



(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)



1-0045031

(51)^{2020.01} H04N 19/33; H04N 19/177; H04N
19/70; H04N 19/597; H04N 19/172;
H04N 19/187

(13) B

(21) 1-2021-06501

(22) 18/09/2020

(86) PCT/US2020/051478 18/09/2020

(87) WO2021/055739 25/03/2021

(30) 62/903,603 20/09/2019 US; 17/019,959 14/09/2020 US

(45) 25/04/2025 445

(43) 27/06/2022 411A

(71) 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).

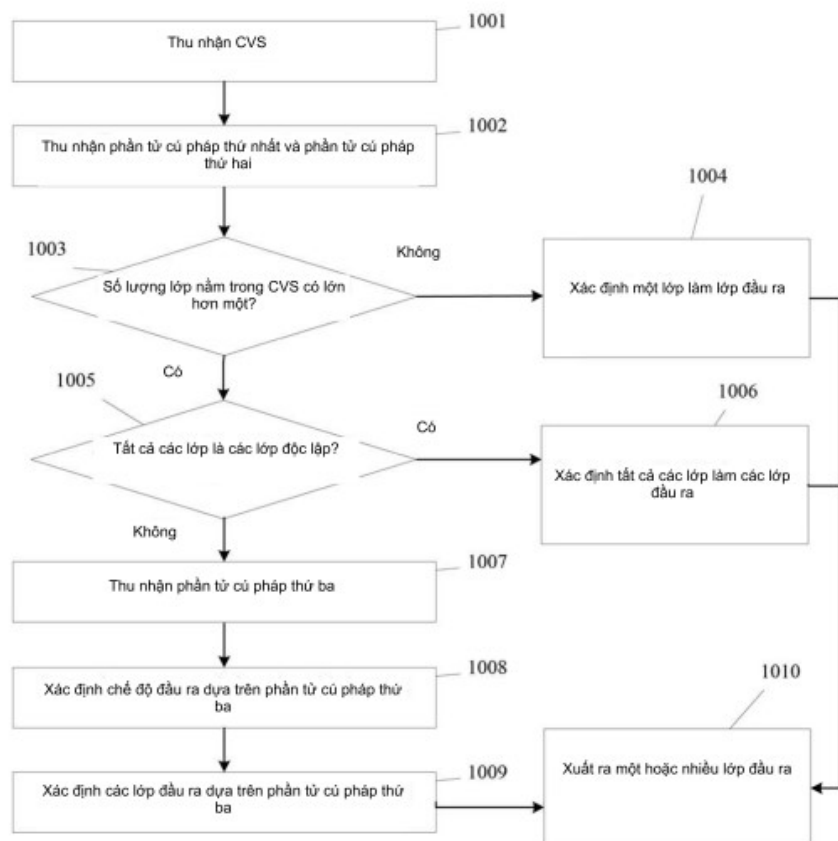
(74) Công ty TNHH một thành viên Sở hữu trí tuệ VCCI (VCCI-IP CO.,LTD)

(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) 1-2021-06501

(57) Sáng chế đề cập đến phương pháp và thiết bị giải mã dòng bit video được mã hóa và phương tiện đọc được bằng máy tính. Phương pháp giải mã dòng bit video được mã hóa bằng cách sử dụng ít nhất một bộ xử lý bao gồm thu nhận chuỗi video được mã hóa từ dòng bit video được mã hóa; thu nhận phần tử cú pháp thứ nhất chỉ báo số lượng các lớp nằm trong chuỗi video được mã hóa; thu nhận phần tử cú pháp thứ hai chỉ báo xem các lớp nằm trong chuỗi video được mã hóa có đều là các lớp độc lập không; dựa trên phần tử cú pháp thứ nhất chỉ báo rằng số lượng các lớp lớn hơn một, và phần tử cú pháp thứ hai chỉ báo rằng các lớp không phải tất cả là các lớp độc lập, thu nhận phần tử cú pháp thứ ba chỉ báo chế độ đầu ra; xác định chế độ đầu ra dựa trên phần tử cú pháp thứ ba; xác định một hoặc nhiều lớp đầu ra từ trong số các lớp dựa trên chế độ đầu ra được xác định; và xuất ra một hoặc nhiều lớp đầu ra.

FIG. 10 1000



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến việc mã hóa và giải mã video, và cụ thể hơn, đề cập đến phương pháp và thiết bị giải mã dòng bit video được mã hóa và phương pháp đọc được bằng máy tính.

Tình trạng kỹ thuật của sáng chế

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

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

dụng phân phối ti-vi. Tỷ lệ nén có thể đạt được có thể chỉ báo rằng: độ méo cho phép/chấp nhận cao hơn có thể đạt được các tỷ lệ nén cao hơn.

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

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

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

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

Bản chất kỹ thuật của sáng chế

Trong phương án, đề xuất phương pháp giải mã dòng bit video được mã hóa bằng cách sử dụng ít nhất một bộ xử lý, phương pháp bao gồm thu nhận chuỗi video được mã hóa từ dòng bit video được mã hóa; thu nhận phần tử cú pháp thứ nhất chỉ báo số lượng các lớp nằm trong chuỗi video được mã hóa; thu nhận phần tử cú pháp thứ hai chỉ báo xem các lớp nằm trong chuỗi video được mã hóa có đều là các lớp độc lập không; dựa trên phần tử cú pháp thứ nhất chỉ báo rằng số lượng các lớp lớn hơn một, và phần tử cú pháp thứ hai chỉ báo rằng các lớp không phải tất cả là các lớp độc lập, thu nhận phần tử cú pháp thứ ba chỉ báo chế độ đầu ra; xác định chế độ đầu ra dựa trên phần tử cú pháp thứ ba; xác định một hoặc nhiều lớp đầu ra từ trong số các lớp dựa trên chế độ đầu ra được xác định; và xuất ra một hoặc nhiều lớp đầu ra.

Trong phương án, đề xuất thiết bị giải mã dòng bit video được mã hóa, thiết bị bao gồm ít nhất một bộ nhớ được tạo cấu hình để lưu trữ mã chương trình; và ít nhất một bộ xử lý được tạo cấu hình để đọc mã chương trình và thao tác khi được lệnh bởi mã chương trình, mã chương trình bao gồm: mã thu nhận thứ nhất được tạo cấu hình để khiến ít nhất một bộ xử lý thu nhận chuỗi video được mã hóa từ dòng bit video được mã hóa; mã thu nhận thứ hai được tạo cấu hình để khiến ít nhất một bộ xử lý thu nhận phần tử cú pháp thứ nhất chỉ báo số lượng các lớp nằm trong chuỗi video được mã hóa; mã thu nhận thứ ba được tạo cấu hình để khiến ít nhất một bộ xử lý thu nhận phần tử cú pháp thứ hai chỉ báo xem các lớp nằm trong chuỗi video được mã hóa có đều là các lớp độc lập không; mã thu nhận thứ tư được tạo cấu hình để khiến ít nhất một bộ xử lý, dựa trên phần tử cú pháp thứ nhất chỉ báo rằng số lượng các lớp lớn hơn một, và phần tử cú pháp thứ hai chỉ báo rằng các lớp không phải là tất cả các lớp độc lập, thu nhận phần tử cú pháp thứ ba chỉ báo chế độ đầu ra; mã xác định thứ nhất

được tạo cấu hình để khiến ít nhất một bộ xử lý xác định chế độ đầu ra dựa trên phần tử cú pháp thứ ba; mã xác định thứ hai được tạo cấu hình để khiến ít nhất một bộ xử lý để xác định một hoặc nhiều lớp đầu ra từ trong số các lớp dựa trên chế độ đầu ra được xác định; và mã đầu ra được tạo cấu hình để khiến ít nhất một bộ xử lý xuất ra một hoặc nhiều lớp đầu ra.

Trong phương án, đề xuất phương tiện bất biến đọc được bằng máy tính lưu trữ các lệnh máy, các lệnh máy này bao gồm một hoặc nhiều lệnh máy mà, khi được thực hiện 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: thu nhận chuỗi video được mã hóa từ dòng bit video được mã hóa; thu nhận phần tử cú pháp thứ nhất chỉ báo số lượng các lớp nằm trong chuỗi video được mã hóa; thu nhận phần tử cú pháp thứ hai chỉ báo xem các lớp nằm trong chuỗi video được mã hóa có đều là các lớp độc lập không; dựa trên phần tử cú pháp thứ nhất chỉ báo rằng số lượng các lớp lớn hơn một, và phần tử cú pháp thứ hai chỉ báo rằng các lớp không phải tất cả là các lớp độc lập, thu nhận phần tử cú pháp thứ ba chỉ báo chế độ đầu ra; xác định chế độ đầu ra dựa trên phần tử cú pháp thứ ba; xác định một hoặc nhiều lớp đầu ra từ trong số các lớp dựa trên chế độ đầu ra được xác định; và xuất ra một hoặc nhiều lớp đầu ra.

Mô tả vắn tắt các hình vẽ

Các đặc điểm, bản chất, và các ưu điểm khác của đối tượng được bộc lộ sẽ được làm rõ hơn với phần mô tả chi tiết và các hình vẽ kèm theo dưới đây mà trong đó:

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

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

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

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

Fig. 5 là hình minh họa giản lược của các tùy chọn để báo hiệu các tham số ARC/RPR theo phương án.

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

Fig. 7 là hình minh họa giản lược của báo hiệu sự phụ thuộc lớp trong VPS theo phương án.

Fig. 8 là hình minh họa giản lược của báo hiệu danh sách các khuôn dạng đại diện trong VPS theo phương án.

Fig. 9 là hình minh họa giản lược của báo hiệu chỉ số của khuôn dạng đại diện trong SPS theo phương án.

Fig. 10 là lưu đồ của ví dụ về quá trình giải mã dòng bit video được mã hóa theo phương án.

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

Mô tả chi tiết sáng chế

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

Fig. 1 minh họa cặp thiết bị đầu cuối thứ hai 130, 140 được bố trí để hỗ trợ việc truyền hai chiều của video được mã hóa mà có thể diễn ra, ví dụ, trong hội nghị video. Đối với việc truyền dữ liệu hai chiều, mỗi thiết bị đầu cuối 130,

140 có thể mã hóa dữ liệu video được thu tại vị trí cục bộ cho việc truyền tới thiết bị đầu cuối khác thông qua mạng 150. Mỗi thiết bị đầu cuối 130, 140 cũng có thể thu dữ liệu video được mã hóa được truyền bởi thiết bị đầu cuối khác, có thể giải mã dữ liệu được mã hóa và có thể hiển thị dữ liệu video được khôi phục tại thiết bị hiển thị cục bộ.

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

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

Hệ thống tạo dòng có thể bao gồm hệ thống thu phụ 213, mà có thể bao gồm nguồn video 201, ví dụ camera số, tạo ra, ví dụ, dòng mẫu video không được nén 202. Dòng mẫu 202 này, được thể hiện như là đường đậm để nhấn mạnh khối lượng dữ liệu cao khi được so sánh với các dòng bit video được mã hóa, có thể được xử lý bởi bộ mã hóa 203 được ghép nối tới camera 201. Bộ mã hóa 203 có thể bao gồm phần cứng, phần mềm, hoặc kết hợp của chúng để cho

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

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

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

điều khiển đủ, hoặc từ mạng đồng bộ, bộ đệm 315 có thể không cần thiết, hoặc có thể là nhỏ. Để sử dụng trên các mạng gói tin nổi lực cao nhất như là mạng Internet, bộ đệm 315 có thể được yêu cầu, có thể là tương đối lớn và có thể có lợi về kích cỡ thích ứng.

Bộ giải mã video 210 có thể bao gồm bộ phân tích 320 để khôi phục các ký tự 321 từ chuỗi video được mã hóa entropy. Các mục của các ký tự này bao gồm thông được sử dụng để quản lý hoạt động của bộ giải mã 210, và thông tin có khả năng để điều khiển thiết bị kết xuất đồ họa như màn hình 212 mà không phải phần liên khối của bộ giải mã mà có thể được ghép nối tới đó, như được thể hiện trong Fig. 3. Thông tin điều khiển dùng cho các thiết bị kết xuất đồ họa có thể nằm trong dạng của thông tin nâng cao bổ sung (các bản tin SEI - Supplementary Enhancement Information) hoặc các phân đoạn tập hợp tham số thông có khả năng sử dụng video (VUI-Video Usability Information) (không được thể hiện). Bộ phân tích 320 có thể phân tích/giải mã entropy chuỗi video được mã hóa thu được. Việc mã hóa chuỗi video được mã hóa có thể tuân theo kỹ thuật hoặc tiêu chuẩn mã hóa video, và có thể tuân theo các nguyên tắc đã được biết đến rộng rãi đối với người có hiểu biết trung bình trong lĩnh vực kỹ thuật, bao gồm mã hóa độ dài biến thiên, mã hóa Huffman, mã hóa số học có hoặc không có cảm nhận ngữ cảnh, và v.v. Bộ phân tích 320 có thể tách từ chuỗi video được mã hóa, tập hợp của các tham số nhóm con đối với ít nhất một trong số các nhóm con của các điểm ảnh trong bộ giải mã video, dựa trên ít nhất một tham số tương ứng với nhóm. Các nhóm con có thể bao gồm các Nhóm của các ảnh (GOPs), các ảnh, các ảnh con, các ô, các lát, các viên, các khối macrô, các đơn vị mã hóa (CU-Coding Unit) các đơn vị cây mã hóa (CTU), các khối, các đơn vị biến đổi (TU-Transform Unit), các đơn vị dự đoán (PU-Prediction Unit) và v.v. Ô có thể chỉ báo vùng chữ nhật của CU/các CTU nằm trong cột và hàng ô cụ thể trong hình ảnh. Viên có thể chỉ báo vùng chữ nhật của các hàng CU/CTU nằm trong ô cụ thể. Lát có thể chỉ báo một hoặc nhiều viên của hình ảnh, mà được chứa trong đơn vị NAL. Ảnh con có thể chỉ báo vùng chữ nhật của một hoặc nhiều lát trong hình ảnh. Bộ giải mã entropy / bộ phân

tích có thể cũng trích xuất từ thông tin chuỗi video được mã hóa như là các hệ số biến đổi, các giá trị tham số lượng tử hóa, các vector chuyển động, và v.v..

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

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

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

Bộ phận thứ nhất là bộ biến đổi tỷ lệ/bộ biến đổi ngược 351. Bộ biến đổi tỷ lệ/bộ biến đổi ngược 351 thu hệ số biến đổi được lượng tử hóa cũng như thông tin điều khiển, bao gồm biến đổi nào được sử dụng, kích cỡ khối, hệ số lượng tử hóa, các ma trận biến đổi tỷ lệ lượng tử hóa, v.v như các ký tự 321 từ bộ phân tích 320. Bộ biến đổi tỷ lệ/bộ biến đổi ngược 351 có thể xuất ra các khối bao gồm các giá trị mẫu, mà có thể được đưa vào bộ kết hợp 355.

Trong một số trường hợp, các mẫu đầu ra của bộ định tỷ lệ / biến đổi ngược 351 có thể thuộc về khối được mã hóa nội bộ; tức là: khối mà không sử dụng thông tin mang tính dự đoán từ các hình ảnh được tái cấu trúc trước đó, nhưng có thể sử dụng thông tin mang tính dự đoán từ các phần được tái cấu trúc trước đó của hình ảnh hiện thời. Thông tin dự đoán này có thể được cấp bởi bộ dự đoán nội ảnh 352. Trong một vài trường hợp, bộ dự đoán nội ảnh 352 tạo ra

khối có cùng dạng và kích cỡ của khối đang được khôi phục, nhờ sử dụng thông tin đã được khôi phục xung quanh được truy xuất từ ảnh hiện tại (được khôi phục một phần) 358. Bộ tổng hợp 355, trong một số trường hợp, thêm, trên cơ sở mỗi mẫu, thông tin dự đoán của bộ phận dự đoán nội bộ 352 được tạo ra cho thông tin mẫu đầu ra như được đưa ra bởi bộ định tỷ lệ / bộ phận biến đổi ngược 351.

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

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

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

Các ảnh được mã hóa định trước, được khôi phục đầy đủ một lần, có thể được sử dụng như là các ảnh tham chiếu cho việc dự đoán tương lai. Một khi ảnh được mã hóa được khôi phục đầy đủ và ảnh được mã hóa đã được nhận dạng là ảnh tham chiếu (ví dụ, bởi bộ phân tích 320), ảnh tham chiếu hiện tại 358 có thể trở thành một phần của bộ đệm ảnh tham chiếu 357, và bộ nhớ ảnh hiện tại làm mới có thể được cấp phát lại trước khi bắt đầu khôi phục ảnh được mã hóa tiếp theo..

Bộ giải mã video 210 có thể thực hiện các hoạt động giải mã theo kỹ thuật nén video định trước mà được mô tả trong tiêu chuẩn, như khuyến nghị ITU-T H.265. Chuỗi video được mã hóa có thể tuân theo cú pháp được chỉ rõ bởi kỹ thuật nén video hoặc tiêu chuẩn đang được sử dụng, trong trường hợp mà tuân theo cú pháp của kỹ thuật hoặc tiêu chuẩn nén video, như được chỉ rõ trong tiêu chuẩn hoặc tài liệu kỹ thuật nén video và một cách cụ thể trong tài liệu mẫu. Điều cần tuân theo có thể là việc độ phức tạp của chuỗi video được mã hóa nằm trong giới hạn đã được xác định bởi cấp độ của kỹ thuật hoặc tiêu chuẩn nén video. Trong một số trường hợp, các cấp độ giới hạn kích cỡ ảnh tối đa, tốc độ khung tối đa, tỷ lệ mẫu tái cấu trúc tối đa (được đo theo, ví dụ các siêu mẫu trên giây), kích cỡ ảnh tham chiếu tối đa, và v.v.. Các giới hạn được thiết lập bởi các cấp độ có thể, trong một số trường hợp, còn được giới hạn thông qua các mô tả kỹ thuật về Bộ giải mã tham chiếu giả thuyết (HRD) và siêu dữ liệu dùng cho việc quản lý bộ đệm HRD được báo hiệu trong chuỗi video được mã hóa.

Trong phương án của sáng chế, bộ thu 310 có thể thu dữ liệu bổ sung (dư thừa) với video được mã hóa. Dữ liệu bổ sung có thể được bao gồm như là một phần của (các) chuỗi video được mã hóa. Dữ liệu bổ sung có thể được sử dụng bởi bộ giải mã video 210 để giải mã một cách thích hợp dữ liệu và/hoặc khôi phục chính xác hơn dữ liệu video gốc. Dữ liệu bổ sung có thể nằm trong dạng

của, ví dụ, các lớp nâng cao tỷ lệ tín hiệu trên tạp âm (SNR) hoặc theo thời gian, không gian, các lát dư thừa, các ảnh dư thừa, các mã sửa lỗi tiến, và loại tương tự.

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

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

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

Theo phương án của sáng chế, bộ mã hóa 203 có thể mã hóa và nén các ảnh của chuỗi video nguồn thành chuỗi video được mã hóa 443 trong thời gian thực hoặc theo bất kỳ điều kiện thời gian khác như được yêu cầu bởi ứng dụng. Cường ép tốc độ mã hóa thích hợp là một chức năng của bộ điều khiển 450. Bộ điều khiển sẽ điều khiển các bộ phận chức năng khác như được mô tả dưới đây và được ghép nối chức năng tới các bộ phận này. Việc ghép nối không được mô tả rõ ràng. Các tham số được thiết lập bởi bộ điều khiển có thể bao gồm các

tham số dành cho điều khiển tốc độ (bỏ qua ảnh, lượng tử hóa, giá trị lamđã của các kỹ thuật tối ưu hóa biến dạng tỷ lệ, ...), kích cỡ ảnh, nhóm bản thiết kế ảnh (GOP), phạm vi tìm kiếm vectơ chuyển động tối đa, và v.v.. Người có hiểu biết trung bình trong lĩnh vực kỹ thuật có thể dễ dàng nhận ra các chức năng khác của bộ điều khiển 450 do chúng có thể thuộc về bộ mã hóa video 203 được tối ưu hóa cho thiết kế hệ thống định trước.

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

Hoạt động bộ giải mã “cục bộ” 433 có thể giống với của bộ giải mã “tù xa” 210, mà đã được mô tả chi tiết nêu trên dựa trên Fig. 3. Viện dẫn vắn tắt đến Fig. 4. tuy nhiên, viện dẫn ngắn gọn tới Fig 4, khi các ký tự là khả dụng và việc mã hóa/giải mã các ký tự đối với chuỗi video được mã hóa bởi bộ mã hóa entropy 445 và bộ phân tích 320 có thể là không tồn hao, các phần giải mã

entrôpy của bộ giải mã 210, bao gồm kênh 312, bộ thu 310, bộ đệm 315, và bộ phân tích 320 có thể không được thực hiện đầy đủ trong bộ giải mã cục bộ 433.

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

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

Bộ giải mã video cục bộ 433 có thể giải mã dữ liệu video được mã hóa của các khung mà có thể được chỉ định là các khung tham chiếu, dựa trên các ký tự đã được tạo ra bởi bộ mã hóa nguồn 430. Các thao tác của cơ cấu mã hóa 432 có thể là các quá trình tổn hao. Khi dữ liệu video được mã hóa có thể được giải mã ở bộ giải mã video (không được thể hiện trên Fig. 4), chuỗi video được tái cấu trúc thường có thể là bản sao của chuỗi video nguồn với một số sai số. Bộ giải mã video cục bộ 433 sao chép các xử lý giải mã mà có thể được thực hiện bởi bộ giải mã video trên các khung tham chiếu và có thể làm cho các khung tham chiếu được khôi phục được lưu trữ trong bộ lưu trữ đệm ảnh tham chiếu 434. Theo cách thức này, bộ mã hóa 203 có thể lưu trữ các bản sao của các khung tham chiếu được khôi phục cục bộ mà có nội dung chung như các khung

tham chiếu được khôi phục mà sẽ thu được bởi bộ giải mã video phía xa (không có các lỗi truyền).

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

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

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

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

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

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

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

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

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

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

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

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

Gần đây, việc tổng hợp hoặc trích xuất miền được nén của các phân hình ảnh độc lập về mặt ngữ nghĩa thành hình ảnh video đơn nhất đã thu hút một vài sự quan tâm. Cụ thể là, trong ngữ cảnh của, ví dụ, việc mã hóa 360 hoặc các ứng dụng giám sát cụ thể, các ảnh nguồn độc lập về mặt ngữ nghĩa (ví dụ như sáu bề mặt khối phỏng lập phương của cảnh được chiếu lên khối phỏng lập phương 360, hoặc các đầu vào camera chuyên biệt trong trường hợp cài đặt giám sát nhiều camera) có thể yêu cầu các thiết lập độ phân giải thích ứng riêng biệt để xử lý theo từng hoạt động của mỗi cảnh khác nhau ở điểm thời gian đã cho. Nói cách khác, các bộ mã hóa, ở điểm thời gian đã cho, có thể chọn để sử

dụng các hệ số tái lấy mẫu khác nhau cho các hình ảnh khác nhau độc lập về mặt ngữ nghĩa mà tạo nên toàn bộ cảnh 360 hoặc cảnh giám sát. Khi được kết hợp vào hình ảnh đơn nhất, theo thứ tự, yêu cầu rằng việc tái lấy mẫu hình ảnh tham chiếu được thực hiện, và báo hiệu mã hóa độ phân giải thích ứng khả dụng, đối với các phần của ảnh được mã hóa.

Sau đây, số ít thuật ngữ sẽ được đưa ra mà được viện dẫn tới phần còn lại của phần mô tả này.

Ảnh con có thể gọi là, trong một vài trường hợp, sự sắp xếp hình chữ nhật của các mẫu, các khối, các khối macrô, các đơn vị mã hóa, hoặc các thực thể tương tự mà được nhóm lại về mặt ngữ nghĩa học, và có thể được mã hóa độc lập trong độ phân giải được thay đổi. Một hoặc nhiều ảnh con có thể 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 thành hình ảnh, và một hoặc nhiều ảnh con có thể được trích xuất từ hình ảnh. Trong các môi trường nhất định, một hoặc nhiều ảnh con được mã hóa có thể được ghép trong miền được nén mà không chuyển mã tới mức mẫu thành ảnh được mã hóa, và trong các trường hợp khác hoặc tương tự, một hoặc nhiều ảnh con được mã hóa có thể được tách từ ảnh được mã hóa trong miền được nén.

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

Trong các phương án mã hóa và giải mã có thể được thực hiện trên ảnh video đơn nhất và được mã hóa độc lập về mặt ngữ nghĩa. Trước khi mô tả ý nghĩa của việc mã hóa/giải mã của các ảnh con với các tham số RPR/ARC độc

lập và tính phức tạp bổ sung của việc này, các tùy chọn để báo hiệu các tham số RPR/ARC sẽ được mô tả.

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

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

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

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

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

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

- các tọa độ “cong” gần giống với các tọa độ này trong Phụ lục P H.263, có cùng mức chi tiết phù hợp như được mô tả ở trên. Phụ lục P H.263 xác định

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

-các tham số bộ lọc lấy mẫu lên hoặc lấy mẫu xuống. Trong các phương án, có thể chỉ là bộ lọc đơn nhất cho việc lấy mẫu lên và/hoặc lấy mẫu xuống. Tuy nhiên, trong các phương án, có thể mong muốn cho phép tính linh hoạt hơn trong thiết kế bộ lọc, và có yêu cầu báo hiệu của các tham số lọc. Các tham số này có thể được lựa chọn thông qua chỉ số trong danh sách của các phương án bộ lọc khả dụng, bộ lọc có thể được chỉ rõ đầy đủ (ví dụ thông qua danh sách của các hệ số lọc, nhờ sử dụng các kỹ thuật mã hóa entropy thích hợp), bộ lọc có thể được lựa chọn ẩn thông qua các tỷ lệ lấy mẫu lên/xuống mà được báo hiệu theo bất kỳ trong số các cơ chế được đề cập nêu trên, và v.v..

Sau đây, phần mô tả giả định việc mã hóa của tập hợp hữu hạn các hệ số lấy mẫu lên/xuống (hệ số giống nhau cần được sử dụng trong cả chiều X và Y), được chỉ báo thông qua từ mã. Từ mã này có thể được mã hóa độ dài biến thiên, ví dụ nhờ sử dụng mã Golomb số mũ dùng chung cho các phần tử cú pháp nhất định trong các tiêu chuẩn kỹ thuật mã hóa video như H.264 và H.265. Một việc ánh xạ thích hợp của các giá trị tới các hệ số lấy mẫu lên/xuống có thể, ví dụ, tuân theo theo Bảng 1:

BẢNG 1

Từ mã	Mã số mũ Golomb	Độ phân giải gốc/đích
0	1	1/1
1	010	1 / 1,5 (biến đổi tỷ lệ lên

		bởi 50%)
2	011	1,5 / 1 (biến đổi tỷ lệ xuống bởi 50%)
3	00100	1 / 2 (biến đổi tỷ lệ lên bởi 100%)
4	00101	2 / 1 (biến đổi tỷ lệ xuống bởi 100%)

Nhiều cách ánh xạ tương tự có thể được suy ra theo các yêu cầu của ứng dụng và các khả năng của các cơ chế biến đổi tỷ lệ lên và xuống khả dụng trong kỹ thuật hoặc tiêu chuẩn nén video. Bảng này có thể được mở rộng tới nhiều giá trị hơn. Các giá trị có thể cũng được thể hiện bởi các cơ chế mã hóa entropy ngoài các mã Ext-Golomb, ví dụ bằng cách sử dụng mã hóa nhị phân. Điều này có thể có các ưu điểm nhất định khi bản thân các hệ số lấy mẫu được quan tâm nằm ngoài cơ cấu xử lý video (bộ mã hóa và bộ giải mã đầu tiên), ví dụ bởi MANE. Lưu ý rằng, đối với các trường hợp trong đó việc thay đổi độ phân giải không được yêu cầu, mã Golomb số mũ có thể được lựa chọn là ngắn; trong Bảng nêu trên, chỉ là một bit. Có thể có ưu điểm về hiệu quả mã hóa qua việc sử dụng các mã nhị phân cho trường hợp chung nhất.

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

Phần sau đây mô tả cách thức hệ số lấy mẫu lên/xuống (thông tin ARC), được mã hóa như được mô tả nêu trên, có thể được chứa trong cú pháp kỹ thuật

hoặc tiêu chuẩn mã hóa video. Các xem xét tương tự có thể áp dụng tới một, hoặc một ít, từ mã mà điều khiển các bộ lọc lấy mẫu lên/xuống. Sau đây là phần mô tả khi lượng tương đối lớn dữ liệu được yêu cầu cho bộ lọc hoặc các cấu trúc dữ liệu khác.

Như được thể hiện trên Fig. 5, H.263 Phụ lục P bao gồm thông tin ARC 502 có dạng bốn tọa độ cong vênh trong tiêu đề hình ảnh 501, cụ thể trong phần mở rộng của tiêu đề H.263 PLUSPTYPE 503. Điều này có thể là sự lựa chọn thiết kế hợp lý khi a) có tiêu đề hình ảnh khả dụng, và b) các thay đổi thường xuyên của thông tin ARC được mong đợi. Tuy nhiên, thông tin tiêu đề khi bằng cách sử dụng báo hiệu kiểu H.263 có thể là cực cao, và các hệ số biến đổi tỷ lệ có thể là không liên quan trong số các biên ảnh như tiêu đề hình ảnh có thể là của tính chất nhất thời.

Trong cùng một phương án hoặc một phương án khác, báo hiệu của các tham số ARC có thể tuân theo ví dụ chi tiết như đã nêu trong các Fig. 6A-6B. Các Fig. 6A-6B mô tả các sơ đồ cú pháp trong kiểu đại diện bằng cách sử dụng ký hiệu mà cơ bản tuân theo lập trình kiểu C, như ví dụ được sử dụng trong các tiêu chuẩn mã hóa video ít nhất là từ năm 1993. Các đường in đậm chỉ báo các phần tử cú pháp thể hiện trong dòng bit, các đường nhạt thường chỉ báo dòng điều khiển hoặc thiết lập của các biến.

Như được thể hiện trên Fig. 6A, tiêu đề nhóm ô 601 như cấu trúc cú pháp mẫu của tiêu đề ứng dụng được cho phần (có thể là chữ nhật) của hình ảnh có thể chứa có điều kiện, độ dài biến, phần tử cú pháp được mã hóa Exp-Golomb `dec_pic_size_idx` 602 (được mô tả in đậm). Sự thể hiện của phần tử cú pháp này trong tiêu đề nhóm ô có thể được giám sát khi sử dụng độ phân giải thích ứng 603 ở đây, giá trị của cờ không được mô tả in đậm, mà có nghĩa rằng cờ được thể hiện trong dòng bit ở điểm mà xuất hiện trong sơ đồ cú pháp. Có hay không việc độ phân giải thích ứng khi sử dụng hình ảnh hoặc các phần này nó có thể được báo hiệu trong cấu trúc cú pháp bậc cao bất kỳ bên trong hoặc bên ngoài

dòng bit. Trong ví dụ được thể hiện, được báo hiệu trong tập tham số chuỗi như được đưa ra dưới đây.

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

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

Trong các ứng dụng nhất định, có thể được đề xuất rằng bộ mã hóa chỉ dẫn bộ giải mã để sử dụng kích cỡ ảnh tham chiếu định trước mà không giả định rằng kích cỡ là kích cỡ ảnh đầu ra. Trong ví dụ này, phần tử cú pháp `reference_pic_size_present_flag` 614 tạo công sự có mặt có điều kiện của các chiều ảnh tham chiếu 615 (một lần nữa, số này liên quan đến các độ rộng và độ cao).

Cuối cùng, được thể hiện là Bảng độ rộng và các độ cao của hình ảnh có thể giải mã. Bảng này có thể được đưa ra, ví dụ, bởi chỉ báo Bảng

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

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

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

Các phương án được đề cập ở đây có thể được sử dụng trong môi trường như vậy. Trong các trường hợp nhất định, trong cùng phương án hoặc phương án khác, giá trị trong thông tin tiêu đề đơn vị NAL, ví dụ trường ID thời gian, có thể được sử dụng để chỉ báo không chỉ thời gian mà còn lớp không gian. Việc này có thể có các ưu điểm nhất định đối với các thiết kế hệ thống định trước; ví dụ, các đơn vị chuyển tiếp được lựa chọn (SFU-Selected Forwarding Unit) hiện tại được tạo ra và được tối ưu hóa đối với lớp thời gian được lựa chọn để chuyển tiếp dựa trên giá trị ID thời gian thông tin tiêu đề đơn vị NAL, có thể được sử dụng mà không cần cải biến, đối với các môi trường có khả năng biến đổi tỷ lệ. Để cho phép điều này, có thể có yêu cầu đối với việc ánh xạ giữa kích cỡ ảnh

được mã hóa và lớp thời gian được chỉ báo bởi trường ID thời gian trong thông tin tiêu đề đơn vị NAL.

Bằng việc tái lấy mẫu ảnh tham chiếu (RPR-Reference Picture Resampling) hoặc thay đổi độ phân giải thích nghi (ARC-Adaptive Resolution Change), tải bổ sung để hỗ trợ khả năng biến đổi tỷ lệ có thể đạt được bằng cách cải biến phần nhỏ cú pháp bậc cao (HLS-high-level syntax). Trong các phương án, việc dự đoán liên lớp được sử dụng trong hệ thống có khả năng biến đổi tỷ lệ để cải thiện hiệu quả mã hóa của các lớp nâng cao. Ngoài các việc dự đoán bù chuyển động theo thời gian và không gian mà khả dụng trong mã hóa-giải mã đơn lớp, việc dự đoán liên lớp có thể sử dụng dữ liệu video được tái lấy mẫu của ảnh tham chiếu được khôi phục từ lớp tham chiếu để dự đoán lớp nâng cao hiện tại. Sau đó, xử lý tái lấy mẫu đối với việc dự đoán liên lớp có thể được thực hiện tại mức khối, bằng cách cải biến xử lý nội suy hiện tại để bù chuyển động. Trong các phương án, không cần thiết có thêm xử lý tái lấy mẫu để hỗ trợ khả năng biến đổi tỷ lệ. Các phương án liên quan đến các phần tử cú pháp mức cao để hỗ trợ khả năng biến đổi tỷ lệ không gian/chất lượng nhờ sử dụng RPR.

Trong cùng một phương án hoặc một phương án khác, viện dẫn tới Fig. 7, phần tử cú pháp `vps_max_layers_minus1` 703 plus 1 có thể chỉ định số lượng lớp được cho phép tối đa trong mỗi CVS liên quan đến VPS 701. `vps_all_independent_layers_flag` 704 bằng 1 có thể chỉ định rằng tất cả các lớp trong CVS được mã hóa độc lập không cần sử dụng dự đoán liên lớp. `vps_all_independent_layers_flag` 704 bằng 0 có thể chỉ định rằng một hoặc nhiều lớp trong CVS có thể sử dụng dự đoán liên lớp. Khi không được thể hiện, giá trị của `vps_all_independent_layers_flag` có thể được xem là bằng 1. Khi `vps_all_independent_layers_flag` bằng 1, giá trị của `vps_independent_layer_flag[ i ]` 706 có thể được coi là bằng 1. Khi `vps_all_independent_layers_flag` bằng 0, giá trị của `vps_independent_layer_flag[ 0 ]` có thể được xem là bằng 1.

Viện dẫn tới Fig. 7, `vps_independent_layer_flag[ i ]` 706 bằng 1 có thể chỉ định rằng lớp với chỉ số `i` không sử dụng dự đoán lớp liên đới. `vps_independent_layer_flag[ i ]` bằng 0 có thể chỉ định rằng lớp với chỉ số `i` có thể sử dụng dự đoán lớp liên đới và `vps_layer_dependency_flag[ i ]` được thể hiện trong VPS. `vps_direct_dependency_flag[ i ][ j ]` 707 bằng 0 có thể chỉ định rằng lớp với chỉ số `j` không phải lớp tham chiếu trực tiếp cho lớp với chỉ số `i`. `vps_direct_dependency_flag[ i ][ j ]` bằng 1 có thể chỉ định rằng lớp với chỉ số `j` là lớp tham chiếu trực tiếp cho lớp với chỉ số `i`. Khi `vps_direct_dependency_flag[ i ][ j ]` có thể không hiện `i` và `j` trong phạm vi 0 đến `vps_max_layers_minus1`, gồm cả hai giá trị đầu và cuối, có thể được coi là bằng 0. Biến `DirectDependentLayerIdx[ i ][ j ]`, mà chỉ rõ lớp phụ thuộc trực tiếp thứ `j` của lớp thứ `i`, có thể được thu nhận như sau:

```
for( i = 1; i < vps_max_layers_minus1; i-- )
    if( !vps_independent_layer_flag[ i ] )
        for( j = i, k = 0; j >= 0; j-- )
            if( vps_direct_dependency_flag[ i ][ j ] )
                DirectDependentLayerIdx[ i ][ k++ ] = j
```

Biến `GeneralLayerIdx[ i ]`, mà chỉ rõ chỉ số lớp của lớp với `nuh_layer_id` bằng `vps_layer_id[ i ]`, có thể được thu nhận như sau:

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

Trong cùng một phương án hoặc một phương án khác, viện dẫn tới Fig. 7, khi `vps_max_layers_minus1` lớn hơn không và giá trị của `vps_all_independent_layers_flag` bằng không, `vps_output_layers_mode` và `vps_output_layer_flags[ i ]` được báo hiệu. `vps_output_layers_mode` 708 bằng 0 có thể chỉ định rằng chỉ lớp cao nhất là đầu ra. `vps_output_layer_mode` bằng 1 có thể chỉ định rằng tất cả các lớp là đầu ra. `vps_output_layer_mode` bằng 2 có thể chỉ định rằng các lớp mà là đầu ra là các lớp với `vps_output_layer_flag[ i ]` 709 bằng 1. Giá trị của `vps_output_layers_mode` có thể nằm trong khoảng từ 0

đến 2. Giá trị 3 của `vps_output_layer_mode` được dành riêng cho việc sử dụng trong tương lai bởi ITU-T | ISO/IEC. Khi có thể không hiện, giá trị của `vps_output_layers_mode` có thể được coi là bằng 1. `vps_output_layer_flag[ i ]` bằng 1 có thể chỉ định rằng lớp thứ `i` là đầu ra. `vps_output_layer_flag[ i ]` bằng 0 có thể chỉ định rằng lớp thứ `i` không là đầu ra. Danh sách `OutputLayerFlag[ i ]`, mà giá trị 1 có thể chỉ định rằng lớp thứ `i` là đầu ra và giá trị 0 được chỉ định rằng lớp thứ `i` không là đầu ra, có thể được thu nhận như sau:

```

OutputLayerFlag[ vps_max_layers_minus1 ] = 1
for( i = 0; i < vps_max_layers_minus1; i++ )
    if( vps_output_layer_mode == 0 )
        OutputLayerFlag[ i ] = 0
    else if( vps_output_layer_mode == 1 )
        OutputLayerFlag[ i ] = 1
    else if( vps_output_layer_mode == 2 )
        OutputLayerFlag[ i ] = vps_output_layer_flag[ i ]

```

Trong cùng một phương án hoặc một phương án khác, viện dẫn tới Fig. 8, `vps_num_rep_formats_minus1` plus 1 có thể chỉ định số lượng các cấu trúc cú pháp `rep_format( )` trong VPS. Giá trị của `vps_num_rep_formats_minus1` có thể nằm trong khoảng 0 đến 255, bao gồm cả hai giá trị đầu và cuối. `vps_rep_format_idx[ i ]` có thể chỉ định chỉ số, vào danh sách của các cấu trúc `rep_format( )` syntax trong VPS, của cấu trúc cú pháp `rep_format( )` mà áp dụng cho lớp có `GeneralLayerIdx[ nuh_layer_id ]` bằng `i`. giá trị của `vps_rep_format_idx[ i ]` có thể nằm trong khoảng 0 đến `vps_num_rep_formats_minus1`, bao gồm cả hai giá trị đầu và cuối. Số lượng bit được sử dụng cho việc đại diện của `vps_rep_format_idx[ i ]` là $\text{Ceil}(\text{Log}_2(\text{vps_num_rep_formats_minus1} + 1))$. `pic_width_max_vps_in_luma_samples`, `pic_height_max_vps_in_luma_samples`, `chroma_format_vps_idc`, `separate_colour_plane_vps_flag`, `bit_depth_vps_luma_minus8` và `bit_depth_vps_chroma_minus8` được sử dụng để suy luận các giá trị của các phân tử cú pháp SPS `pic_width_max_in_luma_samples`,

`pic_height_max_in_luma_samples`, `chroma_format_idc`,
`separate_colour_plane_flag`, `bit_depth_luma_minus8` và
`bit_depth_chroma_minus8`, theo cách tương ứng, cho mỗi SPS mà đề cập đến
VPS. `pic_width_max_vps_in_luma_samples` có thể không bằng 0 và có thể là
bội số nguyên của `max(8, MinCbSizeY)`.
`pic_height_max_vps_in_luma_samples` có thể không bằng 0 và có thể là bội số
nguyên của `max(8, MinCbSizeY)`. Giá trị của `chroma_format_vps_idc` có thể
nằm trong khoảng 0 đến 3, bao gồm cả hai giá trị đầu và cuối.
`bit_depth_vps_luma_minus8` có thể nằm trong khoảng 0 đến 8, bao gồm cả hai
giá trị đầu và cuối. `bit_depth_vps_chroma_minus8` có thể nằm trong khoảng 0
đến 8, bao gồm cả hai giá trị đầu và cuối.

`subpics_vps_present_flag` bằng 1 chỉ báo rằng các tham số ảnh con được
thể hiện trong việc thể hiện trong cú pháp VPS RBSP. `subpics_present_flag`
bằng 0 chỉ báo rằng các tham số ảnh con không được thể hiện trong việc thể
hiện trong cú pháp VPS RBSP. `max_subpics_vps_minus1`,
`subpic_grid_vps_col_width_minus1` `subpic_grid_vps_row_height_minus1`,
`subpic_grid_vps_row_height_minus1`, `subpic_grid_vps_idx[ i ][ j ]`,
`subpic_treated_as_pic_vps_flag[ i ]` và
`loop_filter_across_subpic_enabled_vps_flag[ i ]` được sử dụng để suy luận các
giá trị của các phần tử cú pháp SPS `max_subpics_minus1`,
`subpic_grid_col_width_minus1` `subpic_grid_row_height_minus1`,
`subpic_grid_row_height_minus1`, `subpic_grid_idx[ i ][ j ]`,
`subpic_treated_as_pic_flag[ i ]` và `loop_filter_across_subpic_enabled_flag[ i ]`
theo cách tương ứng, cho mỗi SPS mà đề cập đến VPS.

Trong cùng một phương án hoặc một phương án khác, viện dẫn tới Fig. 9,
`use_rep_format_flag` bằng 1 có thể chỉ định rằng `sps_rep_format_idx` được thể
hiện. `use_rep_format_flag` bằng 0 có thể chỉ định rằng `sps_rep_format_idx`
không được thể hiện, và `pic_width_max_in_luma_samples`,
`pic_height_max_in_luma_samples`, `chroma_format_idc`,
`separate_colour_plane_flag`, `bit_depth_luma_minus8`,

`bit_depth_chroma_minus8,` `max_subpics_minus1,`
`subpic_grid_col_width_minus1` `subpic_grid_row_height_minus1,`
`subpic_grid_row_height_minus1,` `subpic_grid_idx[ i ][ j ],`
`subpic_treated_as_pic_flag[ i ]` và `loop_filter_across_subpic_enabled_flag[ i ]`
 được thể hiện trong SPS. `sps_rep_format_idx` có thể chỉ định rằng các cấu trúc
 cú pháp `sps_rep_format_idx`-th `rep_format( )` trong VPS được tham chiếu áp
 dụng cho các lớp mà viện dẫn đến SPS này. Khi thể hiện,
`pic_width_max_in_luma_samples,` `pic_height_max_in_luma_samples,`
`chroma_format_idc,` `separate_colour_plane_flag,` `bit_depth_luma_minus8,`
`bit_depth_chroma_minus8,` `max_subpics_minus1,`
`subpic_grid_col_width_minus1` `subpic_grid_row_height_minus1,`
`subpic_grid_row_height_minus1,` `subpic_grid_idx[ i ][ j ],`
`subpic_treated_as_pic_flag[ i ]` và `loop_filter_across_subpic_enabled_flag[ i ]`
 được suy ra bằng các giá trị của `pic_width_max_vps_in_luma_samples,`
`pic_height_max_vps_in_luma_samples,` `chroma_format_vps_idc,`
`separate_colour_plane_vps_flag,` `bit_depth_vps_luma_minus8,`
`bit_depth_vps_chroma_minus8,` `max_subpics_vps_minus1,`
`subpic_grid_vps_col_width_minus1` `subpic_grid_vps_row_height_minus1,`
`subpic_grid_vps_row_height_minus1,` `subpic_grid_vps_idx[ i ][ j ],`
`subpic_treated_as_pic_vps_flag[ i ]` và
`loop_filter_across_subpic_enabled_vps_flag[ i ]` trong các cấu trúc cú pháp
`sps_rep_format_idx`-th `rep_format( )` trong VPS được tham chiếu.
`subpics_present_flag` bằng 1 chỉ báo rằng các tham số ảnh con được thể hiện
 trong việc thể hiện trong cú pháp SPS RBSP. `subpics_present_flag` bằng 0 chỉ
 báo rằng các tham số ảnh con không được thể hiện trong việc thể hiện trong cú
 pháp PS RBSP. `max_subpics_minus1` plus 1 có thể chỉ định số lượng ảnh con
 lớn nhất mà có thể thể hiện trong CVS. `max_subpics_minus1` có thể nằm trong
 khoảng 0 đến 254. `subpic_grid_col_width_minus1` plus 1 có thể chỉ định độ
 rộng của mỗi phần tử của lưới nhận dạng ảnh con trong các đơn vị của 4 mẫu.

Độ dài của phần tử cú pháp là các bit $\text{Ceil}(\text{Log2}(\text{pic_width_max_in_luma_samples} / 4))$.

Biến NumSubPicGridCols có thể được lấy như sau:

$$\text{NumSubPicGridCols} = \frac{(\text{pic_width_max_in_luma_samples} + \text{subpic_grid_col_width_minus1} * 4 + 3)}{(\text{subpic_grid_col_width_minus1} * 4 + 4)}$$

subpic_grid_row_height_minus1 plus 1 có thể chỉ định độ cao của mỗi phần tử của lưới ký hiệu nhận dạng ảnh con trong các đơn vị của 4 mẫu. Độ dài của phần tử cú pháp là các bit $\text{Ceil}(\text{Log2}(\text{pic_height_max_in_luma_samples} / 4))$.

Biến NumSubPicGridRows có thể được lấy như sau:

$$\text{NumSubPicGridRows} = \frac{(\text{pic_height_max_in_luma_samples} + \text{subpic_grid_row_height_minus1} * 4 + 3)}{(\text{subpic_grid_row_height_minus1} * 4 + 4)}$$

subpic_grid_idx[i][j] có thể chỉ định chỉ số ảnh con của vị trí lưới (i, j). Độ dài của phần tử cú pháp là các bit $\text{Ceil}(\text{Log2}(\text{max_subpics_minus1} + 1))$.

Các biến SubPicTop[subpic_grid_idx[i][j]], SubPicLeft[subpic_grid_idx[i][j]], SubPicWidth[subpic_grid_idx[i][j]], SubPicHeight[subpic_grid_idx[i][j]], và NumSubPics có thể được lấy như sau:

```

NumSubPics = 0
for( i = 0; i < NumSubPicGridRows; i++ ) {
    for( j = 0; j < NumSubPicGridCols; j++ ) {
        if ( i == 0 )
            SubPicTop[ subpic_grid_idx[ i ][ j ] ] = 0
        else if( subpic_grid_idx[ i ][ j ] != subpic_grid_idx[ i - 1 ][ j ] ) {
            SubPicTop[ subpic_grid_idx[ i ][ j ] ] = i
            SubPicHeight[ subpic_grid_idx[ i - 1 ][ j ] ] =

```

```

i - SubPicTop[ subpic_grid_idx[ i - 1 ][ j ] ]
    }
    if ( j == 0 )
        SubPicLeft[ subpic_grid_idx[ i ][ j ] ] = 0
    else if ( subpic_grid_idx[ i ][ j ] != subpic_grid_idx[ i ][ j - 1 ] ) {
        SubPicLeft[ subpic_grid_idx[ i ][ j ] ] = j
        SubPicWidth[ subpic_grid_idx[ i ][ j ] ] =
j - SubPicLeft[ subpic_grid_idx[ i ][ j - 1 ] ]
    }
    if ( i == NumSubPicGridRows - 1 )
        SubPicHeight[ subpic_grid_idx[ i ][ j ] ] =
i - SubPicTop[ subpic_grid_idx[ i - 1 ][ j ] ] + 1
    if ( j == NumSubPicGridRows - 1 )
        SubPicWidth[ subpic_grid_idx[ i ][ j ] ] =
j - SubPicLeft[ subpic_grid_idx[ i ][ j - 1 ] ] + 1
    if( subpic_grid_idx[ i ][ j ] > NumSubPics )
        NumSubPics = subpic_grid_idx[ i ][ j ]
    }
}

```

subpic_treated_as_pic_flag[i] bằng 1 có thể chỉ định rằng ảnh con thứ i của mỗi hình ảnh được mã hóa trong CVS được coi là hình ảnh trong quá trình giải mã ngoài các thao tác lọc trong vòng. subpic_treated_as_pic_flag[i] bằng 0 có thể chỉ định rằng ảnh con thứ i của mỗi hình ảnh được mã hóa trong CVS không được coi là hình ảnh trong quá trình giải mã ngoài các thao tác lọc trong vòng. Khi có thể không hiện, giá trị của subpic_treated_as_pic_flag[i] có thể được coi là bằng 0.

loop_filter_across_subpic_enabled_flag[i] bằng 1 có thể chỉ định rằng các thao tác lọc trong vòng có thể được thực hiện qua các biên của ảnh con thứ i trong mỗi hình ảnh được mã hóa trong CVS.

`loop_filter_across_subpic_enabled_flag[ i ]` bằng 0 có thể chỉ định rằng các thao tác lọc trong vòng không được thực hiện qua các biên của ảnh con thứ *i* trong mỗi hình ảnh được mã hóa trong CVS. Khi có thể không hiện, giá trị của `loop_filter_across_subpic_enabled_pic_flag[ i ]` có thể được coi là bằng 1.

Fig. 10 là lưu đồ của ví dụ về quá trình 1000 để giải mã dòng bit video được mã hóa. Trong một số cách thực hiện, một hoặc nhiều khối xử lý của Fig. 10 có thể được thực hiện bởi bộ giải mã 210. Trong một số cách thực hiện, một hoặc nhiều khối xử lý của Fig. 10 có thể được thực hiện bởi một thiết bị khác hoặc nhóm thiết bị riêng biệt với hoặc bao gồm bộ giải mã 210, như là bộ mã hóa 203.

Như được thể hiện trên Fig. 10, quá trình 1000 có thể bao gồm thu nhận chuỗi video được mã hóa từ dòng bit video được mã hóa (khối 1001).

Như được thể hiện thêm trên Fig. 10, quá trình 1000 có thể bao gồm thu nhận phần tử cú pháp thứ nhất chỉ báo số lượng các lớp nằm trong chuỗi video được mã hóa và phần tử cú pháp thứ hai chỉ báo xem các lớp nằm trong chuỗi video được mã hóa có đều là các lớp độc lập (khối 1002) hay không.

Như được thể hiện thêm trên Fig. 10, quá trình 1000 có thể bao gồm xác định, dựa trên phần tử cú pháp thứ nhất, xem số lượng các lớp có nằm trong CVS lớn hơn một (khối 1003) không.

Như được thể hiện thêm trên Fig. 10, quá trình 1000 có thể bao gồm, dựa trên phần tử cú pháp thứ nhất chỉ báo rằng số lượng các lớp nằm trong CVS không lớn hơn một (KHÔNG ở khối 1003), quá trình 1000 có thể tiến hành đến khối 1004, mà ở đó một lớp được xác định là lớp đầu ra, và sau đó đến khối 1010, mà ở đó lớp đầu ra là đầu ra.

Như được thể hiện thêm trên Fig. 10, dựa trên phần tử cú pháp thứ nhất chỉ báo rằng số lượng các lớp nằm trong CVS lớn hơn một (CÓ ở khối 1003), quá trình 1000 có thể tiến hành đến khối 1005.

Như được thể hiện thêm trên Fig. 10, quá trình 1000 có thể bao gồm xác định, dựa trên phần tử cú pháp thứ hai, xem các lớp nằm trong chuỗi video được mã hóa có đều là các lớp độc lập (khối 1005) hay không.

Như được thể hiện thêm trên Fig. 10, quá trình 1000 có thể bao gồm, dựa trên phần tử cú pháp thứ hai chỉ báo rằng các lớp nằm trong chuỗi video được mã hóa đều là các lớp độc lập (CÓ ở khối 1005), quá trình 1000 có thể tiến hành đến khối 1006, mà ở đó tất cả các lớp được xác định là các lớp đầu ra, và sau đó đến khối 1010, mà ở đó các lớp đầu ra là đầu ra.

Như được thể hiện thêm trên Fig. 10, dựa trên phần tử cú pháp thứ hai chỉ báo rằng các lớp nằm trong chuỗi video được mã hóa không phải tất cả là các lớp độc lập (KHÔNG ở khối 1003), quá trình 1000 có thể tiến hành đến khối 1007.

Như được thể hiện thêm trên Fig. 10, quá trình 1000 có thể bao gồm thu nhận phần tử cú pháp thứ ba chỉ báo chế độ đầu ra (khối 1007).

Như được thể hiện thêm trên Fig. 10, quá trình 1000 có thể bao gồm xác định chế độ đầu ra dựa trên phần tử cú pháp thứ ba (khối 1008).

Như được thể hiện thêm trên Fig. 10, quá trình 1000 có thể bao gồm xác định một hoặc nhiều lớp đầu ra từ trong số các lớp dựa trên chế độ đầu ra được xác định (khối 1009).

Như được thể hiện thêm trên Fig. 10, quá trình 1000 có thể bao gồm xuất ra một hoặc nhiều lớp đầu ra (khối 1010).

Trong phương án, dựa trên phần tử cú pháp thứ ba chỉ báo rằng chế độ đầu ra là chế độ thứ nhất, một hoặc nhiều lớp đầu ra có thể chỉ bao gồm lớp cao nhất từ trong số các lớp.

Trong phương án, dựa trên phần tử cú pháp thứ ba chỉ báo rằng chế độ đầu ra là chế độ thứ hai, một hoặc nhiều lớp đầu ra có thể bao gồm các lớp.

Trong phương án, quá trình 1000 có thể còn bao gồm thu nhận phần tử cú pháp thứ tư chỉ báo rằng lớp tương ứng của các lớp cần được xuất ra, trong đó

dựa trên phần tử cú pháp thứ ba chỉ báo rằng chế độ đầu ra là chế độ thứ ba, một hoặc nhiều lớp đầu ra có thể bao gồm lớp tương ứng.

Trong phương án, phần tử cú pháp thứ nhất, phần tử cú pháp thứ hai, và phần tử cú pháp thứ ba có thể được báo hiệu trong tập tham số video (VPS).

Trong phương án, VPS có thể còn bao gồm phần tử cú pháp thứ năm chỉ báo số lượng các khuôn dạng hình ảnh được báo hiệu trong VPS, và phần tử cú pháp thứ sáu chỉ báo chỉ số của khuôn dạng hình ảnh từ trong số các khuôn dạng hình ảnh cần được áp dụng cho lớp từ trong số các lớp.

Trong phương án, VPS có thể còn bao gồm phần tử cú pháp thứ bảy chỉ báo xem các tham số ảnh con có được báo hiệu trong VPS hay không.

Trong phương án, tập tham số chuỗi (SPS) viện dẫn tới VPS có thể bao gồm phần tử cú pháp thứ tám chỉ báo rằng SPS bao gồm phần tử cú pháp thứ chín chỉ báo chỉ số của khuôn dạng hình ảnh từ trong số các khuôn dạng hình ảnh được báo hiệu trong VPS cần được áp dụng cho lớp từ trong số các lớp.

Trong phương án, SPS có thể còn bao gồm phần tử cú pháp thứ sáu chỉ báo xem các tham số ảnh con có được báo hiệu trong SPS hay không.

Mặc dù Fig. 10 thể hiện ví dụ về các khối của quá trình 1000, trong một số cách thực hiện, quá trình 1000 có thể bao gồm các khối bổ sung, các khối thấp hơn, các khối khác nhau, hoặc các khối được bố trí khác nhau so với các khối được mô tả trên Fig. 10. Thêm vào đó, hoặc ngoài ra, hai hoặc nhiều khối của quá trình 1000 có thể được thực hiện song song.

Ngoài ra, các phương pháp được đề xuất có thể được thực hiện bởi hệ 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 mà được lưu trữ trong phương tiện bất biến đọc được bằng máy tính để thực hiện một hoặc nhiều phương pháp được đề xuất.

Các kỹ thuật được mô tả nêu trên, có thể được thực hiện như là phần mềm máy tính nhờ sử dụng các lệnh có thể đọc được bởi máy tính và được lưu trữ vật

lý trong một hoặc nhiều phương tiện đọc được bởi máy tính. Ví dụ, Fig. 11 thể hiện hệ thống máy tính 1100 thích hợp để thực hiện các phương án của đối tượng được bộc lộ.

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

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

Các thành phần được thể hiện trên Fig. 11 đối với hệ thống máy tính 1100 là ví dụ về mặt bản chất và không dự định đề xuất bất kỳ giới hạn nào đối với phạm vi sử dụng hoặc chức năng của phần mềm máy tính mà thực hiện các phương án của sáng chế. Cấu trúc của các bộ phận không được hiểu là có bất kỳ sự phụ thuộc hoặc yêu cầu liên quan đến bất kỳ một hoặc kết hợp của các bộ phận được minh họa trong phương án ví dụ của hệ thống máy tính 1100.

Hệ thống máy tính 1100 có thể bao gồm các thiết bị đầu vào giao diện người dùng định trước. Thiết bị đầu vào giao diện người dùng này có thể phản hồi lại việc nhập vào bởi một hoặc nhiều người dùng thông qua, ví dụ, việc nhập xúc giác (như: nhấn phím, trượt, di chuyển bao dữ liệu), nhập audio (như: thoại, vỗ tay), nhập trực quan (như: các cử chỉ), nhập khứu giác (không được thể hiện). Các thiết bị giao diện người dùng có thể cũng được sử dụng để quay nội dung đa phương tiện không cần thiết liên quan trực tiếp đến việc nhập nhận thức bởi người dùng, như audio (như: âm nhạc xung quanh, âm nhạc thoại), các ảnh (như: các ảnh được quét, các ảnh chụp thu nhận từ camera ảnh tĩnh), video (như video hai chiều, video ba chiều bao gồm video lập thể).

Các thiết bị giao diện người dùng đầu vào có thể bao gồm một hoặc nhiều trong số (chỉ một trong số các thiết bị được thể hiện): bàn phím 1101, chuột 1102, bàn di chuột 1103, màn chạm 1110, và bộ điều hợp đồ họa được kết hợp 1150, găng tay dữ liệu, cần điều khiển 1105, microphôn 1106, máy quét 1107, camera 1108.

Hệ thống máy tính 1100 có thể cũng bao gồm các thiết bị đầu ra giao diện người dùng định trước. Các thiết bị đầu ra giao diện người dùng này có thể là sự mô phỏng các giác quan của một hoặc nhiều người dùng thông qua, ví dụ, xúc giác, âm thanh, ánh sáng, và mùi/vị giác. Các thiết bị đầu ra giao diện người dùng này có thể bao gồm các thiết bị đầu ra xúc giác (ví dụ xúc giác phản hồi bởi màn chạm 1110, găng tay dữ liệu, hoặc cần điều khiển 1105, nhưng có thể cũng có các thiết bị phản hồi xúc giác mà không đóng vai trò là các thiết bị đầu vào), các thiết bị xuất audio (như: các loa 1109, các tai nghe (không được thể hiện)), các thiết bị đầu ra trực quan (như các màn hình 1110 bao gồm các màn hình CRT ống tia catot, màn hình LCD hiển thị tinh thể lỏng, màn hình plasma, màn hình điôt phát quang hữu cơ (OLED), mỗi trong số chúng có hoặc không có khả năng nhập màn chạm, có hoặc không có khả năng phản hồi xúc giác—một vài trong số chúng có thể có khả năng xuất ra đầu ra trực quan hai chiều hoặc đầu ra nhiều hơn ba chiều thông qua phương tiện như đầu ra lập thể; các kính thực tế ảo (không được thể hiện), các màn hình toàn ký và bề khối (không được thể hiện)), và các máy in (không được thể hiện).

Hệ thống máy tính 1100 có thể cũng bao gồm các thiết bị lưu trữ có thể truy nhập được bởi người dùng và phương tiện được kết hợp của chúng như phương tiện quang bao gồm CD/DVD ROM/RW 1120 với CD/DVD hoặc phương tiện tương tự 1121, ổ chóp 1122, ổ cứng tháo rời hoặc ổ bán dẫn 1123, phương tiện từ truyền thống như băng và đĩa mềm (không được thể hiện), các thiết bị dựa trên ROM/ASIC/PLD chuyên dụng như các khóa điện tử bảo mật (không được thể hiện), và loại tương tự.

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

Hệ thống máy tính 1100 có thể cũng bao gồm (các) giao diện tới một hoặc nhiều mạng truyền thông 1155. Các mạng có thể ví dụ là mạng không dây, nối dây, quang học. Các mạng có thể là mạng cục bộ, mạng diện rộng, mạng đô thị, mạng phương tiện giao thông và công nghiệp, thời gian thực, dung sai trễ, và v.v. Ví dụ về các mạng bao gồm các mạng cục bộ như là Ethernet, LAN không dây, các mạng di động 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), Phát triển dài hạn (LTE), và loại tương tự, các mạng kỹ thuật số vùng rộng có dây hoặc không dây TV để bao gồm TV cáp, TV vệ tinh, và TV phát sóng trên mặt đất, trong phương tiện công cộng và trong xe cộ bao gồm CANBus, và v.v.. Các mạng nhất định thường yêu cầu các bộ thích ứng giao diện mạng phía ngoài 1154 mà được gắn vào cổng dữ liệu mục đích chung định trước hoặc các kênh truyền ngoại vi 1149 (như, ví dụ các cổng kênh truyền chuỗi vạn năng (USB) của hệ thống máy tính 1100); các phần khác thường được tích hợp vào lõi của hệ thống máy tính 1100 bằng cách gắn tới kênh truyền hệ thống như được mô tả dưới đây (ví dụ giao diện Ethernet vào hệ thống máy tính PC hoặc giao diện mạng tế bào vào hệ thống máy tính điện thoại thông minh). Lấy ví dụ, mạng 1155 có thể được kết nối với kênh truyền ngoại vi 1149 bằng cách sử dụng giao diện mạng 1154. Nhờ sử dụng bất kỳ các mạng này, hệ thống máy tính 1100 có thể truyền thông với các thực thể khác. Việc truyền thông này có thể là chỉ thu đơn hướng (ví dụ, TV quảng bá), chỉ gửi đơn hướng (ví dụ các thiết bị CANbus tới CANbus định trước), hoặc hai chiều, ví dụ tới các hệ thống máy tính khác nhờ sử dụng các mạng số diện rộng hoặc cục bộ. Các giao thức định trước và các ngăn giao thức có thể được sử dụng trên mỗi mạng và giao diện mạng 1154 này như được mô tả nêu trên.

Các thiết bị giao diện người dùng, các thiết bị lưu trữ có thể truy nhập được bởi người dùng, và các giao diện mạng nêu trên có thể được gắn vào lỗi 1140 của hệ thống máy tính 1100.

Lỗi 1140 có thể bao gồm một hoặc nhiều bộ xử lý trung tâm (CPU) 1141, các bộ xử lý đồ họa (GPU) 1142, các bộ xử lý khả trình chuyên dụng dưới dạng của mảng cổng khả trình dạng trường (FPGA) 1143, các bộ tăng tốc phần cứng 1144 đối với các tác vụ định trước, và v.v. Các thiết bị này, cùng với bộ nhớ chỉ đọc (ROM) 1145, bộ nhớ truy nhập ngẫu nhiên (RAM) 1146, bộ lưu trữ khối bên trong như các ổ cứng có thể không truy nhập bởi người dùng bên trong, ổ đĩa thể đặc (SSD), và loại tương tự 1147, có thể được kết nối thông qua kênh truyền hệ thống 1148. Trong một vài hệ thống máy tính, kênh truyền hệ thống 1148 có thể truy nhập được trong dạng của một hoặc nhiều đầu cắm vật lý để cho phép các phần mở rộng bởi các CPU, GPU bổ sung, và loại tương tự. Các thiết bị ngoại vi có thể được gắn trực tiếp tới kênh truyền hệ thống của lỗi 1148, hoặc thông qua kênh truyền ngoại vi 1149. Các cấu trúc của các kênh truyền ngoại vi bao gồm kết nối thành phần ngoại vi (PCI), USB, và loại tương tự.

Các CPU 1141, GPU 1142, FPGA 1143, và các bộ tăng tốc 1144 có thể thực thi các lệnh định trước mà, một cách kết hợp, có thể thực thi mã máy tính nêu trên. Mã máy tính có thể được lưu trữ trong ROM 1145 hoặc RAM 1146. Dữ liệu chuyển tiếp có thể cũng được lưu trữ trong RAM 1146, trong khi dữ liệu cố định có thể được lưu trữ ví dụ, trong bộ nhớ khối bên trong 1147. Việc lưu trữ nhanh và truy xuất tới bất kỳ thiết bị bộ nhớ có thể được cho phép thông qua việc sử dụng của bộ nhớ lưu trữ đệm, mà có thể được kết hợp mật thiết với một hoặc nhiều CPU 1141, GPU 1142, bộ lưu trữ cỡ lớn 1147, ROM 1145, RAM 1146, và loại tương tự.

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

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

Theo ví dụ của sáng chế và không bị giới hạn, hệ thống máy tính có cấu trúc 1100, và cụ thể, lõi 1140 có thể cung cấp chức năng như là kết quả của các bộ xử lý (bao gồm các CPU, GPU, FPGA, các bộ tăng tốc, và loại tương tự) mà thực thi phần mềm trong một hoặc nhiều phương tiện hữu hình đọc được bởi máy tính. Các phương tiện đọc được bởi máy tính này có thể là phương tiện được kết hợp với bộ lưu trữ khối có thể truy nhập bởi người dùng như được giới thiệu nêu trên, cũng như bộ lưu trữ định trước của lõi 1140 mà có bản chất bất biến, như bộ lưu trữ khối bên trong của lõi 1147 hoặc ROM 1145. Phần mềm mà thực hiện các phương án của sáng chế có thể được lưu trữ trong các thiết bị này và được thực thi bởi lõi 1140. Phương tiện đọc được bởi máy tính có thể bao gồm một hoặc nhiều thiết bị bộ nhớ hoặc chip, theo nhu cầu cụ thể. Phần mềm có thể làm cho lõi 1140 và một cách cụ thể, các bộ xử lý trong đó (bao gồm CPU, GPU, FPGA, và loại tương tự) thực hiện các xử lý cụ thể hoặc các phần cụ thể của các xử lý cụ thể được mô tả ở đây, bao gồm xác định các cấu trúc dữ liệu được lưu trữ trong RAM 1146 và cải biến các cấu trúc dữ liệu này theo các xử lý được xác định bởi phần mềm. Ngoài ra hoặc như là tùy chọn khác, hệ thống máy tính có thể cung cấp chức năng như là kết quả của mạng cứng logic hoặc được lắp trong mạch (ví dụ: bộ tăng tốc 1144), mà có thể hoạt động trong vị trí của hoặc cùng với phần mềm để thực hiện các xử lý cụ thể hoặc các phần cụ thể của các xử lý cụ thể được mô tả ở đây. Việc tham chiếu tới phần mềm có thể bao hàm logic, và ngược lại, nếu cần. Việc tham chiếu tới phương tiện đọc được bởi máy tính có thể bao hàm mạch (như mạch tích hợp (IC)) mà lưu trữ phần mềm để thực thi, mạch thực hiện logic để thực thi, hoặc cả hai, nếu cần. Sáng chế bao hàm kết hợp thích hợp của phần cứng và phần mềm.

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

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

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 chuỗi video được mã hóa từ dòng bit video được mã hóa;

thu nhận phần tử cú pháp thứ nhất chỉ báo số lượng các lớp nằm trong chuỗi video được mã hóa;

thu nhận phần tử cú pháp thứ hai chỉ báo xem các lớp nằm trong chuỗi video được mã hóa đều là các lớp độc lập;

dựa trên phần tử cú pháp thứ nhất chỉ báo rằng số lượng các lớp lớn hơn một, và phần tử cú pháp thứ hai chỉ báo rằng các lớp không phải tất cả là các lớp độc lập, thu nhận phần tử cú pháp thứ ba chỉ báo chế độ đầu ra;

xác định chế độ đầu ra dựa trên phần tử cú pháp thứ ba;

xác định một hoặc nhiều lớp đầu ra từ trong số các lớp dựa trên chế độ đầu ra được xác định; và

xuất ra một hoặc nhiều lớp đầu ra.

2. Phương pháp theo điểm 1, trong đó dựa trên phần tử cú pháp thứ ba chỉ báo rằng chế độ đầu ra là chế độ thứ nhất, một hoặc nhiều lớp đầu ra chỉ bao gồm lớp cao nhất từ trong số các lớp.

3. Phương pháp theo điểm 1, trong đó dựa trên phần tử cú pháp thứ ba chỉ báo rằng chế độ đầu ra là chế độ thứ hai, một hoặc nhiều lớp đầu ra bao gồm các lớp.

4. Phương pháp theo điểm 1, còn bao gồm thu nhận phần tử cú pháp thứ tư chỉ báo rằng lớp tương ứng của các lớp cần được xuất ra,

trong đó dựa trên phần tử cú pháp thứ ba chỉ báo rằng chế độ đầu ra là chế độ thứ ba, một hoặc nhiều lớp đầu ra bao gồm lớp tương ứng.

5. Phương pháp theo điểm 1, trong đó phần tử cú pháp thứ nhất, phần tử cú pháp thứ hai, và phần tử cú pháp thứ ba được báo hiệu trong tập tham số video (VPS).

6. Phương pháp theo điểm 5, trong đó VPS còn bao gồm phần tử cú pháp thứ năm chỉ báo số lượng các khuôn dạng hình ảnh được báo hiệu trong VPS, và phần tử cú pháp thứ sáu chỉ báo chỉ số của khuôn dạng hình ảnh từ trong số các khuôn dạng hình ảnh cần được áp dụng cho lớp từ trong số các lớp.

7. Phương pháp theo điểm 5, trong đó VPS còn bao gồm phần tử cú pháp thứ bảy chỉ báo xem các tham số hình ảnh con được báo hiệu trong VPS hay không.

8. Phương pháp theo điểm 5, trong đó tập tham số chuỗi (SPS) đề cập đến VPS bao gồm phần tử cú pháp thứ tám chỉ báo rằng SPS bao gồm phần tử cú pháp thứ chín chỉ báo chỉ số của khuôn dạng hình ảnh từ trong số các khuôn dạng hình ảnh được báo hiệu trong VPS cần được áp dụng cho lớp từ trong số các lớp.

9. Phương pháp theo điểm 8, trong đó SPS còn bao gồm phần tử cú pháp thứ sáu chỉ báo xem các tham số hình ảnh con được báo hiệu trong SPS hay không.

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

ít nhất một bộ nhớ có cấu trúc để lưu trữ mã chương trình; và

ít nhất một bộ xử lý được tạo cấu hình để đọc mã chương trình và thao tác khi được lệnh bởi mã chương trình, mã chương trình bao gồm:

mã thu nhận thứ nhất được tạo cấu hình để khiến ít nhất một bộ xử

lý thu nhận chuỗi video được mã hóa từ dòng bit video được mã hóa;

mã thu nhận thứ hai được tạo cấu hình để khiến ít nhất một bộ xử lý thu nhận phần tử cú pháp thứ nhất chỉ báo số lượng các lớp nằm trong chuỗi video được mã hóa;

mã thu nhận thứ ba được tạo cấu hình để khiến ít nhất một bộ xử lý thu nhận phần tử cú pháp thứ hai chỉ báo xem các lớp nằm trong chuỗi video được mã hóa có đều là các lớp độc lập hay không;

mã thu nhận thứ tư được tạo cấu hình để khiến ít nhất một bộ xử lý, dựa trên phần tử cú pháp thứ nhất chỉ báo rằng số lượng các lớp lớn hơn một, và phần tử cú pháp thứ hai chỉ báo rằng các lớp không phải tất cả là các lớp độc lập, thu nhận phần tử cú pháp thứ ba chỉ báo chế độ đầu ra;

mã xác định thứ nhất được tạo cấu hình để khiến ít nhất một bộ xử lý để xác định chế độ đầu ra dựa trên phần tử cú pháp thứ ba;

mã xác định thứ hai được tạo cấu hình để khiến ít nhất một bộ xử lý để xác định một hoặc nhiều lớp đầu ra từ trong số các lớp dựa trên chế độ đầu ra được xác định; và

mã đầu ra được tạo cấu hình để khiến ít nhất một bộ xử lý xuất ra một hoặc nhiều lớp đầu ra.

11. Thiết bị theo điểm 10, trong đó dựa trên phần tử cú pháp thứ ba chỉ báo rằng chế độ đầu ra là chế độ thứ nhất, một hoặc nhiều lớp đầu ra chỉ bao gồm lớp cao nhất từ trong số các lớp.

12. Thiết bị theo điểm 10, trong đó dựa trên phần tử cú pháp thứ ba chỉ báo rằng chế độ đầu ra là chế độ thứ hai, một hoặc nhiều lớp đầu ra bao gồm các lớp.

13. Thiết bị theo điểm 10, trong đó mã chương trình còn bao gồm

thu nhận lần thứ năm phần tử cú pháp thứ tư chỉ báo rằng lớp tương ứng của các lớp cần được xuất ra,

trong đó dựa trên phần tử cú pháp thứ ba chỉ báo rằng chế độ đầu ra là chế độ thứ ba, một hoặc nhiều lớp đầu ra bao gồm lớp tương ứng.

14. Thiết bị theo điểm 10, trong đó phần tử cú pháp thứ nhất, phần tử cú pháp thứ hai, và phần tử cú pháp thứ ba được báo hiệu trong tập tham số video (VPS).

15. Thiết bị theo điểm 14, trong đó VPS còn bao gồm phần tử cú pháp thứ năm chỉ báo số lượng các khuôn dạng hình ảnh được báo hiệu trong VPS, và phần tử cú pháp thứ sáu chỉ báo chỉ số của khuôn dạng hình ảnh từ trong số các khuôn dạng hình ảnh cần được áp dụng cho lớp từ trong số các lớp.

16. Thiết bị theo điểm 14, trong đó VPS còn bao gồm phần tử cú pháp thứ bảy chỉ báo xem các tham số hình ảnh con được báo hiệu trong VPS hay không.

17. Thiết bị theo điểm 14, trong đó tập tham số chuỗi (SPS) đề cập đến VPS bao gồm phần tử cú pháp thứ tám chỉ báo rằng SPS bao gồm phần tử cú pháp thứ chín chỉ báo chỉ số của khuôn dạng hình ảnh từ trong số các khuôn dạng hình ảnh được báo hiệu trong VPS cần được áp dụng cho lớp từ trong số các lớp.

18. Thiết bị theo điểm 17, trong đó SPS còn bao gồm phần tử cú pháp thứ sáu chỉ báo xem các tham số hình ảnh con được báo hiệu trong SPS hay không.

19. Phương tiện bất biến đọc được bằng máy tính lưu trữ các lệnh máy, các lệnh máy này bao gồm: một hoặc nhiều lệnh mà, khi được thực hiện 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:

thu nhận chuỗi video được mã hóa từ dòng bit video được mã hóa;

thu nhận phần tử cú pháp thứ nhất chỉ báo số lượng các lớp nằm trong chuỗi video được mã hóa;

thu nhận phần tử cú pháp thứ hai chỉ báo xem các lớp nằm trong chuỗi video được mã hóa có đều là các lớp độc lập không;

dựa trên phần tử cú pháp thứ nhất chỉ báo rằng số lượng các lớp lớn hơn một, và phần tử cú pháp thứ hai chỉ báo rằng các lớp không phải tất cả là các lớp độc lập, thu nhận phần tử cú pháp thứ ba chỉ báo chế độ đầu ra;

xác định chế độ đầu ra dựa trên phần tử cú pháp thứ ba;

xác định một hoặc nhiều lớp đầu ra từ trong số các lớp dựa trên chế độ đầu ra được xác định; và

xuất ra một hoặc nhiều lớp đầu ra.

20. Phương tiện bất biến đọc được bằng máy tính theo điểm 19, trong đó phần tử cú pháp thứ nhất, phần tử cú pháp thứ hai, và phần tử cú pháp thứ ba được báo hiệu trong tập tham số video (VPS).

1/11

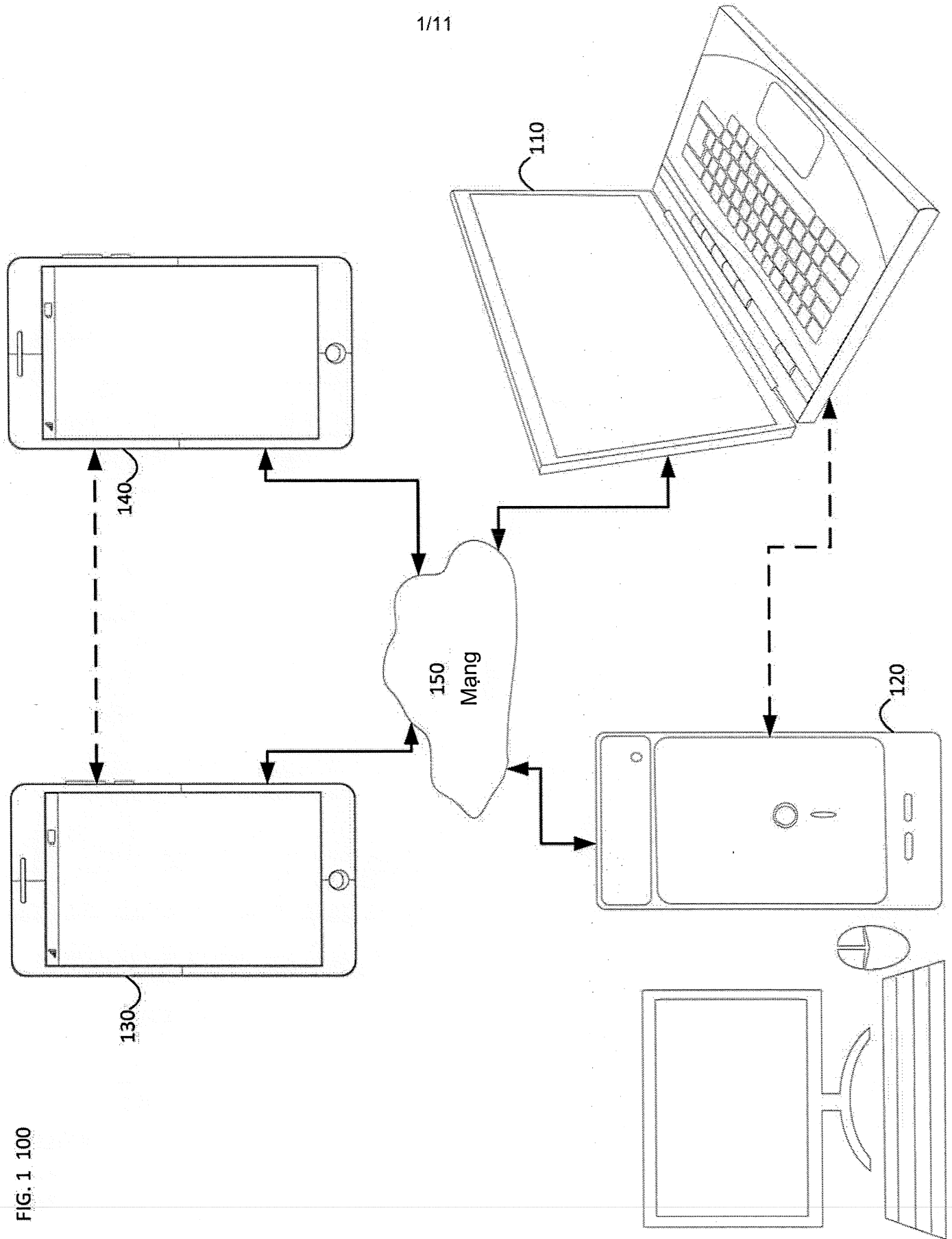


FIG. 1 100

2/11

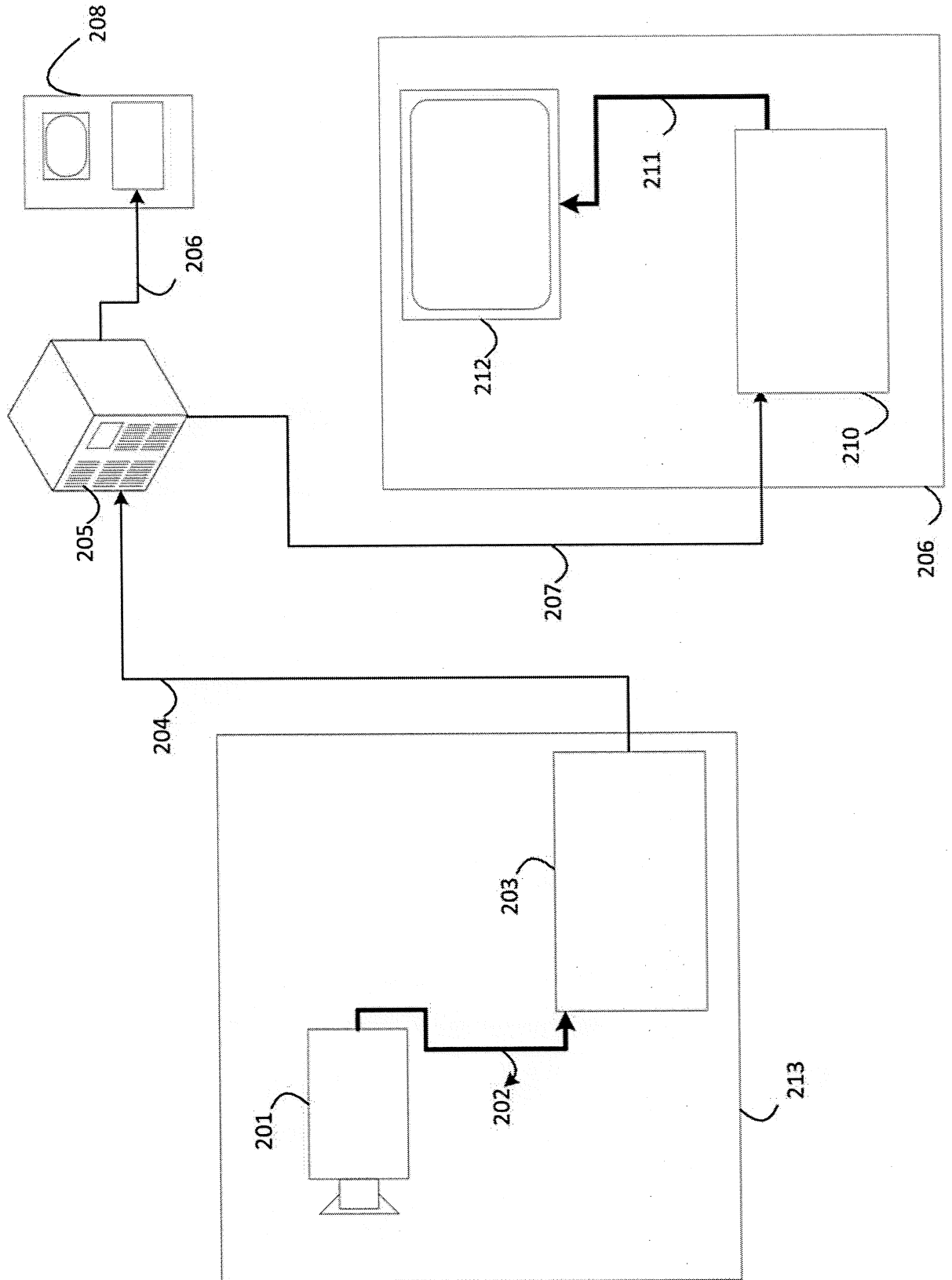
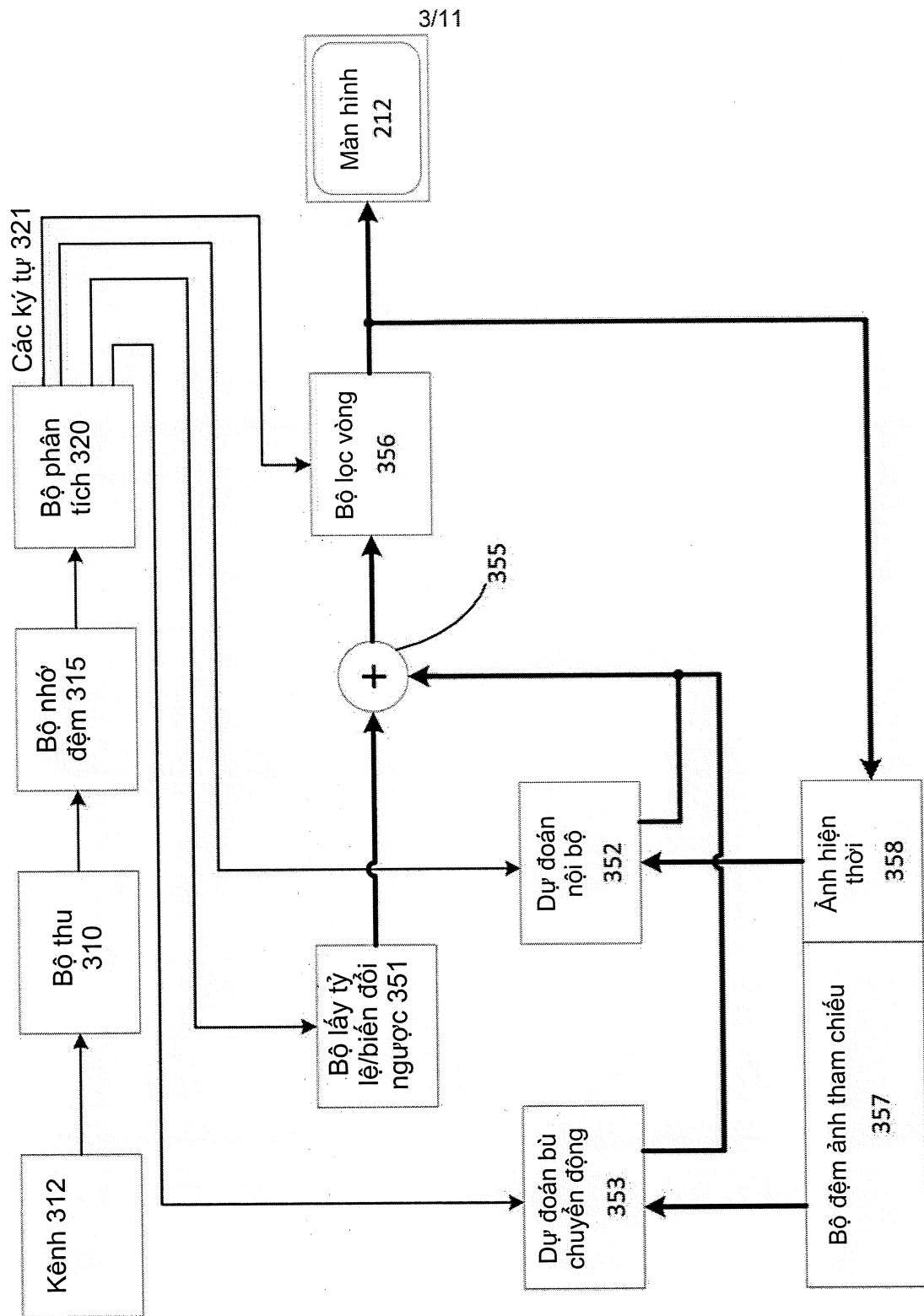


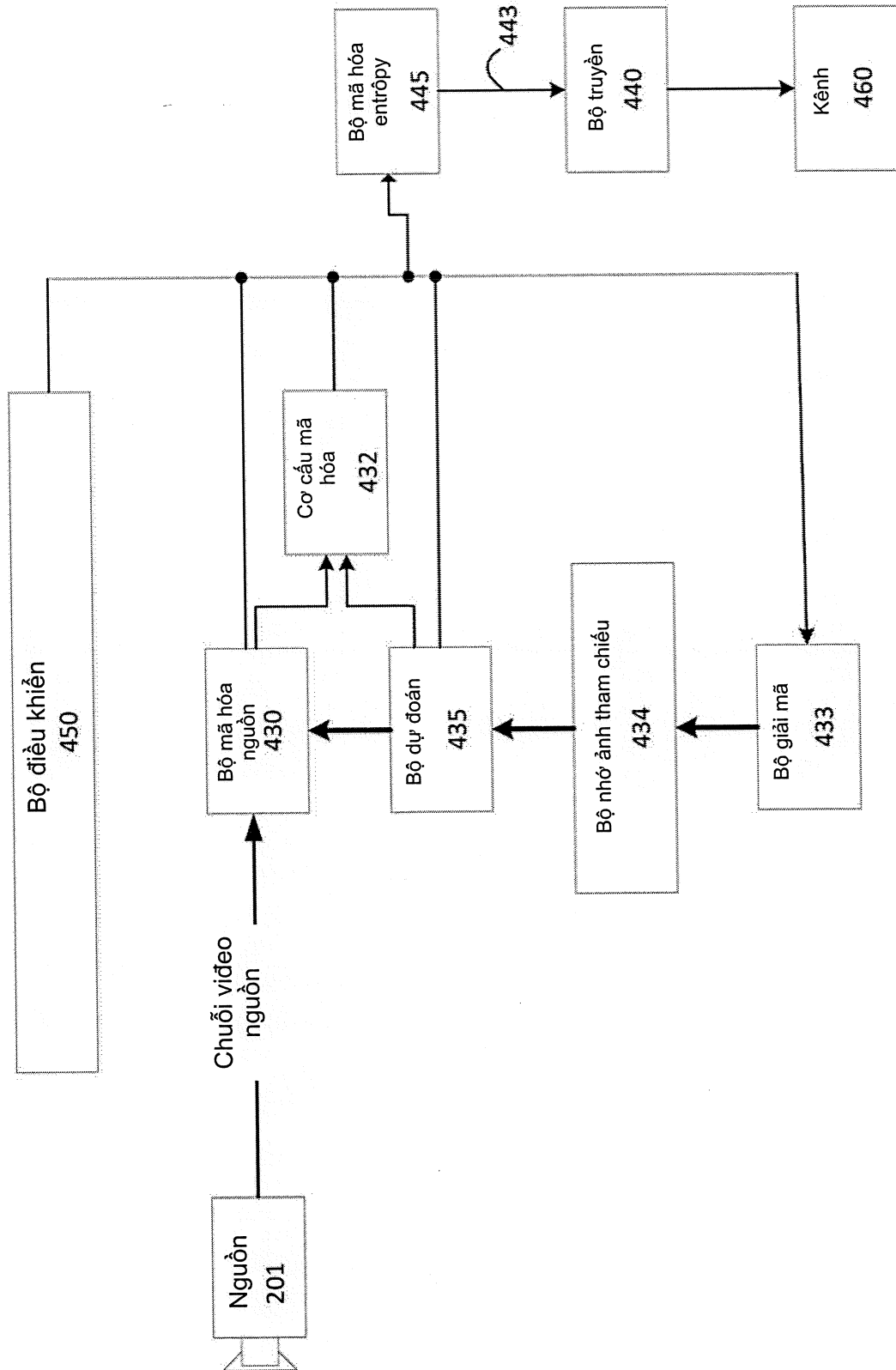
FIG. 2

FIG. 3 Bộ giải mã 210



4/11

FIG. 4 Bộ mã hóa 203



5/11

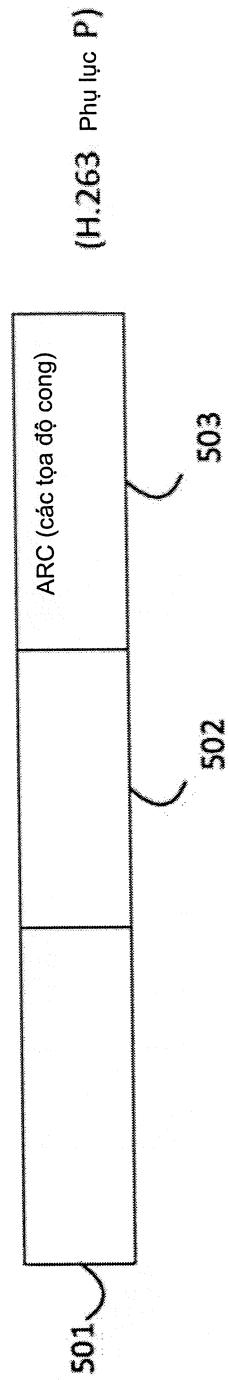


FIG. 5

6/11

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() {	Descriptor
	...	
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)
	}	
	num_dec_pic_size_in_luma_samples_minus1	ue(v)
616	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

7/11

FIG. 7

701	video parameter set rbsp() {	Descriptor
702	vps video parameter set id	u(4)
703	vps max layers minus1	u(6)
704	if(vps max layers minus1 > 0)	
	vps all independent layers flag	u(1)
705	for(i = 0; i <= vps max layers minus1; i++) {	
	vps layer id[i]	u(6)
706	if(i > 0 && !vps all independent layers flag) {	
	vps independent layer flag[i]	u(1)
	if(!vps independent layer flag[i])	
	for(j = 0; j < i; j++)	
707	vps direct dependency flag[i][j]	u(1)
	}	
	}	
708	if(vps max layers minus1 > 0 && !vps all independent layers flag) {	
	vps output layers mode	u(2)
	if(vps output layers mode == 2)	
709	for(i = 0; i < vps max layers minus1; i++)	
	vps output layer flag[i]	u(1)
	}	
	vps constraint info present flag	u(1)
	vps reserved zero 7bits	u(7)
	if(vps constraint info present flag)	
	general constraint info()	
	vps extension flag	u(1)
	if(vps extension flag)	
	while(more rbsp data())	
	vps extension data flag	u(1)
	rbsp_trailing_bits()	
	}	

8/11

FIG. 8

	Descriptor
video_parameter_set_rbsp() {	
...	
vps_num_rep_formats_minus1	ue(v)
for(i = 0; i <= vps_num_rep_formats_minus1; i++)	
rep_format()	
for(i = 0; i < vps_max_layers_minus1; i++)	
vps_rep_format_idx[i]	u(v)
...	
}	

	Descriptor
rep_format() {	
pic_width_max_vps_in_luma_samples	u(16)
pic_height_max_vps_in_luma_samples	u(16)
chroma_format_vps_idc	u(2)
if(chroma_format_vps_idc == 3)	
separate_colour_plane_vps_flag	u(1)
bit_depth_vps_luma_minus8	u(4)
bit_depth_vps_chroma_minus8	u(4)
subpics_vps_present_flag	u(1)
if(subpics_vps_present_flag) {	
max_subpics_vps_minus1	u(8)
subpic_grid_vps_col_width_minus1	u(v)
subpic_grid_vps_row_height_minus1	u(v)
for(i = 0; i < NumSubPicVpsGridRows; i++)	
for(j = 0; j < NumSubPicVpsGridCols; j++)	
subpic_grid_vps_idx[i][j]	u(v)
for(i = 0; i <= NumSubPicsVps; i++) {	
subpic_treated_as_pic_vps_flag[i]	u(1)
loop_filter_across_subpic_enabled_vps_flag[i]	u(1)
}	
}	
}	

9/11

FIG. 9

seq_parameter set rbsp() {	Descriptor
...	
use_rep_format_idx_flag	u(1)
if(use_rep_format_idx_flag)	
sps_rep_format_idx	u(8)
} else {	
pic_width_in_luma_samples	ue(v)
pic_height_in_luma_samples	ue(v)
chroma_format_idc	ue(v)
if(chroma_format_idc == 3)	
separate_colour_plane_flag	u(1)
bit_depth_luma_minus8	ue(v)
bit_depth_chroma_minus8	ue(v)
subpics_present_flag	u(1)
if(subpics_present_flag) {	
max_subpics_minus1	u(8)
subpic_grid_col_width_minus1	u(v)
subpic_grid_row_height_minus1	u(v)
for(i = 0; i < NumSubPicGridRows; i++)	
for(j = 0; j < NumSubPicGridCols; j++)	
subpic_grid_idx[i][j]	u(v)
for(i = 0; i <= NumSubPics; i++) {	
subpic_treated_as_pic_flag[i]	u(1)
loop_filter_across_subpic_enabled_flag[i]	u(1)
}	
}	

10/11

