô tương ứng với AU có cùng một giá trị AUC có thể được kết hợp với cùng một đối tượng thời gian giải mã hoặc đầu ra. Kể từ đây, không cần có bất kỳ dư thừa phân tích liên đới/giải mã qua các ảnh, các lát hoặc các ô trong cùng một AU, tất cả hoặc tập con của các ảnh, các lát hoặc các ô được kết hợp với cùng một AU có thể được giải mã song song, và có thể được xuất ra cùng một đối tượng thời gian.

Trong các phương án giống nhau hoặc các phương án khác, ngay cả khi giá trị của POC của hình ảnh, lát, hoặc ô có thể là khác nhau, hình ảnh, lát, hoặc ô tương ứng với AU có cùng một giá trị AUC có thể được kết hợp với cùng một đối tượng thời gian xử lý/hiển thị. Khi thời gian xử lý được chứa theo khuôn dạng bộ chứa, ngay cả khi các ảnh tương ứng với các AU khác nhau, nếu các ảnh có cùng một thời điểm xử lý, các ảnh có thể được hiển thị cùng một đối tượng thời gian.

Trong các phương án giống nhau hoặc các phương án khác, mỗi hình ảnh, lát, hoặc ô có thể có cùng một ký hiệu nhận dạng theo thời gian (temporal_id)

trong cùng một AU. Tất cả hoặc tập con của các ảnh, các lát hoặc các ô tương ứng với đối tượng thời gian có thể được kết hợp với cùng một lớp con theo thời gian. Trong các phương án giống nhau hoặc các phương án khác, mỗi hình ảnh, lát, hoặc ô có thể có cùng hoặc có id lớp theo thời gian khác nhau (*layer_id*) trong cùng một AU. Tất cả hoặc tập con của các ảnh, các lát hoặc các ô tương ứng với đối tượng thời gian có thể được kết hợp với cùng một lớp hoặc lớp theo không gian khác nhau.

FIG.11 thể hiện ví dụ về dòng video bao gồm video mặt sau CSPS với *layer_id* bằng 0 và nhiều lớp CSPS mặt trước. Trong khi ảnh con được mã hóa có thể gồm có một hoặc nhiều lớp CSPS, vùng mặt sau, mà không thuộc về bất kỳ lớp CSPS mặt trước nào, có thể gồm có lớp nền. Lớp nền có thể chứa vùng mặt sau và các vùng mặt trước, trong khi đó lớp CSPS nâng cao chứa vùng mặt trước. Lớp CSPS nâng cao có thể có chất lượng phân nhìn tốt hơn lớp nền, ở cùng một vùng. Lớp CSPS nâng cao có thể tham chiếu các điểm ảnh được tái cấu trúc và các vector chuyển động của lớp nền, tương ứng với cùng một vùng.

Trong các phương án giống nhau hoặc một phương án khác, dòng bit video tương ứng với lớp nền được chứa trong phần theo dõi, trong khi đó các lớp CSPS tương ứng với mỗi ảnh con được chứa trong phần theo dõi riêng, trong tệp video.

Trong các phương án giống nhau hoặc một phương án khác, dòng bit video tương ứng với lớp nền được chứa trong phần theo dõi, trong khi đó các lớp CSPS với cùng *layer_id* được chứa trong phần theo dõi riêng. Trong ví dụ này, phần theo dõi tương ứng với lớp k chỉ bao gồm các lớp CSPS tương ứng với lớp k .

Trong các phương án giống nhau hoặc một phương án khác, mỗi lớp CSPS của mỗi ảnh con được lưu trữ trong phần theo dõi riêng. Mỗi phần theo dõi có thể hoặc có thể không có bất kỳ phân tích hoặc giải mã phần dư từ một hoặc nhiều phần theo dõi khác.

Trong các phương án giống nhau hoặc một phương án khác, mỗi phần theo dõi có thể chứa các dòng bit tương ứng với lớp i đến lớp j của các lớp CSPS

của tất cả hoặc tập con của ảnh con, trong đó $0 < i \leq j \leq k$, k là lớp cao nhất của CSPA.

Trong các phương án giống nhau hoặc một phương án khác, hình ảnh gồm có một hoặc nhiều dữ liệu truyền thông được liên kết bao gồm ánh xạ độ sâu, ánh xạ an-pha, dữ liệu địa lý 3D, ánh xạ chiếm giữ, v.v..Dữ liệu truyền thông được định thời được liên kết có thể được chia thành một hoặc nhiều dòng con dữ liệu mà mỗi trong số đó tương ứng với một ảnh con.

Trong cùng một phương án hoặc một phương án khác, FIG.12 thể hiện ví dụ của hội thảo video dựa trên phương pháp ảnh con được lấy đa lớp.Trong dòng video, một dòng bit video lớp nền tương ứng với ảnh mặt sau và một hoặc nhiều dòng bit video lớp nâng cao tương ứng với ảnh con mặt trước được chứa.Mỗi dòng bit video lớp nâng cao tương ứng với lớp CSPA.Trong phần hiển thị, hình ảnh tương ứng với lớp nền được hiển thị mặc định.Chứa một hoặc nhiều ảnh của người dùng trong hình ảnh (PIP).Khi người dùng cụ thể được lựa chọn bằng sự điều khiển của máy khách, lớp CSPA nâng cao tương ứng với người dùng được lựa chọn được giải mã và được hiển thị với chất lượng hoặc độ phân giải không gian nâng cao.FIG.13 thể hiện sơ đồ cho thao tác.

Trong các phương án giống nhau hoặc một phương án khác, hộp trung tâm mạng (như là bộ định tuyến) có thể lựa chọn tập con của các lớp để gửi đến người dùng phụ thuộc vào băng thông.Sự sắp xếp ảnh/ảnh con có thể được sử dụng cho việc thích ứng băng thông.Ví dụ, nếu người dùng không có băng thông, các dải bộ định tuyến của các lớp hoặc lựa chọn một số ảnh con do tính quan trọng của chúng hoặc dựa trên cài đặt được sử dụng và việc này có thể được hoàn thành động để thích ứng với băng thông.

FIG.14 thể hiện trường hợp sử dụng của video 360.Khi ảnh 360 hình cầu được chiếu lên ảnh phẳng, ảnh 360 trình chiếu có thể được phân vùng thành nhiều ảnh con như lớp nền.Lớp nâng cao của ảnh con cụ thể có thể được mã hóa và được truyền đến máy khách.Bộ giải mã có thể giải mã cả lớp nền bao gồm tất cả ảnh con và lớp nâng cao của ảnh con được chọn.Khi khung nhìn hiện tại giống với ảnh con

được chọn, hình ảnh được hiển thị có thể có chất lượng cao hơn với ảnh con được giải mã với lớp nâng cao. Ngược lại, hình ảnh được giải mã với lớp nền có thể được hiển thị, với chất lượng thấp.

Trong các phương án giống nhau hoặc một phương án khác, thông tin bố cục bất kỳ để hiển thị có thể thể hiện trong tệp tin, như thông tin bổ sung (như là bản tin SEI hoặc siêu dữ liệu). Một hoặc nhiều ảnh con được giải mã có thể được đặt lại và được hiển thị phụ thuộc vào thông tin bố cục được báo hiệu. Thông tin bố cục có thể được báo hiệu bởi máy chủ phát trực tuyến hoặc máy truyền thanh, hoặc có thể được tạo lại bởi thực thể mạng hoặc máy chủ đám mây, hoặc có thể được xác định bởi thiết lập tùy chỉnh của người dùng.

Trong phương án, khi hình ảnh đầu vào được chia thành một hoặc nhiều vùng con (hình chữ nhật), mỗi vùng con có thể được mã hóa như lớp độc lập. Mỗi lớp độc lập tương ứng với vùng cục bộ có thể có giá trị `layer_id` duy nhất. Đối với mỗi lớp độc lập, kích cỡ ảnh con và thông tin vị trí có thể được báo hiệu. Ví dụ, kích cỡ ảnh (độ rộng, độ cao), thông tin dịch vị của góc trên cùng bên trái (`x_offset`, `y_offset`). FIG.15 thể hiện ví dụ của bố cục của ảnh con được chia, kích cỡ ảnh con và thông tin vị trí và cấu trúc dự đoán hình ảnh tương ứng. Thông tin bố cục bao gồm (các) kích cỡ ảnh con và (các) vị trí ảnh con có thể được báo hiệu trong cấu trúc cú pháp bậc cao, như là (các) tập tham số, tiêu đề của lát hoặc nhóm ô, hoặc bản tin SEI.

Trong cùng một phương án, mỗi ảnh con tương ứng với lớp độc lập có thể có giá trị POC duy nhất trong AU. Khi hình ảnh tham chiếu trong số các ảnh được lưu trữ trong DPB được chỉ báo bằng cách sử dụng (các) phân tử cú pháp trong cấu trúc RPS hoặc RPL, (các) giá trị POC của mỗi ảnh con tương ứng với lớp có thể được sử dụng.

Trong các phương án giống nhau hoặc một phương án khác, để chỉ báo cấu trúc dự đoán (lớp liên đới), `layer_id` có thể không được sử dụng và giá trị (delta) POC có thể được sử dụng.

Trong cùng một phương án, ảnh con với giá trị POC bằng N tương ứng với lớp (hoặc vùng cục bộ) có thể hoặc có thể không được sử dụng làm hình ảnh tham chiếu của ảnh con với giá trị POC bằng $N+K$, tương ứng với cùng một lớp (hoặc cùng vùng cục bộ) cho dự đoán bù chuyển động. Trong hầu hết các trường hợp, giá trị của số K có thể bằng số lượng tối đa của các lớp (độc lập), mà có thể giống với số lượng các vùng con.

Trong cùng một phương án hoặc một phương án khác, FIG.16 thể hiện trường hợp mở rộng của FIG.15. Khi hình ảnh đầu vào được chia thành nhiều (ví dụ như bốn) vùng con, mỗi vùng cục bộ có thể được mã hóa với một hoặc nhiều các lớp. Trong trường hợp, số lượng các lớp độc lập có thể bằng số lượng các vùng con, và một hoặc nhiều các lớp có thể tương ứng với vùng con. Do đó, mỗi vùng con có thể được mã hóa với một hoặc nhiều lớp độc lập và bằng không hoặc nhiều lớp phụ thuộc.

Trong cùng một phương án, trên FIG.16, ảnh đầu vào có thể được chia thành bốn vùng con. Vùng con trên cùng bên phải có thể được mã hóa thành hai lớp, mà là lớp 1 và lớp 4, trong khi đó vùng con dưới cùng bên phải có thể được mã hóa thành hai lớp, mà là lớp 3 và lớp 5. Trong trường hợp này, lớp 4 có thể tham chiếu lớp 1 cho dự đoán bù chuyển động, trong khi đó lớp 5 có thể tham chiếu lớp 3 cho việc bù chuyển động.

Trong các phương án giống nhau và các phương án khác, việc lọc trong vòng (như là lọc giải khối, lọc trong vòng thích ứng, phục vụ, long song hướng hoặc lọc dựa trên học sâu bất kỳ) qua biên lớp có thể (một cách tùy chọn) được vô hiệu.

Trong các phương án giống nhau hoặc một phương án khác, dự đoán bù chuyển động hoặc sao chép nội khối qua biên lớp có thể (một cách tùy chọn) được vô hiệu.

Trong các phương án giống nhau hoặc một phương án khác, việc đệm biên cho dự đoán bù chuyển động hoặc lọc trong vòng ở biên của ảnh con có thể được xử lý một cách tùy chọn. Cờ chỉ báo để xem việc đệm biên được xử lý hay

không có thể được báo hiệu trong cấu trúc cú pháp bậc cao, như là (các) tập tham số (VPS, SPS, PPS, hoặc APS), tiêu đề lát hoặc nhóm ô, hoặc bản tin SEI.

Trong các phương án giống nhau hoặc một phương án khác, thông tin bố cục của (các) ảnh con (hoặc (các) ảnh con) có thể được báo hiệu trong VPS hoặc SPS. FIG.17 thể hiện ví dụ của các phần tử cú pháp trong VPS và SPS. Trong ví dụ này, `vps_sub_picture_dividing_flag` được báo hiệu trong VPS. Còn có thể chỉ báo xem (các) ảnh đầu vào được chia thành nhiều vùng con hay không. Khi giá trị của `vps_sub_picture_dividing_flag` bằng 0, (các) hình ảnh đầu vào trong (các) chuỗi video được mã hóa tương ứng với VPS hiện tại có thể không được chia thành nhiều vùng con. Trong trường hợp này, kích cỡ ảnh đầu vào có thể bằng hình ảnh được mã hóa kích cỡ (`pic_width_in_luma_samples`, `pic_height_in_luma_samples`), mà được báo hiệu trong SPS. Khi giá trị của `vps_sub_picture_dividing_flag` bằng 1, (các) hình ảnh đầu vào có thể được chia thành nhiều vùng con. Trong trường hợp này, các phần tử cú pháp `vps_full_pic_width_in_luma_samples` và `vps_full_pic_height_in_luma_samples` được báo hiệu trong VPS. Các giá trị của `vps_full_pic_width_in_luma_samples` và `vps_full_pic_height_in_luma_samples` có thể bằng độ rộng và độ cao của (các) hình ảnh đầu vào, theo cách tương ứng.

Trong cùng một phương án, các giá trị của `vps_full_pic_width_in_luma_samples` và `vps_full_pic_height_in_luma_samples` có thể không được sử dụng để giải mã, nhưng có thể được sử dụng cho xử lý và hiển thị.

Trong cùng một phương án, khi giá trị của `vps_sub_picture_dividing_flag` bằng 1, các phần tử cú pháp `pic_offset_x` và `pic_offset_y` có thể được báo hiệu trong SPS, mà tương ứng với (các) lớp cụ thể. Trong trường hợp này, hình ảnh được mã hóa kích cỡ (`pic_width_in_luma_samples`, `pic_height_in_luma_samples`) được báo hiệu trong SPS có thể bằng độ rộng và độ cao của vùng con tương ứng với lớp cụ

thể. Cũng vậy, vị trí (`pic_offset_x`, `pic_offset_y`) của góc trên cùng bên trái của vùng con có thể được báo hiệu trong SPS.

Trong cùng một phương án, thông tin vị trí (`pic_offset_x`, `pic_offset_y`) của góc trên cùng bên trái của vùng con có thể không được sử dụng để giải mã, nhưng có thể được sử dụng cho việc xử lý và hiển thị.

Trong cùng một phương án hoặc một phương án khác, thông tin bố cục (kích cỡ và vị trí) của tất cả hoặc (các) vùng con tập con của (các) ảnh đầu vào, thông tin dư thừa giữa (các) lớp có thể được báo hiệu trong tập tham số hoặc bản tin SEI. FIG.18 thể hiện ví dụ của các phần tử cú pháp để chỉ báo thông tin của bố cục của các vùng con, sự dư thừa giữa các lớp, và quan hệ giữa vùng con và một hoặc nhiều các lớp. Trong ví dụ này, phần tử cú pháp `num_sub_region` chỉ báo số lượng các vùng con (hình chữ nhật) trong chuỗi video được mã hóa hiện tại. Phần tử cú pháp `num_layers` chỉ báo số lượng các lớp trong chuỗi video được mã hóa hiện tại. Giá trị của `num_layers` có thể bằng hoặc lớn hơn giá trị của `num_sub_region`. Khi vùng con bất kỳ được mã hóa dưới dạng lớp đơn nhất, giá trị của `num_layers` có thể bằng giá trị của `num_sub_region`. Khi một hoặc nhiều các vùng con được mã hóa dưới dạng đa lớp, giá trị của `num_layers` có thể lớn hơn giá trị của `num_sub_region`. Phần tử cú pháp `direct_dependency_flag[i][j]` chỉ báo sự dư thừa từ lớp thứ `j` đến lớp thứ `i`. `num_layers_for_region[i]` chỉ báo số lượng các lớp được kết hợp với vùng con thứ `i`. `sub_region_layer_id[i][j]` chỉ báo `layer_id` của lớp thứ `j` được kết hợp với vùng con thứ `i`. `sub_region_offset_x[i]` và `sub_region_offset_y[i]` chỉ báo vị trí ngang và dọc của góc trên cùng bên trái của vùng con thứ `i`, theo cách tương ứng. `sub_region_width[i]` và `sub_region_height[i]` chỉ báo độ rộng và độ cao của vùng con thứ `i`, theo cách tương ứng.

Trong một phương án, một hoặc nhiều phần tử cú pháp mà chỉ định tập hợp lớp đầu ra để chỉ báo một trong nhiều lớp cần được xuất ra với hoặc mà không cần có thông tin cấp độ hồ sơ có thể được báo hiệu trong cấu trúc cú pháp bậc cao, ví dụ như VPS, DPS, SPS, PPS, APS hoặc bản tin SEI. Viện dẫn tới FIG.19, phần tử cú pháp `num_output_layer_sets` chỉ báo số lượng tập hợp lớp đầu ra (OLS)

trong chuỗi video được mã hóa viện dẫn tới VPS có thể được báo hiệu trong VPS. Đối với mỗi tập hợp lớp đầu ra, `output_layer_flag` có thể được báo hiệu theo số lượng như số lượng các lớp đầu ra.

Trong cùng một phương án, `output_layer_flag[ i ]` bằng 1 chỉ định rằng lớp thứ `i` là đầu ra. `vps_output_layer_flag[ i ]` bằng 0 chỉ định rằng lớp thứ `i` không phải là đầu ra.

Trong các phương án giống nhau hoặc một phương án khác, một hoặc nhiều phần tử cú pháp mà chỉ định thông tin cấp độ hồ sơ cho mỗi tập hợp lớp đầu ra có thể được báo hiệu trong cấu trúc cú pháp bậc cao, ví dụ như VPS, DPS, SPS, PPS, APS hoặc bản tin SEI. Viện dẫn tiếp tới FIG.19, phần tử cú pháp `num_profile_tile_level` chỉ báo số lượng thông tin cấp độ hồ sơ trên mỗi OLS trong chuỗi video được mã hóa viện dẫn tới VPS có thể được báo hiệu trong VPS. Đối với mỗi tập hợp lớp đầu ra, tập các phần tử cú pháp dùng cho thông tin cấp độ hồ sơ hoặc chỉ số chỉ báo thông tin cấp độ hồ sơ cụ thể trong số các mục nhập trong thông tin cấp độ hồ sơ có thể được báo hiệu với số lượng như số lượng các lớp đầu ra.

Trong cùng một phương án, `profile_tier_level_idx[ i ][ j ]` chỉ định chỉ số, vào danh sách của các cấu trúc cú pháp `profile_tier_level( )` trong VPS, của cấu trúc cú pháp `profile_tier_level( )` mà áp dụng cho lớp thứ `j` của OLS thứ `i`.

Trong cùng một phương án hoặc một phương án khác, viện dẫn tới FIG.20, các phần tử cú pháp `num_profile_tile_level` và/hoặc `num_output_layer_sets` có thể được báo hiệu khi số lượng tối đa của các lớp lớn hơn 1 (`vps_max_layers_minus1 > 0`).

Trong cùng một phương án hoặc một phương án khác, viện dẫn tới FIG.20, phần tử cú pháp `vps_output_layers_mode[ i ]` chỉ báo chế độ báo hiệu lớp đầu ra cho tập hợp lớp đầu ra thứ `i` có thể được thể hiện trong VPS.

Trong cùng một phương án, `vps_output_layers_mode[ i ]` bằng 0 chỉ định rằng chỉ lớp cao nhất là đầu ra với tập hợp lớp đầu ra thứ `i`. `vps_output_layer_mode[ i ]` bằng 1 chỉ định rằng tất cả các lớp là đầu ra với tập

hợp lớp đầu ra thứ i . $vps_output_layer_mode[i]$ bằng 2 chỉ định rằng các lớp mà là đầu ra là các lớp với $vps_output_layer_flag[i][j]$ bằng 1 với tập hợp lớp đầu ra thứ i . Nhiều giá trị hơn có thể được dành riêng.

Trong cùng một phương án, $output_layer_flag[i][j]$ có thể hoặc có thể không được báo hiệu phụ thuộc vào giá trị của $vps_output_layers_mode[i]$ cho tập hợp lớp đầu ra thứ i .

Trong cùng một phương án hoặc một phương án khác, viện dẫn tới FIG.20, cờ $vps_ptl_signal_flag[i]$ có thể được thể hiện cho tập hợp lớp đầu ra thứ i . Phụ thuộc vào giá trị của $vps_ptl_signal_flag[i]$, thông tin cấp độ hồ sơ cho tập hợp lớp đầu ra thứ i có thể hoặc có thể không được báo hiệu.

Trong cùng một phương án hoặc một phương án khác, viện dẫn tới FIG.21, số lượng ảnh con, $max_subpics_minus1$, trong CVS hiện tại có thể được báo hiệu trong cấu trúc cú pháp bậc cao, ví dụ như VPS, DPS, SPS, PPS, APS hoặc bản tin SEI.

Trong cùng một phương án, viện dẫn tới FIG.21, ký hiệu nhận dạng ảnh con, $sub_pic_id[i]$, cho ảnh con thứ i có thể được báo hiệu, khi số lượng các ảnh con lớn hơn 1 ($max_subpics_minus1 > 0$).

Trong các phương án giống nhau hoặc một phương án khác, một hoặc nhiều phần tử cú pháp chỉ báo ký hiệu nhận dạng ảnh con thuộc về mỗi lớp của mỗi tập hợp lớp đầu ra có thể được báo hiệu trong VPS. Viện dẫn tới các FIG.22 và 23, $sub_pic_id_layer[i][j][k]$, mà chỉ báo ảnh con thứ k thể hiện trong lớp thứ j của tập hợp lớp đầu ra thứ i . Với thông tin đó, bộ giải mã có thể nhận biết việc ảnh con có thể được giải mã và được đầu ra cho mỗi lớp của tập hợp lớp đầu ra cụ thể.

Trong phương án, tiêu đề ảnh (PH) là cấu trúc cú pháp mà chứa các phần tử cú pháp mà áp dụng cho tất cả các lát của ảnh được mã hóa. Đơn vị ảnh (PU) là tập hợp của các đơn vị NAL mà được kết hợp với nhau theo quy tắc phân loại được chỉ định, là liên tiếp theo thứ tự giải mã, và chứa đúng một ảnh được mã hóa. PU có thể chứa tiêu đề ảnh (PH) và một hoặc nhiều đơn vị VCL NAL mà đưa ra ảnh được mã hóa.

Trong phương án, SPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, nằm trong ít nhất một AU với TemporalId bằng 0 hoặc được cung cấp qua phương tiện ngoại vi.

Trong phương án, SPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, nằm trong ít nhất một AU với TemporalId bằng 0 trong CVS, mà chứa một hoặc nhiều PPS liên quan đến SPS, hoặc được cung cấp qua phương tiện ngoại vi.

Trong phương án, SPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PPS, nằm trong ít nhất một PU với nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các đơn vị PPS NAL mà liên quan đến đơn vị SPS NAL trong CVS, mà chứa một hoặc nhiều PPS liên quan đến SPS, hoặc được cung cấp qua phương tiện ngoại vi.

Trong phương án, SPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PPS, nằm trong ít nhất một PU với TemporalId bằng 0 và nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các đơn vị PPS NAL mà liên quan đến đơn vị SPS NAL hoặc được cung cấp qua phương tiện ngoại vi.

Trong phương án, SPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PPS, nằm trong ít nhất một PU với TemporalId bằng 0 và nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các đơn vị PPS NAL mà liên quan đến đơn vị SPS NAL trong CVS, mà chứa một hoặc nhiều PPS liên quan đến SPS, hoặc được cung cấp qua phương tiện ngoại vi.

Trong các phương án giống nhau hoặc một phương án khác, pps_seq_parameter_set_id chỉ định giá trị của sps_seq_parameter_set_id cho SPS được tham chiếu. Giá trị của pps_seq_parameter_set_id có thể là bằng nhau trong tất cả các PPS mà được viện dẫn đến bởi các ảnh được mã hóa trong CLVS.

Trong các phương án giống nhau hoặc một phương án khác, tất cả các đơn vị SPS NAL với giá trị cụ thể của sps_seq_parameter_set_id trong CVS có thể có cùng nội dung.

Trong các phương án giống nhau hoặc một phương án khác, bất kể các giá trị `nuh_layer_id` là bao nhiêu, các đơn vị SPS NAL có thể chia sẻ cùng một khoảng giá trị của `sps_seq_parameter_set_id`.

Trong các phương án giống nhau hoặc một phương án khác, giá trị `nuh_layer_id` của đơn vị SPS NAL có thể bằng giá trị `nuh_layer_id` thấp nhất của các đơn vị PPS NAL mà liên quan đến đơn vị SPS NAL.

Trong phương án, khi SPS với `nuh_layer_id` bằng m được viện dẫn tới bởi một hoặc nhiều PPS với `nuh_layer_id` bằng n . Lớp với `nuh_layer_id` bằng m có thể tương tự như lớp với `nuh_layer_id` bằng n hoặc lớp tham chiếu (trực tiếp hoặc gián tiếp) của lớp với `nuh_layer_id` bằng m .

Trong phương án, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, nằm trong ít nhất một AU với `TemporalId` bằng `TemporalId` của đơn vị PPS NAL hoặc được cung cấp qua phương tiện ngoại vi.

Trong phương án, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, nằm trong ít nhất một AU với `TemporalId` bằng `TemporalId` của đơn vị PPS NAL trong CVS, mà chứa một hoặc nhiều PH (hoặc các đơn vị NAL lát được mã hóa) liên quan đến PPS, hoặc được cung cấp qua phương tiện ngoại vi.

Trong phương án, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PH (hoặc các đơn vị NAL lát được mã hóa), nằm trong ít nhất một PU với `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của các đơn vị NAL lát được mã hóa mà liên quan đến đơn vị PPS NAL trong CVS, mà chứa một hoặc nhiều PH (hoặc các đơn vị NAL lát được mã hóa) liên quan đến PPS, hoặc được cung cấp qua phương tiện ngoại vi.

Trong phương án, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PH (hoặc các đơn vị NAL lát được mã hóa), nằm trong ít nhất một PU với `TemporalId` bằng `TemporalId` của đơn vị PPS NAL và `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của các đơn vị NAL lát được mã hóa mà liên quan đến đơn vị PPS NAL trong CVS, mà chứa

một hoặc nhiều PH (hoặc các đơn vị NAL lát được mã hóa) liên quan đến PPS, hoặc được cung cấp qua phương tiện ngoại vi.

Trong các phương án giống nhau hoặc một phương án khác, `ph_pic_parameter_set_id` trong PH chỉ định giá trị của `pps_pic_parameter_set_id` cho PPS được tham chiếu trong sử dụng. Giá trị của `pps_seq_parameter_set_id` có thể là bằng nhau trong tất cả các PPS mà được viện dẫn đến bởi các ảnh được mã hóa trong CLVS.

Trong các phương án giống nhau hoặc một phương án khác, tất cả các đơn vị PPS NAL với giá trị cụ thể của `pps_pic_parameter_set_id` trong PU có thể có cùng nội dung.

Trong các phương án giống nhau hoặc một phương án khác, bất kể các giá trị `nuh_layer_id` là bao nhiêu, các đơn vị PPS NAL có thể chia sẻ cùng một khoảng giá trị của `pps_pic_parameter_set_id`.

Trong các phương án giống nhau hoặc một phương án khác, giá trị `nuh_layer_id` của đơn vị PPS NAL có thể bằng giá trị `nuh_layer_id` thấp nhất của các đơn vị NAL lát được mã hóa mà liên quan đến các đơn vị NAL mà liên quan đến đơn vị PPS NAL.

Trong phương án, khi PPS với `nuh_layer_id` bằng m được viện dẫn tới bởi một hoặc nhiều đơn vị NAL lát được mã hóa với `nuh_layer_id` bằng n . Lớp với `nuh_layer_id` bằng m có thể tương tự như lớp với `nuh_layer_id` bằng n hoặc lớp tham chiếu (trực tiếp hoặc gián tiếp) của lớp với `nuh_layer_id` bằng m .

Trong phương án, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, nằm trong ít nhất một AU với `TemporalId` bằng `TemporalId` của đơn vị PPS NAL hoặc được cung cấp qua phương tiện ngoại vi.

Trong phương án, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, nằm trong ít nhất một AU với `TemporalId` bằng `TemporalId` của đơn vị PPS NAL trong CVS, mà chứa một hoặc nhiều PH (hoặc các đơn vị NAL lát được mã hóa) liên quan đến PPS, hoặc được cung cấp qua

phương tiện ngoại vi.

Trong phương án, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PH (hoặc các đơn vị NAL lát được mã hóa), nằm trong ít nhất một PU với nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các đơn vị NAL lát được mã hóa mà liên quan đến đơn vị PPS NAL trong CVS, mà chứa một hoặc nhiều PH (hoặc các đơn vị NAL lát được mã hóa) liên quan đến PPS, hoặc được cung cấp qua phương tiện ngoại vi.

Trong phương án, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PH (hoặc các đơn vị NAL lát được mã hóa), nằm trong ít nhất một PU với TemporalId bằng TemporalId của đơn vị PPS NAL và nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các đơn vị NAL lát được mã hóa mà liên quan đến đơn vị PPS NAL trong CVS, mà chứa một hoặc nhiều PH (hoặc các đơn vị NAL lát được mã hóa) liên quan đến PPS, hoặc được cung cấp qua phương tiện ngoại vi.

Trong các phương án giống nhau hoặc một phương án khác, ph_pic_parameter_set_id trong PH chỉ định giá trị của pps_pic_parameter_set_id cho PPS được tham chiếu trong sử dụng. Giá trị của pps_seq_parameter_set_id có thể là bằng nhau trong tất cả các PPS mà được viện dẫn đến bởi các ảnh được mã hóa trong CLVS.

Trong các phương án giống nhau hoặc một phương án khác, tất cả các đơn vị PPS NAL với giá trị cụ thể của pps_pic_parameter_set_id trong PU có thể có cùng nội dung.

Trong các phương án giống nhau hoặc một phương án khác, bất kể các giá trị nuh_layer_id là bao nhiêu, các đơn vị PPS NAL có thể chia sẻ cùng một khoảng giá trị của pps_pic_parameter_set_id.

Trong các phương án giống nhau hoặc một phương án khác, giá trị nuh_layer_id của đơn vị PPS NAL có thể bằng giá trị nuh_layer_id thấp nhất của các đơn vị NAL lát được mã hóa mà liên quan đến các đơn vị NAL mà liên quan đến đơn vị PPS NAL.

Trong phương án, khi PPS với `nuh_layer_id` bằng m được viện dẫn tới bởi một hoặc nhiều đơn vị NAL lát được mã hóa với `nuh_layer_id` bằng n . Lớp với `nuh_layer_id` bằng m có thể tương tự như lớp với `nuh_layer_id` bằng n hoặc lớp tham chiếu (trực tiếp hoặc gián tiếp) của lớp với `nuh_layer_id` bằng m .

Lớp đầu ra chỉ báo lớp của tập lớp đầu ra mà được xuất ra. Tập lớp đầu ra (OLS) chỉ báo tập các lớp gồm có tập các lớp được chỉ định, trong đó một hoặc nhiều lớp trong tập các lớp được chỉ định là các lớp đầu ra. Chỉ số lớp tập lớp đầu ra (OLS) là chỉ số, của lớp trong OLS, cho danh sách của các lớp trong OLS.

Lớp con chỉ báo lớp có thể biến đổi tỷ lệ theo thời gian của dòng bit có thể biến đổi tỷ lệ theo thời gian, gồm có các đơn vị VCL NAL với giá trị cụ thể của biến `TemporalId` và các đơn vị không VCL NAL được liên kết. Sự biểu diễn lớp con chỉ báo tập con của dòng bit gồm có các đơn vị NAL của lớp con cụ thể và các lớp con thấp hơn.

VPS RBSP có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, nằm trong ít nhất một AU với `TemporalId` bằng 0 hoặc được cung cấp qua phương tiện ngoại vi. Tất cả các đơn vị VPS NAL với giá trị cụ thể của `vps_video_parameter_set_id` trong CVS có thể có cùng nội dung.

`vps_video_parameter_set_id` cung cấp ký hiệu nhận dạng cho VPS cho việc tham chiếu bởi các phần tử cú pháp khác. Giá trị của `vps_video_parameter_set_id` có thể lớn hơn 0.

`vps_max_layers_minus1` plus 1 chỉ định số lượng các lớp được cho phép tối đa trong mỗi CVS mà liên quan đến VPS.

`vps_max_sublayers_minus1` plus 1 chỉ định số lượng các lớp con thời gian tối đa mà có thể có trong lớp trong mỗi CVS mà liên quan đến VPS. Giá trị của `vps_max_sublayers_minus1` có thể nằm trong khoảng 0 đến 6, bao gồm cả hai giá trị đầu và cuối.

`vps_all_layers_same_num_sublayers_flag` bằng 1 chỉ định rằng số lượng các lớp con thời gian là tương tự cho tất cả các lớp trong mỗi CVS mà liên

quan đến `VPS.vps_all_layers_same_num_sublayers_flag` bằng 0 chỉ định rằng các lớp trong mỗi CVS mà liên quan đến VPS có thể hoặc có thể không có cùng số lượng các lớp con thời gian. Khi không hiện diện, giá trị của `vps_all_layers_same_num_sublayers_flag` được thu nhận bằng 1.

`vps_all_independent_layers_flag` bằng 1 chỉ định rằng tất cả các lớp trong CVS được mã hóa độc lập mà không sử dụng dự đoán lớp liên đới. `vps_all_independent_layers_flag` bằng 0 chỉ định rằng một hoặc nhiều các lớp trong CVS có thể sử dụng dự đoán lớp liên đới. Khi không hiện diện, giá trị của `vps_all_independent_layers_flag` được thu nhận bằng 1.

`vps_layer_id[ i ]` chỉ định giá trị `nuh_layer_id` của lớp thứ *i*. Đối với hai giá trị nguyên không âm bất kỳ của *m* và *n*, khi *m* nhỏ hơn *n*, giá trị của `vps_layer_id[ m ]` có thể nhỏ hơn `vps_layer_id[ n ]`.

`vps_independent_layer_flag[ i ]` bằng 1 chỉ định rằng lớp với chỉ số *i* không sử dụng dự đoán lớp liên đới. `vps_independent_layer_flag[ i ]` bằng 0 chỉ định rằng lớp với chỉ số *i* có thể sử dụng dự đoán lớp liên đới và các phần tử cú pháp `vps_direct_ref_layer_flag[ i ][ j ]` cho *j* nằm trong khoảng 0 đến *i* - 1, bao gồm cả hai giá trị đầu và cuối, được thể hiện trong VPS. Khi không hiện diện, giá trị của `vps_independent_layer_flag[ i ]` được thu nhận bằng 1.

`vps_direct_ref_layer_flag[ i ][ j ]` bằng 0 chỉ định rằng lớp với chỉ số *j* không là lớp tham chiếu trực tiếp cho lớp với chỉ số *i*. `vps_direct_ref_layer_flag[ i ][ j ]` bằng 1 chỉ định rằng lớp với chỉ số *j* là lớp tham chiếu trực tiếp cho lớp với chỉ số *i*. Khi `vps_direct_ref_layer_flag[ i ][ j ]` không được thể hiện cho *i* và *j* nằm trong khoảng 0 đến `vps_max_layers_minus1`, bao gồm cả hai giá trị đầu và cuối, nó được thu nhận bằng 0. Khi `vps_independent_layer_flag[ i ]` bằng 0, có thể là ít nhất một giá trị của *j* nằm trong khoảng 0 đến *i* - 1, bao gồm cả hai giá trị đầu và cuối, như vậy giá trị của `vps_direct_ref_layer_flag[ i ][ j ]` bằng 1.

Các biến `NumDirectRefLayers[ i ]`, `DirectRefLayerIdx[ i ][ d ]`, `NumRefLayers[ i ]`, `RefLayerIdx[ i ][ r ]`, và `LayerUsedAsRefLayerFlag[ j ]` được lấy như sau:

```

for( i = 0; i <= vps_max_layers_minus1; i++ ) {
    for( j = 0; j <= vps_max_layers_minus1; j++ ) {
        dependencyFlag[ i ][ j ] = vps_direct_ref_layer_flag[ i ][ j ]
        for( k = 0; k < i; k++ )
            if( vps_direct_ref_layer_flag[ i ][ k ] && dependencyFlag[ k ][ j ] )
                dependencyFlag[ i ][ j ] = 1
    }
    LayerUsedAsRefLayerFlag[ i ] = 0
}
for( i = 0; i <= vps_max_layers_minus1; i++ ) {
    for( j = 0, d = 0, r = 0; j <= vps_max_layers_minus1; j++ ) {
        if( vps_direct_ref_layer_flag[ i ][ j ] ) {
            DirectRefLayerIdx[ i ][ d++ ] = j
            LayerUsedAsRefLayerFlag[ j ] = 1
        }
        if( dependencyFlag[ i ][ j ] )
            RefLayerIdx[ i ][ r++ ] = j
    }
    NumDirectRefLayers[ i ] = d
    NumRefLayers[ i ] = r
}

```

Biến GeneralLayerIdx[i], chỉ định chỉ số lớp của lớp với nuh_layer_id bằng vps_layer_id[i], được lấy như sau:

```

for( i = 0; i <= vps_max_layers_minus1; i++ )
    GeneralLayerIdx[ vps_layer_id[ i ] ] = i

```

Đối với hai giá trị khác nhau bất kỳ của i và j, cả hai nằm trong khoảng 0 đến vps_max_layers_minus1, bao gồm cả hai giá trị đầu và cuối, khi dependencyFlag[i][j] bằng 1, có yêu cầu về sự phù hợp dòng bit mà các giá trị của chroma_format_idc và bit_depth_minus8 áp dụng cho lớp thứ i có thể bằng

các giá trị của `chroma_format_idc` và `bit_depth_minus8`, theo cách tương ứng, mà áp dụng cho lớp thứ j .

`max_tid_ref_present_flag[ i ]` bằng 1 chỉ định rằng phần tử cú pháp `max_tid_il_ref_pics_plus1[ i ]` hiện diện. `max_tid_ref_present_flag[ i ]` bằng 0 chỉ định rằng phần tử cú pháp `max_tid_il_ref_pics_plus1[ i ]` không hiện diện.

`max_tid_il_ref_pics_plus1[ i ]` bằng 0 chỉ định rằng dự đoán lớp liên đới không được sử dụng bởi các ảnh không IRAP của lớp thứ i . `max_tid_il_ref_pics_plus1[ i ]` lớn hơn 0 chỉ định rằng, để giải mã các ảnh của lớp thứ i , không có ảnh nào có `TemporalId` lớn hơn `max_tid_il_ref_pics_plus1[ i ] - 1` được sử dụng làm ILRP. Khi không hiện diện, giá trị của `max_tid_il_ref_pics_plus1[ i ]` được thu nhận bằng 7.

`each_layer_is_an_ols_flag` bằng 1 chỉ định rằng mỗi OLS chỉ chứa một lớp và mỗi lớp của nó trong CVS mà liên quan đến VPS là OLS với lớp đơn nhất được bao gồm là lớp đầu ra duy nhất. `each_layer_is_an_ols_flag` bằng 0 mà OLS có thể chứa nhiều hơn một lớp. Nếu `vps_max_layers_minus1` bằng 0, giá trị của `each_layer_is_an_ols_flag` được thu nhận bằng 1. Ngược lại, khi `vps_all_independent_layers_flag` bằng 0, giá trị của `each_layer_is_an_ols_flag` được thu nhận bằng 0.

`ols_mode_idc` bằng 0 chỉ định rằng tổng số OLS được chỉ định bởi VPS bằng `vps_max_layers_minus1 + 1`, OLS thứ i bao gồm các lớp với các chỉ số lớp từ 0 đến i , bao gồm cả hai giá trị đầu và cuối, và đối với mỗi OLS chỉ lớp cao nhất trong OLS được xuất ra.

`ols_mode_idc` bằng 1 chỉ định rằng tổng số OLS được chỉ định bởi VPS bằng `vps_max_layers_minus1 + 1`, OLS thứ i bao gồm các lớp với các chỉ số lớp từ 0 đến i , bao gồm cả hai giá trị đầu và cuối, và đối với mỗi OLS tất cả các lớp trong OLS được xuất ra.

`ols_mode_idc` bằng 2 chỉ định rằng tổng số OLS được chỉ định bởi VPS được báo hiệu rõ ràng và đối với mỗi OLS các lớp đầu ra được báo hiệu rõ ràng và

các lớp khác là các lớp mà là các lớp tham chiếu trực tiếp hoặc gián tiếp của các lớp đầu của OLS.

Giá trị của `ols_mode_idc` có thể nằm trong khoảng 0 đến 2, bao gồm cả hai giá trị đầu và cuối. Giá trị 3 của `ols_mode_idc` được dành riêng cho việc sử dụng trong tương lai bởi ITU-T | ISO/IEC.

Khi `vps_all_independent_layers_flag` bằng 1 và `each_layer_is_an_ols_flag` bằng 0, giá trị của `ols_mode_idc` được thu nhận bằng 2.

num_output_layer_sets_minus1 plus 1 chỉ định tổng số OLS được chỉ định bởi VPS khi `ols_mode_idc` bằng 2.

Biến `TotalNumOlss`, chỉ định tổng số của các OLS được chỉ định bởi VPS, được lấy như sau:

```
if( vps_max_layers_minus1 == 0 )
    TotalNumOlss = 1
else if( each_layer_is_an_ols_flag || ols_mode_idc == 0 || ols_mode_idc == 1 )
    TotalNumOlss = vps_max_layers_minus1 + 1
else if( ols_mode_idc == 2 )
    TotalNumOlss = num_output_layer_sets_minus1 + 1
```

ols_output_layer_flag[i][j] bằng 1 chỉ định rằng lớp có `nuh_layer_id` bằng `vps_layer_id[ j ]` là lớp đầu ra của OLS thứ *i* khi `ols_mode_idc` bằng 2. **ols_output_layer_flag[i][j]** bằng 0 chỉ định rằng lớp có `nuh_layer_id` bằng `vps_layer_id[ j ]` không là lớp đầu ra của OLS thứ *i* khi `ols_mode_idc` bằng 2.

Biến `NumOutputLayersInOls[ i ]`, mà chỉ định số lượng các lớp đầu ra trong OLS thứ *i*, biến `NumSubLayersInLayerInOLS[ i ][ j ]`, mà chỉ định số lượng các lớp con trong lớp thứ *j* trong OLS thứ *i*, biến `OutputLayerIdInOls[ i ][ j ]`, mà chỉ định giá trị `nuh_layer_id` của lớp đầu ra thứ *j* trong OLS thứ *i*, và biến `LayerUsedAsOutputLayerFlag[ k ]`, mà chỉ định xem lớp thứ *k* có được sử dụng làm lớp đầu ra trong ít nhất một OLS không, được lấy như sau:

```

NumOutputLayersInOls[ 0 ] = 1
OutputLayerIdInOls[ 0 ][ 0 ] = vps_layer_id[ 0 ]
NumSubLayersInLayerInOLS[ 0 ][ 0 ] = vps_max_sub_layers_minus1 + 1
LayerUsedAsOutputLayerFlag[ 0 ] = 1
for( i = 1, i <= vps_max_layers_minus1; i++ ) {
    if( each_layer_is_an_ols_flag || ols_mode_idc < 2 )
        LayerUsedAsOutputLayerFlag[ i ] = 1
    else /*( !each_layer_is_an_ols_flag && ols_mode_idc == 2 ) */
        LayerUsedAsOutputLayerFlag[ i ] = 0
}
for( i = 1; i < TotalNumOlss; i++ )
    if( each_layer_is_an_ols_flag || ols_mode_idc == 0 ) {
        NumOutputLayersInOls[ i ] = 1
        OutputLayerIdInOls[ i ][ 0 ] = vps_layer_id[ i ]
        for( j = 0; j < i && ( ols_mode_idc == 0 ); j++ )
            NumSubLayersInLayerInOLS[ i ][ j ] = max_tid_il_ref_pics_plus1[ i ]
        NumSubLayersInLayerInOLS[ i ][ i ] = vps_max_sub_layers_minus1 + 1
    } else if( ols_mode_idc == 1 ) {
        NumOutputLayersInOls[ i ] = i + 1
        for( j = 0; j < NumOutputLayersInOls[ i ]; j++ ) {
            OutputLayerIdInOls[ i ][ j ] = vps_layer_id[ j ]
            NumSubLayersInLayerInOLS[ i ][ j ] =
vps_max_sub_layers_minus1 + 1
        }
    } else if( ols_mode_idc == 2 ) {
        for( j = 0; j <= vps_max_layers_minus1; j++ ) {
            layerIncludedInOlsFlag[ i ][ j ] = 0
            NumSubLayersInLayerInOLS[ i ][ j ] = 0
        }
    }
for( k = 0, j = 0; k <= vps_max_layers_minus1; k++ ) (40)

```

```

    if( ols_output_layer_flag[ i ][ k ] ) {
        layerIncludedInOlsFlag[ i ][ k ] = 1
        LayerUsedAsOutputLayerFlag[ k ] = 1
        OutputLayerIdx[ i ][ j ] = k
        OutputLayerIdInOls[ i ][ j++ ] = vps_layer_id[ k ]
        NumSubLayersInLayerInOLS[ i ][ j ] =
vps_max_sub_layers_minus1 + 1
    }
    NumOutputLayersInOls[ i ] = j
    for( j = 0; j < NumOutputLayersInOls[ i ]; j++ ) {
        idx = OutputLayerIdx[ i ][ j ]
        for( k = 0; k < NumRefLayers[ idx ]; k++ ) {
            layerIncludedInOlsFlag[ i ][ RefLayerIdx[ idx ][ k ] ] = 1
            if( NumSubLayersInLayerInOLS[ i ][ RefLayerIdx[ idx ][ k ] ] <
max_tid_il_ref_pics_plus1[ OutputLayerIdInOls[ i ][ j ] ] )
                NumSubLayersInLayerInOLS[ i ][ RefLayerIdx[ idx ][ k ] ] =
=
max_tid_il_ref_pics_plus1[ OutputLayerIdInOls[ i ][ j ] ]
        }
    }
}

```

Đối với mỗi giá trị của i nằm trong khoảng 0 đến $vps_max_layers_minus1$, bao gồm cả hai giá trị đầu và cuối, các giá trị của $LayerUsedAsRefLayerFlag[i]$ và $LayerUsedAsOutputLayerFlag[i]$ có thể đều không bằng 0. Nói cách khác, có thể không có lớp nào mà đều không là lớp đầu ra của ít nhất một OLS và lớp tham chiếu trực tiếp của lớp bất kỳ khác.

Đối với mỗi OLS, có thể là ít nhất một lớp mà là lớp đầu ra. Nói cách khác, đối với mỗi giá trị của i nằm trong khoảng 0 đến $TotalNumOlss - 1$, bao

gồm cả hai giá trị đầu và cuối, giá trị của NumOutputLayersInOls[i] có thể là lớn hơn hoặc bằng 1.

Biến NumLayersInOls[i], mà chỉ định số lượng các lớp trong OLS thứ i, và biến LayerIdInOls[i][j], chỉ định giá trị nuh_layer_id của lớp thứ j trong OLS thứ i, được lấy như sau:

```

NumLayersInOls[ 0 ] = 1
LayerIdInOls[ 0 ][ 0 ] = vps_layer_id[ 0 ]
for( i = 1; i < TotalNumOlss; i++ ) {
    if( each_layer_is_an_ols_flag ) {
        NumLayersInOls[ i ] = 1
        LayerIdInOls[ i ][ 0 ] = vps_layer_id[ i ]
    } else if( ols_mode_idc == 0 || ols_mode_idc == 1 ) {
        NumLayersInOls[ i ] = i + 1
        for( j = 0; j < NumLayersInOls[ i ]; j++ )
            LayerIdInOls[ i ][ j ] = vps_layer_id[ j ]
    } else if( ols_mode_idc == 2 ) {
        for( k = 0, j = 0; k <= vps_max_layers_minus1; k++ )
            if( layerIncludedInOlsFlag[ i ][ k ] )
                LayerIdInOls[ i ][ j++ ] = vps_layer_id[ k ]
        NumLayersInOls[ i ] = j
    }
}

```

Biến OlsLayerIdx[i][j], mà chỉ định chỉ số lớp OLS của lớp có nuh_layer_id bằng LayerIdInOls[i][j], được lấy như sau:

```

for( i = 0; i < TotalNumOlss; i++ )
    for j = 0; j < NumLayersInOls[ i ]; j++ )
        OlsLayerIdx[ i ][ LayerIdInOls[ i ][ j ] ] = j

```

Lớp thấp nhất trong mỗi OLS có thể là lớp độc lập. Nói cách khác, đối với mỗi i nằm trong khoảng 0 đến TotalNumOlss - 1, bao gồm cả hai giá trị đầu

và cuối, giá trị của `vps_independent_layer_flag[ GeneralLayerIdx[ LayerIdInOls[ i ][ 0 ] ] ]` có thể bằng 1.

Mỗi lớp có thể nằm trong ít nhất một OLS được chỉ định bởi VPS. Nói cách khác, đối với mỗi lớp có giá trị cụ thể của `nuh_layer_id` `nuhLayerId` bằng một trong `vps_layer_id[ k ]` với `k` nằm trong khoảng 0 đến `vps_max_layers_minus1`, bao gồm cả hai giá trị đầu và cuối, có thể là ít nhất một cặp giá trị `i` và `j`, trong đó `i` nằm trong khoảng 0 đến `TotalNumOls - 1`, bao gồm cả hai giá trị đầu và cuối, và `j` nằm trong khoảng `NumLayersInOls[ i ] - 1`, bao gồm cả hai giá trị đầu và cuối, như là giá trị của `LayerIdInOls[ i ][ j ]` bằng `nuhLayerId`.

Trong phương án, quá trình giải mã thao tác như sau cho ảnh hiện tại `CurrPic`:

- `PictureOutputFlag` được thiết đặt như sau:
- Nếu một trong các điều kiện sau là đúng, `PictureOutputFlag` được thiết đặt bằng 0:
 - Ảnh hiện tại là ảnh RASL và `NoOutputBeforeRecoveryFlag` của ảnh IRAP được liên kết bằng 1.
 - `gdr_enabled_flag` bằng 1 và ảnh hiện tại là ảnh GDR với `NoOutputBeforeRecoveryFlag` bằng 1.
 - `gdr_enabled_flag` bằng 1, ảnh hiện tại được kết hợp với ảnh GDR với `NoOutputBeforeRecoveryFlag` bằng 1, và `PicOrderCntVal` của ảnh hiện tại nhỏ hơn `RpPicOrderCntVal` của ảnh GDR được liên kết.
- `sps_video_parameter_set_id` lớn hơn 0, `ols_mode_idc` bằng 0 và AU hiện tại chứa ảnh `picA` mà thỏa mãn tất cả các điều kiện sau đây:
 - `PicA` có `PictureOutputFlag` bằng 1.
 - `PicA` có `nuh_layer_id` `nuhLid` lớn hơn của ảnh hiện tại.
 - `PicA` thuộc về lớp đầu ra của OLS (tức là, `OutputLayerIdInOls[ TargetOlsIdx ][ 0 ]` bằng `nuhLid`).
- `sps_video_parameter_set_id` lớn hơn 0, `ols_mode_idc` bằng 2, và

ols_output_layer_flag[TargetOlsIdx][GeneralLayerIdx[nuh_layer_id]] bằng 0.

– Ngược lại, PictureOutputFlag được thiết đặt bằng pic_output_flag.

Sau khi tất cả các lát của ảnh hiện tại được giải mã, ảnh được giải mã hiện thời được đánh dấu là "được sử dụng để tham chiếu ngắn hạn", và mỗi mục nhập ILRP trong RefPicList[0] hoặc RefPicList[1] được đánh dấu là "được sử dụng để tham chiếu ngắn hạn".

Trong cùng phương án hoặc một phương án khác, khi mỗi lớp là tập lớp đầu ra, PictureOutputFlag được thiết đặt bằng pic_output_flag, bất kể giá trị của ols_mode_idc là bao nhiêu.

Trong cùng phương án hoặc một phương án khác, PictureOutputFlag được thiết đặt bằng 0 khi sps_video_parameter_set_id lớn hơn 0, each_layer_is_an_ols_flag bằng 0, ols_mode_idc bằng 0 và AU hiện tại chứa ảnh picA mà thỏa mãn tất cả các điều kiện sau đây: PicA có PictureOutputFlag bằng 1, PicA có nuh_layer_id nuhLid lớn hơn của ảnh hiện tại và PicA thuộc về lớp đầu ra của OLS (tức là, OutputLayerIdInOls[TargetOlsIdx][0] bằng nuhLid).

Trong cùng phương án hoặc một phương án khác, PictureOutputFlag được thiết đặt bằng 0 khi sps_video_parameter_set_id lớn hơn 0, each_layer_is_an_ols_flag bằng 0, ols_mode_idc bằng 2, và ols_output_layer_flag[TargetOlsIdx][GeneralLayerIdx[nuh_layer_id]] bằng 0.

Khi ảnh có thể hoặc có thể không được tham chiếu bởi một hoặc nhiều các ảnh kế tiếp theo thứ tự giải mã, cho bù chuyển động hoặc dự đoán tham số. Còn mà chỉ báo xem ảnh hiện tại có được tham chiếu bởi các ảnh dưới đây hay không có thể được báo hiệu theo cách thông thường hay không trong tiêu đề ảnh hoặc tiêu đề lát.

Ví dụ, trên FIG.24, non_reference_picture_flag được báo hiệu trong tiêu đề ảnh. non_reference_picture_flag bằng 1 chỉ định ảnh được kết hợp với PH không bao giờ được sử dụng làm ảnh tham chiếu. non_reference_picture_flag

bằng 0 chỉ định ảnh được kết hợp với PH có thể hoặc có thể không được sử dụng làm ảnh tham chiếu.

Khi ảnh có thể hoặc có thể không được cắt và được xuất ra để hiển thị hoặc cho các mục đích khác. Còn mà chỉ báo xem ảnh hiện tại có được cắt và xuất ra hay không có thể được báo hiệu theo cách thông thường hay không trong tiêu đề ảnh hoặc tiêu đề lát.

Ví dụ, trên FIG.24, `pic_output_flag` được báo hiệu trong tiêu đề ảnh. `pic_output_flag` bằng 1 chỉ báo ảnh hiện tại có thể được cắt và được xuất ra. `pic_output_flag` bằng 0 chỉ báo ảnh hiện tại có thể không được cắt và được xuất ra.

Khi ảnh hiện tại là ảnh không được tham chiếu mà có thể không được tham chiếu bởi các ảnh kế tiếp theo thứ tự giải mã và giá trị của `non_reference_picture_flag` bằng 1, giá trị của `pic_output_flag` có thể bằng 1, vì ảnh bất kỳ, mà không được tham chiếu bởi các ảnh sau và không được xuất ra, có thể không nằm trong dòng bit video, ở phía bộ giải mã.

Trong cùng phương án hoặc một phương án khác, khi ảnh hiện tại là ảnh không được tham chiếu (tức là `non_reference_picture_flag` bằng 1), `pic_output_flag` không được báo hiệu theo cách thông thường, nhưng được thu nhận bằng 1.

Ở phía bộ mã hóa, ảnh không được tham chiếu mà không được xuất ra có thể không được mã hóa thành dòng bit được mã hóa.

Ở phần tử hệ thống trung gian, ảnh được mã hóa với `non_reference_picture_flag` bằng 1 và `pic_output_flag` bằng 0 có thể bị loại khỏi dòng bit được mã hóa.

Trong khi sáng chế đã mô tả một vài phương án ví dụ, có các phương án khác, các hoán vị, và các thay thế tương đương khác nhau, mà nằm trong phạm vi của sáng chế. Do đó, sẽ được hiểu rằng người có hiểu biết trung bình trong lĩnh vực kỹ thuật sẽ có thể nghĩ ra các hệ thống và phương pháp, mặc dù không được

thể hiện cụ thể hoặc được mô tả ở đây, áp dụng các nguyên tắc của sáng chế và do đó nằm trong bản chất và phạm vi của sáng chế.

Đơn sáng chế này yêu cầu hưởng quyền ưu tiên từ Đơn yêu cầu cấp bằng độc quyền sáng chế tạm thời Mỹ số 63/003,112, được nộp ngày 31 tháng 03 năm 2019, và Đơn yêu cầu cấp bằng độc quyền sáng chế Mỹ số 17/087,865, được nộp ngày 03 tháng 11 năm 2020, nội dung của các đơn này được tích hợp ở đây.

YÊU CẦU BẢO HỘ

1. Phương pháp mã hóa dữ liệu video, thực thi được bởi bộ xử lý, bao gồm:

thu dữ liệu video bao gồm ảnh hiện tại và một hoặc nhiều ảnh khác;

kiểm tra cờ thứ nhất tương ứng với việc ảnh hiện tại có được tham chiếu bởi một hoặc nhiều ảnh khác theo thứ tự giải mã hay không;

kiểm tra cờ thứ hai tương ứng với việc ảnh hiện tại có được xuất ra hay không; và

giải mã dữ liệu video dựa trên các giá trị tương ứng với cờ thứ nhất và cờ thứ hai, trong đó giá trị của cờ thứ hai được thu nhận dựa trên giá trị của cờ thứ nhất.

2. Phương pháp theo điểm 1, trong đó cờ thứ nhất và cờ thứ hai được báo hiệu trong tiêu đề ảnh hoặc tiêu đề lát được kết hợp với dữ liệu video.

3. Phương pháp theo điểm 1, trong đó ảnh hiện tại bị loại dựa trên cờ thứ nhất đang thiết đặt bằng một và cờ thứ hai đang thiết đặt bằng không.

4. Phương pháp theo điểm 1, trong đó cờ thứ nhất tương ứng với việc ảnh hiện tại có được tham chiếu cho bù chuyển động hoặc dự đoán tham số hay không.

5. Phương pháp theo điểm 1, trong đó cờ thứ hai tương ứng với việc ảnh hiện tại có được cắt hoặc được xuất ra để hiển thị hoặc các mục đích khác hay không.

6. Phương pháp theo điểm 1, trong đó giá trị của cờ thứ hai được thu nhận ở phía bộ mã hóa hoặc ở phía bộ giải mã.

7. Hệ thống máy tính mã hóa dữ liệu video, hệ thống máy tính bao gồm:

một hoặc nhiều phương tiện lưu trữ bất biến đọc được bằng máy tính được tạo cấu hình để lưu trữ mã chương trình máy tính; và

một hoặc nhiều bộ xử lý máy tính được cấu hình để truy nhập mã chương trình máy tính và hoạt động như được chỉ dẫn bởi mã chương trình máy tính, mã chương trình máy tính này bao gồm:

mã thu được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính

thu dữ liệu video bao gồm ảnh hiện tại và một hoặc nhiều ảnh khác;

mã kiểm tra thứ nhất được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính kiểm tra cờ thứ nhất tương ứng với việc ảnh hiện tại có được tham chiếu bởi một hoặc nhiều ảnh khác hay không trong thứ tự giải mã;

mã kiểm tra thứ hai được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính kiểm tra cờ thứ hai tương ứng với việc ảnh hiện tại có được xuất ra hay không; và

mã giải mã được tạo cấu hình để khiến một hoặc nhiều bộ xử lý giải mã dữ liệu video dựa trên các giá trị tương ứng với cờ thứ nhất và cờ thứ hai, trong đó giá trị của cờ thứ hai được thu nhận dựa trên giá trị của cờ thứ nhất.

8. Hệ thống máy tính theo điểm 7, trong đó cờ thứ nhất và cờ thứ hai được báo hiệu trong tiêu đề ảnh hoặc tiêu đề lát được kết hợp với dữ liệu video.

9. Hệ thống máy tính theo điểm 7, trong đó ảnh hiện tại bị loại dựa trên cờ thứ nhất đang thiết đặt bằng một và cờ thứ hai đang thiết đặt bằng không.

10. Hệ thống máy tính theo điểm 7, trong đó cờ thứ nhất tương ứng với việc ảnh hiện tại có được tham chiếu cho bù chuyển động hoặc dự đoán tham số hay không.

11. Hệ thống máy tính theo điểm 7, trong đó cờ thứ hai tương ứng với việc ảnh hiện tại có được cắt và được xuất ra để hiển thị hay các mục đích khác hay không.

12. Hệ thống máy tính theo điểm 7, trong đó giá trị của cờ thứ hai được thu nhận ở phía bộ mã hóa hoặc ở phía bộ giải mã.

13. Phương tiện bất biến đọc được bằng máy tính lưu trữ trên đó chương trình máy tính để mã hóa dữ liệu video, chương trình máy tính được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính:

thu dữ liệu video bao gồm ảnh hiện tại và một hoặc nhiều ảnh khác;

kiểm tra cờ thứ nhất tương ứng với việc ảnh hiện tại có được tham chiếu bởi một hoặc nhiều ảnh khác theo thứ tự giải mã hay không;

kiểm tra cờ thứ hai tương ứng với việc ảnh hiện tại có được xuất ra hay

không; và

giải mã dữ liệu video dựa trên các giá trị tương ứng với cờ thứ nhất và cờ thứ hai, trong đó giá trị của cờ thứ hai được thu nhận dựa trên giá trị của cờ thứ nhất.

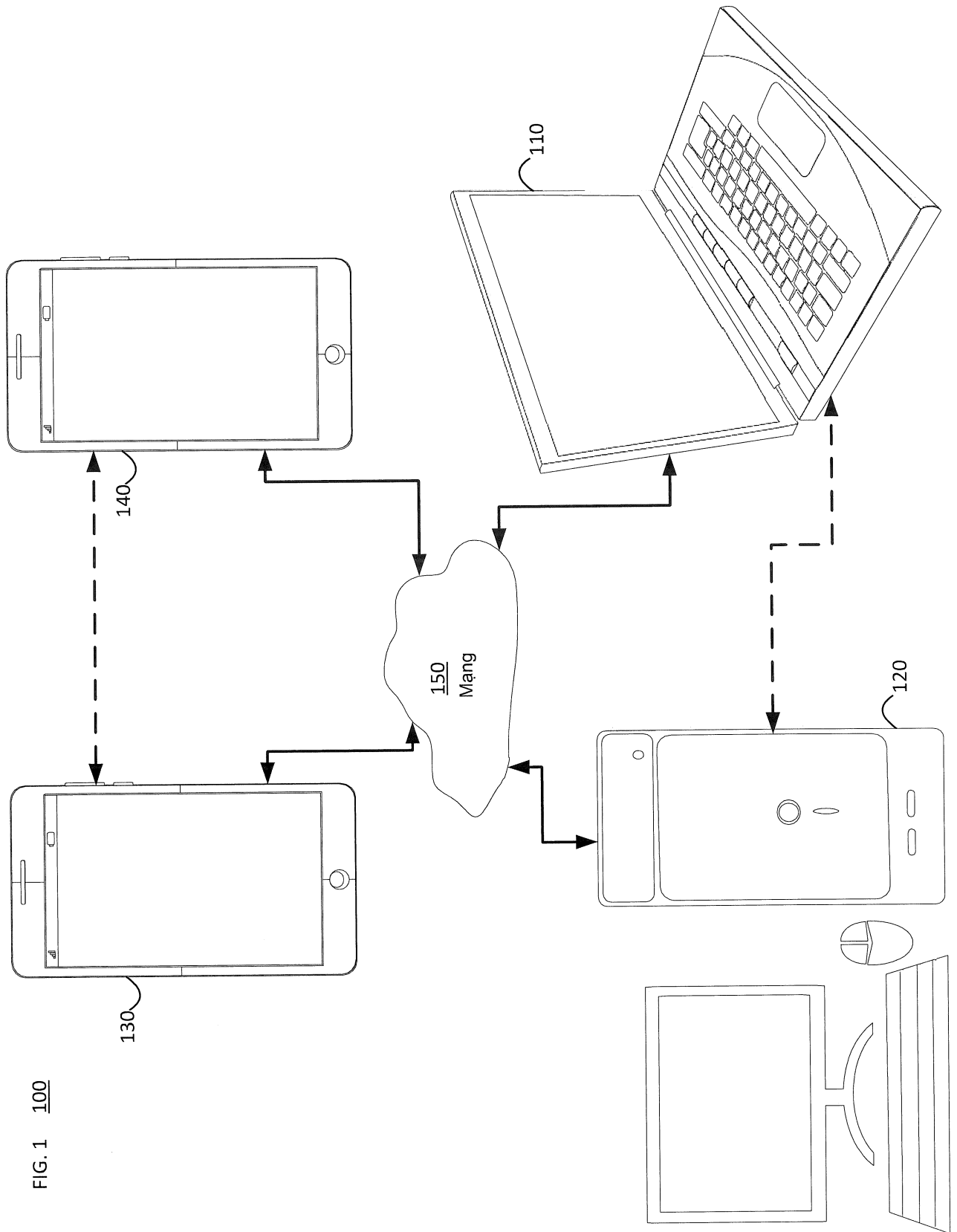
14. Phương tiện đọc được bằng máy tính theo điểm 13, trong đó cờ thứ nhất và cờ thứ hai được báo hiệu trong tiêu đề ảnh hoặc tiêu đề lát được kết hợp với dữ liệu video.

15. Phương tiện đọc được bằng máy tính theo điểm 13, trong đó ảnh hiện tại bị loại dựa trên cờ thứ nhất đang thiết đặt bằng một và cờ thứ hai đang thiết đặt bằng không.

16. Phương tiện đọc được bằng máy tính theo điểm 13, trong đó cờ thứ nhất tương ứng với việc ảnh hiện tại có được tham chiếu cho bù chuyển động hoặc dự đoán tham số hay không.

17. Phương tiện đọc được bằng máy tính theo điểm 13, trong đó cờ thứ hai tương ứng với việc ảnh hiện tại có được cắt và được xuất ra để hiển thị hay các mục đích khác hay không.

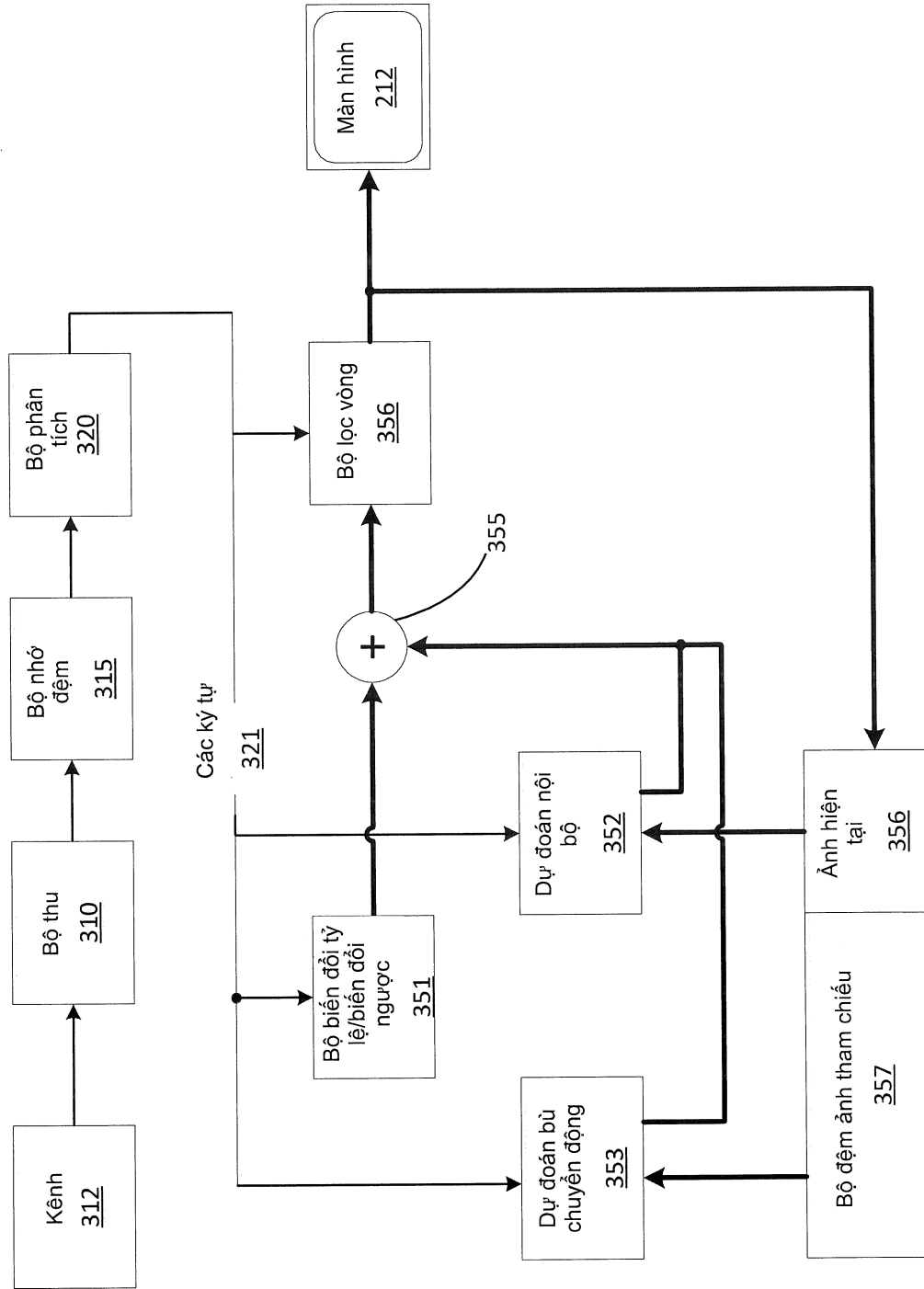
1/24



3/24

Bố giải mã 210

FIG. 3



4/24

FIG. 4 Bộ mã hóa 203

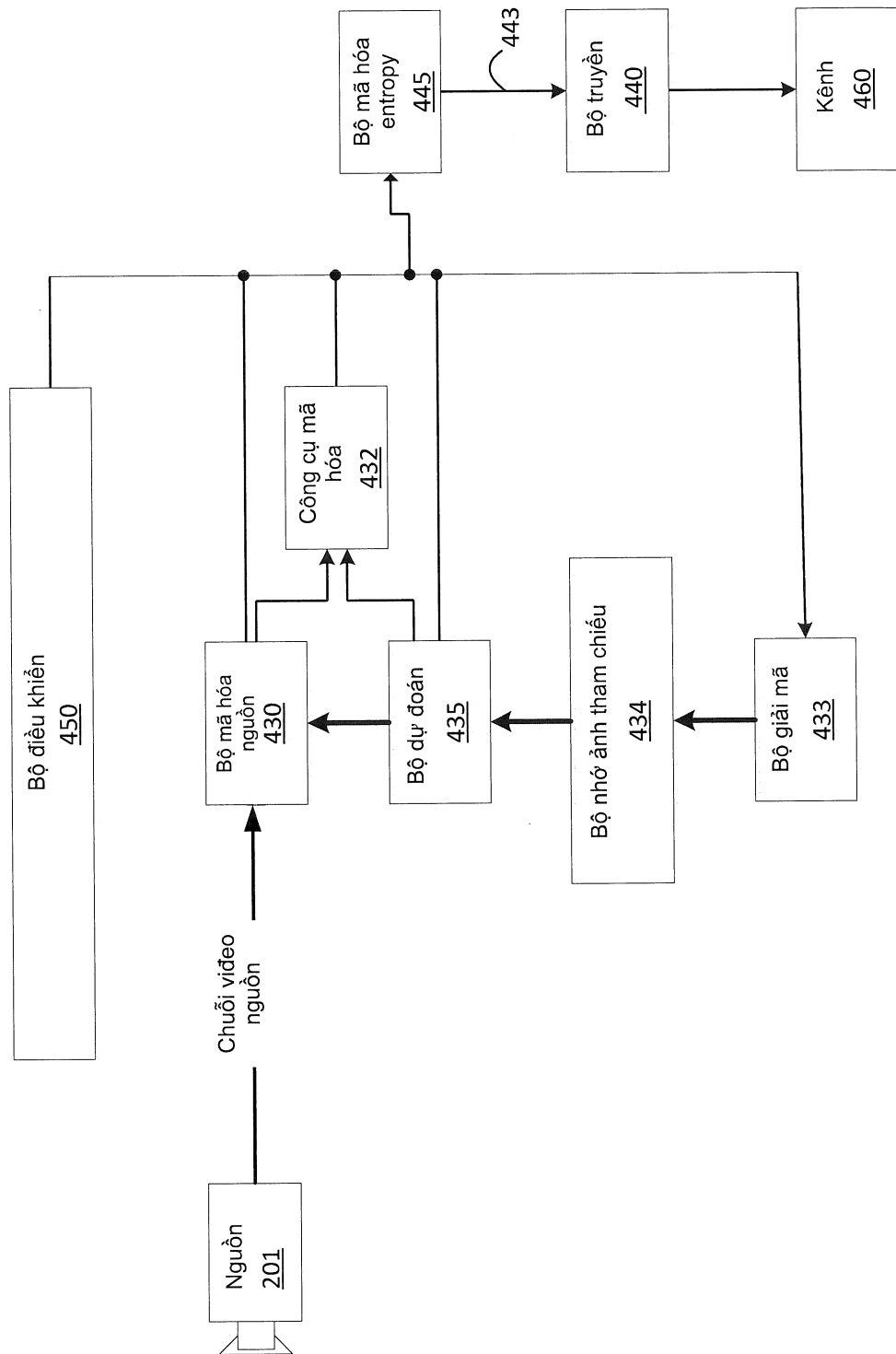


FIG. 5

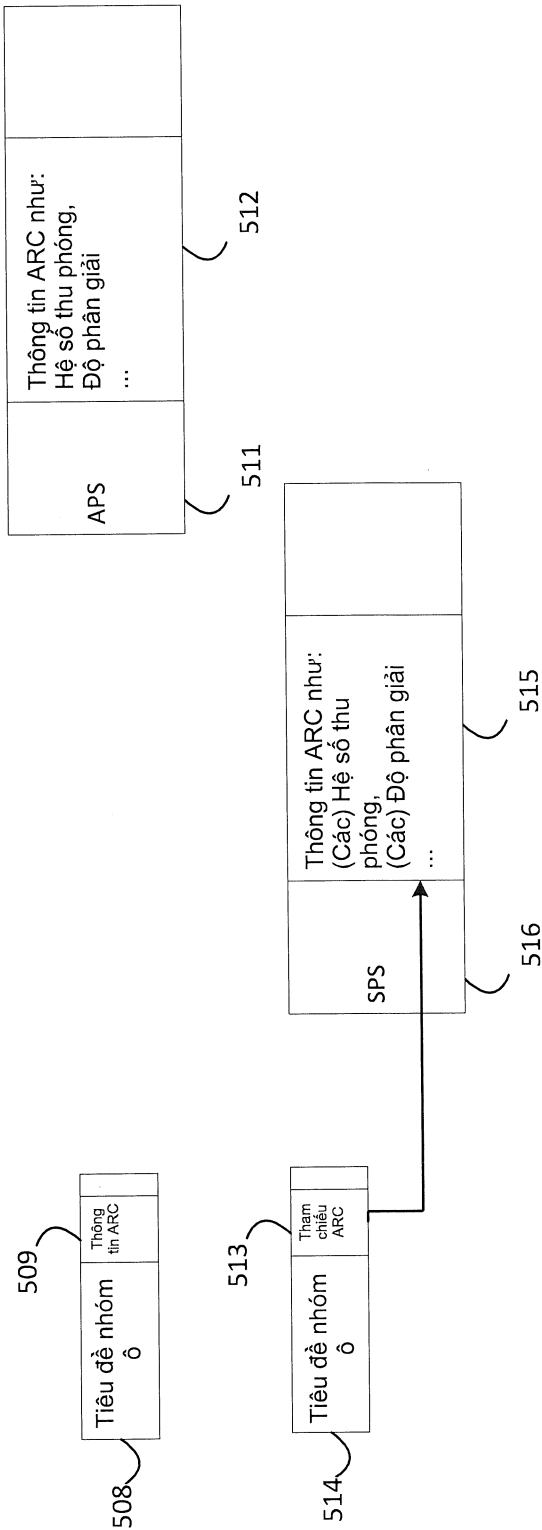
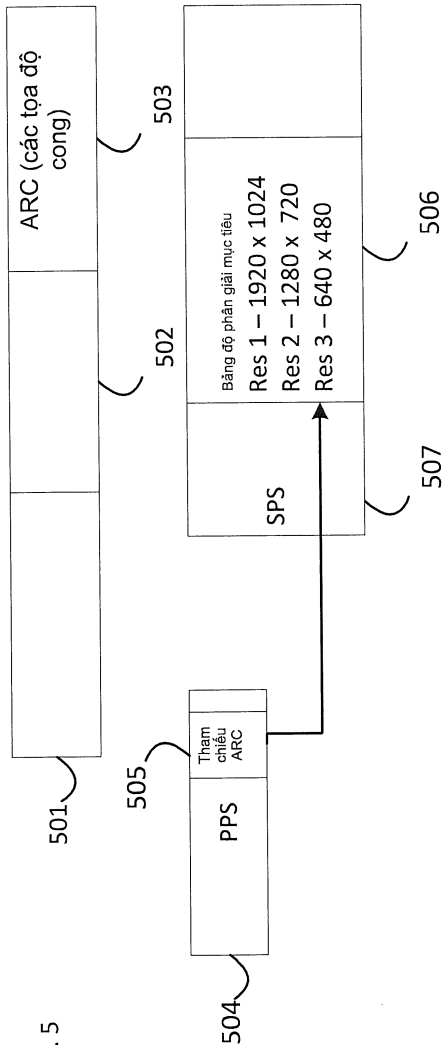
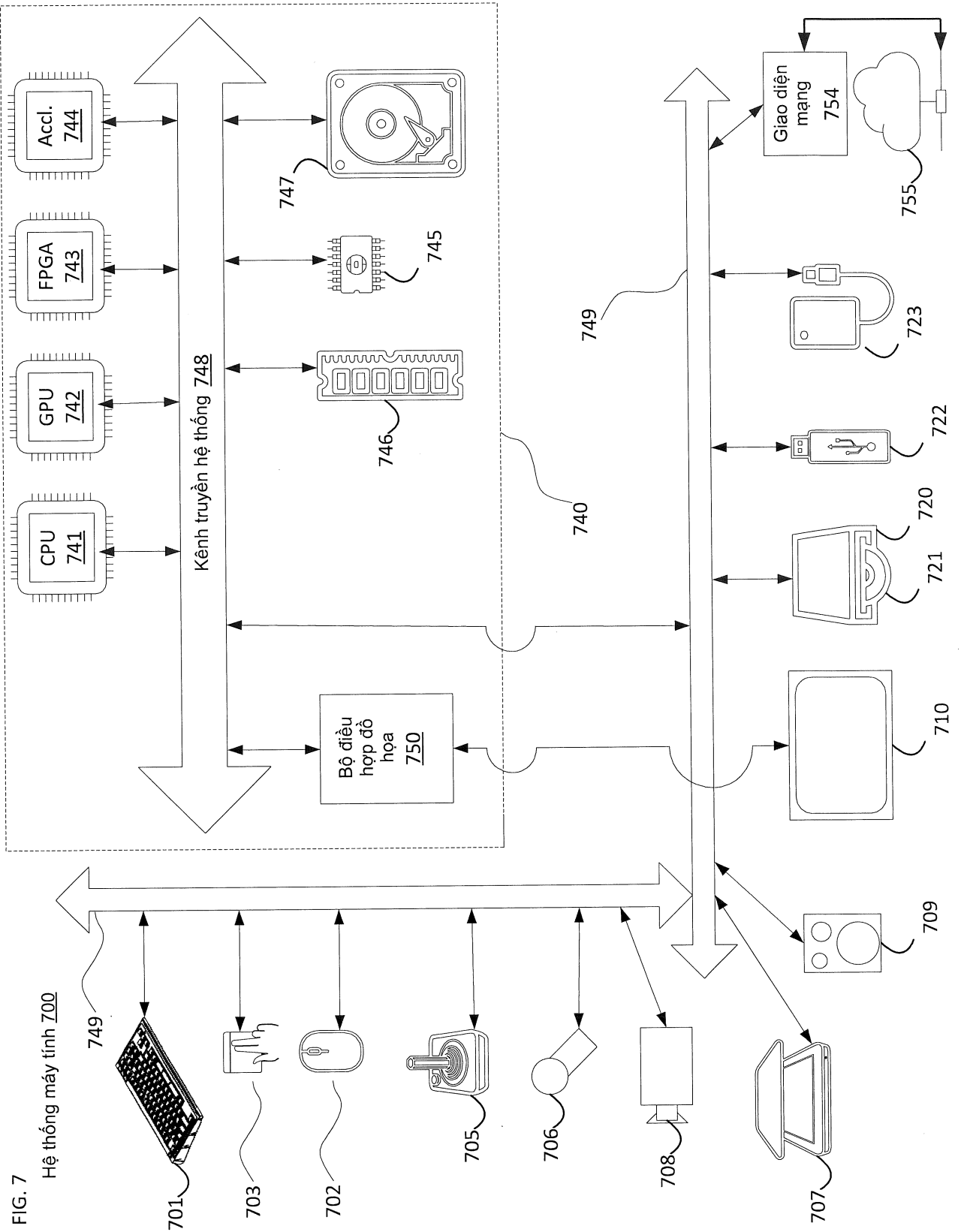


Fig. 6

601	tile_group_header() {	
	...	
603	if(adaptive_pic_resolution_change_flag) {	
602	dec_pic_size_idx	u(1)
	}	
	...	
	}	

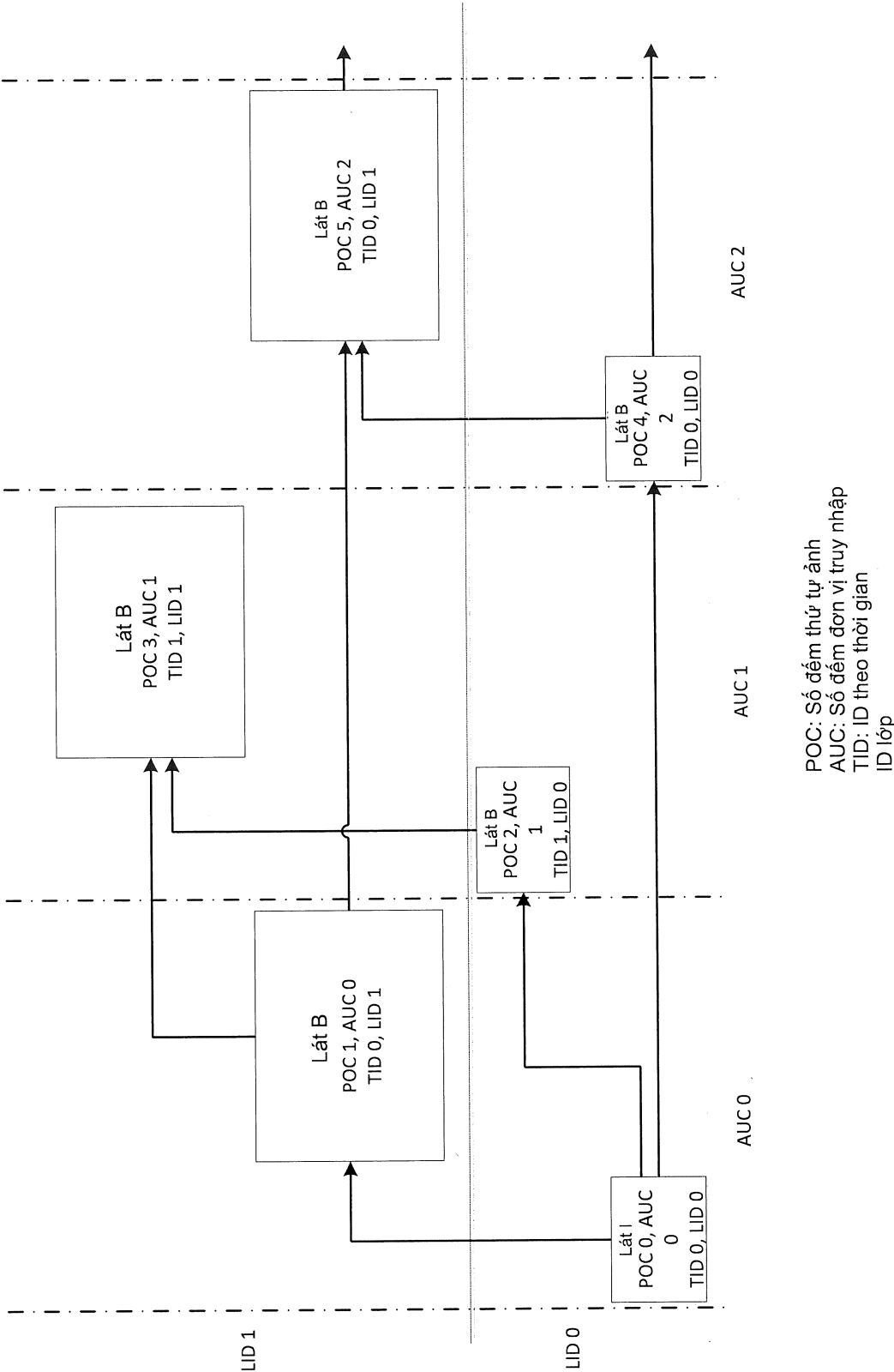
610	seq_parameter_set_rbsp() {	Nhãn mô tả
	...	
611	adaptive_pic_resolution_change_flag	u(1)
612	if(adaptive_pic_resolution_change_flag) {	
	output_pic_width_in_luma_samples	ue(v)
613	output_pic_height_in_luma_samples	ue(v)
	reference_pic_size_present_flag	u(1)
614	if(reference_pic_size_present_flag)	
	{	
	reference_pic_width_in_luma_samples	ue(v)
	reference_pic_height_in_luma_samples	ue(v)
	}	
	num_dec_pic_size_in_luma_samples_minus1	ue(v)
	for(i = 0; i <= num_dec_pic_size_in_luma_samples_minus1; i++) {	
	dec_pic_width_in_luma_samples[i]	ue(v)
	dec_pic_height_in_luma_samples[i]	ue(v)
	}	
	}	
	...	
	}	

7/24



8/24

Fig. 8



9/24

Fig. 9

video_parameter_set_rbsp() { vps_video_parameter_set_id; vps_max_layers_minus1; for(i = 0; i <= vps_max_layers_minus1; i++) { vps_included_layer_id[i]; vps_reserved_zero_bit; } vps_constraint_info_present_flag; vps_constant_poc_cycle_per_au; if(vps_constant_poc_cycle_per_au) vps_poc_cycle_au; ... }	Nhân mô tả u(4) u(8) + u(1) u(1) + u(1) u(1) + u(8) + + +
slice_header() { slice_pic_parameter_set_id; if(rect_slice_flag NumBricksInPic > 1) { slice_address; if(rect_slice_flag && !single_brick_per_slice_flag) { num_bricks_in_slice_minus1; slice_type; if(NalUnitType == GRA_NUT) { recovery_poc_cnt; slice_pic_order_cnt_lsb; ... } if(vps_constant_poc_cycle_per_au) { slice_poc_cycle_au; } ... } }	Nhân mô tả ue(v) + u(v) + ue(v) ue(v) + se(v) u(v) + + + u(8) + + +

10/24

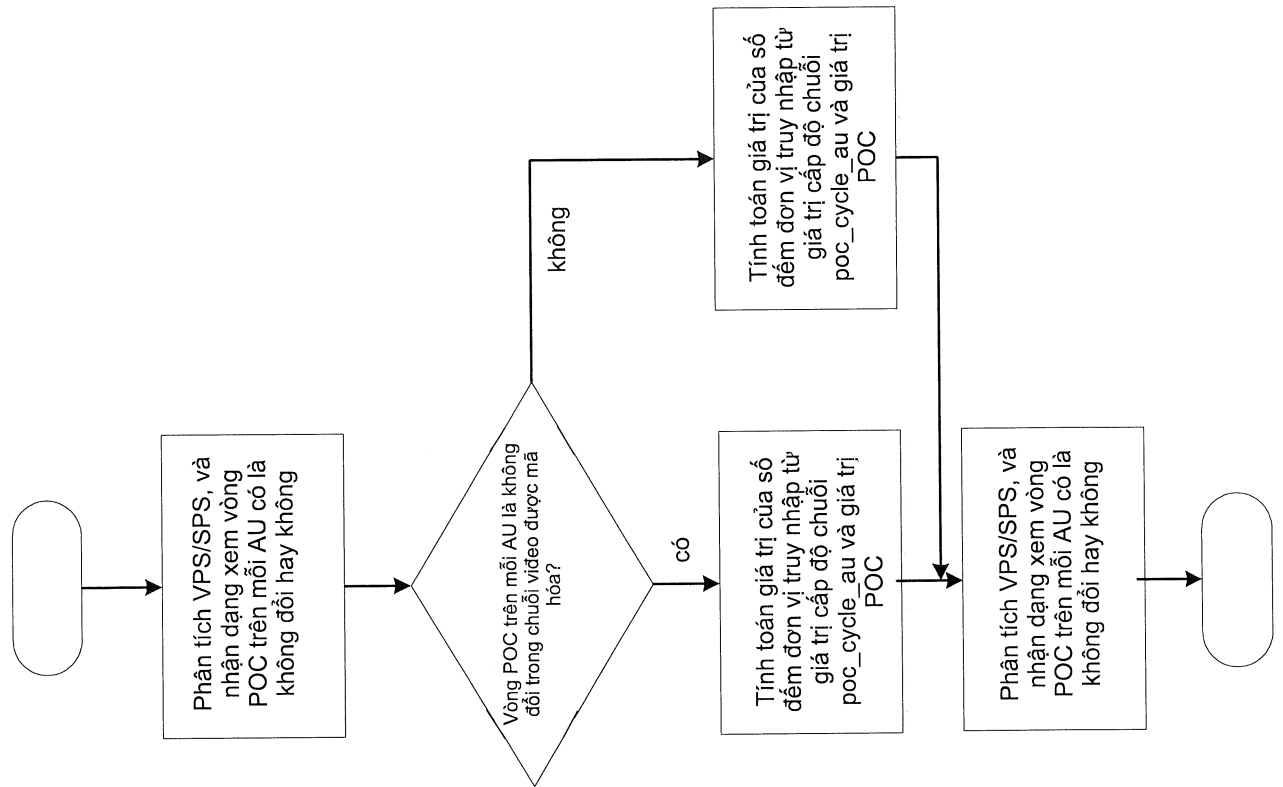
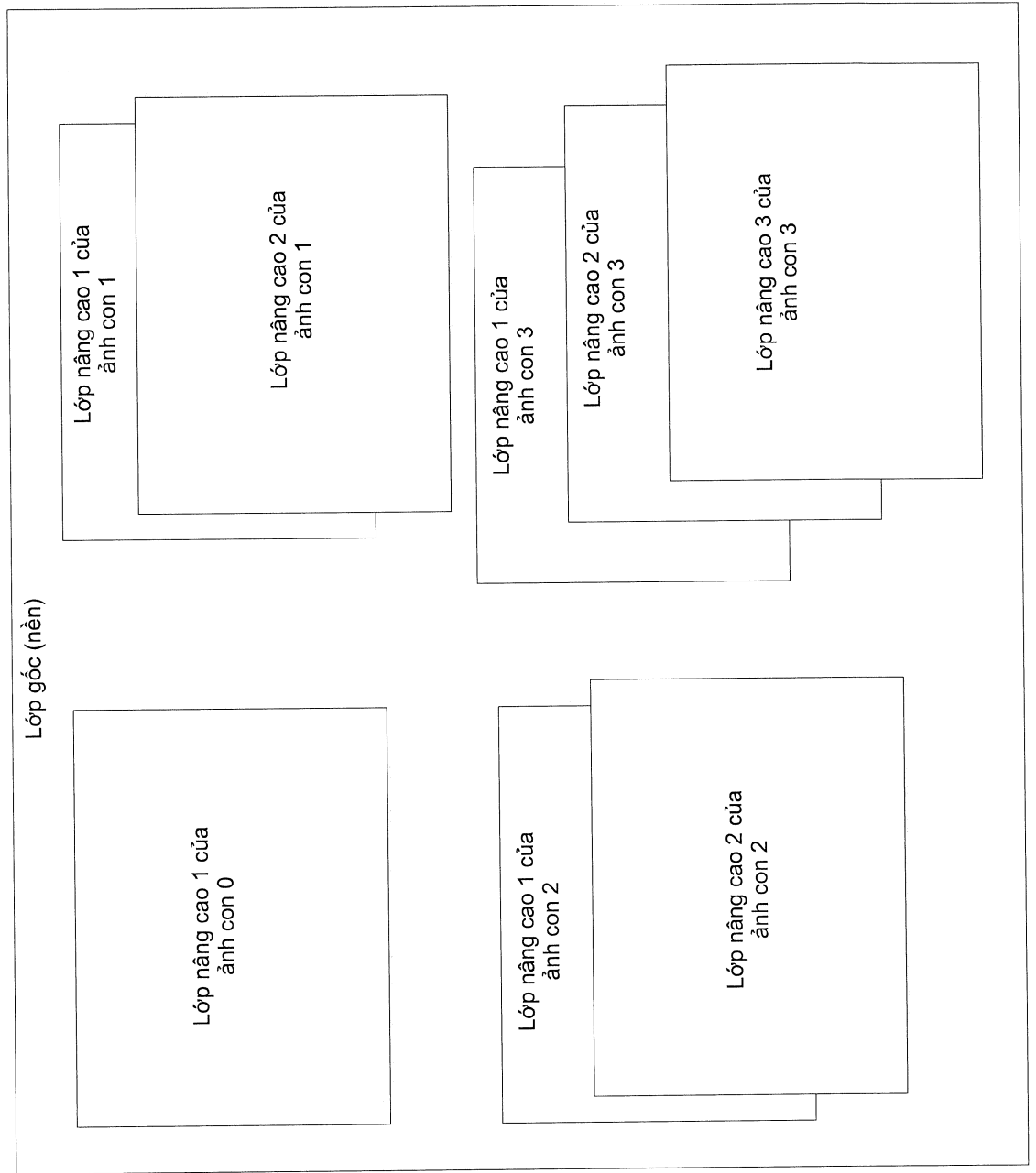


Fig. 10

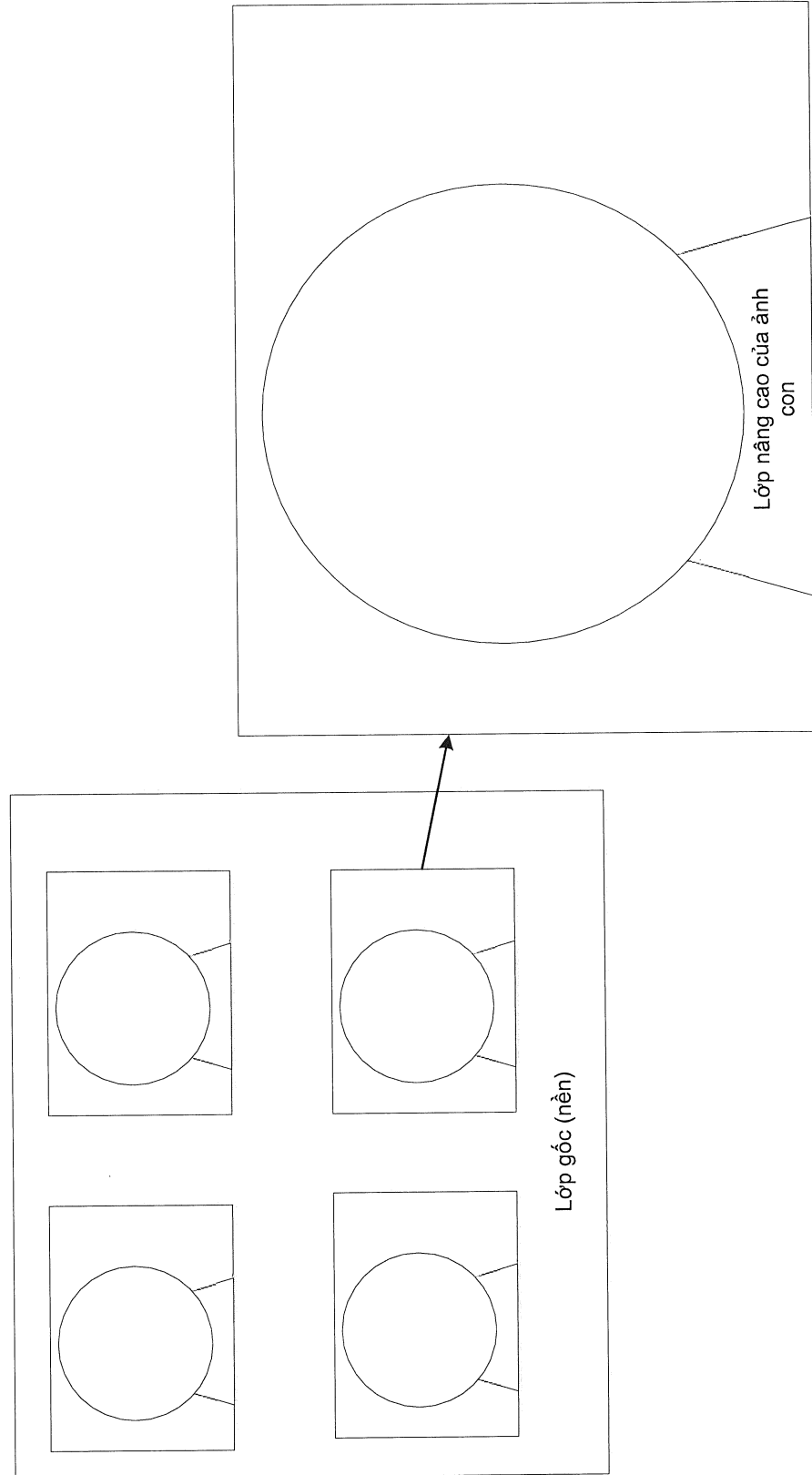
11/24

FIG. 11



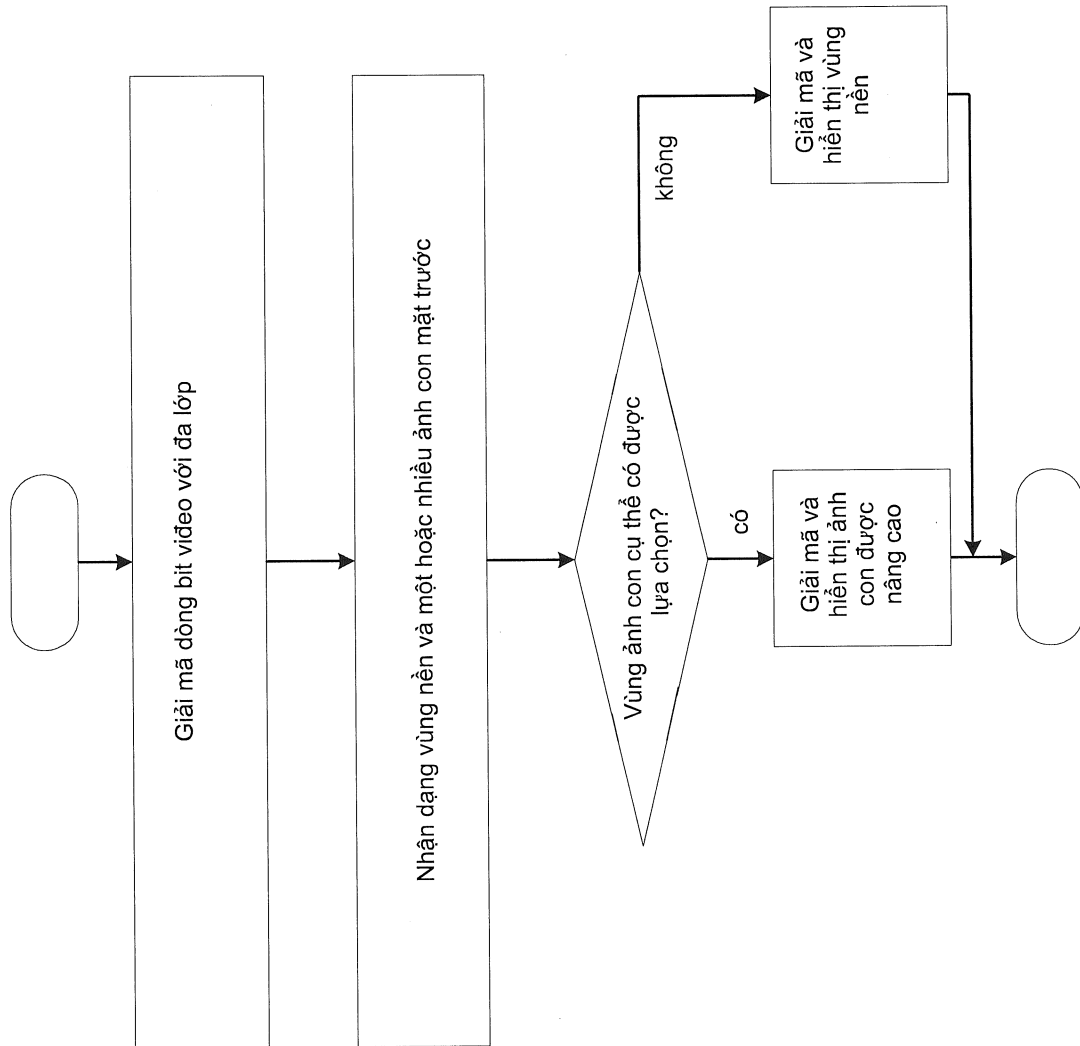
12/24

FIG. 12



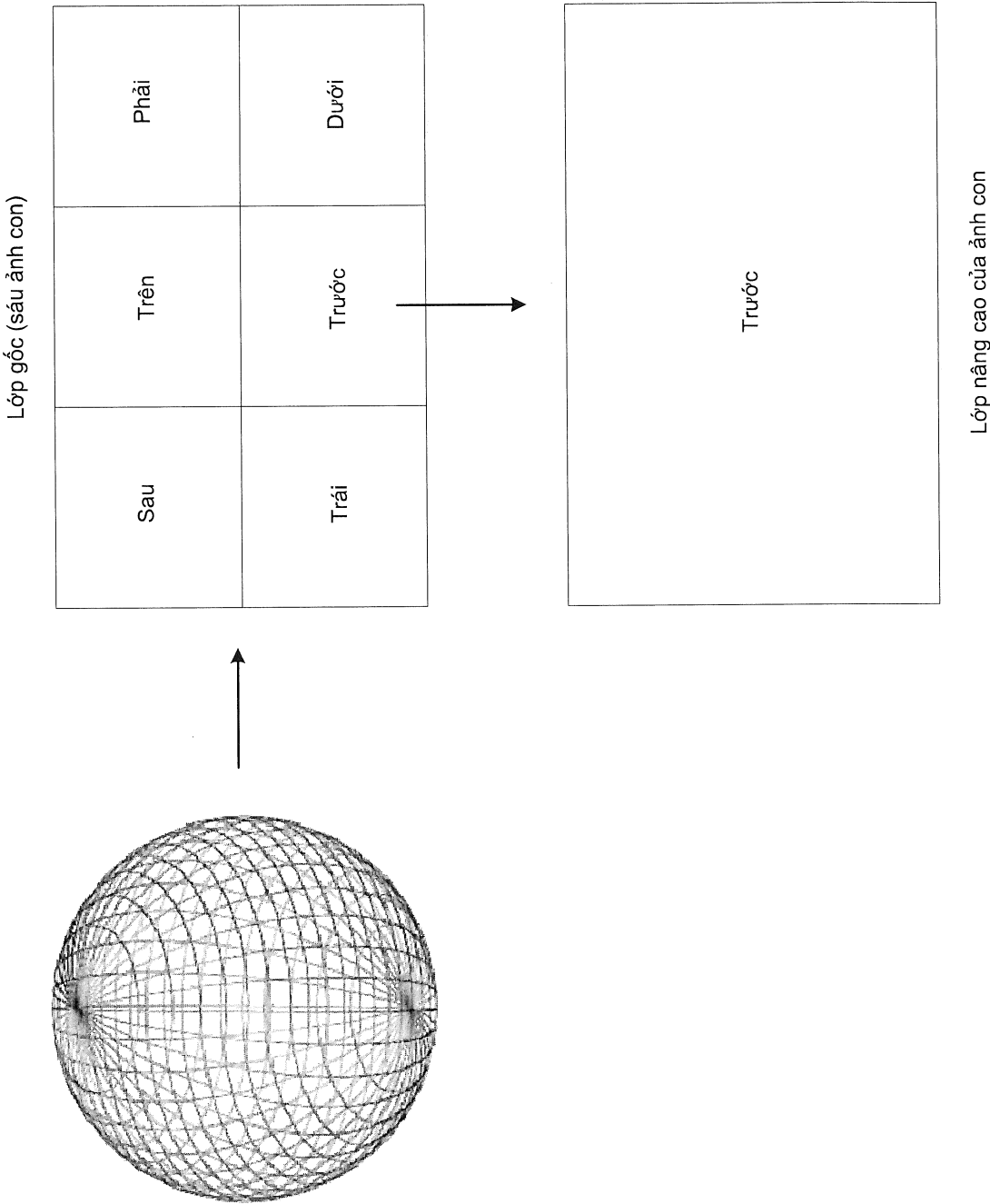
13/24

Fig. 13



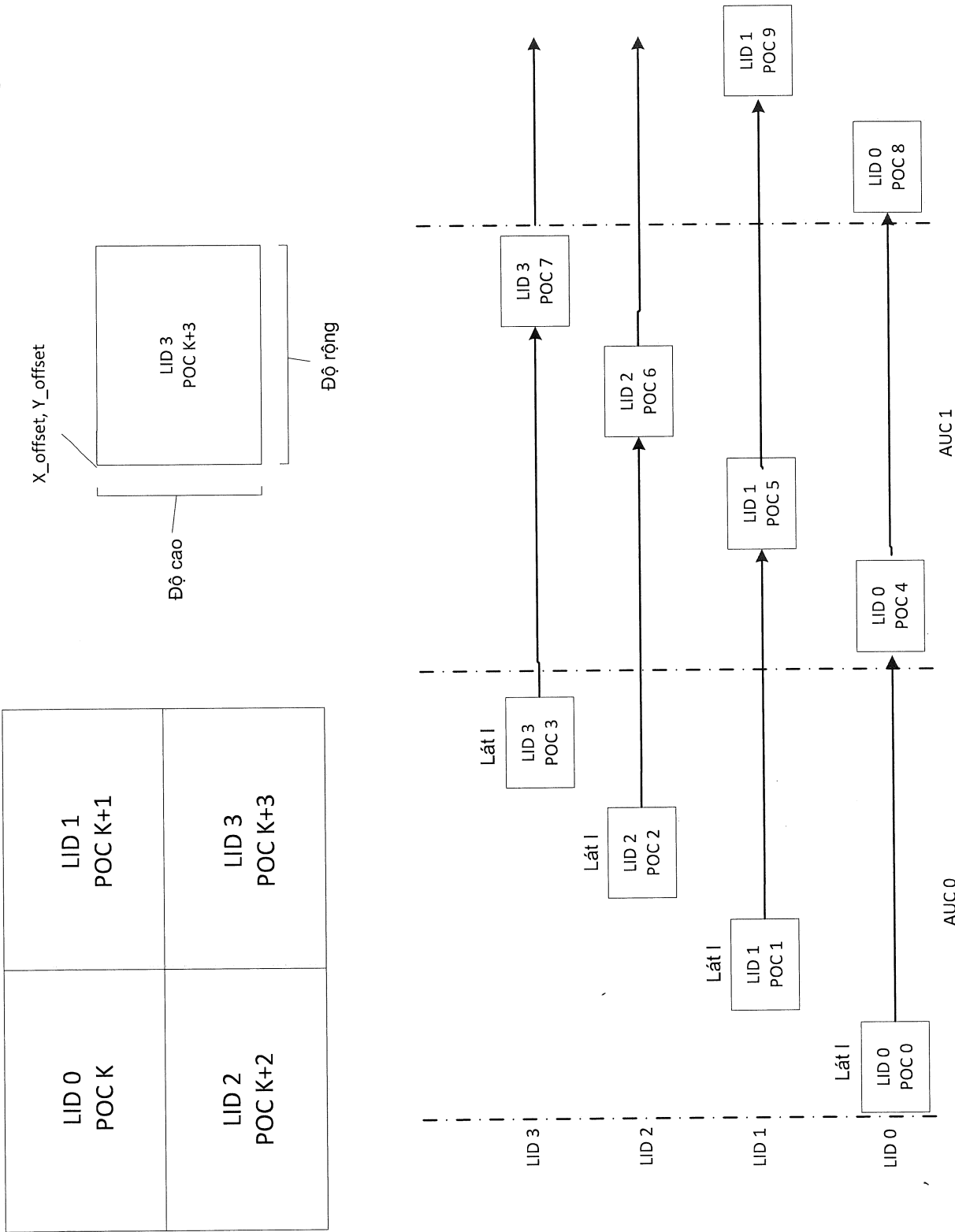
14/24

Fig. 14



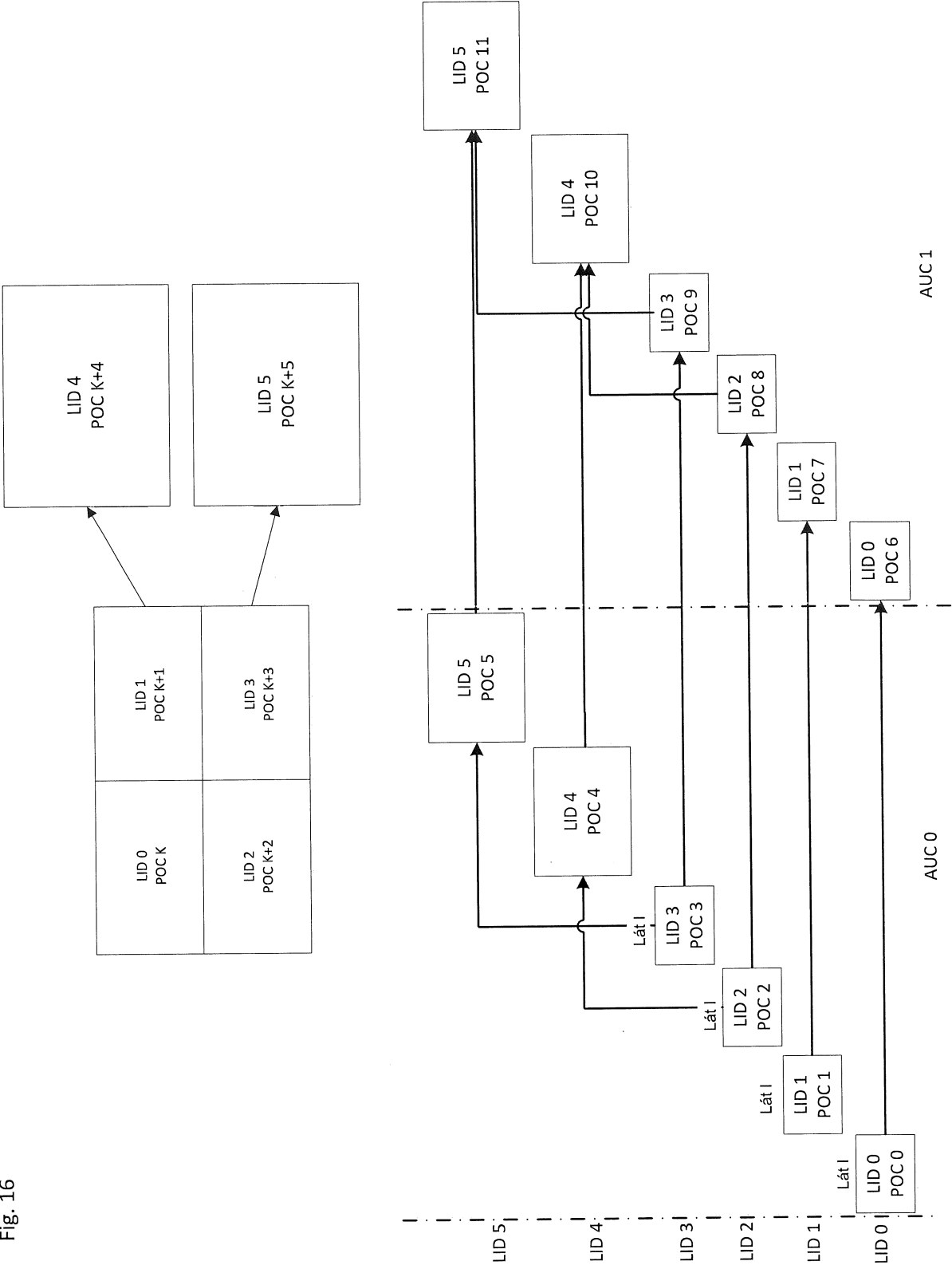
15/24

Fig. 15



16/24

Fig. 16



17/24

Fig. 17

<code>video_parameter_set_rbsp() { vps_video_parameter_set_id; vps_max_layers_minus1; for(i = 0; i <= vps_max_layers_minus1; i++) { vps_included_layer_id[i]; vps_reserved_zero_bit; } ... vps_sub_picture_dividing_flag; if(vps_sub_picture_dividing_flag) { vps_full_pic_width_in_luma_samples; vps_full_pic_height_in_luma_samples; } ... }</code>	Nhãn mô tả $\underline{u}(4)^+$ $u(8)^+$ $^+$ $u(7)^+$ $u(1)^+$ $^+$ $^+$ $u(1)^+$ $^+$ $ue(v)^+$ $ue(v)^+$ $^+$ $^+$ $^+$
<code>seq_parameter_set_rbsp() { sps_decoding_parameter_set_id; sps_video_parameter_set_id; sps_max_sub_layers_minus1; ... pic_width_in_luma_samples; pic_height_in_luma_samples; if(vps_sub_picture_dividing_flag) { pic_offset_x; pic_offset_y; } ... }</code>	Nhãn mô tả $\underline{u}(4)^+$ $\underline{u}(4)^+$ $\underline{u}(3)^+$ $^+$ $ue(v)^+$ $ue(v)^+$ $^+$ $ue(v)^+$ $ue(v)^+$ $^+$ $^+$ $^+$

18/24

Fig. 18

sub_region_partitioning_info(payloadSize) { ⁺	Nhân mô tả
num_sub_region ⁺	us(v) ⁺
num_layers ⁺	us(v) ⁺
for(j = 0; j <= num_layers; i++) ⁺	⁺
layer_id [i] ⁺	us(v) ⁺
for(j = 1; j < num_layers; j++) ⁺	⁺
for(i = 0; j < i; i++) ⁺	⁺
direct_dependency_flag [i] [j] ⁺	<u>u</u> (1) ⁺
for(j = 0; j < num_sub_region; i++) { ⁺	⁺
num_layers_for_region [i] ⁺	us(v) ⁺
for(i = 0; j < num_layers_for_region [i]; i++) ⁺	⁺
sub_region_layer_id [i] [j] ⁺	us(v) ⁺
sub_region_offset_x [i] ⁺	us(v) ⁺
sub_region_offset_y [i] ⁺	us(v) ⁺
sub_region_width [i] ⁺	us(v) ⁺
sub_region_height [i] ⁺	us(v) ⁺
} ⁺	⁺
... ⁺	⁺
} ⁺	⁺

Fig. 19

video parameter set rbsp() {	Nhân mô tả
...	
vps max layers minus1	u(6)
num output layer sets	ue(v)
num profile tier level	ue(v)
for(i = 0; i < num profile tier level; i++)	
profile tier level(vps max sub layers minus1)	
for(i = 0; i < num output layer sets; i++)	
for(j = 0; j < NumLayersInIdList[i]; j++) {	
output layer flag[i][j]	u(1)
profile tier level idx[i][j]	u(v)
}	
...	
}	

Fig. 20

	Nhân mô tả
video parameter set rbsp() {	
...	
vps max layers minus1	u(6)
if(vps max layers minus1 > 0) {	
num output layer sets	ue(v)
num profile tier level	ue(v)
}	
for(i = 0; i < num profile tier level; i++)	
profile tier level(vps max sub layers minus1)	
for(i = 0; i < num output layer sets; i++) {	
vps output layers mode[i]	u(2)
vps ptl signal flag[i]	u(1)
for(j = 0; j < NumLayersInIdList[I]; j++) {	
if(vps output layers mode[i] == 2)	
output layer flag[I][i]	u(1)
if(vps ptl signal flag[i])	
profile tier level idx[I][i]	u(v)
}	
}	
....	
}	

21/24

Fig. 21

video parameter set rbsp() {	Nhân mô tả
...	
vps max layers minus1	u(6)
if(vps max layers minus1 > 0) {	
num output layer sets	ue(v)
num profile tier level	ue(v)
}	
max subpics minus1	u(8)
for(i = 0; i < max subpics minus1; i++) {	
sub pic id[i]	u(8)
}	
for(i = 0; i < num profile tier level; i++)	
profile tier level(vps max sub layers minus1)	
for(i = 0; i < num output layer sets; i++) {	
vps output layers mode[i]	u(2)
vps ptl signal flag[i]	u(1)
for(j = 0; j < NumLayersInIdList[i]; j++) {	
num output subpic layer[i][j]	ue(v)
for(k = 0; k < num output subpic layer[i][j]; k++)	
sub pic id layer[i][j][k]	u(8)
if(vps output layers mode[i] == 2)	
output layer flag[i][j]	u(1)
if(vps ptl signal flag[i])	
profile tier level idx[i][j]	u(v)
}	
}	
}	
...	
}	

22/24

Fig. 22

video_parameter_set_rbsp() {	Nhân mô tả
vps_video_parameter_set_id	u(4)
vps_max_layers_minus1	u(6)
vps_max_sublayers_minus1	u(3)
if vps_max_layers_minus1 > 0 && vps_max_sublayers_minus1 > 0)	
vps_all_layers_same_num_sublayers_flag	u(1)
if vps_max_layers_minus1 > 0)	
vps_all_independent_layers_flag	u(1)
for(i = 0; i <= vps_max_layers_minus1; i++) {	
vps_layer_id[i]	u(6)
if i > 0 && !vps_all_independent_layers_flag) {	
vps_independent_layer_flag[i]	u(1)
if(!vps_independent_layer_flag[i]) {	
for(i = 0; i <= i; i++)	
vps_direct_ref_layer_flag[i][j]	u(1)
max_tid_ref_present_flag[i]	u(1)
if max_tid_ref_present_flag[i])	
max_tid_ref_pics_plus1[i]	u(3)
}	
}	
if vps_max_layers_minus1 > 0) {	
if vps_all_independent_layers_flag)	
each_layer_is_an_ols_flag	u(1)
if !each_layer_is_an_ols_flag) {	
if !vps_all_independent_layers_flag)	
ols_mode_idc	u(2)
if ols_mode_idc == 2) {	
num_output_layer_sets_minus1	u(8)
for(i = 1; i <= num_output_layer_sets_minus1; i++)	
for(j = 0; j <= vps_max_layers_minus1; j++)	
ols_output_layer_flag[i][j]	u(1)
}	
}	
}	
}	
}	

23/24

Fig. 23

video parameter set rbsp() {	Nhân mô tả
vps video parameter set id	u(4)
vps max layers minus1	u(6)
vps max sublayers minus1	u(3)
if(vps max layers minus1 > 0 && vps max sublayers minus1 > 0)	
vps all layers same num sublayers flag	u(1)
if(vps max layers minus1 > 0)	
vps all independent layers flag	u(1)
for(i = 0; i <= vps max layers minus1; i++) {	
vps layer id[i]	u(6)
if(i > 0 && !vps all independent layers flag) {	
vps independent layer flag[i]	u(1)
if(!vps independent layer flag[i]) {	
for(j = 0; j < i; j++)	
vps direct ref layer flag[i][j]	u(1)
max tid ref present flag[i]	u(1)
if(max tid ref present flag[i])	
max tid il ref pics plus1[i]	u(3)
}	
}	
}	
if(vps max layers minus1 > 0) {	
if(vps all independent layers flag)	
each layer is an ols flag	u(1)
if(!each layer is an ols flag) {	
if(!vps all independent layers flag)	
ols mode idc	u(2)
if(ols mode idc == 2 && vps max layers minus1 > 1) {	
num output layer sets minus1	u(8)
for(i = 1; i <= num output layer sets minus1; i++)	
for(j = 0; j <= vps max layers minus1; j++)	
ols output layer flag[i][j]	u(1)
}	
}	
}	
...	
}	

Fig. 24

picture header structure() {	Nhãn mô tả
gdr or irap pic flag	u(1)
if(gdr or irap pic flag)	
gdr pic flag	u(1)
ph inter slice allowed flag	u(1)
if(ph inter slice allowed flag)	
ph intra slice allowed flag	u(1)
non reference picture flag	u(1)
...	
if(output flag present flag && !non reference picture flag)	
pic output flag	u(1)
if(rpl info in ph flag)	
ref pic lists()	
...	
}	