



(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)



1-0049102

(51)^{2006.01} H04B 1/66; H04N 7/26; H04N 7/12;
H04N 11/02; H04N 11/04

(13) B

(21) 1-2021-07222

(22) 09/11/2020

(86) PCT/US2020/059702 09/11/2020

(87) WO 2021/194558 30/09/2021

(30) 63/001,045 27/03/2020 US; 17/000,018 21/08/2020 US

(45) 25/07/2025 448

(43) 26/12/2022 417A

(73) TENCENT AMERICA LLC (US)

2747 Park Boulevard Palo Alto, California 94306, USA

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

(74) Công ty Luật TNHH Phạm và Liên danh (PHAM & ASSOCIATES)

(54) PHƯƠNG PHÁP VÀ THIẾT BỊ GIẢI MÃ, VÀ VẬT GHI MÁY TÍNH

(21)1-2021-07222

(57) Sáng chế đề cập đến phương pháp giải mã có thể bao gồm các bước: nhận dòng bit bao gồm dữ liệu ảnh/video được nén; phân tách hoặc dẫn xuất, từ dòng bit, bộ chỉ báo chế độ tập hợp lớp đầu ra trong tập hợp tham số video (VPS); nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra; nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tập hợp lớp đầu ra được nhận diện; và giải mã một hoặc nhiều lớp đầu ra ảnh được nhận diện.

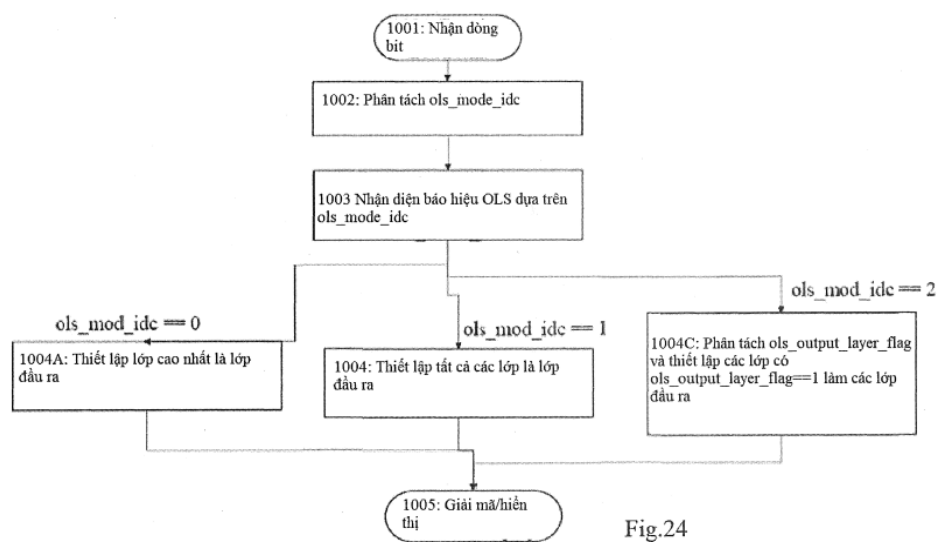


Fig.24

Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập đến các công nghệ nén video và dự báo ngoài và dự báo trong trong codec video cao cấp. Cụ thể là, sáng chế đề cập đến các công nghệ tạo mã video thế hệ tiếp theo bao gồm các công nghệ mã hóa/giải mã video ngoài tạo mã video hiệu suất cao (HEVC), chẳng hạn mã hóa video đa dụng (VVC). Cụ thể hơn, khía cạnh sáng chế đề cập đến phương pháp, thiết bị và vật ghi máy tính đọc được đề cập đến dẫn suất lớp đầu ra được dành cho tập hợp công nghệ mã hóa video cao cấp trong dòng video được mã hóa có nhiều lớp.

Tình trạng kỹ thuật của sáng chế

Kỹ thuật mã hóa và giải mã video nhờ sử dụng dự báo ngoài hoặc trong ảnh có bù chuyển động đã được biết đến trong nhiều thập kỷ. Video số không được nén có thể bao gồm chuỗi ảnh, mỗi ảnh có kích thước không gian, chẳng hạn, 1920 x 1080 mẫu độ sáng và các mẫu sắc độ liên kết. Chuỗi ảnh có thể có tốc độ ảnh cố định hoặc biến thiên (cũng thường được biết đến là tốc độ khung), chẳng hạn 60 ảnh trên giây hoặc 60 Hz. Video không được nén có các yêu cầu tốc độ bit quan trọng. Chẳng hạn, 1080p60 4:2:0 video ở 8 bit trên mẫu (1920x1080 độ phân giải mẫu độ sáng ở tốc độ khung hình 60 Hz) yêu cầu xấp xỉ băng thông 1,5 Gbit/s. Một giờ của video này yêu cầu nhiều hơn 600 gigabyte không gian lưu trữ.

Một mục đích của mã hóa và giải mã video có thể giảm dư thừa trong tín hiệu video đầu vào, thông qua nén. Việc nén có thể giúp giảm các yêu cầu băng thông hoặc không gian lưu trữ nêu trên, trong một số trường hợp đi hai bậc độ lớn hoặc lớn hơn. Cả nén không tổn hao và có tổn hao, cũng như kết hợp của nó có thể được sử dụng. Nén không tổn hao đề cập đến kỹ thuật trong đó bản sao chính xác của tín hiệu ban đầu có thể được tái tạo từ tín hiệu ban đầu được nén. Khi sử dụng nén tổn hao, tín hiệu được tái tạo có thể không giống hệt với tín hiệu ban đầu, nhưng méo dạng giữa tín hiệu ban đầu và được tái tạo đủ nhỏ để khiến tín hiệu được tái tạo hữu ích cho ứng dụng chủ đích. Trong trường hợp video, nén

có tổn hao được sử dụng rộng rãi. Lượng méo dạng được chấp thuận tùy thuộc vào ứng dụng; chẳng hạn, người dùng của các ứng dụng streaming (phát trực tuyến) tiêu dùng có thể chấp thuận méo dạng cao hơn người dùng của các ứng dụng phát sóng TV. Tỷ lệ nén có thể đạt được có thể phản ánh rằng: méo dạng được phép/chịu được cao hơn có thể cho các tỷ lệ nén cao hơn.

Bộ mã hóa video và bộ giải mã video có thể sử dụng các kỹ thuật từ vài phân loại rộng, bao gồm, chẳng hạn, bù chuyển động, biến đổi, lượng tử hóa, và mã hóa entropy, một số phân loại sẽ được nêu dưới đây.

Trong quá khứ, bộ mã hóa và bộ giải mã video sẽ hoạt động trên kích thước ảnh cụ thể đã, trong phần lớn các trường hợp, được định nghĩa và giữ không đổi cho chuỗi video được mã hóa (CVS), nhóm ảnh (GOP), hoặc khung thời gian nhiều ảnh tương tự. Chẳng hạn, trong MPEG-2, các thiết kế hệ thống được biết để thay đổi độ phân giải ngang (và, nhờ đó, kích thước ảnh) tùy thuộc vào các yếu tố, chẳng hạn, mức độ hoạt động của khung hình, nhưng chỉ ở các ảnh trong khung (hoặc khung i, hoặc ảnh i), do vậy thường cho GOP. Việc lấy mẫu lại các ảnh tham chiếu để sử dụng các độ phân giải khác nhau trong CVS được biết đến, chẳng hạn, từ ITU-T Rec. H.263 Annex P. Tuy nhiên, ở đây, kích thước ảnh không thay đổi, chỉ các ảnh tham chiếu được lấy mẫu lại, có thể dẫn đến chỉ các phần của canvas ảnh đang được sử dụng (trong trường hợp giảm kích thước mẫu), hoặc chỉ riêng các phần của khung hình được nhận (trong trường hợp tăng kích thước mẫu). Ngoài ra, H.263 Annex Q cho phép lấy mẫu lại khối lớn riêng rẽ theo hệ số 2 (trong mỗi chiều), chiều lên hoặc chiều xuống. Kích thước ảnh vẫn giữ nguyên. Kích thước của khối lớn được cố định trong H.263, và do vậy không cần được báo hiệu.

Các thay đổi của kích thước ảnh trong các ảnh được dự báo trở nên chủ đạo hơn trong mã hóa video hiện đại. Chẳng hạn, VP9 cho phép lấy mẫu lại ảnh tham chiếu và thay đổi độ phân giải cho toàn bộ ảnh. Tương tự, các kiến nghị cụ thể được thực hiện với mã hóa video đa dụng (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 đưa vào toàn bộ ở đây) cho phép lấy mẫu lại toàn bộ ảnh tham chiếu ở các độ phân giải khác nhau - cao hơn

hoặc thấp hơn. Theo Hendry, các độ phân giải ứng viên khác nhau được đề xuất để được mã hóa trong tập tham số chuỗi (SPS) và được đề cập bởi các phần tử cú pháp theo ảnh trong tập hợp tham số ảnh (PPS).

Bản chất kỹ thuật của sáng chế

Sáng chế đề cập đến các kỹ thuật báo hiệu kích thước ảnh thích ứng trong dòng bit video theo các phương án thực hiện khác nhau.

Theo khía cạnh của sáng chế, phương pháp giải mã có thể bao gồm các bước: nhận dòng bit bao gồm dữ liệu ảnh/video được nén, trong đó dòng bit có nhiều lớp; phân tách hoặc dẫn xuất, từ dòng bit, bộ chỉ báo chế độ tập hợp lớp đầu ra trong tập hợp tham số video (VPS); nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra; nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tập hợp lớp đầu ra được nhận diện; và giải mã một hoặc nhiều lớp đầu ra ảnh được nhận diện.

Việc nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra có thể bao gồm: trong trường hợp mà bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS là giá trị thứ nhất, nhận diện lớp cao nhất trong dòng bit dưới dạng một hoặc nhiều lớp đầu ra ảnh; trong trường hợp mà bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS là giá trị thứ hai, nhận diện tất cả các lớp trong dòng bit dưới dạng một hoặc nhiều lớp đầu ra ảnh; và trong trường hợp mà bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS là giá trị thứ ba, nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tường minh trong VPS.

Giá trị thứ nhất có thể khác với giá trị thứ hai và có thể khác với giá trị thứ ba, và giá trị thứ hai có thể khác với giá trị thứ ba.

Giá trị thứ nhất có thể bằng 0, giá trị thứ hai có thể bằng 1, và giá trị thứ ba có thể bằng 2. Tuy nhiên, các giá trị khác có thể được sử dụng và sáng chế không bị giới hạn ở việc sử dụng 0, 1 và 2 như nêu trên.

Việc nhận diện một hoặc nhiều lớp đầu ra ảnh bằng cách báo hiệu tường minh trong VPS có thể bao gồm: (i) phân tách hoặc dẫn xuất, từ VPS, cờ lớp đầu ra, và (ii) các lớp thiết lập có cờ lớp đầu ra bằng 1 bằng một hoặc nhiều lớp đầu ra ảnh.

Việc nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ lớp đầu ra có thể bao gồm: trong trường hợp mà bộ chỉ báo chế độ lớp đầu ra trong VPS là giá trị định trước, việc báo hiệu tập hợp lớp đầu ra bao gồm nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tường minh trong VPS.

Việc nhận diện một hoặc nhiều lớp đầu ra ảnh bằng cách báo hiệu tường minh trong VPS bao gồm các bước: (i) phân tách hoặc dẫn xuất, từ VPS, cờ lớp đầu ra, và (ii) thiết lập các lớp có cờ lớp đầu ra bằng 1 là một hoặc nhiều lớp đầu ra ảnh, trong đó số lượng nhiều lớp lớn hơn 2.

Việc báo hiệu tập hợp lớp đầu ra có thể bao gồm nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tường minh trong VPS khi bộ chỉ báo chế độ tập hợp lớp đầu ra bằng 2, và số lượng lớp trong nhiều lớp lớn hơn 2.

Việc báo hiệu tập hợp lớp đầu ra có thể bao gồm nhận diện lớp cao nhất trong dòng bit hoặc tất cả các lớp trong dòng bit dưới dạng một hoặc nhiều lớp đầu ra ảnh bằng cách suy ra một hoặc nhiều lớp đầu ra ảnh, khi bộ chỉ báo chế độ tập hợp lớp đầu ra nhỏ hơn 2, và số lượng nhiều lớp bằng 2, và bộ chỉ báo chế độ tập hợp lớp đầu ra thực sự nhỏ hơn 2, và số lượng nhiều lớp thực sự bằng 2.

Theo phương án thực hiện, bộ chỉ báo `number_output_layer_sets_minus1` trong VPS chỉ báo số lượng tập hợp lớp đầu ra.

Theo phương án thực hiện, bộ chỉ báo `VPS_maximum_layer_minus1` trong VPS chỉ báo số lượng lớp trong dòng bit.

Theo phương án thực hiện, cờ tập hợp lớp đầu ra `[i][j]` trong VPS chỉ báo liệu lớp thứ `j` của tập hợp lớp đầu ra thứ `i` có phải là lớp đầu ra hoặc không.

Theo phương án thực hiện, nếu tất cả nhiều lớp là độc lập trong đó cờ `vps_all_independent_layers_flag` của VPS bằng 1, bộ chỉ báo chế độ tập hợp lớp đầu ra không được báo hiệu, và giá trị của bộ chỉ báo chế độ tập hợp lớp đầu ra được suy ra là giá trị thứ hai.

Theo phương án thực hiện, khi mỗi lớp là tập hợp lớp đầu ra, cờ ảnh đầu ra của VPS được thiết lập bằng cờ đầu ra ảnh được báo hiệu trong tiêu đề ảnh, bất kể giá trị của bộ chỉ báo chế độ tập hợp lớp đầu ra.

Lưu ý: ảnh trong lớp đầu ra có thể hoặc không thể có `PictureOutputFlag` bằng 1. Ảnh trong lớp không phải đầu ra có `PictureOutputFlag` bằng 0. Ảnh có

PictureOutputFlag bằng 1 được xuất ra để hiển thị. Ảnh có PictureOutputFlag bằng 0 không được xuất ra để hiển thị.

Theo phương án thực hiện, cờ đầu ra ảnh được thiết lập bằng 0 khi định danh (ID) VPS tập hợp tham số chuỗi (SPS) lớn hơn 0, mà chỉ báo rằng nhiều hơn một lớp hiện có trong dòng bit, `each_layer_is_an_ols_flag` của VPS bằng 0, mà chỉ báo rằng nhiều lớp trong dòng bit đều không độc lập, bộ chỉ báo chế độ tập hợp lớp đầu ra bằng 0 và khối truy nhập hiện tại chứa ảnh thỏa mãn tất cả các điều kiện sau bao gồm: có cờ đầu ra ảnh bằng 1, có ID lớp NUH lớn hơn ID của ảnh hiện tại và thuộc lớp đầu ra của tập hợp lớp đầu ra.

Theo phương án thực hiện, cờ ảnh đầu ra của VPS được thiết lập bằng 0 khi SPS của VPS lớn hơn 0, mỗi lớp là cờ tập hợp lớp đầu ra bằng 0, bộ chỉ báo chế độ tập hợp lớp đầu ra bằng 2, và `output_layer_set_output_layer_flag [Target OLS Index][General Layer Index [nuh layer identifier]]` bằng 0.

Theo phương án thực hiện, phương pháp có thể còn bao gồm: điều khiển bộ hiển thị để hiển thị một hoặc nhiều lớp đầu ra ảnh được giải mã.

Theo khía cạnh của sáng chế, vật ghi máy tính đọc được bất biến lưu trữ các lệnh mà, khi được thực thi, khiến hệ thống hoặc thiết bị bao gồm một hoặc nhiều bộ xử lý: nhận dòng bit bao gồm dữ liệu ảnh/video được nén, trong đó dòng bit có nhiều lớp; phân tách hoặc dẫn xuất, từ dòng bit, bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS; nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra; nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tập hợp lớp đầu ra được nhận diện; và giải mã một hoặc nhiều lớp đầu ra ảnh được nhận diện.

Theo phương án thực hiện, các lệnh còn được tạo cấu hình để khiến hệ thống hoặc thiết bị bao gồm một hoặc nhiều bộ xử lý: điều khiển bộ hiển thị để hiển thị một hoặc nhiều lớp đầu ra ảnh được giải mã.

Theo khía cạnh của sáng chế, thiết bị có thể bao gồm: ít nhất một bộ nhớ được tạo cấu hình để lưu trữ mã chương trình máy tính; và ít nhất một bộ xử lý được tạo cấu hình để truy nhập ít nhất một bộ nhớ và hoạt động theo mã chương trình máy tính. Theo phương án thực hiện, mã chương trình máy tính có thể bao gồm: mã nhận được tạo cấu hình để khiến ít nhất một bộ xử lý nhận dòng bit bao

gồm dữ liệu ảnh/video được nén, trong đó dòng bit có nhiều lớp; mã phân tách hoặc dẫn xuất được tạo cấu hình để khiến ít nhất một bộ xử lý phân tách hoặc dẫn xuất, từ dòng bit, bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS; mã nhận diện báo hiệu lớp đầu ra được tạo cấu hình để khiến ít nhất một bộ xử lý nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra; mã nhận diện lớp đầu ra ảnh được tạo cấu hình để khiến ít nhất một bộ xử lý nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tập hợp lớp đầu ra được nhận diện; và mã giải mã được tạo cấu hình để khiến ít nhất một bộ xử lý: giải mã một hoặc nhiều lớp đầu ra ảnh được nhận diện.

Theo phương án thực hiện, mã chương trình máy tính có thể còn bao gồm: mã điều khiển hiển thị được tạo cấu hình để khiến ít nhất một bộ xử lý điều khiển bộ hiển thị để hiển thị một hoặc nhiều lớp đầu ra ảnh.

Theo khía cạnh của sáng chế, phương pháp báo hiệu kích thước ảnh thích ứng trong dòng bit video có thể bao gồm: nhận dòng bit bao gồm dữ liệu ảnh/video được nén, trong đó dòng bit có nhiều lớp; nhận diện vùng nền sau và một hoặc nhiều ảnh con nền trước; xác định liệu vùng ảnh con cụ thể đã được chọn; và dựa trên việc xác định vùng ảnh con cụ thể đã được chọn: tạo các khối được lượng tử ngược tương ứng với ảnh con được chọn bởi quá trình bao gồm mà không bị giới hạn ở phân tách dòng bit, giải mã dòng bit được mã hóa entropy và lượng tử ngược các khối tương ứng.

Phương pháp có thể còn bao gồm các bước: giải mã và hiển thị vùng nền sau, nếu vùng ảnh con cụ thể đã không được chọn.

Dòng bit có thể bao gồm các phần tử cú pháp mà xác định các lớp nào có thể được xuất ra ở phía bộ giải mã.

Các phần tử cú pháp có thể bao gồm tiêu đề của ảnh chứa phần tử cú pháp được mã hóa Exp-Golomb, chiều dài biến thiên.

Phương pháp có thể còn bao gồm các bước: xác định liệu độ phân giải thích ứng có đang sử dụng cho ảnh hoặc các phần ảnh của nó dựa trên báo hiệu trong tham số chuỗi.

Việc xác định liệu độ phân giải thích ứng có đang sử dụng cho ảnh hoặc các phần ảnh của nó hay không có thể bao gồm: xác định liệu phần tử cú pháp thứ nhất, vốn là cờ, chỉ báo sử dụng độ phân giải thích ứng.

Phương pháp có thể bao gồm bộ mã hóa ra lệnh bộ giải mã sử dụng kích thước ảnh tham chiếu cụ thể thay vì ngầm giả sử kích thước là kích thước ảnh đầu ra bằng cách sử dụng cờ giới hạn sự có mặt có điều kiện của các chiều ảnh tham chiếu.

Các phần tử cú pháp có thể bao gồm bảng chiều rộng và chiều cao ảnh giải mã khả thi.

Theo phương án thực hiện, giá trị trong tiêu đề đơn vị lớp trừu tượng mạng (NAL) có thể được sử dụng để chỉ báo không chỉ lớp thời gian mà cả lớp không gian.

Giá trị trong tiêu đề đơn vị NAL có thể là trường ID thời gian (ID).

Phương pháp có thể còn bao gồm bước: sử dụng đơn vị chuyển tiếp được chọn (SFU) hiện có được tạo và được tối ưu để chuyển tiếp lớp thời gian được chọn dựa trên giá trị ID thời gian tiêu đề đơn vị NAL mà không chỉnh sửa, đối với các môi trường khả mở.

Phương pháp có thể còn bao gồm: ánh xạ giữa kích thước ảnh được mã hóa và lớp thời gian được chỉ báo bởi trường ID thời gian trong tiêu đề đơn vị NAL.

Phương pháp có thể còn bao gồm: nhận dữ liệu bổ sung với video được mã hóa, dữ liệu bổ sung được bao gồm dưới dạng một phần của (các) chuỗi video được mã hóa; và sử dụng dữ liệu bổ sung để giải mã đúng dữ liệu và/hoặc để tái tạo chính xác dữ liệu video ban đầu.

Dữ liệu bổ sung có thể ở dạng một hoặc nhiều trong số: thời gian, không gian, hoặc các lớp tăng cường SNR, các lát dư, các ảnh dư, hoặc các mã sửa lỗi thuận.

Theo phương án thực hiện, vật ghi máy tính đọc được bất biến lưu trữ các lệnh mà, khi được thực thi, có thể khiến hệ thống hoặc thiết bị bao gồm một hoặc nhiều bộ xử lý: nhận dòng bit bao gồm dữ liệu ảnh/video được nén, trong đó dòng bit có nhiều lớp; nhận diện vùng nền sau và một hoặc nhiều ảnh con nền trước; xác định liệu vùng ảnh con cụ thể đã được chọn; và dựa trên việc xác định

vùng ảnh con cụ thể đã được chọn: tạo các khối được lượng tử ngược tương ứng với ảnh con được chọn bởi quá trình bao gồm mà không bị giới hạn ở phân tách dòng bit, giải mã dòng bit được mã hóa entropy và lượng tử ngược các khối tương ứng.

Theo phương án thực hiện, thiết bị có thể bao gồm: ít nhất một bộ nhớ được tạo cấu hình để lưu trữ mã chương trình máy tính; và ít nhất một bộ xử lý được tạo cấu hình để truy nhập ít nhất một bộ nhớ và hoạt động theo mã chương trình máy tính, mã chương trình máy tính bao gồm: mã nhận được tạo cấu hình để khiến ít nhất một bộ xử lý nhận dòng bit bao gồm dữ liệu ảnh/video được nén, trong đó dòng bit có nhiều lớp; mã nhận diện được tạo cấu hình để khiến ít nhất một bộ xử lý nhận diện vùng nền sau và một hoặc nhiều ảnh con nền trước; mã xác định được tạo cấu hình để khiến ít nhất một bộ xử lý xác định liệu vùng ảnh con cụ thể đã được chọn; và mã tạo được tạo cấu hình để khiến ít nhất một bộ xử lý: dựa trên việc xác định vùng ảnh con cụ thể đã được chọn: tạo các khối được lượng tử ngược tương ứng với ảnh con được chọn bởi quá trình bao gồm mà không bị giới hạn ở phân tách dòng bit, giải mã dòng bit được mã hóa entropy và lượng tử ngược các khối tương ứng.

Mô tả vắn tắt các hình vẽ

Các dấu hiệu khác, bản chất, và các ưu điểm khác nhau của sáng chế sẽ trở nên rõ ràng hơn trong phần mô tả chi tiết sau và các hình vẽ đi kèm trong đó:

Fig.1 là hình vẽ minh họa sơ đồ khối được giản lược của hệ thống truyền thông theo phương án thực hiện;

Fig.2 là hình vẽ minh họa sơ đồ khối được giản lược của hệ thống truyền thông theo phương án thực hiện;

Fig.3 là hình vẽ minh họa sơ đồ khối được giản lược của bộ giải mã theo phương án thực hiện;

Fig.4 là hình vẽ minh họa sơ đồ khối được giản lược của bộ mã hóa theo phương án thực hiện;

Fig.5 là hình vẽ minh họa các tùy chọn để báo hiệu các tham số ARC theo giải pháp kỹ thuật đã biết hoặc phương án thực hiện, như được chỉ báo;

Fig.6 là ví dụ của bảng cú pháp theo phương án thực hiện;

Fig.7 là hình vẽ minh họa hệ thống máy tính theo phương án thực hiện;

Fig.8 là ví dụ của cấu trúc dự báo đối với tính khả mở có thay đổi độ phân giải thích ứng;

Fig.9 là ví dụ của bảng cú pháp theo phương án thực hiện;

Fig.10 là hình vẽ minh họa sơ đồ khối được giản lược phân tách và giải mã chu trình POC trên khối truy nhập và giá trị đếm khối truy nhập;

Fig.11 là hình vẽ minh họa cấu trúc dòng bit video bao gồm các ảnh con được tạo nhiều lớp;

Fig.12 là hình vẽ minh họa hiển thị của ảnh con được chọn có độ phân giải được tăng cường;

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

Fig.14 là hình vẽ minh họa hiển thị video 360 với lớp tăng cường của ảnh con;

Fig.15 là ví dụ của thông tin bố trí của các ảnh con và lớp tương ứng của nó và cấu trúc dự báo ảnh;

Fig.16 là ví dụ của thông tin bố trí của các ảnh con và lớp tương ứng của nó và cấu trúc dự báo ảnh, với thể thức khả mở không gian của vùng cục bộ;

Fig.17 là ví dụ của bảng cú pháp đối với thông tin bố trí ảnh con;

Fig.18 là ví dụ của bảng cú pháp của thông điệp SEI đối với thông tin bố trí ảnh con;

Fig.19 là ví dụ của bảng cú pháp để chỉ báo các lớp đầu ra và thông tin tiêu sử/tầng/mức độ cho mỗi tập hợp lớp đầu ra;

Fig.20 là ví dụ của bảng cú pháp để chỉ báo chế độ lớp đầu ra bật cho mỗi tập hợp lớp đầu ra;

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

Fig.22 là ví dụ của bảng cú pháp của VPS RBSP;

Fig.23 là ví dụ của bảng cú pháp để chỉ báo tập hợp lớp đầu ra có bộ chỉ báo chế độ tập hợp lớp đầu ra;

Fig.24 là sơ đồ khối thể hiện quá trình giải mã cho dòng bit bao gồm để chỉ báo tập hợp lớp đầu ra có bộ chỉ báo chế độ tập hợp lớp đầu ra; và

Fig.25A đến Fig.25C là các hình vẽ thể hiện thông tin liên quan đến việc chỉ báo tập hợp lớp đầu ra có bộ chỉ báo chế độ tập hợp lớp đầu ra.

Mô tả chi tiết sáng chế

Khi các ảnh được mã hóa thành dòng bit bao gồm nhiều lớp có các chất lượng khác nhau, dòng bit có thể có các phần tử cú pháp xác định các lớp nào có thể được xuất ra ở phía bộ giải mã. Tập hợp các lớp cần được xuất ra được định nghĩa là tập hợp lớp đầu ra. Trong codec video mới nhất hỗ trợ nhiều lớp và các tính khả mở, một hoặc nhiều tập hợp lớp đầu ra được báo hiệu trong VPS. Các phần tử cú pháp đó xác định tập hợp lớp đầu ra và lệ thuộc của chúng, tiêu sử/tầng/mức độ và các tham số mô hình tham chiếu bộ giải mã giả định cần được báo hiệu hiệu quả trong tập hợp tham số.

Các phương án thực hiện sáng chế giải quyết một hoặc nhiều vấn đề của sáng chế.

Fig.1 minh họa sơ đồ khối được giản lược của hệ thống truyền thông (100) theo phương án thực hiện sáng chế. Hệ thống (100) có thể bao gồm ít nhất hai thiết bị đầu cuối (110, 120) được liên kết qua mạng (150). Để truyền dữ liệu đơn hướng, thiết bị đầu cuối thứ nhất (110) có thể mã hóa dữ liệu video ở vị trí cục bộ để truyền đến thiết bị đầu cuối khác (120) qua mạng (150). Thiết bị đầu cuối thứ hai (120) có thể nhận dữ liệu video được mã hóa của thiết bị đầu cuối khác từ mạng (150), giải mã dữ liệu được mã hóa và hiển thị dữ liệu video được khôi phục. Truyền dữ liệu đơn hướng có thể là chung trong các ứng dụng phục vụ đa phương tiện và tương tự.

Fig.1 minh họa cặp thứ hai của các thiết bị đầu cuối (130, 140) được cấp để hỗ trợ truyền hai hướng video được mã hóa mà có thể xuất hiện, chẳng hạn, trong suốt hội nghị video. Để truyền hai hướng dữ liệu, mỗi thiết bị đầu cuối (130, 140) có thể mã hóa dữ liệu video được chụp ở vị trí cục bộ để truyền đến thiết bị đầu cuối khác qua mạng (150). Mỗi thiết bị đầu cuối (130, 140) cũng có thể nhận dữ liệu video được mã hóa được truyền bởi thiết bị đầu cuối khác, có thể giải mã

dữ liệu được mã hóa và có thể hiển thị dữ liệu video được khôi phục ở thiết bị hiển thị cục bộ.

Trên Fig.1, các thiết bị đầu cuối (110-140) được minh họa dưới dạng máy tính xách tay 110, máy tính cá nhân (PC) 120 và các thiết bị đầu cuối di động 130 và 140, thiết bị đầu cuối (110-140) không bị giới hạn ở đó và thiết bị đầu cuối (110-140) có thể tương ứng với một hoặc nhiều hoặc tổ hợp bất kỳ của: các máy chủ, máy tính cá nhân, thiết bị di động, máy tính bảng và điện thoại thông minh. Các phương án thực hiện sáng chế tìm ra ứng dụng có máy tính cá nhân, máy tính bảng, máy phát đa phương tiện và/hoặc thiết bị hội nghị video chuyên dụng. Mạng (150) là số lượng mạng bất kỳ mà truyền tải dữ liệu video được mã hóa trong số các thiết bị đầu cuối (110-140), bao gồm chẳng hạn mạng truyền thông hữu tuyến và/hoặc không dây. Mạng truyền thông (150) có thể trao đổi dữ liệu trong các kênh chuyển mạch kênh và/hoặc chuyển mạch gói. Mạng đại diện bao gồm mạng viễn thông, mạng cục bộ (LAN), mạng diện rộng (WAN) và/hoặc mạng Internet. Theo sáng chế, kiến trúc và topology của mạng (150) có thể vô hình đối với hoạt động của sáng chế trừ khi được giải thích dưới đây.

Fig.2 minh họa, như là ví dụ cho ứng dụng theo phương án thực hiện sáng chế, bố trí bộ mã hóa video và bộ giải mã trong môi trường streaming. Sáng chế có thể áp dụng cân bằng cho các ứng dụng kích hoạt video khác, bao gồm, chẳng hạn, hội nghị video, TV số, lưu trữ video nén trên phương tiện số bao gồm CD, DVD, gậy nhớ, và tương tự, và v.v.

Hệ thống streaming có thể bao gồm hệ thống phụ chụp (213), có thể bao gồm nguồn video (201), chẳng hạn, camera số, được tạo cấu hình để tạo, chẳng hạn, dòng mẫu video không được nén (202). Dòng mẫu đó (202), được mô tả dưới dạng đường nét đậm trên Fig.2 để nhấn mạnh lưu lượng dữ liệu cao khi được so sánh với các dòng bit video được mã hóa, có thể được xử lý bằng bộ mã hóa (203) được ghép nối với camera (201). Bộ mã hóa (203) có thể bao gồm phần cứng, phần mềm, hoặc kết hợp của nó để có thể hoặc thực thi các khía cạnh của sáng chế như được mô tả chi tiết hơn dưới đây. Video được mã hóa dòng bit (204), được mô tả dưới dạng đường mảnh trên Fig.2 để nhấn mạnh lưu lượng dữ liệu thấp hơn khi được so sánh với dòng mẫu (202), có thể được lưu trữ trên máy

chủ streaming (205) để sử dụng sau này. Một hoặc nhiều máy khách streaming (206, 208) có thể truy nhập máy chủ streaming (205) để truy xuất các bản sao (207, 209) của video được mã hóa dòng bit (204). Máy khách (206) có thể bao gồm bộ giải mã video (210) mà giải mã bản sao đầu vào của video được mã hóa dòng bit (207) và tạo dòng mẫu video ra (211) có thể được kết xuất trên thiết bị hiển thị (212) hoặc thiết bị kết xuất khác. Trong một số hệ thống streaming, các dòng bit video (204, 207, 209) có thể được mã hóa theo các chuẩn nén/mã hóa video cụ thể. Các ví dụ về các chuẩn đó bao gồm ITU-T Recommendation H.265. Chuẩn mã hóa video đang được phát triển thường được biết đến là mã hóa video đa dụng hoặc VVC. Sáng chế có thể được sử dụng trong ngữ cảnh của VVC.

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

Bộ nhận (310) có thể nhận một hoặc nhiều chuỗi video codec cần được giải mã bằng bộ giải mã (210); theo phương án thực hiện tương tự hoặc khác, một chuỗi video được mã hóa tại một thời điểm, trong đó việc giải mã mỗi chuỗi video được mã hóa độc lập với các chuỗi video được mã hóa khác. Chuỗi video được mã hóa có thể được nhận từ kênh (312), mà có thể là liên kết phần cứng/phần mềm với thiết bị lưu trữ mà lưu trữ dữ liệu video được mã hóa. Bộ nhận (310) có thể nhận dữ liệu video được mã hóa với dữ liệu khác, chẳng hạn, dữ liệu audio được mã hóa và/hoặc các dòng dữ liệu phụ thuộc, mà có thể được chuyển tiếp đến các thiết bị tương ứng nhờ sử dụng các thực thể (không được mô tả trên hình vẽ). Bộ nhận (310) có thể tách riêng chuỗi video được mã hóa khỏi dữ liệu khác. Để khắc phục biến động mạng, bộ nhớ đệm (315) có thể được ghép nối giữa bộ nhận (310) và bộ giải mã entropy/bộ phân tách (320) (“bộ phân tách” dưới đây). Khi bộ nhận (310) được nhận dữ liệu từ bộ lưu trữ/chuyển tiếp của đủ băng thông và mức độ điều khiển, hoặc từ mạng đẳng hướng, bộ đệm (315) có thể không cần thiết, hoặc có thể nhỏ. Để sử dụng trên mạng gói nỗ lực tốt nhất chẳng hạn mạng Internet, bộ đệm (315) có thể được yêu cầu, có thể tương đối lớn và có thể có kích thước thích ứng theo cách có lợi.

Bộ giải mã video (210) có thể bao gồm bộ phân tách (320) để tái tạo các ký hiệu (321) từ chuỗi video được mã hóa entropy. Các phân loại của các ký hiệu

đó bao gồm thông tin được sử dụng để quản lý hoạt động của bộ giải mã (210), và thông tin có thể điều khiển thiết bị kết xuất, chẳng hạn, bộ phân hiển thị (212) vốn không phải là phân tích hợp của bộ giải mã nhưng có thể được ghép nối với nó, như được thể hiện trên Fig.2. Thông tin điều khiển cho (các) thiết bị kết xuất có thể ở dạng thông tin tăng cường bổ sung (SEI messages) hoặc các đoạn tập hợp tham số thông tin khả dụng video (VUI). Bộ phân tách (320) có thể phân tách/giải mã entropy chuỗi video được mã hóa được nhận. Việc mã hóa chuỗi video được mã hóa có thể tuân theo công nghệ hoặc chuẩn mã hóa video, và có thể tuân theo các nguyên lý đã quen thuộc với người có hiểu biết trung bình trong lĩnh vực, bao gồm mã hóa chiều dài biến thiên, mã hóa Huffman, mã hóa số học có hoặc không có độ nhảy ngữ cảnh, và v.v. Bộ phân tách (320) có thể trích xuất từ chuỗi video được mã hóa, tập hợp tham số nhóm phụ cho ít nhất một trong các nhóm phụ của các pixel trong bộ giải mã video, dựa trên ít nhất một tham số tương ứng với nhóm. Các nhóm phụ có thể bao gồm GOP, ảnh, ô, lát, khối lớn, đơn vị mã (CU), khối, đơn vị biến đổi (TU), đơn vị dự báo (PU) và v.v.. Bộ giải mã entropy/bộ phân tách cũng có thể trích xuất từ chuỗi video được mã hóa thông tin, chẳng hạn, hệ số biến đổi, giá trị tham số lượng tử (QP), MV, và v.v.

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

Việc tái tạo các ký hiệu (321) có thể bao gồm nhiều đơn vị khác nhau tùy thuộc loại ảnh video được mã hóa hoặc các phần của nó (chẳng hạn: trong và ngoài ảnh, trong và ngoài khối), và các hệ số khác. Những khối nào liên quan, và cách thức, có thể được điều khiển bằng thông tin điều khiển nhóm phụ đã được phân tách từ chuỗi video được mã hóa bằng bộ phân tách (320). Thông tin điều khiển nhóm phụ có thể chạy giữa bộ phân tách (320) và nhiều khối phụ.

Ngoài các khối chức năng đã được đề cập, bộ giải mã 210 có thể được phân chia phụ về mặt khái niệm thành số lượng khối chức năng như được mô tả dưới đây. Trong triển khai thực tế hoạt động theo các ràng buộc thương mại, nhiều khối này tương tác chặt chẽ với nhau và có thể, ít nhất một phần, được tích hợp vào nhau. Tuy nhiên, để mô tả sáng chế, phân chia phụ về mặt khái niệm thành các khối chức năng dưới đây là thích hợp.

Khối thứ nhất là bộ chia tỷ lệ/khối biến đổi ngược (351). Bộ chia tỷ lệ/khối biến đổi ngược (351) nhận hệ số biến đổi được lượng tử hóa cũng như thông tin điều khiển, bao gồm biến đổi nào để sử dụng, kích thước khối, hệ số lượng tử, các ma trận chia tỷ lệ lượng tử, v.v dưới dạng (các) ký hiệu (321) từ bộ phân tách (320). Có thể xuất ra các khối bao gồm các giá trị mẫu, có thể được đưa vào bộ kết tập (355).

Trong một số trường hợp, các mẫu đầu ra của bộ chia tỷ lệ/khối biến đổi ngược (351) có thể gắn với khối được mã hóa trong; tức là: khối không sử dụng thông tin dự báo từ các ảnh được tái tạo trước đó, nhưng có thể sử dụng thông tin dự báo từ các phần được tái tạo trước đó của ảnh hiện tại. Thông tin dự báo này có thể được cấp bởi khối dự báo trong ảnh (352). Trong một số trường hợp, khối dự báo trong ảnh (352) tạo khối có cùng kích thước và hình dạng của khối đang được tái tạo, sử dụng thông tin đã được tái tạo xung quanh được nạp từ ảnh hiện tại (được tái tạo một phần) (356). Trong một số trường hợp, bộ kết tập (355) bổ sung trên cơ sở theo mẫu, thông tin dự báo khối dự báo trong ảnh (352) đã tạo cho thông tin mẫu đầu ra như được tạo bởi bộ chia tỷ lệ/khối biến đổi ngược (351).

Trong các trường hợp khác, các mẫu đầu ra của bộ chia tỷ lệ/khối biến đổi ngược (351) có thể gắn với khối được mã hóa ngoài và bù chuyển động tiềm năng. Trong trường hợp này, khối dự báo bù chuyển động (353) có thể truy nhập bộ nhớ ảnh tham chiếu (357) để nạp các mẫu được sử dụng để dự báo. Sau khi bù chuyển động các mẫu được nạp theo các ký hiệu (321) gắn với khối, các mẫu này có thể được bổ sung bởi bộ kết tập (355) vào đầu ra của bộ chia tỷ lệ/khối biến đổi ngược (trong trường hợp này được gọi là các mẫu dư hoặc tín hiệu dư) để tạo thông tin mẫu đầu ra. Các địa chỉ trong ở dạng bộ nhớ ảnh tham chiếu trong đó khối bù chuyển động nạp các mẫu dự báo có thể được điều khiển bằng các vectơ chuyển động (MV), khả dụng cho khối bù chuyển động ở dạng các ký hiệu (321) mà có thể có, chẳng hạn X, Y, và các thành phần ảnh tham chiếu. Bù chuyển động cũng có thể bao gồm nội suy các giá trị mẫu như được nạp từ bộ nhớ ảnh tham chiếu khi các MV chính xác đến mẫu phụ được sử dụng, các cơ chế dự báo MV, và v.v.

Các mẫu đầu ra của bộ kết tập (355) có thể tùy thuộc vào các kỹ thuật lọc vòng khác nhau trong khối lọc vòng (356). Các công nghệ nén video có thể bao gồm các công nghệ lọc trong vòng được điều khiển bởi các tham số được bao gồm trong dòng bit video được mã hóa và trở nên khả dụng với khối lọc vòng (356) dưới dạng các ký hiệu (321) từ bộ phân tách (320), nhưng cũng có thể đáp ứng với siêu thông tin thu được trong khi giải mã các phần trước đó (theo thứ tự giải mã) của ảnh được mã hóa hoặc chuỗi video được mã hóa, cũng như đáp ứng với các giá trị mẫu được lọc vòng và được tái tạo trước đó.

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

Các ảnh được mã hóa cụ thể, khi được tái tạo hoàn chỉnh, có thể được sử dụng làm các ảnh tham chiếu để dự báo tương lai. Khi ảnh được mã hóa được tái tạo hoàn chỉnh và ảnh được mã hóa đã được nhận diện là ảnh tham chiếu (bằng, chẳng hạn, bộ phân tách (320)), ảnh tham chiếu hiện tại (356) có thể trở thành một phần của bộ đệm ảnh tham chiếu (357), và bộ nhớ ảnh mới hiện tại có thể được phân phối lại trước khi bắt đầu tái tạo tiếp sau ảnh được mã hóa.

Bộ giải mã video (bộ phân tách) 320 có thể thực hiện các hoạt động giải mã theo công nghệ nén video định trước mà có thể được dẫn chứng trong tài liệu theo chuẩn, chẳng hạn ITU-T Rec. H.265. Chuỗi video được mã hóa có thể tuân theo cú pháp được xác định bởi công nghệ hoặc chuẩn nén video đang được sử dụng, theo nghĩa là nó gắn với cú pháp này của công nghệ hoặc chuẩn nén video, như được xác định trong tài liệu hoặc chuẩn công nghệ nén video và cụ thể là trong tài liệu tiêu sử trong đó. Cần thiết cho việc tuân thủ cũng có thể là độ phức tạp của chuỗi video được mã hóa nằm trong các biên như được định nghĩa bởi cấp độ của công nghệ hoặc chuẩn nén video. Trong một số trường hợp, các cấp độ giới hạn kích thước ảnh lớn nhất, tốc độ khung hình lớn nhất, tốc độ mẫu tái tạo lớn nhất (được đo trong, chẳng hạn các mega mẫu trên giây), kích thước ảnh tham chiếu lớn nhất, và v.v. Các giới hạn được thiết lập bằng các cấp độ có thể, trong một số trường hợp, còn bị giới hạn qua các đặc tả bộ giải mã chuẩn giả

định (HRD) và siêu dữ liệu để quản lý bộ đệm HRD được báo hiệu trong chuỗi video được mã hóa.

Theo phương án thực hiện, bộ nhận (310) có thể nhận dữ liệu (dư) bổ sung có video được mã hóa. Dữ liệu bổ sung có thể được bao gồm dưới dạng một phần của (các) chuỗi video được mã hóa. Dữ liệu bổ sung có thể được sử dụng bởi bộ giải mã video (320) để giải mã đúng dữ liệu và/hoặc để tái tạo chính xác hơn dữ liệu video ban đầu. Dữ liệu bổ sung có thể ở dạng, chẳng hạn, thời gian, không gian, hoặc các lớp tăng cường SNR, các lát dư, ảnh dư, mã sửa lỗi thuận, và v.v.

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

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

Nguồn video (201) có thể tạo chuỗi video nguồn cần được mã hóa bằng bộ mã hóa (203) ở dạng dòng mẫu video số có thể có chiều 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 cấu trúc lấy mẫu thích hợp bất kỳ (chẳng hạn Y CrCb 4:2:0, Y CrCb 4:4:4). Trong hệ thống phục vụ đa phương tiện, nguồn video (201) có thể là thiết bị lưu trữ lưu trữ video đã được chuẩn bị trước. Trong hệ thống hội nghị video, nguồn video (203) có thể là camera chụp thông tin ảnh cục bộ dưới dạng chuỗi video. Dữ liệu video có thể được tạo dưới dạng các ảnh riêng rẽ truyền chuyển động khi được xem lần lượt. Chính các ảnh có thể được tổ chức dưới dạng mảng pixel không gian, trong đó mỗi pixel có thể bao gồm một hoặc nhiều mẫu tùy thuộc vào cấu trúc lấy mẫu, không gian màu, v.v đang được sử dụng. Người có hiểu biết trung bình trong lĩnh vực có thể dễ hiểu mối quan hệ giữa các pixel và các mẫu. Các phần mô tả dưới đây tập trung vào các mẫu.

Theo phương án thực hiện, bộ mã hóa (203) có thể mã hóa và nén các ảnh của chuỗi video nguồn thành chuỗi video được mã hóa (443) theo thời gian thực hoặc theo ràng buộc thời gian khác bất kỳ theo yêu cầu của sáng chế. Việc ép buộc tốc độ mã hóa thích hợp là một chức năng của bộ điều khiển (450). Bộ điều khiển (450) có thể điều khiển các khối chức năng khác như được mô tả dưới đây

và được ghép nối về mặt chức năng với các khối này. Các tham số được thiết lập bằng bộ điều khiển (450) có thể bao gồm các tham số liên quan điều khiển tốc độ (bỏ qua ảnh, lượng tử hóa, giá trị lambda của các kỹ thuật tối ưu hóa méo dạng tốc độ, ...), kích thước ảnh, bố trí GOP, khoảng tìm kiếm MV lớn nhất, và v.v.. Người có hiểu biết trung bình trong lĩnh vực có thể dễ nhận diện các chức năng khác của bộ điều khiển (450) do chúng có thể gắn với bộ mã hóa video (203) được tối ưu hóa cho thiết kế hệ thống cụ thể.

Một số bộ mã hóa video hoạt động trong người có hiểu biết trung bình trong lĩnh vực dễ ghi nhận là “vòng mã”. Như là mô tả được giản lược quá mức, vòng mã có thể gồm phần mã hóa của bộ mã hóa (203) (“bộ mã hóa nguồn” dưới đây) (chịu trách nhiệm tạo các ký hiệu dựa trên ảnh đầu vào cần được mã hóa, và (các) ảnh tham chiếu), và bộ giải mã (cục bộ) (433) được nhúng trong bộ mã hóa (203) mà tái tạo các ký hiệu để tạo dữ liệu mẫu bộ giải mã (từ xa) cũng sẽ tạo (do kỹ thuật nén bất kỳ giữa các ký hiệu và dòng bit video được mã hóa không tổn hao trong các công nghệ nén video được xem xét theo sáng chế). Dòng mẫu được tái tạo đó được đưa vào bộ nhớ ảnh tham chiếu (434). Do việc giải mã dòng ký hiệu dẫn đến các kết quả chính xác theo bit độc lập với vị trí bộ giải mã (cục bộ hoặc từ xa), nội dung bộ đệm ảnh tham chiếu cũng chính xác theo bit giữa bộ mã hóa cục bộ và bộ mã hóa từ xa. Nói cách khác, phần dự báo của bộ mã hóa “xem” như là các mẫu ảnh tham chiếu giống hệt các giá trị mẫu như bộ giải mã sẽ “xem” khi sử dụng dự báo trong khi giải mã. Nguyên lý cơ bản của đồng bộ ảnh tham chiếu (và dẫn đến xê dịch, nếu đồng bộ không thể được duy trì, chẳng hạn do các lỗi kênh) quen thuộc với người có hiểu biết trung bình trong lĩnh vực.

Hoạt động của bộ giải mã “cục bộ” (433) có thể giống như của bộ giải mã “từ xa” (210), mà đã được mô tả chi tiết nêu trên cùng với Fig.3. Cũng dựa ngắn gọn vào Fig 3, tuy nhiên, do các ký hiệu là khả dụng và các ký hiệu mã hóa/giải mã đến chuỗi video được mã hóa bởi bộ mã hóa entropy (445) và bộ phân tách (320) có thể là không tổn hao, các phần giải mã entropy của bộ giải mã (210), bao gồm kênh (312), bộ nhận (310), bộ đệm (315), và bộ phân tách (320) có thể không được triển khai đầy đủ trong bộ giải mã cục bộ (433):

Quan sát mà có thể được thực hiện ở điểm này chính là công nghệ giải mã bất kỳ ngoại trừ phân tách/giải mã entropy xuất hiện trong bộ giải mã cũng nhất thiết cần xuất hiện, ở dạng chức năng gần như tương tự, trong bộ mã hóa tương ứng. Do vậy, sáng chế tập trung vào hoạt động bộ giải mã. Phần mô tả của các công nghệ mã hóa có thể được viết tắt dưới dạng chúng là đảo của các công nghệ giải mã được mô tả đầy đủ. Chỉ trong các khu vực cụ thể, mô tả chi tiết hơn được yêu cầu và được nêu dưới đây.

Như là một phần của hoạt động, bộ mã hóa nguồn (203) có thể thực hiện mã hóa dự báo bù chuyển động, mã hóa dự báo khung đầu vào dựa vào một hoặc nhiều khung được mã hóa trước đó từ chuỗi video được chọn làm “các khung tham chiếu.” Theo cách này, máy mã hóa (432) mã hóa các hiệu số giữa các khối pixel của khung đầu vào và các khối pixel của (các) khung tham chiếu mà có thể được chọn làm (các) tham chiếu dự báo đến khung đầu vào.

Bộ giải mã video cục bộ (433) có thể giải mã dữ liệu video được mã hóa của các khung mà có thể được chọn làm các khung tham chiếu, dựa trên các ký hiệu được tạo bởi bộ mã hóa nguồn (430). Các hoạt động của máy mã hóa (432) có thể một cách có lợi là các tiến trình tổn hao. Khi dữ liệu video được mã hóa có thể được giải mã ở bộ giải mã video (không được thể hiện trên Fig.4), chuỗi video được tái tạo thường có thể là bản sao của chuỗi video nguồn với một vài sai số. Bộ giải mã video cục bộ (433) sao chép các quá trình giải mã mà có thể được thực hiện bằng bộ giải mã video trên các khung tham chiếu và có thể khiến các khung tham chiếu được tái tạo được lưu trữ trong cache ảnh tham chiếu (434). Theo cách này, bộ mã hóa (203) có thể lưu trữ các bản sao của các khung tham chiếu được tái tạo cục bộ mà có chung nội dung với các khung tham chiếu được tái tạo mà sẽ thu được bởi bộ giải mã video đầu xa (mà không có các lỗi truyền).

Bộ dự báo (435) có thể thực hiện các tìm kiếm dự báo cho máy mã hóa (432). Tức 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 pixel tham chiếu ứng viên) hoặc metadata cụ thể chẳng hạn các MV ảnh tham chiếu, các hình dạng khối, và v.v, mà có thể dùng làm tham chiếu dự báo thích hợp cho các ảnh mới. Bộ dự báo (435) có thể hoạt động trên cơ sở khối mẫu x khối pixel để

tìm các tham chiếu dự báo thích hợp. Trong một số trường hợp, như được xác định bởi các 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 của các tham số và các tham số nhóm phụ được sử dụng để mã hóa dữ liệu video.

Đầu ra của các khối chức năng nêu trên có thể được mã hóa entropy trong bộ mã hóa entropy (445). Bộ mã hóa entropy biên dịch các ký hiệu như được tạo bởi các khối chức năng khác nhau thành chuỗi video được mã hóa, bằng cách nén không tổn hao các ký hiệu theo các công nghệ quen thuộc với người có hiểu biết trung bình trong lĩnh vực như là, chẳng hạn mã hóa Huffman, mã hóa chiều dài biến thiên, mã hóa số học, và v.v.

Bộ truyền (440) có thể đệm (các) chuỗi video được mã hóa như được tạo bởi bộ mã hóa entropy (445) để chuẩn bị truyền qua kênh truyền thông (460), mà có thể là liên kết phần cứng/phần mềm với thiết bị lưu trữ mà sẽ lưu trữ dữ liệu video được mã hóa. Bộ truyền (440) có thể hợp nhất dữ liệu video được mã hóa từ bộ mã hóa video (430) với dữ liệu khác cần được truyền, chẳng hạn, dữ liệu audio được mã hóa và/hoặc các dòng dữ liệu phụ thuộc/nguồn.

Bộ điều khiển (450) có thể quản lý hoạt động của bộ mã hóa (203). Trong khi mã hóa, bộ điều khiển (450) có thể gán cho mỗi ảnh được mã hóa loại ảnh được mã hóa cụ thể, mà có thể ảnh hưởng các kỹ thuật mã hóa mà có thể được áp dụng cho ảnh tương ứng. Chẳng hạn, các ảnh thường có thể được gán dưới dạng một trong các loại khung sau:

Nội ảnh (ảnh I) có thể là ảnh mà có thể được mã hóa và giải mã không sử dụng khung khác bất kỳ trong chuỗi làm nguồn dự báo. Một số codec video cho phép các loại nội ảnh khác nhau, bao gồm, chẳng hạn, ảnh làm mới bộ giải mã độc lập (IDR). Người có hiểu biết trung bình trong lĩnh vực nhận biết về các biến thể khác nhau của các ảnh I và các ứng dụng và các dấu hiệu tương ứng của nó.

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

Ảnh dự báo hai hướng (ảnh B) có thể là ảnh mà có thể được mã hóa và giải mã sử dụng dự báo trong hoặc dự báo ngoài nhờ sử dụng nhiều nhất hai MV và các 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 metadata liên kết để tái tạo một khối.

Các ảnh nguồn thường có thể được phân chia phụ không gian thành các khối mẫu (chẳng hạn, các khối của mỗi mẫu 4x4, 8x8, 4x8, hoặc 16x16) và được mã hóa dựa trên từng khối. Các khối có thể được mã hóa dự báo dựa vào các khối khác (đã được mã hóa) như được xác định bởi phép gán mã hóa được áp dụng cho các ảnh tương ứng của các khối. Chẳng hạn, các khối của các ảnh I có thể được mã hóa không dự báo hoặc chúng có thể được mã hóa dự báo dựa vào các khối đã được mã hóa của cùng ảnh (dự báo không gian hoặc dự báo trong). Các khối pixel của các ảnh P có thể được mã hóa không dự báo, qua dự báo không gian hoặc qua dự báo thời gian dựa vào một ảnh tham chiếu được mã hóa trước đó. Các khối của các ảnh B có thể được mã hóa không dự báo, qua dự báo không gian hoặc qua dự báo thời gian dựa vào một hoặc hai ảnh tham chiếu được mã hóa trước đó.

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

Theo phương án thực hiện, bộ truyền (440) có thể truyền dữ liệu bổ sung có video được mã hóa. Bộ mã hóa video (430) có thể bao gồm dữ liệu chẳng hạn một phần của chuỗi video được mã hóa. Dữ liệu bổ sung có thể bao gồm thời

gian/không gian/các lớp tăng cường SN, các dạng khác của dữ liệu dư chẳng hạn các ảnh và lát dư, các thông điệp SEI, các đoạn tập hợp tham số VUI, và v.v.

Trước khi mô tả chi tiết hơn các khía cạnh cụ thể của sáng chế, vài cụm từ cần được giới thiệu mà sẽ được nêu trong phần còn lại của bản mô tả.

Ảnh con dưới đây đề cập đến, trong một số trường hợp, nhóm mẫu chữ nhật, các khối, các khối lớn, các CU, hoặc các thực thể tương tự được nhóm về mặt ngữ nghĩa, và mà có thể được mã hóa độc lập trong độ phân giải được thay đổi. Một hoặc nhiều ảnh con có thể dành cho một ảnh. Một hoặc nhiều ảnh con được mã hóa có thể tạo ảnh được mã hóa. Một hoặc nhiều ảnh con có thể được lắp ráp thành ảnh, và một hoặc nhiều ảnh con có thể được trích xuất từ ảnh. Trong các môi trường cụ thể, một hoặc nhiều ảnh con được mã hóa có thể được lắp ráp trong miền được nén mà không chuyển mã thành mức mẫu thành ảnh được mã hóa, và trong các trường hợp tương tự hoặc cụ thể khác, một hoặc nhiều ảnh con được mã hóa có thể được trích xuất từ ảnh được mã hóa trong miền được nén.

Thay đổi độ phân giải thích ứng (ARC) dưới đây đề cập đến các cơ chế mà cho phép thay đổi độ phân giải của ảnh hoặc ảnh con trong chuỗi video được mã hóa, thông qua, chẳng hạn, lấy mẫu lại ảnh tham chiếu. Các tham số ARC dưới đây đề cập đến thông tin điều khiển được yêu cầu để thực hiện ARC, mà có thể bao gồm, chẳng hạn, các tham số lọc, 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 nêu trên tập trung vào mã hóa và giải mã ảnh video đơn được mã hóa độc lập về ngữ nghĩa theo các phương án thực hiện khác nhau. Trước khi mô tả ngụ ý của mã hóa/giải mã nhiều ảnh con với các tham số ARC độc lập và độ phức tạp bổ sung được ngụ ý của nó, các tùy chọn để báo hiệu các tham số ARC có thể được mô tả.

Fig.5 thể hiện vài tùy chọn mới để báo hiệu các tham số ARC. Như được lưu ý với mỗi tùy chọn trong các tùy chọn, chúng có các ưu điểm và các nhược điểm cụ thể từ khía cạnh hiệu suất mã hóa, độ phức tạp, và kiến trúc. chuẩn hoặc công nghệ mã hóa video có thể chọn một hoặc nhiều trong số các tùy chọn này, hoặc các tùy chọn được biết theo giải pháp kỹ thuật đã biết, để báo hiệu các tham số ARC. Các tùy chọn này có thể không loại trừ lẫn nhau, và hiểu được rằng là có

thể trao đổi dựa trên các nhu cầu ứng dụng, các chuẩn công nghệ liên quan, hoặc lựa chọn của bộ mã hóa.

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

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

- các hệ số tăng kích thước mẫu và/hoặc giảm kích thước mẫu, có bổ sung chiều thời gian, chỉ báo thu/phòng có tốc độ không đổi đối với số lượng cụ thể của các ảnh

- Một trong hai lớp nêu trên có thể bao gồm mã hóa một hoặc nhiều phần tử cú pháp giả thiết là ngắn mà có thể trở vào bảng chứa (các) hệ số.

- Độ phân giải, theo chiều X hoặc Y, theo các đơn vị mẫu, các khối, các khối lớn, CU, hoặc độ hạt thích hợp khác bất kỳ, 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 (chẳng hạn, một cho ảnh đầu vào, một cho ảnh tham chiếu) thì, trong các trường hợp cụ thể, một tập hợp giá trị có thể được suy ra từ tập hợp giá trị khác. Điều này có thể bị giới hạn, chẳng hạn, nhờ sử dụng cờ. Để hiểu ví dụ chi tiết hơn, xem dưới đây.

- Các tọa độ “cong” giống như các tọa độ được sử dụng trong H.263 Annex P, lại ở độ hạt thích hợp như nêu trên. H.263 Annex P định nghĩa một cách hiệu quả để mã hóa các tọa độ cong này, nhưng tọa độ khác, các cách hiệu quả hơn có thể cũng được tạo theo cách nhận thức được. Chẳng hạn, mã hóa kiểu “Huffman” chiều dài biến thiên đảo được của các tọa độ cong của Annex P có thể được thay thế bằng mã hóa nhị phân chiều dài thích hợp, trong đó chiều dài của từ mã nhị phân có thể, chẳng hạn, được dẫn xuất từ kích thước ảnh lớn nhất, có thể được nhân với hệ số cụ thể và được lệch bởi giá trị cụ thể, để cho phép “cong” nằm ngoài các đường biên của kích thước ảnh lớn nhất.

- tăng kích thước mẫu hoặc giảm kích thước mẫu các tham số lọc. Trong trường hợp dễ nhất, có thể chỉ có một bộ lọc để tăng kích thước mẫu và/hoặc giảm kích thước mẫu. Tuy nhiên, trong các trường hợp cụ thể, một cách có lợi có thể cho phép linh hoạt hơn trong thiết kế bộ lọc, và mà có thể yêu cầu báo hiệu các tham số lọc. Các tham số này có thể được chọn qua chỉ số trong danh

sách các thiết kế lộ khả thi, bộ lọc có thể được xác định đầy đủ (chẳng hạn, qua danh sách các hệ số lọc, nhờ sử dụng các kỹ thuật mã hóa entropy thích hợp), bộ lọc có thể được chọn ngầm qua các tỷ lệ tăng và/hoặc giảm kích thước mẫu theo đó được báo hiệu theo cơ chế bất kỳ trong các cơ chế nêu trên, và v.v.

Dưới đây, phần mô tả giả sử mã hóa tập hữu hạn các hệ số tăng kích thước mẫu và/hoặc giảm kích thước mẫu (hệ số tương tự được sử dụng trong cả chiều X và Y), được chỉ báo qua từ mã. Từ mã đó có thể được mã hóa có chiều dài biến thiên theo cách có lợi, chẳng hạn sử dụng mã Ext-Golomb chung cho các phần tử cú pháp cụ thể trong các đặc tả mã hóa video, chẳng hạn, H.264 và H.265.

Một phép ánh xạ thích hợp các giá trị đến các hệ số tăng kích thước mẫu và/hoặc giảm kích thước mẫu có thể, chẳng hạn, theo bảng sau.

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

Bảng 1

Nhiều phép ánh xạ giống nhau có thể được tạo theo các nhu cầu của ứng dụng và các khả năng của các cơ chế tăng và giảm tỷ lệ khả dụng trong công nghệ hoặc chuẩn nén video. Bảng có thể được mở rộng sang nhiều giá trị hơn. Các giá trị cũng có thể được biểu diễn bằng các cơ chế mã hóa entropy khác ngoài các mã Ext-Golomb, chẳng hạn sử dụng mã hóa nhị phân. Có thể có các ưu điểm cụ thể khi các hệ số lấy mẫu lại liên quan nằm ngoài chính các máy xử lý video (bộ mã hóa và bộ giải mã trước tiên), chẳng hạn bằng các MANET (mạng ad hoc di động). Cần lưu ý rằng, đối với trường hợp phổ biến nhất (giả sử) trong đó không yêu cầu thay đổi độ phân giải, mã Ext-Golomb có thể được chọn là ngắn; trong bảng nêu trên, chỉ bit đơn. Nó có thể có ưu thế hiệu quả mã hóa so với nhờ sử dụng các mã nhị phân cho trường hợp phổ biến nhất.

Số lượng danh mục trong bảng, cũng như ngữ nghĩa của nó có thể được tạo cấu hình đầy đủ hoặc một phần. Chẳng hạn, phác thảo cơ sở của bảng có thể được truyền trong tập hợp tham số “cao”, chẳng hạn, tập hợp tham số chuỗi hoặc bộ giải mã. Theo cách khác hoặc ngoài ra, một hoặc nhiều bảng này có thể được

định nghĩa trong công nghệ hoặc chuẩn mã hóa video, và có thể được chọn qua, chẳng hạn, tập hợp tham số bộ giải mã hoặc chuỗi.

Dưới đây, mô tả cách thức hệ số tăng kích thước mẫu và/hoặc giảm kích thước mẫu (thông tin ARC), được mã hóa như nêu trên, có thể được bao gồm trong cú pháp công nghệ hoặc chuẩn mã hóa video. Các xem xét tương tự có thể áp dụng cho một, hoặc vài từ mã điều khiển các bộ lọc tăng và/hoặc giảm kích thước mẫu. Xem phần mô tả dưới đây khi 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 ở dạng bốn tọa độ cong vào tiêu đề ảnh 501, cụ thể là trong phần mở rộng tiêu đề H.263 PLUSPTYPE (503). Đây có thể là lựa chọn thiết kế dễ hiểu khi a) có sẵn tiêu đề ảnh, và b) mong đợi các thay đổi thường xuyên của thông tin ARC. Tuy nhiên, phụ tải khi sử dụng báo hiệu kiểu H.263 có thể khá cao, và hệ số tỷ lệ không thể duy trì giữa các đường biên ảnh do tiêu đề ảnh có thể khả biến.

JVCET-M135-v1, được nêu trên, bao gồm thông tin tham số ARC (505) (chỉ số) được đặt trong PPS (504), đánh chỉ số bảng (506) bao gồm các độ phân giải mục tiêu đến lượt được đặt bên trong SPS (507). Việc bố trí độ phân giải khả thi trong bảng (506) trong tập tham số chuỗi (507) có thể, theo mô tả bằng lời của tác giả, được điều chỉnh nhờ sử dụng SPS làm điểm thỏa hiệp liên tác trong khi trao đổi khả năng. Độ phân giải có thể thay đổi, trong các giới hạn được thiết lập bởi các giá trị trong bảng (506) từ ảnh đến ảnh bằng cách tạo tham chiếu PPS thích hợp (504).

Vấn đề cập đến Fig.5, các tùy chọn bổ sung sau có thể có để truyền thông tin ARC trong dòng bit video. Mỗi tùy chọn trong các tùy chọn này có các ưu điểm cụ thể so với giải pháp kỹ thuật đã biết như nêu trên. Các tùy chọn này có thể đồng thời xuất hiện trong chuẩn hoặc công nghệ mã hóa video tương tự.

Theo phương án thực hiện, thông tin ARC (509), chẳng hạn, hệ số lấy mẫu lại (thu phóng) có thể xuất hiện trong tiêu đề lát, tiêu đề nhóm khối (GOB), tiêu đề ô ảnh, hoặc tiêu đề nhóm ô (tiêu đề nhóm ô dưới đây) (508). Có thể đầy đủ nếu thông tin ARC vốn nhỏ chẳng hạn, từ mã có chiều dài biến thiên đơn vị (v) hoặc từ mã chiều dài cố định có vài bit, chẳng hạn, như được thể hiện trên đây.

Việc có thông tin ARC trong tiêu đề nhóm ô có ngay lợi thế bổ sung thông tin ARC có thể áp dụng cho ảnh con được biểu diễn bằng, chẳng hạn, nhóm ô đó, thay vì toàn bộ ảnh. Cũng xem dưới đây. Ngoài ra, thậm chí nếu công nghệ hoặc chuẩn nén video đề xuất chỉ các ARC toàn bộ ảnh (ngược với, chẳng hạn, các ARC dựa trên nhóm ô), đặt thông tin ARC vào tiêu đề nhóm ô tương tự việc đặt nó vào tiêu đề ảnh kiểu H.263 có các ưu điểm cụ thể từ quan điểm khả năng khôi phục lỗi.

Theo phương án thực hiện tương tự hoặc khác, chính thông tin ARC (512) có thể xuất hiện trong tập hợp tham số thích hợp (511) chẳng hạn, PPS, tập hợp tham số tiêu đề, tập hợp tham số ô ảnh, tập hợp tham số thích ứng, và v.v (Tập hợp tham số thích ứng được mô tả). Phạm vi của tập hợp tham số đó một cách có lợi có thể không lớn hơn ảnh, chẳng hạn nhóm ô ảnh. Việc sử dụng thông tin ARC là ngầm thông qua kích hoạt tập hợp tham số liên quan. Chẳng hạn, khi công nghệ hoặc chuẩn mã hóa video xem xét chỉ ARC dựa trên ảnh, sau đó PPS, hoặc tương đương có thể là thích hợp.

Theo phương án thực hiện tương tự hoặc khác, thông tin tham chiếu ARC (513) có thể xuất hiện trong tiêu đề nhóm ô (514) hoặc cấu trúc dữ liệu tương tự. Thông tin tham chiếu đó (513) có thể đề cập đến tập con của thông tin ARC (515) khả dụng trong tập hợp tham số (516) với phạm vi vượt ngoài ảnh đơn, chẳng hạn SPS, hoặc tập hợp tham số bộ giải mã.

Mức độ bổ sung kích hoạt ngầm vô hướng của PPS từ tiêu đề nhóm ô, PPS, SPS, như được sử dụng trong JVET-M0135-v1 dường như là không cần thiết, như là các PPS, đúng như các SPS, có thể (và có trong các chuẩn cụ thể, chẳng hạn, RFC3984) được sử dụng để thương lượng khả năng hoặc thông báo. Tuy nhiên, nếu thông tin ARC có thể áp dụng cho ảnh con được biểu diễn, chẳng hạn, cũng bằng các nhóm ô ảnh, tập hợp tham số có phạm vi kích hoạt bị giới hạn ở nhóm ô ảnh, chẳng hạn tập hợp tham số thích ứng hoặc tập hợp tham số tiêu đề có thể là lựa chọn tốt hơn. Nếu thông tin ARC cũng có kích thước lớn hơn không đáng kể — chẳng hạn, chứa thông tin điều khiển lọc chẳng hạn nhiều hệ số lọc — thì tập hợp tham số có thể là lựa chọn tốt hơn sử dụng tiêu đề (508) trực tiếp từ quan điểm hiệu suất mã hóa, do các thiết lập đó có thể có thể sử dụng lại bằng

các ảnh hoặc các ảnh con sau này bằng cách tạo tham chiếu cùng tập hợp tham số.

Khi sử dụng SPS hoặc tập hợp tham số cao hơn khác có phạm vi trải qua nhiều ảnh, các xem xét cụ thể có thể áp dụng:

1. Tập hợp tham số để lưu trữ bảng thông tin ARC (516) có thể, trong một số trường hợp, là SPS, nhưng trong các trường hợp khác một cách có lợi là tập hợp tham số bộ giải mã. Tập hợp tham số bộ giải mã có thể có phạm vi kích hoạt nhiều CVS, tức là dòng video được mã hóa, tức là, tất cả các bit video được mã hóa từ lúc bắt đầu phiên đến khi kết thúc phiên. Phạm vi này có thể thích hợp hơn do các hệ số ARC khả thi có thể là đặc điểm bộ giải mã, có thể được triển khai trong phần cứng, và các đặc điểm phần cứng sẽ không thay đổi với CVS bất kỳ (mà trong ít nhất vài hệ thống giải trí là GOP, dài một giây hoặc nhỏ hơn). Tức là, đặt bảng vào SPS được bao gồm rõ ràng trong các tùy chọn bố trí được mô tả ở đây, cụ thể cùng với điểm 2 dưới đây.

2. Thông tin tham số ARC (513) có thể một cách có lợi được đặt trực tiếp vào ảnh/lát/ô/GOB/tiêu đề nhóm ô (tiêu đề nhóm ô dưới đây) (514) thay vì vào PPS như trong JVCET-M0135-v1, lý do như sau: khi bộ mã hóa muốn thay đổi giá trị đơn trong PPS, chẳng hạn, thông tin tham số ARC, sau đó, nó phải tạo PPS mới và tạo tham chiếu PPS mới đó. Giả sử rằng chỉ thông tin tham số ARC thay đổi, nhưng thông tin khác, chẳng hạn, thông tin ma trận lượng tử trong PPS duy trì. Thông tin này có thể có kích thước đáng kể, và sẽ cần được truyền lại để tạo PPS mới hoàn toàn.

Do thông tin tham số ARC có thể là từ mã đơn, chẳng hạn chỉ số vào bảng (513) và đó là giá trị duy nhất mà thay đổi, sẽ là công kênh và lãng phí để truyền lại tất cả, chẳng hạn, thông tin ma trận lượng tử. Trong phạm vi, có thể tốt hơn đáng kể từ hiệu suất mã hóa để tránh vô hướng qua PPS, như được đề cập trong JVCET-M0135-v1. Tương tự, đặt thông tin tham số ARC vào PPS có nhược điểm bổ sung là thông tin ARC được tạo tham chiếu bởi thông tin tham số ARC (513) bắt buộc cần áp dụng cho toàn ảnh và không phải cho ảnh con, do phạm vi kích hoạt PPS là ảnh.

Theo phương án thực hiện tương tự hoặc khác, báo hiệu các tham số ARC có thể tuân theo ví dụ chi tiết như được phác thảo trên Fig.6. Fig.6 mô tả các sơ đồ cú pháp trong biểu diễn như được sử dụng trong các chuẩn mã hóa video từ ít nhất 1993. Ký pháp của các sơ đồ cú pháp này gần như tuân theo lập trình kiểu C. Trên Fig.6, các đường in đậm chỉ báo các phần tử cú pháp xuất hiện trong dòng bit, các đường không đậm thường chỉ báo luồng điều khiển hoặc thiết lập của các biến.

Tiêu đề nhóm ô (601) như là cấu trúc cú pháp lấy làm ví dụ của tiêu đề có thể áp dụng cho phần ảnh (có thể là chữ nhật) có thể chứa một cách có điều kiện, chiều dài biến thiên, phần tử cú pháp được mã hóa Exp-Golomb `dec_pic_size_idx` (602) (được mô tả in đậm). Sự xuất hiện của phần tử cú pháp này trong tiêu đề nhóm ô có thể bị giới hạn ở việc sử dụng độ phân giải thích ứng (603) — ở đây, giá trị của cờ không được mô tả trên hình vẽ in đậm, nghĩa là cờ đó xuất hiện trong dòng bit ở điểm trong đó nó xuất hiện trong sơ đồ cú pháp này.

Liệu độ phân giải thích ứng có đang sử dụng cho ảnh này hoặc các phần khác của ảnh hay không có thể được báo hiệu trong cấu trúc cú pháp mức cao bất kỳ bên trong hoặc ngoài dòng bit. Trong ví dụ được thể hiện trên Fig.6, được báo hiệu trong SPS như được phác thảo dưới đây.

Vẫn dựa vào Fig.6, cũng thể hiện trích đoạn của SPS (610). Phần tử cú pháp thứ nhất được thể hiện là `adaptive_pic_resolution_change_flag` (611). Nếu đúng, cờ đó có thể chỉ báo sử dụng độ phân giải thích ứng mà sau đó có thể yêu cầu thông tin điều khiển cụ thể. Trong ví dụ này, thông tin điều khiển này xuất hiện có điều kiện dựa trên giá trị của cờ dựa trên mệnh đề `if()` trong tập hợp tham số (612) và tiêu đề nhóm ô (601).

Khi độ phân giải thích ứng đang sử dụng, trong ví dụ này, độ phân giải đầu ra được mã hóa theo các đơn vị mẫu (613). Số chỉ dẫn 613 đề cập đến cả `output_pic_width_in_luma_samples` và `output_pic_height_in_luma_samples`, đồng thời có thể xác định độ phân giải của ảnh đầu ra. Đầu đó trong công nghệ hoặc chuẩn mã hóa video, các giới hạn cụ thể ở giá trị bất kỳ có thể được định nghĩa. Chẳng hạn, xác định mức có thể giới hạn số lượng tổng số mẫu đầu ra, có

thể là tích số của giá trị của hai phần tử cú pháp đó. Các công nghệ hoặc các chuẩn mã hóa video cụ thể, hoặc các công nghệ hoặc các chuẩn ngoài, chẳng hạn, các chuẩn hệ thống, cũng có thể giới hạn khoảng đánh số (chẳng hạn, một hoặc cả hai chiều phải chia hết cho số lũy thừa 2), hoặc khuôn dạng (chẳng hạn, chiều rộng và chiều cao phải theo quan hệ, chẳng hạn 4:3 hoặc 16:9). Các giới hạn này có thể được đưa vào để tạo thuận tiện lắp đặt phần cứng hoặc vì các nguyên nhân khác, và đã quen thuộc theo giải pháp kỹ thuật.

Trong các ứng dụng cụ thể, bộ mã hóa có thể ra lệnh bộ giải mã sử dụng kích thước ảnh tham chiếu cụ thể thay vì ngầm giả sử kích thước là kích thước ảnh đầu ra. Trong ví dụ này, phần tử cú pháp này `reference_pic_size_present_flag` (614) giới hạn sự xuất hiện có điều kiện của các chiều ảnh tham chiếu (615) (số chỉ dẫn lại chỉ cả chiều rộng lẫn chiều cao).

Cuối cùng thể hiện trên Fig.6 là bảng chiều rộng và chiều cao ảnh giải mã khả thi. Bảng này có thể được biểu diễn, chẳng hạn, bằng ký hiệu chỉ báo bảng (`num_dec_pic_size_in_luma_samples_minus1`) (616). “minus1” có thể hiểu là giá trị của phần tử cú pháp đó. Chẳng hạn, nếu giá trị được mã hóa bằng 0, một danh mục bảng xuất hiện. Nếu giá trị này bằng 5, bảng có sáu mục. Đối với mỗi “dòng” trong bảng, thì chiều rộng và chiều cao của ảnh được giải mã được bao gồm trong cú pháp (617).

Các mục bảng xuất hiện (617) có thể được đánh chỉ số nhờ 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 quả, các hệ số thu phóng — theo nhóm ô.

Các công nghệ hoặc các chuẩn mã hóa video cụ thể, chẳng hạn VP9, hỗ trợ tính khả mở không gian bằng cách thực hiện các dạng cụ thể lấy mẫu lại ảnh tham chiếu (được báo hiệu tương đối khác với sáng chế) cùng với tính khả mở thời gian, để kích hoạt tính khả mở không gian. Cụ thể là, các ảnh tham chiếu cụ thể có thể được tăng kích thước mẫu nhờ sử dụng các công nghệ kiểu ARC đến độ phân giải cao hơn để tạo cơ sở của lớp tăng cường không gian. Các ảnh được tăng kích thước mẫu có thể được tinh lọc, nhờ sử dụng các cơ chế dự báo thông thường ở độ phân giải cao, để bổ sung chi tiết.

Sáng chế có thể được sử dụng trong môi trường này. Trong các trường hợp cụ thể, theo phương án thực hiện tương tự hoặc khác, giá trị trong tiêu đề khối 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. Làm vậy có các ưu điểm cụ thể cho các thiết kế hệ thống cụ thể; chẳng hạn, các khối chuyển tiếp được chọn hiện có (SFU) được tạo và tối ưu hóa để chuyển tiếp theo lớp thời gian được chọn dựa trên giá trị ID thời gian tiêu đề khối NAL có thể được sử dụng mà không chỉnh sửa, cho các môi trường khả mở. Để kích hoạt điều đó, có thể có yêu cầu đối với ánh xạ giữa kích thước ảnh được mã hóa và lớp thời gian được chỉ báo bởi trường ID thời gian trong tiêu đề khối NAL.

Trong một số công nghệ mã hóa video, khối truy nhập (AU) có thể chỉ (các) ảnh được mã hóa, (các) lát, (các) ô ảnh, (các) khối NAL, và v.v, mà được chụp và được tổng hợp thành dòng bit ảnh tương ứng/lát/ô/khối NAL ở thực thể thời gian cụ thể. Thực thể thời gian đó có thể là thời gian tổng hợp.

Trong chuẩn HEVC, và các công nghệ mã hóa video khác cụ thể, giá trị đếm thứ tự ảnh (POC) có thể được sử dụng để chỉ báo ảnh tham chiếu được chọn trong số nhiều ảnh tham chiếu được lưu trữ trong bộ đệm ảnh được giải mã (DPB). Khi AU bao gồm một hoặc nhiều ảnh, lát, hoặc ô, mỗi ảnh, lát, hoặc ô thuộc cùng AU có thể mang cùng giá trị POC, mà từ đó nó có thể được suy ra rằng chúng được tạo từ nội dung của cùng thời gian tổng hợp. Nói cách khác, trong kịch bản trong đó hai ảnh/lát/ô mang cùng giá trị POC cụ thể, mà có thể chỉ báo hai ảnh/lát/ô thuộc cùng AU và có cùng thời gian tổng hợp. Ngược lại, hai ảnh/ô ảnh/lát có các giá trị POC khác nhau có thể chỉ báo các ảnh/các lát/các ô ảnh đó thuộc các AU khác nhau và có các thời gian tổng hợp khác nhau.

Theo phương án thực hiện sáng chế, mối quan hệ rắn chắc nêu trên có thể được nói lỏng trong đó AU có thể bao gồm ảnh, lát, hoặc ô có các giá trị POC khác nhau. Bằng cách cho phép các giá trị POC khác nhau trong AU, trở nên có thể sử dụng giá trị POC để nhận diện các ảnh/các lát/các ô có thể giải mã độc lập có thời gian hiển thị giống hệt. Sau đó, có thể hỗ trợ nhiều lớp khả mở mà không thay đổi báo hiệu lựa chọn ảnh tham chiếu (chẳng hạn, báo hiệu tập hợp ảnh

tham chiếu hoặc báo hiệu danh sách ảnh tham chiếu), như được mô tả chi tiết hơn dưới đây.

Tuy nhiên, vẫn mong muốn có thể nhận diện AU mà có ảnh/lát/ô, dựa vào các ảnh/lát/ô khác có các giá trị POC khác nhau, từ riêng giá trị POC. Điều này có thể đạt được, như được mô tả dưới đây.

Theo các phương án thực hiện giống hết hoặc tương tự, số đếm AU (AUC) có thể được báo hiệu trong cấu trúc cú pháp mức cao, chẳng hạn, Tiêu đề khối NAL, tiêu đề lát, tiêu đề nhóm ô, Thông điệp SEI, tập hợp tham số hoặc dấu tách AU. Giá trị của AUC có thể được sử dụng để nhận diện các khối NAL, ảnh, lát hoặc ô nào thuộc AU cụ thể. Giá trị của AUC có thể tương ứng với thực thể thời gian tổng hợp riêng biệt. Giá trị AUC có thể bằng bội số của giá trị POC. Bằng cách chia giá trị POC cho giá trị nguyên, giá trị AUC có thể được tính toán. Trong các trường hợp cụ thể, các phép chia có thể đặt gánh nặng cụ thể lên các triển khai bộ giải mã. Trong các trường hợp này, các giới hạn nhỏ trong khoảng đánh số của các giá trị AUC có thể cho phép thay thế phép chia bằng phép dịch. Chẳng hạn, giá trị AUC có thể bằng giá trị bit trọng số lớn nhất (MSB) của khoảng giá trị POC.

Theo phương án thực hiện tương tự, giá trị của chu trình POC trên AU (`poc_cycle_au`) có thể được báo hiệu trong cấu trúc cú pháp mức cao, chẳng hạn, tiêu đề khối NAL, tiêu đề lát, tiêu đề nhóm ô, thông điệp SEI, tập hợp tham số hoặc dấu tách AU. `poc_cycle_au` có thể chỉ báo bao nhiêu giá trị POC liên tiếp và khác nhau có thể được liên kết với cùng AU. Chẳng hạn, nếu giá trị của `poc_cycle_au` bằng 4, các ảnh, các lát hoặc các ô có giá trị POC bằng trong khoảng từ 0 đến 3, được liên kết với AU có giá trị AUC bằng 0, và các ảnh, các lát hoặc các ô có giá trị POC nằm trọn trong khoảng từ 4 đến 7, được liên kết với AU có giá trị AUC bằng 1. Do vậy, giá trị của AUC có thể được suy ra bằng cách chia giá trị POC cho giá trị của `poc_cycle_au`.

Theo phương án thực hiện tương tự hoặc khác, giá trị của `poc_cyle_au` có thể được dẫn xuất từ thông tin, được đặt chẳng hạn trong VPS, mà nhận diện số lượng lớp không gian hoặc SNR trong chuỗi video được mã hóa. Mỗi quan hệ khả thi này được mô tả ngắn gọn dưới đây. Trong khi phép dẫn xuất như nêu trên

có thể tiết kiệm vài bit trong VPS và do vậy có thể cải thiện hiệu suất mã hóa, có lợi thế là có thể mã hóa tường minh poc_cycle_au trong cấu trúc cú pháp mức cao thích hợp về mặt phân cấp dưới VPS, để có thể giảm tối đa poc_cycle_au đối với phần nhỏ cụ thể của dòng bit, chẳng hạn, ảnh. Việc tối ưu hóa này có thể tiết kiệm nhiều bit hơn có thể tiết kiệm qua quá trình dẫn xuất nêu trên do các giá trị POC (và/hoặc các giá trị của các phần tử cú pháp đề cập gián tiếp đến POC) có thể được mã hóa trong các cấu trúc cú pháp mức thấp.

Theo phương án thực hiện tương tự hoặc khác, Fig.9 thể hiện ví dụ của các bảng cú pháp để báo hiệu phần tử cú pháp của vps_poc_cycle_au trong VPS (hoặc SPS), mà chỉ báo poc_cycle_au được sử dụng cho tất cả các ảnh/lát trong chuỗi video được mã hóa, và phần tử cú pháp của slice_poc_cycle_au, mà chỉ báo poc_cycle_au của lát hiện tại, trong tiêu đề lát. Nếu giá trị POC tăng đồng đều trên AU, vps_contant_poc_cycle_per_au trong VPS được gán bằng 1 và vps_poc_cycle_au được báo hiệu trong VPS. Trong trường hợp này, slice_poc_cycle_au không được báo hiệu tường minh, và giá trị của AUC cho mỗi AU được tính toán bằng cách chia giá trị của POC cho vps_poc_cycle_au. Nếu giá trị POC không tăng đồng đều trên AU, vps_contant_poc_cycle_per_au trong VPS được gán bằng 0. Trong trường hợp này, vps_access_unit_cnt không được báo hiệu, trong khi slice_access_unit_cnt được báo hiệu trong tiêu đề lát cho mỗi ảnh hoặc lát. Mỗi ảnh hoặc lát có thể có giá trị khác nhau của slice_access_unit_cnt. Giá trị của AUC cho mỗi AU được tính toán bằng cách chia giá trị của POC cho slice_poc_cycle_au. Fig.10 thể hiện sơ đồ khối minh họa luồng công việc tương đương.

Theo các phương án thực hiện giống hệt hoặc tương tự, mặc dù giá trị của POC của ảnh, lát, hoặc ô có thể là khác nhau, ảnh, lát, hoặc ô tương ứng với AU có cùng giá trị AUC có thể được liên kết với thực thể thời gian đầu ra hoặc giải mã tương tự. Do vậy, không có lệ thuộc giải mã/phân tách ngoài bất kỳ giữa các ảnh, các lát hoặc các ô trong cùng AU, tất cả hoặc tập con 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 thực thể thời gian.

Theo các phương án thực hiện giống hệt hoặc tương tự, mặc dù giá trị của POC của ảnh, lát, hoặc ô có thể là khác nhau, ảnh, lát, hoặc ô tương ứng với AU có cùng giá trị AUC có thể được liên kết với thực thể thời gian hiển thị/tổng hợp tương tự. Khi thời gian tổng hợp được chứa trong định dạng phần chứa, mặc dù các ảnh tương ứng với các AU khác nhau, nếu các ảnh có cùng thời gian tổng hợp, các ảnh có thể được hiển thị ở cùng thực thể thời gian.

Theo các phương án thực hiện giống hệt hoặc tương tự, mỗi ảnh, lát, hoặc ô có thể có cùng ID thời gian (`temporal_id`) trong cùng AU. Tất cả hoặc tập con của các ảnh, các lát hoặc các ô tương ứng với thực thể thời gian có thể được liên kết với cùng lớp con thời gian. Theo các phương án thực hiện giống hệt hoặc tương tự, mỗi ảnh, lát, hoặc ô có thể có ID lớp không gian tương tự hoặc khác nhau (`layer_id`) trong cùng AU. Tất cả hoặc tập con của các ảnh, các lát hoặc các ô tương ứng với thực thể thời gian có thể được liên kết với lớp không gian tương tự hoặc khác nhau.

Fig.8 thể hiện ví dụ về cấu trúc chuỗi video có kết hợp của `temporal_id`, `layer_id`, các giá trị POC và AUC có ARC. Trong ví dụ này, ảnh, lát hoặc ô trong AU thứ nhất có $AUC = 0$ có thể có `temporal_id = 0` và `layer_id = 0` hoặc 1, trong khi ảnh, lát hoặc ô trong AU thứ hai có $AUC = 1$ có thể có lần lượt `temporal_id = 1` và `layer_id = 0` hoặc 1. Giá trị của POC được tăng lên 1 trên ảnh bất kể các giá trị của `temporal_id` và `layer_id`. Trong ví dụ này, giá trị của `poc_cycle_au` có thể bằng 2. Tốt hơn nếu giá trị của `poc_cycle_au` có thể được gán bằng số lớp (tính khả mở không gian). Trong ví dụ này, do vậy, giá trị của POC được tăng lên 2, trong khi giá trị của AUC được tăng lên 1.

Theo các phương án thực hiện nêu trên, 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 hiệu RPS hiện tại trong HEVC hoặc báo hiệu RPL. Trong RPS hoặc RPL, ảnh tham chiếu được tạo được chỉ báo bằng cách báo hiệu giá trị của POC hoặc giá trị delta của POC giữa ảnh hiện tại và ảnh tham chiếu được tạo. Theo 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 báo hiệu, nhưng có các giới hạn sau. Nếu giá trị của `temporal_id` của ảnh tham chiếu lớn hơn giá trị của `temporal_id` của ảnh

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

Theo các phương án thực hiện tương tự hoặc khác, phép chia tỷ lệ MV dựa trên hiệu số POC cho dự báo MV thời gian (TMVP) có thể bị vô hiệu hóa giữa nhiều ảnh trong AU. Do vậy, mặc dù mỗi ảnh có thể có giá trị POC khác trong AU, MV không được chia tỷ lệ và được sử dụng cho TMVP trong AU. Đây là do ảnh tham chiếu có POC khác nhau trong cùng AU được xem xét là ảnh tham chiếu có cùng thực thể thời gian. Do vậy, theo phương án thực hiện, hàm chia tỷ lệ MV có thể trả về 1, khi ảnh tham chiếu thuộc AU được liên kết với ảnh hiện tại.

Theo các phương án thực hiện tương tự hoặc khác, phép chia tỷ lệ MV dựa trên hiệu số POC cho TMVP có thể bị vô hiệu hóa một cách tùy chọn giữa nhiều ảnh, khi độ phân giải không gian của ảnh tham chiếu khác với độ phân giải không gian của ảnh hiện tại. Khi phép chia tỷ lệ MV được phép, MV được chia tỷ lệ dựa trên cả hiệu số POC và tỷ lệ độ phân giải không gian giữa ảnh hiện tại và ảnh tham chiếu.

Theo phương án thực hiện tương tự hoặc khác, MV có thể được chia tỷ lệ dựa trên hiệu số AUC thay cho hiệu số POC, đối với TMVP, đặc biệt khi `poc_cycle_au` có giá trị không đồng nhất (khi `vps_contant_poc_cycle_per_au == 0`). Ngược lại (khi `vps_contant_poc_cycle_per_au == 1`), phép chia tỷ lệ MV dựa trên hiệu số AUC có thể giống hệt phép chia tỷ lệ MV dựa trên hiệu số POC.

Theo phương án thực hiện tương tự hoặc khác, khi MV được chia tỷ lệ dựa trên hiệu số AUC, MV tham chiếu trong cùng AU (có cùng giá trị AUC) với ảnh hiện tại không được chia tỷ lệ dựa trên hiệu số AUC và được sử dụng để dự báo MV không chia tỷ lệ hoặc có chia tỷ lệ dựa trên tỷ lệ độ phân giải không gian giữa ảnh hiện tại và ảnh tham chiếu.

Theo các phương án thực hiện tương tự hoặc khác, giá trị AUC được sử dụng để nhận diện biên của AU và được sử dụng cho hoạt động bộ giải mã chuẩn giả định (HRD), cần thời điểm đầu vào và đầu ra có độ hạt AU. Trong phần lớn

trường hợp, ảnh được giải mã có lớp cao nhất trong AU có thể được xuất ra để hiển thị. Giá trị AUC và giá trị layer_id có thể được sử dụng để nhận diện ảnh đầu ra.

Theo phương án thực hiện, ảnh có thể bao gồm một hoặc nhiều ảnh con. Mỗi ảnh con có thể bao phủ vùng cục bộ hoặc toàn bộ vùng của ảnh. Vùng được hỗ trợ bởi ảnh con có thể hoặc không thể được chồng lấp với vùng được hỗ trợ bởi ảnh con khác. Vùng được tổng hợp bởi một hoặc nhiều ảnh con có thể hoặc không thể bao phủ toàn bộ vùng của ảnh. Nếu ảnh bao gồm ảnh con, vùng được hỗ trợ bởi ảnh con giống hệt vùng được hỗ trợ bởi ảnh.

Theo phương án thực hiện tương tự, ảnh con có thể được mã hóa bằng phương pháp mã hóa giống như phương pháp mã hóa được sử dụng cho ảnh được mã hóa. Ảnh con có thể được mã hóa độc lập hoặc có thể được mã hóa phụ thuộc vào ảnh con khác hoặc ảnh được mã hóa. Ảnh con có thể hoặc không thể có lệ thuộc phân tách bất kỳ từ ảnh con khác hoặc ảnh được mã hóa.

Theo phương án thực hiện tương tự, ảnh con được mã hóa có thể được chứa trong một hoặc nhiều lớp. Ảnh con được mã hóa trong lớp có thể có độ phân giải không gian khác. Ảnh con ban đầu có thể được lấy mẫu lại theo không gian (được tăng kích thước mẫu hoặc được giảm kích thước mẫu), được mã hóa với các tham số độ phân giải không gian khác nhau, và được chứa trong dòng bit tương ứng với lớp.

Theo phương án thực hiện tương tự hoặc khác, ảnh con có (W, H) , trong đó W chỉ báo chiều rộng của ảnh con và H chỉ báo chiều cao của ảnh con, một cách lần lượt, có thể được mã hóa và được chứa trong dòng bit được mã hóa tương ứng với lớp 0, trong khi ảnh con được tăng kích thước mẫu (hoặc được giảm kích thước mẫu) từ ảnh con có độ phân giải không gian ban đầu, với $(W^*S_{w,k}, H^*S_{h,k})$, có thể được mã hóa và được chứa trong dòng bit được mã hóa tương ứng với lớp k , trong đó $S_{w,k}, S_{h,k}$ chỉ báo các tỷ lệ lấy mẫu lại, chiều ngang và chiều dọc. Nếu các giá trị của $S_{w,k}, S_{h,k}$ lớn hơn 1, lấy mẫu lại tương đương tăng kích thước mẫu. Trong đó, nếu các giá trị của $S_{w,k}, S_{h,k}$ nhỏ hơn 1, lấy mẫu lại tương đương giảm kích thước mẫu.

Theo phương án thực hiện tương tự hoặc khác, ảnh con được mã hóa trong lớp có thể có chất lượng hình ảnh khác với chất lượng của ảnh con được mã hóa trong lớp khác trong cùng ảnh con hoặc ảnh con khác. Chẳng hạn, ảnh con i trong lớp, n , được mã hóa với tham số lượng tử, $Q_{i,n}$, trong khi ảnh con j trong lớp, m , được mã hóa với tham số lượng tử, $Q_{j,m}$.

Theo phương án thực hiện tương tự hoặc khác, ảnh con được mã hóa trong lớp có thể được giải mã độc lập, mà không có lệ thuộc phân tách hoặc giải mã bất kỳ với ảnh con được mã hóa trong lớp khác của cùng vùng cục bộ. Lớp ảnh con, mà có thể được giải mã độc lập mà không tạo tham chiếu lớp ảnh con khác của cùng vùng cục bộ, là lớp ảnh con độc lập. Ảnh con được mã hóa trong lớp ảnh con độc lập có thể hoặc không thể có lệ thuộc giải mã hoặc phân tách với ảnh con được mã hóa trước đó trong cùng lớp ảnh con, nhưng ảnh con được mã hóa có thể không có lệ thuộc bất kỳ với ảnh được mã hóa trong lớp ảnh con khác.

Theo phương án thực hiện tương tự hoặc khác, ảnh con được mã hóa trong lớp có thể được giải mã phụ thuộc, có lệ thuộc phân tách hoặc giải mã bất kỳ với ảnh con được mã hóa trong lớp khác của cùng vùng cục bộ. Lớp ảnh con, mà có thể được giải mã phụ thuộc mà có tạo tham chiếu lớp ảnh con khác của cùng vùng cục bộ, là lớp ảnh con phụ thuộc. Ảnh con được mã hóa trong ảnh con lệ thuộc có thể tham chiếu ảnh con được mã hóa thuộc cùng ảnh con, ảnh con được mã hóa trước đó trong cùng lớp ảnh con, hoặc cả hai ảnh con tham chiếu.

Theo phương án thực hiện tương tự hoặc khác, ảnh con được mã hóa gồm một hoặc nhiều lớp ảnh con độc lập và một hoặc nhiều lớp ảnh con phụ thuộc. Tuy nhiên, ít nhất một lớp ảnh con độc lập có thể xuất hiện cho ảnh con được mã hóa. Lớp ảnh con độc lập có thể có giá trị của định danh lớp (`layer_id`), có thể xuất hiện trong tiêu đề khối NAL hoặc cấu trúc cú pháp mức cao khác, bằng 0. Lớp ảnh con có `layer_id` bằng 0 là lớp ảnh con cơ sở.

Theo phương án thực hiện tương tự hoặc khác, ảnh có thể bao gồm một hoặc nhiều ảnh con nền trước và một ảnh con nền sau. Vùng được hỗ trợ bởi ảnh con nền sau có thể bằng vùng của ảnh. Vùng được hỗ trợ bởi ảnh con nền trước có thể được chồng lấp với vùng được hỗ trợ bởi ảnh con nền sau. Ảnh con nền sau có thể là lớp ảnh con cơ sở, trong khi ảnh con nền trước có thể là lớp ảnh con

không cơ sở (tăng cường). Một hoặc nhiều lớp ảnh con không cơ sở có thể tham chiếu cùng lớp cơ sở để giải mã. Mỗi lớp ảnh con không cơ sở có `layer_id` bằng a có thể tham chiếu lớp ảnh con không cơ sở có `layer_id` bằng b , trong đó a lớn hơn b .

Theo phương án thực hiện tương tự hoặc khác, ảnh có thể bao gồm một hoặc nhiều ảnh con nền trước có hoặc không có ảnh con nền sau. Mỗi ảnh con có thể có lớp ảnh con cơ sở riêng và một hoặc nhiều lớp ảnh con không cơ sở (tăng cường). Mỗi lớp ảnh con cơ sở có thể được tạo tham chiếu bởi một hoặc nhiều lớp ảnh con không cơ sở. Mỗi lớp ảnh con không cơ sở có `layer_id` bằng a có thể tham chiếu lớp ảnh con không cơ sở có `layer_id` bằng b , trong đó a lớn hơn b .

Theo phương án thực hiện tương tự hoặc khác, ảnh có thể bao gồm một hoặc nhiều ảnh con nền trước có hoặc không có ảnh con nền sau. Mỗi ảnh con được mã hóa trong lớp ảnh con (cơ sở hoặc không cơ sở) có thể được tạo tham chiếu bởi một hoặc nhiều lớp ảnh con không cơ sở thuộc cùng ảnh con và một hoặc nhiều lớp ảnh con không cơ sở, mà không thuộc cùng ảnh con.

Theo phương án thực hiện tương tự hoặc khác, ảnh có thể bao gồm một hoặc nhiều ảnh con nền trước có hoặc không có ảnh con nền sau. Ảnh con trong lớp a có thể còn được phân vùng thành nhiều ảnh con trong cùng lớp. Một hoặc nhiều ảnh con được mã hóa trong lớp b có thể tham chiếu ảnh con được phân vùng trong lớp a .

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

Theo phương án thực hiện tương tự hoặc khác, CSPS có thể được mã hóa và được chứa trong một hoặc nhiều lớp. CSPS có thể gồm một hoặc nhiều lớp CSPS. Giải mã một hoặc nhiều lớp CSPS tương ứng với CSPS có thể tái tạo chuỗi ảnh con tương ứng với cùng vùng cục bộ.

Theo phương án thực hiện tương tự hoặc khác, số lượng lớp CSPS tương ứng với CSPS có thể giống hệt hoặc khác với số lượng lớp CSPS tương ứng với CSPS khác.

Theo phương án thực hiện tương tự hoặc khác, lớp CSPS có thể có độ phân giải thời gian khác (chẳng hạn, tốc độ khung hình) với lớp CSPS khác. Chuỗi ảnh con ban đầu (không được nén) có thể được lấy mẫu lại theo thời gian (được tăng kích thước mẫu hoặc được giảm kích thước mẫu), được mã hóa với các tham số độ phân giải thời gian khác nhau, và được chứa trong dòng bit tương ứng với lớp.

Theo phương án thực hiện tương tự hoặc khác, chuỗi ảnh con có tốc độ khung hình, F , có thể được mã hóa và được chứa trong dòng bit được mã hóa tương ứng với lớp 0, trong khi chuỗi ảnh con được tăng kích thước mẫu (hoặc giảm kích thước mẫu) theo thời gian từ chuỗi ảnh con ban đầu, 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 lệ lấy mẫu theo thời gian cho lớp k . Nếu giá trị của $S_{t,k}$ lớn hơn 1, tiến trình lấy mẫu lại theo thời gian là biến đổi tăng tốc độ khung hình. Trong đó, nếu giá trị của $S_{t,k}$ nhỏ hơn 1, tiến trình lấy mẫu lại theo thời gian là biến đổi giảm tốc độ khung hình.

Theo phương án thực hiện tương tự hoặc khác, khi ảnh con có lớp CSPS a được tạo tham chiếu bởi ảnh con 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 , các pixel được giải mã trong lớp CSPS a được lấy mẫu lại và được sử dụng để tham chiếu. Tiến trình lấy mẫu lại có thể cần lọc tăng kích thước mẫu hoặc lọc giảm kích thước mẫu.

Fig.11 thể hiện dòng video ví dụ bao gồm CSPS video nền sau với layer_id bằng 0 và nhiều lớp CSPS nền trước. Trong khi ảnh con được mã hóa có thể gồm một hoặc nhiều lớp CSPS, vùng nền sau, mà không thuộc lớp CSPS nền trước bất kỳ, có thể gồm lớp cơ sở. Lớp cơ sở có thể chứa vùng nền sau và các vùng nền trước, trong khi lớp CSPS tăng cường chứa vùng nền trước. Lớp CSPS tăng cường có thể có chất lượng hình ảnh tốt hơn lớp cơ sở, ở vùng tương tự. Lớp

CSPS tăng cường có thể tham chiếu các pixel được tái tạo và các MV của lớp cơ sở, tương ứng với vùng tương tự.

Theo phương án thực hiện tương tự hoặc khác, dòng bit video tương ứng với lớp cơ sở được chứa trong rãnh, trong khi các lớp CSPS tương ứng với mỗi ảnh con được chứa trong rãnh được tách riêng, trong tệp tin video.

Theo phương án thực hiện tương tự hoặc khác, dòng bit video tương ứng với lớp cơ sở được chứa trong rãnh, trong khi các lớp CSPS có cùng layer_id được chứa trong rãnh được tách riêng. Trong ví dụ này, rãnh tương ứng với lớp k bao gồm các lớp CSPS tương ứng với riêng lớp k .

Theo phương án thực hiện tương tự hoặc khác, mỗi lớp CSPS của mỗi ảnh con được lưu trữ trong rãnh được tách riêng. Mỗi rãnh có thể hoặc không thể có lệ thuộc phân tách hoặc giải mã bất kỳ với một hoặc nhiều rãnh khác.

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

Theo phương án thực hiện tương tự hoặc khác, ảnh gồm một hoặc nhiều dữ liệu đa phương tiện liên kết bao gồm bản đồ chiều sâu, bản đồ alpha, bản đồ hình học 3D, bản đồ chiếm dụng, v.v.. Dữ liệu đa phương tiện định thời liên kết có thể được phân chia thành một hoặc nhiều dòng con dữ liệu mà mỗi dòng tương ứng với một ảnh con.

Theo phương án thực hiện tương tự hoặc khác, Fig.12 thể hiện ví dụ của hội nghị video dựa trên phương pháp ảnh con nhiều lớp. Trong dòng video, một lớp cơ sở dòng bit video tương ứng với ảnh nền sau và một hoặc nhiều lớp tăng cường các dòng bit video tương ứng với các ảnh con nền trước được chứa. Mỗi lớp tăng cường video dòng bit có thể tương ứng với lớp CSPS. Trong bộ phận hiển thị, ảnh tương ứng với lớp cơ sở được hiển thị mặc định. Nó chứa một hoặc nhiều ảnh trong ảnh của người dùng (PIP). Khi người dùng cụ thể được chọn bởi điều khiển của máy khách, lớp CSPS tăng cường tương ứng với người dùng được chọn được giải mã và được hiển thị có độ phân giải không gian hoặc chất lượng được tăng cường.

Fig.13 thể hiện sơ đồ khối của quá trình giải mã và hiển thị cho dòng bit video bao gồm các ảnh con đa lớp theo phương án thực hiện. Chẳng hạn, tiến trình có thể bao gồm một hoặc nhiều hoạt động sau. Chẳng hạn, trong hoạt động 1301, giải mã dòng bit video có nhiều lớp có thể xuất hiện. Hoạt động 1302 có thể bao gồm nhận diện vùng nền sau và một hoặc nhiều ảnh con nền trước. Hoạt động 1303 có thể bao gồm xác định liệu vùng ảnh con cụ thể đã được chọn. Hoạt động 1304 có thể bao gồm giải mã và hiển thị ảnh con được tăng cường nếu vùng ảnh con cụ thể đã được chọn (tức là, 1303 = Có). Hoạt động 1305 có thể bao gồm giải mã và hiển thị vùng nền sau, nếu vùng ảnh con cụ thể đã không được chọn (tức là, 1303 = Không).

Theo phương án thực hiện tương tự hoặc khác, hộp trung gian của mạng (chẳng hạn bộ định tuyến) có thể chọn ảnh con của các lớp để gửi đến người dùng tùy thuộc vào băng thông của nó. Việc tổ chức ảnh/ảnh con có thể được sử dụng để thích ứng băng thông. Chẳng hạn, nếu người dùng không có băng thông, bộ định tuyến có thể xóa lớp hoặc chọn một số ảnh con theo độ quan trọng hoặc dựa trên thiết lập được sử dụng và điều này có thể được hoàn thành động để sử dụng băng thông.

Fig.14 thể hiện trường hợp sử dụng của video 360. Khi ảnh 360 hình cầu được chiếu lên ảnh mặt phẳng, ảnh hình chiếu 360 có thể được phân vùng thành nhiều ảnh con như là lớp cơ sở. Lớp tăng cường của ảnh con cụ thể có thể được mã hóa và được truyền đến máy khách. Bộ giải mã có thể giải mã cả lớp cơ sở bao gồm tất cả các ảnh con và lớp tăng cường của ảnh con được chọn. Khi công nhìn hiện tại giống hệt ảnh con được chọn, ảnh được hiển thị có thể có chất lượng cao hơn với ảnh con được giải mã với lớp tăng cường. Ngược lại, ảnh được giải mã với lớp cơ sở có thể được hiển thị, có chất lượng thấp.

Theo phương án thực hiện tương tự hoặc khác, thông tin bố trí bất kỳ để hiển thị có thể xuất hiện trong tệp tin, dưới dạng thông tin bổ sung (chẳng hạn, thông điệp SEI hoặc metadata). Một hoặc nhiều ảnh con được giải mã có thể được định vị lại và được hiển thị tùy thuộc vào thông tin bố trí được báo hiệu. Thông tin bố trí có thể được báo hiệu bởi máy chủ streaming hoặc bộ phát quảng bá, hoặc có

thể được tạo lại bởi thực thể mạng hoặc máy chủ đám mây, hoặc có thể được xác định bởi thiết lập tùy chỉnh của người dùng.

Theo phương án thực hiện, khi ảnh đầu vào được phân chia thành một hoặc nhiều vùng con (chữ nhật), mỗi vùng con có thể được mã hóa dưới dạng lớp độc lập. Mỗi lớp độc lập tương ứng với vùng cục bộ có thể có giá trị `layer_id` duy nhất. Đối với mỗi lớp độc lập, kích thước ảnh con và thông tin vị trí có thể được báo hiệu. Chẳng hạn, kích thước ảnh (chiều rộng, chiều cao), thông tin độ lệch của góc trên bên trái (`x_offset`, `y_offset`). Fig.15 thể hiện ví dụ của bố trí của các ảnh con được phân chia, kích thước ảnh con của nó và thông tin vị trí và cấu trúc dự báo ảnh tương ứng của nó. Thông tin bố trí bao gồm (các) kích thước ảnh con và (các) vị trí ảnh con có thể được báo hiệu trong cấu trúc cú pháp mức cao, chẳng hạn (các) tập hợp tham số, tiêu đề của lát hoặc nhóm ô, hoặc thông điệp SEI.

Theo phương án thực hiện tương tự, mỗi ảnh con tương ứng với lớp độc lập có thể có giá trị POC duy nhất trong AU. Khi ảnh tham chiếu trong số các ảnh được lưu trữ trong DPB được chỉ báo bởi nhờ sử dụng (các) phần tử cú pháp trong cấu trúc RPS hoặc RPL, (các) giá trị POC của mỗi ảnh con tương ứng với lớp có thể được sử dụng.

Theo phương án thực hiện tương tự hoặc khác, để chỉ báo cấu trúc dự báo (lớp ngoài), `layer_id` có thể không được sử dụng và giá trị POC (`delta`) có thể được sử dụng.

Theo phương án thực hiện tương tự, ảnh con có giá trị POC bằng N tương ứng với lớp (hoặc vùng cục bộ) có thể hoặc không thể được sử dụng làm ảnh tham chiếu của ảnh con có giá trị POC bằng $N+K$, tương ứng với lớp tương tự (hoặc cùng vùng cục bộ) để dự báo được bù chuyển động. Trong phần lớn trường hợp, giá trị của số lượng K có thể bằng số lượng lớn nhất của các lớp (độc lập), mà có thể giống như số lượng vùng con.

Theo phương án thực hiện tương tự hoặc khác, Fig.16 thể hiện trường hợp mở rộng trên Fig.15. Khi ảnh đầu vào được phân chia thành nhiều (chẳng hạn, bốn) vùng con, mỗi vùng cục bộ có thể được mã hóa với một hoặc nhiều lớp. Trong trường hợp này, số lượng lớp độc lập có thể bằng số lượng vùng con, và

một hoặc nhiều lớp có thể tương ứng với vùng con. Do vậy, mỗi vùng con có thể được mã hóa với một hoặc nhiều lớp độc lập và 0 hoặc nhiều lớp phụ thuộc.

Theo phương án thực hiện tương tự, trên Fig.16, ảnh đầu vào có thể được phân chia thành bốn vùng con. Vùng con trên bên phải có thể được mã hóa như là hai lớp, vốn là lớp 1 và lớp 4, trong khi vùng con dưới bên phải có thể được mã hóa dưới dạng hai lớp, vốn là lớp 3 và lớp 5. Trong trường hợp này, lớp 4 có thể tham chiếu lớp 1 để dự báo được bù chuyển động, trong khi lớp 5 có thể tham chiếu lớp 3 để bù chuyển động.

Theo phương án thực hiện tương tự hoặc khác, lọc trong vòng (chẳng hạn lọc tách khối, lọc trong vòng thích ứng, bộ tạo dạng lại, lọc song phương hoặc lọc dựa trên học sâu) giữa đường biên lớp có thể được vô hiệu hóa (tùy chọn).

Theo phương án thực hiện tương tự hoặc khác, dự báo được bù chuyển động hoặc IBC giữa đường biên lớp có thể được vô hiệu hóa (tùy chọn).

Theo phương án thực hiện tương tự hoặc khác, bước đệm đường biên để dự báo được bù chuyển động hoặc lọc trong vòng ở đường biên của ảnh con có thể được xử lý tùy chọn. Cờ chỉ báo liệu bước đệm đường biên được xử lý hoặc không có thể được báo hiệu trong cấu trúc cú pháp mức cao, chẳng hạn (các) tập hợp tham số (VPS, SPS, PPS, hoặc APS), tiêu đề lát hoặc nhóm ô, hoặc thông điệp SEI.

Theo phương án thực hiện tương tự hoặc khác, thông tin bố trí của (các) vùng con (hoặc (các) ảnh con) có thể được báo hiệu trong VPS hoặc SPS. Fig.17 thể hiện ví dụ của các phần tử cú pháp trong VPS và SPS. Trong ví dụ này, `vps_sub_picture_dividing_flag` được báo hiệu trong VPS. Cờ có thể chỉ báo liệu (các) ảnh đầu vào được phân chia thành nhiều vùng con hoặc không. Khi giá trị của `vps_sub_picture_dividing_flag` bằng 0, (các) ảnh đầu vào trong (các) chuỗi video được mã hóa tương ứng với VPS hiện tại có thể không được phân chia thành nhiều vùng con. Trong trường hợp này, kích thước ảnh đầu vào có thể bằng kích thước ảnh được mã hóa (`pic_width_in_luma_samples`, `pic_height_in_luma_samples`), được báo hiệu trong SPS. Khi giá trị của `vps_sub_picture_dividing_flag` bằng 1, (các) ảnh đầu vào có thể được phân chia thành nhiều vùng con. Trong trường hợp này, các phần tử cú pháp

`vps_full_pic_width_in_luma_samples` và `vps_full_pic_height_in_luma_samples` được báo hiệu trong VPS. Các giá trị của `vps_full_pic_width_in_luma_samples` và `vps_full_pic_height_in_luma_samples` có thể lần lượt bằng chiều rộng và chiều cao của (các) ảnh đầu vào.

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

Theo phương án thực hiện tương tự, thông tin vị trí (`pic_offset_x`, `pic_offset_y`) của góc trên bên trái của vùng con có thể không được sử dụng để giải mã, nhưng có thể được sử dụng để tổng hợp và hiển thị.

Theo phương án thực hiện tương tự hoặc khác, thông tin bố trí (kích thước và vị trí) của tất cả hoặc tập con của (các) vùng con của (các) ảnh đầu vào, thông tin lệ thuộc giữa (các) lớp có thể được báo hiệu trong tập hợp tham số hoặc thông điệp SEI.

Fig.18 thể hiện ví dụ của các phần tử cú pháp để chỉ báo thông tin hoặc bố trí của các vùng con, lệ thuộc giữa các lớp, và mối quan hệ giữa vùng con và một hoặc nhiều lớp. Trong ví dụ này, phần tử cú pháp này `num_sub_region` chỉ báo số lượng vùng con (chữ nhật) trong chuỗi video được mã hóa hiện tại. Theo phương án thực hiện, phần tử cú pháp `num_layers` có thể chỉ báo số lượng lớp trong chuỗi video được mã hóa hiện tại. Giá trị của `num_layers` có thể bằng hoặc lớn hơn giá trị của `num_sub_region`. Khi vùng con bất kỳ được mã hóa dưới dạng

một lớp, giá trị của `num_layers` có thể bằng giá trị của `num_sub_region`. Khi một hoặc nhiều vùng con được mã hóa dưới dạng nhiều lớp, giá trị của `num_layers` có thể lớn hơn giá trị của `num_sub_region`. Phần tử cú pháp `direct_dependency_flag[i][j]` chỉ báo lệ thuộc từ lớp thứ `j` đến lớp thứ `i`. `num_layers_for_region[i]` chỉ báo số lượng lớp được liên kết với vùng con thứ `i`. `sub_region_layer_id[i][j]` chỉ báo `layer_id` của lớp thứ `j` được liên kết với vùng con thứ `i`. `sub_region_offset_x[i]` và `sub_region_offset_y[i]` lần lượt chỉ báo vị trí ngang và dọc của góc trên bên trái của vùng con thứ `i`. `sub_region_width[i]` và `sub_region_height[i]` lần lượt chỉ báo chiều rộng và chiều cao của vùng con thứ `i`.

Theo một phương án thực hiện, một hoặc nhiều phần tử cú pháp xác định tập hợp lớp đầu ra để chỉ báo một trong nhiều lớp cần được xuất ra có hoặc không có thông tin tiểu sử/tầng/mức có thể được báo hiệu trong cấu trúc cú pháp mức cao, chẳng hạn, VPS, DPS, SPS, PPS, APS hoặc thông điệp SEI. Trên Fig.19, phần tử cú pháp `num_output_layer_sets` chỉ báo số của tập lớp đầu ra (OLS) trong CVS chỉ đến VPS có thể được báo hiệu trong VPS. Đối với mỗi tập hợp lớp đầu ra, `output_layer_flag` có thể được báo hiệu càng nhiều số lượng lớp đầu ra.

Theo phương án thực hiện tương tự, `output_layer_flag[i]` bằng 1 xác định rằng lớp thứ `i` được xuất ra. `vps_output_layer_flag[i]` bằng 0 xác định rằng lớp thứ `i` không được xuất ra.

Theo phương án thực hiện tương tự hoặc khác, một hoặc nhiều phần tử cú pháp xác định thông tin tiểu sử/tầng/mức cho mỗi tập hợp lớp đầu ra có thể được báo hiệu trong cấu trúc cú pháp mức cao, chẳng hạn, VPS, DPS, SPS, PPS, APS hoặc thông điệp SEI. Vẫn dựa vào Fig.19, phần tử cú pháp này `num_profile_tile_level` chỉ báo số lượng thông tin tiểu sử/tầng/mức trên OLS trong CVS chỉ đến VPS có thể được báo hiệu trong VPS. Đối với mỗi tập hợp lớp đầu ra, tập hợp phần tử cú pháp cho thông tin tiểu sử/tầng/mức hoặc chỉ số chỉ báo thông tin tiểu sử/tầng/mức cụ thể trong số các mục trong thông tin tiểu sử/tầng/mức có thể được báo hiệu càng nhiều số lượng lớp đầu ra.

Theo phương án thực hiện tương tự, `profile_tier_level_idx[i][j]` xác định chỉ số, vào danh sách của các cấu trúc cú pháp `profile_tier_level()` trong VPS, của cấu trúc cú pháp `profile_tier_level()` mà áp dụng lớp thứ j của OLS thứ i .

Theo phương án thực hiện tương tự hoặc khác, đề cập đến Fig.20, các phần tử cú pháp `num_profile_tile_level` và/hoặc `num_output_layer_sets` có thể được báo hiệu khi số lượng lớp lớn nhất lớn hơn 1 (`vps_max_layers_minus1 > 0`).

Theo phương án thực hiện tương tự hoặc khác, đề cập đến Fig.20, phần tử cú pháp `vps_output_layers_mode[i]` chỉ báo chế độ báo hiệu lớp đầu ra cho tập lớp đầu ra thứ i có thể xuất hiện trong VPS.

Theo phương án thực hiện tương tự, `vps_output_layers_mode[i]` bằng 0 xác định rằng chỉ lớp cao nhất được xuất ra với tập lớp đầu ra thứ i . `vps_output_layer_mode[i]` bằng 1 xác định rằng tất cả các lớp được xuất ra với tập lớp đầu ra thứ i . `vps_output_layer_mode[i]` bằng 2 xác định rằng các lớp mà được xuất ra là các lớp có `vps_output_layer_flag[i][j]` bằng 1 với tập lớp đầu ra thứ i . Nhiều giá trị hơn có thể được dành riêng.

Theo phương án thực hiện tương tự, `output_layer_flag[i][j]` có thể hoặc không thể được báo hiệu tùy thuộc vào giá trị của `vps_output_layers_mode[i]` cho tập lớp đầu ra thứ i .

Theo phương án thực hiện tương tự hoặc khác, đề cập đến Fig.20, cờ `vps_ptl_signal_flag[i]` có thể xuất hiện cho tập lớp đầu ra thứ i . Giá trị phụ thuộc của `vps_ptl_signal_flag[i]`, thông tin tiêu sử/tầng/mức cho tập lớp đầu ra thứ i có thể hoặc không thể được báo hiệu.

Theo phương án thực hiện tương tự hoặc khác, đề cập đến Fig.21, số lượng ảnh con, `max_subpics_minus1`, trong CVS hiện tại có thể được báo hiệu trong cấu trúc cú pháp mức cao, chẳng hạn, VPS, DPS, SPS, PPS, APS hoặc thông điệp SEI.

Theo phương án thực hiện tương tự, đề cập đến Fig.21, ID ảnh con, `sub_pic_id[i]`, đối với ảnh con thứ i có thể được báo hiệu, khi số lượng ảnh con lớn hơn 1 (`max_subpics_minus1 > 0`).

Theo phương án thực hiện tương tự hoặc khác, một hoặc nhiều phần tử cú pháp chỉ báo ID ảnh con thuộc mỗi lớp của mỗi tập hợp lớp đầu ra có thể được

báo hiệu trong VPS. Đề cập đến Fig.22, sub_pic_id_layer[i][j][k], mà chỉ báo ảnh con thứ k hiện có trong lớp thứ j của tập lớp đầu ra thứ i. Với thông tin đó, bộ giải mã có thể ghi nhận ảnh con nào có thể được giải mã và được xuất ra cho mỗi lớp của tập hợp lớp đầu ra cụ thể.

Theo phương án thực hiện, 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. Đơn vị ảnh (PU) là tập hợp khối NAL được liên kết với nhau theo quy tắc phân loại được xác định, là liên tiếp theo thứ tự giải mã, và chứa chính xác một ảnh được mã hóa. PU có thể chứa tiêu đề ảnh (PH) và một hoặc nhiều khối NAL VCL tổng hợp thành ảnh được mã hóa.

Theo phương án thực hiện, SPS (RBSP) có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu, được bao gồm trong ít nhất một AU có TemporalId bằng 0 hoặc được tạo qua phương tiện ngoài.

Theo phương án thực hiện, SPS (RBSP) có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu, được bao gồm trong ít nhất một AU có TemporalId bằng 0 trong CVS, mà chứa một hoặc nhiều PPS đề cập đến SPS, hoặc được tạo qua phương tiện ngoài.

Theo phương án thực hiện, SPS (RBSP) có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu bởi một hoặc nhiều PPS, được bao gồm trong ít nhất một PU có nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các khối NAL PPS đề cập đến khối NAL SPS trong CVS, mà chứa một hoặc nhiều PPS đề cập đến SPS, hoặc được tạo qua phương tiện ngoài.

Theo phương án thực hiện, SPS (RBSP) có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu bởi một hoặc nhiều PPS, được bao gồm trong ít nhất một PU có TemporalId bằng 0 và nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các khối NAL PPS đề cập đến khối NAL SPS hoặc được tạo qua phương tiện ngoài.

Theo phương án thực hiện, SPS (RBSP) có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu bởi một hoặc nhiều PPS, được bao gồm trong ít nhất một PU có TemporalId bằng 0 và nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các khối NAL PPS đề cập đến khối NAL SPS trong CVS, mà chứa

một hoặc nhiều PPS đề cập đến SPS, hoặc được tạo qua phương tiện ngoài hoặc được tạo qua phương tiện ngoài.

Theo phương án thực hiện tương tự hoặc khác, `pps_seq_parameter_set_id` xác định giá trị của `sps_seq_parameter_set_id` đối với SPS được tạo tham chiếu. Giá trị của `pps_seq_parameter_set_id` có thể giống nhau trong tất cả các PPS được tham chiếu bởi các ảnh được mã hóa trong CLVS.

Theo phương án thực hiện tương tự hoặc khác, tất cả các khối SPS NAL có giá trị cụ thể của `sps_seq_parameter_set_id` trong CVS có thể có cùng nội dung.

Theo phương án thực hiện tương tự hoặc khác, bất kể các giá trị `nuh_layer_id`, các khối NAL SPS có thể chia sẻ cùng giá trị của `sps_seq_parameter_set_id`.

Theo phương án thực hiện tương tự hoặc khác, giá trị `nuh_layer_id` của khối NAL SPS có thể bằng giá trị `nuh_layer_id` thấp nhất của các khối NAL PPS đề cập đến khối NAL SPS.

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

Theo phương án thực hiện, PPS (RBSP) có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu, được bao gồm trong ít nhất một AU có TemporalId bằng TemporalId của khối NAL PPS hoặc được tạo qua phương tiện ngoài.

Theo phương án thực hiện, PPS (RBSP) có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu, được bao gồm trong ít nhất một AU có TemporalId bằng TemporalId của khối NAL PPS trong CVS, mà chứa một hoặc nhiều PH (hoặc các khối NAL lát được mã hóa) đề cập đến PPS, hoặc được tạo qua phương tiện ngoài.

Theo phương án thực hiện, PPS (RBSP) có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu bởi một hoặc nhiều PH (hoặc các khối NAL lát được mã hóa), được bao gồm trong ít nhất một PU có `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của các khối NAL lát được mã hóa mà chỉ đến khối

NAL PPS trong CVS, mà chứa một hoặc nhiều PH (hoặc các khối NAL lát được mã hóa) đề cập đến PPS, hoặc được tạo qua phương tiện ngoài.

Theo phương án thực hiện, PPS (RBSP) có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu bởi một hoặc nhiều PH (hoặc các khối NAL lát được mã hóa), được bao gồm trong ít nhất một PU có TemporalId bằng TemporalId của khối NAL PPS và nuh_layer_id bằng giá trị nuh_layer_id thấp nhất của các khối NAL lát được mã hóa mà chỉ đến khối NAL PPS trong CVS, mà chứa một hoặc nhiều PH (hoặc các khối NAL lát được mã hóa) đề cập đến PPS, hoặc được tạo qua phương tiện ngoài.

Theo phương án thực hiện tương tự hoặc khác, ph_pic_parameter_set_id trong PH xác định giá trị của pps_pic_parameter_set_id cho PPS được tạo tham chiếu đang sử dụng. Giá trị của pps_seq_parameter_set_id có thể giống nhau trong tất cả các PPS được tham chiếu bởi các ảnh được mã hóa trong CLVS.

Theo phương án thực hiện tương tự hoặc khác, tất cả khối NAL PPS có giá trị cụ thể của pps_pic_parameter_set_id trong PU có thể có cùng nội dung.

Theo phương án thực hiện tương tự hoặc khác, bất kể giá trị nuh_layer_ids, các khối NAL PPS có thể chia sẻ cùng giá trị của pps_pic_parameter_set_id.

Theo phương án thực hiện tương tự hoặc khác, giá trị nuh_layer_id của khối NAL PPS có thể bằng giá trị nuh_layer_id thấp nhất của các khối NAL lát được mã hóa mà đề cập đến khối NAL mà chỉ đến khối NAL PPS.

Theo phương án thực hiện, khi PPS có nuh_layer_id bằng m được tham chiếu bởi một hoặc nhiều khối NAL lát được mã hóa có nuh_layer_id bằng n . lớp có nuh_layer_id bằng m có thể giống như lớp có nuh_layer_id bằng n hoặc lớp tham chiếu (trực tiếp hoặc gián tiếp) của lớp có nuh_layer_id bằng m .

Theo phương án thực hiện, PPS (RBSP) có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu, được bao gồm trong ít nhất một AU có TemporalId bằng TemporalId của khối NAL PPS hoặc được tạo qua phương tiện ngoài.

Theo phương án thực hiện, PPS (RBSP) có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu, được bao gồm trong ít nhất một AU có TemporalId bằng TemporalId của khối NAL PPS trong CVS, mà chứa một hoặc

nhiều PH (hoặc các khối NAL lát được mã hóa) đề cập đến PPS, hoặc được tạo qua phương tiện ngoài.

Theo phương án thực hiện, PPS (RBSP) có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu bởi một hoặc nhiều PH (hoặc các khối NAL lát được mã hóa), được bao gồm trong ít nhất một PU có `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của các khối NAL lát được mã hóa mà chỉ đến khối NAL PPS trong CVS, mà chứa một hoặc nhiều PH (hoặc các khối NAL lát được mã hóa) đề cập đến PPS, hoặc được tạo qua phương tiện ngoài.

Theo phương án thực hiện, PPS (RBSP) có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu bởi một hoặc nhiều PH (hoặc các khối NAL lát được mã hóa), được bao gồm trong ít nhất một PU có `TemporalId` bằng `TemporalId` của khối NAL PPS và `nuh_layer_id` bằng giá trị `nuh_layer_id` thấp nhất của các khối NAL lát được mã hóa mà chỉ đến khối NAL PPS trong CVS, mà chứa một hoặc nhiều PH (hoặc các khối NAL lát được mã hóa) đề cập đến PPS, hoặc được tạo qua phương tiện ngoài.

Theo phương án thực hiện tương tự hoặc khác, `ph_pic_parameter_set_id` trong PH xác định giá trị của `pps_pic_parameter_set_id` cho PPS được tạo tham chiếu đang sử dụng. Giá trị của `pps_seq_parameter_set_id` có thể giống nhau trong tất cả các PPS được tham chiếu bởi các ảnh được mã hóa trong CLVS.

Theo phương án thực hiện tương tự hoặc khác, tất cả khối NAL PPS có giá trị cụ thể của `pps_pic_parameter_set_id` trong PU có thể có cùng nội dung.

Theo phương án thực hiện tương tự hoặc khác, bất kể giá trị `nuh_layer_ids`, các khối NAL PPS có thể chia sẻ cùng giá trị của `pps_pic_parameter_set_id`.

Theo phương án thực hiện tương tự hoặc khác, giá trị `nuh_layer_id` của khối NAL PPS có thể bằng giá trị `nuh_layer_id` thấp nhất của các khối NAL lát được mã hóa mà đề cập đến khối NAL mà chỉ đến khối NAL PPS.

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

Lớp đầu ra chỉ báo lớp của tập hợp lớp đầu ra được xuất ra. Tập hợp lớp đầu ra (OLS) chỉ báo tập hợp lớp bao gồm tập hợp lớp cụ thể, trong đó một hoặc nhiều lớp trong tập hợp lớp riêng cho các lớp được xuất ra. Chỉ số lớp OLS là chỉ số, của lớp trong OLS, đến danh sách lớp trong OLS.

Lớp con chỉ báo lớp khả mở thời gian của dòng bit khả mở theo thời gian, bao gồm các khối NAL VCL có giá trị cụ thể của biến TemporalId và các khối NAL phi VCL liên kết. Biểu diễn tập con chỉ báo tập con của dòng bit bao gồm các khối NAL của lớp con cụ thể và các lớp con thấp hơn.

VPS RBSP có thể khả dụng cho tiến trình giải mã trước khi được tạo tham chiếu, được bao gồm trong ít nhất một AU có TemporalId bằng 0 hoặc được tạo qua phương tiện ngoài. Tất cả các khối NAL VPS có giá trị cụ thể của vps_video_parameter_set_id trong CVS có thể có cùng nội dung.

vps_video_parameter_set_id tạo ID cho VPS để tham chiếu bởi các phần tử cú pháp khác. Giá trị của vps_video_parameter_set_id có thể lớn hơn 0.

vps_max_layers_minus1 cộng 1 xác định số lượng lớp lớn nhất được phép trong mỗi CVS chỉ đến VPS.

vps_max_sublayers_minus1 cộng 1 xác định số lượng lớn nhất của các lớp con thời gian mà có thể xuất hiện trong lớp trong mỗi CVS chỉ đến VPS. Giá trị của vps_max_sublayers_minus1 có thể nằm trọn trong khoảng từ 0 đến 6.

vps_all_layers_same_num_sublayers_flag bằng 1 xác định rằng số lượng lớp con thời gian là giống nhau trong tất cả các lớp trong mỗi CVS chỉ đến VPS.

vps_all_layers_same_num_sublayers_flag bằng 0 xác định rằng các lớp trong mỗi CVS chỉ đến VPS có thể hoặc không thể có cùng số lượng lớp con thời gian. Khi không xuất hiện, giá trị của vps_all_layers_same_num_sublayers_flag được suy ra bằng 1.

vps_all_independent_layers_flag bằng 1 xác định rằng tất cả các lớp trong CVS được mã hóa độc lập mà không sử dụng dự báo ngoài lớp.

vps_all_independent_layers_flag bằng 0 xác định rằng một hoặc nhiều lớp trong CVS có thể sử dụng dự báo ngoài lớp. Khi không xuất hiện, giá trị của vps_all_independent_layers_flag được suy ra bằng 1.

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

`vps_independent_layer_flag[i]` bằng 1 xác định rằng lớp có chỉ số *i* không sử dụng dự báo ngoài lớp. `vps_independent_layer_flag[i]` bằng 0 xác định rằng lớp có chỉ số *i* có thể sử dụng dự báo ngoài lớp và các phần tử cú pháp.

`vps_direct_ref_layer_flag[i][j]` với *j* nằm trọn trong khoảng từ 0 đến *i* - 1, hiện có trong VPS. Khi không xuất hiện, giá trị của `vps_independent_layer_flag[i]` được suy ra bằng 1.

`vps_direct_ref_layer_flag[i][j]` bằng 0 xác định rằng lớp có chỉ số *j* không phải là lớp tham chiếu trực tiếp đối với lớp có chỉ số *i*. `vps_direct_ref_layer_flag[i][j]` bằng 1 xác định rằng lớp có chỉ số *j* là lớp tham chiếu trực tiếp đối với lớp có chỉ số *i*.

Khi `vps_direct_ref_layer_flag[i][j]` không xuất hiện với *i* và *j* nằm trọn trong khoảng từ 0 đến `vps_max_layers_minus1`, được suy ra bằng 0. Khi `vps_independent_layer_flag[i]` bằng 0, có thể có ít nhất một giá trị của *j* nằm trọn trong khoảng từ 0 đến *i* - 1, sao cho giá trị của `vps_direct_ref_layer_flag[i][j]` bằng 1.

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

```

for(i = 0; i <= vps_max_layers_minus1; i++) {
    for(j = 0; j <= vps_max_layers_minus1; j++) {
        dependencyFlag[i][j] = vps_direct_ref_layer_flag[i][j]
        for(k = 0; k < i; k++)
            if(vps_direct_ref_layer_flag[i][k] && dependencyFlag[k][j])
                dependencyFlag[i][j] = 1
    }
    LayerUsedAsRefLayerFlag[i] = 0
}
for(i = 0; i <= vps_max_layers_minus1; i++) {

```

```

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

```

Biến GeneralLayerIdx[i], xác định chỉ số lớp của lớp có nuh_layer_id bằng vps_layer_id[i], được dẫn xuất như sau:

```

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

```

Đối với hai giá trị khác nhau bất kỳ của i và j, đều nằm trọn trong khoảng từ 0 đến vps_max_layers_minus1, khi dependencyFlag[i][j] bằng 1, tương thích dòng bit yêu cầu rằng các giá trị của chroma_format_idc và bit_depth_minus8 áp dụng cho lớp thứ i có thể lần lượt bằng các giá trị của chroma_format_idc và bit_depth_minus8, áp dụng cho lớp thứ j.

max_tid_ref_present_flag[i] bằng 1 xác định rằng phần tử cú pháp này max_tid_il_ref_pics_plus1[i] xuất hiện. max_tid_ref_present_flag[i] bằng 0 xác định rằng phần tử cú pháp này max_tid_il_ref_pics_plus1[i] không xuất hiện.

max_tid_il_ref_pics_plus1[i] bằng 0 xác định rằng dự báo ngoài lớp không được sử dụng bởi các ảnh P phi IRA của lớp thứ i. max_tid_il_ref_pics_plus1[i] lớn hơn 0 xác định rằng, để giải mã các ảnh của lớp thứ i, không ảnh nào có TemporalId lớn hơn max_tid_il_ref_pics_plus1[i] - 1 được sử dụng làm ILRP. Khi không xuất hiện, giá trị của max_tid_il_ref_pics_plus1[i] được suy ra bằng 7.

each_layer_is_an_ols_flag bằng 1 xác định rằng mỗi OLS chứa chỉ một lớp và chính mỗi lớp trong CVS chỉ đến VPS là OLS với lớp đơn được bao gồm là

lớp đầu ra duy nhất. `each_layer_is_an_ols_flag` bằng 0 mà OLS có thể chứa nhiều hơn một lớp. Nếu `vps_max_layers_minus1` bằng 0, giá trị của `each_layer_is_an_ols_flag` được suy ra bằng 1. Ngược lại, khi `vps_all_independent_layers_flag` bằng 0, giá trị của `each_layer_is_an_ols_flag` được suy ra bằng 0.

`ols_mode_idc` bằng 0 xác định rằng tổng số OLS được xác định bởi VPS bằng `vps_max_layers_minus1 + 1`, OLS thứ *i* bao gồm các lớp có các chỉ số lớp nằm trọn từ 0 đến *i*, và đối với mỗi OLS chỉ lớp cao nhất trong OLS được xuất ra.

`ols_mode_idc` bằng 1 xác định rằng tổng số OLS được xác định bởi VPS bằng `vps_max_layers_minus1 + 1`, OLS thứ *i* bao gồm các lớp có các chỉ số lớp nằm trọn từ 0 đến *i*, và đối với mỗi OLS tất cả các lớp trong OLS được xuất ra.

`ols_mode_idc` bằng 2 xác định rằng tổng số OLS được xác định bởi VPS được báo hiệu tường minh và đối với mỗi OLS, các lớp đầu ra được báo hiệu tường minh và các lớp khác là các lớp mà là các lớp tham chiếu trực tiếp hoặc gián tiếp của các lớp đầu ra của OLS.

Giá trị của `ols_mode_idc` có thể nằm trọn trong khoảng từ 0 đến 2. Giá trị 3 của `ols_mode_idc` được dành riêng để sử dụng sau này b[ir ITU-T | ISO/IEC.

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

`num_output_layer_sets_minus1` cộng 1 xác định tổng số OLS được xác định bởi VPS khi `ols_mode_idc` bằng 2.

Biến `TotalNumOlss`, xác định tổng số OLS được xác định bởi VPS, được dẫn xuất như sau:

```

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

```

`vps_all_layers_same_num_sublayers_flag` bằng 0 xác định rằng các lớp trong mỗi CVS viện dẫn đến VPS có thể hoặc không thể có cùng số lượng lớp phụ thời gian. Khi không xuất hiện, giá trị của `vps_all_layers_same_num_sublayers_flag` được suy ra bằng 1.

`vps_all_independent_layers_flag` bằng 1 xác định rằng tất cả các lớp trong CVS được mã hóa độc lập mà không sử dụng dự báo ngoài lớp.

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

Biến `NumOutputLayersInOls[i]`, xác định số lượng lớp đầu ra trong OLS thứ *i*, biến `NumSubLayersInLayerInOLS[i][j]`, xác định số lượng lớp con trong lớp thứ *j* trong OLS thứ *i*, biến `OutputLayerIdInOls[i][j]`, xác định giá trị `nuh_layer_id` của lớp đầu ra thứ *j* trong OLS thứ *i*, và biến `LayerUsedAsOutputLayerFlag[k]`, xác định liệu lớp thứ *k* được sử dụng làm lớp đầu ra trong ít nhất một OLS, được dẫn xuất như sau:

```

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

```

```

        OutputLayerIdInOls[i][0] = vps_layer_id[i]
        for(j = 0; j < i && (ols_mode_idc == 0); j++)
            NumSubLayersInLayerInOLS[i][j] =
max_tid_il_ref_pics_plus1[i]
            NumSubLayersInLayerInOLS[i][i] =
vps_max_sub_layers_minus1 + 1
    } else if(ols_mode_idc == 1) {
        NumOutputLayersInOls[i] = i + 1
        for(j = 0; j < NumOutputLayersInOls[i]; j++) {
            OutputLayerIdInOls[i][j] = vps_layer_id[j]
            NumSubLayersInLayerInOLS[i][j] =
vps_max_sub_layers_minus1 + 1
        }
    } else if(ols_mode_idc == 2) {
        for(j = 0; j <= vps_max_layers_minus1; j++) {
            layerBao_includedInOlsFlag[i][j] = 0
            NumSubLayersInLayerInOLS[i][j] = 0
        }
        for(k = 0, j = 0; k <= vps_max_layers_minus1; k++)
            if(ols_output_layer_flag[i][k]) {
                layerIncludedInOlsFlag[i][k] = 1
                LayerUsedAsOutputLayerFlag[k] = 1
                OutputLayerIdx[i][j] = k
                OutputLayerIdInOls[i][j++] = vps_layer_id[k]
                NumSubLayersInLayerInOLS[i][j] =
vps_max_sub_layers_minus1 + 1
            }
        NumOutputLayersInOls[i] = j
        for(j = 0; j < NumOutputLayersInOls[i]; j++) {
            idx = OutputLayerIdx[i][j]
            for(k = 0; k < NumRefLayers[idx]; k++) {

```

```

        layerBao includedInOlsFlag[i][RefLayerIdx[idx][k]] =
1
        if(NumSubLayersInLayerInOLS[i][RefLayerIdx[idx][k]] <
            max_tid_il_ref_pics_plus1[OutputLayerIdInOls[i][j]])
            NumSubLayersInLayerInOLS[i][RefLayerIdx[idx][k]] =
                max_tid_il_ref_pics_plus1[OutputLayerIdInOls[i][j]]
    }
}
}

```

Đối với mỗi giá trị của i nằm trọn trong khoảng từ 0 đến $vps_max_layers_minus1$, các giá trị của $LayerUsedAsRefLayerFlag[i]$ và $LayerUsedAsOutputLayerFlag[i]$ có thể đều không bằng 0. Nói cách khác, có thể không có lớp vốn không phải là lớp đầu ra của ít nhất một OLS hoặc lớp tham chiếu trực tiếp của lớp khác bất kỳ.

Đối với mỗi OLS, có thể có ít nhất một lớp vốn là lớp đầu ra. Nói cách khác, đối với giá trị bất kỳ của i nằm trọn trong khoảng từ 0 đến $TotalNumOlss - 1$, giá trị của $NumOutputLayersInOls[i]$ có thể lớn hơn hoặc bằng 1.

Biến $NumLayersInOls[i]$, xác định số lượng lớp trong OLS thứ i , và biến $LayerIdInOls[i][j]$, xác định giá trị nuh_layer_id của lớp thứ j trong OLS thứ i , được dẫn xuất như sau:

```

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

```

```

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

```

Biến `OlsLayerIdx[i][j]`, xác định chỉ số lớp OLS của lớp có `nuh_layer_id` bằng `LayerIdInOls[i][j]`, được dẫn xuất như sau:

```

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

```

Lớp thấp nhất trong mỗi OLS có thể là lớp độc lập. Nói cách khác, đối với mỗi i nằm trọn trong khoảng từ 0 đến $\text{TotalNumOlss} - 1$, giá trị của `vps_independent_layer_flag[GeneralLayerIdx[LayerIdInOls[i][0]]]` có thể bằng 1.

Mỗi lớp có thể được bao gồm trong ít nhất một OLS được xác định bởi VPS. Nói cách khác, đối với mỗi lớp có giá trị cụ thể của `nuh_layer_id` `nuhLayerId` bằng một trong `vps_layer_id[k]` với k nằm trọn trong khoảng từ 0 đến `vps_max_layers_minus1`, có thể có ít nhất một cặp giá trị của i và j , trong đó i nằm trọn trong khoảng từ 0 đến $\text{TotalNumOlss} - 1$, và j nằm trọn trong khoảng từ $\text{NumLayersInOls}[i] - 1$, sao cho giá trị của `LayerIdInOls[i][j]` bằng `nuhLayerId`.

Theo phương án thực hiện, tiến trình giải mã hoạt động như sau đối với ảnh hiện tại `CurrPic`:

- `PictureOutputFlag` được thiết lập như sau:
- Nếu một trong các điều kiện sau là đúng, `PictureOutputFlag` được gán bằng 0:
 - Ảnh hiện tại là ảnh RASL và `NoOutputBeforeRecoveryFlag` của ảnh IRAP liên kết bằng 1.

- `gdr_enabled_flag` bằng 1 và ảnh hiện tại là ảnh GDR có `NoOutputBeforeRecoveryFlag` bằng 1.

- `gdr_enabled_flag` bằng 1, ảnh hiện tại được liên kết với ảnh GDR với `NoOutputBeforeRecoveryFlag` bằng 1, và `PicOrderCntVal` của ảnh hiện tại nhỏ hơn `RpPicOrderCntVal` của ảnh GDR được liên kết.

- `sps_video_parameter_set_id` lớn hơn 0, `ols_mode_idc` bằng 0 và AU hiện tại chứa ảnh `picA` thỏa mãn tất cả các điều kiện sau:

- `PicA` có `PictureOutputFlag` bằng 1.

- `PicA` có `nuh_layer_id` `nuhLid` lớn hơn của ảnh hiện tại.

- `PicA` thuộc lớp đầu ra của OLS (tức là, `OutputLayerIdInOls[TargetOlsIdx][0]` bằng `nuhLid`).

- `sps_video_parameter_set_id` lớn hơn 0, `ols_mode_idc` bằng 2, và `ols_output_layer_flag[TargetOlsIdx][GeneralLayerIdx[nuh_layer_id]]` bằng 0.

- Ngược lại, `PictureOutputFlag` được thiết lập bằng `pic_output_flag`.

Sau khi tất cả các lát của ảnh hiện tại đã được giải mã, ảnh đang được giải mã được đánh dấu là “được sử dụng cho tham chiếu ngắn hạn”, và mỗi danh mục ILRP trong `RefPicList[0]` hoặc `RefPicList[1]` được đánh dấu là “được sử dụng cho tham chiếu ngắn hạn”.

Theo phương án thực hiện tương tự hoặc khác, khi mỗi lớp là tập hợp lớp đầu ra, `PictureOutputFlag` được thiết lập bằng `pic_output_flag`, bất kể giá trị của `ols_mode_idc`.

Theo phương án thực hiện tương tự hoặc khác, `PictureOutputFlag` được gán bằng 0 khi `sps_video_parameter_set_id` lớn hơn 0, `each_layer_is_an_ols_flag` bằng 0, `ols_mode_idc` bằng 0 và AU hiện tại chứa ảnh `picA` thỏa mãn tất cả các điều kiện sau: `PicA` có `PictureOutputFlag` bằng 1, `PicA` có `nuh_layer_id` `nuhLid` lớn hơn của ảnh hiện tại và `PicA` thuộc lớp đầu ra của OLS (tức là, `OutputLayerIdInOls[TargetOlsIdx][0]` bằng `nuhLid`).

Theo phương án thực hiện tương tự hoặc khác, `PictureOutputFlag` được gán bằng 0 khi `sps_video_parameter_set_id` lớn hơn 0, `each_layer_is_an_ols_flag` bằng 0, `ols_mode_idc` bằng 2, và `ols_output_layer_flag[TargetOlsIdx][GeneralLayerIdx[nuh_layer_id]]` bằng 0.

Fig.23 thể hiện ví dụ của bảng cú pháp để chỉ báo tập hợp lớp đầu ra có chế độ OLS theo phương án thực hiện.

Fig.24 thể hiện sơ đồ khối của tiến trình giải mã cho dòng bit, theo khía cạnh của sáng chế. Cụ thể là, Fig.24 thể hiện lưu đồ phía bộ giải mã để chỉ báo tập hợp lớp đầu ra có chế độ OLS theo phương án thực hiện.

Theo khía cạnh của sáng chế, phương pháp giải mã có thể bao gồm: nhận dòng bit bao gồm dữ liệu ảnh/video được nén (Hoạt động 1001 trên Fig.24). Dòng bit có thể có nhiều lớp.

Phương pháp giải mã có thể còn bao gồm hoạt động 1002 phân tách hoặc dẫn xuất, từ dòng bit, bộ chỉ báo chế độ tập hợp lớp đầu ra (chẳng hạn, `ols_mode_idc`) trong VPS.

Phương pháp giải mã có thể còn bao gồm hoạt động 1003 nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra.

Phương pháp giải mã có thể còn bao gồm hoạt động 1004 (chẳng hạn, các hoạt động 1004A, 1004B, hoặc 1004C) nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tập hợp lớp đầu ra được nhận diện.

Phương pháp giải mã có thể còn bao gồm hoạt động 1005 giải mã một hoặc nhiều lớp đầu ra ảnh được nhận diện. Một hoặc nhiều lớp đầu ra ảnh được giải mã có thể được hiển thị.

Việc nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra có thể bao gồm: trong trường hợp mà bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS là giá trị thứ nhất, nhận diện lớp cao nhất trong dòng bit dưới dạng một hoặc nhiều lớp đầu ra ảnh (xem, chẳng hạn, Fig.25A); trong trường hợp mà bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS là giá trị thứ hai, nhận diện tất cả các lớp trong dòng bit dưới dạng một hoặc nhiều lớp đầu ra ảnh (xem, chẳng hạn, Fig.25B); và trong trường hợp mà bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS là giá trị thứ ba, nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tường minh trong VPS (xem, chẳng hạn Fig.25C).

Fig.25A đến Fig.25C thể hiện hình ảnh đầu ra như được gạch chân. Như được thể hiện trên Fig.25A đến Fig.25C, có thể có năm lớp trong dòng bit và chỉ

các lớp cụ thể được xuất ra bộ phận hiển thị. Chẳng hạn, Lớp 3 có thể được xuất ra (chẳng hạn, xem, Fig.25C).

Theo phương án thực hiện, `Ols_mode_idc` trong VPS có thể chỉ báo phương pháp (cơ chế) báo hiệu OLS. Chẳng hạn, nếu bằng 0, lớp cao nhất trong dòng bit có thể là lớp đầu ra duy nhất, nếu bằng 1, tất cả các lớp trong dòng bit có thể là các lớp đầu ra, và nếu bằng 2, một hoặc nhiều lớp đầu ra có thể được báo hiệu tường minh trong VPS. Tức là, đầu ra ảnh có thể được xác định bằng cách báo hiệu tập hợp lớp đầu ra.

Chẳng hạn, đầu ra ảnh của mỗi lớp có thể được xác định bằng cách báo hiệu tập hợp lớp đầu ra, và phương pháp báo hiệu lớp đầu ra có thể được xác định bằng `ols_mode_idc`.

Theo phương án thực hiện, mỗi dòng bit có thể báo hiệu chế độ lớp đầu ra và chế độ lớp đầu ra có thể thay đổi theo thời gian.

Theo phương án thực hiện, như được thể hiện trên Fig.25A, có thể có 5 lớp (Lớp 4 – Lớp 0) trong dòng bit. Như được thể hiện trên Fig.25A, ở thời gian = K, trong đó bộ chỉ báo chế độ tập hợp lớp đầu ra (`ols_mode_idc`) = 0, lớp cao nhất có thể được xuất ra. Do vậy, như được thể hiện trên Fig.25A, ở Time = K, ảnh đầu ra ở lớp 4 được xuất ra. Ở Time = K+1, Lớp 3, vốn là lớp cao nhất có thể được xuất ra, v.v.. với K+2 và K+3.

Như được thể hiện trên Fig.25B, trong đó bộ chỉ báo chế độ tập hợp lớp đầu ra (`ols_mode_idc`) bằng 1, tất cả các lớp có thể được xuất ra, chẳng hạn, ở mỗi thời gian từ K đến K+3.

Như được thể hiện trên Fig.25C, báo hiệu tường minh có thể được thực hiện. Chẳng hạn, như được thể hiện trên Fig.26, lớp ảnh 3 (được thể hiện dưới dạng đầu ra bằng cách gạch chân) là lớp đầu ra trong suốt, chẳng hạn, thời gian từ K đến K+3.

Giá trị thứ nhất có thể khác với giá trị thứ hai và có thể khác với giá trị thứ ba, và giá trị thứ hai có thể khác với giá trị thứ ba.

Giá trị thứ nhất có thể bằng 0, giá trị thứ hai có thể bằng 1, và giá trị thứ ba có thể bằng 2. Tuy nhiên, các giá trị khác có thể được sử dụng và sáng chế không bị giới hạn ở việc sử dụng 0, 1 và 2 như nêu trên.

Việc nhận diện một hoặc nhiều lớp đầu ra ảnh bằng cách báo hiệu tường minh trong VPS có thể bao gồm: (i) phân tách hoặc dẫn xuất, từ VPS, cờ lớp đầu ra, và (ii) các lớp thiết lập có cờ lớp đầu ra bằng 1 be một hoặc nhiều lớp đầu ra ảnh.

Việc nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ lớp đầu ra có thể bao gồm: trong trường hợp mà bộ chỉ báo chế độ lớp đầu ra trong VPS là giá trị định trước, báo hiệu tập hợp lớp đầu ra bao gồm nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tường minh trong VPS.

Việc nhận diện một hoặc nhiều lớp đầu ra ảnh bằng cách báo hiệu tường minh trong VPS bao gồm: (i) phân tách hoặc dẫn xuất, từ VPS, cờ lớp đầu ra, và (ii) các lớp thiết lập có cờ lớp đầu ra bằng 1 là một hoặc nhiều lớp đầu ra ảnh, trong đó số lượng nhiều lớp lớn hơn 2.

Báo hiệu tập hợp lớp đầu ra có thể bao gồm nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tường minh trong VPS khi bộ chỉ báo chế độ tập hợp lớp đầu ra bằng 2, và số lượng lớp trong nhiều lớp lớn hơn 2.

Báo hiệu tập hợp lớp đầu ra có thể bao gồm nhận diện lớp cao nhất trong dòng bit hoặc tất cả các lớp trong dòng bit dưới dạng một hoặc nhiều lớp đầu ra ảnh bằng cách suy ra một hoặc nhiều lớp đầu ra ảnh, khi bộ chỉ báo chế độ tập hợp lớp đầu ra nhỏ hơn 2, và số lượng nhiều lớp bằng 2, và bộ chỉ báo chế độ tập hợp lớp đầu ra thực sự nhỏ hơn 2, và số lượng nhiều lớp thực sự bằng 2.

Theo phương án thực hiện, bộ chỉ báo `number_output_layer_sets_minus1` trong VPS chỉ báo số lượng tập hợp lớp đầu ra.

Theo phương án thực hiện, bộ chỉ báo `VPS_maximum_layer_minus1` trong VPS chỉ báo số lượng lớp trong dòng bit.

Theo phương án thực hiện, cờ tập hợp lớp đầu ra `[i][j]` trong VPS chỉ báo liệu lớp thứ `j` của tập hợp lớp đầu ra thứ `i` có phải là lớp đầu ra hoặc không.

Theo phương án thực hiện, nếu tất cả nhiều lớp là độc lập trong đó `vps_all_independent_layer_flag` của VPS bằng 1, bộ chỉ báo chế độ tập hợp lớp đầu ra không được báo hiệu, và giá trị của bộ chỉ báo chế độ tập hợp lớp đầu ra được suy ra là giá trị thứ hai.

Theo phương án thực hiện, khi mỗi lớp là tập hợp lớp đầu ra, cờ ảnh đầu ra của VPS được thiết lập bằng cờ đầu ra ảnh được báo hiệu trong tiêu đề ảnh, bất kể giá trị của bộ chỉ báo chế độ tập hợp lớp đầu ra.

Lưu ý: ảnh trong lớp đầu ra có thể hoặc không thể có PictureOutputFlag bằng 1. ảnh trong lớp không phải đầu ra có PictureOutputFlag bằng 0. Ảnh có PictureOutputFlag bằng 1 được xuất ra để hiển thị. Ảnh có PictureOutputFlag bằng 0 không được xuất ra để hiển thị.

Theo phương án thực hiện, cờ đầu ra ảnh được thiết lập bằng 0 khi ID SPS VPS lớn hơn 0, mà chỉ báo rằng nhiều hơn một lớp hiện có trong dòng bit, mỗi lớp là cờ tập hợp lớp đầu ra của VPS bằng 0, mà chỉ báo rằng nhiều lớp trong dòng bit đều không độc lập, bộ chỉ báo chế độ tập hợp lớp đầu ra bằng 0 và khối truy nhập hiện tại chứa ảnh thỏa mãn tất cả các điều kiện sau bao gồm: có cờ đầu ra ảnh bằng 1, có ID lớp NUH lớn hơn ID của ảnh hiện tại và thuộc lớp đầu ra của tập hợp lớp đầu ra.

Theo phương án thực hiện, cờ ảnh đầu ra của VPS được thiết lập bằng 0 khi SPS của VPS lớn hơn 0, `each_layer_is_an_ols_flag` bằng 0, bộ chỉ báo chế độ tập hợp lớp đầu ra bằng 2, và `ols_output_layer_flag [Target OLS Index][General Layer Index [nuh layer identifier]]` bằng 0.

Theo phương án thực hiện, phương pháp có thể còn bao gồm: điều khiển bộ hiển thị để hiển thị một hoặc nhiều lớp đầu ra ảnh được giải mã.

Theo khía cạnh của sáng chế, vật ghi máy tính đọc được bất biến lưu trữ các lệnh mà, khi được thực thi, khiến hệ thống hoặc thiết bị bao gồm một hoặc nhiều bộ xử lý: nhận dòng bit bao gồm dữ liệu ảnh/video được nén, trong đó dòng bit có nhiều lớp; phân tách hoặc dẫn xuất, từ dòng bit, bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS; nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra; nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tập hợp lớp đầu ra được nhận diện; và giải mã một hoặc nhiều lớp đầu ra ảnh được nhận diện.

Theo phương án thực hiện, các lệnh còn được tạo cấu hình để khiến hệ thống hoặc thiết bị bao gồm một hoặc nhiều bộ xử lý: điều khiển bộ hiển thị để hiển thị một hoặc nhiều lớp đầu ra ảnh được giải mã.

Theo khía cạnh của sáng chế, thiết bị có thể bao gồm: ít nhất một bộ nhớ được tạo cấu hình để lưu trữ mã chương trình máy tính; và ít nhất một bộ xử lý được tạo cấu hình để truy nhập ít nhất một bộ nhớ và hoạt động theo mã chương trình máy tính. Theo phương án thực hiện, mã chương trình máy tính có thể bao gồm: mã nhận được tạo cấu hình để khiến ít nhất một bộ xử lý nhận dòng bit bao gồm dữ liệu ảnh/video được nén, trong đó dòng bit có nhiều lớp; mã phân tách hoặc dẫn xuất được tạo cấu hình để khiến ít nhất một bộ xử lý phân tách hoặc dẫn xuất, từ dòng bit, bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS; mã nhận diện báo hiệu lớp đầu ra được tạo cấu hình để khiến ít nhất một bộ xử lý nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra; mã nhận diện lớp đầu ra ảnh được tạo cấu hình để khiến ít nhất một bộ xử lý nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tập hợp lớp đầu ra được nhận diện; và mã giải mã được tạo cấu hình để khiến ít nhất một bộ xử lý: giải mã một hoặc nhiều lớp đầu ra ảnh được nhận diện.

Theo phương án thực hiện, mã chương trình máy tính có thể còn bao gồm: mã điều khiển hiển thị được tạo cấu hình để khiến ít nhất một bộ xử lý điều khiển bộ hiển thị để hiển thị một hoặc nhiều lớp đầu ra ảnh.

Theo phương án thực hiện, giá trị của `temporalId` của khối NAL VCL bất kỳ trong chuỗi video được mã hóa có thể không lớn hơn `vps_max_sublayers_minus1` trong VPS được tham chiếu bởi khối VCL NAL.

Theo phương án thực hiện tương tự hoặc khác, `max_tid_il_ref_pics_plus1[i]` có các ngữ nghĩa sau:

`max_tid_il_ref_pics_plus1[i]` bằng 0 xác định rằng dự báo ngoài lớp không được sử dụng bởi các ảnh P phi IRA của lớp thứ *i*. `max_tid_il_ref_pics_plus1[i]` lớn hơn 0 xác định rằng, để giải mã các ảnh của lớp thứ *i*, không ảnh nào có `TemporalId` lớn hơn `max_tid_il_ref_pics_plus1[i] - 1` được sử dụng làm ILRP. Khi không xuất hiện, giá trị của `max_tid_il_ref_pics_plus1[i]` được suy ra bằng `vps_max_sublayers_minus1 + 1`.

Theo phương án thực hiện tương tự hoặc khác, `max_tid_il_ref_pics_plus1[i]` có thể không lớn hơn `vps_max_sublayers_minus1 + 1`.

Theo phương án thực hiện, tập hợp lớp đầu ra theo lớp con có thể không được dẫn xuất cho các lớp độc lập. Do vậy, `max_tid_il_ref_pics_plus1` và `layerIncludedInOlsFlag` có thể không được dẫn xuất cho các lớp độc lập.

Theo phương án thực hiện tương tự khác, biến `NumOutputLayersInOls[i]`, xác định số lượng lớp đầu ra trong OLS thứ `i`, biến `NumSubLayersInLayerInOLS[i][j]`, xác định số lượng lớp con trong lớp thứ `j` trong OLS thứ `i`, biến `OutputLayerIdInOls[i][j]`, xác định giá trị `nuh_layer_id` của lớp đầu ra thứ `j` trong OLS thứ `i`, và biến `LayerUsedAsOutputLayerFlag[k]`, xác định liệu lớp thứ `k` được sử dụng dưới dạng lớp đầu ra trong ít nhất một OLS, được dẫn xuất như sau:

```

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

```

```

        for(j = 0; j < NumOutputLayersInOls[i]; j++) {
            OutputLayerIdInOls[i][j] = vps_layer_id[j]
            NumSubLayersInLayerInOLS[i][j] =
vps_max_sub_layers_minus1 + 1
        }
    } else if(ols_mode_idc == 2) {
        for(j = 0; j <= vps_max_layers_minus1; j++) {
            layerIncludedInOlsFlag[i][j] = 0
            NumSubLayersInLayerInOLS[i][j] = 0
        }
        for(k = 0, j = 0; k <= vps_max_layers_minus1; k++)
            if(ols_output_layer_flag[i][k]) {
                layerIncludedInOlsFlag[i][k] = 1
                LayerUsedAsOutputLayerFlag[k] = 1
                OutputLayerIdx[i][j] = k
                OutputLayerIdInOls[i][j] = vps_layer_id[k]
                NumSubLayersInLayerInOLS[i][j++] =
vps_max_sub_layers_minus1 + 1
            }
        if(!vps_all_independent_layers_flag) {
            NumOutputLayersInOls[i] = j
            for(j = 0; j < NumOutputLayersInOls[i]; j++) {
                idx = OutputLayerIdx[i][j]
                if(!vps_independent_layer_flag[idx])
                    for(k = 0; k < NumRefLayers[idx]; k++) {
                        layerIncludedInOlsFlag[i][RefLayerIdx[idx][k]] = 1
                        if(NumSubLayersInLayerInOLS[i][RefLayerIdx[idx][k]] <
max_tid_il_ref_pics_plus1[OutputLayerIdInOls[i][j]])
                            NumSubLayersInLayerInOLS[i][RefLayerIdx[idx][k]] =
max_tid_il_ref_pics_plus1[OutputLayerIdInOls[i][j]]
                    }
            }
        }
    }
}

```

```

    }
  }
}

```

Fig.23 thể hiện ví dụ của bảng cú pháp SPS RBSP. Trong cấu trúc cú pháp này, `sps_max_sublayers_minus1` plus 1 xác định số lượng lớn nhất của các lớp con thời gian mà có thể hiện có trong mỗi CLVS đề cập đến SPS. Giá trị của `sps_max_sublayers_minus1` có thể nằm trọn trong khoảng từ 0 đến `vps_max_sublayers_minus1`.

Theo phương án thực hiện, khi tất cả các lớp trong chuỗi video được mã hóa có số lượng lớn con tương tự, số lượng lớp con của mỗi lớp có thể bằng số lượng lớn nhất của các lớp con được báo hiệu trong VPS.

Theo phương án thực hiện, khi giá trị của `vps_all_layers_same_num_sublayers_flag` bằng 1, giá trị của `sps_max_sublayers_minus1` trong SPS chỉ đến VPS sẽ bằng giá trị của `vps_max_sublayers_minus1`.

Theo phương án thực hiện, khi `sps_video_parameter_set_id` lớn hơn 0 và `vps_all_layers_same_num_sublayers_flag` bằng 1, `sps_max_sublayers_minus1` sẽ bằng `vps_max_sublayers_minus1`.

Theo phương án thực hiện, tiến trình giải mã hoạt động như sau đối với ảnh hiện tại CurrPic:

- `PictureOutputFlag` được thiết lập như sau:
- Nếu một trong các điều kiện sau là đúng, `PictureOutputFlag` được gán bằng 0:
 - ảnh hiện tại là ảnh RASL và `NoOutputBeforeRecoveryFlag` của ảnh IRAP liên kết bằng 1.
 - `gdr_enabled_flag` bằng 1 và ảnh hiện tại là ảnh GDR có `NoOutputBeforeRecoveryFlag` bằng 1.

- `gdr_enabled_flag` bằng 1, ảnh hiện tại được liên kết với ảnh GDR với `NoOutputBeforeRecoveryFlag` bằng 1, và `PicOrderCntVal` of ảnh hiện tại nhỏ hơn `RpPicOrderCntVal` của ảnh GDR được liên kết.

- `sps_video_parameter_set_id` lớn hơn 0, `ols_mode_idc` bằng 0 và AU hiện tại chứa ảnh `picA` thỏa mãn tất cả các điều kiện sau:

- `PicA` có `PictureOutputFlag` bằng 1.

- `PicA` có `nuh_layer_id` `nuhLid` lớn hơn của ảnh hiện tại.

- `PicA` thuộc lớp đầu ra của OLS (tức là, `OutputLayerIdInOls[TargetOlsIdx][0]` bằng `nuhLid`).

- `sps_video_parameter_set_id` lớn hơn 0, `ols_mode_idc` bằng 2, và `ols_output_layer_flag[TargetOlsIdx][GeneralLayerIdx[nuh_layer_id]]` bằng 0.

- Ngược lại, `PictureOutputFlag` được thiết lập bằng `pic_output_flag`.

Sau khi tất cả các lát của ảnh hiện tại đã được giải mã, ảnh đang được giải mã được đánh dấu là “được sử dụng cho tham chiếu ngắn hạn”, và mỗi danh mục ILRP trong `RefPicList[0]` hoặc `RefPicList[1]` được đánh dấu là “được sử dụng cho tham chiếu ngắn hạn”.

Theo phương án thực hiện tương tự hoặc khác, khi mỗi lớp là tập hợp lớp đầu ra, `PictureOutputFlag` được thiết lập bằng `pic_output_flag`, bất kể giá trị của `ols_mode_idc`.

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

Theo phương án thực hiện tương tự hoặc khác, `PictureOutputFlag` được thiết lập bằng 0 khi `sps_video_parameter_set_id` lớn hơn 0, `each_layer_is_an_ols_flag` bằng 0, `ols_mode_idc` bằng 2, và `ols_output_layer_flag[TargetOlsIdx][GeneralLayerIdx[nuh_layer_id]]` bằng 0.

Fig.23 thể hiện ví dụ của bảng cú pháp để chỉ báo tập hợp lớp đầu ra có chế độ OLS theo phương án thực hiện.

Fig.24 thể hiện sơ đồ khối của tiến trình giải mã cho dòng bit, theo khía cạnh của sáng chế. Cụ thể là, Fig.24 thể hiện lưu đồ phía bộ giải mã để chỉ báo tập hợp lớp đầu ra có chế độ OLS theo phương án thực hiện.

Theo khía cạnh của sáng chế, phương pháp giải mã có thể bao gồm: nhận dòng bit bao gồm dữ liệu ảnh/video được nén (Hoạt động 1001 trên Fig.24). Dòng bit có thể có nhiều lớp.

Phương pháp giải mã có thể còn bao gồm hoạt động 1002 phân tách hoặc dẫn xuất, từ dòng bit, bộ chỉ báo chế độ tập hợp lớp đầu ra (chẳng hạn, `ols_mode_idc`) trong VPS.

Phương pháp giải mã có thể còn bao gồm hoạt động 1003 nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra.

Phương pháp giải mã có thể còn bao gồm hoạt động 1004 (chẳng hạn, các hoạt động 1004A, 1004B, hoặc 1004C) nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tập hợp lớp đầu ra được nhận diện.

Phương pháp giải mã có thể còn bao gồm hoạt động 1005 giải mã một hoặc nhiều lớp đầu ra ảnh được nhận diện. Một hoặc nhiều lớp đầu ra ảnh được giải mã có thể được hiển thị.

Việc nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra có thể bao gồm: trong trường hợp mà bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS là giá trị thứ nhất, nhận diện lớp cao nhất trong dòng bit dưới dạng một hoặc nhiều lớp đầu ra ảnh (xem, chẳng hạn, Fig.25A); trong trường hợp mà bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS là giá trị thứ hai, nhận diện tất cả các lớp trong dòng bit dưới dạng một hoặc nhiều lớp đầu ra ảnh (xem, chẳng hạn, Fig.25B); và trong trường hợp mà bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS là giá trị thứ ba, nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tường minh trong VPS (xem, chẳng hạn Fig.25C).

Fig.25A đến Fig.25C thể hiện hình ảnh đầu ra như được gạch chân. Như được thể hiện trên Fig.25A đến Fig.25C, có thể có năm lớp trong dòng bit và chỉ các lớp cụ thể được xuất ra bộ phận hiển thị. Chẳng hạn, Lớp 3 có thể được xuất ra (chẳng hạn, xem, Fig.25C).

Theo phương án thực hiện, `Ols_mode_idc` trong VPS có thể chỉ báo phương pháp (cơ chế) báo hiệu OLS. Chẳng hạn, nếu bằng 0, lớp cao nhất trong dòng bit có thể là lớp đầu ra duy nhất, nếu bằng 1, tất cả các lớp trong dòng bit có thể là các lớp đầu ra, và nếu bằng 2, một hoặc nhiều lớp đầu ra có thể được báo hiệu tường minh trong VPS. Tức là, đầu ra ảnh có thể được xác định bằng cách báo hiệu tập hợp lớp đầu ra.

Chẳng hạn, đầu ra ảnh của mỗi lớp có thể được xác định bằng cách báo hiệu tập hợp lớp đầu ra, và phương pháp báo hiệu lớp đầu ra có thể được xác định bằng `ols_mode_idc`.

Theo phương án thực hiện, mỗi dòng bit có thể báo hiệu chế độ lớp đầu ra và chế độ lớp đầu ra có thể thay đổi theo thời gian.

Theo phương án thực hiện, như được thể hiện trên Fig.25A, có thể có 5 lớp (Lớp 4 – Lớp 0) trong dòng bit. Như được thể hiện trên Fig.25A, ở thời gian = K, trong đó bộ chỉ báo chế độ tập hợp lớp đầu ra (`ols_mode_idc`) = 0, lớp cao nhất có thể được xuất ra. Do vậy, như được thể hiện trên Fig.25A, ở Time = K, ảnh đầu ra ở lớp 4 được xuất ra. Ở Time = K+1, Lớp 3, vốn là lớp cao nhất có thể được xuất ra, v.v.. với K+2 và K+3.

Như được thể hiện trên Fig.25B, trong đó bộ chỉ báo chế độ tập hợp lớp đầu ra (`ols_mode_idc`) bằng 1, tất cả các lớp có thể được xuất ra, chẳng hạn, ở mỗi thời gian từ K đến K+3.

Như được thể hiện trên Fig.25C, báo hiệu tường minh có thể được thực hiện. Chẳng hạn, như được thể hiện trên Fig.26, lớp ảnh 3 (được thể hiện dưới dạng đầu ra bằng cách gạch chân) là lớp đầu ra trong suốt, chẳng hạn, thời gian từ K đến K+3.

Giá trị thứ nhất có thể khác với giá trị thứ hai và có thể khác với giá trị thứ ba, và giá trị thứ hai có thể khác với giá trị thứ ba.

Giá trị thứ nhất có thể bằng 0, giá trị thứ hai có thể bằng 1, và giá trị thứ ba có thể bằng 2. Tuy nhiên, các giá trị khác có thể được sử dụng và sáng chế không bị giới hạn ở việc sử dụng 0, 1 và 2 như nêu trên.

Việc nhận diện một hoặc nhiều lớp đầu ra ảnh bằng cách báo hiệu tường minh trong VPS có thể bao gồm: (i) phân tách hoặc dẫn xuất, từ VPS, cờ lớp đầu

ra, và (ii) các lớp thiết lập có cờ lớp đầu ra bằng 1 be một hoặc nhiều lớp đầu ra ảnh.

Việc nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ lớp đầu ra có thể bao gồm: trong trường hợp mà bộ chỉ báo chế độ lớp đầu ra trong VPS là giá trị định trước, the báo hiệu tập hợp lớp đầu ra bao gồm nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tường minh trong VPS.

Việc nhận diện một hoặc nhiều lớp đầu ra ảnh bằng cách báo hiệu tường minh trong VPS bao gồm: (i) phân tách hoặc dẫn xuất, từ VPS, cờ lớp đầu ra, và (ii) các lớp thiết lập có cờ lớp đầu ra bằng 1 là một hoặc nhiều lớp đầu ra ảnh, trong đó số lượng nhiều lớp lớn hơn 2.

Báo hiệu tập hợp lớp đầu ra có thể bao gồm nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tường minh trong VPS khi bộ chỉ báo chế độ tập hợp lớp đầu ra bằng 2, và số lượng lớp trong nhiều lớp lớn hơn 2.

Báo hiệu tập hợp lớp đầu ra có thể bao gồm nhận diện lớp cao nhất trong dòng bit hoặc tất cả các lớp trong dòng bit dưới dạng một hoặc nhiều lớp đầu ra ảnh bằng cách suy ra một hoặc nhiều lớp đầu ra ảnh, khi bộ chỉ báo chế độ tập hợp lớp đầu ra nhỏ hơn 2, và số lượng nhiều lớp bằng 2, và bộ chỉ báo chế độ tập hợp lớp đầu ra thực sự nhỏ hơn 2, và số lượng nhiều lớp thực sự bằng 2.

Theo phương án thực hiện, bộ chỉ báo `number_output_layer_sets_minus1` trong VPS chỉ báo số lượng tập hợp lớp đầu ra.

Theo phương án thực hiện, bộ chỉ báo `VPS_maximum_layer_minus1` trong VPS chỉ báo số lượng lớp trong dòng bit.

Theo phương án thực hiện, cờ tập hợp lớp đầu ra `[i][j]` trong VPS chỉ báo liệu lớp thứ `j` của tập hợp lớp đầu ra thứ `i` có phải là lớp đầu ra hoặc không.

Theo phương án thực hiện, nếu tất cả nhiều lớp là độc lập trong đó `vps_all_independent_layer_flag` của VPS bằng 1, bộ chỉ báo chế độ tập hợp lớp đầu ra không được báo hiệu, và giá trị của bộ chỉ báo chế độ tập hợp lớp đầu ra được suy ra là giá trị thứ hai.

Theo phương án thực hiện, khi mỗi lớp là tập hợp lớp đầu ra, `vps_picture_output_flag` được thiết lập bằng cờ đầu ra ảnh được báo hiệu trong tiêu đề ảnh, bất kể giá trị của bộ chỉ báo chế độ tập hợp lớp đầu ra.

Lưu ý: ảnh trong lớp đầu ra có thể hoặc không thể có PictureOutputFlag bằng 1. Ảnh trong lớp không phải đầu ra có PictureOutputFlag bằng 0. Ảnh có PictureOutputFlag bằng 1 được xuất ra để hiển thị. Ảnh có PictureOutputFlag bằng 0 không được xuất ra để hiển thị.

Theo phương án thực hiện, cờ đầu ra ảnh được thiết lập bằng 0 khi ID SPS VPS lớn hơn 0, mà chỉ báo rằng nhiều hơn một lớp hiện có trong dòng bit, mỗi lớp là cờ tập hợp lớp đầu ra của VPS bằng 0, mà chỉ báo rằng nhiều lớp trong dòng bit đều không độc lập, bộ chỉ báo chế độ tập hợp lớp đầu ra bằng 0 và khối truy nhập hiện tại chứa ảnh thỏa mãn tất cả các điều kiện sau bao gồm: có cờ đầu ra ảnh bằng 1, có ID lớp NUH lớn hơn ID của ảnh hiện tại và thuộc lớp đầu ra của tập hợp lớp đầu ra.

Theo phương án thực hiện, picture_output_flag của VPS được gán bằng 0 khi SPS của VPS lớn hơn 0, mỗi lớp là cờ tập hợp lớp đầu ra bằng 0, bộ chỉ báo chế độ tập hợp lớp đầu ra bằng 2, và ols_output_layer_flag [Target OLS Index][General Layer Index [nuh layer identifier]] bằng 0.

Theo phương án thực hiện, phương pháp có thể còn bao gồm: điều khiển bộ hiển thị để hiển thị một hoặc nhiều lớp đầu ra ảnh được giải mã.

Theo khía cạnh của sáng chế, vật ghi máy tính đọc được bất biến lưu trữ các lệnh mà, khi được thực thi, khiến hệ thống hoặc thiết bị bao gồm một hoặc nhiều bộ xử lý: nhận dòng bit bao gồm dữ liệu ảnh/video được nén, trong đó dòng bit có nhiều lớp; phân tách hoặc dẫn xuất, từ dòng bit, bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS; nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra; nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tập hợp lớp đầu ra được nhận diện; và giải mã một hoặc nhiều lớp đầu ra ảnh được nhận diện.

Theo phương án thực hiện, các lệnh còn được tạo cấu hình để khiến hệ thống hoặc thiết bị bao gồm một hoặc nhiều bộ xử lý: điều khiển bộ hiển thị để hiển thị một hoặc nhiều lớp đầu ra ảnh được giải mã.

Theo khía cạnh của sáng chế, thiết bị có thể bao gồm: ít nhất một bộ nhớ được tạo cấu hình để lưu trữ mã chương trình máy tính; và ít nhất một bộ xử lý được tạo cấu hình để truy nhập ít nhất một bộ nhớ và hoạt động theo mã chương

trình máy tính. Theo phương án thực hiện, mã chương trình máy tính có thể bao gồm: mã nhận được tạo cấu hình để khiến ít nhất một bộ xử lý nhận dòng bit bao gồm dữ liệu ảnh/video được nén, trong đó dòng bit có nhiều lớp; mã phân tách hoặc dẫn xuất được tạo cấu hình để khiến ít nhất một bộ xử lý phân tách hoặc dẫn xuất, từ dòng bit, bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS; mã nhận diện báo hiệu lớp đầu ra được tạo cấu hình để khiến ít nhất một bộ xử lý nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra; mã nhận diện lớp đầu ra ảnh được tạo cấu hình để khiến ít nhất một bộ xử lý nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tập hợp lớp đầu ra được nhận diện; và mã giải mã được tạo cấu hình để khiến ít nhất một bộ xử lý: giải mã một hoặc nhiều lớp đầu ra ảnh được nhận diện.

Theo phương án thực hiện, mã chương trình máy tính có thể còn bao gồm: mã điều khiển hiển thị được tạo cấu hình để khiến ít nhất một bộ xử lý điều khiển bộ hiển thị để hiển thị một hoặc nhiều lớp đầu ra ảnh.

Các kỹ thuật để báo hiệu các tham số độ phân giải thích ứng nêu trên, có thể được triển khai dưới dạng phần mềm máy tính sử dụng các lệnh máy tính đọc được và được lưu trữ dạng vật lý trong một hoặc nhiều vật máy tính đọc được. Chẳng hạn, Fig.7 thể hiện hệ thống máy tính 700 thích hợp để thực hiện các phương án thực hiện cụ thể của sáng chế.

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

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

Các thành phần được thể hiện trên Fig.7 cho hệ thống máy tính 700 về bản chất lấy làm ví dụ và không nhằm giới hạn bất kỳ ở phạm vi sử dụng hoặc chức năng của phần mềm máy tính thực hiện các phương án thực hiện sáng chế. Cấu

hình của các thành phần không nên được hiểu như là có lệ thuộc hoặc yêu cầu bất kỳ liên quan đến thành phần bất kỳ hoặc kết hợp của các thành phần được minh họa trên phương án thực hiện lấy làm ví dụ của hệ thống máy tính 700.

Hệ thống máy tính 700 có thể bao gồm các thiết bị đầu vào giao diện người cụ thể. Thiết bị đầu vào giao diện người này có thể đáp lại đầu vào bởi một hoặc nhiều người dùng qua, chẳng hạn, đầu vào xúc giác (chẳng hạn: bấm phím, trượt, di chuyển găng đỡ liệu), đầu vào audio (chẳng hạn: giọng nói, vỗ tay), đầu vào hình ảnh (chẳng hạn: cử chỉ), đầu vào khứu giác (không được mô tả trên hình vẽ). Các thiết bị giao diện người cũng có thể được sử dụng để chụp các media cụ thể không nhất thiết liên quan trực tiếp đến đầu vào nhận thức của người, chẳng hạn audio (chẳng hạn: giọng nói, âm nhạc, âm thanh xung quanh), ảnh (chẳng hạn, ảnh quét, ảnh chụp thu được từ camera ảnh tĩnh), video (chẳng hạn video 2D, video 3D bao gồm video lập thể).

Các thiết bị giao diện người đầu vào có thể bao gồm một hoặc nhiều trong số (chỉ một trong số chúng được mô tả): bàn phím 701, chuột 702, bàn di chuột cảm ứng 703, màn hình cảm ứng 710, găng đỡ liệu 704, joystick 705, micro 706, máy quét 707, camera 708.

Hệ thống máy tính 700 cũng có thể bao gồm các thiết bị đầu ra giao diện người cụ thể. Các thiết bị đầu ra giao diện người này có thể kích hoạt các giác quan của một hoặc nhiều người dùng qua, chẳng hạn, đầu ra xúc giác, âm thanh, ánh sáng, và ngửi/nếm. Các thiết bị đầu ra giao diện người này có thể bao gồm các thiết bị đầu ra xúc giác (chẳng hạn phản hồi xúc giác bởi màn hình cảm ứng 710, găng đỡ liệu 704, hoặc joystick 705, nhưng cũng có thể có thiết bị phản hồi xúc giác không dùng làm thiết bị đầu vào), thiết bị đầu ra audio (chẳng hạn: loa 709, tai nghe (không được mô tả trên hình vẽ)), thiết bị đầu ra hình ảnh (chẳng hạn, màn hình 710 bao gồm màn hình CRT, màn hình LCD, màn hình plasma, màn hình OLED, mỗi màn hình có hoặc không có khả năng chạm màn hình cảm ứng, mỗi màn hình có hoặc không có khả năng phản hồi xúc giác—một số màn hình có thể xuất ra đầu ra hình ảnh 2D hoặc đầu ra nhiều hơn 3D qua các phương tiện chẳng hạn đầu ra lập thể; kính thực tế ảo (không được mô tả trên hình vẽ),

màn hình toàn ký và smoke tanks (không được mô tả trên hình vẽ)), và các máy in (không được mô tả trên hình vẽ).

Hệ thống máy tính 700 cũng có thể bao gồm các thiết bị lưu trữ người truy cập được và media liên kết chẳng hạn các phương tiện quang học bao gồm CD/DVD ROM/RW 720 có CD/DVD hoặc phương tiện tương tự 721, ổ nhớ USB 722, ổ cứng tháo được hoặc ổ trạng thái rắn 723, phương tiện từ tính kế thừa chẳng hạn băng và đĩa mềm (không được mô tả trên hình vẽ), thiết bị dựa trên ROM/ASIC/PLD chuyên dụng chẳng hạn dongle bảo mật (không được mô tả trên hình vẽ), và tương tự.

Người có hiểu biết trung bình trong lĩnh vực cũng cần hiểu rằng cụm từ “các phương tiện máy tính đọc được” như được sử dụng cùng với sáng chế không bao gồm phương tiện truyền, sóng mang, hoặc các tín hiệu khả biến khác.

Hệ thống máy tính 700 cũng có thể bao gồm giao diện với một hoặc nhiều mạng truyền thông. Mạng có thể chẳng hạn là không dây, hữu tuyến, quang học. Mạng có thể còn là cục bộ, diện rộng, đô thị, xe cộ và công nghiệp, thời gian thực, chịu trễ, và v.v. Các ví dụ của mạng bao gồm mạng cục bộ chẳng hạn Ethernet, LAN không dây, mạng tế bào để bao gồm GSM, 3G, 4G, 5G, LTE và tương tự, mạng số diện rộng không dây hoặc hữu tuyến TV để bao gồm TV cáp, TV vệ tinh, và TV phát sóng mặt đất, xe cộ và công nghiệp để bao gồm CANBus, và v.v. Mạng cụ thể thường yêu cầu các bộ điều hợp giao diện mạng bên ngoài gắn với các cổng dữ liệu đa năng hoặc buýt ngoại vi cụ thể (749) (chẳng hạn, chẳng hạn các cổng USB của hệ thống máy tính 700; các thiết bị 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 vào hệ thống máy tính PC hoặc giao diện mạng tế bào vào hệ thống máy tính điện thoại thông minh). Nhờ sử dụng các mạng này, hệ thống máy tính 700 có thể truyền thông với các thực thể khác. Truyền thông này có thể là đơn hướng, chỉ nhận (chẳng hạn, phát sóng TV), đơn hướng chỉ gửi (chẳng hạn CANbus đến các thiết bị CANbus cụ thể), hoặc hai hướng, chẳng hạn đến các hệ thống máy tính khác nhờ sử dụng mạng số cục bộ hoặc diện rộng. Các giao thức và các xếp chồng giao

thức cụ thể có thể được sử dụng trên mỗi trong số các mạng và giao diện mạng như nêu trên.

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

Lõi 740 có thể bao gồm một hoặc nhiều CPU 741, GPU 742, khối xử lý lập trình được chuyên dụng ở dạng mảng công dụng trường lập trình được (FPGA) 743, các bộ gia tốc phần cứng cho các tác vụ cụ thể 744, và v.v. Các thiết bị này, cùng với bộ nhớ chỉ đọc (ROM) 745, bộ nhớ truy nhập ngẫu nhiên (RAM) 746, bộ lưu trữ khối nội bộ chẳng hạn ổ cứng trong người dùng không truy nhập được, ổ trạng thái rắn (SSD), và tương tự 747, có thể được kết nối qua buýt hệ thống 748. Trong một số hệ thống máy tính, buýt hệ thống 748 có thể truy nhập được ở dạng một hoặc nhiều ổ cắm vật lý để kích hoạt các phần mở rộng bởi CPU, GPU bổ sung, và tương tự. Các thiết bị ngoại vi có thể được gắn trực tiếp vào buýt hệ thống của lõi 748, hoặc qua buýt ngoại vi 749. Các kiến trúc cho buýt ngoại vi bao gồm PCI, USB, và tương tự.

CPU 741, GPU 742, FPGA 743, và các bộ gia tốc 744 có thể thực thi các lệnh cụ thể mà, khi kết hợp, có thể tạo thành mã máy tính nêu trên. Mã máy tính đó có thể được lưu trữ trong ROM 745 hoặc RAM 746. Dữ liệu tĩnh tiền cũng có thể được lưu trữ trong RAM 746, trong đó dữ liệu vĩnh cửu có thể được lưu trữ chẳng hạn, trong bộ lưu trữ khối nội bộ 747. Lưu trữ nhanh và truy xuất vào thiết bị bất kỳ trong các thiết bị nhớ có thể được kích hoạt qua việc sử dụng của bộ nhớ cache, mà có thể được liên kết chặt chẽ với một hoặc nhiều CPU 741, GPU 742, lưu trữ khối 747, ROM 745, RAM 746, và tương tự.

Các phương tiện máy tính đọc được có thể có mã máy tính trên đó để thực hiện các hoạt động được triển khai bằng máy tính khác nhau. Media và mã máy tính có thể được thiết kế và xây dựng đặc biệt theo sáng chế, hoặc chúng có thể là loại đã biết và khả dụng cho người có hiểu biết trung bình trong lĩnh vực phần mềm máy tính.

Như là ví dụ và không giới hạn, hệ thống máy tính có kiến trúc 700, và cụ thể là lõi 740 có thể tạo chức năng như là kết quả của (các) bộ xử lý (bao gồm CPU,

GPU, FPGA, các bộ gia tốc, và tương tự) thực thi phần mềm được triển khai trong một hoặc nhiều vật máy tính đọc được hữu hình. Vật máy tính đọc được này có thể là phương tiện được liên kết với bộ lưu trữ khối người dùng truy nhập được như nêu trên, cũng như bộ lưu trữ cụ thể của lõi 740 về bản chất là bất biến, chẳng hạn bộ lưu trữ khối nội bộ lõi 747 hoặc ROM 745. Phần mềm thực hiện các phương án thực hiện khác nhau của sáng chế có thể được lưu trữ trong các thiết bị này và được thực thi bằng lõi 740. Phương tiện máy tính đọc được có thể bao gồm một hoặc nhiều thiết bị nhớ hoặc chip, theo các nhu cầu cụ thể. Phần mềm có thể khiến lõi 740 và cụ thể là các bộ xử lý trong đó (bao gồm CPU, GPU, FPGA, và tương tự) để thực thi các quá trình cụ thể hoặc các phần cụ thể của các quá trình cụ thể được mô tả ở đây, bao gồm định nghĩa các cấu trúc dữ liệu được lưu trữ trong RAM 746 và chỉnh sửa các cấu trúc dữ liệu này theo các quá trình được định nghĩa bằng phần mềm. Ngoài ra, hệ thống máy tính có thể tạo chức năng như là kết quả của nối cứng lô gic hoặc ngược lại được triển khai trong mạch (chẳng hạn: bộ gia tốc 744), 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. Tham chiếu đến phần mềm có thể bao gồm lô gic, và ngược lại, nếu thích hợp. Tham chiếu đến vật máy tính đọc được có thể bao gồm mạch (chẳng hạn, mạch tích hợp (IC)) lưu trữ phần mềm để thực thi, mạch triển khai lô gic để thực thi, hoặc cả hai, nếu thích hợp. Sáng chế bao gồm kết hợp thích hợp bất kỳ của phần cứng và phần mềm.

Trong khi sáng chế mô tả vài phương án thực hiện lấy làm ví dụ, có các thay thế, hoán đổi, và các thay thế tương đương khác nhau, mà nằm trong phạm vi của sáng chế. Do vậy, cần hiểu rằng người có hiểu biết trung bình trong lĩnh vực sẽ có thể phát minh ra các hệ thống và các phương pháp mà, mặc dù không được thể hiện tường minh hoặc được mô tả ở đây, thực hiện các nguyên lý của sáng chế và do vậy nằm trong nguyên lý và phạm vi của sáng chế.

YÊU CẦU BẢO HỘ

1. Phương pháp giải mã, trong đó phương pháp bao gồm các bước:

nhận dòng bit bao gồm dữ liệu ảnh/video được nén, trong đó dòng bit có nhiều lớp;

phân tách hoặc dẫn xuất, từ dòng bit, bộ chỉ báo chế độ tập hợp lớp đầu ra trong tập hợp tham số video (VPS);

nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra;

nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tập hợp lớp đầu ra được nhận diện; và

giải mã một hoặc nhiều lớp đầu ra ảnh được nhận diện, bước giải mã bao gồm bước thiết lập cờ đầu ra ảnh bằng 0 dựa trên việc định danh VPS lớn hơn 0, chỉ báo rằng có nhiều hơn một lớp trong dòng bit, và dựa trên phân tử cú pháp chỉ báo rằng lớp hiện tại không phải là lớp đầu ra.

2. Phương pháp theo điểm 1, trong đó việc nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra bao gồm:

trong trường hợp mà bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS là giá trị thứ nhất, nhận diện lớp cao nhất trong dòng bit dưới dạng một hoặc nhiều lớp đầu ra ảnh;

trong trường hợp mà bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS là giá trị thứ hai, nhận diện tất cả các lớp trong dòng bit dưới dạng một hoặc nhiều lớp đầu ra ảnh; và

trong trường hợp mà bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS là giá trị thứ ba, nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tường minh trong VPS, trong đó:

giá trị thứ nhất khác với giá trị thứ hai và khác với giá trị thứ ba, và

giá trị thứ hai khác với giá trị thứ ba.

3. Phương pháp theo điểm 2, trong đó giá trị thứ nhất bằng 0, giá trị thứ hai bằng 1, và giá trị thứ ba bằng 2.

4. Phương pháp theo điểm 2, trong đó việc nhận diện một hoặc nhiều lớp đầu ra ảnh bằng cách báo hiệu tường minh trong VPS bao gồm: (i) phân tách hoặc dẫn

xuất, từ VPS, cờ lớp đầu ra, và (ii) thiết lập các lớp có cờ lớp đầu ra bằng 1 là một hoặc nhiều lớp đầu ra ảnh.

5. Phương pháp theo điểm 1, trong đó việc nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra bao gồm:

trong trường hợp mà bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS là giá trị định trước, việc báo hiệu tập hợp lớp đầu ra bao gồm nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tường minh trong VPS.

6. Phương pháp theo điểm 5, trong đó việc nhận diện một hoặc nhiều lớp đầu ra ảnh bằng cách báo hiệu tường minh trong VPS bao gồm các bước: (i) phân tách hoặc dẫn xuất, từ VPS, cờ lớp đầu ra, và (ii) thiết lập các lớp có cờ lớp đầu ra bằng 1 là một hoặc nhiều lớp đầu ra ảnh, trong đó số lượng nhiều lớp lớn hơn 2.

7. Phương pháp theo điểm 5, trong đó việc báo hiệu tập hợp lớp đầu ra bao gồm nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tường minh trong VPS khi bộ chỉ báo chế độ tập hợp lớp đầu ra bằng 2, và số lượng lớp trong nhiều lớp lớn hơn 2.

8. Phương pháp theo điểm 1, trong đó:

việc báo hiệu tập hợp lớp đầu ra bao gồm nhận diện lớp cao nhất trong dòng bit hoặc tất cả các lớp trong dòng bit dưới dạng một hoặc nhiều lớp đầu ra ảnh bằng cách suy ra một hoặc nhiều lớp đầu ra ảnh, khi bộ chỉ báo chế độ tập hợp lớp đầu ra nhỏ hơn 2, và số lượng nhiều lớp bằng 2.

9. Phương pháp theo điểm 8, trong đó bộ chỉ báo `number_output_layer_sets_minus1` trong VPS chỉ báo số lượng tập hợp lớp đầu ra.

10. Phương pháp theo điểm 9, trong đó bộ chỉ báo `vps_maximum_layer_minus1` trong VPS chỉ báo số lượng lớp trong dòng bit.

11. Phương pháp theo điểm 2, trong đó nếu tất cả các lớp trong dòng bit là các lớp độc lập mà không có lệ thuộc phân tách và giải mã vào lớp khác, trong đó `vps_all_independent_layers_flag` của VPS bằng 1, bộ chỉ báo chế độ tập hợp lớp đầu ra không được báo hiệu, và giá trị của bộ chỉ báo chế độ tập hợp lớp đầu ra được suy ra là giá trị thứ hai.

12. Phương pháp theo điểm 1, trong đó phương pháp còn bao gồm bước: điều khiển bộ hiển thị để hiển thị một hoặc nhiều lớp đầu ra ảnh được giải mã.

13. Vật ghi máy tính đọc được bất biến lưu trữ các lệnh mà, khi được thực thi, khiến hệ thống hoặc thiết bị bao gồm một hoặc nhiều bộ xử lý:

nhận dòng bit bao gồm dữ liệu ảnh/video được nén, trong đó dòng bit có nhiều lớp;

phân tách hoặc dẫn xuất, từ dòng bit, bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS;

nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra;

nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tập hợp lớp đầu ra được nhận diện; và

giải mã một hoặc nhiều lớp đầu ra ảnh được nhận diện, bước giải mã bao gồm bước thiết lập cờ đầu ra ảnh bằng 0 dựa trên việc định danh VPS lớn hơn 0, chỉ báo rằng có nhiều hơn một lớp trong dòng bit, và dựa trên phần tử cú pháp chỉ báo rằng lớp hiện tại không phải là lớp đầu ra.

14. Vật lưu trữ máy tính đọc được bất biến theo điểm 13, trong đó các lệnh còn được tạo cấu hình để khiến hệ thống hoặc thiết bị bao gồm một hoặc nhiều bộ xử lý:

điều khiển bộ hiển thị để hiển thị một hoặc nhiều lớp đầu ra ảnh được giải mã.

15. Thiết bị giải mã bao gồm:

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

ít nhất một bộ xử lý được tạo cấu hình để truy nhập ít nhất một bộ nhớ và hoạt động theo mã chương trình máy tính, mã chương trình máy tính bao gồm:

mã nhận được tạo cấu hình để khiến ít nhất một bộ xử lý nhận dòng bit bao gồm dữ liệu ảnh/video được nén, trong đó dòng bit có nhiều lớp;

mã phân tách hoặc dẫn xuất được tạo cấu hình để khiến ít nhất một bộ xử lý phân tách hoặc dẫn xuất, từ dòng bit, bộ chỉ báo chế độ tập hợp lớp đầu ra trong VPS;

mã nhận diện báo hiệu lớp đầu ra được tạo cấu hình để khiến ít nhất một bộ xử lý nhận diện báo hiệu tập hợp lớp đầu ra dựa trên bộ chỉ báo chế độ tập hợp lớp đầu ra;

mã nhận diện lớp đầu ra ảnh được tạo cấu hình để khiến ít nhất một bộ xử lý nhận diện một hoặc nhiều lớp đầu ra ảnh dựa trên báo hiệu tập hợp lớp đầu ra được nhận diện; và

mã giải mã được tạo cấu hình để khiến ít nhất một bộ xử lý: giải mã một hoặc nhiều lớp đầu ra ảnh được nhận diện,

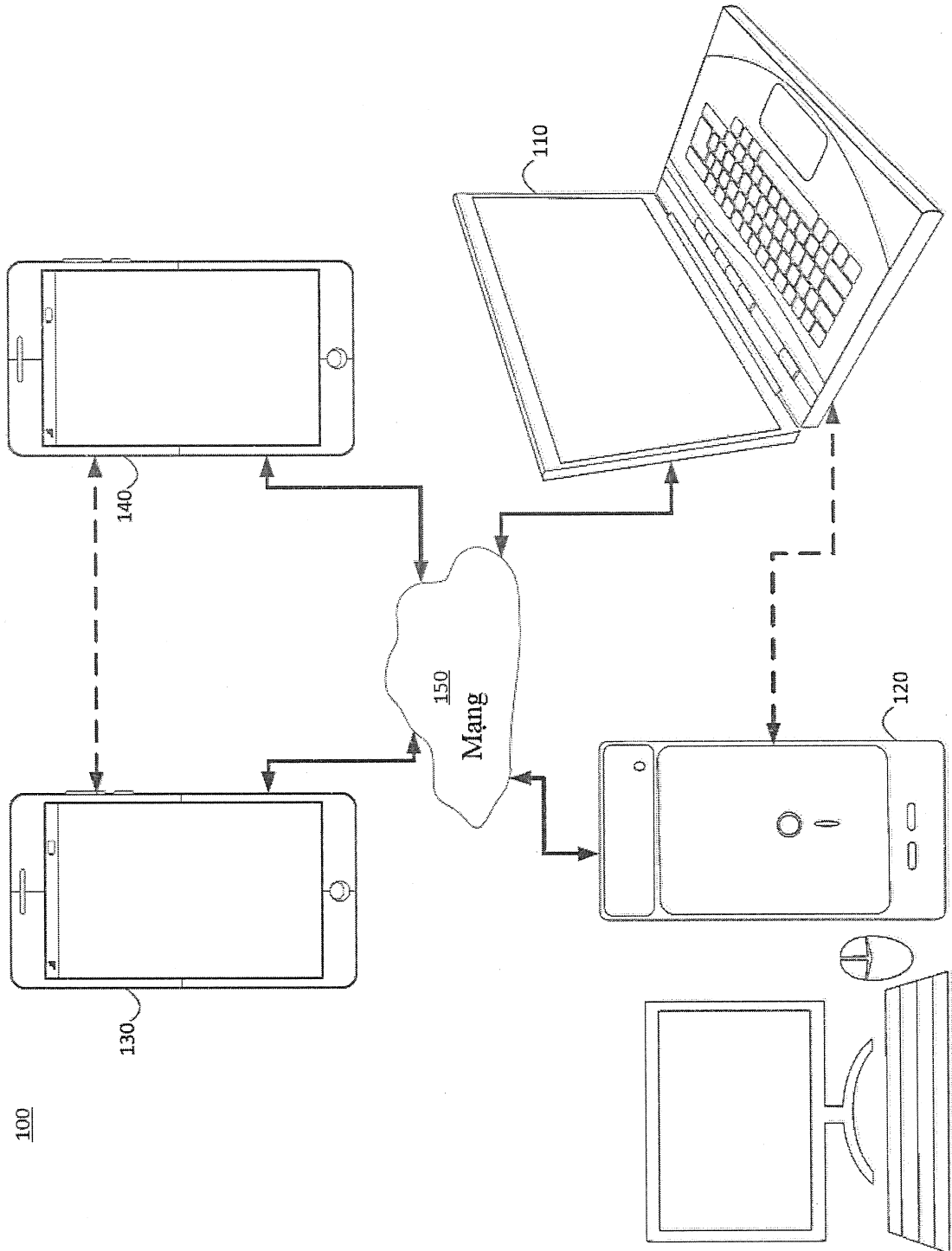
trong đó mã giải mã được tạo cấu hình để khiến ít nhất một bộ xử lý thiết lập cờ đầu ra ảnh bằng 0 dựa trên việc định danh VPS lớn hơn 0, chỉ báo rằng có nhiều hơn một lớp trong dòng bit, và dựa trên phần tử cú pháp chỉ báo rằng lớp hiện tại không phải là lớp đầu ra.

16. Thiết bị theo điểm 15, trong đó mã chương trình máy tính còn bao gồm:

mã điều khiển hiển thị được tạo cấu hình để khiến ít nhất một bộ xử lý điều khiển bộ hiển thị để hiển thị một hoặc nhiều lớp đầu ra ảnh.

17. Thiết bị theo điểm 15, trong đó mã giải mã còn được tạo cấu hình để khiến ít nhất một bộ xử lý thiết lập cờ đầu ra ảnh bằng cờ đầu ra ảnh được báo hiệu trong tiêu đề ảnh, dựa trên việc định danh VPS không lớn hơn 0 hoặc dựa trên phần tử cú pháp chỉ báo rằng lớp hiện tại là lớp đầu ra.

1/27

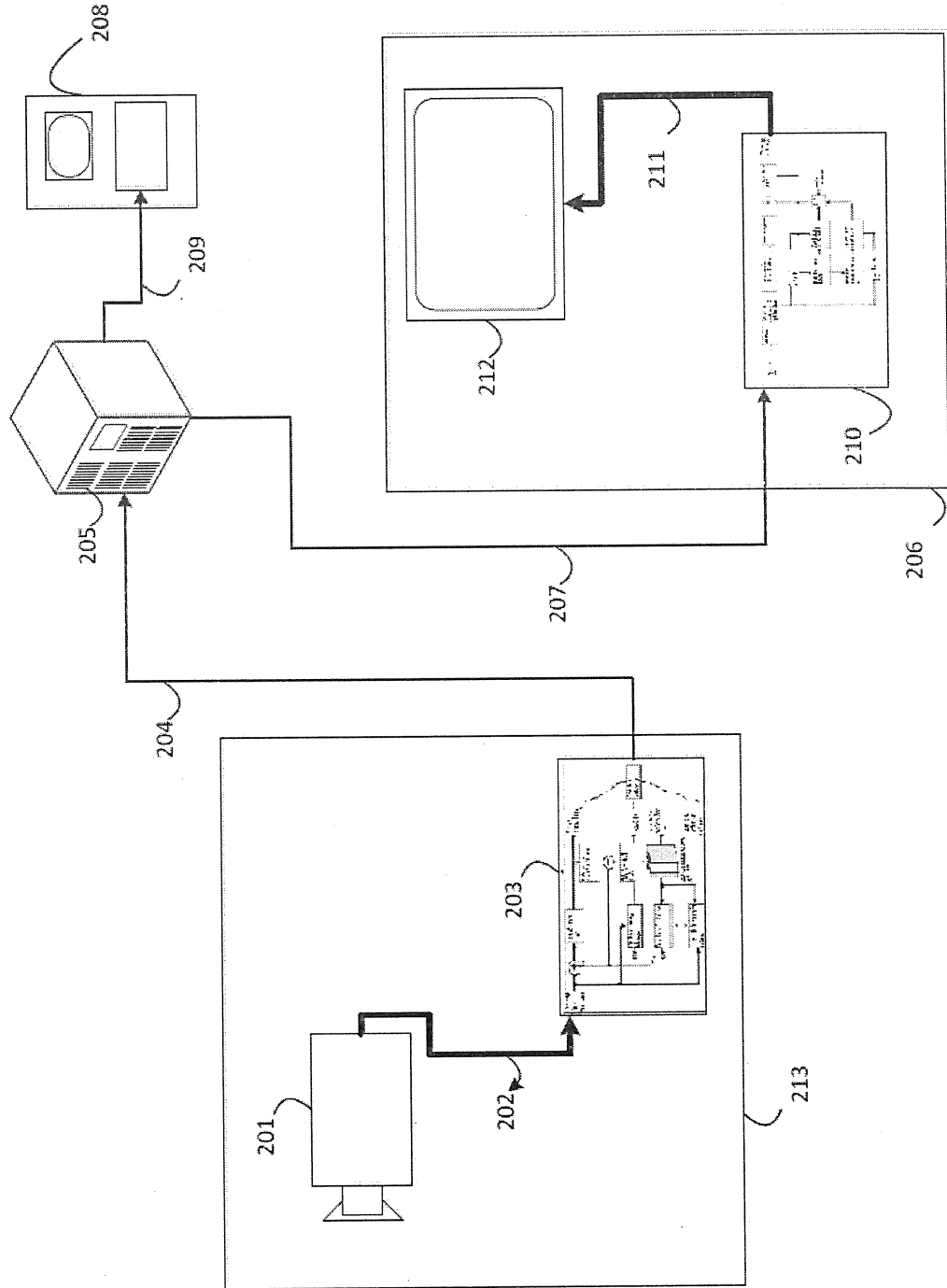


100

Fig.1

2/27

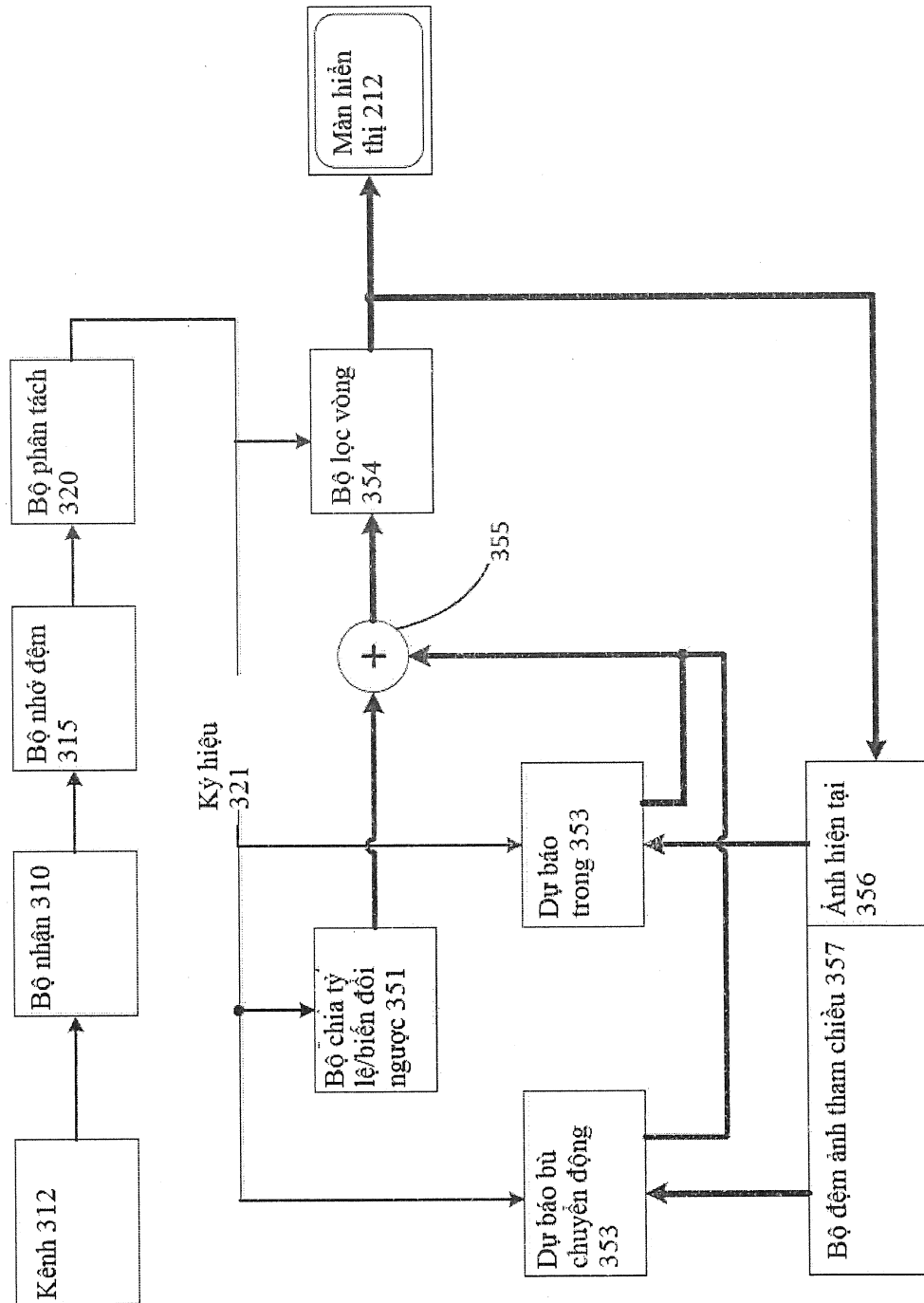
Fig.2
Hệ thống streaming 200



3/27

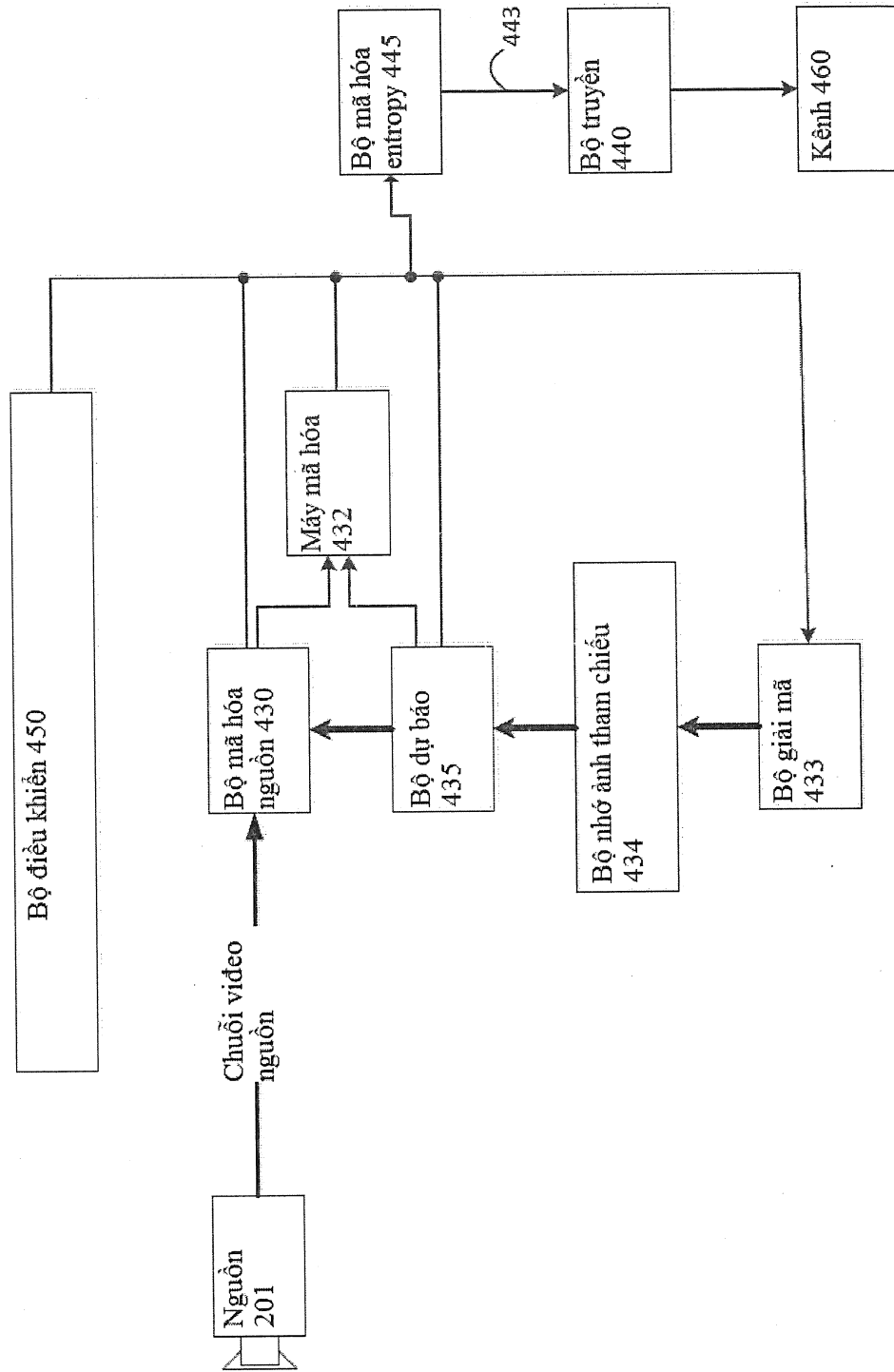
Bộ giải mã 210

Fig.3



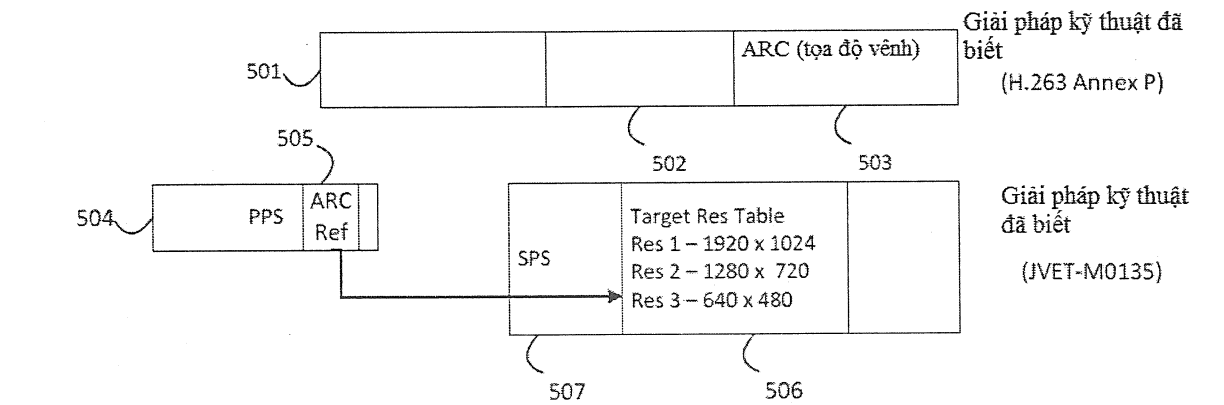
4/27

Fig.4 Bộ mã hóa 203

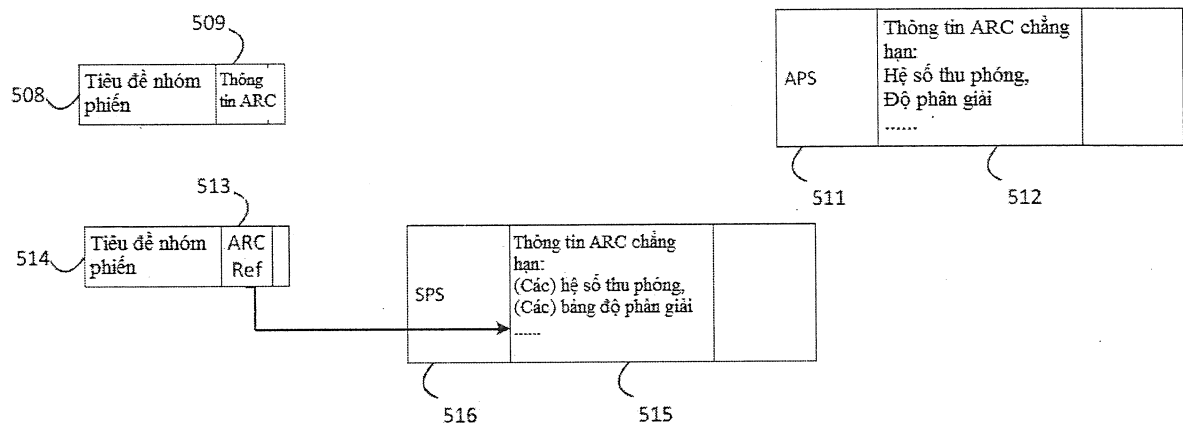


5/27

Fig.5



Sáng chế được bộc lộ



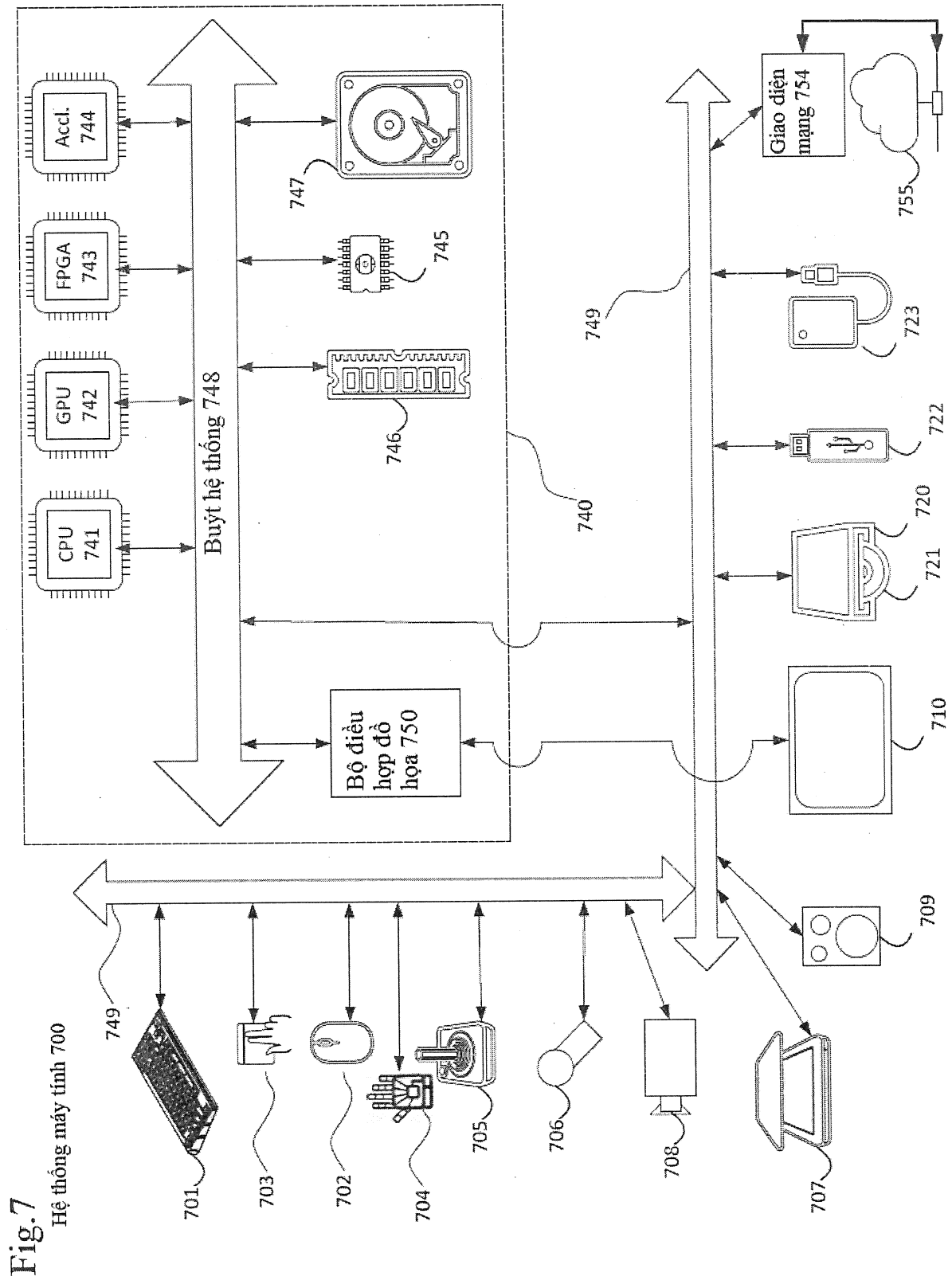
6/27

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

7/27



8/27

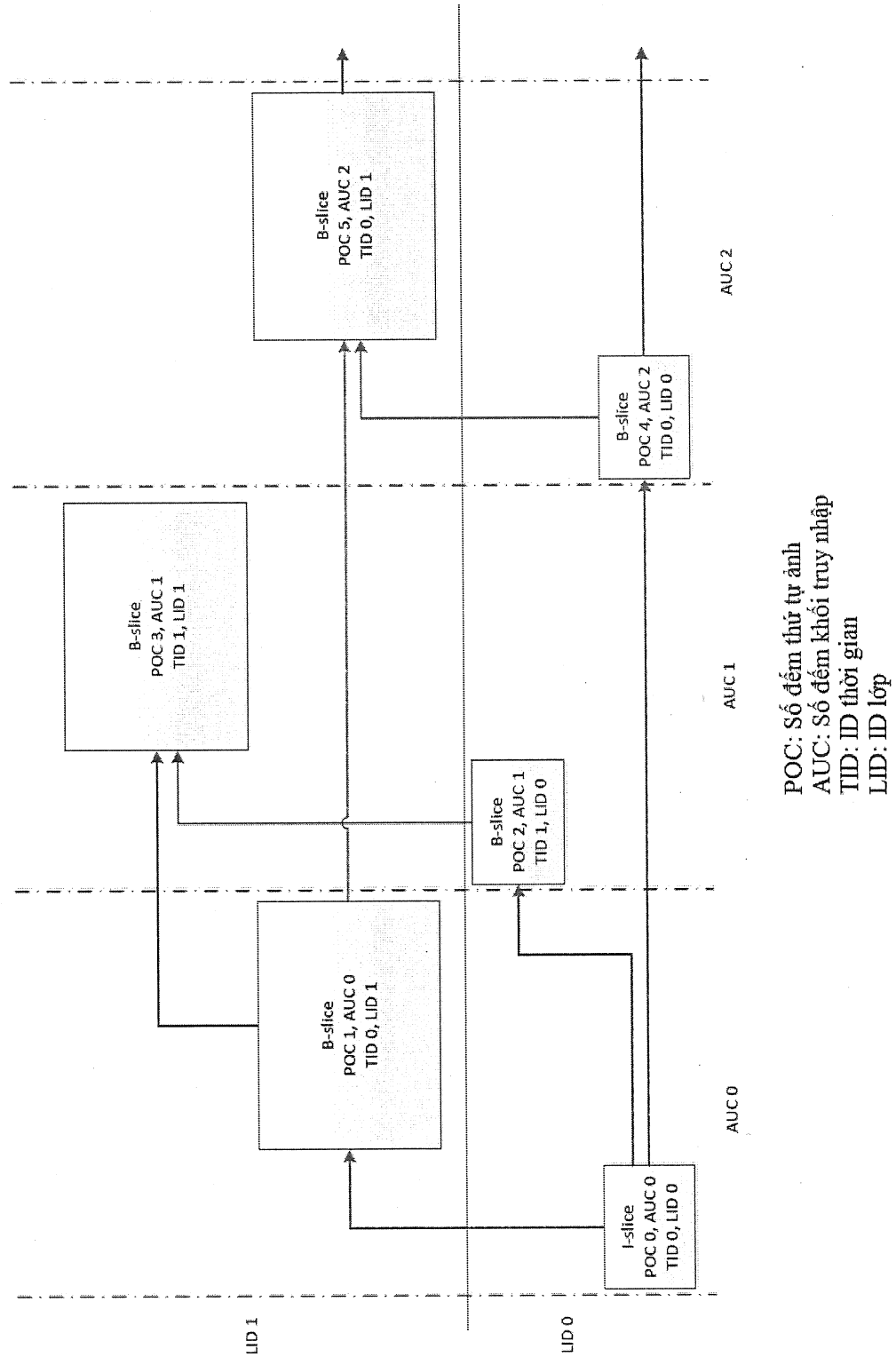
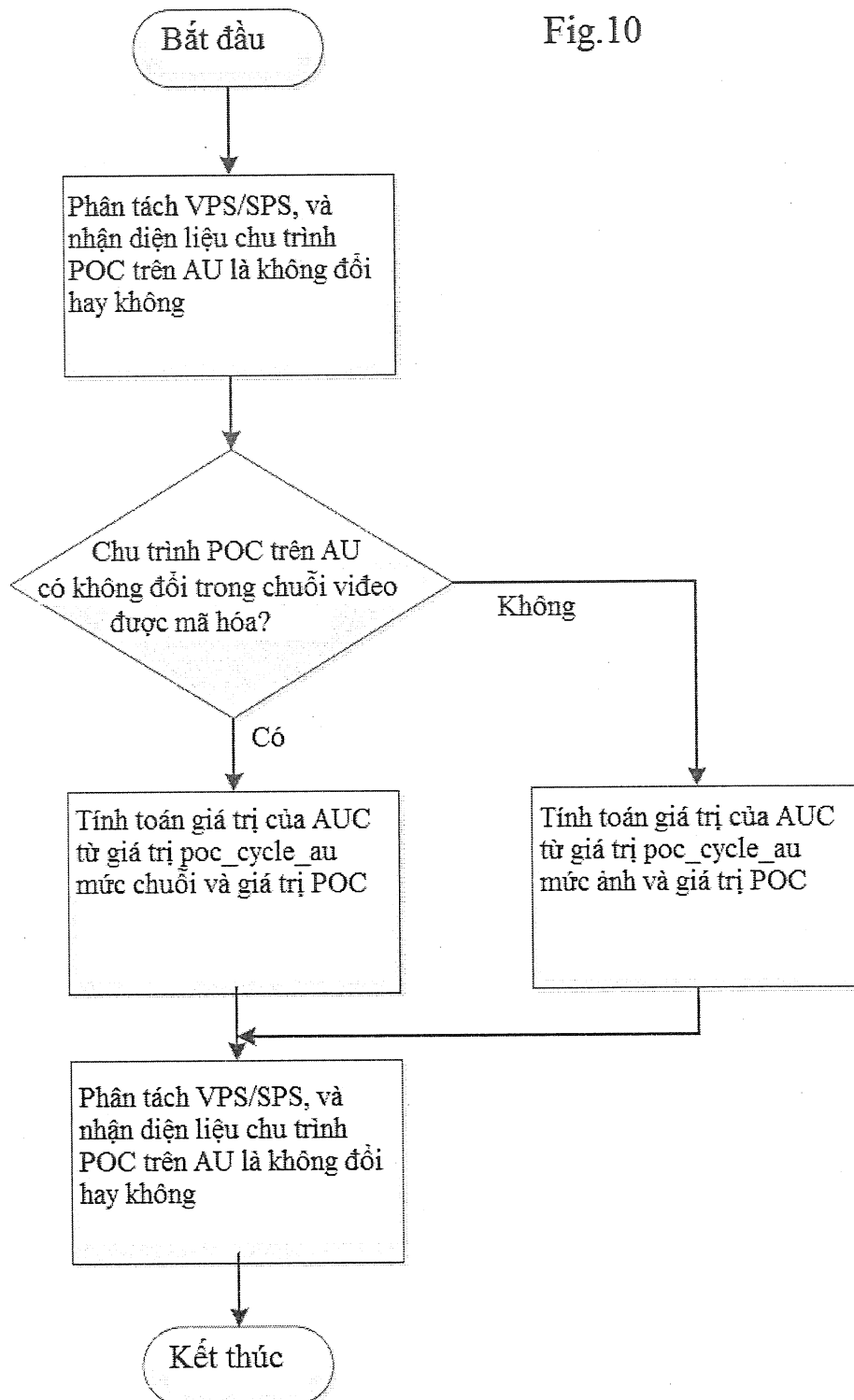


Fig. 9

Descriptor	Code
	video_parameter_set_chap() {
u(4) ³²	vps_video_parameter_set_id ³²
u(8) ³²	vps_max_layers_minus1 ³²
e ³²	for(i = 0; i <= vps_max_layers_minus1; i++) {
u(7) ³²	vps_included_layer_id[i] ³²
u(1) ³²	vps_reserved_zero_bit ³²
e ³²	}
u(1) ³²	vps_constraint_info_present_flag ³²
u(1) ³²	vps_constant_poc_cycle_per_aud ³²
e ³²	if(vps_constant_poc_cycle_per_aud)
u(8) ³²	vps_poc_cycle_aud ³²
e ³²	... ³²
e ³²	}
Descriptor	e ³²
ue(v) ³²	slice_header() {
e ³²	slice_pic_parameter_set_id ³²
u(v) ³²	if(rect_slice_flag NumBricksInPic > 1)
e ³²	slice_address ³²
ue(v) ³²	if(!rect_slice_flag && !single_brick_per_slice_flag)
ue(v) ³²	num_bricks_in_slice_minus1 ³²
ue(v) ³²	slice_type ³²
e ³²	if(NalUnitType == GRA_NUT)
se(v) ³²	recovery_poc_cnt ³²
u(v) ³²	slice_pic_order_cnt_minus1 ³²
e ³²	... ³²
e ³²	if(!vps_constant_poc_cycle_per_aud)
u(8) ³²	slice_poc_cycle_aud ³²
e ³²	e ³²
e ³²	e ³²
e ³²	}

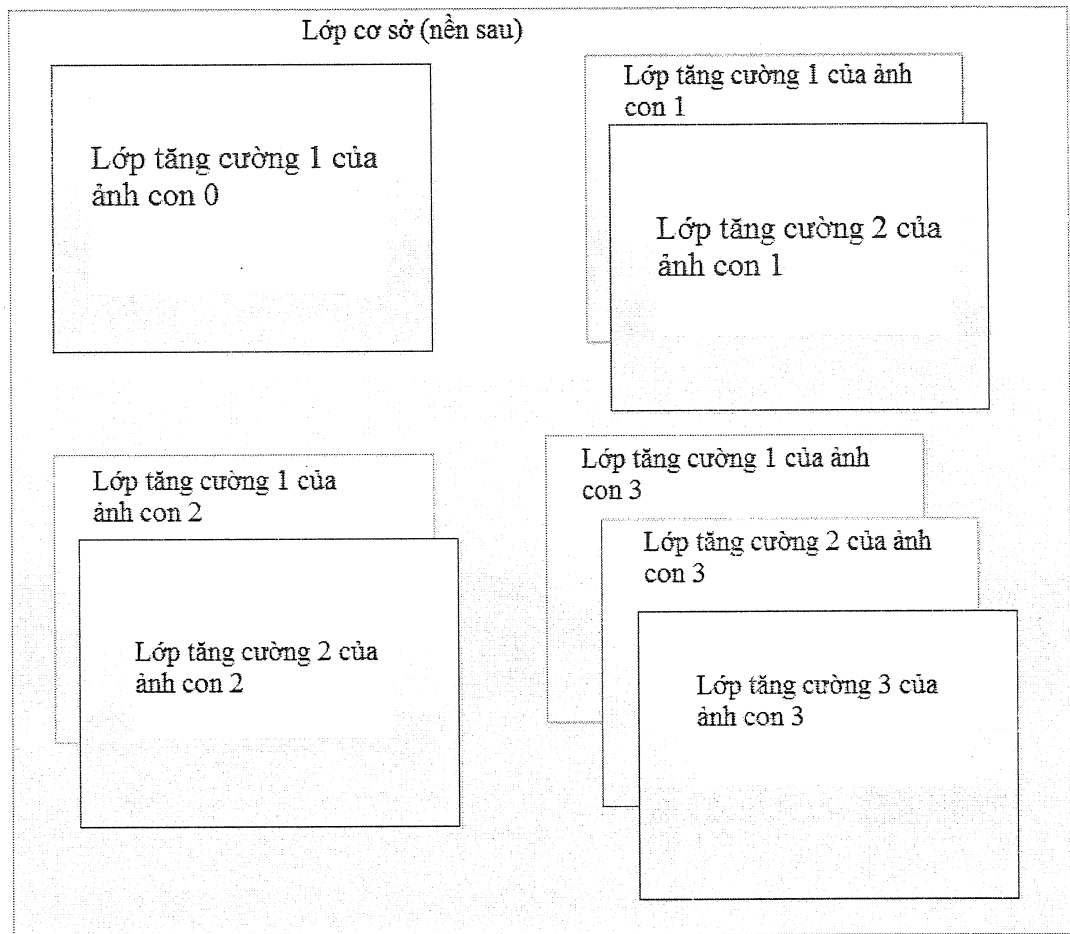
10/27

Fig.10



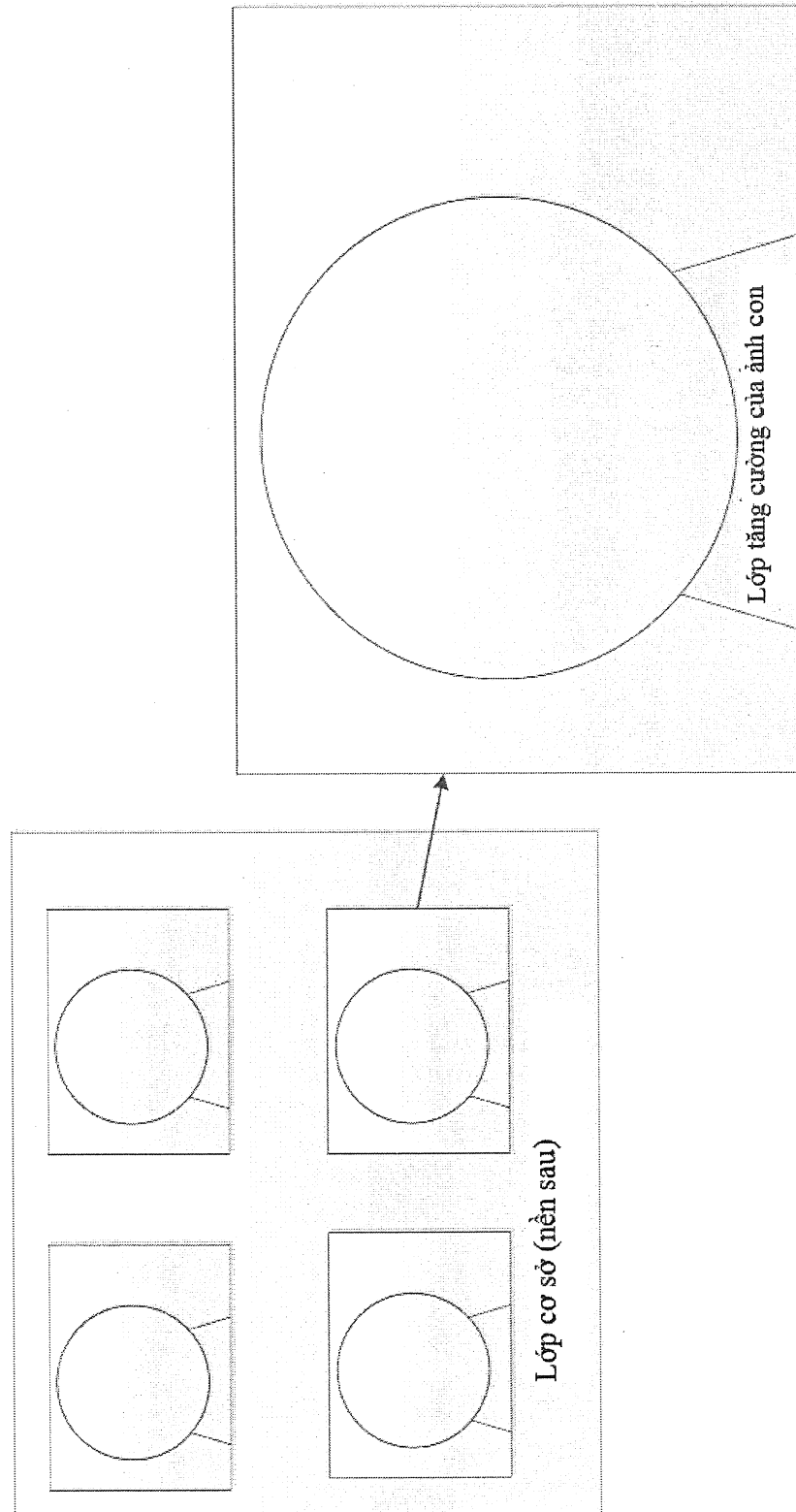
11/27

Fig.11



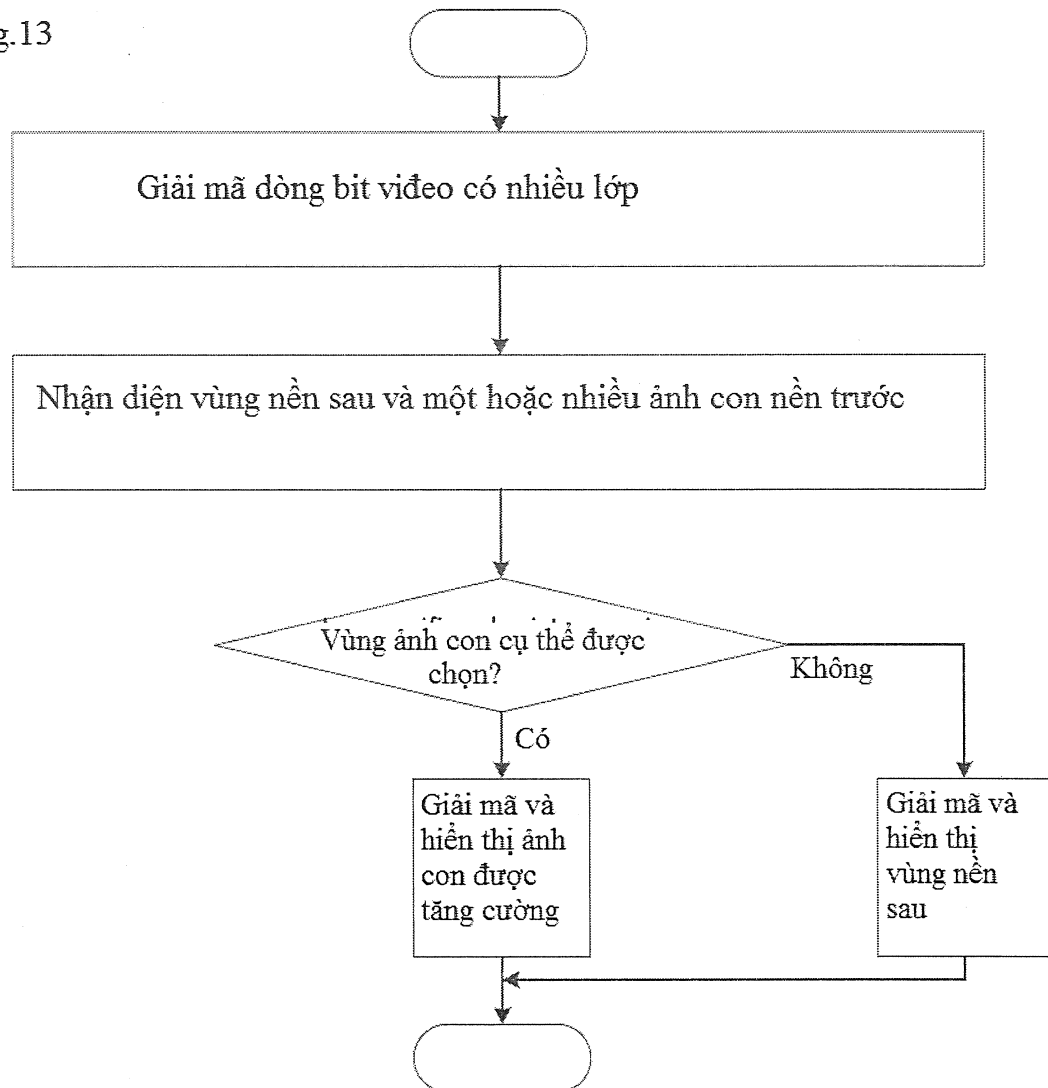
12/27

Fig.12



13/27

Fig.13



14/27

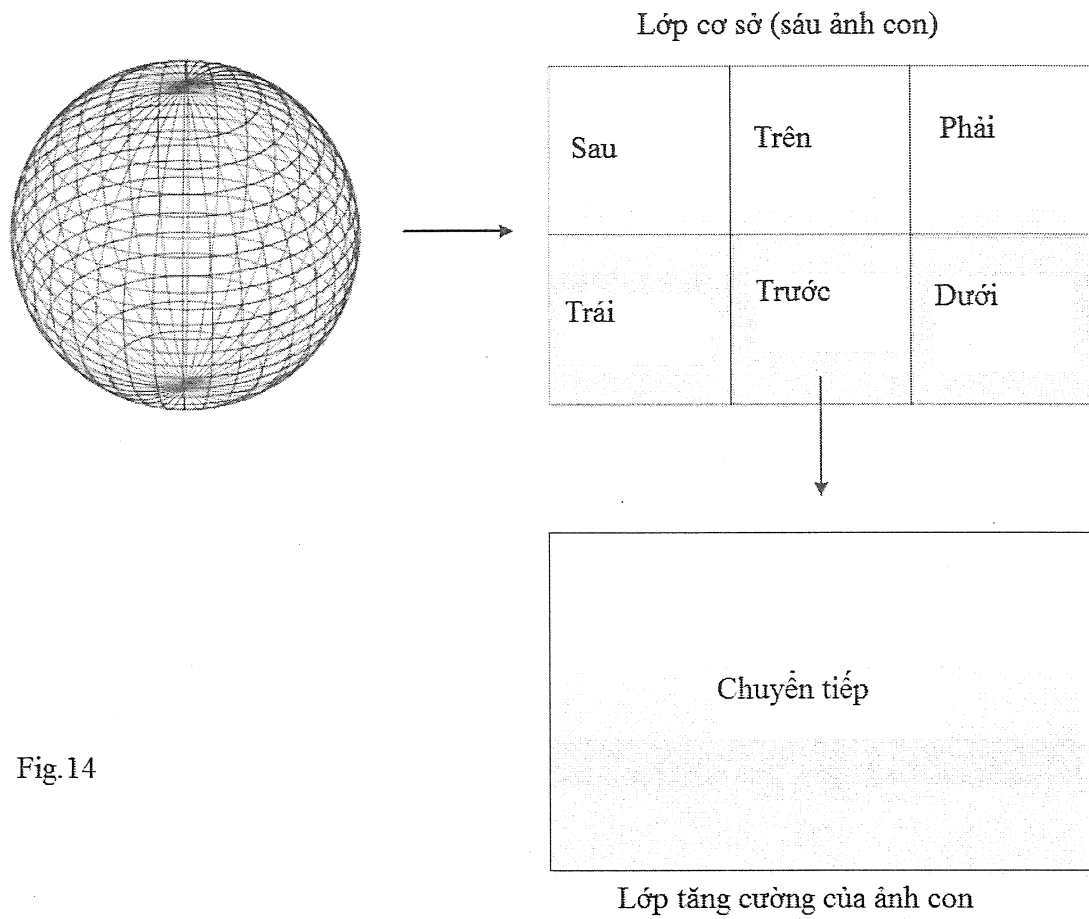
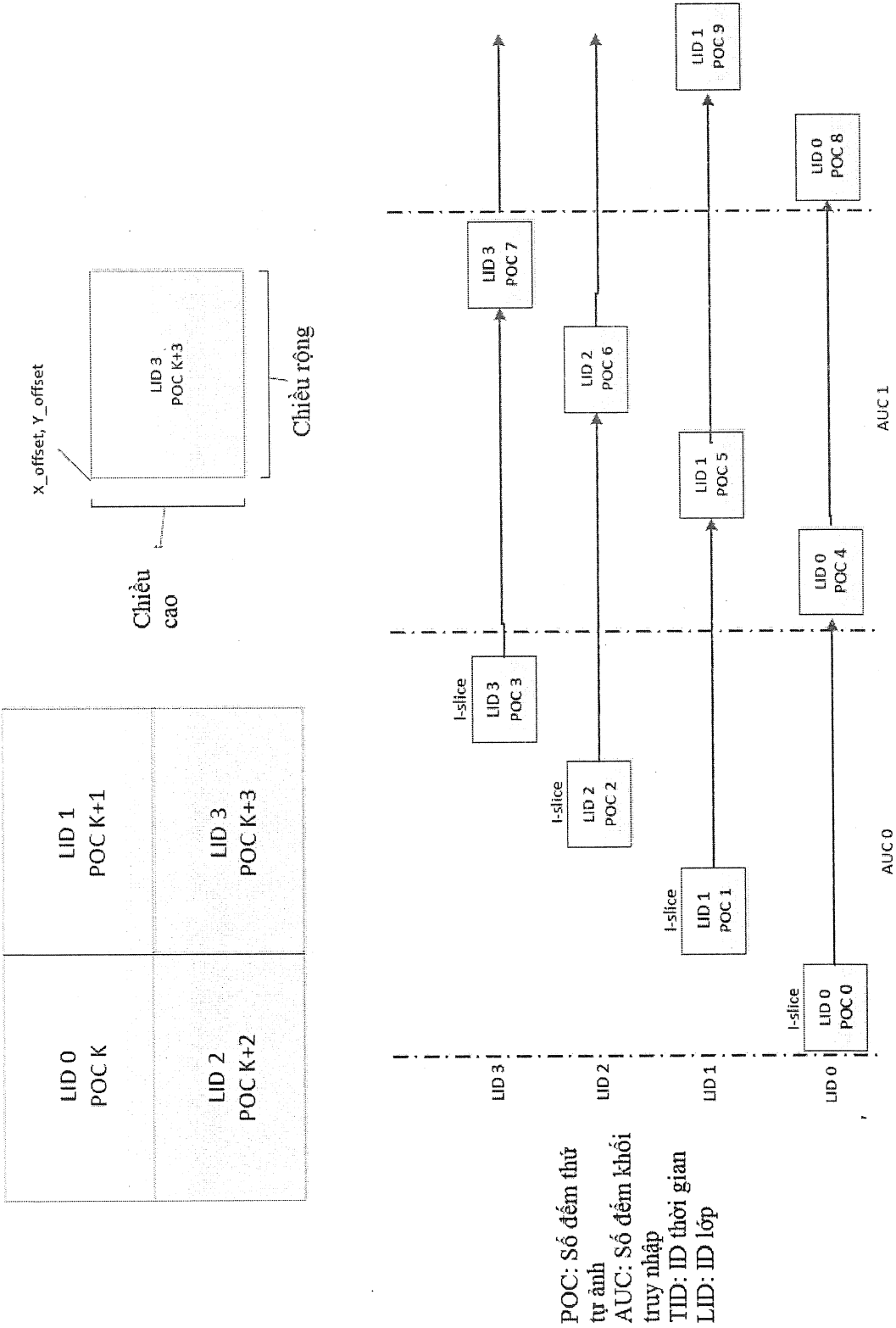


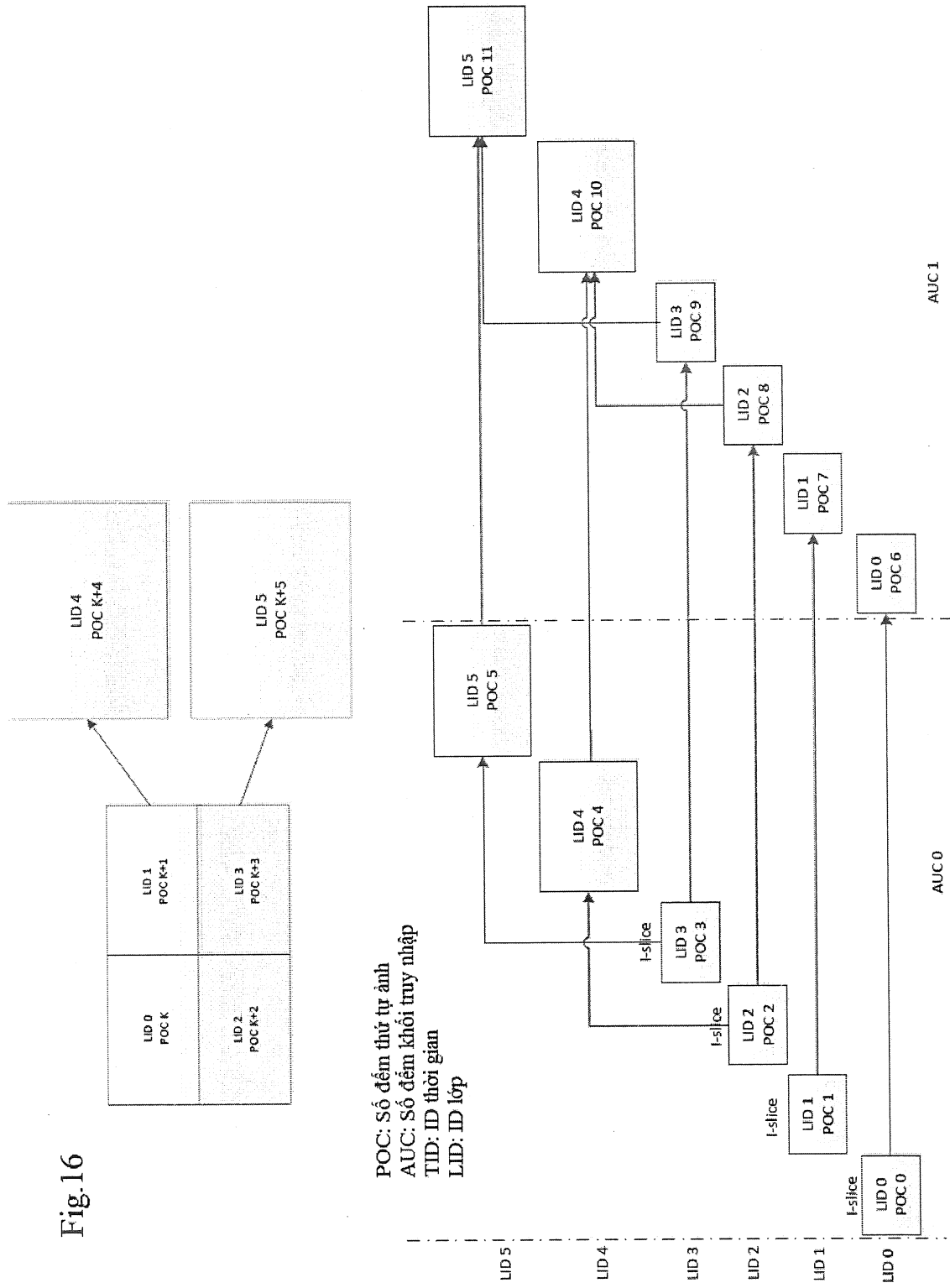
Fig.14

15/27

Fig.15



16/27



Descriptor ¹	
video_parameter_set_chap() { ²	
vps_video_parameter_set_id ³	u(4) ²
vps_max_layers_minus1 ³	u(3) ²
for(i = 0; i <= vps_max_layers_minus1; i++) { ⁴	⁴
vps_included_layer_id[i] ³	u(7) ²
vps_reserved_zero_bit ³	u(1) ²
} ²	²
... ²	²
vps_sub_picture_dividing_flag ³	u(1) ²
if(vps_sub_picture_dividing_flag) { ⁴	⁴
vps_full_pic_width_in_luma_samples ³	ue(v) ²
vps_full_pic_height_in_luma_samples ³	ue(v) ²
} ²	²
... ²	²
} ²	²
²	
seq_parameter_set_rbsp() { ¹	Descriptor ¹
sps_decoding_parameter_set_id ³	u(4) ²
sps_video_parameter_set_id ³	u(4) ²
sps_max_sub_layers_minus1 ³	u(3) ²
... ³	³
pic_width_in_luma_samples ³	ue(v) ²
pic_height_in_luma_samples ³	ue(v) ²
if(vps_sub_picture_dividing_flag) { ⁴	⁴
pic_offset_x ³	ue(v) ²
pic_offset_y ³	ue(v) ²
} ²	²
... ²	²
} ²	²

18/27

Descriptor	
$ug(v)^*$	<code>sub_region_partitioning_info(payloadSize) {</code>
$ug(v)^*$	<code> num_sub_region</code>
$ug(v)^*$	<code> num_layers</code>
e^*	<code> for(j = 0; j <= num_layers; i++)</code>
$ug(v)^*$	<code> layer_id[i]</code>
e^*	<code> for(i = 1; i < num_layers; i++)</code>
e^*	<code> for(j = 0; j < i; j++)</code>
$u(l)^*$	<code> direct_dependency_flag[i][j]</code>
e^*	<code> for(i = 0; i < num_sub_region; i++) {</code>
$ug(v)^*$	<code> num_layers_for_region[i]</code>
e^*	<code> for(j = 0; j < num_layers_for_region[i]; j++)</code>
$ug(v)^*$	<code> sub_region_layer_id[i][j]</code>
$ug(v)^*$	<code> sub_region_offset_x[i]</code>
$ug(v)^*$	<code> sub_region_offset_y[i]</code>
$ug(v)^*$	<code> sub_region_width[i]</code>
$ug(v)^*$	<code> sub_region_height[i]</code>
e^*	<code> }</code>
e^*	<code> ...</code>
e^*	<code>}</code>

Fig.18

19/27

Fig.19

video parameter set rbsp() {	Descriptor
...	
vps_max_lavers_minus1	u(6)
num_output_layer_sets	ue(v)
num_profile_tier_level	ue(v)
for(i = 0; i < num_profile_tier_level; i++)	
profile_tier_level(vps_max_sub_lavers_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)
}	
...	
}	

20/27

Fig.20

video parameter set rbsp() {	Descriptor
...	
vps_max_lavers_minus1	u(6)
if(vps_max_lavers_minus1 > 0) {	
num_output_laver_sets	ue(v)
num_profile_tier_level	ue(v)
}	
for(i=0; i<num_profile_tier_level; i++)	
profile_tier_level(vps_max_sub_lavers_minus1)	
for(i=0; i<num_output_laver_sets; i++) {	
vps_output_lavers_mode[i]	u(2)
vps_ptl_signal_flag[i]	u(1)
for(j=0; j<NumLayersInIdList[I]; i++) {	
if(vps_output_lavers_mode[i] == 2)	
output_laver_flag[I][j]	u(1)
if vps_ptl_signal_flag[i])	
profile_tier_level_idx[I][j]	u(v)
}	
}	
...	
}	

21/27

Fig.21

video parameter set rbsp() {	Descriptor
...	
vps_max_layers_minus1	u(6)
if(vps_max_layers_minus1 > 0) {	
num_output_layer_sets	ue(v)
num_profile_tier_level	ue(v)
}	
max_subpics_minus1	u(8)
for(i=0; i<max_subpics_minus1; i++) {	
sub_pic_id[i]	u(8)
}	
for(i=0; i<num_profile_tier_level; i++)	
profile_tier_level(vps_max_sub_layers_minus1)	
for(i=0; i<num_output_layer_sets; i++) {	
vps_output_layers_mode[i]	u(2)
vps_ptl_signal_flag[i]	u(1)
for(j=0; j<NumLayersInIdList[i]; j++) {	
num_output_subpic_layer[i][j]	ue(v)
for(k=0; k<num_output_subpic_layer[i][j]; k++)	
sub_pic_id_layer[i][j][k]	u(8)
if(vps_output_layers_mode[i] == 2)	
output_layer_flag[i][j]	u(1)
if(vps_ptl_signal_flag[i])	
profile_tier_level_idx[i][j]	u(v)
}	
}	
...	
}	

22/27

Fig.22

video_parameter_set_rbsp() {	Descriptor
vps video parameter set id	u(4)
vps max layers minus1	u(6)
vps max sublayers minus1	u(3)
if vps_max_layers_minus1 > 0 && vps_max_sublayers_minus1 > 0)	
vps_all_layers_same_num_sublayers_flag	u(1)
if vps_max_layers_minus1 > 0)	
vps_all_independent_layers_flag	u(1)
for(i=0; i <= vps_max_layers_minus1; i++) {	
vps_layer_id[i]	u(6)
if i > 0 && !vps_all_independent_layers_flag) {	
vps_independent_layer_flag[i]	u(1)
if !vps_independent_layer_flag[i]) {	
for(i=0; i < i++;)	
vps_direct_ref_layer_flag[i][j]	u(1)
max_tid_ref_present_flag[i]	u(1)
if max_tid_ref_present_flag[i])	
max_tid_ref_pics_plus1[i]	u(3)
}	
}	
}	
if vps_max_layers_minus1 > 0) {	
if vps_all_independent_layers_flag)	
each layer is an ols flag	u(1)
if !each layer is an ols flag) {	
if !vps_all_independent_layers_flag)	
ols_mode_idc	u(2)
if ols_mode_idc == 2) {	
num_output_layer_sets_minus1	u(8)
for(i=1; i <= num_output_layer_sets_minus1; i++)	
for(j=0; j <= vps_max_layers_minus1; j++)	
ols_output_layer_flag[i][j]	u(1)
}	
}	
}	
}	
...	
}	

23/27

Fig.23

	Descriptor
seq parameter set rbsp() {	
sps seq parameter set id	u(4)
sps video parameter set id	u(4)
sps max sublayers minus1	u(3)
sps reserved zero 4bits	u(4)
sps ptl dpb hrd params present flag	u(1)
if(sps ptl dpb hrd params present flag)	
profile tier level(1, sps max sublayers minus1)	
...	
}	

24/27

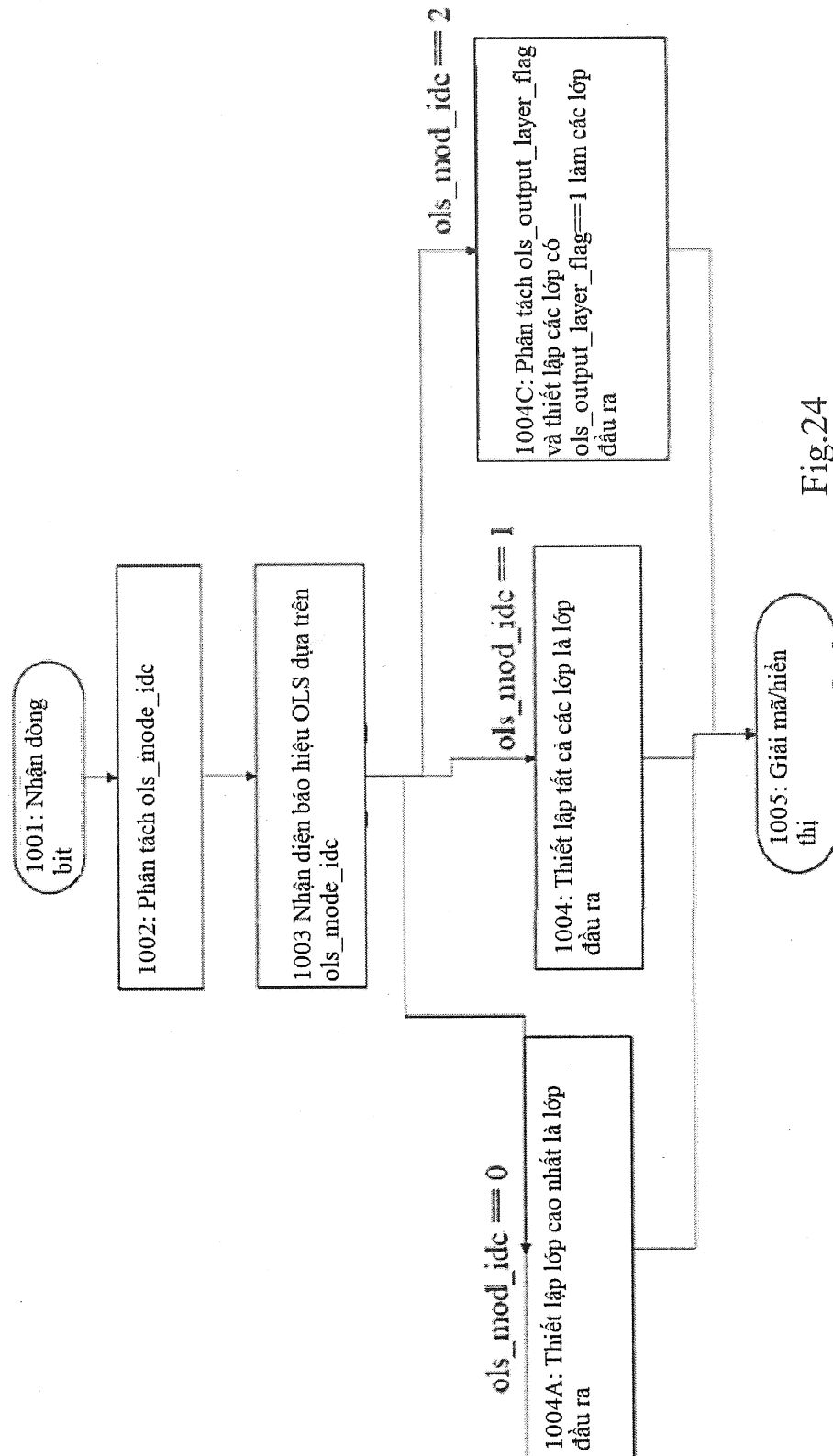


Fig.24

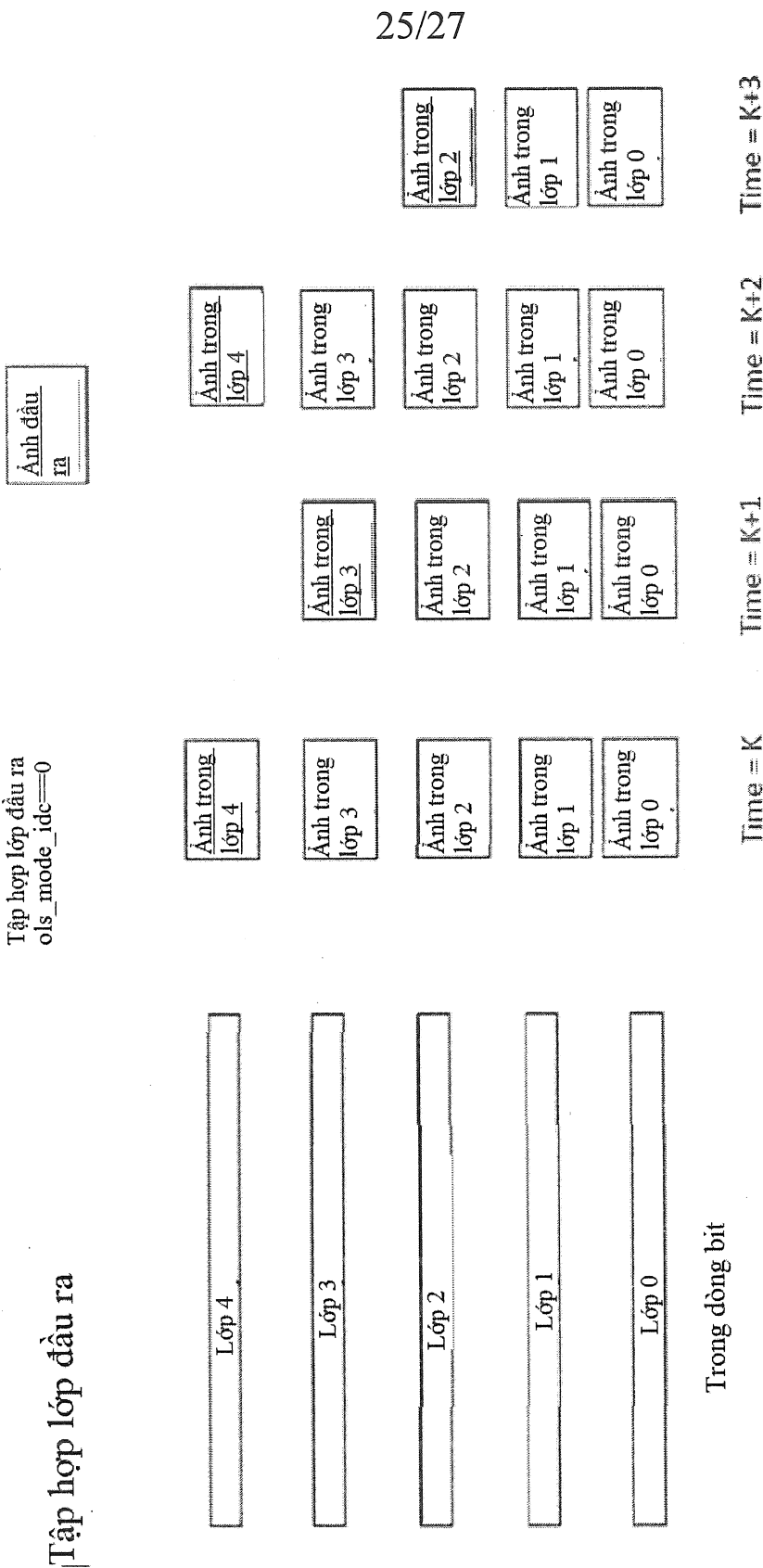


Fig.25A

26/27

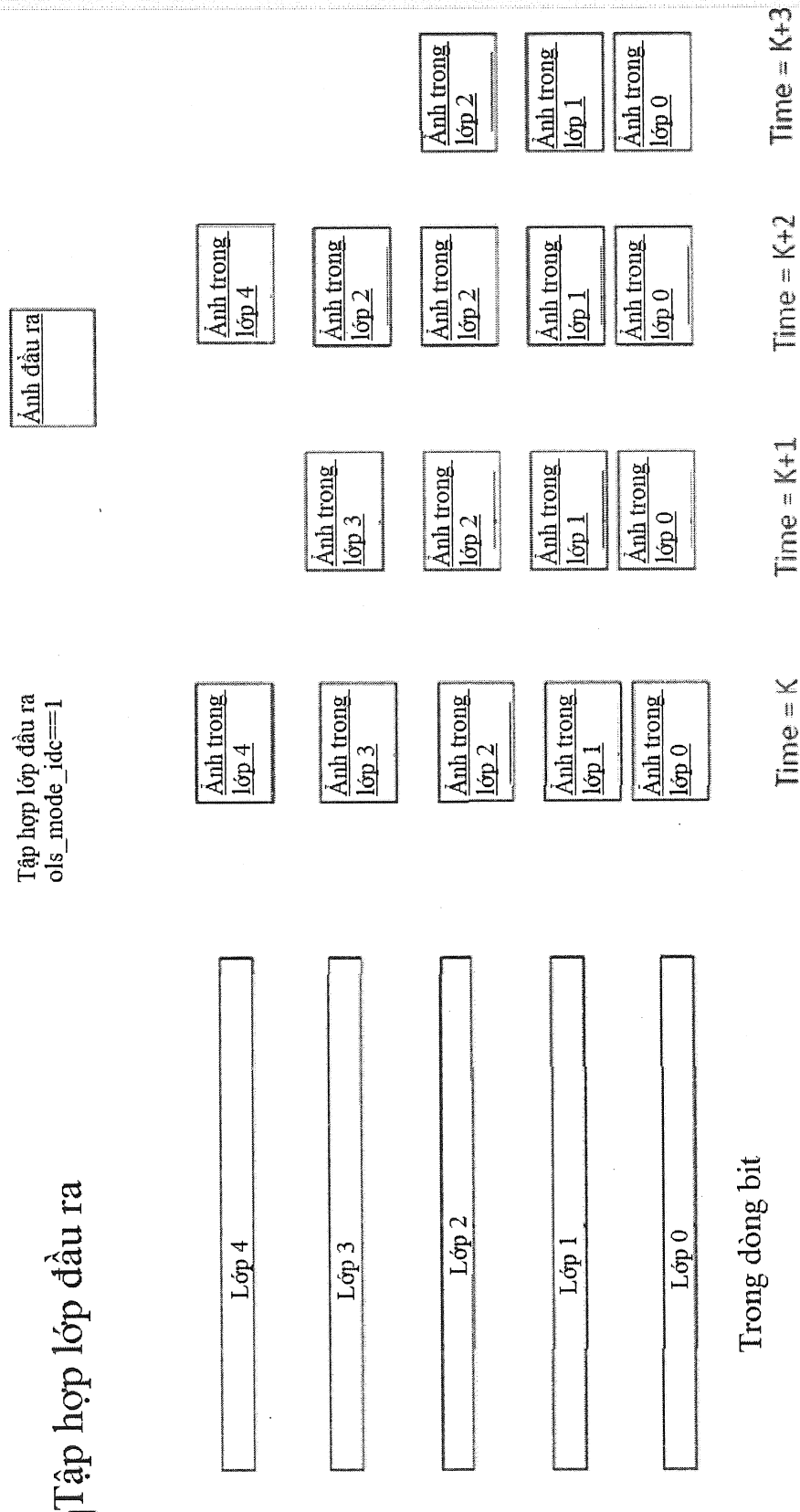


Fig.25B

27/27

