



(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)



1-0048043

(51)^{2020.01} G06K 9/36; H04B 1/66

(13) B

(21) 1-2021-07168

(22) 05/10/2020

(86) PCT/US2020/054244 05/10/2020

(87) WO2021/133452 A1 01/07/2021

(30) 62/954,096 27/12/2019 US; 17/038,541 30/09/2020 US

(45) 25/06/2025 447

(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)

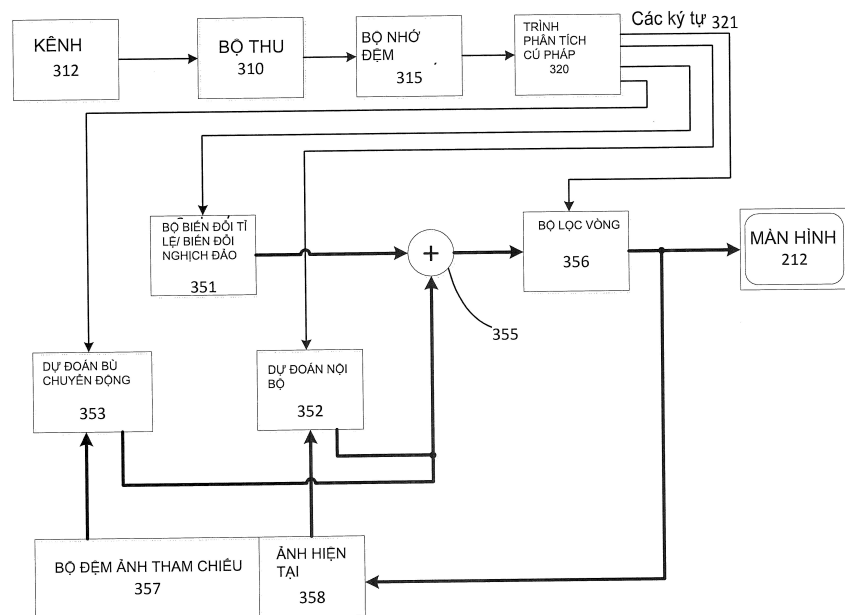
(54) PHƯƠNG PHÁP VÀ THIẾT BỊ GIẢI MÃ DÒNG BIT VIDEO ĐƯỢC MÃ HÓA
VÀ PHƯƠNG TIỆN BẮT BIẾN ĐỌC ĐƯỢC BẰNG MÁY TÍNH

(21) 1-2021-07168

(57) Sáng chế đề cập đến phương pháp giải mã dòng bit video được mã hóa bằng cách sử dụng ít nhất một bộ xử lý, bao gồm: thu nhận từ dòng bit video được mã hóa chuỗi video được mã hóa bao gồm đơn vị ảnh tương ứng với ảnh được mã hóa; thu nhận đơn vị lớp trừu tượng mạng (NAL) tiêu đề ảnh (PH) được bao gồm trong đơn vị ảnh; thu nhận ít nhất một đơn vị lớp trừu tượng mạng (NAL) lớp mã hóa video (VCL) được bao gồm trong đơn vị ảnh; giải mã ảnh được mã hóa dựa trên đơn vị lớp trừu tượng mạng PH, ít nhất một đơn vị lớp trừu tượng mạng VCL, và tập tham số thích ứng (APS) bao gồm trong đơn vị lớp trừu tượng mạng APS thu nhận từ chuỗi video được mã hóa; và xuất ra ảnh đã được giải mã, trong đó đơn vị lớp trừu tượng mạng APS khả dụng cho ít nhất một bộ xử lý trước ít nhất một đơn vị lớp trừu tượng mạng VCL.

FIG. 3

Bộ giải mã 210



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc mã hóa và giải mã video, và cụ thể hơn, liên quan đến tham chiếu tập tham số thích ứng và các ràng buộc trong dòng video được mã hóa.

Tình trạng kỹ thuật của sáng chế

Mã hóa và giải mã video bằng cách sử dụng dự đoán ảnh liên đới có bù chuyển động đã được biết đến. Video kỹ thuật số không nén có thể bao gồm loạt ảnh, mỗi ảnh có kích thước không gian, ví dụ, mẫu độ chói 1920×1080 và mẫu sắc độ liên quan. Loạt ảnh có thể có tốc độ ảnh cố định hoặc thay đổi (thường được biết đến là tốc độ khung hình), của, ví dụ 60 ảnh/giây hoặc 60 Hz. Video không nén có yêu cầu về tốc độ bit đáng kể. Ví dụ, video 1080p60 4: 2: 0 ở 8 bit trên mỗi mẫu (độ phân giải mẫu độ chói 1920×1080 ở tốc độ khung hình 60 Hz) yêu cầu băng thông gần 1,5 Gbit/s. Một giờ video như vậy cần hơn 600 GByte dung lượng lưu trữ.

Một mục đích của mã hóa và giải mã video có thể là giảm phần dư trong tín hiệu video đầu vào, thông qua nén. Nén có thể giúp giảm băng thông đã nói ở trên hoặc yêu cầu không gian lưu trữ, trong một số trường hợp bởi hai bậc độ lớn hoặc nhiều hơn. Cả nén không tổn hao và nén có tổn hao, cũng như kết hợp chúng có thể được sử dụng. Nén không tổn hao đề cập đến các kỹ thuật trong đó bản sao chính xác của tín hiệu gốc có thể được tái cấu trúc từ tín hiệu gốc đã được nén. Khi sử dụng phương pháp nén có tổn hao, tín hiệu được tái cấu trúc có thể không giống với tín hiệu ban đầu, nhưng biến dạng giữa tín hiệu gốc và tín hiệu được tái cấu trúc đủ nhỏ để làm cho tín hiệu được tái cấu trúc hữu ích cho các ứng dụng dự kiến. Trong trường hợp video, nén có tổn hao được sử dụng rộng rãi. Số lượng biến dạng chấp nhận được tùy thuộc vào ứng dụng; ví dụ, người dùng các ứng dụng phát trực tuyến của người tiêu dùng nhất định có thể chịu được biến dạng cao hơn so với người dùng ứng dụng đóng góp truyền hình. Tỷ lệ nén có thể đạt

được có thể phản ánh rằng: biến dạng cho phép/chịu được cao hơn có thể mang lại tỷ lệ nén cao hơn.

Bộ mã hóa và giải mã video có thể sử dụng các kỹ thuật từ một số danh mục rộng, bao gồm, ví dụ, bù chuyển động, biến đổi, lượng tử hóa, và mã hóa entropy, một số trong số đó sẽ được giới thiệu bên dưới.

Về phương diện lịch sử, bộ mã hóa và giải mã video có xu hướng hoạt động trên kích thước ảnh nhất định mà, trong hầu hết các trường hợp, được xác định và giữ không đổi đối với chuỗi video được mã hóa (CVS), Nhóm ảnh (GOP), hoặc khung thời gian đa ảnh tương tự. Ví dụ, trong MPEG-2, các thiết kế hệ thống được biết là thay đổi độ phân giải chiều ngang (và, do đó, kích thước ảnh) phụ thuộc vào các hệ số như hoạt động của cảnh, nhưng chỉ ở ảnh I, do đó thường đối với GOP. Lấy mẫu lại các ảnh tham chiếu để sử dụng các độ phân giải khác nhau trong CVS được biết, ví dụ, từ ITU-T Rec. H.263 Phụ lục P. Tuy nhiên, ở đây kích thước ảnh không thay đổi, chỉ các ảnh tham chiếu đang được lấy mẫu lại, dẫn đến khả năng chỉ 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). Hơn nữa, H.263 Phụ lục Q cho phép lấy mẫu lại của khối macro (macroblock) riêng lẻ theo hệ số hai (trong mỗi chiều), hướng lên hoặc hướng xuống. Một lần nữa, kích thước ảnh vẫn giữ nguyên. Kích thước của khối macro (macroblock) được cố định trong H.263, và do đó không cần phải được báo hiệu.

Bản chất kỹ thuật của sáng chế

Theo phương án, có phương pháp giải mã dòng bit video được mã hóa được đề xuất sử dụng ít nhất một bộ xử lý, bao gồm: thu nhận từ dòng bit video được mã hóa chuỗi video được mã hóa bao gồm đơn vị ảnh tương ứng với ảnh được mã hóa; thu nhận đơn vị lớp trừu tượng mạng (NAL) tiêu đề ảnh (PH) được bao gồm trong đơn vị ảnh; thu nhận ít nhất một đơn vị lớp trừu tượng mạng (NAL) lớp mã hóa video (VCL) được bao gồm trong đơn vị ảnh; giải mã ảnh được mã hóa dựa trên đơn vị lớp trừu tượng mạng PH, ít nhất một đơn vị lớp trừu tượng mạng VCL, và tập tham số thích ứng (APS) được bao gồm trong đơn vị lớp trừu tượng mạng APS thu nhận từ chuỗi video được mã hóa; và xuất ra ảnh đã được giải mã, trong

đó đơn vị lớp trừu tượng mạng APS khả dụng cho ít nhất một bộ xử lý trước ít nhất một đơn vị lớp trừu tượng mạng VCL.

Theo phương án, thiết bị được đề xuất để giải mã dòng bit video được mã hóa, bao gồm: ít nhất một bộ nhớ được cấu hình để lưu trữ mã chương trình; và ít nhất một bộ xử lý được cấu hình để đọc mã chương trình và hoạt động theo lệnh của mã chương trình, mã chương trình bao gồm: mã thu nhận thứ nhất được cấu hình để khiến ít nhất một bộ xử lý thu nhận từ dòng bit video được mã hóa chuỗi video được mã hóa bao gồm đơn vị ảnh tương ứng với ảnh được mã hóa; mã thu nhận thứ hai được cấu hình để khiến ít nhất một bộ xử lý thu nhận đơn vị lớp trừu tượng mạng (NAL) tiêu đề ảnh (PH) được bao gồm trong đơn vị ảnh; mã thu nhận thứ ba được cấu hình để khiến ít nhất một bộ xử lý thu nhận ít nhất một đơn vị lớp trừu tượng mạng (NAL) lớp mã hóa video (VCL) được bao gồm trong đơn vị ảnh; mã giải mã được cấu hình để khiến ít nhất một bộ xử lý giải mã ảnh được mã hóa dựa trên đơn vị lớp trừu tượng mạng PH, ít nhất một đơn vị lớp trừu tượng mạng VCL, và tập tham số thích ứng (APS) bao gồm trong đơn vị lớp trừu tượng mạng APS thu nhận từ chuỗi video được mã hóa; và mã đầu ra được cấu hình để khiến ít nhất một bộ xử lý xuất ra ảnh đã được giải mã, trong đó đơn vị lớp trừu tượng mạng APS khả dụng cho ít nhất một bộ xử lý trước ít nhất một bộ VCL NAL.

Trong phương án, phương tiện đọc được bằng máy tính không nhất thời lưu trữ các lệnh được đề xuất, các lệnh này bao gồm: một hoặc nhiều lệnh, khi được thực thi bởi một hoặc nhiều bộ xử lý của thiết bị để giải mã dòng bit video được mã hóa, khiến một hoặc nhiều bộ xử lý để: thu nhận từ dòng bit video được mã hóa chuỗi video được mã hóa bao gồm đơn vị ảnh tương ứng với ảnh được mã hóa; thu nhận đơn vị lớp trừu tượng mạng (NAL) tiêu đề ảnh (PH) được bao gồm trong đơn vị ảnh; thu nhận ít nhất một đơn vị lớp trừu tượng mạng (NAL) lớp mã hóa video (VCL) bao gồm trong đơn vị ảnh; giải mã ảnh được mã hóa dựa trên đơn vị lớp trừu tượng mạng PH, ít nhất một đơn vị lớp trừu tượng mạng VCL, và tập tham số thích ứng (APS) bao gồm trong đơn vị lớp trừu tượng mạng APS thu nhận từ chuỗi video được mã hóa; và xuất ra ảnh đã được giải mã, trong đó đơn vị lớp trừu tượng mạng APS khả dụng cho ít nhất một bộ xử lý trước ít nhất một đơn vị lớp trừu tượng mạng VCL.

Mô tả vắn tắt các hình vẽ

Các đặc điểm khác, bản chất, và các ưu điểm khác nhau của đối tượng được bộc lộ sẽ rõ ràng hơn từ phần mô tả chi tiết sau đây và các bản vẽ kèm theo trong đó:

FIG.1 là minh họa sơ đồ của sơ đồ khối được đơn giản hóa của hệ thống thông tin liên lạc theo phương án.

FIG.2 là minh họa sơ đồ của sơ đồ khối được đơn giản hóa của hệ thống thông tin liên lạc theo phương án.

FIG.3 là minh họa sơ đồ của sơ đồ khối đơn giản của bộ giải mã theo một phương án.

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

Các FIG.5A-5E là các minh họa sơ đồ về các tùy chọn cho các tham số ARC báo hiệu theo phương án, theo phương án.

Các FIG.6A-6B là minh họa sơ đồ về các ví dụ về bảng cú pháp theo phương án.

FIG.7 là ví dụ về cấu trúc dự đoán cho khả năng mở rộng với sự thay đổi độ phân giải thích ứng, theo phương án.

FIG.8 là ví dụ về bảng cú pháp theo phương án.

FIG.9 là minh họa sơ đồ của sơ đồ khối đơn giản về phân tích cú pháp và giải mã chu trình POC trên mỗi đơn vị truy cập và giá trị đếm đơn vị truy cập, theo phương án.

FIG.10 là minh họa giản đồ của cấu trúc dòng bit video bao gồm các ảnh con nhiều lớp, theo phương án.

FIG.11 là minh họa sơ đồ về việc hiển thị ảnh con đã chọn với độ phân giải nâng cao, theo phương án.

FIG.12 là sơ đồ khối của quá trình giải mã và hiển thị cho dòng bit video bao gồm các ảnh con nhiều lớp, theo phương án.

FIG.13 là minh họa sơ đồ về hiển thị video 360 với lớp nâng cao của ảnh con, theo phương án.

FIG.14 là ví dụ về thông tin bố cục của các ảnh con và cấu trúc dự đoán ảnh và lớp tương ứng của nó, theo phương án.

FIG.15 là ví dụ về thông tin bố cục của các ảnh con và lớp ảnh và cấu trúc dự đoán ảnh tương ứng của nó, với phương thức khả năng biến đổi tỷ lệ không gian của vùng cục bộ, theo phương án.

Các FIG.16A-16B là ví dụ về bảng cú pháp cho thông tin bố cục ảnh con, theo các phương án.

FIG.17 là ví dụ về bảng cú pháp của thông điệp SEI cho thông tin bố cục ảnh con, theo phương án.

FIG.18 là ví dụ về bảng cú pháp để chỉ báo các lớp đầu ra và thông tin biên dạng/cấp/mức cho mỗi tập lớp đầu ra, theo phương án.

FIG.19 là ví dụ về bảng cú pháp để chỉ báo chế độ lớp đầu ra được bật cho từng tập lớp đầu ra, theo phương án.

FIG.20 là ví dụ về bảng cú pháp để chỉ báo ảnh con hiện tại của mỗi lớp cho từng tập lớp đầu ra, theo phương án.

FIG.21 là minh họa giản đồ về yêu cầu tuân thủ dòng bit theo phương án.

FIG.22 là sơ đồ của quy trình ví dụ để giải mã luồng bit video được mã hóa theo phương án.

FIG.23 là minh họa sơ đồ của hệ thống máy tính theo phương án.

Mô tả chi tiết sáng chế

FIG.1 minh họa sơ đồ khối đơn giản hóa của hệ thống thông tin liên lạc 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. Để truyền dữ liệu một chiều, thiết bị đầu cuối thứ nhất 110 có thể mã hóa dữ liệu video tại vị trí cục bộ để truyền đến thiết bị đầu cuối còn lại 120 qua mạng 150. Thiết bị đầu cuối thứ hai 120 có thể nhận 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 đã khôi phục. Truyền dữ liệu một chiều có thể phổ biến trong các ứng dụng phục vụ phương tiện truyền thông và các ứng dụng tương tự.

FIG.1 minh họa cặp thiết bị đầu cuối thứ hai (130, 140) được đề xuất để hỗ trợ truyền hai chiều video được mã hóa có thể xảy ra, ví dụ, trong hội nghị truyền hình. Để truyền dữ liệu hai chiều, mỗi thiết bị đầu cuối (130, 140) có thể mã hóa dữ liệu video được ghi lại tại một vị trí cục bộ để truyền đến thiết bị đầu cuối kia qua mạng 150. Mỗi thiết bị đầu cuối (130, 140) cũng có thể nhận dữ liệu video được mã hóa do thiết bị đầu cuối kia truyền, có thể giải mã dữ liệu được mã hóa và có thể hiển thị dữ liệu video đã khôi phục tại thiết bị hiển thị cục bộ.

Trong FIG.1, các thiết bị đầu cuối (110-140) có thể được minh họa như máy chủ, máy tính cá nhân và điện thoại thông minh nhưng các nguyên tắc của sáng chế có thể không bị giới hạn như vậy. Các phương án theo sáng chế cho thấy ứng dụng với máy tính xách tay, máy tính bảng, trình phát đa phương tiện và/hoặc thiết bị hội nghị truyền hình chuyên dụng. Mạng 150 đại diện cho bất kỳ số lượng mạng nào truyền tải dữ liệu video được mã hóa giữa các thiết bị đầu cuối (110-140), bao gồm ví dụ mạng có dây và/hoặc mạng truyền thông không dây. Mạng truyền thông 150 có thể trao đổi dữ liệu trong các kênh chuyển mạch kênh và/hoặc chuyển mạch gói. Mạng đại diện bao gồm mạng viễn thông, mạng cục bộ, mạng diện rộng và/hoặc Internet. Đối với mục đích của cuộc tham khảo này, kiến trúc và cấu trúc liên kết của mạng 150 có thể không quan trọng đối với hoạt động của sáng chế trừ khi được giải thích ở đây dưới đây.

FIG.2 minh họa, làm ví dụ cho ứng dụng cho đối tượng được bộc lộ, vị trí của bộ mã hóa và giải mã video trong môi trường phát trực tuyến. Chủ đề được bộc lộ có thể áp dụng tương tự cho các ứng dụng hỗ trợ video khác, bao gồm, ví dụ, hội nghị truyền hình, TV kỹ thuật số, lưu trữ video nén trên phương tiện kỹ thuật số bao gồm CD, DVD, thẻ nhớ và các loại tương tự, v.v.

Hệ thống phát trực tuyến có thể bao gồm hệ thống con chụp 213, mà có thể bao gồm nguồn video 201, ví dụ như máy ảnh kỹ thuật số, tạo ra ví dụ như dòng mẫu video không nén 202. Dòng mẫu đó 202, được mô tả như một dòng in đậm

để nhấn mạnh khối lượng dữ liệu cao khi so sánh với dòng bit video được mã hóa, có thể được xử lý bởi một bộ mã hóa 203 được kết hợp với máy ảnh 201. Bộ mã hóa 203 có thể bao gồm phần cứng, phần mềm hoặc sự kết hợp của chúng để kích hoạt hoặc triển khai các khía cạnh của đối tượng được bộc lộ như được mô tả chi tiết hơn bên dưới. Dòng bit video được mã hóa 204, được mô tả như một đường mảnh để nhấn mạnh khối lượng dữ liệu thấp hơn khi so sánh với dòng mẫu, có thể được lưu trữ trên một máy chủ phát trực tuyến 205 để sử dụng trong tương lai. Một hoặc nhiều máy khách phát trực tuyến (206, 208) có thể truy cập máy chủ phát trực tuyến 205 để truy xuất các bản sao (207, 209) của luồng video được mã hóa 204. Máy khách 206 có thể bao gồm bộ giải mã video 210 giải mã bản sao đến của dòng bit video được mã hóa 207 và tạo dòng mẫu video đi 211 có thể được kết xuất trên màn hình 212 hoặc thiết bị kết xuất khác (không mô tả). Trong một số hệ thống phát trực tuyến, các dòng bit video (204, 207, 209) có thể được mã hóa theo các tiêu chuẩn nén/mã hóa video nhất định. Ví dụ về các tiêu chuẩn đó bao gồm Khuyến nghị H.265 của ITU-T. Đang được phát triển là một tiêu chuẩn mã hóa video thường được gọi là Mã hóa video đa năng hoặc VVC. Chủ đề 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 nhận một hoặc nhiều chuỗi video codec sẽ được giải mã bởi bộ giải mã 210; theo cùng phương án hoặc phương án khác, một chuỗi video được mã hóa tại một thời điểm, trong đó việc giải mã từng chuỗi video được mã hóa là độc lập với các chuỗi video được mã hóa khác. Chuỗi video được mã hóa có thể được nhận từ một kênh 312, có thể là liên kết phần cứng/phần mềm tối thiểu bị lưu trữ mà lưu trữ dữ liệu video được mã hóa. Máy thu 310 có thể nhận dữ liệu video được mã hóa cùng với dữ liệu khác, ví dụ, dữ liệu âm thanh được mã hóa và/hoặc các dòng dữ liệu phụ trợ, có thể được chuyển tiếp đến các thực thể tương ứng bằng cách sử dụng (không được mô tả). Bộ thu 310 có thể tách chuỗi video được mã hóa khỏi dữ liệu khác. Để chống lại hiện tượng chập chờn mạng, bộ nhớ đệm 315 có thể được ghép nối giữa bộ thu 310 và bộ giải mã entropy/ bộ phân tích cú pháp 320 (từ nay về sau là “bộ phân tích cú pháp”). Khi bộ thu 310 đang

nhận dữ liệu từ thiết bị lưu trữ/chuyển tiếp có đủ băng thông và khả năng điều khiển, hoặc từ mạng đẳng thời, bộ đệm 315 có thể không cần thiết, hoặc có thể nhỏ. Để sử dụng trên các mạng gói nỗ lực tốt nhất như Internet, bộ đệm 315 có thể được yêu cầu, có thể tương đối lớn và có thể có kích thước thích ứng một cách thuận lợi.

Bộ giải mã video 210 có thể bao gồm bộ phân tích cú pháp 320 để tái cấu trúc các ký hiệu 321 từ chuỗi video được mã hóa entropy. Các danh mục của các ký hiệu đó bao gồm thông tin được sử dụng để quản lý hoạt động của bộ giải mã 210, và thông tin tiềm năng để điều khiển thiết bị kết xuất như màn hình 212 mà không phải là một phần không thể tách rời của bộ giải mã nhưng có thể được kết hợp với nó, như được thể hiện trên FIG.3. Thông tin điều khiển cho (các) thiết bị kết xuất có thể ở dạng Thông tin tăng cường bổ sung (thông báo SEI) hoặc các đoạn bộ tham số Thông tin về khả năng sử dụng video (VUI) (không được mô tả). Bộ phân tích cú pháp 320 có thể phân tích cú pháp và/hoặc giải mã entropy chuỗi video được mã hóa nhận được. Việc mã hóa chuỗi video được mã hóa có thể phù hợp với kỹ thuật hoặc tiêu chuẩn mã hóa video, và có thể tuân theo các nguyên tắc mà một người có hiểu biết trung bình trong lĩnh vực kỹ thuật này biết rõ, bao gồm mã hóa độ dài thay đổi, mã hóa Huffman, mã hóa số học có hoặc không có độ nhảy ngữ cảnh và v.v.. Bộ phân tích cú pháp 320 có thể trích xuất từ chuỗi video được mã hóa, bộ các tham số nhóm con cho ít nhất một trong các nhóm con 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 Nhóm ảnh (GOP), ảnh, ảnh con, ô, lát, khối hình, khối macro (macroblock), Đơn vị mã hóa (CU) Đơn vị cây mã hóa (CTU), khối, Đơn vị biến đổi (TU), Đơn vị dự đoán (PU) và v.v.. Ô có thể chỉ báo vùng hình chữ nhật của CU/CTU bên trong cột và hàng ô cụ thể trong ảnh. Gạch có thể chỉ báo vùng hình chữ nhật của các hàng CU/CTU trong một ô cụ thể. Lát có thể chỉ báo một hoặc nhiều khối hình của một bức tranh, mà được chứa trong đơn vị NAL. Ảnh con có thể chỉ báo vùng hình chữ nhật của một hoặc nhiều lát trong ảnh. Bộ giải mã entropy/bộ phân tích cú pháp cũng có thể trích xuất từ thông tin chuỗi video được mã hóa như hệ số biến đổi, giá trị tham số bộ lượng tử, vectơ chuyển động, v.v..

Bộ phân tích cú pháp 320 có thể thực hiện hoạt động giải mã entropy và/hoặc phân tích cú pháp trên chuỗi video nhận được từ bộ đệm 315, do đó để tạo các ký hiệu 321.

Việc tái tạo các ký hiệu 321 có thể liên quan đến nhiều đơn vị khác nhau tùy thuộc vào loại ảnh video được mã hóa hoặc các phần của chúng (chẳng hạn như: ảnh liên và nội bộ, khối liên và nội bộ), và các hệ số khác. Những đơn vị nào có liên quan, và bằng cách nào, có thể được kiểm soát bởi thông tin điều khiển nhóm con đã được phân tích cú pháp từ chuỗi video được mã hóa bởi bộ phân tích cú pháp 320. Luồng thông tin điều khiển nhóm con như vậy giữa bộ phân tích cú pháp 320 và nhiều đơn vị bên dưới không được mô tả rõ ràng.

Ngoài các khối chức năng đã được đề cập, bộ giải mã 210 có thể được chia nhỏ về mặt khái niệm thành một số đơn vị chức năng như mô tả bên dưới. Trong một triển khai thực tế hoạt động dưới các ràng buộc thương mại, nhiều đơn vị trong số này tương tác chặt chẽ với nhau và có thể, ít nhất phần nào, có thể được tích hợp vào nhau. Tuy nhiên, với mục đích mô tả vấn đề được bộc lộ, việc chia nhỏ khái niệm thành các đơn vị chức năng dưới đây là phù hợp.

Đơn vị thứ nhất là bộ biến đổi tỷ lệ và/hoặc bộ biến đổi nghịch đảo 351. Bộ biến đổi tỷ lệ và/hoặc bộ biến đổi nghịch đảo 351 nhận hệ số biến đổi lượng tử hóa cũng như thông tin điều khiển, bao gồm biến đổi sẽ sử dụng, kích thước khối, hệ số lượng tử hóa, ma trận tỷ lệ lượng tử hóa, v.v. dưới dạng (các) ký hiệu 321 từ bộ phân tích cú pháp 320. Nó có thể xuất ra các khối bao gồm các giá trị mẫu, mà có thể được nhập vào bộ tổng hợp 355.

Trong một số trường hợp, các mẫu đầu ra của bộ biến đổi tỷ lệ và/hoặc biến đổi nghịch đảo 351 có thể liên quan đến khối được mã hóa nội bộ; nghĩa là: khối không sử dụng thông tin dự đoán từ các ảnh được tái cấu trúc trước đó, nhưng có thể sử dụng thông tin dự đoán từ các phần được tái cấu trúc trước đó của ảnh hiện tại. Thông tin dự đoán như vậy có thể được đề xuất bởi đơn vị dự đoán ảnh nội bộ 352. Trong một số trường hợp, đơn vị dự đoán ảnh nội bộ 352 tạo ra khối có cùng kích thước và hình dạng của khối đang được tái cấu trúc, sử dụng thông tin đã được tái cấu trúc xung quanh được lấy từ ảnh hiện tại (được tái cấu trúc một phần)

358. Bộ tổng hợp 355, trong một số trường hợp, bổ sung, trên cơ sở mỗi mẫu, thông tin dự đoán mà đơn vị dự đoán nội bộ 352 đã tạo cho thông tin mẫu đầu ra như được đề xuất bởi bộ biến đổi tỉ lệ/hoặc bộ biến đổi nghịch đảo 351.

Trong các trường hợp khác, các mẫu đầu ra của bộ biến đổi tỉ lệ và/hoặc bộ biến đổi nghịch đảo 351 có thể liên quan đến một khối được mã hóa liên, và có khả năng bù chuyển động. Trong trường hợp như vậy, bộ dự báo bù chuyển động 353 có thể truy cập bộ đệm ảnh tham chiếu 357 để nạp các mẫu được sử dụng để dự đoán. Sau khi bù chuyển động các mẫu được lấy phù hợp với các ký hiệu 321 liên quan đến khối, các mẫu này có thể được thêm vào bởi bộ tổng hợp 355 cho đầu ra của bộ biến đổi tỉ lệ và/hoặc bộ biến đổi nghịch đảo (trong trường hợp này được gọi là các mẫu phần dư hoặc tín hiệu dư) để tạo ra thông tin mẫu đầu ra. Các địa chỉ trong biểu mẫu bộ nhớ ảnh tham chiếu nơi đơn vị bù chuyển động tìm nạp các mẫu dự đoán có thể được điều khiển bằng vector chuyển động, khả dụng cho đơn vị bù chuyển động dưới dạng ký hiệu 321 mà có thể có, ví dụ X, Y và các thành phần ảnh tham chiếu. Bù chuyển động cũng có thể bao gồm nội suy các giá trị mẫu được nạp từ bộ nhớ ảnh tham chiếu khi sử dụng vector chuyển động chính xác của mẫu phụ, cơ chế dự đoán vector chuyển động, và v.v..

Các mẫu đầu ra của bộ tổng hợp 355 có thể tuân theo các kỹ thuật lọc vòng khác nhau trong bộ lọc vòng 356. Kỹ thuật nén video có thể bao gồm kỹ thuật lọc trong vòng được kiểm soát bởi các tham số có trong dòng bit video được mã hóa và được đề xuất cho đơn vị lọc vòng 356 dưới dạng ký hiệu 321 từ bộ phân tích cú pháp 320, nhưng cũng có thể đáp ứng với siêu thông tin thu nhận trong quá trình giải mã các phần trước đó (theo thứ tự giải mã) của chuỗi ảnh được mã hóa hoặc video được mã hóa, cũng như đáp ứng các giá trị mẫu được tái cấu trúc và lọc vòng lặp trước đó.

Đầu ra của đơn vị lọc vòng 356 có thể là một dòng mẫu mà có thể được xuất ra thiết bị kết xuất 212 cũng như được lưu trữ trong bộ nhớ ảnh tham chiếu để sử dụng trong dự đoán ảnh liên đối trong tương lai.

Một số ảnh được mã hóa nhất định, sau khi được tái cấu trúc hoàn chỉnh, có thể được sử dụng làm ảnh tham chiếu để dự đoán trong tương lai. Sau khi ảnh

được mã hóa được tái cấu trúc hoàn chỉnh và ảnh được mã hóa đã được xác định là ảnh tham chiếu (bởi, ví dụ, bộ phân tích cú pháp 320), ảnh tham chiếu hiện tại 358 có thể trở thành một phần của bộ đệm ảnh tham chiếu 357, và bộ nhớ ảnh hiện tại mới có thể được phân bổ lại trước khi bắt đầu tái tạo ảnh được mã hóa sau.

Bộ giải mã video 210 có thể thực hiện các hoạt động giải mã theo kỹ thuật nén video định trước mà có thể được ghi lại trong một tiêu chuẩn, chẳng hạn như ITU-T Rec. H.265. Chuỗi video được mã hóa có thể tuân theo cú pháp được chỉ định bởi kỹ thuật nén video hoặc tiêu chuẩn đang được sử dụng, theo nghĩa mà nó tuân theo cú pháp của kỹ thuật hoặc tiêu chuẩn nén video, như được chỉ định trong tài liệu kỹ thuật nén video hoặc tiêu chuẩn và cụ thể trong tài liệu hồ sơ trong đó. Cũng cần thiết để tuân thủ có thể là mức độ phức tạp của chuỗi video được mã hóa nằm trong giới hạn được xác định bởi cấp độ của kỹ thuật hoặc tiêu chuẩn nén video. Trong một số trường hợp, các cấp hạn chế kích thước ảnh tối đa, tốc độ khung hình tối đa, tốc độ mẫu tái tạo tối đa (được đo bằng, ví dụ như mẫu lớn trên giây), kích thước ảnh tham chiếu tối đa, v.v. Các giới hạn do các cấp thiết lập có thể, trong một số trường hợp, bị hạn chế thêm thông qua các tham số kỹ thuật của Bộ giải mã tham chiếu giả thuyết (HRD) và siêu dữ liệu để quản lý bộ đệm HRD được báo hiệu trong chuỗi video được mã hóa.

Theo phương án, bộ thu 310 có thể nhận dữ liệu bổ sung (dư) với video được mã hóa. Dữ liệu bổ sung có thể được bao gồm như một phần của (các) chuỗi video được mã hóa. Bộ giải mã video 210 có thể sử dụng dữ liệu bổ sung để giải mã dữ liệu một cách chính xác và/hoặc để tái tạo chính xác hơn dữ liệu video gốc. Dữ liệu bổ sung có thể ở dạng, ví dụ, các lớp tăng cường theo thời gian, không gian hoặc SNR, các lát dư, ảnh dư, mã sửa lỗi chuyển tiếp, v.v..

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

Nguồn video 201 có thể cung cấp chuỗi video nguồn được mã hóa bởi bộ mã hóa 203 ở dạng dòng mẫu video kỹ thuật số mà có thể có độ sâu bit phù hợp bất kỳ (ví dụ: 8 bit, 10 bit, 12 bit,...), bất kỳ không gian màu nào (ví dụ, BT.601 Y CrCb, RGB,...) và bất kỳ cấu trúc lấy mẫu phù hợp nào (ví dụ Y CrCb 4: 2: 0, Y CrCb 4: 4: 4). Trong hệ thống phân phát phương tiện, nguồn video 201 có thể là thiết bị lưu trữ lưu trữ video đã chuẩn bị trước đó. Trong hệ thống hội nghị truyền hình, nguồn video 203 có thể là máy ảnh ghi lại thông tin ảnh cục bộ dưới dạng một chuỗi video. Dữ liệu video có thể được đề xuất dưới dạng nhiều ảnh riêng lẻ mà truyền tải chuyển động khi xem theo chuỗi. Bản thân các ảnh có thể được tổ chức dưới dạng mảng điểm ảnh không gian, trong đó mỗi điểm ảnh có thể bao gồm một hoặc nhiều mẫu tùy 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 này có thể dễ dàng hiểu được mối quan hệ giữa điểm ảnh và mẫu. Mô tả bên dưới tập trung vào các mẫu.

Theo phương án, bộ mã hóa 203 có thể mã hóa và nén ảnh của chuỗi video nguồn thành chuỗi video được mã hóa 443 trong thời gian thực hoặc theo bất kỳ ràng buộc thời gian nào khác theo yêu cầu của ứng dụng. Thực thi tốc độ mã hóa thích hợp là một chức năng của Bộ điều khiển 450. Bộ điều khiển điều khiển các đơn vị chức năng khác như được mô tả dưới đây và được ghép nối về mặt chức năng với các đơn vị này. Ghép nối không được mô tả rõ ràng. Các tham số thiết lập bởi bộ điều khiển có thể bao gồm các tham số liên quan đến điều khiển tốc độ (bỏ qua ảnh, bộ định lượng, giá trị lambda của các kỹ thuật tối ưu hóa tốc độ biến dạng...), kích thước ảnh, bố cục nhóm ảnh (GOP), phạm vi tìm kiếm vector chuyển động tối đa, v.v. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật này có thể dễ dàng xác định các chức năng khác của bộ điều khiển 450 vì chúng có thể liên quan đến bộ mã hóa video 203 được tối ưu hóa cho thiết kế hệ thống nhất định.

Một số bộ mã hóa video hoạt động theo cách mà một người có hiểu biết trung bình dễ dàng nhận ra là "vòng lặp mã hóa". Là mô tả đơn giản hóa quá mức, vòng lặp mã hóa có thể bao gồm phần mã hóa của bộ mã hóa 430 ("mã nguồn" từ giờ trở đi) (chịu trách nhiệm tạo các ký hiệu dựa trên ảnh đầu vào được mã hóa và (các) ảnh tham chiếu), và bộ giải mã (cục bộ) 433 được nhúng trong bộ mã hóa

203 sẽ tái tạo các ký hiệu để tạo dữ liệu mẫu mà bộ giải mã (từ xa) cũng sẽ tạo ra (vì bất kỳ sự nén nào giữa các ký hiệu và dòng bit video được mã hóa đều không tồn hao trong kỹ thuật nén video được xem xét trong đối tượng được bộc lộ). Dòng mẫu được tái cấu trúc đó được đưa vào bộ nhớ ảnh tham chiếu 434. Vì việc giải mã dòng ký hiệu dẫn đến kết quả chính xác bit không phụ thuộc vào vị trí của bộ giải mã (cục bộ hoặc từ xa), nội dung bộ đệm ảnh tham chiếu cũng chính xác bit giữa bộ mã hóa cục bộ và bộ mã hóa từ xa. Nói cách khác, phần dự đoán của bộ mã hóa “nhìn thấy” như các mẫu ảnh tham chiếu chính xác các giá trị mẫu giống như bộ giải mã sẽ “nhìn thấy” khi sử dụng dự đoán trong quá trình giải mã. Nguyên tắc cơ bản của tính đồng bộ ảnh tham chiếu này (và dẫn đến trôi dạt, nếu tính đồng bộ không thể duy trì, ví dụ như do lỗi kênh) được biết rõ với người có hiểu biết trung bình trong lĩnh vực kỹ thuật này.

Hoạt động của bộ giải mã “cục bộ” 433 có thể giống như hoạt động của bộ giải mã “từ xa” 210, mà đã được mô tả chi tiết ở trên cùng với FIG.3. Cũng đề cập ngắn gọn đến FIG.4, tuy nhiên, vì các ký hiệu khả dụng và việc mã hóa và/hoặc giải mã các ký hiệu thành chuỗi video được mã hóa bằng bộ mã hóa entropy 445 và bộ phân tích cú pháp 320 có thể không tồn hao, các bộ phận giải mã entropy của bộ giải mã 210, bao gồm cả kênh (312), bộ thu 310, bộ đệm 315, và bộ phân tích cú pháp 320 có thể không được triển khai đầy đủ trong bộ giải mã cục bộ 433.

Quan sát có thể được thực hiện tại thời điểm này là bất kỳ kỹ thuật giải mã nào ngoại trừ phân tích cú pháp và/hoặc giải mã entropy mà có trong bộ giải mã cũng nhất thiết phải có mặt, ở dạng chức năng về cơ bản giống hệt nhau, trong một bộ mã hóa tương ứng. Vì lý do này, đối tượng được bộc lộ tập trung vào hoạt động của bộ giải mã. Mô tả của các kỹ thuật bộ mã hóa có thể được viết tắt vì chúng là nghịch đảo của các kỹ thuật bộ giải mã được mô tả toàn diện. Chỉ ở một số khu vực nhất định mô tả chi tiết hơn mới được yêu cầu và cung cấp bên dưới.

Là một phần của hoạt động, bộ mã nguồn 430 có thể thực hiện mã hóa dự đoán bù chuyển động, mà mã hóa khung đầu theo dự đoán với tham chiếu đến một hoặc nhiều khung được mã hóa trước đó từ chuỗi video mà được chỉ định là

“khung tham chiếu.” Theo cách này, công cụ mã hóa 432 mã hóa sự 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 chọn làm (các) tham chiếu dự đoán cho 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 làm khung tham chiếu, dựa trên các ký hiệu được tạo bởi bộ mã nguồn 430. Các hoạt động của công cụ mã hóa 432 có thể là các quá trình tổn hao. Khi dữ liệu video được mã hóa có thể được giải mã tại bộ giải mã video (không được thể hiện trên Hình 4), chuỗi video được tạo lại thường có thể là bản sao của chuỗi video nguồn với một số lỗi. Bộ giải mã video cục bộ 433 sao chép các quy trình giải mã có thể được thực hiện bởi bộ giải mã video trên các khung tham chiếu và có thể khiến các khung tham chiếu được tái cấu trúc được lưu trữ trong bộ đệm ảnh tham chiếu 434. Theo cách này, bộ mã hóa 203 có thể lưu trữ cục bộ các bản sao của khung tham chiếu được tái cấu trúc có nội dung phổ biến là khung tham chiếu được tái cấu trúc sẽ được bộ giải mã video ở xa thu nhận (không có lỗi truyền).

Bộ dự đoán 435 có thể thực hiện các tìm kiếm dự đoán cho công cụ mã hóa 432. Nghĩa là, đối với khung mới được mã hóa, bộ dự đoán 435 có thể tìm kiếm trong bộ nhớ ảnh tham chiếu 434 để tìm dữ liệu mẫu (dưới dạng khối điểm ảnh tham chiếu ứng cử viên) hoặc siêu dữ liệu nhất định như vector chuyển động ảnh tham chiếu, hình khối, v.v. trên, điều đó có thể đóng vai trò như tham chiếu dự đoán thích hợp cho các ảnh mới. Bộ dự đoán 435 có thể hoạt động trên cơ sở mẫu khối từng điểm ảnh để tìm các tham chiếu dự đoán thích hợp. Trong một số trường hợp, như được xác định bởi kết quả tìm kiếm thu nhận bởi công cụ dự đoán 435, ảnh đầu vào có thể có các tham chiếu dự đoán được lấy từ nhiều ảnh tham chiếu được lưu trữ trong bộ nhớ ảnh tham chiếu 434.

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

Đầu ra của tất cả các đơn vị chức năng nói trên có thể được mã hóa entropy trong bộ mã hóa entropy 445. Bộ mã hóa entropy dịch các ký hiệu do các đơn vị chức năng khác nhau tạo ra thành chuỗi video được mã hóa, bằng cách nén không tổn hao các ký hiệu theo kỹ thuật mà người có hiểu biết trung bình trong lĩnh vực kỹ thuật này biết, ví dụ như mã hóa Huffman, mã hóa độ dài thay đổi, mã số học, và kể từ đó trở đi.

Bộ phát 440 có thể đệm (các) chuỗi video được mã hóa như được tạo bởi bộ mã hóa entropy 445 để chuẩn bị cho việc truyền qua kênh liên lạc 460, 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 phát 440 có thể hợp nhất dữ liệu video được mã hóa từ bộ mã hóa video 430 với dữ liệu khác được truyền, ví dụ, dữ liệu âm thanh được mã hóa và/hoặc các luồng dữ liệu phụ trợ (các nguồn không được hiển thị).

Bộ điều khiển 450 có thể quản lý hoạt động của bộ mã hóa 203. Trong lúc mã hóa, bộ điều khiển 450 có thể gán cho mỗi ảnh được mã hóa loại ảnh được mã hóa nhất định, mà có thể ảnh hưởng đến các kỹ thuật mã hóa có thể được áp dụng cho ảnh tương ứng. Ví dụ, ảnh thường có thể được chỉ định là một trong các loại khung sau đây:

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

Ảnh dự đoán (ảnh P) có thể là ảnh mà được mã hóa và giải mã bằng cách sử dụng dự đoán nội bộ hoặc dự đoán liên 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 kép (ảnh B) có thể là ảnh mà được mã hóa và giải mã bằng cách sử dụng dự đoán nội bộ hoặc dự đoán liên sử dụng nhiều nhất hai vector chuyển động và chỉ số tham chiếu để dự đoán 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 liên quan để tái tạo khối duy nhất.

Ảnh nguồn thường có thể được chia nhỏ theo không gian thành nhiều khối mẫu (ví dụ, mỗi khối 4x4, 8x8, 4x8 hoặc 16x16 mẫu) và được mã hóa trên cơ sở từng khối. Các khối có thể được mã hóa theo dự đoán với tham chiếu đến các khối khác (đã được mã hóa) như được xác định bởi việc gán mã được áp dụng cho các ảnh tương ứng của khối. Ví dụ, các khối ảnh I có thể được mã hóa không theo dự đoán hoặc chúng có thể được mã hóa theo dự đoán với tham chiếu đến các khối đã được mã hóa của cùng ảnh (dự đoán không gian hoặc dự đoán nội bộ). Các khối điểm ảnh của ảnh P có thể được mã hóa không theo dự đoán, thông qua dự đoán không gian hoặc dự đoán thời gian với tham chiếu đến một ảnh tham chiếu đã được mã hóa trước đó. Các khối ảnh B có thể được mã hóa không theo dự đoán, thông qua dự đoán không gian hoặc dự đoán thời gian với tham chiếu đến 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 được xác định trước, chẳng hạn như ITU-T Rec. H.265. Trong hoạt động của nó, bộ mã hóa video 203 có thể thực hiện các hoạt động nén khác nhau, bao gồm các hoạt động mã hóa dự đoán khai thác các phần dư về thời gian và không gian trong chuỗi video đầu vào. Dữ liệu video được mã hóa, do đó, có thể tuân theo cú pháp được chỉ định bởi kỹ thuật hoặc tiêu chuẩn mã hóa video đang được sử dụng.

Theo phương án, bộ phát 440 có thể truyền dữ liệu bổ sung cùng với video được mã hóa. Bộ mã hóa video 430 có thể bao gồm dữ liệu như phần của chuỗi video được mã hóa. Dữ liệu bổ sung có thể bao gồm các lớp tăng cường theo thời gian/không gian/SNR, các dạng dữ liệu dư khác như ảnh và lát dư, thông báo Thông tin nâng cao bổ sung (SEI), các đoạn bộ tham số Thông tin khả năng sử dụng trực quan (VUI), v.v..

Gần đây, bộ miền nén hoặc trích xuất nhiều phần ảnh độc lập về ngữ nghĩa thành ảnh video duy nhất đã thu hút được sự chú ý. Cụ thể, trong bối cảnh, ví dụ, mã hóa 360 hoặc các ứng dụng giám sát nhất định, nhiều ảnh nguồn độc lập về

ngữ nghĩa (ví dụ: bề mặt sáu khối của cảnh 360 được chiếu hình khối, hoặc đầu vào camera riêng lẻ trong trường hợp giám sát nhiều camera thiết lập) có thể yêu cầu cài đặt độ phân giải thích ứng riêng biệt để đối phó với các hoạt động trên mỗi cảnh khác nhau tại thời điểm nhất định. Nói cách khác, các bộ mã hóa, tại thời điểm nhất định, có thể chọn sử dụng các hệ số lấy mẫu lại khác nhau cho các ảnh độc lập về ngữ nghĩa khác nhau tạo nên toàn bộ 360 hoặc cảnh giám sát. Khi được kết hợp thành một ảnh duy nhất, mà, đến lượt, đòi hỏi lấy mẫu lại ảnh tham chiếu được thực hiện, và tín hiệu mã hóa độ phân giải thích ứng khả dụng, cho các phần của ảnh được mã hóa.

Dưới đây, một số thuật ngữ sẽ được giới thiệu mà sẽ được đề cập đến trong phần còn lại của mô tả này.

Ảnh con có thể đề cập đến, trong một số trường hợp, sự sắp xếp hình chữ nhật của các mẫu, khối, khối macro (macroblock), đơn vị mã hóa, hoặc các thực thể tương tự được nhóm theo ngữ nghĩa, và có thể được mã hóa độc lập theo độ phân giải đã thay đổi. Một hoặc nhiều ảnh con có thể tạo thành 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 lại thành ảnh, và một hoặc nhiều ảnh con có thể được trích xuất từ ảnh. Trong một số môi trường nhất định, một hoặc nhiều ảnh con được mã hóa có thể được ghép lại trong miền nén mà không cần chuyển mã đến cấp độ mẫu thành ảnh được mã hóa, và trong cấp độ tương tự hoặc các trường hợp khác, một hoặc nhiều ảnh con được mã hóa có thể được trích xuất từ ảnh được mã hóa trong miền nén.

Thay đổi độ phân giải thích ứng (ARC) có thể đề cập đến các cơ chế mà cho phép thay đổi độ phân giải của ảnh hoặc ảnh con trong chuỗi video được mã hóa, bằng cách, ví dụ, lấy mẫu lại ảnh tham chiếu. Các tham số ARC từ đó đề cập đến thông tin điều khiển cần thiết để thực hiện thay đổi độ phân giải thích ứng, có thể bao gồm, ví dụ, tham số bộ lọc, hệ số tỷ lệ, độ phân giải của đầu ra và/hoặc ảnh tham chiếu, các cờ điều khiển khác nhau, v.v..

Theo các phương án mã hóa và giải mã có thể được thực hiện trên ảnh video được mã hóa duy nhất, độc lập về mặt ngữ nghĩa. Trước khi mô tả hàm ý của việc

mã hóa/giải mã nhiều ảnh con với các tham số ARC độc lập và độ phức tạp bổ sung ngụ ý của nó, các tùy chọn cho báo hiệu tham số ARC sẽ được mô tả.

Đề cập đến Các FIG.5A-5E, được hiển thị là một số phương án cho báo hiệu các tham số ARC. Như đã lưu ý với mỗi phương án, chúng có thể có một số ưu điểm và nhược điểm nhất định theo quan điểm hiệu quả mã hóa, độ phức tạp và kiến trúc. Tiêu chuẩn hoặc kỹ thuật mã hóa video có thể chọn một hoặc nhiều phương án này, hoặc các tùy chọn đã biết từ kỹ thuật liên quan, để báo hiệu các tham số ARC. Các phương án có thể không loại trừ lẫn nhau, và hình dung được có thể được thay thế cho nhau dựa trên nhu cầu ứng dụng, kỹ thuật tiêu chuẩn liên quan, hoặc lựa chọn của bộ mã hóa.

Các loại tham số ARC có thể bao gồm:

- Hệ số lấy mẫu lên và/hoặc hệ số lấy mẫu xuống, tách biệt hoặc kết hợp trong kích thước X và Y.

- Các hệ số lấy mẫu lên và/hoặc lấy mẫu xuống, với việc bổ sung kích thước thời gian, chỉ báo tốc độ phóng to/thu nhỏ không đổi đối với một số ảnh nhất định.

- Một trong hai cách trên có thể liên quan đến việc mã hóa một hoặc nhiều phần tử cú pháp ngắn gọn có thể được mã hóa có thể trở vào bảng chứa (các) hệ số.

- Độ phân giải, theo kích thước X hoặc Y, theo đơn vị mẫu, khối, khối macro (macroblock), đơn vị mã hóa (CUs), hoặc bất kỳ độ chi tiết phù hợp nào khác, của ảnh đầu vào, ảnh đầu ra, ảnh tham chiếu, ảnh được mã hóa, kết hợp hoặc riêng biệt. Nếu có nhiều hơn một độ phân giải (chẳng hạn như, ví dụ, một cho ảnh đầu vào, một cho ảnh tham chiếu) thì, trong một số trường hợp nhất định, một bộ giá trị có thể được suy ra từ một bộ giá trị khác. Như vậy có thể tạo công, ví dụ, bằng cách sử dụng các cờ. Để có ví dụ chi tiết hơn, xem phần bên dưới.

- Tọa độ “cong vênh” tương tự như tọa độ được sử dụng trong H.263 Phụ lục P, một lần nữa độ chi tiết phù hợp như đã mô tả ở trên. H.263 Phụ lục P xác định một cách hiệu quả để mã các tọa độ cong vênh như vậy, nhưng cách khác, cách có tiềm năng có hiệu quả hơn có thể hình dung được cũng được nghĩ ra. Ví dụ, mã hóa theo kiểu “Huffman” có thể đảo ngược độ dài có thể thay đổi của tọa

độ cong vênh của Phụ lục P có thể được thay thế bằng mã hóa nhị phân có độ dài phù hợp, nơi độ dài của từ mã nhị phân có thể, ví dụ, được tính từ kích thước ảnh tối đa, có thể nhân với một hệ số nhất định và phần bù bởi một giá trị nhất định, để cho phép "cong vênh" bên ngoài ranh giới của kích thước ảnh tối đa.

-các tham số bộ lọc lấy mẫu lên và/hoặc lấy mẫu xuống. Trong các phương án, có thể chỉ có bộ lọc duy nhất để lấy mẫu lên và/hoặc lấy mẫu xuống. Tuy nhiên, theo các phương án, có thể mong muốn cho phép linh hoạt hơn trong thiết kế bộ lọc, và điều đó có thể yêu cầu báo hiệu các tham số bộ lọc. Các tham số như vậy có thể được chọn thông qua một chỉ số trong danh sách các thiết kế bộ lọc có thể có, bộ lọc có thể được chỉ định đầy đủ (ví dụ thông qua danh sách các hệ số bộ lọc, sử dụng kỹ thuật mã hóa entropy phù hợp), bộ lọc có thể được chọn ngầm thông qua tỉ lệ lấy mẫu lên và/hoặc lấy mẫu xuống theo mà lần lượt được báo hiệu theo bất kỳ cơ chế nào được đề cập ở trên, v.v..

Từ đó, mô tả giả định mã hóa bộ hữu hạn các hệ số lấy mẫu lên và/hoặc hệ số lấy mẫu xuống (cùng một hệ số được sử dụng trong cả kích thước X và Y), được chỉ báo thông qua từ mã. Từ mã đó có thể được mã hóa với độ dài thay đổi, ví dụ sử dụng mã Ext-Golomb chung cho các phần tử cú pháp nhất định trong tham số kỹ thuật mã hóa video như H.264 và H.265. Ánh xạ phù hợp của các giá trị đến các hệ số lấy mẫu lên và/hoặc lấy mẫu xuống có thể, ví dụ, theo Bảng 1:

BẢNG 1

Từ mã	Mã Ext-Golomb	Nguyên bản/độ phân giải mục tiêu
0	1	1/1
1	010	1/1,5 (lấy tỷ lệ lên 50%)
2	011	1,5/1 (lấy tỷ lệ xuống 50%)
3	00100	1/2 (lấy tỷ lệ lên 100%)

4	00101	2/1 (lấy tỷ lệ xuống 100%)
---	-------	----------------------------

Nhiều ảnh xạ tương tự có thể được tạo ra tùy theo nhu cầu của ứng dụng và khả năng của cơ chế tăng và giảm tỷ lệ khả dụng trong kỹ thuật hoặc tiêu chuẩn nén video. Bảng có thể được mở rộng đến nhiều giá trị hơn. Các giá trị cũng có thể được chỉ báo bằng cơ chế mã hóa entropy khác với mã Ext-Golomb, ví dụ như sử dụng mã nhị phân. Điều đó có thể có những lợi thế nhất định khi các hệ số lấy mẫu lại được quan tâm bên ngoài chính các công cụ xử lý video (trước hết bộ mã hóa và bộ giải mã), ví dụ bởi các MANE. Cần lưu ý rằng, đối với các tình huống không có sự thay đổi độ phân giải nào được yêu cầu, mã Ext-Golomb có thể được chọn mà ngắn; trong bảng trên, chỉ một bit duy nhất. Điều đó mà có thể có lợi thế về hiệu quả mã hóa so với việc sử dụng mã nhị phân cho trường hợp phổ biến nhất.

Số lượng mục nhập trong bảng, cũng như ngữ nghĩa của chúng, có thể được cấu hình đầy đủ hoặc một phần. Ví dụ, đường viền cơ bản của bảng có thể được chuyển tải trong bộ tham số “cao” như bộ tham số chuỗi hoặc bộ giải mã. Theo các phương án, một hoặc nhiều bảng như vậy có thể được xác định trong kỹ thuật hoặc tiêu chuẩn mã hóa video và có thể được chọn thông qua ví dụ như bộ giải mã hoặc bộ tham số chuỗi.

Dưới đây là mô tả cách hệ số lấy mẫu lên và/hoặc lấy mẫu xuống (thông tin ARC), được mã hóa như mô tả ở trên, có thể được bao gồm trong kỹ thuật mã hóa video hoặc cú pháp chuẩn. Các cân nhắc tương tự có thể áp dụng cho một, hoặc một số, từ mã kiểm soát bộ lọc lấy mẫu lên và/hoặc bộ lọc lấy mẫu xuống. Xem bên dưới để tham khảo khi lượng dữ liệu tương đối lớn được yêu cầu cho bộ lọc hoặc các cấu trúc dữ liệu khác.

Như được thể hiện trên FIG.5A, H.263 Phụ lục P bao gồm thông tin ARC 502 ở dạng bốn tọa độ cong vênh vào tiêu đề ảnh 501, cụ thể là trong phần mở rộng tiêu đề H.263 PLUSPTYPE 503. Đây có thể là lựa chọn thiết kế hợp lý khi

a) khả dụng tiêu đề ảnh và b) những thay đổi thường xuyên về thông tin ARC được dự kiến. Tuy nhiên, chi phí khi sử dụng báo hiệu kiểu H.263 có thể khá cao, và các hệ số tỷ lệ có thể không xác định giữa các ranh giới ảnh vì tiêu đề ảnh có thể mang tính chất theo thời gian.

Như được thể hiện trên FIG.5B, JVCET-M135-v1 bao gồm thông tin tham chiếu ARC 505 (chỉ số) nằm trong bộ tham số ảnh 504, lập chỉ số bảng 506 bao gồm độ phân giải mục tiêu mà lần lượt nằm bên trong bộ tham số chuỗi (507). Việc đặt độ phân giải có thể có trong bảng 506 trong bộ tham số chuỗi 507 có thể, theo tuyên bố bằng lời của các tác giả, được chứng minh bằng cách sử dụng SPS như một điểm thương lượng khả năng tương tác trong quá trình trao đổi năng lực. Độ phân giải có thể thay đổi, trong giới hạn được đặt bởi các giá trị trong bảng 506 từ ảnh này sang ảnh khác bằng cách tham khảo bộ tham số ảnh thích hợp 504.

Tham chiếu đến các FIG.5C-5E, các phương án sau có thể tồn tại để truyền tải thông tin ARC trong dòng bit video. Mỗi tùy chọn trong số đó đều có những ưu điểm nhất định so với các phương án được mô tả ở trên. Các phương án có thể có mặt đồng thời trong cùng một kỹ thuật hoặc tiêu chuẩn mã hóa video.

Theo các phương án, ví dụ như phương án được hiển thị trên FIG.5C, thông tin ARC 509 chẳng hạn như hệ số lấy mẫu lại (thu phóng) có thể xuất hiện trong tiêu đề lát, tiêu đề GOP, tiêu đề ô hoặc tiêu đề nhóm ô. FIG.5C minh họa phương án trong đó tiêu đề nhóm ô 508 được sử dụng. Điều này có thể phù hợp nếu thông tin ARC nhỏ, chẳng hạn như từ mã có độ dài thay đổi duy nhất $ue(v)$ hoặc từ mã có độ dài cố định của một vài bit, chẳng hạn như được hiển thị ở trên. Việc có trực tiếp thông tin ARC trong tiêu đề nhóm ô có lợi ích bổ sung của thông tin ARC có thể là khả năng cho ảnh con được đại diện bởi, ví dụ, nhóm ô đó, thay vì toàn bộ ảnh. Xem thêm bên dưới. Ngoài ra, ngay cả khi kỹ thuật nén video hoặc tiêu chuẩn chỉ hình dung ra toàn bộ thay đổi độ phân giải thích ứng ảnh (trái ngược với, ví dụ, thay đổi độ phân giải thích ứng dựa trên nhóm ô), đưa thông tin ARC vào tiêu đề nhóm ô để đưa thông tin đó vào Tiêu đề ảnh kiểu H.263 có một số lợi thế nhất định từ quan điểm về khả năng phục hồi lỗi.

Theo các phương án, ví dụ như phương án được thể hiện trên FIG.5D, bản thân thông tin ARC 512 có thể có trong bộ tham số thích hợp, chẳng hạn như bộ tham số ảnh, bộ tham số tiêu đề, bộ tham số ô, tập tham số thích ứng, v.v. FIG.5D minh họa phương án trong đó tập tham số thích ứng 511 được sử dụng. Phạm vi của bộ tham số đó có thể thuận lợi là không lớn hơn ảnh, ví dụ như một nhóm ô. Việc sử dụng thông tin ARC được ngầm định thông qua việc kích hoạt bộ tham số liên quan. Ví dụ, khi kỹ thuật mã hóa video hoặc tiêu chuẩn chỉ xem xét ARC dựa trên ảnh, thì bộ tham số ảnh hoặc tương đương có thể phù hợp.

Theo các phương án, ví dụ như phương án được thể hiện trên FIG.5E, thông tin tham chiếu ARC 513 có thể có trong tiêu đề Nhóm ô 514 hoặc cấu trúc dữ liệu tương tự. Thông tin tham chiếu đó 513 có thể tham chiếu đến bộ con thông tin ARC 515 khả dụng trong tập tham số 516 với phạm vi vượt ra ngoài ảnh duy nhất, ví dụ bộ tham số chuỗi, hoặc bộ tham số bộ giải mã.

Như được thể hiện trên FIG.6A, tiêu đề nhóm ô 601 như cấu trúc cú pháp mẫu của tiêu đề áp dụng cho (có thể là hình chữ nhật) một phần của ảnh có thể chứa một cách có điều kiện, độ dài thay đổi, phần tử cú pháp được mã hóa Exp-Golomb `dec_pic_size_idx` 602 (được mô tả bằng chữ in đậm). Sự hiện diện của phần tử cú pháp này trong tiêu đề nhóm ô có thể tạo cổng dựa trên việc sử dụng độ phân giải thích ứng 603 —trong đó, giá trị của cờ không được mô tả bằng chữ in đậm, mà có nghĩa là cờ đó hiện diện trong dòng bit tại điểm mà nó xảy ra trong sơ đồ cú pháp. Độ phân giải thích ứng có được sử dụng cho ảnh này hay không hoặc các phần của nó có thể được báo hiệu trong bất kỳ cấu trúc cú pháp bậc cao nào bên trong hoặc bên ngoài dòng bit. Trong ví dụ được hiển thị, nó được báo hiệu trong bộ tham số chuỗi như được nêu bên dưới.

Tham chiếu đến FIG.6B, được hiển thị cũng là đoạn trích của bộ tham số chuỗi 610. Phần tử cú pháp thứ nhất được hiển thị là `adaptive_pic_resolution_change_flag` 611. Khi đúng, cờ đó có thể chỉ báo việc sử dụng độ phân giải thích ứng mà, đến lượt có thể yêu cầu một số thông tin điều khiển nhất định. Trong ví dụ, thông tin điều khiển như vậy hiện diện 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 sử dụng độ phân giải thích ứng, trong ví dụ này, được mã hóa là độ phân giải đầu ra tính bằng đơn vị mẫu 613. Chữ số 613 đề cập đến cả `output_pic_width_in_luma_samples` và `output_pic_height_in_luma_samples`, mà chúng cùng có thể xác định độ phân giải của ảnh đầu ra. Ở những nơi khác trong kỹ thuật hoặc tiêu chuẩn mã hóa video, có thể xác định một số hạn chế nhất định đối với một trong hai giá trị có thể được định nghĩa. Ví dụ, định nghĩa cấp độ có thể giới hạn số lượng tổng số mẫu đầu ra, mà có thể là tích của giá trị của hai phần tử cú pháp đó. Ngoài ra, một số kỹ thuật hoặc tiêu chuẩn mã hóa video nhất định, hoặc các kỹ thuật hoặc tiêu chuẩn bên ngoài, chẳng hạn, ví dụ, như tiêu chuẩn hệ thống, có thể giới hạn phạm vi đánh số (ví dụ, một hoặc cả hai kích thước phải chia hết cho lũy thừa của 2 số), hoặc khía cạnh tỷ lệ (ví dụ, chiều rộng và chiều cao phải theo mối quan hệ như 4: 3 hoặc 16: 9). Những hạn chế như vậy có thể được đưa ra để tạo điều kiện thuận lợi cho việc triển khai phần cứng hoặc vì các lý do khác, và được biết đến nhiều trong lĩnh vực kỹ thuật này.

Trong một số ứng dụng nhất định, có thể khuyến khích bộ mã hóa hướng dẫn bộ giải mã sử dụng kích thước ảnh tham chiếu nhất định thay vì mặc nhiên cho rằng kích thước đó là kích thước ảnh đầu ra. Trong ví dụ này, phần tử cú pháp `reference_pic_size_present_flag` 614 tạo công sự hiện diện có điều kiện của kích thước ảnh tham chiếu 615 (một lần nữa, chữ số đề cập đến cả chiều rộng và chiều cao).

Cuối cùng, được hiển thị là bảng có thể giải mã chiều rộng và chiều cao ảnh. Bảng như vậy có thể được thể hiện, ví dụ, bằng chỉ báo bảng (`num_dec_pic_size_in_luma_samples_minus1`) 616. "Số trừ 1" có thể đề cập đến việc giải thích giá trị của phần tử cú pháp đó. Ví dụ, nếu giá trị được mã hóa là 0, thì một mục nhập bảng sẽ có mặt. Nếu giá trị là năm, sáu mục nhập bảng có mặt. Đối với mỗi "dòng" trong bảng, chiều rộng và chiều cao của ảnh được giải mã sau đó được đưa vào cú pháp 617.

Các mục trong bảng được trình bày 617 có thể được lập chỉ số bằng cách sử dụng phần tử cú pháp `dec_pic_size_idx` 602 trong tiêu đề nhóm ô, do đó cho phép các kích thước được giải mã khác nhau — có hiệu lực, hệ số thu phóng — trên mỗi nhóm ô.

Một số kỹ thuật hoặc tiêu chuẩn mã hóa video, ví dụ VP9, hỗ trợ khả năng biến đổi tỷ lệ không gian bằng cách triển khai một số hình thức lấy mẫu lại ảnh tham chiếu nhất định (được báo hiệu khá khác với đối tượng được bộc lộ) cùng với khả năng mở rộng theo thời gian, do đó, để cho phép khả năng biến đổi tỷ lệ không gian. Đặc biệt, một số ảnh tham chiếu nhất định có thể được lấy mẫu lên bằng cách sử dụng kỹ thuật kiểu ARC lên độ phân giải cao hơn để tạo thành cơ sở của lớp nâng cao không gian. Những ảnh được lấy mẫu lên đó có thể được tinh chỉnh, sử dụng các cơ chế dự đoán thông thường ở độ phân giải cao, do đó, để thêm chi tiết.

Các phương án được tham khảo ở đây có thể được sử dụng trong môi trường như vậy. Trong một số trường hợp nhất định, theo cùng phương án hoặc phương án khác, giá trị trong tiêu đề đơn vị NAL, ví dụ như trường Temporal ID, có thể được sử dụng để chỉ báo không chỉ lớp thời gian mà còn cả lớp không gian. Làm như vậy có thể có những lợi thế nhất định đối với một số thiết kế hệ thống nhất định; ví dụ, Đơn vị chuyển tiếp được chọn (SFU) hiện có được tạo và tối ưu hóa cho chuyển tiếp được chọn lớp theo thời gian dựa trên tiêu đề đơn vị NAL Giá trị ID theo thời gian có thể được sử dụng mà không cần sửa đổi, cho các môi trường có thể mở rộng. Để thực hiện điều đó, có thể có yêu cầu ánh xạ giữa kích thước ảnh được mã hóa và lớp thời gian được chỉ báo bởi trường ID theo thời gian trong tiêu đề đơn vị NAL.

Trong một số kỹ thuật mã hóa video, đơn vị Truy cập (AU) có thể tham chiếu đến (các) ảnh được mã hóa, (các) lát, (các) ô, (các) Đơn vị NAL, v.v., mà đã được ghi lại và tạo thành ảnh, lát, ô và/hoặc dòng bit đơn vị NAL tương ứng tại nấc nhất định trong thời gian. Nấc trong thời gian đó có thể là, ví dụ, thời gian kết cấu.

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

Theo các phương án, mối quan hệ chặt chẽ này có thể được nói lỏng trong đó đơn vị truy cập có thể bao gồm ảnh, lát hoặc ô với các giá trị POC khác nhau. Bằng cách cho phép các giá trị POC khác nhau trong AU, điều này trở nên có thể để sử dụng giá trị POC để xác định các ảnh/lát/ô có khả năng giải mã độc lập với thời gian trình bày giống hệt nhau. Điều đó, đến lượt, có thể cho phép hỗ trợ nhiều lớp có thể mở rộng mà không cần thay đổi báo hiệu lựa chọn ảnh tham chiếu, ví dụ báo hiệu bộ ảnh tham chiếu hoặc báo hiệu danh sách ảnh tham chiếu, như được mô tả chi tiết hơn bên dưới.

Điều này, tuy nhiên, vẫn mong muốn có thể xác định AU mà ảnh /lát/ô thuộc về, đối với các ảnh /lát/ô khác có các giá trị POC khác nhau, chỉ từ giá trị POC. Điều này có thể đạt được, như mô tả bên dưới.

Trong các phương án, số đếm đơn vị truy cập (AUC) có thể được báo hiệu trong cấu trúc cú pháp bậc cao, chẳng hạn như tiêu đề đơn vị NAL, tiêu đề lát, tiêu đề nhóm ô, thông báo SEI, bộ tham số hoặc dấu phân cách AU. Giá trị của AUC có thể được sử dụng để xác định đơn vị NAL, ảnh, lát, hoặc ô nào thuộc về AU nhất định. Giá trị của AUC có thể tương ứng với một phiên bản thời gian kết cấu riêng biệt. Giá trị AUC có thể bằng bội số của giá trị POC. Bằng cách bổ sung giá trị POC cho một giá trị nguyên, giá trị AUC có thể được tính toán. Trong một số trường hợp nhất định, các hoạt động phân chia có thể tạo ra một gánh nặng nhất định cho việc triển khai bộ giải mã. Trong những trường hợp như vậy, những

hạn chế nhỏ trong không gian đánh số của các giá trị AUC có thể cho phép thay thế hoạt động phân chia bằng các hoạt động dịch chuyển. Ví dụ, giá trị AUC có thể bằng giá trị Bit quan trọng nhất (MSB) của phạm vi giá trị POC.

Trong các phương án, giá trị của chu kỳ POC trên mỗi AU (`poc_cycle_au`) có thể được báo hiệu trong cấu trúc cú pháp bậc cao, chẳng hạn như tiêu đề đơn vị NAL, tiêu đề lát, tiêu đề nhóm ô, thông báo SEI, bộ tham số hoặc dấu phân cách AU. `Poc_cycle_au` có thể chỉ báo có bao nhiêu giá trị POC khác nhau và liên tiếp có thể được liên kết với cùng một AU. Ví dụ, nếu giá trị của `poc_cycle_au` bằng 4, thì ảnh, lát hoặc ô có giá trị POC bằng 0 - 3, bao gồm, có thể được liên kết với AU với giá trị AUC bằng 0, và ảnh, lát hoặc các ô có giá trị POC bằng 4 - 7, bao gồm, có thể được liên kết với AU với giá trị AUC bằng 1. Do đó, 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, giá trị của `poc_cyle_au` có thể được lấy từ thông tin, ví dụ: nằm trong tập tham số video (VPS), xác định số lượng lớp không gian hoặc SNR trong chuỗi video được mã hóa. Ví dụ về mối quan hệ khả dĩ như vậy được mô tả vắn tắt dưới đây. Trong khi dẫn xuất như được mô tả ở trên có thể tiết kiệm một vài bit trong VPS và do đó có thể cải thiện hiệu quả mã hóa, trong một số phương án, `poc_cycle_au` có thể được mã hóa rõ ràng theo cấu trúc cú pháp bậc cao thích hợp phân cấp bên dưới tập tham số video, do đó để có thể giảm thiểu `poc_cycle_au` cho một phần nhỏ nhất định của dòng bit chẳng hạn như ảnh. Tối ưu hóa này có thể tiết kiệm nhiều bit hơn mà có thể được tiết kiệm thông qua quá trình dẫn xuất ở trên vì giá trị POC và/hoặc giá trị của các phần tử cú pháp tham chiếu gián tiếp đến POC, có thể được mã hóa trong cấu trúc cú pháp cấp thấp.

Theo các phương án, FIG.8 hiển thị ví dụ về 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 tất cả ảnh/lát trong chuỗi video được mã hóa, và phần tử cú pháp của `slice_poc_cycle_au`, chỉ báo `poc_cycle_au` của lát hiện tại, trong tiêu đề lát. Nếu giá trị POC tăng đồng nhất theo AU, `vps_contant_poc_cycle_per_au` trong VPS có thể được đặt bằng 1 và

vps_poc_cycle_au có thể được báo hiệu trong VPS. Trong trường hợp này, slice_poc_cycle_au có thể không được báo hiệu rõ ràng, và giá trị của AUC cho mỗi AU có thể được tính bằng cách chia giá trị của POC cho vps_poc_cycle_au. Nếu giá trị POC không tăng đồng nhất theo AU, vps_contant_poc_cycle_per_au trong VPS có thể được đặt bằng 0. Trong trường hợp này, vps_access_unit_cnt có thể không được báo hiệu, trong khi slice_access_unit_cnt có thể được báo hiệu trong phần tiêu đề của 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 có thể được tính bằng cách chia giá trị của POC cho slice_poc_cycle_au.

FIG.9 hiển thị sơ đồ khối minh họa ví dụ của quá trình trên. Ví dụ, trong hoạt động S910, VPS (hoặc SPS) có thể được phân tích cú pháp, và có thể được xác định xem chu kỳ POC trên mỗi AU có là không đổi trong chuỗi video được mã hóa ở hoạt động S920 hay không. Nếu chu kỳ POC trên mỗi AU là không đổi (CÓ khi hoạt động S920), thì giá trị của số lượng đơn vị truy cập cho một đơn vị truy cập cụ thể có thể được tính toán từ poc_cycle_au được báo hiệu cho chuỗi video được mã hóa và giá trị POC của đơn vị truy cập cụ thể khi hoạt động S930. Nếu chu kỳ POC trên mỗi AU không phải là không đổi (NO tại hoạt động S920), thì giá trị của số lượng đơn vị truy cập cho một đơn vị truy cập cụ thể có thể được tính từ poc_cycle_au được báo hiệu ở cấp độ ảnh và giá trị POC của đơn vị truy cập cụ thể tại hoạt động S940. Khi hoạt động S950, VPS mới (hoặc SPS) có thể được phân tích cú pháp.

Theo các phương án, mặc dù giá trị POC của ảnh, lát, hoặc ô có thể khác nhau, ảnh, lát hoặc ô tương ứng với AU có cùng giá trị AUC có thể được liên kết với cùng nấc thời gian giải mã hoặc đầu ra. Do đó, không có bất kỳ sự phụ thuộc giữa phân tích cú pháp/giải mã giữa các ảnh, lát hoặc ô trong cùng AU, tất cả hoặc bộ con của ảnh, lát hoặc ô liên kết với cùng AU có thể được giải mã song song và có thể được đầu ra cùng một lúc..

Theo các phương án, mặc dù giá trị của POC cho ảnh, lát hoặc ô có thể khác nhau, ảnh, lát cắt hoặc ô tương ứng với AU có cùng giá trị AUC có thể được liên kết với cùng một kết cấu/nấc thời gian hiển thị ví dụ. Khi thời gian kết cấu

được chứa ở định dạng vùng chứa, mặc dù các ảnh tương ứng với các AU khác nhau, nếu các ảnh có cùng thời gian kết cấu, các ảnh có thể được hiển thị cùng lúc..

Theo các phương án, mỗi ảnh, lát hoặc ô có thể có cùng ký hiệu nhận dạng thời gian (`temporal_id`) trong cùng AU. Tất cả hoặc bộ con của ảnh, lát hoặc ô tương ứng với nấc thời gian có thể được liên kết với cùng lớp con thời gian. Theo các phương án, mỗi ảnh, lát hoặc ô có thể có cùng hoặc khác id lớp không gian (`layer_id`) trong cùng AU. Tất cả hoặc bộ con của ảnh, lát hoặc ô tương ứng với nấc thời gian có thể được liên kết với cùng hoặc lớp không gian khác.

FIG.7 hiển thị ví dụ về cấu trúc chuỗi video với sự kết hợp của `temporal_id`, `layer_id`, POC và giá trị AUC với sự thay đổi độ phân giải thích ứng. Trong ví dụ này, ảnh, lát hoặc ô trong AU đầu tiên có $AUC = 0$ có thể có `temporal_id = 0` và `layer_id = 0` hoặc 1, trong khi ảnh, lát hoặc ô trong AU thứ hai có $AUC = 1$ có thể có `temporal_id = 1` và `layer_id = 0` hoặc 1, tương ứng. Giá trị của POC được tăng lên 1 trên mỗi ảnh bất kể giá trị của `temporal_id` và `layer_id`. Trong ví dụ này, giá trị của `poc_cycle_au` có thể bằng 2. Theo các phương án, giá trị của `poc_cycle_au` có thể được đặt bằng số lớp (khả năng biến đổi tỷ lệ không gian). Trong ví dụ này, do đó, giá trị của POC được tăng lên 2, trong khi giá trị của AUC tăng lên 1.

Trong các phương án trên, tất cả hoặc bộ con của cấu trúc dự đoán ảnh liên đới hoặc liên lớp và chỉ báo ảnh tham chiếu có thể được hỗ trợ bằng cách sử dụng báo hiệu bộ ảnh tham chiếu (RPS) hiện có trong HEVC hoặc báo hiệu danh sách ảnh tham chiếu (RPL). Trong RPS hoặc RPL, ảnh tham chiếu đã chọn có thể được chỉ báo bằng cách báo hiệu giá trị của POC hoặc tam giác giá trị của POC giữa ảnh hiện tại và ảnh tham chiếu đã chọn. Theo các phương án, RPS và RPL có thể được sử dụng để chỉ báo cấu trúc dự đoán ảnh liên đới hoặc liên lớp mà không thay đổi báo hiệu, nhưng với các hạn chế sau. Nếu giá trị `temporal_id` của ảnh tham chiếu lớn hơn giá trị `temporal_id` của ảnh hiện tại, thì ảnh hiện tại có thể không sử dụng ảnh tham chiếu để bù chuyển động hoặc các dự đoán khác. Nếu giá trị `layer_id` của ảnh tham chiếu lớn hơn giá trị `layer_id` của ảnh hiện tại, ảnh

hiện tại có thể không sử dụng ảnh tham chiếu để bù chuyển động hoặc các dự đoán khác.

Trong các phương án, tỷ lệ vector chuyển động dựa trên sự chênh lệch POC đối với dự đoán vector chuyển động theo thời gian có thể bị vô hiệu hóa trên nhiều ảnh trong đơn vị truy cập. Do đó, mặc dù mỗi ảnh có thể có giá trị POC khác nhau trong đơn vị truy cập, vector chuyển động không được chia tỷ lệ và được sử dụng để dự đoán vector chuyển động theo thời gian trong đơn vị truy cập. Điều này là do ảnh tham chiếu có POC khác trong cùng AU được coi là ảnh tham chiếu có cùng nấc thời gian. Do đó, trong phương án này, hàm chia tỷ lệ vector chuyển động có thể trả về 1, khi ảnh tham chiếu thuộc AU được liên kết với ảnh hiện tại.

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

Trong các phương án, vector chuyển động có thể được chia tỷ lệ dựa trên chênh lệch AUC thay vì chênh lệch POC, để dự đoán vector chuyển động theo thời gian, đặc biệt khi `poc_cycle_au` có giá trị không đồng nhất (ví dụ: khi `vps_contant_poc_cycle_per_au == 0`). Mặt khác (ví dụ khi `vps_contant_poc_cycle_per_au == 1`), tỷ lệ vector chuyển động dựa trên sự chênh lệch AUC có thể giống với tỷ lệ vector chuyển động dựa trên sự chênh lệch POC.

Theo các phương án, khi vector chuyển động được chia tỷ lệ dựa trên sự chênh lệch AUC, vector chuyển động tham chiếu trong cùng AU (có cùng giá trị AUC) với ảnh hiện tại không được chia tỷ lệ dựa trên sự chênh lệch AUC và được sử dụng để dự đoán vector chuyển động mà không có chia tỷ lệ hoặc chia tỷ lệ dựa trên tỷ lệ độ phân giải không gian giữa ảnh hiện tại và ảnh tham chiếu.

Trong các phương án, giá trị AUC có thể được sử dụng để xác định ranh giới của AU và được sử dụng cho hoạt động của bộ giải mã tham chiếu giả định (HRD), mà cần định thời đầu vào và đầu ra với mức độ chi tiết của AU. Theo các

phương án, ảnh được giải mã có 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 xác định ảnh đầu ra.

Theo các phương án, ảnh có thể bao gồm một hoặc nhiều ảnh con. Mỗi ảnh con có thể bao phủ vùng cục bộ hoặc toàn bộ vùng của ảnh. Vùng được hỗ trợ bởi ảnh con có thể hoặc không bị chồng lên với vùng được hỗ trợ bởi ảnh con khác. Vùng được bao phủ bởi một hoặc nhiều ảnh con có thể bao phủ hoặc không bao phủ toàn bộ vùng của ảnh. Nếu ảnh bao gồm một ảnh con, vùng được hỗ trợ bởi ảnh con có thể giống với vùng được hỗ trợ bởi ảnh.

Theo các phương án, ảnh con có thể được mã hóa bằng phương pháp mã hóa tương tự như phương pháp mã hóa được sử dụng cho ảnh được mã hóa. Ảnh con có thể được mã hóa độc lập hoặc có thể được mã hóa tùy thuộc vào ảnh con khác hoặc ảnh được mã hóa. Ảnh con có thể có hoặc không có bất kỳ sự phụ thuộc phân tích cú pháp nào từ ảnh con khác hoặc ảnh được mã hóa.

Theo các 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 không gian khác. Ảnh con gốc có thể được lấy mẫu lại theo không gian (ví dụ như lấy mẫu lên hoặc lấy mẫu xuống), được mã hóa bằng các tham số độ phân giải không gian khác nhau và được chứa trong dòng bit tương ứng với một lớp.

Theo các phương án, ảnh con có (W, H), trong đó W chỉ báo chiều rộng của ảnh con và H chỉ báo chiều cao của ảnh con, tương ứng, có thể được mã hóa và chứa trong dòng bit được mã hóa tương ứng đến lớp 0, trong khi ảnh con được lấy mẫu lên (hoặc lấy mẫu xuống) từ ảnh con có độ phân giải không gian ban đầu, với $(W * S_w, k, H * S_h, k)$, có thể được mã hóa và chứa trong dòng bit được mã hóa tương ứng với lớp k, trong đó S_w, k, S_h, k chỉ báo tỷ lệ lấy mẫu lại, theo chiều ngang và chiều dọc. Nếu các giá trị S_w, k, S_h, k lớn hơn 1 thì việc lấy mẫu lại có thể là lấy mẫu lên. Trong khi đó, nếu các giá trị S_w, k, S_h, k nhỏ hơn 1, thì việc lấy mẫu lại có thể là lấy mẫu xuống.

Theo các phương án, ảnh con được mã hóa trong lớp có thể có chất lượng ảnh khác với của ảnh con được mã hóa được mã hóa ở lớp khác trong cùng

ảnh con hoặc ảnh con khác. Ví dụ, ảnh con thứ i trong lớp, n , có thể được mã hóa bằng tham số lượng tử hóa, Q_i, n , trong khi ảnh con thứ j trong lớp, m , có thể được mã hóa bằng tham số lượng tử hóa, $Q_{j,m}$.

Theo các phương án, ảnh con được mã hóa trong lớp có thể được giải mã độc lập, mà không có bất kỳ sự phụ thuộc phân tích cú pháp hoặc giải mã nào từ ảnh con được mã hóa trong lớp khác của cùng một vùng cục bộ. Lớp ảnh con, mà có thể được giải mã độc lập không cần tham chiếu đến lớp ảnh con khác của cùng vùng cục bộ, có thể là lớp ảnh con độc lập. Ảnh con được mã hóa trong lớp ảnh con độc lập có thể có hoặc không có sự phụ thuộc vào việc giải mã hoặc phân tích cú pháp từ một ảnh con đã được mã hóa trước đó trong cùng lớp ảnh con, nhưng ảnh con được mã hóa có thể không có bất kỳ sự phụ thuộc nào từ ảnh được mã hóa trong lớp ảnh con khác.

Theo các phương án, ảnh con được mã hóa trong lớp có thể được giải mã một cách phụ thuộc, với bất kỳ sự phụ thuộc phân tích cú pháp hoặc giải mã nào từ một ảnh con được mã hóa trong lớp khác của cùng vùng cục bộ. Lớp ảnh con, mà có thể được giải mã một cách phụ thuộc với tham chiếu đến lớp ảnh con khác của cùng vùng cục bộ, có thể 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 cùng ảnh con, ảnh con được mã hóa trước đó trong cùng lớp ảnh con, hoặc cả hai ảnh con tham chiếu.

Theo các phương án, ảnh con được mã hóa có thể bao gồm 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 hiện diện cho ảnh con được mã hóa. Giá trị của ký hiệu nhận dạng lớp (`layer_id`), mà có thể có trong tiêu đề đơn vị NAL hoặc một cấu trúc cú pháp bậc cao khác, của lớp ảnh con độc lập có thể bằng 0. Lớp ảnh con có `layer_id` bằng 0 có thể là lớp ảnh con gốc.

Theo các phương án, ảnh có thể bao gồm một hoặc nhiều ảnh con tiền cảnh và ảnh con hậu cảnh. Vùng được hỗ trợ bởi ảnh con hậu cảnh có thể bằng vùng của ảnh. Vùng được hỗ trợ bởi ảnh con tiền cảnh có thể bị chồng lên với vùng được hỗ trợ bởi ảnh con hậu cảnh. Ảnh con hậu cảnh có thể là lớp ảnh con gốc,

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

Theo các phương án, ảnh có thể bao gồm một hoặc nhiều ảnh con tiền cảnh có hoặc không có ảnh con hậu cảnh. Mỗi ảnh con có thể có lớp ảnh con gốc riêng 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 phải cơ sở có layer_id bằng a có thể tham chiếu đến lớp ảnh con không gốc với layer_id bằng b , trong đó a lớn hơn b .

Theo các phương án, ảnh có thể bao gồm một hoặc nhiều ảnh con tiền cảnh có hoặc không có ảnh con hậu cảnh. Mỗi ảnh con được mã hóa trong lớp ảnh con (cơ sở hoặc không phải cơ sở) có thể được tham chiếu bởi một hoặc nhiều ảnh con lớp không gốc thuộc cùng một ảnh con và một hoặc nhiều ảnh con lớp không gốc, mà không thuộc cùng một ảnh con.

Theo các phương án, ảnh có thể bao gồm một hoặc nhiều ảnh con tiền cảnh có hoặc không có ảnh con hậu cảnh. Ảnh con trong lớp a có thể được phân chia thêm thành nhiều ảnh con trong cùng một lớp. Một hoặc nhiều ảnh con được mã hóa trong lớp b có thể tham chiếu đến ảnh con được phân vùng trong lớp a .

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

Theo các phương án, CSPS có thể được mã hóa và chứa trong một hoặc nhiều lớp. CSPS có thể bao gồm một hoặc nhiều lớp CSPS. Việc giải mã một hoặc nhiều lớp CSPS tương ứng với CSPS có thể tái tạo chuỗi các ảnh con tương ứng với cùng vùng cục bộ.

Theo các phương án, số lớp CSPS tương ứng với CSPS có thể giống hoặc khác với số lớp CSPS tương ứng với CSPS khác.

Theo các phương án, lớp CSPS có thể có độ phân giải theo thời gian (ví dụ. tốc độ khung hình) khác với lớp CSPS khác. Chuỗi ảnh con gốc (không nén) có thể được lấy mẫu lại theo thời gian (ví dụ như 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 một lớp.

Theo các phương án, chuỗi ảnh con với tốc độ khung hình, F , có thể được mã hóa và chứa trong dòng bit được mã hóa tương ứng với lớp 0, trong khi chuỗi ảnh con được lấy mẫu lên (hoặc lấy mẫu xuống) theo thời gian từ chuỗi ảnh con ban đầu, với $F \cdot S_{t,k}$, có thể được mã hóa và 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 lấy mẫu lại theo thời gian có thể là chuyển đổi tăng tốc độ khung hình. Trong khi đó, nếu giá trị của $S_{t,k}$ nhỏ hơn 1, quá trình lấy lại mẫu theo thời gian có thể là chuyển đổi giảm tốc độ khung hình.

Theo các phương án, khi ảnh con có lớp CSPS a được tham chiếu bởi ảnh con có lớp CSPS b để bù chuyển động hoặc bất kỳ dự đoán liên lớp nào, nếu độ phân giải không gian của lớp CSPS a khác với độ phân giải không gian của lớp CSPS b , các điểm ảnh được giải mã trong lớp CSPS a được lấy mẫu lại và được sử dụng để tham khảo. Quá trình lấy mẫu lại có thể sử dụng bộ lọc lấy mẫu lên hoặc bộ lọc lấy mẫu xuống.

FIG.10 hiển thị dòng video mẫu bao gồm CSPS video hậu cảnh với `layer_id` bằng 0 và nhiều lớp CSPS tiền cảnh. Mặc dù ảnh con được mã hóa có thể bao gồm một hoặc nhiều lớp CSPS, vùng hậu cảnh, mà không thuộc bất kỳ lớp CSPS tiền cảnh nào, có thể bao gồm lớp gốc. Lớp cơ sở có thể chứa vùng hậu cảnh và các vùng tiền cảnh, trong khi lớp CSPS nâng cao có thể chứa vùng tiền cảnh. Lớp CSPS nâng cao có thể có chất lượng ảnh tốt hơn so với lớp gốc, tại cùng khu vực. Lớp CSPS nâng cao có thể tham chiếu các điểm ảnh được tái cấu trúc và các vector chuyển động của lớp gốc, tương ứng với cùng vùng.

Trong các phương án, dòng bit video tương ứng với lớp gốc được chứa trong rãnh, trong khi các lớp CSPS tương ứng với mỗi ảnh con được chứa trong rãnh riêng biệt, trong tệp video.

Trong các phương án, dòng bit video tương ứng với lớp gốc được chứa trong một rãnh, trong khi các lớp CSPS có cùng layer_id được chứa trong rãnh riêng biệt. Trong ví dụ này, rãnh tương ứng với lớp k bao gồm các lớp CSPS tương ứng với lớp k , chỉ.

In embodiments, each CSPS layer of each sub-picture is stored in a separate track. Each track may or may not have any parsing or decoding dependency from one or more other tracks.

Trong các phương án, mỗi rãnh có thể chứa các dòng bit tương ứng với lớp i đến lớp j của tất cả các lớp CSPS hoặc một bộ con các ảnh con, trong đó $0 < i \leq j \leq k$, k là lớp cao nhất của CSPS.

Trong các phương án, ảnh bao gồm một hoặc nhiều dữ liệu phương tiện được liên kết bao gồm bản đồ độ sâu, bản đồ alpha, dữ liệu hình học 3D, bản đồ tỷ lệ sử dụng, v.v. Dữ liệu phương tiện được định thời liên quan như vậy có thể được chia thành một hoặc nhiều luồng dữ liệu con, mỗi luồng dữ liệu trong số đó tương ứng với một ảnh con.

FIG.11 hiển thị ví dụ về hội nghị truyền hình dựa trên phương pháp ảnh con nhiều lớp. Trong dòng video, một dòng video lớp gốc tương ứng với ảnh hậu cảnh và một hoặc nhiều luồng video lớp nâng cao tương ứng với các ảnh con tiền cảnh được chứa. Mỗi dòng video của lớp nâng cao có thể tương ứng với lớp CSPS. Trong màn hình, ảnh tương ứng với lớp gốc được hiển thị theo mặc định. Nó chứa một hoặc nhiều ảnh trong ảnh của người dùng (PIP). Khi người dùng cụ thể được lựa chọn bởi sự điều khiển của khách hàng, lớp CSPS nâng cao tương ứng với người dùng đã chọn có thể được giải mã và hiển thị với chất lượng nâng cao hoặc độ phân giải không gian.

FIG.12 hiển thị sơ đồ khối minh họa ví dụ về quy trình trên. Ví dụ, trong hoạt động S1210, dòng bit video có nhiều lớp có thể được giải mã. Trong hoạt động S1220, vùng hậu cảnh và một hoặc nhiều ảnh con tiền cảnh có thể được xác định. Tại hoạt động S1230, nó có thể được xác định xem một vùng ảnh con cụ thể, chẳng hạn như một trong các ảnh con tiền cảnh, có được chọn hay không. Nếu một vùng ảnh con cụ thể được chọn (CÓ ở thao tác S1240), ảnh con nâng

cao có thể được giải mã và hiển thị. Nếu vùng ảnh con cụ thể không được chọn (KHÔNG ở thao tác S1240), vùng hậu cảnh có thể được giải mã và hiển thị

Trong các phương án, hộp trung gian mạng (ví dụ bộ định tuyến) có thể chọn một bộ con các lớp để gửi cho người dùng tùy thuộc vào băng thông của nó. Tổ chức ảnh/ảnh con có thể được sử dụng để điều chỉnh băng thông. Ví dụ, nếu người dùng không có băng thông, bộ định tuyến sẽ tách các lớp hoặc chọn một số ảnh con do tầm quan trọng của chúng hoặc dựa trên thiết lập đã sử dụng và điều này có thể được thực hiện động để thông qua băng thông.

FIG.13 cho thấy phương án liên quan đến trường hợp sử dụng của video 360. Khi ảnh 360 độ hình cầu, ví dụ như ảnh 1310, được chiếu lên ảnh phẳng, ảnh 360 độ hình chiếu có thể được phân chia thành nhiều ảnh con làm lớp gốc. Ví dụ, nhiều ảnh con có thể bao gồm ảnh con phía sau, ảnh con trên cùng, ảnh con bên phải, ảnh con bên trái, ảnh con phía trước và ảnh con phía dưới. Lớp nâng cao của ảnh con cụ thể, ví dụ như một ảnh con phía trước, có thể được mã hóa và truyền đến máy khách. Bộ giải mã có thể giải mã cả lớp gốc bao gồm tất cả các ảnh con và lớp nâng cao của ảnh con đã chọn. Khi khung nhìn hiện tại giống hệt với ảnh con đã chọn, ảnh được hiển thị có thể có chất lượng cao hơn với ảnh con được giải mã bằng lớp nâng cao. Nếu không, ảnh được giải mã với lớp gốc có thể được hiển thị, với chất lượng thấp hơn.

Trong các phương án, bất kỳ thông tin bố cục nào để hiển thị đều có thể có trong tệp, dưới dạng thông tin bổ sung (chẳng hạn như thông báo SEI hoặc siêu dữ liệu). Một hoặc nhiều ảnh con đã được giải mã có thể được di chuyển và hiển thị tùy 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 một đài truyền hình, hoặc có thể được tái cấu trúc bởi một thực thể mạng hoặc một máy chủ đám mây hoặc có thể được xác định bởi cài đặt tùy chỉnh của người dùng.

Theo các phương án, khi ảnh đầu vào được chia thành một hoặc nhiều (các) 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 một giá trị `layer_id` duy nhất. Đối với mỗi lớp độc lập, kích thước ảnh con và thông tin vị trí có thể được báo

hiệu. Ví dụ, kích thước ảnh (chiều rộng, chiều cao), thông tin độ lệch chuẩn của góc trên bên trái (x_offset , y_offset). FIG.14 hiển thị ví dụ về bố cục của các ảnh con đã chia, kích thước và thông tin vị trí của ảnh con đó và cấu trúc dự đoán ảnh tương ứng của nó. Thông tin bố cục bao gồm (các) kích thước ảnh con và (các) vị trí ảnh con có thể được báo hiệu trong cấu trúc cú pháp bậc cao, chẳng hạn như (các) bộ tham số, tiêu đề của lát hoặc nhóm lát hoặc thông báo SEI.

Theo các phương án, mỗi ảnh con tương ứng với lớp độc lập có thể có giá trị POC duy nhất của nó trong AU. Khi ảnh tham chiếu trong số các ảnh được lưu trữ trong DPB được chỉ báo bằng cách sử dụng (các) phần 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, để chỉ ra cấu trúc dự đoán (liên lớp), $layer_id$ có thể không được sử dụng và giá trị POC (tam giác) có thể được sử dụng.

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

Theo các phương án, FIG.15 cho thấy trường hợp mở rộng của FIG.14. Khi ảnh đầu vào được chia thành nhiều vùng con (ví dụ: bốn), mỗi vùng cục bộ có thể được mã hóa bằng một hoặc nhiều lớp. Trong trường hợp, số lớp độc lập có thể bằng số vùng con, và một hoặc nhiều lớp có thể tương ứng với một vùng con. Do đó, mỗi vùng con có thể được mã hóa bằng một hoặc nhiều (các) lớp độc lập và không hoặc nhiều (các) lớp phụ thuộc.

Theo các phương án, trên Fig. 15, ảnh đầu vào có thể được chia thành bốn vùng con. Như là ví dụ, vùng con trên cùng bên phải có thể được mã hóa thành hai lớp, đó là lớp 1 và lớp 4, trong khi vùng con dưới cùng bên phải có thể được mã hóa thành hai lớp, đó 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 lớp 5 có thể tham chiếu lớp 3 để bù chuyển động.

Trong các phương án, lọc trong vòng (chẳng hạn như lọc gỡ khối, lọc trong vòng thích ứng, định dạng lại, lọc song phương hoặc bất kỳ lọc dựa trên học tập sâu nào) qua ranh giới lớp có thể bị vô hiệu hóa (tùy chọn).

Trong các phương án, dự đoán bù chuyển động hoặc bản sao trong khối qua ranh giới lớp có thể bị vô hiệu hóa (tùy chọn).

Trong các phương án, phần đệm ranh giới cho dự đoán bù chuyển động hoặc lọc trong vòng ở ranh giới của ảnh con có thể được xử lý theo tùy chọn. Cờ chỉ báo vùng đệm ranh giới có được xử lý hay không có thể được báo hiệu trong cấu trúc cú pháp bậc cao, chẳng hạn như (các) bộ tham số (VPS, SPS, PPS hoặc APS), tiêu đề nhóm lát hoặc ô hoặc thông báo SEI.

Theo các phương án, thông tin bố cục của (các) vùng con (hoặc (các) ảnh con) có thể được báo hiệu trong VPS hoặc SPS. FIG.16A hiển thị ví dụ về các phần tử cú pháp trong VPS và FIG.16B hiển thị ví dụ về các phần tử cú pháp trong SPS. Trong ví dụ này, `vps_sub_picture_dividing_flag` được báo hiệu trong VPS. Cờ có thể chỉ báo liệu (các) ảnh đầu vào có được chia thành nhiều vùng con hay không. Khi giá trị của `vps_sub_picture_dividing_flag` bằng 0, (các) 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 thước ảnh đầu vào có thể bằng kích thước ảnh được mã hóa (`pic_width_in_luma_samples`, `pic_height_in_luma_samples`), được báo hiệu trong SPS. Khi giá trị của `vps_sub_picture_dividing_flag` bằng 1, (các) 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. Giá trị của `vps_full_pic_width_in_luma_samples` và `vps_full_pic_height_in_luma_samples` có thể tương ứng bằng chiều rộng và chiều cao của (các) ảnh đầu vào.

Trong các 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 để kết cấu và hiển thị.

Trong các 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, tương ứng với (a) (các) lớp cụ thể. Trong trường hợp này, kích thước ảnh được mã hóa (`pic_width_in_luma_samples`, `pic_height_in_luma_samples`) được báo hiệu trong SPS có thể bằng chiều rộng và chiều cao của vùng con tương ứng với một lớp cụ thể. Ngoài ra, vị trí (`pic_offset_x`, `pic_offset_y`) của góc trên bên trái của vùng con có thể được báo hiệu trong SPS.

Trong các 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 để kết cấu và hiển thị.

Trong các phương án, thông tin bố cục (kích thước và vị trí) của tất cả hoặc bộ con (các) vùng con của (một) (các) ảnh đầu vào, thông tin phụ thuộc giữa (các) lớp có thể được báo hiệu trong bộ tham số hoặc tin nhắn SEI. FIG.17 hiển thị ví dụ về các phần tử cú pháp để chỉ báo thông tin về cách bố trí của các vùng con, sự phụ thuộc giữa các lớp, và mối quan hệ giữa vùng con với một hoặc nhiều lớp. Trong ví dụ này, phần tử cú pháp `num_sub_region` chỉ báo số vùng con (hình chữ nhật) trong chuỗi video được mã hóa hiện tại. phần tử cú pháp `num_layers` chỉ báo số 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 bất kỳ vùng con nào được mã hóa thành một lớp, giá trị của `num_layers` có thể bằng giá trị của `num_sub_region`. Khi một hoặc nhiều vùng con được mã hóa thành nhiều 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 phụ thuộc từ lớp thứ j đến lớp thứ i. `num_layers_for_region[i]` chỉ báo số lớp được liên kết với vùng con thứ i. `sub_region_layer_id[i][j]` chỉ báo `layer_id` của lớp thứ j được liên kết 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, tương ứng. `sub_region_width[i]` và `sub_region_height[i]` chỉ báo chiều rộng và chiều cao của vùng con thứ i, tương ứng.

Trong các phương án, một hoặc nhiều phần tử cú pháp chỉ định tập lớp đầu ra để biểu thị một trong nhiều lớp sẽ được xuất ra có hoặc không có thông tin cấp tầng tệp tin có thể được báo hiệu trong cấu trúc cú pháp bậc cao, ví dụ: Tin nhắn VPS, DPS, SPS, PPS, APS hoặc SEI. Đề cập đến FIG.18, phần tử cú pháp `num_output_layer_sets` chỉ báo số lượng lớp đầu ra (OLS) trong chuỗi video được mã hóa đề cập đến VPS có thể được báo hiệu trong VPS. Đối với mỗi tập lớp đầu ra, `output_layer_flag` có thể được báo hiệu bằng số lượng lớp đầu ra.

Trong các 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, một hoặc nhiều phần tử cú pháp chỉ định thông tin cấp độ cấu hình cho mỗi tập lớp đầu ra có thể được báo hiệu trong cấu trúc cú pháp bậc cao, ví dụ. Tin nhắn VPS, DPS, SPS, PPS, APS hoặc SEI. Vẫn đề cập đến FIG.18, phần tử cú pháp `num_profile_tile_level` chỉ báo số lượng thông tin cấp tầng tệp tin trên mỗi OLS trong chuỗi video được mã hóa đề cập đến VPS có thể được báo hiệu trong VPS. Đối với mỗi tập lớp đầu ra, bộ các phần tử cú pháp cho thông tin cấp độ cấu hình hoặc chỉ số chỉ báo thông tin cấp tầng tệp tin cụ thể giữa các mục nhập trong thông tin cấp tầng tệp tin có thể được báo hiệu bằng số lượng lớp đầu ra.

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

Theo các phương án, đề cập đến FIG.19, các phần tử cú pháp `num_profile_tile_level` và/hoặc `num_output_layer_sets` có thể được báo hiệu khi số lượng lớp tối đa lớn hơn 1 (`vps_max_layers_minus1 > 0`).

Theo các phương án, đề cập đến FIG.19, phần tử cú pháp `vps_output_layers_mode [i]` chỉ báo chế độ báo hiệu lớp đầu ra cho tập lớp đầu ra thứ *i* có thể có trong VPS.

Trong các phương án, `vps_output_layers_mode [i]` bằng 0 xác định rằng chỉ lớp cao nhất được xuất ra với tập lớp đầu ra thứ *i*. `vps_output_layer_mode [i]`

bảng 1 xác định rằng tất cả các lớp được xuất ra với tập lớp đầu ra thứ i . $vps_output_layer_mode[i]$ bảng 2 xác định rằng các lớp được đầu ra là các lớp có $vps_output_layer_flag[i][j]$ bằng 1 với tập lớp đầu ra thứ i . Nhiều giá trị hơn có thể được bảo lưu.

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

Theo các phương án, đề cập đến FIG.19, cờ $vps_ptl_signal_flag[i]$ có thể hiện diện cho tập lớp đầu ra thứ i . Phụ thuộc vào giá trị của $vps_ptl_signal_flag[i]$, thông tin cấp tầng tệp tin cho tập lớp đầu ra thứ i có thể được báo hiệu hoặc không.

Theo các phương án, đề cập đến FIG.20, số lượng subpicture, $max_subpics_minus1$, trong CVS hiện tại có thể được báo hiệu trong cấu trúc cú pháp bậc cao, ví dụ. Tin nhắn VPS, DPS, SPS, PPS, APS hoặc SEI.

Theo các phương án, đề cập đến FIG.20, ký hiệu nhận dạng ảnh con, $sub_pic_id[i]$, cho ảnh con thứ i có thể được báo hiệu, khi số lượng ảnh con lớn hơn 1 ($max_subpics_minus1 > 0$).

Trong các phương án, một hoặc nhiều phần tử cú pháp chỉ báo ký hiệu nhận dạng ảnh con thuộc từng lớp của mỗi tập lớp đầu ra có thể được báo hiệu trong VPS. Đề cập đến FIG.20, $sub_pic_id_layer[i][j][k]$, mà chỉ báo ảnh con thứ k có trong lớp thứ j của tập lớp đầu ra thứ i . Với thông tin này, bộ giải mã có thể nhận ra ảnh con nào có thể được giải mã và xuất cho từng lớp của tập lớp đầu ra cụ thể.

Trong các phương án, tiêu đề ảnh (PH) có thể là cấu trúc cú pháp chứa các phần tử cú pháp áp dụng cho tất cả các lát của ảnh được mã hóa. Đơn vị ảnh (PU) có thể là bộ các đơn vị NAL được liên kết với nhau theo quy tắc phân loại cụ thể, liên tiếp theo thứ tự giải mã, và chứa chính xác một ảnh được mã hóa. PU có thể chứa một tiêu đề ảnh (PH) và một hoặc nhiều đơn vị NAL lớp mã hóa video (VCL) tạo thành một ảnh được mã hóa.

Trong các phương án, tập tham số thích ứng (APS) có thể là cấu trúc cú pháp chứa các phần tử cú pháp có thể áp dụng cho không hoặc nhiều lát được xác định bởi không hoặc nhiều phần tử cú pháp được tìm thấy trong tiêu đề lát.

Trong các phương án, `adapt_parameter_set_id` được báo hiệu (ví dụ như u(5)) trong APS có thể cung cấp ký hiệu nhận dạng cho APS để tham chiếu bởi các phần tử cú pháp khác. Khi `ps_params_type` bằng `ALF_APS` hoặc `SCALING_APS`, giá trị của `adapt_parameter_set_id` có thể nằm trong khoảng từ 0 đến 7, bao gồm. Khi `aps_params_type` bằng `LMCS_APS`, giá trị của `adapt_parameter_set_id` có thể nằm trong khoảng từ 0 đến 3, bao gồm.

Trong các phương án, `adapt_parameter_set_id` được báo hiệu là `ue(v)` trong APS có thể cung cấp ký hiệu nhận dạng cho APS để tham chiếu bởi các phần tử cú pháp khác. Khi `ps_params_type` bằng `ALF_APS` hoặc `SCALING_APS`, giá trị của `adapt_parameter_set_id` có thể nằm trong khoảng từ 0 đến $7 * (\text{số lớp tối đa trong CVS hiện tại})$. Khi `aps_params_type` bằng `LMCS_APS`, giá trị của `adapt_parameter_set_id` có thể nằm trong khoảng từ 0 đến $3 * (\text{số lớp tối đa trong CVS hiện tại})$, bao gồm.

Trong các phương án, mỗi APS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, được bao gồm trong ít nhất một AU với ký hiệu nhận dạng theo thời gian (ví dụ: `TemporalId`) nhỏ hơn hoặc bằng ký hiệu nhận dạng theo thời gian (ví dụ: `TemporalId`) của đơn vị NAL lát được mã hóa mà đề cập đến nó + hoặc được đề xuất thông qua các phương tiện bên ngoài. Khi đơn vị lớp trừu tượng mạng APS được bao gồm trong AU, giá trị của ký hiệu nhận dạng theo thời gian (ví dụ., `TemporalId`) của đơn vị lớp trừu tượng mạng APS có thể bằng giá trị của ký hiệu nhận dạng theo thời gian (ví dụ: `TemporalId`) của AU bao gồm cả đơn vị lớp trừu tượng mạng APS.

Trong các phương án, mỗi APS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, được bao gồm trong ít nhất một AU với ký hiệu nhận dạng theo thời gian (ví dụ., `TemporalId`) bằng 0 hoặc được đề xuất thông qua các phương tiện bên ngoài. Khi đơn vị lớp trừu tượng mạng APS được bao

gồm trong AU, giá trị của ký hiệu nhận dạng theo thời gian (ví dụ: TemporalId) của đơn vị lớp trừu tượng mạng APS có thể bằng 0.

Trong các phương án, APS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, bao gồm trong ít nhất một AU với ký hiệu nhận dạng theo thời gian (ví dụ., TemporalId) bằng với ký hiệu nhận dạng theo thời gian (ví dụ., TemporalId) của đơn vị lớp trừu tượng mạng APS trong CVS, mà chứa một hoặc nhiều PH hoặc một hoặc nhiều đơn vị NAL lát được mã hóa đề cập đến APS, hoặc được đề xuất thông qua các phương tiện bên ngoài.

Trong các phương án, APS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, được bao gồm trong ít nhất một AU với ký hiệu nhận dạng theo thời gian (ví dụ., TemporalId) bằng 0 trong CVS, mà chứa một hoặc nhiều PH hoặc một hoặc nhiều đơn vị NAL lát được mã hóa tham chiếu đến APS, hoặc được đề xuất thông qua các phương tiện bên ngoài.

Trong các phương án, khi cờ, no_temporal_sublayer_switching_flag được báo hiệu trong DPS, VPS, SPS hoặc PPS, giá trị ký hiệu nhận dạng theo thời gian (ví dụ., TemporalId) của APS tham chiếu đến bộ tham số chứa cờ bằng 1 có thể bằng 0, trong khi giá trị ký hiệu nhận dạng theo thời gian (ví dụ., TemporalId) của APS tham chiếu đến bộ tham số chứa cờ bằng 1 có thể bằng hoặc lớn hơn giá trị ký hiệu nhận dạng theo thời gian (ví dụ., TemporalId) của bộ tham số.

Theo các phương án, mỗi APS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, bao gồm trong ít nhất một AU với ký hiệu nhận dạng theo thời gian (ví dụ., TemporalId) nhỏ hơn hoặc bằng ký hiệu nhận dạng theo thời gian (ví dụ: TemporalId) của đơn vị NAL lát được mã hóa (hoặc đơn vị lớp trừu tượng mạng PH) đề cập đến nó hoặc được đề xuất thông qua các phương tiện bên ngoài. Khi đơn vị lớp trừu tượng mạng APS được bao gồm trong AU trước AU có chứa đơn vị NAL lát được mã hóa tham chiếu đến APS, đơn vị lớp trừu tượng mạng VCL cho phép chuyển đổi lên lớp theo thời gian hoặc đơn vị lớp trừu tượng mạng VCL với nal_unit_type bằng STSA_NUT, mà điều này chỉ báo rằng ảnh trong đơn vị lớp trừu tượng mạng VCL có thể là ảnh truy cập lớp con theo thời gian (STSA) theo từng bước, có thể không xuất hiện sau đơn vị lớp trừu

tượng mạng APS và trước đơn vị NAL lát được mã hóa đề cập đến APS. FIG.21 hiển thị ví dụ về ràng buộc này.

Theo các phương án, đơn vị lớp trừu tượng mạng APS và đơn vị NAL lát được mã hóa (và đơn vị lớp trừu tượng mạng PH của nó) đề cập đến APS có thể được bao gồm trong cùng AU.

Theo các phương án, đơn vị lớp trừu tượng mạng APS và đơn vị lớp trừu tượng mạng STSA có thể được bao gồm trong cùng AU, mà có thể có trước đơn vị NAL lát được mã hóa (và đơn vị lớp trừu tượng mạng PH của nó) đề cập đến APS.

Theo các phương án, đơn vị lớp trừu tượng mạng STSA, đơn vị lớp trừu tượng mạng APS và đơn vị NAL lát được mã hóa (và đơn vị lớp trừu tượng mạng PH của nó) đề cập đến APS có thể có trong cùng AU.

Trong các phương án, giá trị ký hiệu nhận dạng theo thời gian (ví dụ: TemporalId) của đơn vị lớp trừu tượng mạng VCL chứa APS có thể bằng giá trị ký hiệu nhận dạng theo thời gian (ví dụ: TemporalId) của đơn vị lớp trừu tượng mạng STSA trước đó.

Trong các phương án, giá trị số đếm thứ tự ảnh (POC) của đơn vị lớp trừu tượng mạng APS có thể bằng hoặc lớn hơn giá trị POC của đơn vị lớp trừu tượng mạng STSA. Trong FIG.21, giá trị POC M của APS có thể bằng hoặc lớn hơn giá trị POC L của đơn vị lớp trừu tượng mạng STSA.

Theo các phương án, giá trị số đếm thứ tự ảnh (POC) của lát được mã hóa hoặc đơn vị lớp trừu tượng mạng PH, mà dùng để chỉ đơn vị lớp trừu tượng mạng APS, có thể bằng hoặc lớn hơn giá trị POC của đơn vị lớp trừu tượng mạng APS được tham chiếu. Trong FIG.21, giá trị POC M của đơn vị lớp trừu tượng mạng VCL tham chiếu đến APS có thể bằng hoặc lớn hơn giá trị POC L của đơn vị lớp trừu tượng mạng APS.

Theo phương án, APS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PH hoặc một hoặc nhiều đơn vị NAL lát được mã hóa, bao gồm trong ít nhất một PU với nuh_layer_id bằng giá trị thấp

nhất giá trị `nuh_layer_id` của đơn vị NAL lát được mã hóa tham chiếu đến đơn vị lớp trừu tượng mạng APS trong CVS, mà chứa một hoặc nhiều PH hoặc một hoặc nhiều đơn vị NAL lát được mã hóa tham chiếu đến APS hoặc được đề xuất thông qua các phương tiện bên ngoài.

Theo phương án, APS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PH hoặc một hoặc nhiều đơn vị NAL lát được mã hóa, bao gồm trong ít nhất một PU với `TemporalId` bằng `TemporalId` của đơn vị lớp trừu tượng mạng APS và `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của đơn vị NAL lát được mã hóa tham chiếu đến đơn vị lớp trừu tượng mạng APS trong CVS, chứa một hoặc nhiều PH hoặc một hoặc nhiều đơn vị NAL lát được mã hóa tham chiếu đến PPS, hoặc được đề xuất thông qua các phương tiện bên ngoài.

Theo phương án, APS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PH hoặc một hoặc nhiều đơn vị NAL lát được mã hóa, được bao gồm trong ít nhất một PU với `TemporalId` bằng 0 và `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của đơn vị NAL lát được mã hóa tham chiếu đến đơn vị lớp trừu tượng mạng APS trong CVS, mà chứa một hoặc nhiều PH hoặc một hoặc nhiều đơn vị NAL lát được mã hóa tham chiếu đến PPS, hoặc được đề xuất thông qua các phương tiện bên ngoài.

Theo phương án, APS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PH hoặc một hoặc nhiều đơn vị NAL lát được mã hóa, được bao gồm trong ít nhất một PU với `TemporalId` bằng 0 và `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của đơn vị NAL lát được mã hóa tham chiếu đến đơn vị lớp trừu tượng mạng APS trong CVS, mà chứa một hoặc nhiều PH hoặc một hoặc nhiều đơn vị NAL lát được mã hóa tham chiếu đến PPS, hoặc được đề xuất thông qua các phương tiện bên ngoài.

Trong các phương án, tất cả các đơn vị lớp trừu tượng mạng APS có giá trị cụ thể của `adapt_parameter_set_id` và giá trị cụ thể của `aps_params_type` trong PU, bất kể chúng là tiền tổ hay hậu tổ các đơn vị lớp trừu tượng mạng APS, đều có thể có cùng nội dung.

Trong các phương án, bất kể giá trị `nuh_layer_id` là gì, các đơn vị lớp trừu tượng mạng APS có thể chia sẻ cùng không gian giá trị của `adapt_parameter_set_id` và `aps_params_type`.

Trong các phương án, giá trị `nuh_layer_id` của đơn vị lớp trừu tượng mạng APS có thể bằng giá trị `nuh_layer_id` thấp nhất của đơn vị NAL lát được mã hóa tham chiếu đến đơn vị NAL tham chiếu đến đơn vị lớp trừu tượng mạng APS.

Theo phương án, khi APS có `nuh_layer_id` bằng m được tham chiếu bởi một hoặc nhiều đơn vị NAL lát được mã hóa với `nuh_layer_id` bằng n , lớp có `nuh_layer_id` bằng m có thể giống với lớp có `nuh_layer_id` bằng n hoặc lớp tham chiếu (trực tiếp hoặc gián tiếp) của lớp có `nuh_layer_id` bằng m .

FIG.22 là lưu đồ là quy trình ví dụ 2200 để giải mã dòng bit video được mã hóa. Trong một số triển khai, một hoặc nhiều khối quy trình của FIG.22 có thể được thực hiện bởi bộ giải mã 210. Trong một số triển khai, một hoặc nhiều khối quá trình của FIG.22 có thể được thực hiện bởi thiết bị khác hoặc một nhóm thiết bị tách biệt với hoặc bao gồm cả bộ giải mã 210, chẳng hạn như bộ mã hóa 203.

Như trong FIG.22, quy trình 2200 có thể bao gồm việc thu nhận từ luồng video được mã hóa một chuỗi video được mã hóa bao gồm đơn vị ảnh tương ứng với ảnh được mã hóa (khối 2210).

Như được hiển thị thêm trong FIG.22, quy trình 2200 có thể bao gồm việc thu nhận đơn vị lớp trừu tượng mạng PH có trong đơn vị ảnh (khối 2220).

Như được hiển thị thêm trong FIG.22, quy trình 2200 có thể bao gồm việc thu nhận ít nhất một đơn vị lớp trừu tượng mạng VCL có trong đơn vị ảnh (khối 2230).

Như được hiển thị thêm trong FIG.22, quy trình 2200 có thể bao gồm giải mã ảnh được mã hóa dựa trên đơn vị lớp trừu tượng mạng PH, ít nhất một đơn vị lớp trừu tượng mạng VCL, và tập tham số thích ứng (APS) có trong đơn vị lớp trừu tượng mạng APS thu nhận từ chuỗi video được mã hóa, trong đó APS NAL thiết bị khả dụng cho ít nhất một bộ xử lý trước ít nhất một đơn vị lớp trừu tượng mạng VCL (khối 2240).

Như được hiển thị thêm trong FIG.22, quy trình 2200 có thể bao gồm việc xuất ra ảnh đã được giải mã (khởi 2250).

Trong các phương án, đơn vị lớp trừu tượng mạng APS khả dụng cho quá trình giải mã trước khi được tham chiếu bởi một hoặc nhiều tiêu đề ảnh (PH) hoặc một hoặc nhiều đơn vị lớp trừu tượng mạng lát cắt được mã hóa (NAL), được bao gồm trong ít nhất một đơn vị dự đoán (PU) với `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của nhiều hoặc nhiều đơn vị NAL lát được mã hóa tham chiếu đến đơn vị lớp trừu tượng mạng APS trong CVS, chứa một hoặc nhiều PH hoặc một hoặc nhiều đơn vị NAL lát được mã hóa đề cập đến APS.

Theo các phương án, ký hiệu nhận dạng theo thời gian của ít nhất một đơn vị lớp trừu tượng mạng VCL lớn hơn hoặc bằng ký hiệu nhận dạng theo thời gian của đơn vị lớp trừu tượng mạng APS.

Theo các phương án, ký hiệu nhận dạng theo thời gian của đơn vị lớp trừu tượng mạng APS có thể bằng 0.

Theo các phương án, POC của ít nhất một đơn vị lớp trừu tượng mạng VCL có thể lớn hơn hoặc bằng POC của đơn vị lớp trừu tượng mạng APS.

Theo các phương án, ký hiệu nhận dạng lớp của đơn vị lớp trừu tượng mạng PH và ký hiệu nhận dạng lớp của ít nhất một đơn vị lớp trừu tượng mạng VCL lớn hơn hoặc bằng ký hiệu nhận dạng lớp của đơn vị lớp trừu tượng mạng APS.

Trong các phương án, đơn vị lớp trừu tượng mạng PH, ít nhất một đơn vị lớp trừu tượng mạng VCL và đơn vị lớp trừu tượng mạng APS được bao gồm trong một đơn vị truy cập duy nhất

Trong các phương án, chuỗi video được mã hóa còn bao gồm đơn vị lớp trừu tượng mạng STSA tương ứng với ảnh STSA, và đơn vị lớp trừu tượng mạng STSA không nằm giữa đơn vị lớp trừu tượng mạng APS và ít nhất một đơn vị lớp trừu tượng mạng VCL.

Theo các phương án, ít nhất một đơn vị lớp trừu tượng mạng VCL, đơn vị lớp trừu tượng mạng APS, và đơn vị lớp trừu tượng mạng STSA có thể được bao gồm trong một đơn vị truy cập duy nhất.

Theo các phương án, ký hiệu nhận dạng theo thời gian của đơn vị lớp trừu tượng mạng APS có thể lớn hơn hoặc bằng ký hiệu nhận dạng theo thời gian của đơn vị lớp trừu tượng mạng STSA.

Theo các phương án, số đếm thứ tự ảnh (POC) của đơn vị lớp trừu tượng mạng APS có thể lớn hơn hoặc bằng POC của đơn vị lớp trừu tượng mạng STSA.

Mặc dù FIG.22 hiển thị các khối mẫu của quy trình 2200, trong một số triển khai, quy trình 2200 có thể bao gồm các khối bổ sung, ít khối hơn, các khối khác, hoặc các khối được sắp xếp khác với những khối được mô tả trong FIG.22 Ngoài ra, hoặc theo cách khác, hai hoặc nhiều khối của quy trình 2200 có thể được thực hiện song song.

Hơn nữa, các phương pháp được đề xuất có thể được thực hiện bằng mạch xử lý (ví dụ., một hoặc nhiều bộ xử lý hoặc một hoặc nhiều mạch tích hợp). Trong một ví dụ, một hoặc nhiều bộ xử lý thực thi chương trình được lưu trữ trong một phương tiện đọc được bằng máy tính không nhất thời để thực hiện một hoặc nhiều phương pháp được đề xuất.

Các kỹ thuật được mô tả ở trên có thể được thực hiện dưới dạng phần mềm máy tính sử dụng các lệnh đọc được bằng máy tính và được lưu trữ vật lý trong một hoặc nhiều phương tiện đọc được bằng máy tính. Ví dụ, FIG.23 hiển thị hệ thống máy tính 2300 thích hợp để triển khai các phương án nhất định của đối tượng được bộc lộ.

Phần mềm máy tính có thể được mã hóa bằng bất kỳ mã máy hoặc ngôn ngữ máy tính thích hợp nào, có thể phải tuân theo các cơ chế lắp ráp, biên dịch, liên kết hoặc tương tự để tạo mã bao gồm các lệnh mà có thể được thực thi trực tiếp hoặc thông qua diễn giải, vi mã thực thi, và những thứ tương tự, bởi đơn vị xử lý trung tâm máy tính (CPU), Đơn vị xử lý đồ họa (GPU), và những thứ tương tự.

Các lệnh có thể được thực thi trên nhiều loại máy tính hoặc thành phần khác nhau của chúng, bao gồm, ví dụ, máy tính cá nhân, máy tính bảng, máy chủ, điện thoại thông minh, thiết bị chơi game, thiết bị kết nối Internet và những thứ tương tự.

Các thành phần được thể hiện trên FIG.23 đối với hệ thống máy tính 2300 là ví dụ về bản chất và không nhằm mục đích đề xuất bất kỳ giới hạn nào đối với phạm vi sử dụng hoặc chức năng của phần mềm máy tính triển khai các phương án của sáng chế. Cũng như cấu hình của các thành phần không nên được hiểu là có bất kỳ sự phụ thuộc hoặc yêu cầu nào liên quan đến bất kỳ một hoặc sự kết hợp các thành phần nào được minh họa trong phương án được lấy làm ví dụ về hệ thống máy tính 2300.

Hệ thống máy tính 2300 có thể bao gồm một số thiết bị đầu vào giao diện người. Thiết bị đầu vào giao diện người như vậy có thể đáp ứng đầu vào của một hoặc nhiều người dùng thông qua, ví dụ, đầu vào xúc giác (chẳng hạn như: gõ phím, vuốt, chuyển động căng tay dữ liệu), đầu vào âm thanh (chẳng hạn như: giọng nói, vỗ tay), đầu vào bằng ảnh (chẳng hạn như: cử chỉ), đầu vào khứu giác (không được mô tả). Các thiết bị giao diện người cũng có thể được sử dụng để chụp một số phương tiện nhất định không nhất thiết liên quan trực tiếp đến đầu vào có ý thức của con người, chẳng hạn như âm thanh (chẳng hạn như: giọng nói, âm nhạc, âm thanh xung quanh), ảnh (chẳng hạn như: ảnh được quét, ảnh chụp thu nhận từ máy ảnh chụp ảnh tĩnh), video (chẳng hạn như video hai chiều, video ba chiều bao gồm cả video lập thể).

Thiết bị giao diện người đầu vào có thể bao gồm một hoặc nhiều (chỉ một trong số mỗi thiết bị được mô tả): bàn phím 2301, chuột 2302, bàn di chuột 2303, màn hình cảm ứng 2310 và bộ điều hợp đồ họa đi kèm 2350, căng tay dữ liệu, cần điều khiển 2305, micrô 2306, máy quét 2307, máy ảnh 2308.

Hệ thống máy tính 2300 cũng có thể bao gồm một số thiết bị đầu ra giao diện người nhất định. Các thiết bị đầu ra giao diện người như vậy có thể kích thích các giác quan của một hoặc nhiều người dùng thông qua, ví dụ, đầu ra xúc giác, âm thanh, ánh sáng và khứu giác/vị giác. Các thiết bị đầu ra giao diện người như vậy có thể bao gồm các thiết bị đầu ra xúc giác (ví dụ phản hồi xúc giác bằng màn hình cảm ứng 2310, căng tay dữ liệu hoặc cần điều khiển 2305, nhưng cũng có thể có các thiết bị phản hồi xúc giác không đóng vai trò là thiết bị đầu vào), thiết bị đầu ra âm thanh (chẳng hạn như: loa 2309, tai nghe (không được mô tả)), thiết

bị đầu ra ảnh (chẳng hạn như màn hình 2310 để bao gồm màn hình ống tia âm cực (CRT), màn hình tinh thể lỏng (LCD), màn hình plasma, điốt phát sáng hữu cơ (OLED) màn hình, mỗi màn hình có hoặc không có khả năng đầu vào màn hình cảm ứng, mỗi màn hình có hoặc không có khả năng phản hồi xúc giác — mà một số trong số đó có thể có khả năng xuất ra đầu ra ảnh hai chiều hoặc đầu ra nhiều hơn ba chiều thông qua các phương tiện như đầu ra ảnh lập thể; thực tế ảo kính (không được mô tả), màn hình nổi ba chiều và bề khối (không được mô tả)), và máy in (không được mô tả).

Hệ thống máy tính 2300 cũng có thể bao gồm các thiết bị lưu trữ người có thể truy cập và phương tiện liên kết của chúng như phương tiện quang bao gồm CD/DVD ROM/RW 2320 với CD/DVD hoặc phương tiện tương tự 2321, ổ cứng 2322, ổ cứng di động hoặc trạng thái rắn ổ đĩa 2323, phương tiện từ tính cũ như băng và đĩa mềm (không được mô tả), các thiết bị dựa trên ROM/ASIC/PLD chuyên dụng như thiết bị bảo mật (không được mô tả), và những thứ tương tự.

Những người có hiểu biết trung bình trong lĩnh vực kỹ thuật này cũng nên hiểu rằng thuật ngữ "phương tiện đọc được bằng máy tính " được sử dụng liên quan đến đối tượng được bộc lộ hiện tại không bao gồm phương tiện truyền dẫn, sóng mang, hoặc các tín hiệu theo thời gian khác.

Hệ thống máy tính 2300 cũng có thể bao gồm (các) giao diện với một hoặc nhiều mạng truyền thông 955. Mạng có thể ví dụ là không dây, có dây, quang. Các mạng còn có thể là mạng cục bộ, diện rộng, đô thị, xe cộ và công nghiệp, thời gian thực, khả năng chịu trễ, v.v. Ví dụ về mạng bao gồm mạng cục bộ như Ethernet, mạng LAN không dây, mạng di động để bao gồm các hệ thống toàn cầu cho truyền thông di động (GSM), thế hệ thứ ba (3G), thế hệ thứ tư (4G), thế hệ thứ năm (5G), Tiến hóa dài hạn (LTE), và tương tự, truyền hình có dây hoặc mạng kỹ thuật số diện rộng không dây để bao gồm truyền hình cáp, truyền hình vệ tinh và truyền hình phát sóng mặt đất, phương tiện và công nghiệp để bao gồm CANBus, v.v. Một số mạng nhất định thường yêu cầu bộ điều hợp giao diện mạng bên ngoài 954 được gắn với một số cổng dữ liệu mục đích chung nhất định hoặc bus ngoại vi 949 (chẳng hạn như, ví dụ cổng bus nối tiếp chung (USB) của hệ

thống máy tính 2300; những mạng khác thường được tích hợp vào lõi của hệ thống máy tính 2300 bằng cách gắn vào bus hệ thống như mô tả bên dưới (ví dụ: giao diện Ethernet vào hệ thống máy tính PC hoặc giao diện mạng di động vào hệ thống máy tính điện thoại thông minh). Ví dụ, mạng 2355 có thể được kết nối với bus ngoại vi 2349 bằng cách sử dụng giao diện mạng 2354. Sử dụng bất kỳ mạng nào trong số các mạng này, hệ thống máy tính 2300 có thể giao tiếp với các thực thể khác. Giao tiếp như vậy có thể là đơn hướng, chỉ nhận (ví dụ: truyền hình quảng bá), chỉ gửi một hướng (ví dụ CANbus tới một số CANbus nhất định thiết bị), hoặc hai chiều, ví dụ đối với các hệ thống máy tính khác sử dụng mạng số cục bộ hoặc diện rộng. Một số giao thức và ngăn xếp giao thức nhất định có thể được sử dụng trên mỗi networks và giao diện mạng 954 như được mô tả ở trên.

Các thiết bị giao diện người nói trên, thiết bị lưu trữ con người có thể truy cập, và giao diện mạng có thể được gắn vào lõi 2340 của hệ thống máy tính 2300.

Lõi 2340 có thể bao gồm một hoặc nhiều Đơn vị xử lý trung tâm (CPU) 2341, Đơn vị xử lý đồ họa (GPU) 2342, các đơn vị xử lý lập trình chuyên biệt ở dạng Khu vực cổng lập trình trường (FPGA) 2343, bộ tăng tốc phần cứng 2344 cho một số tác vụ nhất định, và kể từ đó trở đi. Các thiết bị này, cùng với bộ nhớ chỉ đọc (ROM) 2345, bộ nhớ truy cập ngẫu nhiên (RAM) 2346, bộ nhớ chung bên trong chẳng hạn như ổ cứng không người dùng có thể truy cập nội bộ, ổ cứng thể rắn (SSD), và các loại tương tự 2347, có thể được kết nối thông qua bus hệ thống 2348. Trong một số hệ thống máy tính, bus hệ thống 2348 có thể được truy cập dưới dạng một hoặc nhiều phích cắm vật lý để kích hoạt các phần mở rộng bằng CPU, GPU bổ sung, v.v. Các thiết bị ngoại vi có thể được gắn trực tiếp vào bus hệ thống 2348 của lõi hoặc thông qua bus ngoại vi 2349. Kiến trúc cho bus ngoại vi bao gồm kết nối thành phần ngoại vi (PCI), USB, v.v..

CPU 2341, GPU 2342, FPGA 2343 và bộ tăng tốc 2344 có thể thực thi một số lệnh nhất định, kết hợp với nhau, có thể tạo nên mã máy tính nói trên. Mã máy tính đó có thể được lưu trữ trong ROM 2345 hoặc RAM 2346. Dữ liệu chuyển tiếp cũng có thể được lưu trữ trong RAM 2346, trong khi dữ liệu vĩnh viễn có thể được lưu trữ, chẳng hạn như trong bộ nhớ chung 2347. Lưu trữ nhanh và truy xuất

vào bất kỳ thiết bị nhớ nào có thể được kích hoạt thông qua việc sử dụng bộ nhớ đệm, có thể liên kết chặt chẽ với một hoặc nhiều CPU 2341, GPU 2342, bộ nhớ chung 2347, ROM 2345, RAM 2346, v.v..

Phương tiện máy tính có thể đọc có thể có mã máy tính trên đó để thực hiện các hoạt động khác nhau do máy tính thực hiện. Phương tiện và mã máy tính có thể là những phương tiện được thiết kế và xây dựng đặc biệt cho các mục đích của sáng chế, hoặc chúng có thể thuộc loại nổi tiếng và sẵn có đối với những người có hiểu biết trung bình trong lĩnh vực kỹ thuật phần mềm máy tính..

Như là ví dụ và không phải giới hạn, hệ thống máy tính có kiến trúc 2300, và cụ thể là lõi 2340 có thể cung cấp chức năng do (các) bộ xử lý (bao gồm CPU, GPU, FPGA, bộ tăng tốc và những thứ tương tự) thực thi phần mềm được thể hiện trong một hoặc nhiều phương tiện hữu hình, đọc được bằng máy tính được. Phương tiện đọc được bằng máy tính được như vậy có thể là phương tiện được liên kết với bộ nhớ chung mà người dùng có thể truy cập như đã giới thiệu ở trên, cũng như bộ lưu trữ nhất định của lõi 2340 không mang tính chất chuyển tiếp, chẳng hạn như bộ lưu trữ nội bộ lõi 2347 hoặc ROM 2345. Phần mềm triển khai các phương án khác nhau của sáng chế có thể được lưu trữ trong các thiết bị như vậy và được thực thi bởi lõi 2340. Phương tiện đọc được bằng máy tính có thể bao gồm một hoặc nhiều thiết bị hoặc chip bộ nhớ, tùy theo nhu cầu cụ thể. Phần mềm có thể khiến lõi 2340 và cụ thể là các bộ xử lý trong đó (bao gồm CPU, GPU, FPGA, v.v.) thực thi các quy trình cụ thể hoặc các phần cụ thể của các quy trình cụ thể được mô tả ở đây, bao gồm xác định cấu trúc dữ liệu được lưu trữ trong RAM 2346 và sửa đổi cấu trúc dữ liệu đó theo các quy trình được xác định bởi phần mềm. Ngoài ra hoặc để thay thế, hệ thống máy tính có thể cung cấp chức năng là kết quả của logic được nổi cứng hoặc được thể hiện bằng cách khác trong một mạch (ví dụ: máy gia tốc 2344), mà có thể hoạt động thay thế hoặc cùng với phần mềm để thực hiện các quy trình cụ thể hoặc các bộ phận cụ thể của các quy trình cụ thể được mô tả ở đây. Tham chiếu đến phần mềm có thể bao gồm logic, và ngược lại, khi thích hợp. Tham chiếu đến phương tiện đọc được bằng máy tính có thể bao gồm một mạch (chẳng hạn như mạch tích hợp (IC)) lưu trữ phần mềm

để thực thi, mạch thể hiện logic để thực thi hoặc cả hai, nếu thích hợp. Sáng chế đề cập đến bất kỳ sự kết hợp thích hợp nào giữa phần cứng và phần mềm.

Mặc dù sáng chế này đã mô tả một số phương án được lấy làm ví dụ, nhưng có những thay thế, hoán vị, và các tương đương thay thế khác nhau, mà nằm trong phạm vi của sáng chế. Do đó, sẽ được đánh giá cao rằng những người có hiểu biết trung bình trong lĩnh vực kỹ thuật này sẽ có thể đưa ra nhiều hệ thống và phương pháp mà, mặc dù không được trình bày hoặc mô tả rõ ràng ở đây, nhưng thể hiện các nguyên tắc của sáng chế và do đó trong bản chất và phạm vi của chúng.

Đơn sáng chế này yêu cầu hưởng quyền ưu tiên từ Đơn xin cấp bằng sáng chế tạm thời Hoa Kỳ số 62/954,096, được nộp vào ngày 27 tháng 12 năm 2019, và Đơn xin cấp bằng sáng chế Hoa Kỳ số 17/038,541, được nộp vào ngày 30 tháng 09 năm 2020, toàn bộ nội dung của chúng được tích hợp ở đây bằng cách tham chiếu.

YÊU CẦU BẢO HỘ

1. Phương pháp giải mã dòng bit video được mã hóa bằng cách sử dụng ít nhất một bộ xử lý, phương pháp này bao gồm:

thu nhận từ dòng bit video được mã hóa chuỗi video được mã hóa (CVS) bao gồm đơn vị ảnh tương ứng với ảnh được mã hóa;

thu nhận đơn vị lớp trù tượng mạng (NAL) tiêu đề ảnh (PH) được bao gồm trong đơn vị ảnh;

thu nhận ít nhất một đơn vị NAL lớp mã hóa video (VCL) được bao gồm trong đơn vị ảnh;

giải mã ảnh được mã hóa dựa trên đơn vị lớp trù tượng mạng PH, ít nhất một đơn vị lớp trù tượng mạng VCL, và tập tham số thích ứng (APS) được bao gồm trong đơn vị lớp trù tượng mạng APS được thu nhận từ chuỗi video được mã hóa; và

xuất ra ảnh đã được giải mã,

trong đó đơn vị lớp trù tượng mạng APS khả dụng cho ít nhất một bộ xử lý trước ít nhất một đơn vị lớp trù tượng mạng VCL,

trong đó đơn vị lớp trù tượng mạng APS khả dụng đối với ít nhất một bộ xử lý trước khi được tham chiếu bởi một hoặc nhiều thông tin tiêu đề ảnh (PH) hoặc một hoặc nhiều đơn vị NAL lát được mã hóa được bao gồm trong CVS, và

trong đó đơn vị lớp trù tượng mạng APS được bao gồm trong ít nhất một đơn vị dự đoán (PU) mà có giá trị `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của một hoặc nhiều đơn vị NAL lát được mã hóa mà tham chiếu tới đơn vị lớp trù tượng mạng APS.

2. Phương pháp theo điểm 1, trong đó ký hiệu nhận dạng theo thời gian của ít nhất một đơn vị lớp trù tượng mạng VCL lớn hơn hoặc bằng ký hiệu nhận dạng theo thời gian của đơn vị lớp trù tượng mạng APS.

3. Phương pháp theo điểm 1, trong đó số đếm thứ tự ảnh (POC) của ít nhất một đơn vị lớp trù tượng mạng VCL lớn hơn hoặc bằng POC của đơn vị lớp trù tượng mạng APS.

4. Phương pháp theo điểm 1, trong đó ký hiệu nhận dạng lớp của đơn vị lớp trù tượng mạng PH và ký hiệu nhận dạng lớp của ít nhất một đơn vị lớp trù tượng

mạng VCL lớn hơn hoặc bằng ký hiệu nhận dạng lớp của đơn vị lớp trừu tượng mạng APS.

5. Phương pháp theo điểm 1, trong đó đơn vị lớp trừu tượng mạng PH, ít nhất một đơn vị lớp trừu tượng mạng VCL, và đơn vị lớp trừu tượng mạng APS được bao gồm trong thiết bị truy cập đơn nhất.

6. Phương pháp theo điểm 1, trong đó chuỗi video được mã hóa còn bao gồm đơn vị NAL truy cập lớp con theo thời gian (STSA) theo từng bước tương ứng với ảnh STSA, và

trong đó đơn vị lớp trừu tượng mạng STSA không nằm giữa đơn vị lớp trừu tượng mạng APS và ít nhất một đơn vị lớp trừu tượng mạng VCL.

7. Phương pháp theo điểm 6, trong đó đơn vị lớp trừu tượng mạng PH, ít nhất một đơn vị lớp trừu tượng mạng VCL, đơn vị lớp trừu tượng mạng APS, và đơn vị lớp trừu tượng mạng STSA được bao gồm trong thiết bị truy cập đơn nhất.

8. Phương pháp theo điểm 6, trong đó ký hiệu nhận dạng theo thời gian của đơn vị lớp trừu tượng mạng APS lớn hơn hoặc bằng ký hiệu nhận dạng theo thời gian của đơn vị lớp trừu tượng mạng STSA.

9. Phương pháp theo điểm 6, trong đó số đếm thứ tự ảnh (POC) của đơn vị lớp trừu tượng mạng APS lớn hơn hoặc bằng POC của đơn vị lớp trừu tượng mạng STSA.

10. Thiết bị giải mã dòng bit video được mã hóa, thiết bị này bao gồm:

ít nhất một bộ nhớ được cấu hình để lưu mã chương trình; và

ít nhất một bộ xử lý được cấu hình để đọc mã chương trình và hoạt động theo hướng dẫn của mã chương trình, mã chương trình bao gồm:

mã thu nhận thứ nhất được cấu hình để khiến ít nhất một bộ xử lý thu nhận từ dòng bit video được mã hóa chuỗi video được mã hóa bao gồm đơn vị ảnh tương ứng với ảnh được mã hóa;

mã thu nhận thứ hai được cấu hình để khiến ít nhất một bộ xử lý thu nhận đơn vị lớp trừu tượng mạng (NAL) tiêu đề ảnh (PH) được bao gồm trong đơn vị ảnh;

mã thu nhận thứ ba được cấu hình để khiến ít nhất một bộ xử lý thu nhận ít nhất một đơn vị NAL lớp mã hóa video (VCL) được bao gồm trong đơn vị ảnh;

mã giải mã được cấu hình để khiến ít nhất một bộ xử lý giải mã ảnh được mã hóa dựa trên đơn vị lớp trừu tượng mạng PH, ít nhất một đơn vị lớp trừu tượng mạng VCL, và tập tham số thích ứng (APS) có trong đơn vị lớp trừu tượng mạng APS thu nhận từ chuỗi video được mã hóa; và

mã đầu ra được cấu hình để khiến ít nhất một bộ xử lý xuất ra ảnh đã được giải mã,

trong đó đơn vị lớp trừu tượng mạng APS khả dụng cho ít nhất một bộ xử lý trước ít nhất một đơn vị lớp trừu tượng mạng VCL,

trong đó đơn vị lớp trừu tượng mạng APS khả dụng đối với ít nhất một bộ xử lý trước khi được tham chiếu bởi một hoặc nhiều thông tin tiêu đề ảnh (PH) hoặc một hoặc nhiều đơn vị NAL lát được mã hóa được bao gồm trong CVS, và

trong đó đơn vị lớp trừu tượng mạng APS được bao gồm trong ít nhất một đơn vị dự đoán (PU) mà có giá trị `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của một hoặc nhiều đơn vị NAL lát được mã hóa mà tham chiếu tới đơn vị lớp trừu tượng mạng APS.

11. Thiết bị theo điểm 10, trong đó ký hiệu nhận dạng theo thời gian của ít nhất một đơn vị lớp trừu tượng mạng VCL lớn hơn hoặc bằng ký hiệu nhận dạng theo thời gian của đơn vị lớp trừu tượng mạng APS.

12. Thiết bị theo điểm 10, trong đó số đếm thứ tự ảnh (POC) của ít nhất một đơn vị lớp trừu tượng mạng VCL lớn hơn hoặc bằng POC của đơn vị lớp trừu tượng mạng APS.

13. Thiết bị theo điểm 10, trong đó ký hiệu nhận dạng lớp của đơn vị lớp trừu tượng mạng PH và ký hiệu nhận dạng lớp của ít nhất một đơn vị lớp trừu tượng mạng VCL lớn hơn hoặc bằng ký hiệu nhận dạng lớp của đơn vị lớp trừu tượng mạng APS.

14. Thiết bị theo điểm 10, trong đó đơn vị lớp trừu tượng mạng PH, ít nhất một đơn vị lớp trừu tượng mạng VCL, và đơn vị lớp trừu tượng mạng APS được bao gồm trong thiết bị truy cập đơn nhất.

15. Thiết bị theo điểm 10, trong đó chuỗi video được mã hóa còn bao gồm đơn vị NAL truy cập lớp con theo thời gian theo từng bước (STSA) tương ứng với ảnh STSA, và

trong đó đơn vị lớp trù tượng mạng STSA không nằm giữa đơn vị lớp trù tượng mạng APS và ít nhất một đơn vị lớp trù tượng mạng VCL.

16. Thiết bị theo điểm 15, trong đó đơn vị lớp trù tượng mạng PH, ít nhất một đơn vị lớp trù tượng mạng VCL, đơn vị lớp trù tượng mạng APS, và đơn vị lớp trù tượng mạng STSA được bao gồm trong thiết bị truy cập đơn nhất.

17. Thiết bị theo điểm 15, trong đó ký hiệu nhận dạng theo thời gian của đơn vị lớp trù tượng mạng APS lớn hơn hoặc bằng ký hiệu nhận dạng theo thời gian của đơn vị lớp trù tượng mạng STSA.

18. Thiết bị theo điểm 15, trong đó số đếm thứ tự ảnh (POC) của đơn vị lớp trù tượng mạng APS lớn hơn hoặc bằng POC của đơn vị lớp trù tượng mạng STSA.

19. Phương tiện bất biến đọc được bằng máy tính lưu trữ các lệnh, các lệnh này bao gồm: một hoặc nhiều lệnh, khi được thực thi bởi một hoặc nhiều bộ xử lý của thiết bị để giải mã dòng bit video được mã hóa, khiến một hoặc nhiều bộ xử lý:

thu nhận từ dòng bit video được mã hóa chuỗi video được mã hóa bao gồm đơn vị ảnh tương ứng với ảnh được mã hóa;

thu nhận đơn vị NAL tiêu đề ảnh (PH) được bao gồm trong đơn vị ảnh;

thu nhận ít nhất một đơn vị NAL lớp mã hóa video (VCL) được bao gồm trong đơn vị ảnh;

giải mã ảnh được mã hóa dựa trên đơn vị lớp trù tượng mạng PH, ít nhất một đơn vị lớp trù tượng mạng VCL và tập tham số thích ứng (APS) có trong đơn vị lớp trù tượng mạng APS thu nhận từ chuỗi video được mã hóa; và

xuất ra ảnh đã được giải mã,

trong đó đơn vị lớp trù tượng mạng APS khả dụng cho ít nhất một bộ xử lý trước ít nhất một đơn vị lớp trù tượng mạng VCL,

trong đó đơn vị lớp trù tượng mạng APS khả dụng đối với ít nhất một bộ xử lý trước khi được tham chiếu bởi một hoặc nhiều thông tin tiêu đề ảnh (PH) hoặc một hoặc nhiều đơn vị NAL lát được mã hóa được bao gồm trong CVS, và

trong đó đơn vị lớp trù tượng mạng APS được bao gồm trong ít nhất một đơn vị dự đoán (PU) mà có giá trị nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của một hoặc nhiều đơn vị NAL lát được mã hóa mà tham chiếu tới đơn vị lớp trù tượng mạng APS.

1/24

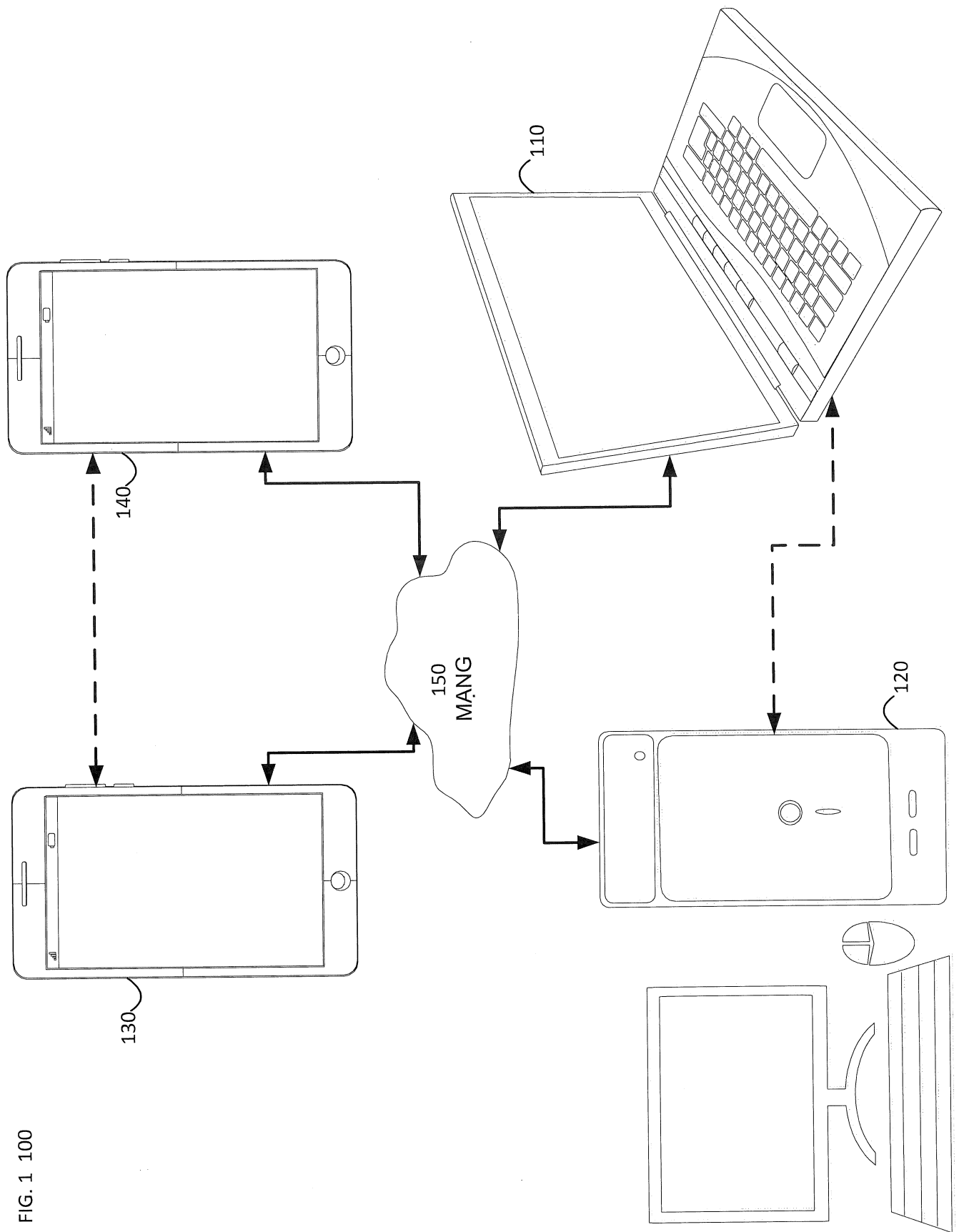
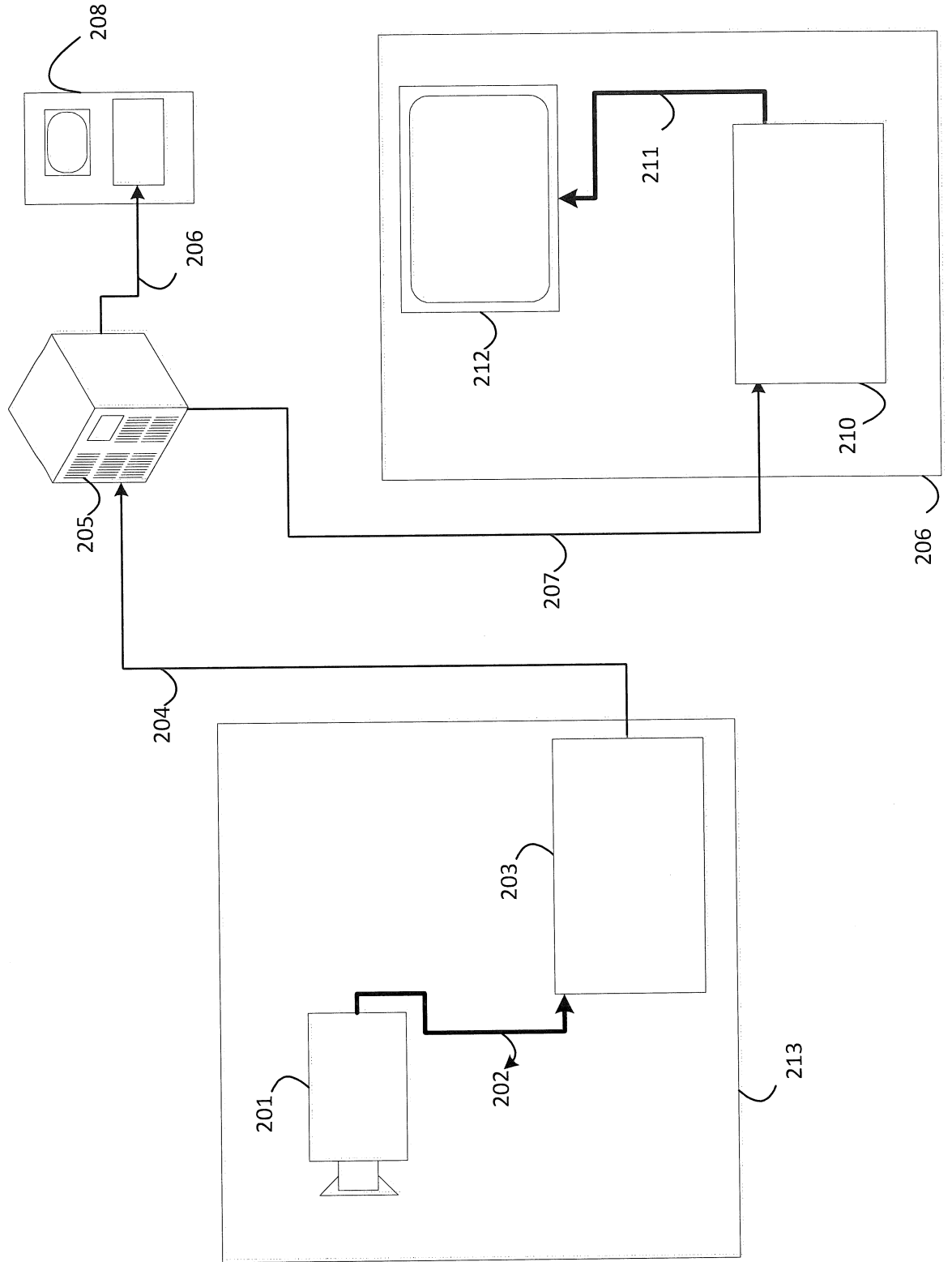


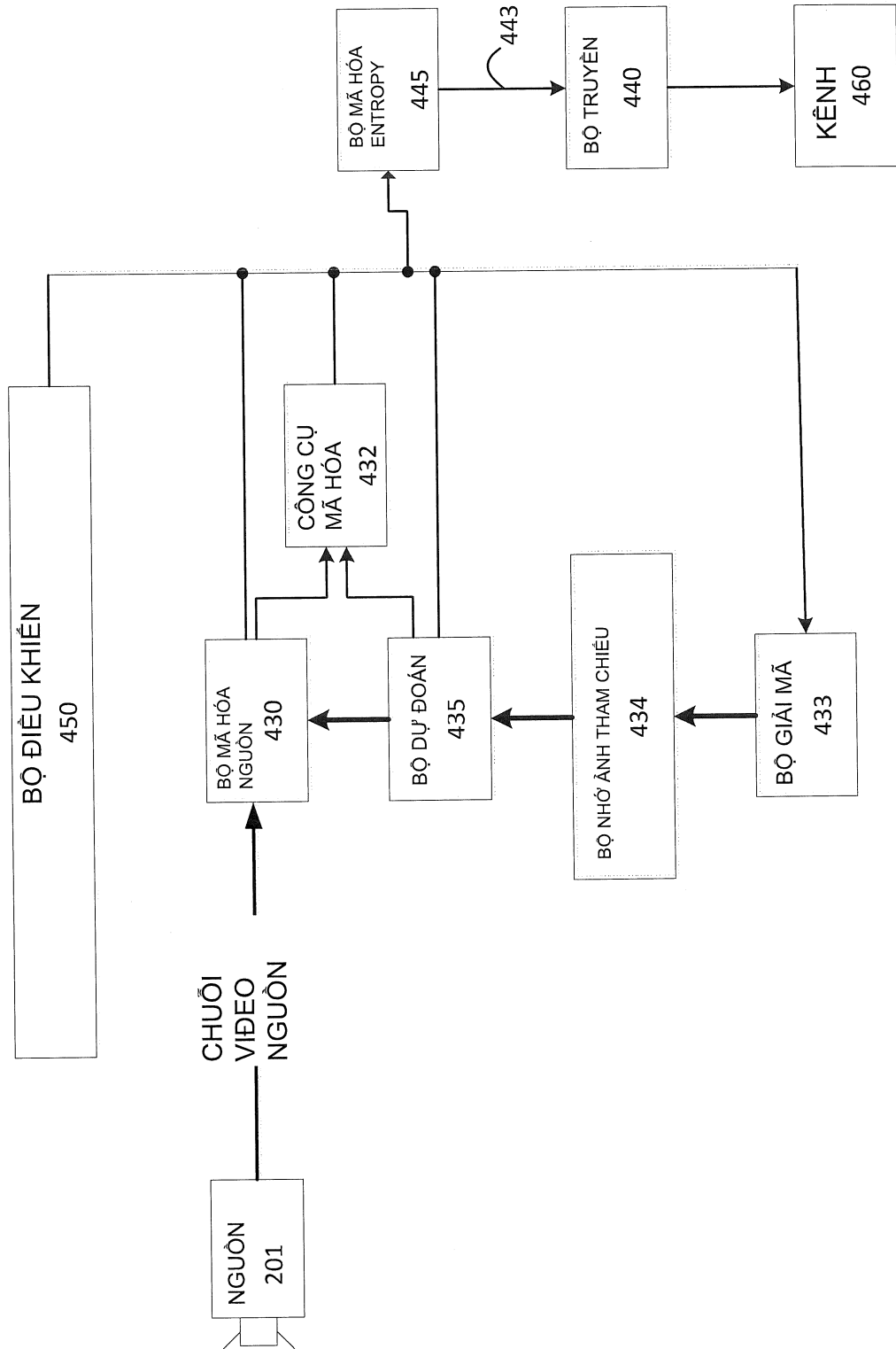
FIG. 2

200

2/24



4/24

FIG. 4
Bộ mã hóa 203

5/24

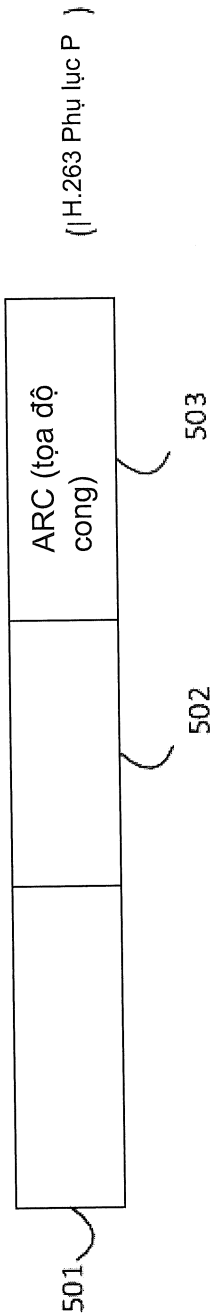


FIG. 5A

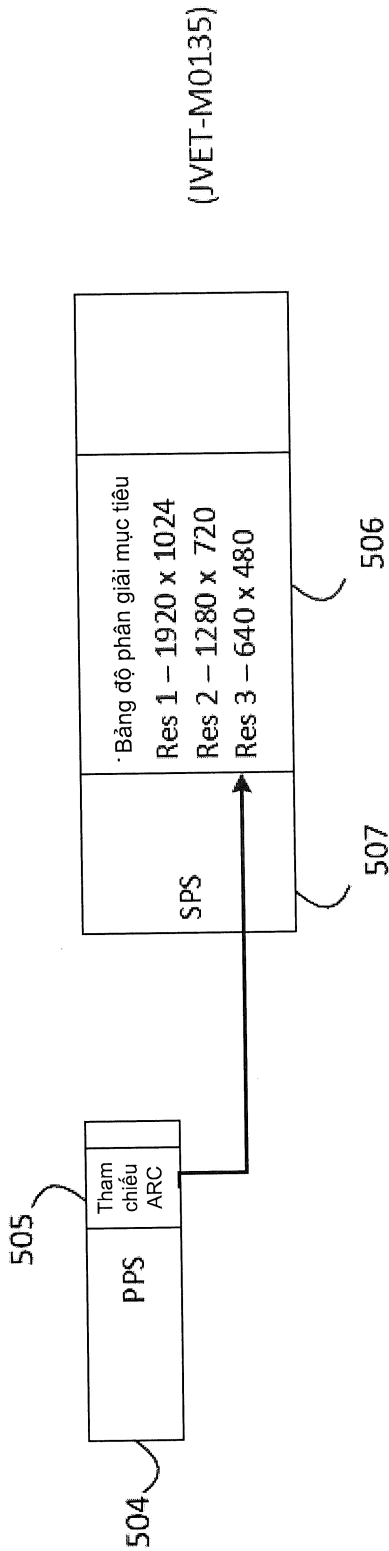


FIG. 5B

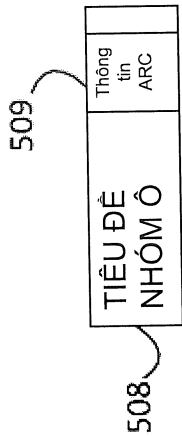


FIG. 5C

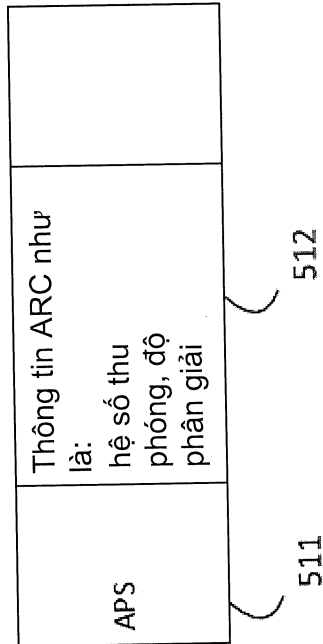


FIG. 5D

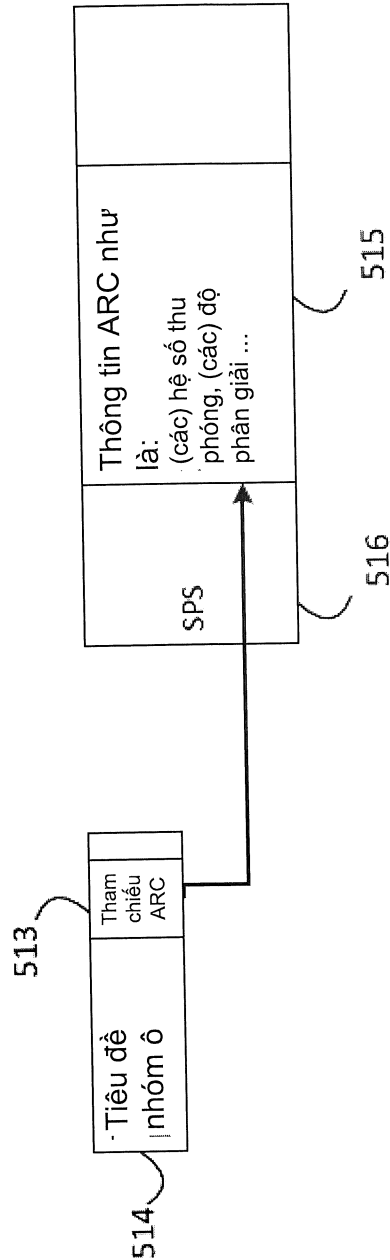


FIG. 5E

7/24

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

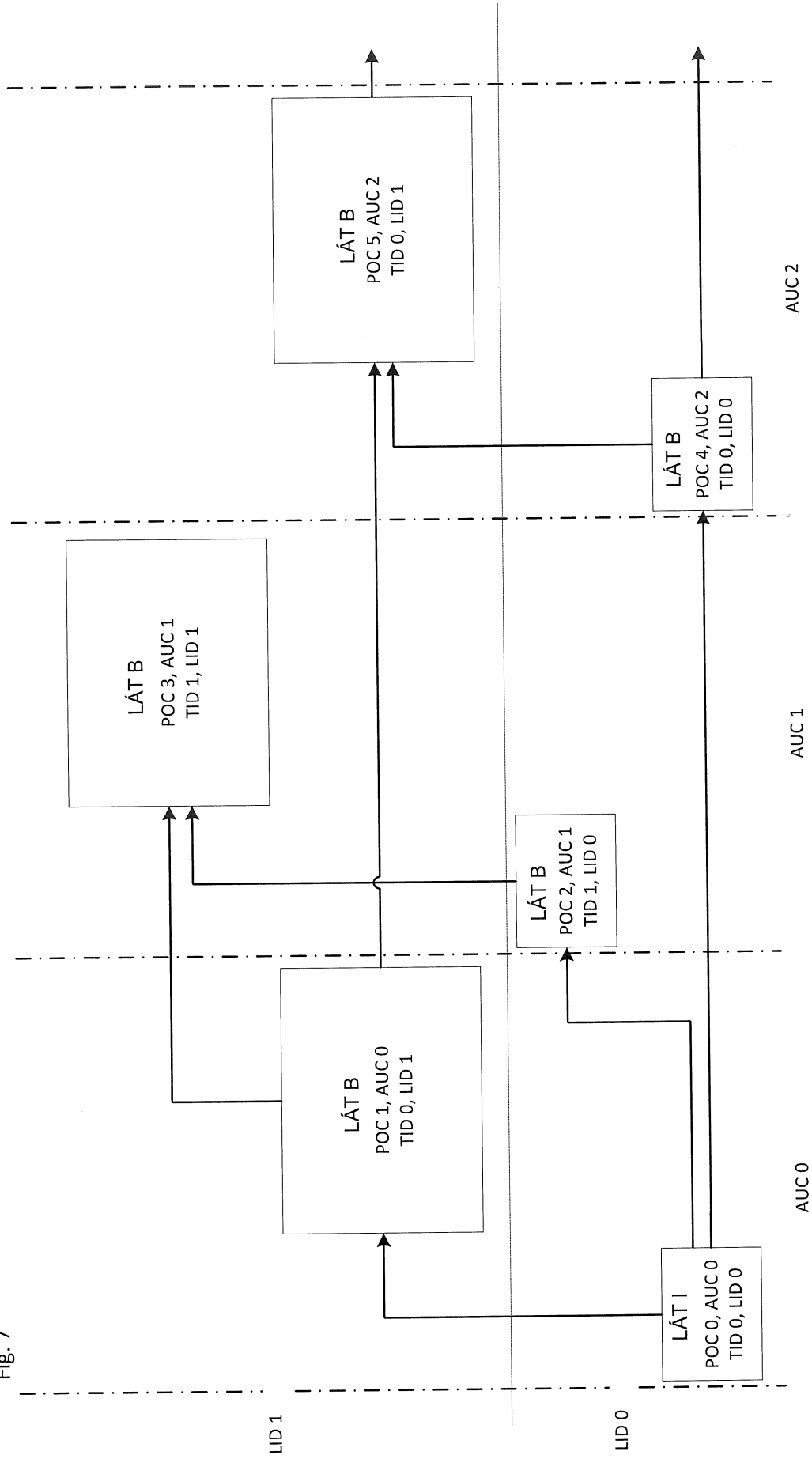
FIG. 6A

610	seq_parameter_set_rbsp() {	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)
615	reference_pic_height_in_luma_samples	ue(v)
	}	
616	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)
617	dec_pic_height_in_luma_samples[i]	ue(v)
	}	
	}	
	...	
	}	

FIG. 6B

8/24

Fig. 7



POC: số đếm thứ tự ảnh
AUC: Số đếm đơn vị truy cập
TID: ID thời gian
LID: ID lớp

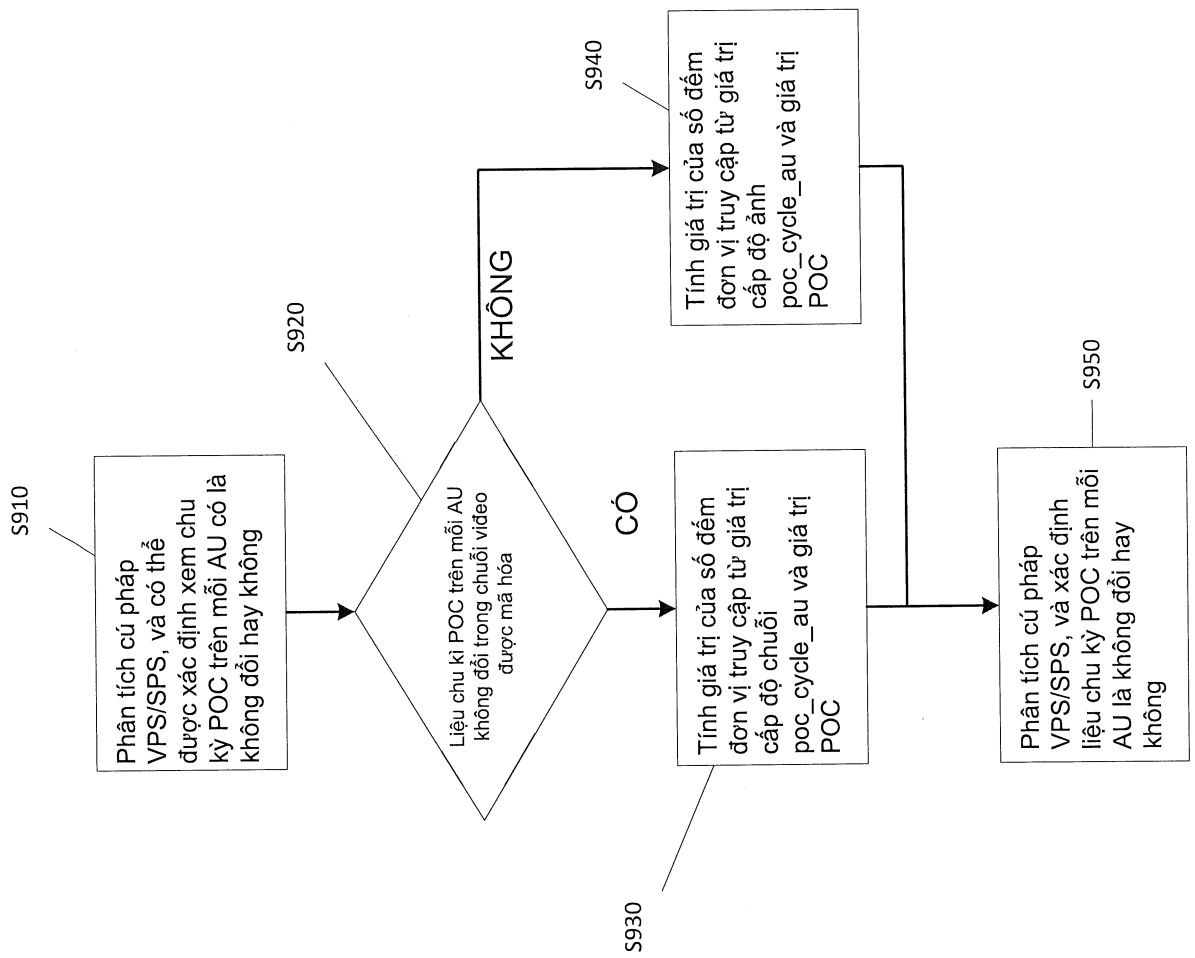
Fig. 8

video_parameter_set_lsb() { ^e	Mô tả ^e
vps_video_parameter_set_id ^e	u(4) ^e
vps_max_layers_minus1 ^e	u(8) ^e
for(i=0; i <= vps_max_layers_minus1; i++) { ^e	^e
vps_included_layer_id[i] ^e	u(7) ^e
vps_reserved_zero_bit ^e	u(1) ^e
} ^e	^e
vps_constraint_info_present_flag ^e	u(1) ^e
vps_constant_poc_cycle_per_aud ^e	u(1) ^e
if(vps_constant_poc_cycle_per_aud ^e)	^e
vps_poc_cycle_aud ^e	u(8) ^e
... ^e	^e
} ^e	^e

^e slice_header() { ^e	Mô tả ^e
slice_pic_parameter_set_id ^e	ue(v) ^e
if(rect_slice_flag NumBricksInPic > 1) ^e	^e
slice_address ^e	u(v) ^e
if(!rect_slice_flag && !single_brick_per_slice_flag) ^e	^e
num_bricks_in_slice_minus1 ^e	ue(v) ^e
slice_type ^e	ue(v) ^e
if(NatUnitType == GRA_NUT) ^e	^e
recovery_poc_cnt ^e	se(v) ^e
slice_pic_order_cnt_lsb ^e	u(v) ^e
... ^e	^e
if(!vps_constant_poc_cycle_per_aud) ^e	^e
slice_poc_cycle_aud ^e	u(8) ^e
... ^e	^e
... ^e	^e
} ^e	^e

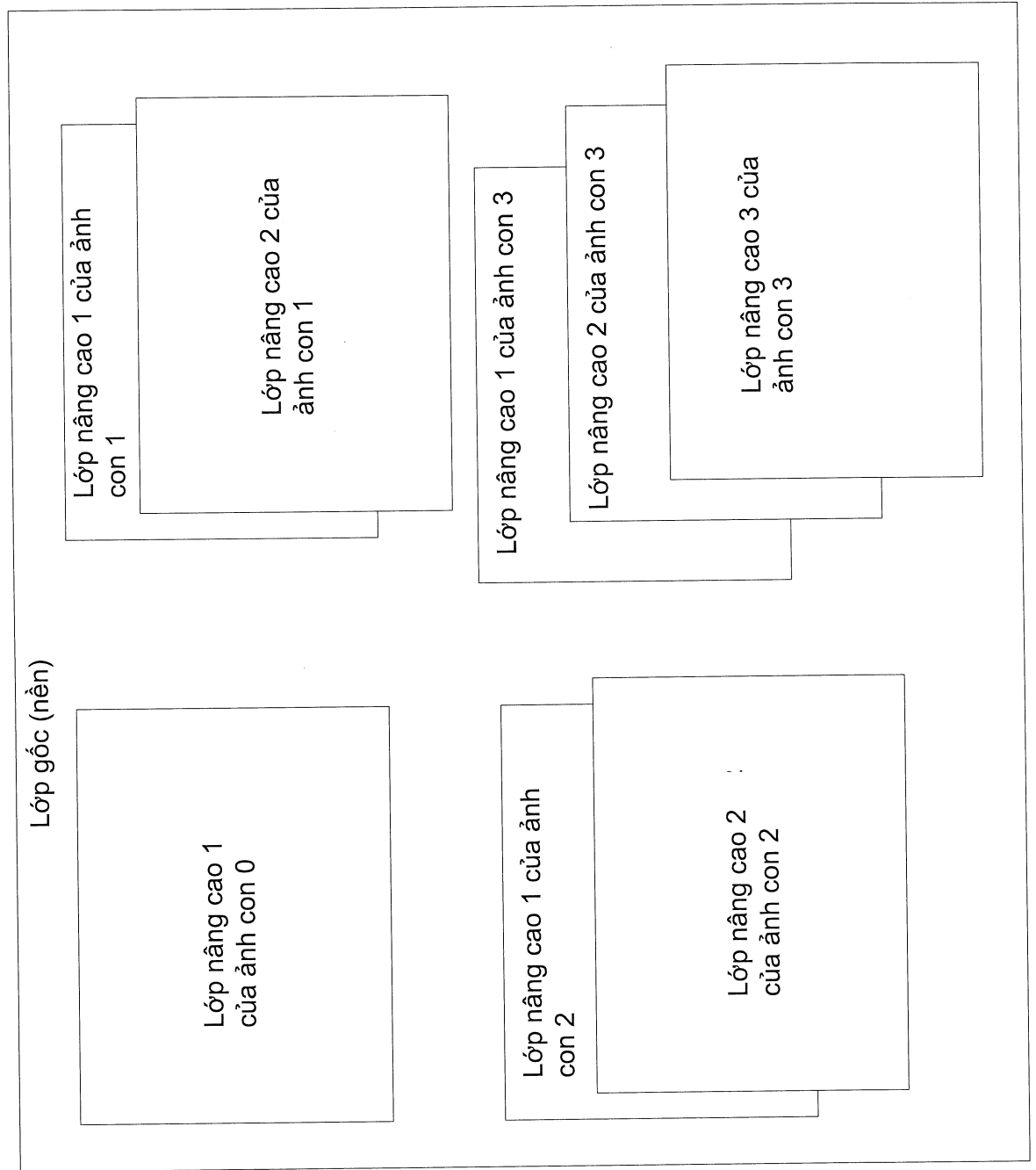
10/24

Fig. 9



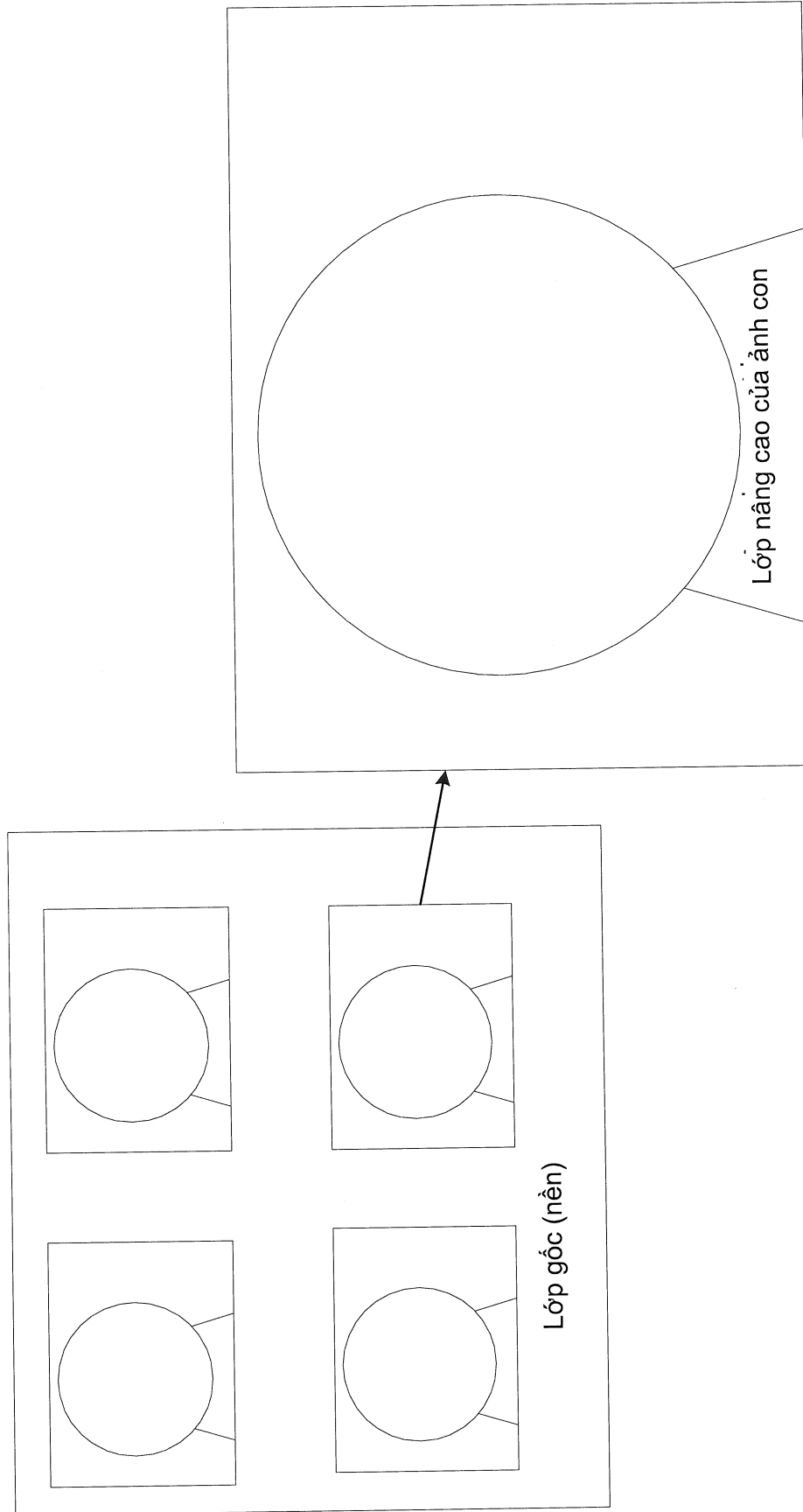
11/24

FIG. 10



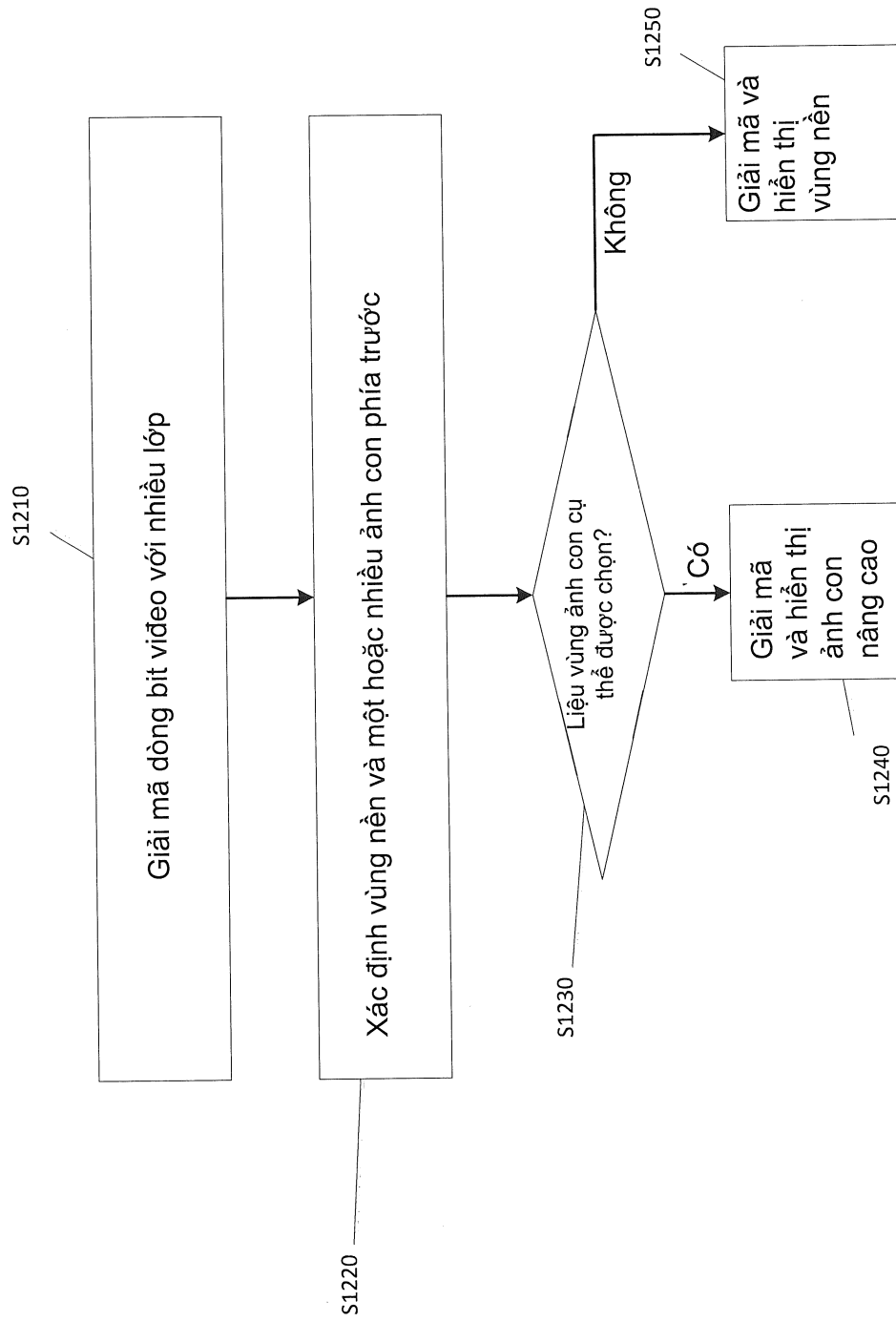
12/24

FIG. 11



13/24

Fig. 12



14/24

Fig. 13

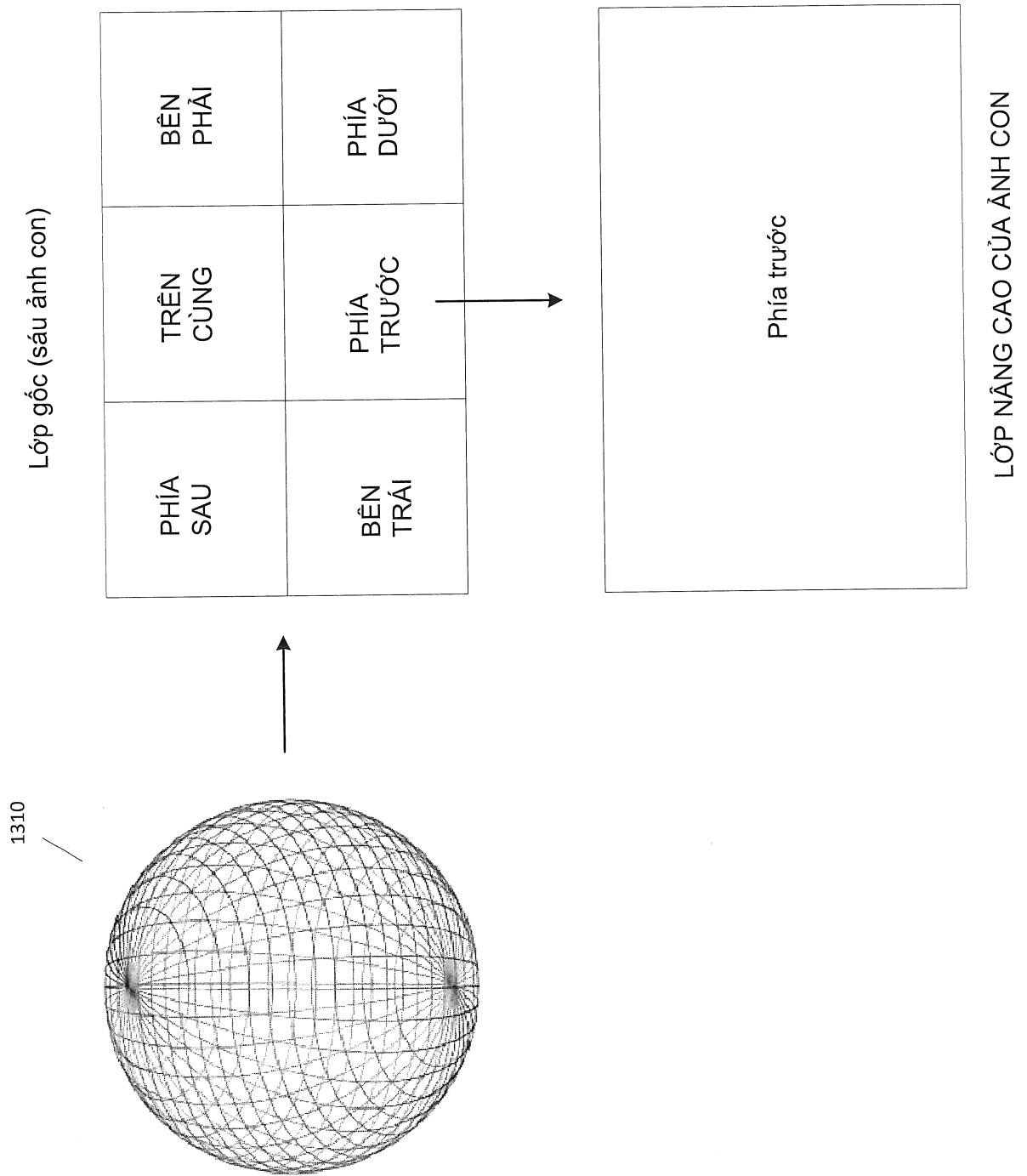
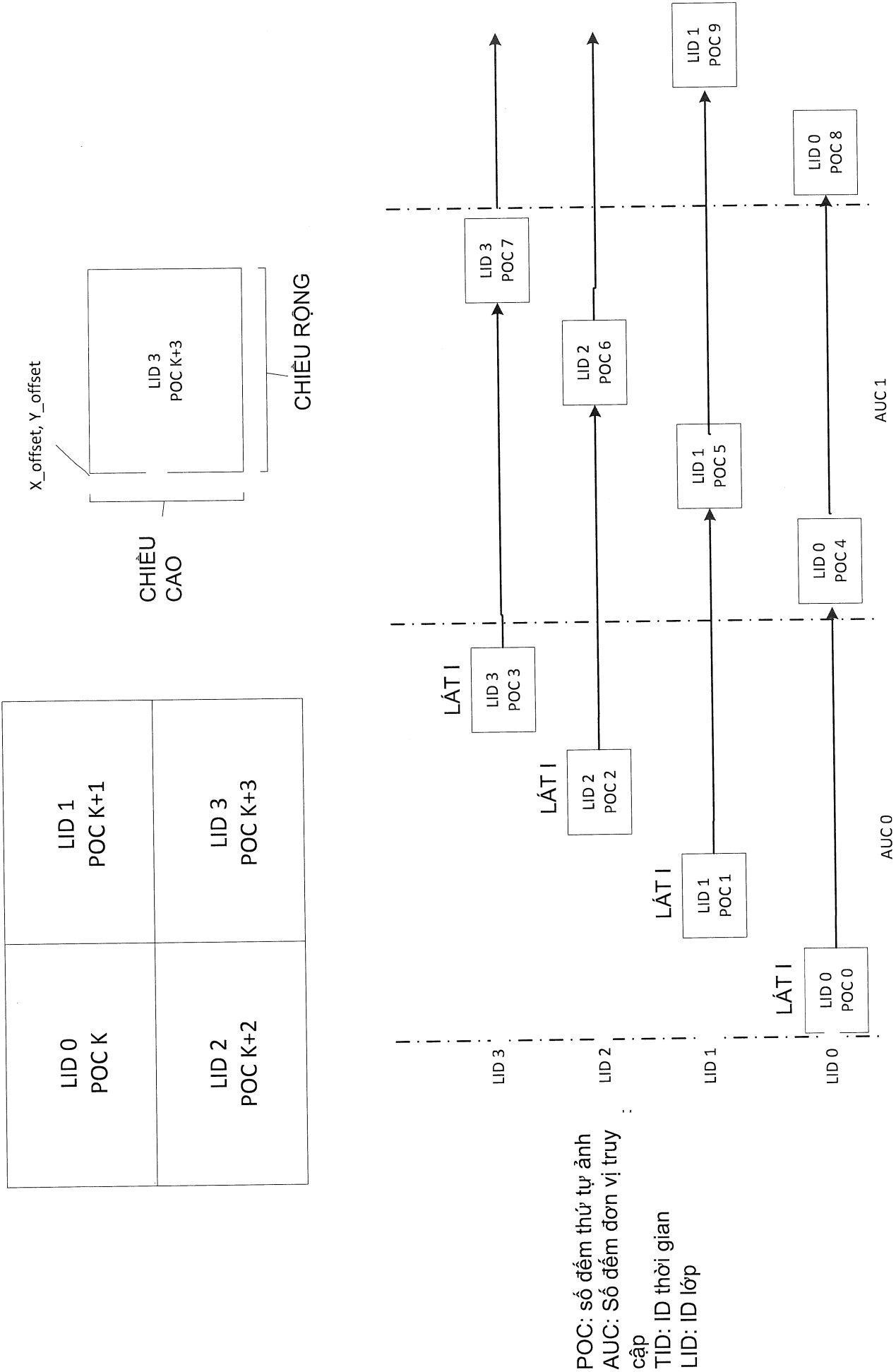


Fig. 14



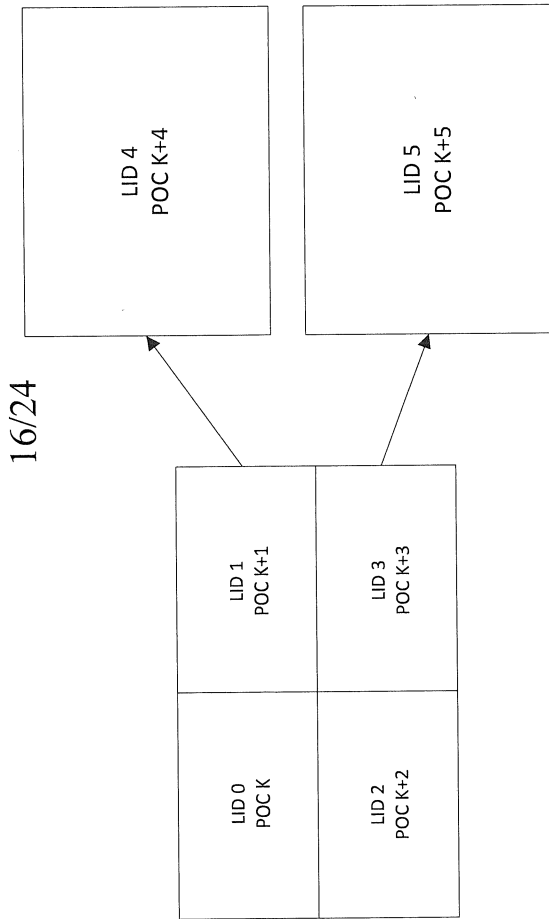
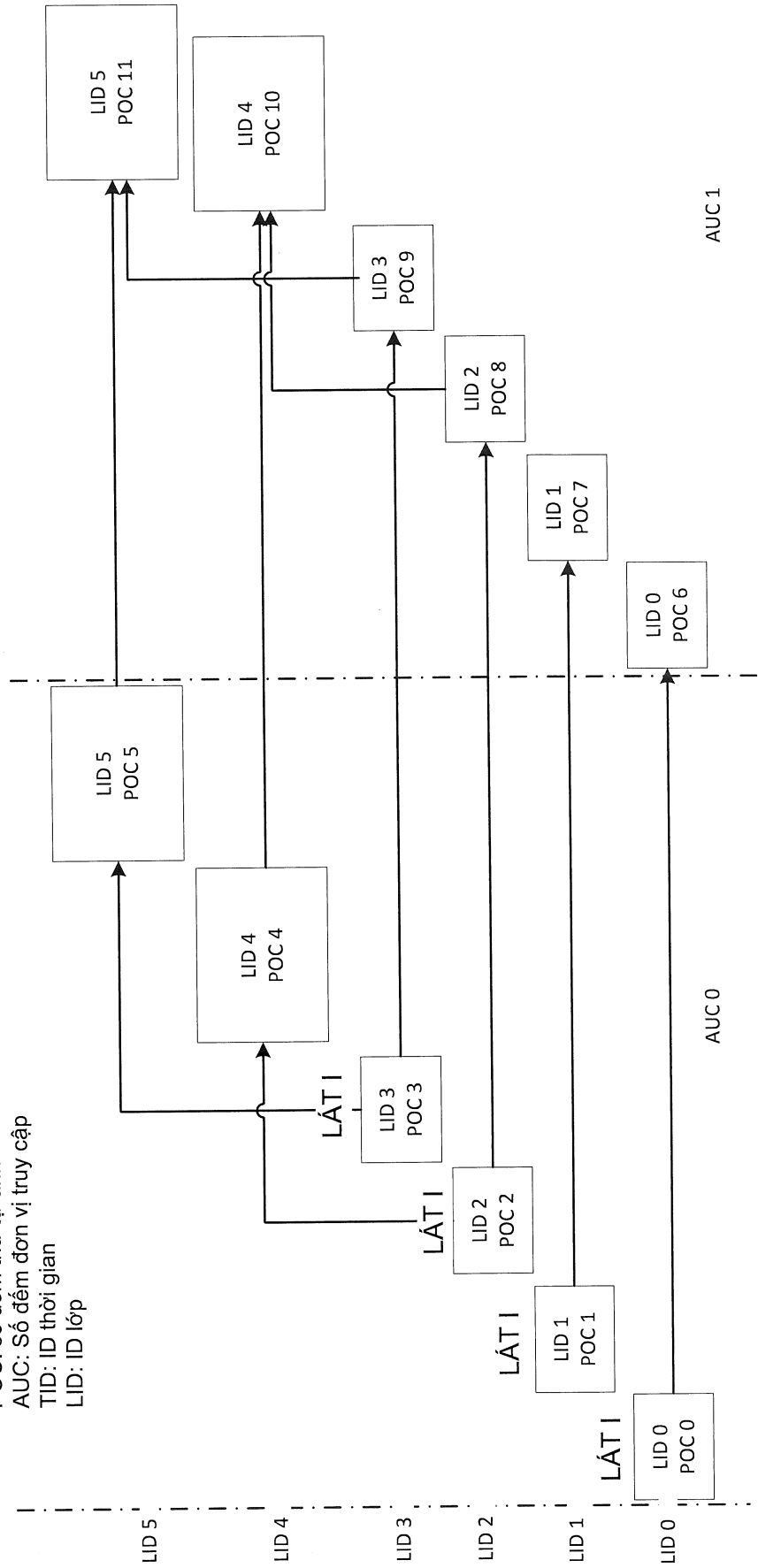


Fig. 15

POC: số đếm thứ tự ảnh
AUC: Số đếm đơn vị truy cập
TID: ID thời gian
LID: ID lớp



<code>video_parameter_set_rbsp() {</code>	Mô tả
<code> xps_video_parameter_set_id</code>	$\underline{u}(4)^{+2}$
<code> xps_max_layers_minus1</code>	$u(8)^{+2}$
<code> for(i = 0; i <= xps_max_layers_minus1; i++) {</code>	$^{+2}$
<code> xps_included_layer_id[i]</code>	$u(7)^{+2}$
<code> xps_reserved_zero_bit</code>	$u(1)^{+2}$
} $^{+2}$	$^{+2}$
... $^{+2}$	$^{+2}$
<code>xps_sub_picture_dividing_flag</code>	$u(1)^{+2}$
if(xps_sub_picture_dividing_flag) { $^{+2}$	$^{+2}$
<code> xps_full_pic_width_in_luma_samples</code>	$ue(v)^{+2}$
<code> xps_full_pic_height_in_luma_samples</code>	$ue(v)^{+2}$
} $^{+2}$	$^{+2}$
... $^{+2}$	$^{+2}$
} $^{+2}$	$^{+2}$

FIG. 16A

<code>seq_parameter_set_rbsp() {</code>	Mô tả
<code>sps_decoding_parameter_set_id</code>	$\underline{u}(4)^{+2}$
<code>sps_video_parameter_set_id</code>	$\underline{u}(4)^{+2}$
<code>sps_max_sub_layers_minus1</code>	$\underline{u}(3)^{+2}$
... $^{+2}$	$^{+2}$
<code>pic_width_in_luma_samples</code>	$ue(v)^{+2}$
<code>pic_height_in_luma_samples</code>	$ue(v)^{+2}$
if(xps_sub_picture_dividing_flag) { $^{+2}$	$^{+2}$
<code> pic_offset_x</code>	$ue(v)^{+2}$
<code> pic_offset_y</code>	$ue(v)^{+2}$
} $^{+2}$	$^{+2}$
... $^{+2}$	$^{+2}$
} $^{+2}$	$^{+2}$

FIG. 16B

Fig. 17

18/24

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

19/24

Fig. 18

video parameter set rsp() {	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 < NumLayersInList[i]; j++) {	
output layer flag[i][j]	u(1)
profile tier level idx[i][j]	u(v)
}	
...	
}	

Fig. 19

20/24

video parameter set rbsp() {	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)
}	
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][j]	u(1)
if(vps ptl signal flag[i])	
profile tier level idx[i][j]	u(v)
}	
}	
...	
}	

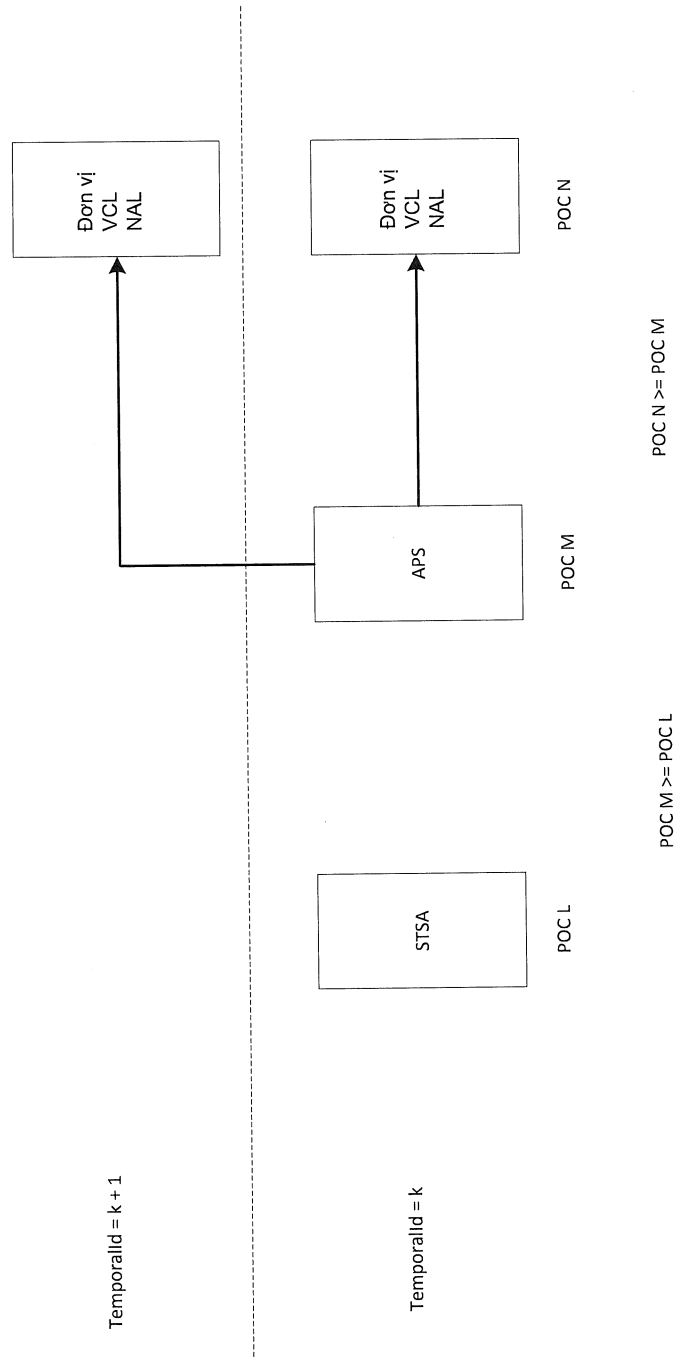
21/24

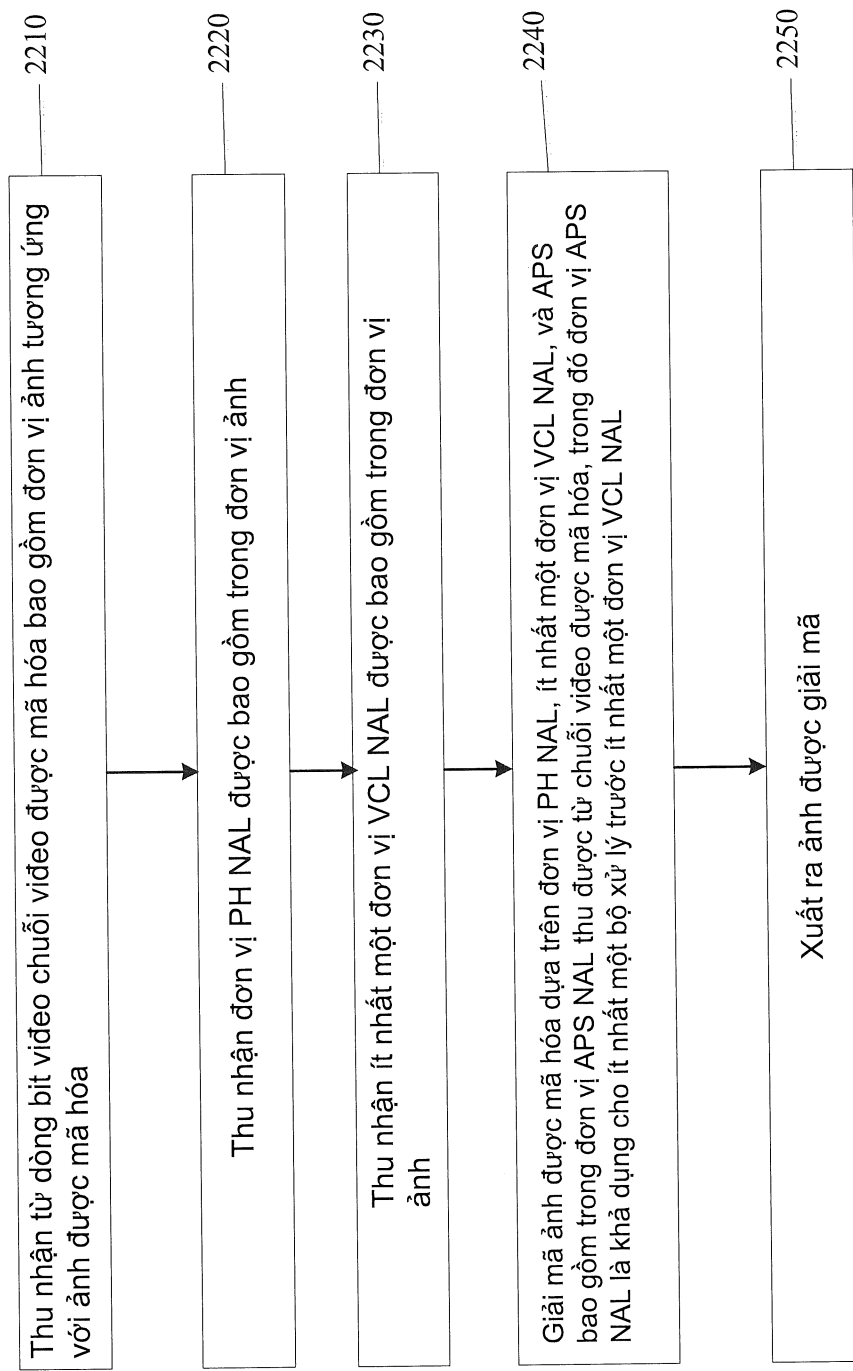
Fig. 20

video_parameter set rsp() {	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. 21





24/24

