



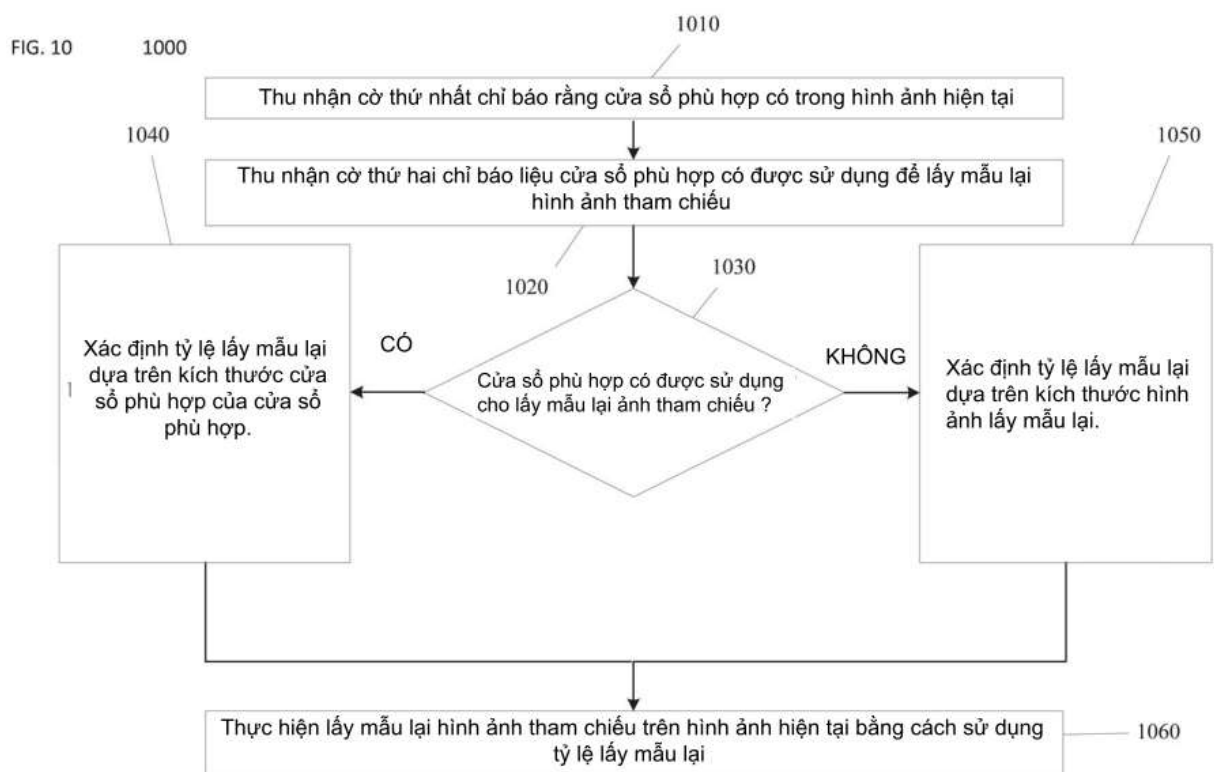
(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ
(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN) (11) 
CỤC SỞ HỮU TRÍ TUỆ
1-0048053
(51)^{2020.01} H04N 19/132; H04N 19/172; H04N 19/463; H04N 19/167 (13) B

(21) 1-2021-06779 (22) 10/09/2020
(86) PCT/US2020/050076 10/09/2020 (87) WO2021/055215 A1 25/03/2021
(30) 62/903,639 20/09/2019 US; 62/905,319 24/09/2019 US; 17/010,163 02/09/2020 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); WENGER, Stephan (DE); LIU, Shan (US).
(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) 1-2021-06779

(57) Sáng chế đề cập đến phương pháp, thiết bị, và phương tiện đọc được bằng máy tính để giải mã dòng bit video được mã hóa bằng cách sử dụng ít nhất một bộ xử lý, bao gồm bước thu nhận cờ thứ nhất chỉ báo rằng cửa sổ tương thích có trong hình ảnh hiện tại; dựa trên cờ thứ nhất chỉ báo rằng cửa sổ tương thích hiện có, thu nhận cờ thứ hai chỉ báo liệu cửa sổ tương thích có được sử dụng để lấy mẫu lại hình ảnh tham chiếu hay không; dựa trên cờ thứ hai chỉ báo rằng cửa sổ tương thích được sử dụng để lấy mẫu lại hình ảnh tham chiếu, xác định tỷ lệ lấy mẫu lại giữa hình ảnh hiện tại và hình ảnh tham chiếu dựa trên kích thước cửa sổ tương thích của cửa sổ tương thích; dựa trên cờ thứ hai chỉ báo rằng cửa sổ tương thích không được sử dụng để lấy mẫu lại ảnh tham chiếu, xác định tỷ lệ lấy mẫu lại dựa trên kích thước ảnh lấy mẫu lại; và thực hiện lấy mẫu lại hình ảnh tham chiếu trên hình ảnh hiện tại bằng cách sử dụng tỷ lệ lấy mẫu lại.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến mã hóa và giải mã video, và cụ thể hơn, là bảo hiệu của việc lấy mẫu lại hình ảnh tham chiếu mà chỉ báo kích thước hình ảnh lấy mẫu lại.

Tình trạng kỹ thuật của sáng chế

Mã hóa và giải mã video bằng cách sử dụng dự đoán hình ảnh liên đới với 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, các mẫu độ chói 1920x1080 và các mẫu sắc độ liên quan. Loạt ảnh có thể có tốc độ ảnh cố định hoặc thay đổi (thường được gọi là tốc độ khung hình), của, ví dụ 60 ảnh mỗi 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 1920x1080 ở tốc độ khung hình 60 Hz) yêu cầu băng thông gần 1,5 Gbit/s. Một giờ của video như vậy cần hơn 600 GByte dung lượng lưu trữ.

Mục đích của việc mã hóa và giải mã video có thể làm giảm sự dư thừa trong tín hiệu video đầu vào, thông qua quá trình nén. Nén có thể giúp giảm các yêu cầu về băng thông hoặc không gian lưu trữ nói trên, trong các trường hợp bằng hai bậc cường độ trở lên. Cả nén không tổn hao và nén có tổn hao đều có thể được sử dụng, cũng như kết hợp của chúng. 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 độ méo 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 chịu được phụ thuộc vào ứng dụng; ví dụ, người dùng các ứng dụng phát trực tuyến của người tiêu dùng nhất định có thể chịu được

độ méo cao hơn so với người dùng ứng dụng đóng góp truyền hình. Tỷ lệ nén có thể đạt được có thể phản ánh rằng: độ méo cho phép/có thể chịu đựng 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 các 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.

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

Những thay đổi về kích thước hình ảnh trong các hình ảnh được dự đoán trở nên phổ biến hơn trong mã hóa video hiện đại. Ví dụ, VP9 cho phép lấy mẫu lại hình ảnh tham chiếu và thay đổi độ phân giải cho toàn bộ hình ảnh. Tương tự, các đề xuất nhất định được đưa ra đối với VVC (bao gồm, ví dụ, Hendry, và cộng sự, “Thay đổi độ phân giải thích ứng (ARC) cho VVC”, tài liệu của Nhóm video chung JVET-M0135-v1, ngày 9-19 tháng 1 năm 2019, được đưa vào tài liệu này toàn bộ) cho phép lấy mẫu lại toàn bộ hình ảnh tham chiếu thành các độ phân giải khác nhau — cao hơn hoặc thấp hơn —. Trong tài liệu đó, các độ phân giải ứng

viên khác nhau được đề xuất được mã hóa trong tập tham số trình tự và được tham chiếu bởi các phân tử cú pháp trên mỗi ảnh trong tập tham số ảnh.

Bản chất kỹ thuật của sáng chế

Theo phương án của sáng chế, sáng chế đề xuất phương pháp giải mã dòng bit video được mã hóa bằng cách sử dụng ít nhất một bộ xử lý bao gồm thu nhận cờ thứ nhất chỉ báo rằng cửa sổ tương thích có trong hình ảnh hiện tại; dựa trên cờ thứ nhất chỉ báo rằng cửa sổ tương thích hiện có, thu nhận cờ thứ hai chỉ báo liệu cửa sổ tương thích có được sử dụng để lấy mẫu lại hình ảnh tham chiếu hay không; dựa trên cờ thứ hai chỉ báo rằng cửa sổ tương thích được sử dụng để lấy mẫu lại hình ảnh tham chiếu, xác định tỷ lệ lấy mẫu lại giữa hình ảnh hiện tại và hình ảnh tham chiếu dựa trên kích thước cửa sổ tương thích của cửa sổ tương thích; dựa trên cờ thứ hai chỉ báo rằng cửa sổ tương thích không được sử dụng để lấy mẫu lại ảnh tham chiếu, xác định tỷ lệ lấy mẫu lại dựa trên kích thước ảnh lấy mẫu lại; và thực hiện lấy mẫu lại hình ảnh tham chiếu trên hình ảnh hiện tại bằng cách sử dụng tỷ lệ lấy mẫu lại.

Theo phương án của sáng chế, sáng chế đề xuất thiết bị 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 nhận được cờ thứ nhất chỉ báo rằng cửa sổ tương thích có trong bức ảnh hiện tại; mã thu nhận thứ hai được cấu hình để khiến ít nhất một bộ xử lý, dựa trên cờ thứ nhất chỉ báo rằng cửa sổ tương thích hiện có, thu nhận cờ thứ hai chỉ báo liệu cửa sổ tương thích có được sử dụng để lấy mẫu lại hình ảnh tham chiếu hay không; mã xác định thứ nhất được cấu hình để khiến ít nhất một bộ xử lý, dựa trên cờ thứ hai chỉ báo rằng cửa sổ tương thích được sử dụng để lấy mẫu lại hình ảnh tham chiếu, xác định tỷ lệ lấy mẫu lại giữa hình ảnh hiện tại và hình ảnh tham chiếu dựa trên kích thước cửa sổ tương thích là cửa sổ tương thích; mã xác định thứ hai được cấu hình để khiến ít nhất một bộ xử lý, dựa trên cờ thứ hai chỉ báo rằng cửa sổ tương thích không được sử dụng cho việc lấy mẫu lại hình ảnh tham chiếu, xác định tỷ lệ lấy mẫu lại dựa

trên kích thước hình ảnh lấy mẫu lại; và thực hiện mã được cấu hình để khiến ít nhất một bộ xử lý thực hiện lấy mẫu lại hình ảnh tham chiếu trên hình ảnh hiện tại bằng cách sử dụng tỷ lệ lấy mẫu lại.

Theo phương án của sáng chế, sáng chế đề xuất phương tiện bất biến đọc được bằng máy tính lưu trữ các lệnh, các lệnh bao gồm một hoặc nhiều lệnh mà, 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 cờ thứ nhất chỉ báo rằng cửa sổ tương thích hiện có trong hình ảnh hiện tại; dựa trên cờ thứ nhất chỉ báo rằng cửa sổ tương thích hiện có, thu nhận cờ thứ hai chỉ báo cửa sổ tương thích có được sử dụng để lấy mẫu lại hình ảnh tham chiếu hay không; dựa trên cờ thứ hai chỉ báo rằng cửa sổ tương thích được sử dụng để lấy mẫu lại hình ảnh tham chiếu, xác định tỷ lệ lấy mẫu lại giữa hình ảnh hiện tại và hình ảnh tham chiếu dựa trên kích thước cửa sổ tương thích của cửa sổ tương thích; dựa trên cờ thứ hai chỉ báo rằng cửa sổ tương thích không được sử dụng để lấy mẫu lại ảnh tham chiếu, xác định tỷ lệ lấy mẫu lại dựa trên kích thước ảnh lấy mẫu lại; và thực hiện lấy mẫu lại hình ảnh tham chiếu trên hình ảnh hiện tại bằng cách sử dụng tỷ lệ lấy mẫu lại.

Mô tả vắn tắt các hình vẽ

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

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

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

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

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

FIG.5 là minh họa giản lược về các tùy chọn để báo hiệu các tham số ARC theo phương án.

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

FIG.7 là minh họa giản lược về báo hiệu của kích thước ảnh và cửa sổ tương thích trong SPS theo phương án.

Các FIG.8A-8B là các minh họa giản lược về báo hiệu của kích thước ảnh và cửa sổ tương thích trong PPS theo các phương án.

FIG.9 là minh họa giản lược về báo hiệu của việc lấy mẫu lại kích thước hình ảnh trong PPS theo phương án.

FIG.10 là lưu đồ của quy trình ví dụ để giải mã dòng bit video được mã hóa theo phương án.

FIG.11 là minh họa giản lược của hệ thống máy tính theo phương án.

Mô tả chi tiết các hình vẽ

FIG.1 minh họa sơ đồ khối đơn giản hóa của hệ thống truyền thông 100 theo phương án của sáng chế. Hệ thống 100 có thể bao gồm ít nhất hai thiết bị đầu cuối (110-120) được kết nối liên 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 kia 120 qua mạng 150. Thiết bị đầu cuối thứ hai 120 có thể thu dữ liệu video được mã hóa của thiết bị đầu cuối khác từ mạng 150, giải mã dữ liệu được mã hóa và hiển thị dữ liệu video 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 cung cấp để hỗ trợ truyền hai chiều video được mã hóa mà 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 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 thu dữ liệu video

được mã hóa được truyền bởi thiết bị đầu cuối khác, có thể giải mã dữ liệu được mã hóa và có thể hiển thị dữ liệu video 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 thể hiện 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 mà 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 và/hoặc chuyên gói. Mạng đại diện bao gồm mạng viễn thông, mạng cục bộ, mạng diện rộng và/hoặc Internet. Đối với mục đích của cuộc thảo luận 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 dưới đây.

FIG.2 minh họa, lấy làm ví dụ cho ứng dụng cho đối tượng sáng chế, vị trí của bộ mã hóa và giải mã video trong môi trường phát trực tuyến. Đối tượng sáng chế 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 ảnh 213, mà có thể bao gồm nguồn video 201, ví dụ máy ảnh kỹ thuật số, tạo ra ví dụ luồng mẫu video không nén 202. Luồng mẫu đó 202, được mô tả là 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 sáng chế như được mô tả chi tiết hơn bên dưới. Dòng bit video được mã hóa 204, được mô tả là đường mảnh để nhấn mạnh khối lượng dữ liệu thấp hơn khi so sánh với dòng mẫu, có thể được lưu trữ trên 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 luồng mẫu video đi bên ngoài 211 có thể được hiển thị trên màn hình 212 hoặc thiết bị kết xuất khác (không mô tả). Trong các hệ thống phát trực tuyến, các dòng bit video (204, 207, 209) có thể được mã hóa theo các tiêu chuẩn mã hóa/nén video nhất định. Ví dụ về các tiêu chuẩn đó bao gồm Khuyến nghị H.265 của ITU-T. Đang được phát triển là tiêu chuẩn mã hóa video thường được gọi là Mã hóa video đa năng hoặc VVC. Đối tượng sáng chế có thể được sử dụng trong ngữ cảnh của VVC.

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

Bộ thu 310 có thể thu một hoặc nhiều chuỗi mã hóa-giải mã video sẽ được giải mã bởi bộ giải mã 210; theo cùng hoặc phương án khác, một chuỗi video được mã hóa được thu tại một thời điểm, mà 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 thu từ kênh 312, mà có thể là liên kết phần cứng/phần mềm tới thiết bị lưu trữ để lưu trữ dữ liệu video được mã hóa. Bộ thu 310 có thể thu dữ liệu video được mã hóa cùng với dữ liệu khác, ví dụ, dữ liệu âm thanh được mã hóa và/hoặc các luồng dữ liệu phụ trợ, mà 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/phân tích cú pháp entropy 320 (“phân tích cú pháp” trở về sau). 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 năng suất cao 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 trình phân tích cú pháp 320 để tái cấu trúc các ký tự 321 từ chuỗi video được mã hóa entropy. Các danh mục của các ký tự 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 phần không thể tách rời của bộ giải mã nhưng có thể được kết hợp với nó, như được hiển thị trong FIG.3. Thông tin điều khiển cho (các) thiết bị kết xuất có thể ở dạng Thông tin tăng cường bổ sung (các bản tin SEI) hoặc các đoạn tập tham số Thông tin về khả năng sử dụng video (VUI) (không được mô tả). Trình phân tích cú pháp 320 có thể phân tích cú pháp/giải mã- entropy chuỗi video được mã hóa nhận được. Việc mã hóa chuỗi video được mã hóa có thể phù hợp với công nghệ hoặc tiêu chuẩn mã hóa video và có thể tuân theo các nguyên tắc mà người có hiểu biết trung bình trong lĩnh vực kỹ thuật biết rõ, bao gồm mã hóa độ dài thay đổi, mã hóa Huffman, mã hóa số học với có hoặc không có độ nhạy ngữ cảnh, và vân vân. Trình phân tích cú pháp 320 có thể trích xuất từ chuỗi video được mã hóa, tập hợp các tham số nhóm phụ cho ít nhất một trong các nhóm phụ của điểm ảnh trong bộ giải mã video, dựa trên ít nhất một tham số tương ứng với nhóm. Các nhóm phụ có thể bao gồm Nhóm ảnh (GOP), ảnh, ảnh phụ, ô xếp, lớp, khối hình, khối macro, Đơn vị cây mã hóa (CTU) Đơn vị mã hóa (CU), khối, Đơn vị biến đổi (TU), Đơn vị dự đoán (PU) và vân vân. Ô có thể chỉ ra vùng hình chữ nhật của CU/CTU với cột và hàng ô cụ thể trong hình ảnh. Viên có thể chỉ ra vùng hình chữ nhật của các hàng CU/CTU trong ô cụ thể. Lớp có thể chỉ ra một hoặc nhiều khối hình của ảnh, mà được chứa trong đơn vị NAL. Hình phụ có thể chỉ ra vùng hình chữ nhật của một hoặc nhiều lớp trong hình. Bộ giải mã/phân tích cú pháp entropy cũng có thể trích xuất từ thông tin chuỗi video được mã hóa như các hệ số biến đổi, các giá trị tham số bộ lượng tử, các vector chuyển động, và v.v..

Trình phân tích cú pháp 320 có thể thực hiện hoạt động giải mã/phân tích cú pháp entropy trên chuỗi video nhận được từ bộ đệm 315, để tạo các ký tự 321.

Việc tái cấu trúc các ký tự 321 có thể bao gồm nhiều bộ phận khác nhau tùy thuộc vào loại hình ảnh video được mã hóa hoặc các phần của chúng (chẳng hạn như: hình ảnh liên đới và nội bộ, khối liên đới và nội bộ), và các yếu tố khác.

Những bộ phận nào có liên quan, và bằng cách nào, có thể được điều khiển bởi thông tin điều khiển nhóm phụ đã được phân tích cú pháp từ chuỗi video được mã hóa bởi trình phân tích cú pháp 320. Luồng thông tin điều khiển nhóm phụ như vậy giữa trình phân tích cú pháp 320 và nhiều bộ phận 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 các bộ phận chức năng như được mô tả bên dưới. Trong triển khai thực tế hoạt động dưới các ràng buộc thương mại, nhiều bộ phận trong số này tương tác chặt chẽ với nhau và có thể, ít nhất một phần, có thể được tích hợp vào nhau. Tuy nhiên, với mục đích mô tả đối tượng sáng chế, việc chia nhỏ khái niệm thành các bộ phận chức năng dưới đây là phù hợp.

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

Trong các trường hợp, các mẫu đầu ra của phép biến đổi tỷ lệ/ 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 hình 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 hình ảnh hiện tại. Thông tin dự đoán như vậy có thể được cung cấp bởi bộ phận dự đoán hình ảnh nội bộ 352. Trong một số trường hợp, bộ phận dự đoán hình 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 các trường hợp, bổ sung, trên cơ sở mỗi mẫu, thông tin dự đoán mà bộ phận dự đoán nội bộ 352 đã tạo cho thông tin mẫu đầu ra do bộ phận biến đổi tỷ lệ/ngịch đảo 351 cung cấp.

Trong các trường hợp khác, các mẫu đầu ra của bộ phận biến đổi tỷ lệ/nghịch đảo 351 có thể liên quan đến khối được mã hóa liên, và có khả năng bù chuyển động. Trong trường hợp như vậy, bộ phận Dự báo bù chuyển động 353 có thể truy cập bộ nhớ hình ảnh tham chiếu 357 để lấy các mẫu được sử dụng để dự đoán. Sau khi bù chuyển động các mẫu được lấy theo các ký tự 321 liên quan đến khối, các mẫu này có thể được bổ sung bởi bộ tổng hợp 355 đến đầu ra của bộ phận biến đổi tỷ lệ / nghịch đảo (trong trường hợp này được gọi là mẫu dư hoặc báo hiệu dư) để tạo ra thông tin mẫu đầu ra. Các địa chỉ trong biểu mẫu bộ nhớ hình ảnh tham chiếu nơi bộ phận bù chuyển động tìm nạp các mẫu dự đoán có thể được điều khiển bằng vector chuyển động, khả dụng cho bộ phận bù chuyển động dưới dạng ký tự 321 mà có thể có, ví dụ X, Y, và các thành phần hình ảnh tham chiếu. Bù chuyển động cũng có thể bao gồm nội suy của các giá trị mẫu được lấy từ bộ nhớ hình 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. Công nghệ nén video có thể bao gồm công nghệ lọc trong vòng được kiểm soát bởi các tham số có trong dòng bit video được mã hóa và được cung cấp cho bộ phận lọc vòng 356 dưới dạng ký tự 321 từ trình 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 hình 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 bộ lọc vòng lặp 356 có thể là 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 hình ảnh liên đới trong tương lai.

Các hình ảnh được mã hóa, sau khi được tái cấu trúc hoàn chỉnh, có thể được sử dụng làm hình ảnh tham chiếu để dự đoán trong tương lai. Sau khi hình ảnh được mã hóa được tái cấu trúc hoàn chỉnh và hình ảnh được mã hóa đã được xác định là hình ảnh tham chiếu (bởi, ví dụ, trình phân tích cú pháp 320), hình ảnh

tham chiếu hiện tại 358 có thể trở thành một phần của bộ đệm hình ảnh tham chiếu 357, và bộ nhớ hình ảnh hiện tại mới có thể được phân bổ lại trước khi bắt đầu tái cấu trúc hình ảnh được mã hóa sau đây ..

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

Theo phương án, máy thu 310 có thể thu dữ liệu bổ sung (dự phòng) với video được mã hóa. Dữ liệu bổ sung có thể được đưa vào như một phần của (các) chuỗi video được mã hóa. Dữ liệu bổ sung có thể được sử dụng bởi bộ giải mã video 210 để giải mã dữ liệu một cách chính xác và / hoặc để tái cấu trúc 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 lớp thừa, hình ảnh dư, mã sửa lỗi chuyển tiếp, và 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 thu các mẫu video từ nguồn video 201 (không phải là một phần của bộ mã hóa) có thể chụp (các) hình ảnh video được bộ mã hóa 203.

Nguồn video 201 có thể cung cấp chuỗi video nguồn được mã hóa bởi bộ mã hóa 203 ở dạng luồng 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 hình ảnh cục bộ dưới dạng một chuỗi video. Dữ liệu video có thể được cung cấp dưới dạng nhiều hình ảnh riêng lẻ truyền tải chuyển động khi xem theo trình tự. 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 các hình 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 chức năng của Bộ điều khiển 450. Bộ điều khiển điều khiển các bộ phận chức năng khác như được mô tả dưới đây và được kết hợp chức năng với các bộ phận này. Khớp nối không được mô tả rõ ràng. Các tham số do bộ điều khiển thiết lập có thể bao gồm các tham số liên quan đến kiểm soát tốc độ (bỏ qua hình 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, và v.v. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật 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.

Các bộ mã hóa video hoạt động mà người có hiểu biết trung bình trong lĩnh vực kỹ thuật 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” trở về sau) (chịu trách nhiệm tạo các ký tự dựa trên hình ảnh đầu vào được mã hóa, và (các) hình ảnh tham chiếu), và bộ giải mã (cục bộ) 433 được nhúng trong bộ mã hóa 203 mà tái cấu trúc các ký tự để tạo dữ liệu mẫu mà bộ giải mã (từ xa) cũng sẽ tạo ra (vì bất kỳ sự nén nào giữa các ký tự và dòng bit video được mã hóa đều không bị mất trong công nghệ nén video được xem xét trong đối tượng sáng chế). Luồng mẫu được tái cấu trúc đó được đưa vào bộ nhớ hình ảnh tham chiếu 434. Vì việc giải mã luồng ký tự dẫn đến kết quả chính xác theo bit không phụ thuộc vào vị trí của bộ giải mã (cục bộ hoặc từ xa), nội dung bộ đệm hình ảnh tham chiếu cũng chính xác từng 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 hình 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 hình ảnh tham chiếu (và lệch kết quả, nếu tính đồng bộ không thể được duy trì, ví dụ do lỗi kênh) được biết rõ bởi người có hiểu biết trung bình trong lĩnh vực kỹ thuật.

Hoạt động của bộ giải mã “cục bộ” 433 có thể giống như hoạt động của bộ giải mã “từ xa” 210, 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 tự khả dụng và việc vi / giải mã các ký tự thành chuỗi video được mã hóa bằng bộ mã hóa entropy 445 và trình phân tích cú pháp 320 có thể không tổn hao, các bộ phận giải mã entropy của bộ giải mã 210, bao gồm kênh 312, bộ thu 310, bộ đệm 315, và trình phân tích cú pháp 320 có thể không được triển khai đầy đủ trong bộ giải mã cục bộ 433.

Nhận xét mà có thể được thực hiện tại thời điểm này là bất kỳ công nghệ giải mã nào ngoại trừ phân tích cú pháp/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 bộ mã hóa tương ứng. Vì lý do này, đối tượng sáng chế tập trung vào hoạt động của bộ giải mã. Mô tả của các công nghệ bộ mã hóa có thể được viết tắt vì chúng là nghịch đảo của các công nghệ bộ giải mã được mô tả toàn diện. Chỉ trong một số khu vực nhất định, mô tả chi tiết hơn mới được yêu cầu và cung cấp bên dưới.

Là một phần hoạt động của nó, bộ mã nguồn 430 có thể thực hiện mã hóa dự đoán bù chuyển động, mà mã hóa dự đoán khung đầu vào 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 khác biệt giữa các khối điểm ảnh của khung đầu vào và các khối điểm ảnh của (các) khung tham chiếu 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 tự được tạo bởi bộ mã nguồn 430. Các hoạt động của công cụ mã hóa 432 có thể có lợi cho các quá trình tổn hao. Khi dữ liệu video được mã hóa có thể được giải mã tại bộ giải mã video (không được hiển thị trong FIG.4), chuỗi video được tái cấu trúc 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 chung là khung tham chiếu được tái cấu trúc sẽ được bộ giải mã video ở xa thu nhận (các lỗi nghi truyền).

Công cụ 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 hình ảnh tham chiếu 434 để tìm dữ liệu mẫu (là khối điểm ảnh tham chiếu ứng viên) hoặc siêu dữ liệu nhất định như vector chuyển động hình ảnh tham chiếu, hình khối, và v.v., điều đó có thể đóng vai trò như tham chiếu dự đoán thích hợp cho các hình ảnh mới. Bộ dự đoán 435 có thể hoạt động trên cơ sở khối mẫu từng điểm ảnh để tìm các tham chiếu dự đoán thích hợp. Trong một số trường hợp, được xác định bởi kết quả tìm kiếm do công cụ dự đoán 435 thu nhận, 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 phụ được sử dụng để mã hóa dữ liệu video.

Đầu ra của tất cả các bộ phận chức năng nói trên có thể được mã hóa entropy trong bộ mã hóa entropy 445. Bộ mã hóa entropy dịch các ký tự do các bộ phận chức năng khác nhau tạo ra thành chuỗi video được mã hóa, bằng cách nén không tổn hao các ký tự theo công nghệ mà người có hiểu biết trung bình trong lĩnh vực kỹ thuật biết, ví dụ mã hóa Huffman, mã hóa độ dài thay đổi, mã hóa số học, và v.v.

Bộ phát 440 có thể lưu trữ đệm (các) chuỗi video được mã hóa như được tạo bởi bộ mã hóa entropy 445 để chuẩn bị cho việc truyền qua kênh truyền thông 460, mà có thể là liên kết phần cứng/phần mềm tới thiết bị lưu trữ. 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 sẽ 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 quá trình mã hóa, bộ điều khiển 450 có thể gán cho mỗi bức ảnh được mã hóa một loại ảnh được mã hóa nhất định, điều này có thể ảnh hưởng đến các kỹ thuật mã hóa có thể được áp dụng cho bức ảnh tương ứng. Ví dụ, ảnh thường có thể được chỉ định là một trong các loại khung sau:

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

Hình ảnh dự đoán (Hình ảnh P) có thể được mã hóa và giải mã bằng cách sử dụng dự đoán nội bộ hoặc dự đoán liên sử dụng nhiều nhất một vectơ chuyển động và chỉ số tham chiếu để dự đoán các giá trị mẫu của mỗi khối.

Hình ảnh dự đoán kép (hình ảnh B) mà có thể được mã hóa và giải mã bằng cách sử dụng dự đoán nội bộ hoặc dự đoán liên bằng cách sử dụng nhiều nhất hai vectơ chuyển động và chỉ số tham chiếu để dự đoán các giá trị mẫu của mỗi khối. Tương tự, các ảnh đa dự đoán có thể sử dụng nhiều hơn hai ảnh tham chiếu và siêu dữ liệu liên quan để tái cấu trúc khối duy nhất.

Ảnh nguồn thường có thể được chia nhỏ theo không gian thành nhiều khối mẫu (ví dụ, các khối 4x4, 8x8, 4x8, hoặc 16x16 mỗi 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 hình ảnh tương ứng của khối. Ví dụ, các khối của hình 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 hình 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 hình ảnh tham chiếu đã được mã hóa trước đó.

Bộ mã hóa video 203 có thể thực hiện các hoạt động mã hóa theo công nghệ 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 dư thừa về thời gian và không gian trong chuỗi video đầu vào. Do đó, dữ liệu video được mã hóa có thể tuân theo cú pháp được chỉ định bởi công nghệ hoặc tiêu chuẩn mã hóa video đang được sử dụng.

Theo phương án, máy 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 thừa khác như hình ảnh và lớp thừa,

bản tin Thông tin Tăng cường Bổ sung (SEI), các đoạn tập tham số Thông tin Khả năng sử dụng Trực quan (VUI), và v.v.

Gần đây, tập hợp miền nén hoặc trích xuất của các phần hình ảnh độc lập về ngữ nghĩa thành hình 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 một thời điểm nhất định, có thể chọn sử dụng các yếu tố lấy mẫu lại khác nhau cho các hình ảnh độc lập về ngữ nghĩa khác nhau tạo nên toàn bộ 360 hoặc cảnh giám sát. Khi được kết hợp thành hình ảnh duy nhất, điều đó, theo thứ tự, yêu cầu phải thực hiện lấy lại mẫu hình ảnh tham chiếu, và báo hiệu mã hóa độ phân giải thích ứng khả dụng, cho các phần của hình ảnh được mã hóa.

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

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

Lấy mẫu lại Hình ảnh Tham chiếu (RPR) hoặc Thay đổi Độ phân giải Thích ứng (ARC) có thể đề cập đến các cơ chế cho phép thay đổi độ phân giải của hình

ảnh hoặc hình ảnh phụ trong chuỗi video được mã hóa, bằng cách, ví dụ, lấy mẫu lại hình ảnh tham chiếu. Từ đó, các tham số RPR / ARC 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 hình ảnh tham chiếu, các lệnh điều khiển khác nhau, và v.v.

Theo các phương án mã hóa và giải mã có thể được thực hiện trên hình 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 RPR / ARC độc lập và độ phức tạp bổ sung ngụ ý của nó, các tùy chọn để báo hiệu các tham số RPR / ARC sẽ được mô tả.

Đề cập đến các FIG.5A-5E, được thể hiện là một số phương án để báo hiệu các tham số RPR / 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 công nghệ mã hóa video có thể chọn một hoặc nhiều trong số các phương án này, hoặc các tùy chọn đã biết từ kỹ thuật liên quan, để báo hiệu các tham số RPR / ARC. Các phương án có thể không loại trừ lẫn nhau, và có thể được thay thế cho nhau dựa trên nhu cầu ứng dụng, công nghệ tiêu chuẩn liên quan, hoặc lựa chọn của bộ mã hóa.

Các loại tham số RPR / ARC có thể bao gồm:

- các yếu tố mẫu lên / xuống, tách biệt hoặc kết hợp theo kích thước X và Y
- các yếu tố mẫu lên / xuống, với việc bổ sung kích thước thời gian, chỉ báo tốc độ phóng to / thu nhỏ liên tục cho 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 trở vào bảng chứa (các) yếu tố.

- độ phân giải, theo kích thước X hoặc Y, theo đơn vị mẫu, khối, khối macro, đơn vị mã hóa (CU), hoặc bất kỳ độ chi tiết phù hợp nào khác, của hình ảnh đầu vào, hình ảnh đầu ra, hình ảnh tham chiếu, hình 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 tập các giá trị có thể được suy ra từ tập các giá trị khác. Như vậy có thể được kiểm soát, ví dụ, bằng cách sử dụng các lệnh. Để có ví dụ chi tiết hơn, hãy xem bên dưới.

- Tọa độ “cong vênh” tương tự như tọa độ được sử dụng trong H.263 Phụ lục P, lại có độ chi tiết phù hợp như đã mô tả ở trên. H.263 Phụ lục P xác định một cách hiệu quả để mã hóa các tọa độ cong vênh như vậy, nhưng mặt khác, có khả năng có nhiều cách hiệu quả cũng có thể được nghĩ ra. Ví dụ, độ dài thay đổi có thể đảo ngược, mã hóa theo kiểu “Huffman” của tọa độ cong vênh của Phụ lục P có thể được thay thế bằng mã hóa nhị phân có độ dài phù hợp, ví dụ, có thể được tính từ kích thước hình ảnh tối đa, có thể được nhân bản với yếu tố nhất định và bù đắp bởi giá trị nhất định, để cho phép "cong vênh" bên ngoài ranh giới của kích thước hình ảnh tối đa.

- tham số bộ lọc mẫu lên hoặc 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 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 tỷ lệ mẫu lên / xuống theo đó được báo hiệu theo bất kỳ cơ chế nào được đề cập ở trên, v.v.

Do đó, mô tả giả định mã hóa tập hợp hữu hạn các yếu tố mẫu lên / xuống (cùng yếu tố 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 như sử dụng mã Ext-Golomb phổ biến cho các phần tử cú pháp nhất định trong các tham số kỹ thuật mã hóa video như H.264 và H.265. Ánh xạ phù hợp của các giá trị với các yếu tố mẫu lên / xuống có thể, ví dụ, theo Bảng 1:

BẢNG 1

Từ mã	Mã Ext-Golomb	Độ phân giải Gốc/Mục tiêu
0	1	1 / 1
1	010	1 / 1,5 (tăng tỷ lệ lên 50%)
2	011	1,5 / 1 (giảm tỷ lệ xuống 50%)
3	00100	1 / 2 (tăng tỷ lệ lên 100%)
4	00101	2 / 1 (giảm tỷ lệ xuống 100%)

Nhiều ánh xạ tương tự có thể được tạo ra theo nhu cầu của ứng dụng và khả năng của cơ chế tăng và giảm tỷ lệ khả dụng trong công nghệ hoặc tiêu chuẩn nén video. Bảng có thể được mở rộng đến nhiều giá trị hơn. Giá trị cũng có thể được biểu thị bằng cơ chế mã hóa entropy khác với mã Ext-Golomb, ví dụ sử dụng mã nhị phân. Điều đó có thể có những lợi thế nhất định khi các yếu tố lấy mẫu lại được quan tâm bên ngoài chính các công cụ xử lý video (bộ mã hóa và bộ giải mã), ví dụ như các MANE. Cần lưu ý rằng, đối với các tình huống không cần thay đổi độ phân giải, mã Ext-Golomb có thể được chọn mà ngắn; trong bảng trên, chỉ bit duy nhất. Điều đó 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 các mục trong bảng, cũng như ngữ nghĩa của chúng, có thể được cấu hình đầy đủ hoặc một phần. Ví dụ, đường biên cơ bản của bảng có thể được chuyển tải trong tập tham số “cao” như tập tham số trình tự hoặc bộ giải mã. Theo các phương án, một hoặc nhiều bảng như vậy có thể được xác định trong công nghệ hoặc tiêu chuẩn mã hóa video và có thể được chọn thông qua ví dụ như bộ giải mã hoặc tập tham số trình tự.

Dưới đây là mô tả cách một yếu tố mẫu lên / xuống (thông tin ARC), được mã hóa như mô tả ở trên, có thể được đưa vào công nghệ 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 mẫu lên / xuống. Xem bên dưới để thảo luận khi cần lượng dữ liệu tương đối lớn cho bộ lọc hoặc các cấu trúc dữ liệu khác.

Như trong FIG.5, H.263 Phụ lục P bao gồm thông tin ARC 502 dưới dạng bốn tọa độ cong vênh vào tiêu đề hình 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 hình ảnh, và b) dự kiến sẽ có những thay đổi thường xuyên về thông tin ARC. Tuy nhiên, chi phí khi sử dụng báo hiệu kiểu H.263 có thể khá cao, và các yếu tố tỷ lệ có thể không xác định giữa các ranh giới hình ảnh vì tiêu đề hình ảnh có thể mang tính chất thoáng qua.

Theo cùng một phương án hoặc phương án khác, báo hiệu của các tham số ARC có thể tuân theo ví dụ chi tiết như được nêu trong các FIG.6A-6B. Các FIG.6A-6B mô tả sơ đồ cú pháp trong kiểu biểu diễn bằng cách sử dụng ký tự mà gần như tuân theo lập trình kiểu C, chẳng hạn như được sử dụng trong các tiêu chuẩn mã hóa video ít nhất từ năm 1993. Các dòng in đậm biểu thị các phần tử cú pháp có trong dòng bit, các dòng không có chữ in đậm thường chỉ ra luồng điều khiển hoặc cài đặt của các biến.

Như trong FIG.6A, tiêu đề nhóm ô 601 như cấu trúc cú pháp mẫu của tiêu đề áp dụng cho một phần (có thể là hình chữ nhật) 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 có mặt của phần tử cú pháp này trong tiêu đề nhóm ô có thể được kiểm soát dựa trên việc sử dụng độ phân giải thích ứng 603 —trong đó, giá trị của lệnh không được mô tả bằng chữ in đậm, mà có nghĩa là lệnh đó có mặt trong dòng bit tại điểm mà nó xảy ra trong sơ đồ cú pháp. Độ phân giải thích ứng có được sử dụng cho bức ảnh này hay không hoặc các phần của nó có thể được báo hiệu trong bất kỳ cấu trúc cú pháp

mức cao nào bên trong hoặc bên ngoài dòng bit. Trong ví dụ được hiển thị, nó được báo hiệu trong tập tham số trình tự như được nêu bên dưới.

Đề cập đến FIG.6B, được hiển thị cũng là trích dẫn của tập tham số trình tự 610. Phần tử cú pháp thứ nhất được hiển thị là `adaptive_pic_resolution_change_flag` 611. Khi đúng, lệnh đó có thể chỉ ra việc sử dụng độ phân giải thích ứng, do đó có thể yêu cầu một số thông tin điều khiển nhất định. Trong ví dụ, thông tin điều khiển như vậy có điều kiện dựa trên giá trị của lệnh 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`, chúng cùng có thể xác định độ phân giải của hình ảnh đầu ra. Ở những nơi khác trong công nghệ hoặc tiêu chuẩn mã hóa video, các hạn chế nhất định đối với giá trị có thể được xác định. Ví dụ, định nghĩa mức 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 công nghệ hoặc tiêu chuẩn mã hóa video nhất định, hoặc các công nghệ hoặc tiêu chuẩn bên ngoài, chẳng hạn như, ví dụ, 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 này.

Trong các ứng dụng nhất định, có thể khuyến khích rằng bộ mã hóa lệnh bộ giải mã để sử dụng kích thước hình ảnh tham chiếu nhất định thay vì mặc nhiên cho rằng kích thước đó là kích thước hình ảnh đầu ra. Trong ví dụ này, phần tử cú pháp `reference_pic_size_present_flag` 614 chỉ báo sự có mặt có điều kiện của kích thước hình 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 hình ảnh. Ví dụ, bảng như vậy có thể được thể hiện 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 mục nhập bảng sẽ có mặt. Nếu giá trị là năm, có sáu mục nhập bảng. Đối với mỗi "dòng" trong bảng, chiều rộng và chiều cao của hình ảnh đã được giải mã sau đó được đưa vào cú pháp 617.

Các mục trong bảng được thể hiện 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.

Các công nghệ hoặc tiêu chuẩn mã hóa video nhất định, ví dụ VP9, hỗ trợ khả năng mở rộng không gian bằng cách triển khai một số hình thức lấy mẫu lại hình ảnh tham chiếu nhất định (được báo hiệu hơi khác so với đối tượng sáng chế) kết hợp với khả năng mở rộng theo thời gian, do đó, để cho phép khả năng mở rộng không gian. Đặc biệt, một số hình ảnh tham chiếu nhất định có thể được lấy mẫu bổ sung bằng cách sử dụng công nghệ kiểu ARC lên độ phân giải cao hơn để tạo thành cơ sở của lớp tăng cường không gian. Những bức ảnh được lấy mẫu đó có thể được tinh chỉnh, sử dụng cơ chế dự đoán thông thường ở độ phân giải cao, để thêm chi tiết.

Các phương án được thảo luận ở đây có thể được sử dụng trong môi trường như vậy. Trong một số trường hợp nhất định, theo cùng hoặc phương án khác, giá trị trong tiêu đề đơn vị NAL, ví dụ như trường ID Thời gian, có thể được sử dụng để chỉ ra 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 thời gian dựa trên tiêu đề đơn vị NAL giá trị ID 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 hình ảnh được mã hóa và lớp tạm thời được chỉ báo bởi trường ID thời gian trong tiêu đề đơn vị NAL.

Gần đây, tập hợp miền nén hoặc trích xuất của nhiều phần hình ảnh độc lập về ngữ nghĩa thành hình 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 các 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 một thời điểm nhất định. Nói cách khác, các bộ mã hóa, tại một thời điểm nhất định, có thể chọn sử dụng các yếu tố lấy mẫu lại khác nhau cho các hình ảnh độc lập về ngữ nghĩa khác nhau tạo nên toàn bộ 360 hoặc cảnh giám sát. Khi được kết hợp thành hình ảnh duy nhất, mà, theo thứ tự, yêu cầu thực hiện lấy mẫu lại hình ảnh tham chiếu và báo hiệu mã hóa độ phân giải thích ứng khả dụng cho các phần của hình ảnh được mã hóa.

Theo các phương án, cửa sổ tương thích có thể là một phần phụ hình chữ nhật của hình ảnh, mà trong đó bộ mã hóa đã chỉ báo rằng các mẫu của phần phụ đó được xuất ra bởi bộ giải mã, trong khi các mẫu bên ngoài cửa sổ có thể bị bỏ qua khỏi đầu ra bởi bộ giải mã.

Vấn đề được giải quyết ở đây là cửa sổ tương thích có thể phục vụ hai mục đích. Mục đích là xác định kích thước đầu ra như mô tả ở trên. Kích thước đầu ra có thể được điều khiển bởi ứng dụng và có thể hoàn toàn khác với kích thước hình ảnh được mã hóa, ví dụ như trong các ứng dụng bản đồ lập thể hoặc 360 hình khối. Mục đích thứ hai có thể là cửa sổ tương thích cũng là đầu vào cho việc lấy mẫu lại hình ảnh tham chiếu được mô tả ở trên.

Khi có sự kết hợp nhất định giữa kích thước cửa sổ tương thích và kích thước hình ảnh cho cả hình ảnh được mã hóa và lấy mẫu lại, có thể rằng các mẫu nằm ngoài vùng đệm có thể được tham chiếu. Mặc dù vấn đề này có thể được giải quyết thông qua ràng buộc dòng bit về kích thước cửa sổ tương thích, nhưng giải

pháp đó sẽ có nhược điểm là cửa sổ tương thích không được lựa chọn tự do bởi ứng dụng nữa. Tách hai chức năng, trong một số ứng dụng có thể có lợi.

Theo phương án, kích thước cửa sổ tương thích có thể được báo hiệu trong PPS. Các tham số cửa sổ tương thích có thể được sử dụng để tính toán tỷ lệ lấy mẫu lại, khi kích thước cửa sổ tương thích của hình ảnh tham chiếu khác với kích thước của hình ảnh hiện tại. Bộ giải mã có thể cần phải nhận ra kích thước cửa sổ tương thích của mỗi hình ảnh, để xác định xem liệu quá trình lấy mẫu lại có cần thiết hay không.

Theo phương án, hệ số tỷ lệ để lấy mẫu lại hình ảnh tham chiếu (RPR) có thể được tính toán dựa trên chiều rộng đầu ra và chiều cao đầu ra, mà có thể được suy ra từ các tham số cửa sổ tương thích, giữa hình ảnh hiện tại và hình ảnh được tham chiếu. Điều này có thể cho phép hệ số tỷ lệ được tính toán chính xác hơn, so với việc sử dụng kích thước hình ảnh đã được giải mã. Điều này có thể hoạt động tốt đối với hầu hết các chuỗi video, mà có kích thước hình ảnh đầu ra gần giống với kích thước hình ảnh được giải mã, với các vùng đệm nhỏ.

Tuy nhiên, điều này cũng có thể gây ra nhiều vấn đề khác nhau. Ví dụ, khi kích thước cửa sổ tương thích hoàn toàn khác với kích thước hình ảnh được giải mã có giá trị bù lớn, đối với các ứng dụng đa phương tiện (ví dụ: 360 cubemap, lập thể, đám mây điểm), việc tính toán hệ số tỷ lệ dựa trên kích thước cửa sổ tương thích có thể không đảm bảo chất lượng của dự đoán liên với các độ phân giải khác nhau. Trong trường hợp cực đoan, vùng sắp xếp của CU hiện tại trong hình ảnh tham chiếu có thể không có. Khi RPR được sử dụng cho khả năng mở rộng với nhiều lớp, các hiệu số của cửa sổ tương thích có thể không được sử dụng để tính toán vùng được tham chiếu giữa các lớp. Lưu ý rằng vùng được tham chiếu của mỗi lớp phụ thuộc trực tiếp có thể được báo hiệu rõ ràng trong PPS-mở rộng, trong Mở rộng khả năng biến đổi tỷ lệ HEVC (SHVC). Khi dòng bit phụ nhắm mục tiêu đến vùng cụ thể (ảnh phụ) được trích xuất từ toàn bộ dòng bit, kích thước cửa sổ tương thích hoàn toàn không khớp với kích thước ảnh. Lưu ý rằng các tham

số cửa sổ tương thích không thể được cập nhật miễn là các tham số được sử dụng để tính toán tỷ lệ, sau khi dòng bit được mã hóa.

Dựa trên các vấn đề tiềm ẩn ở trên, việc tính toán các hệ số tỷ lệ dựa trên kích thước cửa sổ tương thích có thể có các trường hợp góc, mà các tham số thay thế được yêu cầu. Như dự phòng, khi các tham số cửa sổ tương thích có thể không được sử dụng để tính toán các hệ số tỷ lệ, nó được đề xuất để báo hiệu các tham số vùng tham chiếu mà có thể được sử dụng để tính toán tỷ lệ cho RPR và Khả năng mở rộng.

Theo phương án, đề cập đến FIG.7, `conformance_window_flag` có thể được báo hiệu trong PPS. `conformance_window_flag` bằng 1 có thể chỉ báo rằng các tham số độ dịch cửa sổ cắt xén phù hợp tiếp theo trong PPS. `conformance_window_flag` bằng 0 có thể chỉ báo rằng các tham số độ dịch cửa sổ cắt xén phù hợp không có mặt.

Theo phương án, vẫn đề cập đến FIG.7, `conf_win_left_offset`, `conf_win_right_offset`, `conf_win_top_offset` và `conf_win_bottom_offset` chỉ định các mẫu của ảnh tham chiếu đến PPS mà được xuất ra từ quá trình giải mã, dưới dạng vùng hình chữ nhật được chỉ định trong tọa độ ảnh cho đầu ra. Khi `conformance_window_flag` bằng 0, các giá trị của `conf_win_left_offset`, `conf_win_right_offset`, `conf_win_top_offset` và `conf_win_bottom_offset` có thể được suy ra bằng 0.

Theo phương án, cửa sổ cắt xén phù hợp có thể chứa các mẫu độ chói có tọa độ ảnh ngang từ $\text{SubWidthC} * \text{conf_win_left_offset}$ đến $\text{pic_width_in_luma_samples} - (\text{SubWidthC} * \text{conf_win_right_offset} + 1)$ và tọa độ ảnh dọc từ $\text{SubHeightC} * \text{conf_win_left_offset}$ đến $\text{pic_width_in_luma_samples} - (\text{SubHeightC} * \text{conf_win_right_offset} + 1)$, bao gồm.

Giá trị của $\text{SubWidthC} * (\text{conf_win_left_offset} + \text{conf_win_right_offset})$ có thể nhỏ hơn $\text{pic_width_in_luma_samples}$, và giá trị của $\text{SubHeightC} *$

$(\text{conf_win_top_offset} + \text{conf_win_bottom_offset})$ sẽ nhỏ hơn $\text{pic_height_in_luma_samples}$.

Các biến PicOutputWidthL và PicOutputHeightL có thể được dẫn xuất như thể hiện trong Công thức 1 và Công thức 2 bên dưới:

$$\text{PicOutputWidthL} = \text{pic_width_in_luma_samples} -$$

$$\text{SubWidthC} * (\text{conf_win_right_offset} + \text{conf_win_left_offset}) \quad (\text{Công thức 1})$$

$$\text{PicOutputHeightL} = \text{pic_height_in_pic_size_units} -$$

$$\text{SubHeightC} * (\text{conf_win_bottom_offset} + \text{conf_win_top_offset}) \quad (\text{Công thức 2})$$

Khi ChromaArrayType không bằng 0, các mẫu được chỉ định tương ứng của hai mảng màu có thể là các mẫu có tọa độ ảnh ($x / \text{SubWidthC}$, $y / \text{SubHeightC}$), mà trong đó (x, y) là tọa độ ảnh của các mẫu độ chói được chỉ định.

Theo phương án, lệnh có thể xuất hiện trong PPS, hoặc tập tham số khác, và có thể chỉ báo kích thước hình ảnh lấy mẫu lại (chiều rộng và chiều cao) có được báo hiệu rõ ràng hay không trong PPS hoặc, hoặc tập tham số khác. Nếu các tham số kích thước hình ảnh lấy mẫu lại được báo hiệu rõ ràng, tỷ lệ lấy mẫu lại giữa hình ảnh hiện tại và hình ảnh tham chiếu có thể được tính toán dựa trên các tham số kích thước hình ảnh lấy mẫu lại.

Theo phương án, đề cập đến FIG.8A, $\text{use_conf_win_for_rpr_flag}$ bằng 0 có thể chỉ báo rằng $\text{resampled_pic_width_in_luma_samples}$ và $\text{resampled_pic_height_in_luma_samples}$ theo sau tại vị trí thích hợp, ví dụ như tiếp theo trong PPS.

Trong phương án, $\text{use_conf_win_for_rpr_flag}$ bằng 1 có thể chỉ báo rằng $\text{resampling_pic_width_in_luma_samples}$ và $\text{resampling_pic_height_in_luma_samples}$ không có mặt.

Trong phương án, `resampling_pic_width_in_luma_samples` có thể chỉ định chiều rộng của mỗi hình ảnh tham chiếu đề cập đến PPS theo đơn vị mẫu độ chói, để lấy mẫu lại. `resampling_pic_width_in_luma_samples` có thể không bằng 0, có thể là bội số nguyên của `Max (8, MinCbSizeY)`, và có thể nhỏ hơn hoặc bằng `pic_width_max_in_luma_samples`.

Trong phương án, `resampling_pic_height_in_luma_samples` có thể chỉ định chiều cao của mỗi hình ảnh tham chiếu đề cập đến PPS theo đơn vị của mẫu độ chói, để lấy mẫu lại. `resampling_pic_height_in_luma_samples` có thể không bằng 0 và có thể là bội số nguyên của `Max (8, MinCbSizeY)`, và có thể nhỏ hơn hoặc bằng `pic_height_max_in_luma_samples`.

Khi phần tử cú pháp `resampling_pic_width_in_luma_samples` không có, giá trị của `resampling_pic_width_in_luma_samples` có thể bằng `PicOutputWidthL`.

Khi phần tử cú pháp `resampling_pic_height_in_luma_samples` không có, giá trị của `resampling_pic_height_in_luma_samples` có thể bằng `PicOutputHeightL`.

Theo phương án, ví dụ về quy trình nội suy phân số với việc lấy mẫu lại hình ảnh tham chiếu có thể được xử lý như sau:

Đầu vào cho quá trình này có thể là: vị trí độ chói (`xSb`, `ySb`) chỉ định mẫu trên cùng bên trái của khối phụ mã hóa hiện tại so với mẫu độ chói trên cùng bên trái của hình ảnh hiện tại, biến `sbWidth` chỉ định độ rộng của khối con mã hóa hiện tại, biến số `sbHeight` chỉ định chiều cao của khối con mã hóa hiện tại, `mvOffset` độ dịch vector chuyển động, vector chuyển động tinh chỉnh `refMvLX`, mảng mẫu hình ảnh tham chiếu đã chọn `refPicLX`, chỉ số bộ lọc nội suy nửa mẫu `hpellIdx`, lệnh lưu lượng quang kép `bdofFlag`, và biến số `cIdx` xác định chỉ số thành phần màu của khối hiện tại.

Kết quả đầu ra của quá trình này có thể là: $(sbWidth + brdExtSize) \times (sbHeight + brdExtSize)$ mảng `predSamplesLX` của các giá trị mẫu dự đoán.

Kích thước phần mở rộng đường biên khối dự đoán brdExtSize có thể được suy ra theo Công thức 3, như sau:

$$\text{brdExtSize} = (\text{bdofFlag} \quad || \quad (\text{inter_affine_flag}[\text{xSb}][\text{ySb}] \quad \&\& \text{sps_affine_prof_enabled_flag})) ? 2 : 0 \quad (\text{Công thức 3})$$

Biến fRefWidth có thể được đặt bằng resampling_pic_width_in_luma_samples của hình ảnh tham chiếu trong các mẫu độ chói.

Biến fRefHeight có thể được đặt bằng resampling_pic_height_in_luma_samples của hình ảnh tham chiếu trong các mẫu độ chói.

Vector chuyển động mvLX có thể được đặt bằng (refMvLX - mvOffset).

Nếu cIdx bằng 0, các hệ số tỷ lệ và biểu diễn điểm cố định của chúng có thể được xác định theo Công thức 4 và Công thức 5 dưới đây:

$$\begin{aligned} \text{hori_scale_fp} &= \\ &((\text{fRefWidth} \ll 14) + (\text{resampling_pic_width_in_luma_samples} \\ &\gg 1)) / \text{resampling_pic_width_in_luma_samples} \quad (\text{Công thức 4}) \end{aligned}$$

$$\begin{aligned} \text{vert_scale_fp} &= \\ &((\text{fRefHeight} \ll 14) + (\text{resampling_pic_height_in_luma_samples} \\ &\gg 1)) / \text{resampling_pic_height_in_luma_samples} \quad (\text{Công thức 5}) \end{aligned}$$

Gọi (xIntL, yIntL) là vị trí độ chói được đưa ra trong các đơn vị mẫu đầy đủ và (xFracL, yFracL) là độ dịch được đưa ra trong 1/16 đơn vị mẫu. Các biến này có thể được sử dụng để chỉ định vị trí mẫu phân số bên trong mảng mẫu tham chiếu refPicLX.

Tọa độ trên cùng bên trái của khối giới hạn cho đệm mẫu tham chiếu (xSbIntL, ySbIntL) có thể được đặt bằng (xSb + (mvLX[0] >> 4), ySb + (mvLX[1] >> 4)).

Đối với mỗi vị trí mẫu độ chói ($x_L = 0..sbWidth - 1 + brdExtSize$, $y_L = 0..sbHeight - 1 + brdExtSize$) bên trong mảng mẫu độ chói dự đoán $predSamplesLX$, giá trị mẫu độ chói dự đoán tương ứng $predSamplesLX[x_L][y_L]$ có thể được suy ra như sau:

- Gọi ($refxSb_L$, $refySb_L$) và ($refx_L$, $refy_L$) là các vị trí lớn được trỏ tới bởi vector chuyển động ($refMvLX$, $refMvLX$) được cho trong đơn vị mẫu 1/16. Các biến $refxSb_L$, $refx_L$, $refySb_L$, và $refy_L$ có thể được dẫn xuất như thể hiện trong Công thức 6 đến Công thức 9 như được hiển thị bên dưới:

$refxSb_L = ((xSb \ll 4) + refMvLX[0]) * hori_scale_fp$ (Công thức 6)

$refx_L = ((Sign(refxSb) * ((Abs(refxSb) + 128) \gg 8) + x_L * ((hori_scale_fp + 8) \gg 4)) + 32) \gg 6$ (Công thức 7)

$refySb_L = ((ySb \ll 4) + refMvLX[1]) * vert_scale_fp$ (Công thức 8)

$refy_L = ((Sign(refySb) * ((Abs(refySb) + 128) \gg 8) + y_L * ((vert_scale_fp + 8) \gg 4)) + 32) \gg 6$ (Công thức 9)

- Các biến $xInt_L$, $yInt_L$, $xFrac_L$ và $yFrac_L$ có thể được dẫn xuất như thể hiện trong Công thức 10 đến Công thức 13 như hình dưới đây:

$xInt_L = refx_L \gg 4$ (Công thức 10)

$yInt_L = refy_L \gg 4$ (Công thức 11)

$xFrac_L = refx_L \& 15$ (Công thức 12)

$yFrac_L = refy_L \& 15$ (Công thức 13)

Nếu $bdoFlag$ bằng TRUE hoặc ($sps_affine_prof_enabled_flag$ bằng TRUE và $inter_affine_flag[xSb][ySb]$ bằng TRUE), và một hoặc nhiều điều kiện sau là đúng, giá trị mẫu dự đoán là $predSamplesLX[x_L][y_L]$ có thể được bắt nguồn bằng cách gọi quy trình tìm nạp mẫu số nguyên độ chói với ($xInt_L + (xFrac_L \gg 3) - 1$), $yInt_L + (yFrac_L \gg 3) - 1$ và $refPicLX$ làm đầu vào.

- xL bằng 0.
- xL bằng sbWidth + 1.
- yL bằng 0.
- yL bằng sbHeight + 1.

Nếu không, giá trị mẫu độ chói dự đoán predSamplesLX [xL] [yL] được suy ra bằng cách gọi quy trình lọc nội suy 8 chấm mẫu độ chói với (xIntL - (brdExtSize > 0 ? 1 : 0), yIntL - (brdExtSize > 0 ? 1 : 0)), (xFracL, yFracL), (xSbIntL, ySbIntL), refPicLX, hpelIdx, sbWidth, sbHeight và (xSb, ySb) làm đầu vào.

Nếu không (cIdx không bằng 0), điều sau có thể áp dụng:

Gọi (xIntC, yIntC) là vị trí sắc độ được đưa ra trong các đơn vị mẫu đầy đủ và (xFracC, yFracC) là độ dịch được cho trong đơn vị mẫu 1/32. Các biến này có thể được sử dụng để chỉ định vị trí mẫu phân số chung bên trong mảng mẫu tham chiếu refPicLX.

Tọa độ trên cùng bên trái của khối giới hạn cho đệm mẫu tham chiếu (xSbIntC, ySbIntC) có thể được đặt bằng ((xSb / SubWidthC) + (mvLX [0] >> 5), (ySb / SubHeightC) + (mvLX [1] >> 5)).

Đối với mỗi vị trí mẫu màu sắc (xC = 0..sbWidth - 1, yC = 0 .. sbHeight - 1) bên trong mảng mẫu màu sắc dự đoán predSamplesLX, giá trị mẫu màu sắc dự đoán tương ứng predSamplesLX [xC] [yC] có thể được suy ra như sau :

- Gọi (refxSbC, refySbC) và (refxC, refyC) là các vị trí màu được trở tới bởi vector chuyển động (mvLX [0], mvLX [1]) được cho trong đơn vị mẫu 1/32. Các biến refxSbC, refySbC, refxC và refyC có thể được dẫn xuất như thể hiện trong Công thức 14 đến Công thức 17 như được hiển thị bên dưới:

$$\text{refxSb}_C = ((xSb / \text{SubWidth}_C \ll 5) + mvLX[0]) * \text{hori_scale_fp}$$

(Công thức 14)

$$\text{refx}_C = ((\text{Sign}(\text{refxSb}_C) * ((\text{Abs}(\text{refxSb}_C) + 256) \gg 9) + xC * ((\text{hori_scale_fp} + 8) \gg 4)) + 16) \gg 5$$

(Công thức 15)

$$\text{refySb}_C = ((ySb / \text{SubHeight}_C \ll 5) + mvLX[1]) * \text{vert_scale_fp}$$

(Công thức 16)

$$\text{refy}_C = ((\text{Sign}(\text{refySb}_C) * ((\text{Abs}(\text{refySb}_C) + 256) \gg 9) + \text{yC} * ((\text{vert_scale_fp} + 8) \gg 4) + 16) \gg 5) \quad (\text{Công thức 17})$$

- Các biến xInt_C , yInt_C , xFrac_C và yFrac_C có thể được suy ra từ Công thức 18 đến Công thức 21 như hình dưới đây:

$$\text{xInt}_C = \text{refx}_C \gg 5 \quad (\text{Công thức 18})$$

$$\text{yInt}_C = \text{refy}_C \gg 5 \quad (\text{Công thức 19})$$

$$\text{xFrac}_C = \text{refy}_C \& 31 \quad (\text{Công thức 20})$$

$$\text{yFrac}_C = \text{refy}_C \& 31 \quad (\text{Công thức 21})$$

Giá trị mẫu dự đoán $\text{predSamplesLX}[\text{xC}][\text{yC}]$ có thể được lấy bằng cách gọi quy trình với $(\text{xInt}_C, \text{yInt}_C)$, $(\text{xFrac}_C, \text{yFrac}_C)$, $(\text{xSbInt}_C, \text{ySbInt}_C)$, sbWidth , sbHeight và refPicLX làm đầu vào.

Theo phương án, đề cập đến FIG.8B, $\text{use_conf_win_for_rpr_flag}$ bằng 0 có thể chỉ định rằng $\text{resampled_pic_width_in_luma_samples}$ và $\text{resampled_pic_height_in_luma_samples}$ tiếp theo trong PPS. $\text{use_conf_wid_for_rpr_flag}$ bằng 1 chỉ định rằng $\text{resampling_pic_width_in_luma_samples}$ và $\text{resampling_pic_height_in_luma_samples}$ không có mặt.

Trong phương án, $\text{ref_region_left_offset}$ có thể chỉ định độ dịch ngang giữa mẫu độ chói trên cùng bên trái của vùng tham chiếu trong hình ảnh được giải mã. Giá trị của $\text{ref_region_left_offset}$ sẽ nằm trong khoảng từ -2^{14} đến $2^{14} - 1$, bao gồm. Khi không có mặt, giá trị của $\text{ref_region_left_offset}$ có thể được suy ra bằng $\text{conf_win_left_offset}$.

Trong phương án, $\text{ref_region_top_offset}$ có thể chỉ định độ dịch dọc giữa mẫu độ chói trên cùng bên trái của vùng tham chiếu trong hình ảnh được giải mã. Giá trị của $\text{ref_region_top_offset}$ sẽ nằm trong khoảng từ -2^{14} đến $2^{14} - 1$, bao gồm. Khi không xuất hiện, giá trị của $\text{ref_region_top_offset}$ có thể được suy ra bằng $\text{conf_win_right_offset}$.

Theo phương án, `ref_region_right_offset` có thể chỉ định độ dịch ngang giữa mẫu độ chói dưới cùng bên phải của vùng tham chiếu trong hình ảnh được giải mã. Giá trị của `ref_layer_right_offset` sẽ nằm trong khoảng từ -2^{14} đến $2^{14} - 1$, bao gồm. Khi không xuất hiện, giá trị của `ref_region_right_offset` có thể được suy ra bằng `conf_win_top_offset`.

Theo phương án, `ref_region_bottom_offset` có thể chỉ định độ dịch dọc giữa mẫu độ chói phía dưới bên phải của vùng tham chiếu trong hình ảnh được giải mã. Giá trị của `ref_layer_bottom_offset` sẽ nằm trong khoảng từ -214 đến $214 - 1$, bao gồm cả. Khi không có mặt, giá trị của `ref_region_bottom_offset` [`ref_loc_offset_layer_id` [`i`]] có thể được suy ra bằng `conf_win_bottom_offset`.

Các biến `PicRefWidthL` và `PicRefHeightL` có thể được dẫn xuất như thể hiện trong Công thức 22 và Công thức 23, như được hiển thị bên dưới:

$$\text{PicRefWidthL} = \text{pic_width_in_luma_samples} - \text{SubWidthC} * (\text{ref_region_right_offset} + \text{ref_region_left_offset})$$

(Công thức 22)

$$\text{PicRefHeightL} = \text{pic_height_in_pic_size_units} - \text{SubHeightC} * (\text{ref_region_bottom_offset} + \text{ref_region_top_offset})$$

(Công thức 23)

Biến `fRefWidth` có thể được đặt bằng `PicRefWidthL` của hình ảnh tham chiếu trong các mẫu độ chói.

Biến `fRefHeight` có thể được đặt bằng `PicRefHeightL` của hình ảnh tham chiếu trong các mẫu độ chói.

Vectơ chuyển động `mvLX` có thể được đặt bằng $(\text{refMvLX} - \text{mvOffset})$.

Nếu `cIdx` bằng 0, các hệ số tỷ lệ và hiển thị điểm cố định của chúng có thể được xác định như thể hiện trong Công thức 24 và Công thức 25, như được hiển thị bên dưới:

$$\text{hori_scale_fp} = ((\text{fRefWidth} \ll 14) + (\text{PicRefWidthL} \gg 1)) / \text{PicRefWidthL}$$

(Công thức 24)

$$\text{vert_scale_fp} = ((\text{fRefHeight} \ll 14) + (\text{PicRefHeightL} \gg 1)) / \text{PicRefHeight}$$

(Công thức 25)

Tọa độ trên cùng bên trái của khối giới hạn cho đệm mẫu tham chiếu (x_{SbIntL} , y_{SbIntL}) có thể được đặt bằng ($x_{\text{Sb}} + (\text{mvLX}[0] \gg 4)$, $y_{\text{Sb}} + (\text{mvLX}[1] \gg 4)$).

Đối với mỗi vị trí mẫu độ chói ($x_L = 0..sbWidth - 1 + \text{brdExtSize}$, $y_L = 0..sbHeight - 1 + \text{brdExtSize}$) bên trong mảng mẫu độ chói dự đoán predSamplesLX , giá trị mẫu độ chói dự đoán tương ứng $\text{predSamplesLX}[x_L][y_L]$ có thể được suy ra như sau:

- Gọi (refxSbL , refySbL) và (refxL , refyL) là các vị trí độ chói được trỏ tới bởi vector chuyển động (refMvLX , refMvLY) được cho trong đơn vị mẫu 1/16. Các biến refxSbL , refxL , refySbL , và refyL có thể được suy ra như thể hiện trong Công thức 26 đến Công thức 29 như được hiển thị bên dưới:

$$\text{refxSbL} = ((x_{\text{Sb}} + \text{ref_region_left_offset}) \ll 4) + \text{refMvLX}[0] * \text{hori_scale_fp}$$

(Công thức 26)

$$\text{refxL} = ((\text{Sign}(\text{refxSb}) * (\text{Abs}(\text{refxSb}) + 128) \gg 8) + x_L * ((\text{hori_scale_fp} + 8) \gg 4) + 32) \gg 6$$

(Công thức 27)

$$\text{refySbL} = ((y_{\text{Sb}} + \text{ref_region_top_offset}) \ll 4) + \text{refMvLY}[1] * \text{vert_scale_fp}$$

(Công thức 28)

$$\text{refyL} = ((\text{Sign}(\text{refySb}) * (\text{Abs}(\text{refySb}) + 128) \gg 8) + y_L * ((\text{vert_scale_fp} + 8) \gg 4) + 32) \gg 6$$

(Công thức 29)

Theo phương án, đề cập đến FIG.9, khi tỷ lệ lấy mẫu lại để lấy mẫu lại hình ảnh tham chiếu được tính toán dựa trên kích thước của sổ tương thích, bộ giải mã thực hiện quy trình giải mã khối hiện tại với dự đoán bù chuyển động liên hình ảnh có thể truy cập một hoặc nhiều điểm ảnh trong sổ các điểm ảnh được giải mã trong hình ảnh tham chiếu. Trong một số trường hợp, quá trình giải mã có thể sử

dụng các hình ảnh được giải mã không khả dụng trong DPB, do đó có thể dẫn đến sự cố bộ giải mã hoặc hành vi giải mã không mong muốn.

Ví dụ, như minh họa trong FIG.9, dựa trên hình ảnh được giải mã có kích thước hình ảnh được giải mã 903 và kích thước cửa sổ tương thích 904, và hình ảnh tham chiếu có kích thước hình ảnh được giải mã 901 và kích thước cửa sổ tương thích 902, quá trình giải mã cho khối hiện tại 906 có thể sử dụng khối tham chiếu 905 không khả dụng.

Để giải quyết vấn đề như vậy, báo hiệu rõ ràng về kích thước ảnh lấy mẫu lại có thể được sử dụng để tính toán tỷ lệ lấy mẫu lại để lấy mẫu lại ảnh tham chiếu, trong một số trường hợp.

Theo phương án, kích thước cửa sổ tương thích có thể không thích hợp để sử dụng cho việc tính toán tỷ lệ lấy mẫu lại để lấy mẫu lại hình ảnh tham chiếu, khi kích thước cửa sổ tương thích không được căn chỉnh với tỷ lệ lấy mẫu lại thực tế. Ví dụ, kích thước cửa sổ tương thích có thể khác với tỷ lệ lấy mẫu lại mong muốn, khi hình ảnh hiện tại bao gồm các hình ảnh phụ lập thể (trái và phải) hoặc nhiều mặt của hình ảnh 360 (ví dụ, sáu mặt của phép chiếu cubemap). Trong một số trường hợp, kích thước cửa sổ tương thích có thể bao phủ vùng phụ của hình ảnh được giải mã. Trong những trường hợp đó, kích thước hình ảnh lấy mẫu lại có thể được sử dụng để tính toán tỷ lệ lấy mẫu lại.

FIG.10 là lưu đồ là quy trình ví dụ quy trình 1000 để giải mã dòng bit video được mã hóa. Trong một số triển khai, một hoặc nhiều khối quy trình của FIG.10 có thể được thực hiện bởi bộ giải mã 210. Trong một số triển khai, một hoặc nhiều khối quá trình của FIG.10 có thể được thực hiện bởi thiết bị khác hoặc nhóm thiết bị tách biệt với hoặc bao gồm bộ giải mã 210, chẳng hạn như bộ mã hóa 203.

Như trong FIG.10, quy trình 1000 có thể bao gồm việc thu nhận cờ thứ nhất chỉ báo rằng cửa sổ tương thích có trong hình ảnh hiện tại (khối 1010).

Như thể hiện thêm trong FIG.10, quy trình 1000 có thể bao gồm, dựa trên cờ thứ nhất chỉ báo rằng cửa sổ tương thích hiện có, thu nhận cờ thứ hai chỉ báo

liệu cửa sổ tương thích có được sử dụng để lấy mẫu lại hình ảnh tham chiếu (khối 1020).

Như thể hiện thêm trong FIG.10, quy trình 1000 có thể bao gồm việc xác định xem cờ thứ hai có chỉ báo rằng cửa sổ tương thích được sử dụng để lấy mẫu lại hình ảnh tham chiếu hay không (khối 1030). Nếu cờ thứ hai được xác định để chỉ báo rằng cửa sổ tương thích được sử dụng để lấy mẫu lại hình ảnh tham chiếu (CÓ ở khối 1030), quá trình 1000 có thể tiến hành đến khối 1040 và sau đó khối 1060. Ở khối 1040, quá trình 1000 có thể bao gồm việc xác định tỷ lệ lấy mẫu lại giữa hình ảnh hiện tại và hình ảnh tham chiếu dựa trên kích thước cửa sổ tương thích của cửa sổ tương thích.

Nếu cờ thứ hai được xác định là không chỉ báo rằng cửa sổ tương thích được sử dụng để lấy mẫu lại hình ảnh tham chiếu (KHÔNG ở khối 1030), quá trình 1000 có thể tiến hành đến khối 1050 và sau đó khối 1060. Ở khối 1050, quá trình 1000 có thể bao gồm việc xác định tỷ lệ lấy mẫu lại dựa trên kích thước hình ảnh lấy mẫu lại.

Như được thể hiện thêm trong FIG.10, quy trình 1000 có thể bao gồm việc thực hiện lấy mẫu lại hình ảnh tham chiếu trên hình ảnh hiện tại bằng cách sử dụng tỷ lệ lấy mẫu lại (khối 1060).

Theo phương án, cờ thứ nhất và cờ thứ hai có thể được báo hiệu trong tập tham số hình ảnh.

Theo phương án, kích thước cửa sổ tương thích có thể được xác định dựa trên ít nhất một khoảng cách độ dịch so với đường biên của hình ảnh hiện tại.

Theo phương án, ít nhất một khoảng cách độ dịch có thể được báo hiệu trong tập tham số hình ảnh.

Theo phương án, kích thước hình ảnh lấy mẫu lại có thể được chỉ báo trong dòng bit video được mã hóa bằng ít nhất một chiều rộng của kích thước hình ảnh lấy mẫu lại và chiều cao của kích thước hình ảnh lấy mẫu lại.

Theo phương án, ít nhất một trong các chiều rộng và chiều cao có thể được báo hiệu trong tập tham số hình ảnh.

Theo phương án, ít nhất một trong số chiều rộng và chiều cao có thể được biểu thị bằng số mẫu độ chói được bao gồm trong ít nhất một trong số chiều rộng và chiều cao.

Theo phương án, kích thước hình ảnh lấy mẫu lại có thể được xác định dựa trên ít nhất một khoảng cách độ dịch từ mẫu độ chói trên cùng bên trái của vùng tham chiếu của hình ảnh hiện tại.

Theo một phương án, ít nhất một khoảng cách độ dịch có thể được báo hiệu trong tập tham số hình ảnh.

Mặc dù FIG.10 thể hiện các khối mẫu của quy trình 1000, trong một số triển khai, quy trình 1000 có thể bao gồm các khối bổ sung, ít khối hơn, các khối khác nhau, hoặc các khối được sắp xếp khác với những khối được mô tả trong FIG.10. Ngoài ra, hoặc cách khác, hai hoặc nhiều khối của quy trình 1000 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 ví dụ, một hoặc nhiều bộ xử lý thực thi chương trình mà được lưu trữ trong phương tiện có thể 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 máy tính có thể đọc được và được lưu trữ vật lý trong một hoặc nhiều phương tiện đọc được bằng máy tính. Ví dụ, FIG.11 thể hiện hệ thống máy tính 1100 thích hợp để triển khai các phương án nhất định của đối tượng sáng chế.

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 có thể được thực thi trực tiếp,

hoặc thông qua diễn giải, thực thi vi mã, và tương tự, bởi đơn vị xử lý trung tâm máy tính (CPU), Đơn vị xử lý đồ họa (GPU), và 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 hiển thị trong FIG.11 dành cho hệ thống máy tính 1100 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. Cấu hình của các thành phần không được hiểu là có bất kỳ sự phụ thuộc hoặc yêu cầu nào liên quan đến bất kỳ một hoặc tổ hợp các thành phần nào được minh họa trong phương án được lấy làm ví dụ về hệ thống máy tính 1100.

Hệ thống máy tính 1100 có thể bao gồm các thiết bị đầu vào giao diện người nhất định. 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 găng tay dữ liệu), đầu vào âm thanh (chẳng hạn như: giọng nói, vỗ tay), đầu vào bằng hình ảnh (chẳng hạn như: cử chỉ), đầu vào khứu giác (không được mô tả). Các thiết bị giao diện con người cũng có thể được sử dụng để chụp một số phương tiện 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), hình ảnh (chẳng hạn như: hình ảnh được quét, hình ảnh chụp được 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 con người đầu vào có thể bao gồm một hoặc nhiều (chỉ một trong số mỗi thiết bị được mô tả): bàn phím 1101, chuột 1102, bàn di chuột 1103, màn hình cảm ứng 1110 và bộ điều hợp đồ họa đi kèm 1150, găng tay dữ liệu, cần điều khiển 1105, micrô 1106, máy quét 1107, camera 1108.

Hệ thống máy tính 1100 cũng có thể bao gồm các 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 1110, găng tay dữ liệu, hoặc cần điều khiển 1105, nhưng cũng có thể có các thiết bị phản hồi xúc giác không đóng vai trò thiết bị đầu vào), thiết bị đầu ra âm thanh (chẳng hạn như: loa 1109, tai nghe (không được mô tả)), thiết bị xuất hình ảnh (chẳng hạn như màn hình 1110 để bao gồm màn hình ống tia âm cực (CRT), màn hình tinh thể lỏng (LCD), màn hình plasma, màn hình điốt phát sáng hữu cơ (OLED), 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 — một số trong số đó có thể có khả năng xuất ra đầu ra hình ảnh hai chiều hoặc đầu ra nhiều hơn ba chiều thông qua các phương tiện như đầu ra ảnh lập thể; kính thực tế ảo (không được mô tả), màn hình nổi ba chiều và bể khối (không được mô tả), và máy in (không được mô tả).

Hệ thống máy tính 1100 cũng có thể bao gồm các thiết bị lưu trữ có thể truy cập được của con người và phương tiện liên kết của chúng như phương tiện quang bao gồm CD / DVD ROM / RW 1120 với CD / DVD hoặc phương tiện tương tự 1121, ổ cứng 1122, ổ cứng di động hoặc ổ đĩa trạng thái rắn 1123, 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 cũng nên hiểu rằng thuật ngữ "phương tiện có thể đọc được bằng máy tính" được sử dụng liên quan đến đối tượng sáng chế hiện tại không bao gồm phương tiện truyền dẫn, sóng mang, hoặc các báo hiệu tạm thời khác.

Hệ thống máy tính 1100 cũng có thể bao gồm (các) giao diện với một hoặc nhiều mạng truyền thông (1155). Ví dụ, mạng có thể 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 (1154) được gắn với một số cổng dữ liệu mục đích chung nhất định hoặc xe buýt ngoại vi (1149) (chẳng hạn như, ví dụ cổng kênh nối tiếp chung (USB) của hệ thống máy tính 1100; những mạng khác thường được tích hợp vào lõi của hệ thống máy tính 1100 bằng cách gắn vào kênh hệ thống như được 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 1155 có thể được kết nối với kênh ngoại vi 1149 bằng cách sử dụng giao diện mạng 1154. Sử dụng bất kỳ mạng nào trong số các mạng này, hệ thống máy tính 1100 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 các thiết bị CANBus nhất định), hoặc hai chiều, ví dụ cho các hệ thống máy tính khác sử dụng mạng số cục bộ hoặc mạng diện rộng. Một số giao thức và ngăn xếp giao thức nhất định có thể được sử dụng trên mỗi mạng và giao diện mạng (1154) như được mô tả ở trên.

Các thiết bị giao diện người nói trên, thiết bị lưu trữ mà con người có thể truy cập, và giao diện mạng có thể được gắn vào lõi 1140 của hệ thống máy tính 1100.

Lõi 1140 có thể bao gồm một hoặc nhiều Đơn vị xử lý trung tâm (CPU) 1141, Đơn vị xử lý đồ họa (GPU) 1142, các đơn vị xử lý lập trình chuyên biệt ở dạng Khu vực công lập trình trường (FPGA) 1143, bộ tăng tốc phần cứng 1144 cho các tác vụ nhất định, và kể từ đó trở đi. Các thiết bị này, cùng với Bộ nhớ chỉ đọc (ROM) 1145, Bộ nhớ truy cập ngẫu nhiên (RAM) 1146, bộ nhớ chung nội bộ như là ổ cứng truy cập không người dùng nội bộ, ổ cứng thể rắn (SSDs), và những thứ tương tự, có thể được kết nối thông qua kênh hệ thống 1148. Trong một số hệ thống máy tính, kênh hệ thống 1148 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 kênh hệ thống 1148 của lõi hoặc qua kênh ngoại vi 1149. Kiến trúc cho kênh ngoại vi bao gồm kết nối thành phần ngoại vi (PCI), USB, và những thứ tương tự.

CPU 1141, GPU 1142, FPGA 1143, và bộ tăng tốc 1144 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 1145 hoặc RAM 1146. Dữ liệu chuyển tiếp cũng có thể được lưu trữ trong RAM 1146, trong khi dữ liệu vĩnh viễn có thể được lưu trữ cho ví dụ, trong bộ nhớ chung nội bộ 1147. Lưu trữ nhanh và truy xuất vào bất kỳ thiết bị bộ nhớ nào có thể được kích hoạt thông qua việc sử dụng bộ nhớ đệm, mà có thể được liên kết chặt chẽ với một hoặc nhiều CPU 1141, GPU 1142, bộ nhớ chung 1147, ROM 1145, RAM 1146, và những thứ tương tự.

Phương tiện có thể đọc được trên máy tính 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 về kỹ thuật phần mềm máy tính.

Như ví dụ và không phải giới hạn, hệ thống máy tính có kiến trúc 1100, và cụ thể là lõi 1140 có thể cung cấp chức năng là kết quả của (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, máy tính có thể đọc được. Phương tiện đọc được bằng máy tính 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 1140 có tính chất không tạm thời, chẳng hạn như bộ lưu trữ nội bộ lõi 1147 hoặc ROM 1145. 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 1140. 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 1140 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 1146 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 nối cứng hoặc được thể hiện bằng cách khác trong mạch (ví dụ: máy gia tốc 1144), 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, và nếu thích hợp. Tham chiếu đến phương tiện đọc được bằng máy tính có thể bao gồm 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 đổi, hoán vị, và các phương án thay thế khác nhau, 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 sẽ có thể đưa ra nhiều hệ thống và phương pháp, mặc dù không được thể hiện hoặc mô tả rõ ràng ở đây, nhưng thể hiện các nguyên tắc của sáng chế và do đó nằm trong tinh thần và phạm vi của chúng.

Đơn này yêu cầu hưởng quyền ưu tiên từ Đơn tạm thời của Hoa Kỳ số 62/903,639, được nộp vào ngày 20 tháng 9 năm 2019, Đơn tạm thời của Hoa Kỳ số 62/905,319, được nộp vào ngày 24 tháng 9 năm 2019, và Đơn xin cấp bằng sáng chế Hoa Kỳ số 17/010,163, được nộp vào ngày 2 tháng 9 năm 2020, toàn bộ nội dung của các đơn này được đưa vào đây bằng cách viện dẫn.

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 cờ thứ nhất chỉ báo rằng cửa sổ tương thích được hiện diện trong hình ảnh hiện tại;
 dựa trên cờ thứ nhất chỉ báo rằng cửa sổ tương thích được hiện diện, thu nhận cờ thứ hai chỉ báo liệu cửa sổ tương thích có được sử dụng để lấy mẫu lại hình ảnh tham chiếu hay không;
 dựa trên cờ thứ hai chỉ báo rằng cửa sổ tương thích được sử dụng để lấy mẫu lại hình ảnh tham chiếu, xác định tỷ lệ lấy mẫu lại giữa hình ảnh hiện tại và hình ảnh tham chiếu dựa trên kích thước của cửa sổ tương thích của cửa sổ tương thích;
 dựa trên cờ thứ hai chỉ báo rằng cửa sổ tương thích không được sử dụng để lấy mẫu lại ảnh tham chiếu, xác định tỷ lệ lấy mẫu lại dựa trên kích thước ảnh lấy mẫu lại; và
 thực hiện lấy mẫu lại hình ảnh tham chiếu trên hình ảnh hiện tại bằng cách sử dụng tỷ lệ lấy mẫu lại.
 trong đó cờ thứ nhất bao gồm cờ cửa sổ tương thích ,
 cờ cửa sổ tương thích, trong đó khi cờ cửa sổ tương thích bằng 0, các giá trị của `conf_win_left_offset`, `conf_win_right_offset`, `conf_win_top_offset`, và `conf_win_bottom_offset` được suy ra bằng 0.
2. Phương pháp theo điểm 1, trong đó cờ thứ nhất và cờ thứ hai được báo hiệu trong tập tham số hình ảnh.
3. Phương pháp theo điểm 1, trong đó kích thước cửa sổ tương thích được xác định dựa trên ít nhất một khoảng cách độ dịch từ đường biên của hình ảnh hiện tại.
4. Phương pháp theo điểm 3, trong đó ít nhất một khoảng cách độ dịch được báo hiệu trong tập tham số hình ảnh.
5. Phương pháp theo điểm 1, trong đó kích thước hình ảnh lấy mẫu lại được chỉ báo trong dòng bit video được mã hóa là ít nhất một trong số chiều rộng của kích

thước hình ảnh lấy mẫu lại và chiều cao của kích thước hình ảnh lấy mẫu lại.

6. Phương pháp theo điểm 5, trong đó ít nhất một trong số các chiều rộng và chiều cao được báo hiệu trong tập tham số hình ảnh.

7. Phương pháp theo điểm 5, trong đó ít nhất một trong số các chiều rộng và chiều cao được biểu thị là số lượng mẫu độ chói có trong ít nhất một trong số chiều rộng và chiều cao.

8. Phương pháp theo điểm 1, trong đó kích thước ảnh lấy mẫu lại được xác định dựa trên ít nhất một khoảng cách độ dịch từ mẫu độ chói trên cùng - bên trái của vùng tham chiếu của ảnh hiện tại.

9. Phương pháp theo điểm 8, trong đó ít nhất một khoảng cách độ dịch được báo hiệu trong tập tham số hình ảnh.

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

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

ít nhất một bộ xử lý được tạo cấu hình để đọc mã chương trình và hoạt động theo lệnh của mã chương trình để thực hiện phương pháp giải mã dòng bit video được mã hóa theo điểm bất kỳ trong số các điểm từ 1 đến 9.

11. Phương tiện bất biến đọc được bằng máy tính lưu trữ các lệnh, các lệnh này bao gồm: một hoặc nhiều lệnh mà, khi được thực thi bởi một hoặc nhiều bộ xử lý của thiết bị giải mã dòng bit video được mã hóa, khiến một hoặc nhiều bộ xử lý thực hiện phương pháp giải mã dòng bit video được mã hóa theo điểm bất kỳ trong số các điểm từ 1 đến 9.

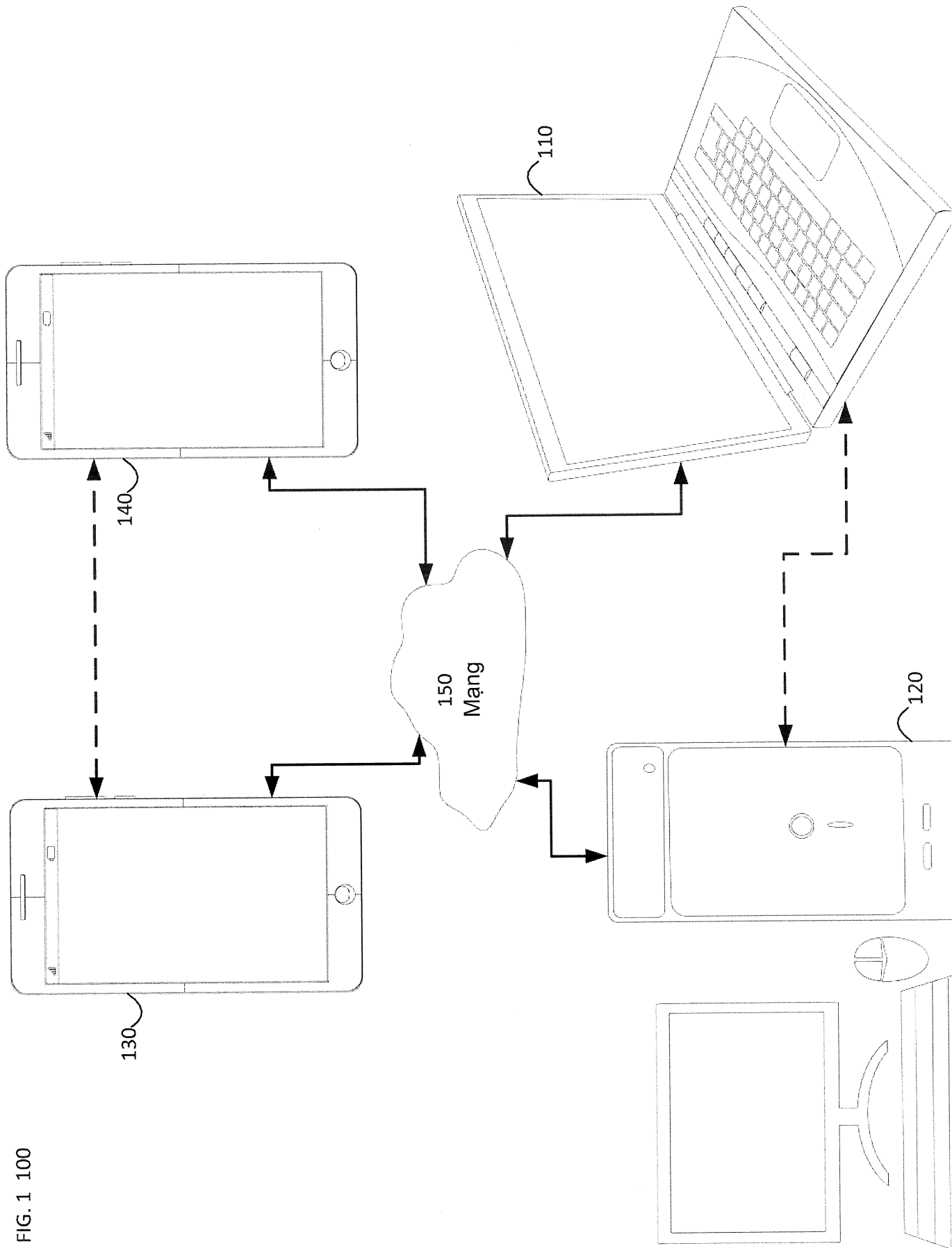
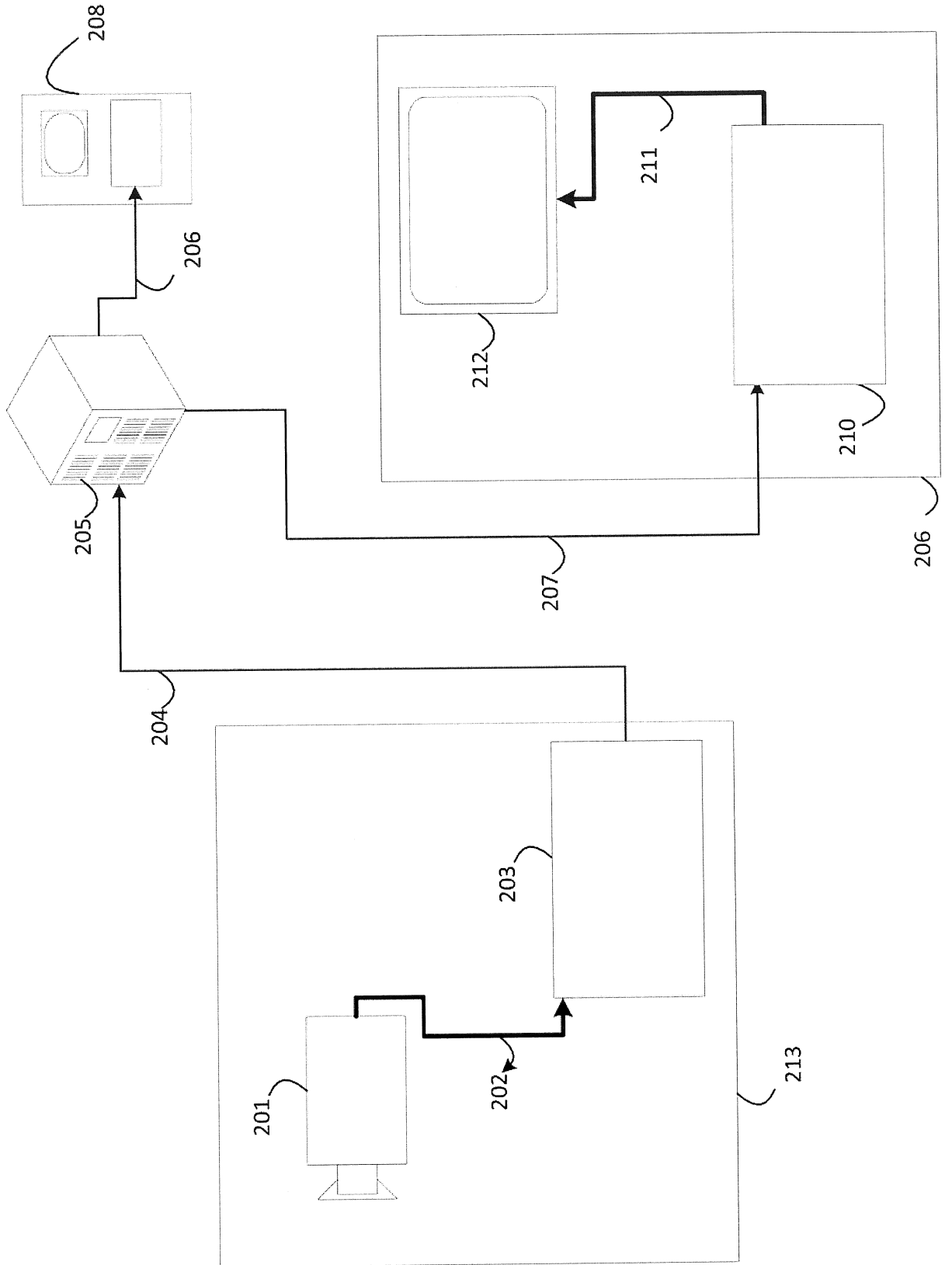


FIG. 1 100

FIG. 2



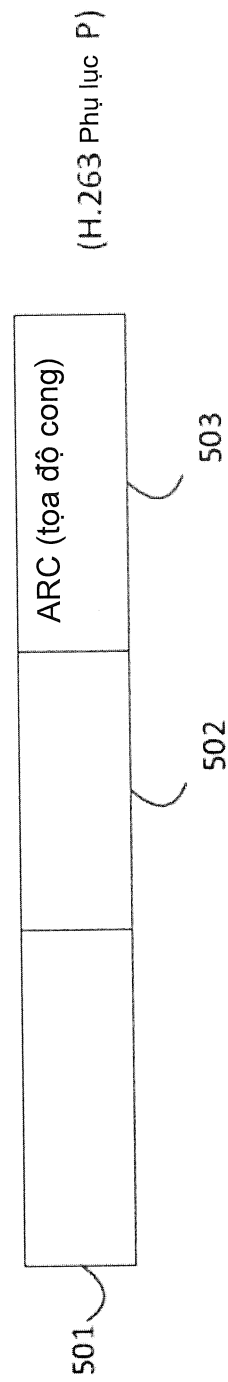


FIG. 5

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)
	reference_pic_height_in_luma_samples	ue(v)
	}	
	num_dec_pic_size_in_luma_samples_minus1	ue(v)
615	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

Fig. 7

	Mô tả
pic parameter set rbsp() {	
...	
pic width in luma samples	ue(v)
pic height in luma samples	ue(v)
conformance window flag	u(1)
if (conformance window flag) {	
conf win left offset	ue(v)
conf win right offset	ue(v)
conf win top offset	ue(v)
conf win bottom offset	ue(v)
}	
...	
}	

FIG. 8A

pic parameter set rbsp() {	Mô tả
...	
pic width in luma samples	ue(v)
pic height in luma samples	ue(v)
conformance window flag	u(l)
if(conformance window flag) {	
conf win left offset	ue(v)
conf win right offset	ue(v)
conf win top offset	ue(v)
conf win bottom offset	ue(v)
}	
use conf win for rpr flag	u(l)
if(luse conf win for rpr flag) {	
resampling pic width in luma samples	ue(v)
resampling pic height in luma samples	ue(v)
}	
....	
}	

FIG. 8B

pic_parameter_set_rbsp() {	Mô tả
....	
pic_width_in_luma_samples	ue(v)
pic_height_in_luma_samples	ue(v)
conformance_window_flag	u(1)
if(conformance_window_flag) {	
conf_win_left_offset	ue(v)
conf_win_right_offset	ue(v)
conf_win_top_offset	ue(v)
conf_win_bottom_offset	ue(v)
}	
use_conf_win_for_rpr_flag	u(1)
if(!use_conf_win_for_rpr_flag) {	
ref_region_left_offset	ue(v)
ref_region_right_offset	ue(v)
ref_region_top_offset	ue(v)
ref_region_bottom_offset	ue(v)
}	
....	
}	

