



(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)



1-0041753

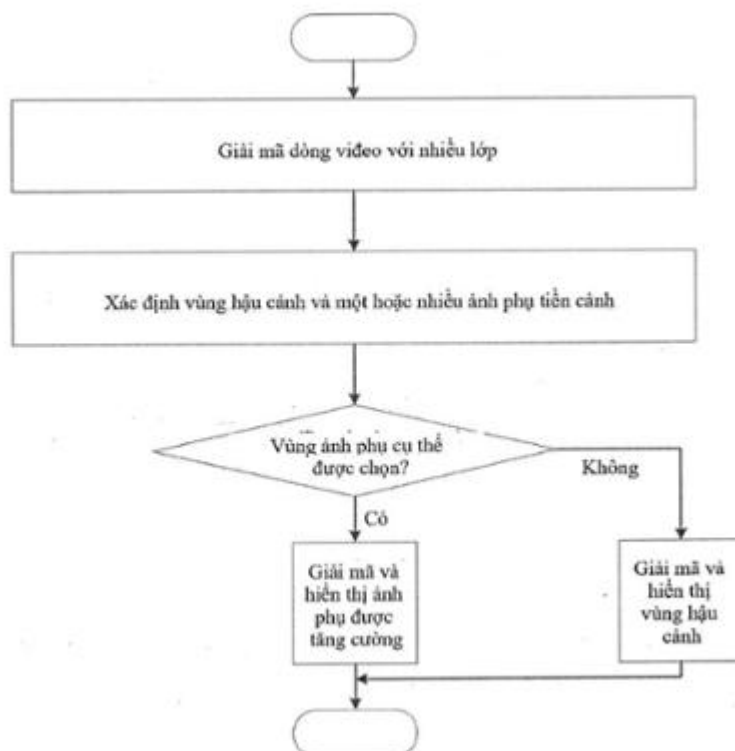
(51)^{2020.01} H04N 19/33; H04N 19/177; H04N
19/70; H04N 19/597; H04N 19/172

(13) B

(21) 1-2021-06411 (22) 18/09/2020
(86) PCT/US2020/051477 18/09/2020 (87) WO2021/055738 25/03/2021
(30) 62/903,660 20/09/2019 US; 17/021,243 15/09/2020 US
(45) 25/11/2024 440 (43) 27/06/2022 411
(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 Luật TNHH Phạm và Liên danh (PHAM & ASSOCIATES)

(54) PHƯƠNG PHÁP VÀ HỆ THỐNG MÁY TÍNH ĐỂ BẢO TÍN HIỆU CÁC TẬP LỚP ĐẦU RA TRONG DÒNG VIDEO ĐƯỢC MÃ HÓA

(57) Sáng chế đề xuất phương pháp, chương trình máy tính, và hệ thống máy tính để bảo tín hiệu các tập lớp đầu ra trong dòng video được mã hóa. Dữ liệu video có nhiều lớp được nhận. Một hoặc nhiều phần tử cú pháp được xác định. Các phần tử cú pháp này chỉ rõ một hoặc nhiều tập lớp đầu ra tương ứng với các lớp đầu ra từ trong số nhiều lớp của dữ liệu video được nhận. Một hoặc nhiều lớp đầu ra này tương ứng với các tập lớp đầu ra được định rõ được giải mã và được hiển thị.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập chung đến lĩnh vực mã hóa và giải mã video, và cụ thể hơn tới việc tham chiếu và phạm vi của tập thông số trong dòng video được mã hóa.

Tình trạng kỹ thuật của sáng chế

Việc mã hóa và giải mã video bằng cách sử dụng dự báo ngoài ảnh có bù chuyển động đã được biết đến hàng chục năm nay. Video kỹ thuật số chưa được nén có thể bao gồm các dãy ảnh, mỗi ảnh có kích thước không gian bằng, chẳng hạn, 1920 x 1080 mẫu độ sáng và mẫu sắc độ liên kết. Các dãy ảnh này có thể có tốc độ ảnh (còn được biết một cách không chính thức là tốc độ khung) cố định hoặc biến thiên, chẳng hạn bằng 60 ảnh trên giây hoặc 60 Hz. Video chưa được nén có các yêu cầu đáng kể về tốc độ bit. Ví dụ, video 1080p60 4:2:0 ở 8 bit trên mẫu (độ phân giải mẫu độ sáng 1920x1080 ở tốc độ khung 60 Hz) cần có băng thông gần 1.5 Gbit/s. Một giờ video như vậy cần trên 600 GByte 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 kỹ thuật nén. Việc nén có thể làm giảm các yêu cầu về băng thông hoặc không gian lưu trữ nêu trên, trong một số trường hợp đến hai bậc về độ lớn hoặc hơn. Cả cách nén không tổn hao lẫn có tổn hao, cũng như sự kết hợp của chúng đều có thể được sử dụng. Nén không tổn hao là các kỹ thuật khi bản sao chính xác của tín hiệu ban đầu có thể được tái tạo từ tín hiệu ban đầu đã được nén. Khi sử dụng kỹ thuật nén có tổn hao, thì tín hiệu được tái tạo có thể không giống hệt tín hiệu ban đầu, nhưng độ méo giữa tín hiệu ban đầu và tín hiệu được tái tạo là đủ nhỏ để làm cho tín hiệu được tái tạo hữu dụng đối với ứng dụng được dự kiến. Trong trường hợp video, kỹ thuật nén có tổn hao được áp dụng rộng rãi. Mức độ méo chấp nhận được phụ thuộc vào ứng dụng; chẳng hạn, người dùng của một số ứng dụng phát trực tuyến (streaming)

cá nhân nhất định có thể chấp nhận độ méo cao hơn người dùng của các ứng dụng tham gia vô tuyến truyền hình. Tỷ số nén đạt được có thể phản ánh rằng: độ méo cho phép/chấp nhận được càng cao thì có thể đạt được tỷ số nén càng cao.

Bộ tạo mã và giải mã video có thể sử dụng các kỹ thuật từ một số chủng loại rộng, bao gồm, chẳng hạn, bù chuyển động, biến đổi, lượng tử hóa, và mã hóa entropi, mà một vài trong số đó sẽ được giới thiệu dưới đây.

Trong quá khứ, các bộ tạo mã và bộ giải mã video có xu hướng hoạt động trên kích thước ảnh cho trước mà trong phần lớn các trường hợp, được xác định và được giữ không đổi đối với dãy video được mã hóa (CVS), nhóm ảnh (Group of Pictures - GOP), hoặc khung thời gian đa ảnh tương tự. Chẳng hạn, trong MPEG-2, được biết rằng các thiết kế hệ thống thay đổi độ phân giải ngang (và, nhờ đó, cả kích thước ảnh) tùy theo các yếu tố như mức độ hoạt động của cảnh, nhưng chỉ tại các ảnh I, nên đặc trưng cho GOP. Đã biết việc lấy mẫu lại các ảnh tham chiếu để dùng các độ phân giải khác nhau bên trong CVS, chẳng hạn, từ ITU-T Rec. H.263 Annex P. Tuy nhiên, tại đây kích thước ảnh không thay đổi, mà chỉ các ảnh tham chiếu được lấy mẫu lại, dẫn đến có khả năng là chỉ trong các phần của canvas ảnh được sử dụng (trong trường hợp giảm mẫu), hoặc chỉ các phần của cảnh được thu (trong trường hợp tăng mẫu). Tiếp theo, H.263 Annex Q cho phép lấy mẫu lại khối macro riêng lẻ theo hệ số hai (theo mỗi chiều kích thước), tăng lên hoặc giảm xuống. Kích thước ảnh lại vẫn giữ nguyên. Kích thước của khối macro được cố định trong H.263, và do đó không cần được báo tín hiệu.

Sự thay đổi kích thước ảnh trong các ảnh được dự báo trở nên chủ đạo hơn trong kỹ thuật mã hóa video hiện đại. Chẳng hạn, VP9 cho phép lấy mẫu lại ảnh tham chiếu và thay đổi độ phân giải cho toàn ảnh. Tương tự, có một số đề xuất hướng tới VVC (bao gồm, chẳng hạn, Hendry, et. al, “On adaptive resolution change (ARC) for VVC”, Joint Video Team document JVET-M0135-v1, Jan 9-19, 2019, được kết hợp toàn bộ vào đây) cho phép lấy mẫu lại toàn bộ các ảnh tham chiếu tới các độ phân giải cao hơn hoặc thấp hơn khác nhau. Trong tài liệu này, các độ phân giải ứng viên khác nhau được đề xuất là được mã hóa

trong tập thông số dãy và được gọi là các phần tử cú pháp theo ảnh trong tập thông số ảnh.

Bản chất kỹ thuật của sáng chế

Các phương án của sáng chế đề cập đến phương pháp, hệ thống, vật ghi máy tính đọc được dành cho việc báo tín hiệu các tập lớp đầu ra trong dữ liệu video được mã hóa. Theo một khía cạnh của sáng chế, phương pháp để báo tín hiệu các tập lớp đầu ra trong dữ liệu video được mã hóa được đề xuất. Phương pháp có thể bao gồm việc nhận được dữ liệu video có nhiều lớp. Một hoặc nhiều phần tử cú pháp được xác định. Các phần tử cú pháp chỉ rõ một hoặc nhiều tập lớp đầu ra tương ứng với các lớp đầu ra từ trong số nhiều lớp dữ liệu video được nhận. Một hoặc nhiều lớp đầu ra tương ứng với các tập lớp đầu ra được định rõ được giải mã và được hiển thị.

Theo một khía cạnh khác, hệ thống máy tính dành cho việc báo tín hiệu các tập lớp đầu ra được đề xuất. Hệ thống máy tính này có thể bao gồm một hoặc nhiều bộ xử lý, một hoặc nhiều bộ nhớ máy tính đọc được, một hoặc nhiều bộ phận lưu trữ hữu hình máy tính đọc được, và các lệnh chương trình được lưu trữ trong ít nhất một trong số một hoặc nhiều bộ phận lưu trữ để thực thi bởi ít nhất một trong số một hoặc nhiều bộ xử lý thông qua ít nhất một trong số một hoặc nhiều bộ nhớ, nhờ đó hệ thống máy tính có thể thực hiện được phương pháp. Phương pháp có thể bao gồm nhận dữ liệu video có nhiều lớp. Một hoặc nhiều phần tử cú pháp được xác định. Các phần tử cú pháp chỉ rõ một hoặc nhiều tập lớp đầu ra tương ứng với các lớp đầu ra từ trong số nhiều lớp dữ liệu video được nhận. Một hoặc nhiều lớp đầu ra tương ứng với các tập lớp đầu ra được định rõ được giải mã và được hiển thị.

Theo một khía cạnh khác nữa, vật ghi máy tính đọc được dành cho việc báo tín hiệu các tập lớp đầu ra được đề xuất. Vật ghi máy tính đọc được có thể bao gồm một hoặc nhiều bộ phận lưu trữ máy tính đọc được và các lệnh chương trình được lưu trữ trong ít nhất một trong số một hoặc nhiều bộ phận lưu trữ hữu hình, các lệnh chương trình có thể được thực thi bởi bộ xử lý. Các lệnh chương trình có thể được thực thi bởi bộ xử lý để thực hiện phương pháp mà có thể một

cách tương ứng bao gồm việc nhận dữ liệu video có nhiều lớp. Một hoặc nhiều phần tử cú pháp được xác định. Các phần tử cú pháp chỉ rõ một hoặc nhiều tập lớp đầu ra tương ứng với các lớp đầu ra từ trong số nhiều lớp của dữ liệu video được nhận. Một hoặc nhiều lớp đầu ra tương ứng với các tập lớp đầu ra được định rõ được giải mã và được hiển thị.

Mô tả vắn tắt các hình vẽ

Các mục đích, dấu hiệu và lợi ích của sáng chế sẽ trở nên rõ ràng từ phần mô tả chi tiết sau đây của các phương án minh họa sáng chế mà cần được xem xét cùng với các hình vẽ kèm theo. Các dấu hiệu khác nhau của các hình vẽ không được định tỷ lệ để việc minh họa trở nên rõ ràng và tạo thuận lợi cho người có kiến thức trung bình trong lĩnh vực hiểu rõ khi xem xét cùng với phần mô tả chi tiết. Trong các hình vẽ:

FIG. 1 là sơ đồ minh họa của sơ đồ khối đơn giản hóa của hệ thống truyền thông theo một phương án của sáng chế;

FIG. 2 là sơ đồ minh họa của sơ đồ khối đơn giản hóa của hệ thống truyền thông theo một phương án của sáng chế;

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

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

FIG. 5 là sơ đồ minh họa của các tùy chọn để báo tín hiệu các thông số ARC theo một phương án của sáng chế;

FIG. 6 là ví dụ về bảng cú pháp theo một phương án của sáng chế;

FIG. 7 là sơ đồ minh họa về hệ thống máy tính theo một phương án của sáng chế;

FIG. 8 là ví dụ về cấu trúc dự báo đối với khả năng định tỷ lệ với thay đổi độ phân giải thích ứng;

FIG. 9 là ví dụ về bảng cú pháp theo một phương án của sáng chế;

FIG. 10 là sơ đồ minh họa của sơ đồ khối đơn giản hóa của việc phân tách và giải mã chu kỳ poc trên bộ truy cập và giá trị đếm bộ truy cập;

FIG. 11 là sơ đồ minh họa của cấu trúc dòng bit video bao gồm các ảnh phụ nhiều lớp;

FIG. 12 là sơ đồ minh họa màn hiển thị của ảnh phụ được chọn với độ phân giải được tăng cường;

FIG. 13 là sơ đồ khối của quá trình giải mã và hiển thị đối với dòng bit video bao gồm các ảnh phụ nhiều lớp;

FIG. 14 là sơ đồ minh họa hiển thị video 360 với lớp tăng cường của ảnh phụ;

FIG. 15 là ví dụ về thông tin bố trí của các ảnh phụ và lớp tương ứng và cấu trúc dự báo ảnh của ảnh này;

FIG. 16 là ví dụ về thông tin bố trí của các ảnh phụ và lớp tương ứng và cấu trúc dự báo ảnh của ảnh này, với phương thức mở rộng không gian của vùng địa phương;

FIG. 17 là ví dụ về bảng cú pháp dành cho thông tin bố trí ảnh phụ;

FIG. 18 là ví dụ về bảng cú pháp của thông điệp SEI dành cho thông tin bố trí ảnh phụ;

FIG. 19 là ví dụ về bảng cú pháp chỉ báo các lớp đầu ra và thông tin tiêu sử/bậc/mức cho mỗi tập lớp đầu ra;

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

FIG. 21 là ví dụ về bảng cú pháp chỉ báo ảnh phụ hiện tại của mỗi lớp đối với mỗi tập lớp đầu ra.

Mô tả chi tiết sáng chế

Các phương án chi tiết của các kết cấu và các phương pháp theo sáng chế được mô tả ở đây; tuy nhiên, cần hiểu rằng các phương án được mô tả chỉ là minh họa cho các kết cấu và các phương pháp theo sáng chế có thể được thực hiện dưới các dạng khác nhau. Tuy nhiên, các kết cấu và các phương pháp này có thể được thực hiện dưới các dạng khác nhau và không được coi là bị giới hạn ở các phương án được nêu làm ví dụ ở đây. Hơn thế nữa, các phương án được nêu làm ví dụ này được đề xuất sao cho phần mô tả này trở nên đầy đủ và kỹ

càng và chuyển tải được phạm vi của sáng chế đến những người có hiểu biết trung bình trong lĩnh vực này. Trong phần mô tả, các chi tiết của các dấu hiệu và kỹ thuật đã biết có thể được bỏ qua để tránh làm lu mờ một cách không cần thiết các phương án được trình bày.

Các phương án đề cập chung đến lĩnh vực xử lý dữ liệu, và cụ thể hơn đến việc xử lý media. Phần sau đây mô tả các phương án ví dụ nhằm đề xuất hệ thống, phương pháp và chương trình máy tính để, ngoài những mục đích khác, cho phép báo tín hiệu các tập lớp đầu ra của dữ liệu video được mã hóa. Do đó, một số phương án có khả năng cải tiến lĩnh vực tính toán bằng việc tạo mã và giải mã video được cải tiến.

Như đã mô tả trên đây, các bộ tạo mã và các bộ giải mã video có xu hướng hoạt động trên kích thước ảnh cho trước mà trong phần lớn các trường hợp, được xác định và được giữ không đổi đối với dãy video được mã hóa (CVS), nhóm ảnh (Group of Pictures - GOP), hoặc khung thời gian đa ảnh tương tự. Chẳng hạn, trong MPEG-2, được biết rằng các thiết kế hệ thống thay đổi độ phân giải ngang (và, nhờ đó, cả kích thước ảnh) tùy theo các yếu tố như mức độ hoạt động của cảnh, nhưng chỉ tại các ảnh I, nên đặc trưng cho GOP. Đã biết việc lấy mẫu lại các ảnh tham chiếu để dùng các độ phân giải khác nhau bên trong CVS, chẳng hạn, từ ITU-T Rec. H.263 Annex P. Tuy nhiên, tại đây kích thước ảnh không thay đổi, mà chỉ các ảnh tham chiếu được lấy mẫu lại, dẫn đến có khả năng là chỉ trong các phần của canvas ảnh được sử dụng (trong trường hợp giảm mẫu), hoặc chỉ các phần của cảnh được thu (trong trường hợp tăng kích thước mẫu). Tiếp theo, H.263 Annex Q cho phép lấy mẫu lại khối macro riêng lẻ theo hệ số hai (theo mỗi chiều kích thước), tăng lên hoặc giảm xuống. Kích thước ảnh lại vẫn giữ nguyên. Kích thước của khối macro được cố định trong H.263, và do đó không cần được báo tín hiệu.

Tuy nhiên, trong ngữ cảnh của, chẳng hạn, mã hóa 360 hoặc một số ứng dụng giám sát, nhiều ảnh nguồn độc lập về ngữ nghĩa (chẳng hạn sáu bề mặt khối lập phương của cảnh 360 chiếu lập phương, hoặc các đầu vào camera riêng lẻ trong trường hợp thiết lập giám sát nhiều camera) có thể yêu cầu các cài đặt độ phân giải thích ứng riêng rẽ để đáp ứng mức độ hoạt động trên từng cảnh khác

nhau tại thời điểm cho trước. Nói cách khác, các bộ tạo mã, tại thời điểm cho trước, có thể chọn sử dụng các hệ số lấy mẫu lại khác nhau cho các ảnh độc lập về ngữ nghĩa khác nhau mà tạo nên 360 toàn cảnh hoặc cảnh giám sát. Khi được kết hợp thành một ảnh, mà đến lượt mình, lại yêu cầu thực hiện việc lấy mẫu lại ảnh tham chiếu, và việc báo tín hiệu mã hóa độ phân giải thích ứng là khả thi, đối với các phần của ảnh được mã hóa. Do đó, có thể có lợi khi sử dụng dữ liệu báo tín hiệu mã hóa độ phân giải thích ứng khả dụng để báo tín hiệu, tạo mã, giải mã, và hiển thị tốt hơn các lớp video.

FIG. 1 thể hiện 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 đầu cuối (110-120) được kết nối qua mạng (150). Để truyền dữ liệu đơn hướng, đầu cuối thứ nhất (110) có thể mã hóa dữ liệu video tại vị trí địa phương để truyền đến đầu cuối kia (120) qua mạng (150). Đầu cuối thứ hai (120) có thể nhận dữ liệu video được mã hóa của đầu cuối kia từ mạng (150), giải mã dữ liệu được mã hóa và hiển thị dữ liệu video được phục hồi. Dữ liệu truyền đơn hướng có thể là thông dụng trong các ứng dụng dịch vụ truyền thông và tương tự.

FIG. 1 thể hiện cặp đầu cuối thứ hai (130, 140) được đề xuất để hỗ trợ truyền hai hướng video được mã hóa mà có thể diễn ra, chẳng hạn, trong khi hội nghị video. Để truyền dữ liệu hai hướng, mỗi đầu cuối (130, 140) có thể mã hóa dữ liệu video được thu ở vị trí địa phương để truyền đến đầu cuối kia qua mạng (150). Mỗi đầu cuối (130, 140) cũng có thể nhận dữ liệu video được mã hóa được truyền bởi đầu cuối kia, có thể giải mã dữ liệu được mã hóa và có thể hiển thị dữ liệu video được phục hồi tại bộ phận hiển thị địa phương.

Trên FIG. 1, các đầu cuối (110-140) có thể được thể hiện dưới dạng các máy chủ, các máy tính cá nhân và điện thoại thông minh nhưng các nguyên lý của sáng chế không thể bị giới hạn ở đó. Các phương án của sáng chế được ứng dụng trong các máy tính laptop, máy tính bảng, bộ phát media và/hoặc thiết bị hội nghị video chuyên dụng. Mạng (150) thể hiện số lượng mạng bất kỳ mà chuyển dữ liệu video được mã hóa giữa các đầu cuối (110-140), bao gồm chẳng hạn các mạng truyền thông có dây dẫn và/hoặc không dây. Mạng truyền thông (150) có thể trao đổi dữ liệu trong các kênh chuyển mạch và/hoặc chuyển gói.

Các mạng tiêu biểu bao gồm các mạng viễn thông, các mạng vùng địa phương, các mạng vùng rộng và/hoặc mạng Internet. Đối với mục đích của phần mô tả này, kiến trúc và cấu hình của mạng (150) có thể không quan trọng đối với sự hoạt động của sáng chế trừ phi được giải thích dưới đây.

FIG.2 thể hiện, như ví dụ về ứng dụng đối với đối tượng được mô tả, việc bố trí bộ tạo mã và giải mã video trong môi trường phát trực tuyến. Đối tượng được bộc lộ có thể được áp dụng như nhau cho các ứng dụng cho phép video khác, bao gồm, chẳng hạn, hội nghị video, TV kỹ thuật số, lưu trữ video được nén trong phương tiện kỹ thuật số bao gồm CD, DVD, bộ nhớ lưu động và tương tự, và v.v..

Hệ thống phát trực tuyến có thể bao gồm hệ thống phụ thu (213), có thể bao gồm nguồn video (201), chẳng hạn camera kỹ thuật số, tạo ra chẳng hạn dòng mẫu video không được nén (202). Dòng mẫu (202) đó, được thể hiện bằng đường nét đậm để nhấn mạnh lượng dữ liệu lớn khi so với các dòng bit video được mã hóa, có thể được xử lý bởi bộ tạo mã (203) được ghép với camera (201). Bộ tạo mã (203) có thể bao gồm phần cứng, phần mềm, hoặc sự kết hợp của chúng để cho phép hoặc thực hiện các khía cạnh của đối tượng được bộc lộ như được mô tả chi tiết hơn dưới đây. Dòng bit video được mã hóa (204), được thể hiện bằng đường nét mảnh để nhấn mạnh lượng dữ liệu ít hơn khi so với dòng mẫu nêu trên, có thể được lưu trữ trong máy chủ phát trực tuyến (205) để dùng trong tương lai. Một hoặc nhiều máy khách phát trực tuyến (206, 208) có thể truy cập máy chủ phát trực tuyến (205) để lấy ra 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) để giải mã bản sao đầu vào của dòng bit video được mã hóa (207) và tạo dòng mẫu video đầu ra (211) mà có thể được thể hiện trên bộ hiển thị (212) hoặc bộ phận trình diễn khác (không được thể hiện). Trong một số hệ thống phát trực tuyến, các dòng bit video (204, 207, 209) có thể được tạo mã hóa theo một số chuẩn mã hóa/nén video. Ví dụ về các chuẩn đó bao gồm ITU-T Recommendation H.265. Chuẩn mã hóa video đang được phát triển được gọi không chính thức là mã hóa video đa năng (Versatile Video Coding) hoặc VVC. Đối tượng được bộc lộ có thể được sử dụng trong ngữ cảnh của VVC.

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

Bộ thu (310) có thể nhận một hoặc nhiều dãy mã hóa video-giải mã cần được giải mã bởi bộ giải mã (210); trong cùng một phương án hoặc phương án khác, một dãy video được mã hóa tại một thời điểm, trong đó việc giải mã mỗi dãy video được mã hóa là độc lập với các dãy video được mã hóa khác. Dãy video được mã hóa có thể được nhận từ kênh (312), mà có thể là liên kết phần cứng/phần mềm tới bộ phận lưu trữ mà lưu trữ dữ liệu video được mã hóa. Bộ thu (310) có thể nhận dữ liệu video được mã hóa với dữ liệu khác, chẳng hạn, dữ liệu âm thanh được mã hóa và/hoặc các dòng dữ liệu phụ trợ, mà có thể được chuyển tiếp đến các thực thể tương ứng sử dụng chúng (không được thể hiện). Bộ thu (310) có thể tách riêng dãy video được mã hóa khỏi dữ liệu khác. Để chống lại sự chập chờn của mạng, bộ nhớ đệm (315) có thể được ghép vào giữa bộ thu (310) và bộ giải mã entropi/bộ phân tách (320) (sau đây gọi là “bộ phân tách”). Khi bộ thu (310) nhận dữ liệu từ bộ phận lưu trữ/chuyển tiếp có đủ băng thông và khả năng điều khiển, hoặc từ mạng đẳng thời, thì bộ đệm (315) có thể không còn cần thiết, hoặc có thể nhỏ. Để sử dụng tốt nhất trong các mạng gói như mạng Internet, có thể cần bộ đệm (315), có thể tương đối lớn và có lợi khi có kích thước thích hợp.

Bộ giải mã video (210) có thể bao gồm bộ phân tách (320) để tái tạo các ký hiệu (321) từ dãy video được mã hóa entropi. Các dạng của các ký hiệu đó bao gồm thông tin được sử dụng để quản lý hoạt động của bộ giải mã (210), và có khả năng là thông tin để điều khiển bộ phận trình diễn như bộ hiển thị (212) mà không phải là một phần tích hợp của bộ giải mã nhưng có thể được ghép vào đó, như được thể hiện trên Fig. 2. Thông tin điều khiển dành cho (các) bộ phận trình diễn có thể dưới dạng thông tin tăng cường bổ sung (Supplementary Enhancement Information – thông điệp SEI) hoặc thông tin tính khả dụng video (Video Usability Information - VUI), các đoạn của bộ thông số (không được thể hiện). Bộ phân tách (320) có thể phân tách/giải mã entropi dãy video được mã hóa được nhận. Việc mã hóa dãy video được mã hóa có thể theo công nghệ hoặc chuẩn mã hóa video, và có thể theo các nguyên lý được biết rõ đối với người có

hiểu biết trung bình trong lĩnh vực, bao gồm mã hóa độ dài biến thiên, mã hóa Huffman, mã hóa số học có hoặc không có độ nhạy cảm ngữ cảnh, và v.v.. Bộ phân tách (320) có thể trích xuất từ dãy video được mã hóa, bộ thông số nhóm phụ dành cho ít nhất một trong các nhóm phụ của các điểm ảnh trong bộ giải mã video, dựa trên ít nhất một thông số tương ứng với nhóm này. Các nhóm phụ có thể bao gồm các nhóm ảnh (Group Of Picture - GOP), các ảnh, các ô, lát, các khối-macro, các bộ mã hóa (Coding Unit - CU), các khối, các bộ biến đổi (Transform Unit - TU), các bộ dự báo (Prediction Unit - PU) và v.v.. Bộ giải mã entropi/bộ phân tách cũng có thể trích xuất từ dãy video được mã hóa thông tin như các hệ số biến đổi, các giá trị thông số lượng tử hóa, các vectơ chuyển động, và v.v..

Bộ phân tách (320) có thể thực hiện các hoạt động giải mã entropi/phân tách trên dãy video được nhận từ bộ đệm (315), để tạo ra các ký hiệu (321).

Việc tái tạo các ký hiệu (321) có thể liên quan đến nhiều bộ phận khác nhau tùy theo loại của ảnh video được mã hóa hoặc các phần của nó (như: trong ảnh và ngoài ảnh, trong khối và ngoài khối), và các yếu tố khác. Việc bộ phận nào có liên quan, và liên quan như thế nào, có thể được điều khiển bởi thông tin điều khiển nhóm phụ mà đã được phân tách từ dãy video được mã hóa bởi bộ phân tách (320). Dòng thông tin điều khiển nhóm phụ này giữa bộ phân tách (320) và nhiều bộ phận không được thể hiện dưới đây để cho rõ ràng.

Ngoài các khối chức năng đã được nêu trên, bộ giải mã 210 có thể được phân chia về mặt khái niệm thành nhiều bộ phận chức năng như được mô tả dưới đây. Trong các phương án thực hiện thực tế hoạt động trong các giới hạn thương mại, nhiều bộ phận trong số này tương tác chặt chẽ với nhau và có thể, ít nhất phần nào, được tích hợp vào nhau. Tuy nhiên, nhằm mục đích mô tả đối tượng của sáng chế, việc phân chia theo khái niệm thành các bộ phận chức năng dưới đây là phù hợp.

Bộ phận thứ nhất là bộ đếm gộp/biến đổi ngược (351). Bộ đếm gộp/biến đổi ngược (351) nhận hệ số biến đổi được lượng tử hóa cũng như thông tin điều khiển, bao gồm sử dụng biến đổi nào, kích thước khối, hệ số lượng tử hóa, các ma trận đếm gộp lượng tử hóa, v.v. như (các) ký hiệu (321) từ bộ phân tách (320).

Bộ phận này có thể xuất ra các khối bao gồm các giá trị mẫu mà có thể được nhập vào bộ gộp (355).

Trong một số trường hợp, các mẫu được xuất ra của bộ đếm gộp/biến đổi ngược (351) có thể thuộc về khối được mã hóa trong; nghĩa là: khối mà không sử dụng thông tin dự báo từ các ảnh được tái tạo trước đây, nhưng có thể sử dụng thông tin dự báo từ các phần được tái tạo trước đây của ảnh hiện tại. Thông tin dự báo này có thể được cấp bởi bộ dự báo trong ảnh (352). Trong một số trường hợp, bộ dự báo trong ảnh (352) tạo khối có cùng kích thước và hình dạng như khối đang được tái tạo, bằng cách sử dụng thông tin đã được tái tạo xung quanh được tìm nạp từ ảnh hiện tại (được tái tạo một phần) (356). Trong một số trường hợp, bộ gộp (355) bổ sung, trên cơ sở từng mẫu, thông tin dự báo mà bộ dự báo trong (352) đã tạo ra, vào thông tin mẫu được xuất ra như được cấp bởi bộ đếm gộp/biến đổi ngược (351).

Trong những trường hợp khác, các mẫu được xuất ra của bộ đếm gộp/biến đổi ngược (351) có thể thuộc về khối được mã hóa ngoài, và có khả năng được bù chuyển động. Trong trường hợp này, bộ dự báo bù chuyển động (353) có thể truy cập bộ nhớ ảnh tham chiếu (357) để tìm nạp các mẫu được sử dụng để dự báo. Sau khi bù chuyển động các mẫu được tìm nạp theo các ký hiệu (321) thuộc về khối, các mẫu này có thể được bổ sung bởi bộ gộp (355) vào đầu ra của bộ đếm gộp/biến đổi ngược (trong trường hợp này được gọi là các mẫu dư hoặc tín hiệu dư) nhằm tạo ra thông tin mẫu được xuất ra. Các địa chỉ bên trong dạng bộ nhớ ảnh tham chiếu nơi mà bộ bù chuyển động tìm nạp các mẫu dự báo có thể được điều khiển bởi các vectơ chuyển động, khả dụng cho bộ bù chuyển động dưới dạng các ký hiệu (321) mà có thể có, chẳng hạn, các thành phần X, Y, và ảnh tham chiếu. Việc bù chuyển động cũng có thể bao gồm việc nội suy các giá trị mẫu như được tìm nạp từ bộ nhớ ảnh tham chiếu khi các vectơ chuyển động chính xác mẫu phụ được sử dụng, cơ chế dự báo vectơ chuyển động, và v.v..

Các mẫu được xuất ra của bộ gộp (355) có thể trải qua các kỹ thuật lọc vòng khác nhau trong bộ lọc vòng (356). Các kỹ thuật nén video có thể bao gồm các kỹ thuật trong bộ lọc vòng được điều khiển bởi các thông số được bao gồm trong dòng bit video được mã hóa và được làm khả dụng cho bộ lọc vòng (356)

dưới dạng các ký hiệu (321) từ bộ phân tách (320), nhưng cũng có thể đáp lại siêu thông tin thu được trong khi giải mã các phần trước (theo trình tự giải mã) của ảnh được mã hóa hoặc dãy video được mã hóa, cũng như đáp lại các giá trị mẫu được tái tạo trước đây và được lọc vòng.

Đầu ra của bộ lọc vòng (356) có thể là dòng mẫu mà có thể được xuất ra tới bộ phận trình diễn (212) cũng như được lưu trữ trong bộ nhớ ảnh tham chiếu (356) để sử dụng trong dự báo ngoài ảnh trong tương lai.

Một số ảnh được mã hóa, khi đã được tái tạo đầy đủ, có thể được sử dụng dưới dạng các ảnh tham chiếu để dự báo trong tương lai. Khi ảnh được mã hóa được tái tạo đầy đủ và ảnh được mã hóa đã được xác định là ảnh tham chiếu (chẳng hạn, bởi bộ phân tách (320)), thì ảnh tham chiếu hiện tại (356) có thể trở thành một phần của bộ đệm ảnh tham chiếu (357), và bộ nhớ ảnh hiện tại mới có thể được tái phân bổ trước khi bắt đầu tái tạo ảnh được mã hóa tiếp theo.

Bộ giải mã video 320 có thể thực hiện các hoạt động giải mã theo công nghệ nén video định trước mà có thể được quy định trong chuẩn, ví dụ ITU-T Rec. H.265. Dãy video được mã hóa có thể tuân theo cú pháp được định rõ bởi công nghệ hoặc chuẩn nén video được sử dụng, theo nghĩa là nó chấp nhận cú pháp của công nghệ hoặc chuẩn nén video, như được định rõ trong tài liệu hoặc chuẩn của công nghệ nén video đó và cụ thể là trong tài liệu lược tả trong đó. Để phù hợp thì cũng có thể là cần thiết việc độ phức tạp của dãy video được mã hóa phải nằm trong phạm vi được xác định bởi cấp độ của công nghệ hoặc chuẩn nén video. Trong một số trường hợp, các cấp độ giới hạn kích thước ảnh tối đa, tốc độ khung tối đa, tốc độ mẫu tái tạo tối đa (được đo bằng, chẳng hạn siêu mẫu trên giây), kích thước ảnh tham chiếu tối đa, và v.v.. Trong một số trường hợp, các giới hạn được thiết lập bởi các cấp độ có thể còn bị hạn chế tiếp thông qua các đặc tả bộ giải mã tham chiếu giả định (Hypothetical Reference Decoder - HRD) và siêu dữ liệu để quản lý bộ đệm HRD được báo tín hiệu trong dãy video được mã hóa.

Theo một phương án, bộ thu (310) có thể nhận dữ liệu bổ sung (dư thừa) với video được mã hóa. Dữ liệu bổ sung có thể được bao gồm như một phần của (các) dãy video được mã hóa. Dữ liệu bổ sung có thể được sử dụng bởi bộ giải

mã video (320) để giải mã đúng dữ liệu và/hoặc để tái tạo chính xác hơn dữ liệu video gốc. Dữ liệu bổ sung có thể dưới dạng, chẳng hạn, thời gian, không gian, hoặc các lớp tăng cường SNR, các lát dư, các ảnh dư, các mã sửa lỗi chuyên tiếp, và v.v..

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

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

Nguồn video (201) có thể cấp dãy video nguồn cần được mã hóa bởi bộ tạo mã (203) dưới dạng dòng mẫu video kỹ thuật số mà có thể có độ sâu bit thích hợp bất kỳ (chẳng hạn: 8 bit, 10 bit, 12 bit, ...), không gian màu bất kỳ (chẳng hạn, BT.601 Y CrCb, RGB, ...) và kết cấu lấy mẫu thích hợp bất kỳ (chẳng hạn Y CrCb 4:2:0, Y CrCb 4:4:4). Trong hệ thống dịch vụ truyền thông, nguồn video (201) có thể là bộ phận lưu trữ để lưu video được chuẩn bị trước đây. Trong hệ thống hội nghị video, nguồn video (203) có thể là camera để thu thông tin hình ảnh địa phương dưới dạng dãy video. Dữ liệu video có thể được cấp dưới dạng nhiều ảnh riêng lẻ mà truyền chuyển động khi được xem theo dãy. Bản thân các ảnh có thể được tổ chức dưới dạng mảng không gian của các điểm ảnh, trong đó mỗi điểm ảnh có thể bao gồm một hoặc nhiều mẫu tùy theo kết cấu lấy mẫu, không gian màu, v.v. được sử dụng. Người có hiểu biết trung bình trong lĩnh vực có thể dễ dàng hiểu được mối quan hệ giữa các điểm ảnh và các mẫu. Phần mô tả dưới đây tập trung vào các mẫu.

Theo một phương án, bộ tạo mã (203) có thể mã hóa và nén các ảnh của dãy video nguồn thành dãy video được mã hóa (443) theo thời gian thực hoặc theo bất kỳ sự ràng buộc về thời gian nào khác mà ứng dụng yêu cầu. Việc buộc theo tốc độ mã hóa thích hợp là một 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 ghép nối theo chức năng với các bộ phận này. Đê rõ ràng thì việc ghép nối không được thể hiện. Các thông số được thiết lập bởi bộ điều khiển có thể bao gồm các thông số liên quan điều khiển tốc độ (bỏ qua ảnh, bộ lượng tử hóa, giá trị lamda

của các kỹ thuật tối ưu độ méo, ...), kích thước ảnh, sơ đồ nhóm ảnh (GOP), phạm vi tìm kiếm vector chuyển động tối đa, và v.v.. Người có hiểu biết trung bình trong lĩnh vực có thể dễ dàng xác định các chức năng khác của bộ điều khiển (450) vì chúng có thể thuộc về bộ tạo mã video (203) được tối ưu hóa cho thiết kế hệ thống nhất định.

Một số bộ tạo mã video hoạt động trong cái mà người có hiểu biết trung bình trong lĩnh vực đã biết dưới dạng “vòng mã hóa”. Theo cách mô tả thô thiển thì vòng mã hóa có thể bao gồm phần tạo mã của bộ tạo mã (430) (sau đây gọi là “bộ mã hóa nguồn”) (chịu trách nhiệm tạo ra các ký hiệu dựa trên ảnh đầu vào cần được mã hóa, và (các) ảnh tham chiếu), và bộ giải mã (địa phương) (433) được nhúng trong bộ tạo mã (203) để tái tạo các ký hiệu nhằm tạo ra dữ liệu mẫu mà bộ giải mã (tù xa) cũng có thể tạo ra (vì phép nén bất kỳ giữa các ký hiệu và dòng bit video được mã hóa là không tổn hao trong các công nghệ nén video được xem xét trong đối tượng được bộc lộ). Dòng mẫu được tái tạo đó được nhập vào bộ nhớ ảnh tham chiếu (434). Do việc giải mã dòng ký hiệu dẫn đến các kết quả bit chính xác không phụ thuộc vào vị trí bộ giải mã (địa phương hoặc từ xa), nên nội dung bộ đệm ảnh tham chiếu cũng là bit chính xác giữa bộ tạo mã địa phương và bộ tạo mã từ xa. Nói cách khác, phần dự báo của bộ tạo mã “nhìn thấy” các mẫu ảnh tham chiếu chính xác giống như các giá trị mẫu mà bộ giải mã có thể “nhìn thấy” khi sử dụng dự báo trong khi giải mã. Nguyên lý nền tảng này của tính đồng bộ ảnh tham chiếu (và độ dịch tạo ra, nếu tính đồng bộ không được duy trì, chẳng hạn do lỗi kênh) được người có hiểu biết trung bình trong lĩnh vực biết đến rộng rãi.

Hoạt động của bộ giải mã “địa phương” (433) có thể giống như của bộ giải mã “tù xa” (210), mà đã được mô tả chi tiết trên đây cùng với Fig. 3. Tuy nhiên, vẫn như được thể hiện trên Fig 3, do các ký hiệu là khả dụng và việc tạo/giải mã các ký hiệu cho dãy video được mã hóa bằng bộ mã hóa entropi (445) và bộ phân tách (320) có thể là không tổn hao, nên các phần giải mã entropi của bộ giải mã (210), bao gồm kênh (312), bộ thu (310), bộ đệm (315), và bộ phân tách (320) có thể không được thực hiện đầy đủ trong bộ giải mã địa phương (433).

Tại đây có thể đưa ra nhận xét rằng công nghệ giải mã bất kỳ trừ việc phân tách/giải mã entropi mà đang có trong bộ giải mã cũng nhất thiết phải có mặt, dưới dạng chức năng gần như giống hệt, trong bộ tạo mã tương ứng. Vì lý do này, đối tượng được bọc lộ tập trung vào hoạt động của bộ giải mã. Phần mô tả các công nghệ của bộ tạo mã có thể được rút gọn vì chúng là nghịch đảo của các công nghệ bộ giải mã đã được mô tả đầy đủ. Chỉ có một số phần mô tả chi tiết hơn là cần thiết và được đề cập dưới đây.

Như một phần hoạt động của mình, bộ mã hóa nguồn (430) có thể thực hiện mã hóa dự báo bù chuyển động, mà mã hóa theo cách dự báo khung đầu vào có tham chiếu đến một hoặc nhiều khung được mã hóa trước đây từ dãy video mà được chỉ định là các “khung tham chiếu”. Theo cách này, máy mã hóa (432) mã hóa những khác biệt giữa các khối điểm ảnh của khung đầu vào và các khối điểm ảnh của (các) khung tham chiếu mà có thể được chọn làm (các) tham chiếu dự báo cho khung đầu vào.

Bộ giải mã video địa phương (433) có thể giải mã dữ liệu video được mã hóa của các khung mà có thể được chỉ định làm các khung tham chiếu, dựa trên các ký hiệu được tạo ra bởi bộ mã hóa nguồn (430). Các hoạt động của máy mã hóa (432) có thể có lợi khi là các quá trình có tổn hao. Khi dữ liệu video được mã hóa có thể được giải mã tại bộ giải mã video (không được thể hiện trên FIG. 4), dãy video được tái tạo thường có thể là bản sao của dãy video nguồn với một số lỗi. Bộ giải mã video địa phương (433) lặp lại các quá trình 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 tái tạo được lưu trữ trong bộ nhớ cache ảnh tham chiếu (434). Theo cách này, bộ tạo mã (203) có thể lưu trữ tại địa phương các bản sao của các khung tham chiếu được tái tạo mà có nội dung giống như các khung tham chiếu được tái tạo mà sẽ thu được bởi bộ giải mã video ở đầu xa (không có lỗi truyền).

Bộ dự báo (435) có thể thực hiện các tìm kiếm dự báo cho máy mã hóa (432). Nghĩa là, đối với khung mới cần được mã hóa, bộ dự báo (435) có thể tìm kiếm bộ nhớ ảnh tham chiếu (434) đối với dữ liệu mẫu (dưới dạng các khối điểm ảnh tham chiếu ứng viên) hoặc một số siêu dữ liệu như các vectơ chuyển động

ảnh tham chiếu, các hình dạng khối, và v.v., mà có thể làm tham chiếu dự báo thích hợp cho các ảnh mới. Bộ dự báo (435) có thể hoạt động trên cơ sở khối-trên-điểm ảnh mẫu để tìm các tham chiếu dự báo thích hợp. Trong một số trường hợp, như được xác định bởi kết quả tìm kiếm thu được bởi bộ dự báo (435), ảnh đầu vào có thể có các tham chiếu dự báo được rút ra 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, chẳng hạn, thiết lập các thông số và các thông số nhóm phụ được sử dụng để mã hóa dữ liệu video.

Đầu ra của tất cả các bộ phận chức năng nêu trên có thể trải qua mã hóa entropi trong bộ mã hóa entropi (445). Bộ mã hóa entropi biên dịch các ký hiệu như được tạo ra bởi nhiều bộ phận chức năng khác nhau thành dãy video được mã hóa, bằng cách nén không tổn hao các ký hiệu này theo các công nghệ mà người có hiểu biết trung bình trong lĩnh vực biết rõ, chẳng hạn như mã hóa Huffman, mã hóa độ dài biến thiên, mã hóa số học, và v.v..

Bộ phát (440) có thể đệm (các) dãy video được mã hóa như được tạo ra bởi bộ mã hóa entropi (445) để chuẩn bị chúng cho việc truyền qua kênh truyền thông (460), mà có thể là liên kết phần cứng/phần mềm tới bộ phận lưu trữ mà sẽ lưu trữ dữ liệu video được mã hóa. Bộ phát (440) có thể hợp nhất dữ liệu video được mã hóa từ bộ mã hóa video (430) với dữ liệu khác cần được truyền, chẳng hạn, dữ liệu âm thanh được mã hóa và/hoặc các dòng dữ liệu phụ trợ (các nguồn không được thể hiện).

Bộ điều khiển (450) có thể quản lý hoạt động của bộ tạo mã (203). Trong quá trình mã hóa, bộ điều khiển (450) có thể gán cho mỗi ảnh được mã hóa một kiểu ảnh được mã hóa nhất định, mà có thể ảnh hưởng đến các kỹ thuật mã hóa mà có thể được áp dụng cho ảnh tương ứng. Chẳng hạn, các ảnh thường có thể được gán một trong các kiểu khung sau:

Ảnh trong (ảnh I) có thể là ảnh mà có thể được mã hóa và được giải mã mà không sử dụng bất kỳ khung nào khác trong dãy làm nguồn dự báo. Một số bộ mã hóa-giải mã video cho phép các kiểu khác nhau của các ảnh trong, bao gồm, chẳng hạn, các ảnh làm mới bộ giải mã độc lập. Người có hiểu biết trung

bình trong lĩnh vực biết rõ về các biến thể của các ảnh I và các ứng dụng tương ứng và đặc điểm của chúng.

Ảnh dự báo (ảnh P) có thể là ảnh mà có thể được mã hóa và được giải mã bằng cách sử dụng dự báo trong hoặc dự báo ngoài mà sử dụng nhiều nhất là một vector chuyển động và chỉ số tham chiếu để dự báo các giá trị mẫu của mỗi khối.

Ảnh dự báo hai chiều (ảnh B) có thể là ảnh mà có thể được mã hóa và được giải mã bằng cách sử dụng dự báo trong hoặc dự báo ngoài mà sử dụng nhiều nhất là hai vector chuyển động và chỉ số tham chiếu để dự báo các giá trị mẫu của mỗi khối. Tương tự, các ảnh đa dự báo có thể sử dụng nhiều hơn hai ảnh tham chiếu và siêu dữ liệu liên kết để tái tạo một khối.

Các ảnh nguồn thường có thể được chia nhỏ theo không gian thành nhiều khối mẫu (chẳng hạn, các khối có 4x4, 8x8, 4x8, hoặc 16x16 mẫu mỗi khối) và được mã hóa trên cơ sở khối-trên- khối. Các khối có thể được mã hóa theo cách dự báo có tham chiếu các khối khác (đã được mã hóa) như được xác định bởi phép gán mã hóa được áp dụng cho các ảnh tương ứng trong khối. Chẳng hạn, các khối của các ảnh I có thể được mã hóa theo cách không dự báo hoặc có thể được mã hóa theo cách dự báo có tham chiếu tới các khối đã được mã hóa của cùng ảnh này (dự báo không gian hoặc dự báo trong). Các khối điểm ảnh của các ảnh P có thể được mã hóa theo cách không dự báo, qua dự báo không gian hoặc qua dự báo thời gian có tham chiếu tới một ảnh tham chiếu được mã hóa trước đây. Các khối của các ảnh B có thể được mã hóa theo cách không dự báo, qua dự báo không gian hoặc qua dự báo thời gian có tham chiếu tới một hoặc hai ảnh tham chiếu được mã hóa trước đây.

Bộ mã hóa video (203) có thể thực hiện các hoạt động mã hóa theo công nghệ hoặc chuẩn mã hóa video được định trước, chẳng hạn ITU-T Rec. H.265. Trong khi hoạt động, bộ mã hóa video (203) có thể thực hiện các hoạt động nén khác nhau, bao gồm các hoạt động mã hóa dự báo mà khai thác độ dư thời gian và không gian trong dãy video đầu vào. Do đó, dữ liệu video được mã hóa có thể tuân theo cú pháp được định rõ bởi công nghệ hoặc chuẩn mã hóa video được sử dụng.

Theo một phương án, bộ phát (440) có thể truyền dữ liệu bổ sung với

video được tạo mã. Bộ mã hóa video (430) có thể bao gồm dữ liệu này như một phần của dãy video được mã hóa. Dữ liệu bổ sung có thể bao gồm các lớp tăng cường thời gian/không gian/SNR, các dạng khác của dữ liệu dư như các ảnh dư và các lát dư, các thông điệp thông tin tăng cường bổ sung (SEI), thông tin khả dụng thị giác (VUI), các đoạn của bộ thông số, và v.v..

Trước khi mô tả một số khía cạnh của đối tượng của sáng chế một cách chi tiết hơn, một vài thuật ngữ cần được định nghĩa mà sẽ được viện dẫn trong phần còn lại của bản mô tả.

Sau đây, trong một số trường hợp, ảnh phụ dùng để gọi bố cục hình chữ nhật của các mẫu, các khối, các khối macro, các gói mã hóa, hoặc các phần tử tương tự mà được nhóm lại theo ngữ nghĩa, và có thể được mã hóa một cách độc lập trong độ phân giải được thay đổi. Một hoặc nhiều ảnh phụ có thể dành cho một ảnh. Một hoặc nhiều ảnh phụ được mã hóa có thể tạo nên ảnh được mã hóa. Một hoặc nhiều ảnh phụ có thể được kết hợp thành một ảnh, và một hoặc nhiều ảnh phụ có thể được trích ra từ một ảnh. Trong một số môi trường, một hoặc nhiều ảnh phụ được mã hóa có thể được kết hợp trong miền được nén mà không chuyển mã tới mức mẫu thành ảnh được mã hóa, và trong cùng trường hợp này hoặc một số trường hợp khác, một hoặc nhiều ảnh phụ được mã hóa có thể được trích ra từ ảnh được mã hóa trong miền được nén.

Sau đây, sự thay đổi độ phân giải (Adaptive Resolution Change - ARC) dùng để gọi cơ chế mà cho phép thay đổi độ phân giải của ảnh hoặc ảnh phụ trong dãy video được mã hóa, bằng cách, chẳng hạn, lấy mẫu lại ảnh tham chiếu. Các thông số ARC sau đây dùng để chỉ thông tin điều khiển cần thiết để thực hiện việc thay đổi độ phân giải thích ứng, mà có thể bao gồm, chẳng hạn, các thông số bộ lọc, các hệ số tỷ lệ, các độ phân giải của đầu ra và/hoặc các ảnh tham chiếu, các cờ điều khiển khác nhau, và v.v..

Phần mô tả trên đây tập trung vào việc mã hóa và giải mã một ảnh video được mã hóa độc lập về ngữ nghĩa. Trước khi mô tả hệ quả của việc mã hóa/giải mã nhiều ảnh phụ với các thông số ARC độc lập và mức độ phức tạp bổ sung kéo theo đó, thì các tùy chọn để báo tín hiệu các thông số ARC sẽ được mô tả.

FIG. 5 thể hiện một vài tùy chọn mới để báo tín hiệu các thông số ARC.

Như được lưu ý đối với mỗi tùy chọn trong số các tùy chọn này, chúng có một số ưu điểm và một số nhược điểm xét từ quan điểm hiệu quả mã hóa, mức độ phức tạp, và kiến trúc. Tiêu chuẩn hoặc công nghệ mã hóa video có thể chọn một hoặc nhiều tùy chọn trong số các tùy chọn này, hoặc các tùy chọn đã được biết đến trong tình trạng kỹ thuật, để báo tín hiệu các thông số ARC. Các tùy chọn có thể không loại trừ lẫn nhau, và có thể thấy được là có thể đổi lẫn cho nhau dựa trên các nhu cầu ứng dụng, công nghệ chuẩn liên quan, hoặc lựa chọn của bộ tạo mã.

Các loại thông số ARC có thể bao gồm:

các hệ số tăng/giảm mẫu, riêng rẽ hoặc được kết hợp trong kích thước X và Y.

các hệ số tăng/giảm mẫu, với việc bổ sung kích thước thời gian, để chỉ báo việc thu nhỏ/phóng to tốc độ không đổi đối với số lượng ảnh cho trước.

Cả hai loại nêu trên đều liên quan đến việc mã hóa của một hoặc nhiều các phần tử cú pháp được giả định là ngắn mà có thể chỉ đến bảng có chứa (các) hệ số này.

độ phân giải, trong kích thước X hoặc Y, trong các gói của các mẫu, các khối, các khối macro, các CU, hoặc độ chi tiết phù hợp bất kỳ nào khác, của ảnh đầu vào, ảnh đầu ra, ảnh tham chiếu, ảnh được mã hóa, được kết hợp hoặc riêng rẽ. Nếu có nhiều hơn một độ phân giải (như, chẳng hạn, một cho ảnh đầu vào, một cho ảnh tham chiếu) thì, trong một số trường hợp, một tập giá trị có thể suy ra từ tập giá trị kia. Điều này có thể bị giới hạn, chẳng hạn, bằng cách sử dụng các cờ. Để ví dụ chi tiết hơn xin xem dưới đây.

Các tọa độ “méo” giống như được sử dụng trong H.263 Annex P, lại vẫn ở mức độ chi tiết phù hợp như đã mô tả ở trên. H.263 Annex P xác định một cách hiệu quả để mã hóa các tọa độ méo này, nhưng các cách khác, có khả năng hiệu quả hơn, cũng có thể được đề xuất. Chẳng hạn, mã hóa kiểu “Huffman” khả nghịch chiều dài biến thiên các tọa độ méo của Annex P có thể được thay bằng mã hóa nhị phân chiều dài phù hợp, khi chiều dài của từ mã nhị phân có thể, chẳng hạn, được rút ra từ kích thước ảnh tối đa, có thể được nhân lên với hệ số nào đó và dịch đi giá trị nào đó, để cho phép “làm méo” bên ngoài của các đường

biên của kích thước ảnh tối đa.

Các thông số bộ lọc tăng hoặc giảm mẫu. Trong trường hợp dễ nhất, có thể chỉ có một bộ lọc để tăng và/hoặc giảm mẫu. Tuy nhiên, trong một số trường hợp, có thể có lợi khi cho phép linh hoạt hơn trong thiết kế bộ lọc, và điều đó có thể yêu cầu báo tín hiệu các thông số bộ lọc. Các thông số này có thể được chọn thông qua chỉ số trong danh sách các thiết kế bộ lọc có khả năng, bộ lọc có thể được chỉ rõ hoàn toàn (chẳng hạn thông qua danh sách các hệ số của bộ lọc, bằng cách sử dụng kỹ thuật mã hóa entropi phù hợp), bộ lọc có thể được chọn theo cách ngầm định thông qua các tỷ số tăng/giảm mẫu mà theo đó, đến lượt mình, lại được báo tín hiệu theo cơ chế bất kỳ trong số các cơ chế nêu trên, và v.v..

Sau đây, phần mô tả giả sử rằng việc mã hóa tập hữu hạn các hệ số tăng/giảm mẫu (chính hệ số này được sử dụng trong cả hai kích thước X và Y), được chỉ báo thông qua từ mã. Theo cách có lợi, từ mã này có thể có chiều dài biến thiên được mã hóa, chẳng hạn bằng cách sử dụng phép mã hóa Ext-Golomb phổ biến đối với một số phần tử cú pháp trong các đặc tả mã hóa video như H.264 và H.265.

Nhiều ánh xạ tương tự có thể được đề xuất theo nhu cầu của ứng dụng và khả năng của các cơ chế tăng và giảm tỷ lệ khả dụng trong công nghệ hoặc tiêu chuẩn nén video. Bảng này có thể được mở rộng tới nhiều giá trị hơn. Các giá trị cũng có thể được biểu diễn bởi các cơ chế mã hóa entropi khác với mã hóa Ext-Golomb, chẳng hạn bằng cách sử dụng mã hóa nhị phân. Điều này có thể có một số ưu điểm khi các hệ số lấy mẫu lại đã được quan tâm bên ngoài bản thân các máy xử lý video (bộ tạo mã và bộ giải mã trước tiên), chẳng hạn bởi MANE. Cần lưu ý rằng, đối với trường hợp (được giả định) phổ biến nhất không yêu cầu thay đổi độ phân giải, thì việc mã hóa Ext-Golomb có thể được chọn mà là ngắn; trong bảng trên, chỉ là một bit. Điều này có lợi thế về hiệu quả mã hóa so với việc sử dụng mã hóa nhị phân đối với trường hợp phổ biến nhất.

Số lượng các mục trong bảng, cũng như ngữ nghĩa của chúng có thể được cấu hình hoàn toàn hoặc riêng phần. Chẳng hạn, tóm tắt cơ bản của bảng có thể truyền tải trong tập thông số “cao” như dây hoặc tập thông số bộ giải mã. Theo cách thay thế hoặc thêm vào đó, một hoặc nhiều bảng này có thể được xác định

trong công nghệ hoặc tiêu chuẩn mã hóa video, và có thể được chọn thông qua chẳng hạn bộ giải mã hoặc tập thông số dãy.

Sau đây, là phân mô tả cách thức hệ số tăng mẫu/giảm mẫu (thông tin ARC), được mã hóa như đã mô tả trên đây, có thể được bao gồm cú pháp công nghệ hoặc tiêu chuẩn mã hóa video. Việc xem xét tương tự có thể áp dụng cho một, hoặc một vài, từ mã điều khiển các bộ lọc tăng/giảm mẫu. Xin xem dưới đây phân trình bày khi những lượng dữ liệu tương đối lớn được yêu cầu cho bộ lọc hoặc các cấu trúc dữ liệu khác.

H.263 Annex P bao gồm thông tin ARC 502 dưới dạng bốn tọa độ méo vào tiêu đề ảnh 501, cụ thể trong phần mở rộng tiêu đề H.263 PLUSPTYPE (503). Điều này có thể là lựa chọn thiết kế nhạy cảm khi a) có tiêu đề ảnh khả dụng, và b) các thay đổi thường xuyên của thông tin ARC được kỳ vọng. Tuy nhiên, thủ tục bổ sung khi sử dụng báo tín hiệu theo kiểu H.263 có thể khá cao, và các hệ số tỷ lệ có thể không nằm giữa các đường biên của ảnh do tiêu đề ảnh có bản chất tạm thời.

JVCET-M135-v1, như đã nêu trên, bao gồm thông tin tham chiếu ARC (505) (chỉ số) nằm trong tập thông số ảnh (504), đánh chỉ số bảng (506) bao gồm các độ phân giải đích mà đến lượt mình được bố trí bên trong tập thông số dãy (507). Việc bố trí độ phân giải tiềm năng trong bảng (506) trong tập thông số dãy (507) có thể là hợp lý, theo lời tuyên bố của các tác giả, nhờ sử dụng SPS dưới dạng điểm thỏa thuận tương kết trong khi trao đổi khả năng. Độ phân giải có thể thay đổi từ ảnh tới ảnh, bên trong các giới hạn được thiết lập bởi các giá trị trong bảng (506) bằng cách viện dẫn tập thông số ảnh (504) thích hợp.

FIG. 5 còn thể hiện các tùy chọn bổ sung sau đây có thể tồn tại để chuyển thông tin ARC trong dòng bit video. Mỗi trong số các tùy chọn đó có một số ưu điểm so với giải pháp đã biết như đã được mô tả trên đây. Các tùy chọn có thể đồng thời có mặt trong cùng một công nghệ hoặc tiêu chuẩn mã hóa video.

Theo một phương án của sáng chế, thông tin ARC (509) như hệ số lấy mẫu lại (zoom) có thể có trong tiêu đề lát, tiêu đề GOB, tiêu đề ô, hoặc tiêu đề nhóm ô (sau đây gọi là tiêu đề nhóm ô) (508). Điều này có thể là thỏa đáng với thông tin ARC nhỏ, như một ue(v) chiều dài biến thiên hoặc từ mã chiều dài cố

định bằng vài bit, ví dụ như thể hiện ở trên. Việc có thông tin ARC trong tiêu đề nhóm ô trực tiếp có ưu điểm bổ sung của thông tin ARC có thể áp dụng cho ảnh phụ được biểu diễn bởi, chẳng hạn, nhóm ô, chứ không phải toàn ảnh. Xin xem dưới đây. Thêm vào đó, ngay cả nếu công nghệ hoặc tiêu chuẩn nén video chỉ xem xét thay đổi độ phân giải thích ứng toàn ảnh (trái ngược với, chẳng hạn, các thay đổi độ phân giải thích ứng dựa trên nhóm ô), thì việc đưa thông tin ARC vào tiêu đề nhóm ô thay vì đưa vào tiêu đề ảnh kiểu H.263 có một số ưu điểm xét trên quan điểm chống chịu lỗi.

Theo cùng phương án của sáng chế hoặc theo phương án khác, bản thân thông tin ARC (512) có thể có trong tập thông số (511) thích hợp như, chẳng hạn, tập thông số ảnh, tập thông số tiêu đề, tập thông số ô, tập thông số thích ứng, và v.v. (tập thông số thích ứng được thể hiện). Phạm vi của tập thông số này có thể có ưu điểm là không lớn hơn ảnh, chẳng hạn nhóm ô. Việc sử dụng thông tin ARC là ngầm định thông qua việc kích hoạt tập thông số liên quan. Chẳng hạn, khi công nghệ hoặc tiêu chuẩn mã hóa video chỉ dự kiến ARC dựa trên ảnh, thì tập thông số ảnh hoặc tương đương có thể thích hợp.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, thông tin tham chiếu ARC (513) có thể có trong tiêu đề nhóm ô (514) hoặc cấu trúc dữ liệu tương tự. Thông tin tham chiếu (513) này có thể viện dẫn đến tập phụ của thông tin ARC (515) khả dụng trong tập thông số (516) với phạm vi vượt quá một ảnh, chẳng hạn tập thông số dãy, hoặc tập thông số bộ giải mã.

Mức bổ sung của sự kích hoạt được ngầm định gián tiếp của PPS từ tiêu đề nhóm ô, PPS, SPS, như được sử dụng trong JVET-M0135-v1 có vẻ không cần thiết, vì các tập thông số ảnh, như là các tập thông số dãy, có thể (và có trong một số chuẩn như RFC3984) được sử dụng để thỏa thuận khả năng hoặc cho các thông báo. Tuy nhiên, nếu thông tin ARC cần được áp dụng cho ảnh phụ được biểu diễn bởi, chẳng hạn, cũng các nhóm ô, thì tập thông số với phạm vi kích hoạt được giới hạn ở nhóm ô, như tập thông số thích ứng hoặc tập thông số tiêu đề có thể là lựa chọn tốt hơn. Cũng vậy, nếu thông tin ARC có kích thước lớn hơn mức có thể bỏ qua, ví dụ chứa thông tin điều khiển bộ lọc, chẳng hạn nhiều hệ số bộ lọc, thì thông số có thể là lựa chọn tốt hơn việc sử dụng tiêu đề (508),

trực tiếp từ quan điểm hiệu quả mã hóa, do các cài đặt đó có thể tái sử dụng bởi các ảnh tương lai hoặc các ảnh phụ bằng cách tham chiếu chính tập thông số đó.

Khi sử dụng tập thông số dãy hoặc tập thông số cao hơn khác với phạm vi trải rộng trên nhiều ảnh, thì một số sự cân nhắc có thể được áp dụng:

Tập thông số lưu giữ bảng thông tin ARC (516) có thể, trong một số trường hợp, là tập thông số dãy, nhưng trong các trường hợp khác, có lợi là tập thông số bộ giải mã. Tập thông số bộ giải mã có thể có phạm vi kích hoạt của nhiều CVS, chính là dòng video được mã hóa, nghĩa là tất cả các bit video được mã hóa từ bắt đầu phiên đến khi ngừng phiên. Phạm vi này có thể thích hợp hơn bởi vì các hệ số ARC tiềm năng có thể là đặc tính của bộ giải mã, có thể được thực hiện trong phần cứng, và các đặc tính phần cứng có xu hướng không thay đổi với CVS bất kỳ (mà ít nhất trong một số hệ thống giải trí là các nhóm ảnh, có độ dài một giây hoặc ít hơn). Nghĩa là, việc đưa bảng vào tập thông số dãy được bao gồm một cách rõ ràng trong các tùy chọn bố trí được mô tả ở đây.

Thông tin tham chiếu ARC (513) có thể được đặt trực tiếp một cách có lợi vào ảnh/ô lát/GOB/tiêu đề nhóm ô (sau đây gọi là tiêu đề nhóm ô) (514) thay vì đặt vào tập thông số ảnh như trong JVCET-M0135-v1, Lý do như sau: khi bộ tạo mã muốn thay đổi một giá trị trong tập thông số ảnh, chẳng hạn như thông tin tham chiếu ARC, thì nó phải tạo PPS mới và tham chiếu PPS mới này. Giả sử rằng chỉ thông tin tham chiếu ARC thay đổi, nhưng thông tin khác như, chẳng hạn, thông tin ma trận lượng tử hóa trong PPS vẫn còn nguyên. Thông tin này có thể là có kích thước đáng kể, và cần được phát lại để hoàn chỉnh PPS mới. Do thông tin tham chiếu ARC có thể là một từ mã, như chỉ số vào trong bảng (513) và có thể là giá trị duy nhất thay đổi, nên sẽ nặng nề và lãng phí khi phát lại tất cả, chẳng hạn, thông tin ma trận lượng tử hóa. Trong một chừng mực, có thể là tốt hơn đáng kể xét từ quan điểm hiệu quả mã hóa để tránh việc gián tiếp thông qua PPS, như được đề xuất trong JVET-M0135-v1. Tương tự, việc đưa thông tin tham chiếu ARC vào PPS có nhược điểm bổ sung là thông tin ARC được tham chiếu bởi thông tin tham chiếu ARC (513) nhất thiết cần được áp dụng cho toàn ảnh và không cho ảnh phụ, do phạm vi kích hoạt tập thông số ảnh là ảnh.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, việc

báo tín hiệu các thông số ARC có thể theo ví dụ chi tiết như được tóm tắt trên FIG. 6. FIG. 6 thể hiện các sơ đồ cú pháp được trình bày như được sử dụng trong các chuẩn mã hóa video ít nhất từ năm 1993. Ký pháp của các sơ đồ cú pháp này đại thể theo lập trình C. Các dòng in đậm chỉ báo các phần tử cú pháp có trong dòng bit, các dòng không in đậm thường chỉ báo luồng điều khiển hoặc cài đặt của các biến số.

Tiêu đề nhóm ô (601), dưới dạng cấu trúc cú pháp được nêu làm ví dụ về tiêu đề có thể áp dụng cho phần (có thể là chữ nhật) của ảnh, có thể chứa một cách có điều kiện phần tử cú pháp chiều dài biến thiên, được mã hóa Exp-Golomb `dec_pic_size_idx` (602) (được thể hiện bằng in đậm). Sự có mặt của phần tử cú pháp này trong tiêu đề nhóm ô có thể bị hạn chế ở việc sử dụng độ phân giải thích ứng (603) ở đây, giá trị của cờ không được thể hiện bằng in đậm, nghĩa là cờ có trong dòng bit ở điểm nơi nó xuất hiện trong sơ đồ cú pháp. Việc liệu độ phân giải thích ứng có được sử dụng cho ảnh này hoặc các phần của nó hay không có thể được báo tín hiệu trong cấu trúc cú pháp mức cao bất kỳ bên trong hoặc bên ngoài dòng bit. Trong ví dụ được thể hiện, điều này được báo tín hiệu trong tập thông số dãy như được tóm lược dưới đây.

FIG. 6 cũng thể hiện đoạn trích của tập thông số dãy (610). Phần tử cú pháp thứ nhất được thể hiện là `adaptive_pic_resolution_change_flag` (611). Nếu đúng, cờ này có thể chỉ báo việc sử dụng độ phân giải thích ứng mà, đến lượt mình, có thể yêu cầu một số thông tin điều khiển. Trong ví dụ này, thông tin điều khiển này được trình bày theo cách có điều kiện dựa trên giá trị của cờ dựa trên mệnh đề `if()` trong tập thông số (612) và tiêu đề nhóm ô (601).

Khi độ phân giải thích ứng được sử dụng, trong ví dụ này, thì độ phân giải đầu ra trong các gói của các mẫu (613) được mã hóa. Số 613 được dùng để chỉ cả `output_pic_width_in_luma_samples` lẫn `output_pic_height_in_luma_samples`, mà cùng nhau có thể xác định độ phân giải của ảnh đầu ra. Một số phần khác trong công nghệ hoặc tiêu chuẩn mã hóa video, một số hạn chế đối với cả hai giá trị có thể được xác định. Chẳng hạn, định nghĩa mức có thể giới hạn số lượng tổng các mẫu đầu ra, mà có thể là tích của giá trị của hai phần tử cú pháp đó. Một số công nghệ hoặc chuẩn mã hóa video, hoặc

công nghệ hoặc chuẩn ngoài, chẳng hạn như các chuẩn hệ thống, có thể giới hạn khoảng đánh số (chẳng hạn, một hoặc cả hai kích thước phải chia hết cho lũy thừa của 2 số), hoặc tỷ số cạnh (chẳng hạn, chiều rộng và chiều cao phải theo tỷ lệ 4:3 hoặc 16:9). Những hạn chế này có thể được đưa vào để tạo thuận lợi cho việc thực hiện phần cứng hoặc vì các lý do khác, và được biết đến rộng rãi trong lĩnh vực này.

Trong một số ứng dụng, có thể có lợi là bộ tạo mã ra lệnh cho bộ giải mã sử dụng một kích thước ảnh tham chiếu nhất định thay vì ngầm định coi rằng kích thước là kích thước ảnh đầu ra. Trong ví dụ này, phần tử cú pháp `reference_pic_size_present_flag` (614) chặn sự có mặt theo điều kiện của các kích thước ảnh tham chiếu (615) (số này vẫn chỉ cả chiều rộng lẫn chiều cao).

Cuối cùng, bảng chiều rộng và chiều cao giải mã ảnh tiềm năng được thể hiện. Bảng này có thể được biểu thị, chẳng hạn, bằng chỉ báo bảng (`num_dec_pic_size_in_luma_samples_minus1`) (616). Ký hiệu “minus1” có thể chỉ sự dịch chuyển của giá trị của phần tử cú pháp. Chẳng hạn, nếu giá trị được mã hóa bằng không, thì một mục của bảng entry được trình bày. Nếu giá trị này là năm, thì sáu mục bảng được trình bày. Đối với mỗi “dòng” trong bảng, thì chiều rộng và chiều cao ảnh được giải mã được bao gồm trong cú pháp (617).

Các mục bảng được trình bày (617) có thể được đánh 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 thước được giải mã khác nhau có hiệu lực, các hệ số phóng to theo nhóm ô.

Các kỹ thuật dành cho việc báo tín hiệu các thông số độ phân giải thích ứng đã được mô tả ở trên có thể được thực hiện dưới dạng phần mềm máy tính bằng cách sử dụng các lệnh máy tính đọc được và được lưu trữ một cách vật lý trong một hoặc nhiều vật ghi máy tính đọc được. Chẳng hạn, FIG. 7 thể hiện hệ thống máy tính 700 phù hợp để thực hiện một số phương án của đối tượng được bộc lộ.

Phần mềm máy tính có thể được mã hóa bằng cách sử dụng bất kỳ mã máy hoặc ngôn ngữ máy tính thích hợp, mà có thể được hợp thể, biên dịch, liên kết, hoặc các thao tác tương tự để tạo ra mã bao gồm các lệnh có thể được thực

thi trực tiếp, hoặc thông qua phiên dịch, thực thi mã micro, và tương tự, bởi các bộ xử lý trung tâm của máy tính (Central Processing Unit - CPU), bộ xử lý đồ họa (Graphics Processing Unit - GPU), và tương tự.

Các lệnh này có thể được thực thi trên nhiều loại máy tính hoặc các thành phần của chúng, bao gồm, chẳng hạn, các máy tính cá nhân, các máy tính bảng, các máy chủ, điện thoại thông minh, máy chơi game, thiết bị internet vạn vật, và tương tự.

Các thành phần được thể hiện trên FIG. 7 cho hệ thống máy tính 700 thực chất là ví dụ và không nhằm đưa ra bất kỳ hạn chế nào đối với phạm vi sử dụng hoặc chức năng của phần mềm máy tính thực hiện các phương án của sáng chế. Cấu hình của các thành phần này cũng không được diễn giải là có sự phụ thuộc hoặc yêu cầu bất kỳ liên quan đến bất kỳ thành phần hoặc sự kết hợp các thành phần được thể hiện trong phương án ví dụ của hệ thống máy tính 700.

Hệ thống máy tính 700 có thể bao gồm một số bộ phận đầu vào giao diện người dùng. Bộ phận đầu vào giao diện người dùng này có thể đáp lại đầu vào bởi một hoặc nhiều người dùng thông qua, chẳng hạn, đầu vào xúc giác (như: nhấn phím, vuốt, các chuyển động gắng đỡ liệu), đầu vào âm thanh (như: giọng nói, vỗ tay), đầu vào thị giác (như: ra hiệu), đầu vào khứu giác (không được thể hiện). Các bộ phận giao diện người dùng này cũng có thể được sử dụng để thu một số phiên truyền thông mà không nhất thiết phải liên quan trực tiếp đến việc nhập dữ liệu có ý thức bởi người dùng, chẳng hạn âm thanh (như: bài nói, nhạc, tiếng động của môi trường), các hình ảnh (như: các hình ảnh được quét, các hình ảnh chụp thu được từ camera ảnh tĩnh), video (như video hai chiều, video ba chiều bao gồm video nổi).

Các bộ phận giao diện người dùng đầu vào có thể bao gồm một hoặc nhiều (chỉ thể hiện một trong mỗi loại): bàn phím 701, chuột 702, bàn di cảm ứng 703, màn hình cảm ứng 710, gắng đỡ liệu 704, cần điều khiển 705, micrô 706, máy quét 707, camera 708.

Hệ thống máy tính 700 có thể còn bao gồm một số bộ phận đầu ra giao diện người dùng. Các bộ phận đầu ra giao diện người dùng này có thể kích thích các giác quan của một hoặc nhiều người dùng thông qua, chẳng hạn, đầu ra xúc

giác, tiếng động, ánh sáng, và mùi/vị. Các bộ phận đầu ra giao diện người dùng này có thể bao gồm các bộ phận đầu ra xúc giác (chẳng hạn phản hồi xúc giác bởi màn hình cảm ứng 710, găng đỡ liệu 704, hoặc cần điều khiển 705, nhưng cũng có thể có các bộ phận phản hồi xúc giác mà không phải là các bộ phận đầu vào), các bộ phận đầu ra âm thanh (như: các loa 709, các tai nghe (không được thể hiện)), các bộ phận đầu ra thị giác (như các màn hình 710 bao gồm màn hình CRT, màn hình LCD, màn hình plasma, màn hình OLED, mỗi loại có hoặc không có khả năng nhập liệu qua màn hình cảm ứng, mỗi loại có hoặc không có khả năng phản hồi xúc giác – một số trong đó có thể có khả năng xuất tín hiệu đầu ra thị giác hai chiều hoặc hơn thế tín hiệu đầu ra ba chiều thông qua phương tiện như đầu ra stereo; kính thực tế ảo (không được thể hiện), các màn hiển thị toàn ký và các thùng khói (không được thể hiện)), và các máy in (không được thể hiện).

Hệ thống máy tính 700 còn có thể bao gồm các bộ phận lưu trữ người dùng truy cập được và các phương tiện liên quan của chúng như phương tiện quang bao gồm CD/DVD ROM/RW 720 với CD/DVD hoặc phương tiện tương tự 721, ổ USB 722, ổ cứng di động hoặc ổ lưu trữ thể rắn 723, các phương tiện lưu trữ từ tính cổ điển như băng từ và đĩa mềm (không được thể hiện), các bộ phận dựa trên ROM/ASIC/PLD chuyên dụng như các khóa an ninh điện tử (không được thể hiện), và tương tự.

Người có hiểu biết trung bình trong lĩnh vực cần hiểu rằng thuật ngữ “vật ghi máy tính đọc được” như được sử dụng liên quan đến đối tượng của sáng chế không bao gồm phương tiện truyền, vật mang sóng, hoặc các tín hiệu chuyển tiếp khác.

Hệ thống máy tính 700 cũng có thể bao gồm giao diện tới một hoặc nhiều mạng truyền thông. Các mạng có thể là, chẳng hạn, không dây, dây dẫn, sợi quang. Các mạng còn có thể là mạng địa phương, mạng diện rộng, mạng đại đô thị, mạng của phương tiện giao thông và mạng công nghiệp, mạng thời gian thực, mạng chấp nhận độ trễ, và v.v.. Các ví dụ về mạng bao gồm các mạng vùng địa phương như mạng Ethernet, mạng LAN vô tuyến, các mạng tế bào bao gồm GSM, 3G, 4G, 5G, LTE và tương tự, mạng TV dây dẫn hoặc các mạng kỹ thuật

số vùng rộng vô tuyến bao gồm TV cáp, TV vệ tinh, và TV phát mặt đất, các mạng phương tiện giao thông và công nghiệp bao gồm CANBus, và v.v.. Một số mạng đều yêu cầu các bộ điều hợp giao diện mạng ngoài mà được gắn vào một số cổng dữ liệu thông dụng hoặc các buýt ngoại vi (749) (như, chẳng hạn các cổng USB của hệ thống máy tính 700; các cổng khác thường được tích hợp vào lõi của hệ thống máy tính 700 bằng cách gắn vào buýt hệ thống như được mô tả dưới đây (chẳng hạn giao diện Ethernet được tích hợp vào hệ thống máy tính PC hoặc giao diện mạng tế bào vào hệ thống máy tính của điện thoại thông minh). Nhờ sử dụng mạng bất kỳ trong số này, hệ thống máy tính 700 có thể truyền thông với các thực thể khác. Việc truyền thông này có thể là đơn hướng, chỉ nhận (chẳng hạn, phát sóng TV), đơn hướng chỉ gửi (chẳng hạn CANbus tới một số thiết bị CANbus), hoặc hai chiều, chẳng hạn các hệ thống máy tính khác bằng cách sử dụng các mạng địa phương hoặc mạng kỹ thuật số vùng rộng. Một số giao thức và ngăn xếp giao thức có thể được sử dụng cho mỗi một trong số các mạng và giao diện mạng này như đã mô tả trên đây.

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

Lõi 740 có thể bao gồm một hoặc nhiều bộ xử lý trung tâm (CPU) 741, bộ xử lý đồ họa (GPU) 742, bộ xử lý lập trình chuyên dụng dưới dạng các mảng cổng lập trình được bằng trường (Field Programmable Gate Area - FPGA) 743, các mạch phân cứng cho một số tác vụ 744, và v.v.. Các bộ phận này, cùng với bộ nhớ chỉ đọc (Read-only Memory - ROM) 745, bộ nhớ truy cập ngẫu nhiên (Random Access Memory) 746, bộ nhớ trong dung lượng cao như ổ cứng bộ nhớ trong người dùng không truy cập được, SSD, và tương tự 747, có thể được kết nối thông qua buýt hệ thống 748. Trong một số hệ thống máy tính, buýt hệ thống 748 có thể truy cập được dưới dạng một hoặc nhiều bộ nối vật lý để cho phép mở rộng bằng cách bổ sung các CPU, GPU, và tương tự. Các thiết bị ngoại vi có thể được gắn trực tiếp vào buýt hệ thống 748 của lõi, hoặc qua buýt ngoại vi 749. Cấu trúc của buýt ngoại vi bao gồm PCI, USB, và tương tự.

Các CPU 841, GPU 842, FPGA 743, và các mạch tăng tốc 744 có thể thực

thi một số lệnh, khi được kết hợp, có thể bổ sung mã máy tính nêu trên. Mã máy tính này có thể được lưu trữ trong ROM 745 hoặc RAM 746. Dữ liệu chuyên tiếp cũng có thể được lưu trữ trong RAM 746, trong khi dữ liệu cố định có thể được lưu trữ, chẳng hạn, trong bộ nhớ trong dung lượng cao 747. Việc lưu trữ và truy tìm nhanh tới phần bất kỳ của các bộ nhớ có thể được cho phép thông qua việc sử dụng bộ nhớ cache, mà có thể được liên kết chặt chẽ với một hoặc nhiều CPU 741, GPU 742, bộ nhớ dung lượng cao 747, ROM 745, RAM 746, và tương tự.

Vật ghi máy tính đọc được có thể có mã máy tính được ghi trên đó để thực hiện nhiều hoạt động được thực hiện bằng máy tính khác nhau. Vật ghi và mã máy tính có thể là loại được thiết kế và chế tạo đặc biệt cho mục đích của sáng chế, hoặc có thể là loại được biết rộng rãi và khả dụng đối với những người có hiểu biết trung bình trong lĩnh vực phần mềm máy tính.

Như là ví dụ và không nhằm mục đích giới hạn, hệ thống máy tính có cấu trúc 700, và nhất là lõi 740 có thể có chức năng như là kết quả của (các) bộ xử lý (bao gồm các CPU, GPU, FPGA, các mạch tăng tốc, và tương tự) thực thi phần mềm được lưu trữ trong một hoặc nhiều vật ghi máy tính đọc được hữu hình. Vật ghi máy tính đọc được này có thể là vật ghi gắn liền với bộ nhớ dung lượng lớn người dùng truy cập được như đã nêu trên, cũng như một số bộ nhớ của lõi 740 mà có bản chất là không khả biến, như bộ nhớ trong dung lượng cao lõi 747 hoặc ROM 745. Phần mềm thực hiện các phương án khác nhau của sáng chế có thể được lưu trữ trong các bộ phận này và được thực thi bởi lõi 740. Vật ghi máy tính đọc được có thể bao gồm một hoặc nhiều thiết bị bộ nhớ hoặc chip, tùy theo nhu cầu cụ thể. Phần mềm này có thể làm cho lõi 740 và cụ thể là các bộ xử lý trong đó (bao gồm CPU, GPU, FPGA, và tương tự) thực thi các quá trình cụ thể hoặc các phần cụ thể các quá trình cụ thể được mô tả ở đây, bao gồm việc xác định các kết cấu dữ liệu được lưu trữ trong RAM 746 và cải biến các kết cấu dữ liệu này theo các quá trình được xác định bởi phần mềm. Như một sự bổ sung hoặc thay thế, hệ thống máy tính có thể có chức năng như là kết quả của mạch cứng logic hoặc được thực hiện theo cách khác trong mạch (chẳng hạn: mạch tăng tốc 744), mà có thể hoạt động thay cho hoặc cùng với phần mềm để

thực thi các quá trình cụ thể hoặc các phần cụ thể của các quá trình cụ thể được mô tả ở đây. Việc viện dẫn đến phần mềm có thể bao gồm mạch logic, và ngược lại, khi thích hợp. Việc viện dẫn đến vật ghi máy tính đọc được có thể bao gồm cả mạch (như mạch tích hợp (IC)) lưu trữ phần mềm để thực thi, mạch thực hiện logic để thực thi, hoặc cả hai, khi thích hợp. Phần mô tả này bao gồm sự kết hợp thích hợp bất kỳ của phần cứng và phần mềm.

Một số công nghệ hoặc chuẩn mã hóa video, chẳng hạn VP9, hỗ trợ khả năng định tỷ lệ không gian bằng cách thực hiện một số dạng lấy mẫu lại ảnh tham chiếu (được báo tín hiệu theo cách khá là khác nhau từ đối tượng của sáng chế) cùng với khả năng định tỷ lệ thời gian, để cho phép khả năng định tỷ lệ không gian. Cụ thể là, một số ảnh tham chiếu có thể được tăng mẫu bằng cách sử dụng các công nghệ kiểu ARC tới độ phân giải cao hơn để tạo nền của lớp tăng cường không gian. Các ảnh được tăng mẫu này có thể được làm mịn, bằng cách sử dụng các cơ chế dự báo thông thường ở độ phân giải cao, nhằm bổ sung chi tiết.

Đối tượng của sáng chế có thể được sử dụng trong môi trường như vậy. Trong một số trường hợp, theo cùng một phương án hoặc theo phương án khác của sáng chế, giá trị trong tiêu đề gói NAL, chẳng hạn trường ID thời gian, có thể được sử dụng để chỉ báo không chỉ lớp thời gian mà cả lớp không gian. Cách làm này có một số ưu điểm đối với một số thiết kế hệ thống; chẳng hạn, các gói chuyển tiếp được chọn (Selected Forwarding Unit - SFU) hiện có được tạo ra và được tối ưu hóa cho việc chuyển tiếp được chọn lớp thời gian dựa trên tiêu đề gói NAL. Giá trị ID thời gian có thể được sử dụng không cần cải biến, đối với các môi trường có thể mở rộng. Để có thể làm được điều này, có thể có yêu cầu đối với việc ánh xạ giữa kích thước ảnh được mã hóa và lớp thời gian được chỉ báo bởi trường ID thời gian trong tiêu đề gói NAL.

Trong một số công nghệ mã hóa video, bộ truy cập (Access Unit - AU) có thể dùng để chỉ (các) ảnh được mã hóa, (các) lát, (các) ô, (các) gói NAL, và v.v., mà đã được thu và được kết hợp vào dòng bit ảnh/lát/ô/gói NAL tương ứng ở thời điểm cho trước. Thời điểm đó có thể là thời gian tổng hợp.

Trong HEVC, và một số công nghệ mã hóa video khác, giá trị đếm thứ tự ảnh (picture order count - POC) có thể được sử dụng để chỉ báo ảnh tham chiếu

được chọn trong số nhiều ảnh tham chiếu được lưu giữ trong bộ nhớ đệm ảnh được giải mã (decoded picture buffer - DPB). Khi bộ truy cập (AU) bao gồm một hoặc nhiều ảnh, lát, hoặc ô, thì mỗi ảnh, lát, hoặc ô thuộc về cùng AU có thể mang cùng giá trị POC, mà từ đó có thể là suy ra là chúng đã được tạo ra từ nội dung của cùng một thời gian tổng hợp. Nói cách khác, trong kịch bản khi hai ảnh/lát/ô mang cùng giá trị POC cho trước, mà có thể chỉ báo về hai ảnh/lát/ô thuộc về cùng AU và có cùng thời gian tổng hợp. Ngược lại, hai ảnh/ô/lát có các giá trị POC khác nhau có thể chỉ báo các ảnh/lát/ô đó thuộc về các AU khác nhau và có các thời gian tổng hợp khác nhau.

Theo một phương án của đối tượng của sáng chế, mỗi quan hệ cứng nhắc nêu trên có thể được nói lỏng ở chỗ bộ truy cập có thể bao gồm các ảnh, lát, hoặc ô với các giá trị POC khác nhau. Nhờ cho phép các giá trị POC khác nhau trong AU mà có thể sử dụng giá trị POC này để có thể nhận dạng các ảnh/lát/ô theo tiềm năng có thể giải mã độc lập với thời gian trình diễn giống nhau. Đến lượt mình, điều này có thể cho phép hỗ trợ nhiều lớp có thể mở rộng mà không cần thay đổi báo tín hiệu chọn ảnh tham chiếu (ví dụ, báo tín hiệu tập ảnh tham chiếu hoặc báo tín hiệu danh sách ảnh tham chiếu), như được mô tả chi tiết hơn dưới đây.

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

Theo cùng phương án hoặc theo phương án khác của sáng chế, giá trị đếm bộ truy cập (access unit count - AUC) có thể được báo tín hiệu trong cấu trúc cú pháp mức cao, như tiêu đề gói NAL, tiêu đề lát, tiêu đề nhóm ô, thông điệp SEI, tập thông số hoặc bộ phân cách AU. Giá trị của AUC có thể được sử dụng để nhận dạng các gói NAL, ảnh, lát, hoặc các ô nào thuộc về AU cho trước. Giá trị của AUC có thể tương ứng với thời điểm thời gian tổng hợp khác biệt. Giá trị AUC có thể bằng nhiều giá trị POC. Bằng cách chia giá trị POC cho giá trị số nguyên thì có thể tính được giá trị AUC. Trong một số trường hợp, các thao tác chia có thể tạo ra gánh nặng nhất định cho việc thực hiện bộ giải mã. Trong các trường hợp như vậy, các hạn chế nhỏ trong không gian đánh số của các giá trị

AUC có thể cho phép thay thế thao tác chia bằng thao tác dịch. Chẳng hạn, giá trị AUC có thể bằng giá trị bit có giá trị cao nhất (Most Significant Bit - MSB) của khoảng giá trị POC.

Theo cùng phương án của sáng chế, giá trị của chu kỳ POC trên AU (`poc_cycle_au`) có thể được báo tín hiệu trong cấu trúc cú pháp mức cao, như tiêu đề gói NAL, tiêu đề lát, tiêu đề nhóm ô, thông điệp SEI, tập thông số hoặc bộ phân cách AU. Giá trị `poc_cycle_au` có thể chỉ báo cách thức mà nhiều giá trị POC khác nhau và liên tiếp có thể được liên kết với cùng AU. Chẳng hạn, nếu giá trị của `poc_cycle_au` bằng 4, thì các ảnh, các lát hoặc các ô với giá trị POC bằng 0 – 3, kể cả 2 giá trị biên, được liên kết với AU có giá trị AUC bằng 0, và các ảnh, các lát hoặc các ô có giá trị POC bằng 4 – 7, kể cả 2 giá trị biên, được liên kết với AU có giá trị AUC bằng 1. Do đó, giá trị của AUC có thể suy ra bằng cách chia giá trị POC cho giá trị của `poc_cycle_au`.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, giá trị của `poc_cyle_au` có thể được rút ra từ thông tin được đặt, chẳng hạn, trong tập thông số video (VPS), mà chỉ báo số lượng các lớp không gian hoặc SNR trong dãy video được mã hóa. Mỗi quan hệ tiềm năng này được mô tả ngắn gọn dưới đây. Trong khi việc rút ra này như đã được mô tả trên đây có thể tiết kiệm vài bit trong VPS và do đó có thể nâng cao hiệu quả mã hóa, có thể có lợi việc mã hóa rõ ràng `poc_cycle_au` trong cấu trúc cú pháp mức cao thích hợp mà theo phân cấp ở dưới tập thông số video, nên có thể giảm thiểu `poc_cycle_au` đối với phần nhỏ cho trước của dòng bit, chẳng hạn ảnh. Việc tối ưu hóa này có thể tiết kiệm nhiều bit hơn là có thể tiết kiệm thông qua quá trình rút ra nêu trên vì các giá trị POC (và/hoặc các giá trị của các phần tử cú pháp gián tiếp tham chiếu đến POC) có thể được mã hóa trong các cấu trúc cú pháp mức thấp.

FIG. 8 thể hiện ví dụ về cấu trúc dãy video với sự kết hợp của `temporal_id`, `layer_id`, POC và các giá trị AUC với thay đổi độ phân giải thích ứng. Trong ví dụ này, ảnh, lát hoặc ô trong AU thứ nhất với $AUC = 0$ có thể có `temporal_id = 0` và `layer_id = 0` hoặc 1, trong khi ảnh, lát hoặc ô trong AU thứ hai với $AUC = 1$ có thể lần lượt có `temporal_id = 1` và `layer_id = 0` hoặc 1. Giá trị của POC được tăng lên thêm 1 trên ảnh không phụ thuộc vào các giá trị của `temporal_id` và

layer_id. Trong ví dụ này, giá trị của poc_cycle_au có thể bằng 2. Tốt hơn là, giá trị của poc_cycle_au có thể được thiết lập bằng số lượng (khả năng định tỷ lệ không gian) các lớp. Do đó, trong ví dụ này giá trị của POC được tăng lên thêm 2, trong khi giá trị của AUC được tăng lên thêm 1.

Theo các phương án nêu trên của sáng chế, tất cả hoặc tập con của cấu trúc dự báo ngoài ảnh hoặc ngoài lớp và chỉ báo ảnh tham chiếu có thể được hỗ trợ nhờ sử dụng báo tín hiệu tập ảnh tham chiếu (RPS) hiện có trong HEVC hoặc báo tín hiệu danh sách ảnh tham chiếu (RPL). Trong RPS hoặc RPL, ảnh tham chiếu được chọn được chỉ báo bằng cách báo tín hiệu giá trị của POC hoặc giá trị delta của POC giữa ảnh hiện tại và ảnh tham chiếu được chọn. Đối với đối tượng của sáng chế, RPS và RPL có thể được sử dụng để chỉ báo cấu trúc dự báo ngoài ảnh hoặc ngoài lớp mà không thay đổi việc báo tín hiệu, nhưng với các hạn chế sau. Nếu giá trị của temporal_id của ảnh tham chiếu lớn hơn giá trị của temporal_id ảnh hiện tại, thì ảnh hiện tại có thể không sử dụng ảnh tham chiếu để bù chuyển động hoặc các dự báo khác. Nếu giá trị của layer_id của ảnh tham chiếu lớn hơn giá trị của layer_id ảnh hiện tại, thì ảnh hiện tại có thể không sử dụng ảnh tham chiếu để bù chuyển động hoặc các dự báo khác.

Theo cùng phương án này và các phương án khác của sáng chế, việc định tỷ lệ vector chuyển động dựa trên sự khác biệt POC đối với dự báo vector chuyển động thời gian có thể được vô hiệu qua nhiều ảnh bên trong bộ truy cập. Do đó, mặc dầu mỗi ảnh có thể có giá trị POC khác nhau trong bộ truy cập, vector chuyển động không được định tỷ lệ và được sử dụng để dự báo vector chuyển động thời gian trong bộ truy cập. Đó là vì ảnh tham chiếu với POC khác nhau trong cùng AU được coi là ảnh tham chiếu có cùng thời điểm. Do đó, theo phương án của sáng chế, chức năng định tỷ lệ vector chuyển động có thể trở lại 1, khi ảnh tham chiếu thuộc về AU được liên kết với ảnh hiện tại.

Theo cùng phương án này và các phương án khác của sáng chế, việc định tỷ lệ vector chuyển động dựa trên sự khác biệt POC để dự báo vector chuyển động thời gian có thể theo tùy chọn được vô hiệu qua nhiều ảnh, khi độ phân giải không gian của ảnh tham chiếu khác với độ phân giải không gian của ảnh hiện tại. Khi việc định tỷ lệ vector chuyển động được cho phép, thì vector chuyển động

được định tỷ lệ dựa trên cả sự khác biệt POC lẫn tỷ số độ phân giải không gian giữa ảnh hiện tại và ảnh tham chiếu.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, vector chuyển động có thể được định tỷ lệ dựa trên sự khác biệt AUC thay vì sự khác biệt POC, để dự báo vector chuyển động thời gian, nhất là khi poc_cycle_au có giá trị không đồng nhất (khi vps_contant_poc_cycle_per_au == 0). Trái lại (khi vps_contant_poc_cycle_per_au == 1), thì việc định tỷ lệ vector chuyển động dựa trên sự khác biệt AUC có thể giống như định tỷ lệ vector chuyển động dựa trên sự khác biệt POC.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, khi vector chuyển động được định tỷ lệ dựa trên sự khác biệt AUC, thì vector chuyển động tham chiếu trong cùng AU (với cùng giá trị AUC) với ảnh hiện tại không được định tỷ lệ dựa trên sự khác biệt AUC và được sử dụng để dự báo vector chuyển động mà không định tỷ lệ hoặc có định tỷ lệ dựa trên tỷ số độ phân giải không gian giữa ảnh hiện tại và ảnh tham chiếu.

Theo cùng phương án này và các phương án khác của sáng chế, giá trị AUC được sử dụng để nhận dạng biên của AU và được sử dụng cho hoạt động bộ giải mã tham chiếu (HRD) giả định, mà cần định thời đầu vào và đầu ra với độ chi tiết AU. Trong đa số các trường hợp, ảnh được giải mã với lớp cao nhất trong AU có thể được xuất ra để hiển thị. Giá trị AUC và giá trị layer_id có thể được sử dụng để nhận dạng ảnh đầu ra.

Theo một phương án của sáng chế, ảnh có thể bao gồm một hoặc nhiều ảnh phụ. Mỗi ảnh phụ có thể bao phủ vùng địa phương hoặc toàn bộ vùng của ảnh. Vùng được hỗ trợ bởi ảnh phụ có thể có hoặc có thể không chồng lấn với vùng được hỗ trợ bởi ảnh phụ khác. Vùng được tạo thành bởi một hoặc nhiều ảnh phụ có thể có hoặc có thể không bao phủ toàn bộ vùng của ảnh. Nếu ảnh chứa ảnh phụ, thì vùng được hỗ trợ bởi ảnh phụ giống như vùng được hỗ trợ bởi ảnh.

Theo cùng phương án của sáng chế, ảnh phụ có thể được mã hóa bởi phương pháp mã hóa tương tự như phương pháp mã hóa được sử dụng cho ảnh được mã hóa. Ảnh phụ có thể được mã hóa một cách độc lập hoặc có thể được

mã hóa phụ thuộc vào ảnh phụ khác hoặc ảnh được mã hóa. Ảnh phụ có thể có hoặc có thể không có sự phụ thuộc phân tách bất kỳ vào ảnh phụ khác hoặc ảnh được mã hóa.

Theo cùng phương án của sáng chế, ảnh phụ được mã hóa có thể được chứa trong một hoặc nhiều lớp. Ảnh phụ được mã hóa trong lớp có thể có độ phân giải không gian khác nhau. Ảnh phụ gốc có thể được lấy mẫu lại theo không gian (được tăng mẫu hoặc được giảm mẫu), được mã hóa với các thông số độ phân giải không gian khác nhau, và được chứa trong dòng bit tương ứng với lớp.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, ảnh phụ với (W, H) , trong đó lần lượt W chỉ báo chiều rộng của ảnh phụ và H chỉ báo chiều cao của ảnh phụ, có thể được mã hóa và được chứa trong dòng bit được mã hóa tương ứng với lớp 0, trong khi ảnh phụ được tăng mẫu (hoặc được giảm mẫu) từ ảnh phụ với độ phân giải không gian gốc, với $(W^*S_{w,k}, H^*S_{h,k})$, có thể được mã hóa và được chứa trong dòng bit được mã hóa tương ứng với lớp k , trong đó $S_{w,k}, S_{h,k}$ chỉ báo các tỷ số lấy mẫu lại, theo chiều ngang và theo chiều dọc. Nếu các giá trị của $S_{w,k}, S_{h,k}$ lớn hơn 1, thì việc lấy mẫu lại bằng tăng mẫu. Còn ngược lại, nếu các giá trị của $S_{w,k}, S_{h,k}$ nhỏ hơn 1, thì việc lấy mẫu lại bằng giảm mẫu.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, ảnh phụ được mã hóa trong lớp có thể có chất lượng hình ảnh khác với chất lượng của ảnh phụ được mã hóa trong lớp khác trong cùng ảnh phụ hoặc ảnh phụ khác. Chẳng hạn, ảnh phụ i trong lớp n , được mã hóa với thông số lượng tử hóa $Q_{i,n}$, trong khi ảnh phụ j trong lớp m , được mã hóa với thông số lượng tử hóa, $Q_{j,m}$.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, ảnh phụ được mã hóa trong lớp có thể có thể được giải mã độc lập, mà không có sự phụ thuộc phân tách hoặc giải mã bất kỳ vào ảnh phụ được mã hóa trong lớp khác của cùng vùng địa phương. Lớp ảnh phụ, mà có thể được giải mã độc lập mà không tham chiếu lớp ảnh phụ khác của cùng vùng địa phương, là lớp ảnh phụ độc lập. Ảnh phụ được mã hóa trong lớp ảnh phụ độc lập có thể có hoặc có thể không có sự phụ thuộc giải mã hoặc phân tách vào ảnh phụ được mã hóa trước đó trong cùng lớp ảnh phụ, nhưng ảnh phụ được mã hóa có thể không có

sự phụ thuộc bất kỳ vào ảnh được mã hóa trong lớp ảnh phụ khác.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, ảnh phụ được mã hóa trong lớp có thể có thể được giải mã một cách phụ thuộc, với sự phụ thuộc phân tách hoặc giải mã bất kỳ vào ảnh phụ được mã hóa trong lớp khác của cùng vùng địa phương. Lớp ảnh phụ, mà có thể được giải mã một cách phụ thuộc với việc tham chiếu lớp ảnh phụ khác của cùng vùng địa phương, là lớp ảnh phụ phụ thuộc. Ảnh phụ được mã hóa trong ảnh phụ phụ thuộc có thể tham chiếu ảnh phụ được mã hóa thuộc về cùng ảnh phụ, ảnh phụ được mã hóa trước đó trong cùng lớp ảnh phụ, hoặc cả hai ảnh phụ tham chiếu.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, ảnh phụ được mã hóa bao gồm một hoặc nhiều lớp ảnh phụ độc lập và một hoặc nhiều lớp ảnh phụ phụ thuộc. Tuy nhiên, ít nhất một lớp ảnh phụ độc lập có thể được trình bày cho ảnh phụ được mã hóa. Lớp ảnh phụ độc lập có thể có giá trị của bộ nhận dạng lớp (*layer_id*), mà có thể có trong tiêu đề gói NAL hoặc cấu trúc cú pháp mức cao khác, bằng 0. Lớp ảnh phụ với *layer_id* bằng 0 là lớp ảnh phụ nền.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, ảnh có thể bao gồm một hoặc nhiều ảnh phụ tiền cảnh và một ảnh phụ hậu cảnh. Vùng được hỗ trợ bởi ảnh phụ hậu cảnh có thể bằng vùng của ảnh. Vùng được hỗ trợ bởi ảnh phụ tiền cảnh có thể chồng lấn với vùng được hỗ trợ bởi ảnh phụ hậu cảnh. Ảnh phụ hậu cảnh có thể là lớp ảnh phụ nền, trong khi ảnh phụ tiền cảnh có thể lớp ảnh phụ không phải nền (tăng cường). Một hoặc nhiều lớp ảnh phụ không phải nền có thể tham chiếu cùng lớp nền để giải mã. Mỗi lớp ảnh phụ không phải nền với *layer_id* bằng *a* có thể tham chiếu lớp ảnh phụ không phải nền với *layer_id* bằng *b*, trong đó *a* lớn hơn *b*.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, ảnh có thể bao gồm một hoặc nhiều ảnh phụ tiền cảnh với hoặc không có ảnh phụ hậu cảnh. Mỗi ảnh phụ có thể có lớp ảnh phụ nền của riêng nó và một hoặc nhiều lớp không phải nền (tăng cường). Mỗi lớp ảnh phụ nền có thể được tham chiếu bởi một hoặc nhiều lớp ảnh phụ không phải nền. Mỗi lớp ảnh phụ không phải nền với *layer_id* bằng *a* có thể tham chiếu lớp ảnh phụ không phải nền với

layer_id bằng b , trong đó a lớn hơn b .

Theo cùng một phương án hoặc theo phương án khác của sáng chế, ảnh có thể bao gồm một hoặc nhiều ảnh phụ tiền cảnh với hoặc không có ảnh phụ hậu cảnh. Mỗi ảnh phụ được mã hóa trong lớp ảnh phụ (nền hoặc không phải nền) có thể được tham chiếu bởi một hoặc nhiều ảnh phụ lớp không phải nền thuộc về cùng ảnh phụ và một hoặc nhiều ảnh phụ lớp không phải nền, mà không thuộc về cùng ảnh phụ.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, ảnh có thể bao gồm một hoặc nhiều ảnh phụ tiền cảnh với hoặc không có ảnh phụ hậu cảnh. Ảnh phụ trong lớp a có thể được phân vùng tiếp thành nhiều ảnh phụ trong cùng lớp. Một hoặc nhiều ảnh phụ được mã hóa trong lớp b có thể tham chiếu ảnh phụ được phân vùng trong lớp a .

Theo cùng một phương án hoặc theo phương án khác của sáng chế, dãy video được mã hóa (CVS) có thể là nhóm ảnh được mã hóa. Dãy CVS có thể bao gồm một hoặc nhiều dãy ảnh phụ được mã hóa (CSPS), trong đó CSPS có thể là nhóm ảnh phụ được mã hóa bao phủ cùng vùng địa phương của ảnh. Dãy CSPS có thể có cùng độ phân giải thời gian hoặc độ phân giải thời gian khác với độ phân giải thời gian của dãy video được mã hóa.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, CSPS có thể được mã hóa và được chứa trong một hoặc nhiều lớp. Dãy CSPS có thể bao gồm một hoặc nhiều lớp CSPS. Việc giải mã một hoặc nhiều lớp CSPS tương ứng với dãy CSPS có thể tái tạo dãy của các ảnh phụ tương ứng với cùng vùng địa phương.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, số lượng các lớp CSPS tương ứng với CSPS có thể giống như hoặc khác với số lượng các lớp CSPS tương ứng với CSPS khác.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, lớp CSPS có thể có độ phân giải thời gian (ví dụ tốc độ khung) khác với lớp CSPS khác. Dãy ảnh phụ gốc (không được nén) có thể được lấy mẫu lại theo thời gian (được tăng mẫu hoặc được giảm mẫu), được mã hóa với các thông số độ phân giải thời gian khác nhau, và được chứa trong dòng bit tương ứng với lớp.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, dãy ảnh phụ với tốc độ khung F , có thể được mã hóa và được chứa trong dòng bit được mã hóa tương ứng với lớp 0, trong khi dãy ảnh phụ được tăng mẫu (hoặc được giảm mẫu) theo thời gian từ dãy ảnh phụ gốc, với $F^* S_{t,k}$, có thể được mã hóa và được chứa trong dòng bit được mã hóa tương ứng với lớp k , trong đó $S_{t,k}$ chỉ báo tỷ số lấy mẫu theo thời gian cho lớp k . Nếu giá trị của $S_{t,k}$ lớn hơn 1, thì quá trình lấy mẫu lại theo thời gian bằng với biến đổi tăng tốc độ khung. Ngược lại, nếu giá trị của $S_{t,k}$ nhỏ hơn 1, thì quá trình lấy mẫu lại theo thời gian bằng với biến đổi giảm tốc độ khung.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, khi ảnh phụ với lớp CSPS a được tham chiếu bởi ảnh phụ với lớp CSPS b để bù chuyển động hoặc dự báo ngoài lớp bất kỳ, nếu độ phân giải không gian của lớp CSPS a khác với độ phân giải không gian của lớp CSPS b , thì các pixel được giải mã trong lớp CSPS a được lấy mẫu lại và được sử dụng để tham chiếu. Quá trình lấy mẫu lại có thể cần đến lọc tăng mẫu hoặc lọc giảm mẫu.

FIG. 9 thể hiện ví dụ về các bảng cú pháp để báo tín hiệu phần tử cú pháp của `vps_poc_cycle_au` trong VPS (hoặc SPS), mà chỉ báo `poc_cycle_au` được sử dụng cho tất cả các ảnh/các lát trong dãy video được mã hóa, và phần tử cú pháp của `slice_poc_cycle_au`, mà chỉ báo `poc_cycle_au` của lát hiện tại, trong tiêu đề lát. Nếu giá trị POC tăng lên một cách đồng đều theo AU, thì `vps_contant_poc_cycle_per_au` trong VPS được đặt bằng 1 và `vps_poc_cycle_au` được báo tín hiệu trong VPS. Trong trường hợp này, `slice_poc_cycle_au` không được báo tín hiệu một cách rõ ràng, và giá trị của AUC đối với mỗi AU được tính toán bằng cách chia giá trị của POC cho `vps_poc_cycle_au`. Nếu giá trị POC không tăng lên một cách đồng đều theo AU, thì `vps_contant_poc_cycle_per_au` trong VPS được đặt bằng 0. Trong trường hợp này, `vps_access_unit_cnt` không được báo tín hiệu, trong khi `slice_access_unit_cnt` được báo tín hiệu trong tiêu đề lát đối với mỗi lát hoặc ảnh. Mỗi lát hoặc ảnh có thể có giá trị khác nhau của `slice_access_unit_cnt`. Giá trị của AUC đối với mỗi AU được tính toán bằng cách chia giá trị của POC cho `slice_poc_cycle_au`. FIG. 10 thể hiện sơ đồ khối minh họa lưu đồ công tác tương

ứng.

Theo cùng phương án hoặc theo các phương án khác của sáng chế, ngay cả nếu giá trị của POC của ảnh, lát, hoặc ô có thể khác nhau, thì ảnh, lát, hoặc ô tương ứng với AU với cùng giá trị AUC có thể được liên kết với cùng thời điểm giải mã hoặc xuất ra. Do đó, không cần sự phụ thuộc phân tách ngoài/giải mã bất kỳ qua các ảnh, các lát hoặc các ô trong cùng AU, tất cả hoặc tập phụ của các ảnh, các lát hoặc các ô được liên kết với cùng AU có thể được giải mã song song, và có thể được xuất ra ở cùng một thời điểm.

Theo cùng phương án hoặc theo các phương án khác của sáng chế, ngay cả nếu giá trị của POC của ảnh, lát, hoặc ô có thể khác nhau, thì ảnh, lát, hoặc ô tương ứng với AU với cùng giá trị AUC có thể được liên kết với cùng thời điểm tổng hợp/hiển thị. Khi thời gian tổng hợp được chứa trong định dạng vật chứa, ngay cả nếu các ảnh tương ứng với các AU khác nhau, thì nếu các ảnh có cùng thời gian tổng hợp, các ảnh này có thể là được hiển thị ở cùng một thời điểm.

Theo cùng phương án hoặc theo các phương án khác của sáng chế, mỗi ảnh, lát, hoặc ô có thể có cùng bộ nhận dạng thời gian (`temporal_id`) trong cùng AU. Tất cả hoặc tập phụ của các ảnh, các lát hoặc các ô tương ứng với thời điểm có thể được liên kết với cùng lớp phụ thời gian. Theo cùng phương án hoặc theo các phương án khác của sáng chế, mỗi ảnh, lát, hoặc ô có thể có lớp không gian (`layer_id`) giống nhau hoặc khác nhau trong cùng AU. Tất cả hoặc tập phụ của các ảnh, các lát hoặc các ô tương ứng với thời điểm có thể được liên kết với lớp không gian giống nhau hoặc khác nhau. .

FIG. 11 thể hiện dòng video ví dụ bao gồm CSPS video hậu cảnh với `layer_id` bằng 0 và nhiều lớp CSPS tiền cảnh. Trong khi ảnh phụ được mã hóa có thể bao gồm một hoặc nhiều lớp CSPS, vùng hậu cảnh, mà không thuộc về lớp CSPS tiền cảnh nào, có thể bao gồm lớp nền. Lớp nền có thể chứa vùng hậu cảnh và các vùng tiền cảnh, trong khi lớp CSPS tăng cường chứa vùng tiền cảnh. Lớp CSPS tăng cường có thể có chất lượng xem hình tốt hơn lớp nền ở cùng vùng. Lớp CSPS tăng cường có thể tham chiếu các pixel được tái tạo và các vector chuyển động của lớp nền, tương ứng với cùng vùng.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, dòng

bit video tương ứng với lớp nền được chứa trong một đường ghi, trong khi các lớp CSPS tương ứng với mỗi ảnh phụ được chứa trong đường ghi được tách riêng, trong tệp video.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, dòng bit video tương ứng với lớp nền được chứa trong đường ghi, trong khi các lớp CSPS với cùng layer_id được chứa trong đường ghi được tách riêng. Trong ví dụ này, đường ghi tương ứng với lớp k chỉ bao gồm các lớp CSPS tương ứng với lớp k .

Theo cùng một phương án hoặc theo phương án khác của sáng chế, mỗi lớp CSPS của mỗi ảnh phụ được lưu giữ trong đường ghi riêng rẽ. Mỗi đường ghi có thể hoặc có thể không có sự phụ thuộc phân tách hoặc giải mã bất kỳ vào một hoặc nhiều đường ghi khác.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, mỗi đường ghi có thể chứa các dòng bit tương ứng với lớp i tới lớp j của các lớp CSPS của tất cả hoặc tập phụ của các ảnh phụ, trong đó $0 < i \leq j \leq k$, k là lớp cao nhất của CSPS.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, ảnh bao gồm một hoặc nhiều dữ liệu media được liên kết bao gồm mức độ sâu, mức alpha, dữ liệu hình học 3D, bản đồ cư trú, v.v.. Dữ liệu media được liên kết có thể được chia cho một hoặc nhiều dòng phụ dữ liệu mà mỗi dòng tương ứng với một ảnh phụ.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, FIG. 12 thể hiện ví dụ về hội nghị video dựa trên phương pháp ảnh phụ nhiều lớp. Trong dòng video có chứa một dòng bit video lớp nền tương ứng với ảnh hậu cảnh và một hoặc nhiều dòng bit video lớp tăng cường tương ứng với các ảnh phụ tiền cảnh. Mỗi dòng bit video lớp tăng cường tương ứng với lớp CSPS. Trên bộ hiển thị, ảnh tương ứng với lớp nền được hiển thị theo mặc định. Nó chứa một hoặc nhiều ảnh của người dùng trong ảnh (PIP). Khi người dùng cụ thể được chọn bởi bộ điều khiển khách, lớp CSPS tăng cường tương ứng với người dùng được chọn được giải mã và được hiển thị với chất lượng hoặc độ phân giải không gian được nâng cao. FIG. 13 thể hiện sơ đồ của hoạt động này.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, hộp trung gian mạng (như bộ định tuyến) có thể lựa chọn tập phụ của các lớp để gửi tới người dùng tùy thuộc vào băng thông của nó. Việc tối chức ảnh/ảnh phụ có thể được sử dụng để làm thích ứng băng thông. Ví dụ, nếu người dùng không có băng thông, thì bộ định tuyến tước bỏ các lớp hoặc lựa chọn một số ảnh phụ theo sự quan trọng của chúng hoặc dựa trên thiết lập được sử dụng và điều này có thể được thực hiện theo cách động để chấp nhận cho băng thông.

FIG. 14 thể hiện trường hợp sử dụng của video 360. Khi ảnh 360 hình cầu được chiếu lên ảnh phẳng, thì ảnh 360 được chiếu có thể được phân vùng thành nhiều ảnh phụ làm lớp nền. Lớp tăng cường của ảnh phụ cụ thể có thể được mã hóa và được truyền tới máy khách. Bộ giải mã có thể giải mã được cả lớp nền bao gồm tất cả các ảnh phụ lẫn lớp tăng cường của ảnh phụ được chọn. Khi công nhìn hiện tại giống như ảnh phụ được chọn, thì ảnh được hiển thị có thể có chất lượng cao hơn với ảnh phụ được giải mã với lớp tăng cường. Ngược lại, ảnh được giải mã với lớp nền có thể được hiển thị, với chất lượng thấp.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, thông tin bố trí bất kỳ để hiển thị có thể có trong tệp tin, dưới dạng thông tin bổ sung (như thông điệp SEI hoặc siêu dữ liệu). Một hoặc nhiều ảnh phụ được giải mã có thể được tái bố trí và được hiển thị tùy theo thông tin bố trí được báo tín hiệu. Thông tin bố trí có thể được báo tín hiệu bởi máy chủ phát trực tuyến hoặc bộ phát quảng bá, hoặc có thể được phục hồi lại bởi thực thể mạng hoặc máy chủ đám mây, hoặc có thể được xác định bởi cài đặt được tùy biến của người dùng.

Theo một phương án của sáng chế, khi ảnh đầu vào được chia thành một hoặc nhiều vùng phụ (chữ nhật), thì mỗi vùng phụ có thể được mã hóa dưới dạng lớp độc lập. Mỗi lớp độc lập tương ứng với vùng địa phương có thể là giá trị `layer_id` duy nhất. Đối với mỗi lớp độc lập, kích thước ảnh phụ và thông tin vị trí có thể được báo tín hiệu. Chẳng hạn, kích thước ảnh (chiều rộng, chiều cao), thông tin dịch của trên góc bên trái (`x_offset`, `y_offset`). FIG. 15 thể hiện ví dụ về bố trí của các ảnh phụ được chia, kích thước ảnh phụ và thông tin định vị và cấu trúc dự báo ảnh tương ứng của nó. Thông tin bố trí bao gồm (các) kích thước ảnh phụ và (các) định vị ảnh phụ có thể được báo tín hiệu trong cấu trúc cú pháp

mức cao, như (các) tập thông số, tiêu đề của lát hoặc nhóm ô, hoặc thông điệp SEI.

Theo cùng phương án của sáng chế, mỗi ảnh phụ tương ứng với lớp độc lập có thể có giá trị POC duy nhất của nó trong AU. Khi ảnh tham chiếu trong số các ảnh được lưu giữ trong DPB được chỉ báo bằng cách sử dụng (các) phần tử cú pháp trong cấu trúc RPS hoặc RPL, thì (các) giá trị POC của mỗi ảnh phụ tương ứng với lớp có thể được sử dụng.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, nhằm mục đích chỉ báo cấu trúc dự báo (ngoài lớp), `layer_id` có thể không được sử dụng và giá trị POC (delta) có thể được sử dụng.

Theo cùng phương án của sáng chế, ảnh phụ với giá trị POC bằng N tương ứng với lớp (hoặc vùng địa phương) có thể có hoặc có thể không được sử dụng làm ảnh tham chiếu của ảnh phụ với giá trị POC bằng N+K, tương ứng với cùng lớp (hoặc cùng vùng địa phương) để dự báo bù chuyển động. Trong đa số các trường hợp, giá trị của số K có thể bằng số lượng tối đa của các lớp (độc lập), mà có thể giống như số lượng các vùng phụ.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, FIG. 16 thể hiện trường hợp mở rộng của FIG. 15. Khi ảnh đầu vào được chia thành nhiều (ví dụ bốn) vùng phụ, thì mỗi vùng địa phương có thể được mã hóa với một hoặc nhiều lớp. Trong trường hợp này, số lượng các lớp độc lập có thể bằng số lượng các vùng phụ, và một hoặc nhiều lớp có thể tương ứng với vùng phụ. Như vậy, mỗi vùng phụ có thể được mã hóa với một hoặc nhiều lớp độc lập và không hoặc nhiều hơn lớp phụ thuộc.

Theo cùng phương án của sáng chế, trên FIG. 16, ảnh đầu vào có thể được chia thành bốn vùng phụ. Vùng phụ trên bên phải có thể được mã hóa thành hai lớp, là lớp 1 và lớp 4, trong khi vùng phụ dưới bên phải có thể được mã hóa thành hai lớp, là lớp 3 và lớp 5. Trong trường hợp này, lớp 4 có thể tham chiếu lớp 1 để dự báo bù chuyển động, trong khi lớp 5 có thể tham chiếu lớp 3 để bù chuyển động.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, lọc trong vòng (như lọc tách khối, lọc trong vòng thích ứng, tạo lại dạng, lọc hai

chiều hoặc việc lọc dựa trên học sâu bất kỳ) qua lớp biên có thể (theo tùy chọn) được vô hiệu.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, dự báo bù chuyển động hoặc sao chép trong khối qua lớp biên có thể (theo tùy chọn) được vô hiệu.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, việc đệm biên để dự báo bù chuyển động hoặc lọc trong vòng tại biên của ảnh phụ có thể được xử lý theo tùy chọn. Cờ để chỉ báo liệu việc biên đệm có được xử lý hoặc hay không có thể được báo tín hiệu trong cấu trúc cú pháp mức cao, như (các) tập thông số (VPS, SPS, PPS, hoặc APS), tiêu đề lát hoặc nhóm ô, hoặc thông điệp SEI.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, thông tin bố trí của (các) vùng phụ (hoặc (các) ảnh phụ) có thể được báo tín hiệu trong VPS hoặc SPS. FIG. 17 thể hiện ví dụ về các phần tử cú pháp trong VPS và SPS. Trong ví dụ này, `vps_sub_picture_dividing_flag` được báo tín hiệu trong VPS. Cờ này có thể chỉ báo liệu (các) ảnh đầu vào có được chia thành nhiều vùng phụ hay không. Khi giá trị của `vps_sub_picture_dividing_flag` bằng 0, thì (các) ảnh đầu vào trong (các) dãy video được mã hóa tương ứng với VPS hiện tại có thể không được chia thành nhiều vùng phụ. Trong trường hợp này, kích thước ảnh đầu vào có thể bằng kích thước ảnh được mã hóa (`pic_width_in_luma_samples`, `pic_height_in_luma_samples`), mà được báo tín hiệu trong SPS. Khi giá trị của `vps_sub_picture_dividing_flag` bằng 1, thì (các) ảnh đầu vào có thể được chia thành nhiều vùng phụ. Trong trường hợp này, các phần tử cú pháp `vps_full_pic_width_in_luma_samples` và `vps_full_pic_height_in_luma_samples` được báo tín hiệu trong VPS. Các giá trị của `vps_full_pic_width_in_luma_samples` và `vps_full_pic_height_in_luma_samples` có thể lần lượt bằng chiều rộng và chiều cao của (các) ảnh đầu vào.

Theo cùng phương án của sáng chế, các giá trị của `vps_full_pic_width_in_luma_samples` và `vps_full_pic_height_in_luma_samples` có thể không được sử dụng để giải mã,

nhưng có thể được sử dụng để tổng hợp và hiển thị.

Theo cùng phương án của sáng chế, khi giá trị của `vps_sub_picture_dividing_flag` bằng 1, thì các phần tử cú pháp `pic_offset_x` và `pic_offset_y` có thể được báo tín hiệu trong SPS, mà tương ứng với (các) lớp cụ thể. Trong trường hợp này, kích thước ảnh được mã hóa (`pic_width_in_luma_samples`, `pic_height_in_luma_samples`) được báo tín hiệu trong SPS có thể bằng chiều rộng và chiều cao của vùng phụ tương ứng với lớp cụ thể. Định vị (`pic_offset_x`, `pic_offset_y`) của góc trên bên trái của vùng phụ cũng có thể được báo tín hiệu trong SPS.

Theo cùng phương án của sáng chế, thông tin định vị (`pic_offset_x`, `pic_offset_y`) của góc trên bên trái của vùng phụ có thể không được sử dụng để giải mã, nhưng có thể được sử dụng để tổng hợp và hiển thị.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, thông tin bố trí (kích thước và định vị) của tất cả hoặc tập con (các) vùng phụ của (các) ảnh đầu vào, thông tin phụ thuộc giữa (các) lớp có thể được báo tín hiệu trong tập thông số hoặc thông điệp SEI. FIG. 18 thể hiện ví dụ về các phần tử cú pháp để chỉ báo thông tin về bố trí của các vùng phụ, sự phụ thuộc giữa các lớp, và mối quan hệ giữa vùng phụ và một hoặc nhiều lớp. Trong ví dụ này, phần tử cú pháp `num_sub_region` chỉ báo số lượng các vùng phụ (chữ nhật) trong dãy video được mã hóa hiện tại. Phần tử cú pháp `num_layers` chỉ báo số lượng các lớp trong dãy video được mã hóa hiện tại. giá trị của `num_layers` có thể bằng hoặc lớn hơn giá trị của `num_sub_region`. Khi vùng phụ bất kỳ được mã hóa thành một lớp, thì giá trị của `num_layers` có thể bằng giá trị của `num_sub_region`. Khi một hoặc nhiều vùng phụ được mã hóa thành nhiều lớp, thì giá trị của `num_layers` có thể lớn hơn giá trị của `num_sub_region`. Phần tử cú pháp `direct_dependency_flag[i][j]` chỉ báo sự phụ thuộc từ lớp `j` tới lớp `i`. `num_layers_for_region[i]` chỉ báo số lượng các lớp được liên kết với vùng phụ `i`. `sub_region_layer_id[i][j]` chỉ báo `layer_id` của lớp `j` được liên kết với vùng phụ `i`. `sub_region_offset_x[i]` và `sub_region_offset_y[i]` lần lượt chỉ báo vị trí ngang và dọc của góc trên bên trái của vùng phụ `i`. `sub_region_width[i]` và `sub_region_height[i]` lần lượt chỉ báo chiều rộng và chiều cao của vùng phụ `i`.

Theo một phương án của sáng chế, một hoặc nhiều các phần tử cú pháp mà chỉ rõ tập lớp đầu ra để chỉ báo một hoặc nhiều lớp được xuất ra có hoặc không có thông tin mức bậc tiêu sử có thể được báo tín hiệu trong cấu trúc cú pháp mức cao, ví dụ VPS, DPS, SPS, PPS, APS hoặc thông điệp SEI. FIG. 19 thể hiện phần tử cú pháp `num_output_layer_sets` để chỉ báo số lượng tập lớp đầu ra (OLS) trong dãy video được mã hóa mà tham chiếu đến VPS có thể được báo tín hiệu trong VPS. Đối với mỗi tập lớp đầu ra, `output_layer_flag` có thể được báo tín hiệu nhiều bằng số lượng các lớp đầu ra.

Theo cùng phương án của sáng chế, `output_layer_flag[i]` bằng 1 chỉ rõ rằng lớp *i* được xuất ra. `vps_output_layer_flag[i]` bằng 0 chỉ rõ rằng lớp *i* không được xuất ra.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, một hoặc nhiều phần tử cú pháp mà chỉ rõ thông tin mức bậc tiêu sử đối với mỗi tập lớp đầu ra có thể được báo tín hiệu trong cấu trúc cú pháp mức cao, ví dụ VPS, DPS, SPS, PPS, APS hoặc thông điệp SEI. FIG. 19 còn thể hiện phần tử cú pháp `num_profile_tile_level` để chỉ báo số lượng thông tin mức bậc tiêu sử trên OLS trong dãy video được mã hóa tham chiếu tới VPS có thể được báo tín hiệu trong VPS. Đối với mỗi tập lớp đầu ra, tập các phần tử cú pháp dành cho thông tin mức bậc tiêu sử hoặc chỉ số để chỉ báo thông tin mức bậc tiêu sử cụ thể trong số các mục trong thông tin mức bậc tiêu sử có thể được báo tín hiệu nhiều bằng số lượng các lớp đầu ra.

Theo cùng phương án của sáng chế, `profile_tile_level_idx[i][j]` chỉ rõ chỉ số, trong danh sách của các cấu trúc cú pháp `profile_tile_level()` trong VPS, của cấu trúc cú pháp `profile_tile_level()` mà áp dụng cho lớp *j* của OLS thứ *i*.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, như được thể hiện trên FIG. 20, các phần tử cú pháp `num_profile_tile_level` và/hoặc `num_output_layer_sets` có thể được báo tín hiệu khi số lượng các lớp tối đa lớn hơn 1 (`vps_max_layers_minus1 > 0`).

Theo cùng một phương án hoặc theo phương án khác của sáng chế, như được thể hiện trên FIG. 20, phần tử cú pháp `vps_output_layers_mode[i]` để chỉ báo chế độ báo tín hiệu lớp đầu ra cho tập lớp đầu ra *i* có thể có trong VPS.

Theo cùng phương án của sáng chế, $vps_output_layers_mode[i]$ bằng 0 chỉ rõ rằng chỉ lớp cao nhất được xuất ra với tập lớp đầu ra i . $vps_output_layer_mode[i]$ bằng 1 chỉ rõ rằng tất cả các lớp được xuất ra với tập lớp đầu ra i . $vps_output_layer_mode[i]$ bằng 2 chỉ rõ rằng các lớp mà được xuất ra là các lớp với $vps_output_layer_flag[i][j]$ bằng 1 với tập lớp đầu ra i . Nhiều giá trị hơn có thể được bảo lưu.

Theo cùng phương án của sáng chế, $output_layer_flag[i][j]$ có thể có hoặc có thể không được báo tín hiệu tùy thuộc vào giá trị của $vps_output_layers_mode[i]$ cho tập lớp đầu ra i .

Theo cùng một phương án hoặc theo phương án khác của sáng chế, như được thể hiện trên FIG. 20, cờ $vps_ptl_signal_flag[i]$ có thể có đối với tập lớp đầu ra i . Tùy thuộc vào giá trị của $vps_ptl_signal_flag[i]$ mà thông tin mức bậc tiểu sử dành cho tập lớp đầu ra i có thể có hoặc có thể không được báo tín hiệu.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, như được thể hiện trên FIG. 21, số lượng ảnh phụ, $max_subpics_minus1$, trong CVS hiện tại có thể được báo tín hiệu trong cấu trúc cú pháp mức cao, ví dụ VPS, DPS, SPS, PPS, APS hoặc thông điệp SEI.

Theo cùng phương án của sáng chế, như được thể hiện trên FIG. 21, bộ nhận dạng ảnh phụ, $sub_pic_id[i]$, đối với ảnh phụ i có thể được báo tín hiệu, khi số lượng ảnh phụ lớn hơn 1 ($max_subpics_minus1 > 0$).

Theo cùng một phương án hoặc theo phương án khác của sáng chế, một hoặc nhiều phần tử cú pháp để chỉ báo bộ nhận dạng ảnh phụ thuộc về mỗi lớp của mỗi tập lớp đầu ra có thể được báo tín hiệu trong VPS. Như được thể hiện trên FIG. 22, $sub_pic_id_lớp[i][j][k]$, mà chỉ báo ảnh phụ k có trong lớp j của tập lớp đầu ra i . Với thông tin đó, bộ giải mã có thể ghi nhận ảnh phụ nào có thể được giải mã và được xuất ra cho mỗi lớp của tập lớp đầu ra cụ thể.

Theo một phương án của sáng chế, tiêu đề ảnh (PH) là cấu trúc cú pháp chứa các phần tử cú pháp mà áp dụng cho tất cả các lát của ảnh được mã hóa. Gói ảnh (picture unit - PU) là tập của các gói NAL mà được liên kết với nhau theo quy tắc phân loại được chỉ rõ, liên tiếp trong trình tự giải mã, và chứa chính xác một ảnh được mã hóa. PU có thể chứa tiêu đề ảnh (PH) và một hoặc nhiều

gói NAL VCL hợp thành ảnh được mã hóa.

Theo một phương án của sáng chế, SPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, được bao gồm trong ít nhất một AU với TemporalId bằng 0 hoặc được cấp thông qua phương tiện bên ngoài.

Theo một phương án của sáng chế, SPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, được bao gồm trong ít nhất một AU với TemporalId bằng 0 trong CVS, mà chứa một hoặc nhiều PPS tham chiếu tới SPS, hoặc được cấp thông qua phương tiện bên ngoài.

Theo một phương án của sáng chế, SPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PPS, được bao gồm trong ít nhất một PU với nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các gói NAL PPS mà tham chiếu tới gói NAL SPS trong CVS, mà chứa một hoặc nhiều PPS tham chiếu tới SPS, hoặc được cấp qua phương tiện bên ngoài.

Theo một phương án của sáng chế, SPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PPS, được bao gồm trong ít nhất một PU với TemporalId bằng 0 và nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các gói NAL PPS mà tham chiếu đến gói NAL SPS hoặc được cấp qua phương tiện bên ngoài.

Theo một phương án của sáng chế, SPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PPS, được bao gồm trong ít nhất một PU với TemporalId bằng 0 và nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các gói NAL PPS mà tham chiếu đến gói NAL SPS trong CVS, mà chứa một hoặc nhiều PPS tham chiếu tới SPS, hoặc được cấp qua phương tiện bên ngoài.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, pps_seq_parameter_set_id chỉ rõ giá trị của sps_seq_parameter_set_id cho SPS được tham chiếu. Giá trị của pps_seq_parameter_set_id có thể giống như trong tất cả các PPS mà được tham chiếu bởi các ảnh được mã hóa trong CLVS.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, tất cả các gói NAL SPS với giá trị cụ thể của sps_seq_parameter_set_id trong CVS có thể có cùng nội dung.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, không phụ thuộc vào các giá trị `nuh_layer_id`, các gói NAL SPS có thể chia sẻ cùng không gian giá trị của `sps_seq_parameter_set_id`.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, giá trị `nuh_layer_id` của gói NAL SPS có thể bằng giá trị `nuh_layer_id` thấp nhất của các gói NAL PPS mà tham chiếu đến gói NAL SPS.

Theo một phương án của sáng chế, khi SPS với `nuh_layer_id` bằng m được tham chiếu bởi một hoặc nhiều PPS với `nuh_layer_id` bằng n , thì lớp với `nuh_layer_id` bằng m có thể giống như lớp với `nuh_layer_id` bằng n hoặc lớp tham chiếu (trực tiếp hoặc gián tiếp) của lớp với `nuh_layer_id` bằng m .

Theo một phương án của sáng chế, PPS (RBSP) là khả dụng cho quá trình giải mã trước khi nó được tham chiếu, được bao gồm trong ít nhất một AU với `TemporalId` bằng `TemporalId` của gói NAL PPS hoặc được cấp qua phương tiện bên ngoài.

Theo một phương án của sáng chế, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, được bao gồm trong ít nhất một AU với `TemporalId` bằng `TemporalId` của gói NAL PPS trong CVS, mà chứa một hoặc nhiều PH (hoặc các gói NAL lát được mã hóa) mà tham chiếu tới PPS, hoặc được cấp qua phương tiện bên ngoài.

Theo một phương án của sáng chế, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PH (hoặc các gói NAL lát được mã hóa), được bao gồm trong ít nhất một PU với `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của các gói NAL lát được mã hóa mà tham chiếu đến gói NAL PPS trong CVS, mà chứa một hoặc nhiều PH (hoặc các gói NAL lát được mã hóa) mà tham chiếu tới PPS, hoặc được cấp qua phương tiện bên ngoài.

Theo một phương án của sáng chế, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PH (hoặc các gói NAL lát được mã hóa), được bao gồm trong ít nhất một PU với `TemporalId` bằng `TemporalId` của gói NAL PPS và `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của các gói NAL lát được mã hóa mà tham chiếu đến gói NAL PPS trong

CVS, mà chứa một hoặc nhiều PH (hoặc các gói NAL lát được mã hóa) tham chiếu tới PPS, hoặc được cấp qua phương tiện bên ngoài.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, `ph_pic_parameter_set_id` trong PH chỉ rõ giá trị của `pps_pic_parameter_set_id` cho PPS được tham chiếu được sử dụng. Giá trị của `pps_seq_parameter_set_id` có thể giống như trong tất cả các PPS mà được tham chiếu bởi các ảnh được mã hóa trong CLVS.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, tất cả các gói NAL PPS với giá trị cụ thể của `pps_pic_parameter_set_id` trong PU sẽ có cùng nội dung.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, không phụ thuộc vào các giá trị `nuh_layer_id`, các gói NAL PPS có thể chia sẻ cùng không gian giá trị của `pps_pic_parameter_set_id`.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, giá trị `nuh_layer_id` của gói NAL PPS có thể bằng giá trị `nuh_layer_id` thấp nhất của các gói NAL lát được mã hóa mà tham chiếu đến gói NAL mà tham chiếu đến gói NAL PPS.

Theo một phương án của sáng chế, khi PPS với `nuh_layer_id` bằng m được tham chiếu bởi một hoặc nhiều gói NAL lát được mã hóa với `nuh_layer_id` bằng n , thì lớp với `nuh_layer_id` bằng m có thể giống như lớp với `nuh_layer_id` bằng n hoặc lớp tham chiếu (trực tiếp hoặc gián tiếp) của lớp với `nuh_layer_id` bằng m .

Theo một phương án của sáng chế, PPS (RBSP) sẽ khả dụng cho quá trình giải mã trước khi nó được tham chiếu, được bao gồm trong ít nhất một AU với `TemporalId` bằng `TemporalId` của gói NAL PPS hoặc được cấp qua phương tiện bên ngoài.

Theo một phương án của sáng chế, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, được bao gồm trong ít nhất một AU với `TemporalId` bằng `TemporalId` của gói NAL PPS trong CVS, mà chứa một hoặc nhiều PH (hoặc các gói NAL lát được mã hóa) mà tham chiếu tới PPS, hoặc được cấp qua phương tiện bên ngoài.

Theo một phương án của sáng chế, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PH (hoặc các gói NAL lát được mã hóa), được bao gồm trong ít nhất một PU với `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của các gói NAL lát được mã hóa mà tham chiếu đến gói NAL PPS trong CVS, mà chứa một hoặc nhiều PH (hoặc các gói NAL lát được mã hóa) mà tham chiếu tới PPS, hoặc được cấp qua phương tiện bên ngoài.

Theo một phương án của sáng chế, PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu bởi một hoặc nhiều PH (hoặc các gói NAL lát được mã hóa), được bao gồm trong ít nhất một PU với `TemporalId` bằng `TemporalId` của gói NAL PPS và `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của các gói NAL lát được mã hóa mà tham chiếu đến gói NAL PPS trong CVS, mà chứa một hoặc nhiều PH (hoặc các gói NAL lát được mã hóa) mà tham chiếu tới PPS, hoặc được cấp qua phương tiện bên ngoài.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, `ph_pic_parameter_set_id` trong PH chỉ rõ giá trị của `pps_pic_parameter_set_id` cho PPS được tham chiếu được sử dụng. Giá trị của `pps_seq_parameter_set_id` có thể giống nhau trong tất cả các PPS mà được tham chiếu bởi các ảnh được mã hóa trong CLVS.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, tất cả các gói NAL PPS với giá trị cụ thể của `pps_pic_parameter_set_id` trong PU sẽ có cùng nội dung.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, không phụ thuộc vào các giá trị `nuh_layer_id`, các gói NAL PPS có thể chia sẻ cùng không gian giá trị của `pps_pic_parameter_set_id`.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, giá trị `nuh_layer_id` của gói NAL PPS có thể bằng giá trị `nuh_layer_id` thấp nhất của các gói NAL lát được mã hóa mà tham chiếu đến gói NAL mà tham chiếu đến gói NAL PPS.

Theo một phương án của sáng chế, khi PPS với `nuh_layer_id` bằng m được tham chiếu bởi một hoặc nhiều gói NAL lát được mã hóa với `nuh_layer_id` bằng

n , thì lớp với `nuh_layer_id` bằng m có thể giống như lớp với `nuh_layer_id` bằng n hoặc lớp tham chiếu (trực tiếp hoặc gián tiếp) của lớp với `nuh_layer_id` bằng m .

Theo một phương án của sáng chế, khi cờ `no_temporal_sublayer_switching_flag` được báo tín hiệu trong DPS, VPS, hoặc SPS, thì giá trị `TemporalId` của PPS tham chiếu tới tập thông số chứa cờ bằng 1 có thể bằng 0, trong khi giá trị `TemporalId` của PPS tham chiếu tới tập thông số chứa cờ bằng 1 có thể bằng hoặc lớn hơn giá trị `TemporalId` của tập thông số.

Theo một phương án của sáng chế, mỗi PPS (RBSP) có thể khả dụng cho quá trình giải mã trước khi nó được tham chiếu, được bao gồm trong ít nhất một AU với `TemporalId` nhỏ hơn hoặc bằng `TemporalId` của gói NAL lát được mã hóa (hoặc gói NAL PH) mà tham chiếu đến nó hoặc được cấp qua phương tiện bên ngoài. Khi gói NAL PPS được bao gồm trong AU trước AU chứa gói NAL lát được mã hóa tham chiếu tới PPS, thì gói NAL VCL cho phép chuyển tầng lớp thời gian hoặc gói NAL VCL với `nal_unit_type` bằng `STSA_NUT`, mà chỉ báo rằng ảnh trong gói NAL VCL có thể là ảnh truy cập lớp phụ thời gian theo từng bước (step-wise temporal sublayer access - STSA), có thể không có mặt tiếp sau gói NAL PPS và trước gói NAL lát được mã hóa mà tham chiếu tới APS.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, gói NAL PPS và gói NAL lát được mã hóa (và gói NAL PH của nó) mà tham chiếu tới PPS có thể được bao gồm trong cùng AU.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, gói NAL PPS và gói NAL STSA có thể được bao gồm trong cùng AU, mà trước gói NAL lát được mã hóa (và gói NAL PH của nó) tham chiếu tới PPS.

Theo cùng một phương án hoặc theo phương án khác của sáng chế, gói NAL STSA, gói NAL PPS và gói NAL lát được mã hóa (và gói NAL PH của nó) mà tham chiếu tới PPS có thể có trong cùng AU.

Theo cùng phương án của sáng chế, giá trị `TemporalId` của gói NAL VCL chứa PPS có thể bằng giá trị `TemporalId` của gói NAL STSA trước đó.

Theo cùng phương án của sáng chế, giá trị đếm thứ tự ảnh (picture order count - POC) của gói NAL PPS có thể bằng hoặc lớn hơn giá trị POC của gói

NAL STSA.

Theo cùng phương án của sáng chế, giá trị đếm thứ tự ảnh (POC) của lát được mã hóa hoặc gói NAL PH, mà tham chiếu tới gói NAL PPS, có thể bằng hoặc lớn hơn giá trị POC của gói NAL PPS được tham chiếu.

Theo một phương án của sáng chế, giá trị của `sps_max_sublayers_minus1` sẽ giống nhau qua tất cả các lớp trong dãy video được mã hóa, bởi vì tất cả các gói NAL VCL trong AU sẽ có cùng giá trị `TemporalId`. Giá trị của `sps_max_sublayers_minus1` sẽ giống nhau trong tất cả các SPS mà được tham chiếu bởi các ảnh được mã hóa trong CVS.

Theo một phương án của sáng chế, giá trị `chroma_format_idc` của SPS được tham chiếu bởi một hoặc nhiều ảnh được mã hóa trong lớp A sẽ bằng giá trị `chroma_format_idc` trong SPS được tham chiếu bởi một hoặc nhiều ảnh được mã hóa trong lớp B, trong đó lớp A là lớp tham chiếu trực tiếp của lớp B. Đó là vì ảnh được mã hóa bất kỳ sẽ có cùng giá trị `chroma_format_idc` với ảnh tham chiếu của nó. Giá trị `chroma_format_idc` của SPS được tham chiếu bởi một hoặc nhiều ảnh được mã hóa trong lớp A sẽ bằng giá trị `chroma_format_idc` trong SPS được tham chiếu bởi một hoặc nhiều ảnh được mã hóa trong lớp tham chiếu trực tiếp của lớp A, trong CVS.

Theo một phương án của sáng chế, các giá trị `subpics_present_flag` và `sps_subpic_id_present_flag` của SPS được tham chiếu bởi một hoặc nhiều ảnh được mã hóa trong lớp A sẽ bằng các giá trị `subpics_present_flag` và `sps_subpic_id_present_flag` trong SPS được tham chiếu bởi một hoặc nhiều ảnh được mã hóa trong lớp B, trong đó lớp A là lớp tham chiếu trực tiếp của lớp B. Đó là vì việc bố trí ảnh phụ cần được sắp thẳng hàng hoặc được liên kết qua các lớp. Nếu không, ảnh phụ với nhiều lớp có thể không được trích ra đúng. Các giá trị `subpics_present_flag` và `sps_subpic_id_present_flag` của SPS được tham chiếu bởi một hoặc nhiều ảnh được mã hóa trong lớp A sẽ bằng các giá trị `subpics_present_flag` và `sps_subpic_id_present_flag` trong SPS được tham chiếu bởi một hoặc nhiều ảnh được mã hóa trong lớp tham chiếu trực tiếp của lớp A, trong CVS.

Theo một phương án của sáng chế, khi ảnh STSA trong lớp A được tham

chiếu bởi ảnh trong lớp tham chiếu trực tiếp của lớp A trong cùng AU, thì ảnh tham chiếu tới STSA sẽ là ảnh STSA. Nếu không, việc chuyển tầng lớp phụ thời gian không thể được đồng bộ hóa qua các lớp. khi gói NAL STSA trong lớp A được tham chiếu bởi gói NAL VCL trong lớp tham chiếu trực tiếp của lớp A trong cùng AU, giá trị `nal_unit_type` của gói NAL VCL mà tham chiếu tới gói NAL STSA sẽ bằng `STSA_NUT`.

Theo một phương án của sáng chế, khi ảnh RASL trong lớp A được tham chiếu bởi ảnh trong lớp tham chiếu trực tiếp của lớp A trong cùng AU, thì ảnh tham chiếu tới RASL sẽ là ảnh RASL. Nếu không, ảnh không thể được giải mã đúng. Khi gói NAL RASL trong lớp A được tham chiếu bởi gói NAL VCL trong lớp tham chiếu trực tiếp của lớp A trong cùng AU, thì giá trị `nal_unit_type` của gói NAL VCL mà tham chiếu tới gói NAL RASL sẽ bằng `RASL_NUT`.

Trong khi phần mô tả trên đây đã mô tả một số phương án ví dụ của sáng chế, vẫn có những thay đổi, hoán vị, và các phương án thay thế tương đương khác nhau, mà nằm trong phạm vi của sáng chế. Cần hiểu rằng những người có hiểu biết trung bình trong lĩnh vực này có thể tạo ra nhiều hệ thống và phương pháp mà, mặc dù không được thể hiện hoặc mô tả một cách rõ ràng ở đây, nhưng áp dụng những nguyên lý của sáng chế và do đó nằm trong phạm vi của sáng chế.

YÊU CẦU BẢO HỘ

1. Phương pháp bảo hiệu các tập hợp lớp đầu ra trong dòng video được tạo mã, có thể thực thi bằng bộ xử lý, trong đó phương pháp bao gồm các bước:

nhận dữ liệu video có nhiều lớp;

nhận diện một hoặc nhiều phần tử cú pháp thứ nhất xác định một hoặc nhiều tập hợp lớp đầu ra tương ứng với các lớp đầu ra từ trong số nhiều lớp của dữ liệu video được nhận; và

giải mã và hiển thị một hoặc nhiều lớp đầu ra tương ứng với các tập hợp lớp đầu ra được xác định,

trong đó một hoặc nhiều phần tử cú pháp được bảo hiệu trong cấu trúc cú pháp mức cao,

trong đó cấu trúc cú pháp mức cao bao gồm một tham số trong số: tập hợp tham số video, tập hợp tham số lệ thuộc, tập hợp tham số chuỗi, tập hợp tham số ảnh, tập hợp tham số thích ứng, và thông điệp thông tin tăng cường bổ sung, và

trong đó tập hợp tham số chuỗi khả dụng để giải mã trước khi tập hợp tham số chuỗi được tham chiếu bởi một hoặc nhiều tập hợp tham số ảnh.

2. Phương pháp theo điểm 1, trong đó phương pháp còn bao gồm các bước:

xác định số lượng của một hoặc nhiều tập hợp lớp đầu ra; và

bảo hiệu số lượng lớp đầu ra bằng cách thiết lập cờ lớp đầu ra.

3. Phương pháp theo điểm 1, trong đó phương pháp còn bao gồm bước nhận diện một hoặc nhiều phần tử cú pháp thứ hai xác định thông tin tiểu sử - tầng - mức đối với mỗi tập hợp trong một hoặc nhiều tập hợp lớp đầu ra.

4. Phương pháp theo điểm 3, trong đó một hoặc nhiều phần tử cú pháp thứ hai được bảo hiệu trong cấu trúc cú pháp mức cao.

5. Phương pháp theo điểm 4, trong đó cấu trúc cú pháp mức cao bao gồm một trong số: tập hợp tham số video, tập hợp tham số lệ thuộc, tập hợp tham số chuỗi, tập hợp tham số ảnh, tập hợp tham số thích ứng, và thông điệp thông tin tăng cường bổ sung.

6. Phương pháp theo điểm 3, trong đó phương pháp còn bao gồm:

xác định số lượng của thông tin tiểu sử - tầng - mức đối với mỗi tập hợp

trong một hoặc nhiều tập hợp lớp đầu ra; và

báo hiệu, đối với mỗi tập hợp trong một hoặc nhiều tập hợp lớp đầu ra, tập hợp của các phần tử cú pháp tương ứng với thông tin tiêu sử - tầng - mức, hoặc chỉ số chỉ báo mục thông tin tiêu sử - tầng - mức từ trong số thông tin tiêu sử - tầng - mức.

7. Phương pháp theo điểm 1, trong đó phương pháp còn bao gồm bước báo hiệu một hoặc nhiều phần tử cú pháp thứ ba chỉ báo định danh ảnh con tương ứng với mỗi lớp đầu ra được liên kết với một hoặc nhiều tập hợp lớp đầu ra.

8. Phương pháp theo điểm 7, trong đó định danh ảnh con được báo hiệu trong cấu trúc cú pháp mức cao bao gồm một trong số: tập hợp tham số video, tập hợp tham số lệ thuộc, tập hợp tham số chuỗi, tập hợp tham số ảnh, tập hợp tham số thích ứng, và thông điệp thông tin tăng cường bổ sung.

9. Hệ thống máy tính báo hiệu các tập hợp lớp đầu ra trong dòng video được tạo mã, trong đó hệ thống máy tính bao gồm:

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

một hoặc nhiều bộ xử lý máy tính được tạo cấu hình để truy nhập mã chương trình máy tính và vận hành như được ra lệnh bởi mã chương trình máy tính, mã chương trình máy tính bao gồm:

mã nhận được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính để nhận dữ liệu video có nhiều lớp;

mã nhận diện được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính nhận diện một hoặc nhiều phần tử cú pháp thứ nhất xác định một hoặc nhiều tập hợp lớp đầu ra tương ứng với các lớp đầu ra từ trong số nhiều lớp của dữ liệu video được nhận; và

mã giải mã và hiển thị được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính giải mã và hiển thị một hoặc nhiều lớp đầu ra tương ứng với các tập hợp lớp đầu ra được xác định,

trong đó một hoặc nhiều phần tử cú pháp được báo hiệu trong cấu trúc cú pháp mức cao,

trong đó cấu trúc cú pháp mức cao bao gồm một trong số: tập hợp tham số

video, tập hợp tham số lệ thuộc, tập hợp tham số chuỗi, tập hợp tham số ảnh, tập hợp tham số thích ứng, và thông điệp thông tin tăng cường bổ sung, và

trong đó tập hợp tham số chuỗi khả dụng để giải mã trước khi tập hợp tham số chuỗi được tham chiếu bởi một hoặc nhiều tập hợp tham số ảnh.

10. Hệ thống máy tính theo điểm 9, trong đó hệ thống còn bao gồm:

mã xác định được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính xác định số lượng của một hoặc nhiều tập hợp lớp đầu ra; và

mã báo hiệu được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính báo hiệu số lượng lớp đầu ra bằng cách thiết lập cờ lớp đầu ra.

11. Hệ thống máy tính theo điểm 9, trong đó hệ thống còn bao gồm mã nhận diện được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính nhận diện một hoặc nhiều phần tử cú pháp thứ hai xác định thông tin tiêu sử - tầng - mức đối với mỗi tập hợp trong một hoặc nhiều tập hợp lớp đầu ra.

12. Hệ thống máy tính theo điểm 11, trong đó một hoặc nhiều phần tử cú pháp thứ hai được báo hiệu trong cấu trúc cú pháp mức cao.

13. Hệ thống máy tính theo điểm 12, trong đó cấu trúc cú pháp mức cao bao gồm một trong số: tập hợp tham số video, tập hợp tham số lệ thuộc, tập hợp tham số chuỗi, tập hợp tham số ảnh, tập hợp tham số thích ứng, và thông điệp thông tin tăng cường bổ sung.

14. Hệ thống máy tính theo điểm 11, trong đó hệ thống còn bao gồm:

mã xác định được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính xác định số lượng thông tin tiêu sử - tầng - mức đối với mỗi tập hợp trong một hoặc nhiều tập hợp lớp đầu ra; và

mã báo hiệu được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính báo hiệu, đối với mỗi tập hợp trong một hoặc nhiều tập hợp lớp đầu ra, tập hợp của các phần tử cú pháp tương ứng với thông tin tiêu sử - tầng - mức, hoặc chỉ số chỉ báo mục thông tin tiêu sử - tầng - mức từ trong số thông tin tiêu sử - tầng - mức.

15. Hệ thống máy tính theo điểm 9, trong đó hệ thống còn bao gồm mã báo hiệu được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính báo hiệu một hoặc nhiều phần tử cú pháp thứ ba chỉ báo định danh ảnh con tương ứng với mỗi lớp

đầu ra được liên kết với một hoặc nhiều tập hợp lớp đầu ra.

16. Vật ghi máy tính đọc được bất biến đã lưu trữ trên đó chương trình máy tính để báo hiệu các tập hợp lớp đầu ra trong dòng video được tạo mã, chương trình máy tính được tạo cấu hình để khiến một hoặc nhiều bộ xử lý máy tính:

nhận dữ liệu video có nhiều lớp;

nhận diện một hoặc nhiều phân tử cú pháp thứ nhất xác định một hoặc nhiều các tập hợp lớp đầu ra tương ứng với các lớp đầu ra từ trong số nhiều lớp của dữ liệu video được nhận; và

giải mã và hiển thị một hoặc nhiều lớp đầu ra tương ứng với các tập hợp lớp đầu ra được xác định,

trong đó một hoặc nhiều phân tử cú pháp được báo hiệu trong cấu trúc cú pháp mức cao,

trong đó cấu trúc cú pháp mức cao bao gồm một trong số: tập hợp tham số video, tập hợp tham số lệ thuộc, tập hợp tham số chuỗi, tập hợp tham số ảnh, tập hợp tham số thích ứng, và thông điệp thông tin tăng cường bổ sung, và

trong đó tập hợp tham số chuỗi khả dụng để giải mã trước khi tập hợp tham số chuỗi được tham chiếu bởi một hoặc nhiều tập hợp tham số ảnh.

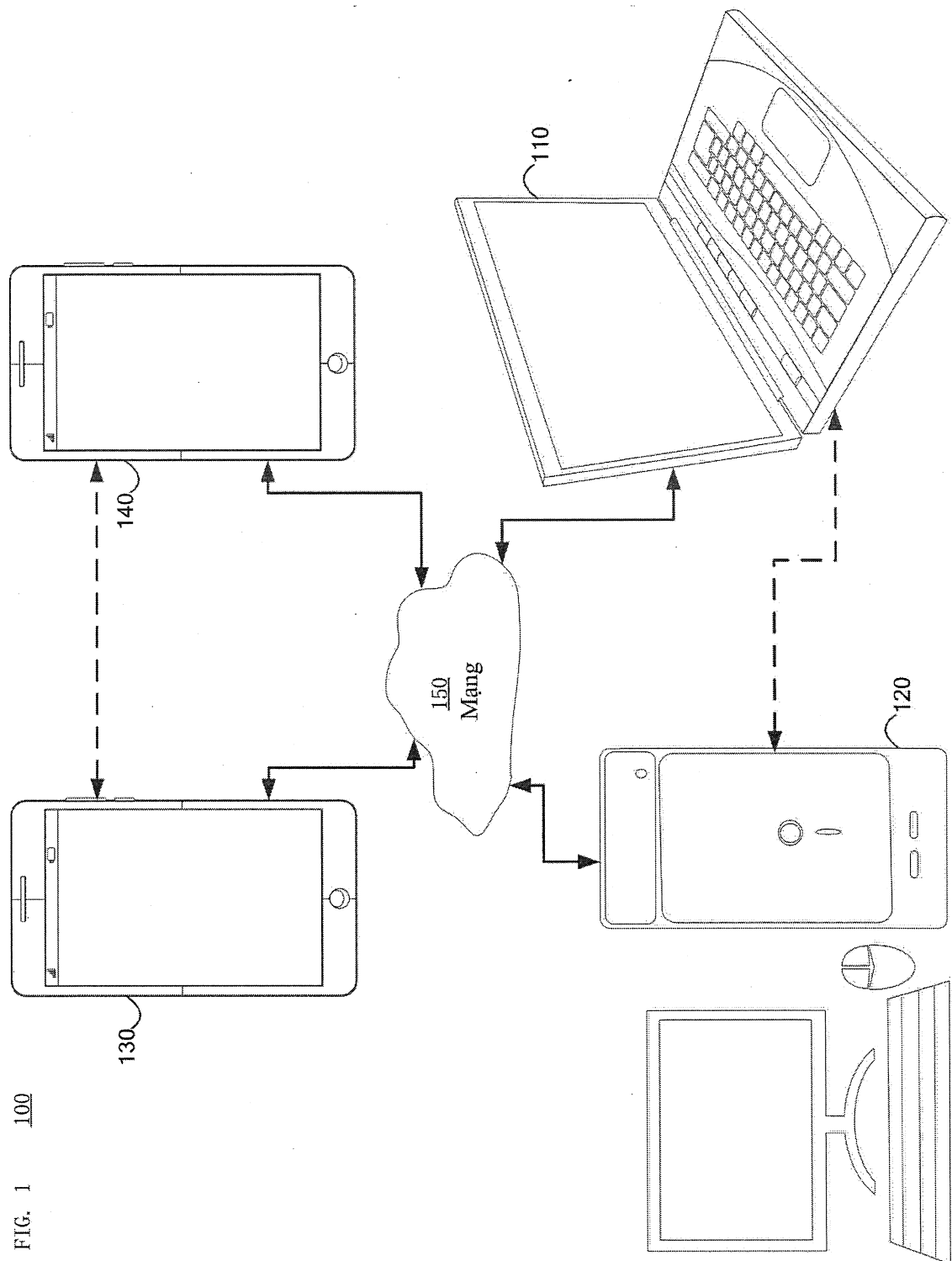


FIG. 2 Hệ thống streaming 200

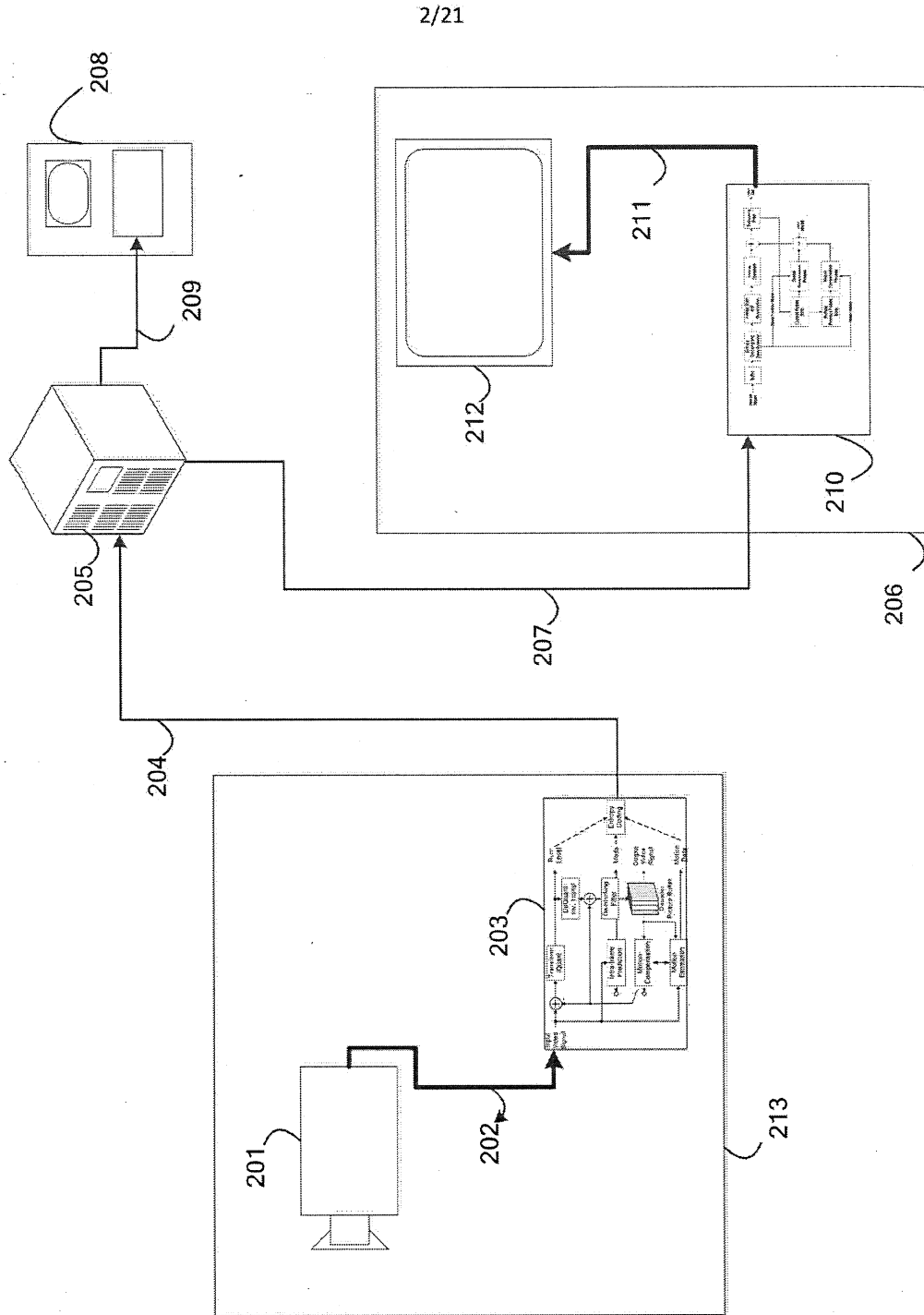
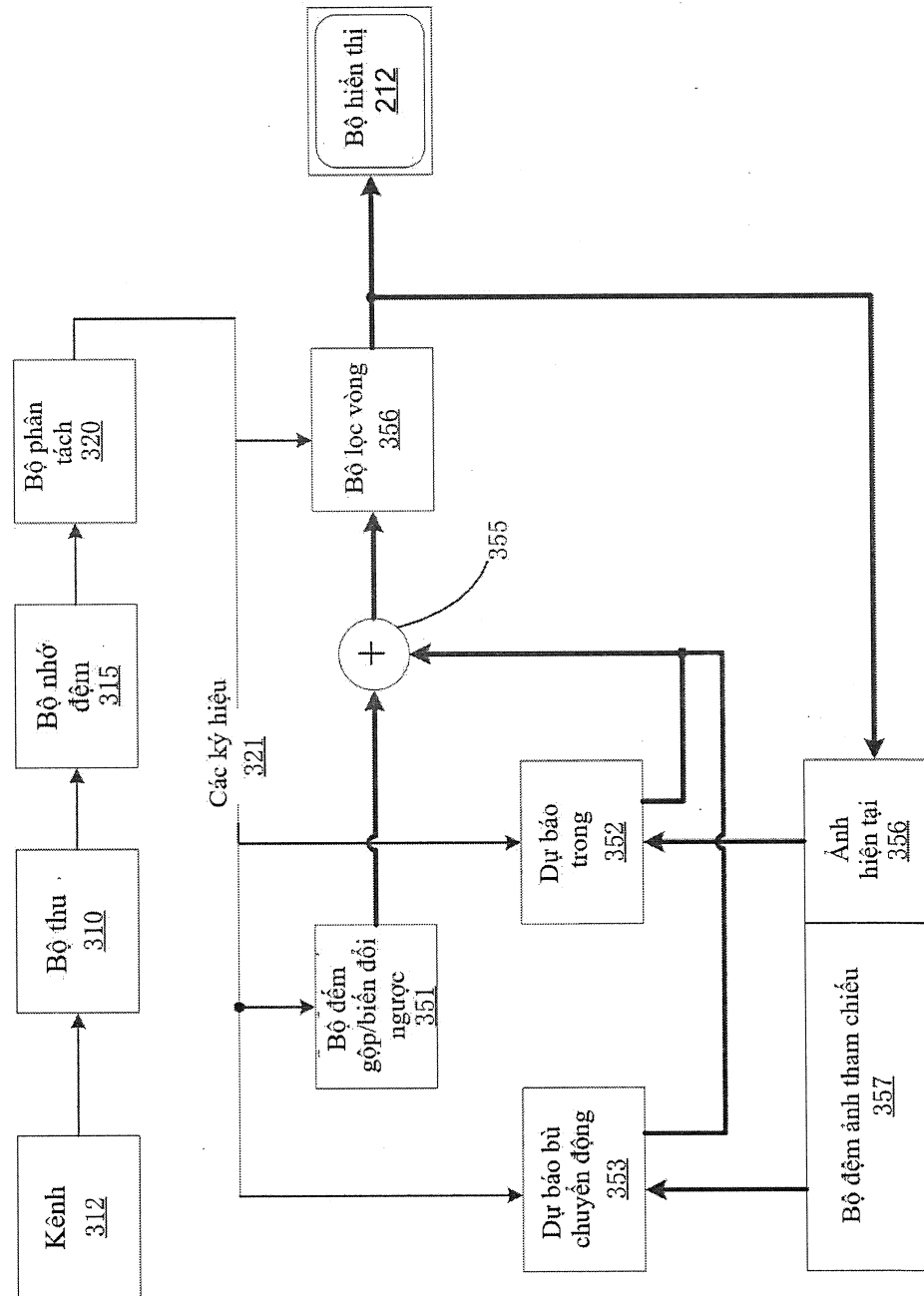


FIG. 3 Bộ giải mã 210



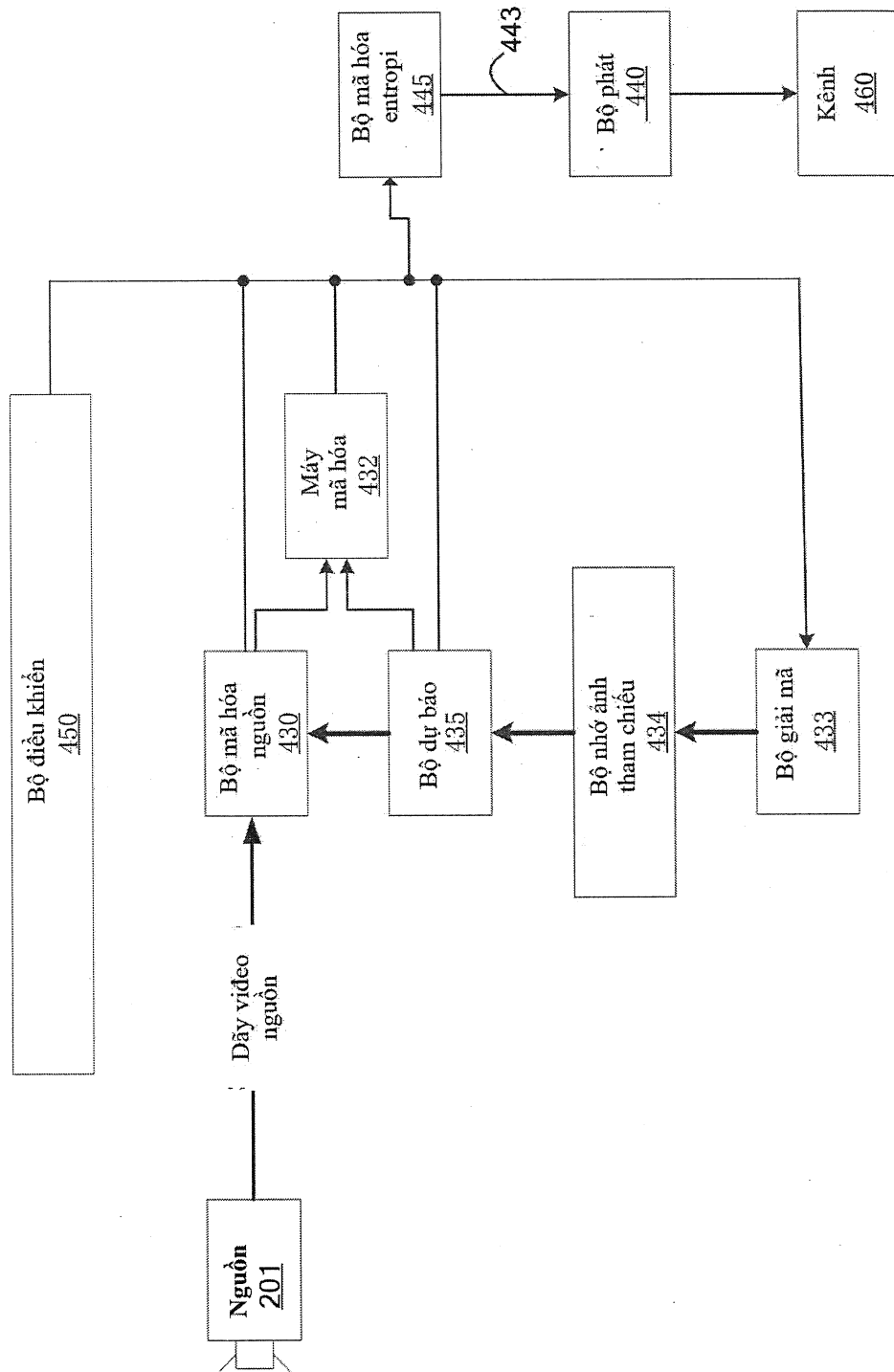
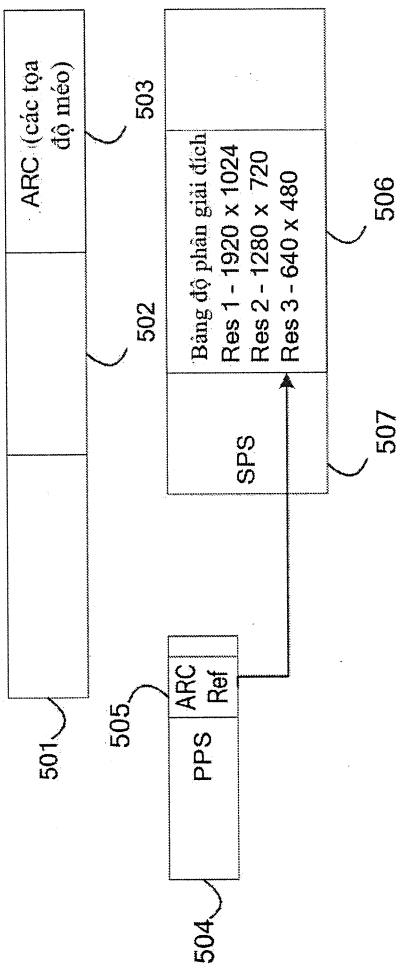
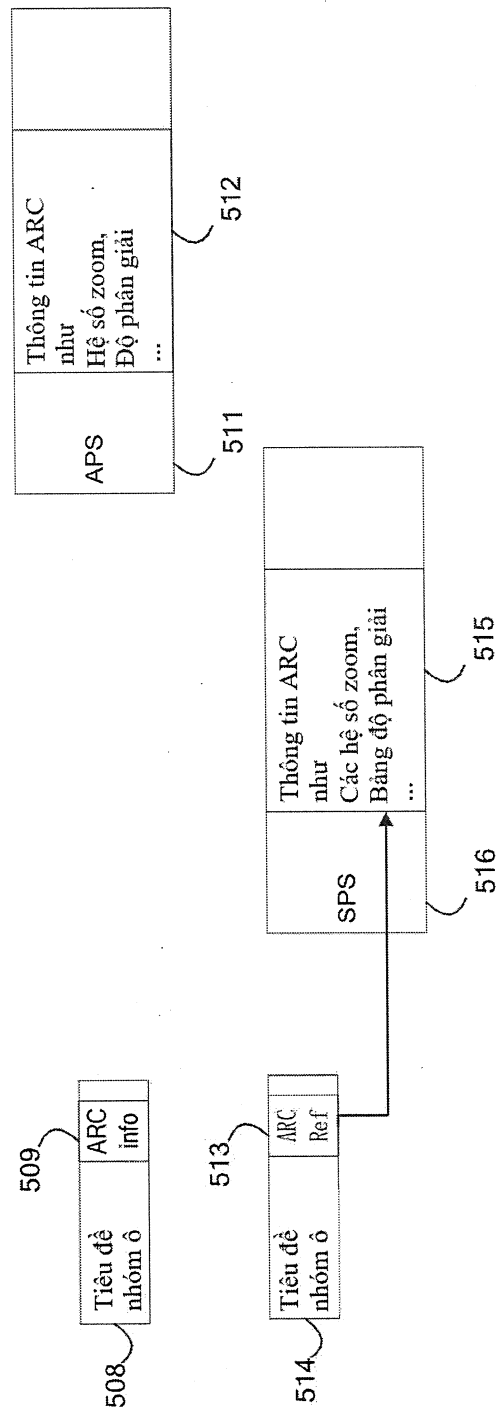


FIG. 4 Bộ tạo mã 203



Đối tượng
được bộc lộ



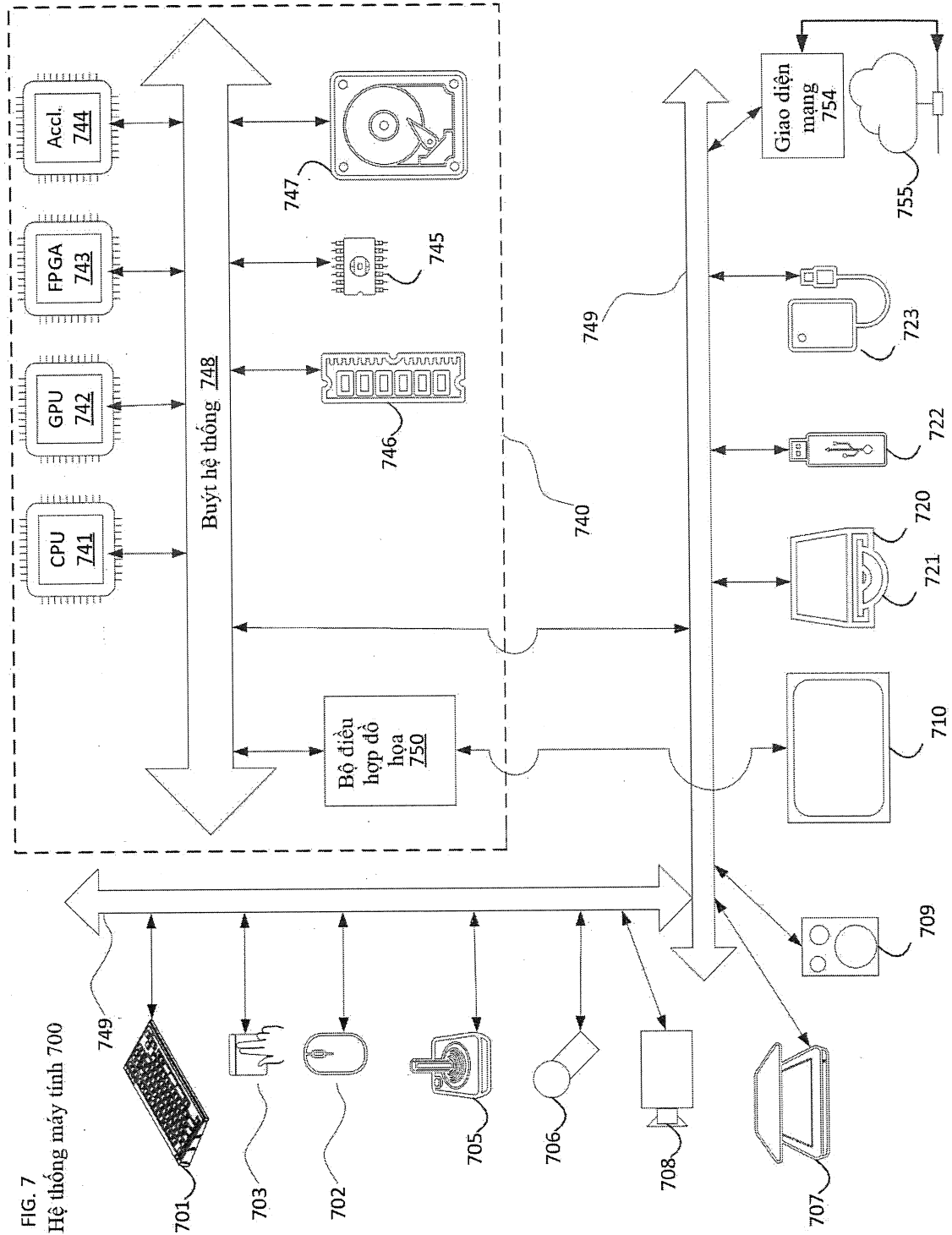
6/21

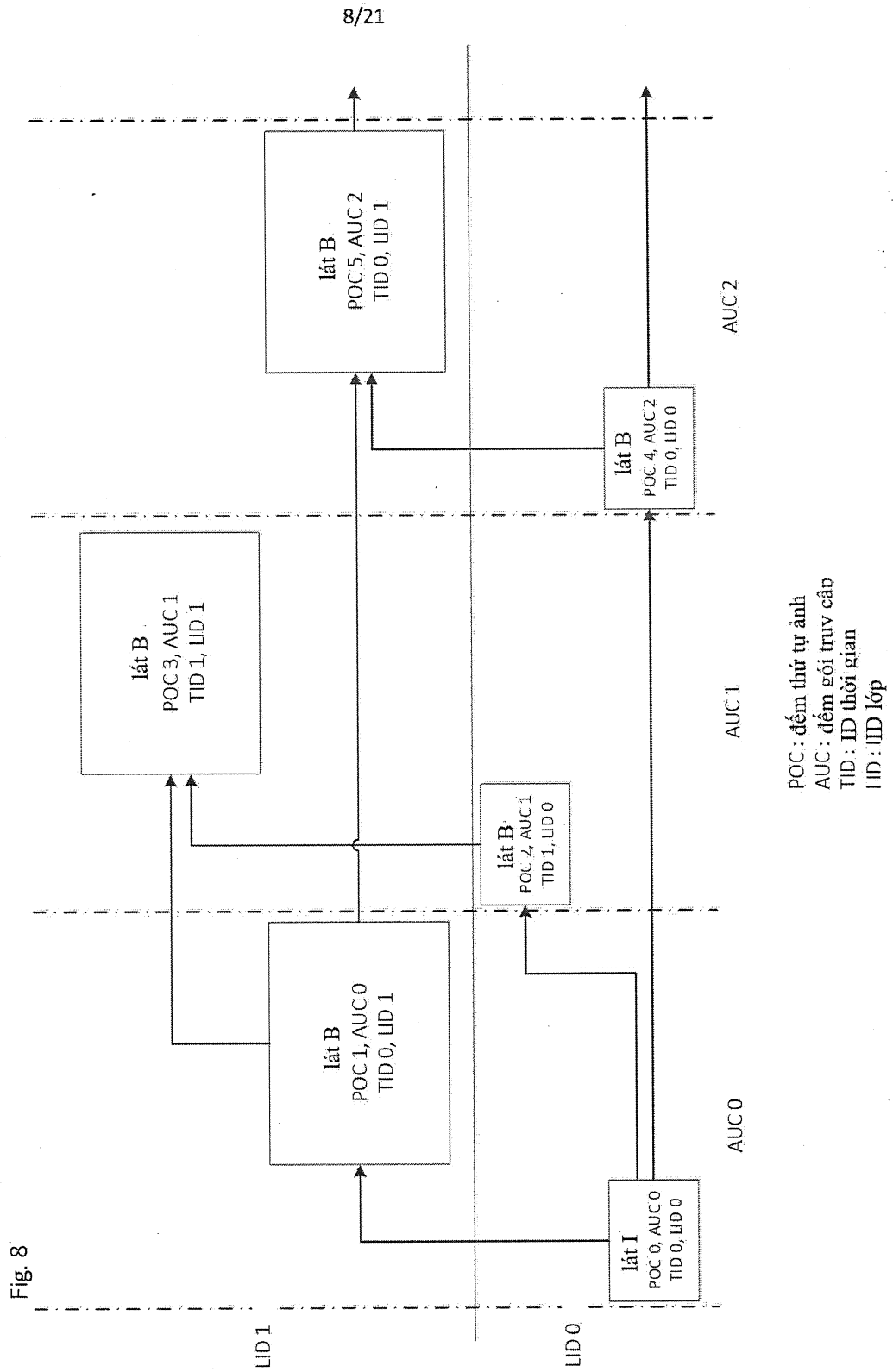
Fig. 6

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

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

7/21





9/21

Fig. 9

Descriptor ⁺	
u(4) ⁺	video_parameter_set_id { ⁺
u(8) ⁺	xps_video_parameter_set_id ⁺
⁺	vps_max_layers_minus1 ⁺
u(7) ⁺	for(i = 0; i <= vps_max_layers_minus1; i++) { ⁺
u(1) ⁺	vps_included_layer_id[i] ⁺
⁺	vps_reserved_zero_bit ⁺
u(1) ⁺	} ⁺
u(1) ⁺	vps_constraint_info_present_flag ⁺
⁺	vps_constant_poc_cycle_per_aud ⁺
u(8) ⁺	if(vps_constant_poc_cycle_per_aud) ⁺
⁺	vps_poc_cycle_aud ⁺
⁺	... ⁺
⁺	} ⁺
Descriptor ⁺	⁺
ue(v) ⁺	slice_header() { ⁺
⁺	slice_pic_parameter_set_id ⁺
u(v) ⁺	if(rect_slice_flag NumBricksInPic > 1) ⁺
⁺	slice_address ⁺
ue(v) ⁺	if(!rect_slice_flag && !single_brick_per_slice_flag) ⁺
ue(v) ⁺	num_bricks_in_slice_minus1 ⁺
⁺	slice_type ⁺
se(v) ⁺	if(NalUnitType == GRA_NUT) ⁺
u(v) ⁺	recovery_poc_cnt ⁺
⁺	slice_pic_order_cnt_lsb ⁺
⁺	... ⁺
⁺	if(!vps_constant_poc_cycle_per_aud) ⁺
u(8) ⁺	slice_poc_cycle_aud ⁺
⁺	... ⁺
⁺	} ⁺

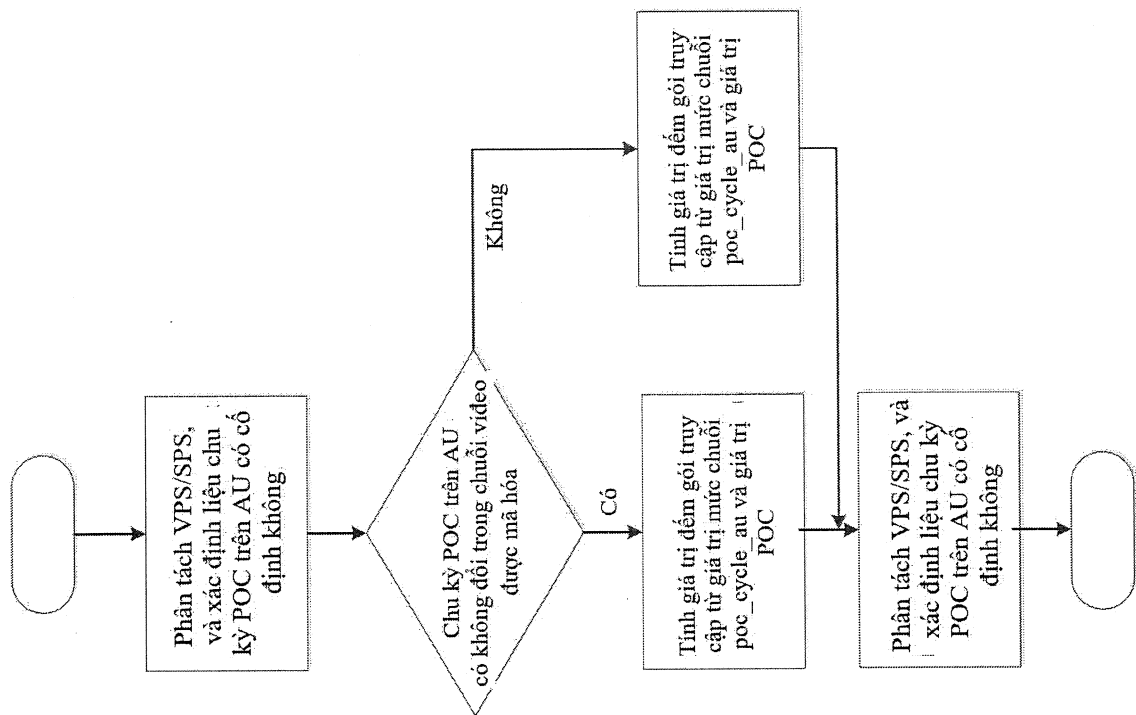
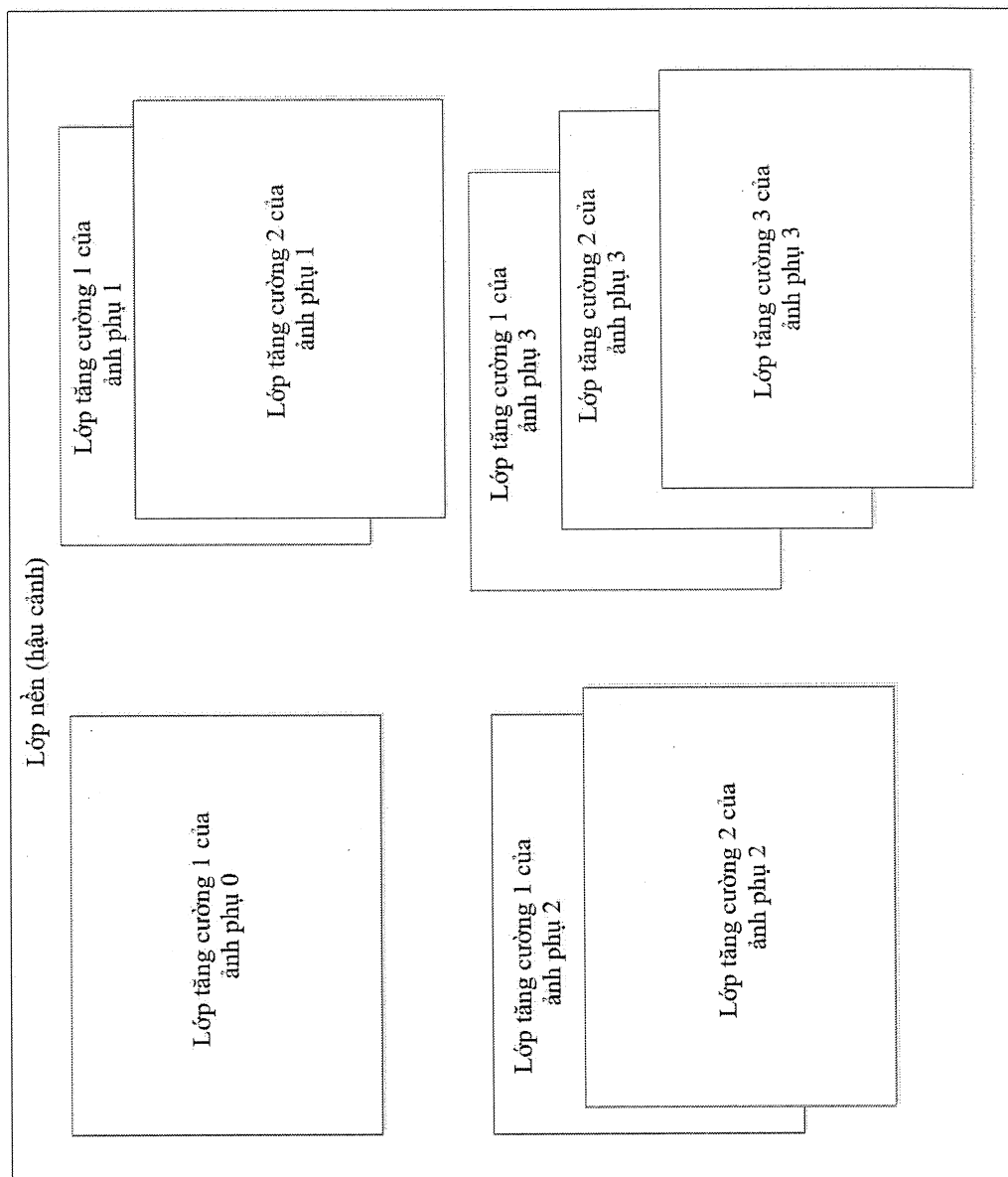


Fig. 10

FIG. 11



12/21

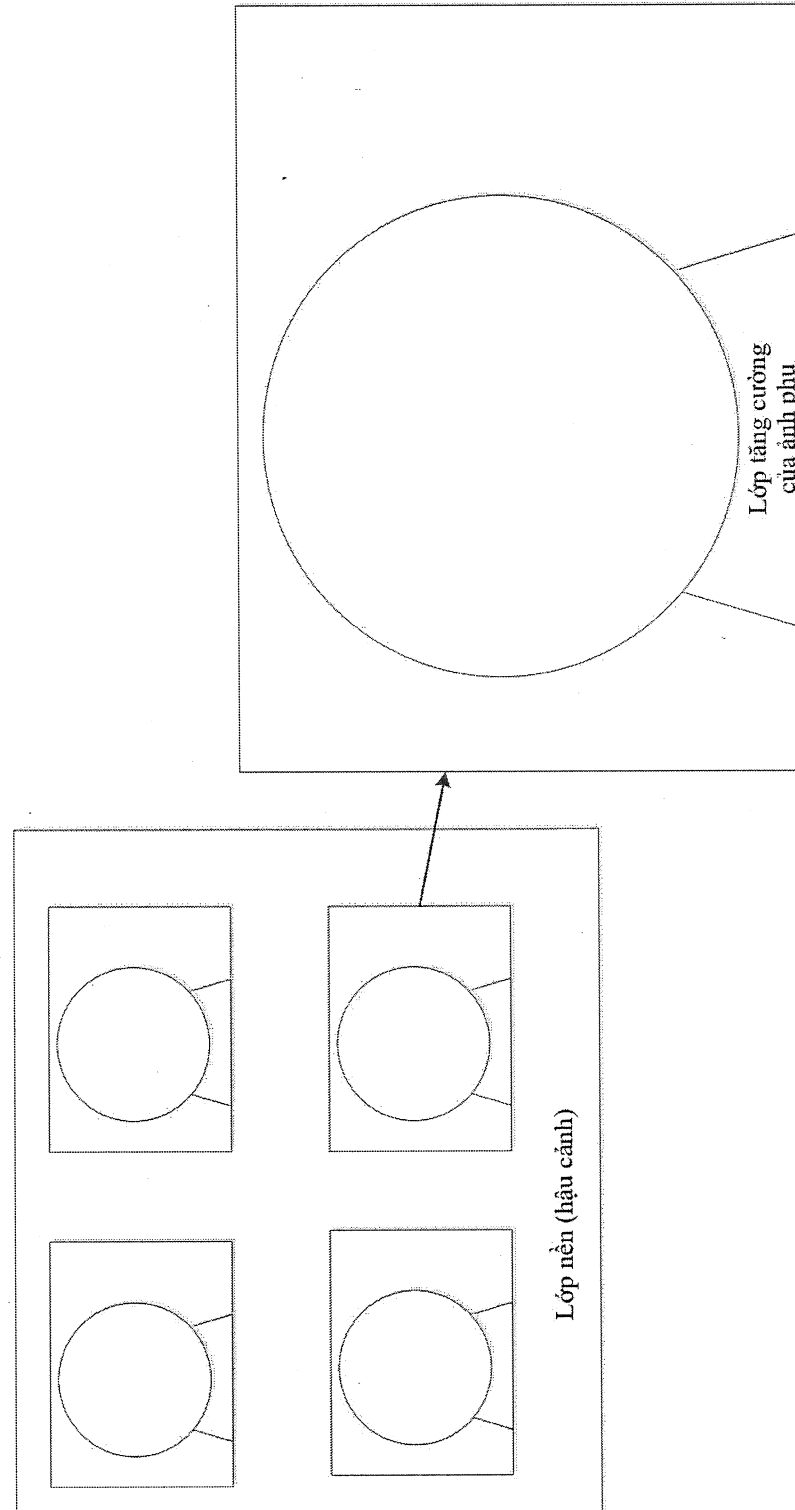
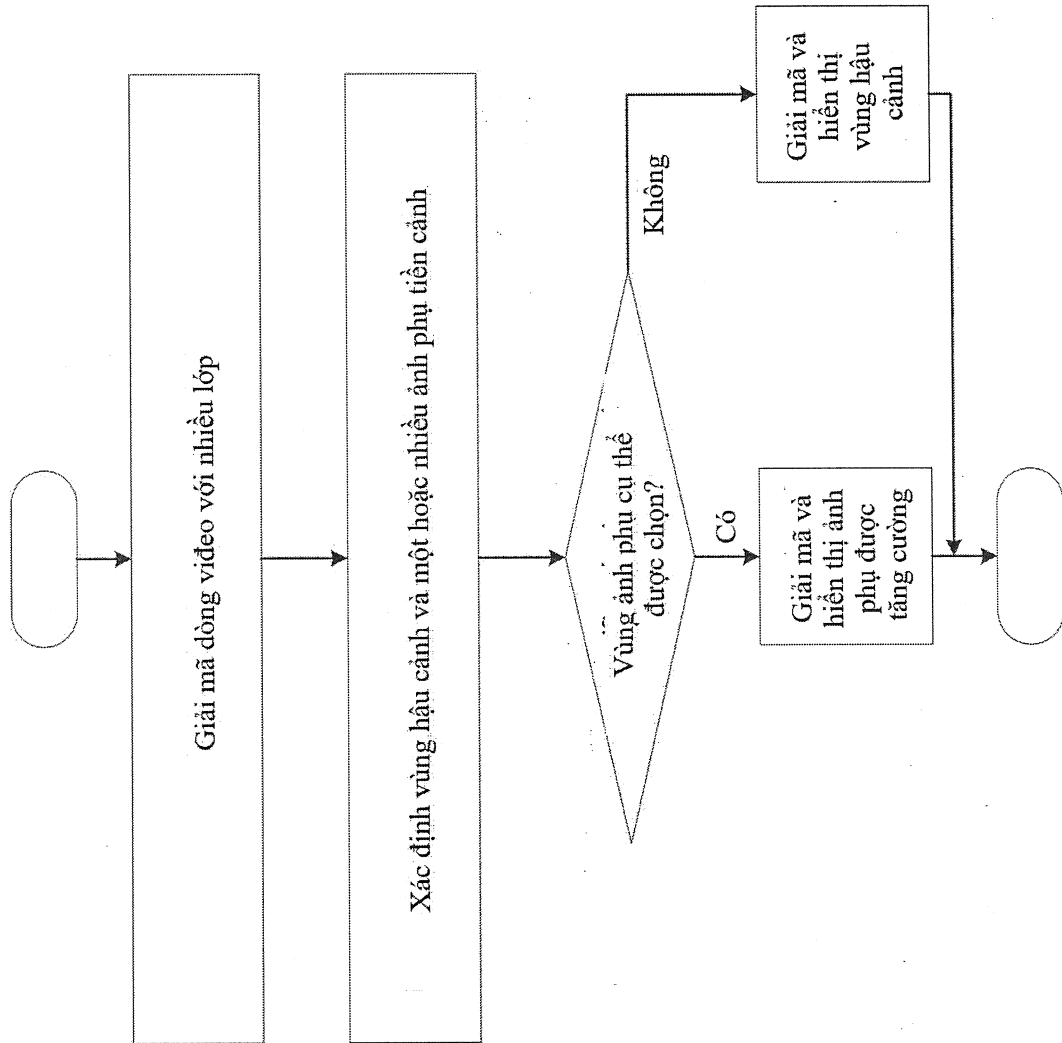


Fig. 13



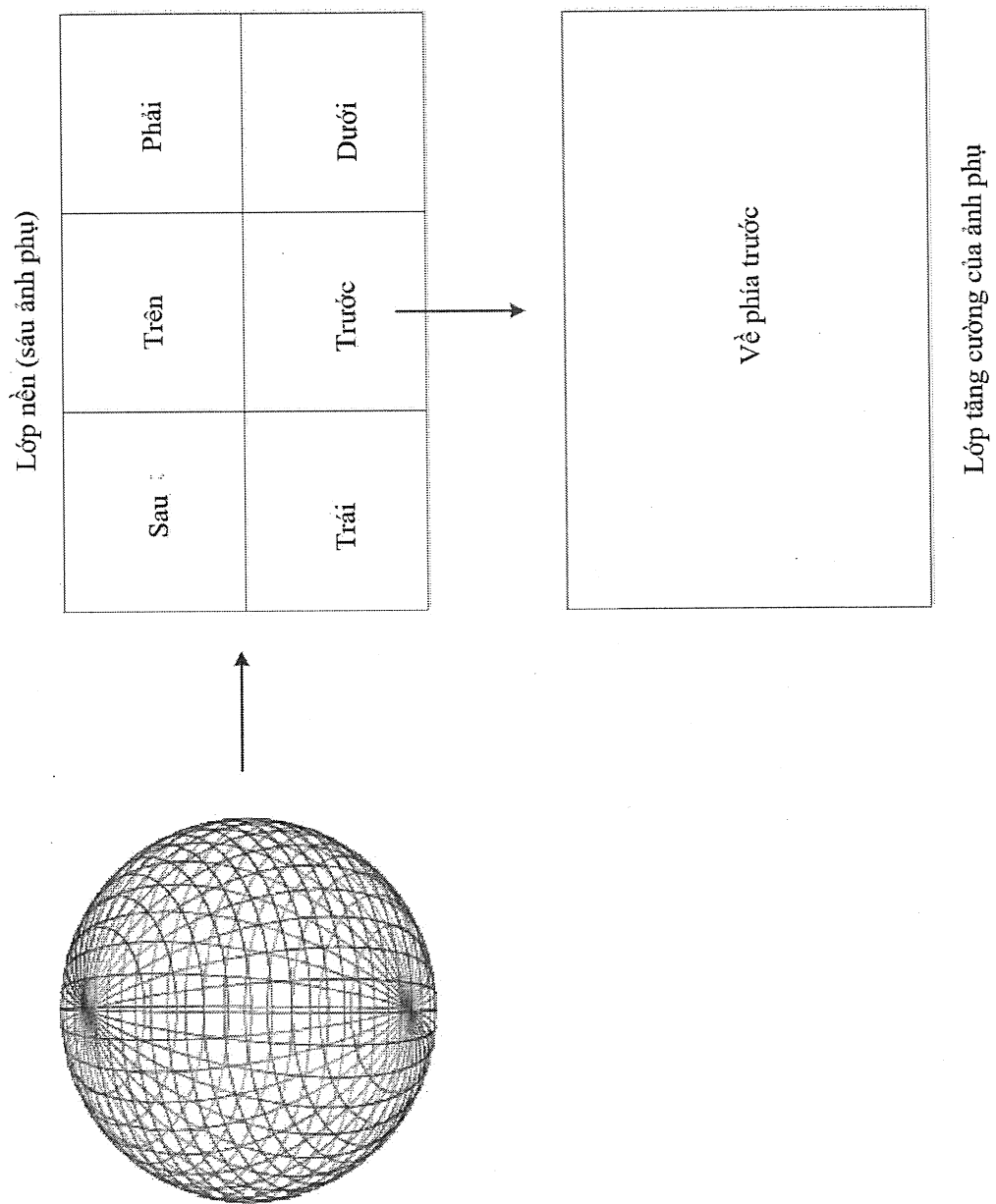


Fig. 15

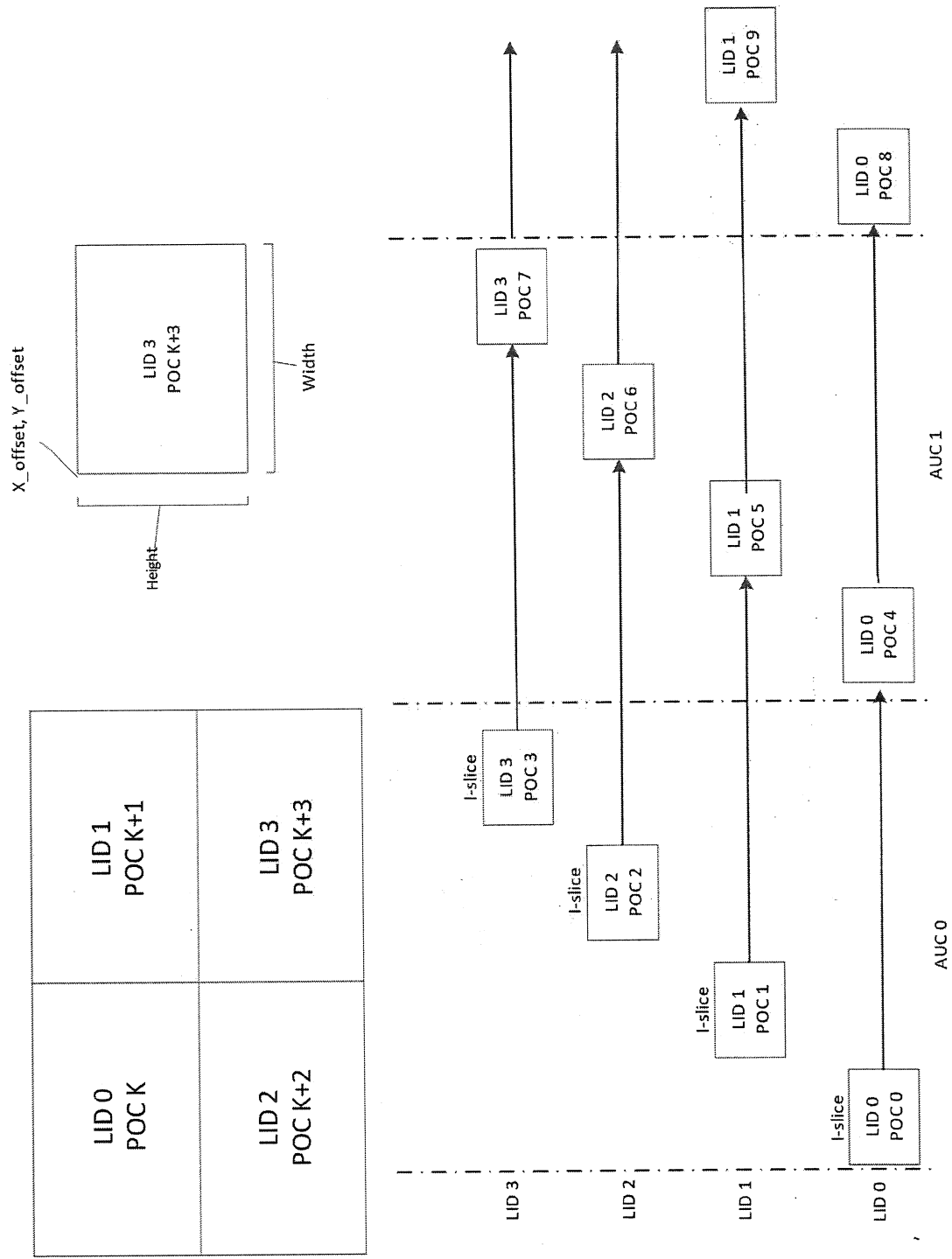


Fig. 16

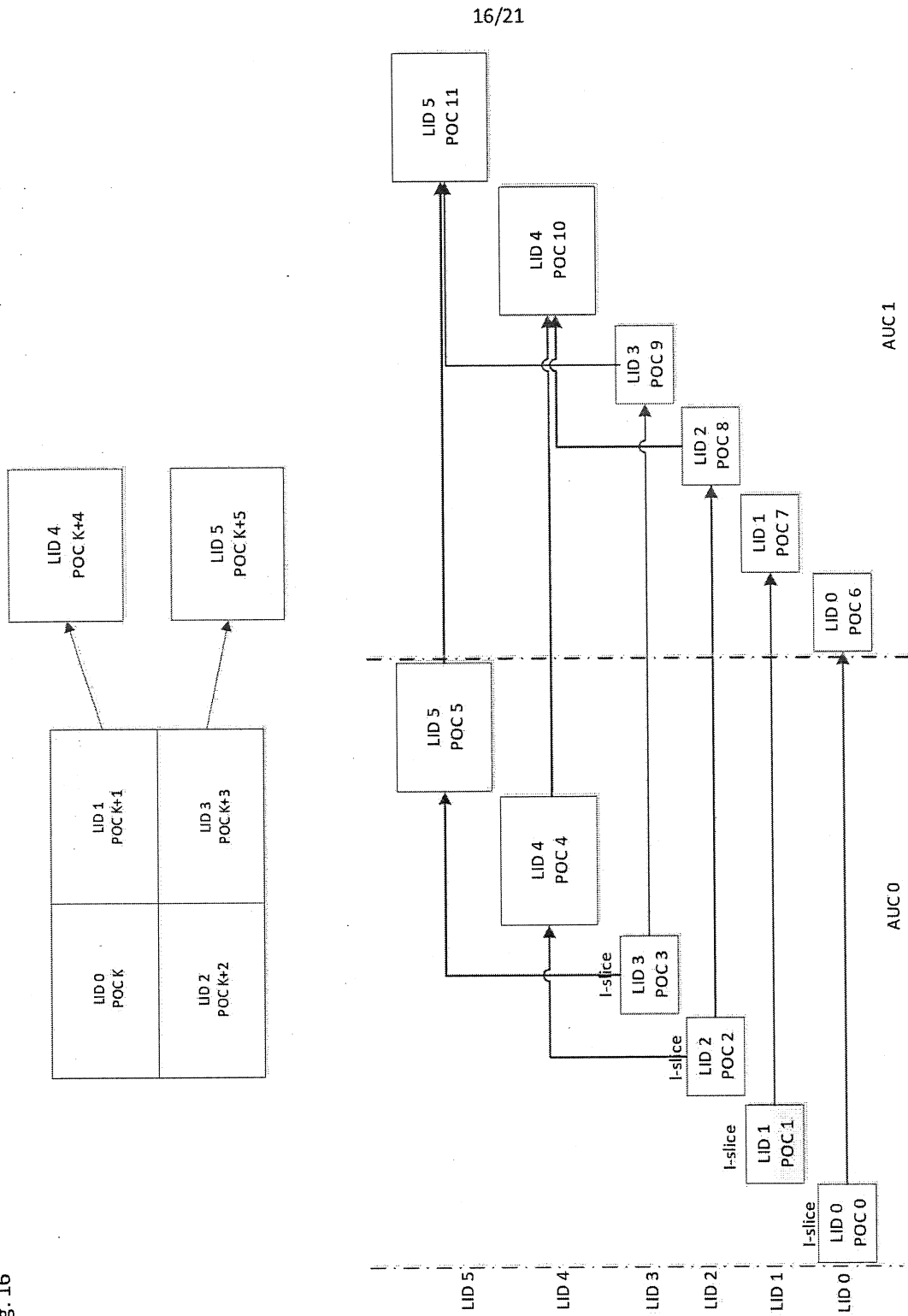


Fig. 17

	Descriptor ⁺
video_parameter_set_rbsp() { ⁺	
vps_video_parameter_set_id ⁺	u(4) ⁺
vps_max_layers_minus1 ⁺	u(8) ⁺
for(i = 0; i <= vps_max_layers_minus1; i++) { ⁺	
vps_included_layer_id[i] ⁺	u(7) ⁺
vps_reserved_zero_bit ⁺	u(1) ⁺
} ⁺	
... ⁺	
vps_sub_picture_dividing_flag ⁺	u(1) ⁺
if(vps_sub_picture_dividing_flag) { ⁺	
vps_full_pic_width_in_luma_samples ⁺	ue(v) ⁺
vps_full_pic_height_in_luma_samples ⁺	ue(v) ⁺
} ⁺	
... ⁺	
} ⁺	
seq_parameter_set_rbsp() { ⁺	
sps_decoding_parameter_set_id ⁺	u(4) ⁺
sps_video_parameter_set_id ⁺	u(4) ⁺
sps_max_sub_layers_minus1 ⁺	u(3) ⁺
... ⁺	
pic_width_in_luma_samples ⁺	ue(v) ⁺
pic_height_in_luma_samples ⁺	ue(v) ⁺
if(vps_sub_picture_dividing_flag) { ⁺	
pic_offset_x ⁺	ue(v) ⁺
pic_offset_y ⁺	ue(v) ⁺
} ⁺	
... ⁺	
} ⁺	

18/21

Descriptor	
sub_region_partitioning_info(payloadSize)	
num_sub_region	ue(v) ²
num_layers	ue(v) ²
for(i = 0; i <= num_layers; i++)	
layer_id[i]	ue(v) ²
for(i = 1; i < num_layers; i++)	
for(j = 0; j < i; j++)	
direct_dependency_flag[i][j]	u(1) ²
for(i = 0; i < num_sub_region; i++)	
num_layers_for_region[i]	ue(v) ²
for(i = 0; i < num_layers_for_region[i]; i++)	
sub_region_layer_id[i][j]	ue(v) ²
sub_region_offset_x[i]	ue(v) ²
sub_region_offset_y[i]	ue(v) ²
sub_region_width[i]	ue(v) ²
sub_region_height[i]	ue(v) ²
}	
...	
}	

Fig. 18

19/21

Fig. 19

	Descriptor
video_parameter_set_rbsp() {	
...	
yps_max_layers_minus1	u(6)
num_output_layer_sets	ue(v)
num_profile_tier_level	ue(v)
for(i = 0; i < num_profile_tier_level; i++)	
profile_tier_level(yps_max_sub_layers_minus1)	
for(i = 0; i < num_output_layer_sets; i++)	
for(j = 0; j < NumLayersInIdList[i]; j++) {	
output_layer_flag[i][j]	u(1)
profile_tier_level_idx[i][j]	u(v)
}	
...	
}	

	Descriptor
video_parameter_set_rbsp() {	
...	
vps_max_layers_minus1	u(6)
if(vps_max_layers_minus1 > 0) {	
num_output_layer_sets	ue(v)
num_profile_tier_level	ue(v)
}	
for(i = 0; i < num_profile_tier_level; i++)	
profile_tier_level(vps_max_sub_layers_minus1)	
for(i = 0; i < num_output_layer_sets; i++) {	
vps_output_layers_mode[i]	u(2)
vps_ptl_signal_flag[i]	u(1)
for(i = 0; i < NumLayersInIdList(I); i++) {	
if(vps_output_layers_mode[i] == 2)	
output_layer_flag[I][i]	u(1)
if(vps_ptl_signal_flag[i])	
profile_tier_level_idx[I][i]	u(v)
}	
}	
...	
}	

Fig. 20

	Descriptor
video_parameter_set_rbsp() {	
...	
vps_max_layers_minus1	u(6)
if(vps_max_layers_minus1 > 0) {	
num_output_layer_sets	ue(v)
num_profile_tier_level	ue(v)
}	
max_subpics_minus1	u(8)
for(i = 0; i < max_subpics_minus1; i++) {	
sub_pic_id[i]	u(8)
}	
for(i = 0; i < num_profile_tier_level; i++)	
profile_tier_level(vps_max_sub_layers_minus1)	
for(i = 0; i < num_output_layer_sets; i++) {	
vps_output_layers_mode[i]	u(2)
vps_ptl_signal_flag[i]	u(1)
for(j = 0; j < NumLayersInList[i]; j++) {	
num_output_subpic_layer_fill[j]	ue(v)
for(k = 0; k < num_output_subpic_layer_fill[j]; k++)	
sub_pic_id_layer_fill[j][k]	u(8)
if(vps_output_layers_mode[i] == 2)	
output_layer_flag[i][j]	u(1)
if(vps_ptl_signal_flag[i])	
profile_tier_level_idx[i][j]	u(v)
}	
}	
...	
}	

Fig. 21